Tools For Spatiotemporally Specific Proteomic Analysis In Multicellular Organisms

by

Kai P. Yuet

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy in Chemical Engineering

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California, United States of America

2016

(Defended May 24, 2016)

To Dave and Paul

Tools For Spatiotemporally Specific Proteomic Analysis In Multicellular Organisms

by
Kai P. Yuet

Abstract

The emergence of mass spectrometry-based proteomics has revolutionized the study of proteins and their abundances, functions, interactions, and modifications. However, in a multicellular organism, it is difficult to monitor dynamic changes in protein synthesis in a specific cell type within its native environment. In this thesis, we describe methods that enable the metabolic labeling, purification, and analysis of proteins in specific cell types and during defined periods in live animals. We first engineered an eukaryotic phenylalanyl-tRNA synthetase (PheRS) to selectively recognize the unnatural L-phenylalanine analog *p*-azido-L-phenylalanine (Azf). Using *Caenorhabditis elegans*, we expressed the engineered PheRS in a cell type of choice (i.e. body wall muscles, intestinal epithelial cells, neurons, pharyngeal muscles), permitting proteins in those cells – and only those cells – to be labeled with azides. Labeled proteins are therefore subject to "click" conjugation to cyclooctyne-functionalized affinity probes, separation from the rest of the protein pool and identification by mass spectrometry. By coupling our methodology with heavy isotopic labeling, we successfully identified proteins – including proteins with previously unknown expression patterns – expressed in targeted subsets of cells. While cell types like body wall or pharyngeal muscles can be targeted with a single promoter, many cells cannot; spatiotemporal selectivity typically results from the combinatorial action of multiple regulators. To enhance spatiotemporal selectivity, we next developed a two-component system to drive overlapping – but not identical – patterns of expression of engineered PheRS, restricting labeling to cells that express both elements. Specifically, we developed a split-intein-based split-PheRS system for highly efficient PheRS-reconstitution through protein splicing. Together, these tools represent a powerful approach for unbiased discovery of proteins uniquely expressed in a subset of cells at specific developmental stages.

Thesis Supervisor: David A. Tirrell

Thesis Supervisor: Paul W. Sternberg

Committee Member: Mark E. Davis

Committee Member: Mikhail G. Shapiro

Acknowledgements

- To my advisors, Profs. David Tirrell and Paul Sternberg: thank you for your guidance and support. A successful doctoral experience requires a journey from the naive first-year who puts everyone and everything on a pedestal to the original thinker who has no qualms about challenging and defying everyone and everything. Thank you for giving me just the right amount of rope for this transition.

- To my committee members, Profs. Mark Davis and Mikhail Shapiro: thank you for your advice and time.

- To Frank Truong, Jim Van Deventer, John Ngo, and Meenakshi Doma: thank you for your friendship and for teaching me everything an effective chemical biologist needs to know.

- To my collaborators: thank you for your assistance. I am forever grateful to Annie Moradian, Bobby Graham, Mike Sweredoski, Roxana Eggleston-Rangel, and Sonja Hess.

- To my colleagues: thank you for making research a fun experience. Thank you for being great labmates, listeners, and teachers.

- To Miguel Gonzalez and Xander Rudelis: thank you for letting me be your mentor. The intein project would not be where it is today without you.

- To Art Larenas, Elisa Brink, Joe Drew, Memo Correa, Mike Vicic, and Steve Gould: thank you for all the work you do for our division.

- To the incoming Chemical Engineering class of 2009, Amy Fu, Brett Babin, (Clint Regan), Devin Wiley, Jeff Bosco, Joseph Ensberg, Tristan Day: thank you for all of the memories.

- To my parents, parents-in-law, and sisters: thank you for your love and encouragement.

- To my wife Amy: thank you for everything.

Published Content and Contributions

Chapter 1 first appeared as an article in *Annals of Biomedical Engineering*: Kai P. Yuet and David A. Tirrell, *Annals of Biomedical Engineering*, February 2014, Volume 42, Issue 2, Pages 299–311 (doi: 10.1007/s10439-013-0878-3).

Chapter 2 first appeared as an article in *Proceedings of the National Academy of Science of the United States of America*: Kai P. Yuet, Meenakshi K. Doma, John T. Ngo, Michael J. Sweredoski, Robert L. J. Graham, Annie Moradian, Sonja Hess, Erin M. Schuman, Paul W. Sternberg, and David A. Tirrell, *Proceedings of the National Academy of Science of the United States of America*, March 2015, Volume 112, Issue 9, Pages 2705–2710 (doi: 10.1073/pnas.1421567112).

Table of Contents

List of Figures

List of Tables

CHAPTER 1

Chemical Tools for Temporally and Spatially Resolved Mass Spectrometry-Based Proteomics

This chapter first appeared as an article in *Annals of Biomedical Engineering*: Kai P. Yuet and David A. Tirrell, *Annals of Biomedical Engineering*, February 2014, Volume 42, Issue 2, Pages 299–311 (doi: 10.1007/s10439-013-0878-3).

1.1 Abstract

Accurate measurements of the abundances, synthesis rates, and degradation rates of cellular proteins are critical for understanding how cells and organisms respond to changes in their environments. Over the past two decades, there has been increasing interest in the use of mass spectrometry for proteomic analysis. In many systems, however, protein diversity as well as cell and tissue heterogeneity limit the usefulness of mass spectrometry-based proteomics. As a result, researchers have had difficulty in systematically identifying proteins expressed within specified time intervals, or low abundance proteins expressed in specific tissues or in a few cells in complex microbial systems. In this chapter, we present recently-developed tools and strategies that probe these two subsets of the proteome: proteins synthesized during well-defined time intervals – *temporally resolved proteomics* – and proteins expressed in predetermined cell types, cells, or cellular compartments – *spatially resolved proteomics* – with a focus on chemical and biological mass spectrometry-based methodologies.

1.2 Introduction

Messenger RNA (mRNA) profiling at the systems level (transcriptomics) with microarray or deep-sequencing technologies offers a high-throughput route to the analysis of gene expression [1, 2]. However, transcriptomic methods are blind to post-transcriptional phenomena such as translational regulation, protein modification, protein-protein interactions, or protein interactions with other molecular components. Furthermore, because mRNA abundance correlates poorly with protein abundance [3, 4, 5], reliable quantitative information about changes in protein abundance cannot be derived from microarray analysis.

Recent advances in genomic sequencing and high-resolution mass spectrometry have enabled rapid progress in the study of proteins and their abundances, modifications, interactions, and functions [6, 7]. Sample analysis workflows that employ high-resolution liquid chromatography-tandem mass spectrometry (LC-MS/MS) use site-specific endopeptidases (e.g., trypsin or Lys-C) to convert mixtures of proteins into peptides [8]. Microscale or nanoscale high-performance liquid chromatography (HPLC) [9] and/or ion exchange (IEX) chromatography [10] separates the peptides and injects them into the mass spectrometer. After an initial ion scan (MS^1), peptide ions fragment in the mass spectrometer and yield secondary "MS/MS" (MS^2) spectra. Peptide search engines like Mascot or Andromeda match the MS/MS spectra with theoretical MS/MS spectra and recover the corresponding peptide sequences from a database [11, 12]. Finally, proteins are identified using two or more unique peptides. Mass spectrometry-based proteomic methods have been applied to a wide range of problems including deciphering the protein composition of organelles [13], systematically mapping protein-protein interactions [14], and large-scale decoding of post-translational events in response to stimuli [15].

As of this writing, no eukaryotic proteome has been mapped with 100% coverage [14, 16]. The dynamic range of protein expression can span many orders of magnitude (e.g., up to 12 orders of magnitude in serum [17]). The resulting variations in protein abundance challenge the dynamic range and sequencing speed of contemporary mass spectrometers. Sequence coverage becomes increasingly important in the analysis of higher eukaryotes where proteins exhibit high levels of sequence homology due to evolution of protein families, alternative splicing, and differential processing [18, 19, 20]. Furthermore, while genome-wide transcriptome analyses are now routinely performed on small samples of RNA, even from single cells [21], low abundance proteins cannot be amplified to improve identification rates. To expand the dynamic range of identified proteins, researchers have focused on information-rich

subsets of the proteome, such as the glycoproteome [22], phosphoproteome [23], and ubiquitome [24]. In this chapter, we highlight recent work that leverages mass spectrometry and chemical biology to examine two additional subsets of the proteome: proteins synthesized during predetermined time intervals – *temporally resolved proteomics* – and proteins expressed in specific cell types, cells, or cellular compartments – *spatially resolved proteomics*.

1.3 Temporally Resolved Proteomic Analysis

1.3.1 Stable-Isotope Labeling with Amino Acids in Cell Culture

Interest in quantitative modeling and analysis of biological processes has motivated the development of tools for quantitative mass spectrometry-based proteomics such as "label-free" and isotopically labeled protein profiling [25, 26]. Comparison of protein abundances in different samples is most accurately accomplished by using stable isotopes [25,26,27]. Stable-isotope labeling of peptides relies on either chemical methods that attach isotopically labeled linkers to peptides (e.g., ICAT [28] or iTRAQ [29]) or metabolic methods that incorporate isotopically labeled amino acids into peptides (e.g., ^{15}N-labeling [30] or SILAC [31]).

Stable-isotope labeling by amino acids in cell culture (SILAC, **Figure 1.1, Top Left**) introduces the label at the amino acid level and overcomes complex isotopic clusters found in alternative metabolic methods like ^{15}N-labeling. Distinct isotopologs of the labeled amino acids – typically "light" arginine ($^{12}C_6{}^{14}N_4{}^{1}H_{14}{}^{16}O_2$) and "light" lysine ($^{12}C_6{}^{14}N_2{}^{1}H_{14}{}^{16}O_2$) in one sample, and "heavy" arginine ($^{13}C_6{}^{15}N_4{}^{1}H_{14}{}^{16}O_2$) and heavy lysine ($^{13}C_6{}^{15}N_2{}^{1}H_{14}{}^{16}O_2$) in the other – are added to the samples of interest. After several generations of growth, all cellular proteins have uniformly incorporated the labeled amino acids. Cells are lysed and subjected to proteolysis, and the resulting peptide pools are mixed. In theory, both the "heavy" and "light" forms

of each peptide should behave identically with respect to processing and LC-MS/MS analysis. However, the origin of each peptide can be determined from the mass of its isotope label, and the relative signal intensities of the "heavy" and "light" peptides indicate the relative abundances of the corresponding proteins in the two samples. Its simplicity and accuracy have made SILAC an increasingly popular method, both for cell culture-based quantitative proteomics [32], and for whole organism quantitative proteomics [3, 33, 34, 35, 36].

1.3.2 Repurposing SILAC for Temporally Resolved Proteomic Analysis

Because "old" and "new" copies of cellular proteins are chemically indistinguishable, selective analysis of the subset of the proteome that is expressed within a specified time interval is challenging. A classic method for quantifying specific protein turnover rates involves pulse labeling cells with a radiolabeled amino acid (typically ^{35}S-methionine or cysteine) [37,38]. Only those proteins synthesized during the pulse incorporate the radioactive element, and their fates can be monitored by standard radioisotope methods. Analogous pulse labeling with stable isotopes can also determine protein turnover or transport (**Figure 1.1, Top Middle**) [39, 40, 41].

Pulse labeling with amino acids and nucleosides can be combined to yield an unbiased and comprehensive picture of protein and mRNA dynamics. To determine if the cellular abundance of proteins is predominantly controlled at the level of transcription or translation, Schwanhäusser *et al.* pulse labeled mammalian cells with the nucleoside analog 4-thiouridine (**Figure 1.2**) and heavy arginine and lysine to quantify absolute mRNA and protein abundances, half-lives, and transcription and translation rates for more than 5000 genes in mouse fibroblast cells [42]. The thio-substituted nucleoside 4-thiouridine is not a natural component of nucleic acids but is incorporated into RNA biosynthetically; only newly synthesized RNAs are thio-labeled and can be tagged and purified using commercially available reagents [43, 44]. Although mRNA

and protein levels correlated better than in previous studies [3, 4, 5, 45, 46, 47], their half-lives showed no overall correlation. Nevertheless, the authors note that genes with similar combinations of mRNA and protein half-lives share common functions, suggesting that half-lives evolved under similar constraints. For example, many genes involved in constitutive processes like translation, respiration, and central metabolism have stable mRNAs and proteins, consistent with the requirement for conservation of resources. On the other hand, genes involved in transcription regulation, signaling, chromatin modification, and cell cycle-specific processes have unstable mRNAs and proteins, consistent with the requirement for rapid regulation. Correlation of protein abundance with mRNA levels and with translation rates showed that protein copy numbers are determined primarily at the level of translation.

1.3.3 Pulsed SILAC

Unlike pulse labeling with a single label to determine protein turnover or transport in the same sample, pulsed stable-isotope labeling with amino acids in cell culture (pSILAC, **Figure 1.1, Top Right**) quantifies differences in protein synthesis between different samples integrated over the measurement time after the pulse. In pSILAC, cells in two different samples previously cultured in light media are transferred to two different pulse media: one containing heavy amino acids (typically arginine-10: $^{13}C_6{}^{15}N_4{}^{1}H_{14}{}^{16}O_2$ and lysine-8: $^{13}C_6{}^{15}N_2{}^{1}H_{14}{}^{16}O_2$) and the other containing medium-heavy amino acids (e.g., arginine-6: $^{13}C_6{}^{14}N_4{}^{1}H_{10}{}^{2}H_4{}^{16}O_2$ and lysine-4: $^{12}C_6{}^{14}N_2{}^{1}H_{10}{}^{2}H_4{}^{16}O_2$) [48, 49]. During labeling, only newly synthesized proteins incorporate either the heavy or the medium-heavy amino acids. Peptides derived from heavy or medium-heavy proteins are distinguishable during the initial ion scan from preexisting, light proteins on the basis of the mass difference introduced by the isotope label. Intensity ratios for heavy and medium-heavy peptides directly indicate the relative abundances of the corresponding newly synthesized proteins in

the samples of interest. Light preexisting peptides are identified but ignored during quantification.

Selbach and coworkers introduced pSILAC in 2008 as a method to measure changes in protein translation involved in cellular iron homeostasis [49]. Since then, researchers have used pSILAC to assess protein dynamics associated with microRNA overexpression [48, 50, 51], monocyte-macrophage differentiation [52], hyperglycemia-induced stress [53] and mammalian target of rapamycin inhibition [54]. Selbach *et al.* utilized pSILAC to measure changes in the synthesis of several thousand proteins in HeLa cells in response to either microRNA transfection or endogenous microRNA knockdown [48]. HeLa cells cultivated in light media were first transfected with miR-1, miR-155, miR-16, miR-30a, or let-7b, or mock transfected. Eight hours later, cells were transferred to media containing either medium-heavy or heavy amino acids for 24 hours. Unexpectedly, pSILAC in this study and subsequent studies revealed that overexpression of a single microRNA can lead to repression of hundreds of proteins; however, the repressive effect was relatively small and rarely exceeded fourfold [50,51]. These results cast microRNAs as general orchestrators that tune cellular physiology and metabolism in subtle ways in response to specific cues.

1.3.4 Bio-Orthogonal Non-Canonical Amino Acid Tagging

Under some circumstances (e.g., for short labeling times), the dynamic range and complexity of the proteome can preclude reliable identification and analysis of low-abundance isotope-labeled proteins in the presence of more abundant preexisting proteins. Moreover, the use of multiple labels increases the number of distinct peptide forms and the computational complexity associated with protein identification. These limitations can be addressed by *selective enrichment* and identification of newly synthesized proteins.

In 2006, Dieterich *et al.* introduced the bio-orthogonal non-canonical amino acid tagging (BONCAT, **Figure 1.1, Bottom Left**) strategy to enable selective enrichment and identification of newly synthesized proteins in cells and tissues [55]. This strategy relies on bio-orthogonal functional groups – functional groups that react rapidly and selectively with each other but remain inert to the functional groups normally found in biological systems [56]. An exemplary bio-orthogonal group, the azide is small, kinetically stable and absent from living systems, yet can be modified easily and selectively through the Staudinger ligation with triarylphosphine reagents [57], the copper(I)-catalyzed "click" cycloaddition with terminal alkynes [58, 59] or the strain-promoted "click" cycloaddition with strained cyclooctynes [60].

BONCAT employs a two-stage procedure. First, pulse labeling of cells or tissues with the azide-bearing methionine (Met, **Figure 1.2**) analog azidohomoalanine (Aha, **Figure 1.2**) or the alkyne-bearing methionine analog homopropargylglycine (Hpg, **Figure 1.2**) enables metabolic incorporation of either azide or alkyne functional groups into the proteome. In bacterial and mammalian systems, wild-type methionyl-tRNA synthetases (MetRSs) are capable of appending the non-canonical amino acid Aha (or Hpg) to cognate transfer RNAs ($tRNA^{Met}$) [55,61]. Aha (or Hpg) is thereby incorporated into proteins made during the Aha (or Hpg) pulse. Second, a bio-orthogonal ligation with the complementary reactive group conjugated to a probe enables detection of Aha- or Hpg-tagged proteins. Treatment of azide-tagged proteins with alkyne-functionalized fluorescent dyes permits visualization of newly synthesized proteins [62, 63, 64]. Alternatively, treatment of azide-tagged proteins with alkyne-functionalized affinity reagents allows selective enrichment of proteins made during the Aha pulse. Enriched proteins can then be identified by LC-MS/MS [65, 66].

BONCAT has been used in mammalian cells [55], in both unicellular [67, 68] and multicellular organisms [69], and even adapted to measure genome-wide nucleosome turnover dynamics in *Drosophila* S2 cells [70]. Here, we highlight a few recent examples with an emphasis on proteomic discovery. Zhang *et al.* reported a tandem labeling (Aha and alkyne-functionalized palmitic acid) and detection method to monitor the dynamic acylation of Lck, an N-myristoylated and S-palmitoylated non-receptor tyrosine kinase required for T-cell activation [71]. Liu *et al.* expanded on this concept by enriching proteins made in Jurkat cells during a specified time interval, then progressively monitoring their post-translational modifications over time by quantitative in-gel fluorescence scanning with Aha/Hpg and eight additional azide- and alkyne-functionalized post-translational modification probes [72]. The up-regulation of myristoylated protein kinase A was found to be intimately linked to butyric acid-induced apoptosis.

As local protein synthesis is critical for long-term functional synaptic changes, BONCAT is an attractive tool for neurobiological studies. Tcherkezian *et al.* examined the colocalization of DCC (Deleted in Colorectal Cancer) transmembrane receptors with newly synthesized proteins in cultured commissural axon growth cones with Aha-labeling [73]. Hodas *et al.* used BONCAT to identify proteins translated in intact hippocampal neuropil sections upon treatment with the selective D1/D5 dopamine receptor agonist SKF81297 [74]. Yoon *et al.* developed DIGE-NCAT – a combination of 2D difference gel electrophoresis (2D-DIGE) and BONCAT – to examine changes in the proteome of *Xenopus* retinal ganglion cell (RGC) axons in response to stimulation with Engrailed-1 [75]. Engrailed-1 belongs to a family of transcription factors previously shown to cause rapid translation-dependent guidance responses in RGC axons [76]. The authors first severed distal portions of the axon bundles from *Xenopus* eyes and stimulated them with Engrailed-1 for 1 hour along with the addition of Aha. Next, Aha-tagged axonally synthesized proteins were treated with an alkynyl dye and

resolved on 2D-DIGE. Compared to control gels, spots with the greatest difference in fluorescence were analyzed by matrix-assisted laser desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS). Surprisingly, this strategy revealed that the intermediate filament protein lamin B2 (LB2) – normally associated with the nuclear membrane – is axonally synthesized in response to stimulation. Coupled with the finding of LB2s association with mitochondria, these results suggest that LB2-promotion of mitochondrial function is needed for axon maintenance.

1.3.5 Quantitative Non-Canonical Amino Acid Tagging

How can we quantitatively study proteomic changes during short time intervals (e.g., in response to a stimulus) without the complications that arise from the abundance of preexisting proteins? As labeling times decrease, co-eluting preexisting peptides increasingly obscure low abundance pSILAC-labeled peptides during the first mass spectrometry scan. To overcome this limitation, pSILAC and BONCAT have been combined in an approach designated Quantitative Non-Canonical Amino Acid Tagging (QuaNCAT, **Figure 1.1, Bottom Right**). In a QuaNCAT experiment, two parallel populations of cells in light media are transferred for a limited time to either medium-heavy or heavy media that also contain Aha [77, 78]. During the labeling period, newly synthesized proteins incorporate either the heavy or medium-heavy amino acid as well as Aha. Through BONCAT enrichment of the combined protein pools, preexisting light peptides are greatly reduced in abundance. As in pSILAC, ratios of intensities of heavy and medium-heavy peptides directly indicate the relative abundances of newly synthesized proteins in the two samples.

To determine the benefit of selectively enriching and quantifying secreted proteins, Eichelbaum *et al.* labeled two human cell lines (PC3 and WPMY-1) with Aha and isotopologs of arginine and lysine for 24 hours [77]. With enrichment via on-bead azide-alkyne cycloaddition and trypsin digestion, 684 secreted proteins were quantified with high correlation (R = 0.96) between biological duplicates. Without enrichment, only 22 proteins were quantified with low correlation (R = 0.02). The authors identified as many as 500 proteins, even at short labeling times (2 hours), including several known lipopolysaccharide (LPS) effector proteins as well as many other previously unassociated proteins during LPS stimulation of mouse macrophages. Howden *et al.* tested QuaNCAT by examining changes in expression of more than 600 proteins enriched from freshly isolated human CD4-positive T cells stimulated by activation with phorbol 12-myristate 13-acetate and ionomycin during 2 and 4 hour pulses [78]. Many transcription factors and transcriptional regulators, cytokines, activation surface markers, protein chaperones, and proteins involved in cytoskeleton dynamics and vesicle transport, were among the proteins found to be substantially increased in expression following stimulation. While less than 1% of the proteome was labeled in a 2 hour pulse, the post-enrichment protein pool consisted of 10-20% heavy and medium-heavy proteins. The ability to resolve protein abundance changes over very short time spans should prove useful when studying systems previously thought to be "unquantifiable" by SILAC such as short-lived progenitor or primary cells [79].

1.3.6 O-Propargyl-Puromycin Labeling

Salic and coworkers have developed an alternative approach to the selective labeling of newly synthesized proteins, in which a puromycin analog (O-propargyl-puromycin, **Figure 1.3, Top**) bearing a terminal alkyne is used to label nascent polypeptide chains [80]. Puromycin mimics aminoacyl-tRNAs and terminates translational elongation after entering the acceptor site of the ribosome. Salic and coworkers showed

that O-propargyl-puromycin not only terminates nascent polypeptide chains but also appends an alkyne label at the C-terminus. Liu and coworkers injected mice intraperitoneally with O-propargyl-puromycin and harvested tissues 1 hour later. Tissues from these mice displayed specific patterns of O-propargyl-puromycin incorporation into nascent proteins. For example, in the small intestine, the most intense labeling occurred in cells in the crypts and at the base of intestinal villi, consistent with the high proliferative and secretory activity of these cells (**Figure 1.3, Bottom**).

1.4 Spatially Resolved Proteomic Analysis

1.4.1 Coupling Flow Cytometry and Mass Spectrometry

Proteomic analysis of rare cells in heterogeneous environments presents important, difficult challenges. Combining flow cytometry with quantitative proteomics provides a solution in some systems of this kind. For example, Rechavi *et al.* combined quantitative proteomics and high-purity cell sorting to discover proteins transferred from human B cells to natural killer cells [81]. These authors developed a strategy called "trans-SILAC", in which one cell type – capable of transferring proteins – is labeled with heavy arginine and lysine and the other cell type – the recipient of transferred proteins – remains unlabeled. To initiate contact- and actin cytoskeleton-dependent protein transfer, the heavy "donor" B cells and freshly isolated light "recipient" natural killer cells were co-incubated for 1.5 hours. Fluorescence-activated cell sorting separated recipient cells from donor cells before the proteomic workflow, and transferred proteins in the recipient cells were identified by their mass shifts. Analysis of the transferred proteins revealed significant enrichment for the annotation term "MHC class II protein complex", as expected.

1.4.2 Cell-Selective BONCAT

Although some cell types can be isolated for proteomic analysis through cell sorting or laser-capture techniques, others cannot. Subsets of neurons or glia in the central nervous system, for example, are difficult to isolate by dissection or dissociation methods. Or perhaps one would like to analyze proteomic responses of pathogenic bacteria hiding inside host cells, or of specific cells in multicellular animals, without interference from proteins derived from the host or from all of the other cells in the animal. Can we perform these kinds of tasks without prior separation of the cells of interest?

Promising solutions to analogous problems in genomics and transcriptomics rely on spatially restricted enzymatic labeling. In 2010, Henikoff and coworkers developed INTACT (isolation of nuclei tagged in specific cell types) – a method that allows affinity-based isolation of nuclei from individual cell types in a tissue [82]. In INTACT, the spatial restriction of the *Escherichia coli* biotin ligase BirA and a nuclear targeting fusion protein results in biotin-labeled nuclei in the cell type of choice. First described in *Arabidopsis*, INTACT has been extended to profile gene expression and histone modifications in adult *Caenorhabditis* muscle [83] and Kenyon cells and octopaminergic neurons in the adult *Drosophila* brain [84].

For RNA, Miller *et al.* developed TU-tagging, a method that allows affinity-based isolation of RNA from individual cell types of a tissue [85]. In TU-tagging, the spatial restriction of the *Toxoplasma gondii* nucleotide salvage enzyme uracil phosphoribosyltransferase enables RNA in the cell type of choice to be labeled with 4-thiouracil (**Figure 1.2**). Unlike 4-thiouridine, 4-thiouracil is not recognized by the endogenous biosynthetic machinery. Recently, Gay *et al.* used TU-tagging to purify transcripts from rare (¡5%) cells, such as *Tie2:Cre+* brain endothelia/microglia in mice [86].

Analogous strategies have been developed for cell-selective proteomic analysis. Because endogenous methionyl-tRNA synthetases charge Aha to cognate tRNAs in all cell types, Aha-based BONCAT is not cell-selective – newly synthesized proteins in all cell types are labeled. To enable cell-selective proteomic analysis, Tirrell and coworkers engineered a family of mutant *E. coli* MetRSs capable of appending the azide-bearing methionine analog azidonorleucine (Anl, **Figure 1.2**) to $tRNA^{Met}$ [87, 88]. Anl is not a good substrate for any of the wild-type aminoacyl-tRNA synthetases in bacteria to mammals; it is excluded from proteins made in wild-type cells but incorporated readily into proteins made in cells that express an appropriately engineered MetRS. Anl-labeling does not require depletion of methionine if the mutant MetRS activates Anl faster than methionine.

Ngo *et al.* first achieved cell type-selectivity by outfitting an *E. coli* strain with the L13N/Y260L/H301L mutant form of the *E. coli* MetRS (NLL-*Ec*MetRS). Proteins made in this strain could be labeled with Anl in co-culture with murine alveolar macrophages, which were not labeled (**Figure 1.4**) [89]. Bacterial proteins were effectively separated from murine proteins by treatment of mixed lysates with alkyne-functionalized biotin reagents and subsequent affinity chromatography on NeutrAvidin resin.

System-wide identification of bacterial proteins expressed or secreted during infection can provide new insight into mechanisms of bacterial pathogenesis. Grammel *et al.* studied cultures of Raw264.7 murine macrophages infected with *Salmonella typhimurium* (a Gram-negative intracellular pathogen) outfitted with NLL-*Ec*MetRS [90]. Cultures were pulse-labeled with 2-aminooctynoic acid (Aoa, **Figure 1.2**), an alkyne analog of Anl, for 1 hour. Of the 218 proteins identified, 185 (85%) were *Salmonella* proteins and 33 were mouse proteins. Five of the *Salmonella* proteins (SodM, SsrB, SseA, PipB2, and PhoP) had been previously described as virulence factors.

In more recent work, Ngo and coworkers demonstrated that heterologous expression of the NLL-*Ec*MetRS enables incorporation of Anl into proteins expressed in human (HEK293) cells, permitting enrichment and visualization of proteins made during various stages of the cell cycle [91]. Interestingly, Anl replaces methionine selectively at N-terminal positions – not at internal sites. Site-selectivity occurs because NLL-*Ec*MetRS catalyzes aminoacylation only of the mammalian initiator $tRNA^{Met}$, not the mammalian elongator $tRNA^{Met}$. Through judicious selection of regulatory elements, systems of this kind will enable cell-selective or "cell-state-selective" [92] interrogation of protein synthesis in co-cultures of mammalian cells, in virally transfected tissues or even in living animals.

To expand the set of tools available for cell-selective BONCAT, Truong *et al.* engineered a MetRS variant capable of activating propargylglycine (Pra, **Figure 1.2**) but not Anl [93]. Pra is an alkynyl amino acid smaller than methionine and is not activated by wild-type aminoacyl-tRNA synthetases or by NLL-*Ec*MetRS. Using directed evolution, Truong et al isolated a MetRS variant (designated propargylglycyl-tRNA synthetase; PraRS, L13P/A256G/P257T/Y260Q/H301F/A331V/Δ548E) capable of near-quantitative replacement of Met by Pra in proteins. In protein mixtures that have been labeled both with Anl and with Pra, treatment with cyclooctyne-functionalized probes selectively tags Anl side chains, and subsequent treatment with azide-functionalized probes selectively tags Pra side chains. By using one promoter to drive expression of NLL-*Ec*MetRS in one cell and a different promoter to drive expression of PraRS in another cell, researchers can perform differential, cell-selective BONCAT in complex multicellular systems without prior separation.

Restricting expression of mutant methionyl-tRNA synthetases by using promoters active only in specific cells or tissues should restrict Anl/Aoa labeling to those cells or tissues. Additional specificity should be possible by exploiting the combinatorial action of multiple regulators. To this end, Mahdavi *et al.* bisected the NLL-*Ec*MetRS

and found several split variants capable of charging Anl to $tRNA^{Met}$ [94]. Because labeling requires expression of both the N- and C-terminal fragments of the synthetase, this system allows the investigator to restrict labeling to cells in which two promoters of interest are active.

1.4.3 Ascorbate Peroxidase Labeling

Ting and coworkers have recently used an engineered variant of ascorbate peroxidase (APEX) to selectively tag proteins that localize to the mitochondrial matrix in living cells [95]. The Ting laboratory had previously used APEX to catalyze the H_2O_2-dependent polymerization of diaminobenzidine to provide contrast in electron microscopy [96]. They genetically targeted APEX to the mitochondrial matrix of human embryonic kidney (HEK) cells, initiated labeling by adding biotin-phenol and H_2O_2, and stopped labeling after 1 minute by cell fixation or lysis (**Figure 1.5**). APEX processes biotin-phenol in a hydrogen peroxide-dependent manner to yield highly reactive products that covalently link biotin to electron-rich amino acids such as tyrosine, tryptophan, histidine, and cysteine. Unlike cell-selective BONCAT, APEX labeling is not time-selective. Any proteins accessible to the phenoxyl radical – including pre-existing proteins made before the biotin-phenol pulse – are biotinylated in the 1 minute "snapshot". Enrichment and mass spectrometry of the biotinylated proteins led to identification of 464 known mitochondrial proteins as well as 31 that were not previously known to localize to mitochondria. Demonstrating APEX-based proteomics as a viable discovery tool, a random subset of these 31 "mitochondrial orphans" were verified by fluorescence imaging to have complete or partial mitochondrial localization.

1.5 Conclusions

This chapter highlights chemical biological strategies that allow researchers to identify, isolate, and quantitatively analyze proteins within defined windows of time and space. Many of these strategies rely on amino acids with isotopic signatures or functional groups that are not found in biological systems. Incorporation of these "tags" permits the investigator to distinguish proteins made in different time intervals or in different cells or organelles. In combination with these tagging methods, mass spectrometry becomes an even more powerful tool for addressing complex biological questions.

1.6 Figures

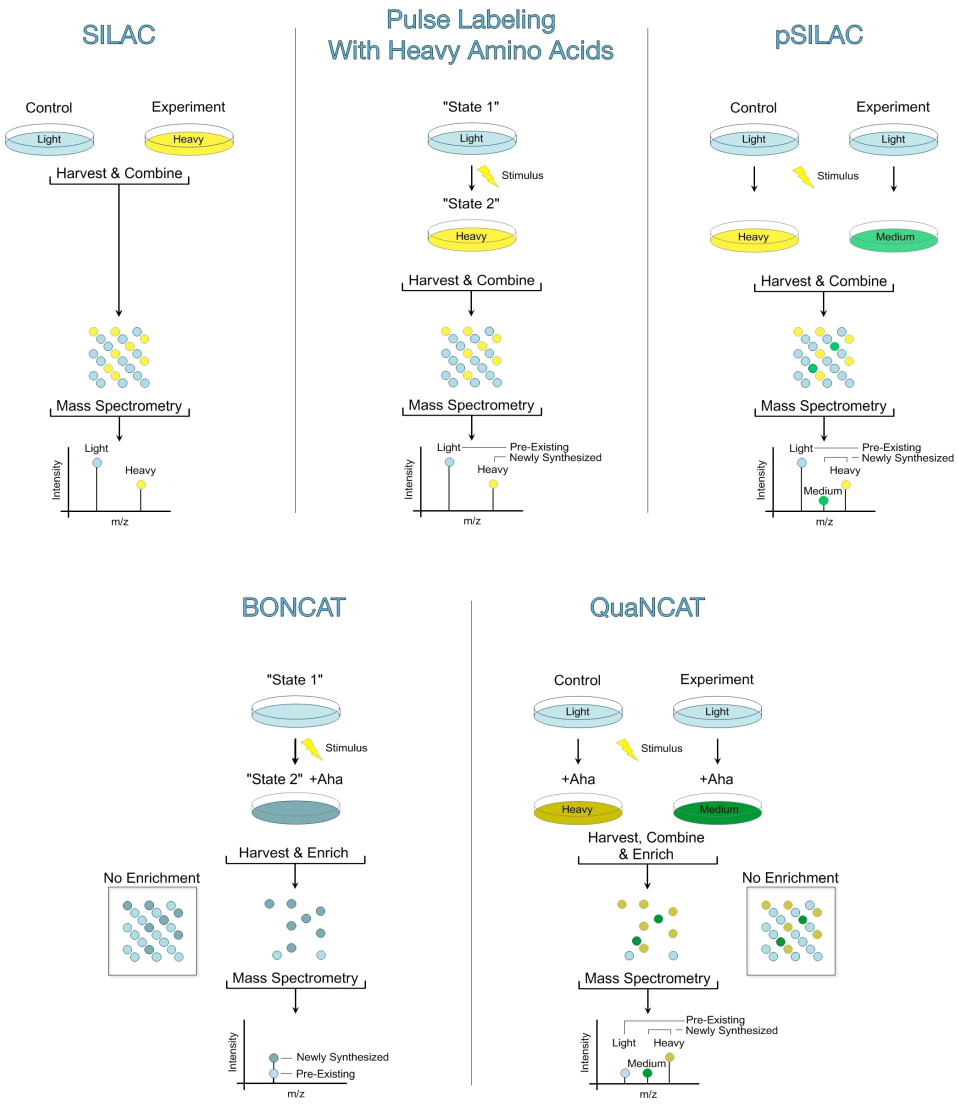


Figure 1.1: *Cells can be metabolically labeled with a combination of Aha and/or stable isotopic variants of arginine and lysine in 5 workflows: standard SILAC (top left), pulse labeling with heavy amino acids (top middle), pSILAC (top right), BONCAT (bottom left), and QuaNCAT (bottom right).*

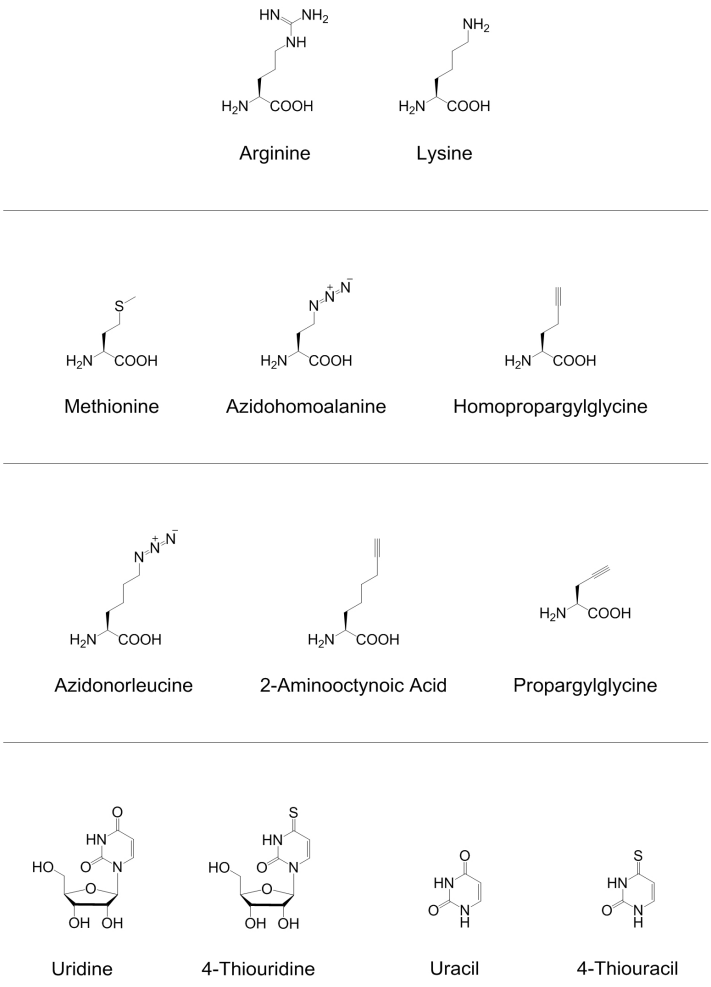
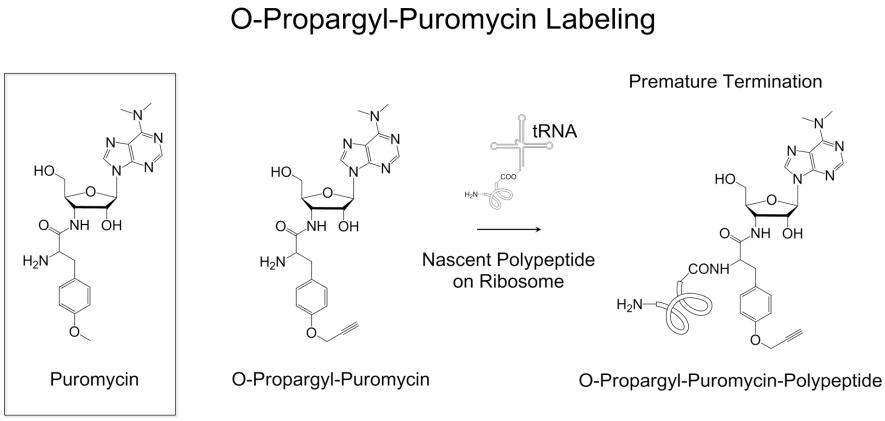


Figure 1.2: *Structures discussed in this chapter: amino acids for stable isotopic labeling (top row), methionine and analogs that are substrates for wild-type methionyl-tRNA synthetases (second row), methionine analogs that require the expression of mutant methionyl-tRNA synthetases for proteomic incorporation (third row), and uridine and uracil as well as their thio-substituted analogs 4-thiouridine and 4-thiouracil (last row).*


O-Propargyl-Puromycin Labeling
Premature Termination
tRNA
Nascent Polypeptide
on Ribosome
Puromycin
O-Propargyl-Puromycin
O-Propargyl-Puromycin-Polypeptide

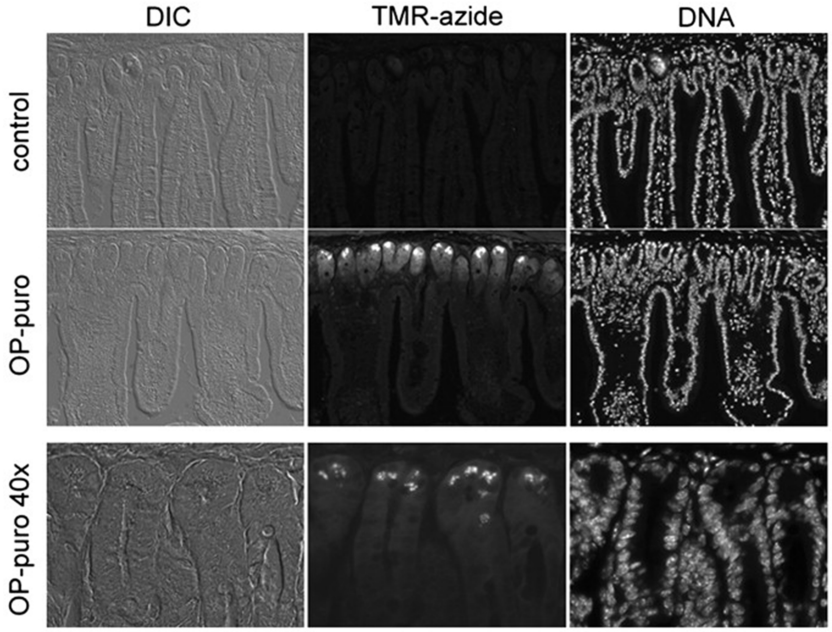
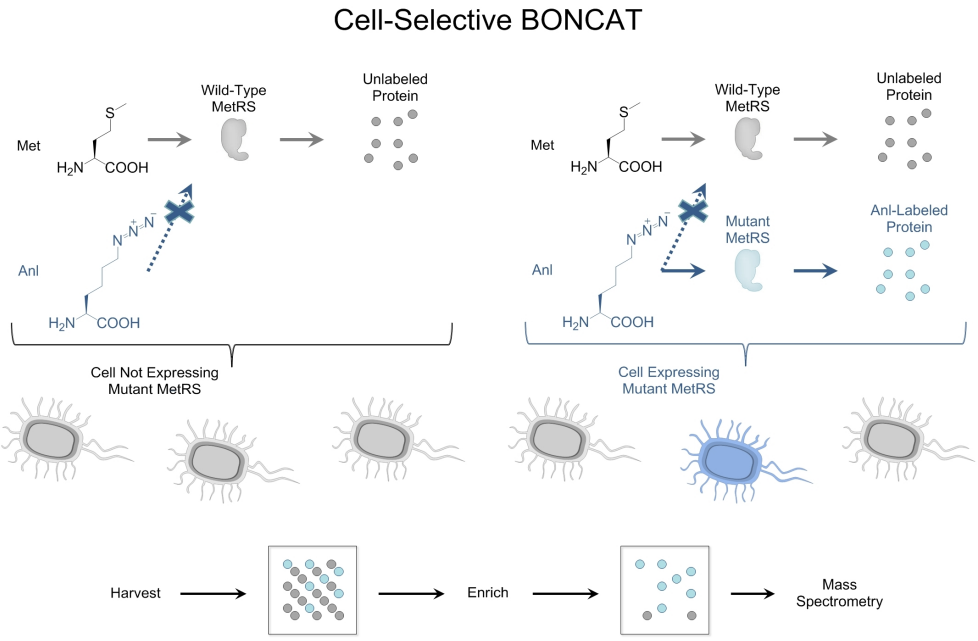

DIC
TMR-azide
DNA
control
OP-puro
OP-puro 40x

Figure 1.3 *(preceding page): (Top) Puromycin and its alkyne analog O-propargyl-puromycin incorporate into nascent polypeptide chains on translating ribosomes, resulting in premature termination of nascent polypeptide chains. (Bottom) Sectioning of mouse small intestine showed that OP-Puro labeling occurred primarily in cells in the crypts and the cells at the base of the villi. (Adapted with permission from Liu et al., Proc. Natl. Acad. Sci. U.S.A.,* ***109****, 413-418, 2012. Copyright 2012 Proceedings of the National Academy of Sciences USA.)*


Cell-Selective BONCAT
Wild-Type MetRS
Unlabeled Protein
Met
Anl
Cell Not Expressing Mutant MetRS
Mutant MetRS
Anl-Labeled Protein
Cell Expressing Mutant MetRS
Harvest
Enrich
Mass Spectrometry

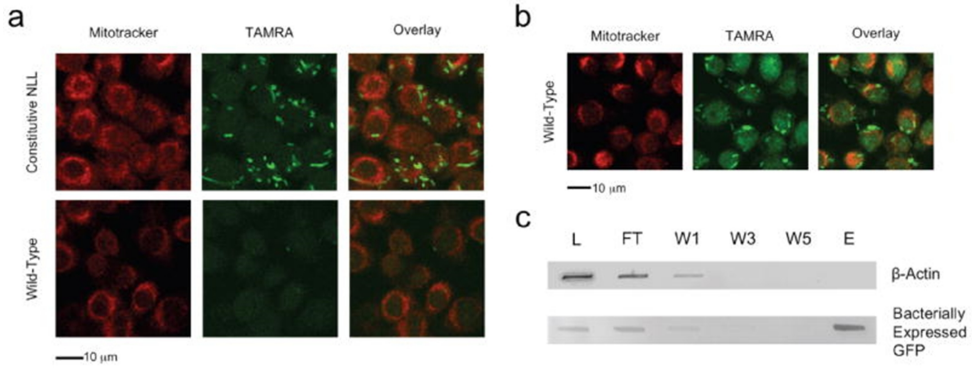

a
Mitotracker
TAMRA
Overlay
Constitutive NLL
Wild-Type
10 μm
b
Mitotracker
TAMRA
Overlay
Wild-Type
10 μm
c
L
FT
W1
W3
W5
E
β-Actin
Bacterially Expressed GFP

Figure 1.4 *(preceding page)*: *(Top) Cell-selective BONCAT performed in a mixture of cells. Restricting expression of a mutant synthetase to a certain cell restricts Anl labeling to that cell (highlighted in blue). Proteins synthesized in cells (highlighted in gray) that do not express the mutant synthetase are neither labeled nor detected following enrichment. (Bottom) Cell-selective labeling in mixtures of bacterial and mammalian cells. (a) In Anl-containing mixed cultures of E. coli and mouse alveolar macrophages, only E. coli cells constitutively expressing the mutant NLL-EcMetRS were labeled by TAMRA-alkyne. Macrophages were labeled with Mitotracker Deep Red and displayed low TAMRA-alkyne background emission. (b) In Aha-containing mixed cultures of E. coli and mouse alveolar macrophages, both wild-type E. coli cells and macrophages exhibited strong TAMRA-alkyne emission; incorporation of Aha occurs in both cell types. (c) Mixed cell lysate was subjected to conjugation with alkyne-functionalized biotin, and labeled proteins were enriched by NeutrAvidin affinity chromatography. Immunoblotting of unbound flow-through (FT), washes (W1, W3, W5) and eluent (E) reveals enrichment of the bacterial marker protein GFP. (Adapted with permission from Ngo et al., Nat. Chem. Biol.,* ***5****, 715-717, 2009. Copyright 2009 Nature Publishing Group.)*

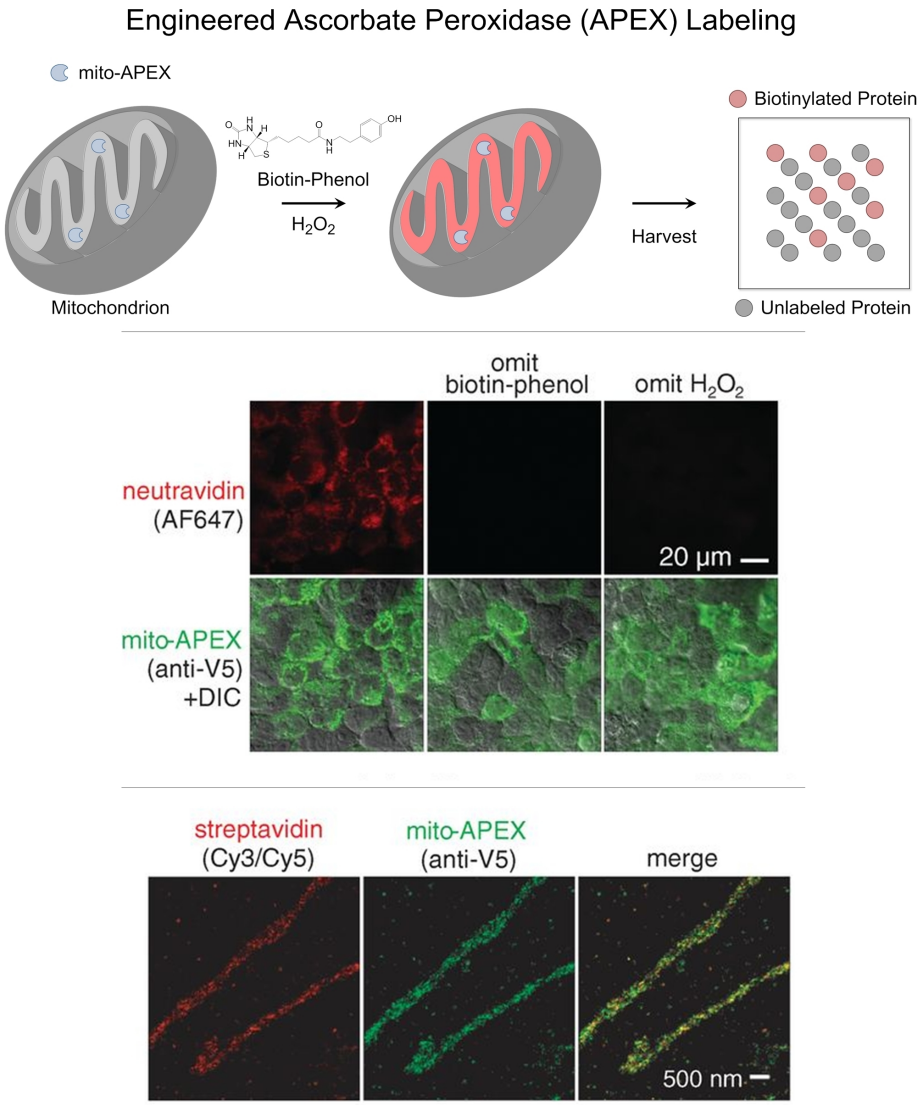

Engineered Ascorbate Peroxidase (APEX) Labeling
mito-APEX
Biotin-Phenol
H_2O_2
Mitochondrion
Harvest
Biotinylated Protein
Unlabeled Protein
omit biotin-phenol
omit H_2O_2
neutravidin (AF647)
20 µm
mito-APEX (anti-V5) +DIC
streptavidin (Cy3/Cy5)
mito-APEX (anti-V5)
merge
500 nm

Figure 1.5 *(preceding page)*: *(Top) Selective labeling of the mitochondrial matrix proteome in living cells requires 1) genetically targeting APEX to the mitochondrial matrix (mito-APEX), 2) initiating biotinylation by adding biotin-phenol and H_2O_2 to the medium, and 3) stopping biotinylation by cell fixation or lysis. (Middle) In human embryonic kidney cells, only mitochondria that expressed mito-APEX and were exposed to both biotin-phenol and H_2O_2 contained biotinylated proteins (stained with NeutrAvidin-Alexa Fluor 647). Both confocal fluorescence imaging (Middle) and stochastic optical reconstruction microscopy (STORM) (Bottom) showed that biotinylated proteins (stained with Streptavidin-Cy3/Cy5 for STORM images) overlapped with mito-APEX only in the mitochondrial matrix. (Adapted with permission from Rhee et al., Science,* ***339****, 1328-1331, 2013. Copyright 2013 The American Association for the Advancement of Science.)*

CHAPTER 2

Cell-Specific Proteomic Analysis in Caenorhabditis elegans

This chapter first appeared as an article in *Proceedings of the National Academy of Science of the United States of America*: Kai P. Yuet, Meenakshi K. Doma, John T. Ngo, Michael J. Sweredoski, Robert L. J. Graham, Annie Moradian, Sonja Hess, Erin M. Schuman, Paul W. Sternberg, and David A. Tirrell, *Proceedings of the National Academy of Science of the United States of America*, March 2015, Volume 112, Issue 9, Pages 2705–2710 (doi: 10.1073/pnas.1421567112).

2.1 Abstract

Proteomic analysis of rare cells in heterogeneous environments presents difficult challenges. Systematic methods are needed to enrich, identify, and quantify proteins expressed in specific cells in complex biological systems including multicellular plants and animals. Here, we have engineered a *Caenorhabditis elegans* phenylalanyl-tRNA synthetase capable of tagging proteins with the reactive non-canonical amino acid *p*-azido-L-phenylalanine. We achieved spatiotemporal selectivity in the labeling of *C. elegans* proteins by controlling expression of the mutant synthetase using cell-selective (body wall muscles, intestinal epithelial cells, neurons, and pharyngeal muscle) or state-selective (heat-shock) promoters in several transgenic lines. Tagged proteins are distinguished from the rest of the protein pool through bio-orthogonal conjugation of the azide side chain to probes that permit visualization and isolation of labeled proteins. By coupling our methodology with stable-isotope labeling of amino acids in cell culture (SILAC), we successfully profiled proteins expressed in pharyngeal muscle cells, and in the process, identified proteins not previously known to be expressed in these cells. Our results show that tagging proteins with spatiotemporal selectivity can be achieved in *C. elegans* and illustrate a convenient and effective approach for unbiased discovery of proteins expressed in targeted subsets of cells.

2.2 Introduction

Caenorhabditis elegans is a small, free-living nematode (roundworm) (**Figure 2.1**) [97]. Since *C. elegans*'s introduction as a model organism to study development, neurobiology, and behavior by Sydney Brenner several decades ago [98], this multicellular organism has been used to study a variety of intricate biological and molecular processes such as ageing [99], apoptosis [100], cell cycle [101], cell polarity [102], cell signaling [103], disease pathogenesis [104], gene regulation [105], host-pathogen

interactions [106], and metabolism [107].

C. elegans biology has several attractive features that contribute to its increasing popularity in research. First, *C. elegans* is easy to maintain in the laboratory. Although the worm normally colonizes various microbe-rich habitats like decaying plant matter, it survives in liquid culture or on the surface of agar plates, feeding on a lawn of *Escherichia coli* [108, 109]. Second, *C. elegans* has a short generation time and reproduces prolifically. Following hatching, it progresses through 4 larval stages (L1 to L4) in about 3 days on route to becoming a mature, egg-bearing adult capable of yielding up to 300 progeny (**Figure 2.2**) [110]. Third, *C. elegans* is small (about 1.3 mm in length and 80 microns in diameter in the adult stage), cell number invariant, and transparent. Hermaphrodites contain exactly 959 and males 1031 somatic cells whose lineage have been completely characterized by light microscopy [111, 112]. Fourth, *C. elegans* is a sophisticated multi-cellular animal. Despite the small number of cells, these cells generate a diverse set of organs and tissues such as an excretory system, gonads, a hypodermis, intestine, muscle, and a nervous system comprising 302 neurons. Last, numerous *C. elegans* genetic and genomic resources are available [113], and *C. elegans* was the first multi-cellular organism whose genome was fully sequenced [114], permitting its genomic and proteomic interrogation.

In a complex eukaryote like *Caenorhabditis elegans*, cell heterogeneity restricts the usefulness of large-scale, mass spectrometry-based proteomic analysis. Enriching for specific cells is challenging, and researchers cannot systematically identify low-abundance proteins expressed in specific cells from whole-organism lysates. Cell-selective bio-orthogonal non-canonical amino acid tagging (cell-selective BONCAT) offers a way to overcome these limitations [115, 116]. We have previously engineered a family of mutant *Escherichia coli* methionyl-tRNA synthetases (MetRSs) capable of appending the azide-bearing L-methionine (Met) analog L-azidonorleucine (Anl) to its cognate tRNA in competition with Met [87, 88]. Because Anl is a poor substrate

for any of the natural aminoacyl-tRNA synthetases, it is excluded from proteins made in wild-type cells but is incorporated readily into proteins made in cells that express an appropriately engineered MetRS. Controlling expression of mutant MetRSs by expression only in specific cells restricts Anl labeling to proteins produced in those cells. Tagged proteins can be distinguished from the rest of the protein pool through bio-orthogonal conjugation of the azide side chain to alkynyl or cyclooctynyl probes that permit facile detection, isolation, and visualization of labeled proteins. This strategy has been used to selectively enrich microbial proteins from mixtures of bacterial and mammalian cells. For example, Ngo et al. [89] found that proteins made in an *E. coli* strain outfitted with a mutant MetRS could be labeled with Anl in coculture with murine alveolar macrophages, which were not labeled. Using similar approaches, Grammel et al. [90] identified virulence factors from *Salmonella typhimurium* that were expressed in the course of infection of murine macrophages, and Mahdavi et al. [117] profiled *Yersinia enterocolitica* proteins that were injected into HeLa cells. In a complementary approach, Chin and coworkers [118] recently reengineered orthogonal *Methanosarcina barkeri* and *Methanosarcina mazei* pyrrolysyl-tRNA synthetase/tRNA pairs for codon-selective incorporation of a cyclopropene lysine derivative into proteins made in *E. coli*, *Drosophila melanogaster* ovaries, and HEK293 cells; however, this technique requires the expression of both exogenous aminoacyl-tRNA synthetases and tRNAs. Here, we configure cell-selective BONCAT for cell-specific proteomic analysis in the nematode *C. elegans* (**Figure 2.3**). We first demonstrate that restricted expression of a mutant *C. elegans* phenylalanyl-tRNA synthetase (*Ce*PheRS) can label proteins with *p*-azido-L-phenylalanine (Azf; **Figure 2.4**) with spatiotemporal selectivity in the live worm. We then show that cell-selective BONCAT combined with stable-isotope labeling of amino acids in cell culture (SILAC) provides a convenient and effective approach for unbiased discovery of proteins uniquely expressed in a subset of cells.

2.3 Results and Discussion

2.3.1 Engineering a C. elegans PheRS Capable of Activating Azf

We focused our attention on the heterotetrameric *Ce*PheRS because we could not prepare healthy transgenic *C. elegans* strains that express mutant *E. coli* MetRSs. Furthermore, we found that *C. elegans* variants of the mutant *E. coli* MetRSs that we had used to activate Anl in our previous experiments showed no activity toward Anl (**Table 2.1**). *Ce*PheRS catalyzes esterification of L-phenylalanine (Phe; **Figure 2.4**) to its cognate tRNA (CetRNA^{Phe}) to form phenylalanyl-tRNA. A conserved "gatekeeper" threonine (Thr412 [the first methionine in the alpha subunit of *Ce*PheRS, isoform A, exon 3 is designated as residue 1 ("Met1")], *C. elegans* numbering; **Figures 2.5 and 2.6**) in the alpha subunit has been proposed to play a key role in determining substrate specificity in both prokaryotic and eukaryotic PheRSs [119, 120]. Therefore, we hypothesized that mutating this residue to smaller residues should enable *Ce*PheRS to activate and charge the larger azide-bearing Phe analog Azf to CetRNA^{Phe}. To screen for such an enzyme, we cultured KY14[pKPY93/pKPY1XX], a phenylalanine-auxotrophic strain of *E. coli* that expresses mutant forms of *Ce*PheRS, in M9 minimal medium supplemented with different concentrations of Phe and Azf (**Figure 2.7**). To assess *Ce*PheRS activity toward Azf, we detected Azf-labeled proteins by conjugation to dibenzocyclooctyne-functionalized tetramethylrhodamine (TAMRA-DBCO; **Figure 2.4**) and subsequent SDS/PAGE-in-gel fluorescence scanning. Although several mutants (Thr412Ser, Thr412Ala, Thr412Gly) showed evidence of labeling with Azf, only the Thr412Gly mutant (Thr412Gly-*Ce*PheRS) displayed robust labeling in cells treated with equimolar amounts of Phe and Azf (**Figure 2.8**). We confirmed by *in vitro* ATP-PP$_i$ exchange assays that Thr412Gly-*Ce*PheRS is highly selective toward Azf: it activates Azf more than 20-fold faster than its canonical substrate Phe (**Table 2.2**). Thus, Azf labeling

does not require depletion of Phe from an animal's diet, making cell-selective labeling feasible in live worms. Although Thr412Gly-*Ce*PheRS also activates tryptophan threefold faster than Phe, MALDI-TOF mass spectrometry measurements of tryptic GFP peptides did not detect misincorporation of tryptophan or any other canonical amino acid when GFP was expressed in media supplemented with either Phe or Azf (**Figure 2.9**). Collectively, these results suggest that Thr412Gly-*Ce*PheRS selectively activates Azf with catalytic efficiency similar to that observed in the activation of Phe by wild-type *Ce*PheRS. Introducing the Thr412Gly mutation into PheRSs of other eukaryotic cells including human also permits Azf activation (**Figure 2.10**). From these observations, we conclude that Thr412Gly-*Ce*PheRS is the best aminoacyl-tRNA synthetase candidate for cell-selective labeling in *C. elegans*. Although we generated transgenic *C. elegans* by DNA injection into the syncytial germ line in this work, inducible or cell-selective genome editing technologies could be used to quickly and efficiently generate transgenic animals because a single mutation in *Ce*PheRS is sufficient for Azf activity [121].

2.3.2 Characterizing Azf Labeling in C. elegans

To evaluate the performance of Thr412Gly-*Ce*PheRS in *C. elegans*, we first generated transgenic *C. elegans* lines that express both the mutant alpha subunit and GFP under control of the *hsp-16.2* promoter. *hsp-16.2* encodes a 16-kDa protein that is induced in multiple tissues in response to heat shock and other stresses [122]. Upon heat shock, we expected that the mutant alpha subunit would be expressed and form a hybrid heterotetramer with the endogenous beta subunit to produce fully active Thr412Gly-*Ce*PheRS. In our initial feeding experiments, we did not detect strong Azf labeling in transgenic animals when Azf was added exogenously to either liquid culture or solid agar plates. However, we found that these lines could be labeled by replacing their normal food source (e.g., *E. coli* OP50) with bacteria whose

proteins contain the non-canonical amino acid of choice in a fashion analogous to isotopic labeling [30, 35, 36]. We first labeled bacteria by culturing KY14[pKPY514], a phenylalanine-auxotrophic strain of *E. coli* that expresses the *E. coli* variant of Thr412Gly-*Ce*PheRS (Thr251Gly-*Ec*PheRS), in M9 minimal medium supplemented with Azf (**Figure 2.11**). The extent of replacement of Phe by Azf in total *E. coli* protein was determined by amino acid analysis to be 50-51% (**Figure 2.12**). We next fed labeled bacteria to *hsp-16.2*::Thr412Gly-*Ce*PheRS worms previously grown on OP50. We induced heat shock in worms grown at 20°C by 1-h exposure to 33°C and cleared external as well as ingested bacteria by washing worms with S medium over a period of 30 min. We could not detect *E. coli* protein (even the overexpressed Thr251Gly-*Ec*PheRS) in processed *C. elegans* lysate (**Figure 2.13**). Both in-gel fluorescence scanning of TAMRA-DBCO-treated lysates and fluorescence microscopy of TAMRA-DBCO-treated fixed animals revealed that only heat-shocked worms exhibited ubiquitous Azf labeling 24 h after heat shock (**Figure 2.14**). We detected labeled proteins as early as 1 h after heat shock (shorter times were not tested). Moreover, we observed no differences in behavior, development, or survival in worms fed with Azf-labeled *E. coli* versus worms fed with unlabeled *E. coli* for up to 72 h after heat shock (longer times were not tested).

2.3.3 Labeling Spatially Defined Protein Subpopulations

The core concept of cell-selective BONCAT is that restricting expression of mutant aminoacyl-tRNA synthetases by using promoters active only in specific cells restricts non-canonical amino acid labeling to those cells. Enrichment of labeled proteins permits examination of proteomic changes in those cells. Encouraged by the performance of *hsp-16.2*::Thr412Gly-*Ce*PheRS worms, we next tested cell-specific expression of Thr412Gly-*Ce*PheRS by generating transgenic *C. elegans* lines that express both the mutant alpha subunit and GFP under control of promoters shown previously to be

active specifically in the 95 body wall muscle cells (*myo-3*; [123]), the 20 intestinal cells (*ges-1*; [124]), neurons (*rab-3*; [125]), and the 20 pharyngeal muscle cells (*myo-2*; [126]) (**Figure 2.15**). We first fed Azf-labeled bacteria to each of these animals. We also observed no differences in behavior, development, or survival in worms fed with Azf-labeled *E. coli* versus worms fed with unlabeled *E. coli* for up to 72 h (longer times were not tested). We then treated fixed worms with TAMRA-DBCO to visualize sites of Azf incorporation. Fluorescence microscopy revealed that labeling in *myo-3*::Thr412Gly-*Ce*PheRS, *ges-1*::Thr412Gly-*Ce*PheRS, *rab-3*::Thr412Gly-*Ce*PheRS, and *myo-2*::Thr412Gly-*Ce*PheRS worms was confined to the body wall muscle, intestine, neurons, and pharyngeal muscle, respectively (**Figures 2.16 and 2.17**).

2.3.4 Identifying Pharyngeal Muscle-Specific Proteins

We next investigated whether proteins isolated from worms with cell-specific Thr412-Gly-*Ce*PheRS fit characteristics of the targeted cell type. We were particularly interested in the *C. elegans* pharynx, a widely used model to study organ formation during embryogenesis [127]. The pharynx is a tube-like muscular pump that concentrates, grinds, and transports bacteria from the mouth to the intestine and comprises 68 cells: 9 epithelial, 4 gland, 9 marginal, 20 muscle, 20 neuronal, and 6 valve cells. We aimed to identify proteins expressed in pharyngeal muscle cells of *myo-2*::Thr412Gly-*Ce*PheRS worms by using a combined cell-selective BONCAT and SILAC approach. We first triply labeled food by culturing KY33[pKPY514], an arginine-, lysine-, and phenylalanine-auxotrophic strain of *E. coli*, in M9 minimal medium supplemented with "heavy" arginine ($^{13}C_6{}^{14}N_4{}^{1}H_{14}{}^{16}O_2$), heavy lysine ($^{13}C_6{}^{15}N_2{}^{1}H_{14}{}^{16}O_2$), and Azf. We next fed these bacteria to fourth larval stage *myo-2*::Thr412Gly-*Ce*PheRS animals previously grown on "light" OP50. According to our model of the cell-selective BONCAT method, all newly synthesized proteins in the animal should contain both

heavy arginine and heavy lysine, but only newly synthesized proteins made in pharyngeal muscle cells – the cells that express Thr412Gly-*Ce*PheRS – should contain Azf. This model gives rise to four classes of proteins: (i) light preexisting proteins synthesized in pharyngeal muscle cells before the shift in food source, (ii) light preexisting proteins synthesized in nontargeted cell types, (iii) heavy newly synthesized proteins – labeled with Azf – made in pharyngeal muscle cells after the shift in food source, and (iv) heavy newly synthesized proteins made in nontargeted cell types. The use of heavy labels enables relative quantitation of newly synthesized (heavy, H) and preexisting (light, L) peptides. Because enrichment concentrates heavy peptides derived from Azf-labeled proteins, cell selectivity can be quantified by comparing H/L ratios for "enriched" and "unenriched" proteins derived from the same worm sample. Proteins with the highest values of $(H/L)_E/(H/L)_U$ should have originated from pharyngeal muscle cells. Following 24 h of labeling through feeding, worm lysates were incubated with sodium dithionite-cleavable dibenzocyclooctyne-functionalized biotin (Diazo Biotin-DBCO; **Figure 2.4**), and Azf-containing proteins were isolated by streptavidin affinity chromatography (**Figures 2.18, 2.19, and 2.20**). Enriched and unenriched samples were resolved by SDS/PAGE and subjected to in-gel proteolytic digestion and liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis.

We identified and quantified 2,270 proteins across triplicate paired (enriched and unenriched) experiments (**Figures 2.21, 2.22, 2.23, and 2.24 and Table 2.3**). Of the quantified proteins, 1,607 (71%) had $(H/L)_E/(H/L)_U$ values greater than one, which indicates that the purification method successfully enriched newly synthesized, cell-specific proteins. Among the enriched proteins were 782 proteins that have expression patterns reported in the literature according to WormBase WS244 [128]; of these proteins, 409 are known to be expressed in the pharynx (**Table 2.4**). We expected that proteins expressed in pharyngeal muscle cells would be overrepresented

among proteins with high $(H/L)_E/(H/L)_U$ values and, indeed, found that of the top 12 proteins quantified, two [TNI-4 [129], a troponin I protein and TNC-2 [130], a troponin C protein] are expressed exclusively in pharyngeal muscle cells. Two proteins [SHL-1 [131], a voltage-gated potassium channel and NCX-2 [132], a sodium-calcium exchanger] are expressed in many muscle cells including pharyngeal muscle cells. A fifth protein F59F4.1 [133] is an acyl-CoA oxidase that is also expressed in the pharynx. Also, three well-known pharyngeal muscle-specific myosin heavy chains were among the top 3% of most highly enriched proteins: MYO-1 (top 1.0%, 22/2,270), MYO-2 (top 1.3%, 29/2,270), and MYO-5 (top 2.6%, 58/2,270). Although the right tail of the $(H/L)_E/(H/L)_U$ distribution contains relatively few members, they represent pharynx-specific proteins: of the 1,100 proteins in our dataset with known expression patterns, 7 of the 18 that are thought to be expressed exclusively in the pharynx have $(H/L)_E/(H/L)_U$ values greater than two ($P = 1.25\ x\ 10^8$; Fisher's exact test).

Three of the remaining seven "top-12" proteins in our dataset (C53C9.2, K03E5.2, and CPN-4) share similarity with Calponin-1 [134], a human protein implicated in the regulation of smooth muscle contraction, but their expression patterns have not been reported. To determine whether they are expressed in pharyngeal muscle cells, we generated transgenic *C. elegans* lines that express GFP under control of each of the respective 5' regulatory regions. We detected strong GFP fluorescence exclusively in pharyngeal muscle cells in *C53C9.2*::*gfp*, *K03E5.2*::*gfp*, and *cpn-4*::*gfp* animals (**Figure 2.25**). Calponin-1 has a single calponin homology (CH) and multiple calponin family repeat (CFR) domains. Notably, like its muscle-specific paralog CPN-3, CPN-4 has a CH domain, but no CFRs (**Figure 2.26**). In contrast, C53C9.2 and K03E5.2 have multiple CFRs but no CH domains. Only four *C. elegans* proteins contain multiple CFRs: C53C9.2, K03E5.2, T25F10.6, and UNC-87 [135], a protein required to maintain structure of myofilaments in muscle cells. T25F10.6 (top 5.3%,

121/2,270) and UNC-87 (top 7.5%, 170/2,270) were also among the top 10% of most highly enriched proteins. Although the characterization of new pharyngeal proteins was beyond the scope of this work, their placement among highly enriched proteins, localization, and similarity to other muscle-specific proteins suggest that C53C9.2, K03E5.2, and CPN-4 are excellent candidates for regulating aspects of pharyngeal muscle biology.

Finally, we asked whether our approach undesirably identified any off-target or "non-pharyngeal" proteins. Miller and coworkers used epitope-tagged (FLAG) polyA-binding protein (PAB-1) under the control of cell-specific promoters to isolate RNA from specific larval cells [136]. This work had larval data sets for body wall muscle, intestine, and neurons, but not pharyngeal muscle. For each body wall muscle, intestine, and neurons dataset, we examined the top 12 most enriched genes versus a reference sample of all cells that also has protein abundance information in PaxDb [137] and therefore had been identified in previous mass spectrometry experiments. SUR-5 appeared in all three lists because SUR-5 is highly expressed in body wall muscle, intestine, and neurons [138]. In this list of 34 "non-pharyngeal" proteins, only three appeared in our dataset: SUR-5 ($(\mathrm{H/L})_E/(\mathrm{H/L})_U = 1.21$), ELPC-3 ($(\mathrm{H/L})_E/(\mathrm{H/L})_U = 0.77$), and C24H12.4 ($(\mathrm{H/L})_E/(\mathrm{H/L})_U = 0.97$). Upon further examination, SUR-5 and ELPC-3 have been found to be also expressed in the pharynx in GFP fusion experiments [138, 139]. C24H12.4 has no expression data reported in the literature. Although the absence of a protein from a proteomic dataset cannot be taken as evidence that the protein is absent from the sample, this result is consistent with the hypothesis that the method described here provides an effective means of enriching pharyngeal proteins.

Together, these results demonstrate that the approach described here can be used to identify proteins (including proteins with previously unknown expression patterns) that are expressed in targeted subsets of cells. We note that in a cell-selective BONCAT experiment, proteins are labeled only after the shift in food source. Although long labeling times can be used to profile the majority of proteins in specific cells, short labeling times can be used to capture rapid changes in protein expression in those cells.

2.4 Conclusions

In summary, by using cell-specific promoters to drive expression of an engineered *Ce*PheRS, we demonstrated that cell-selective BONCAT coupled with SILAC can be used to identify proteins with spatiotemporal selectivity in living *C. elegans*. For future experiments, we suggest several avenues for improvement. First, we incubated worm lysates with Diazo Biotin-DBCO and isolated Azf-labeled proteins by streptavidin affinity chromatography. Because streptavidin affinity chromatography requires mild washing conditions to preserve streptavidin's bioactivity, enrichment quality might be affected by background proteins due to insufficient washing. Alternatively, Azf-labeled proteins can be selectively captured on commercially available alkynyl- or cyclooctynyl-functionalized resins that allow for highly stringent washing conditions to remove nonspecifically bound proteins. Second, we processed samples by SDS/PAGE and in-gel proteolytic digestion before LC-MS/MS analysis, but this approach can introduce contaminants and is time-consuming and laborious. Because chemical tagging of Azf-labeled proteins occurs immediately after animal lysis, our strategy is compatible with new advances in analytical proteomic workflows such as in StageTip-based filter-aided sample preparation (FASP) [140]. Third, in the LC-MS/MS analysis, we normalized H/L ratios of enriched proteins to total unenriched proteins derived from the same worm sample. Because the H/L ratio variability

depends on different intrinsic rates of protein synthesis in different cells, we advise investigators to additionally normalize H/L ratios of enriched proteins from one cell type to total proteins derived from mixed-stage worms or enriched proteins from another cell type for a more comprehensive analysis of cell-specific proteins.

Finally, our methodology should prove useful in multiple contexts. For example, one could easily build cell-specific proteomic atlases because (i) a catalog of cell-specific transcriptional enhancers is readily available and (ii) the creation of transgenic organisms is both rapid and routine. In addition, using regulatory elements to drive intersectional patterns of expression, one could restrict labeling to cells that express both elements and, thus, enhance spatiotemporal selectivity with either a FLP recombinase-based [141] or protein reconstitution-based [94] approach. Furthermore, this technique could be used to study protein-protein interactions in a cell-specific manner because aryl azides like Azf are activated upon UV light irradiation to form covalent adducts with proteins in close proximity [142]. In principle, the methodology described here could be applied to other organisms in which efficient delivery of non-canonical amino acids is feasible, alleviating the need for cell sorting or laser capture techniques to isolate protein from specific cells in intact organisms.

2.5 Figures

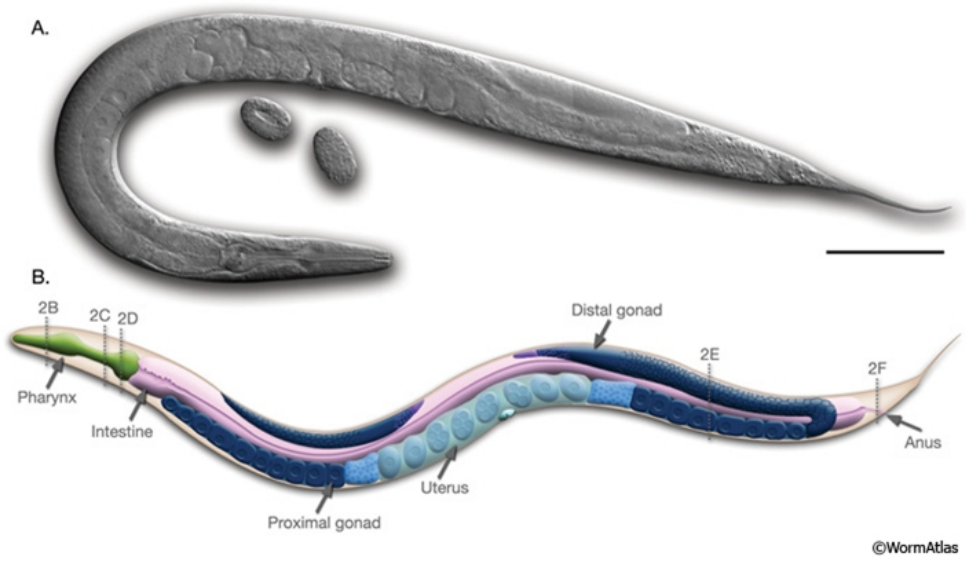


Figure 2.1: *(Top) Differential interference contrast (DIC) micrograph of an adult C. elegans hermaphrodite (Scale Bar: 0.1 mm). (Bottom) Schematic drawing of anatomical structures of an adult hermaphrodite (left lateral side).*

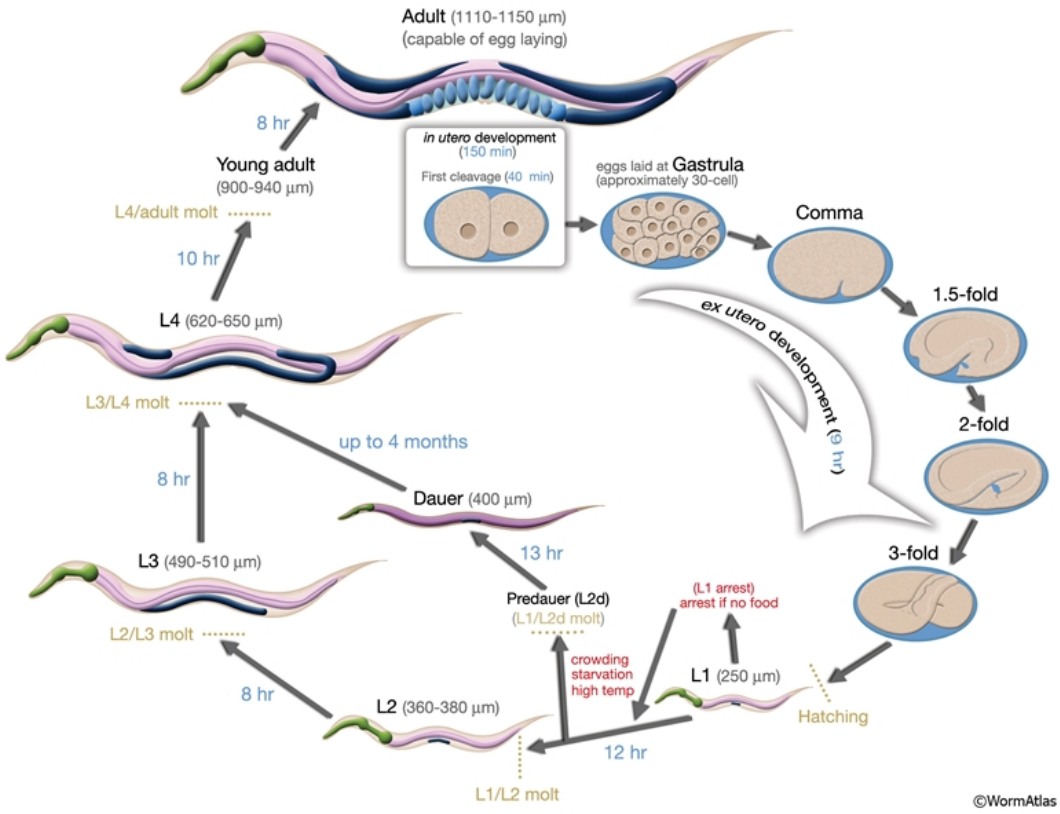


Figure 2.2: *Life cycle of C. elegans at 22° C*

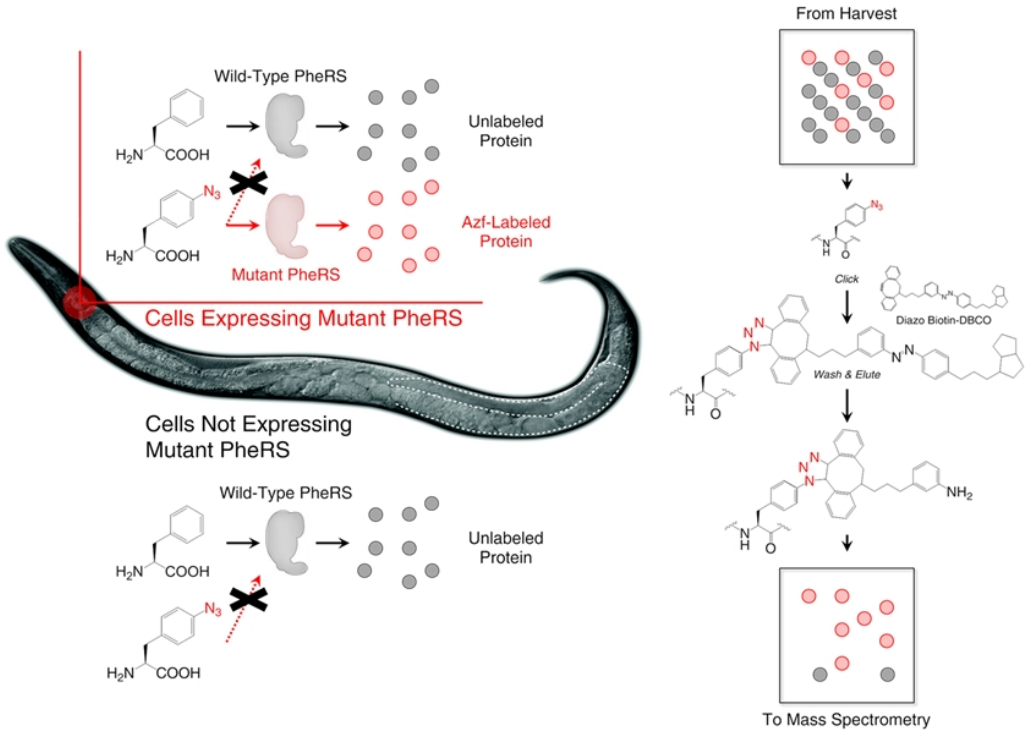


Figure 2.3: *A mutant C. elegans PheRS is capable of tagging proteins with the reactive non-canonical amino acid Azf. Spatiotemporal selectivity is achieved by controlling expression of the mutant synthetase using cell-selective promoters in transgenic lines. Tagged proteins are distinguished from other proteins through conjugation of the azide side chain to probes that permit isolation of the labeled proteins. Worm lysates are reacted with sodium dithionite-cleavable Diazo Biotin-DBCO, and Azf-labeled proteins are isolated by streptavidin affinity chromatography. The structure of the biotin reagent has been simplified; the full structure is shown in* ***Figure 2.4***

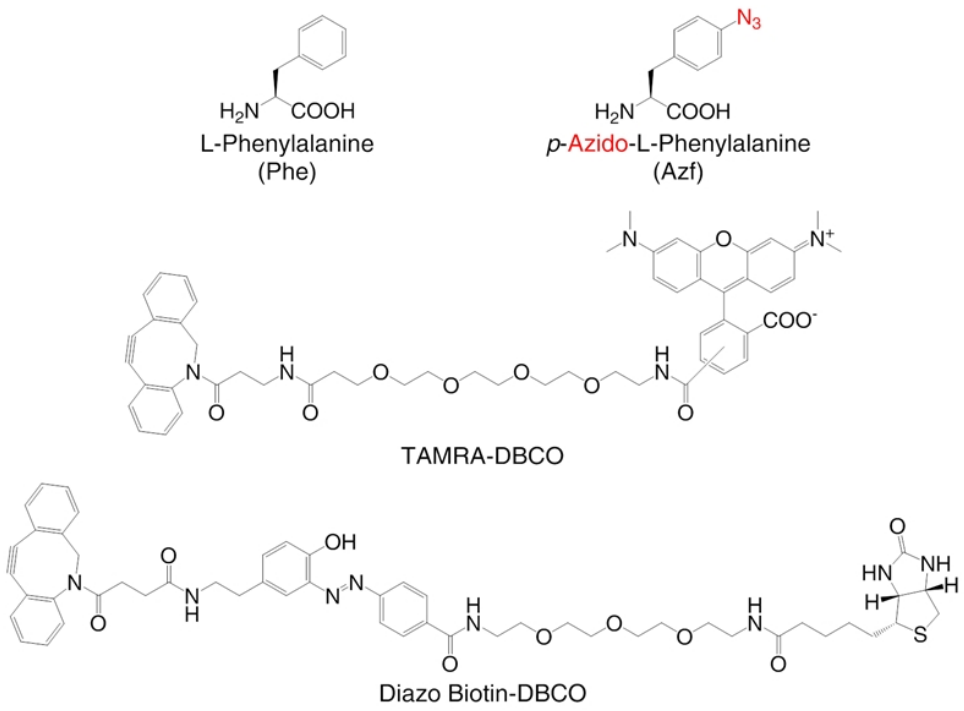


Figure 2.4: *Structures of amino acids and probes used in Chapter 2: L-phenylalanine (Phe), p-azido-L-phenylalanine (Azf), dibenzocyclooctyne tetramethylrhodamine (TAMRA-DBCO), and sodium dithionite-cleavable dibenzocyclooctyne biotin (Diazo Biotin-DBCO).*

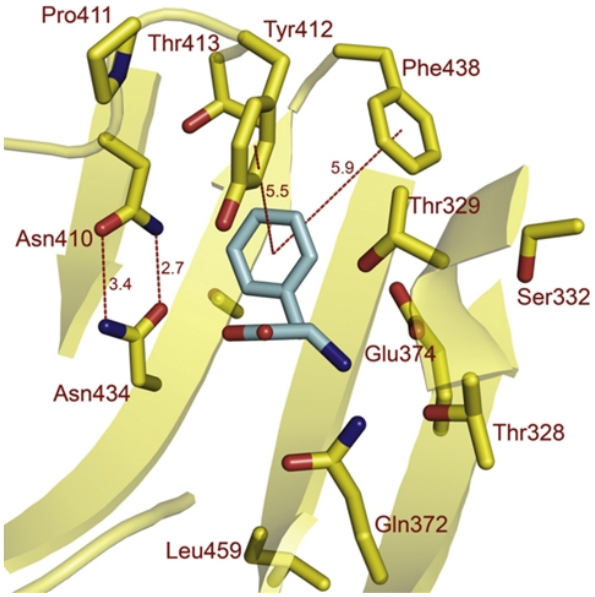


Figure 2.5: *Active site of H. sapiens PheRS with bound Phe (PDB ID code 3L4G).*

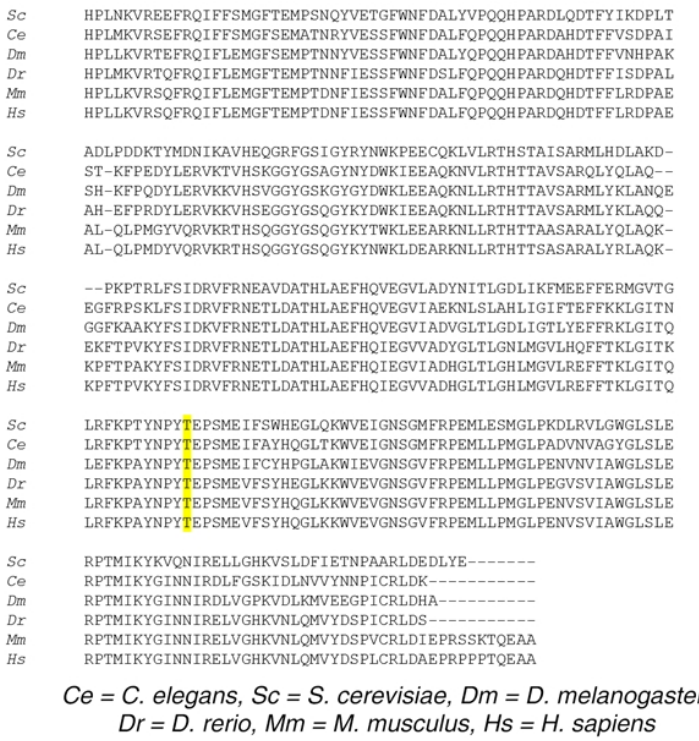


Figure 2.6: *Alignment of S. cerevisiae (Sc, His228-Glu503), C. elegans (Ce, His225-Lys496), D. melanogaster (Dm, His224-Ala498), D. rerio (Dr, His225-Ser497), M. musculus (Mm, His225-Ala508), and H. sapiens (Hs, His225-Ala508) PheRS catalytic cores. A conserved "gatekeeper" threonine (highlighted in yellow) in the alpha subunit has been proposed to play a key role in determining substrate specificity in PheRSs.*

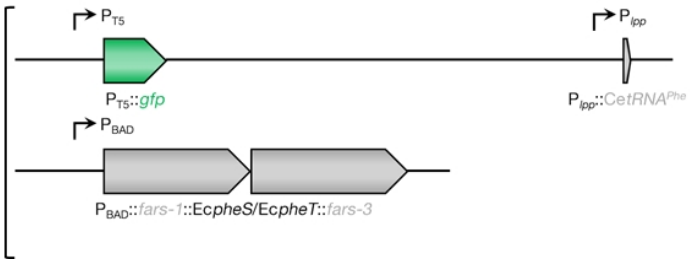


Figure 2.7: *To screen for a mutant CePheRS that activates Azf, we used KY14[pKPY93/pKPY1XX], a phenylalanine-auxotrophic strain of E. coli. This strain houses two compatible plasmids: (i) pKPY93 encodes CetRNA^{Phe} under constitutive E. coli murein lipoprotein (lpp) promoter control and IPTG-inducible (P_{T5}) 6xHis-tagged GFP and (ii) pKPY1XX encodes both CePheRS alpha and beta subunits under arabinose-inducible (P_{BAD}) control.*

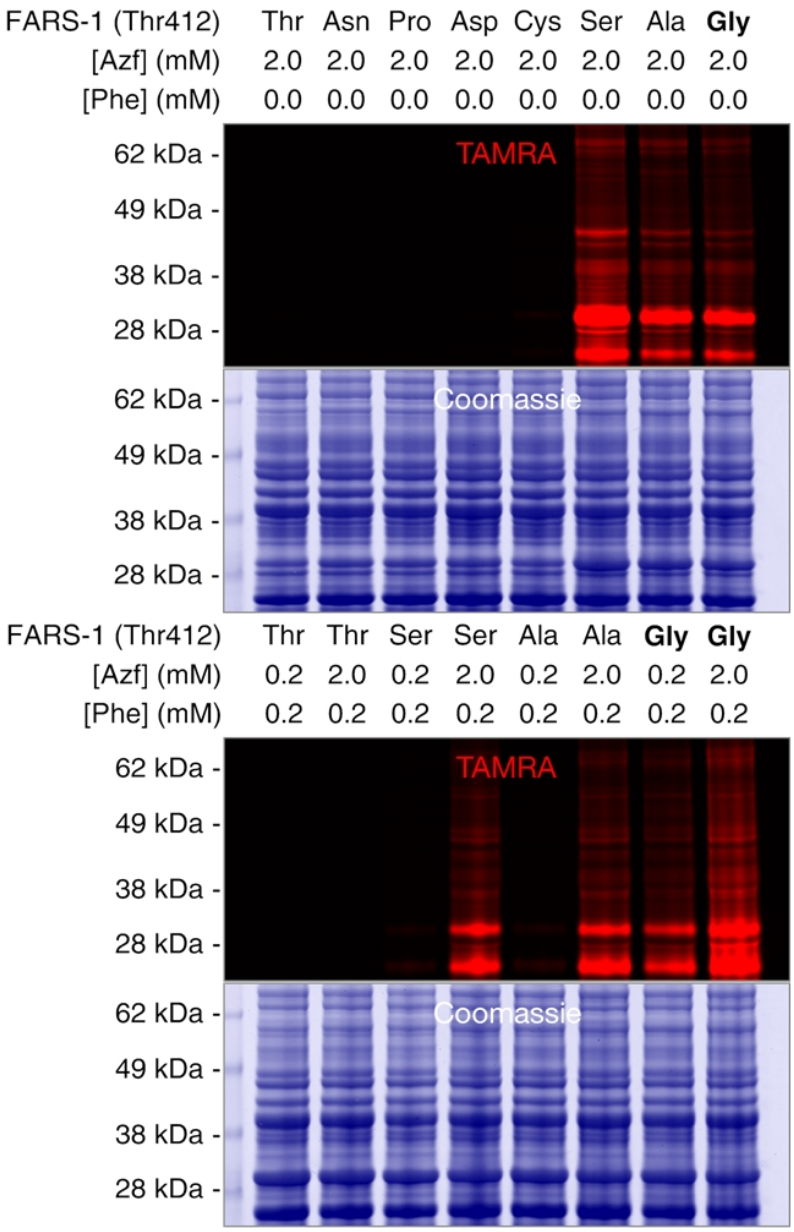


Figure 2.8: *SDS/PAGE and in-gel fluorescence scanning detection of conjugation of TAMRA-DBCO to Azf-labeled proteins derived from KY14[pKPY93/pKPY1XX] lysates.*

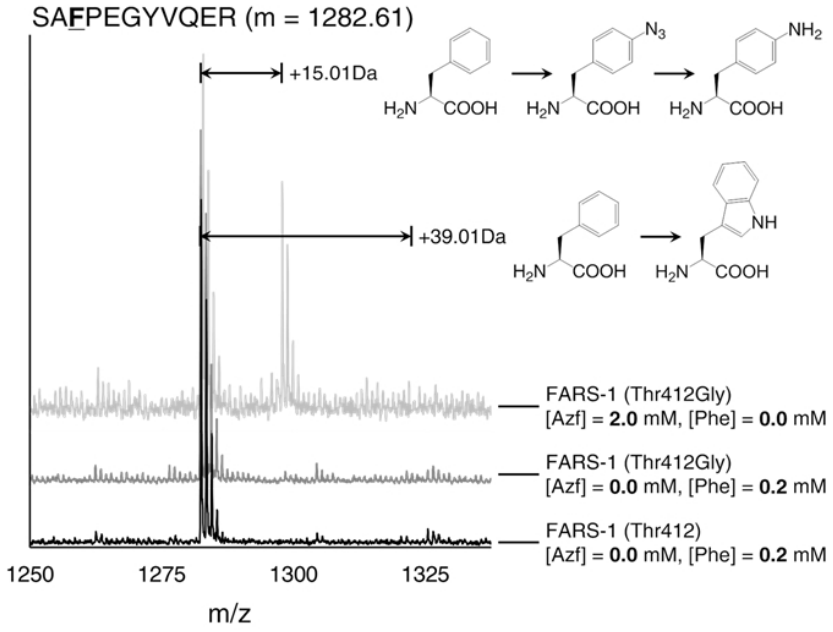


Figure 2.9: *MALDI-TOF mass spectra of the tryptic GFP peptide SAFPEGYVQER. 6xHis-tagged GFP was purified from either KY14[pKPY93/pKPY100] or KY14[pKPY93/pKPY102]. GFP was expressed in cells bearing i) Thr412Gly-CePheRS and treated with 2.0 mM Azf and 0.0 mM Phe, ii) Thr412Gly-CePheRS, treated with 0.0 mM Azf and 0.2 mM Phe or iii) wild-type CePheRS with 0.0 mM Azf and 0.2 mM Phe. Azf was reduced to p-amino-L-phenylalanine during the experiment. Therefore, incorporation of Azf in this peptide results in a +15.01 Da m/z shift. Misincorporation of L-tryptophan in this peptide results in a +39.01 Da m/z shift.*

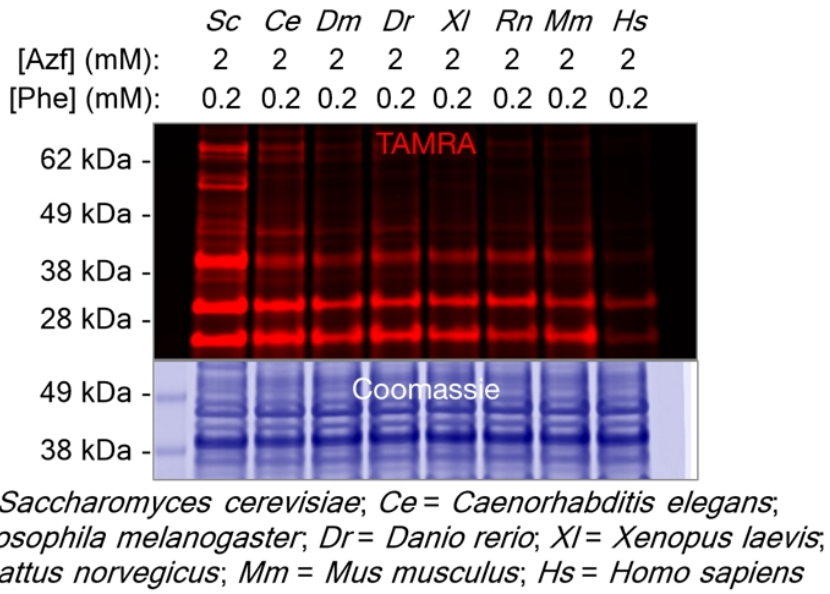


Figure 2.10: *SDS-PAGE and in-gel fluorescence scanning detection of conjugation of TAMRA-DBCO to Azf-labeled protein from KY14 containing pKPY93 and pKPY509, pKPY102, pKPY503, pKPY505, pKPY647, pKPY655, pKPY507, or pKPY137 lysates. These strains house two compatible plasmids: i) pKPY93 encodes CetRNA^{Phe} under constitutive E. coli murein lipoprotein (lpp) promoter control as well as IPTG-inducible (P_{T5}) 6xHis-tagged GFP and ii) pKPY509, pKPY102, pKPY503, pKPY505, pKPY647, pKPY655, pKPY507, and pKPY137 are plasmids that encode Saccharomyces cerevisiae (Sc), Caenorhabditis elegans (Ce), Drosophila melanogaster (Dm), Danio rerio (Dr), Xenopus laevis (Xl), Rattus norvegicus (Rn), Mus musculus (Mm), and Homo sapiens (Hs), respectively, PheRS alpha and beta subunits under arabinose-inducible (P_{BAD}) control.*

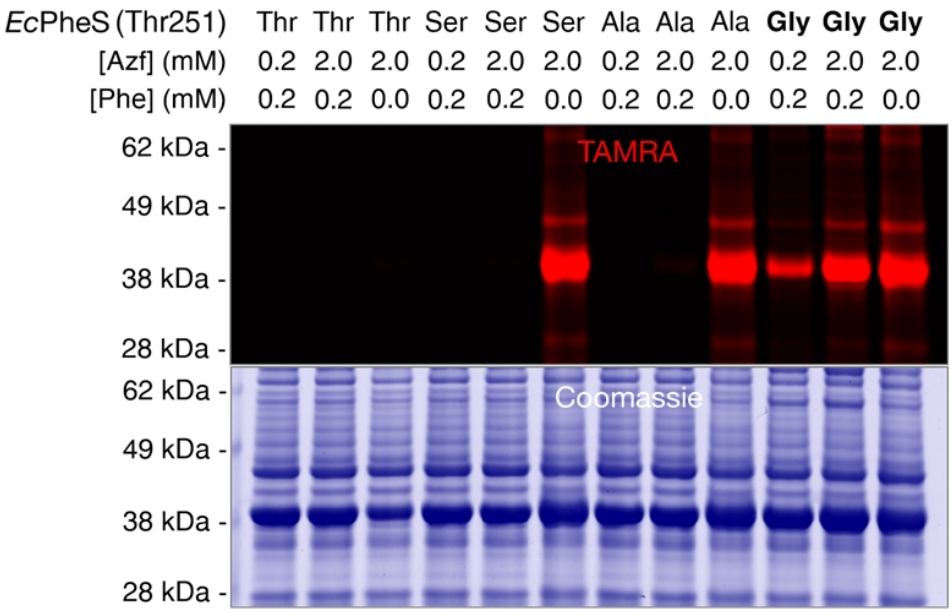


Figure 2.11: *SDS-PAGE and in-gel fluorescence scanning detection of conjugation of TAMRA-DBCO to Azf-labeled proteins from KY14 containing pKPY513, pKPY566, pKPY565, or pKPY514 lysates. pKPY513, pKPY566, pKPY565, and pKPY514 encode the IPTG-inducible (PT5) 6xHis-tagged E. coli variants of wild-type, Thr412Ser, Thr412Ala, or Thr412Gly-CePheRS, respectively.*

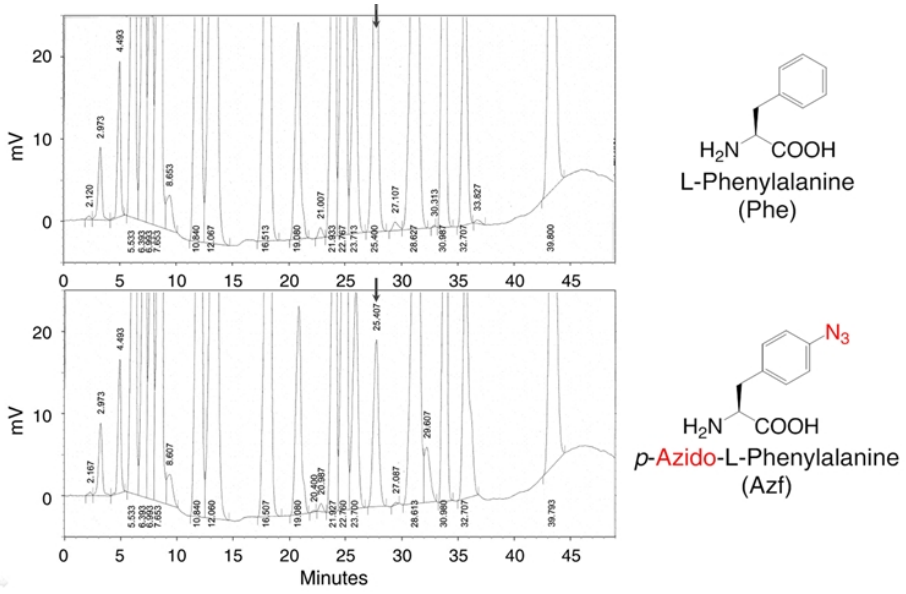


Figure 2.12: *Amino acid analysis of the extent of replacement of Phe by Azf in total E. coli protein of KY14[pKPY514]. Cells bearing Thr251Gly-EcPheRS were treated with i) 0.0 mM Azf and 0.2 mM Phe or ii) 2.0 mM Azf and 0.0 mM Phe. The extent of replacement of Phe by Azf was calculated from the change in the ratio of Phe to other amino acids in cells treated with only Phe versus the ratio of Phe to other amino acids in cells treated with only Azf. Phe elutes around 25.4 minutes as indicated by the black arrow.*

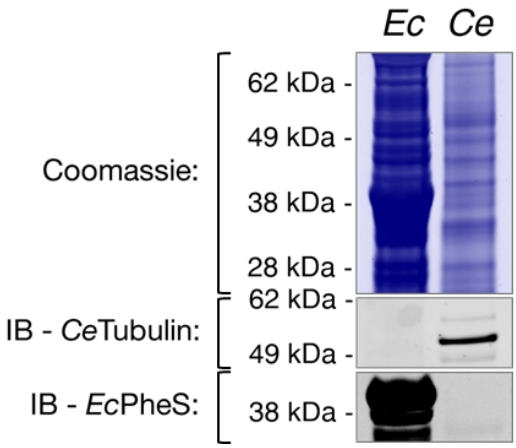


Figure 2.13: *We fed Azf-labeled bacteria (left lane) to hsp-16.2::Thr 412Gly-CePheRS worms previously grown on OP50 and cleared digested bacteria by washing worms with S medium. We could not detect E. coli protein (even the overexpressed Thr251Gly-EcPheRS) in processed C. elegans lysate (right lane). Tubulin is a marker for C. elegans proteins.*

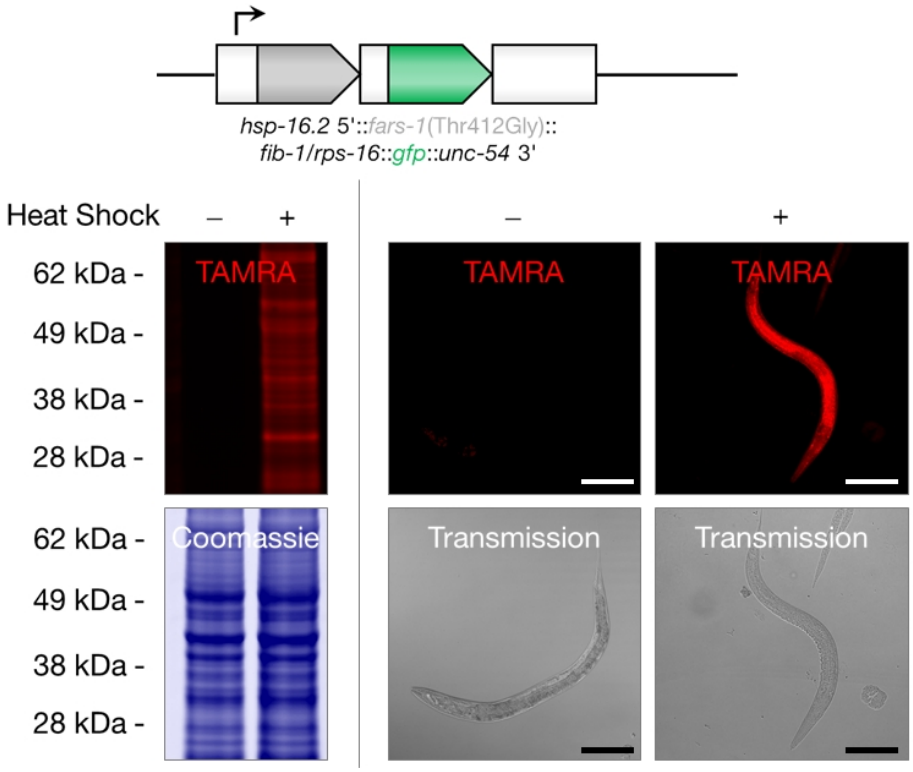


Figure 2.14: *We fed KY14[pKPY514] to hsp-16.2::Thr412Gly-CePheRS animals. This E. coli strain harbors pKPY514, a plasmid encoding Thr251Gly-EcPheRS under IPTG-inducible (P_{T5}) control. hsp-16.2::Thr412Gly-CePheRS C. elegans express both the mutant alpha subunit and GFP under control of the hsp-16.2 promoter. In-gel fluorescence scanning of TAMRA-DBCO-treated lysates (Left) and fluorescence microscopy of TAMRA-DBCO-treated fixed animals (Right) showed that only heat-shocked worms exhibited ubiquitous Azf labeling 24 h after heat shock. (Scale Bars: 100 μm.)*

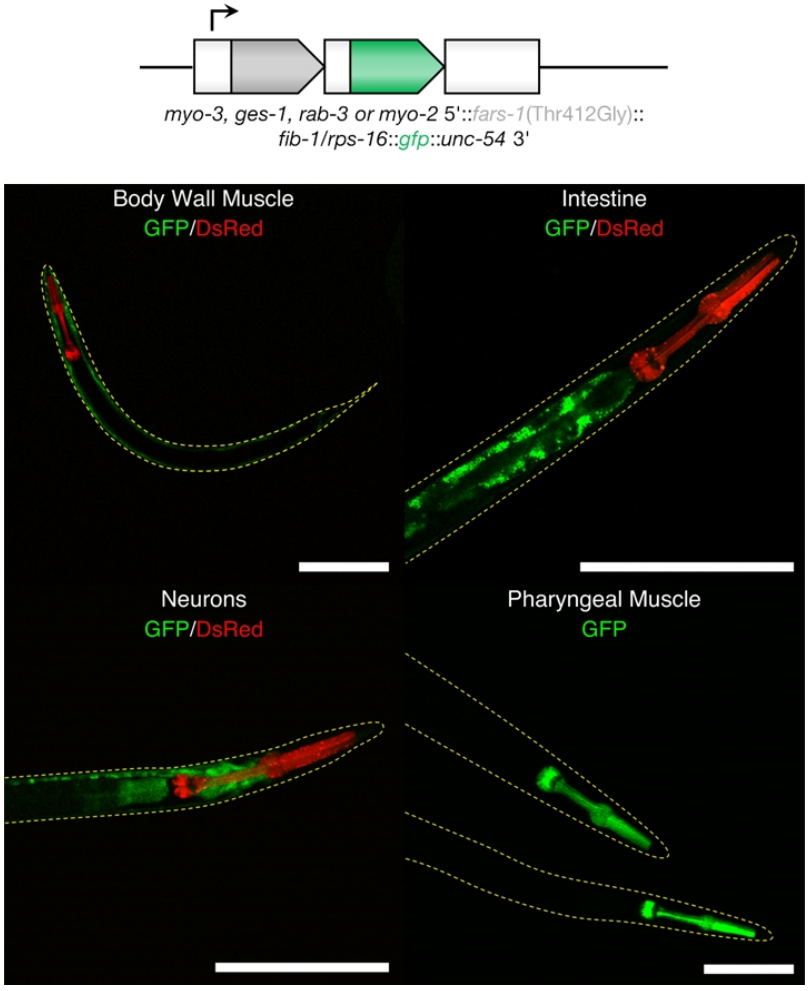


Figure 2.15: *We fed E. coli strain KY14[pKPY514] to myo-3:: (Top Left), ges-1:: (Top Right), rab-3:: (Bottom Left), and myo-2::Thr412Gly-CePheRS (Bottom Right) animals. These transgenic C. elegans strains express both the mutant PheRS alpha subunit and GFP (green) under control of the myo-3, ges-1, rab-3, or myo-2 promoters, respectively. myo-3, ges-1, and rab-3 worms also express the co-injection marker myo-2::dsRed which marks the pharyngeal muscle cells (red). (Scale Bar: 100 µm.)*

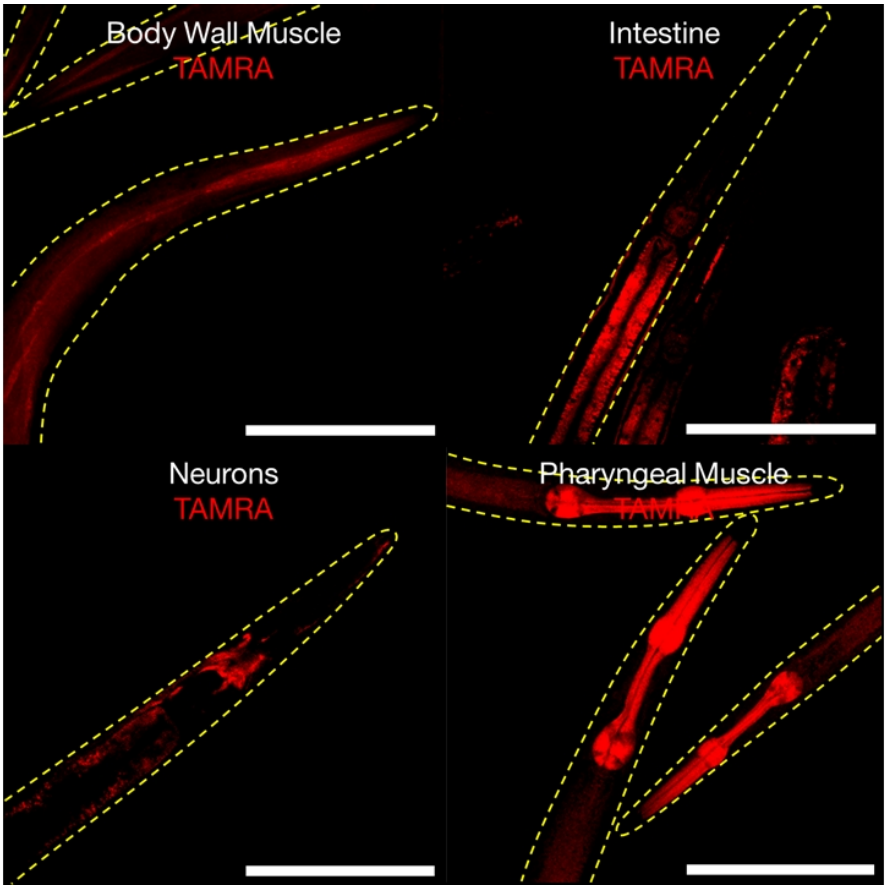


Figure 2.16: *Fluorescence microscopy revealed that labeling in myo-3 (Left), ges-1 (Center Left), rab-3 (Center Right), and myo-2::Thr412Gly-CePheRS (Right) worms was localized to body wall muscle, intestine, neurons, and pharyngeal muscle, respectively; the surrounding tissues were not labeled. (Scale Bar: 100 µm.)*

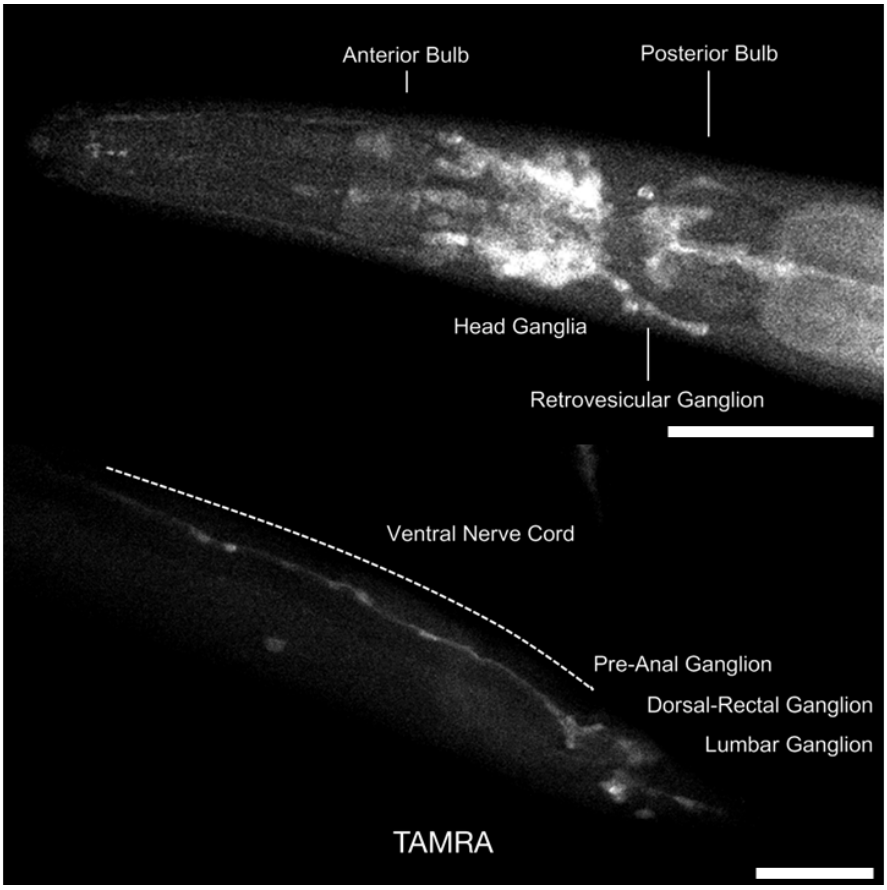


Figure 2.17: *Fluorescence microscopy revealed that worms with rab-3-driven expression of Thr412Gly-CePheRS exhibit selective Azf-labeling in neuronal cell-types. (Scale Bar: 50 μm.)*

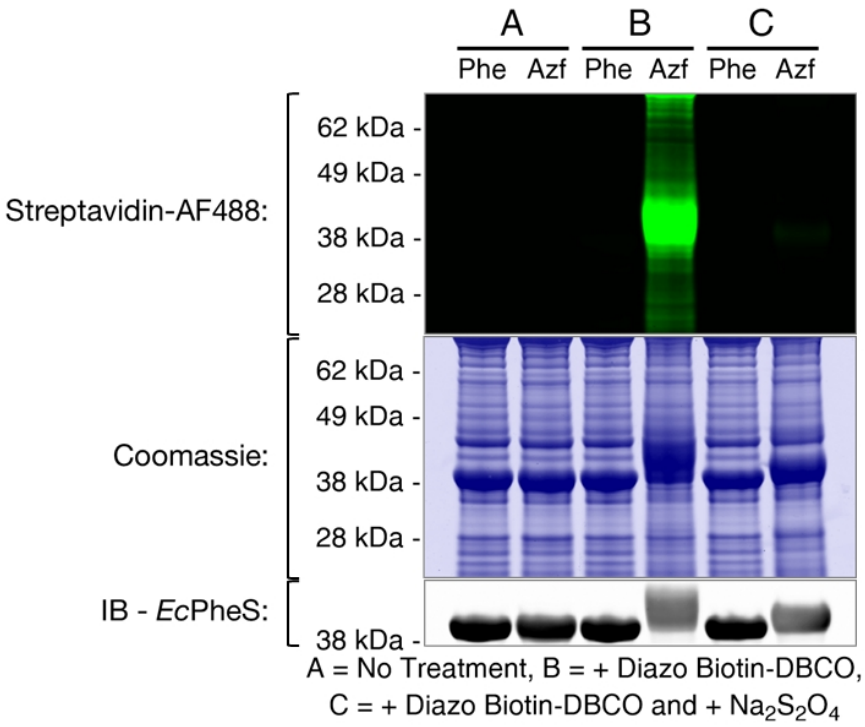


Figure 2.18: *Model labeling of KY14[pKPY514] lysates with Diazo Biotin-DBCO. First, cells bearing Thr251Gly-EcPheRS were treated with i) 0.0 mM Azf and 0.2 mM Phe (lanes marked with "Phe") or ii) 2.0 mM Azf and 0.0 mM Phe (lanes marked with "Azf"). Next, proteins were treated with water ("A" lanes) or Diazo Biotin-DBCO ("B" lanes). Diazo Biotin-DBCO-modified proteins were treated with sodium dithionite ("C" lanes). SDS-PAGE, Western blotting, and in-gel fluorescence scanning detection of Streptavidin-Alexa Fluor 488 (Streptavidin-AF488) revealed that only Azf-labeled proteins are biotinylated after treatment with Diazo Biotin-DBCO and that biotin is removed by treatment with sodium dithionite.*

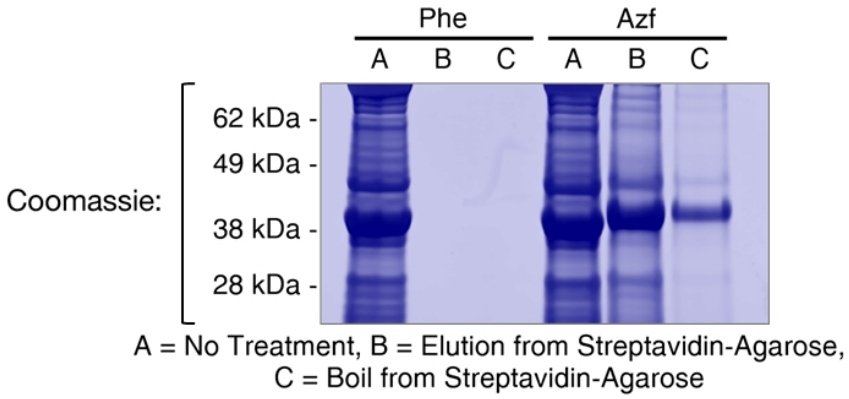


Figure 2.19: *Model enrichment of KY14[pKPY514] lysates with Diazo Biotin-DBCO. First, cells bearing Thr251Gly-EcPheRS were treated with i) 0.0 mM Azf and 0.2 mM Phe (lanes marked "Phe") or ii) 2.0 mM Azf and 0.0 mM Phe (lanes marked "Azf"). Next, proteins were treated with Diazo Biotin-DBCO and incubated with Streptavidin agarose resin. Diazo Biotin-DBCO-modified proteins were eluted by treatment with sodium dithionite. Proteins that remained bound to resin were released after boiling in the presence of SDS. SDS-PAGE revealed that only Azf-labeled proteins are eluted after treatment with sodium dithionite. Unmodified lysates are in "A" lanes, proteins eluted after treatment with sodium dithionite are in "B" lanes, and proteins eluted after boiling are in "C" lanes.*

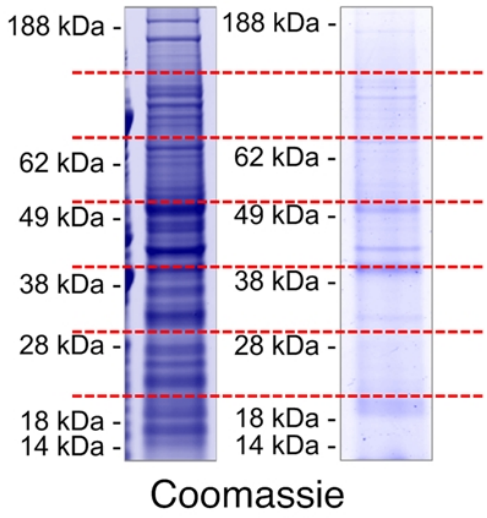


Figure 2.20: *Unenriched (Left) and enriched (Right) samples prepared from labeled myo-2::Thr412Gly-CePheRS worms were subjected to in-gel proteolytic digestion and LC-MS/MS. After SDS-PAGE, gels were sliced as indicated by the dotted red lines.*

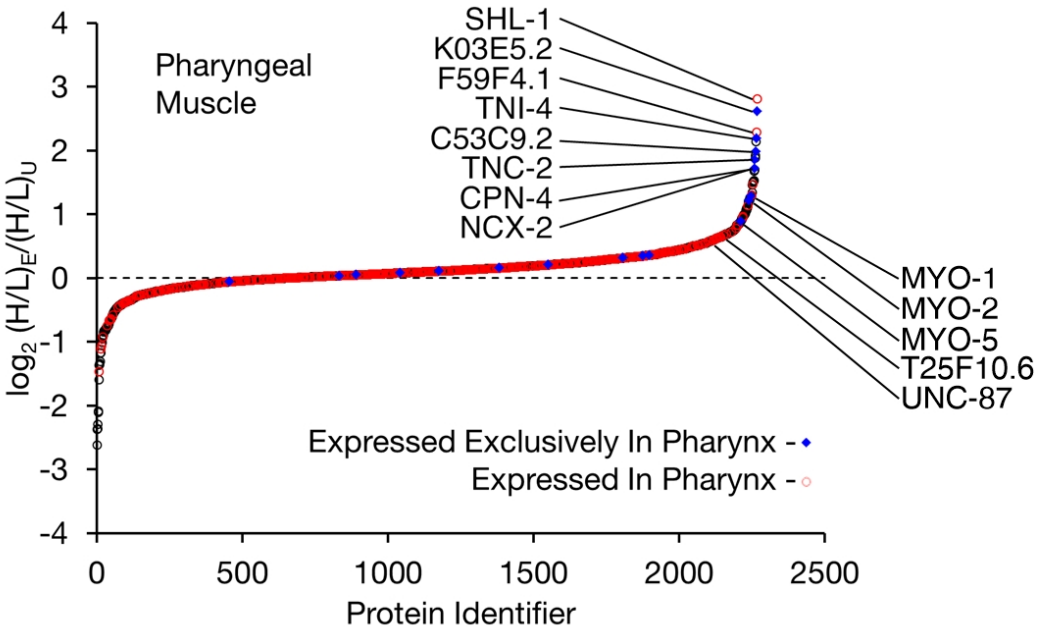


Figure 2.21: *LC-MS/MS analysis of myo-2::Thr412Gly-CePheRS animals fed with triply labeled E. coli (heavy arginine, heavy lysine, and Azf) for 24 h. Solid blue diamond markers represent proteins that are expressed exclusively in the pharynx. Open red circle markers represent proteins that are expressed in the pharynx. Open black circle markers represent proteins that have either unknown expression patterns or expression patterns not associated with the pharynx. The proteins discussed in this work are marked.*

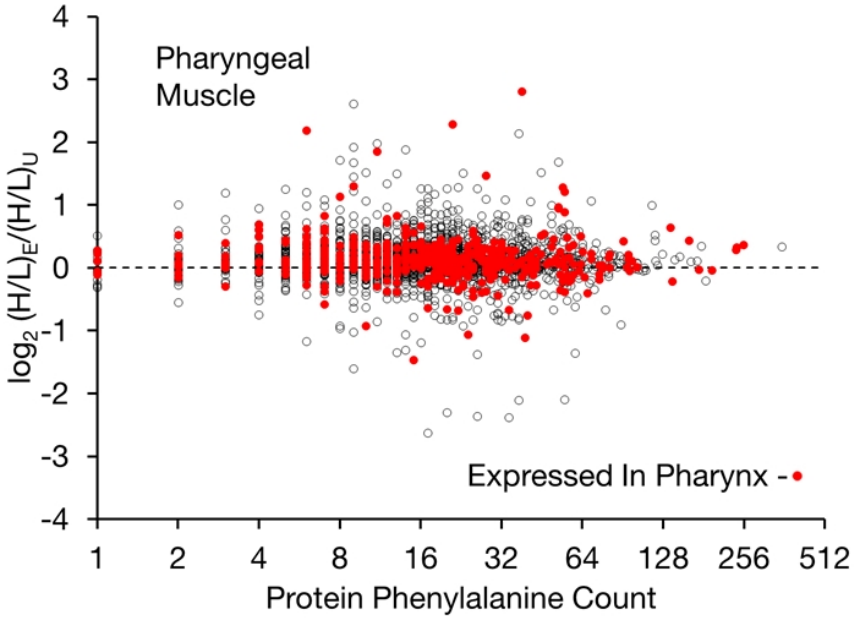


Figure 2.22: *LC-MS/MS analysis of myo-2::Thr412Gly-CePheRS animals fed with triply labeled E. coli ("heavy" arginine, "heavy" lysine, and Azf) for 24 hours: $(H/L)_E/(H/L)_U$ values versus phenylalanine counts. Solid red markers represent proteins that are expressed in the pharynx. Open black markers represent proteins that have either unknown expression patterns or expression patterns not associated with the pharynx. The number of phenylalanine residues in a protein does not influence the extent of enrichment.*

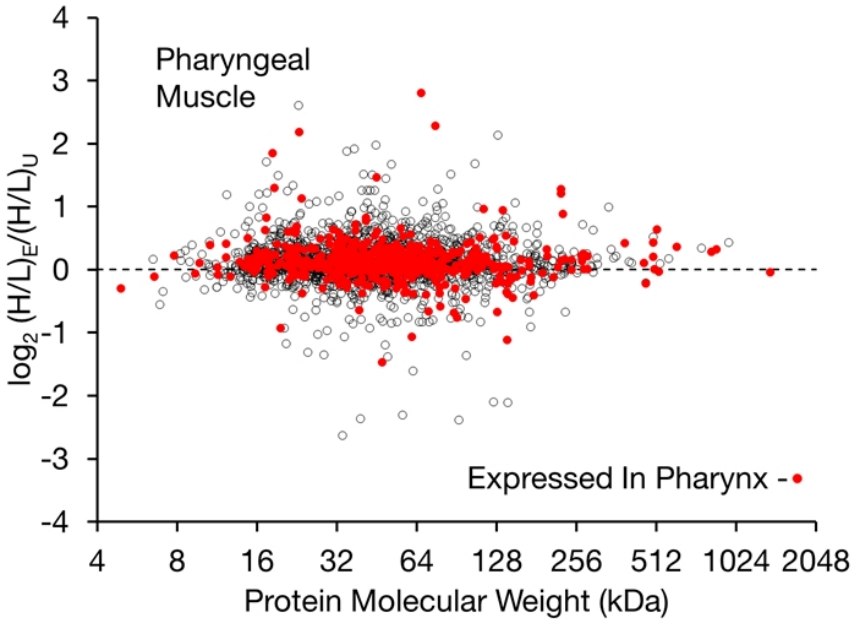


Figure 2.23: *LC-MS/MS analysis of myo-2::Thr412Gly-CePheRS animals fed with triply labeled E. coli ("heavy" arginine, "heavy" lysine, and Azf) for 24 hours: $(H/L)_E/(H/L)_U$ values versus molecular weights. Solid red markers represent proteins that are expressed in the pharynx. Open black markers represent proteins that have either unknown expression patterns or expression patterns not associated with the pharynx. The protein size does not influence the extent of enrichment.*

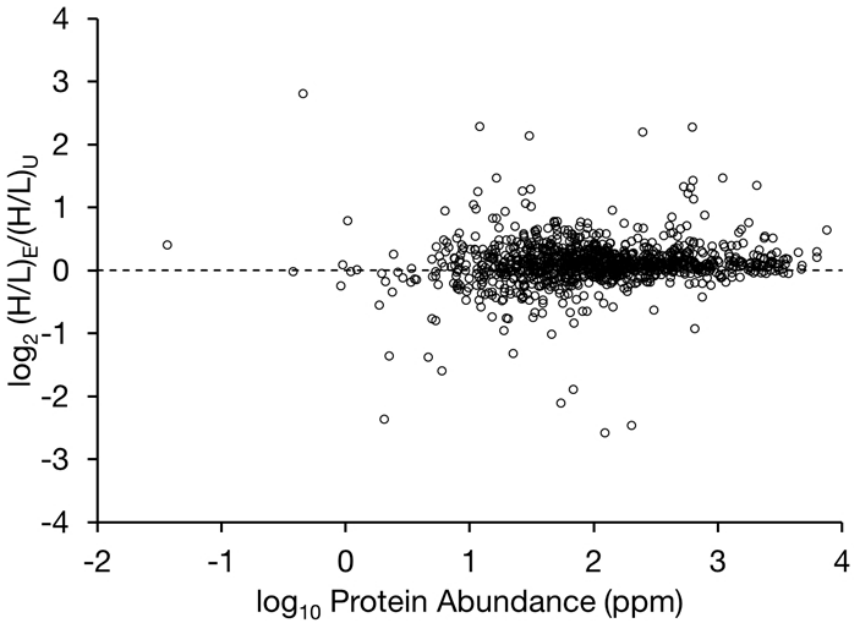


Figure 2.24: *LC-MS/MS analysis of myo-2::Thr412Gly-CePheRS animals fed with triply labeled E. coli ("heavy" arginine, "heavy" lysine, and Azf) for 24 hours: $(H/L)_E/(H/L)_U$ values versus protein abundance values obtained from the PaxDb database (Wang, M., et al., Mol. Cell. Proteomics **11**, 492–500 (2012)). Protein abundance does not influence the extent of enrichment.*

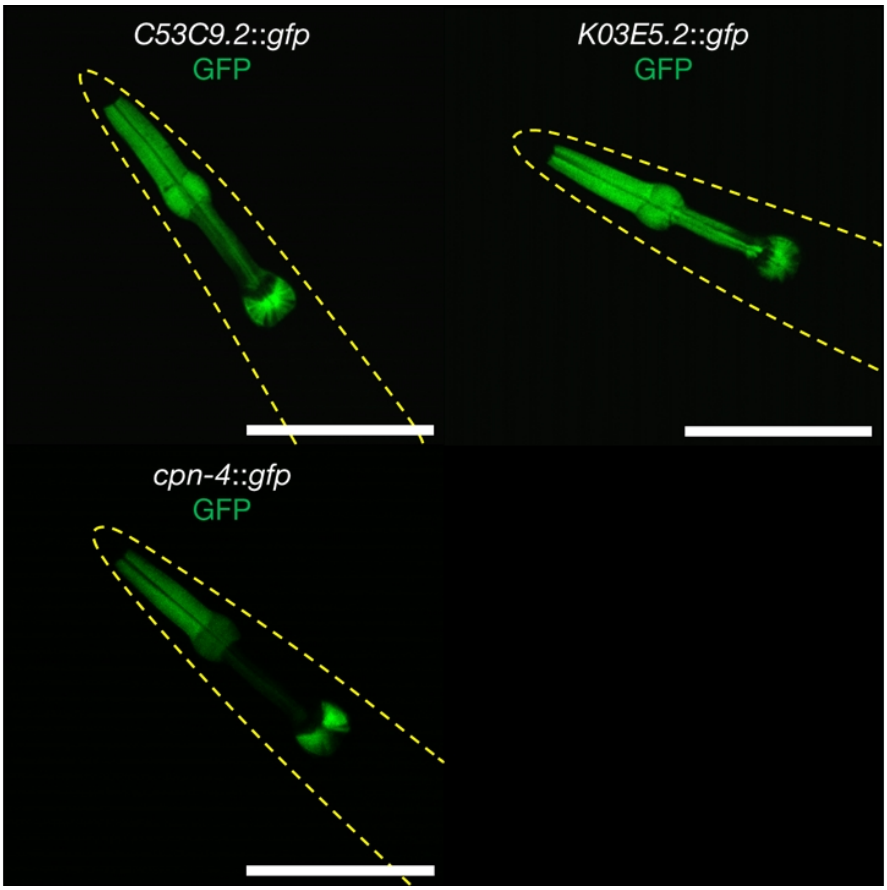


Figure 2.25: *Fluorescence microscopy confirmed GFP fluorescence was exclusively localized to pharyngeal muscle cells in C53C9.2::gfp (Top Left), K03E5.2::gfp (Top Right), and cpn-4::gfp (Bottom Left) animals. (Scale bars: 100 μm.)*

Figure 2.26: *Schematic diagram of human Calponin-1 (Accession Number P51911), C. elegans proteins CPN-4, C53C9.2, K03E5.2 (Isoform A), T25F10.6 (Isoform A), and UNC-87 (Isoform A). Proteins and their domains are drawn in proportion to their length.*

2.6 Tables

Substrate	K_m (10^{-6} M)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
C. elegans Methionyl-tRNA Synthetase			
Met	27.9 ± 3.3	4.0 ± 0.5	140000 ± 24000
Mutant *C. elegans* Methionyl-tRNA Synthetases[1]			
Met	N.D.	N.D.	N.D.
Anl	N.D.	N.D.	N.D.

Table 2.1: *Kinetic parameters (determined by in vitro ATP-PP_i exchange) for activation of amino acids by wild-type and mutant C. elegans methionyl-tRNA synthetases (CeMetRS) (Met: L-Methionine, Anl: L-Azidonorleucine; N.D.: Not Detected).* [1]*Mutant CeMetRS: Leu42Gly; Leu42Cys, Tyr297Leu, His333Leu; Leu42Asn, Tyr297Leu, His333Leu; Leu42Pro, Tyr297Leu, His333Leu; Leu42 Ser, Tyr297Leu, His333Leu.*

Substrate	K_m (10^{-6} M)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)	Selectivity[1]	Relative Activity[2]
C. elegans Phenylalanyl-tRNA Synthetase					
Phe	0.972 ± 0.116	0.301 ± 0.027	312000 ± 46500	-	1
C. elegans Phenylalanyl-tRNA Synthetase (Thr412Gly)					
Phe	39.9 ± 4.1	0.327 ± 0.034	8280 ± 1220	1	1/37.7
Azf	1.95 ± 0.34	0.327 ± 0.026	172000 ± 33000	20.8	1/1.8
Trp	20.6 ± 1.6	0.507 ± 0.039	24600 ± 2690	3.0	1/12.7
Tyr	N.D.	N.D.	N.D.	-	-

Table 2.2: *Kinetic parameters (determined by in vitro ATP-PP_i exchange) for activation of amino acids by wild-type CePheRS and Thr412Gly-CePheRS (Phe: L-Phenylalanine, Azf: p-Azido-L-Phenylalanine, Trp: L-Tryptophan, Tyr: L-Tyrosine; N.D.: Not Detected).* [1]*Selectivity is defined as $(k_{cat}/K_m)_{xxx}$ of substrates of Thr412Gly-CePheRS divided by $(k_{cat}/K_m)_{Phe}$ of Thr412Gly-CePheRS.* [2]*Relative Activity is defined as $(k_{cat}/K_m)_{xxx}$ of substrates of Thr412Gly-CePheRS divided by $(k_{cat}/K_m)_{Phe}$ of wild-type CePheRS.*

Protein	$\log_2$ (H/L)E/(H/L)U
Q95XD1|Q95XD1_CAEEL Protein SHL-1 OS=Caenorhabditis elegans GN=shl-1 PE=4 SV=2	2.81
O61848|O61848_CAEEL Protein K03E5.2, isoform a OS=Caenorhabditis elegans GN=CELE_K03E5.2 PE=2 SV=2; H2KZW0|H2KZW0_CAEEL Protein K03E5.2, isoform b OS=Caenorhabditis elegans GN=CELE_K03E5.2 PE=4 SV=1; H2KZW1|H2KZW1_CAEEL Protein K03E5.2, isoform d	2.61
Q93839|Q93839_CAEEL Acyl-coenzyme A oxidase OS=Caenorhabditis elegans GN=CELE_F59F4.1 PE=3 SV=2	2.29
O44572|TNNI4_CAEEL Troponin I 4 OS=Caenorhabditis elegans GN=tni-4 PE=2 SV=2	2.19
Q18529|Q18529_CAEEL Protein C39D10.7 OS=Caenorhabditis elegans GN=C39D10.7 PE=1 SV=3	2.14
Q09936|YSE2_CAEEL Uncharacterized protein C53C9.2 OS=Caenorhabditis elegans GN=C53C9.2 PE=3 SV=2	1.98
O45577|O45577_CAEEL Protein CPF-2 OS=Caenorhabditis elegans GN=cpf-2 PE=4 SV=1	1.93
P90879|P90879_CAEEL Protein F49C12.9 OS=Caenorhabditis elegans GN=CELE_F49C12.9 PE=1 SV=1	1.88
Q09665|TNNC2_CAEEL Troponin C, isoform 2 OS=Caenorhabditis elegans GN=tnc-2 PE=2 SV=1	1.85
O44727|O44727_CAEEL Protein CPN-4 OS=Caenorhabditis elegans GN=cpn-4 PE=4 SV=3	1.72

Table 2.3: *Proteins identified and quantified from LC-MS/MS analysis of myo-2::Thr412Gly-CePheRS animals fed with triply labeled E. coli ("heavy" arginine, "heavy" lysine, and Azf) for 24 hours. The full table is available as a PDF in the online version of this thesis.*

Protein	$\log_2 (H/L)_E/(H/L)_U$
Q95XD1|Q95XD1_CAEEL Protein SHL-1 OS=Caenorhabditis elegans GN=shl-1 PE=4 SV=2	2.81
Q93839|Q93839_CAEEL Acyl-coenzyme A oxidase OS=Caenorhabditis elegans GN=CELE_F59F4.1 PE=3 SV=2	2.29
O44572|TNNI4_CAEEL Troponin I 4 OS=Caenorhabditis elegans GN=tni-4 PE=2 SV=2	2.19
Q09665|TNNC2_CAEEL Troponin C, isoform 2 OS=Caenorhabditis elegans GN=tnc-2 PE=2 SV=1	1.85
H2KYP4|H2KYP4_CAEEL Protein NCX-2, isoform d OS=Caenorhabditis elegans GN=ncx-2 PE=4 SV=1; H2KYP2|H2KYP2_CAEEL Protein NCX-2, isoform a OS=Caenorhabditis elegans GN=ncx-2 PE=4 SV=1	1.69
Q9U3N4|INX6_CAEEL Innexin-6 OS=Caenorhabditis elegans GN=inx-6 PE=2 SV=1	1.47
P19626|MLR2_CAEEL Myosin regulatory light chain 2 OS=Caenorhabditis elegans GN=mlc-2 PE=1 SV=1; P19625|MLR1_CAEEL Myosin regulatory light chain 1 OS=Caenorhabditis elegans GN=mlc-1 PE=4 SV=1	1.30
P02567|MYO1_CAEEL Myosin-1 OS=Caenorhabditis elegans GN=let-75 PE=2 SV=3	1.28
P12845|MYO2_CAEEL Myosin-2 OS=Caenorhabditis elegans GN=myo-2 PE=2 SV=2	1.21
Q86GU1|Q86GU1_CAEEL Protein HSP-25, isoform b OS=Caenorhabditis elegans GN=hsp-25 PE=2 SV=1; Q17849|Q17849_CAEEL Protein HSP-25, isoform a OS=Caenorhabditis elegans GN=hsp-25 PE=2 SV=1; Q5H9M9|Q5H9M9_CAEEL Protein HSP-25, isoform c OS=Caenorhabdi	1.14

Table 2.4: *Proteins previously known to be expressed in the pharynx from LC-MS/MS analysis of myo-2::Thr412Gly-CePheRS animals fed with triply labeled E. coli ("heavy" arginine, "heavy" lysine, and Azf) for 24 hours. The full table is available as a PDF in the online version of this thesis.*

Body Wall Muscle	Intestine	Neurons
pgp-8	*T08B6.2*	*Y48G10A.6*
C37H5.5	*K10D11.6*	*C09B9.1*
sur-5	*flr-4*	*sur-5*
elpc-3	*fbxa-51*	*daf-11*
sdz-3	*sur-5*	*nlp-5*
unc-93	*inx-17*	*K07C11.10*
T28D9.7	*F01D5.3*	*ZC13.2*
F46C5.6	*clec-166*	*Y42A5A.1*
ugt-39	*Y22D7AL.15*	*nhr-167*
Y48G1A.4	*R05D8.9*	*gpa-2*
sul-1	*clec-66*	*pde-5*
C24H12.4	*ZC196.5*	*fkh-8*

Table 2.5: *We examined a test set of 34 proteins whose genes are highly expressed in body wall muscle, intestinal epithelia, and neuronal cells (Spencer, W.C., et al., Genome Res.* ***21****, 325–341 (2011)).*

2.7 Materials and Methods

2.7.1 ATP-PP$_i$ Exchange Assay

Aminoacylating Solution: In double-distilled water, 50 mM HEPES (Sigma-Aldrich Corporation), pH 7.6, 20 mM magnesium chloride (VWR International), 1 mM DL-dithiothreitol (Sigma-Aldrich Corporation), 2 mM adenosine 5'-triphosphate disodium salt hydrate (Sigma-Aldrich Corporation), 2 mM sodium pyrophosphate (VWR International), 0.1 mg/mL bovine serum albumin (Sigma-Aldrich Corporation), 0.14% (v/v) 2-mercaptoethanol (Sigma-Aldrich Corporation), 3 μCi/200 μL [^{32}P]-sodium pyrophosphate (PerkinElmer)

Quenching Solution: In double-distilled water, 7% (v/v) perchloric acid (EMD Millipore), 3% (w/v) activated charcoal (Sigma-Aldrich Corporation), 200 mM sodium pyrophosphate (VWR International)

Washing Solution: In double-distilled water, 0.5% (v/v) perchloric acid (EMD Millipore), 10 mM sodium pyrophosphate (VWR International)

Wild-type *Ce*PheRS and Thr412Gly-*Ce*PheRS were expressed as described in the **Labeling in *E. coli*** procedure and purified according to the **Isolating 6xHis-Tagged Proteins** procedure. Protein concentrations of purified proteins were determined by absorbance at 280 nm. To assay activation of L-phenylalanine (Sigma-Aldrich Corporation) by wild-type *Ce*PheRS, wild-type *Ce*PheRS (500 nM) and varying concentrations of L-phenylalanine (0.1 to 50 μM) were mixed together in Aminoacylating Solution and incubated for 16 minutes at room temperature. To assay activation of L-phenylalanine by Thr412Gly-*Ce*PheRS, Thr412Gly-*Ce*PheRS (500 nM) and varying concentrations of L-phenylalanine (1 to 500 μM) were mixed together in Aminoacylating Solution and incubated for 16 minutes at room temperature. To assay activation of *p*-azido-L-phenylalanine (Chem-Impex International) by Thr412Gly-*Ce*PheRS, Thr412Gly-*Ce*PheRS (500 nM) and varying concentrations

of *p*-azido-L-phenylalanine (0.1 to 50 μM) were mixed together in Aminoacylating Solution and incubated for 16 minutes at room temperature. To assay activation of L-tryptophan (Sigma-Aldrich Corporation) by Thr412Gly-*Ce*PheRS, Thr412Gly-*Ce*PheRS (500 nM) and varying concentrations of L-tryptophan (1 to 500 μM) were mixed together in Aminoacylating Solution and incubated for 16 minutes at room temperature. To assay activation of L-tyrosine (Sigma-Aldrich Corporation) by Thr412Gly-*Ce*PheRS, Thr412Gly-*Ce*PheRS (500 nM) and varying concentrations of L-tyrosine (1 to 500 μM) were mixed together in Aminoacylating Solution and incubated for 16 minutes at room temperature. Every four minutes, 0.20 volume was removed from the mixtures and blended with charcoal-containing Quenching Solution (2.50 volumes). The charcoal was centrifuged at 4,000 x *g* for five minutes at room temperature, and the top aqueous layer was discarded. Washing Solution (2.50 volumes) was added and mixed. The charcoal was centrifuged at 4,000 x *g* for five minutes at room temperature. The last three steps were repeated twice. The charcoal was transferred to scintillation vials containing Safety-Solve Complete Counting Cocktail (50.00 volumes, Research Products International) and counted with a Liquid Scintillation Analyzer Tri-Carb B2910TR (PerkinElmer) at the Caltech Environmental Health and Safety (EH&S) Office.

2.7.2 Chloroform/Methanol Precipitation

Methanol (EMD Millipore, 2.40 volumes) was added to 1.00 volume of protein solution and vortexed. Chloroform (EMD Millipore, 0.80 volume) was added and vortexed. Double-distilled water (3.20 volumes) was added and vortexed. The mixture was centrifuged at 20,000 x *g* for 20 minutes at room temperature, and the top layer was discarded. Methanol (5.00 volumes) was added and vortexed, and the mixture was centrifuged at 20,000 x *g* for five minutes at room temperature. The last three steps were repeated for a total of five times. The top layer was discarded, and the protein

pellet was air-dried.

2.7.3 Enrichment of Azf-Labeled Proteins

Diazo Biotin-DBCO Solution: 1 mM Diazo Biotin-DBCO (Click Chemistry Tools) in dimethyl sulfoxide, anhydrous (Life Technologies)

Binding/Washing Solution: In double-distilled water, 200 mM Trizma base (Sigma-Aldrich Corporation), 150 mM sodium chloride (VWR International), 0.8% (w/v) sodium dodecyl sulfate (Sigma-Aldrich Corporation), pH 8.0

Eluting Solution: In double-distilled water, 200 mM Trizma base (Sigma-Aldrich Corporation), 150 mM sodium chloride (VWR International), 0.8% (w/v) sodium dodecyl sulfate (Sigma-Aldrich Corporation), pH 8.0, 50 mM sodium dithionite (Sigma-Aldrich Corporation)

Lysing Solution A: In double-distilled water, 200 mM Trizma base (Sigma-Aldrich Corporation), 4% (w/v) sodium dodecyl sulfate (Sigma-Aldrich Corporation), pH 8.0

Lysates (1.00 volume) were incubated with Diazo Biotin-DBCO Solution (0.11 volume) for 20 minutes at room temperature. Excess Diazo Biotin-DBCO was removed by the **Chloroform/Methanol Precipitation** procedure. Protein pellets were re-suspended in Lysing Solution A (0.55 volume). To the protein solutions, Pierce Streptavidin UltraLink Resin (8.33 volumes, 10 μL re-suspended resin/mg protein in Binding/Wash Solution, Thermo Fisher Scientific) was added. The mixtures were mildly agitated for five hours at room temperature, and centrifuged at 1,000 x g for five minutes at room temperature. The top aqueous layer was discarded. To deplete unlabeled proteins, Binding/Wash Solution (75.00 volumes) was added and vortexed. The mixtures were centrifuged at 1,000 x g for five minutes at room temperature. The last three steps were repeated for a total of ten times. To elute bound labeled proteins, the top aqueous layer was discarded, and freshly prepared Eluting Solution (5.55 volumes) was added and mildly agitated for 20 minutes at room temperature.

The mixtures were centrifuged at 1,000 x g for five minutes at room temperature, and the top aqueous layer was removed and saved. Freshly prepared Eluting Solution (5.55 volumes) was added and mildly agitated for 20 minutes at room temperature. The last three steps were repeated for a total of five times. Pooled protein samples were concentrated and buffer exchanged with Lysing Solution A using an Amicon Ultra-0.5 Centrifugal Filter Unit with Ultracel-10 Membrane (EMD Millipore) according to the manufacturer's instructions. Samples were protected from light at all times.

2.7.4 Fluorescence Microscopy of Live C. elegans

Anesthetizing Solution: In double-distilled water, 6.8 g/L sodium phosphate dibasic anhydrous (VWR International), 3 g/L potassium phosphate monobasic (VWR International), 0.5 g/L sodium chloride (VWR International), 1 g/L ammonium chloride (VWR International), 2 mM magnesium sulfate heptahydrate (VWR International), 0.1 mM calcium chloride dihydrate (VWR International), 20 mM D-glucose (VWR International), 25 mM sodium azide (Sigma-Aldrich Corporation)

C. elegans were re-suspended in Anesthetizing Solution. Anesthetized worms were mounted on 3% (w/v) UltraPure Low Melting Point Agarose (Life Technologies) pads and imaged with a LSM 510 META microscope (Carl Zeiss AG) at the Biological Imaging Center in the Beckman Institute at Caltech. Samples were protected from light at all times.

2.7.5 Fluorescence Microscopy of Fixed C. elegans

Alkylating Solution: In double-distilled water, 100 mM Trizma base (Sigma-Aldrich Corporation), pH 8.0, 100 mM 2-chloroacetamide (Sigma-Aldrich Corporation)

C. elegans Fixing Solution: 160 mM potassium chloride (VWR International), 40 mM sodium chloride (VWR International), 20 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid tetrasodium salt (Sigma-Aldrich Corporation), 10 mM spermidine trihydrochloride (Sigma-Aldrich Corporation), 30 mM PIPES, pH 7.4, 50% (v/v) methanol (EMD Millipore), 1.6% (w/v) paraformaldehyde (Electron Microscopy Sciences)

Permeabilizing Solution A: In 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (Life Technologies), 1% (v/v) Triton X-100 (Sigma-Aldrich Corporation), 0.1% (w/v) bovine serum albumin (Sigma-Aldrich Corporation), 0.05% (w/v) sodium azide (Sigma-Aldrich Corporation)

Permeabilizing Solution B: In 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (Life Technologies), 10% (v/v) dimethyl sulfoxide (VWR International), 1% (v/v) Triton X-100 (Sigma-Aldrich Corporation), 0.1% (w/v) bovine serum albumin (Sigma-Aldrich Corporation), 0.05% (w/v) sodium azide (Sigma-Aldrich Corporation)

Reducing Solution: In double-distilled water, 100 mM Trizma base (Sigma-Aldrich Corporation), pH 8.0, 10 mM tris(2-carboxyethyl)phosphine hydrochloride (Thermo Fisher Scientific)

TAMRA-DBCO Solution: 1 mM TAMRA-DBCO (Click Chemistry Tools) in dimethyl sulfoxide, anhydrous (Life Technologies)

Washing Solution A: In double-distilled water, 100 mM Trizma base (Sigma-Aldrich Corporation), pH 8.0

Washing Solution B: In 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (Life Technologies), 20% (v/v) dimethyl sulfoxide (VWR In-

ternational), 1% (v/v) Triton X-100 (Sigma-Aldrich Corporation), 0.1% (w/v) bovine serum albumin (Sigma-Aldrich Corporation), 0.05% (w/v) sodium azide (Sigma-Aldrich Corporation)

Washing Solution C: In 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (Life Technologies), 20% (v/v) methanol (EMD Millipore), 1% (v/v) Triton X-100 (Sigma-Aldrich Corporation), 0.1% (w/v) bovine serum albumin (Sigma-Aldrich Corporation), 0.05% (w/v) sodium azide (Sigma-Aldrich Corporation)

Cell or worm pellets were re-suspended in double-distilled water (1.00 volume) and incubated for five minutes in ice. To fix animals and crack their cuticles, ice-cold *C. elegans* Fixing Solution (1.00 volume) was added to the mixtures of worms and incubated for five minutes in liquid nitrogen. Frozen solutions were thawed for five minutes in room temperature water. The last two freeze-thaw steps were repeated for a total of three times. Thawed mixtures were incubated for an additional hour in ice. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. Washing Solution A (2.00 volumes) was added and mixed, and the mixtures were centrifuged at 1,000 x g for one minute at room temperature. The last three steps were repeated for a total of five times. The top aqueous layer was discarded. To reduce animals' cuticles, Reducing Solution (2.00 volumes) was added and mildly agitated for 30 minutes at room temperature. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. To alkylate animals' cuticles, Alkylating Solution (2.00 volumes) was added and mildly agitated for 30 minutes at room temperature. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (2.00 volumes, Life Technologies) was added and mixed. The mixtures were centrifuged at 1,000 x g for one minute at

room temperature. The last three steps were repeated for a total of five times. The top aqueous layer was discarded. Permeabilizing Solution A (2.00 volumes) was added and mildly agitated for one hour at room temperature. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. Permeabilizing Solution B (1.98 volumes) was added and mixed. TAMRA-DBCO Solution (0.02 volume) was added and mildly agitated for 20 minutes at room temperature. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. Washing Solution B (2.00 volumes) was added and mixed. The mixtures were centrifuged at 1,000 x g for one minute at room temperature. The last three steps were repeated for a total of ten times. The top aqueous layer was discarded. Washing Solution C (2.00 volumes) was added and mixed. The mixtures were centrifuged at 1,000 x g for one minute at room temperature. The last three steps were repeated for a total of ten times. The top aqueous layer was discarded. Permeabilizing Solution A (2.00 volumes) was added and mixed. The mixtures were centrifuged at 1,000 x g for one minute at room temperature. The last three steps were repeated for a total of ten times. Processed worms were mounted on 3% (w/v) UltraPure Low Melting Point Agarose (Life Technologies) pads and imaged with a LSM 510 META microscope (Carl Zeiss AG) at the Biological Imaging Center in the Beckman Institute at Caltech. Samples were protected from light at all times.

2.7.6 In-Gel Fluorescence Scanning of Azf-Labeled Proteins

Lysing Solution A: In double-distilled water, 200 mM Trizma base (Sigma-Aldrich Corporation), 4% (w/v) sodium dodecyl sulfate (Sigma-Aldrich Corporation), pH 8.0
Lysing Solution B: 1X Lysing Solution A, 100 mM 2-chloroacetamide (Sigma-Aldrich Corporation), protease inhibitor cocktail (1 tablet/10 mL cOmplete, Mini, EDTA-free, Roche Diagnostics)

Lysing Solution C: 75% (v/v) Lysing Solution A, 18.75% (v/v) NuPAGE LDS Sample Buffer (Life Technologies), 6.25% (v/v) 2-mercaptoethanol (Sigma-Aldrich Corporation)

Protein Gel Fixing Solution: 40% (v/v) double-distilled water, 50% (v/v) methanol (EMD Millipore), 10% acetic acid, glacial (EMD Millipore)

TAMRA-DBCO Solution: 1 mM TAMRA-DBCO (Click Chemistry Tools) in dimethyl sulfoxide, anhydrous (Life Technologies)

Cell or worm pellets were re-suspended in freshly prepared Lysing Solution B and lysed by sonication. Lysates were clarified by centrifugation at 20,000 x *g* for 20 minutes at room temperature. Protein concentrations of lysates were determined by bicinchoninic acid assay (BCA Protein Assay Kit, Thermo Fisher Scientific) according to the manufacturer's instructions and normalized to 1.11 mg/mL with additional Lysing Solution B. Lysates (1.00 volumes) were incubated with TAMRA-DBCO Solution (0.11 volume) for 20 minutes at room temperature. Excess TAMRA-DBCO was removed according to the **Chloroform/Methanol Precipitation** procedure. Protein pellets were re-suspended in freshly prepared Lysing Solution C (1.11 volumes) and incubated for 20 minutes at 100°C. The samples were electrophoresed on NuPAGE Novex 4-12% Bis-Tris Protein Gels (Life Technologies) according to the manufacturer's instructions. Protein gels were incubated with Protein Gel Fixing Solution for 20 minutes at room temperature and imaged with a Typhoon Trio+ instrument (GE Healthcare Life Sciences). Protein gels were stained using the Colloidal Blue Staining Kit (Life Technologies) according to the manufacturer's instructions. Samples were protected from light at all times.

2.7.7 In-Gel Proteolytic Digestion of Azf-Labeled Proteins

Lysing Solution A: In double-distilled water, 200 mM Trizma base (Sigma-Aldrich Corporation), 4% (w/v) sodium dodecyl sulfate (Sigma-Aldrich Corporation), pH 8.0
Lysing Solution B: 1X Lysing Solution A, 100 mM 2-chloroacetamide (Sigma-Aldrich Corporation), protease inhibitor cocktail (1 tablet/10 mL cOmplete, Mini, EDTA-free, Roche Diagnostics)
Lysing Solution C: 75% (v/v) Lysing Solution A, 18.75% (v/v) NuPAGE LDS Sample Buffer (Life Technologies), 6.25% (v/v) 2-mercaptoethanol (Sigma-Aldrich Corporation)

Cell or worm pellets were re-suspended in freshly prepared Lysing Solution B and lysed by sonication. Lysates were clarified by centrifugation at 20,000 x g for 20 minutes at room temperature. Protein concentrations of lysates were determined by bicinchoninic acid assay (BCA Protein Assay Kit, Thermo Fisher Scientific) according to the manufacturer's instructions and normalized to 5.55 mg/mL with additional Lysing Solution B. The samples are equally divided into two groups, one designated as "unenriched" and the other as "enriched".

"Unenriched" samples were re-normalized to 1.11 mg/mL with additional Lysing Solution B. Lysates (1.11 volumes) were subjected to the **Chloroform/Methanol Precipitation** procedure. Protein pellets were re-suspended in freshly prepared Lysing Solution C (1.11 volumes) and incubated for 20 minutes at 100°C. The samples were electrophoresed on NuPAGE Novex 4-12% Bis-Tris Protein Gels (Life Technologies) according to the manufacturer's instructions. Protein gels were stained using the Colloidal Blue Staining Kit (Life Technologies) according to the manufacturer's instructions.

"Enriched" samples were subjected to the **Enrichment of Azf-Labeled Proteins** procedure and dried by centrifugal evaporation. Protein pellets were re-normalized to 1.11 mg/mL with freshly prepared Lysing Solution C and incubated for 20 minutes at 100°C. The samples were electrophoresed on NuPAGE Novex 4-12% Bis-Tris Protein Gels (Life Technologies) according to the manufacturer's instructions. Protein gels were stained using the Colloidal Blue Staining Kit (Life Technologies) according to the manufacturer's instructions. Peptides were extracted from protein gels according to the in-gel proteolytic digestion procedures described by Shevchenko et al. (Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., and Mann, M. *Nat. Protoc.* **1**, 2856–2860 (2007)) and dried by centrifugal evaporation. Peptides were desalted according to the micro-purification procedures described by Rappsilber et al. (Rappsilber, J., Mann, M., and Ishihama, Y. *Nat. Protoc.* **2**, 1896–1906 (2007)) and dried by centrifugal evaporation.

2.7.8 Isolation of 6xHis-Tagged Proteins

Native Lysing Solution: In double-distilled water, 50 mM sodium phosphate, monobasic monohydrate (VWR International), 300 mM sodium chloride (VWR International), 10 mM imidazole (Sigma-Aldrich Corporation), pH 8.0, 4 mg/mL lysozyme from chicken egg white (Sigma-Aldrich Corporation)

Cell pellets were re-suspended in freshly prepared Native Lysing Solution, incubated for one hour in ice and lysed by sonication. Lysates were clarified by centrifugation at 20,000 x *g* for 20 minutes at 4°C. 6xHis-tagged proteins were extracted from lysates using Ni-NTA Agarose (Qiagen) according to the manufacturers instructions. Proteins were desalted using PD-10 Desalting Columns (GE Healthcare Life Sciences) according to the manufacturers instructions and stored in 100 mM Trizma base (Sigma-Aldrich Corporation), pH 7.4 at 4°C.

2.7.9 Labeling in C. elegans

Potassium Citrate Solution: In double-distilled water, 20 g/L citric acid monohydrate (VWR International), 293.5 g/L potassium citrate monohydrate (VWR International), pH 6.0

S Medium: In double-distilled water, 5.85 g/L sodium chloride (VWR International), 1 g/L potassium phosphate dibasic (VWR International), 6 g/L potassium phosphate monobasic (VWR International), 5 mg/L cholesterol (Sigma-Aldrich Corporation), 1% (v/v) potassium citrate solution, 1% (v/v) Trace Metals Solution, 3 mM calcium chloride dihydrate (VWR International), 3 mM magnesium chloride hexahydrate (VWR International)

Trace Metals Solution: In double-distilled water, 1.86 g/L ethylenediaminetetraacetic acid disodium salt dihydrate (Sigma-Aldrich Corporation), 0.69 g/L iron(II) sulfate heptahydrate (Sigma-Aldrich Corporation), 0.2 g/L manganese(II) chloride tetrahydrate, 0.29 g/L zinc sulfate heptahydrate (Sigma-Aldrich Corporation), 0.025 g/L copper(II) sulfate pentahydrate (Sigma-Aldrich Corporation)

hsp-16.2::Thr412Gly-*Ce*PheRS, *myo-3*::Thr412Gly-*Ce*PheRS, *ges-1*::Thr412Gly-*Ce*PheRS, *rab-3*::Thr412Gly-*Ce*PheRS, and *myo-2*::Thr412Gly-*Ce*PheRS worms: *C. elegans* strains previously maintained in S Medium supplemented with 25 mg/mL *E. coli* OP50 (*Caenorhabditis* Genetics Center) at 20°C with agitation were pelleted by centrifugation at 1000 x *g* for five minutes at room temperature, washed three times with S Medium, and re-suspended in S Medium supplemented with 25 mg/mL *E. coli* KY14[pKPY514] or KY33[pKPY514]. KY14[pKPY514] or KY33[pKPY514] was prepared according to the **Labeling in *E. coli*** procedure. For *hsp-16.2*::Thr412Gly-*Ce*PheRS worms, we induced heat shock of worms grown at 20°C by one hour exposure to 33°C. After 24 hours of agitation at 20°C, worms were harvested by centrifugation at 1000 x *g* for five minutes at room temperature and cleaned according to the sucrose flotation procedures described by Portman (Portman, D.S., *WormBook*,

ed. The *C. elegans* Research Community, *WormBook*, doi/10.1895/wormbook.1.104.1 (2006)). Digested bacterial material was cleared by washing worms with S Medium over a period of 30 minutes. Worms were pelleted by centrifugation at 1000 x g for five minutes at room temperature and frozen in liquid nitrogen. Samples were protected from light at all times.

2.7.10 Labeling in E. coli

M9 Minimal Medium A: In double-distilled water, 6.8 g/L sodium phosphate dibasic Anhydrous (VWR International), 3 g/L potassium phosphate monobasic (VWR International), 0.5 g/L sodium chloride (VWR International), 1 g/L ammonium chloride (VWR International), 2 mM magnesium sulfate heptahydrate (VWR International), 0.1 mM calcium chloride dihydrate (VWR International), 35 mg/L thiamine hydrochloride (Sigma-Aldrich Corporation), 40 mg/L L-alanine (Sigma-Aldrich Corporation), 40 mg/L L-cysteine (Sigma-Aldrich Corporation), 40 mg/L L-aspartic Acid (Sigma-Aldrich Corporation), 40 mg/L L-glutamic Acid (Sigma-Aldrich Corporation), 40 mg/L glycine (Sigma-Aldrich Corporation), 40 mg/L L-histidine (Sigma-Aldrich Corporation), 40 mg/L L-isoleucine (Sigma-Aldrich Corporation), 40 mg/L L-leucine, 40 mg/L L-methionine (Sigma-Aldrich Corporation), 40 mg/L L-asparagine (Sigma-Aldrich Corporation), 40 mg/L L-proline (Sigma-Aldrich Corporation), 40 mg/L L-glutamine (Sigma-Aldrich Corporation), 40 mg/L L-serine (Sigma-Aldrich Corporation), 40 mg/L L-threonine (Sigma-Aldrich Corporation), 40 mg/L L-valine (Sigma-Aldrich Corporation), 40 mg/L L-tryptophan (Sigma-Aldrich Corporation), 40 mg/L L-tyrosine (Sigma-Aldrich Corporation)

M9 Minimal Medium B: In M9 Minimal Medium A, 0.5% (v/v) glycerol (VWR International), 40 mg/L L-phenylalanine (Sigma-Aldrich Corporation), 40 mg/L L-lysine (Sigma-Aldrich Corporation), 40 mg/L L-arginine (Sigma-Aldrich Corporation), 200 mg/L ampicillin sodium salt (Sigma-Aldrich Corporation), 35 mg/L chloramphenicol

(Sigma-Aldrich Corporation)

M9 Minimal Medium C: M9 Minimal Medium A, 0.5% (v/v) glycerol (VWR International), 40 mg/L L-lysine (Sigma-Aldrich Corporation), 40 mg/L L-arginine (Sigma-Aldrich Corporation), 200 mg/L ampicillin sodium salt (Sigma-Aldrich Corporation), 35 mg/L chloramphenicol (Sigma-Aldrich Corporation)

M9 Minimal Medium D: In M9 Minimal Medium A, 20 mM D-glucose (VWR International), 40 mg/L L-phenylalanine (Sigma-Aldrich Corporation), 40 mg/L L-lysine (Sigma-Aldrich Corporation), 40 mg/L L-arginine (Sigma-Aldrich Corporation), 35 mg/L kanamycin sulfate (Sigma-Aldrich Corporation)

M9 Minimal Medium E: In M9 Minimal Medium A, 20 mM D-glucose (VWR International), 40 mg/L L-lysine (Sigma-Aldrich Corporation), 40 mg/L L-arginine (Sigma-Aldrich Corporation), 35 mg/L kanamycin sulfate (Sigma-Aldrich Corporation)

M9 Minimal Medium F: In M9 Minimal Medium A, 20 mM D-glucose (VWR International), 40 mg/L L-phenylalanine (Sigma-Aldrich Corporation), 40 mg/L L-lysine ($^{13}C_6{}^{15}N_2{}^1H_{14}{}^{16}O_2$, Cambridge Isotope Laboratories), 40 mg/L L-arginine ($^{13}C_6{}^{14}N_4{}^1H_{14}{}^{16}O_2$, Cambridge Isotope Laboratories), 35 mg/L kanamycin sulfate (Sigma-Aldrich Corporation)

M9 Minimal Medium G: In M9 Minimal Medium A, 20 mM D-glucose (VWR International), 40 mg/L L-lysine ($^{13}C_6{}^{15}N_2{}^1H_{14}{}^{16}O_2$, Cambridge Isotope Laboratories), 40 mg/L L-arginine ($^{13}C_6{}^{14}N_4{}^1H_{14}{}^{16}O_2$, Cambridge Isotope Laboratories), 35 mg/L kanamycin sulfate (Sigma-Aldrich Corporation)

Terrific Broth: In double-distilled water, 12 g/L Bacto tryptone (BD Biosciences), 24 g/L Bacto yeast extract (BD Biosciences), 0.4% (v/v) glycerol (EMD Millipore), 2.31 g/L potassium phosphate monobasic (VWR International), 12.54 g/L potassium phosphate dibasic (VWR International), 200 mg/L ampicillin sodium salt (Sigma-Aldrich Corporation)

KY14[pKPY66] and KY14[pKPY68]: A single colony was used to inoculate an overnight culture of freshly prepared Terrific Broth (5.0 mL). Overnight cultures were diluted into freshly prepared Terrific Broth (50 mL) and agitated at 37°C until reaching an OD_{600} of 0.5. After another 30 minutes of agitation at 20°C, expression of 6xHis-tagged wild-type and mutant *Ce*PheRSs was induced by addition of 1 M isopropyl β-D-1-thiogalactopyranoside (50 μL, Sigma-Aldrich Corporation). After 16 hours of agitation at 20°C, cells were harvested by centrifugation at 5000 x g for five minutes at 4°C and frozen in liquid nitrogen.

KY14[pKPY93 and pKPY100 or pKPY133 or pKPY127 or pKPY131 or pKPY129 or pKPY125 or pKPY101 or pKPY102 or pKPY509 or pKPY503 or pKPY505 or pKPY647 or pKPY655 or pKPY507 or pKPY137]: A single colony was used to inoculate an overnight culture of freshly prepared M9 Minimal Medium B (5.0 mL). Overnight cultures were diluted into freshly prepared M9 Minimal Medium B (50 mL) and agitated at 37°C until reaching an OD_{600} of 0.5. Expression of wild-type and mutant *Ce*PheRSs was induced by addition of 1 M L-(+)-arabinose (500 μL, Sigma-Aldrich Corporation). After 30 minutes, cells were pelleted by centrifugation at 5000 x g for five minutes at 4°C, washed three times with ice-cold 0.9% (w/v) sodium chloride (VWR International) and re-suspended in freshly prepared M9 Minimal Medium C supplemented with L-phenylalanine (0.0 mM or 0.2 mM final concentration, Sigma-Aldrich Corporation) or *p*-azido-L-phenylalanine (0.0 mM, 0.2 mM or 2.0 mM final concentration, Chem-Impex International). After another 30 minutes of agitation at 37°C, expression of 6xHis-tagged GFPs was induced by addition of 1 M isopropyl β-D-1-thiogalactopyranoside (50 μL, Sigma-Aldrich Corporation). After four hours of agitation at 37°C, cells were harvested by centrifugation at 5000 x g for five minutes at 4°C and frozen in liquid nitrogen. Samples were protected from light at all times.

KY14[pKPY513], KY14[pKPY566], KY14[pKPY565], and KY14[pKPY514]: A single colony was used to inoculate an overnight culture of freshly prepared M9 Minimal Medium D (5.0 mL). Overnight cultures were diluted into freshly prepared M9 Minimal Medium D (50 mL) and agitated at 37°C until reaching an OD_{600} of 0.5. Cells were pelleted by centrifugation at 5000 x g for five minutes at 4°C, washed three times with ice-cold 0.9% (w/v) sodium chloride (VWR International) and re-suspended in freshly prepared M9 Minimal Medium E supplemented with L-phenylalanine (0.0 mM or 0.2 mM final concentration, Sigma-Aldrich Corporation) or p-azido-L-phenylalanine (0.2 mM or 2.0 mM final concentration, Chem-Impex International). After another 30 minutes of agitation at 37°C, expression of wild-type or mutant *Ec*PheRSs was induced by addition of 1 M isopropyl β-D-1-thiogalactopyranoside (50 μL, Sigma-Aldrich Corporation). After four hours of agitation at 37°C, cells were harvested by centrifugation at 5000 x g for five minutes at 4°C and frozen in liquid nitrogen. Extent of replacement of L-phenylalanine by *p*-azido-L-phenylalanine was determined by amino acid analysis at the Protein Sequencing and Amino Acid Analysis Resources of the W.M. Keck Foundation Biotechnology Resource Laboratory at the Yale School of Medicine. Samples were protected from light at all times.

KY14[pKPY514]: A single colony was used to inoculate an overnight culture of freshly prepared M9 Minimal Medium D (5.0 mL). Overnight cultures were diluted into freshly prepared M9 Minimal Medium D (500 mL) and agitated at 37°C until reaching an OD_{600} of 0.5. Cells were pelleted by centrifugation at 5000 x g for 15 minutes at 4°C, washed three times with ice-cold 0.9% (w/v) sodium chloride (VWR International) and re-suspended in freshly prepared M9 Minimal Medium E supplemented with *p*-azido-L-phenylalanine (2.0 mM final concentration, Chem-Impex International). After another 30 minutes of agitation at 37°C, expression of Thr251Gly-*Ec*PheRS was induced by addition of 1 M isopropyl β-D-1-thiogalactopyranoside (500

μL, Sigma-Aldrich Corporation). After four hours of agitation at 37°C, cells were harvested by centrifugation at 5000 x g for 15 minutes at 4°C, re-suspended with freshly prepared M9 Minimal Medium E supplemented with p-azido-L-phenylalanine (2.0 mM final concentration, Chem-Impex International) at a concentration of 250 mg wet cell mass/mL medium and stored at 4°C. Samples were protected from light at all times.

KY33[pKPY514]: A single colony was used to inoculate an overnight culture of freshly prepared M9 Minimal Medium F (5.0 mL). Overnight cultures were diluted into freshly prepared M9 Medium F (500 mL) and agitated at 37°C until reaching an OD_{600} of 0.5. Cells were pelleted by centrifugation at 5000 x g for 15 minutes at 4°C, washed three times with ice-cold 0.9% (w/v) sodium chloride (VWR International) and re-suspended in freshly prepared M9 Minimal Medium G supplemented with p-azido-L-phenylalanine (2.0 mM final concentration, Chem-Impex International). After another 30 minutes of agitation at 37°C, expression of Thr251Gly-*Ec*PheRS was induced by addition of 1 M isopropyl β-D-1-thiogalactopyranoside (500 μL, Sigma-Aldrich Corporation). After four hours of agitation at 37°C, cells were harvested by centrifugation at 5000 x g for 15 minutes at 4°C, re-suspended with freshly prepared M9 Minimal Medium G supplemented with p-azido-L-phenylalanine (2.0 mM final concentration, Chem-Impex International) at a concentration of 250 mg wet cell mass/mL medium and stored at 4°C. Samples were protected from light at all times.

2.7.11 LC-MS/MS of Azf-Labeled Proteins

Solvent A: 0.2% (v/v) formic acid (Sigma-Aldrich Corporation), 2.0% (v/v) LC-MS CHROMASOLV acetonitrile (Sigma-Aldrich Corporation), 97.8% (v/v) LC-MS CHROMASOLV water (Sigma-Aldrich Corporation)

Solvent B: 0.2% (v/v) formic acid (Sigma-Aldrich Corporation), 80.0% (v/v) LC-MS CHROMASOLV acetonitrile (Sigma-Aldrich Corporation), 19.8% (v/v) LC-MS

CHROMASOLV water (Sigma-Aldrich Corporation)

LC-MS/MS experiments were performed on an EASY-nLC1000 (Proxeon Biosystems, now Thermo Fisher Scientific) coupled to a hybrid LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher Scientific) equipped with a nano-electrospray ion source (Thermo Fisher Scientific). Peptides were re-suspended in 0.1% (v/v) formic acid (Sigma-Aldrich Corporation) and loaded on a 15-cm reversed phase analytical column (75 μm ID) packed in-house with 3 μm C18AQ beads (ReproSil-Pur C18AQ, Dr. Maisch-GmbH). The chromatographic separation was achieved using a 90 minute elution gradient from 0% (100% Solvent A) to 30% Solvent B (70% Solvent A) at a flow rate of 350 nL/minute. The mass spectrometer was operated in data-dependent mode to switch automatically between MS and MS/MS scans as described by Kalli and Hess (Kalli, A. and Hess, S. *Proteomics* **12**, 21–31 (2012)). Survey full scan mass spectra were acquired in the Orbitrap (400–1600 m/z) with a resolution of 60,000 at 400 m/z. The top ten most intense ions from the survey scan were isolated and fragmented in the linear ion trap by collisionally induced dissociation (collisional energy 35% and isolation width 2 Da). Precursor ion charge state screening was enabled, and all singly charged and unassigned charge states were rejected. The dynamic exclusion list was set with a maximum retention time of 90 seconds and a relative mass window of 10 ppm. Thermo raw data files were analyzed by MaxQuant (v 1.5.0.0) (Cox, J. and Mann, M. *Nat. Biotechnol.* **26**, 1367–1372 (2008); Cox, J., et al. *J. Proteome Res.* **10**, 1794–1805 (2011)) and were searched against the UniProt *C. elegans* database (26900 sequences) and an in-house contaminant database (259 sequences) including human keratins and proteases. All default options were used except as follows: match between runs was enabled, re-quantify was enabled, and multiplicity of two with heavy labels "Arg6" (+6.0201) and "Lys8" (+8.0142). Variable modifications included oxidation of L-methionine (+15.9949), N-terminal protein acetylation (+42.0106), L-phenylalanine to *p*-azido-L-phenylalanine (+41.0014), L-phenylalanine

to *p*-amino-L-phenylalanine (+15.0109), and cleaved Diazo Biotin-DBCO-modified *p*-azido-L-phenylalanine (+482.2066). Tryptic digest was specified with up to two missed cleavages. Initial precursor mass tolerance was 7 ppm; however, MaxQuant calculates tighter individual precursor tolerances after recalibration. Fragment ion tolerance was 0.5 Da. Peptide, protein and site false discovery rates were fixed at 1% using the target-decoy approach with a reversed database. To account for conversion of heavy labeled L-Lysine and heavy labeled L-Arginine to other amino acids (Park, S.K., Liao, L., Kim, J.Y., and Yates III, J.R. *Nat. Methods* **6**, 184–185 (2009)), a linear regression model was constructed to calculate the effects each amino acid has on the peptide ratio. These coefficients were used to back-calculate peptide ratios without biases from the amino acid content. Overall enrichment factor is the mean of the $\log_2$ biological replicate enrichment factors.

2.7.12 MALDI TOF-MS of 6xHis-Tagged Proteins

α-Cyano-4-Hydroxycinnamic Acid Solution: 10 mg/mL α-cyano-4-hydroxycinnamic Acid (Sigma-Aldrich Corporation) in 10% (v/v) trifluoroacetic acid (Sigma-Aldrich Corporation), 40% (v/v) LC-MS CHROMASOLV water (Sigma-Aldrich Corporation), 50% (v/v) LC-MS CHROMASOLV acetonitrile (Sigma-Aldrich Corporation)

Purified proteins were obtained from the **Isolation of 6xHis-Tagged Proteins** procedure. Protein concentrations of purified proteins were determined by absorbance at 280 nm. NuPAGE LDS Sample Buffer (0.25 volume, Life Technologies) and 2-Mercaptoethanol (0.08 volume, Sigma-Aldrich Corporation) were added to protein solutions (1.00 volume) and incubated for 20 minutes at 100°C. The samples were electrophoresed on NuPAGE Novex 4-12% Bis-Tris Protein Gels (Life Technologies) according to the manufacturers instructions. Protein gels were stained using the Colloidal Blue Staining Kit (Life Technologies) according to the manufacturers instructions. Peptides were extracted from protein gels according to the in-gel proteolytic

digestion procedures described by Shevchenko et al. (Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V., and Mann, M. *Nat. Protoc.* **1**, 2856–2860 (2007)) and dried by centrifugal evaporation. Peptides were desalted according to the micro-purification procedures described by Rappsilber et al. (Rappsilber, J., Mann, M., and Ishihama, Y. *Nat. Protoc.* **2**, 1896–1906 (2007)). The peptides were mixed with α-Cyano-4-Hydroxycinnamic Acid Solution (1-10 pmol peptide/μL solution) and assayed on a Voyager DE Pro instrument (Applied Biosystems) at the Caltech Division of Chemistry and Chemical Engineering Mass Spectrometry Facility.

2.7.13 Plasmids and Strains

pKPY66: The genes encoding wild-type *C. elegans* PheRS were isolated by RNA extraction (RNeasy Mini Kit, Qiagen) from wild-type (N2) *C. elegans*, reverse transcription (SuperScript III Reverse Transcriptase, Life Technologies) and PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies). The PCR product encodes the alpha (isoform A) and beta subunits of *C. elegans* PheRS as well as the intercistronic region from *E. coli pheST*. The purified fragments were ligated into pQE-80L (Qiagen) to generate pKPY66. We have designated the first methionine in the alpha subunit of *Ce*PheRS, isoform A, exon 3 as residue 1 ("Met1).

pKPY68: pKPY66 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr412Gly mutation.

pKPY93: The *lpp* promoter region was isolated by genomic DNA extraction (DNeasy Blood and Tissue Kit, Qiagen) from *E. coli* DH10B (Life Technologies) and PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies). *C. elegans* tRNA^{Phe} was synthesized by primer annealing and extension with Klenow Fragment (3 to 5 exo-, New England Biolabs). The *rrnB* terminator region was isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent

Technologies) of pBAD33 (Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. *J. Bacteriol.* **177**, 4121–4130 (1995)). The purified fragments were ligated into the NdeI site of pJTN4 (Ngo, J.T., et al. *Nat. Chem. Biol.* **5**, 715–717 (2009)) to generate pKPY93.

pKPY100: The genes encoding wild-type *C. elegans* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of pKPY66. The purified fragments were ligated into pBAD33.

pKPY133: pKPY100 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr412Asn mutation.

pKPY127: pKPY100 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr412Pro mutation.

pKPY131: pKPY100 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr412Asp mutation.

pKPY129: pKPY100 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr412Cys mutation.

pKPY125: pKPY100 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr412Ser mutation.

pKPY101: pKPY100 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr412Ala mutation.

pKPY102: pKPY100 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to

generate the Thr412Gly mutation.

pKPY509: The genes encoding wild-type *Saccharomyces cerevisiae* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of plasmids YSC3867-202326582 from Thermo Fisher Scientific (Accession Number YFL022C) and YSC3867-202328360 from Thermo Fisher Scientific (Accession Number YLR060W). The purified fragments were ligated into pBAD33 and subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the equivalent Thr412Gly mutation.

pKPY503: The genes encoding wild-type *Drosophila melanogaster* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of genomic DNA from *Drosophila melanogaster* (0000-1012.01 from *Drosophila* Species Stock Center). The purified fragments were ligated into pBAD33 and subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the equivalent Thr412Gly mutation.

pKPY505: The genes encoding wild-type *Danio rerio* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of plasmids 6806236 from Thermo Fisher Scientific (Accession Numbers BC066611, CA975688) and 7046061 from Thermo Fisher Scientific (Accession Numbers CK126922, BC085625). The purified fragments were ligated into pBAD33 and subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the equivalent Thr412Gly mutation.

pKPY647: The genes encoding wild-type *Xenopus laevis* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of plasmids 5156818 from GE Healthcare Dharmacon (Accession Numbers CA790449 BC078035) and 5048633 from GE Healthcare Dharmacon (Accession Numbers BM262817 BM261881 BC056121). The purified fragments were ligated into pBAD33 and subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the equivalent Thr412 Gly mutation.

pKPY655: The genes encoding wild-type *Rattus norvegicus* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of plasmids 7314191 from GE Healthcare Dharmacon and 7109296 from GE Healthcare Dharmacon. The purified fragments were ligated into pBAD33 and subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the equivalent Thr412Gly mutation.

pKPY507: The genes encoding wild-type *Mus musculus* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of plasmids 3595108 from Thermo Fisher Scientific (Accession Number BC006862) and 4239906 from Thermo Fisher Scientific (Accession Number BC016428). The purified fragments were ligated into pBAD33 and subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the equivalent Thr412Gly mutation.

pKPY137: The genes encoding wild-type *Homo sapiens* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of plasmids 2823551 from Open Biosystems (Accession Numbers AW732158, AW732158.1, BC006495, BC006495.2, BE208518, BE208518.1) and 4664949 from Open Biosystems (Accession Numbers BC017783, BC017783.1, BG504513,

BG504513.1). The purified fragments were ligated into pBAD33 and subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the equivalent Thr412Gly mutation.

pKPY513: The genes encoding wild-type *E. coli* PheRS were isolated by genomic DNA extraction (DNeasy Blood and Tissue Kit, Qiagen) from *E. coli* DH10B (Life Technologies) and PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies). The purified fragments were ligated into pQE-80L-Kan (Qiagen) to generate pKPY513.

pKPY566: pKPY513 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr251Ser mutation.

pKPY565: pKPY513 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr251Ala mutation.

pKPY514: pKPY513 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr251Gly mutation.

pKPY197 (*frs*-1(Thr412Gly)::*fib-1/rps-16*::*gfp*(S65C, SynIVS)::*unc-54*): The gene encoding the mutant alpha subunit of *C. elegans* PheRS was isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of pKPY102. The intercistronic region of the *C. elegans* operon CEOP5428 as described by Lee et al. (Lee, L.W., Lo, H.W., and Lo, S.J. *Gene* **455**, 16–21 (2010)) was isolated by genomic DNA extraction (DNeasy Blood and Tissue Kit, Qiagen) from wild-type (N2) *C. elegans* and PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) The purified fragments were ligated into pPD95.75 (Fire Lab Vector Kit, Addgene) to generate pKPY197.

KY14: The lysine- and phenylalanine-auxotrophic strain of *E. coli* KY14 was made in-house using the red recombinase gene knockout method described by Datsenko and Wanner (Datsenko, K.A. and Wanner, B.L. *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645 (2000)) to eliminate the gene *pheA* from the *E. coli* strain KY2. The lysine-auxotrophic strain of KY2 was made in-house using the red recombinase gene knockout method described by Datsenko and Wanner to eliminate the gene *lysA* from the *E. coli* strain TYJV2 (Van Deventer, J.A., Yuet, K.P., Yoo, T.H., and Tirrell, D.A. *ChemBioChem* **15**, 1777–1781 (2014)).

KY33: The arginine-, lysine- and phenylalanine-auxotrophic strain of *E. coli* KY33 was made in-house using the red recombinase gene knockout method described by Datsenko and Wanner (Datsenko, K.A. and Wanner, B.L. *Proc. Natl. Acad. Sci. USA* **97**, 6640–6645 (2000)) to eliminate the gene *argA* from the *E. coli* strain KY14.

Worm strains were cultured as described by Brenner (Brenner, S. *Genetics* **77**, 71–94 (1974)). The strains used in this study were as follows:

PS6741 (*ges-1*::Thr412Gly-*Ce*PheRS): *pha-1*(*e2123ts*) *III*, *him-5*(*e1490*) *V*; *syEx1341*[*ges-1*::*frs-1*(Thr412Gly)::*fib-1*/*rps-16*::*gfp*(S65C, SynIVS)::*unc-54* (1 ng/μL) + *myo-2*::*dsRed* (10 ng/μL) + *pha-1*(+) (90 ng/μL) + pBluescript (180 ng/μL)]

PS6742 (*myo-2*::Thr412Gly-*Ce*PheRS): *pha-1*(*e2123ts*) *III*, *him-5*(*e1490*) *V*; *syEx1342*[*myo-2*::*frs-1*(Thr412Gly)::*fib-1*/*rps-16*::*gfp*(S65C, SynIVS)::*unc-54* (1 ng/μL) + *unc-122*::*mRFP* (10 ng/μL) + *pha-1*(+) (90 ng/μL) + pBluescript (180 ng/μL)]

PS6743 (*myo-3*::Thr412Gly-*Ce*PheRS): *pha-1*(*e2123ts*) *III*, *him-5*(*e1490*) *V*; *syEx1343*[*myo-3*::*frs-1*(Thr412Gly)::*fib-1*/*rps-16*::*gfp*(S65C, SynIVS)::*unc-54* (1 ng/μL) + *myo-2*::*dsRed* (10 ng/μL) + *pha-1*(+) (90 ng/μL) + pBluescript (180 ng/μL)]

PS6744 (*rab-3*::Thr412Gly-*Ce*PheRS): *pha-1*(*e2123ts*) *III*, *him-5*(*e1490*) *V*; *syEx1344*[*rab-3*::*frs-1*(Thr412Gly)::*fib-1*/*rps-16*::*gfp*(S65C, SynIVS)::*unc-54* (1 ng/μL) + *myo-2*::*dsRed* (10 ng/μL) + *pha-1*(+) (90 ng/μL) + pBluescript (180 ng/μL)]

PS6745 (*hsp-16.2*::Thr412Gly-*Ce*PheRS): *pha-1*(*e2123ts*) *III*, *him-5*(*e1490*) *V*; *syIs266*[*hsp-16.2*::*frs-1*(Thr412Gly)::*fib-1*/*rps-16*::*gfp*(S65C, SynIVS)::*unc-54* (1 ng/μL) + *myo-2*::*dsRed* (10 ng/μL) + *pha-1*(+) (90 ng/μL) + pBluescript (180 ng/μL)]

2.7.14 Western Blotting

Western blots were performed using standard procedures. Antibodies were used at the concentrations suggested by the manufacturers. Alexa Fluor 488-Goat Anti-Rat IgG (H+L) antibody, Alexa Fluor 647-Goat Anti-Rat IgG (H+L) antibody, Streptavidin-Alexa Fluor 488 conjugate and Streptavidin-Alexa Fluor 647 conjugate were from Life Technologies. Penta-His-Alexa Fluor 488 conjugate and Penta-His-Alexa Fluor 647 conjugate were from Qiagen. Rat Monoclonal Antibody to Tubulin [YL1/2] was from Abcam.

CHAPTER 3

Split-Intein-Mediated Split-Aminoacyl-tRNA Synthetase System for Spatiotemporally Specific Proteomic Analysis

3.1 Introduction

Cell-selective metabolic labeling of proteins with functionalizable amino acids is a promising method to precisely interrogate protein synthesis in a living animal without dissection. Because labeled proteins – and only labeled proteins – are subject to conjugation to affinity probes, proteins from targeted cells can be separated from the rest of the organism and identified by mass spectrometry. This potentially powerful technique operates on the premise that restricting expression of engineered aminoacyl-tRNA synthetases with promoters active only in specific cells restricts labeling to those cells. For example, Chin and coworkers used the germline-specific driver GAL4::VP16-nos.UTR in combination with an engineered *Methanosarcina* prroylysyl-tRNA synthetase/tRNA pair to identify proteins synthesized in germ cells of the *Drosophila* ovary [118]. Dieterich and coworkers used the motor neuron-specific driver OK371-GAL4 in combination with an engineered *Drosophila* methionyl-tRNA synthetase to monitor differences in protein synthesis rates in motor neurons in a fly model for Charcot-Marie-Tooth neuropathy [143, 144]. We have used the pharyngeal muscle-specific driver *myo-2* 5 in combination with an engineered *Caenorhabditis* phenylalanyl-tRNA synthetase to discover previously unidentified pharyngeal-muscle-specific calponin-like proteins [145]. However in a multicellular organism, many cells cannot be targeted with a single promoter. Frequently, cell- and time-specific expression arises from the combinatorial action of multiple regulators. To overcome the limitations of single-component systems for cell-selective metabolic labeling of proteins, we have developed a two-component system based on split-intein-mediated reconstitution of a split-aminoacyl-tRNA synthetase (**Figure 3.1**).

Inteins (*in*tervening pro*teins*) are protein segments that excise themselves out from a larger precursor polypeptide and join the remaining portions – exteins (*ex*ternal pro*teins*) – together with a peptide bond in a spontaneous auto-processing event known as protein splicing [146]. A special class of inteins – split inteins – are tran-

scribed and translated as two separate polypeptide fragments: an N-terminal intein fused to an N-terminal extein and a C-terminal intein fused to a C-terminal extein. These intein fragments undergo non-convalent assembly into the canonical intein structure followed by protein splicing *in trans* [147]. Inteins exhibit high splicing efficiencies; consequently, protein chemists and engineers have exploited intein chemistry in a variety of applications including tagless protein purification [148], *in vitro* and *in vivo* protein semi-synthesis [149,150], segmental isotopic labeling [151], protein and peptide cyclization [152], as well as conditional protein splicing [153]. Here in this chapter, we also exploit intein chemistry by demonstrating the reconstitution of aminoacyl-tRNA synthetase activity from individually expressed fragments with the assistance of split inteins first in *E. coli* and then in *C. elegans*. Using a combination of a heat-shock promoter (*hsp-16.2*) and a cell-specific promoter (*myo-2*, pharyngeal-muscle-specific), we show that metabolic labeling of proteins with a functionalizable amino acid can be introduced in a targeted subset of cells at specific times in the live worm.

3.2 Results and Discussion

3.2.1 Engineering Split-Intein-Reconstituted Split-Thr412Gly-CePheRS

We have previously engineered a *C. elegans* phenylalanyl-tRNA synthetase (Thr412 Gly-*Ce*PheRS) capable of appending the azide-bearing L-phenylalanine (Phe; **Figure 3.2**) analog *p*-azido-L-phenylalanine (Azf; **Figure 3.2**) to its cognate tRNA in competition with Phe and achieved spatiotemporal selectivity in the labeling of *C. elegans* proteins by controlling expression of this mutant (Thr412Gly) synthetase using cell-selective promoters [145]. Here, we explore split-intein-mediated protein splicing as a means for efficient reconstitution of the alpha subunit of *Ce*PheRS,

FARS-1, from two separate polypeptides. We first identified the cyanobacteria *Nostoc punctiforme* (*Npu*) *DnaE* intein as a promising split intein candidate as it displays superior splicing kinetics: $t_{1/2}$ of 60 s for the splicing reaction at 37°C (k = 1.1 ± 0.2 x 10^{-2} s^{-1}) [154, 155]. The catalytic subunit of DNA polymerase III DnaE is encoded by two separate genes: *dnaE-n* and *dnaE-c*. *dnaE-n* contains the N-terminal segment (N extein, DnaE(N)) of the protein followed by a 102-amino acid N-intein (Int(N, DnaE)) (**Figure 3.3**), whereas the *dnaE-c* product consists of a 36-residue C-intein (Int(C, DnaE) followed by the C-terminal segment (C extein, DnaE(C)) of the protein (**Figure 3.4**). We exchanged the extein regions with the respective halves of FARS-1 (FARS-1(N) and FARS-1(C)). For the first version of the split-intein-reconstituted split-Thr412Gly-*Ce*PheRS system, we selected the split-sites of FARS-1 between Lys187 and Gln188 in a surface-exposed linker domain bridging the N-terminal tRNA-binding domains and the C-terminal catalytic domains (FARS-1(N, Met1-Lys187)-Int(N, DnaE) and Int(C, DnaE)-Cys-Phe-Asn-FARS-1(C, Gln188-Lys496); **Figures 3.5, 3.6, and 3.7**). We added the native DnaE C-extein tripeptide cysteine-phenylalanine-asparagine between Int(C, DnaE) and FARS-1(C, Gln188-Lys496) as the presence of non-native C-extein residues immediately next to the C-intein can lead to dramatic reductions in splicing efficiency [155].

To evaluate DnaE-intein-driven reconstitution of Thr412Gly-*Ce*PheRS activity in *E. coli*, we designed seven *E. coli* strains. *E. coli* **1** and **2** represent negative and positive controls, respectively, of Azf-labeling resulting from expression of full-length FARS-1. *E. coli* **3**, **4**, and **5** test whether splitting FARS-1 yields variants that require only one fragment that associates with the beta subunit to form an active enzyme. *E. coli* **6** and **7** test Azf-labeling resulting from expression of FARS-1(N, Met1-Lys187)-Int(N, DnaE) and Int(C, DnaE)-Cys-Phe-Asn-FARS-1(C, Gln188-Lys496) and subsequent reconstitution of full-length FARS-1.

- *E. coli* **1** (Lane 1 in **Figure 3.8**) houses the two compatible plasmids pKPY93 and pKPY100: (i) pKPY93 encodes *Ce*tRNA^{Phe} under constitutive *E. coli* murein lipoprotein (*lpp*) promoter control and (ii) pKPY100 encodes both *Ce*PheRS wild-type intact alpha and beta subunits under arabinose-inducible (P_{BAD}) control.

- *E. coli* 2 (Lane 2 in **Figure 3.8**) houses pKPY93 and pKPY102: pKPY102 encodes both *Ce*PheRS mutant intact alpha and beta subunits under P_{BAD} control.

- *E. coli* 3 (Lane 3 in **Figure 3.8**) houses pKPY386 and pKPY273: (i) pKPY386 encodes *Ce*tRNA^{Phe} under *lpp* promoter control as well as FARS-1(N, Met1-Lys187)-Int(N, DnaE) under P_{BAD} control and (ii) pKPY273 encodes just the *Ce*PheRS beta subunit under P_{BAD} control.

- *E. coli* 4 (Lane 4 in **Figure 3.8**) houses pKPY93 and pKPY387: pKPY387 encodes both wild-type Int(C, DnaE)-Cys-Phe-Asn-FARS-1(C, Gln188-Lys496) and the *Ce*PheRS beta subunit under P_{BAD} control.

- *E. coli* 5 (Lane 5 in **Figure 3.8**) houses pKPY93 and pKPY388: pKPY388 encodes both mutant Int(C, DnaE)-Cys-Phe-Asn-FARS-1(C, Gln188-Lys496) and the *Ce*PheRS beta subunit under P_{BAD} control.

- *E. coli* 6 (Lane 6 in **Figure 3.8**) houses pKPY386 and pKPY387.

- *E. coli* 7 (Lane 7 in **Figure 3.8**) houses pKPY386 and pKPY388.

We cultured these strains in M9 minimal medium supplemented with 2 mM Azf and 10 mM arabinose. To assess enzymatic activity toward Azf, we detected Azf-labeled proteins by conjugation to dibenzocyclooctyne-functionalized tetramethylrhodamine (TAMRA-DBCO; **Figure 3.2**) and SDS/PAGE-in-gel fluorescence scanning

(**Figure 3.8**). As expected, wild-type intact (Lane 1) and reconstituted *Ce*PheRS (Lane 6) does not activate Azf. Notably, the expression of the synthetase-intein fragments by themselves does not yield Azf-labeling (Lanes 3, 4, and 5). Only the mutant intact (Lane 2) and reconstituted *Ce*PheRS (Lane 7) displayed robust labeling cells treated with Azf.

These results demonstrate that FARS-1 can be split into two distinct fragments, which form a functional full-length synthetase subunit when brought back together by intein-assisted splicing in *E. coli*. Encouraged by the performance of DnaE-intein-reconstituted split-Thr412Gly-*Ce*PheRS (relative labeling of 50% compared to intact synthetase), we hypothesized that a more judicious placement of the native DnaE C-extein tripeptide Cys-Phe-Asn should enable higher labeling activity. As this tripeptide remains in the product after splicing, the post-splicing variant of FARS-1 may exhibit inferior amino acid activation in relation to its wild-type counterpart. While there is no natural occurrence of Cys-Phe-Asn in FARS-1, the tripeptide Glu26-Phe27-Asn28 exists in a surface-exposed section of the N-terminal tRNA-binding domains. Reassuringly, we mutated this glutamic acid residue to cysteine and observed that bacteria expressing Glu26Cys-Thr412Gly-*Ce*PheRS are Azf-labeled to the same extent as bacteria expressing Thr412Gly-*Ce*PheRS (**Figure 3.9**). Therefore, for the second version of the split-intein-reconstituted split-Thr412Gly-*Ce*PheRS system, we selected the split-sites of FARS-1 between Asn25 and Glu26 and introduced the Glu26Cys mutation (FARS-1(N, Met1-Asn25)-Int(N, DnaE) and Int(C, DnaE)-FARS-1(C, Glu26Cys-Lys496); **Figures 3.5, 3.10, and 3.11**). We designed five new *E. coli* strains in addition to *E. coli* **1** (Lane 1 in **Figure 3.12**) and *E. coli* **2** (Lane 2 in **Figure 3.12**).

- *E. coli* **8** (Lane 3 in **Figure 3.12**) houses pKPY393 and pKPY273: (i) pKPY393 encodes *Ce*tRNA^{Phe} under *lpp* promoter control as well as FARS-1(N, Met1-Asn25)-Int(N, DnaE) under P_{BAD} control.

- *E. coli* **9** (Lane 4 in **Figure 3.12**) houses pKPY93 and pKPY394: pKPY394 encodes both wild-type Int(C, DnaE)-FARS-1(C, Glu26Cys-Lys496) and the *Ce*PheRS beta subunit under P_{BAD} control.

- *E. coli* **10** (Lane 5 in **Figure 3.12**) houses pKPY93 and pKPY395: pKPY388 encodes both mutant Int(C, DnaE)-FARS-1(C, Glu26Cys-Lys496) and the *Ce* PheRS beta subunit under P_{BAD} control.

- *E. coli* **11** (Lane 6 in **Figure 3.12**) houses pKPY393 and pKPY394.

- *E. coli* **12** (Lane 7 in **Figure 3.12**) houses pKPY393 and pKPY395.

Again, only the mutant intact (Lane 2) and reconstituted *Ce*PheRS (Lane 7) displayed robust labeling cells treated with Azf; however, repositioning the split-site from Lys187/Gln188 to Asn25/Glu26Cys did not significantly improve Azf-labeling.

Next, we asked whether replacement of the DnaE intein with an intein that possess higher splicing rates and yields would enable higher labeling activity. To our knowledge, the *Prochlorococcus* cyanomyophage P-SSM2 Gp41-1 is the most active intein reported to date: it gives rise to completed reactions within 20-30 s at 37°C ($k = 1.4 \pm 0.2\ x\ 10^{-1}\ s^{-1}$) [156]. The DNA primase/helicase Gp41-1 is encoded by two separate genes: *gp41-1-n* and *gp41-1-c*. *gp41-1-n* contains the N-terminal segment (N extein, Gp41-1(N)) of the protein followed by an 88-amino acid N-intein (Int(N, Gp41-1)) (**Figure 3.13**), whereas the *gp41-1-c* product consists of a 38-residue C-intein (Int(C, Gp41-1)) followed by the C-terminal segment (C extein, Gp41-1(C)) of the protein (**Figure 3.14**). Fortuitously, the native Gp41-1 C-extein tripeptide serine-serine-aspartic acid exists in another surface-exposed section of the N-terminal

tRNA-binding domains: Ser148-Ser149-Asp150. Therefore, for the third version of the split-intein-reconstituted split-Thr412Gly-*Ce*PheRS system, we selected the split-sites of FARS-1 between Gly147 and Ser148 (FARS-1(N, Met1-Gly147)-Int(N, Gp41-1) and Int(C, Gp41-1)-FARS-1(C, Ser148-Lys496); **Figures 3.5, 3.15, and 3.16**). Again, we designed five new *E. coli* strains in addition to *E. coli* **1** (Lane 1 in **Figure 3.17**) and *E. coli* **2** (Lane 2 in **Figure 3.17**).

- *E. coli* **13** (Lane 3 in **Figure 3.17**) houses pKPY458 and pKPY273: (i) pKPY458 encodes *Ce*tRNA^{Phe} under *lpp* promoter control as well as FARS-1(N, Met1-Gly147)-Int(N, Gp41-1) under P_{BAD} control.

- *E. coli* **14** (Lane 4 in **Figure 3.17**) houses pKPY93 and pKPY459: pKPY459 encodes both wild-type Int(C, Gp41-1)-FARS-1(C, Ser148-Lys496) and the *Ce* PheRS beta subunit under P_{BAD} control.

- *E. coli* **15** (Lane 5 in **Figure 3.17**) houses pKPY93 and pKPY460: pKPY460 encodes both mutant Int(C, Gp41-1)-FARS-1(C, Ser148-Lys496) and the *Ce*Phe RS beta subunit under P_{BAD} control.

- *E. coli* **16** (Lane 6 in **Figure 3.17**) houses pKPY458 and pKPY459.

- *E. coli* **17** (Lane 7 in **Figure 3.17**) houses pKPY458 and pKPY460.

As in the cases of first and second version of the split-intein-reconstituted split-Thr412Gly-*Ce*PheRS system, only the mutant intact (Lane 2) and reconstituted *Ce*PheRS (Lane 7) displayed robust labeling cells treated with Azf. Encouragingly, replacing the DnaE intein with the Gp41-1 intein and repositioning the split-site from Asn25/Glu26Cys to Gly147/Ser148 significantly enhanced Azf-labeling (relative labeling of 80% compared to intact synthetase). From these observations, we conclude that the third version of the split-intein-reconstituted split-Thr412Gly-*Ce*PheRS system is the best candidate for precise spatiotemporal-selective Azf-labeling in a mul-

ticellular organism like *C. elegans*. For simplification, we refer to FARS-1(N, Met1-Gly147)-Int(N, Gp41-1) as FARS-1(N)-Int(N) and mutant Int(C, Gp41-1)-FARS-1(C, Ser148-Lys496) as Int(C)-FARS-1(C) for the remainder of this chapter.

3.2.2 Characterizing Split-Intein-Reconstituted Split-Thr412Gly-CePheRS in C. elegans

To evaluate the performance of the split-intein-reconstituted split-Thr412Gly-*Ce*Phe RS system in *C. elegans*, we generated transgenic *C. elegans* lines that (i) express both FARS-1(N)-Int(N) and GFP under control of the *hsp-16.2* promoter, a promoter shown previously to be active in multiple tissues in response to heat shock and other stresses [122], and (ii) express both Int(C)-FARS-1(C) and mCherry under control of the *myo-2* promoter, a promoter shown previously to be active in the 20 pharyngeal muscle cells [126] (**Figure 3.18**). We envision exploiting the dual component nature of the split-intein-reconstituted split-synthetase to metabolically label proteins in specific cells at specific times. Upon heat shock and only in pharyngeal muscle, we expect that both FARS-1(N)-Int(N) and Int(C)-FARS-1(C) would be expressed, resulting in the recovery of the full-length mutant alpha subunit after intein fragment association and protein splicing *in trans*. Fully active Thr412Gly-*Ce*PheRS is produced after the mutant alpha subunit forms a hybrid heterotetramer with the endogenous beta subunit.

Specifically, we designed four gene cassettes, each consisting of a promoter, a protein-coding segment, and a terminator: (i) *hsp-16.2* 5'::FARS-1(N)-Int(N)::*fib-1/rps-16*::*gfp*::*unc-54* 3', (ii) *myo-2* 5'::Int(C)-FARS-1(C)::*rpl-16/M01F1.8*::*mCherry*::*let-858* 3', (iii) *rps-27* 5'::*neoR*::*unc-54* 3', and (iv) *hsp-16.41* 5'::*peel-1*::*tbb-2* 3' (**Figure 3.19**). The *fib-1/rps-16* and *rpl-16/M01F1.8* intercistronic regions enable the control of FARS-1(N)-Int(N)/GFP and Int(C)-FARS-1(C)/mCherry, respectively, under the same promoter by instructing the genes encoding FARS-1(N)-Int(N)/GFP

and Int(C)-FARS-1(C)/mCherry to be first transcribed as a single polycistronic precursor mRNA and then trans-spliced into monocistronic mature mRNAs [157]. GFP and mCherry act as fluorescent reporters for the promoter activity of the two split-intein split-synthetase cassettes, confirming that the transgene is present and expressing in the expected cells. The *unc-54*, *let-858*, and *tbb-2* 3' regions act as putative transcriptional terminators that prevent misexpression of any downstream sequences. The positive-selection marker cassette *rps-27* 5'::*neoR*::*unc-54* 3' confers resistance to G-418 (Geneticin), a drug that inhibits protein synthesis in prokaryotes and eukaryotes, and therefore allows hands-off maintenance and enrichment of transgenic worms on G-418-containing plates [158]. The negative-selection marker cassette *hsp-16.41* 5'::*peel-1*::*tbb-2* 3' encodes the potent PEEL-1 toxin [159] under control of the *hsp-16.41* promoter, another promoter shown previously to be active in response to heat shock [122]. We assembled the four cassettes in a single plasmid pKPY728 and flanked three of the four cassettes (*hsp-16.2* 5'::FARS-1(N)-Int(N)::*fib-1/rps-16*::*gfp*::*unc-54* 3', *myo-2* 5'::Int(C)-FARS-1(C)::*rpl-16/M01F1.8*::*mCherry*::*let-858* 3', and *rps-27* 5'::*neoR*::*unc-54* 3' but not *hsp-16.41* 5'::*peel-1*::*tbb-2* 3') with modified *Drosophila Mos1* transposon elements (*miniMos*). Jorgensen and coworkers have demonstrated that the *miniMos* transposon can carry large fragments of DNA, even 45-kb fosmids, into the *C. elegans* genome with high insertion frequency and fidelity [160].

To make stable transgenic *C. elegans* lines, we first microinjected young adult *C. elegans* (N2) with a mixture of pKPY728, fluorescent extrachromosomal array marker plasmids pCFJ90 (*myo-2* 5'::*mCherry*::*unc-54* 3', expression of mCherry in pharyngeal muscle), pCFJ104 (*myo-3* 5'::*mCherry*::*unc-54* 3', expression of mCherry in body wall muscle), and pGH8 (*rab-3* 5'::*mCherry*::*unc-54* 3', expression of mCherry in neurons), and a helper plasmid encoding the *Mos1* transposase pCFJ601 (*eft-3* 5'::*Mos*1 Transposase::*tbb-2* 3', expression of *Mos1* transposase in germline) (**Figure 3.20, Top**). Plasmid DNA injected into the syncytial germ line at the distal gonad

(**Figure 3.21**) concatenates to form semistable and multicopy extrachromosomal DNA arrays [161]. Following microinjection, we obtained several independent extrachromosomal G-418-resistant and mCherry-positive (pharyngeal muscle, body wall muscle, and neurons) lines (**Figure 3.22**). Next, we killed array animals by inducing PEEL-1 expression by 1-h exposure to 37°C to distinguish worms with transposase-mediated genomic insertions that have lost the extrachromosomal array from worms still carrying the array (**Figure 3.20, Bottom**). As a result, we identified two independent G-418-resistant and mCherry-positive (pharyngeal muscle) lines with recombinant *miniMos* insertions: PS7055 (*syTi1*) and PS7058 (*syTi2*) (**Figure 3.23**). We mapped the *syTi1* and *syTi2* insertions by inverse PCR [162] to unique insertion sites in the right arm of the X chromosome (**Figure 3.24**) and in the left arm of chromosome II (**Figure 3.26**), respectively. We also verified the insertion fidelity and genotype with chromosome-specific PCRs spanning across the insertion site, genome-insertion junction-specific PCRs, three-primer PCRs and genomic DNA sequencing (**Figures 3.25 and 3.27**).

Next, we investigated whether Azf-labeling of proteins occurs in a temporally resolved (upon heat shock) and a spatially resolved (in pharyngeal muscle) manner in these *C. elegans* lines. We first labeled bacteria by culturing KY33[pKPY514], a phenylalanine-auxotrophic and G-418-resistant strain of *E. coli* that expresses the *E. coli* variant of Thr412Gly-*Ce*PheRS, in M9 minimal medium supplemented with Azf. We next fed Azf-labeled bacteria to PS7055 and PS7058 worms previously maintained on OP50-1[pKPY562], a G-418-resistant strain of *E. coli*, and compared protein labeling after exposure to non-heat-shock and heat-shock conditions (**Figure 3.28**). For both conditions, we transferred worms to solid nematode growth medium without peptone (NGM-peptone) plates containing Azf-labeled bacteria, incubated plates at 24°C for 24 h, and cleared external as well as ingested bacteria by washing worms with S medium over a period of 30 min. For heat-shock conditions, we introduced a

30-min exposure to 37°C immediately after transferring worms to solid NGM-peptone plates. As expected, fluorescence microscopy of live animals revealed that only heat-shocked worms expressed GFP ubiquitously. Following fixation, worms were subjected to strain-promoted conjugation to TAMRA-DBCO to visualize sites of Azf incorporation. Fluorescence microscopy showed that only heat-shocked worms exhibited pharyngeal-muscle-specific Azf-labeling. Importantly, neither non-heat-shocked worms nor the tissues surrounding the pharyngeal muscles in heat-shocked worms exhibited Azf-labeling. These results provided a proof of principle that coupling the combinatorial expression of two promoters with the split-intein-reconstituted split-Thr412Gly-*Ce*PheRS allows precise spatiotemporally selectivity of Azf-labeling of proteins even in living multicellular animals.

3.3 Conclusions

We have developed a previously undescribed way of generating a two-component system involving reconstituted aminoacyl-tRNA synthetase for selective and/or conditional metabolic labeling of targeted cells: individually expressed aminoacyl-tRNA synthetase fragments are brought together by split inteins to produce the active aminoacyl-tRNA synthetase. While this strategy works for the *C. elegans* PheRS, this mode of reconstitution should apply to any aminoacyl-tRNA synthetase. In addition, the finding that split-intein-mediated reconstituted split-Thr412Gly-*Ce*PheRS is active in both *E. coli* and *C. elegans* indicates that this dual-component system will be effective in a broad range of organisms.

3.4 Figures

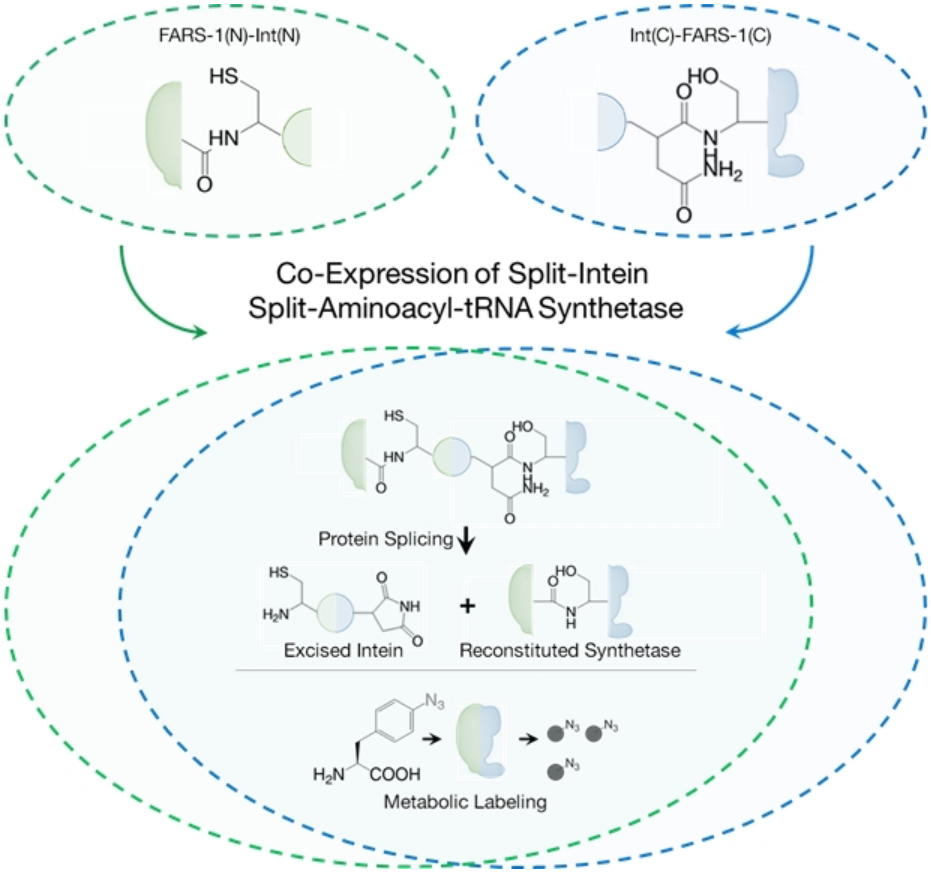


Figure 3.1: *Coupling the combinatorial expression of two promoters with the split-intein-reconstituted split-Thr412Gly-CePheRS allows precise spatiotemporally selectivity of Azf-labeling of proteins.*

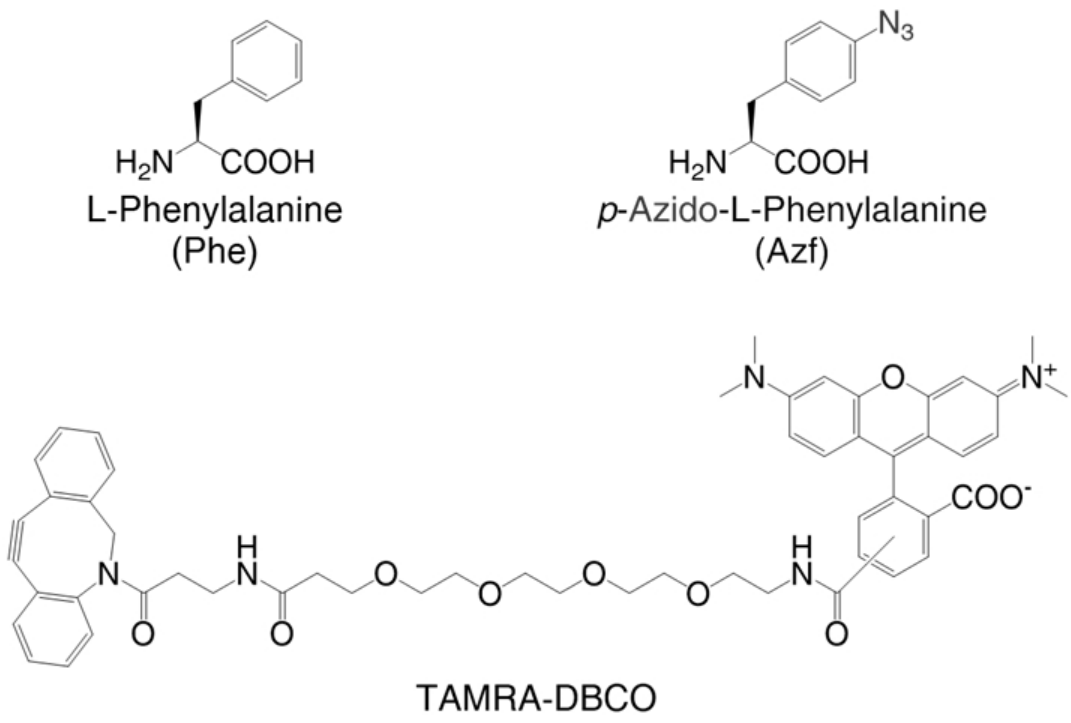


Figure 3.2: *Structures of amino acids and probes used in this study:* L*-phenylalanine (Phe), p-azido-*L*-phenylalanine (Azf), and dibenzocyclooctyne tetramethylrhodamine (TAMRA-DBCO).*

DnaE(N)-Int(N, DnaE);

MSFVPLHIHSDYSLLDGASQLPELVDQAIALGMKAIALT
DHGVMYGAVELIKICRSQNIKPIIGNEMYVINGDIEKQE
RRPKYHQVVLAKNTKGYKNLVKLTTISHLQGVQGKGIFS
RPCINKDLLKQYHEGLIVTSACLGGEVPQAILSNRPDAA
RKVAQWYKDVFGDDYYLEIQDHGSQEDRIVNVEIVKIAR
ELGIKIIATNDSHFISCFDVEAHDALLCIQTGKLIIEDK
RMRYSGTEYLKSGEEMRQLFRDHLPDDVISEAVATTEEV
ADKVEPYHIMGEPQIPTPPIPSGHTADTYAEDVAWNGLL
ERLNRKSRQEVDSVYKERLEYELKMIQQMGFSKYFLVVW
DYIKFARDNNIPVGPGRGSAAGSLVAYAMRITNIDPVHH
GLLFERFLNPERKSMPDIDTDFCIEQRDKVIEYVTEKYG
ADRVAQIITFNRLTSKAVLKDVARVLNIPYGEADKMAKL
IPVVRGKPTKLKVMVSDKTPEPEFKEKYDKEPHVRHWLD
MAMRIEGTNKTFGVHAAGVVISDEPLDEIVPLQKNNDGS
VITQYFMEDLESMGLLKMDFLGLRNLTLIQKTVDLIQET
RGYRVDPDEIPRQERKAQKILAKGEHSSLPKDVQKTYEL
LEAGELEGIFQLESSGMRQIVRDLKPSNIEDISSILALY
RPGPLDAGLIPKFINRKHGRENIDYQHTVLEPILDETYG
IMVYQEQIMKIAQDMAGYSLGQADLLRRAMGKKKVSEMQ
KQREKFVDGAAKNGVPKKVADELFEQMLKFAEYCLSYET
EILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVA
QWHDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPI
DEIFERELDLMRVDNLPN;

Figure 3.3: *dnaE-n contains the N-terminal segment (DnaE(N)) of the DnaE protein followed by a 102-amino acid N-intein (Int(N, DnaE).*

Int(C, DnaE)-DnaE(C);

MIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASNCFN
KSHSTAYGYVTYQTAYLKANYPLEYMAALLTANSGDTDK
VQRYITNCTNMGISIDPPDINRSGVDFTPTLGKILFGFS
AVRNVGQNAIACILEARNETGEFKSLADFCDRVDLRAVN
RRTLESLIYCGAFDKIESNRQQLINDSELVYDWAQSRAK
DRASGQGNLFDLLGDGFSSTQNKRVNNAFETAPKSKPVT
DLPPQKKLQMEKELLGFYVSDHPLKSLRQIAPLLTPINL
SQLGEQREDTRLCAVVMLNNVKKVVTKKGDQMAILQIED
LTTQSEAVVFPKTYERISSLLQVDTRLIIWGKVDRRDEQ
TQFIVEDAEPVETVQMVMVELNPQQAGDMEKLHLLKTIL
QEHSVDKEKAKMPVIGIIQTEKSRKLVRLGWQFSVQDSR
ITVQALQNASFPAHIKSLTGS;

Figure 3.4: *dnaE-c contains a 36-residue C-intein (Int(C, DnaE) followed by the C-terminal segment (C extein, DnaE(C)) of the DnaE protein.*

FARS-1 (Thr412Gly);

MTAETDVRSTENLPQQILDFLQESNEFNSIQLAQQWNLD
HQKVIGAIKSLLANEGVLTTKDVTEKRLELTNEGVQFAN
EGSPEYLVFEFVGTDGAAQADIQKKPFGKIGMAKAMQFK
WVSVDKGRVVRQATEVTDSTRKQLESLRIGSSDVSENEK
KELKKRKLISEVNIKALVVSKGTSFTTSLAKQEADLTPE
MIASGSWKDMQFKKYNFDSLGVVPSSGHLHPLMKVRSEF
RQIFFSMGFSEMATNRYVESSFWNFDALFQPQQHPARDA
HDTFFVSDPAISTKFPEDYLERVKTVHSKGGYGSAGYNY
DWKIEEAQKNVLRTHTTAVSARQLYQLAQEGFRPSKLFS
IDRVFRNETLDATHLAEFHQVEGVIAEKNLSLAHLIGIF
TEFFKKLGITNLRFKPTYNPYGEPSMEIFAYHQGLTKWV
EIGNSGMFRPEMLLPMGLPADVNVAGYGLSLERPTMIKY
GINNIRDLFGSKIDLNVVYNNPICRLDK;

Figure 3.5: *C. elegans FARS-1 Sequence.*

FARS-1(N, Met1-Lys187)-Int(N, DnaE);

MTAETDVRSTENLPQQILDFLQESNEFNSIQLAQQWNLD
HQKVIGAIKSLLANEGVLTTKDVTEKRLELTNEGVQFAN
EGSPEYLVFEFVGTDGAAQADIQKKPFGKIGMAKAMQFK
WVSVDKGRVVRQATEVTDSTRKQLESLRIGSSDVSENEK
KELKKRKLISEVNIKALVVSKGTSFTTSLAKCLSYETEI
LTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQW
HDRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDE
IFERELDLMRVDNLPN;

Figure 3.6: *FARS-1(N, Met1-Lys187)-Int(N, DnaE) Sequence.*

Int(C, DnaE)-*Cys-Phe-Asn*-FARS-1(C,
Gln188-Lys496, Thr412Gly);

MIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN*CFN*
QEADLTPEMIASGSWKDMQFKKYNFDSLGVVPSSGHLHP
LMKVRSEFRQIFFSMGFSEMATNRYVESSFWNFDALFQP
QQHPARDAHDTFFVSDPAISTKFPEDYLERVKTVHSKGG
YGSAGYNYDWKIEEAQKNVLRTHTTAVSARQLYQLAQEG
FRPSKLFSIDRVFRNETLDATHLAEFHQVEGVIAEKNLS
LAHLIGIFTEFFKKLGITNLRFKPTYNPYGEPSMEIFAY
HQGLTKWVEIGNSGMFRPEMLLPMGLPADVNVAGYGLSL
ERPTMIKYGINNIRDLFGSKIDLNVVYNNPICRLDK;

Figure 3.7: *Int(C, DnaE)-Cys-Phe-Asn-FARS-1(C, Gln188-Lys496) Sequence.*

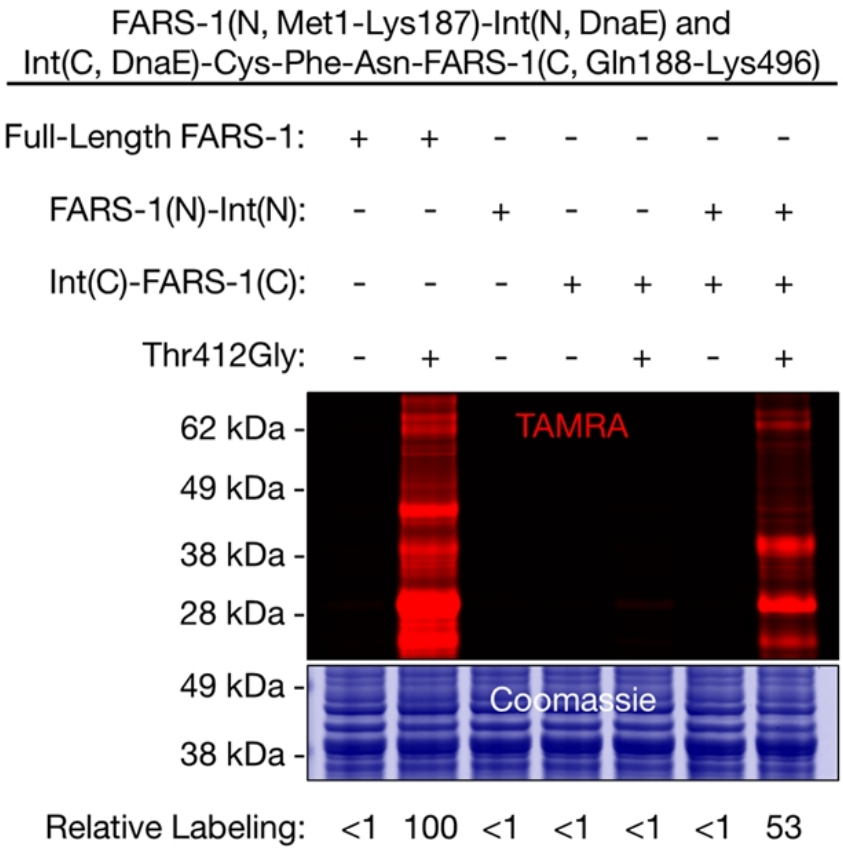


Figure 3.8: *SDS-PAGE and in-gel fluorescence scanning detection of strain-promoted conjugation of TAMRA-DBCO to Azf-labeled protein from E. coli* ***1***, ***2***, ***3***, ***4***, ***5***, ***6***, *and* ***7*** *lysates.*

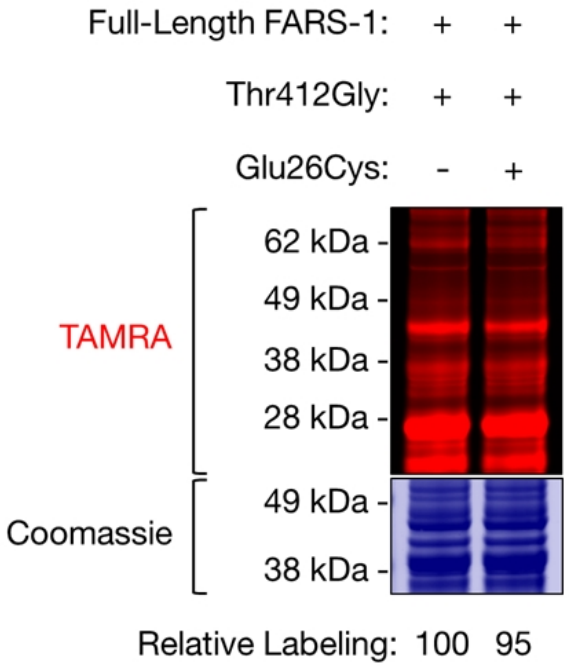


Figure 3.9: *SDS-PAGE and in-gel fluorescence scanning detection of strain-promoted conjugation of TAMRA-DBCO to Azf-labeled protein from E. coli expressing either Thr412Gly-CePheRS or Glu26Cys-Thr412Gly-CePheRS.*

FARS-1(N, Met1-Asn25)-Int(N, DnaE);

MTAETDVRSTENLPQQILDFLQESNCLSYETEILTVEYG
LLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQ
EVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEIFEREL
DLMRVDNLPN;

Figure 3.10: *FARS-1(N, Met1-Asn25)-Int(N, DnaE) Sequence.*

Int(C, DnaE)-FARS-1(C, Glu26Cys-Lys496, Thr412Gly);

MIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASNCFN
SIQLAQQWNLDHQKVIGAIKSLLANEGVLTTKDVTEKRL
ELTNEGVQFANEGSPEYLVFEFVGTDGAAQADIQKKPFG
KIGMAKAMQFKWVSVDKGRVVRQATEVTDSTRKQLESLR
IGSSDVSENEKKELKKRKLISEVNIKALVVSKGTSFTTS
LAKQEADLTPEMIASGSWKDMQFKKYNFDSLGVVPSSGH
LHPLMKVRSEFRQIFFSMGFSEMATNRYVESSFWNFDAL
FQPQQHPARDAHDTFFVSDPAISTKFPEDYLERVKTVHS
KGGYGSAGYNYDWKIEEAQKNVLRTHTTAVSARQLYQLA
QEGFRPSKLFSIDRVFRNETLDATHLAEFHQVEGVIAEK
NLSLAHLIGIFTEFFKKLGITNLRFKPTYNPYGEPSMEI
FAYHQGLTKWVEIGNSGMFRPEMLLPMGLPADVNVAGYG
LSLERPTMIKYGINNIRDLFGSKIDLNVVYNNPICRLDK
;

Figure 3.11: *Int(C, DnaE)-FARS-1(C, Glu26Cys-Lys496) Sequence.*

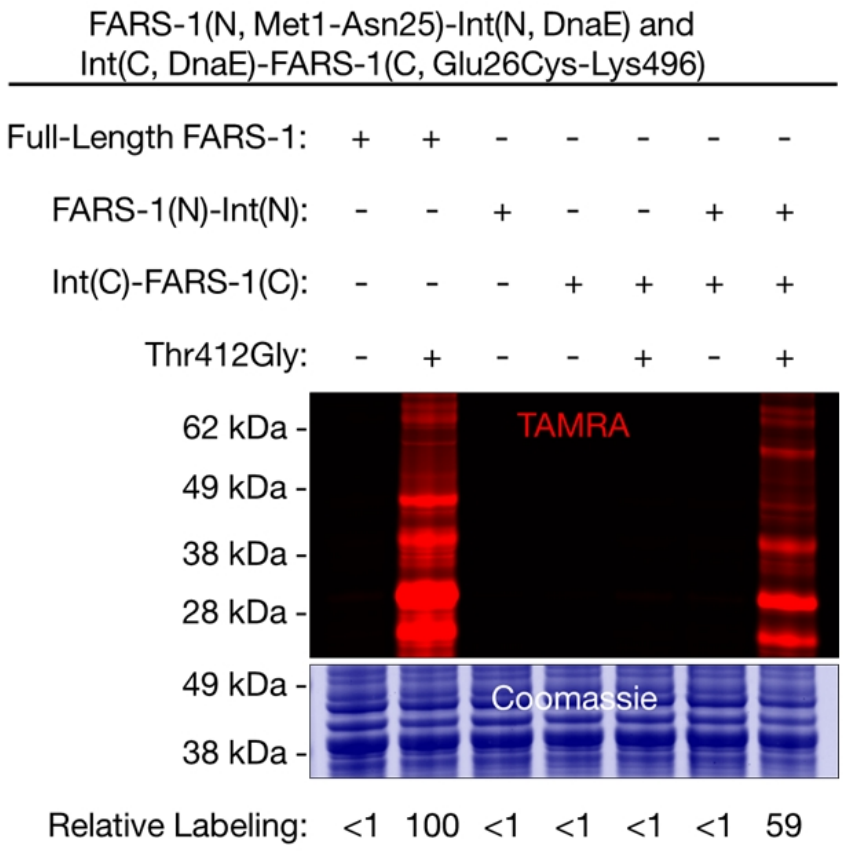


Figure 3.12: *SDS-PAGE and in-gel fluorescence scanning detection of strain-promoted conjugation of TAMRA-DBCO to Azf-labeled protein from E. coli* ***1***, ***2***, ***8***, ***9***, ***10***, ***11***, *and* ***12*** *lysates.*

Gp41-1(N)-Int(N, Gp41-1);

MERIETTILRNLVFNEDFSRKVIPFIEPDYFEERKEKII
FEEVTKFIVKYGSAITVEALNIEIENRTDLNESEIKETR
DISNTLHDSAVEPQWLLDTTEKWCRDRAIYLALMESIHI
ADGEDEQKNRDAIPSILSDALAVSFDSHIGHDYLNDYEE
RYESYHRKEDKIPFDLEYFDKITKGGLPNKTLNIALAGT
GVGKSLFMCHMASSVLLQGKNVLYITLEMAEEKIAERID
ANLLNVNIQNITDLPKPMFENKVSSLTKKTQGSLIIKEY
PTASAHSGHFKSLLQELALKKSFRPDIIFIDYLNICASS
RYRQNASVNSYSFIKAIAEELRGLAVEANLPIVSATQTT
RSGFACLDLKTQVQTPQGMKEISNIQVGDLVLSNTGYNE
VLNVFPKSKKKSYKITLEDGKEIICSEEHLFPTQTGEMN
ISGGLKEGMCLYVKE;

Figure 3.13: *gp41-1-n contains the N-terminal segment (N extein, Gp41-1(N)) of the Gp41-1 protein followed by an 88-amino acid N-intein (Int(N, Gp41-1)).*

Int(C, Gp41-1)-Gp41-1(C);

MMLKKILKIEELDERELIDIEVSGNHLFYANDILTHNSS
SDVDLTDTSESFGLPATADLMFALISTEELEGLNQIMVK
QLKNRYNDPTIFKRFIVGIDRAKMRLYDVEQKAQEDILD
SGKEEEYDPHEEKKPKKSFAGFKFN;

Figure 3.14: *gp41-1-c contains a 38-residue C-intein (Int(C, Gp41-1)) followed by the C-terminal segment (C extein, Gp41-1(C)) of the Gp41-1 protein.*

FARS-1(N, Met1-Gly147)-Int(N, Gp41-1);

MTAETDVRSTENLPQQILDFLQESNEFNSIQLAQQWNLD
HQKVIGAIKSLLANEGVLTTKDVTEKRLELTNEGVQFAN
EGSPEYLVFEFVGTDGAAQADIQKKPFGKIGMAKAMQFK
WVSVDKGRVVRQATEVTDSTRKQLESLRIGCLSYETEIL
TVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWH
DRGEQEVFEYCLEDGSLIRATKDHKFMTVDGQMLPIDEI
FERELDLMRVDNLPN;

Figure 3.15: *FARS-1(N, Met1-Gly147)-Int(N, Gp41-1) Sequence.*

Int(C, Gp41-1)-FARS-1(C, Ser148-Lys496, Thr412Gly);

MMLKKILKIEELDERELIDIEVSGNHLFYANDILTHNSS
SDVSENEKKELKKRKLISEVNIKALVVSKGTSFTTSLAK
QEADLTPEMIASGSWKDMQFKKYNFDSLGVVPSSGHLHP
LMKVRSEFRQIFFSMGFSEMATNRYVESSFWNFDALFQP
QQHPARDAHDTFFVSDPAISTKFPEDYLERVKTVHSKGG
YGSAGYNYDWKIEEAQKNVLRTHTTAVSARQLYQLAQEG
FRPSKLFSIDRVFRNETLDATHLAEFHQVEGVIAEKNLS
LAHLIGIFTEFFKKLGITNLRFKPTYNPYGEPSMEIFAY
HQGLTKWVEIGNSGMFRPEMLLPMGLPADVNVAGYGLSL
ERPTMIKYGINNIRDLFGSKIDLNVVYNNPICRLDK;

Figure 3.16: *Int(C, Gp41-1)-FARS-1(C, Ser148-Lys496) Sequence.*

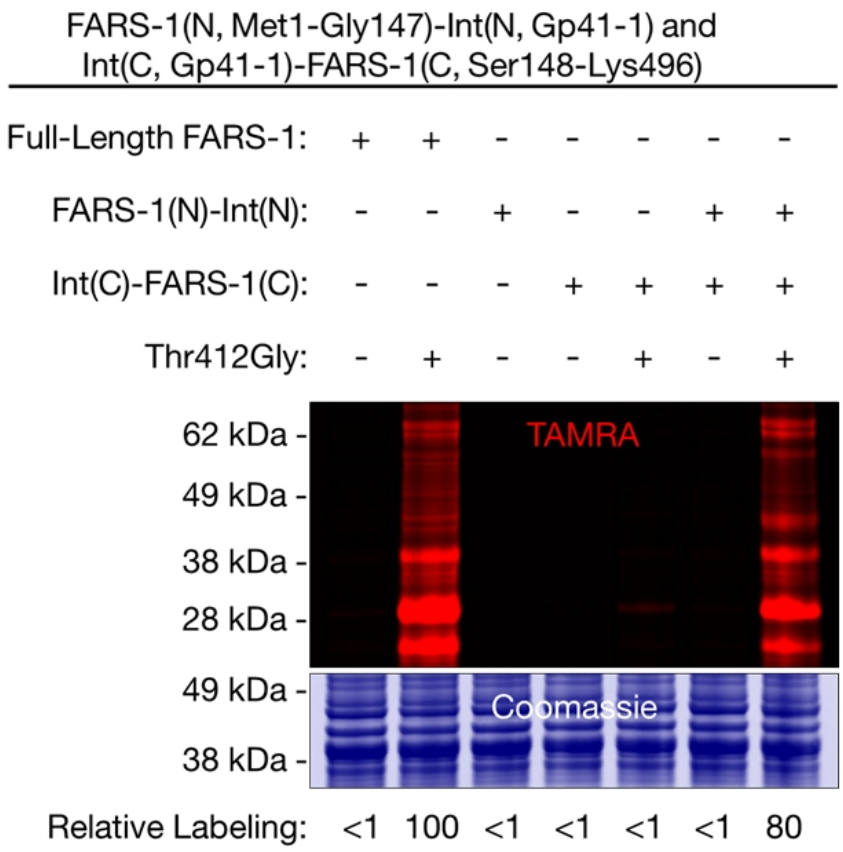


Figure 3.17: *SDS-PAGE and in-gel fluorescence scanning detection of strain-promoted conjugation of TAMRA-DBCO to Azf-labeled protein from E. coli* ***1***, ***2***, ***13***, ***14***, ***15***, ***16***, *and* ***17*** *lysates.*

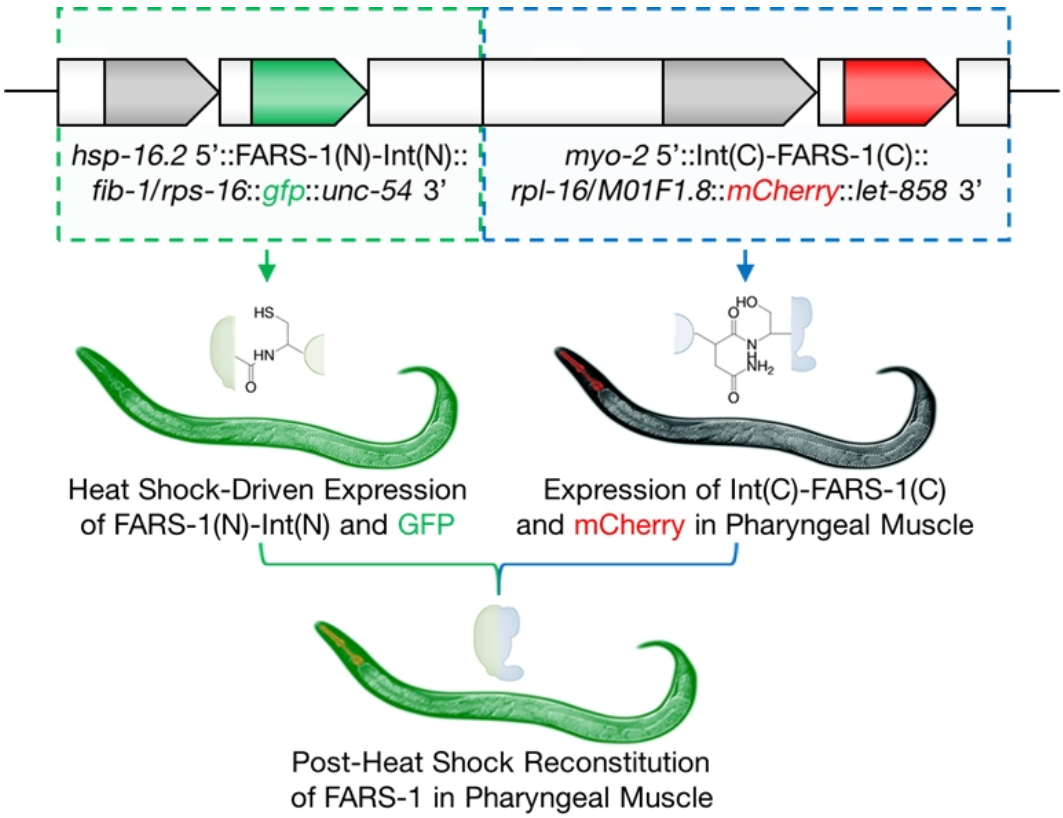


Figure 3.18: *Upon heat shock and only in pharyngeal muscle, we expect that both FARS-1(N)-Int(N) and Int(C)-FARS-1(C) would be expressed, resulting in the recovery of the full-length mutant alpha subunit after intein fragment association and protein splicing in trans.*

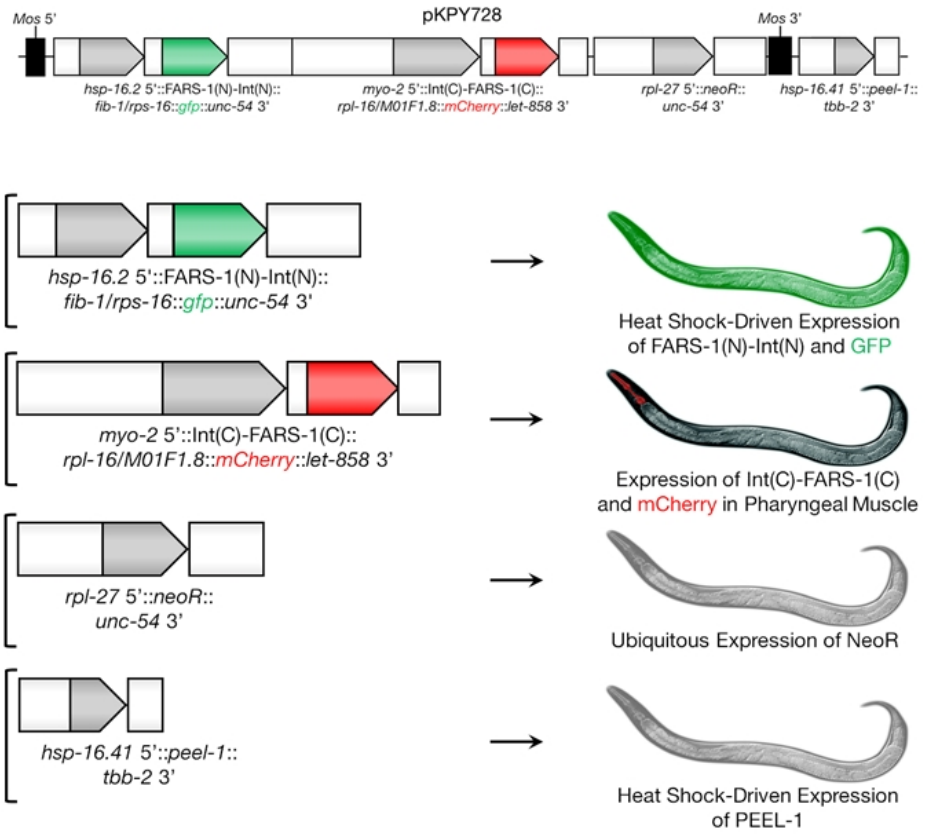


Figure 3.19: *pKPY728 contains four gene cassettes, each consisting of a promoter, a protein-coding segment, and a terminator: (i) hsp-16.2 5'::FARS-1(N)-Int(N)::fib-1/rps-16::gfp::unc-54 3', (ii) myo-2 5'::Int(C)-FARS-1(C)::rpl-16/M01F1.8::mCher ry::let-858 3', (iii) rps-27 5'::neoR::unc-54 3', and (iv) hsp-16.41 5'::peel-1::tbb-2 3'.*

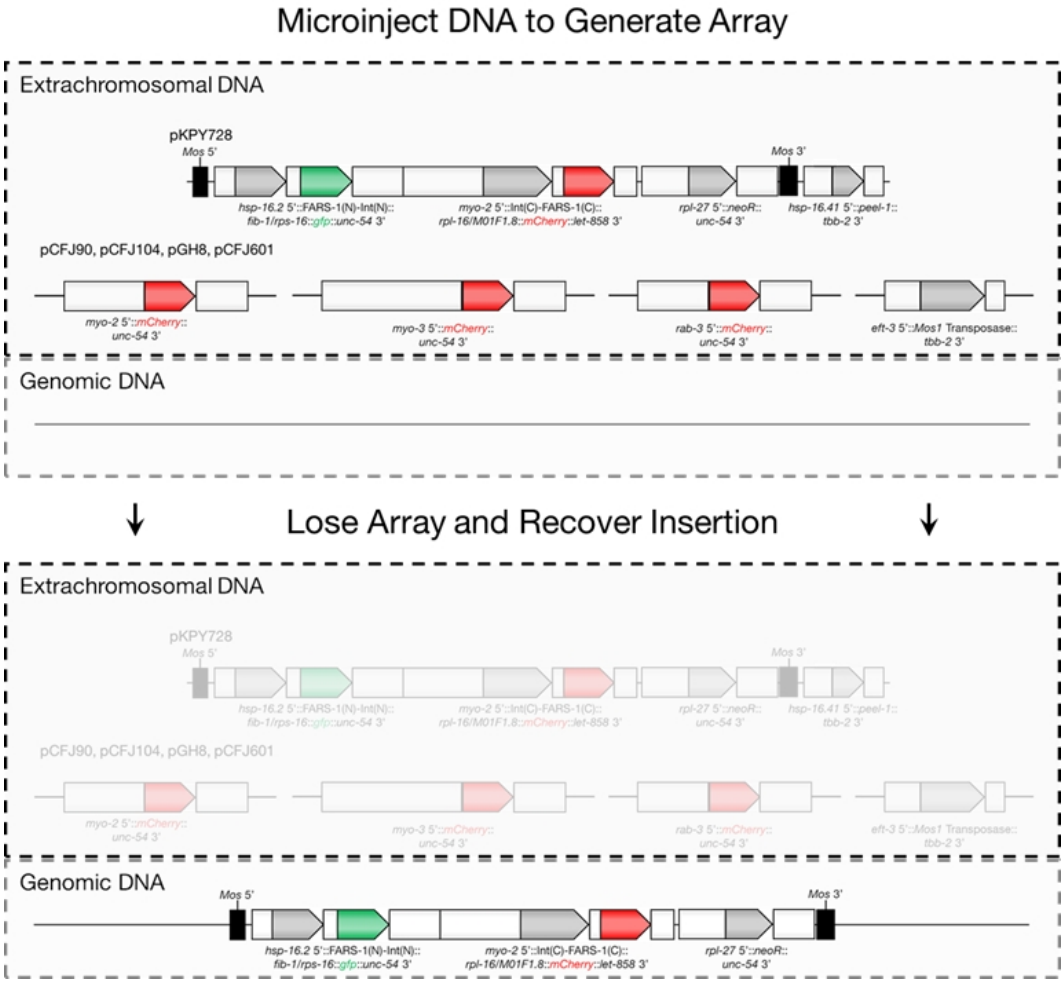


Figure 3.20: *We microinjected worms with a mixture of pKPY728, pCFJ90, pCFJ104, pGH8, and pCFJ601 to generate worms with these plasmids concatenated to semistable and multicopy extrachromosomal DNA arrays. Array animals are killed by inducing PEEL-1 expression by heat shock to distinguish worms with transposase-mediated genomic insertions that have lost the extrachromosomal array from worms still carrying the array.*

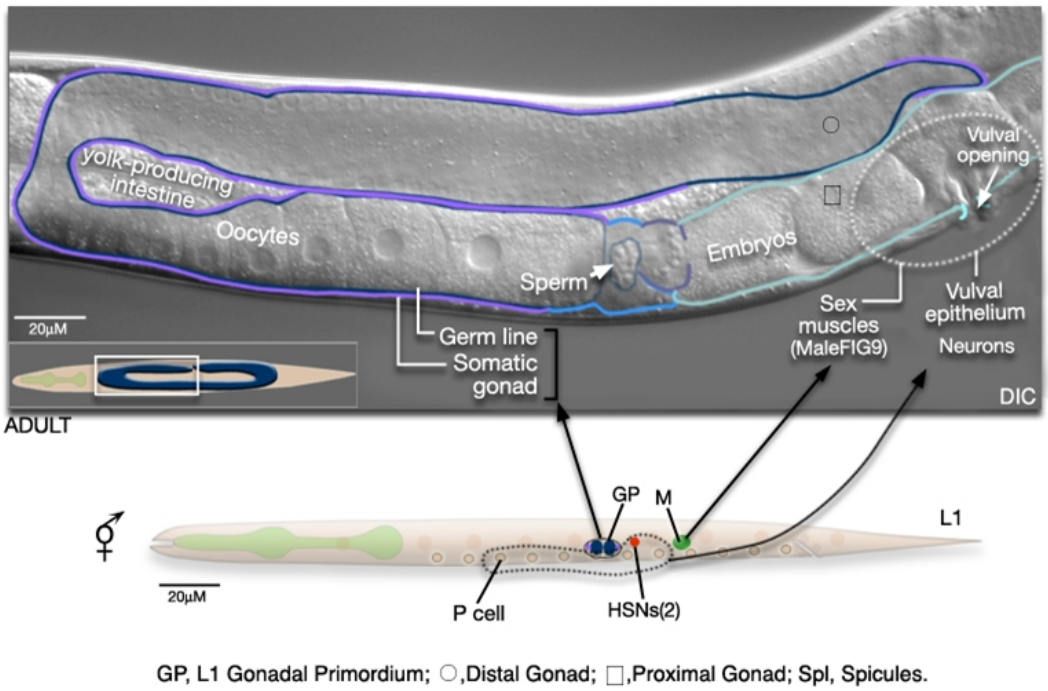


Figure 3.21: *C. elegans Hermaphrodite Gonad.*

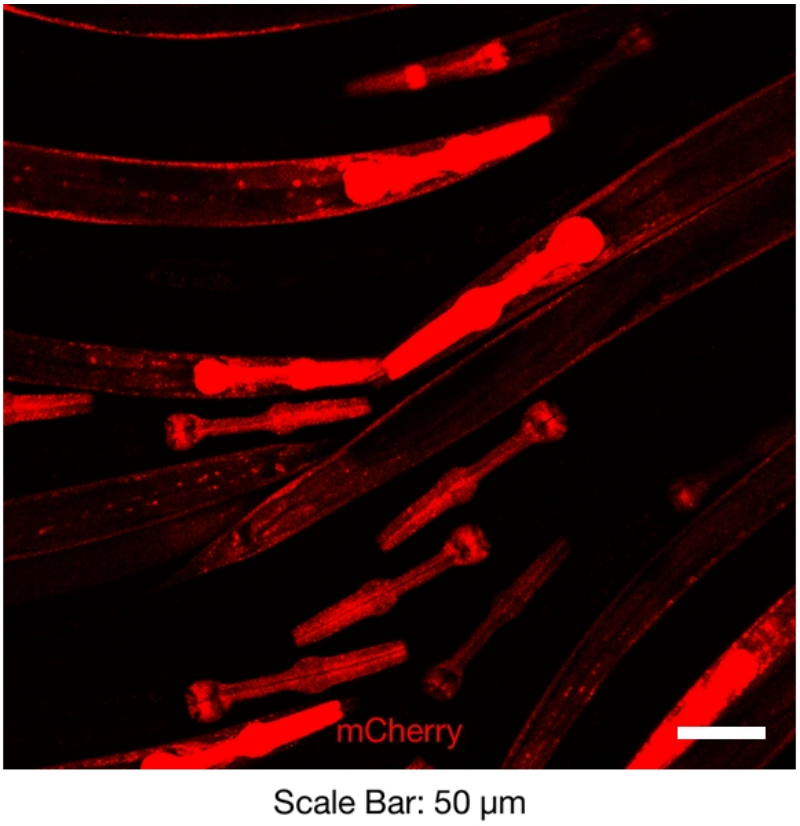


Figure 3.22: *Fluorescence microscopy of an extrachromosomal G-418-resistant and mCherry-positive (pharyngeal muscle, body wall muscle, and neurons; red) line.*

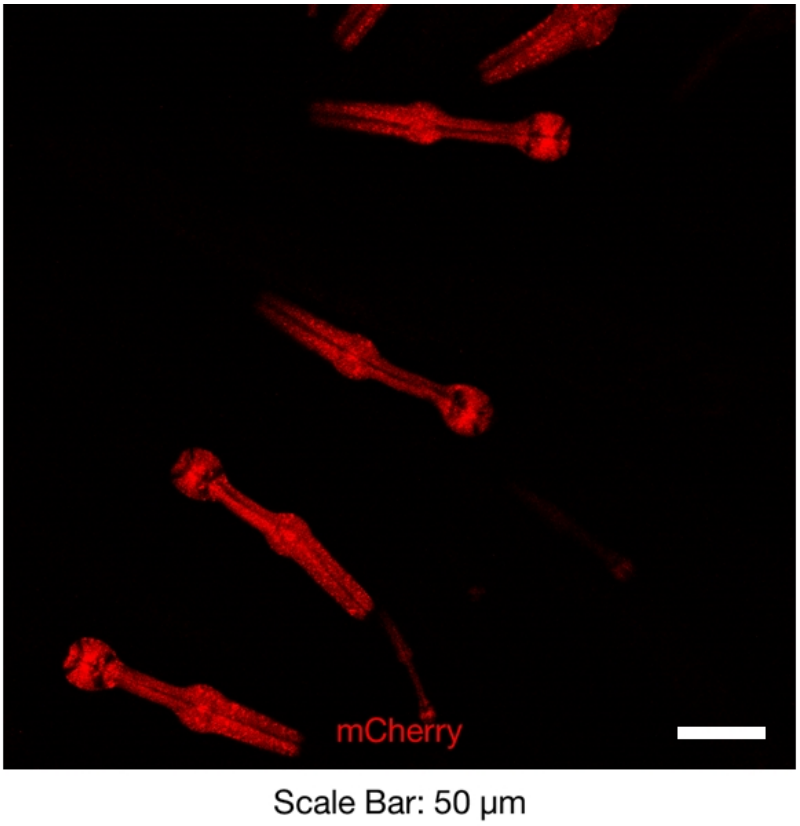


Figure 3.23: *Fluorescence microscopy of the G-418-resistant and mCherry-positive (pharyngeal muscle; red) line PS7055 containing the recombinant miniMos insertions syTi1.*

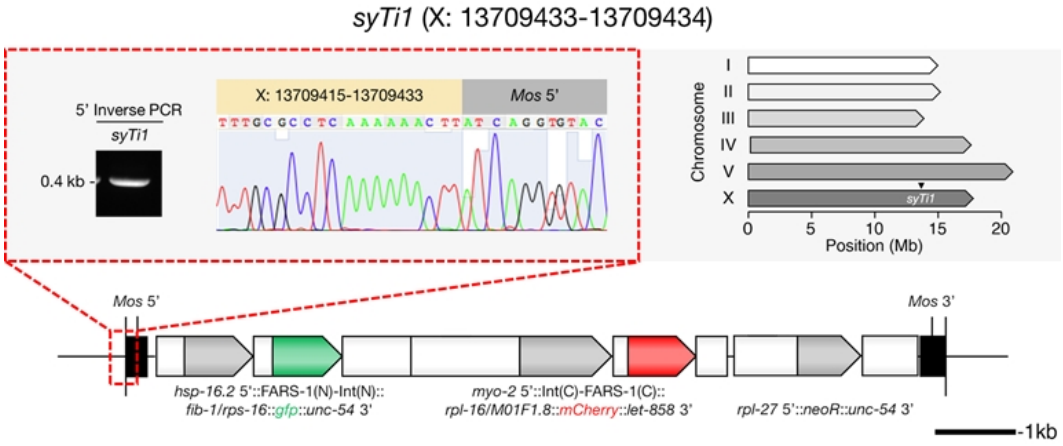


Figure 3.24: *We mapped the syTi1 insertion by inverse PCR to a unique insertion sites in the right arm of the X chromosome.*

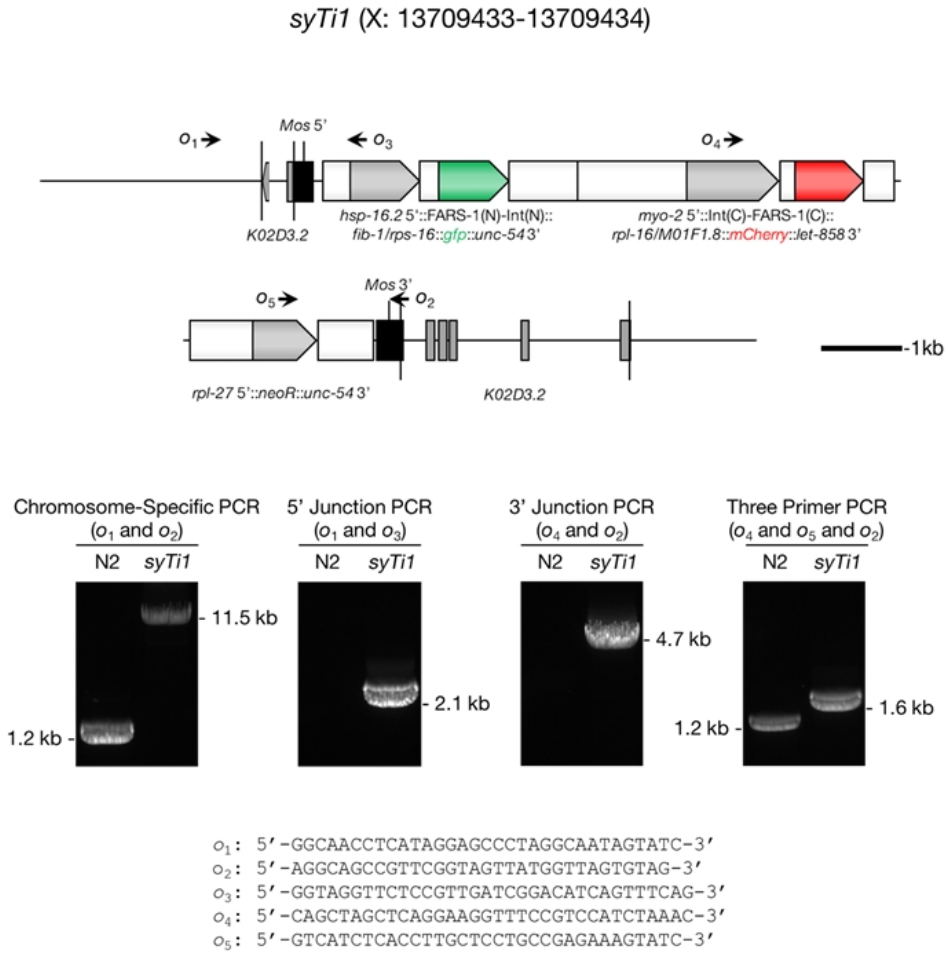


Figure 3.25: *We verified the syTi1 insertion fidelity and genotype with chromosome-specific PCRs spanning across the insertion site, genome-insertion junction-specific PCRs, and three-primer PCRs.*

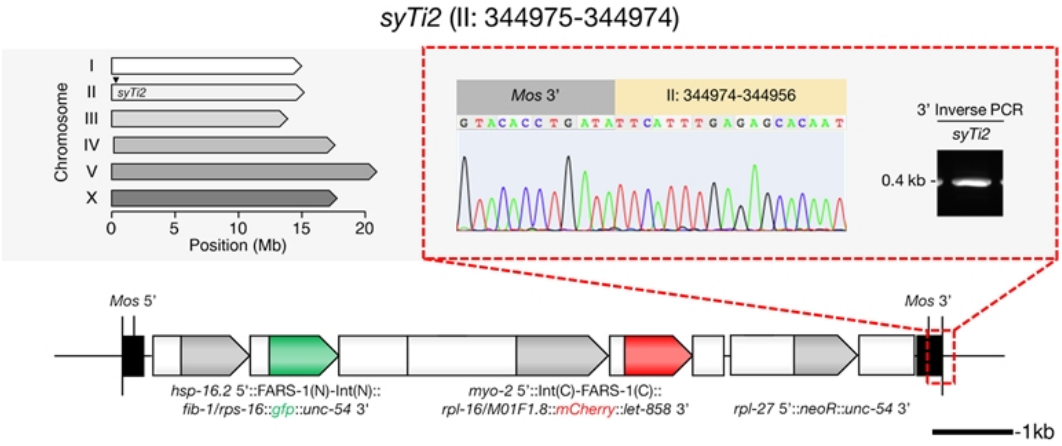


Figure 3.26: *We mapped the syTi2 insertion by inverse PCR to a unique insertion sites in the left arm of chromosome II.*

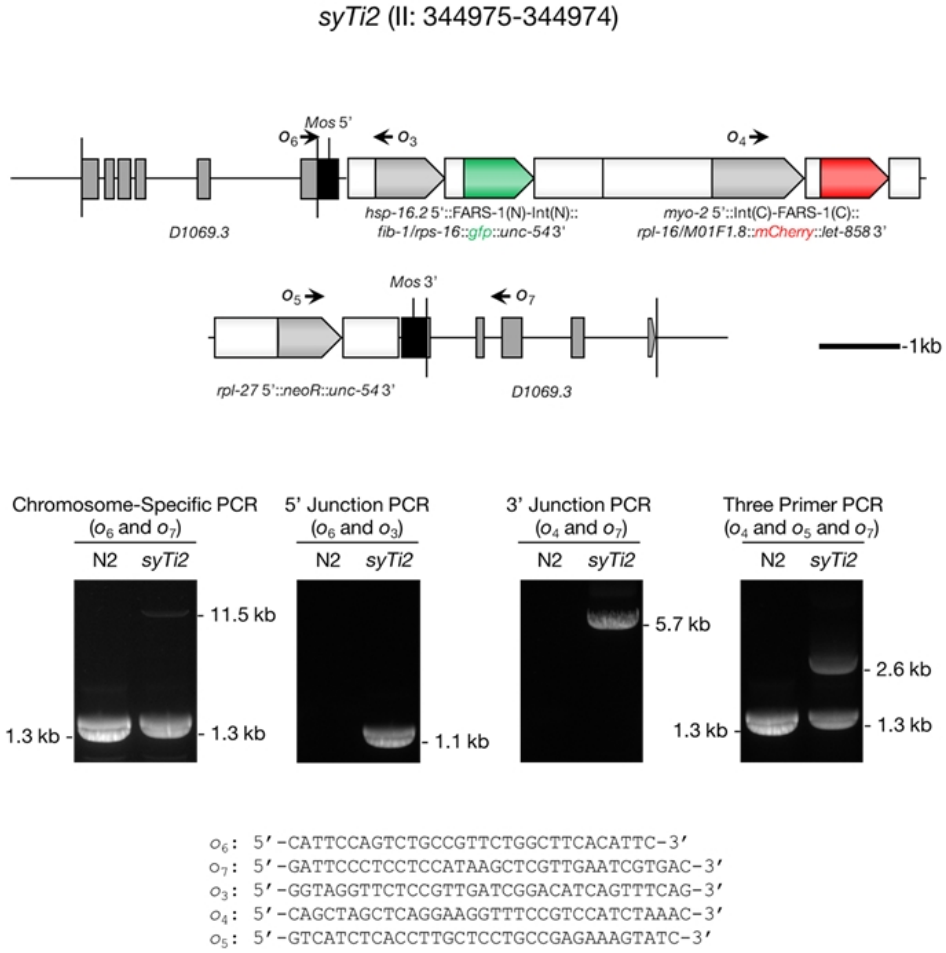


Figure 3.27: *We verified the syTi2 insertion fidelity and genotype with chromosome-specific PCRs spanning across the insertion site, genome-insertion junction-specific PCRs, and three-primer PCRs.*

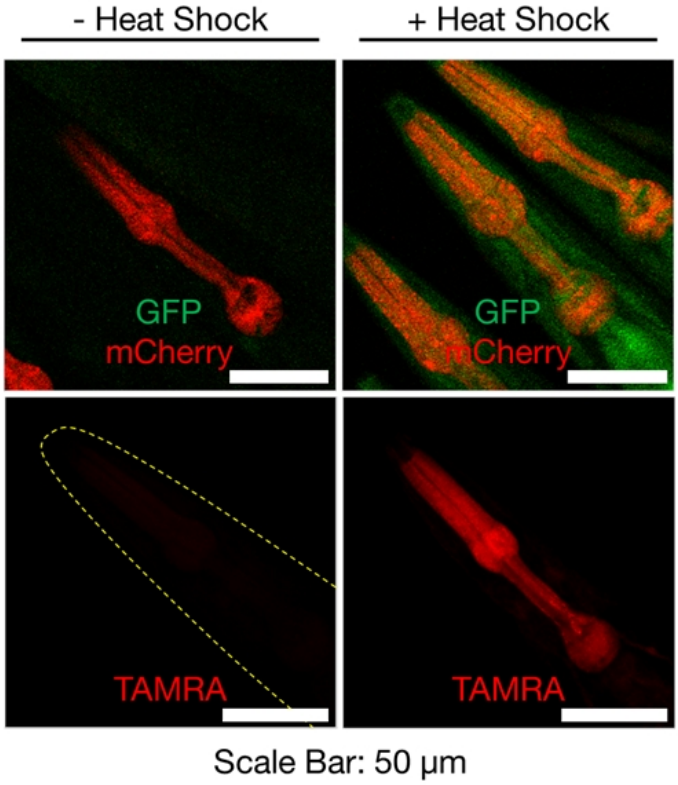


Figure 3.28: *We fed Azf-labeled E. coli strain KY33[pKPY514] to PS7055 and PS7058. These transgenic C. elegans strains (i) express both FARS-1(N)-Int(N) and GFP (green) under control of the hsp-16.2 promoter and (ii) express both Int(C)-FARS-1(C) and mCherry (red) under control of the myo-2 promoter. (Top) Only heat-shocked worms express GFP ubiquitously (PS7055 shown). (Bottom) Azf-labeling in PS7055 and PS7058 were localized to pharyngeal muscle only upon heat shock; the surrounding tissues were not labeled (PS7055 shown).*

3.5 Materials and Methods

3.5.1 Chloroform/Methanol Precipitation

Methanol (EMD Millipore, 2.40 volumes) was added to 1.00 volume of protein solution and vortexed. Chloroform (EMD Millipore, 0.80 volume) was added and vortexed. Double-distilled water (3.20 volumes) was added and vortexed. The mixture was centrifuged at 20,000 x g for 20 minutes at room temperature, and the top layer was discarded. Methanol (5.00 volumes) was added and vortexed, and the mixture was centrifuged at 20,000 x g for five minutes at room temperature. The last three steps were repeated for a total of five times. The top layer was discarded, and the protein pellet was air-dried.

3.5.2 Fluorescence Microscopy of Live C. elegans

Anesthetizing Solution: In double-distilled water, 6.8 g/L sodium phosphate dibasic anhydrous (VWR International), 3 g/L potassium phosphate monobasic (VWR International), 0.5 g/L sodium chloride (VWR International), 1 g/L ammonium chloride (VWR International), 2 mM magnesium sulfate heptahydrate (VWR International), 0.1 mM calcium chloride dihydrate (VWR International), 20 mM D-glucose (VWR International), 25 mM sodium azide (Sigma-Aldrich Corporation)

C. elegans were re-suspended in Anesthetizing Solution. Anesthetized worms were mounted on 3% (w/v) UltraPure Low Melting Point Agarose (Life Technologies) pads and imaged with a LSM 510 META microscope (Carl Zeiss AG) at the Biological Imaging Center in the Beckman Institute at Caltech. Samples were protected from light at all times.

3.5.3 Fluorescence Microscopy of Fixed C. elegans

Alkylating Solution: In double-distilled water, 100 mM Trizma base (Sigma-Aldrich Corporation), pH 8.0, 100 mM 2-chloroacetamide (Sigma-Aldrich Corporation)

C. elegans Fixing Solution: 160 mM potassium chloride (VWR International), 40 mM sodium chloride (VWR International), 20 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid tetrasodium salt (Sigma-Aldrich Corporation), 10 mM spermidine trihydrochloride (Sigma-Aldrich Corporation), 30 mM PIPES, pH 7.4, 50% (v/v) methanol (EMD Millipore), 1.6% (w/v) paraformaldehyde (Electron Microscopy Sciences)

Permeabilizing Solution A: In 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (Life Technologies), 1% (v/v) Triton X-100 (Sigma-Aldrich Corporation), 0.1% (w/v) bovine serum albumin (Sigma-Aldrich Corporation), 0.05% (w/v) sodium azide (Sigma-Aldrich Corporation)

Permeabilizing Solution B: In 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (Life Technologies), 10% (v/v) dimethyl sulfoxide (VWR International), 1% (v/v) Triton X-100 (Sigma-Aldrich Corporation), 0.1% (w/v) bovine serum albumin (Sigma-Aldrich Corporation), 0.05% (w/v) sodium azide (Sigma-Aldrich Corporation)

Reducing Solution: In double-distilled water, 100 mM Trizma base (Sigma-Aldrich Corporation), pH 8.0, 10 mM tris(2-carboxyethyl)phosphine hydrochloride (Thermo Fisher Scientific)

TAMRA-DBCO Solution: 1 mM TAMRA-DBCO (Click Chemistry Tools) in dimethyl sulfoxide, anhydrous (Life Technologies)

Washing Solution A: In double-distilled water, 100 mM Trizma base (Sigma-Aldrich Corporation), pH 8.0

Washing Solution B: In 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (Life Technologies), 20% (v/v) dimethyl sulfoxide (VWR In-

ternational), 1% (v/v) Triton X-100 (Sigma-Aldrich Corporation), 0.1% (w/v) bovine serum albumin (Sigma-Aldrich Corporation), 0.05% (w/v) sodium azide (Sigma-Aldrich Corporation)

Washing Solution C: In 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (Life Technologies), 20% (v/v) methanol (EMD Millipore), 1% (v/v) Triton X-100 (Sigma-Aldrich Corporation), 0.1% (w/v) bovine serum albumin (Sigma-Aldrich Corporation), 0.05% (w/v) sodium azide (Sigma-Aldrich Corporation)

Cell or worm pellets were re-suspended in double-distilled water (1.00 volume) and incubated for five minutes in ice. To fix animals and crack their cuticles, ice-cold *C. elegans* Fixing Solution (1.00 volume) was added to the mixtures of worms and incubated for five minutes in liquid nitrogen. Frozen solutions were thawed for five minutes in room temperature water. The last two freeze-thaw steps were repeated for a total of three times. Thawed mixtures were incubated for an additional hour in ice. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. Washing Solution A (2.00 volumes) was added and mixed, and the mixtures were centrifuged at 1,000 x g for one minute at room temperature. The last three steps were repeated for a total of five times. The top aqueous layer was discarded. To reduce animals' cuticles, Reducing Solution (2.00 volumes) was added and mildly agitated for 30 minutes at room temperature. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. To alkylate animals' cuticles, Alkylating Solution (2.00 volumes) was added and mildly agitated for 30 minutes at room temperature. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. 1X Dulbecco's phosphate buffered saline, no calcium chloride, no magnesium chloride (2.00 volumes, Life Technologies) was added and mixed. The mixtures were centrifuged at 1,000 x g for one minute at

room temperature. The last three steps were repeated for a total of five times. The top aqueous layer was discarded. Permeabilizing Solution A (2.00 volumes) was added and mildly agitated for one hour at room temperature. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. Permeabilizing Solution B (1.98 volumes) was added and mixed. TAMRA-DBCO Solution (0.02 volume) was added and mildly agitated for 20 minutes at room temperature. The mixtures were centrifuged at 1,000 x g for one minute at room temperature, and the top aqueous layer was discarded. Washing Solution B (2.00 volumes) was added and mixed. The mixtures were centrifuged at 1,000 x g for one minute at room temperature. The last three steps were repeated for a total of ten times. The top aqueous layer was discarded. Washing Solution C (2.00 volumes) was added and mixed. The mixtures were centrifuged at 1,000 x g for one minute at room temperature. The last three steps were repeated for a total of ten times. The top aqueous layer was discarded. Permeabilizing Solution A (2.00 volumes) was added and mixed. The mixtures were centrifuged at 1,000 x g for one minute at room temperature. The last three steps were repeated for a total of ten times. Processed worms were mounted on 3% (w/v) UltraPure Low Melting Point Agarose (Life Technologies) pads and imaged with a LSM 510 META microscope (Carl Zeiss AG) at the Biological Imaging Center in the Beckman Institute at Caltech. Samples were protected from light at all times.

3.5.4 In-Gel Fluorescence Scanning of Azf-Labeled Proteins

Lysing Solution A: In double-distilled water, 200 mM Trizma base (Sigma-Aldrich Corporation), 4% (w/v) sodium dodecyl sulfate (Sigma-Aldrich Corporation), pH 8.0
Lysing Solution B: 1X Lysing Solution A, 100 mM 2-chloroacetamide (Sigma-Aldrich Corporation), protease inhibitor cocktail (1 tablet/10 mL cOmplete, Mini, EDTA-free, Roche Diagnostics)

Lysing Solution C: 75% (v/v) Lysing Solution A, 18.75% (v/v) NuPAGE LDS Sample Buffer (Life Technologies), 6.25% (v/v) 2-mercaptoethanol (Sigma-Aldrich Corporation)

Protein Gel Fixing Solution: 40% (v/v) double-distilled water, 50% (v/v) methanol (EMD Millipore), 10% acetic acid, glacial (EMD Millipore)

TAMRA-DBCO Solution: 1 mM TAMRA-DBCO (Click Chemistry Tools) in dimethyl sulfoxide, anhydrous (Life Technologies)

Cell or worm pellets were re-suspended in freshly prepared Lysing Solution B and lysed by sonication. Lysates were clarified by centrifugation at 20,000 x *g* for 20 minutes at room temperature. Protein concentrations of lysates were determined by bicinchoninic acid assay (BCA Protein Assay Kit, Thermo Fisher Scientific) according to the manufacturer's instructions and normalized to 1.11 mg/mL with additional Lysing Solution B. Lysates (1.00 volumes) were incubated with TAMRA-DBCO Solution (0.11 volume) for 20 minutes at room temperature. Excess TAMRA-DBCO was removed according to the **Chloroform/Methanol Precipitation** procedure. Protein pellets were re-suspended in freshly prepared Lysing Solution C (1.11 volumes) and incubated for 20 minutes at 100°C. The samples were electrophoresed on NuPAGE Novex 4-12% Bis-Tris Protein Gels (Life Technologies) according to the manufacturer's instructions. Protein gels were incubated with Protein Gel Fixing Solution for 20 minutes at room temperature and imaged with a Typhoon Trio+ instrument (GE Healthcare Life Sciences). Protein gels were stained using the Colloidal Blue Staining Kit (Life Technologies) according to the manufacturer's instructions. Samples were protected from light at all times.

3.5.5 Labeling in E. coli

M9 Minimal Medium A: In double-distilled water, 6.8 g/L sodium phosphate dibasic Anhydrous (VWR International), 3 g/L potassium phosphate monobasic (VWR International), 0.5 g/L sodium chloride (VWR International), 1 g/L ammonium chloride (VWR International), 2 mM magnesium sulfate heptahydrate (VWR International), 0.1 mM calcium chloride dihydrate (VWR International), 35 mg/L thiamine hydrochloride (Sigma-Aldrich Corporation), 40 mg/L L-alanine (Sigma-Aldrich Corporation), 40 mg/L L-cysteine (Sigma-Aldrich Corporation), 40 mg/L L-aspartic Acid (Sigma-Aldrich Corporation), 40 mg/L L-glutamic Acid (Sigma-Aldrich Corporation), 40 mg/L glycine (Sigma-Aldrich Corporation), 40 mg/L L-histidine (Sigma-Aldrich Corporation), 40 mg/L L-isoleucine (Sigma-Aldrich Corporation), 40 mg/L L-leucine, 40 mg/L L-methionine (Sigma-Aldrich Corporation), 40 mg/L L-asparagine (Sigma-Aldrich Corporation), 40 mg/L L-proline (Sigma-Aldrich Corporation), 40 mg/L L-glutamine (Sigma-Aldrich Corporation), 40 mg/L L-serine (Sigma-Aldrich Corporation), 40 mg/L L-threonine (Sigma-Aldrich Corporation), 40 mg/L L-valine (Sigma-Aldrich Corporation), 40 mg/L L-tryptophan (Sigma-Aldrich Corporation), 40 mg/L L-tyrosine (Sigma-Aldrich Corporation)

M9 Minimal Medium B: In M9 Minimal Medium A, 0.5% (v/v) glycerol (VWR International), 40 mg/L L-phenylalanine (Sigma-Aldrich Corporation), 40 mg/L L-lysine (Sigma-Aldrich Corporation), 40 mg/L L-arginine (Sigma-Aldrich Corporation), 200 mg/L ampicillin sodium salt (Sigma-Aldrich Corporation), 35 mg/L chloramphenicol (Sigma-Aldrich Corporation)

M9 Minimal Medium C: M9 Minimal Medium A, 0.5% (v/v) glycerol (VWR International), 40 mg/L L-lysine (Sigma-Aldrich Corporation), 40 mg/L L-arginine (Sigma-Aldrich Corporation), 200 mg/L ampicillin sodium salt (Sigma-Aldrich Corporation), 35 mg/L chloramphenicol (Sigma-Aldrich Corporation)

M9 Minimal Medium D: In M9 Minimal Medium A, 20 mM D-glucose (VWR Inter-

national), 40 mg/L L-phenylalanine (Sigma-Aldrich Corporation), 40 mg/L L-lysine (Sigma-Aldrich Corporation), 40 mg/L L-arginine (Sigma-Aldrich Corporation), 35 mg/L kanamycin sulfate (Sigma-Aldrich Corporation)

M9 Minimal Medium E: In M9 Minimal Medium A, 20 mM D-glucose (VWR International), 40 mg/L L-lysine (Sigma-Aldrich Corporation), 40 mg/L L-arginine (Sigma-Aldrich Corporation), 35 mg/L kanamycin sulfate (Sigma-Aldrich Corporation)

M9 Minimal Medium F: In M9 Minimal Medium A, 20 mM D-glucose (VWR International), 40 mg/L L-phenylalanine (Sigma-Aldrich Corporation), 40 mg/L L-lysine ($^{13}C_6{}^{15}N_2{}^{1}H_{14}{}^{16}O_2$, Cambridge Isotope Laboratories), 40 mg/L L-arginine ($^{13}C_6{}^{14}N_4{}^{1}H_{14}{}^{16}O_2$, Cambridge Isotope Laboratories), 35 mg/L kanamycin sulfate (Sigma-Aldrich Corporation)

M9 Minimal Medium G: In M9 Minimal Medium A, 20 mM D-glucose (VWR International), 40 mg/L L-lysine ($^{13}C_6{}^{15}N_2{}^{1}H_{14}{}^{16}O_2$, Cambridge Isotope Laboratories), 40 mg/L L-arginine ($^{13}C_6{}^{14}N_4{}^{1}H_{14}{}^{16}O_2$, Cambridge Isotope Laboratories), 35 mg/L kanamycin sulfate (Sigma-Aldrich Corporation)

E. coli **1**, **2**, **3**, **4**, **5**, **6**, **7**, **8**, **9**, **10**, **11**, **12**, **13**, **14**, **15**, **16**, and **17**: A single colony was used to inoculate an overnight culture of freshly prepared M9 Minimal Medium B (5.0 mL). Overnight cultures were diluted into freshly prepared M9 Minimal Medium B (50 mL) and agitated at 37°C until reaching an OD_{600} of 0.5. Expression of wild-type and mutant *Ce*PheRSs was induced by addition of 1 M L-(+)-arabinose (500 μL, Sigma-Aldrich Corporation). After 30 minutes, cells were pelleted by centrifugation at 5000 x *g* for five minutes at 4°C, washed three times with ice-cold 0.9% (w/v) sodium chloride (VWR International) and re-suspended in freshly prepared M9 Minimal Medium C supplemented with L-phenylalanine (0.0 mM or 0.2 mM final concentration, Sigma-Aldrich Corporation) or *p*-azido-L-phenylalanine (0.0 mM, 0.2 mM or 2.0 mM final concentration, Chem-Impex International). After another 30 minutes of agitation at 37°C, expression of 6xHis-tagged GFPs was induced by

addition of 1 M isopropyl β-D-1-thiogalactopyranoside (50 μL, Sigma-Aldrich Corporation). After four hours of agitation at 37°C, cells were harvested by centrifugation at 5000 x g for five minutes at 4°C and frozen in liquid nitrogen. Samples were protected from light at all times.

KY33[pKPY514] 1: A single colony was used to inoculate an overnight culture of freshly prepared M9 Minimal Medium D (5.0 mL). Overnight cultures were diluted into freshly prepared M9 Minimal Medium D (500 mL) and agitated at 37°C until reaching an OD_{600} of 0.5. Cells were pelleted by centrifugation at 5000 x g for 15 minutes at 4°C, washed three times with ice-cold 0.9% (w/v) sodium chloride (VWR International) and re-suspended in freshly prepared M9 Minimal Medium E supplemented with p-azido-L-phenylalanine (2.0 mM final concentration, Chem-Impex International). After another 30 minutes of agitation at 37°C, expression of Thr251Gly-*Ec*PheRS was induced by addition of 1 M isopropyl β-D-1-thiogalactopyranoside (500 μL, Sigma-Aldrich Corporation). After four hours of agitation at 37°C, cells were harvested by centrifugation at 5000 x g for 15 minutes at 4°C, re-suspended with freshly prepared M9 Minimal Medium E supplemented with p-azido-L-phenylalanine (2.0 mM final concentration, Chem-Impex International) at a concentration of 250 mg wet cell mass/mL medium and stored at 4°C. Samples were protected from light at all times.

KY33[pKPY514] 2: A single colony was used to inoculate an overnight culture of freshly prepared M9 Minimal Medium F (5.0 mL). Overnight cultures were diluted into freshly prepared M9 Medium F (500 mL) and agitated at 37°C until reaching an OD_{600} of 0.5. Cells were pelleted by centrifugation at 5000 x g for 15 minutes at 4°C, washed three times with ice-cold 0.9% (w/v) sodium chloride (VWR International) and re-suspended in freshly prepared M9 Minimal Medium G supplemented with p-azido-L-phenylalanine (2.0 mM final concentration, Chem-Impex International). After another 30 minutes of agitation at 37°C, expression of Thr251Gly-*Ec*PheRS was

induced by addition of 1 M isopropyl β-D-1-thiogalactopyranoside (500 μL, Sigma-Aldrich Corporation). After four hours of agitation at 37°C, cells were harvested by centrifugation at 5000 x g for 15 minutes at 4°C, re-suspended with freshly prepared M9 Minimal Medium G supplemented with p-azido-L-phenylalanine (2.0 mM final concentration, Chem-Impex International) at a concentration of 250 mg wet cell mass/mL medium and stored at 4°C. Samples were protected from light at all times.

3.5.6 Plasmids and Strains

pKPY93: The *lpp* promoter region was isolated by genomic DNA extraction (DNeasy Blood and Tissue Kit, Qiagen) from *E. coli* DH10B (Life Technologies) and PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies). *C. elegans* tRNA^{Phe} was synthesized by primer annealing and extension with Klenow Fragment (3 to 5 exo-, New England Biolabs). The *rrnB* terminator region was isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of pBAD33 (Guzman, L.M., Belin, D., Carson, M.J., and Beckwith, J. *J. Bacteriol.* **177**, 41214130 (1995)). The purified fragments were ligated into the NdeI site of pJTN4 (Ngo, J.T., et al. *Nat. Chem. Biol.* **5**, 715717 (2009)) to generate pKPY93.

pKPY100: The genes encoding wild-type *C. elegans* PheRS were isolated by PCR amplification (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies) of pKPY66. The purified fragments were ligated into pBAD33.

pKPY102: pKPY100 was subjected to site-directed mutagenesis (PfuUltra II Fusion HS DNA Polymerase, Agilent Technologies; DpnI, New England Biolabs) to generate the Thr412Gly mutation.

KY33: The arginine-, lysine- and phenylalanine-auxotrophic strain of *E. coli* KY33 was made in-house using the red recombinase gene knockout method described by Datsenko and Wanner (Datsenko, K.A. and Wanner, B.L. *Proc. Natl. Acad. Sci.*

USA **97**, 66406645 (2000)) to eliminate the gene *argA* from the *E. coli* strain KY14. The lysine- and phenylalanine-auxotrophic strain of *E. coli* KY14 was made in-house using the red recombinase gene knockout method described by Datsenko and Wanner (Datsenko, K.A. and Wanner, B.L. *Proc. Natl. Acad. Sci. USA* **97**, 66406645 (2000)) to eliminate the gene *pheA* from the *E. coli* strain KY2. The lysine-auxotrophic strain of KY2 was made in-house using the red recombinase gene knockout method described by Datsenko and Wanner to eliminate the gene *lysA* from the *E. coli* strain TYJV2 (Van Deventer, J.A., Yuet, K.P., Yoo, T.H., and Tirrell, D.A. *ChemBioChem* **15**, 17771781 (2014)).

Worm strains were cultured as described by Brenner (Brenner, S. *Genetics* **77**, 7194 (1974)) and generated as described by Frøkjær-Jensen et al. (Frøkjær-Jensen et al. *Nat. Methods* **11**, 529-534 (2014)). The strains used in this study were as follows:

C. elegans PS7052: *syEx1516[pKPY728 (mos 5-hsp-16.2 5::fars-1(A, N, M1-G203)::gp41-1(N, C1-E88)::fib-1/rps-16::gfp::unc-54 3-myo-2 5::gp41-1(C, M1-S38)::fars-1(A, C, S204-K552, T468G)::rpl-16/M01F1.8::mCherry::let-858 3-rpl-27 5::neoR::unc-54 3-mos 3; hsp-16.41 5::peel-1::tbb-2 3) (5.0 ng/uL) + pCFJ90 (myo-2 5::mCherry::unc-54 3) (2.5 ng/uL) + pCFJ104 (myo-3 5::mCherry::unc-54 3) (5.0 ng/uL) + pGH8 (rab-3 5::mCherry::unc-54 3) (10.0 ng/uL) + pCFJ601 (eft-3 5::mos1::tbb-2 3) (50.0 ng/uL) + pBluescript II SK (127.5 ng/uL)]*

Notes: Extrachromosomal Array Carrying *miniMos* Construct pKPY728 (Injected at 5.0 ng/uL); Maintain with 0.5 mg/mL G-418 (Geneticin);

C. elegans PS7053: *syEx1517[pKPY728 (mos 5-hsp-16.2 5::fars-1(A, N, M1-G203)::gp41-1(N, C1-E88)::fib-1/rps-16::gfp::unc-54 3-myo-2 5::gp41-1(C, M1-S38)::fars-1(A, C, S204-K552, T468G)::rpl-16/M01F1.8::mCherry::let-858 3-rpl-27 5::neoR::unc-54 3-mos 3; hsp-16.41 5::peel-1::tbb-2 3) (10.0 ng/uL) + pCFJ90 (myo-2 5::mCherry::unc-54 3) (2.5 ng/uL) + pCFJ104 (myo-3 5::mCherry::unc-54 3) (5.0 ng/uL) + pGH8 (rab-3 5::mCherry::unc-54 3) (10.0 ng/uL) + pCFJ601 (eft-*

3 5::mos1::tbb-2 3) (50.0 ng/uL) + pBluescript II SK (72.5 ng/uL)]

Notes: Extrachromosomal Array Carrying *miniMos* Construct pKPY728 (Injected at 10.0 ng/uL); Maintain with 0.5 mg/mL G-418 (Geneticin);

C. elegans PS7055: *syTi1[mos 5-hsp-16.2 5::fars-1(A, N, M1-G203)::gp41-1(N, C1-E88)::fib-1/rps-16::gfp::unc-54 3-myo-2 5::gp41-1(C, M1-S38)::fars-1(A, C, S204-K552, T468G)::rpl-16/M01F1.8::mCherry::let-858 3-rpl-27 5::neoR::unc-54 3-mos 3]*

Chromosome: X;

Notes: Derived from *C. elegans* PS7052, Mapped by Inverse PCR to Chromosome X (X: 13709433-13709434), Homozygosity Confirmed by Three-Primer PCR, Maintain w/0.5 mg/mL G-418 (Geneticin);

C. elegans PS7058: *syTi2[mos 5-hsp-16.2 5::fars-1(A, N, M1-G203)::gp41-1(N, C1-E88)::fib-1/rps-16::gfp::unc-54 3-myo-2 5::gp41-1(C, M1-S38)::fars-1(A, C, S204-K552, T468G)::rpl-16/M01F1.8::mCherry::let-858 3-rpl-27 5::neoR::unc-54 3-mos 3]*

Chromosome: II;

Notes: Derived from *C. elegans* PS7053, Mapped by Inverse PCR to Chromosome II (II: 344975-344974), Maintain w/0.5 mg/mL G-418 (Geneticin);

Bibliography

[1] D. Gresham, M. J. Dunham, and D. Botstein, "Comparing whole genomes using DNA microarrays," *Nature Reviews Genetics*, vol. 9, no. 4, pp. 291–302, 2008.

[2] Z. Wang, M. Gerstein, and M. Snyder, "RNA-Seq: a revolutionary tool for transcriptomics," *Nature Reviews Genetics*, vol. 10, no. 1, pp. 57–63, 2009.

[3] L. M. De Godoy, J. V. Olsen, J. Cox, M. L. Nielsen, N. C. Hubner, F. Fröhlich, T. C. Walther, and M. Mann, "Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast," *Nature*, vol. 455, no. 7217, pp. 1251–1254, 2008.

[4] Y. Taniguchi, P. J. Choi, G.-W. Li, H. Chen, M. Babu, J. Hearn, A. Emili, and X. S. Xie, "Quantifying *E. coli* Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells," *Science*, vol. 329, no. 5991, pp. 533–538, 2010.

[5] C. Vogel and E. M. Marcotte, "Insights into the regulation of protein abundance from proteomic and transcriptomic analyses," *Nature Reviews Genetics*, vol. 13, no. 4, pp. 227–232, 2012.

[6] R. Aebersold and M. Mann, "Mass spectrometry-based proteomics," *Nature*, vol. 422, no. 6928, pp. 198–207, 2003.

[7] J. R. Yates, C. I. Ruse, and A. Nakorchevsky, "Proteomics by Mass Spectrometry: Approaches, Advances, and Applications," *Annual Review of Biomedical Engineering*, vol. 11, pp. 49–79, 2009.

[8] J. V. Olsen, S.-E. Ong, and M. Mann, "Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues," *Molecular & Cellular Proteomics*, vol. 3, no. 6, pp. 608–614, 2004.

[9] G. Mitulović and K. Mechtler, "HPLC techniques for proteomics analysisa short overview of latest developments," *Briefings in Functional Genomics & Proteomics*, vol. 5, no. 4, pp. 249–260, 2006.

[10] A. S. Essader, B. J. Cargile, J. L. Bundy, and J. L. Stephenson, "A comparison of immobilized pH gradient isoelectric focusing and strong-cation-exchange chromatography as a first dimension in shotgun proteomics," *Proteomics*, vol. 5, no. 1, pp. 24–34, 2005.

[11] R. G. Sadygov, D. Cociorva, and J. R. Yates, "Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book," *Nature Methods*, vol. 1, no. 3, pp. 195–202, 2004.

[12] J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, and M. Mann, "Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment," *Journal of Proteome Research*, vol. 10, no. 4, pp. 1794–1805, 2011.

[13] F.-M. Boisvert, Y. W. Lam, D. Lamont, and A. I. Lamond, "A Quantitative Proteomics Analysis of Subcellular Proteome Localization and Changes Induced by DNA Damage," *Molecular & Cellular Proteomics*, vol. 9, no. 3, pp. 457–470, 2010.

[14] P. Picotti, B. Bodenmiller, L. N. Mueller, B. Domon, and R. Aebersold, "Full Dynamic Range Proteome Analysis of *S. cerevisiae* by Targeted Proteomics," *Cell*, vol. 138, no. 4, pp. 795–806, 2009.

[15] F. Golebiowski, I. Matic, M. H. Tatham, C. Cole, Y. Yin, A. Nakamura, J. Cox, G. J. Barton, M. Mann, and R. T. Hay, "System-Wide Changes to SUMO Modifications in Response to Heat Shock," *Science Signaling*, vol. 2, no. 72, pp. ra24–ra24, 2009.

[16] P. Picotti, M. Clément-Ziza, H. Lam, D. S. Campbell, A. Schmidt, E. W. Deutsch, H. Röst, Z. Sun, O. Rinner, L. Reiter, *et al.*, "A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis," *Nature*, vol. 494, no. 7436, pp. 266–270, 2013.

[17] G. L. Corthals, V. C. Wasinger, D. F. Hochstrasser, and J.-C. Sanchez, "The dynamic range of protein expression: A challenge for proteomic research," *Electrophoresis*, vol. 21, pp. 1104–1115, 2000.

[18] D. L. Black, "Protein Diversity from Alternative Splicing: A Challenge for Bioinformatics and Post-Genome Biology," *Cell*, vol. 103, pp. 367–370, 2000.

[19] J. Rappsilber and M. Mann, "What does it mean to identify a protein in proteomics?," *Trends in Biochemical Sciences*, vol. 27, pp. 74–78, 2002.

[20] C. H. Ahrens, E. Brunner, E. Qeli, K. Basler, and R. Aebersold, "Generating and navigating proteome maps using mass spectrometry," *Nature Reviews Molecular Cell Biology*, vol. 11, pp. 789–801, 2010.

[21] D. Ramsköld, S. Luo, Y.-C. Wang, R. Li, Q. Deng, O. R. Faridani, G. A. Daniels, I. Khrebtukova, J. F. Loring, L. C. Laurent, *et al.*, "Full-length mRNA-Seq from single cell levels of RNA and individual circulating tumor cells," *Nature Biotechnology*, vol. 30, no. 8, pp. 777–782, 2012.

[22] N. Khidekel, S. B. Ficarro, E. C. Peters, and L. C. Hsieh-Wilson, "Exploring the O-GlcNAc proteome: Direct identification of O-GlcNAc-modified proteins from the brain," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 36, pp. 13132–13137, 2004.

[23] T. E. Thingholm, O. N. Jensen, and M. R. Larsen, "Analytical strategies for phosphoproteomics," *Proteomics*, vol. 9, no. 6, pp. 1451–1468, 2009.

[24] E. Tomlinson, N. Palaniyappan, D. Tooth, and R. Layfield, "Methods for the purification of ubiquitinated proteins," *Proteomics*, vol. 7, no. 7, pp. 1016–1022, 2007.

[25] M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, and B. Kuster, "Quantitative mass spectrometry in proteomics: a critical review," *Analytical and Bioanalytical Chemistry*, vol. 389, pp. 1017–1031, 2007.

[26] M. Bantscheff, S. Lemeer, M. M. Savitski, and B. Kuster, "Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present," *Analytical and Bioanalytical Chemistry*, vol. 404, no. 4, pp. 939–965, 2012.

[27] H. Steen and M. Mann, "The abc's (and xyz's) of peptide sequencing," *Nature Reviews Molecular Cell Biology*, vol. 5, pp. 699–711, 2004.

[28] S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, and R. Aebersold, "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags," *Nature Biotechnology*, vol. 17, pp. 994–999, 1999.

[29] S. Wiese, K. A. Reidegeld, H. E. Meyer, and B. Warscheid, "Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research," *Proteomics*, vol. 7, pp. 340–350, 2006.

[30] J. Krijgsveld, R. F. Ketting, T. Mahmoudi, J. Johansen, M. Artal-Sanz, C. P. Verrijzer, R. H. A. Plasterk, and A. J. R. Heck, "Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics," *Nature Biotechnology*, vol. 21, pp. 927–931, 2003.

[31] S.-E. Ong and M. Mann, "A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC)," *Nature Protocols*, vol. 1, pp. 2650–2660, 2007.

[32] M. Mann, "Functional and quantitative proteomics using SILAC," *Nature Reviews Molecular Cell Biology*, vol. 7, pp. 952–958, 2006.

[33] M. Krüger, M. Moser, S. Ussar, I. Thievessen, C. A. Luber, F. Forner, S. Schmidt, S. Zanivan, R. Fässler, and M. Mann, "SILAC Mouse for Quantitative Proteomics Uncovers Kindlin-3 as an Essential Factor for Red Blood Cell Function," *Cell*, vol. 134, pp. 353–364, 2008.

[34] M. D. Sury, J.-X. Chen, and M. Selbach, "The SILAC Fly Allows for Accurate Protein Quantification in Vivo," *Molecular and Cellular Proteomics*, vol. 9, pp. 2173–2183, 2010.

[35] J. Fredens, K. Engholm-Keller, A. Giessing, D. Pultz, M. R. Larsen, P. Hojrup, J. Moller-Jensen, and N. J. Faergeman, "Quantitative proteomics by amino acid labeling in *C. elegans*," *Nature Methods*, vol. 8, pp. 845–847, 2011.

[36] M. Larance, A. P. Bailly, E. Pourkarimi, R. T. Hay, G. Buchanan, S. Coulthurst, D. P. Xirodimas, A. Gartner, and A. I. Lamond, "Stable Isotope Labeling with Amino acids in Nematodes," *Nature Methods*, vol. 8, pp. 849–851, 2011.

[37] K. Larrabee, J. Phillips, G. Williams, and A. Larrabee, "The relative rates of protein synthesis and degradation in a growing culture of *Escherichia coli*," *Journal of Biological Chemistry*, vol. 255, pp. 4125–4130, 1980.

[38] R. Mosteller, R. Goldstein, and K. Nishimoto, "Metabolism of individual proteins in exponentially growing *Escherichia coli*," *Journal of Biological Chemistry*, vol. 255, pp. 2524–2532, 1980.

[39] J. M. Pratt, J. Petty, I. Riba-Garcia, D. H. Robertson, S. J. Gaskell, S. G. Oliver, and R. J. Beynon, "Dynamics of Protein Turnover, a Missing Dimension in Proteomics," *Molecular & Cellular Proteomics*, vol. 1, no. 8, pp. 579–591, 2002.

[40] E. Milner, E. Barnea, I. Beer, and A. Admon, "The Turnover Kinetics of Major Histocompatibility Complex Peptides of Human Cancer Cells," *Molecular & Cellular Proteomics*, vol. 5, no. 2, pp. 357–365, 2006.

[41] Y. W. Lam, A. I. Lamond, M. Mann, and J. S. Andersen, "Analysis of Nucleolar Protein Dynamics Reveals the Nuclear Degradation of Ribosomal Proteins," *Current Biology*, vol. 17, no. 9, pp. 749–760, 2007.

[42] B. Schwanhäusser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, "Global quantification of mammalian gene expression control," *Nature*, vol. 473, pp. 337–342, 2011.

[43] C. C. Friedel, L. Dölken, Z. Ruzsics, U. H. Koszinowski, and R. Zimmer, "Conserved principles of mammalian transcriptional regulation revealed by RNA half-life," *Nucleic Acids Research*, vol. 37, no. 17, pp. e115–e115, 2009.

[44] M. Rabani, J. Z. Levin, L. Fan, X. Adiconis, R. Raychowdhury, M. Garber, A. Gnirke, C. Nusbaum, N. Hacohen, N. Friedman, *et al.*, "Metabolic labeling

of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells," *Nature Biotechnology*, vol. 29, no. 5, pp. 436–442, 2011.

[45] R. de Sousa Abreu, L. O. Penalva, E. M. Marcotte, and C. Vogel, "Global signatures of protein and mRNA expression levels," *Molecular BioSystems*, vol. 5, no. 12, pp. 1512–1526, 2009.

[46] T. Maier, M. Güell, and L. Serrano, "Correlation of mRNA and protein in complex biological samples," *FEBS Letters*, vol. 583, no. 24, pp. 3966–3973, 2009.

[47] C. Vogel, R. de Sousa Abreu, D. Ko, S.-Y. Le, B. A. Shapiro, S. C. Burns, D. Sandhu, D. R. Boutz, E. M. Marcotte, and L. O. Penalva, "Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line," *Molecular Systems Biology*, vol. 6, no. 1, p. 400, 2010.

[48] M. Selbach, B. Schwanhusser, N. Thierfelder, Z. Fang, R. Khanin, and N. Rajewsky, "Widespread changes in protein synthesis induced by microRNAs," *Nature*, vol. 455, pp. 58–63, 2008.

[49] B. Schwanhäusser, M. Gossen, G. Dittmar, and M. Selbach, "Global analysis of cellular protein translation by pulsed SILAC," *Proteomics*, vol. 9, pp. 205–209, 2009.

[50] M. Kaller, S.-T. Liffers, S. Oeljeklaus, K. Kuhlmann, S. Röh, R. Hoffmann, B. Warscheid, and H. Hermeking, "Genome-wide Characterization of *miR-34a* Induced Changes in Protein and mRNA Expression by a Combined Pulsed SILAC and Microarray Analysis," *Molecular & Cellular Proteomics*, vol. 10, no. 8, pp. M111–010462, 2011.

[51] J. P. Patron, A. Fendler, M. Bild, U. Jung, H. Müller, M. Ø. Arntzen, C. Piso, C. Stephan, B. Thiede, H.-J. Mollenkopf, *et al.*, "MiR-133b Targets Antiapoptotic Genes and Enhances Death Receptor-Induced Apoptosis," *PLoS One*, vol. 7, no. 4, p. e35345, 2012.

[52] S. D. Kraft-Terry, I. L. Engebretsen, D. K. Bastola, H. S. Fox, P. Ciborowski, and H. E. Gendelman, "Pulsed Stable Isotope Labeling of Amino Acids in Cell Culture Uncovers the Dynamic Interactions between HIV-1 and the Monocyte-Derived Macrophage," *Journal of Proteome Research*, vol. 10, no. 6, pp. 2852–2862, 2011.

[53] L. Zhang, H. Zhao, B. S. Blagg, and R. T. Dobrowsky, "A C-Terminal Heat Shock Protein 90 Inhibitor Decreases Hyperglycemia-induced Oxidative Stress and Improves Mitochondrial Bioenergetics in Sensory Neurons," *Journal of Proteome Research*, vol. 11, no. 4, pp. 2581–2593, 2012.

[54] Y. Huo, V. Iadevaia, Z. Yao, I. Kelly, S. Cosulich, S. Guichard, L. J. Foster, and C. G. Proud, "Stable isotope-labelling analysis of the impact of inhibition of the mammalian target of rapamycin on protein synthesis," *Biochemical Journal*, vol. 444, no. 1, pp. 141–151, 2012.

[55] D. C. Dieterich, A. J. Link, J. Graumann, D. A. Tirrell, and E. M. Schuman, "Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT)," *Proceedings of the National Academy of Science of the United States of America*, vol. 103, pp. 9482–9487, 2006.

[56] J. Prescher and C. Bertozzi, "Chemistry in Living Systems," *Nature Chemical Biology*, vol. 1, pp. 13–21, 2005.

[57] E. Saxon and C. Bertozzi, "Cell Surface Engineering by a Modified Staudinger Reaction," *Science*, vol. 287, pp. 2007–2010, 2000.

[58] V. Rostovtsev, L. Green, V. Fokin, and K. Sharpless, "A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective Ligation of Azides and Terminal Alkynes," *Angewandte Chemie International Edition*, vol. 41, pp. 2596–2599, 2002.

[59] C. W. Tornøe, C. Christensen, and M. Meldal, "Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides," *The Journal of Organic Chemistry*, vol. 67, no. 9, pp. 3057–3064, 2002.

[60] N. Agard, J. Prescher, and C. Bertozzi, "A Strain-Promoted [3+2] Azide-Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems," *Journal of the American Chemical Society*, vol. 126, pp. 15046–15047, 2004.

[61] K. L. Kiick, E. Saxon, D. A. Tirrell, and C. R. Bertozzi, "Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation," *Proceedings of the National Academy of Science of the United States of America*, vol. 99, pp. 19–24, 2002.

[62] K. Beatty, F. Xie, Q. Wang, and D. Tirrell, "Selective Dye-Labeling of Newly Synthesized Proteins in Bacterial Cells," *Journal of the American Chemical Society*, vol. 127, pp. 14150–14151, 2005.

[63] K. Beatty, J. Liu, F. Xie, D. Dieterich, E. Schuman, Q. Wang, and D. Tirrell, "Fluorescence Visualization of Newly Synthesized Proteins in Mammalian Cells," *Angewandte Chemie International Edition*, vol. 45, pp. 7364–7367, 2006.

[64] D. C. Dieterich, J. Hodas, G. Gouzer, I. Shadrin, J. Ngo, A. Triller, D. Tirrell, and E. Schuman, "In situ visualization and dynamics of newly synthesized pro-

teins in rat hippocampal neurons," *Nature Neuroscience*, vol. 13, pp. 897–905, 2010.

[65] M. A. Nessen, G. 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, *et al.*, "Selective Enrichment of Azide-Containing Peptides from Complex Mixtures," *Journal of Proteome Research*, vol. 8, no. 7, pp. 3702–3711, 2009.

[66] J. Szychowski, A. Mahdavi, J. J. Hodas, J. D. Bagert, J. T. Ngo, P. Landgraf, D. C. Dieterich, E. M. Schuman, and D. A. Tirrell, "Cleavable Biotin Probes for Labeling of Biomolecules via the Azide-Alkyne Cycloaddition," *Journal of the American Chemical Society*, vol. 132, no. 51, pp. 18351–18360, 2010.

[67] G. Kramer, R. R. Sprenger, J. Back, H. L. Dekker, M. A. Nessen, J. H. van Maarseveen, L. J. de Koning, K. J. Hellingwerf, L. de Jong, and C. G. de Koster, "Identification and Quantitation of Newly Synthesized Proteins in *Escherichia coli* by Enrichment of Azidohomoalanine-labeled Peptides with Diagonal Chromatography," *Molecular & Cellular Proteomics*, vol. 8, no. 7, pp. 1599–1611, 2009.

[68] G. Kramer, R. R. Sprenger, M. A. Nessen, W. Roseboom, D. Speijer, L. de Jong, M. J. T. de Mattos, J. Back, and C. G. de Koster, "Proteome-wide Alterations in *Escherichia coli* Translation Rates upon Anaerobiosis," *Molecular & Cellular Proteomics*, vol. 9, no. 11, pp. 2508–2516, 2010.

[69] F. I. Hinz, D. C. Dieterich, D. A. Tirrell, and E. M. Schuman, "Noncanonical Amino Acid Labeling in Vivo to Visualize and Affinity Purify Newly Synthesized Proteins in Larval Zebrafish," *ACS Chemical Neuroscience*, vol. 3, no. 1, pp. 40–49, 2012.

[70] R. B. Deal, J. G. Henikoff, and S. Henikoff, "Genome-Wide Kinetics of Nucleosome Turnover Determined by Metabolic Labeling of Histones," *Science*, vol. 328, no. 5982, pp. 1161–1164, 2010.

[71] M. M. Zhang, L. K. Tsou, G. Charron, A. S. Raghavan, and H. C. Hang, "Tandem fluorescence imaging of dynamic S-acylation and protein turnover," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 19, pp. 8627–8632, 2010.

[72] K. Liu, P.-Y. Yang, Z. Na, and S. Q. Yao, "Dynamic Monitoring of Newly Synthesized Proteomes: Up-Regulation of Myristoylated Protein Kinase A During Butyric Acid Induced Apoptosis," *Angewandte Chemie International Edition*, vol. 123, no. 30, pp. 6908–6913, 2011.

[73] J. Tcherkezian, P. A. Brittis, F. Thomas, P. P. Roux, and J. G. Flanagan, "Transmembrane Receptor DCC Associates with Protein Synthesis Machinery and Regulates Translation," *Cell*, vol. 141, no. 4, pp. 632–644, 2010.

[74] J. J. Hodas, A. Nehring, N. Höche, M. J. Sweredoski, R. Pielot, S. Hess, D. A. Tirrell, D. C. Dieterich, and E. M. Schuman, "Dopaminergic modulation of the hippocampal neuropil proteome identified by bioorthogonal noncanonical amino acid tagging (BONCAT)," *Proteomics*, vol. 12, no. 15-16, pp. 2464–2476, 2012.

[75] B. C. Yoon, H. Jung, A. Dwivedy, C. M. O'Hare, K. H. Zivraj, and C. E. Holt, "Local Translation of Extranuclear Lamin B Promotes Axon Maintenance," *Cell*, vol. 148, no. 4, pp. 752–764, 2012.

[76] I. Brunet, C. Weinl, M. Piper, A. Trembleau, M. Volovitch, W. Harris, A. Prochiantz, and C. Holt, "The transcription factor Engrailed-2 guides retinal axons," *Nature*, vol. 438, no. 7064, pp. 94–98, 2005.

[77] K. Eichelbaum, M. Winter, M. B. Diaz, S. Herzig, and J. Krijgsveld, "Selective enrichment of newly synthesized proteins for quantitative secretome analysis," *Nature Biotechnology*, vol. 30, no. 10, pp. 984–990, 2012.

[78] A. J. Howden, V. Geoghegan, K. Katsch, G. Efstathiou, B. Bhushan, O. Boutureira, B. Thomas, D. C. Trudgian, B. M. Kessler, D. C. Dieterich, *et al.*, "QuaNCAT: quantitating proteome dynamics in primary cells," *Nature Methods*, vol. 10, no. 4, pp. 343–346, 2013.

[79] C. Hughes and J. Krijgsveld, "Developments in quantitative mass spectrometry for the analysis of proteome dynamics," *Trends in Biotechnology*, vol. 30, no. 12, pp. 668–676, 2012.

[80] J. Liu, Y. Xu, D. Stoleru, and A. Salic, "Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 2, pp. 413–418, 2012.

[81] O. Rechavi, M. Kalman, Y. Fang, H. Vernitsky, J. Jacob-Hirsch, L. J. Foster, Y. Kloog, and I. Goldstein, "Trans-SILAC: sorting out the non-cell-autonomous proteome," *Nature Methods*, vol. 7, pp. 923–927, 2010.

[82] R. B. Deal and S. Henikoff, "A Simple Method for Gene Expression and Chromatin Profiling of Individual Cell Types within a Tissue," *Developmental Cell*, vol. 18, no. 6, pp. 1030–1040, 2010.

[83] F. A. Steiner, P. B. Talbert, S. Kasinathan, R. B. Deal, and S. Henikoff, "Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling," *Genome Research*, vol. 22, no. 4, pp. 766–777, 2012.

[84] G. L. Henry, F. P. Davis, S. Picard, and S. R. Eddy, "Cell type-specific genomics of *Drosophila* neurons," *Nucleic Acids Research*, vol. 40, no. 19, pp. 9691–9704, 2012.

[85] M. R. Miller, K. J. Robinson, M. D. Cleary, and C. Q. Doe, "TU-tagging: cell type specific RNA isolation from intact complex tissues," *Nature Methods*, vol. 6, no. 6, p. 439, 2009.

[86] L. Gay, M. R. Miller, P. B. Ventura, V. Devasthali, Z. Vue, H. L. Thompson, S. Temple, H. Zong, M. D. Cleary, K. Stankunas, *et al.*, "Mouse TU tagging: a chemical/genetic intersectional method for purifying cell type-specific nascent RNA," *Genes & Development*, vol. 27, no. 1, pp. 98–115, 2013.

[87] A. J. Link, M. K. S. Vink, N. J. Agard, J. A. Prescher, C. R. Bertozzi, and D. A. Tirrell, "Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, pp. 10180–10185, 2006.

[88] I. C. Tanrikulu, E. Schmitt, Y. Mechulam, W. A. Goddard, III, and D. A. Tirrell, "Discovery of *Escherichia coli* methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine *in vivo*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, pp. 15285–15290, 2009.

[89] J. T. Ngo, J. A. Champion, A. Mahdavi, I. C. Tanrikulu, K. E. Beatty, R. E. Connor, T. H. Yoo, D. C. Dieterich, E. M. Schuman, and D. A. Tirrell, "Cell-selective metabolic labeling of proteins," *Nature Chemical Biology*, vol. 5, pp. 715–717, 2009.

[90] M. Grammel, M. M. Zhang, and H. C. Hang, "Orthogonal Alkynyl Amino Acid Reporter for Selective Labeling of Bacterial Proteomes during Infection," *Angewandte Chemie International Edition*, vol. 49, pp. 5970–5974, 2010.

[91] J. T. Ngo, E. M. Schuman, and D. A. Tirrell, "Mutant methionyl-tRNA synthetase from bacteria enables site-selective N-terminal labeling of proteins expressed in mammalian cells," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 13, pp. 4992–4997, 2013.

[92] J. T. Ngo, B. M. Babin, J. A. Champion, E. M. Schuman, and D. A. Tirrell, "State-Selective Metabolic Labeling of Cellular Proteins," *ACS Chemical Biology*, vol. 7, no. 8, pp. 1326–1330, 2012.

[93] F. Truong, T. H. Yoo, T. J. Lampo, and D. A. Tirrell, "Two-Strain, Cell-Selective Protein Labeling in Mixed Bacterial Cultures," *Journal of the American Chemical Society*, vol. 134, no. 20, pp. 8551–8556, 2012.

[94] A. Mahdavi, T. H. Segall-Shapiro, S. Kou, G. A. Jindal, K. G. Hoff, S. Liu, M. Chitsaz, R. F. Ismagilov, J. J. Silberg, and D. A. Tirrell, "A Genetically Encoded AND Gate for Cell-Targeted Metabolic Labeling of Proteins," *Journal of the American Chemical Society*, vol. 135, pp. 2979–2982, 2013.

[95] H.-W. Rhee, P. Zou, N. D. Udeshi, J. D. Martell, V. K. Mootha, S. A. Carr, and A. Y. Ting, "Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging," *Science*, vol. 339, no. 6125, pp. 1328–1331, 2013.

[96] J. D. Martell, T. J. Deerinck, Y. Sancak, T. L. Poulos, V. K. Mootha, G. E. Sosinsky, M. H. Ellisman, and A. Y. Ting, "Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy," *Nature Biotechnology*, vol. 30, no. 11, pp. 1143–1148, 2012.

[97] D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess, *C. elegans II*. Cold Spring Harbor Laboratory Press, 1997.

[98] S. Brenner, "The genetics of *Caenorhabditis elegans*," *Genetics*, vol. 77, pp. 71–94, 1974.

[99] C. J. Kenyon, "The genetics of ageing," *Nature*, vol. 464, pp. 504–512, 2010.

[100] G. Lettre and M. O. Hengartner, "Developmental apoptosis in *C. elegans*: a complex CEDnario," *Nature Reviews Molecular Cell Biology*, vol. 7, pp. 97–108, 2006.

[101] E. T. Kipreos, "*C. elegans* cell cycles: invariance and stem cell divisions," *Nature Reviews Molecular Cell Biology*, vol. 6, pp. 766–776, 2005.

[102] P. Gönczy, "Mechanisms of asymmetric cell division: flies and worms pave the way," *Nature Reviews Molecular Cell Biology*, vol. 9, pp. 355–366, 2008.

[103] A. Pires-daSilva and R. J. Sommer, "The evolution of signalling pathways in animal development," *Nature Reviews Genetics*, vol. 4, pp. 39–49, 2003.

[104] T. Kaletta and M. O. Hengartner, "Finding function in novel targets: *C. elegans* as a model organism," *Nature Reviews Drug Discovery*, vol. 5, pp. 387–399, 2006.

[105] V. Ambros, "The functions of animal microRNAs," *Nature*, vol. 431, pp. 350–355, 2004.

[106] J. E. Irazoqui, J. M. Urbach, and F. M. Ausubel, "Evolution of host innate defence: insights from *Caenorhabditis elegans* and primitive invertebrates," *Nature Reviews Immunology*, vol. 10, pp. 47–58, 2010.

[107] M. Artal-Sanz and N. Tavernarakis, "Mechanisms of Aging and Energy Metabolism in *Caenorhabditis elegans*," *IUBMB Life*, vol. 60, pp. 315–322, 2008.

[108] M.-A. Félix and C. Braendle, "The natural history of *Caenorhabditis elegans*," *Current Biology*, vol. 20, pp. R965–R969, 2010.

[109] T. Steiernagle, *Maintenance of C. elegans*, pp. 51–67. Oxford University Press, 1999.

[110] Z. F. Altun and D. H. Hall, "Introduction," *WormAtlas*, 2009.

[111] J. E. Sulston and H. R. Horvitz, "Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*," *Developmental Biology*, vol. 56, pp. 110–156, 1977.

[112] J. E. Sulston, E. Schierenberg, J. G. White, and J. N. Thomson, "The embryonic cell lineage of the nematode *Caenorhabditis elegans*," *Developmental Biology*, vol. 100, pp. 64–119, 1983.

[113] I. Antoshechkin and P. W. Sternberg, "The versatile worm: genetic and genomic resources for *Caenorhabditis elegans* research," *Nature Reviews Genetics*, vol. 8, pp. 518–532, 2007.

[114] *C. elegans* Sequencing Consortium, "Genome sequence of the nematode *C. elegans*: a platform for investigating biology," *Science*, vol. 282, pp. 2012–2018, 1998.

[115] J. T. Ngo and D. A. Tirrell, "Non-Canonical Amino Acids in the Interrogation of Cellular Protein Synthesis," *Accounts of Chemical Research*, vol. 44, pp. 677–685, 2011.

[116] K. P. Yuet and D. A. Tirrell, "Chemical Tools for Temporally and Spatially Resolved Mass Spectrometry-Based Proteomics," *Annals of Biomedical Engineering*, vol. 42, pp. 299–311, 2014.

[117] A. Mahdavi, J. Szychowski, J. T. Ngo, M. J. Sweredoski, R. L. J. Graham, S. Hess, O. Schneewind, S. K. Mazmanian, and D. A. Tirrell, "Identification

of secreted bacterial proteins by noncanonical amino acid tagging," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, pp. 433–438, 2014.

[118] T. S. Elliott, F. M. Townsley, A. Bianco, R. J. Ernst, A. Sachdeva, S. J. Elsässer, L. Davis, K. Lang, R. Pisa, S. Greiss, K. S. Lilley, and J. W. Chin, "Proteome labeling and protein identification in specific tissues and at specific developmental stages in an animal," *Nature Biotechnology*, vol. 32, pp. 465–472, 2014.

[119] I. Mermershtain, I. Finarov, L. Klipcan, N. Kessler, H. Rozenberg, and M. G. Safro, "Idiosyncrasy and identity in the prokaryotic phe-system: Crystal structure of *E. coli* phenylalanyl-tRNA synthetase complexed with phenylalanine and AMP," *Protein Science*, vol. 20, pp. 160–167, 2011.

[120] I. Finarov, N. Moor, N. Kessler, L. Klipcan, and M. G. Safro, "Structure of Human Cytosolic Phenylalanyl-tRNA Synthetase: Evidence for Kingdom-Specific Design of the Active Sites and tRNA Binding Patterns," *Structure*, vol. 18, pp. 343–353, 2010.

[121] Z. Shen, X. Zhang, Y. Chai, Z. Zhu, P. Yi, G. Feng, W. Li, and G. Ou, "Conditional Knockouts Generated by Engineered CRISPR-Cas9 Endonuclease Reveal the Roles of Coronin in *C. elegans* Neural Development," *Developmental Cell*, vol. 30, pp. 625–636, 2014.

[122] E. G. Stringham, D. K. Dixon, D. Jones, and E. P. M. Candido, "Temporal and Spatial Expression Patterns of the Small Heat Shock (*hsp16*) Genes in Transgenic *Caenorhabditis elegans*," *Molecular Biology of the Cell*, vol. 3, pp. 221–233, 1992.

[123] D. M. M. III, I. Ortiz, G. C. Berliner, and H. F. Epstein, "Differential Localization of Two Myosins within Nematode Thick Filaments," *Cell*, vol. 34, pp. 477–490, 1983.

[124] C. R. Egan, M. A. Chung, F. L. Allen, M. F. P. Heschl, C. L. V. Buskirk, and J. D. McGhee, "A Gut-to-Pharynx/Tail Switch in Embryonic Expression of the *Caenorhabditis elegans ges-1* Gene Centers on Two GATA Sequences," *Developmental Biology*, vol. 170, pp. 397–419, 1995.

[125] M. L. Nonet, J. E. Staunton, M. P. Kilgard, T. Fergestad, E. Hartwieg, H. R. Horvitz, E. M. Jorgensen, and B. J. Meyer, "*Caenorhabditis elegans rab-3* Mutant Synapses Exhibit Impaired Function and Are Partially Depleted of Vesicles," *The Journal of Neuroscience*, vol. 17, pp. 8061–8073, 1997.

[126] J. P. Ardizzi and H. E. Epstein, "Immunochemical Localization of Myosin Heavy Chain Isoforms and Paramyosin in Developmentally and Structurally Diverse Muscle Cell Types of the Nematode *Caenorhabditis elegans*," *The Journal of Cell Biology*, vol. 105, pp. 2763–2770, 1987.

[127] S. E. Mango, "The *C. elegans* pharynx: a model for organogenesis," *WormBook*, 2007.

[128] WormBase, "Wormbase, Release WS244," 2014.

[129] R. Ruksana, K. Kuroda, H. Terami, T. Bando, S. Kitaoka, T. Takaya, Y. Sakube, and H. Kagawa, "Tissue expression of four troponin I genes and their molecular interactions with two troponin C isoforms in *Caenorhabditis elegans*," *Genes to Cells*, vol. 10, pp. 261–276, 2005.

[130] A. Higashibata, A. Higashitani, R. Adachi, H. Kagawa, S. Honda, Y. Honda, N. Higashitani, Y. Sasagawa, Y. Miyazawa, N. J. Szewczyk, C. A. Conley,

N. Fujimoto, K. Fukui, T. Shimazu, K. Kuriyama, and N. Ishioka, "Biochemical and Molecular Biological Analyses of space-flown nematodes in Japan, the First International *Caenorhabditis elegans* Experiment (ICE-First)," *Microgravity Science and Technology*, vol. 19, pp. 159–163, 2007.

[131] G. L. Fawcett, C. M. Santi, A. Butler, T. Harris, M. Covarrubias, and L. Salkoff, "Mutant Analysis of the Shal (Kv4) Voltage-gated Fast Transient K+ Channel in *Caenorhabditis elegans*," *The Journal of Biological Chemistry*, vol. 281, pp. 30725–30735, 2006.

[132] V. Sharma, C. He, J. Sacca-Schaeffer, E. Brzozowski, D. E. Martin-Herranz, Z. Mendelowitz, D. A. Fitzpatrick, and D. M. OHalloran, "Insight into the Family of Na+/Ca2+ exchangers of *Caenorhabditis elegans*," *Genetics*, vol. 195, pp. 611–619, 2013.

[133] S. J. McKay, R. Johnsen, J. Khattra, J. Asano, D. L. Baillie, S. Chan, N. Dube, L. Fang, B. Goszczynski, E. Ha, E. Halfnight, R. Hollebakken, P. Huang, K. Hung, V. Jensen, S. J. M. Jones, H. Kai, D. Li, A. Mah, M. Marra, J. McGhee, R. Newbury, A. Pouzyrev, D. L. Riddle, E. Sonnhammer, H. Tian, D. Tu, J. R. Tyson, G. Vatcher, A. Warner, K. Wong, Z. Zhao, and D. G. Moerman, "Gene Expression Profiling of Cells, Tissues, and Developmental Stages of the Nematode *C. elegans*," *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 68, pp. 159–170, 2003.

[134] M. Gimona, I. Kaverina, G. P. Resch, E. Vignal, , and G. Burgstaller, "Calponin Repeats Regulate Actin Filament Stability and Formation of Podosomes in Smooth Muscle Cells," *Molecular Biology of the Cell*, vol. 14, pp. 2482–2491, 2003.

[135] S. Goetinck and R. H. Waterston, "The *Caenorhabditis elegans* Muscle-affecting Gene *unc-87* Encodes a Novel Thin Filament-associated Protein," *The Journal of Cell Biology*, vol. 127, pp. 79–93, 1994.

[136] W. C. Spencer, G. Zeller, J. D. Watson, S. R. Henz, K. L. Watkins, R. D. McWhirter, S. Petersen, V. T. Sreedharan, C. Widmer, J. Jo, V. Reinke, L. Petrella, S. Strome, S. E. V. Stetina, M. Katz, S. Shaham, G. Rätsch, and D. M. M. III, "A spatial and temporal map of *C. elegans* gene expression," *Developmental Cell*, vol. 21, pp. 325–341, 2011.

[137] M. Wang, M. Weiss, M. Simonovic, G. Haertinger, S. P. Schrimpf, M. O. Hengartner, and C. von Mering, "PaxDb, a Database of Protein Abundance Averages Across All Three Domains of Life," *Annals of Biomedical Engineering*, vol. 11, pp. 492–500, 2012.

[138] T. Gu, S. Orita, and M. Han, "*Caenorhabditis elegans* SUR-5, a Novel but Conserved Protein, Negatively Regulates LET-60 Ras Activity during Vulval Induction," *Molecular and Cellular Biology*, vol. 18, pp. 4556–4564, 1998.

[139] C. Chen, S. Tuck, and A. S. Byström, "Defects in tRNA Modification Associated with Neurological and Developmental Dysfunctions in *Caenorhabditis elegans* Elongator Mutants," *The Journal of Cell Biology*, vol. 5, p. e1000561, 2009.

[140] N. A. Kulak, G. Pichler, I. Paron, N. Nagaraj, and M. Mann, "Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells," *Nature Methods*, vol. 11, pp. 319–324, 2014.

[141] M. W. Davis, J. J. Morton, D. Carroll, and E. M. Jorgensen, "Gene Activation Using FLP Recombinase in *C. elegans*," *PLoS Genetic*, vol. 4, p. e1000028, 2008.

[142] G. W. J. Fleet, R. R. Porter, and J. R. Knowles, "Affinity Labelling of Antibodies with Aryl Nitrene as Reactive Group," *Nature*, vol. 224, pp. 511–512, 1969.

[143] I. Erdmann, K. Marter, O. Kobler, S. Niehues, J. Abele, A. Müller, J. Bussmann, E. Storkebaum, T. Ziv, U. Thomas, *et al.*, "Cell-selective labelling of proteomes in *Drosophila melanogaster*," *Nature Communications*, vol. 6, 2015.

[144] S. Niehues, J. Bussmann, G. Steffes, I. Erdmann, C. Köhrer, L. Sun, M. Wagner, K. Schäfer, G. Wang, S. N. Koerdt, *et al.*, "Impaired protein translation in *Drosophila* models for Charcot-Marie-Tooth neuropathy caused by mutant tRNA synthetases," *Nature Communications*, vol. 6, 2015.

[145] K. P. Yuet, M. K. Doma, J. T. Ngo, M. J. Sweredoski, R. L. Graham, A. Moradian, S. Hess, E. M. Schuman, P. W. Sternberg, and D. A. Tirrell, "Cell-specific proteomic analysis in *Caenorhabditis elegans*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 9, pp. 2705–2710, 2015.

[146] N. H. Shah and T. W. Muir, "Inteins: Nature's gift to protein chemists," *Chemical Science*, vol. 5, no. 2, pp. 446–461, 2014.

[147] B. Dassa, N. London, B. L. Stoddard, O. Schueler-Furman, and S. Pietrokovski, "Fractured genes: a novel genomic arrangement involving new split inteins and a new homing endonuclease family," *Nucleic Acids Research*, p. gkp095, 2009.

[148] M. W. Southworth, K. Amaya, T. C. Evans, M.-Q. Xu, and F. B. Perler, "Purification of proteins fused to either the amino or carboxy terminus of the Mycobacterium xenopi gyrase A intein," *Biotechniques*, vol. 27, no. 1, pp. 110–4, 1999.

[149] T. W. Muir, D. Sondhi, and P. A. Cole, "Expressed protein ligation: A general method for protein engineering," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 12, pp. 6705–6710, 1998.

[150] I. Giriat and T. W. Muir, "Protein Semi-Synthesis in Living Cells," *Journal of the American Chemical Society*, vol. 125, no. 24, pp. 7180–7181, 2003.

[151] S. Züger and H. Iwai, "Intein-based biosynthetic incorporation of unlabeled protein tags into isotopically labeled proteins for NMR studies," *Nature biotechnology*, vol. 23, no. 6, pp. 736–740, 2005.

[152] J. A. Camarero and T. W. Muir, "Biosynthesis of a Head-to-Tail Cyclized Protein with Improved Biological Activity," *Journal of the American Chemical Society*, vol. 121, no. 23, pp. 5597–5598, 1999.

[153] E. C. Schwartz, L. Saez, M. W. Young, and T. W. Muir, "Post-translational enzyme activation in an animal via optimized conditional protein splicing," *Nature Chemical Biology*, vol. 3, no. 1, pp. 50–54, 2007.

[154] H. Iwai, S. Züger, J. Jin, and P.-H. Tam, "Highly efficient protein *trans*-splicing by a naturally split DnaE intein from *Nostoc punctiforme*," *FEBS Letters*, vol. 580, no. 7, pp. 1853–1858, 2006.

[155] J. Zettler, V. Schütz, and H. D. Mootz, "The naturally split *Npu* DnaE intein exhibits an extraordinarily high rate in the protein *trans*-splicing reaction," *FEBS Letters*, vol. 583, no. 5, pp. 909–914, 2009.

[156] P. Carvajal-Vallejos, R. Pallissé, H. D. Mootz, and S. R. Schmidt, "Unprecedented Rates and Efficiencies Revealed for New Natural Split Inteins from Metagenomic Sources," *Journal of Biological Chemistry*, vol. 287, no. 34, pp. 28686–28696, 2012.

[157] L.-W. Lee, H.-W. Lo, and S. J. Lo, "Vectors for co-expression of two genes in *Caenorhabditis elegans*," *Gene*, vol. 455, no. 1, pp. 16–21, 2010.

[158] R. Giordano-Santini, S. Milstein, N. Svrzikapa, D. Tu, R. Johnsen, D. Baillie, M. Vidal, and D. Dupuy, "An antibiotic selection marker for nematode transgenesis," *Nature Methods*, vol. 7, no. 9, pp. 721–723, 2010.

[159] H. S. Seidel, M. Ailion, J. Li, A. van Oudenaarden, M. V. Rockman, and L. Kruglyak, "A Novel Sperm-Delivered Toxin Causes Late-Stage Embryo Lethality and Transmission Ratio Distortion in *C. elegans*," *PLoS Biology*, vol. 9, no. 7, p. e1001115, 2011.

[160] C. Frøkjær-Jensen, M. W. Davis, M. Sarov, J. Taylor, S. Flibotte, M. LaBella, A. Pozniakovsky, D. G. Moerman, and E. M. Jorgensen, "Random and targeted transgene insertion in *Caenorhabditis elegans* using a modified *Mos1* transposon," *Nature Methods*, vol. 11, no. 5, pp. 529–534, 2014.

[161] C. C. Mello, J. M. Kramer, D. Stinchcomb, and V. Ambros, "Efficient gene transfer in *C. elegans*: extrachromosomal maintenance and integration of transforming sequences," *The EMBO Journal*, vol. 10, no. 12, pp. 3959–3970, 1991.

[162] T. Boulin and J.-L. Bessereau, "*Mos1*-mediated insertional mutagenesis in *Caenorhabditis elegans*," *Nature Protocols*, vol. 2, no. 5, pp. 1276–1287, 2007.