Engineered Aminoacyl-tRNA Synthetase for Cell-Selective Analysis of Mammalian Protein Synthesis

Alborz Mahdavi,†,‡ Graham D. Hamblin,† Granton A. Jindal,†,⊥ John D. Bagert,†,‡⊥ Cathy Dong,‡ Michael J. Sweredoski,§ Sonja Hess,§ Erin M. Schuman,∥ and David A. Tirrell*†

†Division of Chemistry and Chemical Engineering, ‡Division of Biology and Biological Engineering, and §Proteome Exploration Laboratory, Beckman Institute, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, United States
∥Max Planck Institute for Brain Research, Frankfurt am Main 60438, Germany

Supporting Information

ABSTRACT: Methods for cell-selective analysis of proteome dynamics will facilitate studies of biological processes in multicellular organisms. Here we describe a mutant murine methionyl-tRNA synthetase (designated L274G MmMetRS) that charges the noncanonical amino acid azidonorleucine (Anl) to elongator tRNAMet in hamster (CHO), monkey (COS7), and human (HeLa) cell lines. Proteins made in cells that express the synthetase can be labeled with Anl, tagged with dyes or affinity reagents, and enriched on affinity resin to facilitate identification by mass spectrometry. The method does not require expression of orthogonal tRNAs or depletion of canonical amino acids. Successful labeling of proteins with Anl in several mammalian cell lines demonstrates the utility of L274G MmMetRS as a tool for cell-selective analysis of mammalian protein synthesis.

Metabolic labeling of cellular proteins with noncanonical amino acids (ncAAs) has emerged as an important tool for the study of proteome dynamics.1−5 Bio-orthogonal noncanonical amino acid tagging (BONCAT)6 uses metabolic incorporation of ncAAs into newly made proteins to enable their enrichment and subsequent identification by mass spectrometry. Newly synthesized proteins can be detected in fixed cells or cell lysates by [3 + 2]-cycloaddition reactions between the azide side-chain of an ncAA and alkynyl probes7−9 (Scheme 1a) or tagged for affinity enrichment and identified by tandem mass spectrometry (Scheme 1b). Recent studies have combined BONCAT with SILAC (stable isotope labeling with amino acids in cell culture)10 to acquire quantitative information about protein synthesis in mammalian cell lines11−13 and to identify secreted proteins in serum-supplemented media.14 These studies illustrate the utility of the BONCAT method in studies of proteome dynamics in mammalian cells.

The temporal resolution provided by metabolic labeling methods can be complemented by cell-selectivity through the controlled expression of mutant aminoacyl-tRNA synthetases. Initial studies of this kind were implemented in bacterial systems,15,16 and used to identify proteins injected into mammalian cells in the course of infection by the bacterial pathogen Yersinia enterocolitica.17 More recently, cell-selective methods have been developed for use in Caenorhabditis elegans18

Scheme 1. Targeted Expression of Mutant Aminoacyl-tRNA Synthetases (aaRS) for Cell-Selective BONCAT in Mammalian Cells

(a) The azide side chain of Anl in proteins (R1) is chemoselectively tagged through azide–alkyne cycloaddition reactions. (b) Controlled expression of mutant aaRS in cells of interest restricts labeling to those cells and enables the detection and identification of proteins in a cell-selective manner. (c) The endogenous mammalian aaRS charges Met but not Anl. A mutant aaRS is used to charge Anl to tRNA^Met.

Received: August 24, 2015
Published: March 18, 2016

DOI: 10.1021/jacs.5b08980
synthetase (which we designated NLL-MetRS) to activate the methionine surrogate 1-azidonorleucine (Anl). Because wild-type MetRS does not activate Anl, BONCAT labeling occurs only in cells that express the NLL-MetRS. Although one might imagine using the NLL-MetRS for cell-selective Anl labeling of mammalian proteins (Scheme 1c), the E. coli enzyme charges only the initiator tRNA in mammalian cells; it does not charge elongator tRNA\textsuperscript{Met}\textsuperscript{E. coli} \textsuperscript{20} and Anl labeling is restricted to N-terminal positions in mammalian proteins. This observation prompted us to develop new MetRS variants to tag mammalian cell proteins with Anl.

We investigated two complementary approaches to cell-selective incorporation of Anl into mammalian proteins. The first approach aimed to identify mutant mammalian aaRSs that would charge Anl to endogenous tRNA\textsuperscript{Met}. The second approach involved coexpression of E. coli NLL-MetRS and E. coli elongator tRNA\textsuperscript{Met} in mammalian cells. Reports by RajBhandary and co-workers showing successful concomitant expression of an active E. coli glutaminyl-tRNA synthetase and a functional amber suppressor tRNA that is not aminocylated in vivo by any mammalian aaRS suggested to us that the mammalian translation machinery might co-opt an E. coli elongator tRNA\textsuperscript{Met} that is charged by NLL-MetRS.\textsuperscript{21,22} To identify mutant mammalian aaRSs that charge Anl to endogenous tRNA\textsuperscript{Met}, we made a small mutant aaRS library. We had previously observed that residues 13, 260, and 301 in the Met-binding pocket of E. coli MetRS\textsuperscript{23} can be mutated to enable activation of Anl.\textsuperscript{24} Because there are no reported structures of mammalian MetRSs, we used sequence alignment to determine the corresponding positions in the E. coli and eukaryotic MetRS enzymes (Figure 1a). Conservation of the residues that align with E. coli residues 13, 260, and 301 suggested that site-directed mutagenesis at the corresponding positions in the mouse MetRS (MmMetRS) would yield variants that charge Anl. Therefore, we constructed a small mutant library in which residues L274, Y527 and H562 of MmMetRS (‘LYH’) were mutated to SSL, CIL, PLL, and NLL (Figure S1–S6). We also made an L274G mutant, reflecting our earlier observation that the corresponding L13G variant of the E. coli MetRS activates Anl.\textsuperscript{25} Finally, we overexpressed wild-type MmMetRS to determine whether overexpression of the wild-type enzyme causes observable levels of incorporation of Anl into cellular proteins.

MetRS variants were expressed under control of a CMV promoter (Figure 1b). To express the E. coli elongator tRNA\textsuperscript{Met}, we adopted the approach of the RajBhandary group and introduced 5′- and 3′-flanking sequences derived from human initiator tRNA to enable proper RNA editing \textsuperscript{21} (Figure 1c). Notably, efficient transcription by mammalian RNA polymerase III requires a purine at position 9 of the tRNA coding sequence. In this case, adenine was used. This adenine lies in the A box of the internal promoter of eukaryotic tRNA genes.\textsuperscript{26} In mammalian cells, the 3′-CCA tail is added by a nucleotidyl-transferase;\textsuperscript{27} therefore, we tested tRNA sequences with and without CCA tails (Figure S7–S9). We measured the activities of these constructs with respect to Anl incorporation by transiently transfecting several widely used mammalian cell lines, including Chinese hamster ovary (CHO) cells, monkey kidney cells (COS7), and an immortalized human cervical cancer cell line (HeLa). Cells were incubated in serum-containing medium supplemented with 1.5 mM Anl for 10 h and then lysed. Total cellular proteins were treated with alkylene-tetramethylrhodamine (TAMRA) (Figure S10) for labeling via copper-catalyzed azide–alkyne cycloaddition (reaction 1, Scheme 1a).\textsuperscript{26,27} After SDS-PAGE, TAMRA-labeled proteins were detected by in-gel fluorescence imaging (Figure 1d). As expected, Anl was not charged to any appreciable extent by the wild-type translational machinery. Anl incorporation was detected in cells carrying the E. coli NLL-MetRS, but we did not observe a significant increase in Anl incorporation as a function of tRNA expression. The variability of incorporation across different cell lines suggests that the expression levels of the tRNA and the synthetase may need to be optimized for each cell type, thereby limiting the general applicability of this approach. Nevertheless, variants of this strategy have already proven useful for cell-selective protein labeling: for example, Chin and co-workers have described successful proteome-wide labeling through the action of a genetically encoded pyrrolysyl-tRNA synthetase/tRNA pair in Drosophila.\textsuperscript{28}

In contrast to our results with the E. coli NLL-MetRS/tRNA\textsuperscript{Met} pair, expression of the L274G mouse MetRS, henceforth called L274GMMmMetRS, resulted in robust incorporation of Anl across all three cell lines. Thus, a single mutation is sufficient to enable Anl incorporation into mammalian cell proteins. Surprisingly, we found no evidence of significant Anl labeling with any of the triple mutants we examined (Figure 1d).

An important determinant of the utility of a mutant synthetase in ncAA tagging is the selectivity of the synthetase with respect to activation of the ncAA and its canonical counterpart. To provide a practical measure of the selectivity of L274GMMmMetRS, we examined protein labeling in CHO cells as a function of Anl concentration in serum-containing media. After Anl incorporation, proteins from cell lysates were labeled with alkylene-TAMRA (Figure 2a). We also performed Met competition assays by incubating CHO cells with 1.5 mM Anl in media supplemented with Met at different concentrations (Figure
The ability to restrict L274GMmMetRS expression to cells of interest should enable the cell-selective labeling of proteins for targeted analysis of protein synthesis. To validate this concept, we developed a reporter construct expressing both L274GMmMetRS and mCherry under control of a CMV promoter (pMaRSC). A T2A linker was used to express two disjointed proteins in the same open reading frame; we reasoned that fusing L274GMmMetRS and mCherry under control of a CMV promoter (pMaRSC) (Figures S13 and S14). When transfected into CHO cells, both constructs enabled labeling of cellular proteins at comparable levels, indicating that L274GMmMetRS was functional in both contexts (Figures S14 and S15). We used the mCherry signal to track expression of L274GMmMetRS in pMaRSC-transfected cells. CHO cells transfected with pMaRSC and incubated with Anl (1.5 mM, 6 h) were fixed, permeabilized, and treated with alkyne-TAMRA. We observed TAMRA fluorescence exclusively in the mCherry-expressing cells, consistent with cell-selective incorporation of Anl into cellular proteins (Figure 3b). These cells were transiently transfected with a second (activating) construct bearing a strong Ef1α promoter that drives expression of an eGFP-Cre fusion protein.

We used the eGFP signal to track expression of Cre and followed the mCherry signal to monitor recombination of the loxP sites. As expected, in the subset of cells that expressed Cre (as determined by the GFP signal), we observed both mCherry expression indicative of loxP recombination and TAMRA fluorescence from Anl-tagged cellular proteins, confirming that cell-selective proteomic labeling can be achieved through regulated transgene expression. These results suggest that targeted insertion of loxP-L274GMmMetRS cassettes into the ROSA26 locus of mouse embryonic stem (ES) cells should facilitate engineering of Cre-Lox mice in which tissue-selective proteomic studies can be performed.

To confirm the utility of L274GMmMetRS as a tool for proteomic analysis in mammalian systems, we used azidobenzocyclooctyne-functionalized beads to capture (reaction 2, Scheme 1a) and enrich Anl-labeled proteins from lysates of CHO cells that constitutively express the synthetase under CMV promoter.
control. After on-resin tryptic digestion of purified proteins, we analyzed the resulting tryptic peptides by tandem mass spectrometry (Table S1). A total of 996 proteins were identified in Anl-treated samples in short (120 min) analyses on a linear trap quadrupole (LTQ) Orbitrap mass spectrometer. Of these proteins, 786 were not found in Met-treated controls. Enrichment was confirmed by the acquisition of 19-fold more spectral counts in analyses of Anl-treated samples as compared to controls. We detected proteins associated with a wide variety of cellular components, including nuclear, cytoskeletal, mitochondrial, and membrane proteins (Figure 3d, S16). MS analysis of the proteome of HeLa cells expressing L274G identified 161 sites of Anl incorporation, 150 (93%) at internal Met positions (Figure S18, Table S2). This ratio mirrors the distribution of Met sites across the proteome.

In summary, we find that the L274GMMnMetRS functions in a variety of mammalian cell lines and charges Anl at a rate sufficient for cell-selective proteome analysis. Surprisingly, the single-site mutant is superior to all of the triple mutants we evaluated, including the murine orthologue of the E. coli NLL-MetRS used previously in cell-selective studies of protein synthesis. A variant of L274GMMnMetRS has recently been used for cell-selective labeling of proteins in Drosophila, and we suggest that targeted expression of this enzyme will provide a broadly useful approach to proteome and secretome analysis in subpopulations of mammalian and other eukaryotic cells.

**ASSOCIATED CONTENT**

1. **Supporting Information**
   - The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b08980.
   - Experimental details, plasmid maps and determination of selectivity of L274GMMnMetRS (PDF)
   - List of proteins identified by shotgun proteomics (XLSX)
   - List of sites of incorporation of Anl into cellular proteins (XLSX)

**AUTHOR INFORMATION**

* Corresponding Author
  *tirrell@caltech.edu

**Present Address**

†Department of Chemistry, Princeton University, Princeton, New Jersey 08540, United States.

**Notes**

The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**

We are grateful for financial support by National Institutes of Health grant NIH R01 GM062523 and the Programmable Molecular Technology Initiative of the Gordon and Betty Moore Foundation, the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from U.S. Army Research Office. M.S. and S.H. were supported by the Gordon and Betty Moore Foundation through grant GBMF775 and NIH grant 1S10RR029594-01A1. A.M. was supported by a scholarship from the National Science and Engineering Research Council of Canada and by a postgraduate scholarship from the Donna and Benjamin M. Rosen Center for Bioengineering at Caltech. G.D.H. was supported by a postdoctoral fellowship from the National Science and Engineering Research Council of Canada.

**REFERENCES**

(7) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. Angew. Chem., Int. Ed. 2002, 41, 2596.