

A Covalent Approach for Site-Specific RNA Labeling in Mammalian Cells**

Fahui Li, Jianshu Dong, Xiaosong Hu, Weimin Gong, Jiasong Li, Jing Shen, Huifang Tian, and Jiangyun Wang*

Abstract: Advances in RNA research and RNA nanotechnology depend on the ability to manipulate and probe RNA with high precision through chemical approaches, both *in vitro* and in mammalian cells. However, covalent RNA labeling methods with scope and versatility comparable to those of current protein labeling strategies are underdeveloped. A method is reported for the site- and sequence-specific covalent labeling of RNAs in mammalian cells by using tRNA^{lle2}-agmatidine synthetase (Tias) and click chemistry. The crystal structure of Tias in complex with an azide-bearing agmatine analogue was solved to unravel the structural basis for Tias/substrate recognition. The unique RNA sequence specificity and plastic Tias/substrate recognition enable the site-specific transfer of azide/alkyne groups to an RNA molecule of interest *in vitro* and in mammalian cells. Subsequent click chemistry reactions facilitate the versatile labeling, functionalization, and visualization of target RNA.

It has recently emerged that RNA has functions as diverse and important as those of proteins.^[1] To unravel the dynamic trafficking, localization, and interactions of RNA in living cells, numerous methods have been developed for the non-covalent labeling of RNAs in mammalian cells, and dramatic progress has been made.^[2] Compared to the noncovalent

approaches,^[2] bioorthogonal chemistry^[3] relies upon the specific and covalent attachment of probe molecules to the biomacromolecule of interest (BOI). As a result, it offers several advantages: 1) the affinity can be considered to be infinity, therefore stringent wash conditions can be applied to rigorously purify the BOI; 2) the high affinity allows for the visualization of low-abundance BOIs, which is particularly important for non-coding RNAs;^[1a] 3) this method is extremely versatile; probe molecules harboring fluorescence, NMR, IR, or EPR functional groups can be conveniently conjugated to the BOI.^[3] Because of these unique features, bioorthogonal chemistry has been successfully employed in protein and glycan functional studies to visualize their expression, localization, and interaction partners in mammalian cells.^[3]

While engineering of the translational^[4] or post-translational machinery^[5] in combination with bioorthogonal chemistry has enabled protein labeling with unprecedented precision and versatility,^[6] comparable methods for the site- and sequence-specific tagging of RNAs in mammalian cells is currently underdeveloped.^[7] To achieve this, we need to engineer unique components of the post-transcriptional machinery and employ bioorthogonal chemistry. First, an RNA-modifying enzyme must recognize a specific site in a unique RNA sequence in the transcriptome. Second, this enzyme must be able to transfer a small molecule bearing a unique chemical handle (such as an azide or alkyne group) to the specific site. Subsequent covalent labeling through bioorthogonal chemistry then leads to the site-specific conjugation of reporter groups, thus enabling fluorescence, NMR, EPR, or IR spectroscopy measurements (Figure 1).

To search for a post-transcriptional modification module that can be used for precise RNA labeling in mammalian cells, we turned to the tRNA-modifying enzymes, which are known to catalyze around 100 kinds of tRNA modification.^[8] Faithful translation of the genetic code relies on the chemical modification of tRNA,^[9] especially at the wobble position 34 of the anticodon.^[10] Many tRNA-modifying enzymes acting on this position are found only in specific domains, thus implying that they may have evolved independently.^[11] Notably, the C34 position of the AUA-decoding tRNA^{lle2} is modified with lysine by the enzyme lysidine synthetase (Tils) in bacteria,^[12] but with agmatine (AGM, **1**; Scheme 1) by the enzyme tRNA^{lle2}-agmatidine synthetase (Tias) in archaea.^[13,14] In eukaryotes, tRNA^{lle} bearing pseudouridine (Ψ) or inosine (I) at the wobble position decodes the AUA codon.^[9,15] As shown previously, *Archaeoglobus fulgidus* (Af) Tias recognition of tRNA^{lle2} requires seven nucleotides: G1, G2, C34, U36, A37, C71, and C72 (Figure S13A in the

[*] F. H. Li,^[†] J. S. Dong,^[†] W. M. Gong, J. S. Li, J. Y. Wang
Laboratory of RNA Biology and Laboratory of Quantum Biophysics
Institute of Biophysics, Chinese Academy of Sciences
15 Datun Road, Chaoyang District, Beijing, 100101 (China)
E-mail: jwang@ibp.ac.cn

J. Y. Wang
Beijing National Laboratory for Molecular Sciences (BNLMS)
Institute of Chemistry, Chinese Academy of Sciences
Beijing, 100190 (China)

X. S. Hu^[†]
Department of Chemistry, School of Chemistry
Chemical Engineering and Life Sciences
Wuhan University of Technology, Wuhan 430070 (China)

J. Shen, H. F. Tian
The Laboratory of Carcinogenesis and Translational Research
(Ministry of Education)
Core Laboratory, Peking University School of Oncology
Beijing Cancer Hospital & Institute, Beijing, 100142 (China)

[†] These authors contributed equally to this work.

[**] We gratefully acknowledge the Major State Basic Research Program of China (2015CB856203), the CAS grant (KJZD-EW-L01), and the National Science Foundation of China (91440116, 21325211, 91313301, 31370016, 21102172). We thank Yan Teng for help with fluorescence imaging.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201410433>.

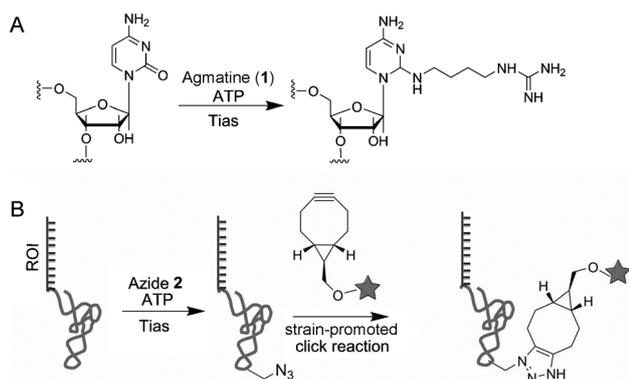
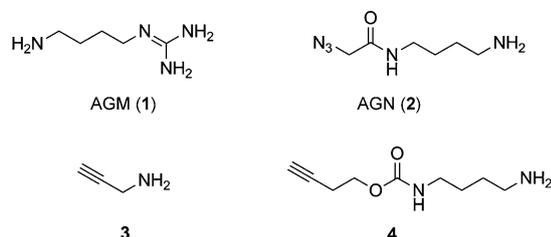


Figure 1. A) In the presence of ATP, Tias catalyzes the conjugation of agmatine to C34 of tRNA^{Ile2}, thereby resulting in the formation of agmatidine. B) Labeling of an RNA of interest (ROI). Tias facilitates the site-specific transfer of an agmatine analogue bearing an azide group to the ROI-tRNA^{Ile2} fusion RNA. Subsequent strain-promoted azide-alkyne cycloaddition allows the conjugation of biophysical probes.



Scheme 1. Structure of agmatine (AGM) and agmatine analogues.

Supporting Information).^[13,14] Since neither Tias nor tRNA bearing this set of identity elements is present in mammalian cells (Figure S13),^[16] we envisioned that the introduction of Tias and archaeal tRNA^{Ile2} into mammalian cells could facilitate the sequence- and site-specific labeling of RNA if Tias can transfer small molecule probes bearing azide or alkyne functional groups to an RNA of interest.

As a first step to developing a covalent method for labeling RNAs, we synthesized three AGM analogues: *N*-(4-aminobutyl)-2-azidoacetamide (AGN, **2**), 2-propynylamine (**3**), and (but-3-yn-1-yl)(4-aminobutyl)carbamate (**4**; Scheme 1). Modification of Af tRNA^{Ile2} was performed with Tias (8 μM), Af tRNA^{Ile2} (1 μM), Tris-HCl (100 mM), pH 8.0, KCl (10 mM), MgCl₂ (5 mM), DTT (5 mM), ATP (1 mM), and substrates 1, 2, 3, or 4 (1 mM). AGM (**1**) was used as a positive control. The modified and unmodified RNAs were then separated by 6.5% acid-urea polyacrylamide gel electrophoresis (PAGE), as previously described.^[17] When tRNA^{Ile2} was modified, the modified RNA exhibited lower mobility (Figure 2 A). To our delight, we found that in the presence of substrates **2**, **3**, or **4**, tRNA^{Ile2} was covalently modified, thus resulting in lower mobility in the acid-urea polyacrylamide gel. Perhaps owing to the lower molecular weight of compound **3**, tRNA^{Ile2} modified with compound **3** exhibited higher mobility than tRNA^{Ile2} modified with compound **2**. In the presence of compound **4**, two major bands were present, thus indicating that tRNA^{Ile2} was only partially modified with

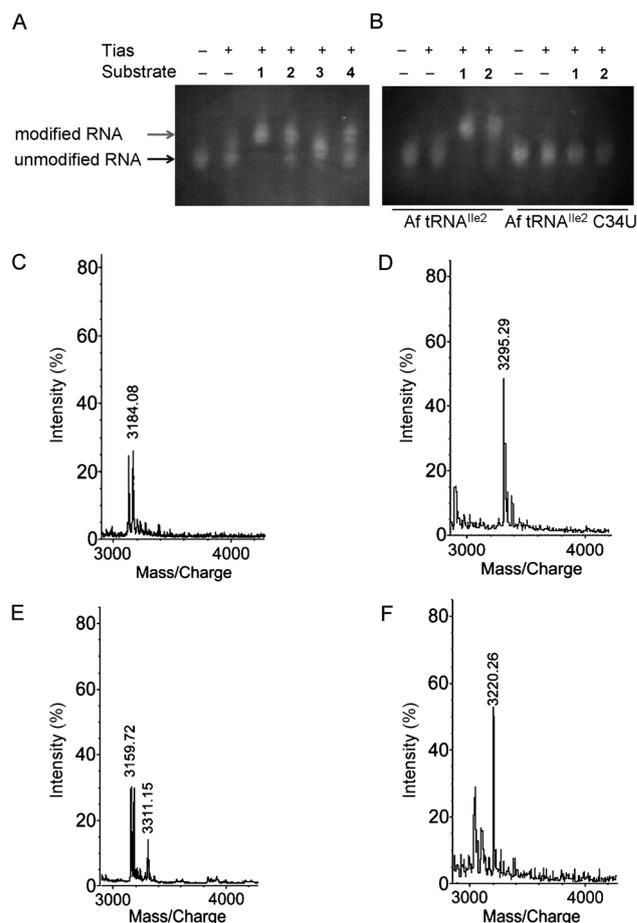


Figure 2. A) Acid-urea PAGE with tRNA^{Ile2} modified by using Tias and AGM analogues. The gel was stained by ethidium bromide (EB). B) While Tias catalyzes the efficient conjugation of AGM (**1**) and AGN (**2**) to Af tRNA^{Ile2}, the reaction does not occur for the Af tRNA^{Ile2} C34U mutant, thus indicating that the small-molecule probes are transferred specifically to the C34 position. The reaction mixtures were analyzed by acid-urea PAGE. MALDI-TOF analyses are shown for unlabeled Af tRNA^{Ile2} (C), Af tRNA^{Ile2} modified with **1** (D), Af tRNA^{Ile2} modified with **2** (E), and Af tRNA^{Ile2} modified with **3** (F).

compound **4**. The much bulkier carbamate and alkyne functional groups may result in inefficient recognition of compound **4** by Tias. We therefore choose compounds **2** and **3** for further investigation.

We then investigated whether Tias catalyzes the specific modification of nucleotide C34 of Af tRNA^{Ile2}. While Tias catalyzes the efficient ligation of compounds **1** and **2** to Af tRNA^{Ile2}, this activity was abolished for the C34U mutant (Figure 2 B). These results indicate that Tias can target both the natural substrate AGM and the unnatural AGM analogues specifically to nucleotide C34 of Af tRNA^{Ile2}. To further demonstrate that compounds **2** and **3** conjugate precisely with nucleotide C34 of Af tRNA^{Ile2}, we digested Af tRNA^{Ile2} with RNase T1 (Figure S15)^[18] and then analyzed the RNA fragment containing the anticodon loop by MALDI-TOF. The unmodified tRNA^{Ile2} fragment gave an observed average mass of 3184.08 Da, which closely matches the calculated mass of 3184.4 Da (Figure 2). After AGM or 2-propynylamine modification, the tRNA^{Ile2} fragment gave

observed average masses of 3295.3 and 3220.3 Da, respectively, which match the calculated masses of 3295.5 and 3220.3 Da. After AGN modification, the tRNA^{Ile2} fragment gave an observed average mass of 3311.2 Da, while the calculated mass is 3337.5 Da. This can be explained by the loss of two nitrogen atoms and the addition of two hydrogen atoms during the laser irradiation used for MALDI-TOF, as previously observed.^[19] Taken together, these results demonstrate that Tias can facilitate the conjugation of azide/alkyne-bearing AGM analogues **2** and **3** to nucleotide C34 in Af tRNA^{Ile2}.

To unravel the molecular basis for AGN recognition by Tias, we solved the structure of Tias with AGN bound to its active site. The main chain carbonyl group of Asp193 and the side chain of Asn194 form two hydrogen bonds with the azide group of AGN, while the guanidine group of Arg217 forms a hydrogen bond with the carbonyl group of AGN (Figure 3).

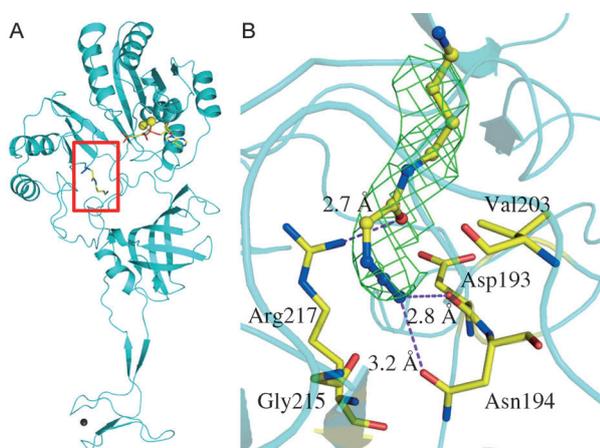


Figure 3. A) Overall structure of Tias in complex with AGN. The overall structure of Tias resembles a tennis racket. The AGN binding site is highlighted by a red box. B) Structure of the AGN binding site of Tias. The Fo-Fc electron density map of AGN was contoured at 2.7 σ . Nitrogen atoms are shown in blue, oxygen atoms in red, and carbon atoms in yellow.

Meanwhile, the side chain of Val203 stabilizes the carbon chain of agmatine through hydrophobic interactions. Similar hydrophobic interactions may also be needed to facilitate Tias recognition of 2-propynylamine. Since Tias can also use compounds **3** and **4** as substrates, although these are either much smaller or much bulkier than compound **2**, it appears that Tias has a somewhat plastic binding pocket for primary amine substrates. We are currently testing whether Tias can also use primary amines bearing cyclopropene groups as substrates, which may facilitate photoclick reactions or inverse electron demand Diels–Alder reactions.^[20]

We then tested whether Tias can facilitate RNA fluorescence labeling through bioorthogonal chemistry. tRNA^{Ile2} with or without AGN (**2**) modification was incubated with BCN-FITC and then analyzed by 6.5% acid-urea PAGE, followed by fluorescent imaging. tRNA^{Ile2} modified with AGN reacted with BCN-FITC selectively, whereas tRNA^{Ile2} without AGN modification did not react (Figure 4 A and 4B).

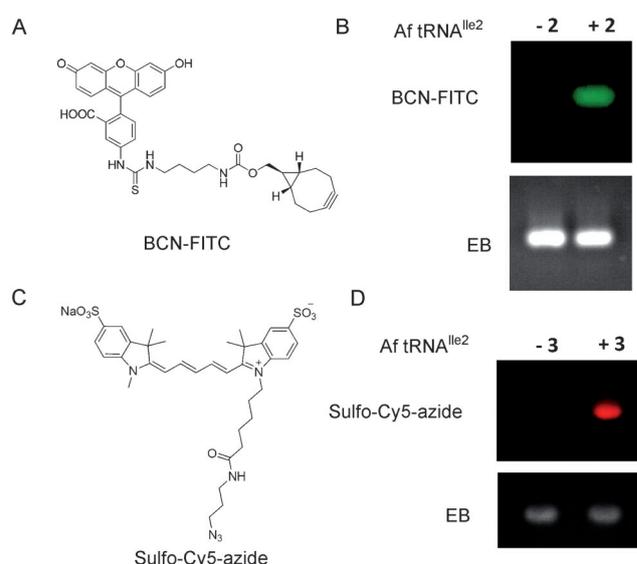


Figure 4. A) Structure of BCN-FITC. B) AGN-modified tRNA^{Ile2} was visualized through strain-promoted click reaction with BCN-FITC. The bottom panel (EB staining) indicates that equivalent amounts of tRNA^{Ile2} with and without AGN modification were loaded. C) The structure of Sulfo-Cy5-azide. D) tRNA^{Ile2} modified with 2-propynylamine was visualized through Cu^I-catalyzed click reaction with Sulfo-Cy5-azide. The bottom panel (EB staining) shows that equivalent amounts of tRNA^{Ile2} with and without 2-propynylamine modifications were loaded.

These results indicate that BCN-FITC reacts specifically with the azide group present on tRNA^{Ile2}. Under similar conditions and in the presence of Cu^I catalyst (1 mM), Sulfo-Cy5-azide reacts specifically with tRNA^{Ile2} modified with 2-propynylamine (**3**), but not with tRNA^{Ile2} without the modification (Figure 4 C and 4D).

Once we had demonstrated that Tias can facilitate the site-specific incorporation of azide and alkyne groups into tRNA^{Ile2}, we then asked whether this strategy could be used to label any RNA of interest (ROI). Through in vitro run-off transcription, we synthesized tRNA^{Ile2} harboring 3' and 5' extensions (termed tRNA^{Ile2}-3-5, Figure 5 A). We then incubated tRNA^{Ile2}-3-5 with Tias in the presence or absence of AGN. After removing excess AGN, BCN-FITC was added. Acid-urea PAGE was then used to analyze the reaction. Only in the presence of AGN did BCN-FITC selectively label tRNA^{Ile2}-3-5 (Figure 5B), thus indicating that Tias can catalyze the conjugation of AGN to tRNA^{Ile2}-3-5 in the presence of 3' and 5' extensions. This result is consistent with previous reports that the agmatidine modification occurs early in the tRNA maturation process to ensure high-fidelity translation.^[13-14]

We then asked whether our method can be used to label a specific RNA in the mammalian cell transcriptome. 293T cells expressing Tias with or without tRNA^{Ile2}-5S fusion RNA expression were grown in medium containing AGN, followed by cell lysis. Subsequently, total cellular RNA was incubated with BCN-Cy5. In the absence of tRNA^{Ile2}-5S expression, BCN-Cy5 did not label any RNA in the 293T transcriptome (Figure 5 C and 5D). When cells expressed tRNA^{Ile2}-5S,

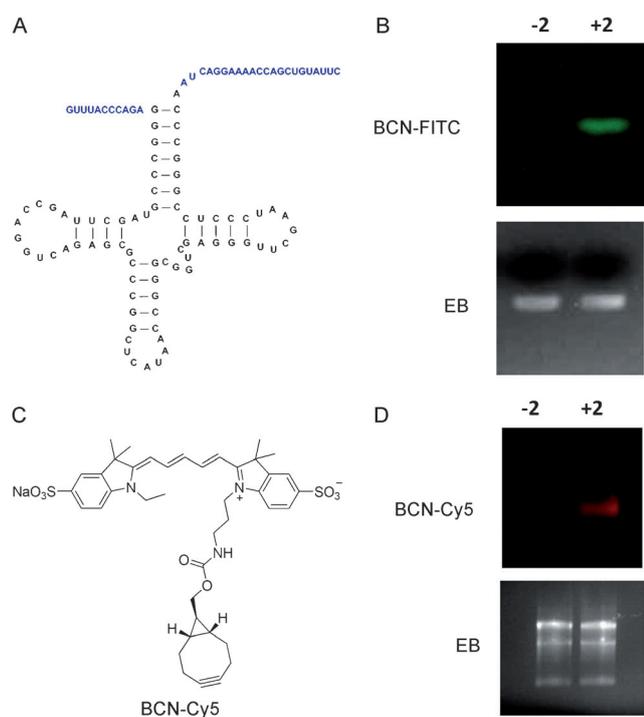


Figure 5. A) Sequence and secondary structure of tRNA^{Ile2} in the presence of 3' and 5' extensions (denoted tRNA^{Ile2}-3-5). B) AGN-modified tRNA^{Ile2}-3-5 is visualized through strain-promoted click reaction with BCN-FITC. The bottom panel (EB staining) shows that equivalent amounts of tRNA^{Ile2}-3-5 with and without AGN modification were loaded. C) Structure of BCN-Cy5. D) Tias specifically modify tRNA^{Ile2}-5S with AGN in the 293T cells transcriptome, which allows subsequent strain-promoted click reaction with BCN-Cy5. The bottom panel (EB staining) shows that equivalent amounts of total RNA with and without tRNA^{Ile2}-5S expression were loaded.

however, BCN-Cy5 specifically labeled tRNA^{Ile2}-5S. Reverse transcription polymerase chain reaction (RT PCR) experiments confirmed that tRNA^{Ile2}-5S existed as a full-length fusion RNA in the 293T cells (Figure S15). To further demonstrate the utility of our method for RNA imaging in mammalian cells, U2OS cells expressing Tias with or without tRNA^{Ile2}-5S fusion RNA expression were grown in the presence of 2-propynylamine and then fixed, permeabilized, and incubated with Sulfo-Cy5-azide and a Cu^I catalyst. We found that only the cells expressing tRNA^{Ile2}-5S fusion RNA could be labeled and imaged with Sulfo-Cy5-azide (Figure 6). These results demonstrate that in mammalian cells, Tias can recognize the unique Af tRNA^{Ile2} sequence in the transcriptome, thus enabling the sequence- and site-specific conjugation of an azide/alkyne group to an RNA of interest bearing the tRNA^{Ile2} tag. Strain-promoted or Cu^I-catalyzed click chemistry then enables the specific labeling and imaging of the RNA of interest in the cells.

A great deal of modern RNA biology and RNA biotechnology involves the site-specific incorporation of modified nucleotides into RNA molecules through chemical approaches.^[1] However, the length of synthetic RNA is typically less than 30–50 nucleotides owing to limitations in coupling efficiency. Our new method relies on the RNA transcription machinery and the high sequence specificity of

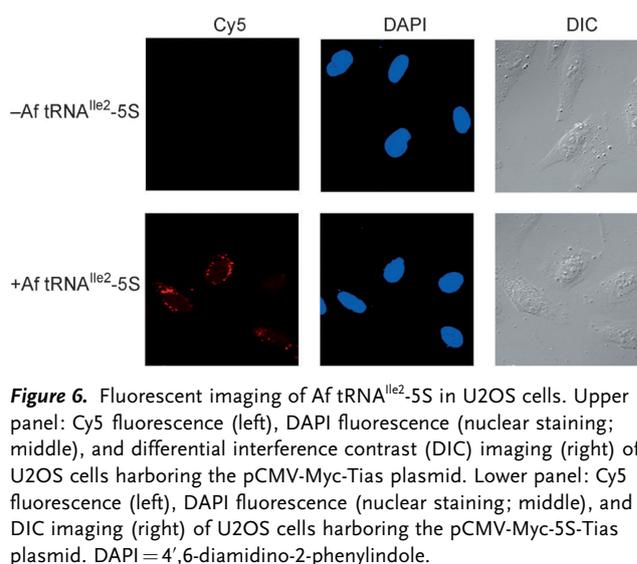


Figure 6. Fluorescent imaging of Af tRNA^{Ile2}-5S in U2OS cells. Upper panel: Cy5 fluorescence (left), DAPI fluorescence (nuclear staining; middle), and differential interference contrast (DIC) imaging (right) of U2OS cells harboring the pCMV-Myc-Tias plasmid. Lower panel: Cy5 fluorescence (left), DAPI fluorescence (nuclear staining; middle), and DIC imaging (right) of U2OS cells harboring the pCMV-Myc-5S-Tias plasmid. DAPI = 4',6-diamidino-2-phenylindole.

the RNA modifying enzyme Tias to enable the highly specific conjugation of biophysical probes to any RNA of interest in mammalian cells. This new technique should be useful for investigating RNA folding, RNA/protein interactions, RNA transportation, and RNA modification, as well as for constructing RNA-based nanomaterials, in vitro and in mammalian cells.^[1]

In comparison to noncovalent RNA imaging methods such as use of the Spinach aptamer,^[2] the major advantages of the Tias strategy are its specificity, highly affinity, and versatility. The Tias labeling strategy relies on the selective recognition of Af tRNA^{Ile2} by Tias (which are both absent in mammalian cells). Since a covalent approach is then employed after an azide or alkyne group is introduced, a covalent bond is created between the fluorophore and target RNA. This would potentially allow the imaging of low-abundance RNA. By contrast, the noncovalent interaction between the GFP chromophore analogue and the Spinach aptamer has affinity only in the micromolar range.^[2] Finally, the Tias strategy enables versatile site-specific RNA labeling with fluorescent/IR/NMR/EPR agent through click chemistry, without changing the Tias protein or tRNA^{Ile2} sequence. The challenge, however, is to develop fluorogenic cyclooctyne dyes that are membrane permeable, reasonably hydrophobic, and suitable for live cell RNA imaging. We are currently working to achieve this goal.

Experimental Section

In vitro modification of Af tRNA^{Ile2}: Modification of Af tRNA^{Ile2} was performed with Tias (8 μM), Af tRNA^{Ile2} (1 μM), Tris-HCl (100 mM), pH 8.0, KCl (10 mM), MgCl₂ (5 mM), DTT (5 mM), ATP (1 mM), and substrates **1**, **2**, **3**, or **4** (1 mM). AGM (**1**) was used as a positive control. Each reaction mixture was incubated at 37 °C for 2 h. A 5 μL aliquot from each reaction mixture was analyzed by 6.5 % acid-urea PAGE (0.1M sodium acetate, pH 5.0, 8M urea). In-gel fluorescence images were taken at a KF BIO-2800 gel imager and the gel was then stained with ethidium bromide (EB).

Specific fluorescent labeling of Af tRNA^{Ile2} in vitro: The AGN (2)-modified Af tRNA^{Ile2} (1 μM) was mixed with BCN-FITC (100 μM)

- Kramer, J. Back, J. M. Baskin, L. E. Smeenk, L. J. de Koning, J. H. van Maarseveen, L. de Jong, C. R. Bertozzi, H. Hiemstra, C. G. de Koster, *J. Proteome Res.* **2009**, *8*, 3702–3711.
- [20] Z. Yu, Q. Lin, *J. Am. Chem. Soc.* **2014**, *136*, 4153–4156.
- [21] C. Besanceney-Webler, H. Jiang, T. Zheng, L. Feng, D. S. del Amo, W. Wang, L. M. Klivansky, F. L. Marlow, Y. Liu, P. Wu, *Angew. Chem. Int. Ed.* **2011**, *50*, 8051–8056; *Angew. Chem.* **2011**, *123*, 8201–8206.
-

Communications

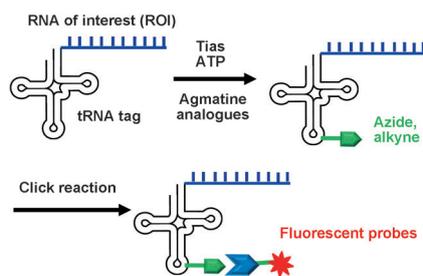
VIP

RNA Labeling



F. H. Li, J. S. Dong, X. S. Hu, W. M. Gong,
J. S. Li, J. Shen, H. F. Tian,
J. Y. Wang*     

A Covalent Approach for Site-Specific
RNA Labeling in Mammalian Cells



Tag and click: The ability to specifically label RNAs in vitro and in mammalian cells would be highly significant for RNA research, however, covalent RNA labeling methods with scope and versatility comparable to those for protein labeling have not been reported. A method was developed for the site- and sequence-specific covalent labeling of RNAs in mammalian cells, based on the action of tRNA^{Ile2}-agmatidine synthetase (Tias) and click chemistry.