nature biotechnology

Article

https://doi.org/10.1038/s41587-024-02174-7

Branched chemically modified poly(A) tails enhance the translation capacity of mRNA

Received: 25 April 2023

Accepted: 15 February 2024

Published online: 22 March 2024

Check for updates

Hongyu Chen[1,2], Dangliang Liu[1,2], Jianting Guo[1,2], Abhishek Aditham[2,3], Yiming Zhou[1,2,4], Jiakun Tian[1,2], Shuchen Luo[1,2], Jingyi Ren[1,2], Alvin Hsu[5,6,7], Jiahao Huang[1,2], Franklin Kostas[1,2], Mingrui Wu[1,2], David R. Liu[5,6,7] & Xiao Wang[1,2,4]

Although messenger RNA (mRNA) has proved effective as a vaccine, its potential as a general therapeutic modality is limited by its instability and low translation capacity. To increase the duration and level of protein expression from mRNA, we designed and synthesized topologically and chemically modified mRNAs with multiple synthetic poly(A) tails. Here we demonstrate that the optimized multitailed mRNA yielded ~4.7–19.5-fold higher luminescence signals than the control mRNA from 24 to 72 h post transfection in cellulo and 14 days detectable signal versus <7 days signal from the control in vivo. We further achieve efficient multiplexed genome editing of the clinically relevant genes *Pcsk9* and *Angptl3* in mouse liver at a minimal mRNA dosage. Taken together, these results provide a generalizable approach to synthesize capped branched mRNA with markedly enhanced translation capacity.

Messenger RNA (mRNA) therapeutics have recently gained prominence owing to their programmability, scalability and ease of design and synthesis[1–3], as demonstrated by the rapid development of mRNA vaccines for coronavirus disease 2019 (COVID-19)[4,5]. While mRNA vaccines have proven highly effective in generating a transient expression of antigens in small quantities, mRNA has potential as a general therapeutic approach for a wide range of diseases through larger-scale production of therapeutic proteins in vivo. For instance, mRNA has been tested in early clinical trials to produce paracrine vascular endothelial growth factor, a protein used to treat heart failure, and Cas9 protein to correct mutations that cause heterozygous familial hypercholesterolemia[6–8]. mRNA therapies for applications such as enzyme replacement, antibody therapy and gene editing require sustained high-level protein production, but traditional mRNA drugs suffer from RNA instability and low efficiency of expression, necessitating high doses that may lead to cytotoxicity[9–12]. To address these challenges, efforts to improve the translation duration and the overall protein production of mRNA vectors are necessary.

Numerous efforts have been made to expand the utility of mRNA drugs. Replacement of uridine (U) with N^1-methylpseudouridine ($m^1\psi$) decreases immune toxicity by blocking Toll-like receptor recognition[13–15]. Codon optimization and sequence manipulation in the untranslated regions (UTRs) have been shown to increase protein expression[16,17]. A strategy that promises to broaden the pharmacokinetic window of mRNA is the use of circular RNA (circRNA), which is resistant to exonuclease cleavage, the primary pathway by which linear mRNA is degraded[18,19]. However, circRNA technology faces several limitations in its current state. Translation initiation in circRNAs relies on internal ribosome entry sites (IRES), which are less efficient than the 7-methylguanylate (m^7G) cap-dependent translation initiation mechanism. Additionally, circRNAs limit the usage of chemical modifications, including $m^1\psi$, as they block RNA splicing required for circRNA synthesis and eliminate IRES-based translation initiation[20–22]. Similar limitations apply for self-amplifying RNA, where the replication-responsive conserved sequence elements are incompatible with $m^1\psi$ and the

[1]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. [2]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [3]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. [4]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [5]Merkin Institute of Transformative Technologies in Healthcare, Broad Institute of MIT and Harvard, Cambridge, MA, USA. [6]Howard Hughes Medical Institute, Harvard University, Cambridge, MA, USA. [7]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA, USA. e-mail: xwangx@mit.edu

generated replicons are uncapped and immunogenic[23]. Therefore, alternative strategies are needed to protect the mRNA transcript from rapid exonuclease-dependent decay, particularly the deadenylation of the 3′ poly(A) tail, while preserving the efficient cap-dependent translation initiation mechanism.

In the most widely accepted model of cap-dependent translation mechanism, the rate-limiting step is the assembly of the translation initiation complex centered around the m^7G cap by eukaryotic translation initiation factors (eIFs) including eIF4E and eIF4G[24]. The translation initiation complex is further stabilized through the interaction of eIF4G with cytoplasmic poly(A)-binding proteins (PABPCs), which recognizes the 3′ poly(A) tail, forming a multimeric complex and a closed-loop structure (Fig. 1a)[25,26]. PABPCs serve a dual role of initiating translation through engagement with eIFs and extending mRNA stability by cloaking the poly(A) tail to prevent deadenylation[27–30]. We have previously demonstrated that site-specific introduction of exonuclease-resistant modifications at the end of poly(A) tails efficiently increases mRNA stability and protein production in cellulo, possibly through stabilization of the nascent poly(A) tail and maintenance of poly(A)–PABPC1 interactions[31]. Structural studies have revealed that PABPC1 functions as a multimeric protein complex where individual PABPC1 molecules binding the poly(A) tail form a stabilized macrostructure[32,33]. In this Article, we reasoned that multimerization of the poly(A) tail via a branched topology, with each individual poly(A) tail bearing extensive nuclease-resistant modifications (Fig. 1b), would protect poly(A) tail integrity against RNA decay and preserve multimeric poly(A)-PABPC1 interactions for prolonged translation.

To achieve the aforementioned goals, we developed chemically modified, capped, branched mRNA-oligo conjugates bearing multimeric poly(A) tails, termed multitailed (multitail) mRNA. We systematically evaluated how different chemical and topological patterns on the poly(A) tails affected mRNA translation capacity and their mechanisms of action. We have successfully constructed branched mRNAs with prolonged protein expression for up to 14 days and improved clustered regularly interspaced short palindromic repeat (CRISPR) editing efficiency at minimal doses in cell lines and in mice.

Results

Design, synthesis and screening of multitail mRNA

We have developed the methodology for synthesizing the multitail mRNA based on our previously reported approach of ligating modified, chemically synthesized oligonucleotides to the 3′ end of in vitro transcription (IVT) mRNA using T4 RNA ligase I (ref. 31). We started by evaluating various synthetic strategies for generating the branched poly(A) oligonucleotides. While the naturally occurring topology of branched RNA occurs in the form of lariat intron RNA, a byproduct of pre-mRNA splicing with a 'branched' linkage between the 2′-hydroxyl group of an internal nucleotide and the 5′-end of the spliced intron[34], it is rapidly cleared by debranching enzymes[35]. Therefore, we focused on screening unnatural branching linkages that cannot be hydrolyzed inside cells and have fast reaction kinetics at the micromolar concentration for the ease of large-scale synthesis. We tested several types of oligonucleotide crosslinking chemistry, including thiol-ene/thiol-yne reactions[36], phosphate–amine coupling[37–39] and click reactions like tetrazine*trans*-cyclooctene (Tz-TCO)[40,41] and copper-catalyzed azide-alkyne cycloaddition (CuAAC)[42,43] (Extended Data Fig. 1a). Thiol-ene/thiol-yne reactions had low conversion under nonradical conditions and led to uncontrolled oligomerization when radical initiators were used (Extended Data Fig. 1b). The phosphate–amine reaction produced trace amounts of product even with catalysis by 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide/imidazole. The Tz-TCO and CuAAC reactions both worked well with high conversion. We selected CuAAC for our following demonstrations due to the ease of introducing alkyne handles via 5-octadiynyl deoxyuridine (OU) during solid-phase oligo synthesis, whereas Tz-TCO requires extra steps of functionalization and purification.

To construct poly(A) oligos with multiple branches, we designed a 30 nt 'stem' oligo containing a 5′-phosphate for subsequent enzymatic ligation and multiple internal OU modification branching sites spaced by 12 nt for PABPC1 binding[44,45], each of which was conjugated to a 30 nt 'branch' poly(A) oligo containing a 5′-azide handle. Both the stem and the branch poly(A) contained a chain-terminating dideoxycytidine at the end to prevent self-ligation and phosphorothioate (PS) modifications on the last six bases to block deadenylation. CuAAC conjugation resulted in a mixture of seven products with one to three branches comprising various topologies as shown on gel electrophoresis (Extended Data Fig. 1c). A preliminary test by ligating the crude branched products to a firefly luciferase (FLuc) reporter showed a 7.25 ± 0.04-fold increase in bioluminescence in comparison with the unligated control at 72 h ($P < 0.0001$, analysis of variance (ANOVA); Extended Data Fig. 1d). In comparison, reversing the direction of the branching poly(A)s completely abolished the increase of luminescence, indicating a normal 5′ to 3′ directionality is required on the branched tails.

We next established a synthesis and purification pipeline for branched RNAs by incorporating high-performance liquid chromatography (HPLC) purification. To systematically screen various chemical modifications, branched poly(A) oligos with one, two or three tails were individually separated from the precursors using HPLC (Extended Data Fig. 2a–c). We first asked whether the branched poly(A) oligos could serve as a functional poly(A) tail to drive mRNA translation or if they were simply blocking deadenylation. When branched poly(A) or nonpoly(A) sequences with the same nuclease-resistant modification patterns were ligated to reporter mRNA without an IVT-incorporated poly(A) tail, the branched poly(A) effectively drove protein production, while a nonpoly(A) sequence on the branch did not have a significant effect (Extended Data Fig. 2d–f). This observation suggests that the branching poly(A) oligos are functioning as additional poly(A) tails.

Encouraged by these preliminary results, we then systematically screened different chemical and topological patterns. Specifically, synthetic poly(A) oligos with different numbers of branches (zero to three) and different types of chemical modifications (PS, DNA and 2′-*O*-methoxyethyl (2MOE)) were chemically assembled, HPLC purified and enzymatically ligated onto a FLuc reporter mRNA with optimized 5′/3′-UTRs and template-encoded 100-A split poly(A) tails, as used in the Pfizer severe acute respiratory syndrome coronavirus 2 vaccine[46] (Fig. 1c). After performing a quality control step measuring the ligation efficiency using an RNase H assay (Extended Data Fig. 2g), the modified mRNAs were cotransfected into HeLa cells with a Renilla luciferase (RLuc) mRNA as an internal transfection control, where the effects of modified, branched poly(A) tails on protein production were quantified by the ratio of FLuc to RLuc bioluminescence. Among all constructs tested, construct 8, with three branches and synergistic incorporation of 2MOE and PS on both the stem and branched poly(A) oligos, had the highest enhancements, yielding 4.7 ± 1.5, 10.7 ± 2.3 and 19.5 ± 2.1-fold higher luminescence signals than the mock ligation (mock lig.) control of linear mRNA at 24, 48 and 72 h post transfection, respectively ($P = 0.0055$, ANOVA; Fig. 1d, construct 8).

Modified poly(A) tails increased mRNA stability

To gain mechanistic insights into how the branched poly(A) tails increased mRNA translation, we dissected the overall effects of protein translation enhancement into two factors, mRNA stability (half-life, $t_{1/2}$) and translation efficiency (TE). Given that the endosomal escape rate for lipid transfection was reported to be as low as 1.4% in HeLa cells[47], bulk methods such as quantitative polymerase chain reaction with reverse transcription (RT–qPCR) would mostly reflect the decay kinetics of endosome-trapped mRNA rather than active translating cytosolic mRNA. We thus performed in situ profiling to quantify RNA copy numbers at subcellular resolution, with STARmap detecting all transcripts of the target sequence[48,49] and RIBOmap detecting only the ribosome-bound fractions of mRNA copies[50] (Fig. 2a). Modified FLuc

https://doi.org/10.1038/s41587-024-02174-7

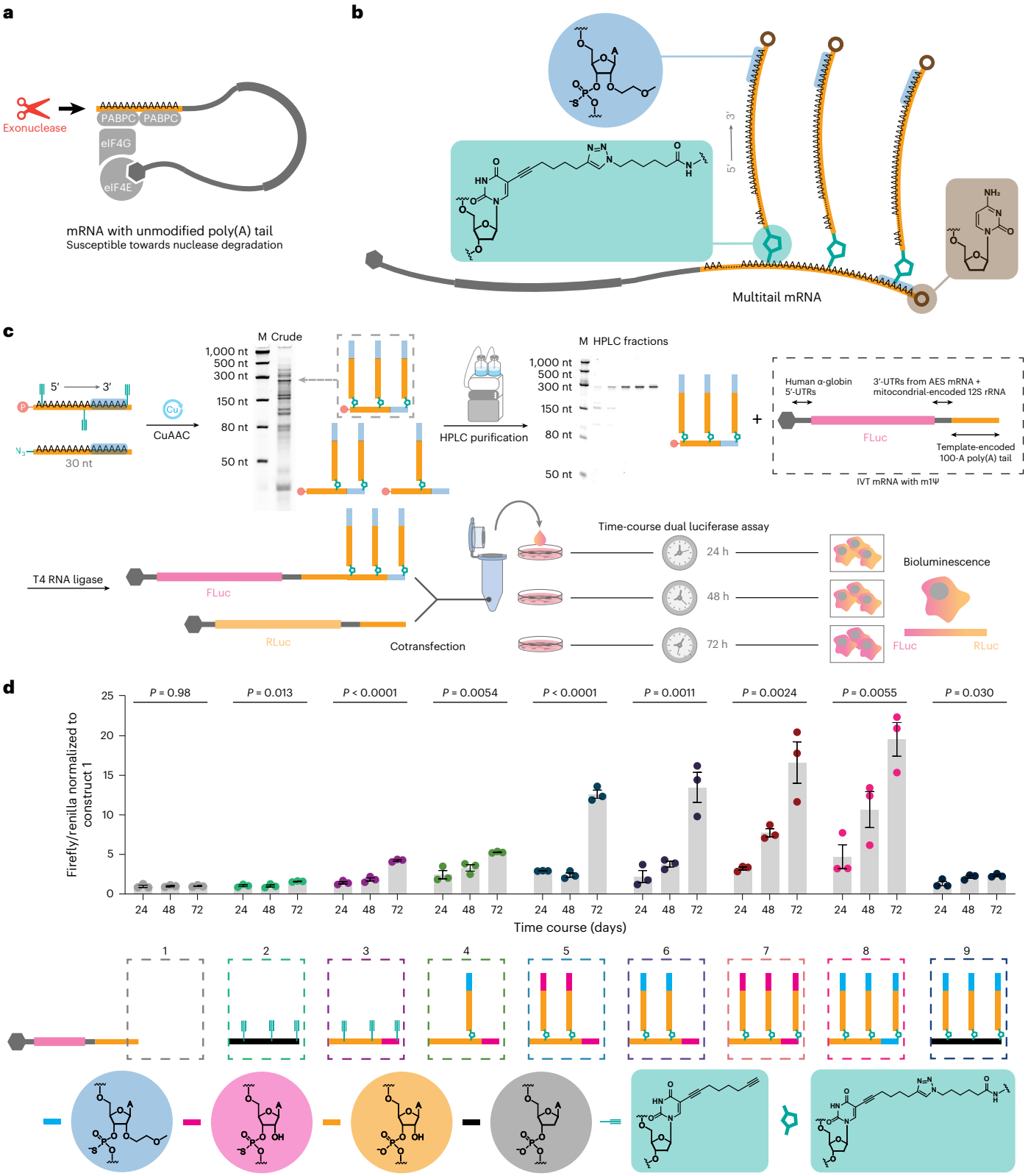


Fig. 1 | Conceptualization of chimeric mRNA bearing multiple poly(A) tails. **a**, Pseudo-closed-loop model of translation initiation. **b**, Chemical structures of engineered branched poly(A) tails. Branching oligos are conjugated through triazole linkages. Nuclease-resistant modifications (phosphorothioate and 2′-*O*-methoxyethyl) are incorporated on the last six nucleotides of each branch. Chain-terminating dideoxycytidines are introduced at 3′-end to prevent self-ligation. **c**, Synthesis and screening pipeline of branched mRNA. Alkyne and azide containing oligos were chemically conjugated, HPLC purified and enzymatically ligated to 3′-end of mRNA prepared from IVT with 100% (m1ψ) substitution to produce a branched mRNA-encoding FLuc reporter. 5′-UTRs were adopted from human α-globin 5′-UTR and 3′-UTR from the amino-terminal enhancer of split (AES) mRNA and the mitochondrial-encoded 12S rRNA as used in the Pfizer COVID-19 vaccine, followed by a template-encoded 100-A split poly(A) tail. Constructs were screened in a dual luciferase assay with RLuc as transfection control over a 72-h time course. M, markers. **d**, Barplots of time-course dual luciferase assay screening chemical modifications on multimerized poly(A) tails. Protein expression was measured by FLuc luminescence normalized to RLuc luminescence and the linear mRNA control at 24, 48 and 72 h post transfection. Mean ± standard error of the mean (s.e.m.). $n = 3$ independent transfections for each time point. The P values were calculated by an ordinary two-sided one-way ANOVA (comparison of means across time points).

https://doi.org/10.1038/s41587-024-02174-7

mRNAs were profiled with both STARmap and RIBOmap for TE quantification, whereas regular RLuc mRNAs were profiled by STARmap in both cases as internal control (Fig. 2a and Extended Data Fig. 3a). The combination of in situ profiling methods allowed filtration of large clusters of endosome-trapped nontranslating transcripts in lipid transfection vesicles (Fig. 2a, white granules), providing more accurate quantification of cytosolic mRNAs than traditional bulk northern blotting or RT–qPCR methods.

Both STARmap and RIBOmap results recapitulated the trends observed in our cellular luciferase assay, with cells transfected with multitail FLuc mRNA yielding observably higher numbers of cytosolic amplicons than their linear counterparts throughout the entire cell population at all time points (Extended Data Fig. 3a–c). The TE of FLuc mRNAs were approximated using the ratio of FLuc RIBOmap and STARmap signals of cells reseeded from the same biological conditions, with each normalized to the internal transfection-control RLuc STARmap signals (Fig. 2b). Interestingly, the difference in TE was insignificant regardless of chemical modification and topology at 24 h, indicating that multiple poly(A) branches did not hamper or enhance TE. However, at 48 h, both modification-only (mod-only) and multitail mRNA exhibited higher TE (by 1.53 ± 0.35 and 1.53 ± 0.38 fold, respectively) than the mock lig. control, and at 72 h, mod-only had similar (1.02 ± 0.27 fold) TE as the control, while multitail mRNA still exhibited 1.84 ± 0.44-fold higher TE, indicating that multitail mRNA preserved a more functional poly(A) tail at later time points to better sustain the cap/tail-dependent translation initiation[51].

To further characterize the kinetics of RNA translation at different time points, we synthesized mRNA transcripts encoding a degron-tagged FLuc (Firefly-PEST), which effectively reduced luciferase protein $t_{1/2}$ in HeLa cells from 20.4 h to around 0.92 h (ref. 52). Luminescence decay kinetics showed that chemical modification alone increased luminescence $t_{1/2}$ from an estimated 7.0 ± 0.2 h to 9.7 ± 0.4 h, and multitail mRNA further increased $t_{1/2}$ to 16.8 ± 1.2 h, whereas the decay kinetics of the internal RLuc control remained the same, confirming prolonged active translation for mRNAs with chemically modified and branched tails (Fig. 2c and Extended Data Fig. 3d).

Chemical and topological modifications preserve poly(A) tail

To confirm whether the functionality of branched poly(A) tails was due to their engagement with PABPC1[53], we conducted an electrophoretic mobility shift assay to quantify the dissociation constant (K_d) of fluorescently labeled linear versus branched poly(A) oligos with purified full-length recombinant PABPC1 protein[33,54] (Extended Data Fig. 4a,b). The K_d values of the first shift of linear and branched poly(A) were 10.7 and 6.4 nM, respectively, which agree with the previously reported K_d of PABPC1-poly(A) tail in the nanomolar range[33,54]. While the linear oligo exhibited a second shift at 21 nM, the multitail poly(A) oligo exhibited its second shift at 11 nM as well as additional third/fourth shifts at 15 and 35 nM, suggesting that each branching poly(A) tail also binds PABPC1 protein (Fig. 2d). Given that the cytoplasmic concentration of PABPC1 is ~1 μM (ref. 55), which is three orders of magnitude higher than the K_d, we propose that all three branched tails are saturated with PABPC1 inside cells as functional poly(A) motifs.

To gain further insight into how chemical and topological modifications increased poly(A) stability, we evaluated the decay kinetics of branched/linear tails using in vitro reconstituted conditions. PAN2–PAN3 and CAF1–CCR4 are the two major deadenylase complexes, which degrade long poly(A) tails (>60 nt) or poly(A) tails of all lengths, respectively[56,57]. Because our branched tails are 30 nt in length, which is below the reported activity window of PAN2–PAN3 (ref. 55), we tested the decay kinetics of four different fluorophore-labeled oligos against the recombinant CAF1–CCR4 complex: unmodified linear poly(A), modified linear poly(A), unmodified branch poly(A) and modified branch poly(A). The results showed that nuclease-resistant chemical modifications on either the linear stem oligo or the branched poly(A)s completely blocked the deadenylation, indicating that chemical modifications successfully protected the tail from CAF1–CCR4-mediated deadenylation (Extended Data Fig. 5a–d).

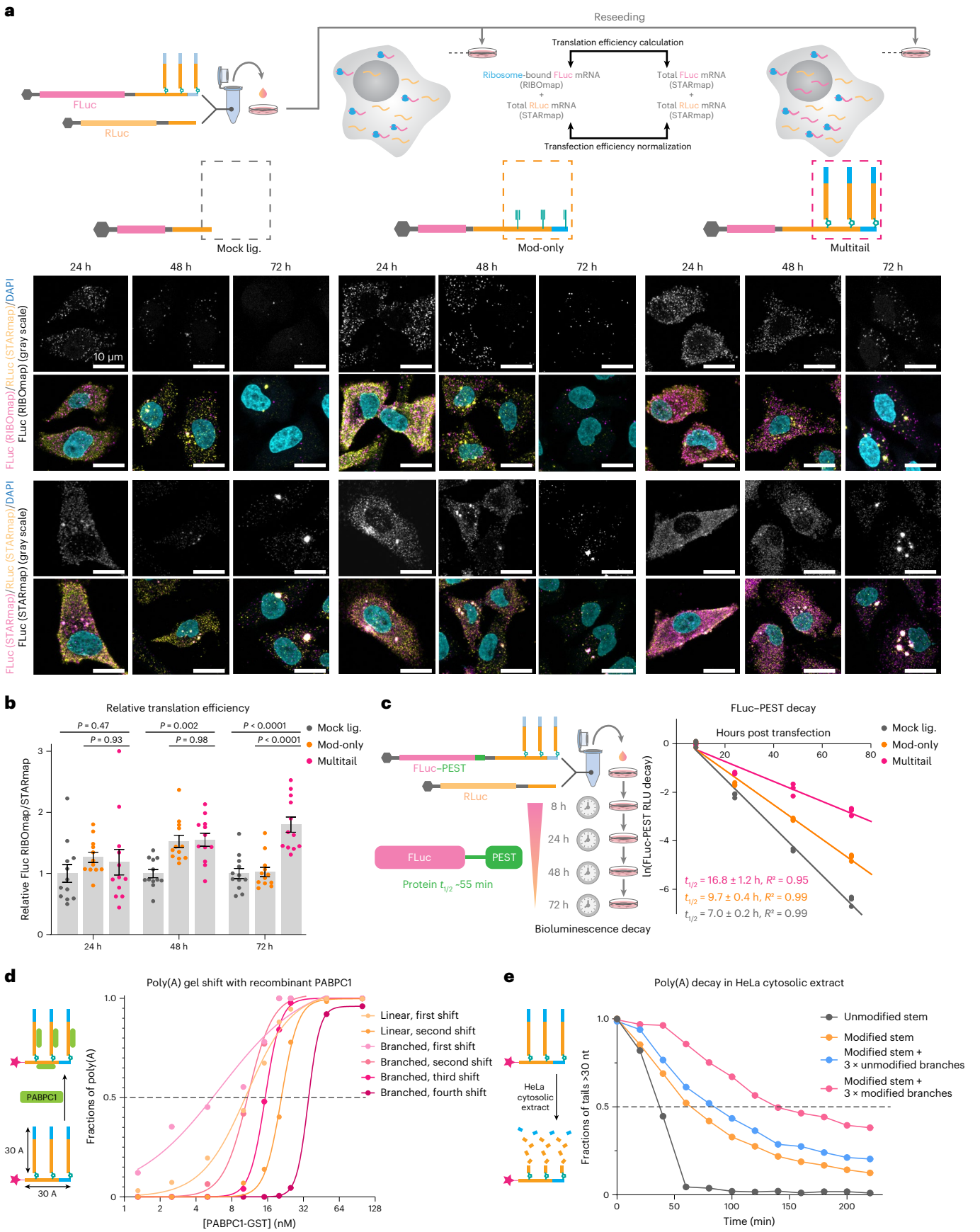
To further evaluate the effects of other RNA degradation pathways than CAF1–CCR4 deadenylation in the presence of PABPC1, the same set of four poly(A) oligos were treated with HeLa cell cytosolic lysate[58] to evaluate their decay rates (Extended Data Fig. 5e–h). We used 30 nt (at least one PABPC1 occupancy site as an estimation of a functional poly(A) tail) as a cut-off for quantifying the integrity of functional poly(A) tail (Fig. 2e). Unmodified linear poly(A) tails quickly decayed, exhibiting a $t_{1/2}$ of only 40 min, whereas adding chemical end modifications extended the $t_{1/2}$ to 60 min (Extended Data Fig. 5e,f). Unmodified poly(A) branches on the stabilized stem were also rapidly degraded at the initial stage, but remained as short tails, extending the $t_{1/2}$ to ~80 min. Finally, adding chemical stabilization to all branches further lengthened the $t_{1/2}$ to 130 min (Extended Data Fig. 5g,h). In conclusion, both chemical modifications and RNA topology contribute to poly(A) functionality and stabilization.

Multitail mRNA uses eIF4–eIF3 for translation initiation

Since canonical mRNA translation initiation depends on interactions between the poly(A)-PABPC1 complex at the 3′ end and the eIF4E–eIF4G complex at the 5′ end, we further tested whether the translation of multitail mRNAs is also dependent on eIF4. Interestingly, siRNA knockdown (KD) of eIF4E1 and eIF4G1 had insignificant impacts on the protein expression of transfected mRNA expression across all constructs (Extended Data Fig. 6a–c). We reasoned that this might be due to the existence of homologous genes like eIF4E2[59] and eIF4G2/eIF4G3 (refs. 60,61), which normally initiate translation for a small subset of genes but might functionally compensate for eIF4E1/eIF4G1 upon KD. Another explanation could be the activation of noncanonical translation initiation when eIF4E1/eIF4G1 levels are low. To rule out such effects, we performed an in vitro translation assay in the presence of

Fig. 2 | Mechanistic characterization of mRNA stabilization by branched poly(A) tails. a, Schematic illustration of pairwise in situ STARmap + RIBOmap characterization of transfected mRNA transcripts. Cells were reseeded into 2 wells per condition per time point 6 h after cotransfection of FLuc and RLuc mRNAs. RIBOmap and STARmap characterization of engineered FLuc transcripts in each duplicately measured ribosome-bound mRNA or total mRNA, respectively. Unmodified RLuc luciferase mRNA was profiled by STARmap in all replicates as the transfection control for normalization. Representative STARmap/RIBOmap images were acquired under the same confocal imaging settings (DAPI (blue), nuclei; FLuc amplicons (magenta); RLuc amplicons (yellow)). Colocalized FLuc/RLuc amplicons in STARmap (white dots) were lipid transfection vesicles and excluded from downstream quantification. Scale bar, 10 μm. **b**, Comparison of TE of mock lig./modification-only/multitail mRNA across different time points. For each single cell, total FLuc mRNA abundance was quantified by FLuc STARmap/RLuc STARmap, and ribosome-bound FLuc mRNA abundance was quantified by FLuc RIBOmap/RLuc STARmap. TE was calculated by the ratio of ribosome-bound FLuc/total FLuc and average values of each FOV were normalized against mock lig. FLuc mRNA. Mean ± s.e.m. $n = 12$, with three independent transfections, and four FOVs were taken from each quadrant for each condition. The P values were calculated by an ordinary two-sided one-way ANOVA. **c**, Kinetic characterization of Firefly-degron (PEST) reporter mRNAs. RLU decay was fitted with first-order kinetics $\ln y = -kx + b$ (least squares), where y is normalized FLuc expressions and x is hours post transfection, and $\ln y$ was fitted against time for k. $t_{1/2}$ was calculated using $\ln 2/k$. **d**, Quantification of gel-shifting of mod-only and multitail oligo at increasing concentrations of recombinant human PABPC1. **e**, Quantification of fractions of >30 nt poly(A) tails remained in HeLa cytosolic lysate digestion assay by fluorescence intensity integration and normalization to 0 min. The average values from two replicated experiments are plotted.

https://doi.org/10.1038/s41587-024-02174-7



4EGI-1, a pan-homolog inhibitor for eIF4E–eIF4G interactions after cap binding by eIF4E[62]. 4EGI-1 significantly decreased the translation of all constructs, indicating that multitail mRNA still undergoes translation initiation through the canonical eIF4E–eIF4G pathway (Extended Data Fig. 6d). Notably, the degree of translation reduction for multitail mRNA (50%) was lower than regular mRNA (85%). We speculate

that this difference is due to differential translatability of multitail mRNA under eIF4E–eIF4G-limiting conditions. Finally, KD of eIF3D, an essential component of eIF3 complex downstream of eIF4 for 40S recruitment during translation initiation, led to dramatic translation reduction (>90%) for all constructs to the same extent, suggesting the necessity of eIF3 in the translation of the multitail mRNA similar as regular mRNA (Extended Data Fig. 6e). Through this set of experiments, we concluded that multitail mRNA goes through the canonical eIF4–eIF3 initiation mechanism.

Prolonged protein production by multitail mRNA in cellulo and in vivo

To compare multitail mRNA with state-of-the-art circRNA technology[18,19], we synthesized linear, branched and circular mRNA vectors encoding a secreted nanoluciferase (NLuc), which allowed us to repeatedly sample media and measure protein expression over a longer time course. The circular mRNA vectors were designed to circularize through backsplicing and include an IRES from human rhinovirus (iHRV) where circularization was confirmed by slower mobility during gel electrophoresis (Extended Data Fig. 7a). Both linear and branched mRNA bearing the optimized UTRs and m1ψ outperformed the circRNA construct by more than three orders of magnitude, even at higher dosage of circRNA (Extended Data Fig. 7b), suggesting that cap-dependent translation initiation with optimized UTRs is still more efficient than IRES in circRNA and circRNA may be further improved by adding chemical modifications. While linear and branched mRNAs produced similar amounts in the first 24 h, the protein signal of unligated linear mRNA (mock lig.) quickly dropped after day 1, while the mod-only and multitail mRNA architectures exhibited dramatically slower decay, with observable protein signal at day 14 (Extended Data Fig. 7c).

Encouraged by the in cellulo results, we examined the multitail mRNA construct in mice. Synthetic mRNAs were encapsulated in lipid nanoparticles (LNPs) and administered through retro-orbital (RO) injection (Fig. 3a). A single dose of 40 μg kg^{-1} of multitail mRNA yielded a luminescence signal 3.3–5.0-fold higher than that obtained with mock lig. mRNA and 1.6–2.2-fold higher than that obtained with mod-only mRNA from 6 h to day 2, and significantly higher luminescence signals were detected at later time points with multitail mRNA expressed up to 7.5-fold higher than mod-only mRNA and 47-fold higher than mock lig. mRNA (Fig. 3b,c). Immunogenicity was evaluated by serum level of tumor necrosis factor (TNF) and toxicity was measured by aspartate transaminase (AST) and alanine transaminase (ALT), as intravenous delivery readily targeted the liver[63,64]. Levels of all biomarkers were indistinguishable among mice treated by multitail mRNA, mock lig. mRNA and poly(C) negative control and significantly lower than that of mice treated with poly (I:C), a known RNA immunosensitizer recognized by Toll-like receptor 3 that induces release of interferon-β (Fig. 3d–f)[65,66]. Taken together, our results indicate that the branched poly(A) topology and unnatural triazole chemical linkages do not result in long-term immunotoxicity in vivo while enabling high protein expression at low doses, justifying their suitability for use as therapeutic RNAs.

To directly assess the stability and translatability of modified mRNA in vivo, we also performed parallel STARmap/RIBOmap experiments using RO injection and profiling of adjacent mouse liver sections to examine mRNA stability and translatability following previously published procedure[50]. After benchmarking the linear dynamic range and reproducibility of the method using coinjection of regular capped, $m^1\Psi$-modified NLuc/FLuc mRNA across the five liver lobes (Extended Data Fig. 8), we subsequently profiled mRNA with additional chemical and topological modifications (Fig. 3g). Consistent with our observations in cell-based assays and by in vivo luminescence, both mod-only and multitail mRNA of 24 h post injection were significantly stabilized in comparison with mock lig. control as profiled by STARmap (Fig. 3h,i). Despite the similar mRNA level measured by STARmap, the multitail mRNA resulted in ~1.7-fold more signals in RIBOmap than the mod-only mRNA, which is consistent with the ~2.0-fold increase in NLuc protein quantification at this time point, indicating multiple chemically stabilized tails better preserved mRNA translatability for a longer duration of protein production in vivo (Fig. 3j,k).

Multitail mRNA for enhanced CRISPR-based therapeutics

Given the potency of multitail mRNA, we then sought to demonstrate its promise as a platform for functional genomics and therapeutic applications. Nonviral LNP-mediated delivery of mRNAs encoding CRISPR-based therapeutic constructs have been shown to have lower off-target editing and decreased risk of immunogenicity compared to delivery by viral DNA vectors, such as an adeno-associated virus[67–69]. However, this potential has been hampered by the requirement for high and repeated dosage of mRNA–LNP needed to observe significant genome editing[70,71]. We hypothesize that multitail mRNA could further enhance the efficacy of mRNA-based CRISPR–Cas9 systems, complementary to chemically modified and stabilized guide RNAs.

To test this hypothesis, we transfected a reporter HEK cell line expressing an unstable degron-tagged green fluorescent protein (uGFP) with branched/linear Cas9 mRNA and two single guide RNAs (sgRNAs) targeting the uGFP CDS for gene deletion[72] (Fig. 4a). Higher levels of Cas9 protein expression were observed for multitail mRNA compared with the mod-only mRNA and mock lig. mRNA controls (Extended Data Fig. 9a). While observable levels of uGFP ablation were observed for all constructs, we found that the multitail Cas9 mRNA construct led to the largest decrease (89.1 ± 0.9%) in uGFP levels across the whole transfected cell population (Fig. 4b and Extended Data Fig. 9b,c). The uGFP fluorescence also faithfully reflected genome editing events as quantified by next-generation sequencing (NGS), with the percentage of uGFP-positive cells correlating linearly and negatively (Pearson $R = 0.94$) with editing efficiency (Fig. 4c).

Encouraged by these preliminary in vitro results, we used multitail Cas9 mRNA coformulated with sgRNAs in LNPs for in vivo gene editing in murine liver. We postulate that the enhanced translatability and stability of multitail mRNA would enable effective expression of Cas9 protein at lowered mRNA doses. Chemically modified sgRNAs were designed to target *Pcsk9* and *Angptl3*, genes implicated in lipoprotein homeostasis and whose KD has shown promise as a therapeutic strategy for treatment of familial hypercholesterolemia (Fig. 4d)[67,70,71]. Similar to the results obtained in our cell-based tests, we detected 4.1 ± 0.7-fold higher levels of Cas9 protein expression for the multitail construct compared to mod-only mRNA (Fig. 4e). A single dose of 0.7 mg kg^{-1} Cas9 mRNA was sufficient to induce detectable editing levels, with our chemical and topological mRNA modifications yielding significantly higher editing efficiency, as quantified by NGS, at both *Pcsk9* (46%) and *Angptl3* (54%) compared to that obtained with mock lig. mRNA (14% and 6%, respectively; Fig. 4f,g). Phenotypic analyses revealed significantly lower serum levels of both targeted proteins for mice treated with multitail mRNA, with a 71 ± 2.5% decrease for Pcsk9 protein and 66 ± 3.7% for Angptl3 protein at week 4. By contrast, mice treated with mock lig. mRNA had only a 45 ± 3.0% reduction in Pcsk9 and a 27 ± 8.9% reduction in Angptl3 (Fig. 4h,i and Extended Data Fig. 9d,e). Collectively, these protein level reductions translated to downstream decrease of serum lipids, with the multitail group showing a 63 ± 3.5% decrease in serum free cholesterol, a 51 ± 1.7% decrease in serum total cholesterol and a 49 ± 3.4% decrease in TG, more efficient than the 26 ± 7.7% reduction in serum free cholesterol, 18 ± 4.6% reduction in total cholesterol and 9 ± 9% reduction in TGs in the mock lig. group at week 4 (Fig. 4j–l and Extended Data Fig. 9f–h).

Discussion

We developed topologically and chemically modified mRNAs with multimeric poly(A) tails, or multitail mRNAs, using a hybrid synthetic–organic and enzymatic approach. While current mRNA technologies encompass linear, circular and self-amplifying RNAs,

https://doi.org/10.1038/s41587-024-02174-7

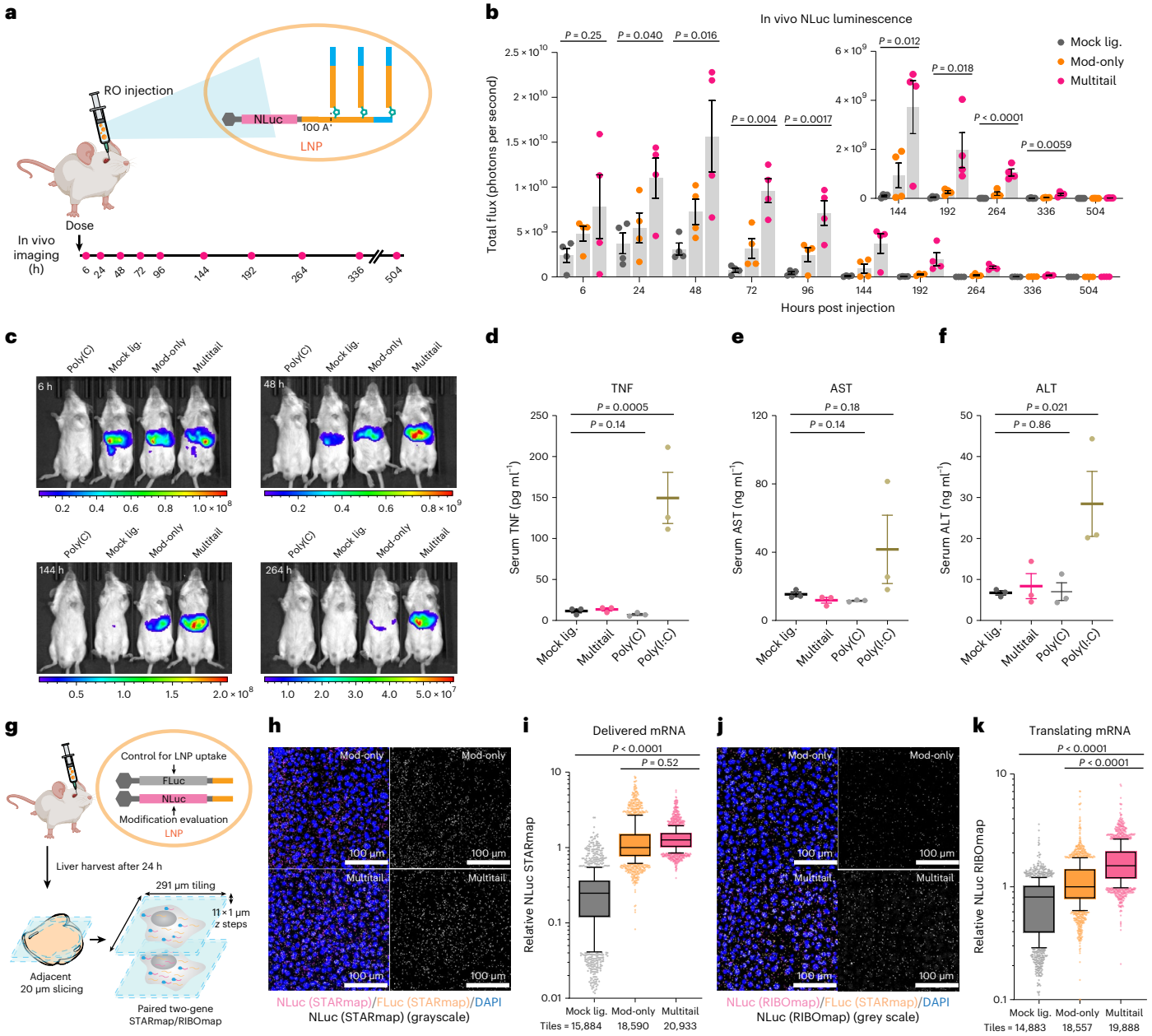


Fig. 3 | Multitail mRNA exhibited prolonged protein expression in vivo. **a**, Schematics of in vivo luminescence quantification over 3 weeks. Mock lig./mod-only/multitail NLuc mRNAs were formulated in LNPs and equal molar quantities of mRNA–LNP complexes were administered through RO injection to male adult BALB/cJ mice. **b**, In vivo luminescence signals from NLuc mRNAs. Luminescence was measured by integration of total flux for each mouse normalized by subtraction of background (poly(C)–LNP-treated mice). Mean ± s.e.m. $n = 4$ mice for each condition. **c**, Exemplary in vivo luminescence images of mice treated with poly(C)/mock lig./modification-only/multitail mRNA at 6, 48, 144 and 264 h post mRNA–LNP complex administration. Color scale of the heatmaps, radiance (photons s^{-1} cm^{-2} sr^{-1}). **d**–**f**, Serum level of immunogenicity and liver toxicity biomarkers. Serum concentrations of TNF (**d**), AST (**e**) and ALT (**f**) were quantified by ELISA. Mean ± s.e.m. $n = 3$ biological replicates. The *P* values were calculated by an ordinary two-sided one-way ANOVA. **g**, Schematics for evaluating mRNA stability and TE with two-gene STARmap/RIBOmap experiments. Different constructs of NLuc and transfection-control FLuc mRNA were coencapsulated in LNP and administered by RO injection. A total of 24 h post mRNA injection, the caudate lobes of mouse livers were collected, and each tissue was sectioned into two adjacent slices, with one slice profiled by NLuc STARmap and FLuc STARmap and the other slice profiled by NLuc RIBOmap and Fluc STARmap. Whole sections were profiled in 291 × 291 $μm^2$ tiles over a 10-μm *z*-stack with a 1 μm step size (11 × 1 μm *z* steps). **h**, Representative NLuc STARmap/FLuc STARmap image for mod-only and multitail mRNA, with NLuc STARmap shown in grayscale. **i**, Ratio of NLuc STARmap to FLuc STARmap in each tile, normalized to the median value in the mod-only construct (~10–90% box and whisker, line at median). $n = 15{,}884/18{,}590/20{,}933$ FOVs from three mice for each condition. The *P* values were calculated by an ordinary two-sided one-way ANOVA. **j**, Representative of NLuc RIBOmap/FLuc STARmap image for mod-only and multitail mRNA, with NLuc RIBOmap shown in grayscale. **k**, Ratio of NLuc RIBOmap to FLuc STARmap in each tile, normalized to the median value in the mod-only construct (~10–90% box and whisker, line at median). $n = 14{,}883/18{,}557/19{,}888$ FOVs from three mice for each condition. The *P* values were calculated by an ordinary two-sided one-way ANOVA.

we offer an alternative RNA topology that preserves cap-dependent translation initiation while increasing RNA stability to boost the overall translation capacity of single RNA molecules. Modified branched poly(A) tails can form multimeric interactions with PABPC1 as natural linear poly(A) tail, and both the chemical and topological modifications effectively protect poly(A) from degradation[73],

https://doi.org/10.1038/s41587-024-02174-7

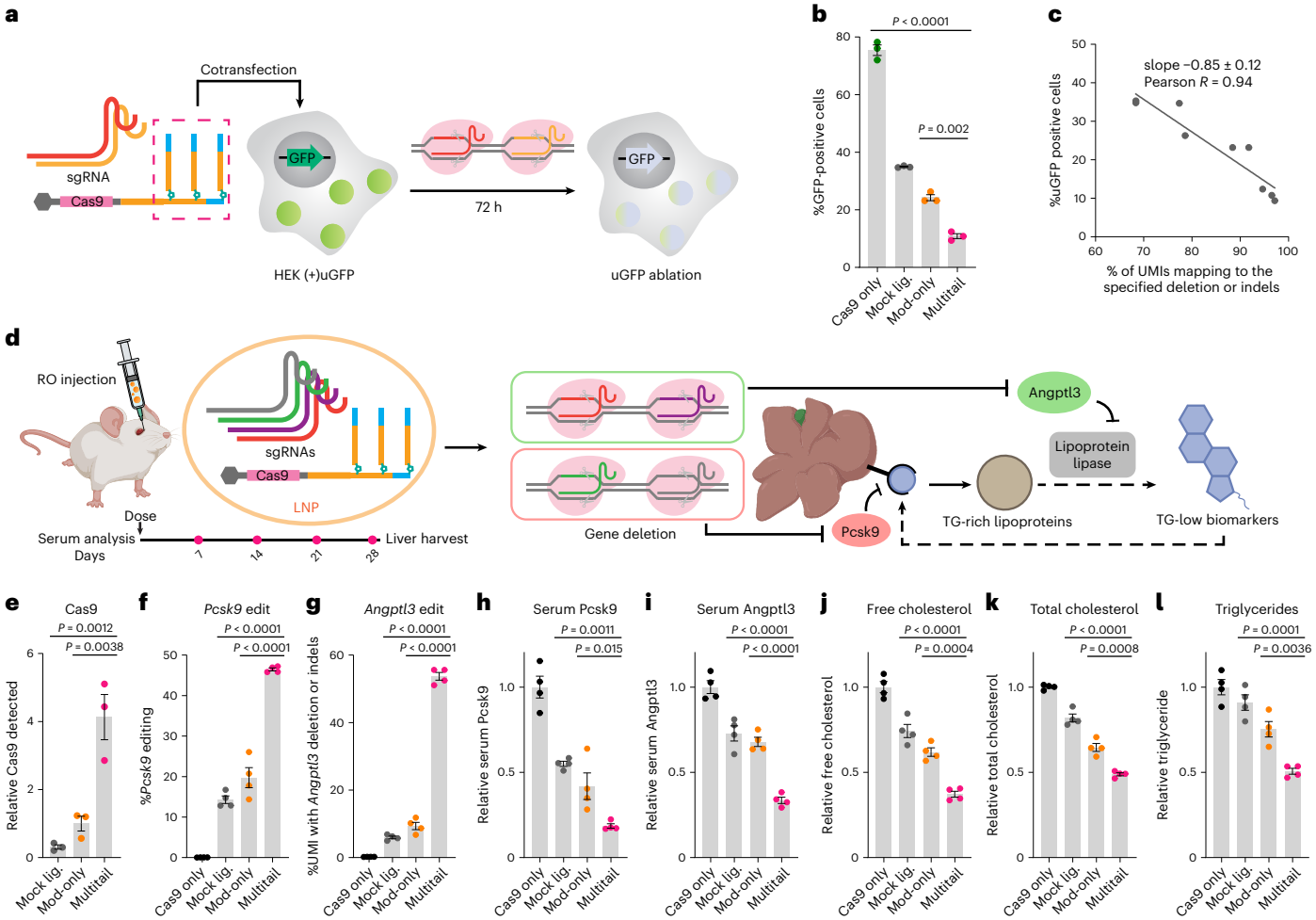


Fig. 4 | Multiplexed genome editing in vivo using stabilized Cas9 mRNA. **a**, Schematics of unstable d2EGFP (uGFP) ablation by two guide RNA induced deletion. HEK293T cells expressing uGFP were transfected with modified multitail Cas9 mRNA with two sgRNAs flanking the uGFP open reading frame. **b**, Comparison of uGFP ablation rates using mock lig., mod-only or multitail Cas9 mRNA. Mean ± s.e.m. $n = 3$ different transfections with each data point representing percent green fluorescent protein (%GFP)-positive cells from fluorescence-activated cell sorting as exemplified in Extended Data Fig. 10c. The P values were calculated by an ordinary two-sided one-way ANOVA. **c**, %uGFP-positive cell reduction correlated linearly with editing efficiency quantified by NGS. **d**, Schematics for in vivo editing of *Pcsk9* and *Angptl3* for homeostasis of triglyceride (TG)-rich lipoproteins. Two sgRNAs targeting each gene were designed to result in ~50 bp (*Pcsk9*) or ~270 bp (*Angptl3*) deletion and encapsulated with different Cas9 mRNA constructs in LNP. Serums were collected and analyzed weekly, and mouse livers were collected 4 weeks post injection. **e**, Characterization of Cas9 expression by ELISA from mouse liver lysate 24 h post injection. Values normalized to the mod-only construct. Mean ± s.e.m. $n = 3$ technical replicates. The P values were calculated by an ordinary two-sided one-way ANOVA. **f**,**g**, Editing efficiency of *Pcsk9* (**f**) and *Angptl3* (**g**) quantified by NGS. Mean ± s.e.m. $n = 4$ biological replicates. The P values were calculated by an ordinary two-sided one-way ANOVA. **h**–**l**, Relative serum Pcsk9 (**h**), Angptl3 (**i**), free cholesterol (**j**), total cholesterol (**k**) and triglycerides (**l**) at week 4 normalized to the Cas9-only (mock lig. Cas9 mRNA without sgRNA) negative control. Mean ± s.e.m. $n = 4$ biological replicates. The P values were calculated by an ordinary two-sided one-way ANOVA.

leading to prolonged protein expression in vivo without increasing immunogenicity and allowing delivery of functional proteins such as Cas9 at lowered mRNA dosage. From a clinical perspective, we envision that multitail mRNA holds promise for improving mRNA-based therapeutic platforms, such as gene therapy[70], multivalent vaccine formulation[74] and personalized T cell therapy[75], as well as numerous other biomedical applications.

The yields and scalability of current multitail mRNA synthesis is limited by the inefficiency of multisite click reactions at a submicromolar concentration, requirement for HPLC purification and superstoichiometric nature of the ligation reaction. The effects of chemical linkage for branching, the spacing of branched tails and patterns of chemical modifications were not exhaustively investigated due to aforementioned challenges and remained to be optimized. Alternatively, yields of branched tails may be increased by construction of the branching topology directly during solid-phase synthesis[76,77] and current efforts are ongoing on this end. Moreover, the observation of significant degradation of end-modified poly(A) tails in HeLa cytosolic extracts (Extended Data Fig. 5h) indicates active involvement of endonuclease in RNA decay. Therefore, introducing nuclease-resistant internal modifications without hampering translation may be beneficial to further block endonuclease-based RNA decay for RNAs of all topologies (linear, branched and circular).

From a broader perspective of RNA biochemistry, we provide a generalizable method to introduce defined but diverse modifications onto poly(A) tails and showed that these chemical and topological patterns were well-tolerated by mRNA translation machinery. While our study focused on mainly RNA stability, the ligation-based synthesis of branched RNAs allows for the modular assembly and combinatorial optimization of oligonucleotide fragments and mRNA constructs. Our results shed light on previously unexplored chemical space to design and develop next-generation mRNA therapeutics not only with

enhanced stability but also with improved translatability, specificity, delivery and pharmacokinetics.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-024-02174-7.

References

1. Sahin, U., Karikó, K. & Türeci, Ö. mRNA-based therapeutics—developing a new class of drugs. *Nat. Rev. Drug Discov.* **13**, 759–780 (2014).
2. Weng, Y. et al. The challenge and prospect of mRNA therapeutics landscape. *Biotechnol. Adv.* **40**, 107534 (2020).
3. Rohner, E., Yang, R., Foo, K. S., Goedel, A. & Chien, K. R. Unlocking the promise of mRNA therapeutics. *Nat. Biotechnol.* **40**, 1586–1600 (2022).
4. Baden, L. R. et al. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *N. Engl. J. Med.* **384**, 403–416 (2021).
5. Walsh, E. E. et al. Safety and immunogenicity of two RNA-based COVID-19 vaccine candidates. *N. Engl. J. Med.* **383**, 2439–2450 (2020)
6. Collén, A. et al. VEGFA mRNA for regenerative treatment of heart failure. *Nat. Rev. Drug Discov.* **21**, 79–80 (2022).
7. Mullard, A. mRNA-based drug approaches phase I milestone. *Nat. Rev. Drug Discov.* **15**, 595 (2016).
8. A study of VERVE-101 in patients with familial hypercholesterolemia and cardiovascular disease. *Clinicaltrials.gov* https://clinicaltrials.gov/ct2/show/NCT05398029?term=verve101&draw=2&rank=1 (2023).
9. Rybakova, Y. et al. mRNA delivery for therapeutic anti-HER2 antibody expression in vivo. *Mol. Ther.* **27**, 1415–1423 (2019).
10. Gillmore, J. D. et al. CRISPR–Cas9 in vivo gene editing for transthyretin amyloidosis. *N. Engl. J. Med.* **385**, 493–502 (2021).
11. Ramaswamy, S. et al. Systemic delivery of factor IX messenger RNA for protein replacement therapy. *Proc. Natl Acad. Sci. USA* **114**, E1941–E1950 (2017).
12. Jiang, L. et al. Dual mRNA therapy restores metabolic function in long-term studies in mice with propionic acidemia. *Nat. Commun.* **11**, 5339 (2020).
13. Karikó, K. et al. Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Mol. Ther.* **16**, 1833–1840 (2008).
14. Karikó, K., Buckstein, M., Ni, H. & Weissman, D. Suppression of RNA recognition by Toll-like receptors: the impact of nucleoside modification and the evolutionary origin of RNA. *Immunity* **23**, 165–175 (2005).
15. Kormann, M. S. D. et al. Expression of therapeutic proteins after delivery of chemically modified mRNA in mice. *Nat. Biotechnol.* **29**, 154–157 (2011).
16. Leppek, K. et al. Combinatorial optimization of mRNA structure, stability, and translation for RNA-based therapeutics. *Nat. Commun.* **13**, 1536 (2022).
17. Asrani, K. H. et al. Optimization of mRNA untranslated regions for improved expression of therapeutic mRNA. *RNA Biol.* **15**, 756–762 (2018).
18. Wesselhoeft, R. A., Kowalski, P. S. & Anderson, D. G. Engineering circular RNA for potent and stable translation in eukaryotic cells. *Nat. Commun.* **9**, 2629 (2018).
19. Chen, R. et al. Engineering circular RNA for enhanced protein production. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-022-01393-0 (2022).
20. Schlake, T., Thess, A., Thran, M. & Jordan, I. mRNA as novel technology for passive immunotherapy. *Cell. Mol. Life Sci.* **76**, 301–328 (2019).
21. Thess, A. et al. Sequence-engineered mRNA without chemical nucleoside modifications enables an effective protein therapy in large animals. *Mol. Ther.* **23**, 1456–1464 (2015).
22. Koch, A., Aguilera, L., Morisaki, T., Munsky, B. & Stasevich, T. J. Quantifying the dynamics of IRES and cap translation with single-molecule resolution in live cells. *Nat. Struct. Mol. Biol.* **27**, 1095–1104 (2020).
23. Bloom, K., van den Berg, F. & Arbuthnot, P. Self-amplifying RNA vaccines for infectious diseases. *Gene Ther.* **28**, 117–129 (2021).
24. Sonenberg, N. & Hinnebusch, A. G. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* **136**, 731–745 (2009).
25. Kahvejian, A., Roy, G. & Sonenberg, N. The mRNA closed-loop model: the function of PABP and PABP-interacting proteins in mRNA translation. *Cold Spring Harb. Symp. Quant. Biol.* **66**, 293–300 (2001).
26. Hinnebusch, A. G. The scanning mechanism of eukaryotic translation initiation. *Annu. Rev. Biochem.* **83**, 779–812 (2014).
27. Eisen, T. J. et al. The dynamics of cytoplasmic mRNA metabolism. *Mol. Cell* **77**, 786–799.e10 (2020).
28. Wang, Z., Day, N., Trifillis, P. & Kiledjian, M. An mRNA stability complex functions with poly(A)-binding protein to stabilize mRNA in vitro. *Mol. Cell. Biol.* **19**, 4552–4560 (1999).
29. Mangus, D. A., Evans, M. C. & Jacobson, A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* **4**, 223 (2003).
30. Bernstein, P., Peltz, S. W. & Ross, J. The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro. *Mol. Cell. Biol.* **9**, 659–670 (1989).
31. Aditham, A. et al. Chemically modified mocRNAs for highly efficient protein expression in mammalian cells. *ACS Chem. Biol.* **17**, 3352–3366 (2022).
32. Sawazaki, R. et al. Characterization of the multimeric structure of poly(A)-binding protein on a poly(A) tail. *Sci. Rep.* **8**, 1455 (2018).
33. Kühn, U. & Pieler, T. Xenopus poly(A) binding protein: functional domains in RNA binding and protein-protein interaction. *J. Mol. Biol.* **256**, 20–30 (1996).
34. Coombes, C. E. & Boeke, J. D. An evaluation of detection methods for large lariat RNAs. *RNA* **11**, 323–331 (2005).
35. Katolik, A. et al. Regiospecific solid-phase synthesis of branched oligoribonucleotides that mimic intronic lariat RNA intermediates. *J. Org. Chem.* **79**, 963–975 (2014).
36. Escorihuela, J. et al. Direct covalent attachment of DNA microarrays by rapid thiol-ene "click" chemistry. *Bioconjug. Chem.* **25**, 618–627 (2014).
37. Greenberg, M. M. Attachment of reporter and conjugate groups to the 3′ termini of oligonucleotides. *Curr. Protoc. Nucleic Acid Chem.* https://doi.org/10.1002/0471142700.nc0405s02 (2001).
38. El-Sagheer, A. H. & Brown, T. Single tube gene synthesis by phosphoramidate chemical ligation. *Chem. Commun.* **53**, 10700–10702 (2017).
39. Kalinowski, M. et al. Phosphoramidate ligation of oligonucleotides in nanoscale structures. *ChemBioChem* **17**, 1150–1155 (2016).
40. Ehret, F., Zhou, C. Y., Alexander, S. C., Zhang, D. & Devaraj, N. K. Site-specific covalent conjugation of modified mRNA by tRNA guanine transglycosylase. *Mol. Pharm.* **15**, 737–742 (2018).
41. Zhang, D. et al. Site-specific and enzymatic cross-linking of sgRNA enables wavelength-selectable photoactivated control of CRISPR gene editing. *J. Am. Chem. Soc.* **144**, 4487–4495 (2022).

42. Fantoni, N. Z., El-Sagheer, A. H. & Brown, T. A hitchhiker's guide to click-chemistry with nucleic acids. *Chem. Rev.* **121**, 7122–7154 (2021).
43. Warminski, M., Kowalska, J. & Jemielity, J. Solid-phase synthesis of RNA 5′-azides and their application for labeling, ligation, and cyclization via click chemistry. *Curr. Protoc. Nucleic Acid Chem.* **82**, e112 (2020).
44. Kühn, U. & Wahle, E. Structure and function of poly(A) binding proteins. *Biochim. Biophys. Acta* **1678**, 67–84 (2004).
45. Deo, R. C., Bonanno, J. B., Sonenberg, N. & Burley, S. K. Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* **98**, 835–845 (1999).
46. Vogel, A. B. et al. BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature* **592**, 283–289 (2021).
47. Gilleron, J. et al. Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. *Nat. Biotechnol.* **31**, 638–646 (2013).
48. Wang, X. et al. Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361**, eaat5691 (2018).
49. Zeng, H. et al. Integrative in situ mapping of single-cell transcriptional states and tissue histopathology in a mouse model of Alzheimer's disease. *Nat. Neurosci.* **26**, 430–446 (2023).
50. Zeng, H. et al. Spatially resolved single-cell translatomics at molecular resolution. *Science* **380**, eadd3067 (2023).
51. Xiang, K. & Bartel, D. P. The molecular basis of coupling between poly(A)-tail length and translational efficiency. *eLife* **10**, e66493 (2021).
52. Li, X. et al. Generation of destabilized green fluorescent protein as a transcription reporter. *J. Biol. Chem.* **273**, 34970–34975 (1998).
53. Nicholson-Shaw, A. L., Kofman, E. R., Yeo, G. W. & Pasquinelli, A. E. Nuclear and cytoplasmic poly(A) binding proteins (PABPs) favor distinct transcripts and isoforms. *Nucleic Acids Res.* **50**, 4685–4702 (2022).
54. Perzanowska, O., Smietanski, M., Jemielity, J. & Kowalska, J. Chemically modified poly(A) analogs targeting PABP: structure activity relationship and translation inhibitory properties. *Chemistry* **28**, e202201115 (2022).
55. Görlach, M., Burd, C. G. & Dreyfuss, G. The mRNA poly(A)-binding protein: localization, abundance, and RNA-binding specificity. *Exp. Cell. Res.* **211**, 400–407 (1994).
56. Schäfer, I. B. et al. Molecular basis for poly(A) RNP architecture and recognition by the Pan2–Pan3 deadenylase. *Cell* **177**, 1619–1631.e21 (2019).
57. Webster, M. W. et al. mRNA deadenylation is coupled to translation rates by the differential activities of Ccr4–Not nucleases. *Mol. Cell* **70**, 1089–1100.e8 (2018).
58. Dehlin, E., Wormington, M., Körner, C. G. & Wahle, E. Cap-dependent deadenylation of mRNA. *EMBO J.* **19**, 1079–1086 (2000).
59. Ruud, K. A., Kuhlow, C., Goss, D. J. & Browning, K. S. Identification and characterization of a novel cap-binding protein from *Arabidopsis thaliana*. *J. Biol. Chem.* **273**, 10325–10330 (1998).
60. Shestakova, E. D., Smirnova, V. V., Shatsky, I. N. & Terenin, I. M. Specific mechanisms of translation initiation in higher eukaryotes: the eIF4G2 story. *RNA* **29**, 282–299 (2023).
61. Ho, J. J. D. et al. Systemic reprogramming of translation efficiencies on oxygen stimulus. *Cell Rep.* **14**, 1293–1300 (2016).
62. Moerke, N. J. et al. Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell* **128**, 257–267 (2007).
63. Jang, D.-I. et al. The role of tumor necrosis factor alpha (TNF-α) in autoimmune disease and current TNF-α inhibitors in therapeutics. *Int. J. Mol. Sci.* **22**, 2719 (2021).
64. Giannini, E. G., Testa, R. & Savarino, V. Liver enzyme alteration: a guide for clinicians. *CMAJ* **172**, 367–379 (2005).
65. Kumar, A., Zhang, J. & Yu, F.-S. X. Toll-like receptor 3 agonist poly(I:C)-induced antiviral response in human corneal epithelial cells. *Immunology* **117**, 11–21 (2006).
66. Okahira, S. et al. Interferon-beta induction through Toll-like receptor 3 depends on double-stranded RNA structure. *DNA Cell Biol.* **24**, 614–623 (2005).
67. Rothgangl, T. et al. In vivo adenine base editing of PCSK9 in macaques reduces LDL cholesterol levels. *Nat. Biotechnol.* **39**, 949–957 (2021).
68. Robson, A. Three different therapies to target PCSK9. *Nat. Rev. Cardiol.* **18**, 541 (2021).
69. Nelson, C. E. et al. Long-term evaluation of AAV–CRISPR genome editing for Duchenne muscular dystrophy. *Nat. Med.* **25**, 427–432 (2019).
70. Musunuru, K. et al. In vivo CRISPR base editing of PCSK9 durably lowers cholesterol in primates. *Nature* **593**, 429–434 (2021).
71. Qiu, M. et al. Lipid nanoparticle-mediated codelivery of Cas9 mRNA and single-guide RNA achieves liver-specific in vivo genome editing of *Angptl3*. *Proc. Natl Acad. Sci. USA* **118**, e2020401118 (2021).
72. Zhou, J. et al. Dual sgRNAs facilitate CRISPR/Cas9-mediated mouse genome targeting. *FEBS J.* **281**, 1717–1725 (2014).
73. Park, J. et al. Short poly(A) tails are protected from deadenylation by the LARP1–PABP complex. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-023-00930-y (2023).
74. Arevalo, C. P. et al. A multivalent nucleoside-modified mRNA vaccine against all known influenza virus subtypes. *Science* **378**, 899–904 (2022).
75. Foy, S. P. et al. Non-viral precision T cell receptor replacement for personalized cell therapy. *Nature* **615**, 687–69 (2023).
76. Dong, Y. et al. DNA functional materials assembled from branched DNA: design, synthesis, and applications. *Chem. Rev.* **120**, 9420–9481 (2020).
77. Horn, T., Chang, C. A. & Urdea, M. S. Chemical synthesis and characterization of branched oligodeoxyribonucleotides (bDNA) for use as signal amplifiers in nucleic acid quantification assays. *Nucleic Acids Res.* **25**, 4842–4849 (1997).

Methods

Plasmid cloning, characterization and purification

The mRNA expression vectors were generated as described before[30]. Briefly, the CDS of FLuc, Nano luciferase and Cas9 were inserted into an optimized backbone containing (in order) an SP6/T7 promoter sequence, a 5′ human alpha globin UTR, a CDS, a 3′ human alpha globin UTR, a 100 × Å template-encoded poly(A) tail and an Esp3I linearization site. The CDS-containing plasmid/gene blocks were PCR amplified, gel-purified and assembled into the optimized backbone using NEBuilder HiFi DNA Assembly Master Mix (NEB, E2621S), transformed into stable cells and sequence verified with Sanger sequencing. The FLuc construct was obtained from pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega, E1330). The Cas9 construct was obtained from Addgene (plasmid #52962) deposited by Dr. Feng Zhang's laboratory. The nano luciferase constructs were obtained by gene synthesis from Genewiz. FLuc-PEST was generated previously by our laboratory (plasmid #200396). RLuc constructs were obtained from pmirGLO without cloning into the optimized vector.

mRNA synthesis and characterization

The DNA plasmids were linearized by Esp3I (NEB, R0734S). The linearized plasmids are purified with the DNA Clean and Concentrator-25 kit from Zymo Research (D4033) and characterized with agarose gel electrophoresis. mRNA constructs were synthesized by IVT using the mMESSAGE mMACHINE SP6 Transcription Kit (ThermoFisher Scientific, AM1340, for SP6 promoter constructs) or the HiScribe T7 High Yield RNA Synthesis Kit (NEB, E2040S; for T7 promoter constructs) per the manufacturer's protocol except with 100% replacement of UTP with N^1-methylpseudouridine-5′-triphosphate (Trilink, N-1081-1) and addition of 1:50 SUPERase-In RNase inhibitor (ThermoFisher Scientific, AM2694). All linear mRNA constructs were cotranscriptionally capped by replacing GTP to a final concentration of 2 mM GTP and 8 mM 3′-*O*-Me-m^{7}G(5′)ppp(5′)G RNA antireversed cap analog (NEB, S1411S) for the SP6 promoter constructs. The T7 constructs were cotranscriptionally capped with CleanCap Reagent AG (Trilink, N-7113). The DNA templates were digested by TURBO DNase and purified using the MEGAclear Transcription Clean-Up Kit (ThermoFisher Scientific, AM1908). The mRNA concentrations were quantified using the Qubit RNA HS Assay (ThermoFisher Scientific, Q32852).

Chemical conjugation of modified synthetic poly(A) oligonucleotides

Chemically modified poly(A) oligonucleotides were synthesized from IDT. Azide/alkyne-containing oligonucleotides were mixed with indicated molar ratio (1:1 for the one-branched oligo, 2:1 for the dibranched oligo and 3:1 for the tribranched oligo). The oligonucleotide mix was diluted in a modified 1.5× click chemistry buffer (Lumiprobe, 61150, with 5% RNasin, 5% dimethyl sulfoxide and 5% 10 mM dNTP mix) that was degassed by argon. 1:25 of 100 mM L-ascorbic acid solution (Sigma-Aldrich, A5960) was added immediately before the reaction. The mixture was incubated at 37 °C for 1 h, and the reaction was stopped by addition of ethylenediaminetetraacetic acid (EDTA). The crude products were purified using RNase-free HPLC on an Agilent 1260 Infinity II HPLC with acetonitrile/hexylamine/acetic acid (pH 7.0, with 10% urea w/v) mobile phase and PLRP-S stationary phase. The HPLC fractions were analyzed by gel, stained by 1× SYBR Gold (ThermoFisher Scientific, S11494) and visualized using a Bio-Rad ChemiDoc MP Imaging System (12003154). The desired fractions were then pooled, desalted and concentrated.

Enzymatic ligation of modified oligonucleotides to mRNAs

The ligation reactions were performed as previously described[30]. The products were purified with HPLC, and residual modified oligos were recovered. Ligation products were characterized by RNase H assays as described before[30]. For each batch of mRNA ligation, a 'mock ligation' condition was included where all reaction set-ups were the same, except the addition of the synthetic oligo, and the recovered product was used as the control.

Cell culture and time-course luciferase assays

HeLa cells (CCL-2, ATCC) were maintained in Dulbecco's modified Eagle medium culture media (ThermoFisher Scientific, 119951) containing 10% fetal bovine serum and 1% penicillin–streptomycin (ThermoFisher Scientific, 15070063) in a 37 °C incubator with 5% CO_2 and passaged at a ratio of 1:10 every 3 days. FluoroBrite Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum and 4 mM L-glutamine was used for culturing for the secreted NanoLuc assay. mRNA were transfected with Lipofectamine MessengerMAX Transfection Reagent (ThermoFisher Scientific, LMRNA003) at equal molar along with an internal control mRNA per the manufacturer's protocols for 6 h. For Firefly/Renilla assays, cells were lysed, and luciferase activity was measured using the Promega Dual Glo Luciferase Assay System (Promega, E2920) at indicated time points and normalized to seeding ratios. For secreted NanoLuc, every day the old media were all replaced with a new aliquot of 150 µl FluoroBrite media for each well, and 5 µl of the old media was used for bioluminescence measurement using the Nano-Glo Luciferase Assay System (Promega, N1110). In case of constructs having significant difference in expression, secreted NanoLuc from different conditions was measured on separate plates to avoid crosstalk.

mRNA quantification in transfected cell culture using STARmap/RIBOmap

Transfected HeLa cells were reseeded into six (two for each time point) glass-bottom 96-well plates (MatTek, PBK96G-1.5-5-F, poly-D-lysine (Sigma-Aldrich, A-003-M) coated) at a ratio of 6:4:3, respectively, for in situ sequencing quantification at 24, 48 and 72 h after transfection. Each well of transfected cells were reseeded equally onto two 96-well plates for each time point with one measured by STARmap (Firefly) + STARmap (Renilla) and the other measured by RIBOmap (Firefly) + STARmap (Renilla).

For in situ mRNA profiling, we followed the STARmap/RIBOmap procedure for cell cultures as published[48,50] (see Supplementary Table 1 for probe sequences). Confocal imaging stacks were taken with a Leica Stellaris 8 using Leica LAS-X software with a 60× oil objective at a pixel size of 283 × 283 nm^2. A 14-µm stack was imaged with 2 µm per step times seven steps. Excitation/detection wavelengths were as follows: 4′,6-diamidino-2-phenylindole (DAPI), diode 405 nm/~(420–489) nm; Alexa546, white light laser (WLL) 557 nm/~(569–612) nm; and Alexa647, WLL 653 nm/~(668–738) nm.

MATLAB 2021a and CellProfiler 4.0.7 were used for the amplicon count-based fluorescence image analysis. First, the centroids of amplicons in each fluorescent channel were identified by finding extended maxima on images. Then a 3 × 3 × 3 voxel volume centering the centroid of each fluorescent dot was defined. Within each voxel volume, the integrated intensities in the Firefly and Renilla channels were calculated, and the ratio between Firefly intensity and Renilla intensity was used for amplicon classification. For STARmap quantification, all the measurements were pooled together and the distribution of log(Firefly/Renilla) values were plotted. The corresponding ratio values at the nadirs (local minimum) on the distribution plot were identified as cut-off values. The first cut-off value smaller than 0 was noted as cut-off1, and the first cut-off value greater than 0 was noted as cut-off2. For RIBOmap quantification, cut-off1 was assigned as −0.3 and cut-off2 was assigned as 0.3. Any amplicon with a log(Firefly/Renilla) value smaller than cut-off1 was identified as a Renilla amplicon. Any amplicon with a log(Firefly/Renilla) value larger than cut-off2 was identified as a Firefly amplicon. Any amplicon with a log(Firefly/Renilla) value between cut-off1 and cut-off2 was identified as a granule. Amplicon classification information and locations of every amplicon were recorded. In each field of

view (FOV), the ratio between the number of Firefly amplicons and the number of Renilla amplicons were calculated and used to reflect the amount of FLuc mRNAs. In single-cell STARmap quantification, cell segmentation was performed by DAPI-stained nuclei to be identified as primary objects. The DAPI and AF647 channels were merged and subsequently converted to a grayscale image. Cells were identified as secondary objects in this grayscale image for segmentation. The segmentation masks were saved as uint16 images. The amplicons were assigned to cells according to their locations on the masks.

Gel shift assay

A codon-optimized DNA sequence encoding full-length human PABPC1 (NP_002559.2; GST tagged) was cloned into pBluescript II KS(−) vector from GenScript. Transformed *Escherichia coli* cells (New England Biolabs, BL21 Competent *E. coli*) were grown in 1 liter Luria Bertani medium with 50 mg ampicillin to an absorbance at 600 nm of 0.6–0.8 at 37 °C. The expressions of GST-fusion proteins were induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside, and cells were grown overnight at 18 °C. The cells were collected by centrifuge and disrupted by sonication in lysis buffer (10 mM Na_2HPO_4 (pH 7.4), 1.8 mM KH_2PO_4, 637 mM NaCl, 2.7 mM KCl, 10% glycerol, 1 mM phenylmethylsulfonyl fluoride, 0.5 mM 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride, 0.15 μM aprotinin, 1 μM E-64 and 1 μM leupeptin) with 0.15% polyethylenimine for removal of nucleic acids in lysate. Cell debris was removed by centrifuge. The supernatants were applied to Glutathione Sepharose 4B column (GE Healthcare) equilibrated with PBS at 4 °C and washed using PBS at 4 °C. The GST-fusion proteins were eluted with the elution buffer (100 mM Tris-HCl (pH 7.8), 300 mM NaCl and 20 mM glutathione (reduced form)). Eluted solution of isolated PABPC1-GST was concentrated using spin filters and quantified using Qubit Protein Broad Range Assay Kits (Invitrogen, Q33211). Branched/linear poly(A) oligonucleotides labeled at 5′-end with Alexa Fluor 546 were incubated in binding buffer (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl, 5 mM $MgSO_4$ and 1 mM dithiothreitol (DTT)) with varying PABPC1-GST concentrations prepared by sequential dilutions from the stock. The mixtures were incubated at 4 °C and characterized by electrophoretic mobility shift assay with Novex TBE Gels, 4–20% (Invitrogen, EC6225BOX) at 4 °C. Alexa546-labeled oligos were imaged with Bio-Rad ChemiDoc MP Imaging System (12003154), and band intensities were quantified using ImageJ. For K_d calculation, the percentage of poly(A) oligos bonded to *x*-mers of PABP were calculated as the total percentages of oligos bonded to >*x*-mers of PABP, and K_d was calculated by plotting percentage of PABP-bonded poly(A) oligos against PABP concentration using GraphPad Prism (specific binding with Hill slope).

In vitro poly(A) digestion

Codon-optimized DNA sequences encoding full-length CCR4 (NP_001273719)/CAF1 (NP_001309021) were cloned into the pET-15b vector from GenScript. Plasmids encoding CAF1 and CCR4 were coexpressed in *E. coli* BL21(DE3) cells (New England Biolabs) in Luria Bertani medium at 37 °C as fusion proteins carrying N-terminal His_6 tags. When the optical density value reaches 0.7, isopropyl β-d-1-thiogalactopyranoside (IPTG) was added to the media to a final concentration of 1 mM. The cells were lysed in a buffer containing 50 mM HEPES pH 7.5, 300 mM NaCl and 25 mM imidazole supplemented with complete EDTA-free protease inhibitors, 5 μg ml^{-1} DNaseI, 1 mg ml^{-1} lysozyme and 2 mM DTT. The proteins were isolated from the cleared lysate by binding to Ni-NTA resin (Qiagen), washed with a buffer containing 40 mM HEPES pH 7.5, 200 mM NaCl and 25 mM imidazole supplemented with 0.5 mM PMSF and 2 mM DTT, and eluted from the column by a buffer containing 40 mM HEPES pH 7.5, 300 mM NaCl and 500 mM imidazole supplemented with 0.5 mM PMSF and 2 mM DTT. The eluted proteins were dialyzed against a buffer containing 10 mM HEPES pH 7.5, 200 mM NaCl supplemented with 0.5 mM PMSF and 2 mM DTT. Linear/branched oligos labeled with 5′ Alexa546/Alexa647 were digested by HeLa cell cytosolic extract or 100 nM CCR4/CAF1 heterodimer in 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT and 2 U μl^{-1} RNasin at 37 °C. At each 20 min interval, 5 μl of the crude reactions were taken and stopped by addition of an equal volume of loading buffer containing 95% (w/v) formamide, 0.01% (w/v) bromophenol blue, 10% (w/v) urea and 10 mM EDTA, then denatured at 70 °C for 5 min. The RNAs were resolved on 15% polyacrylamide denaturing gel and directly visualized with 5′ Alexa546/Alexa647.

uGFP ablation

uGFP-expressing HEK293 cells (GenTarget, SC058) were transfected with Cas9 mRNA and sgRNAs (5′-CCGTCCAGCTCGACCAGGAT and 5′-CAAGACCCGCCACAACATCG), resuspended in PBS and analyzed with a Beckman CytoFLEX LX Flow Cytometer with 100,000-event records for each sample.

CircRNA synthesis

The sequences of circRNA encoding secretive NanoLuc with translation-enhancing UTRs (iHRVB3 + PABPv3) were adopted from previous studies[19]. The DNA templates were synthesized from Genewiz, PCR amplified and gel purified to be used as IVT templates. The circRNA were synthesized as described in literature by using the HiScribe T7 High Yield RNA Synthesis Kit (New England Biolabs, E2040S). Post IVT, DNA templates were digested with Turbo DNase (ThermoFisher, AM2238). The reaction mixture was heated to 70 °C for 5 min and then immediately placed on ice for 3 min, after which GTP was added to a final concentration of 2 mM, and the reaction mixture was incubated at 55 °C for 15 min. The circRNA was enriched by treatment with RNase R (Lucigen Corporation, RNR07250) for 1.5 h. The circRNA products were characterized by gel electrophoresis, and in cases of remaining linear RNA precursors, another round of RNase R digestion was performed.

Mice

All animal procedures followed animal care guidelines approved by the Institutional Animal Care and Use Committee of the Broad Institute of the Massachusetts Institute of Technology (MIT) and Harvard under animal protocol no. 0255-08-19. The animal experiments were conducted in compliance with Institutional Animal Care and Use Committee policies and the National Institutes of Health (NIH) guidelines. The BALB/cJ (male, ~6–8 weeks old) mice used for in vivo NanoLuc assay and toxicity assay in this study were purchased from The Jackson Laboratory. The C57BL/6 (male, 3–4 weeks old) mice used for Pcsk9 and Angptl3 editing were also purchased from The Jackson Laboratory. The mice were housed two to four per cage and kept on a 12 h light–dark cycle with ad libitum food and water at the temperature of 65–75 °F (~18–23 °C) with 40–60% humidity.

In vivo delivery and imaging of NanoLuc

NanoLuc mRNA in 50 mM citrate buffer (pH 4) and lipid mixture (37.5 mM in ethanol) containing 50 mol% SM-102, 10 mol% DSPC, 38.5 mol% cholesterol and 1.5 mol% DMG-PEG 2000 were assembled into lipid nanoparticle with NanoAssemblr Spark (Precision Nanosystems). mRNA–LNPs were then diluted in PBS and the buffer was exchanged by concentration using 30 kDa spin filters (MilliporeSigma, UFC901008) for ethanol removal. Concentrations and encapsulation efficiency of mRNA were determined using the Quant-it RiboGreen RNA Assay Kit (ThermoFisher, R11490). Equal molar of mock lig./modification-only/multitailed NanoLuc mRNA or poly(C) RNA was injected through retro-orbital injection. The in vivo activity of NanoLuc was measured using a Competent IVIS-Perkin Elmer IVIS Spectrum CT System. The fluorofurimazine substrate was freshly reconstituted in PBS. Mice were anesthetized and imaged using default settings. The luminescent activity was quantified using Aura 4.0 imaging software. The background was determined by including mice

injected with poly(C)–LNP negative control and subtracted from each measurement.

In vivo toxicity assay

Peripheral blood was collected and proceeded to serum from mice on day 15 post mRNA–LNP administration with poly(C)/poly (I:C)/NanoLuc mRNA formulated with NanoAssemblr Spark Kit–Hepato9 mRNA (Precision Nanosystems, NWS0016). AST, ALT and TNF were measured using enzyme-linked immunosorbent assay (ELISA) kits for AST (G-Biosciences, IT5530), ALT (G-Biosciences, IT5508) and TNF (R&D Systems, MTA00B) per the manufacturer's protocols.

Western blot

The total protein was extracted from cells or homogenized tissue in RIPA buffer with protease inhibitor and quantified with Qubit BR protein assay. Equal amounts of total protein for each condition were separated by 4–20% SDS–polyacrylamide gel electrophoresis (PAGE) gel (Bio-Rad) and transferred to Immun-Blot low fluorescence polyvinylidene difluoride membrane (Bio-Rad). The membranes were blocked and incubated with corresponding primary antibodies (anti-ACTB-HRP (Biolegned, 664803); anti-Cas9 antibody (Cell Signaling, 14697), eIF3D–rabbit IgG (ProteinTech, 10219-1-Ab); anti-eIF4E–rabbit IgG (MBL, RN00lP); anti-eIF4GI–rabbit IgG (MBL, RN002P)) at 4 °C overnight, followed by secondary antibodies (anti-mouse HRP (Santa Cruz Bio, sc-516102), antirabbit HRP (Sigma-Aldrich, 12-348)) at room temperature for 1 h.

mRNA quantification in mouse liver with STARmap/RIBOmap

FLuc and NLuc mRNA were coencapsulated in LNP and injected through RO injection. At 24 h post injection, mice were anesthetized with isoflurane and rapidly decapitated. Corresponding liver lobes were collected, placed in Tissue-Tek O.C.T. Compound, and frozen in liquid nitrogen. Liver tissues in O.C.T. was transferred to a cryostat [Leica CM1950] and sliced into 20 µm sections at −20 °C. The slices were attached to glass-bottom 24-well plates pretreated with 3-(trimethoxysilyl)propyl methacrylate and poly-D-lysine. STARmap/RIBOmap was performed following the previous report[47,49] (see Supplementary Table 1 for probe information). The detection and imaging procedures were the same as cell culture experiments except using a 40× oil.

The raw images collected from tile scan microscopy are stitched using the FIJI stitching plugin (version 1.2), during which overlapping areas of adjacent tiles are recalculated. Amplicon spots are first identified from the *z*-stack three-dimensional volume images of individual tiles using scikit-image. Then duplicate spots within the overlapping areas are removed. Nucleus counts are calculated using Stardist (version 0.8.5) segmentation of the DAPI images. We used the pipeline implementation by Squidpy (version 1.3.0). Nuclear staining regions with an area of less than 500 square pixels were excluded from nucleus counts. Nucleus areas were calculated using scikit-image without excluding small patches.

In vivo editing of *Angptl3*/*Pcsk9*

sgRNA sequences targeting *Angptl3*/*Pcsk9* were adopted from published reports[66,70]. Mock lig./mod-only/multitail mRNAs encoding Cas9 protein coformulated with two pairs of synthetic sgRNAs targeting *Angptl3* (5′-TTAGTAATTGCATCCAGAGT and 5′-TACACTACAAGTTAAAAACG)[71] and *Pcsk9* (5′-CCCATACCTTGGAGCAACGG and 5′-GCTCGCCCTCCCGTCCCAGG)[67] into LNPs through RO injection. Mice administered with only mock lig. Cas9 mRNA (that is, no sgRNAs) were treated as the negative control. Blood withdrawal was performed at the same time during the day after a 12-h fast the night before. Circulating Angptl3 and Pcsk9 protein levels were quantified using ELISA kits for mouse Pcsk9 (R&D Systems, MPC900) and Angptl3 (R&D Systems, MANL30). The serum free cholesterol was measured using the Cholesterol/Cholesterol Ester-Glo Assay (Promega, J3190). Serum TGs were measured using the Triglyceride Colorimetric Assay Kit (Cayman Chemical, 10010303).

Quantification of gene editing efficiency

Mouse livers were collected and genomic DNA (gDNA) was extracted using DNeasy Blood and Tissue Kits (Qiagen, 69504). For *uGFP* and *Angptl3*, DNA flanking the targeted regions was tagged with unique molecular identifiers (UMIs) and deduplicated to reduce the effects of PCR bias following a previously published workflow[78]. Briefly, gDNA was first linearly amplified with a primer containing a 15 nt randomized UMI and the P5 Illumina adapter sequence, then cleaned up with Ampure XP beads (Beckman Coulter, A63882). The UMI-tagged linear amplification products were then amplified with a constant forward primer containing the P5 adapter sequence and a target-specific reverse primer containing the P7 adapter sequence. Samples were barcoded by PCR, pooled, then separated by agarose gel electrophoresis. Amplicons corresponding to gDNA with/without the large deletions were separated by agarose gel electrophoresis and sequenced on separate paired-end Illumina MiSeq runs to avoid bias due to differential flow cell binding. After standard demultiplexing with bcl2fastq2, paired-end reads corresponding to deletion products were merged with FLASh 1.2.11, while paired-end reads corresponding to full-length amplicons were merged by simple concatenation with SeqKit 2.6.1. Merged reads were concatenated into a single FASTQ file and UMI-deduplicated using AmpUMI. The deduplicated sequencing files were analyzed with CRISPResso2 with the discard_indel_reads option and editing efficiency was quantified as the percent of reads either discarded (indels) or mapping exactly to the specified deletion, as previously reported[79].

Reporting summary

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

Data availability

NGS data were deposited to the NCBI Sequence Read Archive database under the accession code PRJNA1072971[78]. All data supporting the findings of the presented study are listed in the article and Supplementary Information is available upon reasonable request.

References

78. Chen, H. et al. Branched, chemically modified poly(A) tails enhance the translation capacity of mRNA. *Sequence Read Archive* https://www.ncbi.nlm.nih.gov/sra/PRJNA1072971 (2024).
79. Anzalone, A. V. et al. Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat. Biotechnol.* **40**, 731–740 (2022).

Acknowledgements

We thank the MIT Department of Chemistry Instrument Facility for providing instrument access and N. Ye (MIT) for assistance on protein purification. We thank H. Zeng (Broad Institute) for his help in designing the STARmap/RIBOmap probes. We also thank other members of X.W.'s laboratory for helpful discussion throughout the project. X.W. acknowledges the support from the Searle Scholars Program, Thomas D. and Virginia W. Cabot Professorship, E. Scolnick Professorship, Ono Pharma Breakthrough Science Initiative Award, Merkin Institute Fellowship, and NIH DP2 New Innovator Award (1DP2GM146245-01). A.H. is a National Science Foundation Graduate Research Fellow. A.H. and D.R.L. were supported by NIH U01AI142756, R35GM118062, RM1HG009490 and the Howard Hughes Medical Institute (HHMI). This article is subject to HHMI's Open Access to Publications policy. HHMI lab heads have previously granted a nonexclusive CC BY 4.0 license to the public and a sublicensable license to HHMI in their research articles. Pursuant to those licenses,

Author contributions

X.W. conceived the project. H.C. and X.W. designed experiments. H.C., D.L., J.G., A.A., Y.Z., J.T., J.R., A.H., F.K. and M.W. performed the experiments. H.C., J.G., S.L., A.H. and J.H. performed data analysis. X.W. supervised the work. H.C. and X.W. wrote the paper with input from all authors.

Competing interests

X.W., H.C., A.A. and J.G. are inventors on patent applications related to branched RNA. X.W. is a scientific cofounder, consultant and equity holder of Stellaromics and Convergence Therapeutics. D.R.L. is a cofounder, consultant and equity holder of Beam Therapeutics, Prime Medicine, Pairwise Plants, Chroma Medicine, Exo Therapeutics and Nvelop Therapeutics. The other authors declare no competing interests.

Additional information

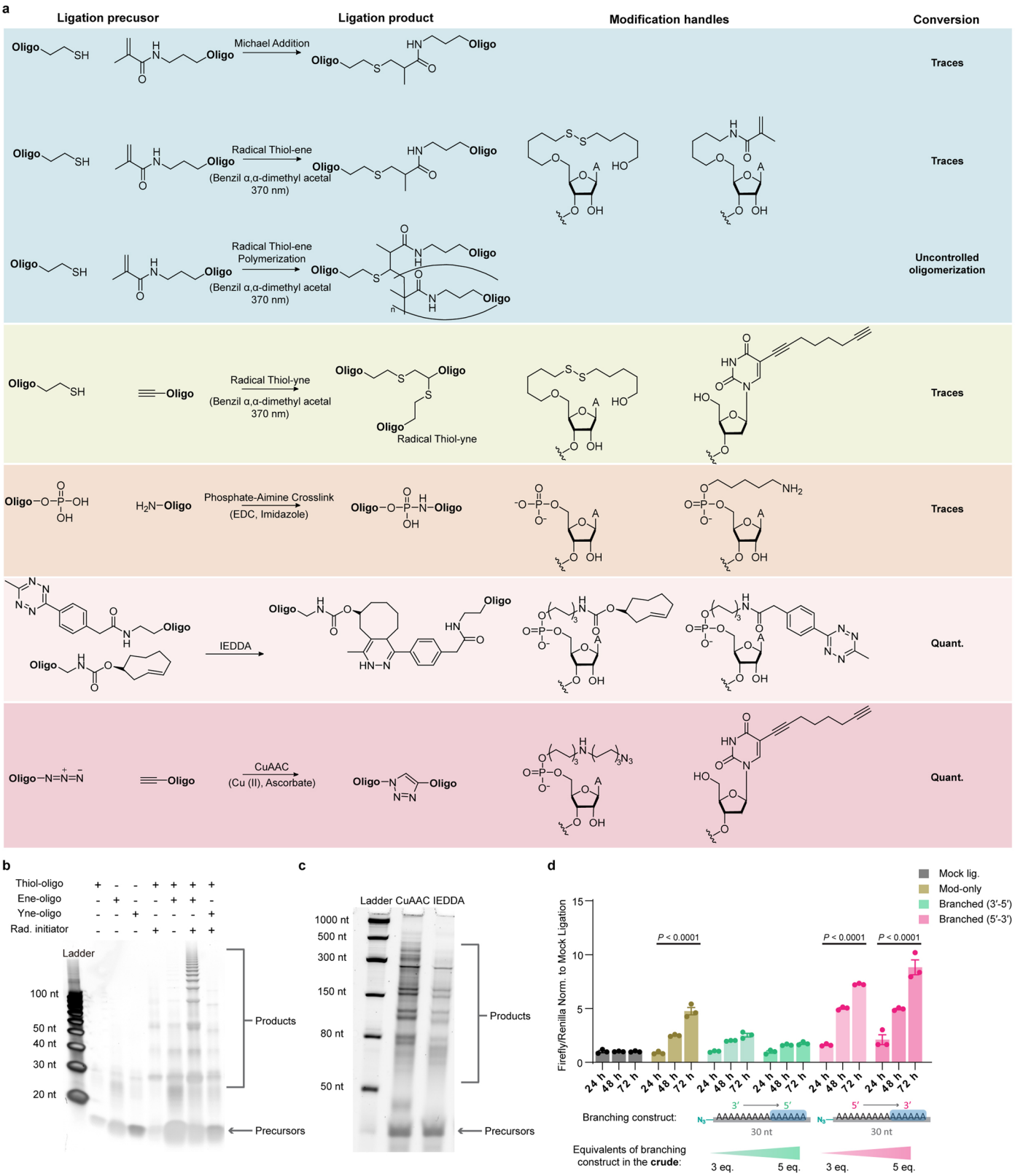
Extended data is available for this paper at https://doi.org/10.1038/s41587-024-02174-7.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-024-02174-7.

Correspondence and requests for materials should be addressed to Xiao Wang.

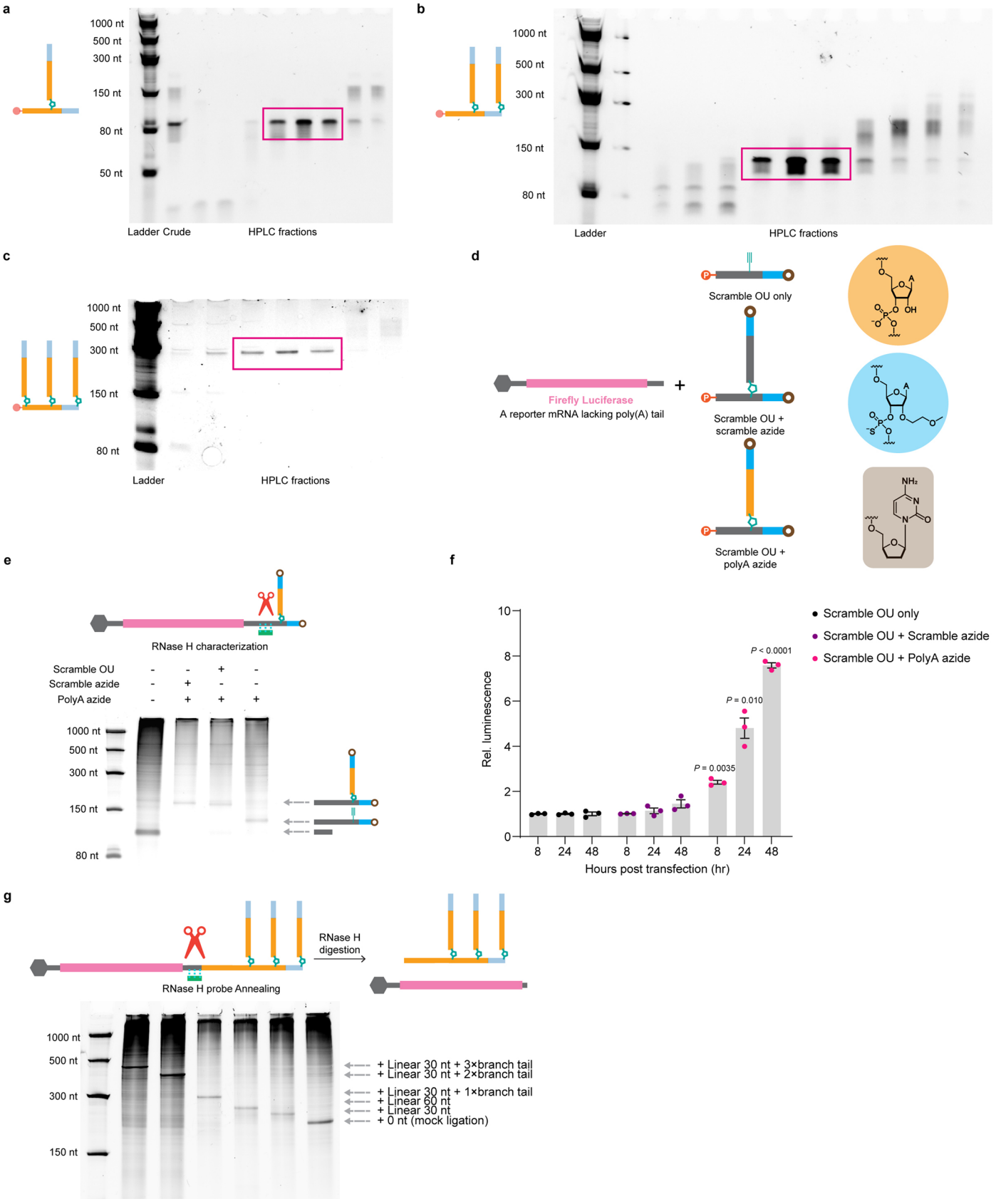
Peer review information *Nature Biotechnology* thanks the anonymous reviewers for their contribution to the peer review of this work.

https://doi.org/10.1038/s41587-024-02174-7



Extended Data Fig. 1 | Screening of oligonucleotide crosslinking chemistry. (a) Summary of oligonucleotide chemical conjugation methods. Screening was performed using 15-nt poly-deoxyadenosine model substrates at micromolar concentrations. Modification handles were incorporated through oligonucleotide solid phase synthesis, followed by amine-NHS labeling and HPLC purification. (b) Gel electrophoresis of crude thiol-ene/yne oligonucleotide conjugation of 15-nt model substrates containing only one conjugation handle. (c) Gel electrophoresis of crude CuAAC and IEDDA 30-nt oligonucleotides bearing three alkyne/trans-cyclooctene handles reacting with 30-nt azide/tetrazine modified oligo. (d) Preliminary dual luciferase assay using branched mRNA prepared from CuAAC product mixture without HPLC purification using different equivalents of branched oligos versus stem oligo. Branching oligos contained modifications on the last six bases with either natural (5′ to 3′) or reversed (3′ to 5′) directionality. Time-course assay was performed as outlined in Fig. 1. $n = 3$ independent transfections for each construct. Mean ± s.e.m. P values were calculated by ordinary two-sided one-way ANOVA (comparison of means across time points). Gels are representative of at least two experiments.

https://doi.org/10.1038/s41587-024-02174-7

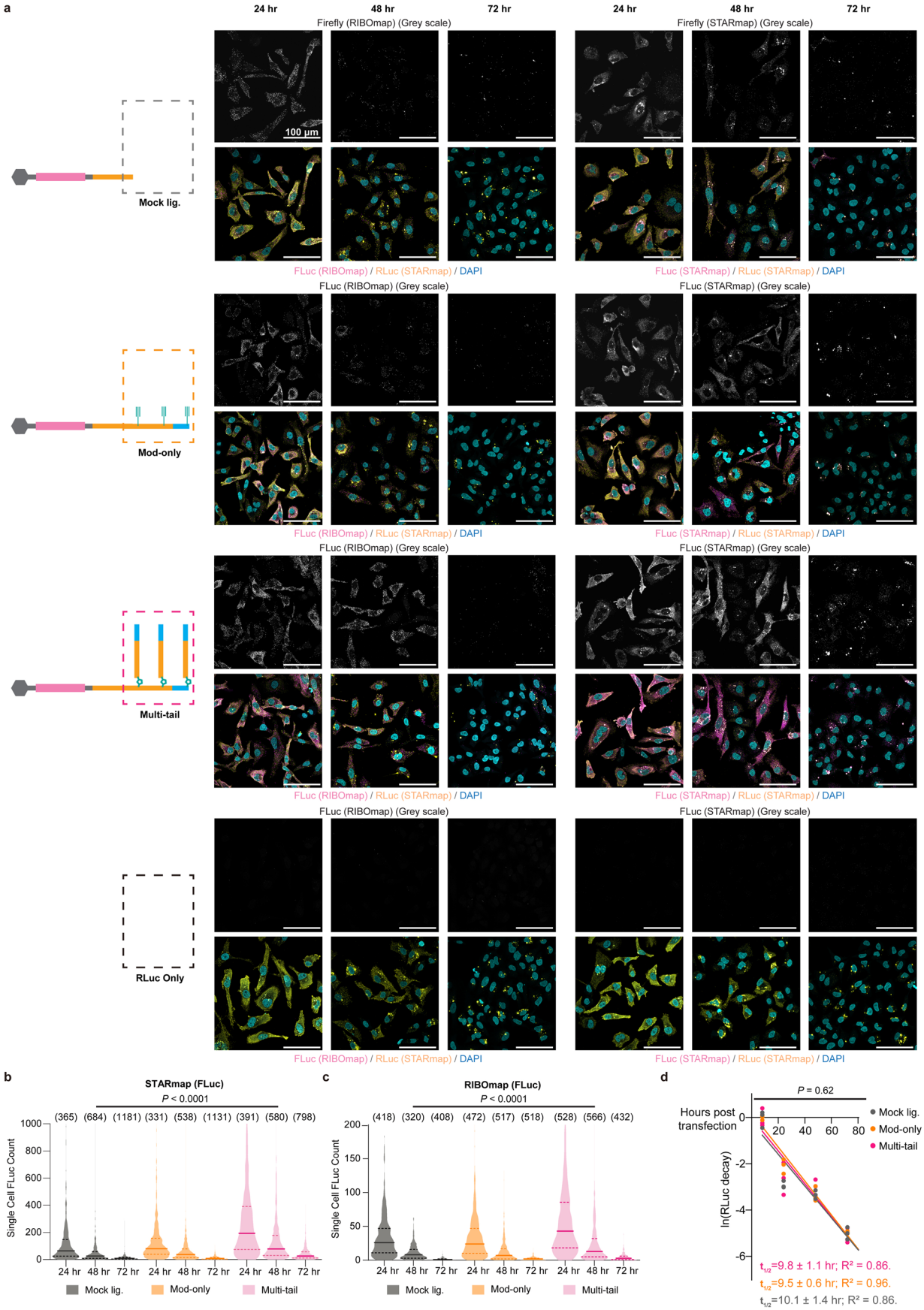


Extended Data Fig. 2 | See next page for caption.

https://doi.org/10.1038/s41587-024-02174-7

Extended Data Fig. 2 | Synthesis and characterization of topologically augmented branched mRNA. (a-c) Representative HPLC purification and gel electrophoresis characterizations of branched oligonucleotides (oligo) containing 1-3 branching poly(A) tails. Fractions containing the desired products (boxed) were pooled and isolated. (d) Schematics of preparation of firefly luciferase (FLuc) mRNA constructs lacking poly(A) tail before ligation. FLuc mRNA without template encoded poly(A) tail was ligated to a scramble (non-polyA) stem oligo with 3′ end modifications (six phosphorothioate/2MOE at 3′ and terminal dideoxycytidine modifications) and internal alkyne (5-octadiynyl deoxyuridine or OU), with or without conjugation to 5′ azide labeled scramble or poly(A) branches with 3′ end modifications (the same as the stem oligo). (e) RNase H characterization of branched mRNA-oligo conjugates with branched poly(A) or scramble tails. The branching topology was confirmed by further band shift on TBU gel. (f) Branched poly(A) sequence, rather than chemical modifications alone, conferred enhanced protein expression over time. Relative FLuc luminescence was normalized to the scramble OU only oligo ligated mRNA at indicated time points. $n = 3$ independent transfections for each construct. Mean ± s.e.m. P values were calculated by two-sided unpaired t-test (with Welch's correction) against the scramble OU + scramble azide construct at corresponding time points. (g) RNase H characterization of branched mRNA-oligo conjugates with multiple ploy(A) tails. mRNA with full-length hemoglobin UTRs and template encoded 100A-tail was ligated to 0 (mock ligation), 30 A, 60 A, or oligos with one, two, or three branched poly(A) tails and characterized by RNase H assay. The branching topology was confirmed by further band shift on TBU gel. Gels are representative of at least two experiments.

https://doi.org/10.1038/s41587-024-02174-7

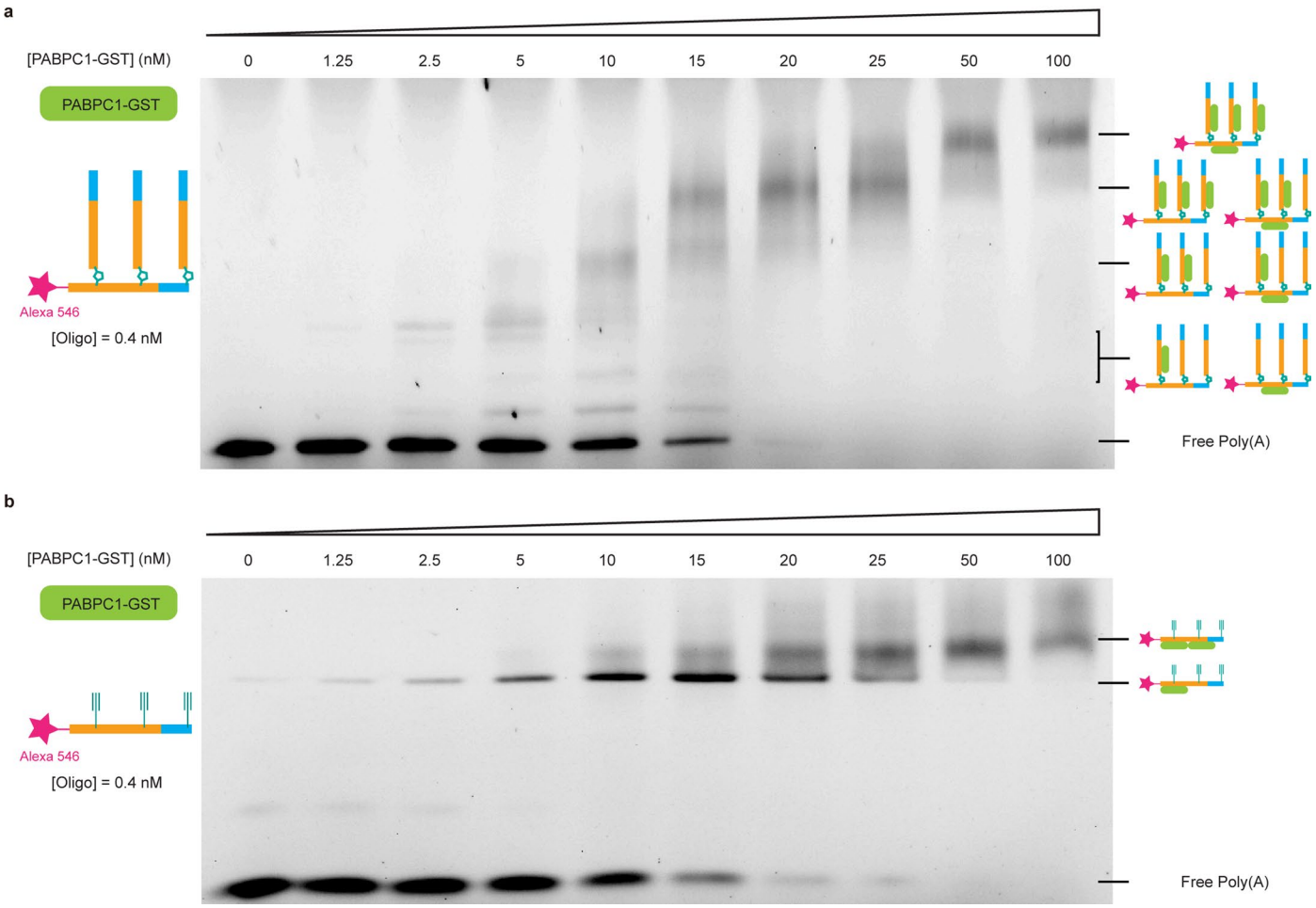


Extended Data Fig. 3 | See next page for caption.

https://doi.org/10.1038/s41587-024-02174-7

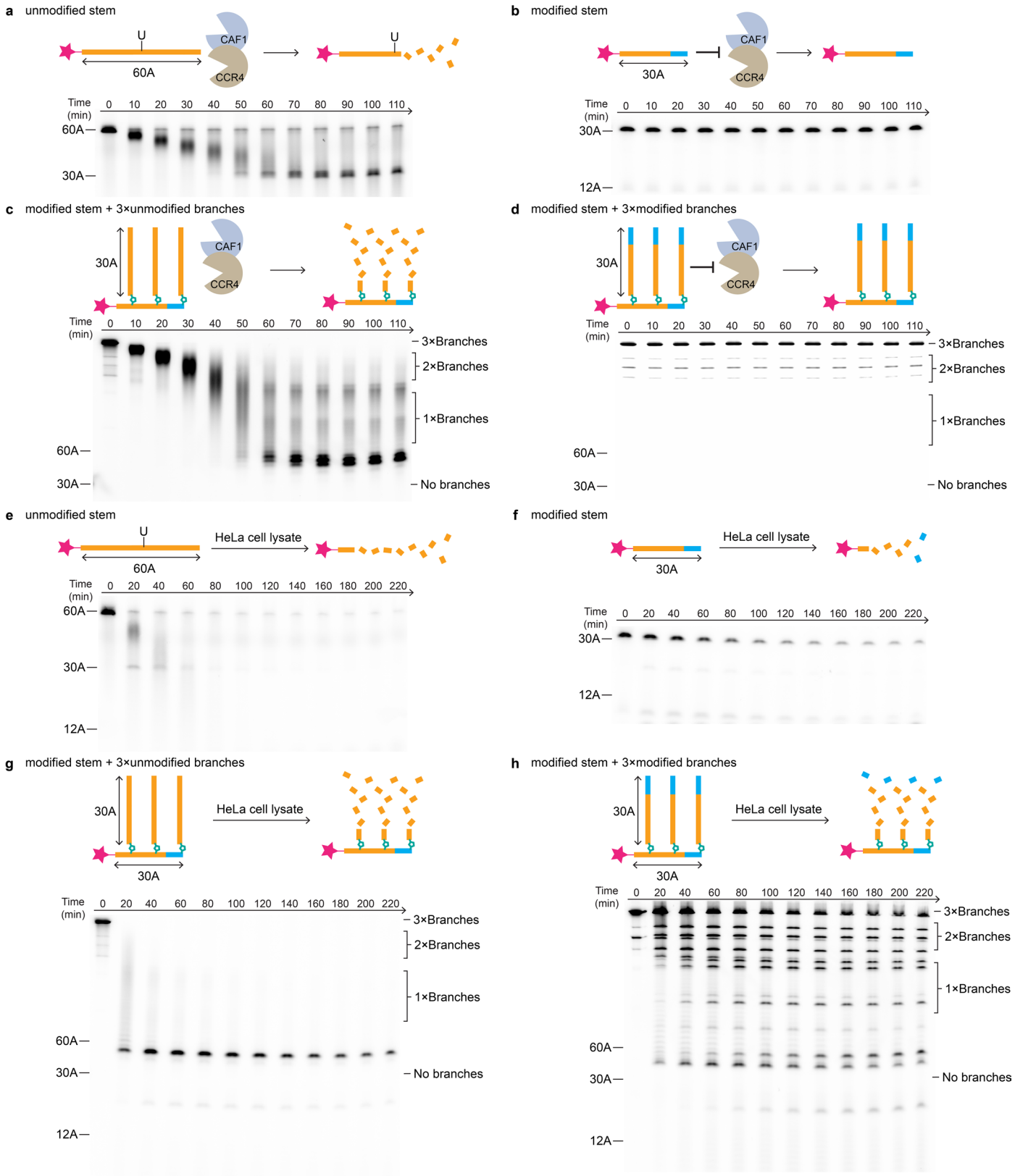
Extended Data Fig. 3 | Dissecting the effects of branched poly(A) tails on mRNA stability and translation efficiency. (a) Representative STARmap/RIBOmap images were acquired under the same confocal imaging settings from three independent experiments for each condition. STARmap versus RIBOmap characterization of different modified Firefly luciferase (FLuc) mRNA constructs and STARmap characterization of internal control Renilla luciferase (RLuc) mRNA were performed as outlined in Fig. 2. DAPI (blue), nuclei; FLuc amplicons (magenta); RLuc amplicons (yellow). Colocalized FLuc/RLuc amplicons in STARmap (white dots) were lipid transfection vesicles and were excluded from downstream quantification. Scale bar = 100 µm. (b-c) Violin plots of single-cell quantification of FLuc STARmap/RIBOmap amplicons. *P* values were calculated by ordinary two-sided two-way ANOVA. (d) Decay kinetics of internal control RLuc luminescence outlined in Fig. 2c normalized to luminescence at 8 hrs post transfection. $n = 3$ independent transfections. *P* values were calculated by ordinary two-sided two-way ANOVA (with Geisser-Greenhouse correction, decay across time points).

https://doi.org/10.1038/s41587-024-02174-7



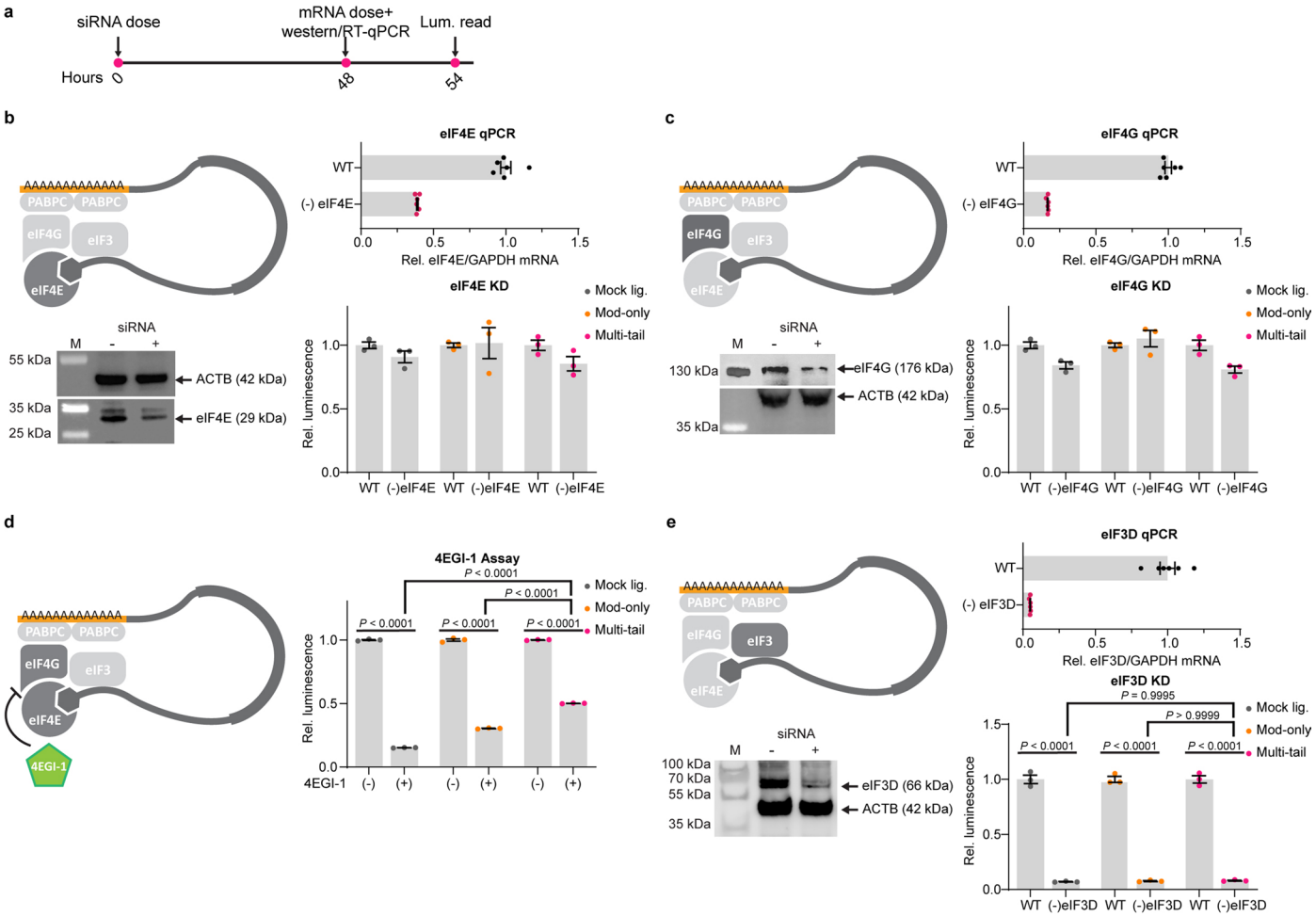
Extended Data Fig. 4 | Branched poly(A) tails bind PABPC1 protein. (a) Electrophoretic mobility shift assay of three branched poly(A) oligo with varying concentrations of GST-tagged recominant human PABPC1 protein (PABPC1-GST). The stem oligo was labeled with Alexa Fluore 546 at 5′end. Both stem and branching poly(A)'s were 30 nt and modified with PS-2MOE at the last six bases and ddC at the 3′ end. (b) Gel shift assay of the same modified poly(A) oligo without branching. Gels are representative of at least two experiments.

https://doi.org/10.1038/s41587-024-02174-7



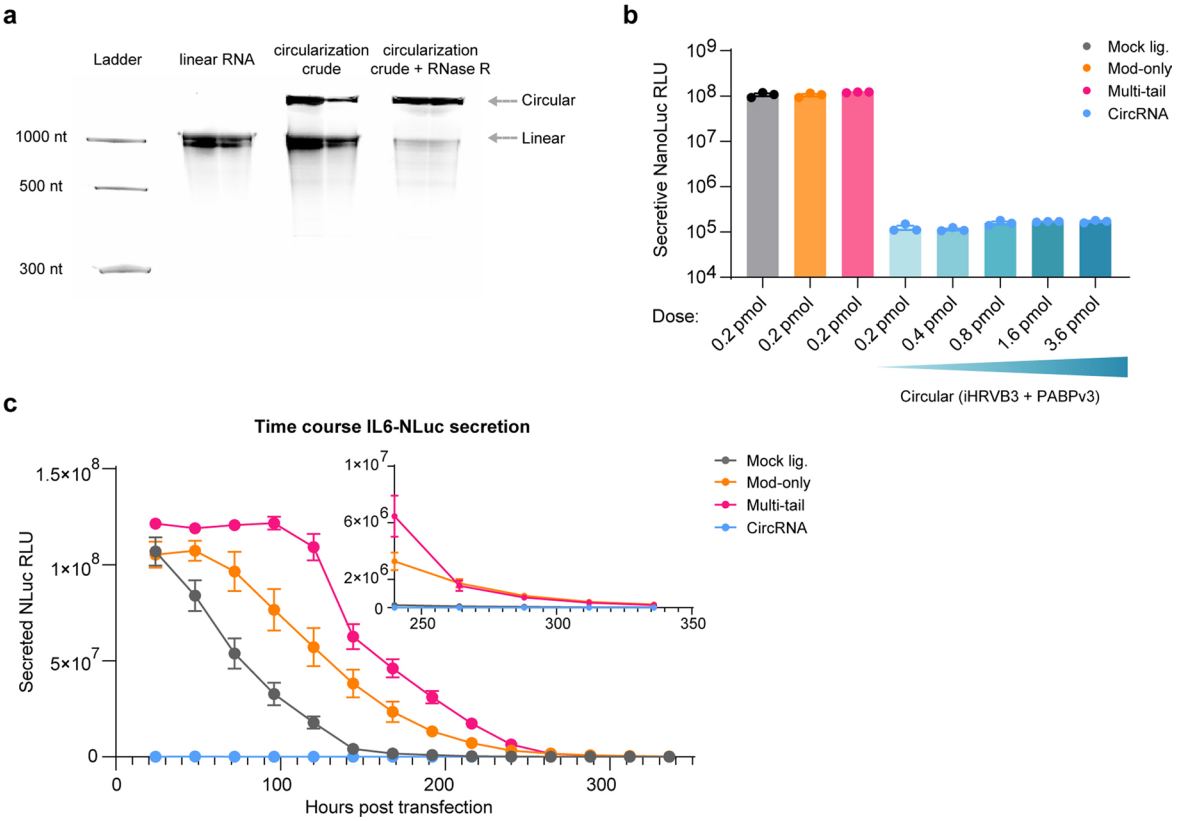
Extended Data Fig. 5 | Chemical and topological augmentation stabilized poly(A) tails against nuclease degradation in vitro. (a-d) Evaluation of chemically and topologically modified poly(A) tails in CAF1-CCR4 deadenylation assay. Four different Alexa-546 labeled poly(A) oligos were subjected to deadenylation using recombinant CAF1/CCR4 protein complex over the course of 110 min: linear unmodified 60 A oligo (a), end modified (six PS-2MOE at 3′ and terminal dideoxycytidine, the same set of modifications for all other oligos) 30 A oligo (b), end modified 30 A stem oligo + unmodified 30 A branches (c), and end modified stabilized 30 A stem oligo + end modified 30 A branches (d). (e-h) Evaluation of chemically and topologically modified poly(A) tails in HeLa cell lysate. Alexa-546 labeled poly(A) oligos (the same four constructs used in a-d) were subjected to digestion in HeLa cytosolic lysate for 220 min: linear unmodified 60 A oligo (e), end modified 30 A oligo (f), end modified 30 A stem oligo + unmodified 30 A branches (g), and end modified 30 A stem oligo + end modified 30 A branches (h). Gels were representative of two experiments.

https://doi.org/10.1038/s41587-024-02174-7



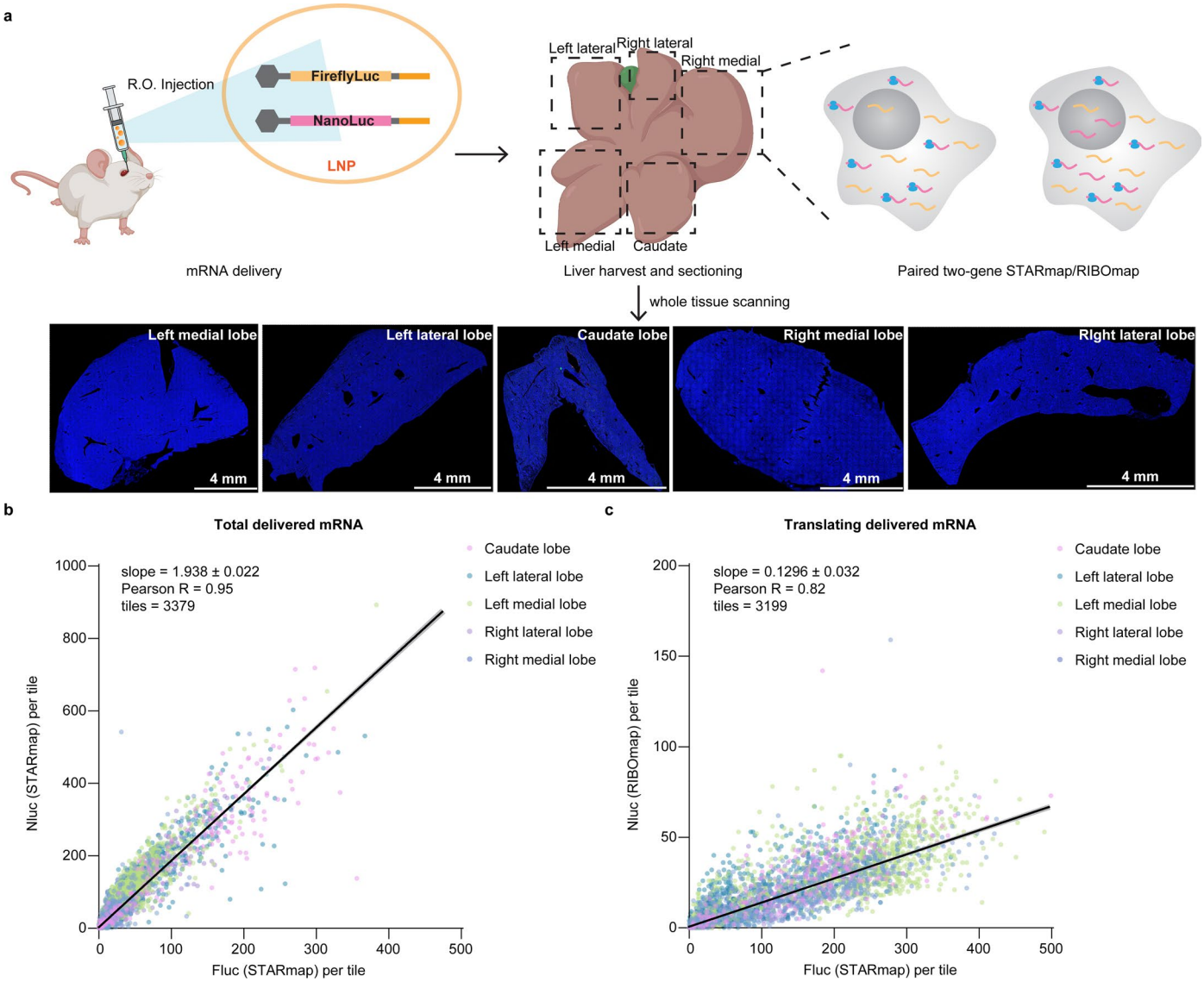
Extended Data Fig. 6 | Multi-tailed mRNA depends on the canonical eIF4-eIF3 translation initiation mechanism. (a) Timeline of the knockdown (KD) experiments. HeLa cells were treated with siRNA cocktails targeting the corresponding eIFs 48 hrs post siRNA transfection, cells were reseeded and transfected with mRNA followed by protein quantification after 6 hrs. Successful knockdowns were confirmed by western blots and RT-qPCR at 48 hrs. (b,c,e) KD experiments for eIF4E/eIF4G/eIF3D. (d) Comparison of multi-tailed mRNA to regular mRNA in 4EGI-1 treated in vitro translation assay using rabbit reticulocyte lysate (RRL). Luciferase expressions were normalized to the wild-type condition for each construct. Mean± s.e.m. $n = 3$. P values were calculated by two-sided unpaired t-test for intra-construct comparison (KD vs WT) and by two-sided one-way ANOVA for cross-construct comparison. Western blots were representative of two experiments.

https://doi.org/10.1038/s41587-024-02174-7



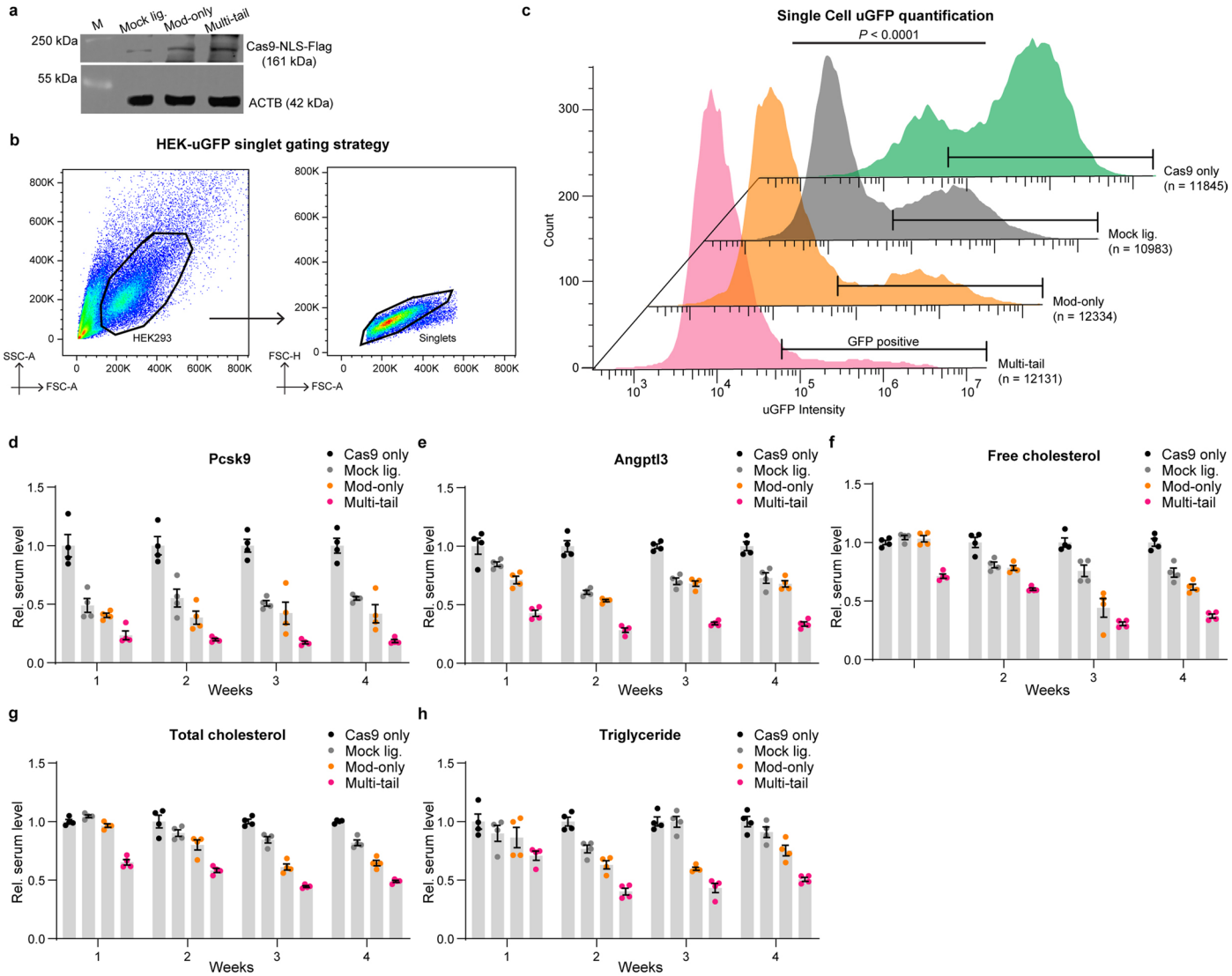
Extended Data Fig. 7 | Comparison of UTR optimized circRNA and multi-tailed mRNA using secreted NanoLuc reporter. (a) Gel electrophoresis of circRNA encoding secreted Nanoluc (IL6-Nluc) generated through IVT, backsplicing, and enriched by RNase R treatment. Successful circularization was confirmed by RNase R resistance and slower mobility on TBU gel compared to corresponding intron-free linear RNA. (b) Comparison of UTR optimized circRNA and multi-tailed mRNA using IL6-NLuc 1 day after transfection with indicated amounts of RNA. Mock lig./mod-only/multi-tail mRNAs contained optimized UTRs, full m1ψ replacement and 100 A through IVT with mock ligation (mock lig.), ligated to modification-only oligo (mod-only), or modified multi-tailed oligo (multi-tail). CircRNAs were designed to contain optimized HRV IRES (with proximal loop Apt-eIF4G insertion) and 3′-PABP binding motif without addition of modified nucleotides. Cells were cultured in phenol red-free media. On each day, media was completely harvested and renewed with 150 µL fresh media and 5 µL of old media was diluted and used for luciferase assay. $n = 3$ independent transfections in each biological condition. Mean ± s.e.m. (c) IL6-Nluc signal over 14 days. $n = 3$ independent transfections in each biological condition. Mean ± s.e.m.

https://doi.org/10.1038/s41587-024-02174-7



Extended Data Fig. 8 | Paired STARmap/RIBOmap effectively characterize the quantity and translatability of mRNA constructs delivered by LNP to murine liver. (a) Schematics for evaluating two-gene STARmap/RIBOmap experiments in different lobes of murine liver. 2:1 NLuc/FLuc mRNA were co-encapsulated in LNP and administered by retro-orbital (R.O.) injection. 6 hrs post mRNA injection, the mouse liver was harvested in five lobes (caudate, left/right medial and lateral lobes). Each lobe was sections into two adjacent 10-μm slices, with one slice profiled by NLuc STARmap and FLuc STARmap and the other slice profiled by NLuc RIBOmap and Fluc STARmap. Whole sections were profiled in 291×291 μm^2 tiles. (b, c) Quantification of numbers of NLuc and FLuc amplicons in each tile. n = 3379/3199 tiles. Slopes (NLuc/FLuc) were calculated by linear regression (least-square, Q = 5% for outlier rejection).

https://doi.org/10.1038/s41587-024-02174-7



Extended Data Fig. 9 | Delivery of CRISPR editing using chemically and topologically modified mRNA. (a) Comparison of Cas9 expression by western blot 48 hrs post transfection in HEK293T-uGFP cells. NLS, nuclear localization signal. Western blots were representative of two experiments. (b) FACS gating strategy for uGFP positive HEK293T cells. (c) Exemplary single-cell uGFP quantification for cells treated with unligated Cas9 mRNA only (Cas9-only) without sgRNAs, or with mock lig./mod-only/multi-tail Cas9 mRNA co-transfected with two sgRNAs. FACS were performed 72 hrs post mRNA transfection. $n = 11845/10983/12334/12131$ single cells. P value was calculated by two-sided one-way ANOVA. (d-h) Serum levels of Pcsk9 protein (d), Angptl3 protein (e), free cholesterol (f), total cholesterol (g), and triglyceride (h) over 4 weeks normalized to the Cas9-only group. Mean ± s.e.m. $n = 4$, biological replicates.

nature portfolio

Corresponding author(s): Xiao Wang

Last updated by author(s): 02/04/2024

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	No code was generated for data collection by this study. Software used to collect data include Leica LAS-X microscope imaging software. 3.5.5.19976, Image Lab 3.0.1.14, EnSpireManager 4.13.3005.1482, Agilent Openlab CDS, Aura 4.0, Illumina MiSeq2, Beckman CytoFLEX LX Flow Cytometer.
Data analysis	Software used to analyze data include MATLAB 2021a, CellProfiler 4.0.7, ImageJ, Prism 9, Microsoft Excel 16, Agilent Openlab CDS, Aura 4.0, FIJI stitching plugin (version 1.2), Stardist (Version 0.8.5), Squidpy (Version 1.3.0), scikit-image, bcl2fastq2, FLASh 1.2.11, SeqKit 2.6.1, Crispresso2 (http:// crispresso2.pinellolab.org/submission). Other customized scripts are listed and available on GitHub (https://github.com/wanglab-broad).

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

Data

Policy information about availability of data

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

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Source data for all figures are provided in the manuscript. NGS data were deposited to the NCBI Sequence Read Archive database under the accession code PRJNA1072971. All other data supporting the findings of the presented study are listed in the article and supplementary information.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

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

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

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

Life sciences study design

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

Sample size	At least 3 independent biological replicates per condition were used in all experiments. Specific sample sizes for all experiments were specified in figure legends/supplementary tables. Sample sizes were determined based on pilot experiments or previous publication with similar experimental designs.
Data exclusions	For cell based STARmap experiments, amplicons identified as transfection granules were excluded from data analysis. For in vivo STARmap/ RIBOmap linear regression, outliers were excluded by Q test (Q=5%).
Replication	All experiments were successfully replicated at least twice (one pilot experiment and one presented experiment).
Randomization	Samples and mice were randomly assigned for all experiments.
Blinding	Blinding is not applicable to this study and all conditions were known to the person throughout experimentation and data analyses.

Reporting for specific materials, systems and methods

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

Materials & experimental systems

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

Methods

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

Antibodies

Antibodies used

anti-ACTB-HRP [Biolegned, 664803]; anti-cas9 antibody [Cell Signaling, 14697]; anti-mouse HRP [Santa Cruz Bio, sc-516102]; anti-eIF3D-Rabbit IgG [ProteinTech, 10219-1-Ab]; anti-eIF4E-Rabbit IgG [MBL, RN00lP]; anti-eIF4GI-Rabbit IgG [MBL, RN002P]; anti-rabbit HRP [SigmaAldrich, 12-348].

Validation

All commercial antibodies were validated by the manufacturer or previous report for western blot purposes as listed below.
anti-ACTB-HRP [Biolegned, 664803]: https://www.biolegend.com/en-ie/products/direct-blot-hrp-anti-beta-actin-antibody-15655?GroupID=GROUP756
anti-cas9 antibody [Cell Signaling, 14697]: https://www.cellsignal.com/products/primary-antibodies/cas9-s-pyogenes-7a9-3a3-mouse-mab/14697
anti-mouse HRP [Santa Cruz Bio, sc-516102]: https://www.scbt.com/p/m-igg-kappa-bp-hrp
anti-eIF3D-Rabbit IgG [ProteinTech, 10219-1-Ab]: https://www.ptglab.com/products/EIF3D-Antibody-10219-1-AP.htm
anti-eIF4E-Rabbit IgG [MBL, RN00lP]: https://www.mblbio.com/bio/g/dtl/A/?pcd=RN001P
anti-eIF4GI-Rabbit IgG [MBL, RN002P]: https://products.mblintl.com/products/rn002p/
anti-rabbit HRP [SigmaAldrich, 12-348]: Lin CC, Chiang TH, Sun YY, Lin MS. Protective Effects of CISD2 and Influence of Curcumin on CISD2 Expression in Aged Animals and Inflammatory Cell Model. Nutrients. 2019 Mar 25;11(3):700. doi: 10.3390/nu11030700. PMID: 30934593; PMCID: PMC6470567.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

HeLa cells (CCL-2) were obtained from ATCC. GFP expressing HEK293 cells were obtained from Gen Target Inc.

Authentication

Cell lines were not authenticated.

Mycoplasma contamination

Cell lines were confirmed to be mycoplasma-free by DAPI staining during microscope imaging experiments. Cell lines were not tested for mycoplasma contamination in experiments not involving imaging.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified lines were used in the study.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals

6~8 weeks old male BALB/cj mice used for in vivo NanoLuc assay in this study is purchased from The Jackson Laboratory (JAX). Male wild-type C57BL/6 mice, aged 3~4 wks, used for in vivo genome editing in this study were obtained from The Jackson Laboratory (JAX). Animals were kept in humidity-controlled rooms at room temperature with 12-hour light/dark cycles.

Wild animals

No wild animals were used in this study.

Reporting on sex

Male mice were used for in vivo NanoLuc assay and in vivo genome editing due to reagent limitation.

Field-collected samples

No field-collected samples were involved in this study.

Ethics oversight

All animal procedures followed animal care guidelines approved by the Institutional Animal Care and Use Committee (IACUC) of the Broad Institute of MIT and Harvard under animal protocol# 0255-08-19. Animal experiments were conducted in compliance with IACUC policies and NIH guidelines.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	n/a
Novel plant genotypes	n/a
Authentication	n/a

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	For GFP ablation analysis, cells were digested from the plate at indicated time points to prepare single cell suspension in PBS and directly analyzed with flow cytometry.
Instrument	Flow cytometry was performed on a Beckman CytoFLEX LX Flow Cytometer.
Software	Data was analyzed using FlowJo 10
Cell population abundance	Abundances are displayed in Extended Data Fig.9 as a percentage of GFP positive cells, which constituted the vast majority of cells.
Gating strategy	Gating strategy is shown in Extended Data Fig.9b. Cells were gated using FSC/SSC to exclude debris and doublets and then gated for GFP positive cells.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.