Chemistry
A European
Journal

Chemistry Europe
European Chemical Societies Publishing

Accepted Article

Title: RNA Ligation for Mono and Dually Labeled RNAs

Authors: Anaïs Depaix, Agnieszka Mlynarska-Cieslak, Marcin Warminski, Pawel J Sikorski, Jacek Jemielity, and Joanna Kowalska

To be cited as: *Chem. Eur. J.* 10.1002/chem.202101909

Link to VoR: https://doi.org/10.1002/chem.202101909

WILEY-VCH

10.1002/chem.202101909

RESEARCH ARTICLE

WILEY-VCH

RNA Ligation for Mono and Dually Labeled RNAs

Anaïs Depaix,[a] Agnieszka Mlynarska-Cieslak,[a] Marcin Warminski,[a] Pawel J. Sikorski,[b] Jacek Jemielity,[b] and Joanna Kowalska*[a]

[a] Dr. A. Depaix, M. Sc. A. Mlynarska-Cieslak, Dr. M. Warminski, Dr. J. Kowalska
Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland
E-mail: jkowalska@fuw.edu.pl, Twitter handle: @JoannaK_UW
[b] Dr. P. J. Sikorski, Prof. Dr. J. Jemielity
Centre of New Technologies
University of Warsaw
Banacha 2c, 02-097 Warsaw, Poland

Abstract: Labeled RNAs are invaluable probes for investigation of RNA function and localization. However, mRNA labeling remains challenging. Here, we developed an improved method for 3′-end labeling of in vitro transcribed RNAs. We synthesized novel adenosine 3′,5′-bisphosphate analogs modified at the N6 or C2 position of adenosine with an azide-containing linker, fluorescent label, or biotin and assessed these constructs as substrates for RNA labeling directly by T4 ligase or via postenzymatic strain-promoted alkyne-azide cycloaddition (SPAAC). All analogs were substrates for T4 RNA ligase. Analogs containing bulky fluorescent labels or biotin showed better overall labeling yields than postenzymatic SPAAC. We successfully labeled uncapped RNAs, NAD-capped RNAs, and 5′-fluorescently labeled $m^7Gp_3A_m$-capped mRNAs. The obtained highly homogenous dually labeled mRNA was translationally active and enabled fluorescence-based monitoring of decapping. This method will facilitate the use of various functionalized mRNA-based probes.

Introduction

RNAs play crucial roles in living organisms and are a promising class of modern therapeutics. Single- or dual-color fluorescently labeled or biologically tagged RNA probes have paved the way for studying the intracellular fates and biological roles of RNAs using advanced microscopic and analytical methods.[1] Fluorescently labeled oligoribonucleotides and RNAs are valuable probes for studying specific ligand or protein binding and for investigating RNA dynamics, function, and localization.[2] These molecules have been used for many applications, including deciphering the biological roles of RNAs, studying viral mRNA replication pathways,[3] and detecting cancer.[2c; 4] Fluorescent labeling of RNA enables the use of multiple techniques (fluorescence intensity changes, anisotropy, fluorescence resonance energy transfer, or fluorescence lifetime) to visualize and monitor enzymatic processes with high sensitivity.[5] Biological tagging with molecules such as biotin enables the application of various RNA capture techniques, relying on the high affinity of the tag to another entity (e.g., biotin and (strept)avidin).[6] Biotin-(strept)avidin technology-based assays are valuable tools for protein affinity purification, pull-down experiments, in vivo visualization, and cellular delivery, among others.[7]

The synthesis of short fluorescent or biologically tagged RNA probes can be achieved via full chemical synthesis, whereas longer RNA probes require different enzymatic[8] and chemoenzymatic methods.[9] Among currently available enzymatic methods, two main approaches have been used. The first is cotranscriptional labeling, in which a labeled nucleotide is incorporated into RNA during an *in vitro* transcription reaction catalyzed by DNA-dependent RNA polymerase. The second is post-transcriptional labeling, typically performed using an RNA ligase or template-independent polymerase,[9f] which incorporates a labeled nucleotide at the 3′-end of *in vitro* transcribed (IVT) RNA.[10] Ribozymatic posttranscriptional approaches for internal, sequence-specific RNA labeling have also been developed.[11]

One of the first developed RNA labeling strategies uses T4 RNA ligase 1, an enzyme known to form a phosphodiester bond between an (oligo)nucleotide containing a 5′-phosphate group and an oligonucleotide possessing a free 3′-hydroxyl group.[10; 12] The minimum substrate that have been attached to mRNA using this method is radioactively labeled cytidine or adenosine 5′,3′-bisphosphate (pCp or pAp, respectively). Different pNp derivatives have been later developed for the purpose of RNA labeling, including fluorescently tagged pCp analogs, photocaged nucleotides and isotopically labeled pAp (^{13}C,^{15}N-pAp).;[13] Chemoenzymatic methods for RNA labeling include post-transcriptional modification of prefunctionalized RNA using reactions such as esterification,[14] amide condensation, inverse-electron demand Diels-Alder cycloaddition, or azide-alkyne cycloaddition (AAC), which is particularly convenient for introducing various labels in a highly efficient and biocompatible manner.[9e; 15] Although the copper-catalyzed version of AAC (CuAAC) has been extensively used to label RNA, copper ions are cytotoxic, making this method incompatible with in vivo imaging.[16] This inconvenience has been circumvented by the development of dibenzocyclooctylamine derivatives that are reactive in strain-promoted azide-alkyne cycloaddition, which, in contrast to CuAAC, does not require any catalyst.[15b; 17]

Various methods have been employed to afford monolabeled RNAs for versatile applications;[2e; 7a; 7c; 9f; 13b; 15c; 18] however, there is still room for improvement and for the development of new methods. Dually labeled RNA probes have also recently attracted interest owing to their advanced application in single-molecule spectroscopy and conformational studies. Labeling of IVT mRNAs is particularly challenging because of the large size and heterogeneity of the mRNAs and the tendency of mRNAs to adopt complex tertiary conformations.[9g] Moreover, mRNAs have atypical 5′ ends, usually modified by a 7-methylguanosine cap, whereas their 3′ ends are polyadenylated. Both of these terminal structures are crucial for the activity of mRNAs in translation; therefore, a

desirable mRNA labeling method should not disrupt the functionalities of these elements. Recently, other types of unconventional mRNA caps with unknown biological roles have also been discovered, and molecular tools to investigate them are also needed.[19] To study the fates of IVT mRNAs and mRNA-based therapeutics, it is necessary to apply site-specific, controllable, and scalable procedures that yield a highly homogenous product. Despite recent progress in the development of mRNA-based vaccines and therapies, methods for achieving this goal are limited.

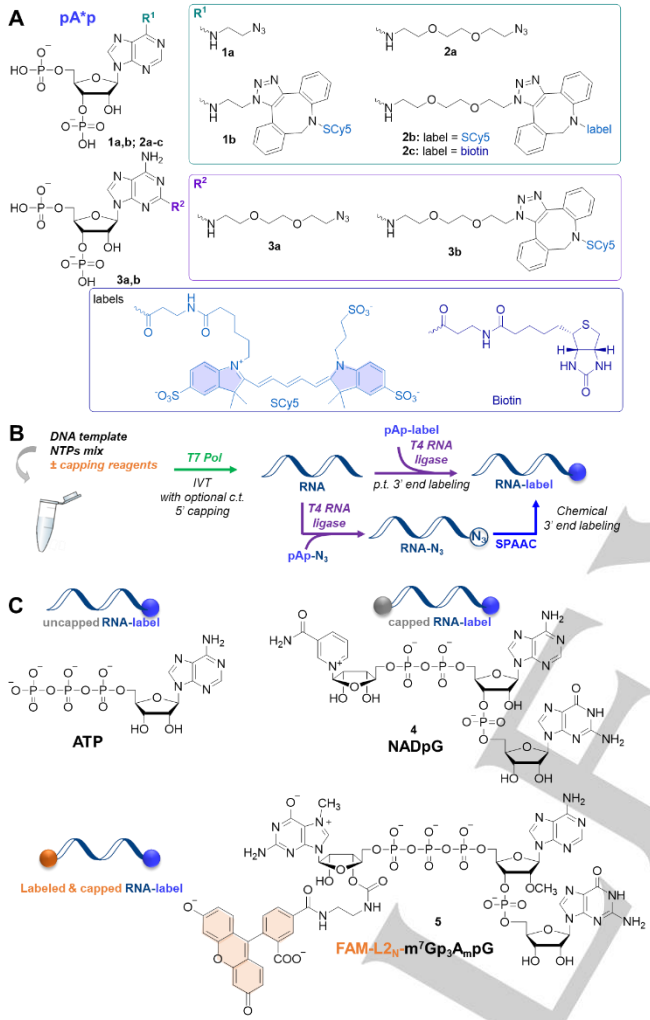


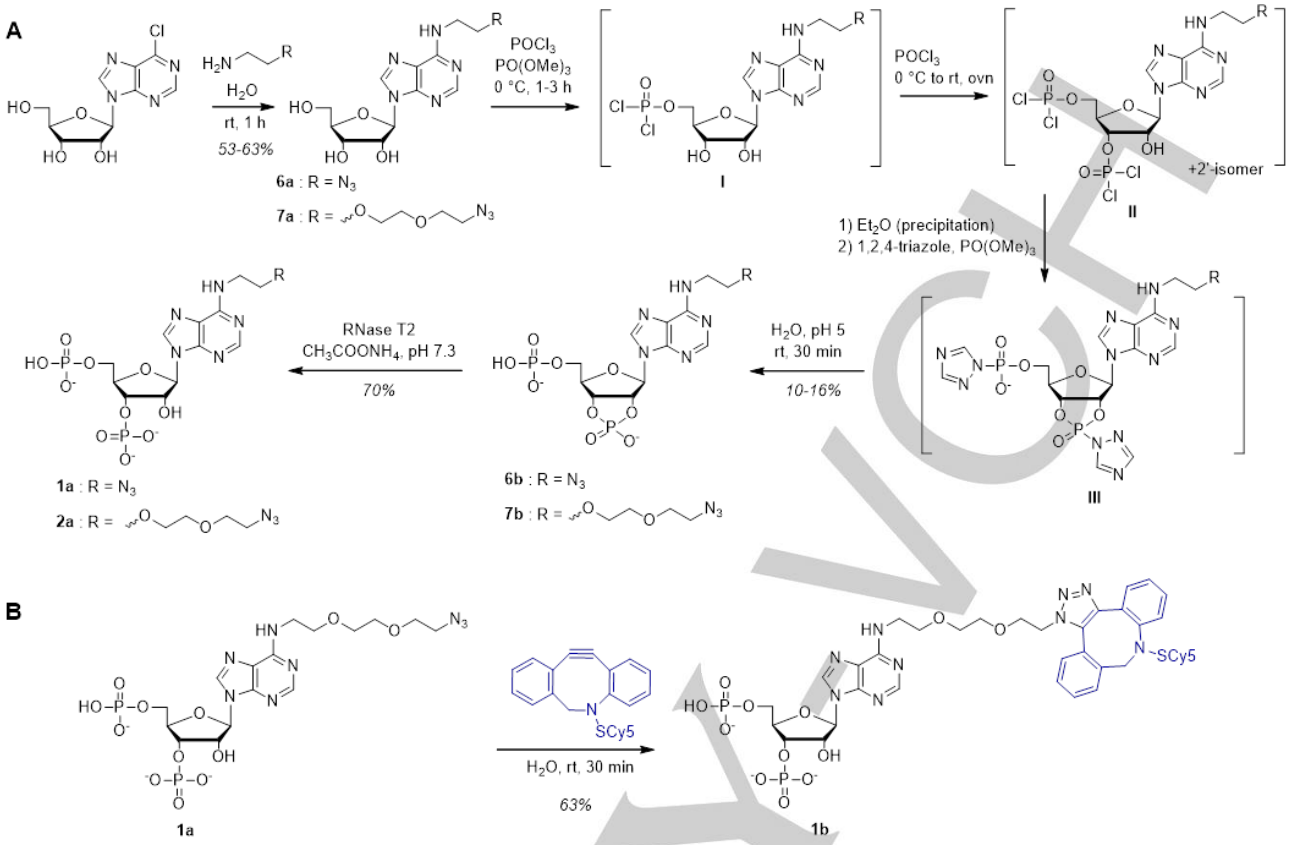
Figure 1. A. Adenosine 5′,3′-bisphosphate analogs designed and used in the ligation step. B. General strategy for generation of 3′-labeled RNA. C. Derivatives used in IVT as first transcribed nucleotides affording uncapped (incorporating ATP at the 5′ end) and capped RNAs (incorporating NADpG **4** or FAM-L2$_N$-m^{7}Gp$_3$A$_m$pG **5**).

In this study, we revisited the long-known strategy for RNA 3′-end labeling with T4 RNA ligase (LT4) to develop a widely applicable method suitable for RNA, including mRNA and labeling. We sought to design new reagents that could serve as efficient substrates for LT4 applicable to copper-free functionalization, bio-tagging, and fluorescent labeling of RNA, including mRNA. Because mRNAs contain polyA tails at the 3' end, we focused on the development of pAp derivatives. To that end, we obtained a set of pAp analogs substituted at different positions (C2 or N6 adenine) with different azide-containing linkers (ethylene or polyethylene glycol; Figures 1A and S1) and used them in chemoenzymatic reactions to access RNA with various modifications at their 3′ ends (Figure 1B). The labeling method was first optimized for short in vitro transcribed RNAs (35 nucleotides long) and later adapted to longer, variously capped RNAs, including mRNAs (Figure 1C). We identified compounds that enabled straightforward incorporation of fluorescent labels or biotin at the 3′ ends of RNAs, providing access to structurally diverse RNA probes as valuable research tools, including 3′-mono- and 3′,5′-dually labeled mRNAs that retained translational activity.

Results and Discussion

Synthesis of nucleotides for RNA modification

All pAp analogs were synthesized starting from unprotected nucleosides in a one-pot three-step reaction yielding the corresponding 2′,3′-cyclophosphonucleoside 5′-phosphates as the key intermediates (Scheme 1A). These intermediates were then subjected to RNAse T2 to achieve regioselective cleavage leading to the target 3′-phosphonucleoside 5′-phosphates. N6-substituted pAp analogues **1a** and **2a** were synthesized starting from unprotected 6-chloropurineriboside (Scheme 1A). This latter was substituted at position 6 with an appropriate amino-azido linker, either 8-azido-3,6-dioxaundecan-1-amine or 2-azidoethylamine. The resulting N6-functionalized nucleosides (**6a** and **7a**, respectively) were converted into the corresponding 5′,3′-diphosphates using a one-pot three-step procedure. To this end, nucleosides were treated with phosphoryl chloride (3 equiv.) to afford the corresponding nucleoside 5′-dichlorophosphates **I**, as intermediates. When high-performance liquid chromatography (HPLC) monitoring showed complete conversion of the starting material, a second portion of phosphoryl chloride was added, yielding the dichlorophosphonucleoside 5′-dichlorophosphate intermediates **II** as two (3′ and 2′) regioisomers. The latter were precipitated with diethyl ether and treated with heterocyclic nucleophiles to induce cyclization of the 2′/3′-dichlorophosphate with concomitant substitution of the chlorine atoms, followed by hydrolysis to the desired 2′,3′-cyclophosphonucleoside 5′-phosphate. We tested different nucleophiles and bases as cyclization-inducing reagents, including triethylamine, diisopropylamine, imidazole, 3-amino-1,2,4-triazole, and 1,2,4-triazole. The first two did not induce cyclization, whereas, imidazole and triazole derivatives were effective yielding corresponding phosphorimidazolide or -triazolide intermediates. We aimed to select a reagent that enables hydrolysis of all P-N bonds at neutral and mildly acidic pH (7–3) because at lower pH values the hydrolysis of the P-N bond was accompanied by undesired opening of the 2',3'-cyclophophate back to 2'- and 3'-phosphates. 1,2,4-Triazole was the optimal cyclization-inducing agent, as the resulting 5′-phosphorotriazolide (**III**) was hydrolyzed within 30 min at pH 5, affording compounds **6b** and **7b** without any observable formation of by-products resulting from the 2′,3′-cyclophosphate opening. Recently, an alternative method providing direct access to 2',3'-cyclophophonucleoside 5'-phosphates from nucleosides has been reported.[20] Finally, **6b** and **7b** were selectively converted to the corresponding 3′-phosphates (**1a** and **2a**) by RNAse T2 cleavage (Figure S2A-D,



Scheme 1. A. Representative pAp synthesis for N6 derivatives **1a** and **2a**. B. Representative click reaction with DIBAC-SCy5 yielding compound **1b.** The structure of SCy5 is shown in Figure 1.

~70% yield) with an overall yield from unprotected nucleoside of 4-7%.

The synthesis of the pAp derivative modified at position C2 (**3a**) was attempted in a similar way starting from 2-fluoroadenosine and included aromatic nucleophilic substitution, followed by exhaustive phosphorylation, 2′,3′-cyclization, and C2-substituted adenosine was not susceptible to hydrolysis by RNAse T2, even at much higher concentrations than initially tested. Therefore, the synthesis was completed by chemical hydrolysis of 2′,3′-cyclophosphate under acidic conditions, followed by separation of the resulting 2′- and 3′-isomers by reverse phase (RP)-HPLC (Figure S2E, F, 8% yield), with an overall 0.3% yield from 2-fluoroadenosine. Finally, all azido-functionalized derivatives (**1a–3a**) were labeled in a strain-promoted azide-alkyne cycloaddition reaction with DIBAC-functionalized Cy5 or biotin (Scheme 1B, DIBAC-SCy5) to give the resulting labeled pAp derivatives (**1b**, **2b**, **2c**, and **3b**; Figure S1) with good yields.

Next, we tested the utility of the synthesized compounds in LT4-mediated RNA labeling in two different variants. The first variant involved direct labeling of RNA at the 3′ end using a fluorescently labeled or biotinylated pAp analog (Figure 1B, pathway A). The second variant was a one-pot two-step labelingprocedure, wherein an azido pAp derivative was ligated to the RNA 3′ end followed by copper-free click chemistry (SPAAC reaction with DIBAC-SCy5) with the resulting RNA-N_3 (Figure 1B, pathway B). In the initial experiments, we aimed to determine the excess pAp analog required for effective ligation to the RNA 3′ end. To this end, we used a fluorescently labeled (sulfoCy5, SCy5) pAp analog (**2b**), which enabled straightforward visualization of the labeled product after gel electrophoresis. The labeling efficiency was tested for 35 nt IVT RNA (at 1 µM) and different excess amounts (from 1- to 200-fold) of **2b** in the presence of 1 U/µL LT4. The RNAs were purified using a commercially available kit and analyzed by high-resolution polyacrylamide gel electrophoresis (HRPAGE, 15% polyacrylamide in the presence of 7 M urea and 1× TBE buffer) to assess the efficiency of the ligation step. The gels were first scanned for Cy5 fluorescence (to visualize labeled RNA), stained with ethidium bromide (EtBr), and scanned for EtBr fluorescence (to visualize total RNA). The analysis revealed that the unlabeled RNA substrate was heterogeneous and, in addition to the expected 35 nt main product, contained transcripts 1 nt shorter or longer, which were visible on the gel below and above the most intense band, respectively (Figure 2, ref). Under the tested conditions, the RNA 3′-end labeling efficiency was highly dependent on the concentration of **2b** (Figure 2, lanes 1–7, and Figure S3); that is, the labeling efficiency gradually improved as the concentration of **2b** increased from 1 to 100 µM. Surprisingly, when using a higher excess (200 equiv., Figure S3), the ligation efficiency was not improved. To assess the lower limit of RNA concentration, we explored the applicability of these conditions to the labeling of RNA at lower concentrations. As shown in Figure 2, lanes 8–11, the conditions were suitable for 0.5 µM RNA, although the efficiency of the ligation step was slightly lower, and using 0.1 µM RNA resulted in an even lower efficiency (Figure S3),

suggesting that the RNA concentration should be maintained higher than 0.5 μM during the labeling procedure.

Next, we compared the labeling using compound **2b** with similar procedures using other pAp analogs (**1b** and **3b**) and with a two-step procedure consisting of labeling RNA with **1a**, **2a**, and **3a**, followed by attachment of sulfo-cyanine 5 (SCy5) using copper-free click chemistry (SPAAC; Figure 3). The conditions optimized earlier for **2b** were used to compare the different labeling variants. All azido derivatives were incorporated into RNA (Figure 3A, lanes 1–3); however, the structure of the linker unexpectedly influenced the incorporation of the labeled derivative. Ligation of the N6-substituted labeled pAp containing a shorter linker (**1b**; Figure 3A, lane 4) was less efficient than that with longer linker (**2b**; Figure 3A, lane 5). Moreover, we also found that the compound substituted at the C2 position (**3b**) afforded a better labeling yield than the corresponding N6-derivative (**2b**). RNA-N_3 obtained by ligation with compounds **1a**, **2a**, or **3a** were purified by commercial kit and subjected to SPAAC with DIBAC-SCy5. The efficiency of the click reaction depended on the linker length; in this case, the reaction was most efficient for the derivative with the shorter linker

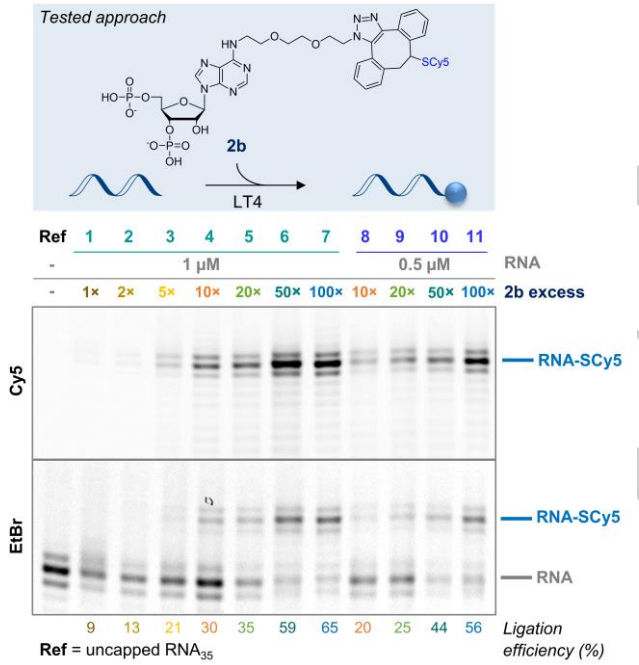


Figure 2. Optimization of LT4-mediated RNA (35 nt) labeling with compound **2b**. HRPAGE (15% PAA, 1× TBE) of the resulting RNAs (100 ng) after ligation with pAp **2b** on uncapped RNA_{35}.

(**1a**; Figure 3A, lane 7 versus lanes 8 and 9). Furthermore, because the C2-substituted compound **3b** was most efficiently incorporated at the RNA 3′ end, we tested whether the compound was more suitable for labeling RNA at lower concentrations (Figure 3B). Even under the most restrictive conditions, that is, 0.1 μM RNA and 10 μM **3b** (100-fold excess, lane 1), **3b** was incorporated with high efficiency (~70%, Figure 3B, lane 1), which was more than 2-fold higher than that of compound **2b** (~30%, Figure S3). These results suggested that C2-modified pAp analogs, despite being synthetically more challenging, may be analogs of choice for labeling procedures performed at low RNA concentrations (e.g., if the amount of RNA material is limited).

Overall, SCy5 labeling was more efficient when performed directly via one-step ligation, than via the two-step procedure. The latter, albeit slightly more complex, may deliver superior results in cases in which the label is not accepted by LT4, owing to its structure or electrostatic properties, or if RNA is to be labeled in biological mixtures, e.g., for pull-down experiments. Indeed, the RNA functionalized with derivative **1a** containing an azidoethylene linker could be useful for experiments requiring labeling of RNA by SPAAC in various biological settings.

The abovementioned labeling reactions were performed on uncapped (5′-triphosphorylated) RNAs. However, endogenous RNA molecules are often modified at their 5′ ends by different canonical and noncanonical cap structures. Therefore, we next tested whether the labeling procedure was compatible with 5′ capped RNAs. As the first model RNA, we chose NAD-linked RNA. NAD was recently discovered as a tag of the 5′ ends of some RNAs in bacteria and eukaryotes.[21] The biological role of this noncanonical cap is not yet well understood, and developing NAD-RNA probes is crucial to elucidating its function. Therefore, we aimed to synthesize NAD-capped RNA modified at the 3′ end with either a fluorescent label or biotin to provide NAD-RNA-based fluorescent or affinity probes, respectively. NAD-RNA was obtained by IVT using NADpG trinucleotide (**4**), as previously reported (Scheme S1, Figure S4).[22] The resulting NAD-RNA_{35} was subjected to reactions with compounds **1b**, **2b**, **2c**, **3b**, **1a**, **2a**, and **3a** following SPAAC and analyzed by HRPAGE (Figure 4). The two biotinylated regioisomers of **2c** were separated during HPLC purification (named **2c-I** and **2c-II** according to the elution order) and were therefore used separately in the ligation experiment.

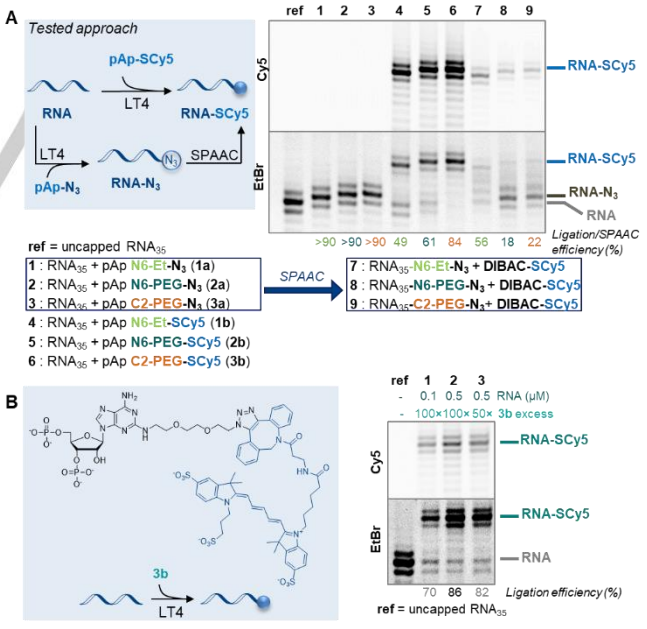


Figure 3. A. HRPAGE (15% PAA, 7 M urea, 1× TBE)-labeling of uncapped RNA_{35}, direct ligation versus ligation followed by SPAAC. B. HRPAGE (15% PAA, 7 M urea, 1× TBE)-labeling of uncapped RNA_{35} using pAp **3b**.

All azido-modified pAp analogs (**1a**, **2a**, and **3a**) were efficiently incorporated at the 3′-end of NAD-RNA (Figure 4, lanes 1–3), and incorporation of the SCy5-labeled derivatives was found to be dependent on the linker length and the position of the substitution, similar to uncapped RNAs (Figure 4, lanes 4–

6). We also found that the nature of the label (SCy5 or biotin) did not influence the efficacy of the ligation (Figure 4, lanes 10 and 11 versus lane 5). The SPAAC reactions on the resulting azido-substituted NAD-RNAs were again the most efficient for RNA derivatives with shorter linkers (Figure 4, lane 7 versus lanes 8 and 9). In agreement with the findings for uncapped RNA, labeling was more efficient in the case of direct ligation using a labeled adenosine analog and most efficient when using the C2-substituted derivative **3b** (Figure S5). Overall, the results were similar to those for uncapped RNA, indicating that the labeling procedure using pAp derivatives was compatible with RNAs modified at the 5′ end.

Because the 3′-end labeling procedure was compatible with 5′ capped RNAs, we envisaged that this approach could also be combined with mRNA 5′-end labeling to afford dually labeled RNAs. To demonstrate this, we first performed labeling on a 35-nt RNA substrate carrying a fluorescently labeled m^7G cap. m^7G is a canonical cap present at the 5′ end of eukaryotic mRNAs. m^7G-capped mRNAs can be efficiently prepared by IVT in the presence of trinucleotide cap analogs, such as m^7GpppA_mpG.[23] Here, we designed a fluorescent analog of this trinucleotide to obtain a 5′-capped and labeled RNA substrate, which could be converted into a dually labeled 5′-capped RNA probe by ligation with a fluorescent pAp analog.

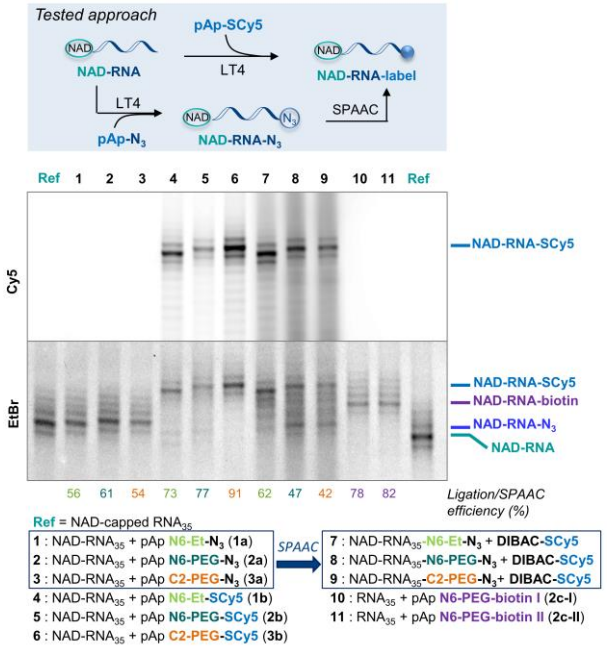


Figure 4. HRPAGE (15% PAA, 7 M urea, 1× TBE) of NAD-RNAs. Direct labeling/ligation versus ligation followed by SPAAC. The ligation/SPAAC efficiency was calculated based on a densitometric analysis (CLIQ software) of EtBr stained gel, as average of 2 or 3 replicates.

Trinucleotide **5** (FAM-L_2N-m^7GpppA_mpG, Figure 1C, Scheme S2) was incorporated into the RNA by IVT with T7 polymerase (Figure S4).[24] The resulting m^7G-RNA_{35} was directly labeled with the fluorescent pAp analog **2b**. The labeling reactions were visualized by HRPAGE and RP-HPLC (Figure 5). The RNA analyzed by PAGE was visualized for both Cy5 fluorescence and FAM fluorescence before EtBr staining. We found that LT4-mediated ligation was very efficient, yielding the desired dually labeled RNA as the major product. HPLC analysis revealed that unlabeled, monolabeled, and dually labeled RNAs were easily separated, allowing us to isolate the dually labeled RNA with good yield and high purity. The separation was efficient enough to allow for the recovery of unreacted 5′-labeled RNA.

Dually labeled RNAs are promising biological tools because they can be visualized using two different fluorescence channels. Such probes could benefit several intensively investigated research areas, such as RNA decapping, intracellular transport of different types of RNA, RNA localization, storage, and conformational dynamics.

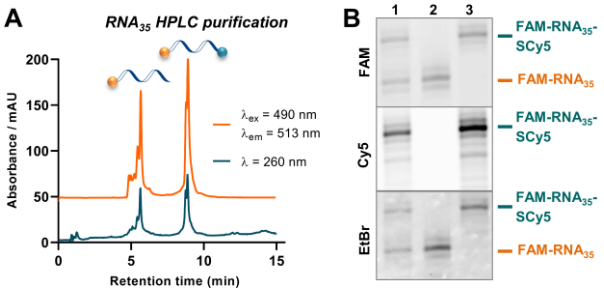


Figure 5. A. HPLC profile of dually labeled RNA_{35} purification; B. HRPAGE analysis (15% PAA, 7 M urea, 1× TBE). 1: RNA_{35} after ligation with **2b** and before HPLC purification. 2: First collected fraction (retention time, 5.7 min; recovered substrate). 3: Second collected fraction (retention time, 8.8 min; purified dually labeled RNA_{35}).

Therefore, we next applied our methodology, i.e., in vitro cotranscriptional 5′-capping/labeling followed by 3′-end post-transcriptional ligation/labeling, to mRNA. Using the appropriate DNA plasmid as a template, the same fluorescently labeled trinucleotide **5** and T7 polymerase was used to obtain FAM-L_2N-$m^7Gp_3A_m$-mRNA, coding for Gaussia luciferase.

In contrast to RNA_{35}, the maximum concentration of labeled mRNA was 0.1 µM, and HPLC analysis of IVT mRNA showed that the capping was not full (signals partially overlapped). Therefore, we purified the monolabeled mRNA by HPLC to remove the main portion of uncapped mRNA in order to perform labeling on high-purity monolabeled mRNA. mRNA labeling performed in the presence of either 500- or 1000-fold excess **2b** gave similarly high conversion into dually labeled mRNA (86% yield from triplicate experiments). Remarkably, HPLC purification allowed the separation of mono- and dually labeled mRNAs in a very efficient way, and the two mRNA isomers, presumably corresponding to the two regioisomers of 4,5-disubstituted triazole in the **2b** derivative, were also separated (Figure 6A-B).

To demonstrate the potential of this dually labeled RNA for mRNA research, we first performed an mRNA decapping assay using human decapping protein 2 (hDcp2). Typically, RNA decapping can be precisely monitored for short RNAs, for which capped and uncapped RNAs can be separated during gel electrophoresis and independently quantified. For longer RNAs, such separation is not possible, and other methods need to be employed; however, these methods are not fully quantitative.[23b; 25] Therefore, we tested whether the formation of uncapped mRNAs catalyzed by *h*Dcp2 can be directly visualized using dually labelled mRNA. As shown in Figure 6C, almost complete decapping occurred within 30 min, as demonstrated by the decrease in fluorescein signal over time as a result of 5′-cap loss, whereas the signal from the SCy5 label (attached at the 3′ end of mRNA) corresponding to total RNA remained constant. The fluorescein signal was stable in the absence of *h*Dcp2 (Figure S6), proving that this dually labeled mRNA was a

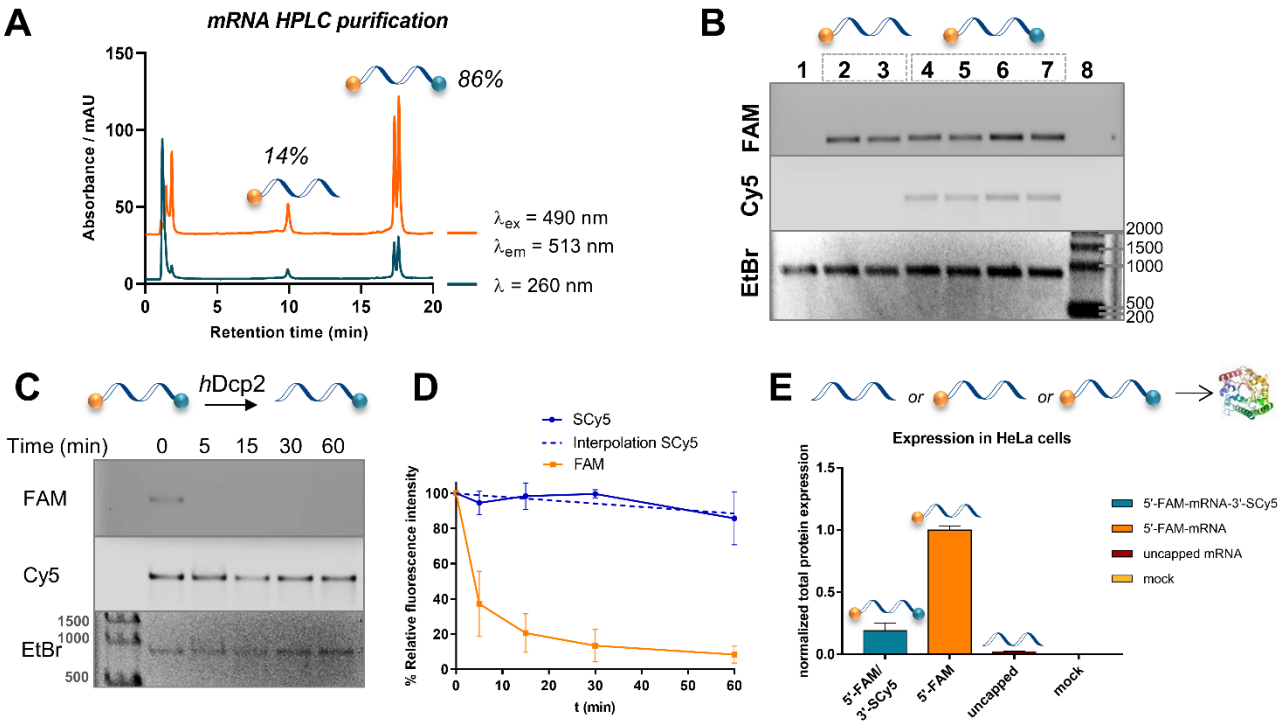


Figure 6. A. HPLC profile of dually labelled mRNA purification. B. Agarose electrophoresis (1.5%, TBE). 1: uncapped mRNA, 2: FAM-m^{7}Gp$_3$A$_m$-mRNA (substrate), 3: FAM-m^{7}Gp$_3$A$_m$-mRNA after HPLC (retention time, 9.9 min), 4 and 6: two different batches of FAM-m^{7}Gp$_3$Am-mRNA-SCy5 isomer 1 (retention time, 17.3 min), 5 and 7: two different batches of FAM-m^{7}Gp$_3$A$_m$-mRNA-SCy5 isomer 2 (retention time, 17.7 min), 8: Riboruler high range Thermofisher. C. Decapping assay monitored by agarose electrophoresis (1.5%, TBE). D. Relative fluorescence intensity as a function of time from 3 replicates of mRNA decapping assay; E. Translation efficiency of unlabeled, monolabeled, or dually labeled mRNA in HeLa cells, the data represent mean values from 6 replicates ± SEM. .

valuable tool for visualizing mRNA and for directly monitoring and quantifying decapping (Figure 6D).

Finally, we asked whether the dually labeled mRNA was suitable for experiments in living cells. We speculated that for this purpose, the mRNA would have to retain its basic biological function, i.e., undergo translation despite the 5′ and 3′ end modifications. Therefore, we studied the expression of dually labeled mRNAs in cultured HeLa cells. As references, we used monolabeled 5′-capped mRNA and uncapped mRNA (the latter of which was expected to be translationally inactive). As shown in Figure 6E, the dually labeled mRNA was translationally active, although the protein expression level was lower than that of monolabeled mRNA. The effect of decreased expression may result either from the presence of a bulky label at the 3′ end or the presence of 3′-phosphate at the RNA terminus. Further experiments are required to clarify this. However, the expression was significantly higher than that of uncapped mRNA, confirming the utility of our mRNA probes for mRNA research.

Conclusion

In conclusion, we developed a straightforward chemo-enzymatic method to provide monolabeled or dually labeled RNA in an efficient and rapid manner based on LT4 and novel pAp analogs. The compounds enabled either direct or two-step labeling of RNA at the 3′ end and were also compatible with cotranscriptional 5′ capping and labeling. Ligation could be applied to short RNAs and long RNAs, including full-length mRNAs. Labeled and biotinylated RNAs obtained using our protocol may serve as excellent probes for studying RNA fate and function in vitro and in vivo. Notably, the RNA$_{35}$ probes developed here have been recently employed to study the activation of decapping in *in vitro* reconstituted biomolecular condensates.[26] The full length, translationally active mRNAs carrying two different labels at the 5′ and 3′ ends, respectively, are better model molecules to monitor the decapping process *in vitro* and are also potentially applicable to more advanced experiments in cell extracts or living cells directed towards finding links among mRNA localization, mRNA turnover, and protein expression and for following the fate of mRNA-based therapeutics.

Acknowledgements

Financial support from the National Science Centre (2015/18/E/ST5/00555 to J.K. and 2019/33/B/ST4/01843 to J.J.) is gratefully acknowledged.

Keywords: RNA labeling • 3'-phosphoadenosine 5'-phosphate • decapping • mRNA probe • T4 ligase •

[1] a) Y. Su, M. C. Hammond, *Curr. Opin. Biotechnol.* **2020**, *63*, 157-166; b) S. Ray, J. R. Widom, N. G. Walter, *Chem Rev* **2018**, *118*, 4120-4155.

[2] a) S. Pitchiaya, L. A. Heinicke, T. C. Custer, N. G. Walter, *Chem. Rev.* **2014**, *114*, 3224-3265; b) G. Bao, W. J. Rhee, A. Tsourkas, *Annu. Rev. Biomed. Eng.* **2009**, *11*, 25-47; c) Y. Xia, R. Zhang, Z. Wang, J. Tian, X. Chen, *Chem. Soc. Rev.* **2017**, *46*, 2824-2843; d) Z. Wang, W. Liu, C. Fan, N. Chen, *Methods* **2019**, *161*, 16-23; e) A. Mamot, P. J. Sikorski, M. Warminski, J. Kowalska, J. Jemielity, *Angew. Chem. Int. Ed. Engl.*

2017, *56*, 15628-15632; *Angew. Chem*. **2017**, *129*, 15834-15838; f) B. Juskowiak, *Anal. Bioanal. Chem.* **2011**, *399*, 3157-3176.

[3] A. Mor, A. White, K. Zhang, M. Thompson, M. Esparza, R. Munoz-Moreno, K. Koide, K. W. Lynch, A. Garcia-Sastre, B. M. Fontoura, *Nat. Microbiol.* **2016**, *1*, 16069.

[4] a) A. Saady, M. Wojtyniak, E. Varon, V. Bottner, N. Kinor, Y. Shav-Tal, C. Ducho, B. Fischer, *Bioconjugate Chem.* **2020**, *31*, 1188-1198; b) R. Verbeke, I. Lentacker, L. Wayteck, K. Breckpot, M. Van Bockstal, B. Descamps, C. Vanhove, S. C. De Smedt, H. Dewitte, *J. Controlled Release* **2017**, *266*, 287-300.

[5] a) J. L. Kirschman, S. Bhosle, D. Vanover, E. L. Blanchard, K. H. Loomis, C. Zurla, K. Murray, B. C. Lam, P. J. Santangelo, *Nucleic Acids Res.* **2017**, *45*, e113; b) N. Muthmann, K. Hartstock, A. Rentmeister, *Wiley Int. Rev.: RNA* **2020**, *11*, e1561.

[6] E. P. Diamandis, T. K. Christopoulos, *Clin. Chem.* **1991**, *37*, 625-636.

[7] a) J. T. Rodgers, P. Patel, J. L. Hennes, S. L. Bolognia, D. P. Mascotti, *Anal. Biochem.* **2000**, *277*, 254-259; b) M. Wilchek, E. A. Bayer, O. Livnah, *Immunol. Lett.* **2006**, *103*, 27-32; c) H. A. Vincent, J. O. Phillips, C. A. Henderson, A. J. Roberts, C. M. Stone, C. E. Mardle, L. E. Butt, D. M. Gowers, A. R. Pickford, A. J. Callaghan, *PLoS One* **2013**, *8*, e79142; d) S. Bednarek, V. Madan, P. J. Sikorski, R. Bartenschlager, J. Kowalska, J. Jemielity, *Philos. Trans. R. Soc., B* **2018**, *373*; e) P. C. Trippier, *Chemmedchem* **2013**, *8*, 190-203.

[8] E. Paredes, M. Evans, S. R. Das, *Methods* **2011**, *54*, 251-259.

[9] a) J. T. George, S. G. Srivatsan, *Methods* **2017**, *120*, 28-38; b) K. Krell, D. Harijan, D. Ganz, L. Doll, H. A. Wagenknecht, *Bioconjugate Chem.* **2020**, *31*, 990-1011; c) O. A. Krasheninina, J. Thaler, M. D. Erlacher, R. Micura, *Angew. Chem. Int. Ed. Engl.* **2021**; *Angew. Chem.* **2021**, *133*, 7046-7050; d) L. Anhauser, A. Rentmeister, *Curr. Opin. Biotechnol.* **2017**, *48*, 69-76; e) K. J. Westerich, K. S. Chandrasekaran, T. Gross-Thebing, N. Kueck, E. Raz, A. Rentmeister, *Chem. Sci.* **2020**, *11*, 3089-3095; f) L. Anhauser, S. Huwel, T. Zobel, A. Rentmeister, *Nucleic Acids Res.* **2019**, *47*, e42; g) M. Zhao, F. D. Steffen, R. Borner, M. F. Schaffer, R. K. O. Sigel, E. Freisinger, *Nucleic Acids Res.* **2018**, *46*, e13; h) S. Croce, S. Serdjukow, T. Carell, T. Frischmuth, *ChemBioChem.* **2020**, *21*, 1641-1646.

[10] O. C. Uhlenbeck, R. I. Gumport, in *The enzymes*, (Ed.: P. D. Boyer), New York: Academic Press, **1982**, p. 31.

[11] a) M. Ghaem Maghami, S. Dey, A. K. Lenz, C. Hobartner, *Angew. Chem. Int. Ed. Engl.* **2020**, *59*, 9335-9339; *Angew. Chem*. **2020**, *132*, 9421 –9425; b) M. Ghaem Maghami, C. P. M. Scheitl, C. Hobartner, *J. Am. Chem. Soc.* **2019**, *141*, 19546-19549; c) A. K. Sharma, J. J. Plant, A. E. Rangel, K. N. Meek, A. J. Anamisis, J. Hollien, J. M. Heemstra, *ACS Chem. Biol.* **2014**, *9*, 1680-1684.

[12] A. G. Bruce, O. C. Uhlenbeck, *Nucleic Acids Res.* **1978**, *5*, 3665-3677.

[13] a) R. Cosstick, L. W. McLaughlin, F. Eckstein, *Nucleic Acids Res.* **1984**, *12*, 1791-1810; b) S. Keyhani, T. Goldau, A. Blumler, A. Heckel, H. Schwalbe, *Angew. Chem. Int. Ed. Engl.* **2018**, *57*, 12017-12021; *Angew. Chem.* **2018**, *130*, 12193-12197.

[14] a) J. H. Ren, D. J. Goss, *Nucleic Acids Res.* **1996**, *24*, 3629-3634; b) M. Ziemniak, M. Szabelski, M. Lukaszewicz, A. Nowicka, E. Darzynkiewicz, R. E. Rhoads, Z. Wieczorek, J. Jemielity, *RSC Adv.* **2013**, *3*.

[15] a) S. Nainar, M. Kubota, C. McNitt, C. Tran, V. V. Popik, R. C. Spitale, *J. Am. Chem. Soc.* **2017**, *139*, 8090-8093; b) N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo, C. R. Bertozzi, *ACS Chem. Biol.* **2006**, *1*, 644-648; c) J. M. Holstein, L. Anhauser, A. Rentmeister, *Angew. Chem. Int. Ed. Engl.* **2016**, *55*, 10899-10903; *Angew. Chem.* **2016**, *128*, 11059-11063; d) U. Pradere, J. Hall, *Bioconjugate Chem.* **2016**, *27*, 681-687; e) M. Jezowska, D. Honcharenko, A. Ghidini, R. Stromberg, M. Honcharenko, *Bioconjugate Chem.* **2016**, *27*, 2620-2628.

[16] D. Ganz, D. Harijan, H.-A. Wagenknecht, *RSC Chem. Biol.* **2020**, *1*, 86-97.

[17] a) N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047; b) M. F. Debets, S. S. van Berkel, S. Schoffelen, F. P. Rutjes, J. C. van Hest, F. L. van Delft, *Chem. Commun.* **2010**, *46*, 97-99.

[18] A. R. Kore, I. Charles, L. Yang, S. Kuersten, *Nucleosides Nucleotides Nucleic Acids* **2009**, *28*, 292-302.

[19] C. Julius, Y. Yuzenkova, *Wiley Int. Rev. RNA* **2019**, *10*, e1512.

[20] T. M. Haas, P. Ebensperger, V. B. Eisenbeis, C. Nopper, T. Durr, N. Jork, N. Steck, C. Jessen-Trefzer, H. J. Jessen, *Chem. Commun.* **2019**, *55*, 5339-5342.

[21] a) Y. G. Chen, W. E. Kowtoniuk, I. Agarwal, Y. Shen, D. R. Liu, *Nat. Chem. Biol.* **2009**, *5*, 879-881; b) D. J. Luciano, J. G. Belasco, *Trends Biochem. Sci.* **2015**, *40*, 245-247; c) H. Cahova, M. L. Winz, K. Hofer, G. Nubel, A. Jaschke, *Nature* **2015**, *519*, 374-377; d) X. Jiao, S. K. Doamekpor, J. G. Bird, B. E. Nickels, L. Tong, R. P. Hart, M. Kiledjian, *Cell* **2017**, *168*, 1015-1027.e1010.

[22] A. Mlynarska-Cieslak, A. Depaix, E. Grudzien-Nogalska, P. J. Sikorski, M. Warminski, M. Kiledjian, J. Jemielity, J. Kowalska, *Org. Lett.* **2018**, *20*, 7650-7655.

[23] a) M. Ishikawa, R. Murai, H. Hagiwara, T. Hoshino, K. Suyama, *Nucleic Acids Symp. Ser. (Oxf.)* **2009**, 129-130; b) P. J. Sikorski, M. Warminski, D. Kubacka, T. Ratajczak, D. Nowis, J. Kowalska, J. Jemielity, *Nucleic Acids Res.* **2020**, *48*, 1607-1626.

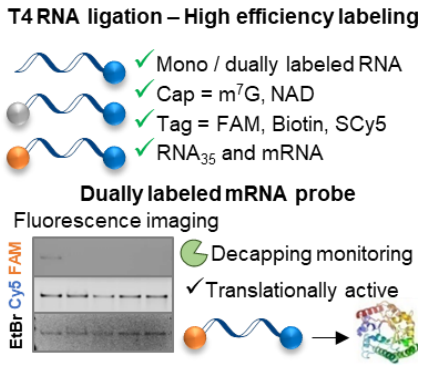
[24] T. M. Coleman, G. Wang, F. Huang, *Nucleic Acids Res.* **2004**, *32*, e14.

[25] a) M. Warminski, P. J. Sikorski, Z. Warminska, M. Lukaszewicz, A. Kropiwnicka, J. Zuberek, E. Darzynkiewicz, J. Kowalska, J. Jemielity, *Bioconjugate Chem.* **2017**, *28*, 1978-1992; b) F. Picard-Jean, C. Brand, M. Tremblay-Letourneau, A. Allaire, M. C. Beaudoin, S. Boudreault, C. Duval, J. Rainville-Sirois, F. Robert, J. Pelletier, B. J. Geiss, M. Bisaillon, *PLoS One* **2018**, *13*, e0193804; c) M. V. Deshmukh, B. N. Jones, D. U. Quang-Dang, J. Flinders, S. N. Floor, C. Kim, J. Jemielity, M. Kalek, E. Darzynkiewicz, J. D. Gross, *Mol Cell* **2008**, *29*, 324-336.

[26] R. W. Tibble, A. Depaix, J. Kowalska, J. Jemielity, J. D. Gross, *Nat. Chem. Biol.* **2021**.

Entry for the Table of Contents

Insert graphic for Table of Contents here.



Labelling of long RNAs, protein-coding RNAs in particular, still poses a challenge. We developed new adenosine 3',5'-bisphophate analogs as reagents for RNA 3'-end labelling by T4 RNA ligase. The compounds were efficient T4 substrates yielding 3'-mono-labelled RNAs with various 5'-ends (triphosphate, NAD). Combination of 3'- and 5'-end labelling yielded dually labeled m^7G-capped RNAs that enabled quantitative analysis of decapping and were translationally active. The proposed method enables the creation of versatile probes for investigation of RNA function and localization.

Institute and/or researcher Twitter usernames: @JoannaK98008579