Published online 30 September 2021

Nucleic Acids Research, 2022, Vol. 50, No. 1 ***e3***
https://doi.org/10.1093/nar/gkab867

Ethylenediamine derivatives efficiently react with oxidized RNA 3′ ends providing access to mono and dually labelled RNA probes for enzymatic assays and *in vivo* translation

Adam Mamot[1,2], Pawel J. Sikorski[1], Aleksandra Siekierska[3], Peter de Witte[3], Joanna Kowalska[2,*] and Jacek Jemielity[1,*]

[1]Centre of New Technologies, University of Warsaw, Banacha 2c Street, 02-097 Warsaw, Poland, [2]Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Pasteura 5 Street, 02-093 Warsaw, Poland and [3]Laboratory for Molecular Biodiscovery, KU Leuven, Campus Gasthuisberg, Herestraat 49, O&N II, 3000 Leuven, Belgium

Received June 15, 2021; Revised August 31, 2021; Editorial Decision September 14, 2021; Accepted September 15, 2021

ABSTRACT

Development of RNA-based technologies relies on the ability to detect, manipulate, and modify RNA. Efficient, selective and scalable covalent modification of long RNA molecules remains a challenge. We report a chemical method for modification of RNA 3′-end based on previously unrecognized superior reactivity of *N*-substituted ethylenediamines in reductive amination of periodate-oxidized RNA. Using this method, we obtained fluorescently labelled or biotinylated RNAs varying in length (from 3 to 2000 nt) and carrying different 5′ ends (including m^7G cap) in high yields (70–100% by HPLC). The method is scalable (up to sub-milligrams of mRNA) and combined with label-facilitated HPLC purification yields highly homogeneous products. The combination of 3′-end labelling with 5′-end labelling by strain-promoted azide-alkyne cycloaddition (SPAAC) afforded a one-pot protocol for site-specific RNA bifunctionalization, providing access to two-colour fluorescent RNA probes. These probes exhibited fluorescence resonance energy transfer (FRET), which enabled real-time monitoring of several RNA hydrolase activities (RNase A, RNase T1, RNase R, Dcp1/2, and RNase H). Dually labelled mRNAs were efficiently translated in cultured cells and in zebrafish embryos, which combined with their detectability by fluorescent methods and scalability of the synthesis, opens new avenues for the investigation of mRNA metabolism and the fate of mRNA-based therapeutics.

INTRODUCTION

Functionalized and labelled RNA molecules are invaluable tools to study RNA function by advanced biophysical and biological methods. RNA can be modified in an enzymatic reaction with a catalytic nucleic acid (DNAzyme or RNAzyme) or a chemical reaction (1). Maintaining high selectivity, efficiency, product purity, and stability, are common concerns of both users and developers of these methods. The problems escalate as the molecular size of the target RNA increases. As such, post-transcriptional modification of messenger RNA (mRNA) is challenging and often leads to a mixture of substrates and products (2–5). However, fluorescent labelling of mRNA plays crucial role in developmental and structural studies, as well as investigation of gene expression, cellular immune responses, and delivery of mRNA based therapeutics (6–10), thus creating the drive to improve and expand the current toolbox of chemoenzymatic RNA modification. Low efficiency, limited substrate specificity, and high cost of upscaling are common limitations among methods based on reactions catalysed by proteins or nucleic acids. Direct chemical RNA modification is an efficient alternative to enzymatic approaches. Unfortunately, only few chemical methods that go beyond the scope of solid-phase nucleic acids synthesis have been established (11–14). One of the most unique methods for direct chemical RNA modification is based on periodate ring-opening oxidation reaction. During the reaction, 2′, 3′-cis diol (vicinal cis-diol) moiety present in RNA molecules (within the 3′ terminal ribose) is transformed into an acyclic 1,5-dialdehyde derivative, which can be subsequently functionalized by a reaction with an appropriate N-nucleophile. In the earliest studies (around seventy years ago), periodate oxidation was used for the analysis of 3′ end structure of

*To whom correspondence should be addressed. Tel: +48 22 55 43774; Fax: +48 22 55 43771; Email: j.jemielity@cent.uw.edu.pl
Correspondence may be also addressed to Joanna Kowalska. Email: jkowalska@fuw.edu.pl

isolated RNA (15–17). It was later found that in presence of primary amines and hydrazides, the 1,5-dialdehyde is prone to amination reaction, leading to formation of imines (Schiff bases) (18–20). Since then, this phenomenon was utilized for coupling of periodate-oxidized RNA with amine or hydrazide derivatives of fluorescent dyes, biotin, nucleic acids, proteins, and resins (Supplementary Table S1). Overall, two different approaches have been used for the subsequent reaction of oxidized RNA with N-nucleophiles. The first approach involves the reaction of dialdehyde with a hydrazine derivative and its subsequent isolation (13,18,21–26). However, this type of modification is reversible and the isolated imines are unstable. To address this problem, another approach has been developed involving the reductive amination (i.e. amination in presence of reducing agents such as sodium borohydride or sodium cyanoborohydride) of oxidized RNA, which leads to formation of a stable morpholine derivative product (19,20,27–34). When primary amines or their derivatives were used, the yield of reductive amination was significantly higher than that of amination (20,28,31,34). Interestingly, while reviewing these studies, we found a great diversity of conditions applied during the periodate oxidation and reductive amination steps (pH (5–10), reaction times (0.3 h–7 days), and temperature (0–37°C) concentration of periodate (0.3 mM–1 M), length (1–350 nt) and concentration (5 μM–100 mM) of the RNA substrate, and the type and concentration of the N-nucleophile). Consequently, the yields (if reported) also varied significantly (26–99%). Although parameters affecting the reaction course and yield were optimized in some of those studies to some extent, relative reactivity of different N-nucleophiles, impact of RNA length and concentration, and other key factors were never systematically investigated.

Thus, in this study, we revisited the RNA 3′ end modification method relying on 3′-*cis*-diol oxidation and subsequent reductive amination to develop reliable and scalable RNA modification protocols that are adaptable to RNAs of different lengths (including mRNAs). We developed efficient and scalable 3′ end modification protocols, which could be combined with concomitant RNA 5′ end labelling to obtain either mono- or dually labelled mRNAs. Moreover, we demonstrated the usefulness of the dual-modification methodology to construct functional RNA-based FRET probes, enabling monitoring of various enzymatic activities. Finally, we examined the translational properties of mono- and dually labelled mRNAs and demonstrated their potential in studying mRNA localization and expression in human cells and zebrafish embryos.

MATERIALS AND METHODS

General information

Detailed information concerning chemical synthesis, experiments with GMP-dial and pU_3, RNA sequences, preparation of DNA templates for *in vitro* transcription, *in vitro* transcription procedure, chemical labelling and biotinylation of RNA, HPLC purification and FRET experiments are provided in the Supplementary Data.

Table 1. RNA labelling conditions and HPLC-determined conversions (%)

			HPLC conversion[a]		
			Labeling type		
RNA length (nt)	Final RNA concentration (μM)[b]	Cond.	3′[c]	5′[d]	5′+3′[e]
3	100–1000	A	85–99	nd	nd
3	10–100	B	75–85	nd	nd
30–300	1–30	B	75–85	73–95	74–88
1000–2000	0.2–3	B	75–85	74–88	46–56

[a]HPLC conversions were calculated by dividing the peak area of the labelled product by sum of peak areas of substrate and reaction product(s). Peak areas were measured for absorbance at 260 nm.
[b]To determine molar concentration, absorbance at 260 nm of RNA substrate solutions were measured and divided by individual molar extinction coefficient calculated for each RNA sequence. Concentration of long RNAs (>300 nt) should not exceed 3 μM for efficient labelling.
[c]RNA containing 3′ ribose is oxidized (1.5 mM $NaIO_4$, 30 min at 25°C in dark) and subsequently subjected to reductive amination in presence of an R-EDA derivative (2 h incubation at 25°C). Conditions A: 10 mM R-EDA, 100 mM $NaBH_3CN$, 500 mM KH_2PO_4 pH 6.0; Condition B: 1 mM R-EDA, 20 mM $NaBH_3CN$, 100 mM KH_2PO_4 pH 6.0.
[d]RNA containing 5′ azide moiety is subjected to SPAAC reaction (incubation with 2 mM DIBAC derivative for 2 h at 25°C).
[e]RNA containing 5′ azide and 3′ ribose moieties is oxidized and subsequently subjected to simultaneous SPAAC and reductive amination reaction (Conditions B + 2 mM DIBAC derivative). nd - not determined.

Cy3 3′-labelling of RNA and mRNA

To a solution of RNA (12.0 μl, form 30 nt to 2000 nt, at concentration form 0.3 μM to 167 μM, for details see Table 1), fresh solution of sodium periodate ($NaIO_4$, 2.00 μl, 10 mM, Chempur) was added, followed by incubation in dark at 25°C for 30 min. Next, potassium hydrogen phosphate buffer (KH_2PO_4, 2.00 μl, 1.00 M pH 6.0), solution of Cy3-EDA (2.00 μl, 10 mM), and fresh solution of sodium cyanoborohydride ($NaBH_3CN$ 2.00 μl, 200 mM, Sigma) were added and the reaction mixture (20.0 μl total volume) was incubated at 25°C for 120 min in the dark. The mixture was diluted with water (160 μl) and sodium acetate (20 μl, 3.0 M, pH 5.2), and ethanol (0.60 ml, 99.8%) were added (with mixing after every addition). The solution was cooled for 30 min at −80°C. The pellet was centrifuged (30 min, 14 000 g, 4°C), washed with ethanol (0.80 ml, 80%), centrifuged (10 min, 14 000 g, 4°C), dried *in vacuo* at room temperature, and suspended in water. The concentration of the resulting solution was measured using Nanodrop. The sample was purified by HPLC, and the collected fractions were concentrated and analysed with agarose or PAGE electrophoresis. Fractions containing desired products (such as ppp-RNA_{35}-Cy3, ppp-RNA_{237}-Cy3, N_3-m^7GRNA_{237}-Cy3, ppp-RNA_{276}-Cy3, N_3-m^7GRNA_{276}-Cy3, ppp-RNA_{gluc}-Cy3, ARCA-RNA_{gluc}-Cy3, N_3-m^7GRNA_{gluc}-Cy3, N_3-m^7GRNA_{egfp}-Cy3, ARCA-RNA_{rluc}-Cy3, N_3-m^7GRNA_{rluc}-Cy3 and ARCA-RNA_{fluc}-Cy3) were combined and applied in experiments featured in Figures 3, 5, 6 and Supplementary Figures S7, S11, S14-S17, S21-S24.

Cy5 and Cy3 5′,3′-dual labelling of RNA and mRNA

To a solution of RNA containing 5′ azide modification (10.0 μl, form 30 nt to 2000 nt, at concentration form 0.4 μM to 200 μM, for details see Table 1) fresh solution of $NaIO_4$ (2.00 μl, 10 mM, Chempur) was added, followed by incubation in the dark at 25°C for 30 min. Next, KH_2PO_4 buffer (2.00 μl, 1.00 M pH 6.0), solution of sulfo-Cyanine5 DBCO (2.00 μl, 20 mM, Lumiprobe) in 50% DMSO, solution of Cy3-EDA (2.00 μl, 10 mM), and fresh solution of $NaBH_3CN$ (2.00 μl, 200 mM, Sigma) were added and the reaction mixture (20.0 μl total volume) was incubated at 25°C for 120 min in the dark. The dually labelled RNA was isolated by subsequent precipitation, HPLC purification, concentration, and gel analysis as described above for Cy3 3′-labelling. Fractions containing desired products (such as Cy5-RNA_{35}-Cy3, Cy5-m^7GRNA_{35}-Cy3, Cy5-m^7GRNA_{237}-Cy3, Cy5-m^7GRNA_{276}-Cy3, Cy5-m^7GRNA_{gluc}-Cy3, Cy5-m^7GRNA_{egfp}-Cy3 and Cy5-m^7GRNA_{rluc}-Cy3) were combined and applied in experiments featured in Figures 3, 5, 6 and Supplementary Figures S14-S16, S19-20, and S22-24.

HPLC purification of RNA

All HPLC separations were performed on modular, low-pressure gradient HPLC apparatus equipped with thermostated column holder or column oven, DAD detector, and fluorescence detector. Chromatograms were recorded with detection of absorbance at 260 nm, absorbance spectrum (220–700 nm), and fluorescence of interest (Cy3 – 550/565, Cy5 – 650/665, FAM – 490/520). RNAs were resolved using Phenomenex Clarity 3 μm OligoRP C18 150 × 4.6 mm column (for 30–300 nt RNAs) or RNASept^{TM}Prep C18 50 × 7.8 mm 2 μm (for 1000–2000 mRNAs) using mobile phase solvents A (100 mM TEAA pH 7.0) and B (200 mM TEAA pH 7.0 / MeCN 1:1) at 50–55°C.

Enzyme activity monitoring using RNA FRET probes

A solution of dually labelled RNA-FRET probe Cy5-RNA_{35}-Cy3 (50.0 μl, 100 nM) in buffer (4 mM Tris–HCl pH 7.5, 15 mM NaCl, 0.1 mM EDTA) was first subjected to refolding/folding procedure, by applying a step-gradient of temperature (5 min at 95°C then 95–25°C over 1 h, −5°C/4 min). The solution was then diluted with degassed buffer (150 μl, 4 mM Tris–HCl pH 7.5, 15 mM NaCl, 0.1 mM EDTA) and transferred to a quartz fluorescence cuvette (1 × 1 × 350 mm). Emission spectra were recorded on a Cary Eclipse spectrofluorometer (Agilent), equipped with a xenon lamp (excitation at 500 nm, emission 510–800 nm, 10 mm slit) at 5°C. After the spectrum stabilized (5–15 min) the enzyme was added: RNase A (1.00 μl, 10 ng/ml, Thermo), RNase T1 (1.00 μl, 10 U/μl, Thermo), RNase R (1.00 μl, 10 U/μl, ABM). Emission spectra were recorded, and samples for PAGE analysis were taken at specified time points.

For Dcp1/2 assay probe Cy5-m^7GRNA_{35}-Cy3 was used. Before transferring to a quartz fluorescence cuvette, the solution of the probe (40.0 μl, 100 nM) was diluted with degassed buffer (150 μl, 4 mM Tris–HCl pH 7.5, 15 mM NaCl, 6.5 mM $MgCl_2$). After the spectrum stabilized *S. pombe* Dcp1/2 complex (10 μl, 7 μM) was added (43).

For RNase H assay probes Cy5-RNA_{35}-Cy3 and Cy5-m^7GRNA_{276}-Cy3 were used. Prior refolding/folding procedure, solution of the probe (36.0 μl, 100 nM) was supplemented with complementary DNA sequence of choice (1.6 μl, 5 μM, 1.2 eq, CTTCCCTTGATCGG for Cy5-RNA_{35}-Cy3, TGCTCGGGGTCGTACACCTT, TC ATTTGCTTGCAGCGAGCC, or CGTGATATCTCTCC CGTGCCTCCACAGGTA for Cy5-m^7GRNA_{276}-Cy3). Emission spectra were recorded at 35°C. After the spectrum stabilized RNase H (2.00 μl, 0.1 mg/ml) was added (44).

CNOT7 deadenylase assay

60 nanograms of an RNA probe (N_3-m^7GRNA_{237}, N_3-m^7GRNA_{237}-Cy3, Cy5-m^7GRNA_{237}, Cy5-m^7GRNA_{237}-Cy3) was dissolved in buffer (10 mM Tris–HCl pH 8, 5 mM $MgCl_2$, 10 mM DTT, 50 mM KCl), and incubated in presence of CNOT7 deadenylase (45) (0.2 mg/ml) at 37°C for 30 min (final reaction volume 5 μl). The reaction was quenched by addition of loading dye and thermal denaturation (5 min 65°C). The products were resolved using 8% denaturing PAGE.

Translation in rabbit reticulocyte lysates

The reaction mixture (8 μl) contained reticulocyte lysate (4 μl, Promega), amino acid mixture without leucine (25 μM, Promega), amino acid mixture without methionine (25 μM Promega), potassium acetate (250 mM), and $MgCl_2$ (1.25 mM). After 1 h of incubation at 30°C, 2 μl of the appropriate *Renilla* luciferase-coding mRNA solution (0.64, 0.51, 0.41, 0.33, 0.16 or 0.00 ng/μl) was added and the incubation of the reaction mixture was continued at 30°C for 1 h. The reaction was stopped by freezing in liquid nitrogen. Next, 50 μl of 10 ng/ml h-coelenterazine (NanoLight) in PBS was added to 10 μl of the lysate and the luminescence was measured on Synergy H1 (BioTek) microplate reader. The luminescence was plotted as a function of mRNA concentration and analysed using linear regression model. The linear regression coefficient (slope) for each replicate was normalized to the value determined for ARCA-RNA_{rluc} to obtain relative translation efficiency parameter. The presented values are mean of tree independent experiments ± standard error (SEM).

Expression of mRNA in cultured cells studied by luminometry

HeLa (human cervical epithelial carcinoma, ATCC CCL-2) cells were grown in DMEM (Gibco) supplemented with 10% FBS (Sigma), GlutaMAX (Gibco) and 1% penicillin/streptomycin (Gibco) at 5% CO_2 and 37°C. For all experiments the passage number of studied cells was between 5 and 25. In a typical experiment, on the day of transfection 104 HeLa cells were seeded in 100 μl medium per well of 96-well plate. Cells in each well were transfected for 16 h using a mixture of 0.3 μl Lipofectamine MessengerMAX Transfection Reagent (Invitrogen) and 25 ng *Gaussia* luciferase-encoding mRNA (ppp-RNA_{gluc}, ppp-RNA_{gluc}-Cy3, VCE-RNA_{gluc}, VCE-RNA_{gluc}-Cy3, ARCA-RNA_{gluc},

ARCA-RNA$_{gluc}$-Cy3, N_3-m^{7}GRNA$_{gluc}$, N_3-m^{7}GRNA$_{gluc}$-mock, N_3-m^{7}GRNA$_{gluc}$-Cy3, Cy5-m^{7}GRNA$_{gluc}$, or Cy5-m^{7}GRNA$_{gluc}$-Cy3) in 10 μl Opti-MEM (Gibco). In order to assess *Gaussia* luciferase expression at multiple time points, medium was fully removed and replaced with the fresh one at each time point. To detect luminescence, 50 μl of 10 ng/ml h-coelenterazine (NanoLight) in PBS was added to 10 μl of cell cultured medium and the luminescence was measured on Synergy H1 (BioTek) microplate reader. Total protein expression (cumulative luminescence) for each mRNA over 4 days was reported as a mean value ± SD normalized to ARCA-capped mRNA.

Flow cytometry studies

HeLa cells were cultured as above. In a typical experiment, 24 h before transfection 5×10^5 cells were seeded in 3 ml medium per well of 6-well plate. Cells in each well were transfected for 17 h using 9 μl Lipofectamine MessengerMAX Transfection Reagent, 0.75 μg mRNA encoding eGFP (N_3-m^{7}GRNA$_{egfp}$, N_3-m^{7}GRNA$_{egfp}$-Cy3, Cy5-m^{7}GRNA$_{egfp}$, or Cy5-m^{7}GRNA$_{egfp}$-Cy3) in 300 μl Opti-MEM. After transfection, medium was removed, cells were washed with PBS and subjected to trypsinization. Detached cells were analysed using LSR Fortessa flow cytometer with FACSDiva software (BD Biosciences). Data was analysed using FlowJo software v10 (Tree Star).

Time-lapse fluorescent imaging of mRNA in living cells

HeLa cells were cultured as above. In a typical experiment 24 h before transfection, 2×10^4 cells were seeded in 200 μl medium per well of 8-well chambered coverglass. Cells were transfected using 0.6 μl Lipofectamine MessengerMAX Transfection Reagent and 50 ng of mRNA encoding eGFP (N_3-m^{7}GRNA$_{egfp}$, N_3-m^{7}GRNA$_{egfp}$-Cy3, Cy5-m^{7}GRNA$_{egfp}$ or Cy5-m^{7}GRNA$_{egfp}$-Cy3) in 20 μl Opti-MEM. One hour after the start of the transfection, images started to be acquired. Time-lapse images were constantly acquired at 21 min 2 s intervals for 21 h. Cells were imaged using Olympus Fluoview FV10i laser scanning microscope, using a 60x/1.2 water objective. eGFP, Cy3 and Cy5 emission were detected at emission spectra of 490–590, 570–670 and 660–760 nm, respectively, after extrication at 473 nm for EGFP, 559 nm for Cy3, and 635 nm for Cy5. Data was analysed using ImageJ software.

Microinjection and fluorescent imaging of mRNA in zebrafish embryos

Adult zebrafish (*Danio rerio*) of the AB strain were maintained at 28.5°C on a 14-h light/10-h dark cycle under standard aquaculture conditions. Fertilized eggs were collected via natural spawning. Embryos were raised in embryo medium, containing 1.5 mM HEPES, pH 7.6, 17.4 mM NaCl, 0.21 mM KCl, 0.12 mM $MgSO_4$ and 0.18 mM $Ca(NO_3)_2$, in an incubator on a 14-h light/10-h dark cycle at 28.5°C. For all experiments described, larvae at 0–28 h post fertilization (hpf) were used. All experiments performed at the University of Leuven were approved by the Ethics Committee of the University of Leuven (approval number 150/2015) and by the Belgian Federal Department of Public Health, Food Safety and Environment (approval number LA1210199). 300 pg of mRNA encoding eGFP (N_3-m^{7}GRNA$_{egfp}$, N_3-m^{7}GRNA$_{egfp}$-Cy3, Cy5-m^{7}GRNA$_{egfp}$ or Cy5-m^{7}GRNA$_{egfp}$-Cy3) resuspended in sterile RNA free water was injected into one-cell stage AB strain zebrafish embryos (1-nl volume) with glass capillaries (WPI, TW100F-4) pulled with a micropipette puller (Sutter Instruments) using a M3301R Manual Micromanipulator (WPI) and a FemtoJet 4i pressure microinjector (Eppendorf). Before imaging, injected zebrafish embryos of 8 and 28 hpf were dechorionated, anesthetized with 0.4 mg/mL tricaine and immobilized in 0.1% agarose on a cover glass. Confocal microscopy images were recorded with a Zeiss LSM 780 – SP Mai Tai HP DS confocal microscope equipped with an LD LCI Plan Apo 25×/0.8 objective. Cy5, Cy3 and EGFP markers were excited at 633, 561 and 488 nm, respectively. The images were visualized with ZEN Lite software and ImageJ. The supplementary microscopy images of whole embryos were taken using Leica MZ10F microscope with a Leica DFC310 FX digital colour camera and Leica Application Suite LASV 4.13 software.

RESULTS AND DISCUSSION

Ethylenediamine shows superior reactivity in reductive amination of oxidized nucleotides

We began our investigation using guanosine monophosphate (GMP) as a model molecule, hoping to gain insight into ribose transformations and dialdehyde reactivity. Aqueous GMP (50 mM) was quantitatively converted into a dialdehyde derivative GMP-dial (Figure 1) upon incubation with potassium periodate (65 mM). The dialdehyde was precipitated with acetone, resuspended in water (at 1 mM), and subjected to reductive amination. We tested several N-nucleophiles, including monoamines, diamines, thioamines, aminoalcohols, and hydrazine (each at the concentration of 1 or 2 mM, i.e. 2 equivalents of NH_2 group per GMP-dial) in the presence of sodium cyanoborohydride (10 mM $NaBH_3CN$) in phosphate buffer (KH_2PO_4 pH 6.0) at 30°C. Reaction progress was monitored by HPLC at 20 min intervals and UV-absorbing products were analysed by mass spectrometry (MS, Figure 1). To our surprise, little or no product formed after 40 min of reaction in majority of the tested nucleophiles (i.e. methylamine, butylamine, ethanolamine, cysteine, cystine, cystamine, and 1,4-diaimnobutane). In contrast, cysteamine, hydrazine, and ethylenediamine (EDA) reacted with GMP-dial robustly. Incubation of GMP-dial with hydrazine resulted in partial conversion into the desired morpholine-containing product (∼50% after 60 min, Supplementary Figure S1B, compound **IV**). Reaction with cysteamine led to a different product, presumably containing a thiazolidine ring (Supplementary Figure S1A, compound **III**). Only during reaction with EDA, the GMP-dial was almost completely converted to the desired product GMP-EDA (Figure 1, S2). These findings encouraged us to identify the cause of superior reactivity of EDA and further optimize the reac-

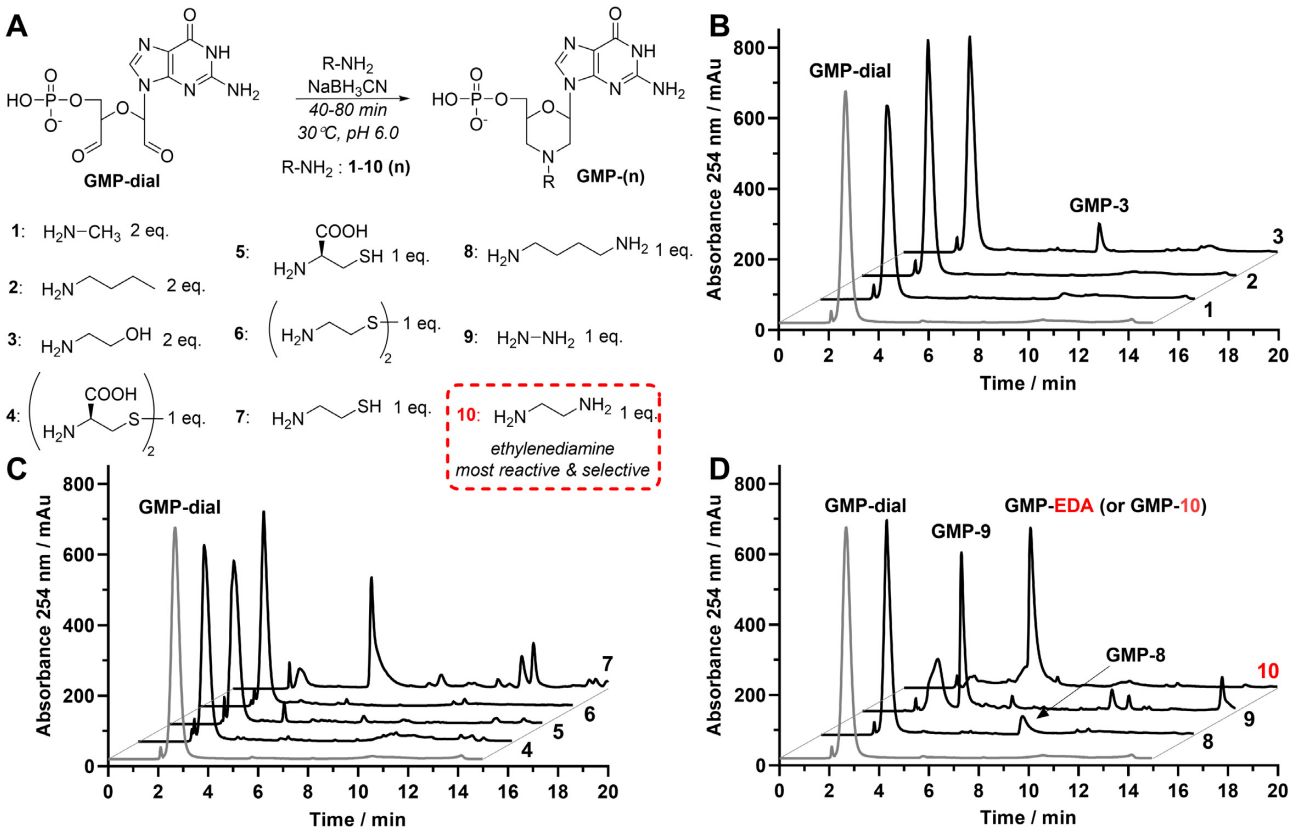

Figure 1. Reductive amination of GMP-dial with different N-nucleophiles (**1**–**10**). (**A**) Reaction scheme with general structure of reductive amination products GMP-(n). (**B**–**D**) HPLC chromatograms of crude reaction products after 40 min (**1**–**6**, **8**) or 60 min (**7**, **9**, **10**) after addition of the nucleophile. During reaction between GMP-dial and butylamine (**2**) or cysteamine (**7**) indicated morpholine products (GMP-2 and GMP-7) were not detected. Product of reductive amination between GMP-dial and ethylenediamine (**10**, EDA) is designated as GMP-10 or GMP-EDA.

tion conditions. Reactions between GMP-dial (1 mM) and EDA (1 mM) were carried out at different pH (4.5, 6.0, and 8.0) and $NaBH_3CN$ concentrations (1 and 10 mM, Supplementary Figure S2). At pH 4.5, GMP-dial converted selectively into GMP-EDA. The reaction at pH 4.5 proceeds faster after increasing $NaBH_3CN$ concentration, maintaining high selectivity. If pH was elevated from 4.5 to 6, the yields and the rate of the reactions further increased (GMP-dial conversions after 60 min in the presence of 1 mM $NaBH_3CN$, ∼30% vs. ∼90%; and ∼50% vs. ∼100% at 10 mM $NaBH_3CN$). At pH 8, conversion of GMP-dial was similar to that at pH 6; however, instead of the desired GMP-EDA product, partially reduced imine derivative was formed in significant amounts (Supplementary Figure S2, compound **IV**). Overall, we established that reductive amination of GMP-dial with EDA at pH 6 in the presence of 10 mM $NaBH_3CN$ yielded the most promising results in the context of RNA labelling. Interestingly, at a lower $NaBH_3CN$ concentration (1 mM), two isomeric intermediates, represented by two partially overlapping peaks (Rt = 4.5 min and 4.7 min, Supplementary Figure S2), formed in a pH-dependent ratio. Monoisotopic mass of the isomers corresponded to imine or imidazolidine intermediates (Supplementary Figure S2, compounds **II** and **III**), with structure analogous to the product of reaction between GMP-dial and cysteamine (Supplementary Figure S1A, compound **III**). Notably, in contrast to the reaction between GMP-dial and hydrazine, no morpholinodiol intermediate (Supplementary Figure S1B; compound **II**) was observed during the reaction with EDA. Therefore, we hypothesized that superior reactivity of EDA towards GMP-dial arises from the intramolecular formation of imidazolidine derivative, which both stabilizes the imine intermediate and facilitates the nucleophilic attack on the adjacent aldehyde (Supplementary Figure S3). Both formation and reactivity of imidazolidine ring under such conditions is consistent with the previously reported mechanisms of reactions between aldehydes and amino acids (35,36) or 1,5-dialdehydes with tris (hydroxymethyl) aminomethane (37).

N-Substituted ethylene diamines (R-EDAs) efficiently modify oxidized RNA 3′-ends

We envisaged that the proposed mechanism for reductive amination of GMP-dial (Supplementary Figure S3) does not preclude similar reactivity for the *N*-substituted EDA derivatives (R-EDA). In such a scenario, an extension of EDA motif by functional groups or labels would provide a robust approach for RNA 3′ end labelling. To test this possibility, uridine trinucleotide 5′-monophosphate (pU_3) and *N*-propargyl ethylenediamine (PEDA) were used as model compounds (Figure 2 and Supplementary Figure S4). Periodate-mediated oxidation of pU_3 occurred rapidly, but the obtained dialdehyde pU_3-dial was prone to decom-

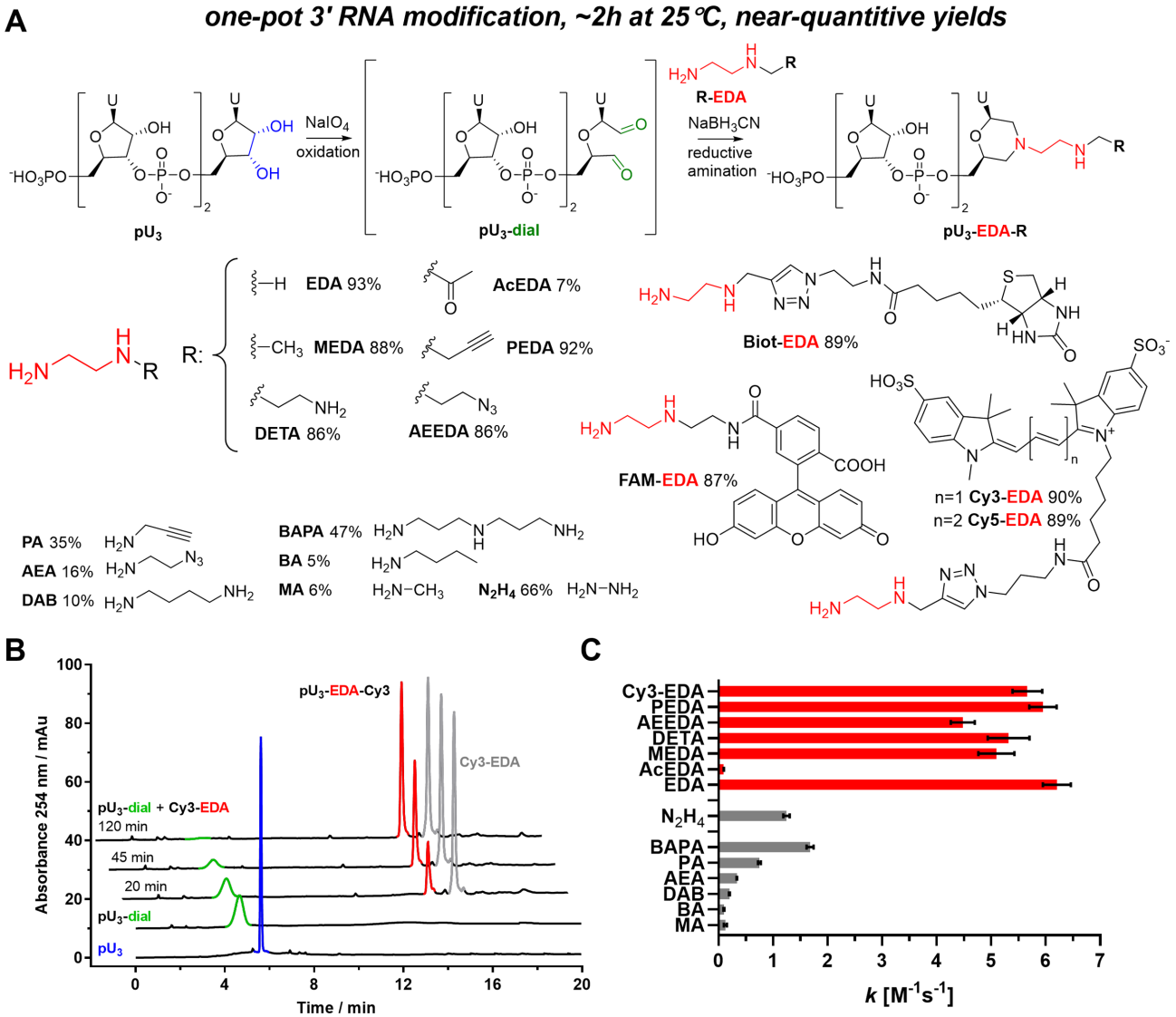

Figure 2. One-pot 3'-end RNA labelling. (**A**) Reaction scheme. Conditions: i) pU_3 (140 μM), aqueous $NaIO_4$ (1.4 mM), 30 min, and 25°C in the dark. ii) Addition of R-EDA (1.0 mM) and $NaBH_3CN$ (20 mM) in KH_2PO_4 buffer (0.1 M pH 6.0). HPLC-determined yields after two hours of reductive amination are presented for each N-nucleophile. (**B**) HPLC chromatograms of reaction with Cy3-EDA at different stages. (**C**) Second-order reaction rate constant k values determined for reductive amination reaction between pU3-dial (100 μM) and different N-nucleophiles (1 mM).

position upon isolation. To minimize the number of RNA handling operations and isolation of the RNA-dial intermediate, we designed the labelling protocol as a one-pot, two-step procedure (Figure 2A and Supplementary Figure S4A). The oxidation of pU_3 (4 mM) was complete within 20 min at 30°C in KH_2PO_4 buffer (pH 6.0) over solid KIO_4 (~30 mM), and even milder conditions of oxidation were sufficient, if pU_3 was diluted to 20 μM (20 min at 25°C in 0.5–1.5 mM KIO_4; Supplementary Figure S4D). To find the optimal conditions of reductive amination, the product of the oxidation, pU_3-dial, was diluted with buffer (to a final concentration of 10, 100, or 1000 μM) and subsequently reacted with PEDA (0.1, 1.0, or 10 mM) in the presence of $NaBH_3CN$ (5, 50, or 100 mM; Supplementary Figure S4; Supplementary Table S2). If the concentration of PEDA was 1 mM or higher, the desired product, pU_3-EDA-P (Supplementary Figures S4A), formed selectively and efficiently (68–99% yield by HPLC). At 0.1 mM of PEDA, the side reactions of elimination (to pU_2p) and reduction (to pU_3-diol) drastically lowered the yields of pU_3-EDA-P (6–15%; Supplementary Figure S4A). Considering these findings, we formulated two variants of the labelling protocol. The first protocol, optimal for millimolar (0.1–1 mM) pU_3 concentrations, consisted of an oxidation step (30 min of incubation in 1.0–1.5 mM periodate at 25°C) followed by a reductive amination step (incubation at 25°C for 2 h), initiated by addition of R-EDA (10 mM), $NaBH_3CN$ (200 mM), and KH_2PO_4 buffer (500 mM pH 6). The second protocol, suitable for pU_3 concentration below 100 μM, required milder conditions of reductive amination (1 mM R-EDA, 20 mM $NaBH_3CN$, 100 mM KH_2PO_4 pH 6). These protocols allowed us to transform pU_3 into pU_3-EDA-R products, con-

taining 3′-amine, -azide, -alkyne, -sulfo-cyanine3 (Cy3), -sulfo-cyanine5 (Cy5), -biotin or -carboxyfluorescein (FAM) moieties, with high yields (75–99%; Figure 2 and Supplementary Figure S5). To compare R-EDA derivatives with other N-nucleophiles, several reactions were performed according to the optimized protocol (for micromolar RNA concentration). Progress of these reactions was monitored by HPLC, which allowed us to calculate second-order reaction rate constants (Figure 2C and Supplementary Figure S6; Supplementary Table S3). The rate constants were significantly lower for all tested aliphatic amines ($k = 0.1$–$1.7\ M^{-1}s^{-1}$) and hydrazine ($1.25 \pm 0.06\ M^{-1}s^{-1}$) than for EDA ($6.15 \pm 0.27\ M^{-1}s^{-1}$) and R-EDA derivatives (4.5–$6.0\ M^{-1}s^{-1}$). As an exception, the *N*-acetyl ethylenediamine (AcEDA) proved to be one of the least reactive nucleophiles ($k = 0.10 \pm 0.01\ M^{-1}s^{-1}$). These findings further emphasize the superior reactivity of the EDA motif (consisting of two nucleophilic nitrogen atoms separated by C2-linker) towards oxidized RNA 3′ ends and support the proposed mechanism of the reductive amination with R-EDA.

Labelling with R-EDA provides easy access to 3′-labelled short and long IVT RNAs

High molecular weight RNA, such as mRNA, can be synthesized exclusively by means of in vitro transcription (IVT). Thus, further development of our method was devoted to 3′ modification of IVT RNA (Figure 3A). 5′-triphosphorylated 35 nt RNA (ppp-RNA_{35}) was subjected to labelling with Cy3-EDA. Crude products were isolated from the reaction mixture with high recovery (90–100%) by precipitation with ethanol. The products were analysed using RP-HPLC in ion-pair mode under denaturing conditions (C18 column, triethylammonium acetate/acetonitrile as mobile phase; 50–55°C) with UV detection at 260 nm and fluorescence detection at 550/565 nm (Supplementary Figure S7A). The retention times of labelled and unlabelled RNAs differed significantly, allowing us to quantify the labelling yields based on HPLC peak areas. Labelling of ppp-RNA_{35} (10 μM) was efficient at 1 mM Cy3-EDA (82%) and the increase of the dye concentration up to 4 mM had little impact on the reaction yield (77–81%, Supplementary Figure S7). Encouraged by the result, we increased the molecular size of the RNA substrate and sought conditions of its efficient labelling. 5′-triphosphorylated 276 nt RNA (ppp-RNA_{276}) was subjected to 3′ Cy3 labelling at various pH (5.5, 6.0, 6.5, 7.0, or 7.6), temperatures (10, 20, 30, or 40°C), and incubation times (1 h or 3 h; Supplementary Figure S8). Additionally, to determine the extent of nonspecific RNA degradation, mock labelling reactions were performed in parallel, during which the periodate oxidation was omitted (Supplementary Figure S9). From all 80 tested conditions, best results were observed if incubation was carried at 20°C for 3 h or at 30°C for 1 h (Supplementary Figure S10). We concluded that reductive amination is optimally performed at pH 6.0 and 25°C for 120–140 min. Finally, RNA and mRNA sequences (random or protein coding), ranging between 35 and 2098 nt, were synthesized and subjected to Cy3 labelling at the 3′ end. The yields of all performed reactions were quantified by HPLC, as even in the case of longest tested mRNA, the unlabelled and labelled entities had different chromatographic mobility (Figure 3B). The reaction yields ranged between 83 and 85% for shorter (35–276 nt) and 75–84% for longer (993–2089 nt) RNA at concentration ranges 1–30 μM, and 0.2–2 μM, respectively (Table 1). Degradation of RNA molecules was not observed in either case (Figure 3, Supplementary Figures S11, S16). Similar results were obtained if the protocol was applied to 3′ biotinylation of RNA (Figure 3D, Supplementary Figure S12).

The labelling protocol based on R-EDA is adaptable to 3′-labelling of 5′ capped IVT mRNA and site-selective 5′,3′-dual labelling of mRNA

The site-selectivity of the 3′ end labelling relies on the fact that only the 3′ terminal nucleoside of the RNA molecule contains a *cis*-diol group. This is indeed true in the case of both endogenous and *in vitro* transcribed prokaryotic RNAs, including mRNAs, which usually contain a triphosphate moiety at the 5′ end. However, eukaryotic mRNA is modified at the 5′ end by addition of 7-methylguanosine cap, which also contains a *cis*-diol group. Therefore, eukaryotic (capped) mRNA is prone to periodate oxidation at both 5′ and 3′ ends. This can be beneficial, as it enables one-step bifunctionalization of mRNA molecules, but poses a problem if 3′-end selective labelling is required. Hence, we demonstrated two ways in which the proposed protocol can be adjusted to provide access to 5′-capped and 3′-labelled IVT mRNA. One option is to obtain 5′-triphosphorylated IVT product, subject it to selective labelling at the 3′ end, and subsequently add 5′ cap in a posttranscriptional manner (e.g. using *Vaccina* capping enzyme, VCE; Supplementary Figure S11). Alternatively, transcription can be performed in the presence of a cap analogue bearing a modification at either the 2′ or 3′ position of 7-methylguanosine (Supplementary Figures S11 and S21) to produce 5′ mRNA caps that are not susceptible to oxidation. Examples of such caps are commercially available anti-reverse cap analogue (ARCA), which is as dinucleotide cap 0 analogue ($m_2{}^{7,3'\text{-}O}$GpppG) (38,39), or a trinucleotide cap 1 analogue ($m_2{}^{7,3'\text{-}O}$Gppp$^{m2'\text{-}O}$ApG), which is used in the production of BioNTech/Pfizer mRNA COVID-19 vaccine (40). To demonstrate this possibility, we performed Cy3 labelling of 5′ ARCA-capped mRNA encoding firefly luciferase (Supplementary Figure S11). Furthermore, in an earlier study, we demonstrated that IVT reaction with cap analogues functionalized at the ribose of 7-methylguanosine with azido-linker, such as N_3-m^7GpppG (Figure 3A) yields RNA containing 5′ azide group, which can be utilized for labelling of translationally active mRNA in living eukaryotic cells using strain-promoted azide-alkyne cycloaddition (SPAAC) (41). Hence, we envisaged that both of our labelling methods are mutually orthogonal and can be used for site-specific two-colour labelling of RNA. Using previously developed cap analogue N_3-m^7GpppG or a newly designed dinucleotide mimicking uncapped RNA, N_3-AG (Figure 3A and Supplementary Figure S13) as transcription initiators, we obtained capped (N_3-m^7GRNA) and uncapped (N_3-RNA) RNA mimics containing a 5′ azido group. The RNAs were subsequently subjected to conditions in which only 5′ end was fluorescently labelled with

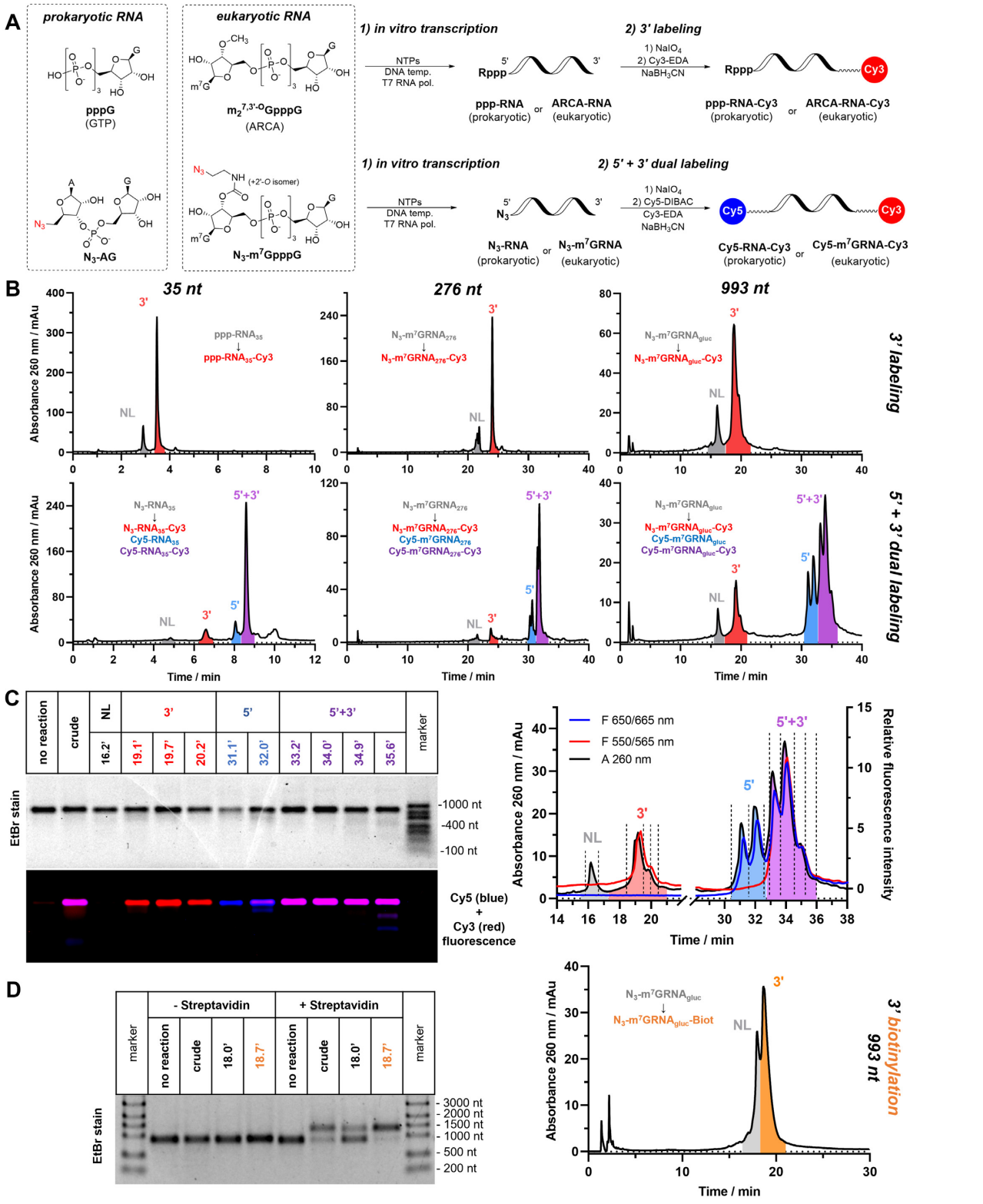

Figure 3. Mono- and dual labelling of IVT RNA. (**A**) Scheme of the process: NTPs mix in the presence or absence of initiating dinucleotide (ARCA, N_3-AG or N_3-m^{7}GpppG) are incubated with gene-coding DNA template in presence of T7 RNA polymerase. After isolation from IVT mixture, RNA products are subjected to 3′ labelling (top) or simultaneous 5′+3′ dual labelling (bottom). (**B**) Representative HPLC chromatograms of crude labelling products of Cy3 3′ labelling (top) or Cy5/Cy3 5′+3′ dual labelling (bottom) performed on 35, 276 and 993 nucleotide-long RNAs (ppp-RNA$_{35}$, N_3-RNA$_{35}$, N_3-m^{7}GRNA$_{276}$ and N_3-m^{7}GRNA$_{gluc}$). (**C**) Absorbance and fluorescence HPLC profiles of crude dual labelling product Cy5-m^{7}GRNA$_{gluc}$-Cy3 with designated retention times of collected fractions (right) and agarose electrophoresis of concentrated HPLC fractions (left); (**D**) HPLC chromatogram of crude 3′-biotinylated mRNA N_3-m^{7}GRNA$_{gluc}$-Biot (right) and electrophoretic mobility shift assay (EMSA) of crude and HPLC-isolated products of 3′ biotinylation (left).

Table 2. Yields of HPLC-purified RNA

	Isolated yield (%)[a]			
	Labelling type[b]			
RNA length (nt)	NL[c]	3′	5′	5′+3′
30–300	53–63	nd	nd	37–58
1000–2000	53–63	34–50	32–53	26–34

[a]Yield of HPLC isolation was measured by dividing total optical density (OD, 260 nm absorbance of solution multiplied by its volume) of collected HPLC fractions (after concentration) by total OD of the injected sample.
[b]RNA after labelling reaction (as in Table 1) was isolated by precipitation in ethanol or using commercially available purification kits. Crude product was resolved using HPLC accordingly to its length, number, and type of modifications (see supplementary experimental section for details). HPLC fractions containing desired product were concentrated by precipitation in isopropanol or freeze-drying.
[c]NL: recovery of unlabelled RNA after HPLC purification. nd, not determined.

Cy5, 3′ end was labelled with Cy3, or both 5′ and 3′ ends were simultaneously labelled with Cy5 and Cy3, respectively. Products were isolated from the reaction mixtures and subsequently analysed by HPLC (Figure 3B-C, Supplementary Figures S12, S14–S16). HPLC retention times of different RNA molecules depended mostly on the structure and number of modifications, rather than RNA length itself. In most cases, addition of Cy5-DIBAC to the 5′ end of RNA almost doubled the retention time and caused peak splitting, which we attributed to formation of two regioisomers of substituted triazole during SPAAC reaction (Figure 3B-C, Supplementary Figures S12, S16). Yields of labelling at the 5′ and 3′ ends were 73–95% and 70–85%, respectively. During simultaneous labelling of 5′ and 3′ ends, both chemical reactions remained highly efficient, which resulted in formation of dually labelled RNA with yields ranging between 74–88% for shorter (35–276 nt) and 46–56% for longer (993–2089 nt) products (Table 1). The labelled products were isolated by HPLC, to yield highly homogenous mono- or dually labelled mRNA probes with isolated yields of 26–53% (Table 2). The whole RNA labelling protocol was very robust, requiring about two days to complete the whole procedure (RNA in vitro transcription and isolation, 3–5 h; labelling reaction and crude product isolation, 3–4 h; and HPLC separation and fraction work-up, 1 day).

The mRNA 3′-end labelling procedure is scalable

One of the advantages of chemical reactions over enzymatic reactions is that they do not require expensive reagents and are easier to up-scale. To demonstrate this, we performed an up-scaling experiment of 3′ Cy3 labelling of mRNA. The ppp-RNA_{egfp} mRNA (5–79 μg) was processed by the labelling protocol at different concentrations (0.4–6.3 μM in 22 μl of reaction mixture) and isolated by precipitation in ethanol. The labelling yields were assayed by HPLC, as previously mentioned (Supplementary Figure S17, Supplementary Table S4). We found that labelling reactions remained efficient (75–84%) below 3.2 μM concentration of mRNA (1.8 μg/μl). Further increase of mRNA concentration was met with a decrease of labelling yield (down to 52% at 6.3 μM). This effect appears to be specific for large mRNA molecules (Table 1 and S4) and could be potentially caused by molecular crowding and increased viscosity of concentrated mRNA solutions. Equipped with this knowledge, we preformed labelling of 500 μg of ppp-RNA_{egfp} in a 440 μl reaction mixture. After ethanol precipitation we recovered 376 μg (75%) of crude product, that contained 77% (∼290 μg) of 3′-labeled mRNA (Supplementary Figure S18).

5′,3′-Dual labelling of RNA affords FRET probes enabling monitoring of various RNA hydrolases

It is hypothesized that eukaryotic translation occurs with mRNA circularization, due to mRNA binding proteins and mRNA secondary structure, during which the 5′ and 3′ ends are in proximity (42). The RNA secondary structure and end-to-end distance may have significant contribution in many other biological processes; however, their investigation is limited due to the lack of appropriate experimental techniques. Since Cy3 and Cy5 dyes have overlapping emission and absorption bands, we investigated fluorescence properties of three dually labelled RNAs differing in length and capping status (Cy5-RNA_{35}-Cy3, Cy5-m^7GRNA_{35}-Cy3, and Cy5-m^7GRNA_{276}-Cy3, Figure 4A) in the context of Förster Resonance Energy Transfer (FRET). All tree RNA probes demonstrated strong Cy5 emission (665 nm) upon excitation of Cy3 at 500 nm, which indicates that FRET occurs in these molecules (Figure 4B, C and Supplementary Figure S19B, C). Although the predicted (with RNAfold web server) secondary structures of their sequences did not indicate any direct contact of terminal nucleosides (i.e. by base-pairing with two adjacent nucleosides, Figure 4A), the observation of FRET suggests that 5′ and 3′ ends are in proximity. This finding is consistent with recently reported work on secondary structures of various RNA sequences, which showed that the average distance between 5′ and 3′ ends is smaller than 10 nm, regardless of RNA length (2). We next tested if we can utilize the FRET phenomenon observed for the RNA probes to monitor enzyme activity in real-time. To that end, select dually labelled RNA probes were hydrolysed by several nucleolytic enzymes. We expected that the hydrolysis of RNA would lead to a loss of RNA secondary structure and consequently to a decrease of the FRET signal (i.e. decrease of the Cy5 emission accompanied by an increase of Cy3 emission). This approach was tested with six different hydrolases differing in substrate specificity and reaction mechanism: RNase A, RNase T1, RNase H1 from *E. coli* (endonucleases), RNase R (3′ exonuclease), decapping complex Dcp1/2 from *S. pombe* (Nudix hydrolase), and CCR4-Not transcription complex subunit 7 (CNOT7) deadenylase. Rapid and complete decrease of the FRET signal was observed for uncapped Cy5-RNA_{35}-Cy3 probe after addition of RNAses A, T1, and R, indicating complete hydrolysis of the probe (Figure 4B). Specificity of the signal loss was confirmed by treating the probe with RNase A in the presence of a selective inhibitor (RiboLock; Figure 4B). The decapping complex Dcp1/2 hydrolyses the triphosphate bridge in mRNA cap to release 7-methylguanosine 5′-diphosphate (m^7GDP) (43). To monitor the activity of Dcp1/2 we used an m^7G-capped probe (Cy5-m^7GRNA_{35}-

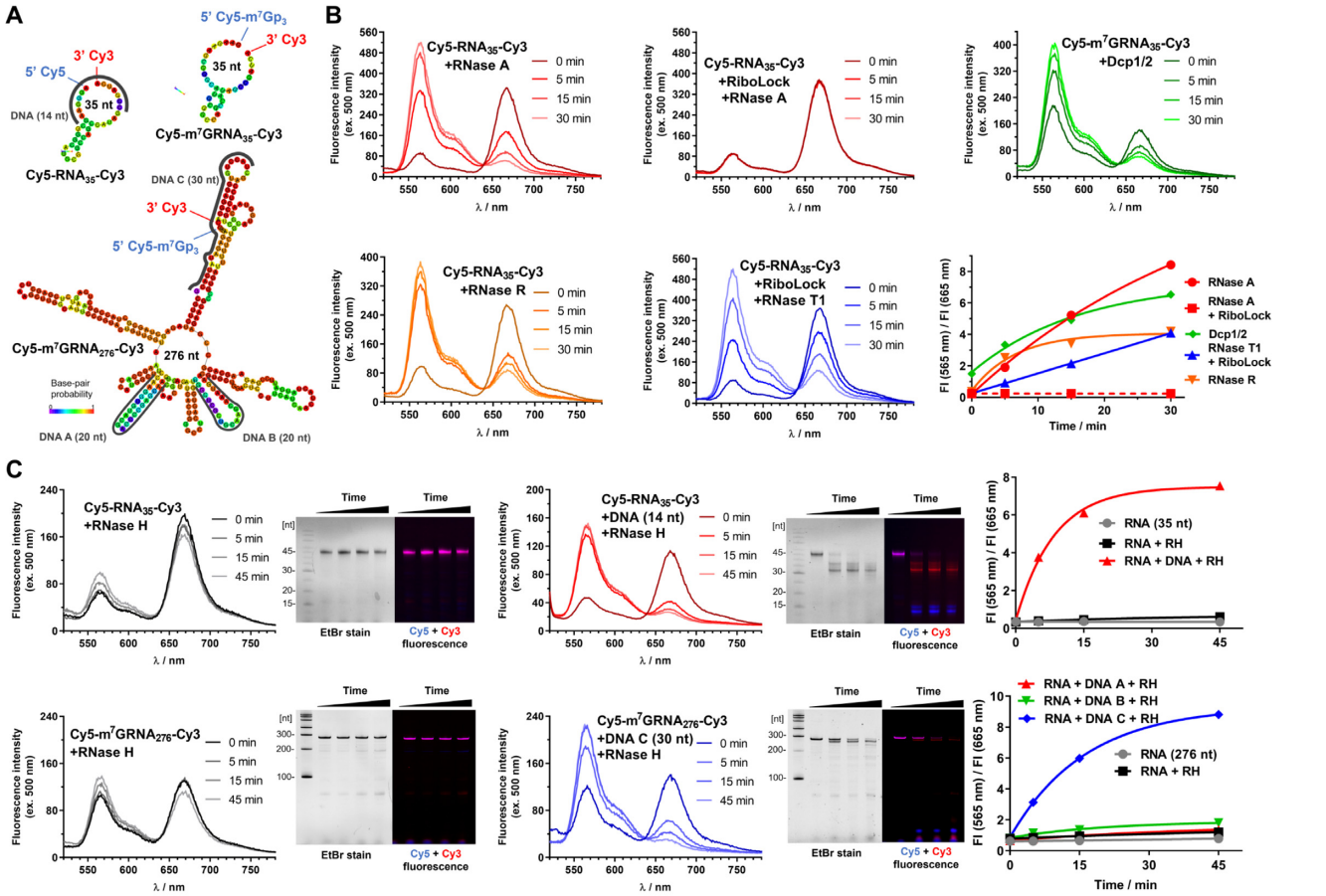

Figure 4. Application of dually labelled RNA probes for enzymatic activity monitoring. (**A**) Sequence and theoretical secondary structures of Cy5-RNA$_{35}$-Cy3, Cy5-m^{7}GRNA$_{35}$-Cy3 and Cy5-m^{7}GRNA$_{276}$-Cy3 probes and illustration of studied DNA sequences complementary to regions of the probes. (**B**) Time-dependent changes of emission spectra and Cy3/Cy5 fluorescence intensity ratio after addition of RNase A, RiboLock and RNase A, RNase T1, RNase R or Dcp1/2 to Cy5-RNA$_{35}$-Cy3 or Cy5-m^{7}GRNA$_{35}$-Cy3 probes. (**C**) Time-dependent changes of emission spectra, PAGE of reaction products, and Cy3/Cy5 fluorescence intensity ratio after addition of RNase H to probe (Cy5-RNA$_{35}$-Cy3 or Cy5-m^{7}GRNA$_{276}$-Cy3) or DNA-probe duplex.

Cy3). The hydrolysis of the probe was observed by the fluorescence spectroscopy and additionally confirmed by gel electrophoresis of the reaction products (Figure 4B and Supplementary Figure S19A). RNase H1 targets RNA in RNA–DNA heteroduplexes (44). To test RNAse H1 activity using our probes, we first designed several DNA complementary sequences and investigated their influence on fluorescence properties of the probe (Figure 4A). Interestingly, hybridization of the DNA did not interfere with the FRET signal unless it was fully complementary to the RNA sequence (Figure 4C and Supplementary Figure S19B). If the annealed DNA was complementary to both 5′ and 3′ RNA terminal sequences (creating a DNA splint), the fluorescence changes correlated with the reaction progress (Figure 4C). RNase H cleaved Cy5-RNA$_{35}$-Cy3 and Cy5-m^{7}GRNA$_{276}$-Cy3 probes at 5′ splinted site. If the Cy5-m^{7}GRNA$_{276}$-Cy3 probe was annealed with DNA that targeted internal, theoretically less structured regions, minor fluoresce changes were observed, while polyacrylamide gel electrophoresis (PAGE) of the reaction products shows complete cleavage of the probe (Figure 4C, Supplementary Figure S19C). We speculate in the latter case, the elements of RNA secondary structure that contribute to approximation of 5′ and 3′ ends remain intact despite internal RNA cleavage. In order to determine if the 3′ end modification inhibits deadenylation by CNOT 7, fluorescent RNA probes containing a 3′ polyA (N$_3$-m^{7}GRNA$_{237}$, N$_3$-m^{7}GRNA$_{237}$-Cy3, Cy5-m^{7}GRNA$_{237}$, Cy5-m^{7}GRNA$_{237}$-Cy3) were prepared. In agreement with previous findings (2), the addition of 3′ polyA disrupted the FRET signal, which compelled us to apply conventional electrophoretic assays (45). After CNOT7 treatment fluorescent RNA probes were resolved using denaturing PAGE (Supplementary Figure S20). We found that 3′ Cy3 modified RNAs were not susceptible to deadenylation by CNOT7 under conditions ensuring complete deadenylation of unmodified RNAs.

Mono- and dually labelled fluorescent mRNAs are efficiently translated in cells and enable RNA visualisation

Next, we assessed the impact of the chemical processing and modifications on biological properties of mRNA. mRNA encoding *Gaussia* luciferase containing 5′ azido-cap (N$_3$-m^{7}GRNA$_{gluc}$) was fluorescently labelled with Cy3 and Cy5 dyes (yielding N$_3$-m^{7}GRNA$_{gluc}$-Cy3, Cy5-m^{7}GRNA$_{gluc}$, and Cy5-m^{7}GRNA$_{gluc}$-Cy3 mRNAs). After

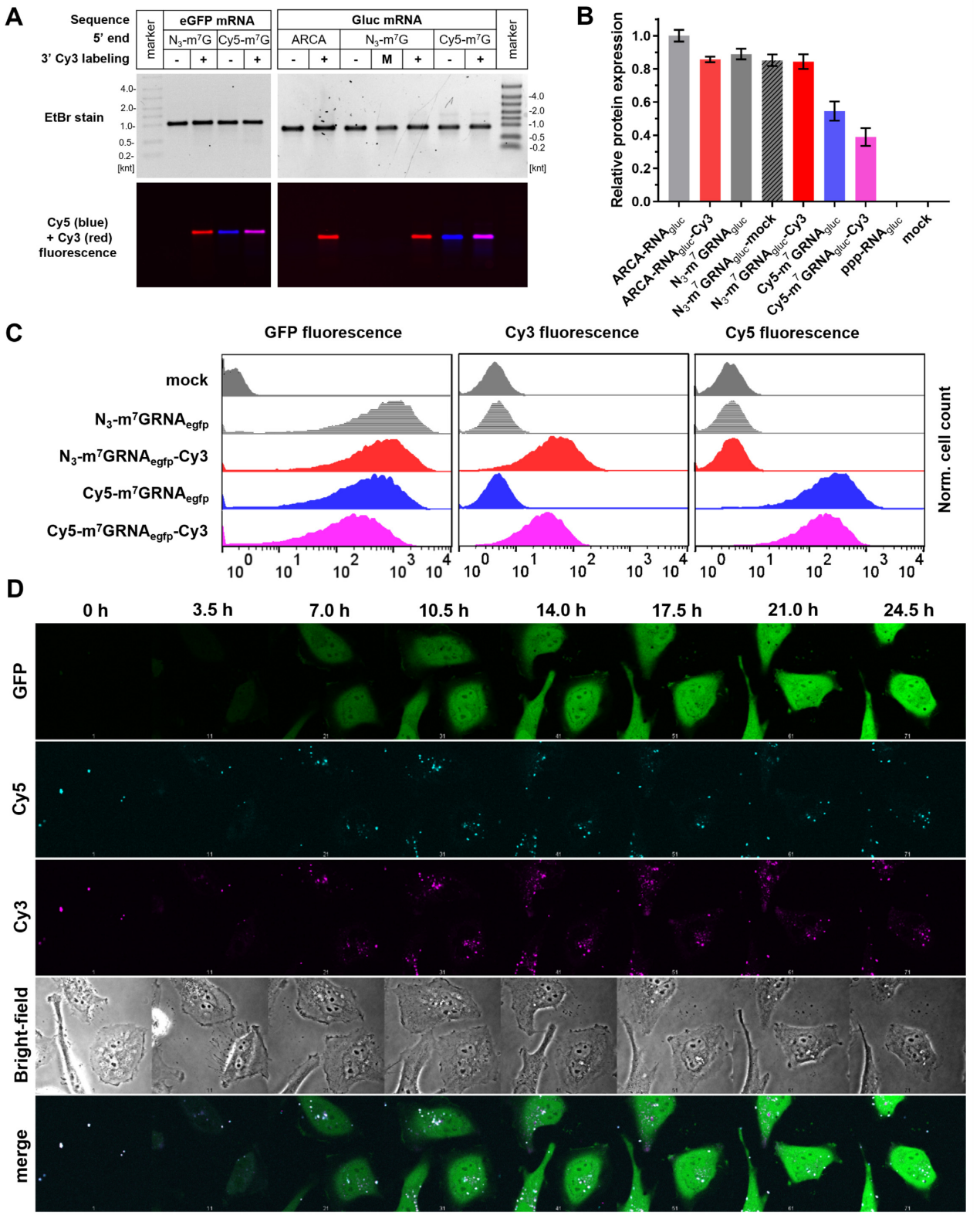

Figure 5. Fluorescent mRNA transfection into HeLa cells. (**A**) Electrophoretic resolution of HPLC-purified mRNAs encoding eGFP (N_3-m^{7}GRNA$_{egfp}$, N_3-m^{7}GRNA$_{egfp}$-Cy3, Cy5-m^{7}GRNA$_{egfp}$ and Cy5-m^{7}GRNA$_{egfp}$-Cy3) or *Gaussia* luciferase (ARCA-RNA$_{gluc}$, ARCA-RNA$_{gluc}$-Cy3, N_3-m^{7}GRNA$_{gluc}$, N_3-m^{7}GRNA$_{gluc}$-mock, N_3-m^{7}GRNA$_{gluc}$-Cy3, Cy5-m^{7}GRNA$_{gluc}$ and Cy5-m^{7}GRNA$_{gluc}$-Cy3) used for HeLa transfections. (**B**) Relative expression of *Gaussia* luciferase after mRNA transfection as a function of the presence of 3′ Cy3 modification and 5′ cap structure. (**C**) Flow cytometry readouts after transfection of fluorescent mRNA encoding eGFP. (**D**) Time-lapse microscopy images of HeLa cells transfected with Cy5-m^{7}GRNA$_{egfp}$-Cy3 mRNA. Time scale set for 0 h at the beginning of recording of the images (1 h after transfection start).

HPLC isolation, fluorescent mRNAs were transfected into HeLa cells and protein expression was measured (Figure 5). We also assessed the translational activity for ARCA-capped and post-transcriptionally capped (using VCE) mRNAs (Supplementary Figure S21). All 3′ labelled mRNAs were translationally active. Importantly, 3′-end modifications did not hamper protein expression, suggesting that the labelling procedure does not compromise mRNA integrity. Protein expression was also observed when cells were transfected with mock labelled mRNA (N_3-m^7GRNA_{gluc}-mock), that was a product of a labelling reaction lacking the oxidation step. This further indicates that chemical modification undergoes mildly and selectively, without compromising the integrity of mRNA body. As expected, based on our previous observations (41), 5′ cap labelling of mRNA slightly decreased overall protein expression (ca. 40%). We next evaluated whether fluorescent mRNA is suitable for the detection by flow cytometry. Additionally, to confirm cap-dependent mechanism of translation for both 5′ and 3′-modified mRNAs, translation of *Renilla* luciferase-coding mRNAs was measured in rabbit reticulocyte lysates (Supplementary Figure S22). While 5′ triphosphorylated and ApppG-capped mRNAs (ppp-RNA_{rluc} and ApppG-RNA_{rluc}) were virtually untranslated, both fluorescently labelled and unlabelled mRNAs (ARCA-RNA_{rluc}, ARCA-RNA_{rluc}-Cy3, N_3-m^7GRNA_{rluc}, N_3-m^7GRNA_{rluc}-Cy3 and Cy5-m^7GRNA_{rluc}-Cy3) maintained similar translation efficiency.

Next, expression of a different reporter, eGFP, was studied in HeLa cells. After mRNA transfection (N_3-m^7GRNA_{egfp}-Cy3, Cy5-m^7GRNA_{egfp} or Cy5-m^7GRNA_{egfp}-Cy3), cells were examined for GFP fluorescence and for Cy3 and Cy5 fluorescence arising from labelled mRNA (Figure 5C). To our delight, flow cytometry analysis not only demonstrated robust detection of eGFP expression, but also precise recognition of cells that successfully internalized fluorescent mRNA. To determine whether fluorescent mRNA can be localized in living cells, we performed time-lapse microscopy imaging. To that end, HeLa cells were transfected with fluorescent mRNA encoding eGFP. Images were acquired at 21 min intervals one hour after the transfection started (Figure 5D, Supplementary Figure S23, Supplementary videos). Strong fluorescence of Cy3 and/or Cy5 arising from the appropriate mRNA was present since the beginning of the observations. Within few hours, as enough protein was biosynthesized, eGFP fluorescence followed.

Fluorescent mRNA probes are compatible with in vivo applications

Finally, to demonstrate the compatibility of our mRNA labelling method with *in vivo* applications, we investigated localization and expression of fluorescently labelled mRNA in zebrafish larvae. Zebrafish (*Danio rerio*) is a simple vertebrate widely used in biological research due to several unique advantages. The optical transparency of zebrafish embryos makes them suitable for imaging of fluorescent proteins and mRNA probes. Moreover, injections at early one-cell stage allow for equal distribution of mRNA molecules to all diving cells during embryonic and larval development (46). Four different eGFP-coding mRNAs (N_3-m^7GRNA_{egfp}, N_3-m^7GRNA_{egfp}-Cy3, Cy5-m^7GRNA_{egfp} and Cy5-m^7GRNA_{egfp}-Cy3), in amounts from 10 to 300 pg, were microinjected into one-cell stage zebrafish embryos and their presence and expression were followed over time (Figure 6A). In case of all tested mRNAs, we observed bright fluorescent signal from eGFP, as early as about 5 hours post fertilization (hpf, Supplementary Figure S24). The signal was present even if only small amount of mRNA (10 pg) was injected and could still be detectable at 48 hpf (for 300 pg), indicating a robust translational activity of the injected mRNA. Using confocal microscopy and fine-tuning of the Cy3 and Cy5 channels, we observed fluorescence signals from mRNA in the cells of 8-hour old embryos (Figure 6B). Cy5 and Cy3 signals were present only if fluorescent mRNA was injected (no signals in case of N_3-m^7GRNA_{egfp}), and were specific for mRNA labelled with particular dye(s). Additionally, if dually labelled mRNA (Cy5-m7GRNAegfp-Cy3) was injected, the Cy3 and Cy5 signals colocalized. At 28 hpf the signals persisted, being more visible as punctate staining throughout various tissues of the embryo, such as the somite boundaries and characteristic chevron structures in the muscle segments of the posterior region of the trunk (Figure 6C) and retinal cell layers of the eye (Figure 6D). Importantly, even after injection with the highest mRNA dose (300 μg mRNA), the embryos developed normally, indicating that the chemically labelled mRNAs are fully biocompatible.

CONCLUSION

By revisiting the chemical basis of a seventy-year-old methodology, we developed a significantly improved protocol for RNA modification. Hydrazine derivatives are commonly used for 3′ RNA labelling, despite difficulties in their synthesis and instability of RNA-label conjugates. We found that R-EDA derivatives are not only more selective and reactive during reductive amination of RNA, but also more convenient in preparation (e.g. using commercially available NHS esters and diethylenetriamine DETA). Combination of this labelling protocol with RNA purification by HPLC provides unprecedented access to highly homogenous mRNA probes. One of the most surprising discoveries was the behaviour of labelled mRNA during HPLC separation. Significant changes of chromatographic mobility, and in consequence peak shape and retention time, were caused by minute changes in the chemical structure of the RNA polymer. HPLC is currently one of the best methods of RNA purification for *in vivo* applications, because it effectively removes immunogenic impurities, such as double-stranded RNAs (dsRNAs) (47). In our case, it additionally enabled isolation of highly homogenous fractions of mono- or dually labelled RNA. This was crucial for isolation of dually labelled RNA suitable for FRET experiments and intracellular localization of mRNA *in vitro* and *in vivo*. The dual-labelling protocol developed here, based on the combination of the reductive amination of RNA 3′ end with 5′ end labelling by SPAAC, allows for carrying out both reactions at the same time, which minimizes the time-dependent RNA degradation. Efficient and site-specific modification (i.e. multiple, orthogonal modification and labelling) of IVT

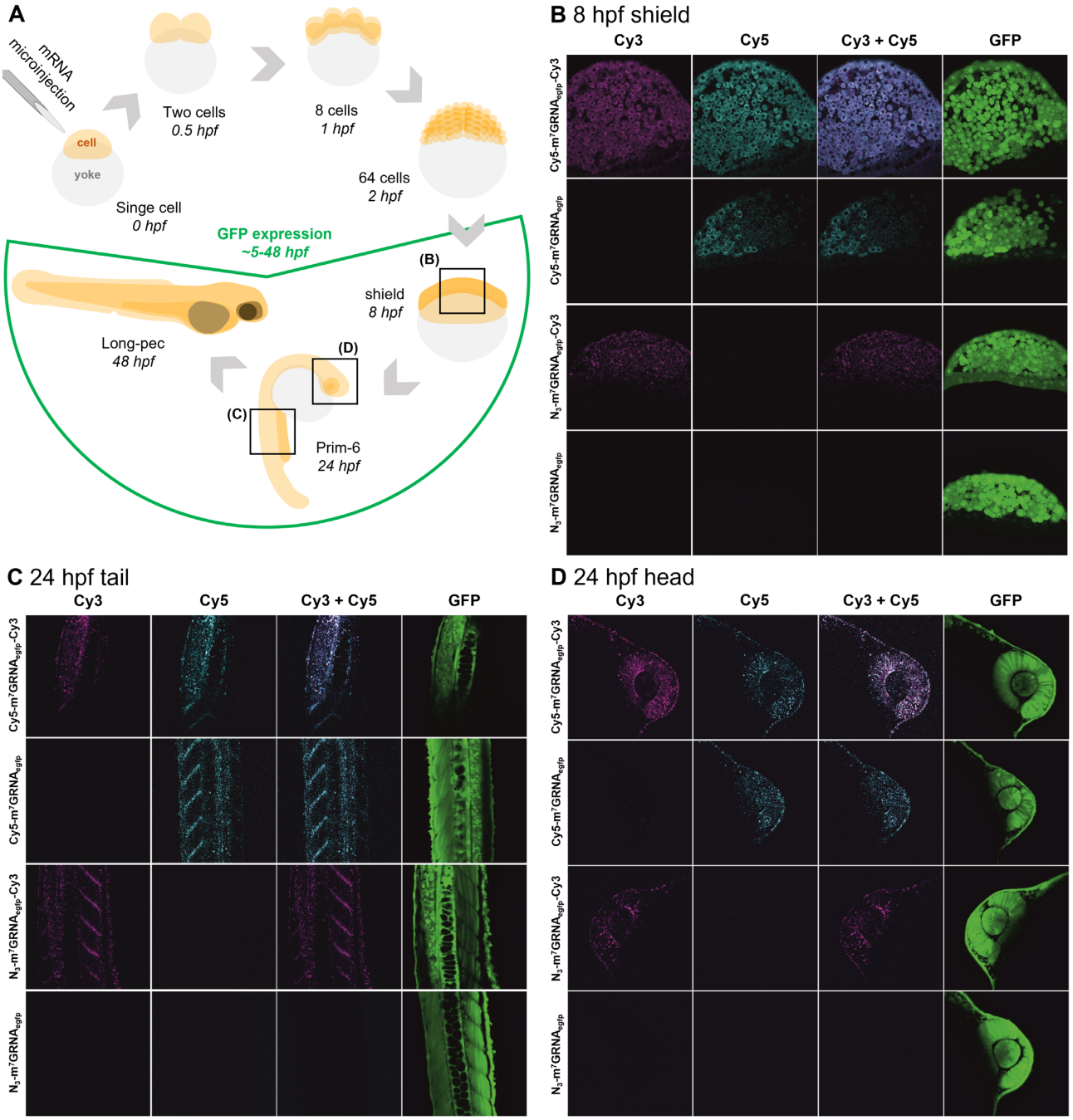

Figure 6. Microinjection of zebrafish embryos with modified mRNA. (**A**) Schematic representation of the experimental set-up and injection of mRNA in the course of zebrafish development during the first 48 h post fertilization (hpf). (**B**–**D**) Confocal microscopy images of embryos injected with 300 pg of N_3-m^{7}GRNA$_{egfp}$, N_3-m^{7}GRNA$_{egfp}$-Cy3, Cy5-m^{7}GRNA$_{egfp}$ or Cy5-m^{7}GRNA$_{egfp}$-Cy3 mRNA captured at 8 hpf (**B**) or 24 hpf in tail (**C**) or head (**D**) sections. Bright field and GFP fluorescence images of whole embryos are presented in Supplementary Figure S24.

RNA is generally a challenging task. It becomes even more difficult, if it concerns mRNA modification, as the final product must retain its biological activity in translation, which involves numerous interactions with proteins.

In this study, we demonstrate that mRNAs modified by our method can serve as substrates for various enzymes and are accepted by cellular translational machinery. However, we also identify that the proposed 3′ end modification may alter susceptibility to some cellular enzymes, as demonstrated in vitro by assessing susceptibility to recombinant CNOT7 deadenylase. Therefore, the exact impact of the proposed mRNA modifications on mRNA decay, localisation, or cellular immune responses requires further investigation. To the best of our knowledge, this is the first report describing the synthesis of highly homogenous dually labelled, translationally active mRNA in an efficient and scalable manner. Synthesis of dually labelled mRNA at high scales (10–300 μg), which are comparable to doses

of mRNA-based vaccines and therapeutics (48), paves the way for novel *in vivo* applications, even in clinical studies. We believe that the reported method is robust and versatile enough to find its application both in academic and industrial environments.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We gratefully thank Marcin Nowotny's Laboratory of Protein Structure (International Institute of Molecular and Cell Biology) for providing RNase H samples and Lukasz S. Borowski (University of Warsaw) for assistance in time laps imaging microscopy.

FUNDING

Foundation for Polish Science [TEAM/2016-2/13 to J.J.]; National Science Centre Poland [UMO-2018/31/B/ST5/03821 to J.K. and UMO-2018/31/D/NZ1/03526 to P.S.]; time laps imaging microscopy was performed using the CePT infrastructure, financed by the European Union: European Regional Development Fund [Innovative Economy 2007-13, agreement no. POIG.02.02.00-14-024/08-00]; zebrafish embryo images were recorded on a Zeiss LSM 780 – SP Mai Tai HP DS (Cell and Tissue Imaging Cluster (CIC)), supported by Hercules [AKUL/11/37 and FWO G.0929.15 to P.V.B.]; University of Leuven. Funding for open access charge: Foundation for Polish Science [TEAM/2016-2/13].
Conflict of interest statement. None declared.

REFERENCES

1. Hanspach,G., Trucks,S. and Hengesbach,M. (2019) Strategic labelling approaches for RNA single-molecule spectroscopy. *RNA Biol.*, **16**, 1119–1132.
2. Lai,W.-J.C., Kayedkhordeh,M., Cornell,E.V., Farah,E., Bellaousov,S., Rietmeijer,R., Salsi,E., Mathews,D.H. and Ermolenko,D.N. (2018) mRNAs and lncRNAs intrinsically form secondary structures with short end-to-end distances. *Nat. Commun.*, **9**, 4328.
3. Zhang,D., Zhou,C.Y., Busby,K.N., Alexander,S.C. and Devaraj,N.K. (2018) Light-activated control of translation by enzymatic covalent mRNA labeling. *Angew. Chem. Int. Ed.*, **57**, 2822–2826.
4. Anhäuser,L., Hüwel,S., Zobel,T. and Rentmeister,A. (2019) Multiple covalent fluorescence labeling of eukaryotic mRNA at the poly(A) tail enhances translation and can be performed in living cells. *Nucleic Acids Res.*, **47**, e42.
5. Croce,S., Serdjukow,S., Carell,T. and Frischmuth,T. (2020) Chemoenzymatic preparation of functional click-labeled messenger RNA. *ChemBioChem*, **21**, 1641–1646.
6. Tyagi,S. (2009) Imaging intracellular RNA distribution and dynamics in living cells. *Nat. Methods*, **6**, 331–338.
7. Buxbaum,A.R., Haimovich,G. and Singer,R.H. (2015) In the right place at the right time: visualizing and understanding mRNA localization. *Nat. Rev. Mol. Cell. Biol.*, **16**, 95–109.
8. Kirschman,J.L., Bhosle,S., Vanover,D., Blanchard,E.L., Loomis,K.H., Zurla,C., Murray,K., Lam,B.C. and Santangelo,P.J. (2017) Characterizing exogenous mRNA delivery, trafficking, cytoplasmic release and RNA–protein correlations at the level of single cells. *Nucleic Acids Res.*, **45**, e113.
9. Koch,A., Aguilera,L., Morisaki,T., Munsky,B. and Stasevich,T.J. (2020) Quantifying the dynamics of IRES and cap translation with single-molecule resolution in live cells. *Nat. Struct. Mol. Biol.*, **27**, 1095–1104.
10. Baladi,T., Nilsson,J.R., Gallud,A., Celauro,E., Gasse,C., Levi-Acobas,F., Sarac,I., Hollenstein,M.R., Dahlén,A., Esbjörner,E.K. *et al.* (2021) Stealth fluorescence labeling for live microscopy imaging of mRNA delivery. *J. Am. Chem. Soc.*, **143**, 5413–5424.
11. Gupta,V., Cherkassky,A., Chatis,P., Joseph,R., Johnson,A.L., Broadbent,J., Erickson,T. and DiMeo,J. (2003) Directly labeled mRNA produces highly precise and unbiased differential gene expression data. *Nucleic Acids Res.*, **31**, e13.
12. Gampe,C.M., Hollis-Symynkywicz,M. and Zécri,F. (2016) Covalent chemical 5′-functionalization of RNA with diazo reagents. *Angew Chem. Int. Ed.*, **55**, 10283–10286.
13. Zhao,M., Steffen,F.D., Börner,R., Schaffer,M.F., Sigel,R.K.O. and Freisinger,E. (2018) Site-specific dual-color labeling of long RNAs for single-molecule spectroscopy. *Nucleic Acids Res.*, **46**, e13.
14. Bhoge,B.A., Mala,P., Kurian,J.S., Srinivasan,V. and Saraogi,I. (2020) Selective functionalization at N2-position of guanine in oligonucleotides via reductive amination. *Chem. Commun.*, **56**, 13832–13835.
15. Markham,R. and Smith,J.D. (1952) The structure of ribonucleic acids. 2. The smaller products of ribonuclease digestion. *Biochem J.*, **52**, 558–565.
16. Whitfeld,R. and Markham,R. (1953) Natural configuration of the purine nucleotides in ribonucleic acids; chemical hydrolysis of the dinucleoside phosphates. *Nature*, **171**, 1151–1152.
17. Brown,D.M., Fried,M. and Todd,A.R. (1955) Nucleotides. Part XXXI. The stepwise degradation of polyribonucleotides: model experiments. *J. Chem. Soc.*, 2206–2210.
18. Zamecnik,P.C., Stephenson,M.L. and Scott,J.F. (1960) Partial purification of soluble RNA. *PNAS*, **46**, 811.
19. Khym,J. X. and Cohn,W. E. (1961) Amine-induced cleavage of periodate-oxidized nucleotide residues. *J. Biol. Chem.*, **236**, PC9–PC10.
20. Khym,J.X. (1963) The reaction of methylamine with periodate-oxidized adenosine 5′-phosphate. *Biochemistry*, **2**, 344–350.
21. Robberson,D.L. and Davidson,N. (1972) Covalent coupling of ribonucleic acid to agarose. *Biochemistry*, **11**, 533–537.
22. Hansske,F., Sprinzl,M. and Cramer,F. (1974) Reaction of the ribose moiety of adenosine and AMP with periodate and carboxylic acid hydrazides. *Bioorg. Chem.*, **3**, 367–376.
23. Reines,S.A. and Cantor,C.R. (1974) New fluorescent hydrazide reagents for the oxidized 3′-terminus of RNA. *Nucleic Acids Res.*, **1**, 767–786.
24. Lentzen,G., Dobberstein,B. and Wintermeyer,W. (1994) Formation of SRP-like particle induces a conformational change in E. coli 4.5 S RNA. *FEBS Lett.*, **348**, 233–238.
25. Griffiin,E.A. Jr, Qin,Z., Michels,W.J. Jr and Pyle,A.M. (1995) Group II intron ribozymes that cleave DNA and RNA linkages with similar efficiency, and lack contacts with substrate 2′-hydroxyl groups. *Chem. Biol.*, **2**, 761–770.
26. Busch,S., Kirsebom,L.A., Notbohm,H. and Hartmann,R.K. (2000) Differential role of the intermolecular base-pairs G292-C75 and G293-C74 in the reaction catalyzed by Escherichia coli RNase P RNA. *J. Mol. Biol.*, **299**, 941–951.
27. Easterbrook-Smith,S.B., Wallace,J.C. and Keech,D.B. (1976) Pyruvate carboxylase affinity labelling of the magnesium adenosine triphosphate binding site. *Eur. J. Biochem.*, **62**, 125–130.
28. Rayford,R., Anthony,D.D. Jr, O'Neill,R.E. Jr and Merrick,W.C. (1985) Reductive alkylation with oxidized nucleotides. Use in affinity labeling or affinity chromatography. *J. Biol. Chem.*, **260**, 15708–15713.
29. Oh,B.-K. and Pace,N.R. (1994) Interaction of the 3′-end of tRNA with ribonuclease P RNA. *Nucleic Acids Res.*, **22**, 4087–4094.
30. von Ahsen,U. and Noller,H.F. (1995) Identification of bases in 16S rRNA essential for tRNA binding at the 30S ribosomal P site. *Science*, **267**, 234–237.
31. Bellon,L., Workman,C., Scherrer,J., Usman,N. and Wincott,F. (1996) Morpholino-linked ribozymes: a convergent synthetic approach. *J. Am. Chem. Soc.*, **118**, 3771–3772.
32. Rossbach,O., Hung,L.-H., Schreiner,S., Grishina,I., Heiner,M., Hui,J. and Bindereif,A. (2009) Auto-and cross-regulation of the hnRNP L proteins by alternative splicing. *Mol. Cell Biol.*, **29**, 1442–1451.

33. Li,D., Meyer,M.H.F., Willkomm,D.K., Keusgen,M. and Hartmann,R.K. (2010) Analysis of bacterial RNase P RNA and protein interaction by a magnetic biosensor technique. *Biochimie*, **92**, 772–778.
34. Yamamoto,J., Ebisuda,S., Kong,L., Yamago,H. and Iwai,S. (2017) Post-synthetic modification of 3′ terminus of RNA with propargylamine: a versatile scaffold for RNA labeling through copper-catalyzed azide-alkyne cycloaddition. *Chem. Lett.*, **46**, 767–770.
35. Reischl,R.J., Bicker,W., Keller,T., Lamprecht,G. and Lindner,W. (2012) Occurrence of 2-methylthiazolidine-4-carboxylic acid, a condensation product of cysteine and acetaldehyde, in human blood as a consequence of ethanol consumption. *Anal. Bioanal. Chem.*, **404**, 1779–1787.
36. Song,Z.-C., Ma,G.-Y. and Zhu,H.-L. (2015) Synthesis, characterization and antibacterial activities of N-tert-butoxycarbonyl-thiazolidine carboxylic acid. *RSC Adv.*, **5**, 24824–24833.
37. Kicsák,M., Mándi,A., Varga,S., Herczeg,M., Batta,G., Bényei,A., Borbás,A. and Herczegh,P. (2018) Tricyclanos: conformationally constrained nucleoside analogues with a new heterotricycle obtained from ad-ribofuranose unit. *Org. Biomol. Chem.*, **16**, 393–401.
38. Stepinski,J., Waddell,C., Stolarski,R., Darzynkiewicz,E. and Rhoads,R.E. (2001) Synthesis and properties of mRNAs containing the novel "anti-reverse" cap analogs 7-methyl (3′-O-methyl) GpppG and 7-methyl (3′-deoxy) GpppG. *RNA*, **7**, 1486–1495.
39. Jemielity,J., Kowalska,J., Rydzik,A.M. and Darzynkiewicz,E. (2010) Synthetic mRNA cap analogs with a modified triphosphate bridge - synthesis, applications and prospects. *New J. Chem.*, **34**, 829–844.
40. Sahin,U., Muik,A., Derhovanessian,E., Vogler,I., Kranz,L.M., Vormehr,M., Baum,A., Pascal,K., Quandt,J. and Maurus,D. (2020) COVID-19 vaccine BNT162b1 elicits human antibody and TH 1 T cell responses. *Nature*, **586**, 594–599.
41. Mamot,A., Sikorski,P.J., Warminski,M., Kowalska,J. and Jemielity,J. (2017) Azido-functionalized 5′ cap analogues for the preparation of translationally active mRNAs suitable for fluorescent labeling in living cells. *Angew Chem. Int. Ed.*, **56**, 15628–15632.
42. Ermolenko,D.N. and Mathews,D.H. (2021) Making ends meet: new functions of mRNA secondary structure. *Wiley Interdiscipl. Rev.: RNA*, **12**, e1611.
43. Floor,S.N., Jones,B.N., Hernandez,G.A. and Gross,J.D. (2010) A split active site couples cap recognition by Dcp2 to activation. *Nat. Struct. Mol. Biol.*, **17**, 1096.
44. Nowotny,M., Gaidamakov,S.A., Ghirlando,R., Cerritelli,S.M., Crouch,R.J. and Yang,W. (2007) Structure of human RNase H1 complexed with an RNA/DNA hybrid: insight into HIV reverse transcription. *Mol. Cell.*, **28**, 264–276.
45. Strzelecka,D., Smietanski,M., Sikorski,P.J., Warminski,M., Kowalska,J. and Jemielity,J. (2020) Phosphodiester modifications in mRNA poly(A) tail prevent deadenylation without compromising protein expression. *RNA*, **26**, 1815–1837.
46. Choi,T.-Y., Choi,T.-I., Lee,Y.-R., Choe,S.-K. and Kim,C.-H. (2021) Zebrafish as an animal model for biomedical research. *Exp. Mol. Med.*, **53**, 310–317.
47. Kariko,K., Muramatsu,H., Ludwig,J. and Weissman,D. (2011) Generating the optimal mRNA for therapy: HPLC purification eliminates immune activation and improves translation of nucleoside-modified, protein-encoding mRNA. *Nucleic Acids Res.*, **39**, e142.
48. Pardi,N., Hogan,M.J., Porter,F.W. and Weissman,D. (2018) mRNA vaccines—a new era in vaccinology. *Nat. Rev. Drug Discov.*, **17**, 261.