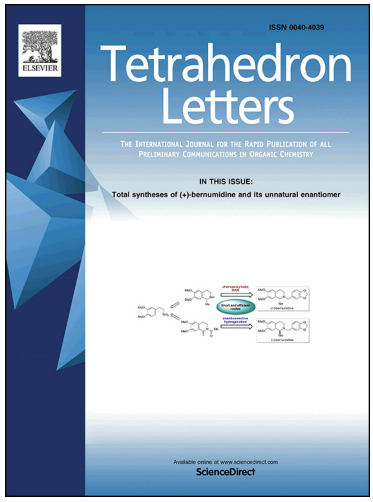


Protected pyrimidine nucleosides for cell-specific metabolic labeling of RNA

Samantha Beasley, Kim Nguyen, Michael Fazio, Robert C. Spitale

PII: S0040-4039(18)31130-4
DOI: https://doi.org/10.1016/j.tetlet.2018.09.040
Reference: TETL 50279

To appear in: *Tetrahedron Letters*

Received Date: 9 July 2018
Revised Date: 13 September 2018
Accepted Date: 14 September 2018

Please cite this article as: Beasley, S., Nguyen, K., Fazio, M., Spitale, R.C., Protected pyrimidine nucleosides for cell-specific metabolic labeling of RNA, *Tetrahedron Letters* (2018), doi: https://doi.org/10.1016/j.tetlet.2018.09.040

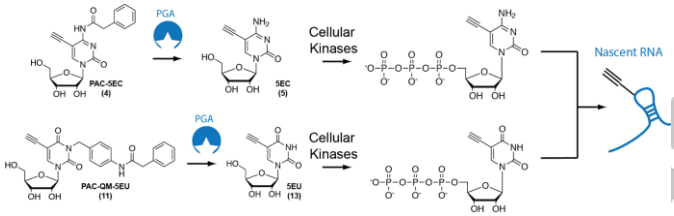
Graphical Abstract

To create your abstract, type over the instructions in the template box below.
Fonts or abstract dimensions should not be changed or altered.

Protected Pyrimidine Nucleosides for Cell-Specific Metabolic Labeling of RNA

Leave this area blank for abstract info.

Samantha Beasley, Kim Nguyen, Michael Fazio, and Robert C. Spitale



Tetrahedron Letters
journal homepage: www.elsevier.com

Protected pyrimidine nucleosides for cell-specific metabolic labeling of RNA

Samantha Beasley[a], Kim Nguyen[a], Michael Fazio[a], and Robert C. Spitale[a,b,*]

[a]*Department of Pharmaceutical Sciences, University of California, Irvine. Irvine, California 92697*
[b]*Department of Chemistry, University of California, Irvine. Irvine, California 92697*
**Corresponding author. e-mail address: rspitale@uci.edu*

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted
Available online

Keywords:
RNA
Metabolic Labeling
Alkynyl Nucleosides

ABSTRACT

RNA molecules can perform a myriad of functions, from the regulation of gene expression to providing the genetic blueprint for protein synthesis. Characterizing RNA expression dynamics, in a cell-specific manner, still remains a great challenge in biology. Herein we present a new set of protected alkynyl nucleosides for cell-specific metabolic labeling of RNA. We anticipate these analogs will find wide spread utility toward the goal of understanding RNA expression in complex cellular and tissue environments, even within living animals.

.

1. Introduction

Cells exist in complex environments in Nature. Understanding the molecular signatures of individual cell types has emerged as a key challenge in biology and medicine. However, precise chemical tools to do so are still overall lacking.[1] RNA molecules, once regarded as transient messengers between the genome and proteome, have now emerged as key players in regulating cell-fate decisions, development, cancer and neurological disorders.[2-4]

Cell-specific profiling of RNA molecules, in the form of gene expression analysis, is extremely challenging. Such analyses rely on cell-sorting techniques, which have been demonstrated to be both notoriously dirty and often detrimental to the stability of RNA molecules.[5, 6] Such obstacles manifest themselves in high false-positive rates in data analysis that can often mislead biologists aiming to understand changes in gene expression or the function of RNA.[7] Cell-specific metabolic labeling has emerged as a powerful technique in understanding the proteome and metabolome, yet, with few exceptions, analogous methods for RNA are still lacking. We recently reported the utility of protected nucleoside and nucleobase analogs for cell-specific metabolic labeling of RNA (Fig. 1, A).[8, 9] In such a case an exogenous enzyme can be expressed, within a cell-type of interest, to reveal RNA metabolic intermediates containing bio-orthogonal chemical handles. This approach allows the imaging and isolation of nascent and cell-type specific RNA. Such an approach is extremely powerful as it opens the door for cell-specific RNA profiling with the potential to reveal novel functions and gene expression signatures.

We previously demonstrated the enzyme uracil phosphoribosyl transferase (UPRT) can be used to incorporate 5-ethynyluracil into nascent RNA.[8] Although useful, UPRT is expressed in many lower organisms and as such the host pyrimidine salvage pathway compromises its utility inside these animals.[10] As a viable alternative, we matched a *N*6-phenylacetyladenosine analog with a 2'-azido handle (2'-AzA), with an enzyme, penicillin G amidase (PGA).[8] PGA is a bacterial-specific enzyme that cleaves an amide bond.[11] In our work, cellular RNA is endowed with azido functionality only when cells express PGA. While useful, we reasoned extending the toolset of the PGA enzymatic removal paradigm could further reveal additional nucleosides that can be used for metabolic labeling of RNA.

We reasoned that the azido moiety of 2'-AzA could be prone to reduction inside the cell, and more complicated environments *in vivo*, which would negate its ability to be utilized for affinity and purification experiments. Azide reduction has been observed with many other azido-containing analogs, and modified nucleosides, and can have a gross effect on their activity in cells and animals.[12] We also reasoned that adenosine analogs would be highly biased towards RNAs that have long polyA tails, a hallmark of messenger RNAs and long non-coding RNAs. Our own work with 2'-Az-A has demonstrated that it is largely incorporated into polyA tails.[13] This is a well-documented problem when understanding RNA expression levels through RNA sequencing – the current state of the art method of characterizing RNA expression.[14] Together, these matters underscore the need to expand the scope of novel analogs for cell-specific metabolic labeling of RNA.

2. Results and Discussion

We sought to utilize a different nucleobase structure for metabolic labeling of RNA. Penicillin G amidase is our current enzyme of choice and as such we decided to design and synthesize an alkynyl cytidine and uridine, which would be refractory to incorporation into nascent RNA due to the placement of the PGA-specific protecting group (Fig. 1, B). Cytidine and uridine are more reasonable choices for metabolic labeling as they are evenly distributed throughout the genome, and as we have seen before bulky substituents can prevent processing by endogenous kinase enzymes, such as uridine/cytidine kinase 1, thus preventing production of the triphosphate.[8] Furthermore, functionalizing the 5-position with an alkyne is not predicted to alter the sugar orientation or the *anti*-position of the nucleobase to disrupt RNA structure.

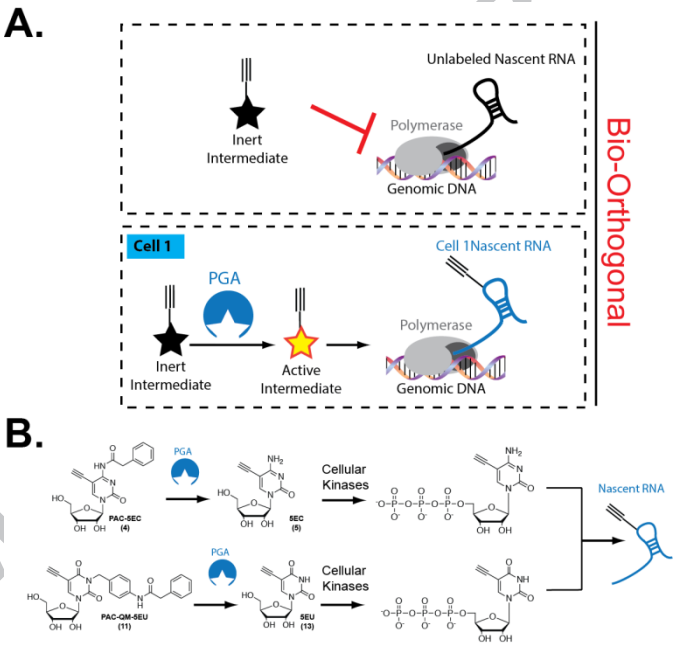


Figure 1. Protected analogs for cell-specific metabolic labeling of RNA. (A) Schematic of cell-specific metabolic labeling of RNA. (B) Structures of analogs synthesized herein and their proposed reactions with PGA.

PAC-5EC was synthesized in five steps starting from commercially available 5-iodocytidine (5-IC; ESI). 5-IC was subjected to a Sonogashira coupling to install the trimethylsilyl alkyne handle at the 5-position, followed by protection of the sugar alcohols with *tert*-butyldimethylsilyl groups. To synthesize the PAC portion, 2-phenylacetic acid was treated with DCC to yield 2-phenylacetic anhydride. The exocyclic amine of the silyl-protected cytidine was then reacted with 2-phenylacetic anhydride to install the phenylacetyl group, followed by tetrabutylammonium fluoride deprotection of the silyl groups to yield the final product, PAC-5EC, in an overall yield of 19%. PACQM-5EU was synthesized in 13 steps starting from commercially available uridine (ESI). Iodination of uridine yielded 5-iodouridine, which was then subjected to a Sonogashira coupling to install the trimethylsilyl alkyne handle at the 5 position. To synthesize the PAC portion, *para*-aminobenzylalcohol was coupled to phenylacetic acid using 2-ethoxy-1-ethoxycarbonyl-1,2-dihyrdoquinoline to generate *N*-(4-(hydroxymethyl)phenyl)-2-phenylacetamide. The benzyl alcohol was then brominated with PBr_3, followed by Finkelstein reaction to install the iodine. The iodo-intermediate was then coupled with the TMS-protected uridine using potassium carbonate, followed by tetrabutylammonium fluoride deprotection of the TMS group to yield the final product, PACQM-5EU, in an overall yield of 11%.

With PAC-5EC and PACQM-5EU in hand, we next focused on testing whether these analogs would display selectivity for metabolic labeling inside cells (Fig. 2). Briefly, HEK293T cells were grown in culture either with or without PGA transfected with plasmid. To the cells, 200uM of PAC-5EC (5EC as positive control) or PACQM-5EU (5EU as positive control) was added.

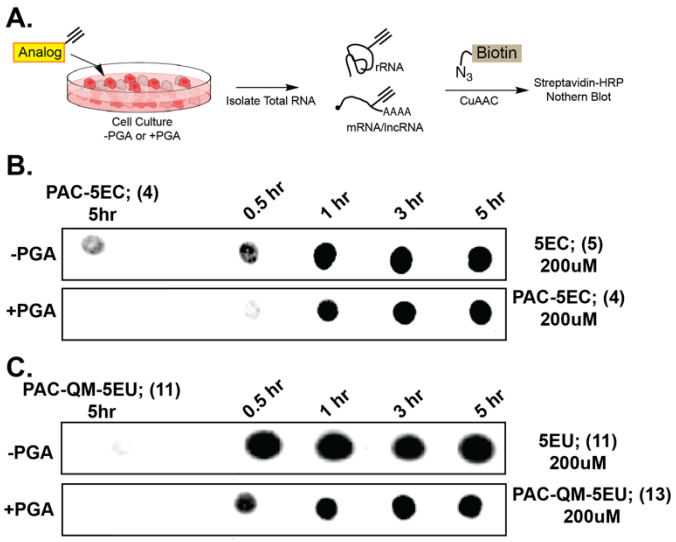


Figure 2. Incorporation of protected analogs into total RNA in HEK cells expressing PGA. (A) Schematic of experiments to test incorporation. (B) Demonstration of PGA-dependent incorporation of PAC-5EC. (C) Demonstration of PGA-dependent incorporation of PAC-QM-5EU.

Total RNA was isolated using TriZol. RNA was then appended with bitoin using Cu mediated Alkyne-Azide cycloaddition (CuAAC) and biotin-azide. Dot blot analysis demonstrated that PAC-5EC and PACQM-5EU can both be robustly deprotected within 1 h in PGA-expressing cells. In contrast, when PAC-5EC and PACQM-5EU were incubated in cells not expressing transfected PGA there was nearly undetectable background incorporation (Fig. 2, B and C). These results demonstrated that PAC-5EC and PACQM-5EU are selectively incorporated into cells expressing PGA.

A major advantage to using analogs that contain alkyne functionality is the ability to image them and their incorporation into cellular RNA; this is in contrast to utilizing 4-thiouridine, which cannot be imaged. We wanted to test whether our novel analogs could be imaged for RNA incorporation in a PGA-dependent manner. As shown in Figure 3, we observed punctate staining within the nucleolus only in PGA transfected cells with the addition of PAC-5EC to the media. A similar labeled pattern was observed with PACQM-5EU (Fig. S1, ESI). This staining is consistent with what has been observed with other analogs and is known to represent rRNA synthesis within the nucleolus. Staining in the nucleolus nicely demonstrates that our analogs are being robustly incorporated into RNA.

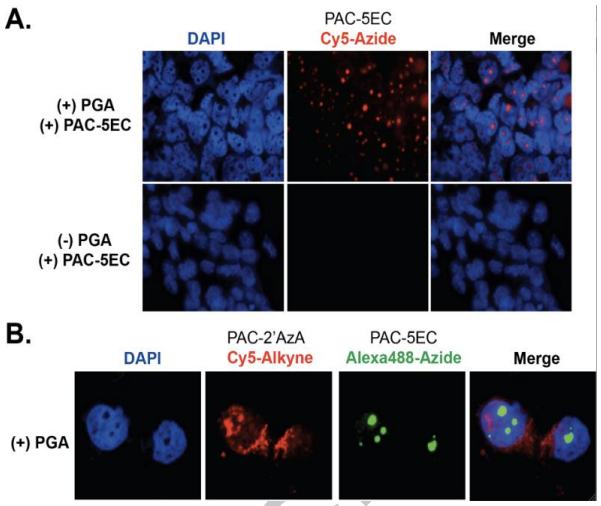


Figure 3. Imaging of sub-RNA population-specific incorporation in HEK cells. (A) Imaging demonstrates that PAC-5EC is incorporated into cellular RNA in a PGA-dependent manner. (B) Imaging demonstrates differential RNA incorporation for PAC-5EC and PAC-2'-AzA mediated by PGA.

To test our analogs further, we aimed to demonstrate that multiple analogs could be differentiated from each other using orthogonal chemistries, and we reasoned our work herein with PAC-5EC and PGA could be used to satisfy this goal. In our previous work we demonstrated that PAC-2'-AzA is incorporated into cellular RNA only in the presence of PGA, but that was mostly incorporated into polyA tails, by polyA polymerase. As such, PAC-2'-AzA staining is almost exclusively cytoplasmic. We compared PAC-2'-AzA and PAC-5EC staining within the same cells and demonstrated their localization patterns are indeed different, further demonstrating that the two analogs have different distributions in RNA. Together, these data further suggest that PAC-5EC (or PACQM-5EU) are likely better analogs for analyzing incorporation into all types of RNAs.

In conclusion, we have reported the design, synthesis, and utilization of protected alkynyl analogs for cell-specific metabolic labeling of RNA. We have demonstrated that PAC-protected analogs are incorporated into cellular RNA only in the presence of PGA. Furthermore, we have utilized imaging to further demonstrate their utility in imaging RNA, and also that the pyrimidine nucleoside analogs herein are incorporated into a different pool of RNA from our previously reported PAC-2'-AzA. Importantly, we envision these analogs will be used widely for understanding changes in RNA expression in a cell-specific manner, and could also find utility for such efforts *in vivo*.

Acknowledgments

We thank members of the Spitale lab for their careful reading of the manuscript. This work is supported by the NIH (1DP2GM11916 and 1R21MH113062 to RCS). RCS is a Pew Biomedical Scholar. Samantha Beasley is supported through an NIH Training Grant (5T32CA009054)

References

1. Handley, A.; Schauer, T.; Ladurner, A. G.; Margulies, C. E. *Mol. Cell* **2015**, *58*, 621.
2. Multherjee, N.; Calviello, L.; Hirsekorn, A.; de Pretis, S.; Pelizzola, M.; Ohler, U. *Nat. Struct. Mol. Biol.* **2017**, *24*, 86.
3. Ribeiro, D. M.; Zanzoni, A.; Cipriano, A.; Delli Ponti, R.; Spinelli, L.; Ballarino, M.; Bozzoni, I.; Tartaglia, G. G.; Brun, C.

Nucleic Acids Res. **2018**, *46*, 917.
4. Schmitt, A. M.; Chang, H. Y. *Cancer Cell* **2016**, *29*, 452.
5. Hrvatin, S.; Deng, F.; O'Donnell, C. W.; Gifford, D. K.; Melton, D. A. *PLoS One* **2014**, *9*.
6. Mack, E.; Neubauer, A.; Brendel, C. *Cytometry A* **2007**, *71a*, 404.
7. Feng, H. J.; Zhang, X. G.; Zhang, C. L. *Nature Communications* **2015**, *6*.
8. Nguyen, K.; Fazio, M.; Kubota, M.; Nainar, S.; Feng, C.; Li, X.; Atwood, S. X.; Bredy, T. W.; Spitale, R. C. *J. Am. Chem. Soc.* **2017**, *139*, 2148.
9. Hida, N.; Aboukilila, M. Y.; Burow, D. A.; Paul, R.; Greenberg, M. M.; Fazio, M.; Beasley, S.; Spitale, R. C.; Cleary, M. D. *Nucleic Acids Res.* **2017**, *45*.
10. Ghosh, A. C.; Shimell, M.; Leof, E. R.; Haley, M. J.; O'Connor, M. B. *Scientific Reports* **2015**, *5*.
11. Forney, L. J.; Wong, D. C. L. *Appl. Environ. Microbiol.* **1989**, *55*, 2556.
12. Handlon, A. L.; Oppenheimer, N. J. *Pharm. Res.* **1988**, *5*, 297.
13. Nainar, S.; Beasley, S.; Fazio, M.; Kubota, M.; Dai, N.; Correa, I. R.; Spitale, R. C. *ChemBioChem* **2016**, *17*, 2149.
14. Sigurgeirsson, B.; Emanuelsson, O.; Lundeberg, J. *PLoS One* **2014**, *9*.

Supplementary Material

Supplementary data associated with this article can be found in the online version.

- A chemical-genetic strategy to enable cell-specific metabolic labeling of RNA.
- Two small molecule-enzyme pairs that enable cell-specific RNA labeling.
- Cell-specific labeling can be imaged with orthogonal “click” chemistry pairs.