Chemo-enzymatic treatment of RNA to facilitate analyses

Nils Muthmann © | Katja Hartstock © | Andrea Rentmeister ©

Abstract
Labeling RNA is a recurring problem to make RNA compatible with state-of-the-art methodology and comes in many flavors. Considering only cellular applications, the spectrum still ranges from site-specific labeling of individual transcripts, for example, for live-cell imaging of mRNA trafficking, to metabolic labeling in combination with next generation sequencing to capture dynamic aspects of RNA metabolism on a transcriptome-wide scale. Combining the specificity of RNA-modifying enzymes with non-natural substrates has emerged as a valuable strategy to modify RNA site- or sequence-specifically with functional groups suitable for subsequent bioorthogonal reactions and thus label RNA with reporter moieties such as affinity or fluorescent tags. In this review article, we will cover chemo-enzymatic approaches (a) for in vitro labeling of RNA for application in cells, (b) for treatment of total RNA, and (c) for metabolic labeling of RNA.

This article is categorized under:
- RNA Processing < RNA Editing and Modification
- RNA Methods < RNA Analyses in vitro and In Silico
- RNA Methods < RNA Analyses in Cells

KEYWORDS
chemo-enzymatic, click chemistry, metabolic labeling, RNA labeling, RNA modification

1 | INTRODUCTION

RNA metabolism involves dynamic processes constituting the basis of cellular homeostasis, where a dynamic state of equilibrium is the condition for optimal functioning of the organism. RNA turnover from transcription to decay is usually tightly regulated. Moreover, many RNA species also regulate gene expression as their main or side function, necessitating studies of the different types of RNA in all their facets. This requires reliable methods to label RNAs for studies in cells as the minimal unit of a living system or to selectively target and isolate certain RNA species for subsequent analysis, for example, in next generation sequencing (NGS) or mass spectrometry (MS).

Investigating the trafficking, localization and interactions of RNA in living cells is of key importance to understanding how RNA molecules orchestrate gene expression pathways and their deleterious functions in diseases. Subcellular localization of RNA has been shown to provide a mechanism for controlling gene expression with spatio-temporal precision. RNA modifications have emerged as an additional mode of gene regulation with writer, eraser and reader proteins being discovered for at least some of them (Zhao, Roundtree, & He, 2017). Finally, understanding cells as functional systems requires studies of the...
entirety of nascent RNAs, their interaction partners and following their fate on a transcriptome-wide level to bring us closer to ultimately understand life as such.

Studying RNA is at the same time simple and complicated. There are only four canonical building blocks—a limited set compared to the diversity of amino acids in proteins. However, chemical treatment alone does not yield the required specificity for targeting a dedicated nucleotide—even in a single RNA. This becomes even more evident in total RNA—a mixture of all RNAs isolated from cells.

On the other hand, it is well known that nature is able to modify different types of RNA at specific sites. Enzymes are able to distinguish between different transcripts based on structure and sometimes even based on sequence, if working in complex with guide RNAs. Furthermore, nature uses spatial separation or co-localization to target certain RNAs. Most notably, transcripts of the RNA polymerase II are equipped co-transcriptionally with a 5′ cap using enzymes that are held in place by the C-terminal domain of the RNA polymerase II itself (Ghosh & Lima, 2010).

Researchers can benefit from nature’s tricks to achieve selectivity but also go beyond nature’s repertoire by harnessing non-natural reactions that are orthogonal to the functional groups existing in the cell (i.e., bio-orthogonal). The combination of chemical and enzymatic treatment of RNA provides a powerful strategy to target an RNA of interest (ROI) at specific sites and make it accessible to visualization, isolation or sequence analyses.

In this review article, we highlight chemo-enzymatic methods developed to make RNA accessible to analyses. We will start with in vitro treatment of RNA for the purpose of labeling or manipulating a single RNA species in cells. The challenge here is to achieve site-specificity within transcripts consisting of hundreds of units of the four different nucleotides. In the second part, we will focus on total RNA from cells and highlight chemo-enzymatic approaches enabling enrichment of certain RNA species as well as detection of RNA modifications. Finally, we will turn our attention to chemo-enzymatic approaches to modify RNA in living cells. These approaches are based on feeding cells with a metabolic precursor and aim to study dynamic aspects on a transcriptome-wide level.

2 | IN VITRO TREATMENT OF RNA FOR APPLICATION IN CELLS

This section focuses on covalent labeling of RNA and features chemo-enzymatic labeling approaches for living cells. To put this into context, we will first outline non-covalent strategies in section 2.1. Since covalent labeling strategies mostly rely on click chemistry for the attachment of a reporter molecule, a short introduction into click chemistry will be given in section 2.2.1. After the main part, the overview of covalent RNA labeling strategies, we will present chemo-enzymatic labeling approaches in section 2.2.3. At the end of section 2, we will take a step beyond labeling and discuss how RNA behavior can be altered in the cellular context by chemo-enzymatic treatment.

2.1 | Non-covalent labeling strategies

In this section, we will provide a brief overview of non-covalent RNA labeling strategies based on either hybridization, coordination of a fluorophore by an aptamer or genetically encoded RNA-binding proteins fused to fluorescent reporters.

Hybridization-based approaches were among the first methods for sequence-specific detection of RNA, first reported by Pardue and Gall in 1969 (in situ hybridization [ISH]) (Gall & Pardue, 1969). These strategies rely on radiolabeled (ISH) or fluorescently labeled (FISH) antisense oligonucleotides. Both ISH and FISH are limited to fixed cells, as unbound probe needs to be washed away to reduce background signal (Lawrence & Singer, 1986). Improving on this, molecular beacons (MBs) were developed (Monroy-Contreras & Vaca, 2011; Tyagi & Kramer, 1996). Here, a hairpin-shaped oligonucleotide contains a terminal fluorophore and a quencher at the other end. Upon binding its target, the probe becomes fluorescent. These probes were successfully used for RNA detection in living cells, although delivery remains a bottleneck (Bratu, Cha, Mhlanga, Kramer, & Tyagi, 2003). Additional developments comprise photo-caged MBs for spatio-temporal control of RNA imaging (Joshi, Vlachos, Mikat, Deller, & Heckel, 2012; Qiu et al., 2013). A newer fluorogenic hybridization-based method is forced intercalation (FIT). Here, an antisense oligonucleotide is equipped with one or two cyanine dyes as nucleobase surrogates, showing a marked increase in fluorescence upon hybridization (up to 200-fold) (Hövelmann, Gaspar, Ephrussi, & Seitz, 2013). Dual imaging of two different RNAs in fixed Drosophila egg chambers by super-resolution microscopy has been achieved using this wash-free FISH-method (Hövelmann et al., 2016). Hybridization-based labeling techniques are reviewed in greater detail elsewhere (Li, Ke, & Spitale, 2019). While hybridization-based approaches have proven to be valuable tools for RNA imaging, they give a static picture of RNA localization. However, RNA localization is a dynamic process and thus benefits from tracking studies in living cells. Furthermore, a common disadvantage of all hybridization-based detection
methods is that the probes cannot be produced in the cell but instead require chemical synthesis and transfection or permeabilization for delivery.

Genetically encoded systems circumvent delivery issues and—to date—rely on non-covalent interactions of RNA binding proteins or aptamers, produced in cellula, with a specific region on the ROI. In the field of aptamers, numerous methods have been developed and successfully employed for RNA labeling in living cells. These strategies rely on appending of an aptamer to the ROI, capable of binding a fluorogenic dye or disrupting a fluorophore-quencher pair and thus emitting fluorescence only upon binding by the aptamer. Alternatively, an aptamer specifically designed to bind the ROI in trans hybridizes with the target RNA and binds the respective fluorescent dye only upon hybridization. The most prominent aptamers Spinach, Broccoli, and Mango have been successfully used for labeling the highly abundant 5S-rRNA or U6-snRNAs in living cells (Autour et al., 2018; Dolgosheina et al., 2014; Filonov, Moon, Svensen, & Jaffrey, 2014; Paige, Wu, & Jaffrey, 2011). A disadvantage of aptamer-based approaches is the tag, which has to be genetically fused to the ROI, rendering it no longer endogenous. This tag may perturb the ROI's endogenous interactions, for example, by providing unwanted binding sites for proteins or affecting the fate or turnover of the RNA. In protein-based genetically encoded approaches, RNA-binding proteins fused to fluorescent proteins are harnessed for detection. The well-established MS2 coat-protein has been extensively used for live cell imaging as reviewed in excellent articles (Lee, Bae, Shim, Park, & Park, 2016). Alternative examples comprise Pumilio or PPR (pentatricopeptide repeat) protein based approaches or—as of recently—utilization of Cas proteins for the specific detection of an ROI (Abil, Denard, & Zhao, 2014; Abudayyeh et al., 2017, 2016; Barkan et al., 2012; Coquille et al., 2014; Gootenberg et al., 2017; Kellermann, Rath, & Rentmeister, 2013; Kellermann & Rentmeister, 2016; Nelles et al., 2016; Shen et al., 2015). The RCas9 and Cas13 systems were successfully used in live cell imaging of different mRNAs. Cas-based methods are straightforward, merely requiring the design of a sgRNA (and a PAMmer in the case of dCas9). However, low sensitivity—just a single Cas protein binds to one RNA molecule unless multiple guide RNAs targeting different parts of a transcript are used—and unspecific background of unbound proteins pose additional problems that need to be overcome (Abudayyeh et al., 2017; Nelles et al., 2016). While harnessing RNA-binding proteins is appealing, one drawback remains the size of the appendage which, once the protein-complex is bound, can exceed the size of the ROI itself. It has been shown that appending multiple aptamer sequences to RNA alters its half-life and cellular localization (Haimovich et al., 2016). Additionally, the reversible, non-covalent coordination means that the reporter is subject to an equilibrium between bound and unbound state, with unbound fluorescent protein causing background signal. For a more detailed insight into aptamer- or RNA-binding protein-based RNA detection, we refer the reader to dedicated review articles (Li et al., 2019; Rau & Rentmeister, 2017).

The aforementioned non-covalent RNA-labeling methods have significantly advanced RNA imaging, giving unprecedented insights into RNA localization and trafficking. On the other hand, covalent RNA labeling strategies show several features that render them interesting for RNA imaging, as discussed in the following section.

### 2.2 Covalent RNA labeling approaches

Covalent RNA labeling approaches establish a stable link between the reporter molecule and the ROI. Compared to non-covalent RNA labeling strategies, an equilibrium between bound and unbound probe does not have to be considered, allowing more stringent washing in fixed cells and making covalent labeling attractive for live-cell imaging. Additionally, covalent modifications are commonly smaller in size, minimizing alterations to the endogenous RNAs and thus their biological functions. Fluorescent dyes are small molecules, comparable to the size of a single nucleotide and can have outstanding quantum yields. In the following section, we will focus on covalent labeling strategies applicable to RNA-labeling in living cells with emphasis on chemo-enzymatic modifications.

#### 2.2.1 Introduction into click chemistry

Most covalent RNA-labeling strategies rely on click chemistry to attach the reporter-molecule, justifying a brief introduction into the most commonly used click reactions to give the reader insight into their respective advantages and disadvantages concerning their application on RNA and in living cells. For more comprehensive reading we refer to excellent in-depth review articles (Lang & Chin, 2014; Shieh & Bertozzi, 2014; Shih, Kamber, & Prescher, 2014).

Click chemistry is defined as a set of reactions that are fast, stereoselective and high-yielding in aqueous solution, rendering them ideal for the labeling of biomolecules (Kolb, Finn, & Sharpless, 2001). Application of click chemistry in living cells demands these reactions to not only be biocompatible (non-toxic and functional under biological conditions) but to be bio-orthogonal (i.e., to not interfere with biological processes) (Sletten & Bertozzi, 2011). Additionally, fast reaction kinetics are
desirable to label low abundance biomolecules in vivo (characterized by the second order rate constant $k_2 \text{[M}^{-1}\text{s}^{-1}]$). The most commonly employed click reaction is the copper(I)-catalyzed azide-alkyne cycloaddition (CuAAC) due to easily accessible functional groups and its fast kinetics (rate constants of 10–200 M$^{-1}$s$^{-1}$, Figure 1a) and has been employed for labeling of glycans, proteins and nucleic acids (Presolski, Hong, Cho, & Finn, 2010). Despite recent efforts in the development of Cu(I)-stabilizing ligands and reports of CuAAC in living cells (Bevilacqua et al., 2014; Li et al., 2017), the cytotoxicity of copper(I) and radical-induced RNA degradation by Cu(I) oxidation are challenges toward its application for live-cell imaging. This shortcoming has inspired researchers to identify truly bioorthogonal copper-free click reactions. To date, the most promising alternatives comprise the strain-promoted azide-alkyne cycloaddition (SPAAC, Figure 1b) and the inverse electron-demand Diels-Alder cycloaddition (iEDDA, Figure 1c). In the SPAAC reaction, an azide is reacted with a strained alkyne. Here, dibenzocyclooctyne derivatives offer the fastest kinetics (Dommerholt, Rutjes, & van Delft, 2016). While enzymatic transfer of the cyclooctyne derivatives is difficult due to their size, azides suffer from limited stability under reducing conditions as found in cells (Cartwright, Hutchinson, & Armstrong, 1976). Furthermore, recent studies point toward cross-reaction of the cyclooctyne moiety with intracellular thiols (Fairbanks, Sims, Anseth, & Bowman, 2010). The iEDDA reaction—first described in 2008—is currently the most promising bioorthogonal reaction for RNA labeling due to its fast kinetics. Here, a tetrazine derivative reacts with a strained alkene. The reaction kinetics depend on the respective alkene and tetrazine derivatives, with trans-cyclooctene being most reactive. A further advantage of the iEDDA reaction with respect to cellular applications is the availability of several turn-on fluorophores. While the unreacted tetrazine can act as a quencher of the dye, its quenching properties are lost upon reaction, allowing for a fluorescence turn-on of up to 1,600-fold. Numerous turn-on dyes have been reported to date, including far-red-emitting fluorophores (Meimetis, Carlson, Giedt, Kohler, & Weissleder, 2014; Wieczorek, Werther, Euchner, & Wombacher, 2017). An important property to consider for their application in a cellular context is the water-solubility of the dye. While multiple click reactions have successfully been used on RNA, the SPAAC and iEDDA reactions are the most promising ones for in vivo labeling of RNA as exemplified by research presented in the following sections.

| Rate constant [M$^{-1}$s$^{-1}$] | Slow | Fast |
|----------------------------------|------|------|
| $10^{-2}$                        |      |      |
| $10^{-1}$                        |      |      |
| $10$                             |      |      |
| $10^2$                           |      |      |
| $10^3$                           |      |      |
| $10^4$                           |      |      |
| $10^5$                           |      |      |
| $10^6$                           |      |      |

**FIGURE 1** Schematic overview of a select set of most commonly employed click reactions, sorted according to their second order rate constants ($k_2 \text{[M}^{-1}\text{s}^{-1}]$). (a) Copper-catalyzed azide-alkyne cycloaddition (CuAAC), exhibiting rate constants from 10–200 M$^{-1}$s$^{-1}$ (Lang & Chin, 2014). (b) Reaction scheme of the strain-promoted azide-alkyne cycloaddition (SPAAC; upper panel) and examples for recently described cyclooctynes used in SPAAC reactions (lower panel). Fastest cyclooctynes such as aza-dibenzocyclooctyne (DIBAC) or biarylazacyclooctyne (BARAC) show rate constants from 0.1–4 M$^{-1}$s$^{-1}$ (Lang & Chin, 2014). (c) Reaction scheme of the inverse-electron demand Diels-Alder cycloaddition (iEDDA, upper panel) and examples of dienophiles (lower panel). For iEDDA reactions, reaction rates generally increase from norbornenes, methycycloprenes, bicyclononynes to trans-cyclooctenes and range from 10–10$^6$ M$^{-1}$s$^{-1}$ (Blackman, Royzen, & Fox, 2008; Kamber et al., 2013; Lang et al., 2012; Schoch, Staadt, Samanta, Wiessler, & Jäschke, 2012; Taylor, Blackman, Dmitrenko, & Fox, 2011; Vrabel et al., 2013). [4 + 2] DA, [4 + 2] Diels-Alder-reaction; [4 + 2] retro-DA, [4 + 2] retro-Diels-Alder reaction
2.2.2 | Incorporation of non-natural building blocks

In this section, RNA labeling approaches based on the chemical or enzymatic incorporation of non-natural nucleotides and strategies based on ribozymes are discussed. For chemical integration of non-natural nucleotides, the respective phosphoramidites are incorporated during solid-phase synthesis (Figure 2a) into the ROI. As chemical synthesis is not limited by the size of the nucleobases, they can either be equipped with a functional handle for subsequent derivatization with a reporter molecule or bear the desired properties intrinsically (e.g., fluorescing nucleobases). One noteworthy example is the incorporation of a norbornene-modified uridine analog into RNA via solid-phase synthesis and subsequent tetrazine ligation in living cells (Pyka, Domnick, Braun, & Kath-Schorr, 2014). A more detailed discussion of chemical incorporation of non-natural nucleotides is subject of a recent review (Holstein & Rentmeister, 2016). While chemical synthesis allows for high flexibility regarding the position and nature of the non-natural nucleotide, synthesis of RNAs longer than 50 nt remains challenging, necessitating enzymatic ligation (Keyhani, Goldau, Blümler, Heckel, & Schwalbe, 2018). Enzymatic production of modified RNA does not face length restrictions (Figure 2b). However, for efficient enzymatic incorporation during in vitro transcription (IVT), the size and position of the non-natural nucleotides functional group is a concern in order to be accepted by the polymerase (detailed review in references Anhäuser & Rentmeister, 2017; Holstein & Rentmeister, 2016). Consequently, the functional moiety is most often a reactive handle for post-transcriptional functionalization in a click reaction. As RNA containing non-natural nucleotides cannot be produced in cellula, delivery via transfection is required.

Metabolic labeling of all RNA species is achieved if suitable non-natural nucleosides are fed to the cell (Figure 2c). Key requirement here is the uptake of the non-natural nucleoside into the nucleoside salvage pathway for production of the modified nucleotide in the cell. However, for site-specific labeling of RNA this approach is not useful, as these non-natural nucleotides are incorporated statistically into all nascent RNAs (see section 4).

Transcript-specific labeling of RNA has been achieved via non-natural base pairs (UBPs). Originally developed for the expansion of the genetic code as a tool for synthetic biology (as reviewed in Hamashima, Kimoto, & Hirao, 2018), UBPs were recently adopted for the labeling and imaging of RNA. Non-natural nucleotides are incorporated into RNA during transcription at the site of the UBP in the DNA-template, rendering it a promising strategy for site-specific RNA labeling. These non-natural nucleotides were equipped with a norbornene or a cyclopropene and could be labeled in tetrazine ligations with a fluorogenic dye in vitro and also in living cells (Domnick, Eggert, & Kath-Schorr, 2015; Eggert & Kath-Schorr, 2016). An important step toward application of unnatural base pairs for RNA labeling in vivo was presented by the Romesberg group in 2017.

![Figure 2](image-url)
Here, for the first time, transcription of an unnatural base-pair into RNA, containing the respective unnatural nucleotide, by the T7-polymerase and its translation into a modified protein containing a non-canonical amino acid were demonstrated in prokaryotic cells (Zhang et al., 2017). Importantly, both the dNTPs and NTPs of the UBPs have to be delivered into Escherichia coli, which was achieved by recombinant expression of the nucleoside triphosphate transporter from Phaeodactylum tricornutum (PtNTT2) (Feldman et al., 2018). For application in eukaryotic cells, a system for efficient delivery of the non-natural (d)NTPs will have to be established.

A promising approach for covalent site-specific labeling that does not require (protein) enzymes or click reactions are self-alkylating ribozymes that can be attached to the ROI or DNazymes used in trans (Baum & Silverman, 2007; Büttner, Javadi-Zarnaghi, & Höbartner, 2014). Thus far, three self-alkylating ribozymes have been reported. One example is the de novo selection of a ribozyme able to transfer a fluorophore—fluorescein iodoacetamide—onto itself. Fusion of this ribozyme to a ROI allowed for specific labeling in cell lysate (Sharma et al., 2014). In an alternative approach, a 41 nt long RNA was selected, capable of transferring 2,3-disubstituted epoxides specifically onto a guanosine, forming a C-N-bond at the N7-position. The range of epoxide substrates included biotinylated, azide- and alkyne-modified substrates and one directly equipped with a tetramethylrhodamine fluorescent dye (TAMRA) (McDonald et al., 2014). The ability to specifically modify a ROI was demonstrated by transfecting HEK cells with plasmid DNA coding for 5S-rRNA fused to the ribozyme. Isolated total RNA was incubated with the azido-containing epoxide substrate and subsequently clicked in a SPAAC reaction with a DBCO-TAMRA dye. The 5S-rRNA-ribozyme construct was found to be specifically labeled within the total RNA, with the fluorescence intensity directly correlating to the number of ribozymes attached to the 5S-rRNA. A third example exploits the chemical promiscuity of a polymerase ribozyme, capable of transferring multiple functionalized or fluorophore-modified nucleotide analogs onto the 3'-end of RNA in a template-directed manner. However, for its application in cells, further selection is required to reduce competition with endogenous nucleotides and allow for exclusive transfer of a single nucleotide analog (Samanta, Horning, & Joyce, 2018). None of the ribozyme-based labeling reactions are fluorogenic, necessitating the removal of unreacted reporter. Consequently, they are likely limited to fixed cells and—at least to date—their applicability for in-cell labeling has not been demonstrated.

### 2.2.3 Chemo-enzymatic labeling approaches

In this section, chemo-enzymatic labeling strategies are presented. Herein, a small functional moiety is first transferred enzymatically from a non-natural cosubstrate onto the biomolecule of interest and subsequently reacted with a reporter molecule via click chemistry. Chemo-enzymatic approaches are widely used for site-specific labeling of biomolecules, with numerous examples in the fields of small molecule (Law, Struck, Bennett, Wilkinson, & Micklefield, 2015; Singh et al., 2014), protein (Willnow, Martin, Lüscher, & Weinhold, 2012) and DNA labeling (Dalhoff, Lukinavičius, Klimašauskas, & Weinhold, 2006; Lukinavičius et al., 2007). This strategy combines the high specificity of enzymes (allowing for site-specific labeling) with the chemical repertoire of high-yielding and fast bioorthogonal click reactions and flexibility regarding the label. Furthermore, enzymatic modifications of the ROI take advantage of the fact that enzymes are predestined to function under physiological conditions, an asset for intracellular labeling. Recent improvements in reaction speed and biocompatibility of bioorthogonal reactions have paved the way for labeling of low abundance biomolecules in living cells. In order to be appropriate for the site-specific labeling of a specific transcript in cells, an ideal labeling approach should meet the following criteria:

1. To eliminate the need for transfection, all components should be taken up or produced by the cell.
2. To ensure the specificity of the labeling approach, the enzyme should only modify a single class of RNAs or even a specific transcript.
3. To achieve specific labeling of the ROI, an orthogonal pair of modifying enzyme and non-natural cosubstrate is required. The enzyme should be engineered to specifically transfer the non-natural moiety and not the natural cosubstrate while the cosubstrate analog should not be transferred by other endogenous enzymes.
4. To avoid deleterious effects of the click reaction, bio- and RNA compatibility of the click reaction are required, excluding copper-catalyzed click reactions (CuAAC).

To date, chemo-enzymatic RNA labeling is based on either methyltransferases (MTases) or tRNA-modifying enzymes. As tRNAs are subject to numerous post-transcriptional modifications, the corresponding tRNA-modifying enzymes bear potential for structure-specific RNA labeling. Thus, a minimal recognition motif based on tRNA typically has to be appended to the
ROI. Methyltransferases often recognize very small structural entities as a highly specific target moiety. In this section, we will present current approaches for chemo-enzymatic labeling of RNA and discuss their applicability to live-cell imaging.

One example for specific labeling of a ROI is the utilization of the tRNA[^ile2]-agmatidine synthetase (TIAS) transferring agmatidine (AGM) onto the tRNA[^ile2] in archaea, which does not exist in mammals. The cosubstrate promiscuity of TIAS was exploited and three different synthetic AGM analogs bearing functional handles for click chemistry were successfully transferred onto tRNA[^ile2]. Subsequently they were labeled with a fluorescent reporter via SPAAC or CuAAC (Li et al., 2015). 5S-RNA fused to tRNA[^ile2] was detected in fixed and permeabilized cells via CuAAC.

Similarly, in an enzymatic approach to label RNA in cells, the bacterial tRNA-modifying enzyme tRNA guanine transglycosylase (TGT) from *E. coli* was used (Figure 3a). Prokaryotic TGTs naturally perform a transglycosylation reaction in which a guanine is exchanged with an amine-containing queuine precursor, PreQ₁. TGT was found to exhibit a remarkable cosubstrate promiscuity, directly transferring fluorescent or biotin-modified PreQ₁ analogs specifically onto a 17 nt long minimal structural hairpin-motif and therefore eliminating the need for a subsequent chemical reaction (Alexander, Busby, Cole, Zhou, & Devaraj, 2015). Rational design of a thiazole orange PreQ₁-analog allowed for fluorogenic labeling of an mCherry construct containing the TGT-recognition element, with 100-fold fluorescence increase, likely due to intercalation into the hairpin structure of the 17 nt motif (Zhou, Alexander, & Devaraj, 2017). With this probe, wash-free labeling in fixed CHO cells was achieved. However, labeling of RNA in living cells, co-transfected with the TGT- and mCherry-construct plasmids was unsuccessful, likely due to difficulties concerning delivery of the PreQ₁-analog at the required concentration. Both

![FIGURE 3](https://example.com/figure3.png)

**FIGURE 3** Chemo-enzymatic labeling strategies for RNA in fixed and/or living cells. (a) TGT from *Escherichia coli* transfers fluorescent PreQ₁ analogs onto a small 17 nt hairpin motif allowing for labeling in fixed cells. (b) Poly(A)-polymerase transfers azido-ATP onto the 3′-end of mRNA. This strategy allowed for live and fixed cell labeling of RNA. (c) Methyltransferase (MTase) based labeling. MTases GlaTgs and Ecm1 site-specifically modify the 5′-cap. Here, labeling in fixed and living cells has been achieved. CD small RNP RNA 2′-O-MTase directs template-dependent sequence-specific internal modification. HEN1 MTase from *Drosophila melanogaster* modifies the 3′-end of all RNAs. (d) The tRNA MTase TRM1 modifies an internal position of tRNA[^phe]. (e) HEN1 MTase from *Arabidopsis thaliana* modifies the 3′-termini of dsRNA.
Aftementioned strategies require the appendage of a specific RNA sequence (or rather structural element) to the ROI, rendering it no longer endogenous. Nevertheless, these tags are markedly smaller than the established MS2-MCP-system.

As basis for another enzymatic modification strategy, methyltransferases (MTases) catalyze the transfer of a methyl group from the highly abundant cosubstrate S-adenosyl-L-methionine (AdoMet or SAM)—the second most ubiquitous cosubstrate after ATP and the major source of methylations in biomolecules—onto various biomolecules such as proteins, DNA and RNA with great specificity (Fontecave, Atta, & Mulliez, 2004). As some MTases show promiscuous activities on AdoMet analogs, which can be readily synthesized, they are attractive for site-specific post-synthetic modification of complex biomolecules, as established for DNA and protein labeling (Lukinavičius et al., 2007; Peters et al., 2010; Wang, Zheng, Yu, Deng, & Luo, 2011). AdoMet analogs with various functional moieties for MTase-based labeling strategies have been reported (Dalhoff et al., 2006; Law et al., 2015; Lukinavičius et al., 2007; Singh et al., 2014; Willnow et al., 2012). Important structural motifs for AdoMet analogs include an unsaturated bond in β-position to the sulfonium center for efficient transfer in an SN2 mechanism (Muttach, Muthmann, Reichert, Anhäuser, & Rentmeister, 2017). Replacing the sulfur by a selenium atom enhances stability toward epimerization while simultaneously the reactivity in MTase-based reactions is increased for certain AdoMet analogs due to the lengthened chalcogen-carbon bond, as demonstrated for the propargylic AdoMet analog SeAdoYn (Bothwell & Luo, 2014; Iwig & Booker, 2004; Willnow et al., 2012).

The MTase Trm1 catalyzes methyl transfer to the N2-atom of guanine 26 in tRNA_{Phe}. A pentenyl group from the corresponding AdoMet analog was transferred onto the tRNA and then fluorescently labeled via CuAAC allowing for single-molecule experiments (Figure 3d) (Motorin et al., 2011). As in the aforementioned strategies, this approach necessitates appendage of a tRNA to the target RNA. Sequence-specific RNA labeling has been achieved with box C/D ribonucleoprotein complexes (C/D RNPs). C/D RNPs guide AdoMet-dependent site-specific 2′-O-methylation to various biological targets in rRNA, tRNA and spliceosomal RNA. Specific design of the guide RNA allowed for sequence-specific transfer of a propargyl group from SeAdoYn onto in vitro transcribed pre-mRNA (Figure 3c) (Tomkuvienė, Clouet-d-Cerniauskas, Weinhold, & Klimašauskas, 2012). To date both, Trm1- and box C/D RNP-based labeling depend on terminal alkyne groups and therefore CuAAC, precluding application in cells.

Terminal labeling of double stranded RNA (dsRNA) was successful with the MTase HEN1, which naturally acts on miRNAs or siRNAs in plants, where the 3′-ends of double-stranded RNAs are methylated at the 2′-hydroxy group. HEN1 now enabled modification at the 3′-ends with amino and alkyne groups from the respective AdoMet analogs (Figure 3d). Subsequent labeling was performed using N-hydroxysuccinimide (NHS)-ester or CuAAC chemistry. HEN1 even accepted bulky side chains containing a biotin moiety, allowing for direct and efficient labeling, followed by capture and enrichment of the targeted RNA species—obviating a dedicated chemical functionalization step (Plotnikova, Osipenko, Masevicius, Vilkaitis, & Klimašauskas, 2014). Recently, similar results have been reported for Drosophila and human HEN1 MTases, which methylate all ssRNAs (Figure 3c) (Mickutė et al., 2018). Remarkably, DmHEN1 showed no preference for AdoMet over an azido-modified AdoMet analog. Furthermore, direct one-step labeling with a fluorophore or biotin was achieved from suitable AdoMet analogs. Importantly, labeling in total RNA was successful, however, the suitability of this approach for RNA labeling in cells was not tested. Potential RNA enrichment applications for sequencing will be discussed in section 3.

Our lab presented a chemo-enzymatic strategy for the modification of the 5′-cap—a hallmark of eukaryotic mRNAs. Trimethylguanosine synthases (Tgs) methylate the 5′-cap at their N2 position. Engineered variants of the Tgs from the organisms Giardia lamblia and Giardia intestinalis (GlaTgs2-V34A and GinTgs-T34A) exhibited substrate-promiscuity toward AdoMet analogs bearing allyl and pentenynyl moieties (Holstein, Stummer, & Rentmeister, 2015a). Labeling of the 5′-cap was achieved using CuAAC and thiol-ene click chemistry (Figure 3c) (Schulz, Holstein, & Rentmeister, 2013). Improving on this, transfer of functional groups for copper-free chemistry was achieved, allowing for RNA labeling by SPAAC, photoclick and iEDDA reactions Holstein, Schulz, & Rentmeister, 2014; Holstein, Stummer, & Rentmeister, 2015b). Importantly, SPAAC labeling of RNA was achieved in living HeLa cells after transfection (Holstein, Anhäuser, & Rentmeister, 2016). Transfer from AdoMet analogs with small side-chains limits the scope of available click chemistry for labeling. The N7-cap-methyltransferase from the parasite Encephalitozoon cuniculi (Ecm1) shows high co-substrate promiscuity, transferring even benzylic side-chains from corresponding AdoMet analogs with comparable efficiency to natural AdoMet. Using Ecm1, transfer of a norbornene attached to a benzylic linker was successful, allowing labeling in a fast and fluorogenic iEDDA reaction in eukaryotic cell lysate (Muttach, Muthmann, Reichert, et al., 2017). A limiting factor for all MTase-based RNA-labeling strategies toward their application in cells is the poor cell permeability and stability of AdoMet and its analogs. We addressed this problem by producing AdoMet biosynthetically. In nature, AdoMet is produced by methionine-adenosyltransferases (MAT) from methionine and ATP (Markham & Pajares, 2009). A previously engineered human MAT variant can produce enantiopure AdoMet analogs from corresponding methionine analogs, whereas chemical synthesis yields a diastereomeric
mixture with only the sulfonium-(S) epimer being bioactive (Wang et al., 2013). This strategy was adopted for the labeling of mRNA. AdoMet analogs were produced in situ from the respective methionine analogs by a MAT variant and then directly converted by the MTase GlaTgs2-V34A to yield modified 5'-caps. Subsequent labeling was achieved in eukaryotic cell lysate in a one-pot reaction (Muttach & Rentmeister, 2016a). The strategy also forms the basis of metabolic labeling starting from methionine analogs (see section 4).

Introduction of azido-modified ATP into the poly(A) tail—another hallmark of eukaryotic mRNAs—allowed for subsequent labeling in a SPAAC reaction (Figure 3b) (Anhäuser, Hüwel, Zobel, & Rentmeister, 2019; Winz, Samanta, Benzing, & Jäschke, 2012). Depending on the ATP analog and the poly(A)-polymerase used, either one or multiple labels could be attached to the ROI. HeLa cells were transfected with poly(A) tail modified RNA and imaged at different time points in living cells. RNA imaging was not only possible after transfection with the fluorescently modified RNA but also after transfection with azido-modified RNA and SPAAC reaction in living cells. Colocalization studies with a FISH-probe in fixed cells unambiguously showed specific labeling of eGFP-mRNA. Importantly, in addition to labeling it was shown that 3'-azido ATP as well as 3'-fluorescently modified mRNA increased the translation efficiency. This feature will be discussed in more detail in the following section.

2.3 Chemo-enzymatic treatment for altering properties

In the previous section, we elaborated on the versatility and potential of chemo-enzymatic approaches for RNA labeling in living cells. Now, we will focus on applications beyond labeling, namely, how chemo-enzymatic modification can be harnessed to alter and control the properties of RNA in the cellular context. This emerging field includes effects on all aspects of RNA turnover and functions as demonstrated already for RNA stability, translation and interaction with binding partners.

Hybridization-based strategies for translational control traditionally consist of microRNAs that bind to either the 3'- or 5'-untranslated region (Fabian, Sonenberg, & Filipowicz, 2010). Further developments include modification of these antisense oligonucleotides with photocleavable groups for temporal control of translation (Yamazoe, Liu, McQuade, Deiters, & Chen, 2010). Recently, covalent strategies for direct translational control of an mROI have emerged. One example for using chemo-enzymatic reactions for altering properties of the target RNA was presented by Holstein et al. Here, various functional moieties including benzylic, allylic, propargylic, and azido-groups were transferred to either the enzymatic reactions for altering properties of the target RNA was presented by Holstein et al. Here, various functional moieties into the triphosphate bridge or the respective 2'-positions of the cap analogs allowed for upregulation of translation (Kore, Shanmugasundaram, Charles, Vlassov, & Barta, 2009; Rydzik et al., 2009). Introduction of different chemical moieties into the triphosphate bridge or the respective 2'/3'-positions of the cap analogs allowed for upregulation of translation in cells, increased resistance against decapping enzymes and modulation of eIF4E binding (Figure 4e) (Kocmik et al., 2018; Rydzik et al., 2017; Wojtczak et al., 2018). Translational control by chemo-enzymatic modifications as demonstrated by these examples has the potential to become a suitable tool for non-invasive and spatio-temporal regulation of gene expression at the post-transcriptional level with prospects for therapeutic applications (potential therapeutic applications of artificial cap-analogs are reviewed in detail [Ziemniak, Strenkowska, Kowalska, & Jemieliity, 2013]). This is highlighted by the fact that artificially capped mRNAs containing a 5'-phosphorothioate cap are investigated in clinical trials against melanoma (Ziemniak et al., 2013).

Besides altering translation efficiency and stability of mRNAs, modifications installed at the RNA cap also allowed for photo-cross-linking with interacting proteins or nucleic acids. The above-mentioned MTase-based cap-modifying system has been used to capture transient cap-protein interactions. Ecm1 is able to transfer various photo-cross-linkers onto the 5'-cap.
which could then be cross-linked to cap-interacting proteins upon irradiation (Muttach, Mäsing, Studer, & Rentmeister, 2017).

Similarly, a synthetic cap analog equipped with a 6-thioguanosine that could be incorporated via IVT was synthesized and used for photo-cross-linking (Figure 4d) (Nowakowska et al., 2014).

### 3 | ENRICHMENT AND DETECTION OF RNA MODIFICATIONS IN ISOLATED TOTAL RNA

Total RNA from cells contains various RNA species differing in length, structure, abundance, stability, and modification status. Mass spectrometry (MS) is a sensitive technique for direct detection of RNA modifications, however, sequence-information and—depending on the efficiency of separation—even RNA species information is lost. While direct (nanopore sequencing) and indirect sequencing (NGS) of total RNA is possible, tRNAs and rRNAs dominate those reads as they are the most abundant RNA species. To gain information about low abundant RNAs like mRNA, enrichment is required prior to analyses. To enrich a certain RNA species or even a specific RNA from total RNA, several methods are available.

Most notably, eukaryotic mRNA is routinely isolated from total RNA via its poly(A) tail using hybridization to a carrier-bound poly(dT) oligonucleotide in various commercialized set-ups. Typically, this procedure is performed twice to efficiently reduce the amount of rRNAs and tRNAs (Rio, Ares, Hannon, & Nilsen, 2010).

Similarly, hybridization with sequence-specific oligonucleotides can be used to enrich or deplete other sequences, as exemplified by the purification of viral RNA from SV40 infected cells in 1976 (based on specific DNA linked onto a cellulose
matrix) or the more recent rRNA depletion (based on rRNA-specific oligodeoxynucleotides) (Pang et al., 2004; Shih & Khoury, 1976; Su & Sordillo, 1998).

Another way to isolate specific RNA is size exclusion. Here, a silica-based stationary phase is used to immobilize RNA species within a certain size range via non-covalent interactions under carefully adjusted buffer conditions (e.g., miRNA and siRNA with commercially available kits). All these methods must be optimized for the requirements of the targeted RNA

FIGURE 5 Legend on next page.
species, rendering the isolation of specific RNAs especially challenging. In section 3.1, we will discuss how chemo-enzymatic strategies complement the existing toolbox of RNA isolation strategies, enabling the isolation of specifically modified RNA species inaccessible by the renowned techniques (Figure 5).

### 3.1 RNA-modifying enzymes for chemo-enzymatic modification in total RNA

As highlighted in the previous section, RNA-modifying enzymes provide a versatile toolbox, allowing to specifically target a broad variety of RNAs, ranging from specific RNA species like miRNA over RNA cap structures to internal RNA modifications. A plethora of different modifications can be transferred onto the RNA of choice from non-natural cosubstrates, allowing subsequent derivatization tailored to the desired downstream application (e.g., biotin for enrichment, labeling with fluorophores or photocaging groups). In this section, we focus on approaches where RNA-modifying enzymes have been used to target specific RNA species in the complex mixture of total RNA.

One successful example is the NAD⁺ (nicotinamide adenine dinucleotide) captureSeq strategy. While the presence of NAD⁺-caps in prokaryotic RNA was proven via LC–MS (Chen, Kowtoniuk, Agarwal, Shen, & Liu, 2009), identification of specific RNA sequences and therefore studies of their function was not possible. To identify NAD⁺-capped RNAs, the ability of adenosine diphosphate ribosyl cyclase (ADPRC) to catalyze a trans-glycosylation, replacing nicotinamide of NAD⁺-RNA with an alkynyl alcohol, was employed. The enzyme acts on the labile N-glycosidic bond of NAD⁺ but not on the canonical purine and pyrimidine nucleotides. Further functionalization via CuAAC with a biotin moiety allowed enrichment of the target RNAs using streptavidin beads. Subsequent NGS analysis revealed specific regulatory small RNAs (sRNA) and certain purine and pyrimidine nucleotides. Further functionalization via CuAAC with a biotin moiety allowed enrichment of the target RNAs using streptavidin beads. Subsequent NGS analysis revealed specific regulatory small RNAs (sRNA) and certain mRNA fragments as NAD⁺-cap carriers. Additional in vitro experiments with the identified RNAs showed that the cap-like NAD⁺ structure influenced stability of the respective RNA and revealed a decapping mechanism formerly unknown for prokaryotic RNAs (Cahova, Winz, Hofer, Nubel, & Jaschke, 2015; Winz et al., 2016). Recently, an adaption of this strategy for nanopore sequencing was published (Zhang et al., 2019). Following the same trans-glycosylation step as the original technique, an RNA sequence is introduced via copper click reaction, serving as label for subsequent enrichment with the complementary DNA probe. This allows for direct RNA sequencing without the need for reverse transcription into cDNA.

Similarly, some of the RNA methyltransferases mentioned in section 2 have been used together with AdoMet analogs to target RNA species in total RNA. As described above (section 2.2.3), the plant MTase HEN1 was shown to transfer different functional groups onto the 3’ end of miRNA/miRNA* and siRNA/siRNA* homoduplexes using synthetic AdoMet analogs (Plotnikova et al., 2014). Here, transcript-specific RNA enrichment was achieved by adding an antisense DNA. Biotin moieties were then either installed directly from corresponding AdoMet analogs or linked to the transferred moiety in a click reaction. Addition of isolated E. coli total RNA did not alter the modification efficiency, demonstrating that this DNA-guided strategy presents a way to label and enrich specific RNA from complex RNA mixtures. However, background reactions from native miRNA/siRNA homoduplexes might arise. In contrast to the isolation via substrate-bound, complementary oligonucleotides, the 3’-end of ssRNA is particularly interesting for sequencing applications. While the wild-type dmHen1 enzyme preferred a uridine in 3’ position, a truncated version showed no bias—a prerequisite for NGS applications (Mickuté et al., 2018). Attaching a biotin handle to all ssRNA molecules shortens sample preparation for NGS, where numerous purification steps are required for library preparation.

**FIGURE 5** Scheme for enrichment and detection of specific RNAs or RNA modifications from total RNA. (a) NAD⁺ captureSeq: Trans-glycosylation of nicotinamide adenine dinucleotide (NAD⁺) modified RNA by adenosine diphosphate ribosyl cyclase (ADPRC) with an alkynyl alcohol in total RNA. The introduced alkynyl group can be conjugated in a click reaction to introduce a biotin label for subsequent enrichment and determination of NAD⁺-capped RNA. (b) Identification of N⁶-methyladenosine (m⁶A): Incorporation of non-natural N⁶-modifications at free METTL3-METTL14 target sites using AdoMet analogs. The introduced modification can be used to either perform cyclization via iodination or to add a biotin label via CuAAC for subsequent enrichment. After reverse transcription (RT) the introduced modifications result in a specific RT pattern (either termination or mutation) which allows to identify the modified position in sequencing. (c) Identification of pseudouridine (Ψ): Treatment with N⁺-CMC installs a clickable moiety on N3 of many pyrimidines and N1 of many purines. Alkaline treatment removes all moieties which are not installed on the Ψ. Remaining moieties are further functionalized with biotin via click reaction, enabling enrichment. RT of the treated sample results in a specific termination pattern while untreated samples result in full length product enabling the detection of pseudouridines. (d) Identification of N1-methyl adenosine (m¹A): Alkaline treatment of m¹A triggers Dimroth rearrangement resulting in the formation of m¹A. RT of the untreated sample results in a specific termination pattern while rearranged samples result in full length product enabling the detection of m¹A.
The promiscuity of MTases can also help detect their target sites in RNA. This is especially important for methylation sites that do not interfere with base pairing during reverse transcription (RT) and thus do not lead to misincorporations or specific termination in the NGS readout. These RT-silent modifications include m^3C (Squires et al., 2012), 2′-O-methylations (Marchand, Blanloeil-Oillo, Helm, & Motorin, 2016) and m^6A—the most abundant internal modifications in eukaryotic mRNA.

The methylation in N^6-position of adenosine has been associated with a variety of biological functions like RNA splicing (Louloupi, Ntini, Conrad, & Örom, 2018), translation (Wang et al., 2015) and degradation (Guo, Liu, Zheng, Huang, & Chen, 2017; Wang et al., 2014), inciting interest in the precise localization and stoichiometry of this methylation on a transcriptome-wide level. While the most widely employed techniques to detect m^6A (namely miCLIP or m^6A-CLIP) are based on immunoprecipitation and crosslinking, alternative direct techniques have been developed recently to avoid possible bias by the employed antibodies. In a chemo-enzymatic approach, RNA was modified using recombinant METTL3/METTL14 and synthetic AdoMet analogs (Hartstock et al., 2018; Shu et al., 2017). Specifically, the AdoMet analogs allyl-SAM and SeAdoYn were shown to be substrates of METTL3/METTL14 for transfer of an alkene or alkyne to the 6-position of unmodified adenosine at the preferred target sites. In the case of N^6-allyladenosine, iodination of the allyl double bond led to spontaneous cyclization between N1 and N6 of the adenosine. This modification caused up to 70% of mutation during RT, enabling its identification in NGS by comparison with an untreated sample. The method was applied to total RNA from HeLa cells and showed ~17% of mutation after RT at a reported m^6A site in ACTB in sequencing (Shu et al., 2017). Incorporation of a propargyl group at the N^6-position can even be used to attach a biotin label, allowing enrichment of successfully modified RNA species. The streptavidin-bound RNA caused up to 65% of termination in the positions directly before the modification (Hartstock et al., 2018). As only unmethylated adenosines can be modified by the enzyme, the amount of incorporated allyl modifications could—in principle—be used to quantify the methylation status of the investigated positions. However, here the efficiency of allyl incorporation needs to be improved. A more promising strategy might be the metabolic labeling of methylation sites via AdoMet analog precursors which will be discussed in section 4.

3.2 Chemical conversion of existing RNA modifications

Even though the enzymatic installation of non-natural modifications onto the unoccupied target sites allows to indirectly identify the modification status, chemical conversion of the existing RNA modifications into a “readable” moiety would be preferable, as it represents a direct way to label and detect them. This strategy was successfully realized in a number of recent mapping techniques, most notably for transcriptome-wide detection and quantification of m^1C using bisulfite m^5C-Seq (Squires et al., 2012). Additional examples for naturally RT silent modifications which can be detected after chemical conversion in NGS-based approaches are the 2′-O-methylations (Marchand et al., 2016) and m^7G (Alexandrov, Martzen, & Phizicky, 2002), as well as the two prominent examples pseudouridine (Ψ) and m^1A, which shall be introduced in detail in the following part.

Ψ does not impair base pairing and is read as its isomer uridine in RT, precluding assignment in regular NGS. The carbodiimide N-cyclohexyl-N′-β-(4-methylmorpholinium)methylcarbodiimide (CMC) was shown to react with N3 of many pyrimidines and N1 of many purines respectively, but only the CMC-Ψ adduct proved to be stable in alkaline conditions (Gilham & Ho, 1971). The resulting moiety was bulky enough to arrest RT downstream of the modified position. This unique behavior was used in primer extension assays to map Ψ positions in rRNA of E. coli (Bakin & Ofengand, 1993) as well as in a number of organisms like Drosophila melanogaster, Mus musculus, Homo sapiens, Bacillus subtilis, and Halobacter halobium (Bakin & Ofengand, 1998; Ofengand & Bakin, 1997). The strategy of combined CMC and alkaline treatment was adapted for NGS and allowed to identify and map Ψ with single nucleotide resolution in mRNA for the first time (Carlile et al., 2014; Lovejoy, Riordan, & Brown, 2014; Schwartz et al., 2014). To enhance sensitivity of this approach and enable detection of the Ψ status in low abundance RNAs, introduction of a chemical handle utilizing azide-labeled CMC (N3-CMC) was proposed. After formation of the N3-CMC-Ψ, SPAAC allowed coupling with biotin and enrichment of labeled RNA via streptavidin beads (Li et al., 2015). Taken together, the development of these techniques allowed to precisely identify and map Ψ in the RNA of many different species in a transcriptome-wide scale and revealed its existence even in low abundant mRNA (Carlile et al., 2014; Li, Zhu, et al., 2015; Lovejoy et al., 2014; Schwartz et al., 2014).

While in the previous case, an originally RT silent modification is chemically treated to display an identifiable RT pattern, the opposite strategy is employed in the next example. Methylation in 1-position of adenosine (m^1A) causes a mutational pattern during RT (Kirpekar, Hansen, Rasmussen, Poehlsgaard, & Vester, 2005). Nevertheless, unambiguous identification of m^1A sites in mRNA via NGS has proven difficult, as the combination of truncation and mutation is sequence-context-
dependent and m1A abundance is relatively low (Hauenschild et al., 2015). This uncertainty can be overcome by chemically
triggering rearrangement of the endocyclic and exocyclic nitrogen of m1A under alkaline conditions (Dimroth
rearrangement [Macon & Wolfenden, 1968]), to yield the RT silent m6A instead (Dominissini et al., 2016). Vanishing of the RT signal in a
sample after sequencing then allows to identify the m1A positions. A similar approach, utilizing the depletition of the m1A sig-
nal via enzymatic demethylation has been published in the same year (Li et al., 2016).

4 | METABOLIC LABELING TO FACILITATE ANALYSES

A third strategy to label RNA is based on feeding cells with metabolic precursors bearing small modifications that do not
interfere with metabolic processing but facilitate analyses. The archetype of this strategy is isotope labeling based on nucleo-
sides or amino acids to support MS-based analysis (Bruckl, Globisch, Wagner, Muller, & Carell, 2009; Heiss, Reichle, &
Kellner, 2017) and NMR (Dutta, Saxena, Schwalbe, & Klein-Seetharaman, 2012; Fürtig, Richter, Wöhnert, & Schwalbe,
2003). Here, we will focus on metabolic labeling of RNA allowing for subsequent chemical derivatization to facilitate ana-
lyses. As such, we will not cover isotope labeling or 5-bromo-uridine (BrU) labeling but instead focus on modifications for
which selective chemistries have been developed, in particular click chemistry (Figure 1) and thiol-reactive chemistry.

Three basic strategies can be distinguished for metabolic RNA labeling based on the metabolic precursor. Strategy I uses
nucleosides (non-natural nucleoside analogs or non-canonical nucleosides) that are cell permeable. Inside the cell, these nucleo-
sides need to be able to enter the so-called “salvage pathway” where nucleotides are not synthesized de novo but by recycling
of the basic components derived from the catabolism during normal RNA turnover. The salvage pathway is energetically less
expensive and involves 5’-phosphorylation of the nucleoside to NMP, followed by formation of the respective di- and triphos-
phates (NDP and NTP). Importantly, the respective NDPs should not be substrates of ribonucleotide reductase catalyzed deox-
xygenation to dNDPs, which would then become dNTPs to be used in DNA synthesis. This strategy can be harnessed to label
and selectively enrich all newly transcribed (i.e., nascent) RNA in vivo by (random) incorporation of NTPs from the nucleo-
side feeding experiment. A distinction between different types of RNA typically cannot be made during incorporation but only
after separation. The most widely used nucleosides for metabolic RNA labeling and subsequent functionalization are ethynyl-
uridine (EU), 4-thiouridine (4SU), 5-aza-cytidine (5-aza-C), N6-propargyl-adenosine (N6pA), and 2-ethynyl-adenosine (EA),
all shown in Figure 6 and all commercially available (Curanovic et al., 2013; Grammel, Hang, & Conrad, 2012; Jao & Salic,
2008; Khoddami & Cairns, 2013; Melvin, Milne, Slater, Allen, & Keir, 1978).

**FIGURE 6** Nucleoside and methionine analogs used for co- and posttranscriptional metabolic labeling. EU, 5-ethynyl-uridine; 4SU, 4-thiouridine;
5-aza-C, 5-aza-cytidine; PSH, propargyl-selenohomocysteine; N6pA, N6-propargyl-adenosine; EA, 2-ethynyl-adenosine; s6G, 6-thioguanosine; TU, 4-thiouracil; DTU,
2,4-dithiouracil; 5EC, 5-ethynylcytosine; 6-TG, 6-thioguanine
Strategy II starts from the nucleobase and thus requires a different salvage pathway, namely the enzyme phosphoribosyltransferase (PRPT) to generate the nucleoside monophosphate directly (and subsequently the triphosphate). The result is co-transcriptional labeling in cells expressing the required PRPT. This strategy has been realized for cell-specific labeling using thiouracil (TU), dithiouracil (DTU), and 5-ethylcytosine (5EC) shown in Figure 6 (Cleary, Meiering, Jan, Guymon, & Boothroyd, 2005; Hida et al., 2017).

Strategy III starts from analogs of the amino acid methionine which are converted to corresponding AdoMet analogs by cellular MAT, as outlined for labeling and enrichment strategies in section 3.1 (Hartstock et al., 2018). The chemical handle (e.g., the terminal alkyne from propargyl-selenohomocysteine [PSH], Figure 6) is then transferred onto RNAs by cellular RNA MTases and can be reacted in a click reaction allowing for enrichment of RNAs with MTase target sites.

The main benefit of co-transcriptional metabolic labeling (strategies I and II) is the ability to monitor nascent RNA during transcription. In traditional transcriptional analyses relying on total cellular RNA, changes in RNA levels cannot be attributed to RNA synthesis or decay, and temporal resolution is poor (Dolken et al., 2008). Efficient metabolic labeling in combination with mild, selective and efficient chemistry provides a means to separate total cellular RNA into newly transcribed and pre-existing RNA (Dolken et al., 2008). Metabolic RNA labeling is therefore useful to understand RNA dynamics without perturbing transcription or RNA processing. This is important because rapid global changes in regulated transcription on the timescale of minutes to hours occur in numerous cellular mechanisms, including circadian rhythm or immune response. The benefit of post-transcriptional metabolic labeling (strategy III) is the ability to enrich and assign MTase target sites. This relatively new strategy also bears potential to look at the dynamics of RNA modifications. In general, metabolic labeling of RNA together with high-throughput technologies and quantitative modeling can help us reveal the biological roles and mechanisms of gene regulation at the systems level (Yamada & Akimitsu, 2019).

### 4.1 Feeding nucleosides as precursors for co-transcriptional RNA modifications followed by post-synthetic chemical modifications

#### 4.1.1 Clickable nucleosides

5-Ethynyl-uridine (EU) has become the primary clickable nucleoside analog for metabolic labeling, because it is straightforward to synthesize and to date commercially available, including kits for derivatization with biotin azide and streptavidin-coated magnetic beads. EU is incorporated into transcripts generated by RNA polymerases I, II, and III in cells, but not into DNA (Jao & Salic, 2008). HPLC-analysis showed that 2.8% of uridines were substituted with EU in total RNA and varying amounts from 1–6% in different RNA species (rRNA, short RNAs, mRNA) after a 24 hr pulse with 1 mM EU in HEK293T cells. In a second step, the newly synthesized cellular RNA can be detected with high sensitivity by a CuAAC reaction with fluorescent azides (Figures 1 and 7). As the click reaction is carried out in fixed and permeabilized cells, Cu(I)-induced degradation is not a concern. EU incorporation was also demonstrated in whole animals by injecting a 3-week old mouse intraperitoneal with 2 mg of EU and harvesting organs 5 hr later (Jao & Salic, 2008). The organs were fixed, embedded in paraffin and sectioned, followed by staining with a fluorophore-azide (TMR-azide). Total RNA synthesis could be visualized in different tissues (spleen, liver, intestine, and kidney) and different levels of bulk transcriptional activities were observed for certain cell types within a tissue.

In addition to labeling, EU facilitates enrichment of nascent RNAs via CuAAC. It is therefore useful for investigating RNA regulation, for example, in conjunction with RNA synthesis, stability, decay, and transcriptional regulation (Abe et al., 2012). As such, it has become an alternative to established radioactive or BrU-based assays that, together with NGS, have been used to map and quantify transcriptionally engaged polymerase density genome-wide (Core, Waterfall, & Lis, 2008) or other dynamic aspects of RNA metabolism (Paulsen et al., 2014; Roberts et al., 2015) as recently reviewed in (Yamada & Akimitsu, 2019).

In addition to EU, a number of alternative clickable nucleosides have been successfully used for metabolic labeling. The adenosine analog N6-propargyladenosine (N6pA, Figure 6) is also taken up by mammalian cells and efficiently incorporated into RNA by poly(A)-polymerase and all three mammalian RNA polymerases (Pol I, Pol II, Pol III), but not DNA (Grammel et al., 2012). Similar to EU, N6pA can be used for fluorescent labeling and for affinity enrichment after CuAAC with appropriate azido-conjugates. Visualization was performed by clicking an azido-rhodamine dye in cells fed with 10–1,000 μM N6pA for 12 hr (Grammel et al., 2012).

One advantage of modified adenosines compared to more commonly used EU is that they can be used to visualize poly(A) tail dynamics of RNA. Indeed, N6-propargyl-ATP and also other modified ATPs are substrates for various poly(A)-polymerases and have become commercially available (Anhäuser et al., 2019; Grammel et al., 2012; Winz et al., 2012).
Transcriptome-wide poly(A) tail profiling was convincingly shown using 2-ethynyladenosine (EA, Figure 6) for metabolic labeling (Curanovic et al., 2013). To achieve selective poly(A) tail labeling (and not co-transcriptional labeling), transcription was blocked. Upon addition of actinomycin D to cells, EA was still incorporated into the poly(A) tail, facilitating poly(A) tail analysis (Figure 7b). EA was also used in frog oocytes, whose maturation can be induced by applying progesterone. This induces polyadenylation of stored, deadenylated mRNAs, while at the same time transcription was inhibited by treatment with actinomycin D. This experiment in combination with subsequent isolation of EA-containing transcripts via CuAAC and NGS identified mRNAs that become polyadenylated on a transcriptome-wide level in response to progesterone (Curanovic et al., 2013). These transcripts targeted for polyadenylation participate in diverse cellular processes. The biological relevance was demonstrated for a lysine MTase that was knocked down and shown to be required for efficient oocyte maturation. Furthermore, known regulatory elements in the 3′-UTR that are associated with polyadenylation were confirmed (the so-called cytoplasmic polyadenylation elements CPE) and new elements were identified (U-rich elements, UREs). These findings showed that the 3′-UTRs of transcripts that undergo progesterone-induced polyadenylation are characterized by a pronounced enrichment of CPEs as well as UREs within 100 nucleotides proximal to the 3′-end of the transcripts. The function of UREs was validated by reporter constructs (Curanovic et al., 2013).

Similar to N6pA, N6-allyladenosine is also taken up by mammalian cells and incorporated into total RNA, as validated by LC–MS (Shu et al., 2017). It was shown on synthetic RNA in vitro that N6-allyladenosine can be cyclized after addition of iodine and that the resulting modified adenosine increases the miscircorporation opposite the modified A to ~70% (compared to 5% for opposite N6-allyl-adenosine) using HIV reverse transcriptase (Figure 5b) (Shu et al., 2017). This is interesting because the N6-position can be modified with methyl groups but also with other alkyl groups if suitable AdoMet analogs are used (Hartstock et al., 2018; Hartstock & Rentmeister, 2018; Shu et al., 2017). It will be interesting to see whether the metabolic labeling with N6-allyladenosine can be combined with the chemical reaction and how this can be used to learn about the MTase target sites. An alternative way would be to start metabolic labeling from the AdoMet precursors, that is, methionine analogs as outlined below (Hartstock & Rentmeister, 2018; Muttach, Muthmann, & Rentmeister, 2017; Muttach & Rentmeister, 2016a, 2016b).

### 4.1.2 | 4-Thiouridine (4SU)

In addition to non-natural clickable nucleosides, the non-canonical thiol-containing uracil derivative 4-thiouridine (4SU, Figure 6) has long been used for metabolic labeling in combination with subsequent chemistry for analyses and/or enrichment.
Similar to EU, 4SU can enter cells and become metabolized to the corresponding triphosphate, which is incorporated into nascent RNA (Melvin et al., 1978). The extent of 4SU incorporation and its toxicity are highly cell-type specific. The 4-thiogroup gives access to a range of selective chemistries that have been exploited long before the term click chemistry was coined (Figure 8). Originally, enrichment of 4SU-containing RNA was based on organomercurial affinity matrices (Kenzelmann et al., 2007; Melvin et al., 1978; Woodford, Schlegel, & Pardee, 1988). These toxic compounds were then replaced by reversible disulfide chemistry—often in combination with biotin-conjugated reagents to enrich 4SU-containing RNA via streptavidin beads. This was soon replaced by commercially available biotin-HPDP (the acronym for N-[6-biotinamido]hexyl)-3′-(2′-pyridyldithio)propionamidine) to improve isolation and enrichment of 4SU-labeled RNA (Cleary et al., 2005; Dolken et al., 2008). This thiol-reactive biotinylation reagent forms a reversible disulfide linkage (Figure 8a) and has been used to label cysteine residues in proteins as well as other substrates containing sulphydryl groups. The HPDP-biotin approach enabled detection of newly transcribed RNAs after only 10 min and in combination with microarrays, qPCR, and in particular NGS stimulated research on various aspects of RNA expression dynamics (Dolken et al., 2008; Friedel & Dolken, 2009). The downstream sample preparation after metabolic labeling with 4SU and biotin-HPDP-based enrichment was further improved, resulting in a method called transient transcriptome sequencing (TT-seq) (Michel et al., 2017; Schwabl et al., 2016). In TT-seq, the nascent transcripts are fragmented prior to affinity enrichment, allowing to measure only the newly transcribed 4SU-containing RNA. This improvement of the downstream sample preparation circumvents a major problem of short metabolic labeling, namely that reference transcripts are not uniformly mapped because only a short 3′ region of nascent transcript is labeled during a transient exposure to 4SU and the long pre-existing 5′ regions dominate the sequencing data.

In recent years, not only the downstream processing but also the chemistry for 4SU derivatization was further explored and improved for the purpose of RNA analysis. Duffy et al. showed that the biotin-HPDP chemistry was inefficient (<20%), produced biased results and could be improved by using an activated disulfide instead (>95%) (Figure 8b,c) (Duffy et al., 2015). Specifically, they used methane thiosulfonate (MTS) to selectively and reversibly react with the thiol-group of 4SU. Here, a disulfide bond is formed with the MTS-reagent, which can be biotin-conjugated or directly tethered to a solid support and can be cleaved under reducing conditions (Duffy et al., 2015; Duffy, Canzio, Maniatis, & Simon, 2018; Duffy & Simon, 2016). The approach was used to track different populations of RNA as demonstrated by studying global miRNA turnover in proliferating cultured human cells.

SLAM seq stands for thiol (SH)-linked alkylation for the metabolic labeling of RNA (Herzog et al., 2017) and is also based on metabolic labeling with 4SU (Figure 8d). In SLAM seq, the primary thiol-reactive compound iodoacetamide (IAA) is used to alkylate the thiol by nucleophilic substitution in high yields (98% under optimized conditions). Importantly, the resulting carboxamidomethyl group at the 4 position leads to a T > C conversion with high yields (>94%). This is 8.5-fold improved compared to 4SU alone, which gave 10% of T > C conversion when untreated. SLAM seq was used to quantify 4SU-labeled transcripts in mouse embryonic stem cells (mESCs) by feeding 100 μM 4SU for 24 hr in combination with 3′-end mRNA sequencing and to measure the polyadenylated transcriptional output in mESCs. SLAM seq enabled the direct quantification of newly synthesized mRNAs. Therefore, it allowed to dissect the regulatory effects of inhibitors on different transcriptional targets in a transcriptome-wide manner and thus understand the effects obtained in chemical genetics approaches (Muhar et al., 2018). It should be noted that IAA is a general alkylation agent and might react with other nucleophilic sites (e.g., N7 of guanosine) to a low degree.

TimeLapse-seq uses alternative chemistry for 4SU-derivatization after metabolic labeling (Schofield, Duffy, Kiefer, Sullivan, & Simon, 2018). Similar to SLAM seq and TUC-seq, the metabolically 4SU-labeled RNA is not enriched but chemically converted to recode the hydrogen bonding pattern of 4SU to cytidine, causing mutation in sequencing. Specifically, in TimeLapse-seq 2,2,2-trifluorethylamine (TFEA) was used in combination with either meta-chloroperoxybenzoic acid (mCPBA) or sodium periodate (NaIO4) to convert 4SU mostly to the trifluoroethylated cytidine with little degradation of the oligoribonucleotide (Figure 8f). The RT showed ~80% of T-to-C mutation. Recently, TimeLapse-seq was extended to 6-thioguanine (6-TG) and 6-thioguanosine (s6G) (Figure 6) (Kiefer, Schofield, & Simon, 2018). Both were taken up by cells and metabolized, leading to incorporation of the respective 6-thio-G into nascent RNA, albeit at lower rates than 4SU (1.5 vs. 4.5%). After application of the TimeLapse-seq chemistry, conversion 6-thio-G to 6-amino-substituted adenosine increased the level of G-to-A substitutions.

TUC-seq stands for thioruridine-to-cytidine-sequencing and uses a different chemical reaction to obtain a T to C conversion for 4SU (Lusser et al., 2019; Riml et al., 2017). Specifically, 4SU is directly converted to cytosin (C) by osmium tetroxide (OsO4) and ammonia before sequencing (Burton, 1967; Riml et al., 2017). The conversion is also nearly quantitative (98%) at conditions that do not degrade the RNA and exceptionally clean, meaning that unwanted modification of canonical nucleobases (e.g., alkylation, hydrolysis) is not observed. The big advantage of this approach is that the chemical reaction...
FIGURE 8  Thiol-selective chemistry for derivatization of 4SU. (a) HPDP-bio leads to reversible formation of a disulfide bond between 4SU and biotin (bio-4SU). (b) MTS-bio is methane thiosulfonate-biotin, an activated disulfide for more efficient biotinylation of 4SU. (c) The MTS-reagent can also be tethered directly to a solid support (gray ball). (d) Iodacetamide modifies 4SU and is the basis of SLAM seq, where modified 4SUs in RNA lead to mutations in the sequencing readout. (e) 4SU is directly converted to cytosine (C) by osmium tetroxide (OsO₄) and ammonia before sequencing in TUC-seq. (f) 4SU is mostly converted to the trifluoroethylated cytidine using 2,2,2-trifluorethylamine (TFEA) in combination with meta-chloroperoxybenzoic acid (mCPBA) or sodium periodate (NaIO₄). The trifluoroethylated cytidine is the basis of TimeLapse-seq and also results in mutations in the sequencing readout at sites originally containing 4SU. (a–c) Are used for enrichment of 4SU-containing RNA. (d–f) Are used for RT, library preparation and NGS where 4SU-containing sites are detected as mutations.
leads to formation of a native C (and not an analog that is recognized as C). This improves the sequencing results independent of RT conditions (e.g., concentration of RNA, dNTPs and type of polymerase) and also ensures maximum efficiency of RT. To date, TUC-seq was used to detect and identify naturally occurring 4SU sites in bacterial RNA and metabolically labeled inducible reporter genes in HEK293 cells. All three methods can be used to capture temporal information about RNA directly in a sequencing experiment without biochemical enrichment. TUC-seq generates native cytidine and should therefore give unambiguous sequencing results, however it was not yet used on a transcriptome-wide scale. All three methods explained above share the advantage that they eliminate the need for tag-based separation of labeled and unlabeled RNA. A side-by-side analysis of all three approaches would be desirable to determine whether one is superior.

In addition to downstream processing and chemistry, bioinformatics represents a complementary route to improve the results for transcriptome-wide datasets. In metabolic labeling, the relatively infrequent incorporation events and significant sequencing error rates make differentiation between old and new RNA highly challenging (Jurges et al., 2018). A statistical approach termed GRAND-SLAM was developed to estimate the proportion of old and new RNA in a SLAM-seq experiment (Jurges et al., 2018). However, further details on bioinformatics go beyond the scope of this review.

In contrast to feeding with clickable nucleoside analogs, metabolic labeling with 4SU is not compatible with cellular imaging due to competing free thiols in cells (Grammel et al., 2012). Although 4SU has been widely used for metabolic labeling, it should be noted that 4SU inhibits rRNA synthesis and causes a nucleolar stress response. Feeding 4SU at elevated concentrations (>50 μM) therefore inhibits production and processing of 47S-rRNA (Burger et al., 2013). A common disadvantage of the metabolic labeling approaches mentioned above is a lack of specificity, because the produced NTP analogs are used by all three RNA polymerases, as well as poly(A)-polymerase in the case of ATP analogs. In attempts to overcome this limitation, enzyme-specific approaches have been developed.

4.2 Enzyme-specific approaches

In strategy II for metabolic labeling of RNA, cells are fed with a nucleobase analog or a non-canonical nucleobase as shown for 4-thiouracil (TU) and 5-ethynylcytosine (5EC) (Figure 6). In “TU-tagging” uracil phosphoribosyltransferase (UPRT) is required to generate the respective nucleoside monophosphates directly (Miller, Robinson, Cleary, & Doe, 2009). Mammals lack this enzyme activity whereas protozoan parasites can incorporate uracil into their nucleic acids (Cleary et al., 2005). Cleary et al. took advantage of this orthogonality and showed that by tissue-specific expression of UPRT from Toxoplasma gondii, 5EC can metabolically label RNA in the respective cells (Figure 9a) (Miller et al., 2009). This elegant strategy was picked up by many researchers and implemented in a number of cell types and model organisms, including Drosophila and mouse (Figure 9a) (Burow et al., 2015; Erickson & Nicolson, 2015; Gay et al., 2013; Miller et al., 2009; Rinn et al., 2008; Tallafuss et al., 2015; Tallafuss, Washbourne, & Postlethwait, 2014).

In analogy to TU-tagging, a method based on clickable 5-ethynylcytosine (5EC, Figure 6) was developed. This methodology requires the activity of two enzymes, namely cytosine deaminase (CD) and UPRT (uracil phosphoribosyltransferase) to convert 5EC to 5-ethynyl-uridine monophosphate that is subsequently incorporated into nascent RNAs (Figure 9b). The ethynyl group allows efficient detection and purification of tagged RNAs and thus cell type-specific gene expression data can be obtained from intact Drosophila larvae, including transcriptome measurements from a small population of central brain neurons (Hida et al., 2017). EU- and EC-tagging for tissue specific RNA labeling are great additions to the toolbox of biologists, who commonly use eGFP expression under certain promoters/enhancers to label specific cell types and isolate them via fluorescence-activated cell sorting (FACS). The tagging methods are particularly useful to obtain information about RNA from cell types that are difficult to isolate by dissection or dissociation methods, such as subsets of neurons or glia in the central nervous system.

An elegant way to capture specific enzymatic activities in a metabolic labeling approach is to harness a particular mechanism. This strategy works for a subset of MTases—the m5C-RNA MTases—that follow a mechanism involving formation of a covalent enzyme-substrate intermediate (Figure 10) (Gabella, Sheluho, & Bhagwat, 1995; Motorin, Lyko, & Helm, 2010; Santi, Garrett, & Barr, 1983; Wu & Santi, 1987). The covalent bond results from the attack of a cysteine at the C6 atom of the target cytidine in RNA (Figure 10a). The MTase then transfers a methyl group from SAM/AdoMet to the C5 of the target cytidine, releasing the enzyme. If cytidine is replaced with 5-azacytidine (5-aza-C), the covalent enzyme-substrate intermediate can still form, however, release of the free enzyme is impeded (Figure 10b). 5-Aza-C can be used for metabolic labeling and leads to a chemical crosslink between the RNA and an MTase following that mechanism. Thus, the target RNA spectrum of an m5C-RNA MTase of interest can be investigated following immunoprecipitation, RNA purification and NGS (Figure 10b). This “Aza-IP” (5-azacytidine-mediated RNA immunoprecipitation) was applied in human cell lines to different MTases
(e.g., NSUN2, NSUN3, NSUN6, DNMT2) and revealed their known targets but also RNA targets previously not associated with these m^5C-RNA MTases (Haag et al., 2016, 2015; Khoddami & Cairns, 2013).

Alternatively, a covalent m^5C-RNA MTase-RNA product can be obtained by targeted mutations—without the need for chemical modifications. Herein, one of the MTase's two cysteins responsible for covalent link formation and release is substituted for alanine resulting in a variant that forms the covalent bond but does not release the RNA accordingly (Hussain et al., 2013). This alternative strategy also allows to identify the targets of an m^5C-MTase as demonstrated for NSUN2, in a mechanism-based way. It should be mentioned that crosslinking methods can also be performed independent of the mechanism and—for NSUN2—have been shown to yield results that are in line with Aza-IP (Hussain et al., 2013).

A conceptually different strategy of metabolic labeling in combination with chemo-enzymatic crosslinking has been used to facilitate RNA localization studies in cells. Utilizing the engineered ascorbate peroxidase APEX2, which allows to convert a cell permeable biotin-phenol substrate into a reactive radical, protein sidechains can be covalently labeled with biotin. By directing APEX2 to specific subcellular locations nearby proteins were biotinylated and subsequently crosslinked via formaldehyde treatment to RNA in close proximity. After streptavidine-based enrichment, these RNAs were identified by NGS. This method works well for membrane enclosed regions, whereas more exposed regions are more challenging to target.

**FIGURE 9** Enzyme and tissue-specific metabolic labeling via TU and 5EC (Miller et al., 2009)
(Kaewsapsak, Shechner, Mallard, Rinn, & Ting, 2017). Recently, an approach for direct biotinylation of RNA in specific localizations, termed APEX-seq, was proposed by two individual groups. Here the biotin radical, produced by APEX2, directly reacts with electron rich nucleobases of RNA. Sequencing of the streptavidin enriched RNAs is then used to capture the patterns of RNA localization across the cell (Fazal et al., 2019; Padròn, Iwasaki, & Ingolia, 2018).

4.3 | Metabolic labeling in combination with photo-crosslinking

In general, studying RNA-protein interactions in the cellular context is important to get a comprehensive picture of RNA processing and effects. Unfortunately, the transient nature of most RNA-protein interactions limits the potential of simple immunoprecipitation experiments. Therefore, chemical crosslinking for genome-wide mapping of cellular protein-RNA interactions has become a widely used approach and deserves a review on its own (Li, Song, & Yi, 2014). Herein, covalent bonds between RNA and the interacting protein stabilize the complex and facilitate co-purification for downstream analysis by sequencing (RNA) or mass spectrometry (protein). In this section, we want to briefly highlight how metabolic labeling with non-canonical or non-natural nucleosides has been used to improve photo-crosslinking strategies. These methods build on previous work of the Darnell group, who used UV crosslinking and immunoprecipitation to purify protein-RNA complexes from mouse brain (Ule et al., 2003). In general, longer wavelength (i.e., 365 nm instead of 254 nm) are considered milder and thus advantageous for photo-crosslinking of cellular components. However, in certain cases 312 nm yielded the best compromise between sensitivity and specificity (He et al., 2016).

Metabolic labeling with 4SU and sG (Figure 6) has been used to facilitate photo-crosslinking with RNA-binding proteins (RBPs) at 365 nm in cells (PAR-CLIP, photoactivatable ribonucleoside enhanced crosslinking, Figure 11b) (Hafner et al., 2010). After lysis, immunoprecipitation of the proteins and partial RNase digest, the crosslinked RNA-proteins are further purified. Finally, the protein is degraded and NGS libraries are prepared to identify RNA regions that bound to the RNA-binding protein under investigation. The advantages conferred by 4SU and sG are the higher and thus less damaging wavelength for photo-crosslinking (365 nm instead of 254 nm) and an increase in the crosslinking efficiency. Furthermore, 4SU can lead to T to C transition in the sequenced cDNA (Figure 8), enabling identification of crosslinking sites at single-nucleotide resolution by analyzing the mutations in cDNA sequences. PAR-CLIP is particularly suitable for studies using cultured cells, because of their high uptake efficiency for 4SU (incorporation efficiency is up to 4% for 4SU relative to uridine at 100 μM for 16 hr) (Hafner et al., 2010). In addition to studying the sequences that are interacting with an RNA-binding
protein, 4SU-aided crosslinking can also be used for the identification of RNA-binding proteins on transcriptome-wide scale (for excellent review see: [Li et al., 2014]). Using 4SU-mediated photo-crosslinking and quantitative MS, He et al. mapped RNA-binding regions in hundreds of known and unknown RNA-binding proteins in nuclei of ESCs (He et al., 2016). In combination with poly(A)-enrichment, 4SU-aided crosslinking provided data for an entire atlas of mRNA-binding proteins (Figure 11b) in mammalian cells (Baltz et al., 2012; Castello et al., 2012) and yeast (Beckmann et al., 2015). A 4SU feeding approach was also used to improve the resolution of m6A-sequencing (originally developed by [Dominissini et al., 2012; Meyer et al., 2012]) by photo-crosslinking to the anti-m6A-antibody binding in proximity to incorporated 4SU (Chen et al., 2015). However, for m6A-detection, UV crosslinking (at 254 nm) without 4SU feeding has become the dominating technique in the field (Linder et al., 2015), which may be due to the low abundance of m6A (one to three sites per transcript but modified substoichiometrically) compromising the crosslinking yield in a metabolic labeling approach. In summary, 4SU-aided crosslinking has been used in many variations and can be considered a standard method for identification of RNA-binding proteins or sequences binding to an RBP.

Recently, also EU feeding has been combined with photo-crosslinking at 254 nm to identify interacting proteins of newly made RNAs based on MS (Figure 11a) (Bao et al., 2018) and termed RICK (capture of the newly transcribed RNA interactome using click chemistry). RICK captures all newly transcribed RNAs if metabolic labeling is performed for short periods of time (0.5–2 hr), as well as proteins that are photo-crosslinked to them, and therefore yielded new insights into non-polyadenylated RNAs that were hard to access before (polyadenylated RNAs were isolated simply via oligo(dT) beads). Bao and Guo applied RICK to HeLa cells and mouse embryonic stem cells (mESCs). They identified mitotic regulators with preferential affinity for nonpolyadenylated RNAs and showed a connection between metabolic enzymes and nascent RNAs. Furthermore, the known RNA-bound proteome of mouse embryonic stem cells was expanded.

4.4 | Feeding amino acids as precursors

Strategies I and II for metabolic labeling rely on formation of NTP analogs from cell permeable precursors and hence label nascent RNAs during transcription or poly(A)-tailing. However, although RNAs consist only for the four basic nucleotides it
is well known that a plethora of modifications exist (currently 170 are listed in the MODOMICs database) (Boccaletto et al., 2018), out of which methylation at various positions is the most abundant modification. As outlined in section 2.2.3, SAM or AdoMet is the most abundant methyl donor in nature and generated from the reaction of methionine with ATP. This reaction is catalyzed by methionine adenosyltransferase (MAT), an enzyme that was also shown to tolerate methionine analogs, in particular if a variant with more space in the active site is generated (Muttach & Rentmeister, 2016a, 2016b; Wang et al., 2013). Recently, we showed that the methionine analog propargyl-selenohomocysteine (PSH, Figure 6) is taken up by mammalian cells and efficiently converted to the AdoMet analog SeAdoYn (Figure 12) using the cellular wildtype MAT. Propargyl modifications were detected in total RNA isolated from these cells (Hartstock et al., 2018). Furthermore, the CuAAC reaction with biotin azide was used to isolate and enrich RNAs harboring MTase target sites from total RNA of PSH-fed cells. To date, NGS confirmed enrichment of MTase target sites in rRNA (Figure 11b). The method has potential to be transferred to other RNA species, which would be particularly interesting for mRNA and also to investigate dynamic aspects of methylation under different conditions.

Similar to the metabolic labeling strategies I and II, the drawback of this strategy is also a lack in specificity for a particular MTase. As most MTases use AdoMet as cosubstrate and are also able to efficiently convert SeAdoYn, many methylation sites will be targeted. Therefore, the strength of metabolic labeling lies in the investigation under cellular conditions and dynamic aspects.

5 CONCLUSIONS AND OUTLOOK

5.1 Conclusions and outlook toward RNA imaging in cells

Live cell imaging of RNA promises to shed light on the dynamics of RNA localization and will help understand its implications in function and regulation. The required toolbox for live cell labeling has grown and matured considerably in the last decades. Hybridization-based probes—originally limited to visualization in fixed cells—were soon improved in terms of fluorogenicity and are now amenable for live cell imaging (Autour et al., 2018; Dolgosheina et al., 2014; Filonov et al., 2014; Paige et al., 2011). The seminal development of non-natural base pair systems has opened the door to site-specific postsynthetic modification of enzymatically produced RNA without extra tags, especially now that their function in living cells has been demonstrated (Zhang et al., 2017). For dynamic visualization of the whole life cycle of RNA, chemo-enzymatic labeling methods are attractive as the label is covalently attached and the resulting linkage stable, and because the ROI is only

**FIGURE 12** Metabolic labeling with the amino acid propargyl-selenohomocysteine (PSH) as representative for methionine analogs. (a) The wildtype enzyme methionine adenosyltransferase (MAT) uses PSH and ATP to generate the AdoMet analog SeAdoYn that is used by many MTases as cosubstrate. Here transfer of the propargyl group to RNA is depicted schematically. (b) Mammalian cells take up PSH leading to intracellular generation of SeAdoYn and transfer of propargyl groups to RNA. After isolation of total RNA, the MTase targets can be enriched via CuAAC reaction with biotin azide. The enriched sequences can be found in NGS (Hartstock et al., 2018)
minimally altered. The development of novel cosubstrate analogs and the engineering of corresponding enzymes has already allowed for site-specific modification of RNA, including RNA labeling via click chemistry in living cells (Holstein et al., 2016). These two-step approaches provide great flexibility regarding the choice of fluorophores and click reactions. Implementation and further development of fluorogenic reactions with dyes emitting in the far-red range or suitable for two-photon excitation are of special interest for live cell imaging. Performing the enzymatic modification in the cell would represent a major advancement in the field of MTase-based RNA labeling strategies. To date, all RNAs labeled in living cells required delivery by lipofection or genetic fusion of tags. MTases do not require tags, however, delivery of the AdoMet analog and specificity of the MTase represent challenges. Delivery of the non-permeable AdoMet analog can be circumvented by feeding methionine analogs as metabolic precursor and intracellular generation of the AdoMet analog by wildtype or engineered MATs. This strategy has been previously demonstrated in vivo, as propargyl-selenohomocysteine (PSH) was fed to HeLa cells and converted with ATP into SeAdoYn by the endogenous MAT (Hartstock et al., 2018). To date, this approach allowed for identification of known MTase target sites, as discussed in a section 4. However, this broad acceptance by cellular MTases currently prevents implementation of this approach for site-specific labeling of a ROI.

5.2 Conclusions and outlook for chemo-enzymatic modification of total RNA

The specificity of some RNA-modifying enzymes is suitable for post-synthetic modification of their RNA targets in total RNA. In combination with non-natural substrates and click chemistry, this approach is attractive to isolate, enrich and identify subclasses of RNAs or targets of these RNA-modifying enzymes. Impressive examples are NAD-captureSeq and different Hen1-MTases for enrichment for RNA-species (Cahova et al., 2015; Mickuté et al., 2018; Osipenko et al., 2017).

Furthermore, natural RNA modifications can be assigned with the help of chemo-enzymatic approaches. However, mapping RNA modifications in a transcriptome-wide manner can be prone to many complications. Secondary structures and other native modifications, as well as low abundance and limited stability of the RNA can complicate the resulting RT patterns. To identify a specific modified position with high confidence, results must be compared between different techniques to exclude biases and false-positive results. As these emerging chemo-enzymatic strategies offer alternative methods with likely different bias compared to approaches based on immunoprecipitation of RNA modifications, comparison between them should ultimately lead to identification with greater confidence. Additionally, mapping strategies for previously inaccessible modifications, like Ψ, can now be performed on a transcriptome-wide scale even in low abundance RNA species, thanks to the attachment of handles used for specific enrichment. Until now, many of these strategies lack a reliable way to quantify the modification status at a specific RNA position, due to varying efficiencies of either the enzymatic or chemical step. Improvement of both steps, allowing for the precise detection of the modified position as well as quantification of its abundance in a transcriptome-wide manner would offer insights into RNA modification levels in biological systems and thus greatly improve the understanding of the biological role of RNA modifications.

5.3 Conclusions and outlook for metabolic labeling

Metabolic labeling starting from non-canonical or non-natural nucleosides that enable selective chemical modification compatible with RNA represents an elegant strategy to capture dynamic aspects of RNA metabolism in living systems. An obvious drawback of the approach is a lack of specificity, meaning that all transcripts are modified depending on the speed of transcription of the respective polymerase and its tolerance for the substrate analog. In an attempt to solve this issue, a photocaged version of EU was developed and used for metabolic labeling (Feng, Li, & Spitale, 2017). Combining metabolic labeling and spatial control with the help of light presents a promising strategy for metabolic labeling in a cell specific manner, although uncaging will have to be improved to become useful for cellular studies (currently 5 min UV light irradiation).

Another limitation that is inherent to most chemo-enzymatic approaches but particularly pronounced in metabolic labeling approaches is the chemical repertoire that is compatible with various enzymes required for generating the precursor analogs (here: enzymes of the salvage pathway). The cellular machinery, especially native RNA polymerases, will only accept certain comparatively small modifications (e.g., short alkynes, but not cyclooctynes). As a consequence, nucleosides modified with linear alkynes (which are substrates for copper-catalyzed but not for copper-free azide–alkyne cycloaddition) have been mainly used to date, and the chemical labeling step could not be performed in living cells. Thus, owing to the need for cell fixation, only static snapshots of transcription, RNA turnover, and poly(A) tail formation, rather than a complete picture of the dynamics of these processes, were obtained (Schulz & Rentmeister, 2014).
Undoubtedly, technical advances in high throughput methods have enabled the generation and interpretation of transcriptome-wide data yielding a more comprehensive picture of RNA metabolism at the cellular level. However, none of the chemical reactions used on a transcriptome-wide scale are perfect, meaning that conversions (especially of RNA at the low cellular concentrations) are not quantitative and side-reactions will occur (especially with more abundant cellular components). Bias is also an issue, as non-natural or non-canonical analogs are converted by many enzymes of a class, but likely with different efficiencies. Many of these subtle factors are not considered in the final readout of transcriptome-wide data. It is also interesting to see that the set of nucleoside analogs used to date is quite limited, relying mostly on the well-known 4SU and on EU. The emerging interest in transcriptome-wide data has led to improved thiol-selective chemistries resulting in efficient conversion of 4SU. The terminal alkyne group in EU (but also EA, N6pA, EA, etc) is still converted with the efficient but non-bioorthogonal CuAAC, which—despite advances in Cu(I) ligands—damages RNA and cells. Therefore, bioorthogonal reactions compatible with small reactive handles are still a major bottleneck for more widespread approaches. Cellular incorporation of azides can help, but azides show limited stability in the reducing cellular environment and therefore also face limitations. Strained alkenes might be a solution if they are small enough for incorporation.

Finally, it is interesting to state that metabolic labeling of RNA—although established as a chemical biology concept—might also exist in nature. Of course, the common notion is that RNA modifications are installed post-transcriptionally by RNA-modifying enzymes. Many of these RNA-modifying enzymes have been identified to date, particularly for the highly modified (both in terms of number of nucleotides and stoichiometry) RNA species like tRNAs and rRNAs. Interestingly, the fraction and stoichiometry of modifications is much lower in the case of mRNA—specifically for internal mRNA modifications (apart from the 5’ cap). These modifications are present at <1 modification per transcript (m^A) and in substoichiometric manner (20–80%) (Liu et al., 2013). For some of them the respective RNA-modifying enzymes have been identified (METTL3-METTL14; METTL16) while for others the responsible enzymes remain elusive. Alternative options would be off-target effects of RNA-modifying enzymes acting on tRNAs or rRNAs (e.g., PUS for pseudouridine) or the co-transcriptional incorporation from modified NTPs. Indeed, modified NTPs were recently shown to exist in cultured mammalian cells, like HeLa, Hek293T, and Jurkat-Tcells (Jiang et al., 2018), suggesting that natural RNA modifications may in part be incorporated statistically during transcription.

**CONFLICT OF INTEREST**

The authors have declared no conflicts of interest for this article.

**RELATED WIREs ARTICLE**

The evolution of RNA structural probing methods: From gels to next-generation sequencing

**ORCID**

Nils Muthmann https://orcid.org/0000-0003-0313-4242
Katja Hartstock https://orcid.org/0000-0002-4815-2367
Andrea Rentmeister https://orcid.org/0000-0002-3107-4147

**REFERENCES**

Abe, K., Ishigami, T., Shyu, A. B., Ohno, S., Unemura, S., & Yamashita, A. (2012). Analysis of interferon-beta mRNA stability control after poly(I:C) stimulation using RNA metabolic labeling by ethynyluridine. *Biochemical and Biophysical Research Communications*, 428(1), 44–49. https://doi.org/10.1016/j.bbrc.2012.09.144

Abil, Z., Denard, C. A., & Zhao, H. (2014). Modular assembly of designer PUF proteins for specific post-transcriptional regulation of endogenous RNA. *Journal of Biological Engineering*, 8, 1–11. https://doi.org/10.1186/1754-1611-8-7

Abudayeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., … Zhang, F. (2017). RNA targeting with CRISPR-Cas13. *Nature*, 550, 280–284. https://doi.org/10.1038/nature24049

Abudayeh, O. O., Gootenberg, J. S., Konermann, S., Joung, J., Slaymaker, I. M., Cox, D. B. T., … Zhang, F. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*, 353, aaf5573. https://doi.org/10.1126/science.aaf5573

Alexander, S. C., Busby, K. N., Cole, C. M., Zhou, C. Y., & Devaraj, N. K. (2015). Site-specific covalent labeling of rna by enzymatic transglycosylation. *Journal of the American Chemical Society*, 137, 12756–12759. https://doi.org/10.1021/jacs.5b07286
Alexandroff, A., Martzen, M. R., & Phizicky, E. M. (2002). Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA. RNA, 8(10), 1253–1266.

Anhäuer, L., Hülwel, S., Zobel, T., & 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 Research, 47, 1–12. https://doi.org/10.1093/nar/gkz2084

Anhäuer, L., & Rentmeister, A. (2017). Enzyme-mediated tagging of RNA. Current Opinion in Biotechnology, 48, 69–76. https://doi.org/10.1016/j.copbio.2017.03.013

Autour, A., Jeng S. C. Y., Cawte A. D., Abdolahzadeh, A., Galli, A., Panchapakesan, S. S. S., … Unrau, P. J. (2018). Fluorogenic RNA mango aptamers for imaging small non-coding RNAs in mammalian cells. Nature Communications, 9, 656. https://doi.org/10.1038/s41467-018-02993-8

Bakin, A., & Ofengand, J. (1993). Four newly located pseudouridylate residues in Escherichia coli 23S ribosomal RNA are all at the peptidyltransferase center: Analysis by the application of a new sequencing technique. Biochemistry, 32(37), 9754–9762.

Bakin, A. V., & Ofengand, J. (1998). Mapping of pseudouridine residues in RNA to nucleotide resolution. Methods in Molecular Biology, 77, 297–309. https://doi.org/10.1385/0-89603-397-x:297

Baltz, A. G., Munschauer, M., Schwanhauser, B., Vasile, A., Murakawa, Y., Schueler, M., … Landthaler, M. (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Molecular Cell, 46(5), 674–690. https://doi.org/10.1016/j.molcel.2012.05.021

Bao, X., Guo, X., Yin, M., Tariq, M., Lai, Y., Kanwal, S., … Esteban, M. A. (2018). Capturing the interactome of newly transcribed RNA. Nature Methods, 15(3), 213–220. https://doi.org/10.1038/nmeth.4595

Barkan, A., Rojas, M., Fuji, S., Yap, A., Chong, Y. S., Bond, C. S., & Small, I. (2012). A combinatorial amino acid code for RNA recognition by Pentatropic peptide repeat proteins. PLoS Genetics, 8, e1002910. https://doi.org/10.1371/journal.pgen.1002910

Baum, D. A., & Silverman, S. K. (2007). Deoxyribozyme-catalyzed Labeling of RNA. Angewandte Chemie International Edition, 46(19), 3502–3504. https://doi.org/10.1002/anie.200700357

Beckmann, B. M., Horos, R., Fischer, B., Castello, A., Eichelbaum, K., Alleaume, A. M., … Hentze, M. W. (2015). The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. Nature Communications, 6, 10127. https://doi.org/10.1038/ncomms10127

Benvlacqua, V., King, M., Chaumontet, M., Nothisen, M., Gabillet, S., Buisson, D., … Taran, F. (2014). Copper-chelating Azides for efficient click conjugation reactions in complex media. Angewandte Chemie (International Ed. in English), 53, 5872–5876. https://doi.org/10.1002/anie.201310671

Blackman, M. L., Royzen, M., & Fox, J. M. (2008). Tetratetazincation: Fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. Journal of the American Chemical Society, 130, 13518–13519. https://doi.org/10.1021/ja8053805

Boccaletto, P., Machnicka, M. A., Purta, E., Pietkowski, P., Baginski, B., Wirecki, T. K., … Bujnicki, J. M. (2018). MORDMICS: A database of RNA modification pathways. 2017 update. Nucleic Acids Research, 46(D1), D303–D307. https://doi.org/10.1093/ngr/gkx1030

Bothwell, I. R., & Luo, M. (2014). Large-scale, protection-free synthesis of se-adenosyl-l-selenomethionine analogues and their application as cofactor surrogates of methyltransferases. Organic Letters, 16, 3056–3059. https://doi.org/10.1021/ol501169y

Bratu, D. P., Cha, B.-J., Mhlanga, M. M., Kramer, F. R., & Tyagi, S. (2003). Visualizing the distribution and transport of mRNAs in living cells. Proceedings of the National Academy of Sciences, 100, 13308–13313. https://doi.org/10.1073/pnas.2233244100

Bruckl, T., Globisch, D., Wagner, M., Muller, M., & Carell, T. (2009). Parallel isotope-based quantification of modified tRNA nucleosides. Angewandte Chemie (International Ed. in English), 48(42), 7932–7934. https://doi.org/10.1002/anie.200902740

Burger, K., Muhl, B., Kellner, M., Rohrmoser, M., Gruber-Eber, A., Windhager, L., … Eick, D. (2013). 4-thiouridine inhibits rRNA synthesis and causes a nucleolar stress response. RNA Biology, 10(10), 1623–1630. https://doi.org/10.4161/rna.26214

Burrow, D. A., Umeh-Garcia, M. C., True, M. B., Bakaj, C. D., Ardell, D. H., & Cleary, M. D. (2015). Dynamic regulation of mRNA decay during neural development. Neural Development, 10, 11. http://doi.org/10.1186/s13064-015-0038-6

Burton, K. (1967). Oxidation of pyrimidine nucleosides and nucleotides by osmium tetroxide. Biochemical Journal, 104(2), 686–694. https://doi.org/10.1042/bj1040868

Büttner, L., Javadi-Zarnaghi, F., & Höbartner, C. (2014). Site-specific Labeling of RNA at internal ribose hydroxyl groups: Terbium-assisted deoxyribozymes at work. Journal of the American Chemical Society, 136(22), 8131–8137. https://doi.org/10.1021/ja503864v

Cahova, H., Winz, M. L., Hofer, K., Nubel, G., & Jaschke, A. (2015). NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs. Nature, 519(7543), 374–377. https://doi.org/10.1038/nature14020

Carlile, T. M., Rojas-Duran, M. F., Zinshteyn, B., Shin, H., Bartoli, K. M., & Gilbert, W. V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature, 515(7525), 143–146. https://doi.org/10.1038/nature13802

Cartwright, I. L., Hutchinson, D. W., & Armstrong, V. W. (1976). The reaction between thiols and 8-azidoadenosine derivatives. Nucleic Acids Research, 3, 2331–2340. https://doi.org/10.1093/nar/3.9.2331

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B. M., Strein, C., … Hentze, M. W. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell, 149(6), 1393–1406. https://doi.org/10.1016/j.cell.2012.04.031

Chen, K., Lu, Z. K., Wang, X., Fu, Y., Luo, G. Z., Liu, N., … He, C. (2015). High-resolution N-6-methyladenosine (m(6)A) mapping using photo-crosslinking-assisted m(6)A sequencing. Angewandte Chemie (International Ed. in English), 54(5), 1587–1590. https://doi.org/10.1002/anie.201410647

Chen, Y. G., Kowtoniu, W. E., Agarwal, I., Shen, Y., & Liu, D. R. (2009). LC/MS analysis of cellular RNA reveals NAD-linked RNA. Nature Chemical Biology, 5(12), 879–881. https://doi.org/10.1038/nchembio.235

Cleary, M. D., Meiering, C. D., Jan, E., Guymon, R., & Boothroyd, J. C. (2005). Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay. Nature Biotechnology, 23, 232. https://doi.org/10.1038/nbt1061
Osipenko, A., Plotnikova, A., Naintye, M., Masevicuis, V., Klimasauskas, S., & Vilkaitis, G. (2017). Oligonucleotide-addressed covalent 3'-terminal derivatization of Small RNA strands for enrichment and visualization. *Angewandte Chemie (International Ed. in English)*, 56(23), 6507–6510. https://doi.org/10.1002/anie.201701448

Padrón, A., Iwasaki, S., & Ingolia, N. T. (2018). Proximity RNA labeling by APEX-seq reveals the organization of translation initiation complexes and repressive RNA granules. *bioRxiv*, 454066. https://doi.org/10.1101/454066

Paige, J. S., Wu, K. Y., & Jaffrey, S. R. (2011). RNA mimics of green fluorescent protein. *Science*, 333, 642–646. https://doi.org/10.1126/science.1207339

Pang, X., Zhou, D., Song, Y., Pei, D., Wang, J., Guo, Z., & Yang, R. (2004). Bacterial mRNA purification by magnetic capture-hybridization method. *Microbiology and Immunology*, 48(2), 91–96.

Paulsen, M. T., Veloso, A., Prasad, J., Bedi, K., Ljungman, E. A., Magnuson, B., … Ljungman, M. (2014). Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods*, 67(1), 45–54. https://doi.org/10.1016/j.ymeth.2013.08.015

Peters, W., Willnow, S., Dusiksen, M., Kleine, H., Macherey, T., Duncan, K. E., … Weinhold, E. (2010). Enzymatic site-specific functionalization of protein methyltransferase substrates with alkynes for click labeling. *Angewandte Chemie (International Ed. in English)*, 49, 5170–5173. https://doi.org/10.1002/anie.201001240

Plotnikova, A., Osipenko, A., Masevicuis, V., Vilkaitis, G., & Klimasauskas, S. (2014). Selective covalent labeling of miRNA and siRNA duplexes using HEN1 methyltransferase. *Journal of the American Chemical Society*, 136(39), 13550–13553. https://doi.org/10.1021/ja507390s

Presolski, S. I., Hong, V., Cho, S. H., & Finn, M. G. (2010). Tailored ligand acceleration of the cu-catalyzed azide-alkyne cycloaddition reaction: Practical and mechanistic implications. *Journal of the American Chemical Society*, 132, 14570–14576. https://doi.org/10.1021/ja105743g

Pyka, A. M., Domnick, C., Braun, F., & Kath-Schorr, S. (2014). Diels-alder cycloadditions on synthetic RNA in mammalian cells. *Bioconjugate Chemistry*, 25, 1438–1443. https://doi.org/10.1021/bc500302y

Qiu, L., Wu, C., You, M., Han, D., Chen, T., Zhu, G., … Tan, W. (2013). A targeted, self-delivered, and photocontrolled molecular beacon for mRNA detection in living cells. *Journal of the American Chemical Society*, 135, 12952–12955. https://doi.org/10.1021/ja406252w

Rau, K., & Rentmeister, A. (2017). Making the message clear: Concepts for mRNA imaging. *ACS Central Science*, 3, 701–707. https://doi.org/10.1021/acscentsci.7b00251

Riml, C., Amort, T., Rieder, D., Gasser, C., Lusser, A., & Micura, R. (2017). Osmium-mediated transformation of 4-thiouridine to cytidine as key to study RNA dynamics by sequencing. *Angewandte Chemie (International Ed. in English)*, 56(43), 13479–13483. https://doi.org/10.1002/anie.201707465

Rinn, J. L., Wang, J. K., Allen, N., Brugmann, S. A., Mikels, A. J., Liu, H., … Jemielity, J. (2008). A dermal HOX transcriptional program regulates site-specific epidermal fate. *Genes & Development*, 22(3), 303–307.

Rio, D. C., Ares, M., Hannon, G. J., & Nilsen, T. W. (2010). Enrichment of poly(A)+ mRNA using immobilized oligo(dT). *Cold Spring Harbor Protocols*, 2010(7), pdb.prot5454. https://doi.org/10.1101/pdb.prot5454

Roberts, T. C., Hart, J. R., Kaikkonen, M. U., Weinberg, M. S., Vogt, P. K., & Morris, K. V. (2015). Quantification of nascent transcription by bromouridine immunocapture nuclear run-on RT-qPCR. *Nature Protocols*, 10(8), 1198–1211. https://doi.org/10.1038/nprot.2015.076

Rzydzik, A. M., Lukaszewicz, M., Zuberek, J., Kowalska, J., Darzynkiewicz, Z. M., Darzynkiewicz, E., & Jemielity, J. (2009). Synthetic dinucleotide mRNA cap analogs with tetraphosphate 5’5’ bridge containing methylenebis(phosphonate) modification. *Organic & Biomolecular Chemistry*, 7, 4763. https://doi.org/10.1039/b911347a

Rzydzik, A. M., Warminska, M., Sikorski, P. J., Baranowski, M. R., Walczak, S., Kowalska, J., … Jemielity, J. (2017). mRNA cap analogues substituted in the tetraphosphate chain with CX2: Identification of O-to-CC12 as the first bridging modification that confers resistance to decapping without impairing translation. *Nucleic Acids Research*, 45, 8661–8675. https://doi.org/10.1093/nar/gkx569

Samanta, B., Horning, D. P., & Joyce, G. F. (2018). 3’-End labeling of nucleic acids by a polymerase ribozyme. *Nucleic Acids Research*, 46, e103–e103. https://doi.org/10.1093/nar/gky513

Santi, D. V., Garrett, C. E., & Barr, P. J. (1983). On the mechanism of inhibition of DNA-cytosine methyltransferases by cytosine analogs. *Cell*, 33 (1), 9–10. https://doi.org/10.1016/0092-8674(83)90327-6

Schoch, J., Staudt, M., Samanta, A., Wiessler, M., & Jäschke, A. (2012). Site-specific one-pot dual labeling of DNA by orthogonal cyclodiaddition chemistry. *Bioconjugate Chemistry*, 23, 1382–1386. https://doi.org/10.1021/bc300181n

Schofield, J. A., Duffy, E. E., Kiefer, L., Sullivan, M. C., & Simon, M. D. (2018). TimeLapse-seq: Adding a temporal dimension to RNA sequencing through nucleoside recoding. *Nature Methods*, 15(3), 221–225. https://doi.org/10.1038/s41592-018-0145-z

Schulz, D., Holstein, J. M., & Rentmeister, A. (2013). A chemo-enzymatic approach for site-specific modification of the RNA cap. *Angewandte Chemie (International Ed. in English)*, 52, 7874–7878. https://doi.org/10.1002/anie.201302874

Schulz, D., & Rentmeister, A. (2014). Current approaches for RNA labeling in vitro and in cells based on click reactions. *Chembiochem*, 15(16), 2342–2347. https://doi.org/10.1002/cbic.201402240

Schwalb, B., Michel, M., Zacher, B., Frühauf, K., Demel, C., Tresch, A., … Cramer, P. (2016). TT-seq maps the human transient transcriptome. *Science*, 352(6290), 1225–1228. https://doi.org/10.1126/science.aad9841

Schwartz, S., Bernstein, D. A., Mumbach, M. R., Jovanovic, M., Herbst, R. H., Leon-Ricardo, B. X., … Regev, A. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell*, 159(1), 148–162. https://doi.org/10.1016/j.cell.2014.08.028

Sharma, A. K., Plant, J. J., Rangel, A. E., Meek, K. N., Anamisis, A. J., Hollien, J., & Heemstra, J. M. (2014). Fluorescent RNA labeling using self-alkylating ribozymes. *ACS Chemical Biology*, 9, 1680–1684. https://doi.org/10.1021(cb5002119

MUTHMANN ET AL.
Shen, C., Wang, X., Liu, Y., Li, Q., Yang, Z., Yan, N., … Yin, P. (2015). Specific RNA recognition by designer pentacriopeptide repeat protein. *Molecular Plant*, 8, 667–670. https://doi.org/10.1016/j.molp.2015.01.001

Shieh, P., & Bertozzi, C. R. (2014). Design strategies for bioorthogonal smart probes. *Organic & Biomolecular Chemistry*, 12, 9307–9320. https://doi.org/10.1039/c4ob01632g

Shih, H.-W., Kamber, D. N., & Prescher, J. A. (2014). Building better bioorthogonal reactions. *Current Opinion in Chemical Biology*, 21, 103–111. https://doi.org/10.1016/j.cbpa.2014.07.002

Shih, T. Y., & Khoury, G. (1976). Isolation of viral specific RNA from SV40 infected cells by viral DNA chemically linked to a cellulose matrix. *Biochemistry*, 15(3), 487–493.

Shu, X., Dai, Q., Wu, T., Bothwell, I. R., Yue, Y., Zhang, Z., … Liu, J. (2017). N(6)-allyladenosine: A new small molecule for RNA labeling identified by mutation assay. *Journal of the American Chemical Society*, 139(48), 17213–17216. https://doi.org/10.1021/jacs.7b06837

Singh, S., Zhang, J., Huber, T. D., Sunkara, M., Hurley, K., Goff, R. D., … Thorson, J. S. (2014). Facile chemoenzymatic strategies for the synthesis and utilization of S-adenosyl-L-methionine analogues. *Angewandte Chemie (International Ed. in English)*, 53, 3965–3969. https://doi.org/10.1002/anie.201308272

Sletten, E. M., & Bertozzi, C. R. (2011). From mechanism to mouse: A tale of two bioorthogonal reactions. *Accounts of Chemical Research*, 44, 666–676. https://doi.org/10.1021/ar200148z

Squires, J. E., Patel, H. R., Nousch, M., Sibbritt, T., Humphreys, D. T., Parker, B. J., … Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Research*, 40(11), 5023–5033. https://doi.org/10.1093/nar/gks144

Su, C., & Sordillo, L. M. (1998). A simple method to enrich mRNA from total prokaryotic RNA. *Molecular Biotechnology*, 10(1), 83–85. https://doi.org/10.1007/bf02745865

Tallafuss, A., Kelly, M., Gay, L., Gibson, D., Batzel, P., Karfilis, K. V., … Washbourne, P. (2015). Transcriptomes of post-mitotic neurons identify the usage of alternative pathways during adult and embryonic neuronal differentiation. *BMC Genomics*, 16, 1100. http://doi.org/10.1186/s12864-015-2215-8

Tallafuss, A., Washbourne, P., & Postlethwait, J. (2014). Temporally and spatially restricted gene expression profiling. *Current Genomics*, 15(4), 278–292.

Taylor, M. T., Blackman, M. L., Dmitrenko, O., & Fox, J. M. (2011). Design and synthesis of highly reactive dienophiles for the tetrazine–cyclooctene ligation. *Journal of the American Chemical Society*, 133, 9646–9649. https://doi.org/10.1021/ja201844c

Tomkuvienė, M., Clouet-d’Orval, B., Černiauskas, I., Weinhold, E., & Klimažauskas, S. (2012). Programmable sequence-specific click-labeling of RNA using arachael box C/D RNP methyltransferases. *Nucleic Acids Research*, 40, 6765–6773. https://doi.org/10.1093/nar/gks381

Tyagi, S., & Kramer, F. R. (1996). Molecular beacons: Probes that fluoresce upon hybridization. *Nature Biotechnology*, 14, 303–308. https://doi.org/10.1038/nbt0996-765

Ule, J., Jensen, K. B., Ruggiu, M., Mele, A., Ule, A., & Darnell, R. B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science*, 302(5648), 1212–1215. https://doi.org/10.1126/science.1090095

Vrabel, M., Krūle, P., Brunner, K. M., Gattner, M. J., López-Carrillo, V., De Vivie-Riedle, R., & Carell, T. (2013). Norbornenes in inverse electron-demand diels-alder reactions. *Chemistry - A European Journal*, 19, 13309–13312. https://doi.org/10.1002/chem.201301838

Wang, R., Islam, K., Liu, Y., Zheng, W., Tang, H., Lailler, N., … Luo, M. (2013). Profiling genome-wide chromatin methylation with engineered postranslation apparatus within living cells. *Journal of the American Chemical Society*, 135, 1048–1056. https://doi.org/10.1021/ja309412s

Wang, R., Zheng, W., Yu, H., Deng, H., & Luo, M. (2011). Labeling substrates of protein arginine methyltransferase with engineered enzymes and matched S-adenosyl-l-methionine analogues. *Journal of the American Chemical Society*, 133, 7648–7651. https://doi.org/10.1021/ja2066719

Wang, X., Lu, Z., Gomez, A., Hon, G. C., Yue, Y., Han, D., … He, C. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature*, 505(7481), 117–120. https://doi.org/10.1038/nature12730

Wang, X., Zhao, B. S., Roundtree, I. A., Lu, Z., Han, D., Ma, H., … He, C. (2015). N(6)-methyladenosine modulates messenger RNA translation efficiency. *Cell*, 161(6), 1388–1399. https://doi.org/10.1016/j.cell.2015.05.014

Wieczorek, A., Werther, P., Euchner, J., & Wombacher, R. (2017). Green- to far-red-emitting fluorogenic tetrazine probes – Synthetic access and no-wash protein imaging inside living cells. *Chemical Science*, 8, 1506–1510. https://doi.org/10.1039/C6SC03879D

Willnow, S., Martin, M., Lüscher, B., & Weinhold, E. (2012). A selenium-based click AdoMet analogue for versatile substrate labeling with wild-type protein methyltransferases. *ChemBioChem*, 13, 1167–1173. https://doi.org/10.1002/cbic.201100781

Winz, M.-L., Cahová, H., Nübel, G., Frindt, J., Höfer, K., & Jäschke, A. (2016). Capture and sequencing of NAD-capped RNA sequences with NAD captureSeq. *Nature Protocols*, 12, 122. https://doi.org/10.1038/nprot.2016.163

Winz, M.-L., Samanta, A., Benzinger, D., & Jäschke, A. (2012). Site-specific terminal and internal labeling of RNA by poly(a) polymerase tailing and copper-catalyzed or copper-free strain-promoted click chemistry. *Nucleic Acids Research*, 40, 1–13. https://doi.org/10.1093/nar/gks062

Wojtczak, B. A., Sikorski, P. J., Fac-Dabrowska, K., Nowicka, A., Warminski, M., Kubacka, D., … Jemielity, J. (2018). 5'-Phosphorothiolate dinucleotide cap analogues: Reagents for messenger RNA modification and potent small-molecular inhibitors of decapping enzymes. *Journal of the American Chemical Society*, 140, 5987–5999. https://doi.org/10.1021/jacs.8b02597

Woodford, T. A., Schlegel, R., & Pardee, A. B. (1988). Selective isolation of newly synthesized mammalian messenger-RNA after invivo labeling with 4-thiouridine or 6-thioguanosine. *Analytical Biochemistry*, 171(1), 166–172.

Wu, J. C., & Santi, D. V. (1987). Kinetic and catalytic mechanism of HhaI methyltransferase. *Journal of Biological Chemistry*, 262(10), 4778–4786.

Yamada, T., & Akimitsu, N. (2019). Contributions of regulated transcription and mRNA decay to the dynamics of gene expression. *Wiley Interdisciplinary Reviews. RNA*, 10(1), e1508. https://doi.org/10.1002/wrna.1508
Yamazoe, S., Liu, Q., McQuade, L. E., Deiters, A., & Chen, J. K. (2014). Sequential gene silencing using wavelength-selective caged morpholino oligonucleotides. *Angewandte Chemie (International Ed. in English)*, 53, 10114–10118. https://doi.org/10.1002/anie.201405355

Zhang, D., Zhou, C. Y., Busby, K. N., Alexander, S. C., & Devaraj, N. K. (2018). Light-activated control of translation by enzymatic covalent mRNA Labeling. *Angewandte Chemie (International Ed. in English)*, 57, 2822–2826. https://doi.org/10.1002/anie.201710917

Zhang, H., Zhong, H., Zhang, S., Shao, X., Ni, M., Cai, Z., … Xia, Y. (2019). NAD tagSeq reveals that NAD$^+$-capped RNAs are mostly produced from a large number of protein-coding genes in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 116(24), 12072–12077. https://doi.org/10.1073/pnas.1903683116

Zhang, Y., Ptacin, J. L., Fischer, E. C., Aerni, H. R., Caffaro, C. E., San Jose, K., … Romesberg, F. E. (2017). A semi-synthetic organism that stores and retrieves increased genetic information. *Nature*, 551, 644–647. https://doi.org/10.1038/nature24659

Zhao, B. X. S., Roundtree, I. A., & He, C. (2017). Post-transcriptional gene regulation by mRNA modifications. *Nature Reviews Molecular Cell Biology*, 18(1), 31–42. https://doi.org/10.1038/nrm.2016.132

Zhou, C. Y., Alexander, S. C., & Devaraj, N. K. (2017). Fluorescent turn-on probes for wash-free mRNA imaging via covalent site-specific enzymatic labeling. *Chemical Science*, 8, 7169–7173. https://doi.org/10.1039/C7SC03150E

Ziemniak, M., Strenkowska, M., Kowalska, J., & Jemielity, J. (2013). Potential therapeutic applications of RNA cap analogs. *Future Medicinal Chemistry*, 5, 1141–1172. https://doi.org/10.4155/fmc.13.96

**How to cite this article:** Muthmann N, Hartstock K, Rentmeister A. Chemo-enzymatic treatment of RNA to facilitate analyses. *WIREs RNA*. 2020;11:e1561. [https://doi.org/10.1002/wrna.1561](https://doi.org/10.1002/wrna.1561)