## Title

Multiple Click-Selective tRNA Synthetases Expand Mammalian Cell-Specific Proteomics.



## Journal

Journal of the American Chemical Society, 140(23)

## Authors

Yang, Andrew
du Bois, Haley
Olsson, Niclas
et al.

## Publication Date

2018-06-13

## DOI

10.1021/jacs.8b03074

Peer reviewed

HHS Public Access

Author manuscript

*J Am Chem Soc*. Author manuscript; available in PMC 2019 June 28.

Published in final edited form as:

*J Am Chem Soc*. 2018 June 13; 140(23): 7046–7051. doi:10.1021/jacs.8b03074.

# Multiple Click-Selective tRNA Synthetases Expand Mammalian Cell-Specific Proteomics

**Andrew C. Yang**[†,§,‡], **Haley du Bois**[‡], **Niclas Olsson**[⊥], **David Gate**[‡], **Benoit Lehallier**[‡], **Daniela Berdnik**[‡,¶], **Kyle D. Brewer**[§,‡], **Carolyn R. Bertozzi**[§,ǁ,‖], **Joshua E. Elias**[⊥], and **Tony Wyss-Coray**[*,‡,¶,§]

[†]Department of Bioengineering, Stanford University, Stanford, California 94305, United States

[§]Chemistry, Engineering, and Medicine for Human Health (ChEM-H), Stanford University, Stanford, California 94305, United States

[‡]Department of Neurology and Neurological Sciences, Stanford University, Stanford, California 94305, United States

[⊥]Department of Chemical and Systems Biology, Stanford University, Stanford, California 94305, United States

[ǁ]Department of Chemistry, Stanford University, Stanford, California 94305, United States

[‖] Howard Hughes Medical Institute, Stanford University, Stanford, California 94305, United States

[¶]Center for Tissue Regeneration, Repair and Restoration, V.A. Palo Alto Healthcare System, Palo Alto, California 94304, United States

## Abstract

Bioorthogonal tools enable cell-type-specific proteomics, a prerequisite to understanding biological processes in multicellular organisms. Here we report two engineered aminoacyl-tRNA synthetases for mammalian bioorthogonal labeling: a tyrosyl (*Sc*Tyr$_{Y43G}$) and a phenylalanyl (*Mm*Phe$_{T413G}$) tRNA synthetase that incorporate azide-bearing noncanonical amino acids specifically into the nascent proteomes of host cells. Azide-labeled proteins are chemoselectively tagged via azide−alkyne cycloadditions with fluorophores for imaging or affinity resins for mass spectrometric characterization. Both mutant synthetases label human, hamster, and mouse cell line proteins and selectively activate their azido-bearing amino acids over 10-fold above the canonical. *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ label overlapping but distinct proteomes in human cell lines, with broader proteome coverage upon their coexpression. In mice, *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ label the melanoma tumor proteome and plasma secretome. This work furnishes new tools for mammalian residue-specific bioorthogonal chemistry, and enables more robust and comprehensive cell-type-specific proteomics in live mammals.

[*]**Corresponding Author** twc@stanford.edu.

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b03074.

Experimental methods and supplemental figures (PDF) List of proteins identified by shotgun proteomics (XLSX)

Notes

The authors declare no competing financial interest.

Organisms and tissues are composed of heterogeneous cell types. Robust and comprehensive cell-type-specific proteomics are foundational to understanding the biological processes underlying health and disease. For example, cell-secreted signaling proteins in blood have been shown to not only correlate with but also modulate organismal and brain aging.[1–3] Current efforts to characterize cell-type-specific proteomes and secretomes rely on various cell isolation techniques before acute analysis or primary cell culture; however, these techniques likely perturb the in *vivo* proteome, completely lose the secretome, and lack temporal resolution.[4–6]

Strategies to label the proteins of target cells with bioorthogonal moieties could enable subsequent enrichment and cell-type-specific proteomics.[7–12] Labeling occurs via the metabolic incorporation of noncanonical amino acids (ncAAs) containing azide, alkyne, or other bioorthogonal side chains.[13–15] Mutant aminoacyl-tRNA synthetases (aaRS) recognize ncAAs that are ignored by endogenous aaRSs. By expressing aaRSs under the control of cell-type-specific promotors or inducible genetic tools, one can achieve cell-type- and temporally restricted metabolic protein labeling in *vivo*.

The recent development of the L274G mouse methionyltRNA synthetase ($MmMet_{L274G}$), and its rapid adoption in multicellular organisms from Drosophila to mice, has enabled the characterization of specific neuronal proteomes.[10,16] However, $MmMet_{L274G}$ is currently the only mutant aaRS available for mammalian cell-type-specific proteomics and is limited to charging its single cognate $tRNA^{Met}$ (ATG) with the methionine surrogate azidonorleucine. As a result, reliance on $MmMet_{L274G}$ potentially constricts proteome coverage and skews subsequent analysis.[17] Depending on the ncAA and targeted codon, labeling could perturb protein stability, folding, and trafficking; be performed at discordant rates; and be masked by steric effects or post-translational processing, such as N-terminal cleavage.[18] Thus, mutant aaRSs likely label many proteins differentially, each preferring a subset of the whole proteome. To enable more robust and broader mammalian cell-type-specific proteomics, we identified and characterized two mutant aaRSs: a tyrosyl ($ScTyr_{Y43G}$) that charges the ncAA 3-azido-L-tyrosine (AzY) onto $tRNA^{Tyr}$ and a phenylalanyl ($MmPhe_{T413G}$) that charges p-azido-L-phenylalanine (AzF) onto $tRNA^{Phe}$ for incorporation into nascent proteins of host cells (Scheme 1a). Proteins can be labeled via either of the $tRNA^{Tyr}$ (TAT, TAC) and $tRNA^{Phe}$ (TTT, TTC) cognate codons. Incorporation does not require depletion of canonical amino acids or the strong coexpression of exogenous tRNAs, as exemplified by complementary approaches.[12] Proteins labeled with AzY and AzF are chemoselectively tagged via azide−alkyne cycloadditions with fluorophores for imaging and flow cytometry or affinity resins for mass spectrometric identification and quantification (Scheme 1b).

To begin, we sought to convert microbial or metazoan aaRSs to be more broadly useful in mammals. We hypothesized that aaRSs and ncAAs developed in earlier genetic code expansion efforts could be adopted without their paired tRNAs: if species-specific tRNA-binding/aminoacylation determinants were properly deduced, the aaRS could be reoriented to recognize endogenous tRNAs for residue-specific proteome labeling.[19–21] For example, we noticed that the *M. jannaschii* Tyr aaRS (TyrRS) used for site-specific ncAA incorporation in *E. coli* possesses eukaryotic tRNA-binding determinants (recognizes tRNA acceptor stem C1-G72).[22] This informed the testing of an engineered *Mj* TyrRS without its

tRNA pair for residue-specific incorporation of AzF.[20] We further spliced a 39 amino acid aminoacylation determinant from human TyrRS into an *E. coli* TyrRS reported to incorporate AzF ("CP1 switch").[23] Given the conserved archaeal/eukaryotic aminoacylation determinants of *Mj* TyrRS, we reasoned that aminoacylation would be preserved across eukaryotic TyrRS and adopted a yeast TyrRS (*Sc*Tyr$_{Y43G}$) reported to incorporate AzY (Figures 1a, S1).[24] Finally, the efficient labeling of C. *elegans* proteomes with an engineered *Ce*PheRS prompted us to develop human and mouse variants: *Hs*Phe$_{T413G}$ and *Mm*Phe$_{T413G}$ (Figures 1b, S1).[9]

*Sc*Tyr$_{Y43G}$, *Hs*Phe$_{T413G}$, and *Mm*Phe$_{T413G}$ exhibited strong labeling in human HEK293T, hamster CHO, and mouse B16-F10 cell lines after transient transfection and incubation with high concentrations (2 mM) of AzY and AzF (Figures 1c, S2). Cells were lysed, pretreated with iodoacetamide (IAM) to block background thiol-yne additions,[25] and treated with dibenzocyclooctyne (DIBO)-Alexa Fluor 647 dye for copper-free, strain-promoted azide −alkyne cycloaddition.

The selectivity of *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ for AzY and AzF respectively over their endogenous counterparts Tyr and Phe is a critical determinant of their utility, especially *in vivo*. Minimizing the amount of exogenous ncAA required for proteome labeling likely reduces toxicities that may perturb the proteome. Adopting previously derived equations for ncAA activation,[11] we quantified the extent of protein labeling in HEK293T cells as a function of AzY and AzF concentrations in serum-containing media (0 to 125 $\mu$M, 24 h); and performed the corollary Tyr and Phe competition assays with 15 $\mu$M AzY and AzF (Figures 1d,e, S3). Taking Alexa Fluor 647 (AF647) in-gel fluorescence as a measure of ncAA proteome incorporation, the line of best fit yields the rate of ncAA activation and the specificity constants ($k_{cat}/K_M$) for the ncAA and canonical amino acid. Interestingly, and in contrast with *Mm*Met$_{L274G}$, *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ exhibit high selectivity for their ncAA: the former activates AzY nearly 250-fold faster than Tyr at equimolar concentrations of the two amino acids; and the latter activates AzF over 11-fold faster than Phe, consistent with prior CePheRST413G *in vitro* measurements.[9]

We next determined whether proteome labeling by *Sc*Ty$_{Y43G}$ and *Mm*Phe$_{T413G}$ was compatible with in situ fluorescence imaging, a critical modality for studying complex biology. Proteome labeling was visualized by tagging azide-bearing proteins with alkyne AF647 in HEK293T cells transfected with *Sc*Tyr$_{Y43G}$ or *Mm*Phe$_{T413G}$ and exposed to 125 $\mu$M AzY or AzF (Figures 2a, S4). Mutant aaRSs were transfected in vectors coexpressing GFP: comparison of GFP$^+$ and AF647$^+$ areas indicated ubiquitous labeling across cells. To assess fluorescent proteome labeling across a cell population via flow cytometry, transfected and ncAA-exposed HEK293T cells were live/dead discriminated, fixed, IAM-treated, and tagged with DIBOAF647. An AF647$^+$ subpopulation emerged even in the GFP$^{lo}$ regime, suggesting that even low *Sc*Tyr$_{Y43G}$ or *Mm*Phe$_{T413G}$ expression is sufficient for maximal proteome labeling (Figures 2c, S5).

Hypothesizing that each mutant aaRS preferentially labels a subset of the full cell proteome, we transfected HEK293T cells with equal total amounts of plasmid containing *Mm*Met$_{L274G}$, *Mm*Phe$_{T413G}$, *Sc*Tyr$_{Y43G}$, or all three aaRSs; and exposed cells to 125 μM of

their corresponding ncAAs or at least as much of endogenous amino acids. This amino acid concentration was informed by aaRS selectivity measurements (Figure 1d,e) and typical, nontoxic concentrations of exogenous agents in live mice. Labeled proteomes from cell lysates were enriched on azadibenzocyclooctyne (DBCO) beads, stringently washed, Trypsin/LysC digested, and conjugated to tandem mass tags (TMT) for multiplexed mass spectrometric characterization. We sought to (1) identify proteins uniquely enriched by each mutant aaRS; (2) quantify labeling effiencies for each aaRS across commonly identified proteins; and (3) determine whether coexpression of aaRSs facilitates broader and more confident proteomics compared to single aaRS expression.

We found each mutant aaRS proteome sufficiently distinct to spatially segregate via principal component analysis (Figure 3a). After filtering for proteins significantly labeled over nonspecific background ($P$-value < 0.05 and $\log_2$(FC) > 1, Figure S6), we observed that each mutant aaRS labeled a distinct set of proteins: *Mm*Met$_{L274G}$ alone identified 46 proteins overlooked by the other two mutant aaRSs, *Mm*Phe$_{T413G}$ uniquely identified 54, and *Sc*Tyr$_{Y43G}$ uniquely identified 138 (Figure 3b). On the other hand, most identified proteins were common across pairs of, if not all three, aaRSs (66%). Within the 463 proteins identified by at least two aaRSs, proteomic differences could arise from each mutant aaRS's labeling biases. Indeed, TMT-enabled quantification of protein abundances revealed that each mutant aaRS enriched different proteins with different efficiencies (Figures 3c, S8). Labeling efficiencies across proteins are consistent with prior selectivity measurements (Figure 1d,e). *Mm*Phe$_{T413G}$ and *Sc*Tyr$_{Y43G}$ label proteins through both their cognate codons (TTT, TTC; and TAT, TAC), with *Mm*Phe$_{T413G}$ exhibiting a preference for TTT (Figure S9). We next wondered whether aaRS coexpression could yield a more comprehensive and confidently identified cell proteome. The triple coexpression of *Mm*Met$_{L274G}$, *Mm*Phe$_{T413G}$, and *Sc*Tyr$_{Y43G}$ expanded proteome coverage and identified proteins more confidently (assessed by lower P-values) compared to single aaRS expression (Figures 3d, S7). Together, these data establish that the introduction of *Mm*Phe$_{T413G}$ and *Sc*Tyr$_{Y43G}$ can improve mammalian cell-type-specific proteomics over *Mm*Met$_{L274G}$ alone.

To determine whether *Mm*Phe$_{T413G}$ and *Sc*Tyr$_{Y43G}$ could label proteomes *in vivo*, we stably integrated each mutant aaRS into B16-F10 melanoma cells before subcutaneous implantation in 12-week-old C57BL/6 mice. Sixteen days after implantation, we administered saturating amounts of the corresponding ncAA intraperitoneally (1 mmol/kg) and intratumorally (~5 mM) daily for 3 days. Confocal fluorescence imaging of tumor sections revealed AzF or AzY proteome incorporation, assessed by chemoselective conjugation to alkyne AF647 (Figures 4a, S10). The restriction of AF647$^+$ signal to implanted GFP$^+$ melanoma cells amidst a wild-type C57BL/6 background suggests *Mm*Phe$_{T413G}$ and *Sc*Tyr$_{Y43G}$ are suitable for additional in *vivo* tissue- and cell-type-specific proteomic applications. Ingel fluorescence of tumor lysates revealed labeling across the proteome (Figures 4b, S11). The stronger *Mm*Phe$_{T413G}$ signal is consistent with Figure 1c, where saturating amounts of ncAA delivered via intratumoral injection neutralized the selectivity advantages of *Sc*Tyr$_{Y43G}$. However, high ncAA selectivity could be advantageous in most in *vivo* applications where target cells are not directly accessible by needle.

We next sought to identify the labeled melanoma proteome. Tumor cell lysates were collected, 3 mg incubated with DBCO beads, and processed for label-free mass spectrometric characterization. As in HEK293T cells, *Mm*Phe$_{T413G}$ and *Sc*Tyr$_{Y43G}$ labeled distinct B16-F10 tumor proteins *in vivo* (Figures 4c, S12). Though *Sc*Tyr$_{Y43G}$ yielded fewer proteins in total, it identified 108 proteins *Mm*Phe$_{T413G}$ did not, further evidence of mutant aaRSs labeling preferred proteome subsets. For both mutant aaRSs, proteins were detected across a wide variety of cellular components, spanning the nucleus, mitochondria, and cell surface (Figures 4d, S13). Labeled proteomes were significantly enriched for canonical pathways implicated in tumor progression, metabolism, and apoptosis (Figures 4e, S14).[26,27] Despite the rapid adoption of *Mm*Met$_{L274G}$ *in vivo*, to our knowledge, cell-type-specific secretome labeling has not yet been demonstrated. We click- enriched 3 mg of plasma from tumor-bearing mice for mass spectrometry, finding that *Mm*Phe$_{T413G}$ and *Sc*Tyr$_{Y43G}$ labeled distinct subsets of the tumor plasma secretome. Surprisingly, secretome labeling was comparable (Figures 4f, S12). This may arise from several factors, such as tyrosine's greater solvent accessibility; but does not correlate with the relative abundance of tyrosine to phenylalanine in the mouse secretome (UniProt:3.6% Phe vs 2.9% Tyr in frequency), consistent with prior work.[17] Several labeled secretome proteins, including the 14-3-3 family of proteins and proteasome subunits, have been validated in human cancer xenograft studies, but many are novel.[28,29] Like the labeled tumor proteome, the secretome was enriched for pathways implicated in cancer, including glycolysis, ubiquitination, and pentose phosphate pathway signaling (Figures 4g, S14).[30–32] Unlike prior studies, cell-type-specific bioorthogonal labeling via mutant aaRSs can distinguish tumor-and host-secreted plasma proteins across a wide variety of immunocompetent mouse models. We report a list of B16-F10 melanoma secreted plasma proteins, to be expanded upon in dedicated, follow-up biological studies (Table S1).

In summary, we find that *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ label proteins across mammalian cell lines and in live mice. These mutant aaRSs demonstrate high selectivity for activating AzY and AzF over endogenous Tyr and Phe, respectively. *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ label overlapping but distinct proteomes in HEK293T cells, and their coexpression yields a fuller proteome. *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ enable a first application of bioorthogonal labeling to a tumor model in mice and to the identification of plasma factors secreted from specific cell types.

Interest in adopting bioorthogonal labeling tools for cell- and tissue-specific proteomics in mammals is growing.[16,33,34] We suggest that targeted coexpression of *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ alongside the existing *Mm*Met$_{L274G}$ via 2A or IRES elements may enhance cell-specific proteome coverage and confidence, and capture a hitherto undetected richness in proteome spatial and temporal dynamics. Multiple engineered synthetases enable the multiplexed incorporation of diverse chemistries into a given mammalian proteome or the simultaneous labeling of different cell types in mice. This work also informs the engineering of additional mutant aaRSs for mammalian proteomics, as the three mutant aaRSs were consistently developed by expanding their amino acid binding pockets via single substitutions to glycine. And as aryl azides, proteome-incorporated AzY and AzF could be used as photo-cross-linkers to investigate protein interactions,[20,24] with other compatible ncAAs introducing additional chemistries. In general, *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ open

new opportunities for *in vivo* cataloguing, tracking, and modulation of proteomes from specific mammalian cells.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGMENTS

We thank V. Mathur, M. E. Zardeneta, R. T. Vest, D. Lee, H. Zhang, J. V. Pluvinage, C. W. Chiang, and P. S. Huang for helpful discussions. We are grateful for financial support by National Institutes of Health grants DP1 AG053015 (T.W.-C.) and R37 GM058867 (C.R.B.), the NOMIS Foundation (T.W.-C.), and the Stanford Neurosciences Institute (T.W.-C., C.R.B., J.E.E.). A.C.Y. was supported by the National Science Foundation Graduate Research Fellowship. N.O. was supported by the Knut and Alice Wallenberg Foundation Postdoctoral Fellowship.

## REFERENCES

(1). Conboy IM; Conboy MJ; Wagers AJ; Girma ER; Weissman IL; Rando TA Nature 2005, 433 (7027), 760–764. [PubMed: 15716955]

(2). Villeda SA; Plambeck KE; Middeldorp J; Castellano JM; Mosher KI; Luo J; Smith LK; Bieri G; Lin K; Berdnik D; Wabl R; Udeochu J; Wheatley EG; Zou B; Simmons DA; Xie XS; Longo FM; Wyss-Coray T Nat. Med 2014, 20 (6), 659–663. [PubMed: 24793238]

(3). Wyss-Coray T Nature 2016, 539, 180–186. [PubMed: 27830812]

(4). Nedergaard M; Verkhratsky A Glia 2012, 60 (7), 1013–1023. [PubMed: 22228580]

(5). Sharma K; Schmitt S; Bergner CG; Tyanova S; Kannaiyan N; Manrique-Hoyos N; Kongi K; Cantuti L; Hanisch UK; Philips MA; Rossner MJ; Mann M; Simons M Nat. Neurosci 2015, 18 (12), 1819–1831. [PubMed: 26523646]

(6). Stone SE; Glenn WS; Hamblin GD; Tirrell DA Curr. Opin. Chem. Biol 2017, 36, 50–57. [PubMed: 28088696]

(7). Ngo JT; Champion JA; Mahdavi A; Tanrikulu IC; Beatty KE; Connor RE; Yoo TH; Dieterich DC; Schuman EM; Tirrell DA Nat. Chem. Biol 2009, 5 (10), 715–717. [PubMed: 19668194]

(8). Ngo JT; Schuman EM; Tirrell DA Proc. Natl. Acad. Sci. U.S. A 2013, 110 (13), 4992–4997. [PubMed: 23479642]

(9). Yuet KP; Doma MK; Ngo JT; Sweredoski MJ; Graham RLJ; Moradian A; Hess S; Schuman EM; Sternberg PW; Tirrell DA Proc. Natl. Acad. Sci. U. S. A 2015, 112 (9), 2705–2710. [PubMed: 25691744]

(10). Erdmann I; Marter K; Kobler O; Niehues S; Abele J; Müller A; Bussmann J; Storkebaum E; Ziv T; Thomas U; Dieterich DC Nat. Commun 2015, 6, 7521. [PubMed: 26138272]

(11). Mahdavi A; Hamblin GD; Jindal GA; Bagert JD; Dong C; Sweredoski MJ; Hess S; Schuman EM; Tirrell DA J. Am. Chem. Soc 2016, 138 (13), 4278–4281. [PubMed: 26991063]

(12). Elliott TS; Townsley FM; Bianco A; Ernst RJ; Sachdeva A; Elsasser SJ; Davis L; Lang K; Pisa R; Greiss S; Lilley KS; Chin JW Nat. Biotechnol 2014, 32 (5), 465–472. [PubMed: 24727715]

(13). Sletten EM; Bertozzi CR Angew. Chem., Int. Ed 2009, 48, 6974–6998.

(14). Dieterich DC; Link AJ; Graumann J; Tirrell DA; Schuman EM Proc. Natl. Acad. Sci. U. S. A 2006, 103 (25), 9482–9487. [PubMed: 16769897]

(15). Kiick KL; Saxon E; Tirrell DA; Bertozzi CR Proc. Natl. Acad. Sci. U. S. A 2002, 99 (1), 19–24. [PubMed: 11752401]

(16). Alvarez-Castelao B; Schanzenbacher CT; Hanus C; Glock C; tom Dieck S; Dörrbaum AR; Bartnik I; Nassim-Assir B; Ciirdaeva E; Mueller A; Dieterich DC; Tirrell DA; Langer JD; Schuman EM Nat. Biotechnol 2017, 35 (12), 1196. [PubMed: 29106408]

(17). Elliott TS; Bianco A; Townsley FM; Fried SD; Chin JW Cell Chem. Biol 2016, 23 (7), 805–815. [PubMed: 27447048]

(18). Howden AJM; Geoghegan V; Katsch K; Efstathiou G; Bhushan B; Boutureira O; Thomas B; Trudgian DC; Kessler BM; Dieterich DC; Davis BG; Acuto O Nat. Methods 2013, 10(4), 343–346. [PubMed: 23474466]

(19). Wang L; Brock A; Herberich B; Schultz PG Science (Washington, DC, U. S.) 2001, 292 (5516), 498–500.

(20). Chin JW; Santoro SW; Martin AB; King DS; Wang L; Schultz PG J. Am. Chem. Soc 2002, 124 (31), 9026–9027. [PubMed: 12148987]

(21). Deiters A; Cropp TA; Mukherji M; Chin JW; Anderson JC; Schultz PG J. Am. Chem. Soc 2003, 125 (39), 11782–11783. [PubMed: 14505376]

(22). Tsunoda M; Kusakabe Y; Tanaka N; Ohno S; Nakamura M; Senda T; Moriguchi T; Asai N; Sekine M; Yokogawa T; Nishikawa K; Nakamura KT Nucleic Acids Res. 2007, 35 (13), 4289–4300. [PubMed: 17576676]

(23). Wakasugi K; Quinn CL; Tao N; Schimmel P EMBO J 1998, 17 (1), 297–305. [PubMed: 9427763]

(24). Yokogawa T; Ohno S; Nishikawa K Methods Mol. Biol 2010, 607, 227–242. [PubMed: 20204861]

(25). van Geel R; Pruijn GJM; van Delft FL; Boelens WC Bioconjugate Chem. 2012, 23 (3), 392–398.

(26). Gogvadze V; Orrenius S; Zhivotovsky B Trends Cell Biol. 2008, 18, 165–173. [PubMed: 18296052]

(27). Chalkiadaki A; Guarente L Nat. Rev. Cancer 2015, 15, 608–624. [PubMed: 26383140]

(28). Schiarea S; Solinas G; Allavena P; Scigliuolo GM; Bagnati R; Fanelli R; Chiabrando CJ Proteome Res. 2010, 9 (9), 4376–4392.

(29). Jansen FH; Krijgsveld J; van Rijswijk A; van den Bemd G-J; van den Berg MS; van Weerden WM; Willemsen R; Dekker LJ; Luider TM; Jenster G Mol. Cell. Proteomics 2009, 8 (6), 1192–1205. [PubMed: 19204029]

(30). Gatenby RA; Gillies RJ Nat. Rev. Cancer 2004, 4, 891–899. [PubMed: 15516961]

(31). Mani A; Gelmann EP J. Clin. Oncol 2005, 23 (21), 4776–4789. [PubMed: 16034054]

(32). Yi H; Zheng X; Song J; Shen R; Su Y; Lin D Int. J. Clin. Exp. Pathol 2015, 8 (12), 15719–15728. [PubMed: 26884841]

(33). Liu Y; Conboy MJ; Mehdipour M; Liu Y; Tran TP; Blotnick A; Rajan P; Santos TC; Conboy IM Nat. Commun 2017, 8 (1), DOI: DOI: 10.1038/s41467-017-00698-y.

(34). Krogager TP; Ernst RJ; Elliott TS; Calo L; Beranek, V; Ciabatti E; Spillantini MG; Tripodi M; Hastings MH; Chin JW Nat. Biotechnol 2017, 36 (2), 156. [PubMed: 29251727]

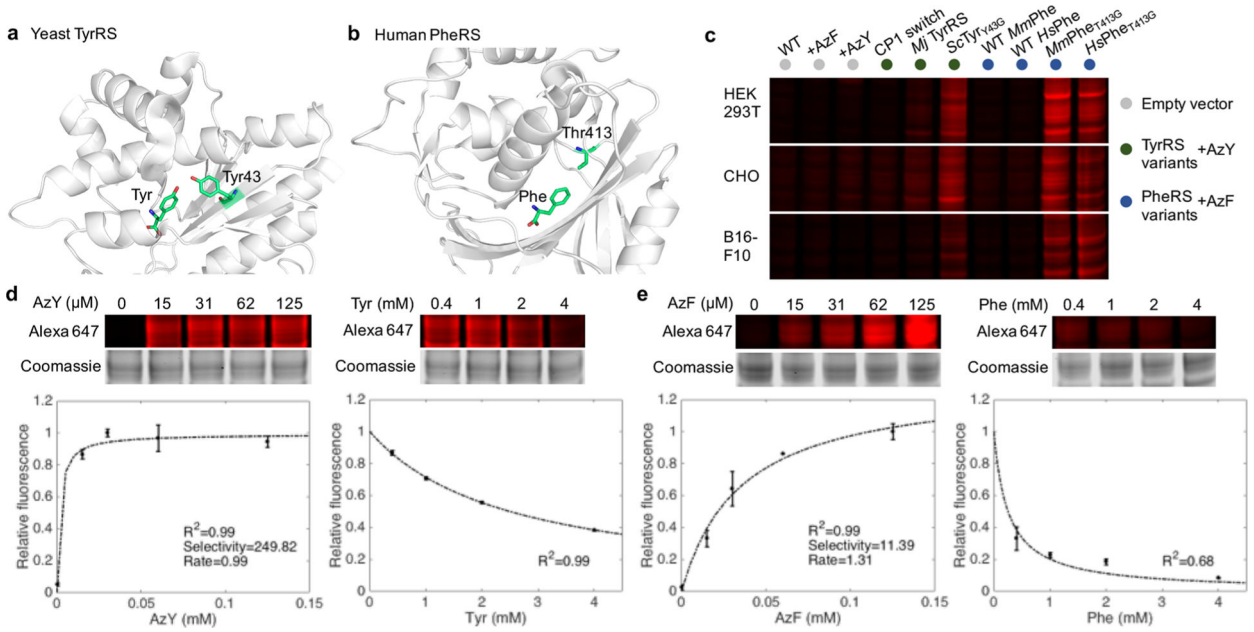



**Figure 1.**

Identification and characterization of TyrRS and PheRS variants for bioorthogonal labeling of mammalian proteomes with AzY and AzF.(a) A single substitution (Y43G) in the amino acid binding site of yeast tyrosyl-tRNA synthetase (*Sc*$Tyr_{Y43G}$) enables charging of the azido tyrosine analog AzY onto $tRNA^{Tyr}$ before incorporation into nascent proteins of host cells. (b) A single substitution (T413G) in the amino binding site of human or mouse phenylalanyl-tRNA synthetase (*Mm*$Phe_{T413G}$) enables charging of the azido phenylalanine analog AzF onto $tRNA^{Phe}$. (c) In-gel fluorescence image of Alexa Fluor 647 labeling corresponding to AzY or AzF incorporation into mammalian cell proteomes. “WT”: no exogenous amino acid added to media. “CP1 switch”: an *E. coli* TyrRS with the human CP1 peptide engrafted. “*Mj* TyRS”: TyrRS from the species *M. jannaschii*, without its accompanying tRNA. (d) High selectivity of *Sc*$Tyr_{Y43G}$ for AzY over Tyr. In-gel Alexa Fluor 647 fluorescence and coomassie blue staining of whole gel lanes were quantified to assess the degree of proteome labeling normalized to total protein content. The selectivity and rate were estimated from the line of best fit to standard Michaelis−Menten kinetics at increasing AzY concentrations with 0.4 mM Tyr, and validation via increasing Tyr concentrations with 15 $\mu$M AzY. Error bars indicate standard deviation. $n = 3$ biological replicates. (e) As in panel d but for *Mm*$Phe_{T413G}$, demonstrating high selectivity for AzF over Phe.

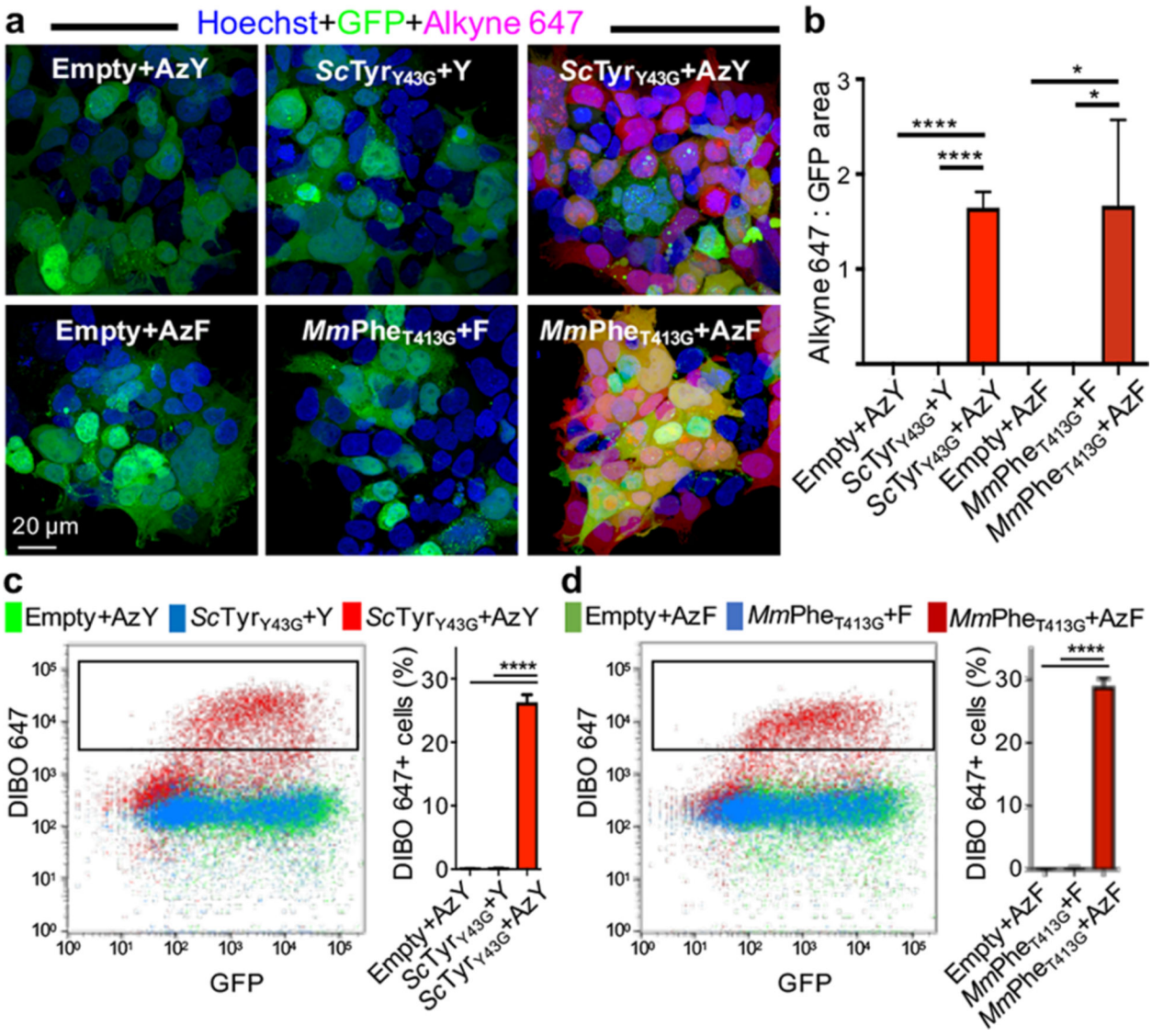



**Figure 2.**

Fluorescence imaging and analysis of azide-labeled proteomes. Incorporation of the azide-bearing ncAA AzY or AzF into proteins by *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ enables chemoselective conjugation to alkyne or DIBO-Alexa Fluor 647. Transfected HEK293T cells expressed mutant aaRS constructs or empty vector controls that coexpressed GFP, and were exposed to 125 $\mu$M AzY, AzF, tyrosine (Y), or phenylalanine (F). (a) Imaging reveals proteome labeling specific to cells expressing the mutant aaRS and exposed to AzF or AzY. Proteome labeling is pervasive across each cell. (b) Ratio of quantified Alexa Fluor 647 and GFP areas ($n$ = 3 images). (c and d) Flow cytometry of fixed HEK293T cells (after live/dead staining), with conditions as in imaging. Only cells with mutant aaRSs and exposed to ncAAs have DIBO-AF647$^{+}$ populations ($n$ = 3 biological replicates). Error bars indicate standard deviation. *$P$ < 0.05, ****$P$ <0.0001.

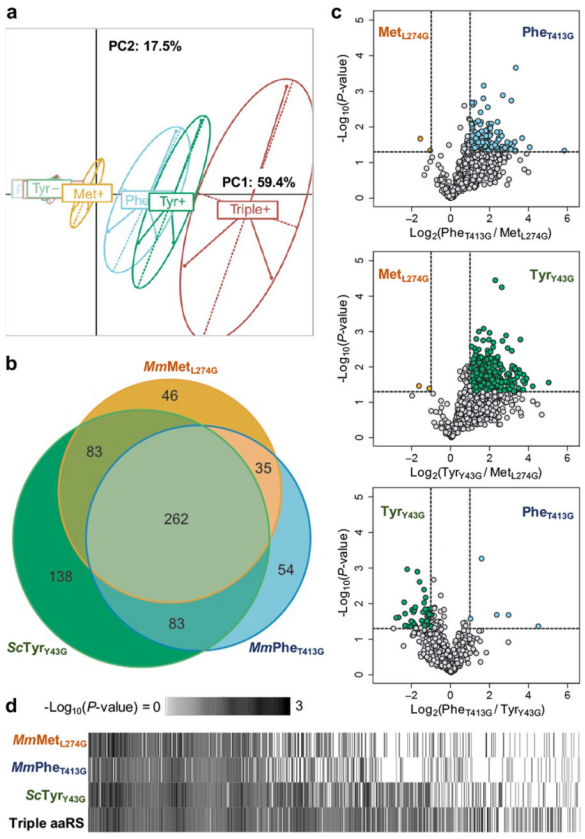



**Figure 3.**

Distinct proteomes labeled by each mutant aaRS. HEK293T cells transfected with equal total amounts of *Mm*Met$_{L274G}$, *Sc*Tyr$_{Y43G}$, and *Mm*Phe$_{T413G}$, or all three aaRSs were incubated with 125 $\mu$M of their corresponding azide-bearing ncAA or endogenous amino acids. Lysates were click-enriched on DBCO beads, washed, digested, and TMT labeled before mass spectrometry. $n = 3$ biological replicates, in technical duplicate, for each condition. (a) Each mutant aaRS differentially labels the HEK293T proteome, as spatially represented by PCA. (b) Each mutant aaRS labels and identifies its own unique set of cell proteins. Most labeled proteins (66%, 463 of 701) are commonly detected across singly expressed mutant aaRSs, increasing confidence in their identification. (c) Among the 463 proteins identified by at least 2 aaRSs, each mutant aaRS exhibits different labeling efficiencies for different proteins. (d) The coexpression of multiple mutant aaRSs ("Triple aaRS") enhances proteome coverage and detects more proteins with greater confidence, as assessed by *P*-value. Proteins significantly identified by at least one mutant aaRS were ordered by average -Log$_{10}$(*P*-value).

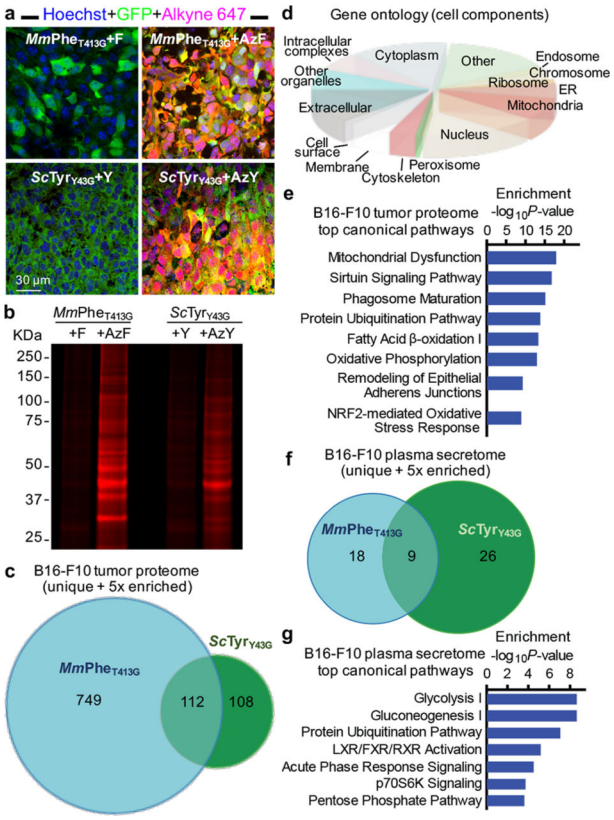
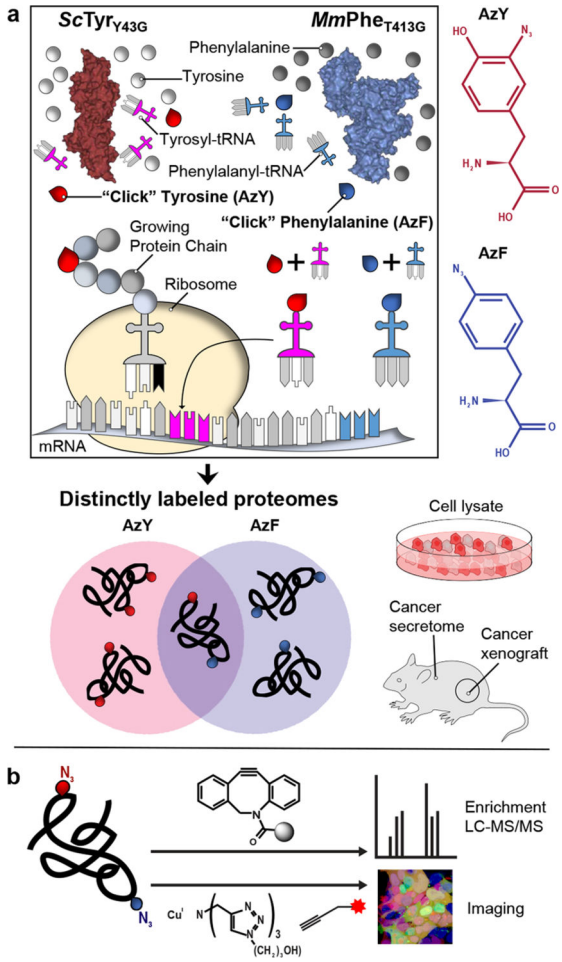



**Figure 4.**

Cell-type-specific proteome and secretome labeling *in vivo*. B16-F10 mouse melanoma cells stably expressing *Sc*Tyr$_{Y43G}$ or *Mm*PheT413G alongside GFP were implanted subcutaneously in wild-type mice, and exposed to saturating amounts of AzF, AzY, Phe, or Tyr amino acids. $n = 3$ mice except *Sc*Tyr$_{Y43G}$+Y and *Mm*Phe$_{T413G}$+F, $n = 2$ mice. (a) In situ fluorescence confocal microscopy reveals AF647$^{+}$ proteome labeling in GFP$^{+}$ melanoma cells. Alkyne AF647 reacts chemoselectively to proteome-incorporated AzF and AzY. (b) Proteome-wide labeling detected via in-gel fluorescence of tumor lysates. (c) *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ label distinct melanoma proteomes. Lysates were click-enriched on DBCO beads, washed, and digested into peptides for label-free mass spectrometry. Labeled proteomes were comprised of proteins unique to or over 5-fold more abundant than in Phe- or Tyr-exposed tumors. (d) Annotation of the labeled melanoma proteome by cellular component (STRAP), *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ proteomes combined. (e) Ingenuity Pathway Analysis. Top pathways enriched in the melanoma proteome, *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ combined. Multiple pathways are implicated in tumor biology. (f) *Sc*Tyr$_{Y43G}$ and *Mm*Phe$_{T413G}$ label distinct melanoma plasma secretomes. (g) Pathway analysis as in panel e but for the melanoma plasma secretome.




**Scheme 1. Integrating Identifications across Mutant tRNA Synthetases Yields More Complete and Confident Proteomics**[a]

[a](a) Engineered tRNA synthetases incorporate their azido amino acids preferentially across proteins of mammalian host cells. (b) Incorporated azide side chains are chemoselectively reacted with alkyne derivatives for protein identification and imaging.