

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)

© 2016

Kai P. Yuet

All Rights Reserved

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

Abstract	v
Chapter 1 Chemical Tools for Temporally and Spatially Resolved Mass Spectrometry-Based Proteomics	1
1.1 Abstract	2
1.2 Introduction	2
1.3 Temporally Resolved Proteomic Analysis	4
1.3.1 <i>Stable-Isotope Labeling with Amino Acids in Cell Culture</i>	4
1.3.2 <i>Repurposing SILAC for Temporally Resolved Proteomic Analysis</i>	5
1.3.3 <i>Pulsed SILAC</i>	6
1.3.4 <i>Bio-Orthogonal Non-Canonical Amino Acid Tagging</i>	7
1.3.5 <i>Quantitative Non-Canonical Amino Acid Tagging</i>	10

1.3.6	<i>O</i> -Propargyl-Puromycin Labeling	11
1.4	Spatially Resolved Proteomic Analysis	12
1.4.1	<i>Coupling Flow Cytometry and Mass Spectrometry</i>	12
1.4.2	<i>Cell-Selective BONCAT</i>	13
1.4.3	<i>Ascorbate Peroxidase Labeling</i>	16
1.5	Conclusions	17
1.6	Figures	18
 Chapter 2 Cell-Specific Proteomic Analysis in <i>C. elegans</i>		27
2.1	Abstract	28
2.2	Introduction	28
2.3	Results and Discussion	31
2.3.1	<i>Engineering a <i>C. elegans</i> PheRS Capable of Activating Azf</i>	31
2.3.2	<i>Characterizing Azf Labeling in <i>C. elegans</i></i>	32
2.3.3	<i>Labeling Spatially Defined Protein Subpopulations</i>	33
2.3.4	<i>Identifying Pharyngeal Muscle-Specific Proteins</i>	34
2.4	Conclusions	38
2.5	Figures	40
2.6	Tables	66
2.7	Materials and Methods	71
2.7.1	<i>ATP-PP_i Exchange Assay</i>	71
2.7.2	<i>Chloroform/Methanol Precipitation</i>	72
2.7.3	<i>Enrichment of Azf-Labeled Proteins</i>	73
2.7.4	<i>Fluorescence Microscopy of Live <i>C. elegans</i></i>	74
2.7.5	<i>Fluorescence Microscopy of Fixed <i>C. elegans</i></i>	75
2.7.6	<i>In-Gel Fluorescence Scanning of Azf-Labeled Proteins</i>	77
2.7.7	<i>In-Gel Proteolytic Digestion of Azf-Labeled Proteins</i>	79
2.7.8	<i>Isolation of 6xHis-Tagged Proteins</i>	80

2.7.9	<i>Labeling in C. elegans</i>	81
2.7.10	<i>Labeling in E. coli</i>	82
2.7.11	<i>LC-MS/MS of Azf-Labeled Proteins</i>	86
2.7.12	<i>MALDI TOF-MS of 6xHis-Tagged Proteins</i>	88
2.7.13	<i>Plasmids and Strains</i>	89
2.7.14	<i>Western Blotting</i>	95
Chapter 3 Split-Intein Split-Aminoacyl-tRNA Synthetase System		97
3.1	Introduction	98
3.2	Results and Discussion	99
3.2.1	<i>Engineering Split System in E. coli</i>	99
3.2.2	<i>Characterizing Split System in C. elegans</i>	105
3.3	Conclusions	108
3.4	Figures	109
3.5	Materials and Methods	137
3.5.1	<i>Chloroform/Methanol Precipitation</i>	137
3.5.2	<i>Fluorescence Microscopy of Live C. elegans</i>	137
3.5.3	<i>Fluorescence Microscopy of Fixed C. elegans</i>	138
3.5.4	<i>In-Gel Fluorescence Scanning of Azf-Labeled Proteins</i>	140
3.5.5	<i>Labeling in E. coli</i>	142
3.5.6	<i>Plasmids and Strains</i>	145
Bibliography		149

List of Figures

1.1	Temporally Resolved Proteomic Analysis	18
1.2	Structures Discussed in Chapter 1	19
1.3	O-Propargyl-Puromycin Labeling	21
1.4	Cell-Selective BONCAT Performed in a Mixture of Cells	23
1.5	Ascorbate Peroxidase Labeling	25
2.1	<i>C. elegans</i> Adult Hermaphrodite	40
2.2	Life Cycle of <i>C. elegans</i>	41
2.3	Cell-Selective Proteomic Analysis in <i>C. elegans</i>	42
2.4	Structures of Phe, Azf, TAMRA-DBCO, and Diazo Biotin-DBCO	43
2.5	Active Site of <i>H. sapiens</i> PheRS	44
2.6	Alignment of Eukaryotic PheRSs	45
2.7	<i>E. coli</i> KY14 and Plasmids pKPY93/pKPY1XX	46

2.8	<i>CePheRS</i> : SDS/PAGE and In-Gel Fluorescence Scanning Detection of Azf-Labeled Proteins	47
2.9	MALDI-TOF Analysis of GFP Peptide SAFPEGYVQER	48
2.10	Eukaryotic PheRS: SDS/PAGE and In-Gel Fluorescence Scanning Detection of Azf-Labeled Proteins	49
2.11	<i>E. coli</i> PheRS: SDS/PAGE and In-Gel Fluorescence Scanning Detection of Azf-Labeled Proteins	50
2.12	Amino Acid Analysis of Whole <i>E. coli</i> Protein	51
2.13	Western Blot Detection of <i>E. coli</i> and <i>C. elegans</i> Proteins	52
2.14	In-Gel Fluorescence Scanning and Fluorescence Microscopy of <i>hsp-16.2::Thr412Gly-CePheRS C. elegans</i>	53
2.15	Fluorescence Microscopy of Live Worms	54
2.16	Fluorescence Microscopy of Labeled Worms	55
2.17	Fluorescence Microscopy of Labeled <i>rab-3::Thr412Gly-CePheRS</i> Worms	56
2.18	Model Labeling of <i>E. coli</i> Lysates with Diazo Biotin-DBCO	57
2.19	Model Enrichment of <i>E. coli</i> Lysates with Diazo Biotin-DBCO	58
2.20	Unenriched and Enriched Samples Prepared from Labeled Worms	59
2.21	LC-MS/MS Analysis	60
2.22	LC-MS/MS Analysis: Phenylalanine Count	61
2.23	LC-MS/MS Analysis: Phenylalanine Count	62
2.24	LC-MS/MS Analysis: Protein Abundance	63
2.25	Fluorescence Microscopy of Live <i>C53C9.2::gfp</i> , <i>K03E5.2::gfp</i> , and <i>cpn-4::gfp</i> Animals	64
2.26	Schematic of Calponin-1, CPN-4, C53C9.2, K03E5.2, T25F10.6, and UNC-87	65
3.1	Split-Intein Mediated Split-Aminoacyl-tRNA Synthetase	109
3.2	Structures of Phe, Azf, and TAMRA-DBCO	110

3.3	<i>dnaE-n</i> Sequence	111
3.4	<i>dnaE-c</i> Sequence	112
3.5	<i>C. elegans</i> FARS-1 Sequence	113
3.6	FARS-1(N, Met1-Lys187)-Int(N, DnaE) Sequence	114
3.7	Int(C, DnaE)-Cys-Phe-Asn-FARS-1(C, Gln188-Lys496) Sequence	115
3.8	First Version of Split-Intein Split-Synthetase System	116
3.9	Evaluating Labeling Activity of Glu26Cys-Thr412Gly- <i>CePheRS</i>	117
3.10	FARS-1(N, Met1-Asn25)-Int(N, DnaE) Sequence	118
3.11	Int(C, DnaE)-FARS-1(C, Glu26Cys-Lys496) Sequence	119
3.12	Second Version of Split-Intein Split-Synthetase System	120
3.13	<i>gp41-1-n</i> Sequence	121
3.14	<i>gp41-1-c</i> Sequence	122
3.15	FARS-1(N, Met1-Gly147)-Int(N, Gp41-1) Sequence	123
3.16	Int(C, Gp41-1)-FARS-1(C, Ser148-Lys496) Sequence	124
3.17	Third Version of Split-Intein Split-Synthetase System	125
3.18	Evaluating Split-Intein Split-Synthetase System in <i>C. elegans</i>	126
3.19	pKPY728	127
3.20	Generating PS7055 and PS7058	128
3.21	<i>C. elegans</i> Hermaphrodite Gonad	129
3.22	Fluorescence Microscopy of PS7055 Precursor	130
3.23	Fluorescence Microscopy of PS7055	131
3.24	Mapping <i>syTi1</i>	132
3.25	Verifying <i>syTi1</i>	133
3.26	Mapping <i>syTi2</i>	134
3.27	Verifying <i>syTi2</i>	135
3.28	Fluorescence Microscopy of Live and Labeled PS7055	136

List of Tables

2.1	<i>C. elegans</i> Methionyl-tRNA Synthetases Activation of Amino Acids	66
2.2	<i>C. elegans</i> Phenylalanyl-tRNA Synthetases Activation of Amino Acids	67
2.3	Proteins Identified and Quantified from LC-MS/MS Analysis	68
2.4	Pharyngeal Proteins Identified and Quantified from LC-MS/MS Analysis	69
2.5	Abundant “Non-Pharyngeal” Proteins	70