

Selective RNA pseudouridinylation in situ by circular gRNAs in designer organelles

Received: 28 October 2023

Accepted: 11 October 2024

Published online: 24 October 2024



Lukas Schartel^{1,2}, Cosimo Jann^{1,3}, Anna Wierczeiko⁴, Tamer Butto⁵, Stefan Mündnich⁵, Virginie Marchand⁶, Yuri Motorin⁷, Mark Helm⁵, Susanne Gerber⁴ & Edward A. Lemke^{1,8}✉

RNA modifications play a pivotal role in the regulation of RNA chemistry within cells. Several technologies have been developed with the goal of using RNA modifications to regulate cellular biochemistry selectively, but achieving selective and precise modifications remains a challenge. Here, we show that by using designer organelles, we can modify mRNA with pseudouridine in a highly selective and guide-RNA-dependent manner. We use designer organelles inspired by concepts of phase separation, a central tenet in developing artificial membraneless organelles in living mammalian cells. In addition, we use circular guide RNAs to markedly enhance the effectiveness of targeted pseudouridinylation. Our studies introduce spatial engineering through optimized RNA editing organelles (OREO) as a complementary tool for targeted RNA modification, providing new avenues to enhance RNA modification specificity.

RNA function is controlled by over hundreds of modifications¹. While the tools for editing, writing, and analyzing RNA modifications are still evolving, they offer immense untapped potential for deepening our understanding of cellular functions controlled by RNA and advancing the field of synthetic biology.

One of the most abundant and functionally versatile RNA modification is pseudouridine (5-ribosyluracil, Ψ), an isomer of uridine. The structural differences to uridine, in particular the isomerization from a C–N glycosidic bond to a C–C bond, result in pseudouridine having an additional possibility to form a hydrogen bond and contribute to RNA stability². In the cell, pseudouridines fulfill critical functions within a diverse array of small RNAs, ranging from regulating the function of rRNAs and stabilizing tRNAs to regulating splicing within small nuclear RNAs³. In the context of in vitro-transcribed mRNAs, pseudouridine plays a pivotal role as a potent enhancer of translation⁴. Additionally, when situated within stop codons, it possesses the unique ability to mediate stop-codon suppression^{4,5}. Applications of engineering pseudouridine into RNAs focus on modifying stop codons within mRNAs, which represents an exciting opportunity to target many heritable

diseases caused by premature stop codons, such as variants of cystic fibrosis⁶. These methods either rely on endogenous factors limiting their efficiency⁷ or overexpression of pseudouridine synthases that could potentially cause unwanted off-target effects⁸.

Cells are equipped with numerous enzymes capable of synthesizing pseudouridine in RNAs. Most of these enzymes function as standalone pseudouridine synthases, each exhibiting specific affinities for particular structural features⁹. We used the RNA-guided pseudouridine synthase dyskerin (DKC1) to construct an RNA-modifying organelle that maximizes target versatility. H/ACA box small nucleolar RNAs (snoRNAs) form complexes with DKC1 to function as guide RNAs (gRNAs). These H/ACA box RNAs possess a common secondary structure with characteristic pockets harboring antisense guide sequences¹⁰. These pseudouridinylation pockets can be engineered to complement any target RNA, making gRNA-dependent DKC1 the ideal candidate to be used in a programmable pseudouridine-dependent RNA-modifying organelle⁷.

Here, we demonstrate that a film-like organelle architecture enables mRNA-selective pseudouridinylation as a proof-of-principle

¹Biocenter, Johannes Gutenberg University, Mainz, Germany. ²Departments of Biology and Chemistry, IMPRS on Cellular Biophysics, Mainz, Germany.

³Institute of Molecular Biology (IMB) Postdoc Programme, Mainz, Germany. ⁴Institute of Human Genetics, University Medical Center of the Johannes Gutenberg University, Mainz, Germany. ⁵Institute of Pharmaceutical and Biomedical Sciences, Johannes Gutenberg-University Mainz, Mainz, Germany.

⁶Université de Lorraine CNRS, INSERM, UAR2008/US40 IBSLor, EpiRNA-Seq Core Facility, Nancy, France. ⁷Université de Lorraine CNRS, UMR7365 IMoPA, Nancy, France. ⁸Institute of Molecular Biology, Mainz, Germany. ✉e-mail: edlemke@uni-mainz.de

for a particular mRNA-modifying organelle. The concept behind film-like organelles is to create a unique biochemical environment within a living mammalian cell by borrowing principles from 2D phase separation, where the film-like structure still stays in exchange with the cytoplasm^{11,12}. Primarily, mRNAs targeted at the thin film should be selectively modified, whereas mRNA located elsewhere in the cytoplasm should get less frequently modified. As the environment of the thin film is in equilibrium with the cytoplasm, there are no physical boundaries to be passed, which abrogates the need for complex transport machinery. Furthermore, components can be shared between the host cell and the designer organelle so that film-like organelles can be simple in design and built from a few critical components while borrowing others from the cytoplasm.

We also show that mRNA pseudouridylation can be significantly enhanced using circular gRNAs. Our results show the functionalities of film-like organelles for RNA editing and highlight the power of circular gRNAs for targeted stop-codon suppression via pseudouridine, offering new possibilities ranging from synthetic biology to RNA therapeutics.

Results

The unique ability of pseudouridine to function as a stop-codon suppressor allows us to design a simple mRNA reporter encoding for a fluorescent protein with a premature stop codon. When stop-codon suppression occurs, the mRNA produces a fluorescent protein that can be detected using flow cytometry. For this purpose, we designed a gRNA derived from H/ACA box RNA pugU2-34/44 of *Xenopus laevis*, which has recently been used to mediate targeted stop-codon suppression¹³. We customized the base pairs within the pseudouridylation pocket of the gRNA to target an iRFP-eGFP^{39TAG} reporter, which should produce eGFP only if pseudouridylation of the premature stop codon occurs.

We built upon our previous design of orthogonally translating film-like organelles¹², which consist of an N-terminal domain of the rodent LCK tyrosine kinase (LCK¹⁻¹⁰, which we from here on refer to as LCK), serving as a plasma membrane (PM) anchor (Fig. 1) an intrinsically disordered region (IDR, aa 2–267) derived from the protein FUS, and major capsid protein (MCP), which serves as a recruitment domain. These parts and various control constructs lacking either the LCK, IDR, or MCP domain were N-terminally fused with dyskerin under the control of a CMV promoter. We utilized a truncated version of dyskerin (aa 22–424), which we refer to hereafter as DKC1. This version lacks NLS/NoLS signals at the N or C termini and we added a nuclear export signal (NES) at the N terminus, which is present in all DKC1

fusion constructs throughout this study. It has also recently been shown that only isoform 3 of dyskerin, which is N/C-terminally truncated, can mediate efficient pseudouridylation of an mRNA target when overexpressed⁸. Airyscan resolution images of a DKC1 fusion construct highlight the differential localization at the PM when fused to LCK (Fig. 2a and Supplementary Fig. 1).

To first assess if such complex fusion constructs can still mediate stop-codon suppression through pseudouridylation, we used an iRFP-eGFP^{39TAG} reporter harboring two RNA stem loops derived from the ms2 phage in the 3' UTR, which we term the efficiency reporter (Fig. 2b). These RNA loops bind specifically to the MCP domain present in our organelle constructs, allowing selective recruitment of our efficiency reporter. We then transiently transfected HEK293T cells with our efficiency reporter, as well as different constructs of DKC1 and a custom-designed gRNA under the control of a U6 promoter. Next, we analyzed our cells with flow cytometry (Supplementary Fig. 2) and calculated the percentage of GFP-positive cells, indicating successful stop-codon pseudouridylation. A LCK-FUS²⁻²⁶⁷-MCP-DKC1 fusion construct in which the efficiency reporter is actively recruited via the ms2/MCP tagging system mediated efficient nonsense suppression when co-transfected with a gRNA, albeit with lower efficiency than NES-DKC1. LCK-DKC1 constructs missing either the IDR or MCP showed lower efficiency in comparison (Fig. 2b, c).

Having established the functionality of our RNA editing organelle (REO) construct, we investigated whether REOs can mediate selective stop codon pseudouridylation. For this purpose, we switched to a bidirectional expression plasmid encoding mCherry^{190TAG, 2xms2} and eGFP^{39TAG}, where both reporters are separately transcribed under the control of a CMV promoter, which we refer to as the selectivity reporter. We then transfected HEK293T cells with our organelle construct and two specific gRNAs targeting the premature stop codons in our fluorescent reporters. In this scenario, we expect mCherry fluorescence to dominate in cells that are co-transfected with our REO because the mCherry reporter mRNA is actively recruited via the ms2/MCP tagging system. In the flow cytometric analysis, we calculated the relative fold change of selectivity, which we define as the ratio of the median mCherry and eGFP intensities. Indeed, we observed a greater than 2.5-fold increase in selectivity in cells hosting our designer organelle compared to the cells expressing NES-DKC1 (Fig. 2d, e). This effect is dependent on ms2 loops as a selectivity reporter without ms2 loops cannot be selectively pseudouridylated (Supplementary Fig. 3). Conversely, inverting the selectivity reporter design (eGFP^{39TAG, 2xms2}/mCherry) qualitatively inverted this effect (Supplementary Fig. 4).

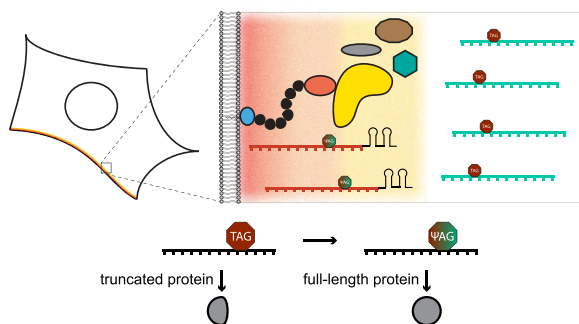
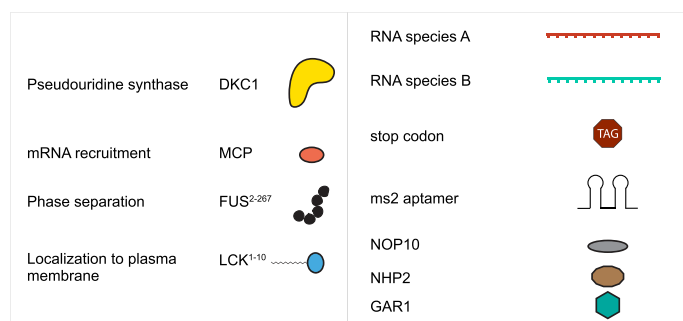
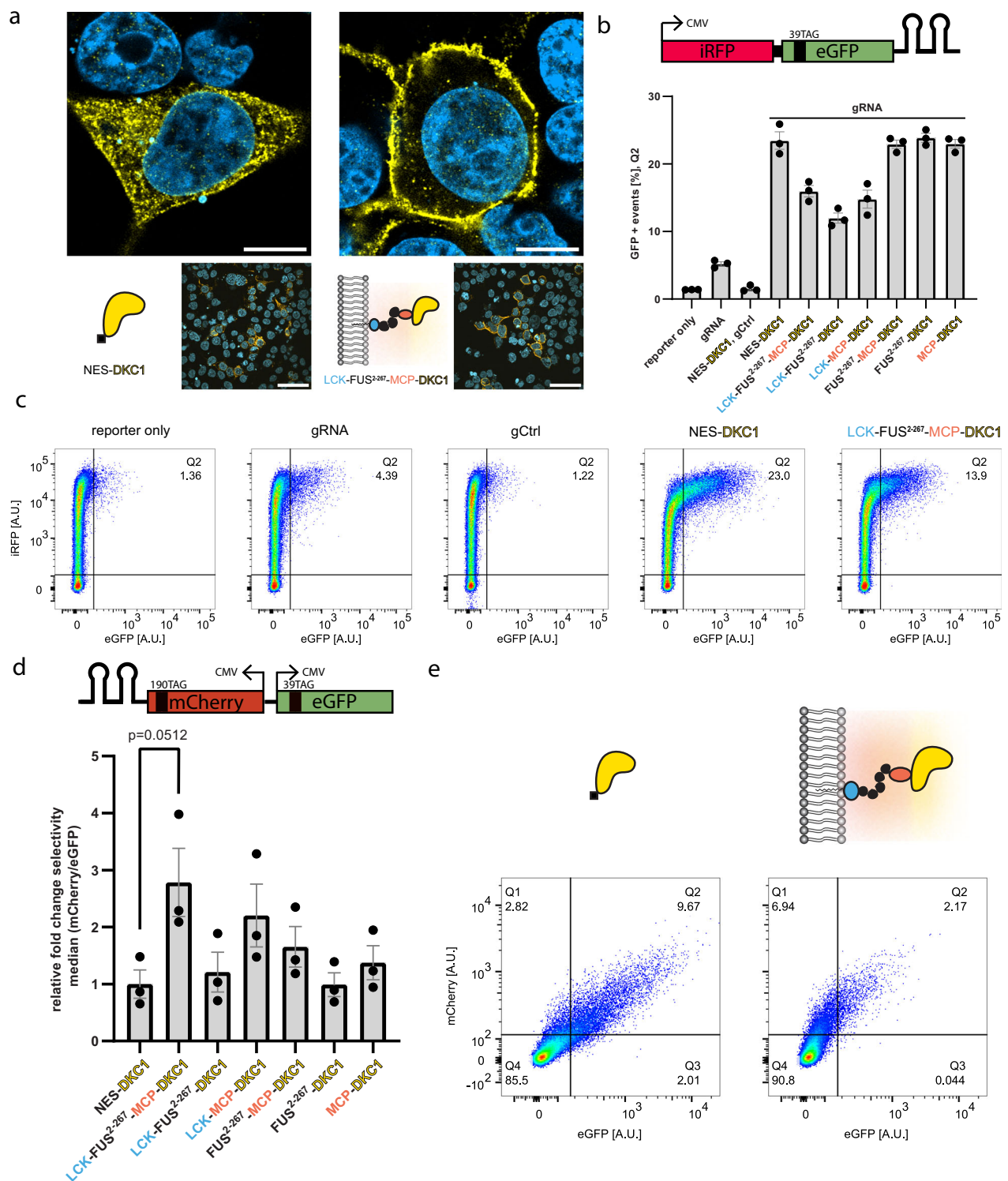


Fig. 1 | Film-like organelles for selective RNA modification. Cartoon showing a film-like organelle mounted on the plasma membrane. The colored shading indicates the organelle microenvironment, RNA is recruited through ms2 loops into the organelle, and modified stop codons result in the full-length production of protein.



gRNAs are not shown to keep the illustration simple. DKC1 is active as a tetrameric complex with NOP10, NHP2, and GAR1, which are also drawn here for reference. Our method enabled us to utilize the endogenous pool of those, which alleviated the need to engineer any of those.



Circular gRNAs markedly enhanced pseudouridine-mediated stop-codon suppression

To optimize the efficiency of our organelle for pseudouridine-dependent stop-codon suppression, we were inspired by an elegant approach for the circularization of small aptamers¹⁴. We chose to express our gRNA that targets the efficiency reporter (Fig. 2b) under the control of a U6 promoter flanked by twister ribozymes and a 10-nt polyadenine linker (Fig. 3a). After expression, twister ribozymes cleave themselves, leaving an RNA stem that is ligated by the endogenous RNA ligase RtcB. Excitingly, circular gRNAs markedly increased the

population of GFP-positive cells, allowing up to 10% more GFP-expressing cells in the case of our organelle (Fig. 3b, c). Circularization of our gRNAs was confirmed by RT-PCR with inverse primer pairs where only a circular gRNA results in a PCR product (Fig. 3d). We suspect that the increase arises from the greater resistance of circular gRNAs to exonucleases, allowing them to accumulate at much higher levels than their linear counterparts as it has been shown in the past for circular aptamers and circular RNAs for ADAR mediated RNA editing^{14–16}. In parallel, a recent study has demonstrated that circular RNAs are preferentially exported from the nucleus¹⁷. This finding offers

Fig. 2 | Film-like organelles mediate stop-codon suppression by selective pseudouridylation. **a** Immunofluorescence of HEK293T cells transfected with NES-DKCI (left) and LCK-FUS²⁻²⁶⁷-MCP-DKCI (right), alongside a cartoon depicting DKCI (yellow), MCP (pink), FUS²⁻²⁶⁷ (black), and LCK (blue). Cells were stained against a FLAG tag with Alexa-488; the nuclei were stained with DAPI. Scale bars = 10/50 μ m (top/bottom images). Top images have been taken with an LSM900 Airyscan2, bottom images have been taken with an BC43 spinning disk confocal microscope. Imaging was repeated once with similar results. **b** Pseudouridine-mediated stop-codon suppression was measured by co-transfecting HEK293T cells with an iRFP-GFP^{39TAG, 2xms2} reporter, a specific gRNA targeting eGFP, and various fusion constructs of DKCI. 5×10^4 live cells were analyzed using flow cytometry and GFP + events [%] were plotted. A control gRNA (gCtrl), in which target sequences have been randomized, served as a control. The

standard error of the mean (SEM) from three independent experiments are shown. **c** Representative flow cytometry scatter plots shown for key samples in **(b)**. Q2 indicates the eGFP positive population. **d** A reporter plasmid carrying mCherry^{190TAG, 2xms2} and eGFP^{39TAG} under the control of two independent CMV promoters served as a selectivity reporter. HEK293T cells were co-transfected with the selectivity reporter as well as two specific gRNAs targeting mCherry/eGFP (both transcribed from the same plasmid, each under the control of a U6 promoter) and various fusion constructs of DKCI. 5×10^4 live HEK293T cells were analyzed using flow cytometry. The relative fold change in selectivity, which is defined as the ratio of median mCherry/eGFP intensities, was normalized to NES-DKCI and plotted. The SEMs from three independent experiments are shown. An unpaired two-sided *t*-test was performed ($n = 3$; * $p < 0.05$) **(e)** Representative flow cytometry scatter plots for data displayed in **(d)**. Source data are provided as a Source Data file.

another potential explanation for the observed higher pseudouridine modification efficiencies with circular gRNAs since all our DKCI constructs are located in the cytoplasm.

We speculated that circular gRNAs could be further engineered and perhaps also actively recruited to our REOs. To test this, we added different versions of ms2 loops 5' and 3' of our gRNAs (Supplementary Fig. 5). In parallel, we designed LCK-FUS²⁻²⁶⁷-4x λ N22-FUS²⁻²⁶⁷-MCP-DKCI organelles, which comprise four repeats of the λ N22 peptide. Similar to the ms2/MCP system, λ N22 peptides facilitate the recruitment of RNAs tagged with boxB RNA loops¹⁸. In addition, we added another FUS²⁻²⁶⁷ linker between the two RNA-binding proteins, hoping the construct would retain flexibility and allow full access to both recruitment domains. Because we tagged our circular gRNAs with ms2 loops, we switched to a bidirectional mCherry^{190TAG, 4xboxB}/eGFP^{39TAG} reporter to avoid competition between circular gRNAs and the reporter for binding MCP. After comparing cells co-transfected with ms2 tagged to unlabeled circular gRNAs, we observed a further optimization in selectivity up to 4-fold when using these gRNA-optimized RNA editing organelles (OREO).

To establish the generality of our approach, we also tested three more previously used membrane anchors¹² to form OREOs (Supplementary Figs. 6 and 7). We show data for additional Golgi, ER, and mitochondrial-targeted OREOs that perform similarly to the LCK, i.e., the PM anchoring organelles. In all cases, we observed a beneficial effect of inserting the FUS²⁻²⁶⁷ LCD in our constructs.

Selective pseudouridylation with circular gRNAs

Having established the enhancing effect of circular gRNAs, we tested their effectiveness in our dual-color reporter experiment introduced above in Fig. 2d. We transfected HEK293T cells with our selectivity reporter and circular gRNAs for mCherry^{190TAG, 2xms2} and eGFP^{39TAG}, as well as our OREO. We observed that cells harboring organelles maintained high selectivity (Fig. 4a), preferentially producing mCherry. The enhancing effect of circular gRNAs was more apparent when we looked at the percentage of mCherry-positive cells, which was, on average, more than doubled compared to REOs that use linear gRNAs (Fig. 4b, c).

Direct methods to validate RNA pseudouridylation are still a contemporary research field^{19,20}. To validate RNA pseudouridylation qualitatively but directly, we chose to use direct RNA-sequencing (DRS; Oxford Nanopore Technologies). DRS allows natural RNA to be sequenced without prior amplification steps. Specific current changes during the sequencing process can then be detected and used to identify the underlying canonical bases. It has been shown that DNA or RNA base modifications can differ in their current signal pattern from their canonical counterparts, allowing those modifications to be identified^{20,21}. The most commonly observed pattern for pseudouridine modifications is a systematic uridine–cytosine mismatch, i.e., pseudouridine is detected as a C^{22,23}. This facilitates the classification of individual reads into unmodified and modified reads based on basecalling error information, allowing quantification of pseudouridine modifications at specific sites. We verified the applicability of DRS by sequencing fully modified

synthetic RNA oligos, mimicking our target sequences within mCherry and eGFP (Supplementary Figs. 8, 10, and 11). We extracted poly-A RNA from cells used in our selectivity assay (Fig. 4c) and from cells in which a non-targeting circular gRNA was expressed with our organelle and selectivity reporter, followed by the direct RNA-sequencing protocol from Oxford Nanopore Technologies. After mapping the base called reads onto the eGFP and mCherry reporter sequences, the basecalling error patterns, including U–C mismatches, were extracted for all reads mapping on the mCherry and eGFP reporter sequence from both the NES-DKCI- and designer organelle-transfected cells. Relative changes are consistent with our flow cytometry data; that is, OREO-targeted mCherry is preferentially pseudouridylated at the desired position with mCherry/eGFP fold changes increasing from 2.9-fold to 25.1-fold (Fig. 4d, top) (Supplementary Fig. 9).

In parallel, we employed BID-seq (bisulfite-induced deletion sequencing) to cross-validate our DRS results. BID-seq allowed quantification of pseudouridine at single base resolution and was also able to confirm the observed modification trend albeit with lower overall fold changes (mCherry/eGFP ratios increasing from 3.4-fold to 15.7-fold) (Fig. 4d, bottom) as observed in DRS.

OREOs modify disease-relevant mRNA

To test OREOs beyond our standard fluorescent reporters, we designed circular gRNAs for AldoB^{148TGA}, a mutation that renders the AldoB gene non-functional, causing hereditary fructose intolerance. We now expressed from a bidirectional promoter the AldoB(148TGA) mutant gene fused to ms2 loops as well as the eGFP gene to track off-target effects. Western blot analysis shows an efficient high on-target modification leading to expression of full-length AldoB gene while substantially reducing off-target eGFP levels (Fig. 4e, f).

Discussion

One of the challenges of synthetic biology is to engineer biochemical reactions with high specificity in complex cellular systems. Our study contributes to this challenge by showing that artificial film-like organelles can mediate selective RNA modifications. Our designer organelle approach allows the spatial engineering of pseudouridine, thereby lowering the chance for unwanted off-target modifications. This was demonstrated using the dual-color reporter (Fig. 1), where full-length GFP is made when STOP codons get altered in the cytoplasm, while full-length mCherry is made when the OREO alters the STOP codon.

Additionally, by using circular gRNAs, our organelles show increased levels of pseudouridylation indicated by pseudouridine-mediated stop-codon suppression. Because pseudouridines also influence other cellular processes, such as splicing, applying the OREO concept to the regulation of splicing might be attractive. We believe this designer organelle approach to engineering RNA modifications can be broadened to encompass other RNA modifications that state-of-the-art assays can accurately detect.

We show that using a ribozyme-mediated circularization approach can enhance the efficiency of RNA-guided

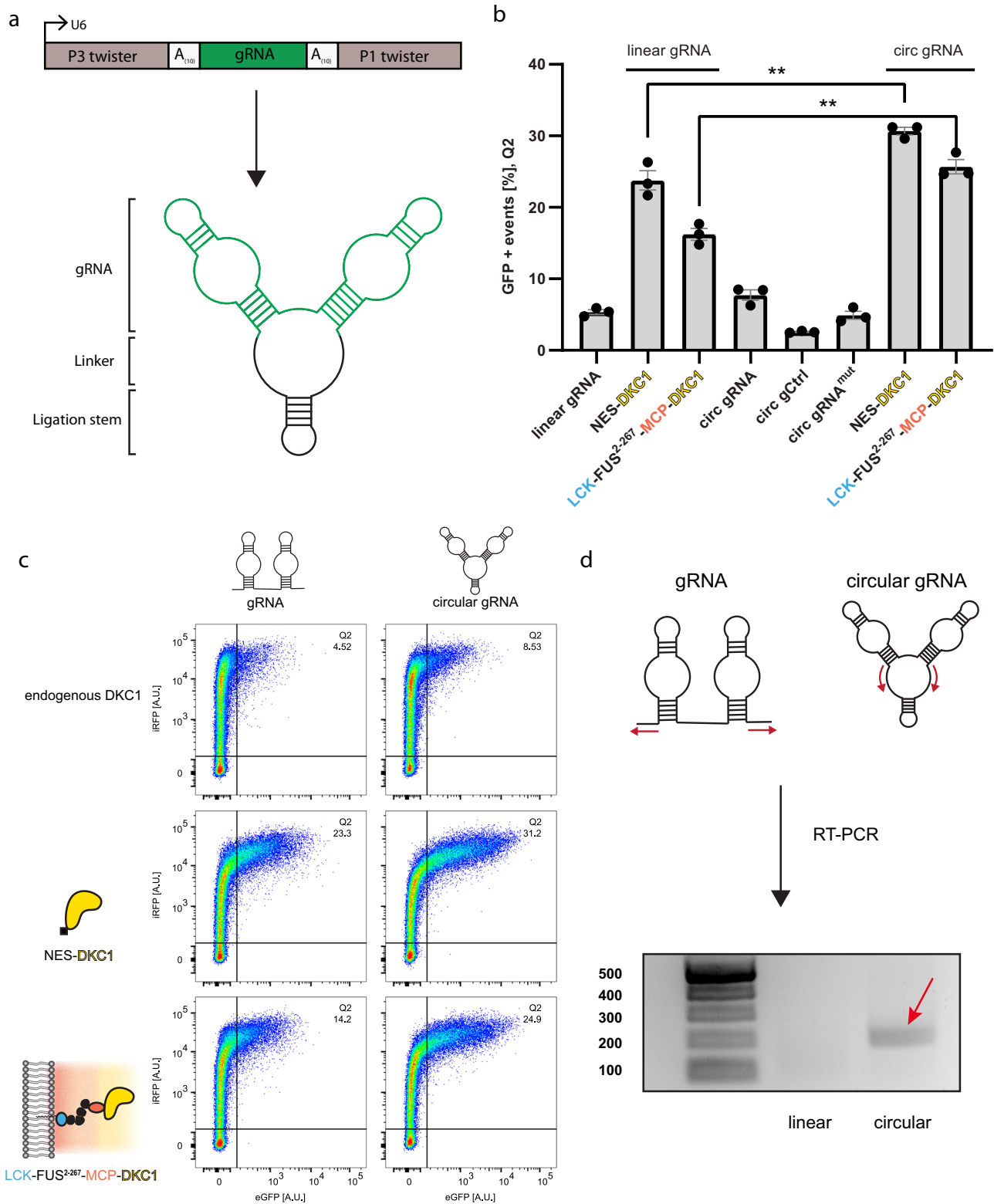
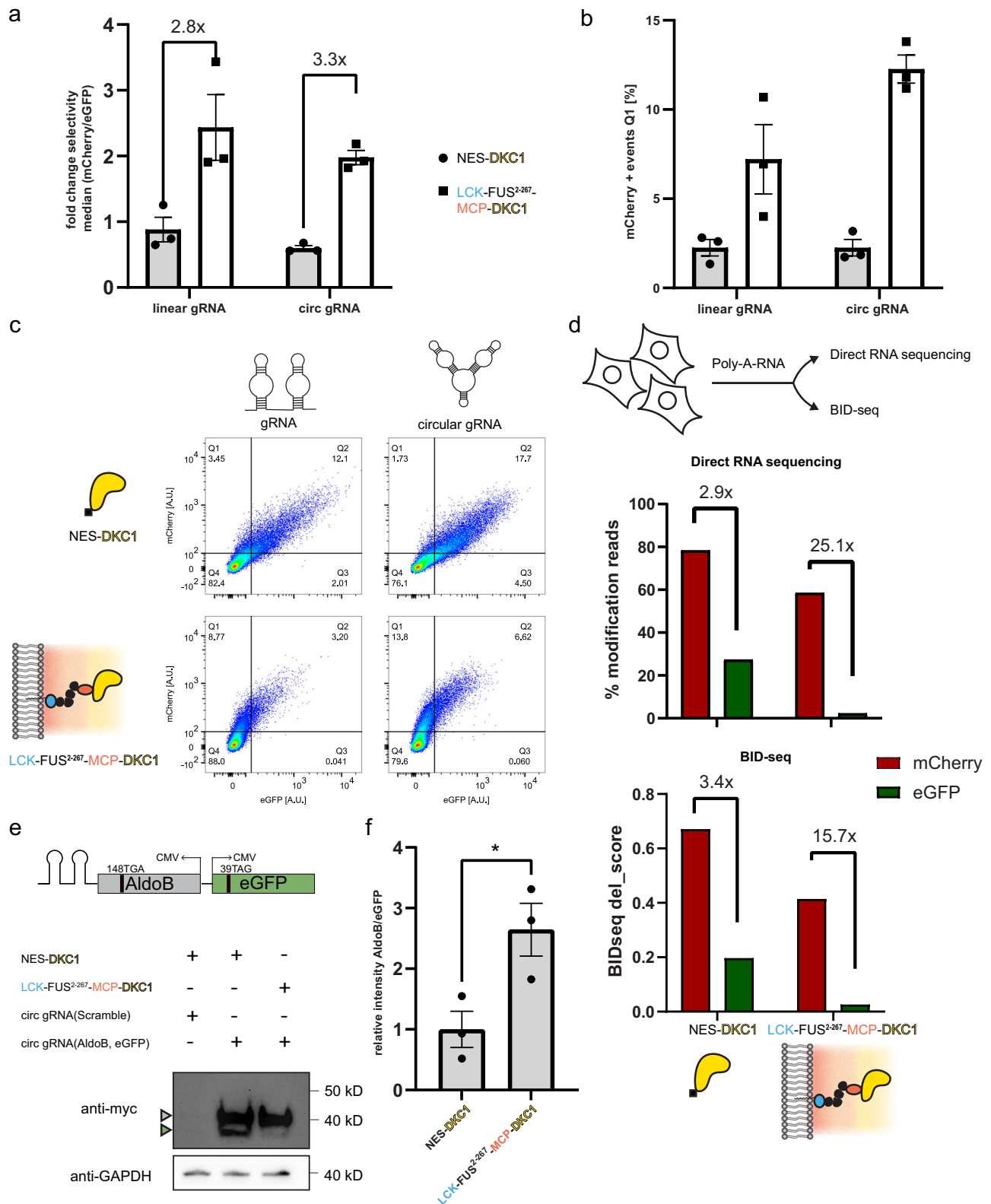


Fig. 3 | Circular gRNAs allow enhanced pseudouridine-mediated stop-codon suppression. **a** Cartoon depicting the generation of circular gRNAs from a U6 promoter using the Tornado system. A 10-nt polyadenine linker was incorporated 5' and 3' of the gRNAs to maintain flexibility. **b** Circular gRNAs outperform linear gRNAs in HEK293T cells co-transfected with the efficiency reporter (Fig. 1b) and DKC1 constructs. The SEMs of three independent measurements are shown, in which 5×10^4 live cells were measured using flow cytometry, an unpaired two-sided

t-test was performed ($n = 3$; $p = 0.0089$ for NES-DKC1, $p = 0.0019$ for LCK-FUS²⁻²⁶⁷-MCP-DKC1, $**p < 0.01$). **c** Flow cytometry scatter plots shown for key samples in (b). **d** circularization of gRNAs was confirmed by RT-PCR using inverse primers, which only result in a PCR product if the gRNA is circularized. The expected product size is 199 bp. Purified PCR samples were run on an agarose gel (ladder is given in base pair units). RT-PCR was repeated once with similar results. Source data are provided in Source Data file.



pseudouridylation. Through further optimization of linker sequences and RNA folding kinetics, circular H/ACA box RNAs can possibly be optimized to mediate even higher rates of pseudouridylation. Additionally, circular gRNAs can potentially be further functionalized with RNA loops, allowing active recruitment into our designer organelles.

More than 10% of genetic diseases are caused by premature stop codons. Stop codon suppression by pseudouridine offers great

potential in treating these conditions²⁴. The enhancing efficiency of circularized H/ACA box-derived snoRNAs holds the potential for targeting such diseases in the future. In a recent study, overexpression of near cognate tRNAs has been shown to increase pseudouridine-mediated stop codon suppression. Our enhanced pseudouridylation system, developed through this approach, holds the potential to synergize with near-cognate tRNAs, further facilitating enhanced stop codon suppression by pseudouridine. Currently, our ability to detect

Fig. 4 | Circular gRNAs increase designer organelle-mediated pseudouridylation. **a** HEK293T cells were co-transfected with the selectivity reporter (Fig. 1c), gRNAs, and either NES_DKCI or LCK_FUS²⁻²⁶⁷_MCP_DKCI. 5×10^4 live cells were measured and the fold change in selectivity was calculated for linear and circular gRNAs. **b** Simultaneously, cells were gated and analyzed for mCherry + cells Q1, as depicted in (c). SEMs of three independent replicates are shown for (a, b). **c** Representative flow cytometry scatter plots shown for (a, b). **d** Poly-A RNA was extracted from transfected HEK293T cells and analyzed via direct RNA sequencing using nanopore sequencing via BID-seq. The percentage of modified reads was plotted and the fold change between reporter modifications was compared. Analogous modifications from BID-seq levels (BIDseq del_score: 0–1 refers to unmodified and fully modified, respectively) were compared between NES-DKCI and

OREO transfected cell reporters. **e** Selective pseudouridylation of AldoB(148TGA) expressed from a modified selectivity reporter with both proteins tagged with 2 \times myc tag on the c-terminus. Transfected HEK293T cells were lysed and enriched for full-length AldoB and eGFP via myc-trap beads. A representative western blot of the eluted proteins is shown (top), triangles (gray/AldoB; green/eGFP) indicate the respective protein bands. A small aliquot of lysate was used for myc-trap pull-down and used as a loading control (bottom). Unprocessed blots are available in Source Data file. **f** Protein bands were quantified and the relative fold change in AldoB/eGFP intensities was calculated. Three independent experiments were performed. An unpaired two-sided *t*-test was performed ($n = 3$; $p = 0.0359$, $*p < 0.05$) and SEM was plotted. Source data are provided as a Source Data file.

RNA modifications makes transcriptome-wide detection of off-target effects a challenge on its own. Nevertheless, we can show that when overexpressing a disease-relevant gene with an “off-target” gene, we can selectively modify our target RNA using OREOs. The use of our reporter design enables us to visualize how selectivity for RNA editing can be achieved when using film-like OREOs, which greatly expands the use of such technologies for enzymatic engineering in general. Specificity and limited off-target effects will always be the principal reasons for selecting the optimal modification tool. Because the film-like organelle is accessible, i.e., open, from the cytoplasm, a major benefit remains the simplicity, as also, in this case, only one of the four components of the DKCI machinery had to be spatially concentrated to the organelle. At the same time, such film-like systems are in equilibrium with the cytoplasm, and thus, biomolecules can still diffuse out of the OREO and into the cytoplasm, and an infinitely high selectivity cannot be expected.

Previous orthogonally translating organelles (OTOs), have been used to alter the genetic code of translation to introduce non-canonical amino acids into proteins, and selectivity increases of ~10 fold have been found¹¹. The “open” architecture of an OREO or an OTO might even enable their combination to reprogram RNA function and/or protein synthesis holistically beyond current single methods. This open architecture is inspired by concepts from phase separation, where organelle-like structures in equilibrium with the cytoplasm can form. We also note that even without using the FUS LCD, all our organelles outcompete the cytoplasmic distributed NES-DKCI system in terms of selectivity. However, understanding such highly concentrated 2D systems’ biophysical nature is still part of contemporary research even for single protein systems^{25,26}. Thus, here, we focus on the engineered function and the membrane-targeted appearance when referring to a film-like OREO.

The open design might also facilitate combining film-like organelle technology with other clever and elegant methods, such as CRISPR and ADAR-based RNA engineering^{27,28}. OREOs thus do not compete with existing methods but might provide another layer of tuning specificity. Thus, the combination of tools might realize true off-target-free RNA editing in the future.

Methods

Cloning

Constructs in this work were cloned using either restriction cloning, golden gate assembly, or Gibson assembly. The efficiency reporter was cloned into a pCI vector (E1731, Promega). The dual fluorescent protein reporter (selectivity reporter) was cloned in a pBI-CMV1 vector (Clontech 631630), with ms2 tagged fluorescent protein (mRNA) version in one multiple cloning site and ms2 free version in the other¹¹. The DKCI constructs were assembled in pcDNA3.1 backbone via Gibson assembly. gRNAs were constructed by amplifying forward and reverse primers and cloned into a puc57 backbone under the control of a U6 promoter using a golden gate assembly. Circular gRNAs were cloned into a puc57 backbone containing P3 twister and P1 twister ribozymes using golden gate assembly.

Cell culture

HEK293T cells (ATCC CRL-3216) were maintained in Dulbecco’s modified Eagle’s medium (DMEM, Gibco 41965-039) supplemented with 1% penicillin-streptomycin (Sigma-Aldrich P0781), 1% L-glutamine (Sigma-Aldrich G7513), 1% sodium pyruvate (Life Technologies 11360), and 10% FBS (Sigma-Aldrich F7524). Cells were cultured at 37 °C in a 5% CO₂ atmosphere and passaged every 2–3 days up to 20 passages.

Flow cytometry

1.1×10^5 HEK293T cells were seeded in 24-well plates and transfected roughly 17 h later using jetPrime (Polyplus) as indicated by the manufacturer. Plasmids for organelle/gRNA/reporter constructs were transfected in a 1:1:1 ratio for experiments in Fig. 1b with a total of 1.2 μ g DNA per well. The media was exchanged after 4 h and cells were grown for additional 48 h. For flow cytometry measurements cells were washed once with PBS and trypsinized. Cells were resuspended in 900 μ l resuspension buffer 1 (1 \times PBS, 10% FBS, 2 mM sodium azide, 2 mM EDTA) and centrifuged for 5 min at 400 \times g at 4 °C. Supernatant was poured off and cells were resuspended in 900 μ l resuspension buffer 2 (1 \times PBS, 3% BSA, 2 mM sodium azide, 2 mM EDTA) and centrifuged again for 5 min at 400 \times g and 4 °C. Supernatant was again poured off and cells were resuspended in 300 μ l resuspension buffer 2 and cooled on ice until the measurement. Data was collected with a LSRFortessa SORP (BD Biosciences). DAPI (50 μ g/mL) was added before measurement to distinguish live from dead cells. 1.1×10^5 live cells were measured for each sample. Data was analyzed using FlowJo (v10.7.1.) HEK293T cells were distinguished based on FSC-A and SSC-A. Single cells were gated based on SSC-W and SSC-A. Live cells were distinguished based on SSC-W and 405–450/50 channels. For experiments using the efficiency reporter cells were then divided in four quadrants based on the reporter-only control (Supplementary Fig. 2). To estimate transfection efficiency from iRFP expressing cells, cells were gated based on SSC-F and 640–730/45 channel. Final GFP + percentages (Q2) were calculated by normalizing to iRFP + percentages of NES-DKCI. Normalized GFP + percentages (Q2) have been plotted in Figs. 1b and 3c. Data were plotted using GraphPad Prism software (v9.1.1). For experiments using the selectivity reporter eGFP fluorescence was acquired using a 488 nm laser and a 530/30 bandpass filter and for mCherry, a 561 nm laser with a 610/20 bandpass filter. Cells were gated similar to efficiency reporter experiments and quadrants were adjusted to a reporter-only control in Q4. For analysis of selectivity experiments a NOT gate for Q4 was created in FlowJo to collect total eGFP and mCherry intensities from Q1 to Q3. Relative fold changes in selectivity were calculated by dividing median mCherry intensities by median eGFP intensities. Calculated fold changes were normalized to NES-DKCI.

Microscopy

3×10^4 HEK293T cells were seeded in eight well ibidi dishes coated with poly-L lysine. Cells were transfected after 17 h using jetPrime (Polyplus). Therefore, 600 μ g of total DNA was added to 25 μ l JetPrime Buffer. DNA:jet Prime reagent ratio was 1:2. After 48 h cells were washed with

PBS and then fixed for 10 min at RT in 2% paraformaldehyde. Cells were once washed 2× with PBS. Cells were permeabilized with 0.5% Triton-X-100 in PBS for 5 min at RT and then washed 2× with PBS. Blocking solution (5% donkey serum in PBS/0.1% Tween-20) was added for 10 min at RT. First antibody (mouse anti-FLAG 1:500, 9A3, CellSignalling) in blocking solution was added overnight at 4 °C. Second antibody (donkey anti-mouse conjugated with Alexa Fluor™ 480, A-31570 Thermo Fisher Scientific) was added in blocking buffer for 1 h at RT. Cells were again washed 3× with PBS/0.1% Tween for 5–10 min and then stained with DAPI (0.5 µg/ml, 1:2000) in PBS for 5 min at RT. Afterwards, cells were washed 2× with PBS and kept at 4 °C until imaging. Cells were imaged using a LSM 900 microscope (Zeiss) using the Airyscan 2 setup. Images were processed using Fiji.

RT-PCR

RNA from HEK293T cells transfected with gRNA/circ gRNA constructs was extracted using TRizol (15596026, Thermo Fisher Scientific) and treated with DNaseI before performing reverse transcription. RNA was reverse transcribed using gene-specific primers and Maxima -H reverse transcriptase (EP0751, Thermo Fisher Scientific). One microliter of cDNA was used for a second PCR (using inverse Primers) and products were run on a 1.8% agarose gel.

Western blotting

5.5×10^5 HEK293T cells were seeded in 6well plates and transfected roughly 16 h later with a total of 6 µg DNA (1:1:1 ratio of plasmids encoding DKC1-constructs/reporters/gRNAs) using 12 µl jetPrime reagent (Polyplus) according to the manufacturer's instructions. Media was changed after 4 h and cells were grown for additional 48 h. Cells were harvested by adding 100 µl RIPA buffer (Chromotek) supplemented with cComplete protease inhibitor (Roche), 1 mM PMSF, 1 mM MgCl₂, and SM nuclease directly onto the 6well dish. The dishes were incubated for 10 min on ice. Afterward, the lysate was collected, thoroughly resuspended, and centrifuged for 10 min at 20,000 × *g* and 4 °C. Supernatant was collected and diluted with 150 µl dilution buffer supplemented with cComplete and 1 mM PMSF. Next 12.5 µl equilibrated myc trap beads (Chromotek) were combined with 225 µl lysate (25 µl were aliquoted and used for loading control) and incubated for at least 1 h at 4 °C on a rolling shaker. Pull-down and wash steps were performed according to the manufacturer's instructions. Protein was eluted by adding 40 µl 2× leammli buffer (120 mM Tris/Cl pH 6.8, 20% glycerol, 4% SDS, 0.04% bromophenol blue, 10% β-mercaptoethanol) to beads and incubating for 10 min at 95 °C. Supernatant was then subjected to SDS-PAGE on a 4–12% Bis-Tris gel (NuPAGE, Thermo Fisher Scientific). Proteins were transferred to a nitrocellulose membrane (Trans-Blot Turbo Midi 0.2 µm Nitrocellulose Transfer Packs, Bio-Rad) using Trans-Blot Turbo Transfer System (Bio-Rad). Membranes were submerged in blocking buffer (5% milk powder in TBST) on a rolling shaker at RT for 1 h. The first antibody (anti-myc, 1:1000, Cell Signaling Technology, 9B11) was incubated at 4 °C overnight. Blots were washed three times 5–10 min with TBST on a rolling shaker. Next blots were incubated with second antibody, anti-mouse HRP (1:10,000, Jackson ImmunoResearch, 715-035-150) in blocking buffer for 1 h at RT. Before imaging blots were again washed four times 5–10 min with TBST and then incubated with 2 ml ECL Western Blotting Detection Reagent (Cytiva) for 5 min. Blots were imaged on a ChemiDoc imaging system (BioRad) imager and band intensities were analyzed using Fiji.

RNA extraction

1.5×10^6 HEK293T cells were seeded in P10 dishes and transfected 17–18 h later. For transfection, 48 µl of jetPrime reagent were mixed with 24 µg total DNA (plasmids were kept 1:1:1 ratio) in 1 ml jetPrime Buffer. Medium was exchanged after 4 h and cells were grown for 48 h. Cells were washed 1× with PBS and then trypsinized. After neutralization with DMEM, cells were collected in 15 ml centrifugation tubes. Cells were centrifuged for 5 min at 400 × *g* at 4 °C. Supernatant was

poured off and cells were resuspended in 2 ml PBS. Cells were centrifuged again for 5 min at 400 × *g* at 4 °C. Supernatant was poured off and the cell pellet was flash-frozen in liquid nitrogen and stored at –20 °C for further analysis. Poly-A RNA was isolated using mRNA Direct Kit (Invitrogen, 61012) according to the manufacturer's instructions. After two sequential rounds of oligo dT isolation poly-A RNA was eluted in 20 µl EB supplied with the kit, flash frozen in liquid nitrogen, and stored until later use. Eluted RNA was split and one half was used for DRS and the other for BID-seq.

BID-seq

To assess the presence/absence of pseudouridine in oligo and polyA RNAs, a derived version of BID-Seq protocol²⁹ was performed. Briefly, 100 ng of RNA were subjected to RNA fragmentation followed by bisulfite treatment and desulphonation as described in ref. 30 Treated RNAs were end-repaired, purified, and subjected to NEBNext® small RNA library following manufacturer's protocol. The quality and quantity of each library were assessed using a high-sensitivity DNA Chip on a Bioanalyzer 2100 and a Qubit 2.0 fluorometer. High-throughput sequencing of the multiplexed libraries was performed on an Illumina NextSeq 2000 instrument in a 75 nt single-end read mode. Raw sequencing reads were inspected with FastQC and adapter sequence was removed by trimmomatic v0.39³¹. Alignment to the reference sequence was done by Bowtie2. v2.4.2³² with “relaxed” alignment stringency, allowing to retain 1–3 nt gapped reads. Further analysis was done by samtools mpileup utility and counting the deletions at every position in the reference.

Preparation and splinted ligation of RNA oligos

Custom 40 bp long RNA oligos (modified/unmodified at the targeted base) mimicking the target regions with our mCherry and eGFP reporters were ordered from Dharmacon. We ligated oligos in order to be long enough to be analyzed by DRS. RNA oligos were 5'-phosphorylated enabling them to be used in later ligations. Phosphorylation was performed using T4 Polynucleotide Kinase by NEB (M0201) following manufacturer instructions, incubating the reaction for 3 h @ 37 °C before terminating the reaction @ 65 °C for 20 min. Phosphorylated oligos were purified using Oligo Clean and Concentrator Kit by Zymo research (D4060) following manufacturer instructions. Purified oligos were ligated using T4 RNA Ligase 2 by NEB (M0239), combining equimolar amounts of oligos to be ligated, 1×T4 RNA Ligase Buffer, 10% PEG8000, 10 U T4 RNA Ligase 2 as well as a complementary cDNA splint ensuring correct order of ligation. Digestion of the cDNA splint was performed through DNase I by ThermoFisher Scientific (EN0525) following manufacturer's instructions and purified through Oligo Clean and Concentrator Kit. Poly(A) tailing of the purified ligation constructs was performed by Escherichia coli Poly(A) Polymerase by NEB (M0276) following manufacturer instructions and purified again using Oligo Clean and Concentrator Kit before proceeding to library preparation. Oligo1 consists of the modified target region of eGFP and the unmodified target region of mCherry. Oligo2 consists of the unmodified target region of eGFP and the modified target region of mCherry. Sequences are listed in Supplementary Table 5

Direct RNA Nanopore-seq library preparation

Library preparation was performed using Direct RNA Sequencing Kit (ONT, SQK-RNA004) following the manufacture's protocol (DRS_9195_v4_revB_20Sep2023). Briefly, 100 ng of poly(A)-tailed RNA or poly(A)-tailed constructs were ligated to ONT RT Adaptor (RTA) using T4 DNA Ligase (NEB, M0202T) and was reverse transcribed using SuperScript III RT (Thermo Fisher Scientific, 18080044). The products were purified using 1.8X Agencourt RNAClean XP beads (Thermo Fisher Scientific, NCO068576) and RNA adapter (RLA) was ligated onto the RNA:DNA hybrid. For PolyA sample, the mix was purified using 0.6X Agencourt RNAClean XP beads whereas for the ligated construct

the mix was purified using 2× beads. Both were washed twice using wash buffer. The samples were then eluted in elution buffer (EB) and mixed with Sequencing Buffer and Library Solution before loading onto primed RNA chemistry PromethION flow cells.

Preprocessing of direct RNA sequencing data

Raw direct RNA sequencing reads were base called using the Dorado base caller (v0.7.0) in super accuracy mode for direct RNA sequencing on RNA004 flowcells (Dorado: Oxford Nanopore Technologies, <https://github.com/nanoporetech/dorado>). Additionally, the recently integrated pseudouridine detection mode was executed. Subsequently, base-called reads were mapped onto the custom reference sequences for the two reporters (eGFP and mCherry) for the LCK-FUS²⁻²⁶⁷-MCP-DKCl, NES-DKCl, and Control samples and onto the custom oligo sequences for the synthetic oligo samples using minimap2 v2.28 with the suggested parameters for direct RNA reads: `-ax splice -uf -k142433`. The resulting bam files were sorted, indexed, and filtered by samtools v1.16 using the `-F 256` flag and specifying the regions of interest to be covered (Samtools: <https://doi.org/10.1093/gigascience/giab008>). Quality control of sequencing and mapping was done by NanoComp v1.23.125³⁴. All sequencing metrics can be found in Supplementary Table 4.

Pseudouridine detection using direct RNA sequencing reads

All subsequent analyses were executed in Python v3.10 using the filtered BAM files with the pysam package v0.22.1, a Python wrapper for HTSLib and Samtools (Pysam: <https://github.com/pysam-developers/pysam>)³⁵. For accuracy metrics, the match and mismatch percentages for each reference base across all runs were extracted using Pysam's `pileup` function with parameters `max_depth = 10,000` and `min_base_quality = 13` and plotted as heatmaps using the seaborn³⁵ v0.13.2 and matplotlib³⁶ v3.9.0 packages. All runs revealed a mean base accuracy >98%.

Previous studies on direct RNA sequencing have shown that pseudouridine modification sites can cause systematic basecalling errors, specifically uridine to cytosine mismatches using the older RNA002 flowcell. To verify whether this error persists with the newer RNA004 flowcell and chemistry, we sequenced two synthetic oligos containing identical 10-mer motifs to the target sites of the eGFP and mCherry reporters, with either 100% or 0% pseudouridine modification rates at one uridine position.

Using Pysam's `pileup` function with `max_depth = 100,000` to extract basecalling information, our positive control regions (100% modified) showed U > C mismatches in 88.3% of mapped bases for eGFP and 34.2% for mCherry specific motifs. The unmodified regions revealed U > C mismatches in only 3.9% and 1.3% of bases for eGFP and mCherry-specific motifs, respectively. These results indicate that despite advancements in direct RNA sequencing chemistry, flowcells, and basecalling models, the U > C mismatch remains indicative of pseudouridine modification, at least for the investigated motifs.

For reads that did not show a U > C mismatch, we used the newly developed Dorado pseU calling model to classify reads as modified or unmodified. Modification calling results from Dorado were extracted from the ML/MM tags in the BAM files, with probability values represented as $P = (p \cdot 256) - 1$, where p is the probability from 0.0 to 1.0. We used a probability threshold of $p > 0.95$ to extract only high-confidence sites. Combining reads with either a U > C mismatch or a Dorado classification, the synthetic oligo samples showed that the percentage of modified reads ranged between 96.6% and 91.8% for fully modified and 4.4% and 1.4% for completely unmodified sites of the eGFP and mCherry target motifs, respectively.

We then calculated the pseU percentage for both mCherry and eGFP motifs in LCK-FUS²⁻²⁶⁷-MCP-DKCl, NES-DKCl, and Control samples following the same approach. To remove standard error for specific sequence motifs in the pseudouridylated samples, we

subtracted the percentages of modified reads observed at the eGFP and mCherry target positions of the control sample from the pseudouridine percentages observed for LCK-FUS²⁻²⁶⁷-MCP-DKCl and NES-DKCl samples, respectively.

The U > C mismatch and Dorado-based modification percentages, both significant ($p < 0.95$) and insignificant ($p > 0.95$), were plotted for the sequence context -5 to +5 bases around the target positions of eGFP and mCherry in both the samples and synthetic oligos using ggplot2³⁷ v 3.4.4 in R v4.2.3 (Supplementary Figs. 7 and 8).

Statistical analysis

In each flow cytometry measurement data from 50,000 live cells were collected. Median values were used for further analysis as they are more robust in transient transfections where population distributions are skewed. For each experiment, unless noted otherwise, three independent replicates were performed. All statistical analysis was done using Graph Pad Prism. We performed an unpaired two-sided t -test. We note that those tests are not defined for very low sample numbers and are thus used here instead as an indicator of substantial difference.

Reporting summary

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

Data availability

The sequencing data generated in this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession code [PRJEB76145](https://www.ebi.ac.uk/ena/record/PRJEB76145). The raw data generated in this study are provided in the Source Data file. Source data are provided with this paper.

Code availability

All scripts are available in the following GitHub repository: https://github.com/AnWiercze/Pseudouridine_detection_Schartel_et_al_2024.

References

1. Delaunay, S., Helm, M. & Frye, M. RNA modifications in physiology and disease: towards clinical applications. *Nat. Rev. Genet.* **25**, 104–122 (2023).
2. Kierzek, E. et al. The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Res.* **42**, 3492–3501 (2014).
3. Cerneckis, J., Cui, Q., He, C., Yi, C. & Shi, Y. Decoding pseudouridine: an emerging target for therapeutic development. *Trends Pharmacol. Sci.* **43**, 522–535 (2022).
4. Anderson, B. R. et al. Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. *Nucleic Acids Res.* **38**, 5884–5892 (2010).
5. Karjilovich, J. & Yu, Y.-T. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* **474**, 395–398 (2011).
6. De Boeck, K., Zolin, A., Cuppens, H., Olesen, H. V. & Viviani, L. The relative frequency of CFTR mutation classes in European patients with cystic fibrosis. *J. Cyst. Fibros.* **13**, 403–409 (2014).
7. Adachi, H. et al. Targeted pseudouridylation: an approach for suppressing nonsense mutations in disease genes. *Mol. Cell* **83**, 637–651.e9 (2023).
8. Song, J. et al. CRISPR-free, programmable RNA pseudouridylation to suppress premature termination codons. *Mol. Cell* **83**, 139–155.e9 (2023).
9. Rintala-Dempsey, A. C. & Kothe, U. Eukaryotic stand-alone pseudouridine synthases—RNA modifying enzymes and emerging regulators of gene expression? *RNA Biol.* **14**, 1185–1196 (2017).
10. Garus, A. & Autexier, C. Dyskerin: an essential pseudouridine synthase with multifaceted roles in ribosome biogenesis, splicing, and telomere maintenance. *RNA* **27**, 1441–1458 (2021).

11. Reinkemeier, C. D., Girona, G. E. & Lemke, E. A. Designer membraneless organelles enable codon reassignment of selected mRNAs in eukaryotes. *Science* **363**, eaaw2644 (2019).
12. Reinkemeier, C. D. & Lemke, E. A. Dual film-like organelles enable spatial separation of orthogonal eukaryotic translation. *Cell* **184**, 4886–4903.e21 (2021).
13. Adachi, H. et al. Targeted pseudouridylation: a novel approach for suppressing nonsense mutations in disease genes. *Mol. Cell.* **83**, 637–651.e9 (2021).
14. Litke, J. L. & Jaffrey, S. R. Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. *Nat. Biotechnol.* **37**, 667–675 (2019).
15. Yi, Z. et al. Engineered circular ADAR-recruiting RNAs increase the efficiency and fidelity of RNA editing in vitro and in vivo. *Nat. Biotechnol.* **40**, 946–955 (2022).
16. Katrekar, D. et al. Efficient in vitro and in vivo RNA editing via recruitment of endogenous ADARs using circular guide RNAs. *Nat. Biotechnol.* **40**, 938–945 (2022).
17. Ngo, L. H. et al. Nuclear export of circular RNA. *Nature* **627**, 212–220 (2024).
18. Daigle, N. & Ellenberg, J. N-GFP: an RNA reporter system for live-cell imaging. *Nat. Methods* **4**, 633–636 (2007).
19. Jain, M., Abu-Shumays, R., Olsen, H. E. & Akeson, M. Advances in nanopore direct RNA sequencing. *Nat. Methods* **19**, 1160–1164 (2022).
20. Leger, A. et al. RNA modifications detection by comparative Nanopore direct RNA sequencing. *Nat. Commun.* **12**, 7198 (2021).
21. Furlan, M. et al. Computational methods for RNA modification detection from nanopore direct RNA sequencing data. *RNA Biol.* **18**, 31–40 (2021).
22. Begik, O. et al. Quantitative profiling of pseudouridylation dynamics in native RNAs with nanopore sequencing. *Nat. Biotechnol.* **39**, 1278–1291 (2021).
23. Tavakoli, S. et al. Semi-quantitative detection of pseudouridine modifications and type I/II hypermodifications in human mRNAs using direct long-read sequencing. *Nat. Commun.* **14**, 334 (2023).
24. Martins-Dias, P. & Romão, L. Nonsense suppression therapies in human genetic diseases. *Cell. Mol. Life Sci.* **78**, 4677–4701 (2021).
25. Jülicher, F. & Weber, C. A. Droplet physics and intracellular phase separation. *Annu. Rev. Condens. Matter Phys.* **15**, 237–261 (2024).
26. Zhao, X., Bartolucci, G., Honigmann, A., Jülicher, F. & Weber, C. A. Thermodynamics of wetting, prewetting and surface phase transitions with surface binding. *N. J. Phys.* **23**, 123003 (2021).
27. Cox, D. B. T. et al. RNA editing with CRISPR-Cas13. *Science* **358**, 1019–1027 (2017).
28. Merkle, T. et al. Precise RNA editing by recruiting endogenous ADARs with antisense oligonucleotides. *Nat. Biotechnol.* **37**, 133–138 (2019).
29. Dai, Q. et al. Quantitative sequencing using BID-seq uncovers abundant pseudouridines in mammalian mRNA at base resolution. *Nat. Biotechnol.* **41**, 344–354 (2023).
30. Zhang, L.-S. et al. BID-seq for transcriptome-wide quantitative sequencing of mRNA pseudouridine at base resolution. *Nat. Protoc.* **19**, 517–538 (2024).
31. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
32. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
33. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
34. De Coster, W. & Rademakers, R. NanoPack2: population-scale evaluation of long-read sequencing data. *Bioinformatics* **39**, btad311 (2023).
35. Waskom, M. Seaborn: statistical data visualization. *JOSS* **6**, 3021 (2021).
36. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
37. Wickham, H. *Ggplot2* 2nd edn, (Springer International Publishing Cham, 2016).

Acknowledgements

We thank the flow cytometry and microscopy core facilities of the IMB and JGU for their assistance. We thank all members of the Lemke lab for helpful discussions. We acknowledge funding of the ERC ADG Multi-OrganelleDesign and the Volkswagen Stiftung. We also acknowledge funding from the SFB1551 “Polymer concepts on Cellular function” of the Deutsche Forschungsgemeinschaft (DFG project number 464588647).

Author contributions

L.S. and E.A.L. conceived the project. L.S. did flow cytometry, western blotting, and microscopy experiments, T.B. prepared the RNA library, S.M. prepared synthetic RNA oligos for sequencing, and A.W. performed bioinformatic analysis. C.J., M.H., and S.G. provided analytical tools and reagents. V.M. and Y.M. performed BID-seq. L.S. and E.A.L. co-wrote the manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

L.S. and E.A.L. have filed a patent related to the manuscript. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-53403-1>.

Correspondence and requests for materials should be addressed to Edward A. Lemke.

Peer review information *Nature Communications* thanks Alexander Hilditch, Xiaocheng Weng, and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024