Chemical and enzymatic tools to study proteins in their native cellular environment

ARCHIVES
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
JUL 01 2013
LIBRARIES

By

Jennifer Zhengzheng Yao

B.S. in Chemical Biology (2008)
University of California, Berkeley

Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

at the

Massachusetts Institute of Technology

June 2013

© 2013 Massachusetts Institute of Technology
All rights reserved

Signature of the Author: ______________________
Department of Chemistry
May 24, 2013

Certified by: ______________________
Alice Y. Ting
Ellen Swallow Richards Associate Professor of Chemistry
Thesis supervisor

Accepted by: ______________________
Robert W. Field
Chairman, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a committee of the Department of Chemistry as follows:

JoAnne Stubbe
Novartis Professor of Chemistry
Professor of Biology
Thesis committee chair

Alice Y. Ting
Ellen Swallow Richards Associate Professor of Chemistry
Thesis advisor

Alexander M. Klibanov
Firmenich Professor of Chemistry and Bioengineering
Thesis committee member

Chemical and enzymatic tools to study proteins in their native cellular environment

By

Jennifer Zhengzheng Yao

Submitted to the Department of Chemistry
on May 24, 2013 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

ABSTRACT

A detailed understating of living systems requires methods to probe molecular processes in cells and whole organisms. A set of technologies that combines chemical and genetic probes have been developed to address the need for dynamic and noninvasive assay of biological processes. In addition to be able to visualize the localization, trafficking, and turnover of *individual* proteins, strategies that allow the tagging, and imaging, and identification of entire proteomes have also offered valuable insights into disease biology.

Since protein visualization serves as a complement to protein identification, this thesis first describes the development of a protein labeling technique that is able to specifically target diverse fluorophores to proteins inside live cells. The methodology uses the *E. coli* lipoic acid ligase (LplA) that we have engineered to accept and ligate an azide functional handle onto a 13-amino acid LplA acceptor peptide (LAP). Subsequent derivatization of the azide with fluorophores functionalized with cyclooctyne via strain-promoted azide-alkyne cycloaddition allowed us to target many bright and photostable fluorophores that could be used in super resolution imaging. Due to the numerous applications to which cyclooctynes are being applied, our observation of the behavior of different cyclooctynes inside cells should also prove useful to the protein labeling community and beyond.

For protein identification, we describe our work of engineering and using LplA to site-specifically target a benzophenone photocrosslinker. Our observations led us to the conclusion that although benzophenone is generally regarded as the more efficient and specific photocrosslinker than aryl azide and diazirine, its high geometric constraint to its proximal crosslinkable C-H bonds may decrease its crosslinking yield. Knowing the protein structure and amino acid environment surrounding benzophenone could help in choosing the most optimal position for the photocrosslinker. Finally, in a different crosslinking approach, we discuss our effort towards using a promiscuous peroxidase enzyme that generates biotin-phenol radicals to study membrane protein topology.

Thesis Advisor: Alice Y. Ting
Title: Ellen Swallow Richards Associate Professor of Chemistry

Acknowledgements

I would like to thank my advisor Professor Alice Ting for her support and inspiration over the past five years. Her passion, relentlessness, and perseverance towards tackling some of the most challenging problems in science have always been a source of motivation during my PhD study. I would like to thank her for her patience and encouragement during times when nothing was working. Her creative ideas and solutions have always inspired and encouraged me to think outside of the box to gain different perspectives.

I am thankful to the members of my thesis committee: Professor Alex Klibanov and Professor JoAnne Stubbe. I've learned a lot from Alex's biological tutorial class during my first year as a graduate student. I appreciate his encouragement during my PhD study. As for my thesis committee chair Professor JoAnne Stubbe, her deep passion and dedication to science are truly inspiring. I enjoyed our annual thesis meetings where she provided me not only with insightful scientific advice on my projects, but also guidance in my career development. During the development of this thesis, JoAnne provided very detailed comments and careful edits. This thesis would not be possible without her support and contribution.

I would also like to thank my undergraduate research advisor Professor Dirk Trauner (University of Munich) and the graduate student I worked with in his lab, Dr. Matthew Banghart. It was in Dirk's lab where I first became interested in developing chemical tools to study biological events. It was Dirk and Matt who first instilled in me the notion of carrying out organic chemical reactions inside a live cell or animal and what implications it could bring. Throughout my PhD career, I was able to keep contact with Matt who is doing a postdoc at Harvard Medical School. I enjoyed to our celebratory dinner when we published at the same time, I, for my first paper during graduate school and he, for his first publication during postdoc. I am thankful to Matt who is always understanding and encouraging.

For the members of the Ting lab, past and present, I am lucky and honored to have the opportunity to work with such a talented and diverse group of scientists. I would like to thank Dan, for being a source of entertainment that keeps us together and for always offering insightful advice; Tao, for training me when I first started working, for our enjoyable collaboration on the fluorophore targeting project, and for being a wonderful bench mate; Katie, for being the lab mother; Stephanie, for our scientific and non-scientific conversations and for being my lab rock. To all other members of the Ting lab: Jeff, Ken, Vicky, Kurt, Chai, Wenjing, Philipp, Jake, Carolyn, Oom, Monica, Kayvon and Betty Lou, thank you for making lab such a special and enjoyable work place, for supporting me through difficult times, and for sharing laughter and tears.

During the last two years of my PhD journey I feel blessed to have Philip by my side. I would like to thank him for always being supportive and encouraging, and for keeping me sane with his often insane ideas and fun projects. I am excited and look forward to many adventures in the future together. At last I am deeply grateful to my parents, who have supported me in every way they could and have sacrificed a lot so I can study worry free. I know I may not have called home as often as I should, but they were the ones on my mind when I had to work late into the night. This thesis is dedicated to my parents, without whom I would not be the person I am today.

Table of Contents

Title Page 1

Signature Page 2

Abstract 3

Acknowledgements 4

Table of Contents 6

List of Figures 9

List of Tables 14

List of Abbreviations 15

Chapter 1. Introduction: specific and non-specific enzyme-mediated small molecule labeling methods 18

Introduction 19

Enzymes and peptides as protein tags to target small molecule probes 20

Using bioorthogonal chemistries to label proteins with small molecules 24

Site-specific targeting and crosslinking with photocrosslinkers to identify protein-protein interactions 26

Promiscuous enzymes 28

Conclusion 31

References 32

Chapter 2. Development of intracellular protein labeling with diverse fluorophores using enzyme-mediated azide ligation and strain-promoted cycloaddition 37

Introduction 38

Screening for an efficient LplA ligase and azide pair for site-specific protein labeling inside live cells........42
Characterization of an improved LplA enzyme and azide probe pair........48
Exploring strain-promoted azide-alkyne cycloaddition reaction inside live cells using alternative source of azide........57
Conclusion........61
Experimental methods........62
References........67

Chapter 3. Optimization and demonstration of diverse fluorophore targeting inside live cells........70
Introduction........71
Comparison of LplA mutant enzymes and azide pairs using a live-cell labeling assay........72
Comparison of cyclooctyne structures........77
Demonstration of intracellular protein labeling with ADIBO-fluorescein diacetate...84
Challenges and optimization for targeting large fluorophores inside live cells........87
Extension of site-specific two-step labeling method to diverse fluorophore structures with labeling inside cells and on the cell surface........97
Measurement of overall two-step ligation yield in cells........100
Conclusion........102
Experimental methods........107
References........116

Chapter 4. Development of a photocrosslinker ligase for protein-protein interaction detection........120
Introduction to methods of protein-protein interaction detection........121
Classes of photocrosslinkers........126
Screening for a benzophenone ligase........130
Detection of ligated benzophenone via reaction with fluorophore-hydroxylamine....133

In vitro photocrosslinking of LAP-tagged FKBP and FRB proteins using benzophenone probe 142
Optimizations to improve *in vitro* photocrosslinking yield 147
Synthesis and LplA(W37G) incorporation of alkyl benzophenone 153
In vitro photocrosslinking with new alkyl benzophenone probe 157
Quantify benzophenone ligation yield inside live mammalian cells 162
Photocrosslinking inside live mammalian cells 166
LAP-actin photocrosslinking with drug treatment inside live mammalian cells 177
Conclusion 181
Experimental methods 183
References 192

Chapter 5. Effort towards membrane protein topology mapping with an ascorbate peroxidase enzyme 199
Introduction 200
EF-hand domain-containing protein 1 - LETM1 205
Restricted and diffusive labeling patterns of localized APEX 208
Biotin-tyramide labeling patterns of APEX-fused LETM1 211
Controlling expression levels of LETM1 fusion proteins 218
Conclusion and future direction 222
Experimental methods 224
References 228

Curriculum Vitae 231

List of Figures

Chapter 1: Introduction: specific and non-specific enzyme-mediated small molecule labeling methods

Chapter 2. Development of intracellular protein labeling with diverse fluorophores using enzyme-mediated azide ligation and strain-promoted cycloaddition

Figure 2-1. Natural ligation of lipoic acid catalyzed by wild-type LplA.................... 38

Figure 2-2. Small and large probes to be incorporated by LplA 40

Figure 2-3. Two-step LplA ligation using functional group handle and chemoselective bioorthogonal reaction inside live cells 41

Figure 2-4. Small molecule substrate binding pocket of *E. coli* lipoic acid ligase (LplA) with bound lipoyl-AMP 45

Figure 2-5. Screening to identify the best LplA mutant/azide substrate pair 47

Figure 2-6. *In vitro* characterization of azide 9 ligation catalyzed by LplA(W37I) 49

Figure 2-7. Fixed cell labeling on nuclear localized LAP-YFP with cy5-alkyne via CuAAC 51

Figure 2-8. Characterization of intracellular azide ligation yields by native gel-shift assay 53

Figure 2-9. Schematic representation of streptavidin gel-shift assay to measure azide ligation yield 54

Figure 2-10. Characterization of intracellular azide ligation yields by streptavidin gel-shift assay 55

Figure 2-11. Fixed cell detection of metabolically incorporated azido sugar analog azido-GalNAc with cy3-alkyne via CuAAC 58

Figure 2-12. Live cell detection of metabolically incorporated azido sugar analog azido-GalNAc with strain-promoted azide-alkyne cycloaddition 60

Chapter 3. Optimization and demonstration of diverse fluorophore targeting inside live cells

Figure 3-1. Determination of azide 9 washout time in live cells 72

Figure 3-2. Identification of the best LplA mutant/azide substrate pair for intracellular protein labeling 74
Figure 3-3. Quantitation of data shown in Figure 3-2 75
Figure 3-4. Immunofluorescence with anti-FLAG to compare labeling against expression of all enzymes 76
Figure 3-5. Evaluation of various cyclooctyne structures for site-specific intracellular protein labeling 80
Figure 3-6. Analysis of background labeling by cyclooctyne-fluorescein conjugates 82
Figure 3-7. Two-dimensional optimization of ADIBO-fluorescein diacetate loading concentration and washout time 84
Figure 3-8. Two-step fluorophore targeting via strain-promoted azide-alkyne cycloaddition 85
Figure 3-9. ADIBO−fluorescein labeling of three localized LAP−BFP fusions, LAP−β-actin, and LAP−MAP2 86
Figure 3-10. Comparison of MOFO-ATTO 647N and ADIBO-ATTO 647N nuclear labeling specificities 88
Figure 3-11. Comparison of two different linker structures for MOFO-ATTO 647N 89
Figure 3-12. Suppression of mitochondrial background labeling with FCCP 91
Figure 3-13. Suppression of mitochondrial background labeling with QSY21 92
Figure 3-14. Suppression of mitochondrial background labeling with FCCP and QSY21 94
Figure 3-15. Comparison of ADIBO-ATTO 647N labeling of LAP-β-actin with and without suppression of mitochondrial background using QSY21 quencher 96
Figure 3-16. Intracellular protein labeling with diverse fluorophore structures 97
Figure 3-17. Cell surface and intracellular protein labeling with DIBO-Alexa Fluor 647 and DIBO-biotin 99
Figure 3-18. Streptavidin gel-shift analysis of two-step labeling yield in live cells.... 101

Chapter 4. Development of a photocrosslinker ligase for protein-protein interaction detection

Figure 4-1. Three classes of photocrosslinkers 126
Figure 4-2. Lipoic acid ligase (LplA) with either bound lipoic acid or a benzophenone unnatural small molecule docked into the enzyme active site. 130
Figure 4-3. Synthetic scheme of benzophenone photocrosslinkers 131
Figure 4-4. Screening to identify LplA mutant/benzophenone substrate pair 132
Figure 4-5. Unsuccessful attempt to detect *in vitro* ligation of benzophenone to E2p protein using biotin-hydrazide 134
Figure 4-6. Hydrolysis of hydrazones and oximes 135
Figure 4-7. Reaction between benzophenone and biotin-hydroxylamine to form oxime product. 137
Figure 4-8. Detection of *in vitro* ligated benzophenone using AF488-hydroxylamine 139
Figure 4-9. Detection of benzophenone chemically conjugated to FRB-YFP using biotin-hydroxylamine 141
Figure 4-10. Crystal structure of FKBP-rapamycin-FRB ternary complex................. 142
Figure 4-11. *In vitro* photocrosslinking with purified proteins FKBP-LAP and FRB-YFP 146
Figure 4-12. Crystal structure of FRB-rapamycin-FKBP ternary complex with methionine residues in ball-and-stick 148
Figure 4-13. *In vitro* photocrosslinking with methionine to alanine mutant proteins of FKBP-LAP.. 149
Figure 4-14. *In vitro* photocrosslinking of myc-FKBP-LAP and HA-FRB 151
Figure 4-15. Scheme of benzophenone insertion into nearby C-H bond after UV irradiation 153
Figure 4-16. Benzophenone and benzophenone derivatives with different functional groups attached to the fourth position of the benzophenone core 155
Figure 4-17. Synthetic scheme of alkyl benzophenone probes with methylene unit attached to the benzophenone core 155

Figure 4-18. *In vitro* characterization of para alkyl benzophenone ligation by LplA(W37G) 156
Figure 4-19. *In vitro* photocrosslinking comparison between alkyl benzophenone and benzophenone 159
Figure 4-20. Reducing concentrations of FKBP-LAP(F) and FRB helps in eliminating background photocrosslinking 161
Figure 4-21. Gel-shift analysis of benzophenone ligation yield in cells 163
Figure 4-22. Native gel-shift assay to measure benzophenone ligation yield over time 164
Figure 4-23. Expression of FKBP-LAP and FRB in HEK 293T cells 167
Figure 4-24. Live cell benzophenone ligation and photocrosslinking of myc-FKBP-LAP and HA-FRB 168
Figure 4-25. Benzophenone photocrosslinker labeling scheme of LAP-actin in live cells 169
Figure 4-26. Benzophenone ligation and photocrosslinking on HA-LAP/LAP(F)-actin in live mammalian cells 170
Figure 4-27. Western blot stained with anti-FLAG antibody to detect FLAG-tagged LplA(W37G) after live cell photocrosslinking of HA-LAP/LAP(F)-actin 172
Figure 4-28. Live cell photocrosslinking of HA-LAP-actin with alkyl benzophenone incubation time course 174
Figure 4-29. Benzophenone ligation and photocrosslinking performed on HA-LAP-MAP2 in live mammalian cells 175
Figure 4-30. Schematic representations of different LAP-actin crosslinked complexes with and without depolymerization drug treatment 177
Figure 4-31. Live cell photocrosslinking of LAP-actin with various drug treatments 179

Chapter 5. Effort towards membrane protein topology mapping with an ascorbate peroxidase enzyme

Figure 5-1. Scheme depicting labeling with APEX fused to either matrix or IMS side of an IMM protein 204
Figure 5-2. Amino acid sequence of human LETM1 206
Figure 5-3. Restricted and diffusive biotin-tyramide labeling patterns in HEK 293T and COS-7 cells 209
Figure 5-4. Constructs with APEX fused to different regions of LETM1 211
Figure 5-5. APEX115-LETM1 fusion protein shows cleavage products on Western blot 212
Figure 5-6. Biotin-tyramide labeling with APEX-LETM1 fusion constructs in HEK 293T cells 217
Figure 5-7. Western blot analysis of APEX-LETM1 fusion proteins expressed in HEK 293T cells 219
Figure 5-8. Western blot analysis of COS-7 cells lipofected (L2k) or viral infected with LETM1 or LETM1-mCherry construct 220

List of Tables

Table 4-1. Optimized UV photocrosslinking conditions for each instrument tested...144

List of Abbreviations

aa	amino acids
ACP	acyl-carrier protein
ATP	adenosine triphosphate
azide 7	8-azidooctanoic acid
azide 8	9-azidononanoic acid
azide 9	10-azidodecanoic acid
azide 10	11-azidoundecanoic acid
BFP	blue fluorescent protein
BSA	bovine serum albumin
BT	biotin-tyramide
CFDA	carboxyfluorescein diacetate
CHO	Chinese hamster ovarian cells
CuAAC	copper-catalyzed azide-alkyne cycloaddition
DAB	diaminobenzidine
DCM	dichloromethane
DHFR	dihydrofolate reductase
DIC	differential interference contrast
DIFO	difluorinated cyclooctyne
DIMAC	6,7-dimethoxyazacyclooct-4-yne
DMEM	Dulbecco's modified eagle medium
DMF	N,N-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DPBS	Dulbecco's phosphate buffered saline
DTT	dithiothreitol
E2p	9 kDa hybrid lipoyl domain derived from E.coli PDH complex
E. coli	Escherichia coli
EDTA	ethylenediaminetetracetic acid
ER	endoplasmic reticulum
FP	fluorescent protein
ER	endoplasmic reticulum
ESI	electrospray ionization
F-actin	filamentous actin
FKBP12	FK506 binding protein 12
FlAsH	fluorescein arsenical hairpin binder
FRB	FKBP-rapamycin binding domain of the mTOR protein
FRET	fluorescence resonance energy transfer
FP	fluorescent protein
G-actin	globular actin
GFP	green fluorescent protein
hAGT	O^6-alkylguanine-DNA alkyltransferase
HA	human influenza hemagglutinin epitope
Halo-tag	haloalkane dehalogenase

HDDA	N,N'-dimethyl-1,6-diaminohexane
HEK 293T	human embryonic kidney cells 293T
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HP1	heterochromatin protein 1
HPLC	high-performance liquid chromatography
IMM	inner mitochondrial membrane
IMS	intermembrane space
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LAP1	22-amino acid rationally designed LplA acceptor peptide
LAP	the best 13-amino acid LplA acceptor peptide evolved from yeast display
LAP(F)	LAP peptide with Trp downstream of catalytic Lys residue changed to Phe
LC-MS/MS	liquid chromatography, tandem mass spectrometry
LplA	Escherichia coli lipoic acid ligase
MAP2	microtubule-associated protein 2
MeCN	acetonitrile
MEM	minimal eagle medium
MOFO	monofluorinated cyclooctyne
MS	mass spectrometry
nAv	neutravidin
NES	nuclear export signal
NHS	*N*-hydroxysuccinimide
NLS	nuclear localization sequence
NMR	nuclear magnetic resonance
OMM	outer mitochondrial membrane
PAGE	polyacrylamide gel electrophoresis
PAL	photoaffinity labeling
PALM	photoactivated localization microscopy
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein database
PFA	paraformaldehyde
PKC	phosphokinase C
PPTase	phosphopantetheine transferase
QY	quantum yield
ROS	reactive oxygen species
SA	streptavidin
SA-HRP	streptavidin-conjugated horseradish peroxidase
s.d.	standard deviation
SDS	sodium dodecyl sulfate
STED	stimulated emission depletion microscopy
STORM	stochastic optical reconstruction microscopy
TA	transient absorption
TBS-T	Tris-buffered saline supplemented with Tween-20
TBTA	tris-(benzyltriazolylmethyl)amine
TEA	triethylamine
TLC	thin-layer chromatography

TM	transmembrane
TMR	tetramethylrhodamine
Tris	tris(hydroxymethyl)aminomethane
UAA	unnatural amino acid
WB	western blotting
WCL	whole cell lysate
YFP	yellow fluorescent protein

Chapter 1. Introduction:
specific and non-specific enzyme-mediated small molecule labeling methods

Introduction

Scientists have long adapted enzymes for fundamental biochemistry research. The unifying theme of this thesis is to utilize the natural and engineered abilities of enzymes as robust biochemical tools that can be applied to study biological processes. This is accomplished in parallel with the design and synthesis of organic small molecule probes, such as fluorophores or probes baring a chemoselective functional handle, to be recognized and incorporated by the enzymes.

Green fluorescent protein (GFP) and its multicolored variants (1) are perhaps the most famous and widely used biochemical tools that have revolutionized cell biologists' ability to gain insight into cellular processes in living cells. Fluorescent proteins (FPs) serve as imaging protein reporter tags to visualize protein movement and localization. FPs offer perfect targeting specificity through genetic fusion and are generally stable upon chromophore maturation in all cellular compartments. Aside from routine imaging applications, scientists have engineered FPs for various applications such as photoactivatable FPs for super-resolution imaging (2) and split-GFP (3) to monitor protein trafficking, interaction and maturation. However, despite the many advantages and utilities of FPs, they carry a payload of 27 kDa when fused to the protein of interest, and are generally less bright and photostable compare to small organic fluorophores (4). We and others have observed interference caused by FP fusion to protein of interest (5). For example, mCherry-tagged actin is excluded from the nucleus of the cell, likely because the mCherry is interfering with the binding of cofilin, a nuclear importer of actin.

Small molecule organic fluorophores are brighter, more photostable, and far smaller than FPs (4). However, they do not have the ease and perfect specificity of genetic targeting offered by FPs. Scientists have therefore came up with strategies to target small molecules fluorophores via the use of enzymes.

Enzymes and peptides as protein tags to target small molecule probes

Covalent enzymatic tags (self-modifying enzymes) and non-covalent enzymatic tags

To exploit using enzymes as protein tags to target small organic molecules, the group of Kai Johnsson employed the ubiquitous human DNA repair enzyme O^6-alkylguanine-DNA alkyltransferase (hAGT) (6) which transfers alkyl groups from the O^6-position of guanine onto a cysteine residue in its active site in a one-turnover self-modification reaction. The Johnsson group has derivatized the para position of the benzyl group of O^6-benzyl guanine with a variety of probes, like fluorophores and biotin, to achieve specific targeting (7,8) of these molecules by fusing hAGT to protein of interest. Further evolution based on phase display led to a more active quadruple mutant hAGT enzyme whose activity is 20 times faster than the wild-type hAGT, which eliminated the need for hAGT deficient cell lines (9). This more active hAGT, termed SNAP-tag, has been used to target benzyl guanines derivatized with fluorescein, biotin, and digoxigenin probes. Later, the Johnsson group expanded the substrate specificity of hAGT through directed evolution to report another hAGT mutant that exhibits 100 times greater preference for cytosine over guanine derivatives. The evolved cytosine-specific enzyme, termed CLIP-tag (10), has been applied together with SNAP-tag to orthogonally label two populations of proteins with different fluorophores inside the same cell.

Another example of self-modifying enzyme that has been adapted as enzymatic tag is a mutant form of the enzyme haloalkane dehalogenase (11). The wild-type haloalkane dehalogenase hydrolyzes haloalkanes by forming a covalent bond between the asparagine residue in its active site and the alkyl group in an S_N2 reaction. A nearby histidine residue in the active site then hydrolyzes the covalent bond to recycle the enzyme. In an attempt to convert this enzyme to a covalent labeling tag, researchers at Promega installed a His289 mutation that eliminated the second hydrolysis step (12). This mutant form of haloalkane dehalogenase is named HaloTag. Substrates for HaloTag include haloalkane conjugated fluorophores and functionalized agarose surface resin for protein purification (11). The HaloTag labeling technology possesses high specificity and low background labeling due to the lack of endogenous dehalogenases in eukaryotic cells. Recently it has been applied in purification of the protein cannabinoid receptor CB2 (13), fused to p75 neurotrophin receptor and tubulin to assess the retrograde axonal transport of the receptor (14), as well as specific labeling of xenograft tumors in living animals (15).

In addition to covalent modification on self-labeling enzymatic tags, several proteins with tight binding to their substrates have also been exploited as non-covalent labeling methods. One of the first examples of non-covalent tagging is the *E. coli.* dihydrofolate reductase (eDHFR), which has high affinity to its antibiotic small molecule ligand trimethoprim (TMP) with K_D as low as 1 nM (16). TMP conjugates to various fluorescent probes can be synthesized by introducing the fluorophores to the para-methoxy position of TMP. Modifications introduced at this position of TMP raises the K_D only slightly to ~35 nM, which is still a lot lower than the affinity between TMP and mammalian DHFR. The eDHFR-TMP receptor-ligand pair has been demonstrated in labeling proteins in the nucleus, plasma membrane, and cytosol in various cell types (16,17).

Another example of non-covalent protein tag based on the high binding affinity between enzyme and its ligand is the FK506 binding protein 12 (FKBP12), and its natural ligand rapamycin, a system pioneered by the Schreiber group (18,19). Clarkson and coworkers developed a nonnative FKBP12-rapamycin pair by engineering a FKBP12(F36V) mutant which has high affinity (K_D~0.094 nM) to a synthetic derivative of rapamycin that can be further conjugated to fluorophores and other probes (20,21). Robers *et al* demonstrated labeling of numerous intracellular proteins fused to FKBP12(F36V) with different rapamycin derivative conjugates (21). However, even though the two non-covalent enzymatic tagging strategies discussed do have very high affinities between receptor and ligand, they are not suitable for long-term studies compared to covalent enzymatic tags.

Overall, enzymatic protein tags, either based on covalent self-labeling or high-affinity non-covalent interaction, provide a way to target small molecule organic fluorophores inside live mammalian cells with high specificity. A significant drawback is that most enzymatic tags are still fairly large. The smallest enzymatic tag that has been demonstrated for specific probe targeting *inside* live cells mentioned above is the 12 kDa FKBP12(F36V) enzyme but its non-covalent binding limits its application. In the next section we review enzymatic labeling techniques that are based on small peptide tag fusions to address this issue.

Enzymatic labeling of peptide tags

The ideal labeling method would combine the exquisite specificity conferred by enzymes, the superior photophysical property of small molecule fluorophores, together with minimal perturbation to the native protein. While specific targeting to endogenous proteins is an extremely challenging feat, scientists have developed enzyme-based techniques that only require fusing a small peptide to the protein of interest. Enzyme-mediated small peptide tags are recognized by specific enzymes, which then selectively modify the tags with small molecule substrates.

Zhou *et al.* used phage display evolution to evolve two orthogonal 12–amino acid peptides A1 and S6 as efficient substrates for Sfp and AcpS phosphopantetheinyl transferases (PPTases) (22). The PPTases catalyze the covalent linkage of a coenzyme A-derivatized probe to a serine residue in the peptide substrates. However, the charged phosphopantetheine arm makes the probe conjugates difficult to cross the cell membranes, limiting the method to labeling only on the cell surface (22,23).

Our lab has been a pioneer in this area of enzyme-mediated peptide tag labeling methods. We developed probe targeting techniques based on lipoic acid ligase (LplA) and biotin ligase (BirA) together with their respective peptide tags. We engineered the *E.coli* enzyme lipoic acid ligase (LplA) though site-directed mutagenesis to incorporate unnatural small molecules such as fluorophores (5,24,25), functional groups (26–28), and a photocrosslinker (29) site-specifically onto the lysine residue in a 13-amino-acid LplA acceptor peptide (LAP). To develop this technology we now call PRIME (PRobe Incorporation Mediated by Enzymes), we explored the plasticity of the small molecule binding pocket of LplA and also reduced the size of its protein substrate from the 9 kDa E2p subunit of pyruvate dehydrogenase to the 13- amino acid LAP through yeast-display evolution (30). By examining the crystal structure of *E. coli* LplA with bound lipoyl-AMP, we engineered LplA to accept unnatural probes through alanine scanning of amino acid residues in the active site (29) and identifying a key "gatekeeper" residue, Trp37, that when mutated to smaller amino acids can enlarged the LplA active site to enable it to bind to larger unnatural probes, such as different derivatives of coumarin (5,24,25), aryl azide (29), alkyl azide (26,27), and *trans*-cyclooctene (28). Dr. Chayasith Uttamapinant and coworkers demonstrated specific labeling of coumarin fluorophore by LplA(W37V) onto many LAP-tagged proteins inside live mammalian cells (5). More importantly, through this peptide-tagging method, they were able to detect the presence of actin in the nucleus, which was shown to preclude actin when it was fused to a large fluorescent protein, further highlighting and confirming the importance of a small tag.

The *E. coli* biotin ligase (BirA) covalently attaches biotin to a lysine residue of a 15-amino acid biotin acceptor peptide (AP). A ketone derivative of biotin was synthesized for subsequent chemoselective reaction with hydrazide or hydroxylamine-conjugated biophysical probes in a two-step process (31). Dr. Sarah Slavoff and coworkers also explored BirA from yeast and *Pyrococcus horikoshii* to label azide- or alkyne-derivatized biotin analogs, which allowed for downstream detection using click chemistry (32). However, one limitation of the BirA labeling method is that since endogenous biotin is still preferred as the substrate, the labeling method either requires biotin starvation of the cells prior to BirA intracellular labeling or is restricted to the cell surface.

The Ploegh group harvested the enzyme sortase A from *Staphylococcus aureus* which recognizes a only 5-amino acid peptide sequence LPXTG, the shortest peptide tag available (33,34). The sortase A enzyme performs a transpeptidation reaction by first cleaving the peptide bond between the T and G residues, and at the same time forming a new peptide bond to a polyG-derivatized small molecule. However, this method is also restricted to the cell surface because the polyG substrates are not cell permeable.

Using bioorthogonal chemistries to label proteins with small molecules

So far we have discussed using enzymes that are able to target fluorophores functionalized with ligands they recognize – haloalkane for Halo-Tag, guanine or cytosine for hAGT and its variant, TMP for eDHFR, so the labeling reaction is complete in a single step by the specific enzyme. In another approach, bioorthogonal chemistries between two functional groups that are not found inside mammalian cells have been extensively explored as a way to target small molecules with high specificity. Chemical labeling techniques that utilize bioorthogonal reactions (35) to target small organic fluorophores inside live cells have seen much advancement over the years. Aside from harnessing specificity from Nature, chemists have began to develop reactions and engineer functional group handles that can specifically recognize and react with each other not only in the exogenous environment of an isolated flask, but inside the complex milieu of cells and living organisms. To develop bioorthogonal reactions as a way to target fluorophores in a biological environment, the functional groups, or reactive "handles" involved in the bioorthogonal reaction must stay inert until they meet to selectively to react with each other; the reaction condition must be biocompatible in terms of pH and temperature and nontoxic to the biological system under study; the reaction must have fast kinetics so that the product is formed at a reasonable rate on the scale of protein turnover and the progression of biological processes, even when the reactant concentrations are preferably low. The bioorthogonal product formed must be stable to proteases and not be degraded so to have a consistent readout to be useful as chemical reporters. Excess reactants should be removed from the system to minimize unwanted side reaction, or masked to reduce possible background signal. The versatility of bioorthogonal reactions has been demonstrated by their ability to tag proteins and glycans. The Staudinger ligation between azide and triarylphosphine was the first bioorthogonal reaction used to label cell-surface azide-derivatized glycans on live cells (36), and subsequently in live animals (37). Despite much effort in mechanistic modifications to increase the reaction kinetics, reduce phosphine oxidation, and aza-ylide hydrolysis, its slow second-order rate constant (10^{-3} $M^{-1}s^{-1}$) leaves much more to be desired (38,39).

Even though the Staudinger ligation is slow, it brought to chemists' attention the potential of organic azide as an excellent functional group handle. The small size of the azide, its absence from living organisms, its kinetic stability inside cells, and its nontoxic profile as demonstrated by early approved drugs like azidothymidine all hinted to the potential of organic azide as an excellent functional group handle. The copper-free or strain-promoted azide-alkyne cycloaddition was adapted from the Huisgen [3+2] cycloaddition that initially used heat and pressure to drive the reaction (40), as well as the later developed Cu(I)-catalyzed azide-alkyne cycloaddition demonstrated by Sharpless and Meldal (41). Instead of using heat, pressure, or Cu(I) which are incompatible inside cells, the strain-promoted azide-alkyne cycloaddition uses energy stored in the ring strain of cyclooctyne to drive the reaction. Since 2000, numerous generations of cyclooctynes were synthesized with improved reactivity, kinetics, hydrophilicity, and overall biocompatibility. The highest second-order rate constant of a strain-promoted azide-alkyne cycloaddition reaction that has been demonstrated on live cell-surface was reported to be 0.96 $M^{-1}s^{-1}$ (42), twenty times higher than the initial Staudinger ligation. In the first part of this thesis, strain-promoted azide-alkyne cycloaddition is used for the targeting of diverse fluorophores to specific proteins inside live cells.

Yet another recent addition to the toolbox of bioorthogonal reactions is the reaction between tetrazine and *trans*-cyclooctene (43). Developed by the Fox group, this inverse Diels-Alder reaction is the fastest bioorthogonal reaction known to date, optimized to achieve an impressive rate constant of 22,000 $M^{-1}s^{-1}$. Our lab has demonstrated two-step fluorophore targeting using a lipoic acid ligase enzyme in combination with Diels-Alder cycloaddition to perform intracellular protein labeling (28). The very fast cycloaddition kinetics yields substantial improvements in signal to background ratio. Furthermore, the tetrazine-fluorophore conjugate is also fluorogenic, which is an additional benefit that masks the background caused by unreacted reagents.

Site-specific targeting and crosslinking with photocrosslinkers to identify protein-protein interactions

So far we have discussed a few probe labeling techniques that capitalize on the specificity of enzymes and bioorthogonal reactions to target small molecules to protein of interest inside cells. We have seen many enzyme-based fluorophore or probe targeting methods that have been demonstrated to label a specific protein among the plethora of proteins inside live mammalian cells. High specificity is a critical requirement since the target is often a single protein. In a different approach, scientists sought methods to label and identify multiple proteins associated with a protein of interest or within a particular cellular compartment. Identifying the many proteins of a certain proteome, within a certain compartment of the cell, or under an altered cellular state requires a different labeling approach.

One of the conventional approaches to capture groups or complexes of proteins is affinity-based methods. Methods such as co-immunoprecipitation (44) that uses an antibody against the protein of interest requires first lysing the cells, then immunocapture the protein of interest together with its interacting proteins. Subsequent Western blots or mass spectrometry (MS) analysis is used to identify proteins that are pulled down. However, using cell lysates as starting material already introduces numerous artifacts. First, the proteins are removed from their native environment. In cell lysate, without cellular compartmentalization that maintains various pHs and effective concentrations of biomolecules, complexes that form depending on those factors could dissociate or aggregate with other proteins. Second, disruption of protein complexes during cell lysis also presents a serious problem. Weakly bound proteins could be lost during washing, or non-specific binding of proteins that are not endogenous interacting partners could wrongly associate. Last but not least, the limited availability of high quality of antibodies introduces additional constraint.

As an alternative, chemical crosslinking prior to immunocapture is employed to covalently link interacting proteins while the proteins are in their endogenous states. Covalent linkage allows more stringent washing conditions to remove non-specifically interacting proteins without fear of losing weakly bound proteins (45). Traditional chemical crosslinking reagents include formaldehyde or amine-reactive moieties such as

N-hydroxysuccinimide (NHS) esters. Formaldehyde can readily react with lysines, tryptophans, or N-termini of proteins after minutes of exposure; however, it does not have any spatial resolution or protein specificity once it permeates cell membranes. To address this issue, photoaffinity labeling (PAL) (46) that uses site-specifically conjugated photoreactive crosslinkers, or photocrosslinkers, enables identifying interacting proteins specific to a protein of interest. Photocrosslinkers such as aryl azides, benzophenones, and diazirines can be site-specifically introduced to proteins via *in vitro* cysteine-maleimide chemistry or incorporated as unnatural amino acids using the unnatural amino acid (UAA) mutagenesis technique developed by the Schultz group (47). Crosslinking with site-specific photocrosslinkers offers high temporal and spatial resolution to capture proteins in their native states and has been demonstrated inside live mammalian cells. However, as discussed previously, techniques to site-specifically target these small molecule photocrosslinkers are still limited. In Chapter 4, we introduce an enzyme-based ligation technique for the benzophenone photocrosslinker.

Promiscuous enzymes

Achieving location specificity of photocrosslinkers or other reactive species to form covalent linkage with proteins inside live cells is a challenging task. By site-specifically attaching a photocrosslinker to a protein of interest allows us to pull down one or multiple interacting proteins, but would not enable us to identify, for example, the entire proteome of the mitochondrial matrix under drug treatment. To identify the proteome of an entire cell or a specific organelle under its native environment without contamination from cell lysate or organelle purification, scientists have yet again looked to enzymes for solutions. One approach is to use enzymes that can convert small molecule probes to diffusive reactive intermediates that can form covalent bonds with its nearby proteins. Having diffusive reactive species versus site-specifically attached crosslinkers also obviates the bottleneck of crosslinker length and shape, which limits its range of reactivity. Several enzymes that are able to take a chemical probe and generate a reactive intermediate have been applied in this type of promiscuous protein labeling. Horseradish peroxidase (HRP) was first applied as a reporter enzyme for electron microscopy (EM) (48) by reacting with diaminobenzidine (DAB) in the presence of H_2O_2 to form an insoluble tar-like product that is electron-opaque and can serve as an EM contrast agent. Later, HRP was adapted by Kotani *et al.* to catalyze the generation of nitrenes from aryl azides (49). In this enzyme-mediated activation of radical source (EMARS) method, cell membrane-targeted HRP generated radicals of aryl azide-biotin conjugate that crosslinked to nearby proteins, which can be pulled down and analyzed using either streptavidin-coated beads or membrane protein specific antibody. EMARS labeling has been applied to study cell-surface molecular clusters of β1-integrin, epidermal growth factor receptor (EGFR), insulin-like growth factor-I receptor (IGF1R), and EphA2 receptor (50,51). Immunoelectron microscopy of the EMARS-labeled HRP-conjugated antibody showed 96% of biotinylated proteins were within 300 nm of the antibody-coated particles (50). However, one limitation of HRP is that the enzyme is not active in the cytosol, perhaps because its disulfide bonds were reduced under the reducing environment.

Promiscuous biotin ligase is another example of an enzyme that can generate reactive small molecules. Through a point mutation (R118G) (52) in the wild-type biotin ligase which has exquisite specificity towards its endogenous protein substrate, promiscuous biotin ligase is able to prematurely release the adenylated biotinoyl-5'-AMP from its active site without the protein substrate present. Once released, biotinyl-5'- AMP can react with amine functional groups on nearby proteins forming an amide bond. Biotinylated proteins are then enriched using streptavidin binding and analyzed on Western blot. However, the time of the reaction is especially long, requiring longer than 6 hours of biotin incubation to saturated labeling. Even though the labeling is performed in intact live cells, but the prolonged labeling time could cause artifacts in endogenous protein interactions. Biotinylated proteins could diffuse away leading to poor spatial resolution or even cause subsequent unwanted cellular changes. Furthermore, the lack of temporal resolution would not allow differentiation and detection of rapid changes in proteomes. Labeling with promiscuous biotin ligase was demonstrated by tagging the enzyme to nuclear lamina protein, lamin A, and the construct was expressed and labeled in HEK 293 cells (53). Roux *et al* identified 122 proteins, some of which are known proteins associated with the nuclear envelope, while others are novel, such as the protein FAM169A (later called SLAP75) which does not have any predicted transmembrane region or any sequence motifs that would hint to its localization. FAM169A was also confirmed by immunostaining to be located in the nuclear envelope.

To develop an enzyme that can be specifically targeted to a region within the cell, works in different organelle environments, and has high temporal and spatial resolution, our lab has engineered a plant ascorbate peroxidase called APEX to catalyze the formation of biotin-tyramide radicals in the presence of H_2O_2 (54). The short lifetime (less than 1 μs) of the radical together with its membrane-impermeant nature ensure tight labeling radius and high temporal resolution. Rhee *et al.* targeted APEX to the mitochondrial matrix and the inner mitochondrial space (IMS), two regions that are especially difficult to purify through subcellular fractionation and organelle purification (54). Using matrix-targeted APEX, biotin-tyramide and H_2O_2 labeling, followed by streptavidin enrichment and tandem mass spectrometry analysis, 495 proteins were identified in the matrix proteome, with 464 proteins having prior mitochondrial annotation (54). Notably, proteins such as CPOX and PPOX which were previously assigned to the mitochondrial intermembrane space were detected in the matrix proteome, leading to novel insights to their trafficking and functional roles in the heme biosynthesis pathway. Currently APEX is being applied towards identifying the proteome of especially challenging regions of the cell that are impossible to purify through traditional means, such as the synaptic space or the junction between mitochondria and the endoplasmic reticulum (ER). In the final chapter of this thesis, we describe extending APEX labeling to probe the topology of membrane proteins.

Conclusion

Despite the advances in cell biology brought by the fluorescent proteins, their large size and poor photophysical properties prompted scientists to look for alternatives. Compared to fluorescent protein tags, chemical tags such as small molecule fluorophore or other biophysical probes are smaller, brighter and more versatile. To target small molecules site-specifically to a protein of interest inside the cell, scientists have utilized the exquisite specificity of enzymes as well as developed bioorthogonal chemical reactions that can work efficiently within the complex environment inside the cell. In Chapter 2 and 3, we describe a fluorophore targeting method that uses only a small peptide tag in combination with enzymatic labeling and azide-alkyne cycloaddition to label numerous proteins inside cells with various fluorophores.

While being able to visualize the localization, movement, and interaction of a specific protein inside its native environment has brought tremendous insights to cell biology regarding the how a cell functions, the ability to identify with high temporal resolution pools or complexes of proteins associated with a single protein, of a particular cellular organelle, or even of the entire cell under a specific cellular condition has also revolutionized the way scientists study and understand biological processes at the molecular and cellular level. In Chapter 4, we present a method to site-specifically ligate a photocrosslinker to identify protein-protein interactions. Aside from targeted photocrosslinkers, scientists have developed enzymes that can generate reactive species, such as radicals, which then react nonspecifically with proximal proteins. In Chapter 5, we discuss current effort towards developing two parallel but distinct approaches to use enzyme-generated radical species to study membrane protein topology.

References

1. Sample V, Newman RH, Zhang J. The structure and function of fluorescent proteins. Chem Soc Rev. 2009 Oct;38(10):2852–64.

2. Wiedenmann J, Gayda S, Adam V, Oswald F, Nienhaus K, Bourgeois D, et al. From EosFP to mIrisFP: structure-based development of advanced photoactivatable marker proteins of the GFP-family. J Biophotonics. 2011 Jun;4(6):377–90.

3. Ozawa T. Protein reconstitution methods for visualizing biomolecular function in living cells. Yakugaku Zasshi. 2009 Mar;129(3):289–95.

4. Lavis LD, Raines RT. Bright ideas for chemical biology. ACS Chem. Biol. 2008 Mar 20;3(3):142–55.

5. Uttamapinant C, White KA, Baruah H, Thompson S, Fernández-Suárez M, Puthenveetil S, et al. A fluorophore ligase for site-specific protein labeling inside living cells. PNAS. 2010 Jun 15;107(24):10914–9.

6. Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K. A general method for the covalent labeling of fusion proteins with small molecules in vivo. Nat Biotech. 2003 Jan;21(1):86–9.

7. Keppler A, Pick H, Arrivoli C, Vogel H, Johnsson K. Labeling of fusion proteins with synthetic fluorophores in live cells. Proc. Natl. Acad. Sci. U.S.A. 2004 Jul 6;101(27):9955–9.

8. Keppler A, Kindermann M, Gendreizig S, Pick H, Vogel H, Johnsson K. Labeling of fusion proteins of O6-alkylguanine-DNA alkyltransferase with small molecules in vivo and in vitro. Methods. 2004 Apr;32(4):437–44.

9. Juillerat A, Gronemeyer T, Keppler A, Gendreizig S, Pick H, Vogel H, et al. Directed evolution of O6-alkylguanine-DNA alkyltransferase for efficient labeling of fusion proteins with small molecules in vivo. Chem. Biol. 2003 Apr;10(4):313–7.

10. Gautier A, Juillerat A, Heinis C, Corrêa IR Jr, Kindermann M, Beaufils F, et al. An engineered protein tag for multiprotein labeling in living cells. Chem. Biol. 2008 Feb;15(2):128–36.

11. Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, et al. HaloTag: a novel protein labeling technology for cell imaging and protein analysis. ACS Chem. Biol. 2008 Jun 20;3(6):373–82.

12. Pries F, Kingma J, Krooshof GH, Jeronimus-Stratingh CM, Bruins AP, Janssen DB. Histidine 289 is essential for hydrolysis of the alkyl-enzyme intermediate of haloalkane dehalogenase. J. Biol. Chem. 1995 May 5;270(18):10405–11.

13. Locatelli-Hoops S, Sheen FC, Zoubak L, Gawrisch K, Yeliseev AA. Application of HaloTag technology to expression and purification of cannabinoid receptor CB2. Protein Expr. Purif. 2013 May;89(1):62–72.

14. Mok S-A, Lund K, Lapointe P, Campenot RB. A HaloTag® method for assessing the retrograde axonal transport of the p75 neurotrophin receptor and other proteins in compartmented cultures of rat sympathetic neurons. J. Neurosci. Methods. 2013 Mar 30;214(1):91–104.

15. Tseng J-C, Benink HA, McDougall MG, Chico-Calero I, Kung AL. In Vivo Fluorescent Labeling of Tumor Cells with the HaloTag® Technology. Curr Chem Genomics. 2012;6:48–54.

16. Miller LW, Cai Y, Sheetz MP, Cornish VW. In vivo protein labeling with trimethoprim conjugates: a flexible chemical tag. Nat Meth. 2005 Apr;2(4):255–7.

17. Calloway NT, Choob M, Sanz A, Sheetz MP, Miller LW, Cornish VW. Optimized fluorescent trimethoprim derivatives for in vivo protein labeling. Chembiochem. 2007 May 7;8(7):767–74.

18. Bierer BE, Somers PK, Wandless TJ, Burakoff SJ, Schreiber SL. Probing immunosuppressant action with a nonnatural immunophilin ligand. Science. 1990 Oct 26;250(4980):556–9.

19. Spencer DM, Wandless TJ, Schreiber SL, Crabtree GR. Controlling signal transduction with synthetic ligands. Science. 1993 Nov 12;262(5136):1019–24.

20. Clackson T, Yang W, Rozamus LW, Hatada M, Amara JF, Rollins CT, et al. Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity. Proc. Natl. Acad. Sci. U.S.A. 1998 Sep 1;95(18):10437–42.

21. Robers M, Pinson P, Leong L, Batchelor RH, Gee KR, Machleidt T. Fluorescent labeling of proteins in living cells using the FKBP12 (F36V) tag. Cytometry A. 2009 Mar;75(3):207–24.

22. Zhou Z, Cironi P, Lin AJ, Xu Y, Hrvatin S, Golan DE, et al. Genetically encoded short peptide tags for orthogonal protein labeling by Sfp and AcpS phosphopantetheinyl transferases. ACS Chem. Biol. 2007 May 22;2(5):337–46.

23. Yin J, Straight PD, McLoughlin SM, Zhou Z, Lin AJ, Golan DE, et al. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. Proc. Natl. Acad. Sci. U.S.A. 2005 Nov 1;102(44):15815–20.

24. Jin X, Uttamapinant C, Ting AY. Synthesis of 7-Aminocoumarin by Buchwald–Hartwig Cross Coupling for Specific Protein Labeling in Living Cells. ChemBioChem. 2011;12(1):65–70.

25. Cohen JD, Thompson S, Ting AY. Structure-Guided Engineering of a Pacific Blue Fluorophore Ligase for Specific Protein Imaging in Living Cells. Biochemistry. 2011 Sep 27;50(38):8221–5.

26. Fernández-Suárez M, Baruah H, Martínez-Hernández L, Xie KT, Baskin JM, Bertozzi CR, et al. Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. Nature Biotechnology. 2007;25(12):1483–7.

27. Yao JZ, Uttamapinant C, Poloukhtine A, Baskin JM, Codelli JA, Sletten EM, et al. Fluorophore Targeting to Cellular Proteins via Enzyme-Mediated Azide Ligation and Strain-Promoted Cycloaddition. J. Am. Chem. Soc. 2012 Feb 29;134(8):3720–8.

28. Liu DS, Tangpeerachaikul A, Selvaraj R, Taylor MT, Fox JM, Ting AY. Diels–Alder Cycloaddition for Fluorophore Targeting to Specific Proteins inside Living Cells. J. Am. Chem. Soc. 2012 Jan 18;134(2):792–5.

29. Baruah H, Puthenveetil S, Choi Y-A, Shah S, Ting AY. An Engineered Aryl Azide Ligase for Site-Specific Mapping of Protein–Protein Interactions through Photo-Cross-Linking. Angewandte Chemie International Edition. 2008;47(37):7018–21.

30. Puthenveetil S, Liu DS, White KA, Thompson S, Ting AY. Yeast Display Evolution of a Kinetically Efficient 13-Amino Acid Substrate for Lipoic Acid Ligase. J. Am. Chem. Soc. 2009 Nov 18;131(45):16430–8.

31. McNeill E, Chen I, Ting AY. Synthesis of a Ketone Analogue of Biotin via the Intramolecular Pauson−Khand Reaction. Org. Lett. 2006 Sep 1;8(20):4593–5.

32. Slavoff SA, Chen I, Choi Y-A, Ting AY. Expanding the Substrate Tolerance of Biotin Ligase through Exploration of Enzymes from Diverse Species. J. Am. Chem. Soc. 2008 Jan 1;130(4):1160–2.

33. Popp MW, Antos JM, Grotenbreg GM, Spooner E, Ploegh HL. Sortagging: a versatile method for protein labeling. Nat. Chem. Biol. 2007 Nov;3(11):707–8.

34. Popp MW-L, Antos JM, Ploegh HL. Site-specific protein labeling via sortase-mediated transpeptidation. Curr Protoc Protein Sci. 2009 Apr;Chapter 15:Unit 15.3.

35. Sletten EM, Bertozzi CR. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. Angewandte Chemie International Edition. 2009;48(38):6974–98.

36. Saxon E, Bertozzi CR. Cell surface engineering by a modified Staudinger reaction. Science. 2000 Mar 17;287(5460):2007–10.

37. Prescher JA, Dube DH, Bertozzi CR. Chemical remodelling of cell surfaces in living animals. Nature. 2004 Aug 19;430(7002):873–7.

38. Lin FL, Hoyt HM, Van Halbeek H, Bergman RG, Bertozzi CR. Mechanistic investigation of the staudinger ligation. J. Am. Chem. Soc. 2005 Mar 2;127(8):2686–95.

39. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. A Comparative Study of Bioorthogonal Reactions with Azides. ACS Chem. Biol. 2006 Nov 1;1(10):644–8.

40. Huisgen R. Centenary Lecture - 1,3-Dipolar Cycloadditions. Proc. Chem. Soc. London. 1961;(OCT):357–96.

41. Meldal M, Tornøe CW. Cu-catalyzed azide-alkyne cycloaddition. Chem. Rev. 2008 Aug;108(8):2952–3015.

42. Jewett JC, Sletten EM, Bertozzi CR. Rapid Cu-Free Click Chemistry with Readily Synthesized Biarylazacyclooctynones. J Am Chem Soc. 2010 Mar 24;132(11):3688–90.

43. Blackman ML, Royzen M, Fox JM. The Tetrazine Ligation: Fast Bioconjugation based on Inverse-electron-demand Diels-Alder Reactivity. J Am Chem Soc. 2008 Oct 15;130(41):13518–9.

44. Co-immunoprecipitations revisited: an update on experimental concepts and their implementation for sensitive interactome investigations of endogenous proteins - Springer. [cited 2013 Apr 19]; Available from: http://link.springer.com/article/10.1007/s00216-007-1385-x/fulltext.html

45. Phizicky EM, Fields S. Protein-protein interactions: methods for detection and analysis. Microbiol. Rev. 1995 Mar 1;59(1):94–123.

46. Hatanaka Y, Sadakane Y. Photoaffinity Labeling in Drug Discovery and Developments: Chemical Gateway for Entering Proteomic Frontier. Current Topics in Medicinal Chemistry. 2002 Mar 1;2(3):271–88.

47. Liu W, Brock A, Chen S, Chen S, Schultz PG. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. Nat Meth. 2007 Mar;4(3):239–44.

48. Hopkins C, Gibson A, Stinchcombe J, Futter C. Chimeric molecules employing horseradish peroxidase as reporter enzyme for protein localization in the electron microscope. In: Jeremy Thorner SDE and JNA, editor. Methods in Enzymology [Internet]. Academic Press; 2000 [cited 2013 Mar 31]. p. 35–45. Available from: http://www.sciencedirect.com/science/article/pii/S0076687900272650

49. Kotani N, Honke K. [Novel approach for the cell surface molecular interactome using enzyme-mediated activation of radical sources (EMARS) reaction]. Seikagaku. 2011 Aug;83(8):754–8.

50. Kotani N, Gu J, Isaji T, Udaka K, Taniguchi N, Honke K. Biochemical visualization of cell surface molecular clustering in living cells. PNAS [Internet]. 2008 May 21

[cited 2013 Mar 31]; Available from: http://www.pnas.org/content/early/2008/05/20/0710346105

51. Honke K, Kotani N. Identification of Cell-Surface Molecular Interactions under Living Conditions by Using the Enzyme-Mediated Activation of Radical Sources (EMARS) Method. Sensors. 2012 Nov 22;12(12):16037–45.

52. Kwon K, Streaker ED, Ruparelia S, Beckett D. Multiple disordered loops function in corepressor-induced dimerization of the biotin repressor. J. Mol. Biol. 2000 Dec 15;304(5):821–33.

53. Roux KJ, Kim DI, Raida M, Burke B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol. 2012 Mar 19;196(6):801–10.

54. Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA, et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science. 2013 Mar 15;339(6125):1328–31.

Chapter 2. Development of intracellular protein labeling with diverse fluorophores using enzyme-mediated azide ligation and strain-promoted cycloaddition

Part of the work discussed in this chapter has been published in: J. Yao, C. Uttamapinant, A. Poloukhtine, J.M. Baskin, J.A. Codelli, E.M. Sletten, C.R. Bertozzi, V.V. Popik, A.Y. Ting, "Fluorophore targeting to cellular proteins via enzyme-mediated azide ligation and strain-promoted cycloaddition", *J Am Chem Soc* **2012** (134), 3720-3728.

Introduction

In the previous chapter, we described various enzyme-based methods of fluorophore targeting. We seek to develop a fluorophore targeting method that combines high labeling specificity with small noninvasive peptide tag, works for intracellular protein labeling inside live cells, and is generalizable towards a wide array of fluorophores with different photophysical functionalities. The Ting lab has developed enzyme-based fluorophore labeling techniques based on the *Escherichia coli* enzyme lipoic acid ligase (LplA) (1,2). The natural function of LplA in *E. coli* is to catalyze the ATP-dependent ligation of the cofactor lipoic acid onto the lysine residue in one of its three substrates: E2p subunit of pyruvate dehydrogenase complex, E3 subunit of 2-oxoglutarate dehydrogenase complex, and H-protein of glycine cleavage system (Figure 2-1) (1). The mechanism of lipoic acid transfer proceeds in two steps. First, the enzyme activates the carboxylic acid of the lipoic acid cofactor using ATP and releasing pyrophosphate in the process. In the second step, the reactive adenylate ester intermediate, lipoyl-AMP, is attacked by the ε-amino group of a lysine residue in the lipoate-acceptor protein (e.g., E2p) to effect lipoic acid transfer.

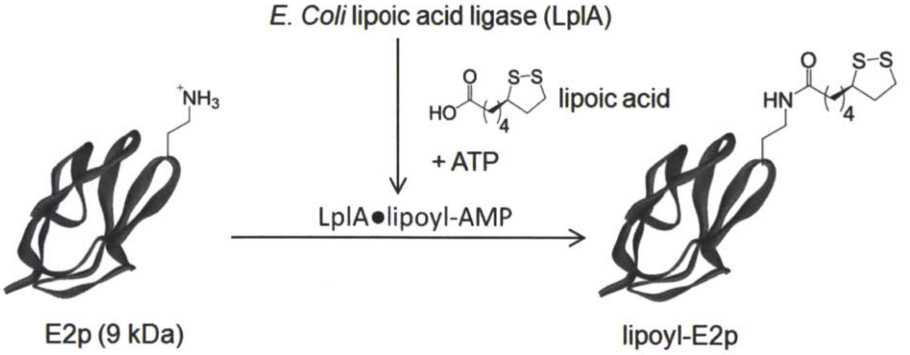


Figure 2-1. Natural ligation of lipoic acid catalyzed by wild-type LplA.

The enzyme has high specificity towards its protein substrates, but exhibits much more promiscuity for its small molecule substrate (2). Over the years, the Ting lab has exploited this small molecule plasticity of LplA and engineered LplA, as well as its protein substrate for the application of protein labeling over several stages. First, in the work of Dr. Sujiet Puthenveetil, he used iterative cycles of rational design (3) and later through yeast display evolution to evolve one of the natural acceptors of LplA, the 86 amino acid (8.9 kDa) lipoyl domain of the full-length E2p subunit of pyruvate dehydrogenase to a small 13-amino acid peptide, which we call LplA acceptor peptide (LAP) (4). The LAP peptide can be genetically fused to any protein of interest and confers small tag size while maintaining high specificity. The k_{cat}/K_m of LplA for the LAP peptide is only eight-fold lower than its k_{cat}/K_m for the H-protein (4). Second, through site-directed mutagenesis of the lipoic acid binding pocket of LplA, our lab has shown that the engineered LplA can incorporate a plethora of unnatural small molecules including 7-hydroxycumarin (5,6), 7-aminocoumarin (7), Pacific Blue (8), aryl azide (9), alkyl azide (3), and *trans*-cyclooctene (10) (Figure 2-2). Using this LplA enzyme–catalyzed probe ligation platform which we call PRIME (PRobe Incorporation Mediated by Enzymes), we have labeled LAP-fused proteins inside live cells and on the cell surface, such as cytoskeletal proteins actin and mitogen-associated protein 2 (MAP2), FK506 binding protein (FKBP), and neuronal adhesion protein neurexin.

The application of PRIME for intracellular fluorophore targeting was first demonstrated by Dr. Tao Uttamapinant and co-workers (5). Through screening of many LplA mutants, LplA(W37V) was identified as a coumarin ligase. While coumarin labeling already offers advantages over other labeling methods in its small fusion tag and high labeling specificity inside mammalian cells, coumarin has excitation and emission maxima of 387 nm and 448 nm, which are not optimal for live cell imaging. The strict restriction to only one class of blue fluorophores limits the applicability of PRIME. We wished to extend PRIME labeling towards other more red-shifted fluorophores; to further increase detection sensitivity compared to coumarin-PRIME via using dyes with better photophysical properties and greater separation from cellular autofluorescence; to reduce photodamage to cells during imaging by using longer-wavelength excitation light; and also to provide additional colors to complement coumarin-PRIME labeling.

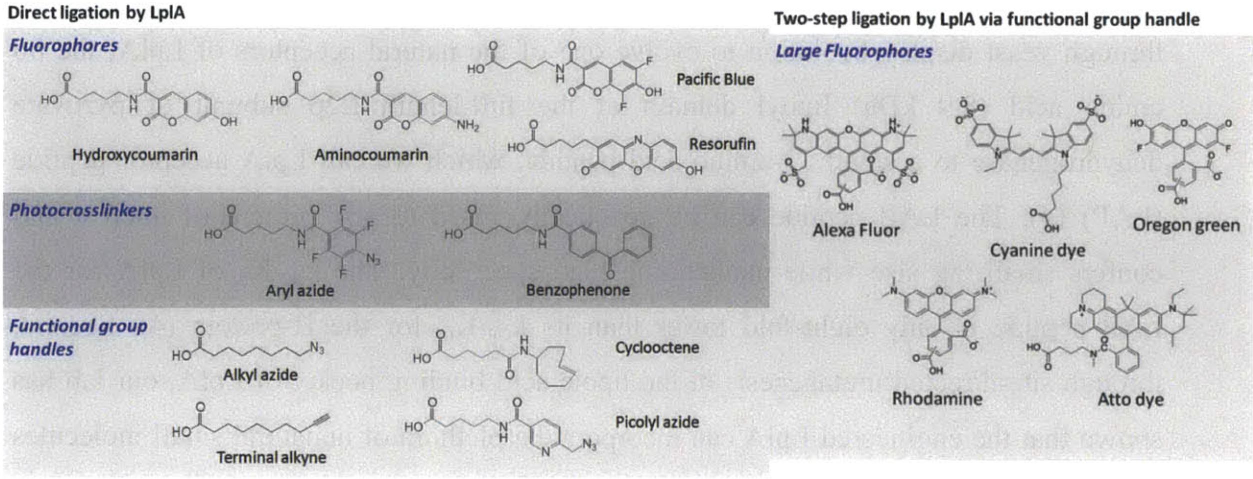

Direct ligation by LplA
Fluorophores
Hydroxycoumarin
Aminocoumarin
Pacific Blue
Resorufin
Photocrosslinkers
Aryl azide
Benzophenone
Functional group handles
Alkyl azide
Cyclooctene
Terminal alkyne
Picolyl azide
Two-step ligation by LplA via functional group handle
Large Fluorophores
Alexa Fluor
Cyanine dye
Oregon green
Rhodamine
Atto dye

Figure 2-2. Small and large probes to be incorporated by LplA. (left) Chemical probes directly ligated by LplA mutants. (right) Large fluorophores that could be ligated by LplA via functional group handle.

We desire a strategy for targeting to intracellular LAP-fused proteins any fluorophore that can cross the cellular membrane and exhibit minimal nonspecific binding to endogenous proteins. Looking at the structures of large fluorophores such as cyanine dyes and ATTO dyes (Figure 2-2), all of them are far larger than coumarin. Although a direct, one-step labeling approach is attractive in that it offers efficient ligation yield and ease of use, engineering a specific LplA to accept each large fluorophore would be very challenging and require extensive protein engineering of the LplA active site. Examining of the crystal structure *E. coli* LplA with lipoate-AMP bound in the active site shows that the dithiolane ring of lipoate-AMP is completely surrounded by active site residues. Even though by mutagenesis of some active site residues we were able to enlarge the lipoate binding pocket to incorporate larger, linear, bi- or tricyclic fluorophore structures such as coumarin (5) and resorufin (Daniel Liu, unpublished), to further mutate the binding pocket to bind to nonlinear, tetracyclic, and charged dyes like fluorescein and tetramethylrhodamine seemed unlikely. Random mutation of active site residues might be fruitless or hurt the folding or specificity of LplA. Instead of one-step fluorophore targeting, we envisioned a two-step ligation process: first using LplA to target a bioorthogonal functional group handle onto LAP, then chemoselectively derivatizing the functional group handle with any fluorophore of interest (Figure 2-3). In the first phase, we screened LplA mutants against a panel of alkyl azides to search for an efficient azide ligase. Second, we compared the efficiency of azide ligation inside live cells. Third, we examined five different cyclooctynes in their abilities to react with azide inside the complex environment of live cells. Finally, we demonstrated using two-step labeling PRIME to label a variety of LAP fusion proteins inside live cells with various fluorophores whose emission wavelengths span the visible spectrum.

In this chapter, we describe our initial efforts to extend PRIME labeling technology to be generalizable towards diverse fluorophores structures. We screened a panel of LplA mutants towards the ligation of an alkyl azide small molecule. We characterized this new LplA and azide pair in terms of its ligation kinetics, live cell ligation yield, and ligation specificity. We also developed a method using a nuclear localized azido-sugar to assay for the efficiency and specificity of the strain-promoted azide-alkyne cycloaddition reaction inside live cells.

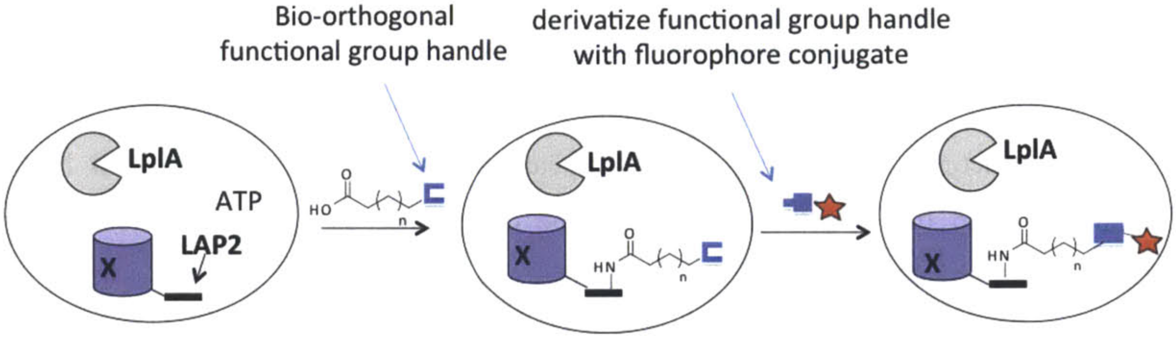

Bio-orthogonal
functional group handle
derivatize functional group handle
with fluorophore conjugate
LplA
ATP
X
LAP2
HO
O
n
LplA
X
HN
O
n
LplA
X
HN
O
n

Figure 2-3. Two-step LplA ligation using functional group handle and chemoselective bioorthogonal reaction inside live cells.

Screening for an efficient LplA ligase and azide pair for site-specific protein labeling inside live cells

The first step to generalize two-step PRIME towards targeting of diverse fluorophore structures was to engineer LplA to ligate a small molecule bioorthogonal functional handle specifically to LAP fusion proteins inside live cells. The functional handle can then be chemoselectively derivatized with fluorophore of interest via a bioorthogonal reaction. At the time of the start of the project, in the fast growing area of potentially bioorthogonal reactions, very few have been tested rigorously in a complex biological context, let alone inside mammalian cells. Among the top candidates are the Staudinger ligation (11–14) and the strain-promoted azide-alkyne cycloaddition (15–17). Despite its impressive biological applications on live cell surface (18) and in mice (13,14), the Staudinger ligation has long been plagued by phosphine oxidation, aza-ylide hydrolysis, and slow reaction rate (15,19). The strain-promoted azide-alkyne cycloaddition is the most promising, especially given its previous demonstrations on the cell surface (3).

Previously, in Dr. Marta Fernández-Suárez's work (3), two-step fluorophore targeting via LplA azide ligation was accomplished on cell-surface proteins LDLR and EGFR, which were fused to the 22-amino acid peptide called LAP1 that was derived through structure comparison and site-directed mutagenesis from the 9 kDa lipoyl domain of the full length E2p protein. The cell-surface LAP1-fused proteins were exogenously ligated with azide using purified wild-type LplA protein. The azide functional group on the cell surface was then chemoselectively derivatized through the strain-promoted cycloaddition reaction with cyclooctyne that was conjugated to a cyanine dye. Dr. Marta Fernández-Suárez screened a panel of alkyl azides with different alkyl chain lengths against wild-type LplA in an HPLC assay. Through comparison of *in vitro* labeling kinetics, 8-azido-octanoic acid, or azide 7, was most efficiently incorporated by wild-type LplA onto the E2p substrate with a k_{cat} of 6.66 min^{-1} and K_m of 127 μM (3). The K_m value was 75-fold increased, while the k_{cat} value was 2-fold decreased compared to the analogous reaction performed with lipoic acid.

While wild-type LplA and azide 7 worked well for live-cell cell surface labeling, several problems appeared when we tried to apply the ligation inside live cells. First, after azide ligation inside live HEK 293T cells expressing wild-type LplA and the 22-mer LAP1, we fixed the cells and labeled them with alkyne functionalized cyanine dye through CuAAC to detect ligated azides. Even though strong signal was observed only in cells expressing both LplA and LAP1, negative control reactions with cells expressing the Lys→Ala mutant of LAP1 or only the LplA enzyme also showed a transfection-specific signal. Second, when labeling LAP1 localized to only the nucleus, background in the cytosol specific to transfected cells was also observed. Third, the washout of the probe, monofluorinated cyclooctyne conjugated to fluorescein, was very poor or never washed out if the probe was loaded more than 50 μM. Furthermore, labeling was inconsistent among several cell lines tested, HEK 293T, HeLa, and COS-7 cells, with the strongest signal in HEK 293T cells, and very weak or no signal in COS-7 cells. There are several possible explanations for the observed intracellular labeling problems. If unligated azides were not washed out completely and stick nonspecifically to cellular organelles and hydrophobic surfaces, they would also be free to react with cyclooctyne-fluorophore in the second step to give rise to unwanted background. This was problematic not only because it would give high background, unligated azides can also compete with specifically ligated azides to react with cyclooctyne, reducing specific signal. With respect to the monofluorinated cyclooctyne-fluorophore, the difficulty with probe washout could be caused by its own hydrophobic nature, nonspecific reaction with excess azides inside cells, or reaction with cellular nucleophiles such as thiols. Although the monofluorinated cyclooctyne has not been experimentally proven to react with thiols, its derivative, the difluorinated cyclooctyne DIFO, showed evidence of reacting with Cys and glutathione *in vitro* at the strained alkyne (20). Chang *et al* also reports DIFO sticks heavily to mouse serum albumin likely because of a combination of hydrophobic and covalent interactions (21). Therefore, the sequential nature of the protocol and various unknown factors inside the cell caused the analysis and troubleshooting of two-step labeling results especially complicated, making it difficult to catch the culprits that were causing problems like nonspecific signal and high background.

Furthermore, since the development of the LAP1 peptide and the application of it to cell-surface protein labeling (3), Dr. Sujiet Puthenveetil engineered an even smaller 13-mer peptide substrate called LAP2 (or simply LAP in this thesis) through yeast display evolution (4). The LAP peptide has a k_{cat}/K_m of 0.99 $\mu M^{-1}min^{-1}$ for wild-type LplA catalyzed lipoic acid ligation reaction, about 70-fold better than the same reaction performed previously on the 22-mer LAP1. In light of the new LAP, as well as new gained insights regarding LplA engineering and mutagenesis, we wished to screen for an improve LplA ligase and azide substrate pair to use for fluorophore labeling inside live mammalian cells.

To extend the two-step fluorophore labeling scheme inside cells, we decided to optimize each step independently to obtain more clear results without the confounding factors associated with the other step. In the first step of azide ligation, we needed the LplA enzyme to specifically and efficiently ligate the alkyl azide small molecule onto LAP-fused protein substrates with high ligation yield in the presence of other intracellular proteins. We wished to avoid azide-dependent background possibly caused by excess unligated azides remaining inside cells. Even though wild-type LplA can catalyze the conjugation of azide 7 onto LAP1-fused proteins expressed on the cell surface, where the exogenous azide probe can be loaded at high concentrations then washed away cleanly without worrying of sticking to cellular organelles, for intracellular azide ligation, we preferred to deliver the azide probe at lower concentrations so that less residual azide remains after ligation reaction, and to minimize interference with the subsequent cycloaddition reaction with cyclooctyne. We would also like the new LplA enzyme and azide pair to have an improved k_{cat} of LplA azide ligation to achieve efficient ligation yield within a short labeling time. Therefore, to use lower azide concentrations without sacrificing azide ligation yield, we needed to engineer the LplA-catalyzed azide ligation reaction to improve its catalytic efficiency.

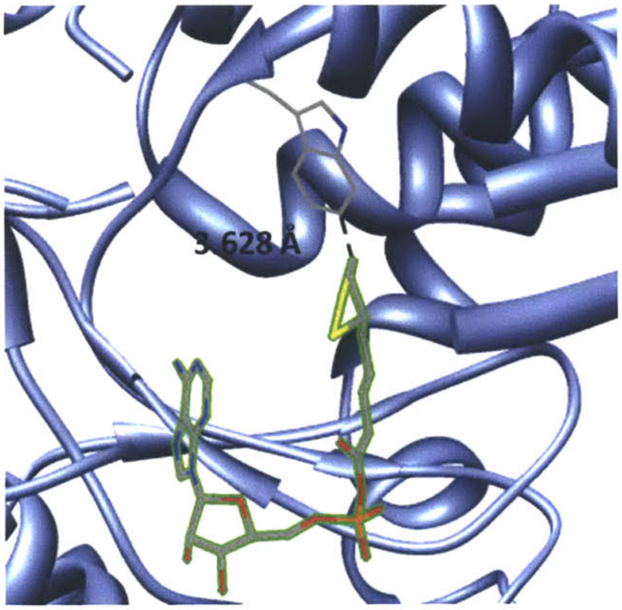


Figure 2-4. Small molecule substrate binding pocket of *E. coli* lipoic acid ligase (LplA) with bound lipoyl-AMP (PDB 3A7R). Distance between the dithiolane ring of lipoyl-AMP and the "gatekeeper" residue W37 of LplA is 3.628 Å.

To screen for a new LplA azide ligase, we consulted the previous azide ligation screen by wild-type LplA completed by Dr. Marta Fernández-Suárez as well as examined the lipoic acid binding pocket of LplA (Figure 2-4). In Dr. Fernández-Suárez's work, she observed that 8-azido-octanoic acid (azide 7) with 7 methylenes as the alkyl chain length was the best substrate for wild-type LplA, while other azides with either shorter or longer alkyl chains did not get ligated as well (3). This data indicated that perhaps alkyl azides longer than azide 7 were too larger to fit in the LplA binding pocket; while alkyl azides shorter than azide 7 were probably too small and could not form enough hydrophobic interactions with the active site residues of wild-type LplA to enable ideal binding and ligation. In addition, previous lab member Dr. Hemanta Baruah laid the groundwork in identifying amino acid residues in the LplA binding pocket that are important in modulating the enzyme's plasticity towards its small molecule substrates. Through alanine screening mutagenesis of residues that are within 7.5 Å of the dithiolane ring of lipoic acid, Dr. Baruah found W37 of lipoic acid ligase to be a "gatekeeper" residue that when mutated to smaller residues like V and S, enabled the LplA mutant to accept a larger aryl azide substrate (9). Having the W37 residue as a lead, Dr. Chayasith Uttamapinant and co-workers later identified W37V mutant of LplA to be an efficient coumarin ligase (5). To find an alkyl azide ligase, we engineered the binding pocket of LplA by mutating W37 residue to six other smaller and mainly hydrophobic amino acids – G, A, V, I, L, S – and screened the mutants against a panel of alkyl azide substrates of increasing alkyl chain lengths, with the shortest azide being azide 7 (Figure 2-5).

Alkyl azide small molecules were synthesized using the corresponding bromoalkanoic acid and sodium azide as described in Experimental Methods. LplA mutant proteins were expressed and purified from *E. coli.* We set up *in vitro* ligation reactions using four different alkyl azide small molecules, six purified LplA proteins, and the 13-amino acid LAP peptide as the substrate. An HPLC assay was used to determine the percent conversion of LAP into LAP-azide conjugate under two different reaction conditions: 500 μM probe for 10 minutes, and 20 μM probe for 20 minutes. We found that for the shortest azide 7 probe, wild-type LplA and LplA(W37V) were the top two ligases (Figure 2-5). As the length of the alkyl chain increases, wild-type LplA with the large W37 "gatekeeper" residue was no longer active, instead LplA(W37V) and LplA(W37I) mutants were the best. The top four best LplA ligase/alkyl azide probe pairs are starred in Figure 2-5.

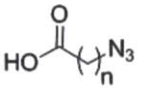
Probe		**% Conversion to product with LplA variants**						
		WT	**W37G**	**W37A**	**W37V**	**W37I**	**W37L**	**W37S**
	n=7 Azide 7	* 37%	<3%	<3%	* 44%	15%	<3%	N.D.
	n=8 Azide 8	7%	N.D.	N.D.	11%	6%	<3%	N.D.
	n=9 Azide 9	N.D.	6%	9%	* 100%	* 59%	<3%	N.D.
	n=10 Azide 10	N.D.	30%	12%	42%	10%	N.D.	N.D.

Figure 2-5. Screening to identify the best LplA mutant/azide substrate pair. The table shows percent conversions of LAP to the LAP−azide product conjugate. Wild-type LplA and six W37 point mutants were screened against four azidoalkanoic acid substrates of various lengths. N.D. indicates that product was not detected. Screening was performed with 100 nM ligase, 600 μM LAP, and 20 μM azide substrate for 20 min at 30 °C. Conversions were measured in duplicate. Note that LplA(W37S) was active with the natural substrate, lipoic acid (data not shown). The starred entries in the table were further evaluated in Figure 3-2.

Characterization of an improved LplA enzyme and azide probe pair

Kinetics properties of LplA(W37I) ligation of azide 9 onto LAP peptide

Even though from our *in vitro* HPLC assay LplA(W37V) and azide 9 pair gave the highest azide ligation yield, we later discovered through a live-cell fluorophore targeting imaging assay that LplA(W37I) and azide 9 actually performed better inside live cells (described later in Chapter 3; Figure 3-2). Therefore, we picked LplA(W37I) and azide 9 as our best enzyme/azide pair and proceeded to compare the kinetic properties of LplA(W37I) and azide 9 pair to the previous wild-type LplA and azide 7 pair used in cell-surface protein labeling. Figure 2-6 A shows an HPLC analysis of LplA(W37I)-catalyzed ligation of azide 9 onto LAP peptide. The identity of the LAP-azide 9 product peak was confirmed by mass spectrometry (Figure 2-6 B). Negative control reactions with ATP omitted or wild-type LplA in place of LplA(W37I) were also analyzed and showed no product formation. We also used HPLC to quantify product amounts in order to measure k_{cat} and K_m values. Figure 2-6 C shows the Michaelis-Menten plot giving a k_{cat} of 3.62 min^{-1} and a K_m of 35 μM for azide 9 ligation catalyzed by LplA(W37I). Compared to our previously reported azide 7 ligation catalyzed by wild-type LplA on E2p, this improved K_m is 4-fold lower. The k_{cat} is 1.8-fold reduced, giving an overall improvement in k_{cat}/K_m of 2-fold.

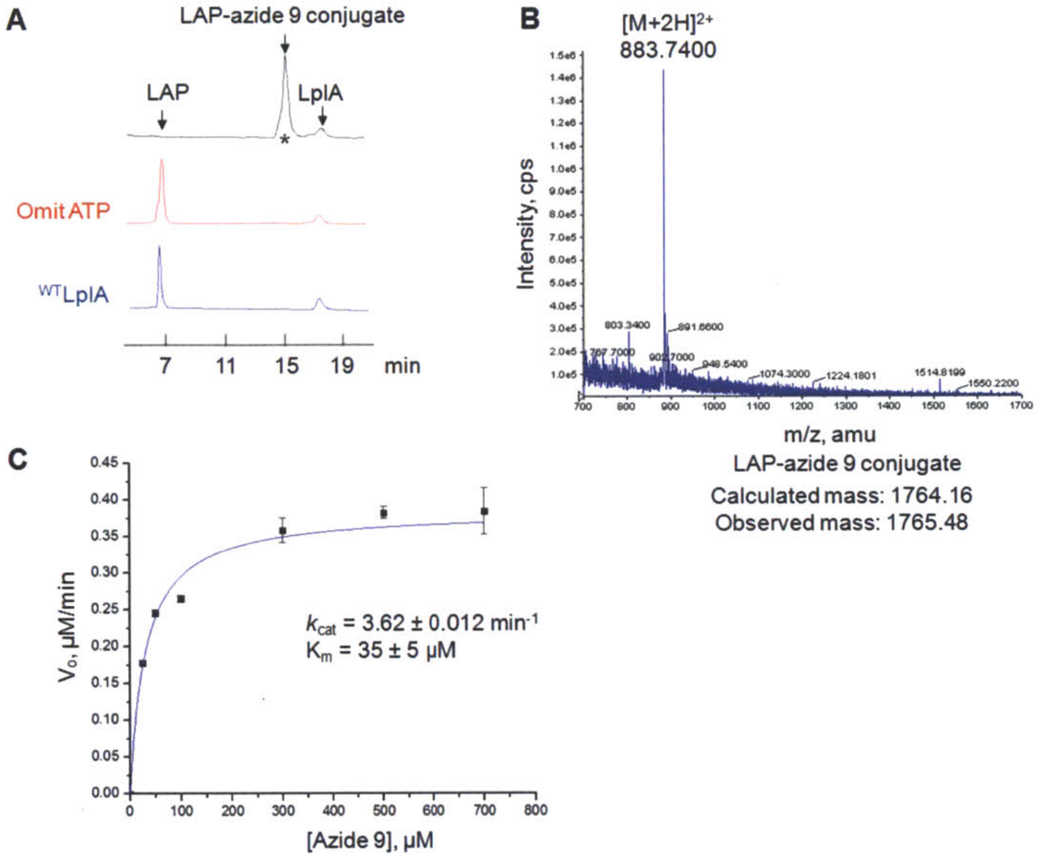


Figure 2-6. *In vitro* characterization of azide 9 ligation catalyzed by LplA(W37I). **(A)** HPLC trace showing LAP conversion to LAP-azide 9 conjugate, catalyzed by LplA(W37I) (black trace). Negative controls are shown with ATP omitted (red trace) and LplA(W37I) replaced by wild-type LplA (blue trace). Reactions were performed for 2 hours with 1 µM LplA, 300 µM LAP, and 500 µM azide 9. **(B)** ESI mass spectrometric analysis of LAP-azide 9 covalent adduct (starred product peak in (A)). **(C)** Michaelis-Menten curve and kinetic parameters for azide 9 ligation onto LAP, catalyzed by 100 nM LplA(W37I). Initial rates (V_o) were measured using the HPLC assay in (A) with 600 µM LAP and 25-700 µM azide 9. Each V_o value was measured in triplicate. Error bars, ±1 s.d.

CuAAC detection of azide ligation in fixed cells

After finding a kinetically improved LplA enzyme and azide pair, we evaluated the specificity of azide ligation using CuAAC as detection method. In this assay, live-cell azide ligation is followed by fixing and permeabilizing the cells for reaction with terminal alkyne conjugated fluorophore via CuAAC. This "live-fix" protocol provided us with several advantages that were helpful in evaluating live-cell azide ligation. Performing enzymatic azide ligation in live cells allowed us to critically exam the ligation specificity and efficiency in its natural environment. Having fixed and permeabilized the cells after azide labeling, we washed out all excess azides, except LAP-ligated azides as well as any unspecific ligation or other background reaction between azide and cellular molecules. Finally, detection with terminal alkyne-fluorophore conjugate via CuAAC has high sensitivity that enabled us to uncover any background reaction that could potentially interfere with the strain-promoted azide-alkyne cycloaddition step.

To perform azide 9 ligation inside live cells, HEK 293T cells were transfected with LplA(W37I) and nuclear-targeted, LAP-tagged YFP. After incubating the cells with 200 μM azide 9 for 1 hour at 37C, unligated azide 9 was washed out of cells by incubating cells in fresh growth media for 1 hour. Following live cells azide 9 ligation and washout, cells were fixed using paraformaldehyde, permeabilized using cold methanol, blocked with BSA, and reacted with cy5-alkyne conjugate via CuAAC. Figure 2-7 shows cy5 signal co-localizing with LAP-YFP nuclear expression, indicating that LplA(W37I) azide 9 ligation is specific to LAP.

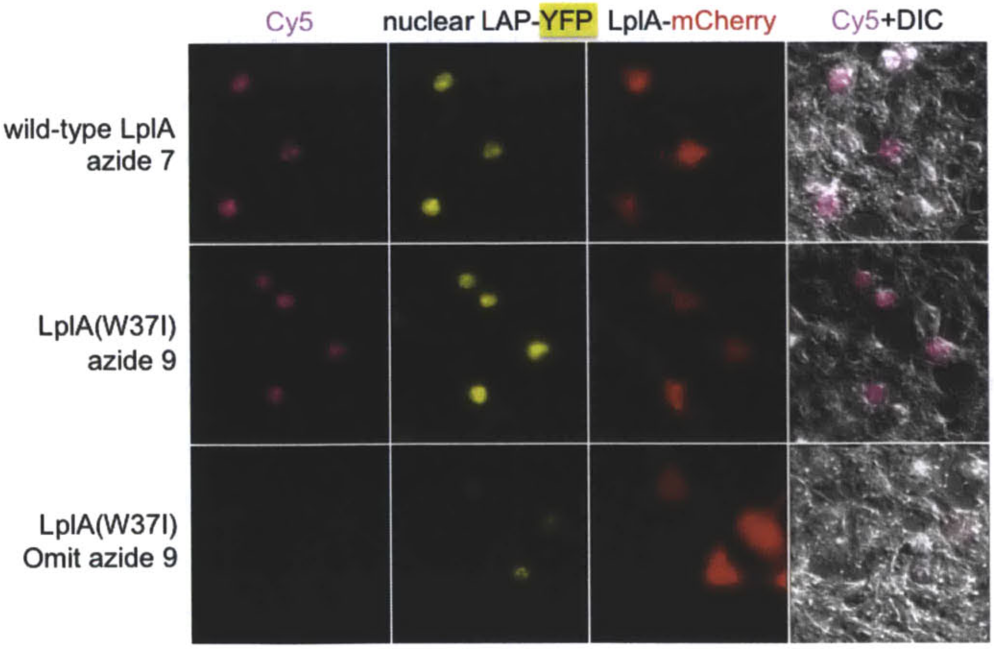


Figure 2-7. Fixed cell labeling on nuclear localized LAP-YFP with cy5-alkyne via CuAAC. HEK 293T cells expressing nuclear LAP-YFP with either wild-type LplA or LplA(W37I) were labeled with azide 7 or azide 9, respectively. After azide washout, cells were fixed and permeabilized. Fixed cells were reacted with cy5-alkyne overnight, washed to remove excess cy5-alkyne and imaged on the confocal microscope.

Estimating azide ligation yield in live cells

To have high labeling signal in the two-step fluorophore targeting scheme, we needed the azide ligation step and the strain-promoted cycloaddition both to achieve quantitative yield. A high azide ligation yield is especially instrumental to achieving this goal. We have shown that the new LplA(W37I) and azide 9 pair has superior kinetic properties compare to wild-type LplA and azide 7 pair. We wanted to see if the *in vitro* result were transferred to better ligation efficiency inside live cells. Two methods were developed to evaluate the yield of live-cell azide ligation. Both methods are based on the shift in migrating position of a LAP-tagged FP protein in an acrylamide gel. The first method is based on the difference in charge between free and azide-ligated LAP-protein. Ali and Guest first observed that lipoic acid modified E2p moves faster than the apo E2p in a non-denaturing native gel owing to the loss of the positively charged lysine (1). We applied the same assay to azide-ligated LAP-protein and free LAP-protein. The loss of the positive charge on the catalytic lysine residue in the azide-ligated LAP-protein causes it to run faster comparing to unligated LAP-protein in native PAGE gel. Hence, we call the method native gel-shift assay. Since azide-ligated LAP-FP protein effectively separates itself from unmodified LAP-FP protein by its faster migration on a native gel, the yield of azide ligation onto LAP-protein can be determine by the ratio of densitometry of each protein band. The second method is based on the selective binding of a streptavidin molecule onto azide ligated LAP-FP protein, causing it to migrate much slower because of the increased molecular weight of 56 kDa than free LAP-FP protein on an SDS-PAGE gel. We named the second method streptavidin gel-shift assay.

- Native Gel-shift Assay

To compare the live-cell azide ligation efficiencies of the newly improved LplA(W37I) and azide 9 pair and the wild-type LplA and azide 7 pair using native gel-shift assay, we incubated HEK 293T cells expressing LAP-YFP and either wild-type LplA or LplA(W37I) with azide 7 or azide 9 for 30 minutes or 1 hour. Following azide incubation, cells were washed by incubating in growth media to remove excess unligated azides. After excess azide washout, we lysed the cells in non-detergent containing lysis buffer and the lysates were ran on a native PAGE gel. LAP-YFP was detected by YFP in-gel fluorescence. The results shown in Figure 2-8 demonstrated that the unmodified LAP-YFP protein (black arrow) runs at an apparent molecular weight of approximately 42 kDa (Figure 2-8, left, lane 2; right, lane 4). Upon azide modification, we observed that the entire LAP-YFP-azide complex migrated faster on the native PAGE gel with the positively charged lysine residue of LAP being converted to a neutral amide (Figure 2-8, left, lane 1; right, lane 1-3). The apparently molecular weight of the azide-ligated LAP-YFP (red arrow) is about 38 kDa. We further confirmed that the migration pattern is indeed caused by the presence or absence of the positive charge on the lysine residue by performing an additional negative control reaction with LAP (K→A)-YFP where the lysine residue is mutated to an alanine. LAP (K→A)-YFP also migrated to around 38 kDa (Figure 2-8, left, lane 3), same as azide-ligated LAP-YFP. The improved LplA(W37I) and azide 9 pair gave nearly quantitative ligation after only 30 minutes of azide 9 incubation (Figure 2-8, right, lane 2), while wild-type LplA and azide 7 pair only gave 73% ligation yield after 1 hour of labeling in live cells, estimated by densitometry (Figure 2-8, right, lane 1). We were glad to see that LplA(W37I) and azide 9 can achieve quantitative LAP ligation yield with short azide incubation time.

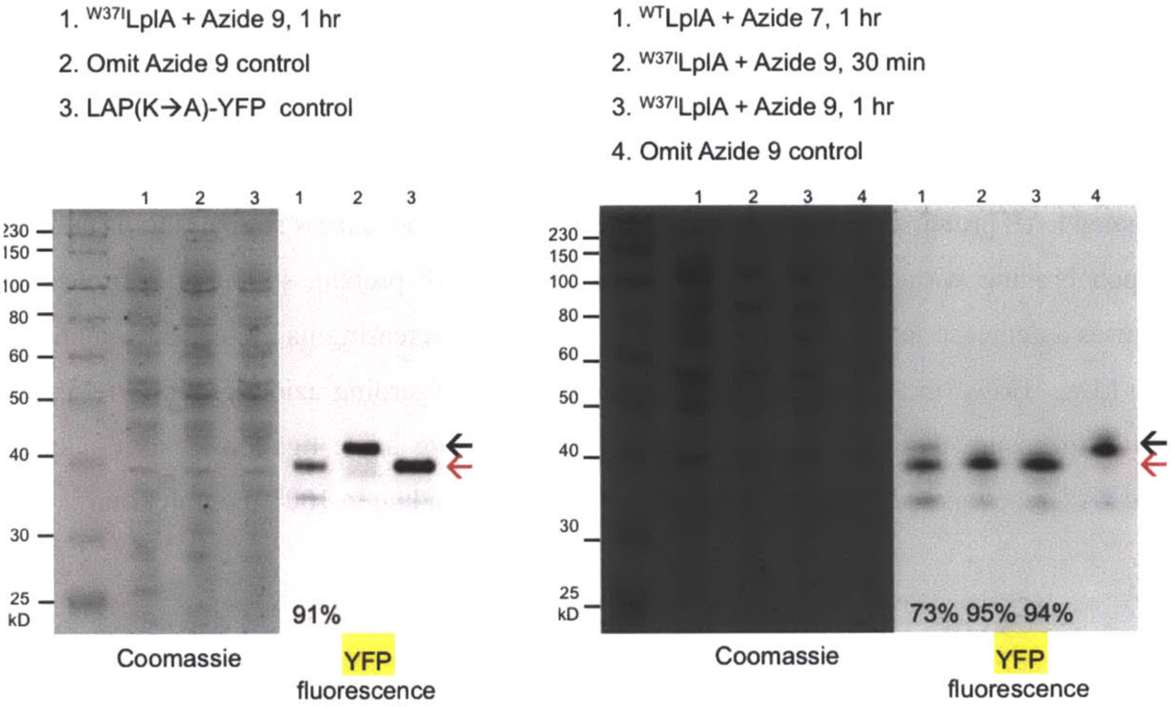

1. W37ILplA + Azide 9, 1 hr
2. Omit Azide 9 control
3. LAP(K→A)-YFP control
1 2 3 1 2 3
230
150
100
80
60
50
40
30
25
kD
91%
Coomassie
YFP fluorescence
1. WTLplA + Azide 7, 1 hr
2. W37ILplA + Azide 9, 30 min
3. W37ILplA + Azide 9, 1 hr
4. Omit Azide 9 control
1 2 3 4 1 2 3 4
230
150
100
80
60
50
40
30
25
kD
73% 95% 94%
Coomassie
YFP fluorescence

Figure 2-8. Characterization of intracellular azide ligation yields by native gel-shift assay. Labeling protocol: HEK 293T cells co-expressing LplA and LAP-tagged YFP were labeled with azide 7 or azide 9 for 30 minutes or 1 hour, washed for 1 hour. Cells were lysed and run on a native 12% polyacrylamide gel. Labeled LAP-YFP runs faster than unlabeled LAP−YFP due to removal of a positive charge. (Left) Gel-shift analysis of azide ligation in cells. Lane 1 is the same as lane 3 on the right. Lane 2 shows a negative control with azide 9 omitted. Lane 3 shows a negative control with the lysine in LAP mutated to alanine. This mutation causes LAP-YFP to shift downward in the 12% polyacrylamide native gel because it removes one positive charge from LAP-YFP. Similarly, labeling of the lysine of LAP causes a downward shift because the positive charge of lysine is converted to a neutral amide. (Right) HEK 293T cells were prepared and labeled with either azide 7 or azide 9 as described previously. Percent conversions to azide-LAP-YFP product are given at the bottom of the YFP fluorescence gel image. Lane 4 shows a negative control with azide 9 omitted. Estimated percent conversions to product are given at the bottom of the YFP fluorescence gel image.

-Streptavidin gel-shift assay

The second approach to estimate the yield of azide ligation on LAP is based on the attachment of a streptavidin molecule selectively onto azide-derivatized population of LAP. As depicted in Figure 2-9, after live-cell azide ligation, cells are lysed and lysate is derivatized with biotin-alkyne via CuAAC. Biotin-alkyne reacts specifically with azide-ligated LAP proteins. After removal of excess biotin-alkyne, excess streptavidin is added. Upon binding with biotin-alkyne derivatized azide-LAP protein, streptavidin molecule causes a decrease in motility of azide-LAP protein by increasing its molecular weight by 56 kDa. However, in order to make any conclusion regarding azide ligation yield, we have to ensure that the two subsequent steps following azide ligation - biotin-alkyne CuAAC reaction and streptavidin binding - are all proceeding to 100% completion.

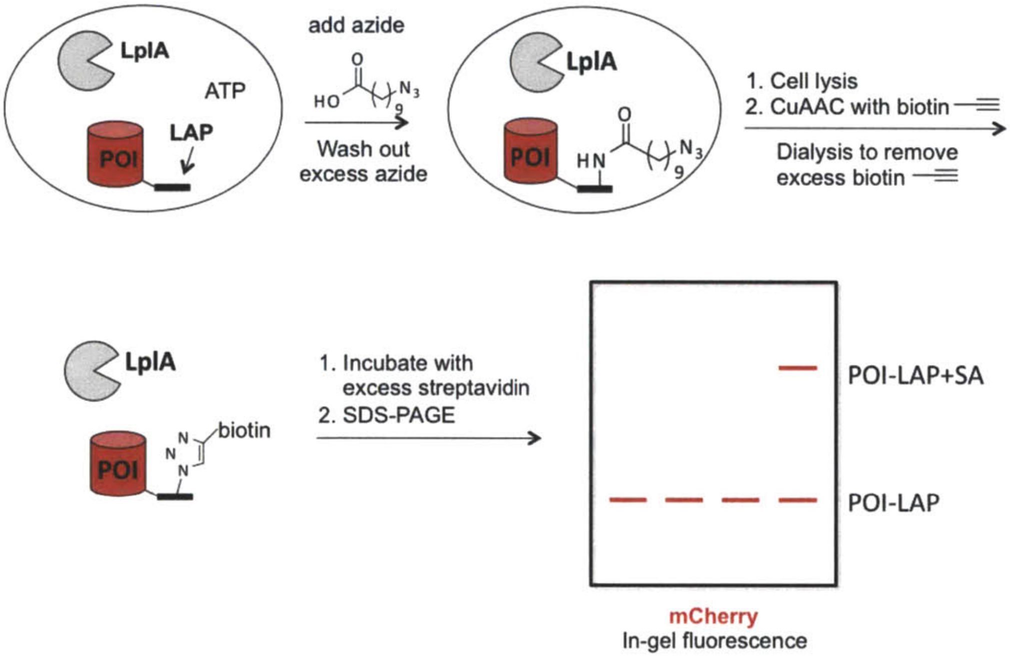

LplA
ATP
LAP
POI
add azide
HO
N3
9
Wash out
excess azide
LplA
HN
N3
9
POI
1. Cell lysis
2. CuAAC with biotin
Dialysis to remove
excess biotin
LplA
POI
N
N
N
biotin
1. Incubate with
excess streptavidin
2. SDS-PAGE
POI-LAP+SA
POI-LAP
mCherry
In-gel fluorescence

Figure 2-9. Schematic representation of streptavidin gel-shift assay to measure azide ligation yield.

HEK 293T cells expressing LAP-mCherry and LplA(W37I) were incubated with azide 9. Following excess azide washout, cells were lysed and the lysate was treated with biotin-alkyne, $CuSO_4$, sodium ascorbate and the tris-triazole ligand for CuAAC - tris-(benzyltriazolylmethyl)amine (TBTA). Under CuAAC, biotin-alkyne would react with azide-ligated LAP-mCherry proteins. After 16 hours incubation with biotin-alkyne, cell lysates were dialyzed to remove excess biotin-alkyne reagent. Lysates were then incubated with excess streptavidin for 1 hour at room temperature. Lysates were separated and analyzed on SDS-PAGE gel. LAP-mCherry was detected by mCherry in-gel fluorescence. Figure 2-10 shows unligated LAP-mCherry appearing to streak on gel but nonetheless migrated to about 30 kDa. We checked in a separate experiment that the streakiness of the bands was caused by copper in the CuAAC reaction (data not shown). The streptavidin-bound azide-ligated LAP-mCherry appeared at around 80 kDa. Negative control reactions with either azide 9 omitted, biotin-alkyne omitted, or streptavidin omitted did not show increase in molecular weight to LAP-mCherry. The azide ligation yield is high, leaving only small amount of LAP-mCherry showing weak fluorescence at 30 kDa (Figure 2-10, lane 3).

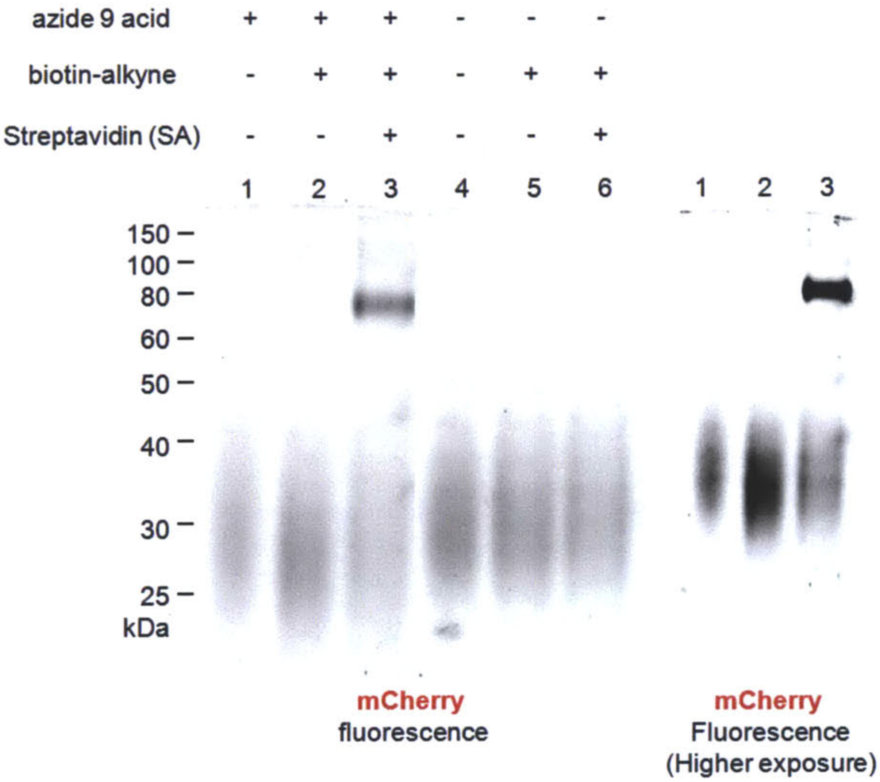


Figure 2-10. Characterization of intracellular azide ligation yields by streptavidin gel-shift assay. Labeling protocol: HEK 293T cells co-expressing LAP-tagged mCherry wild-type LplA or LplA(W37I) and were labeled with azide 9 for 1 hour, washed for 1 hour to remove excess azide. Cells were lysed using hypotonic lysis. Click reaction was

performed in cell lysate with conditions as follows: Biotin-alkyne 100 µM, $CuSO_4$ 1mM, tris-(benzyltriazolylmethyl)amine (TBTA) ligand 100 µM, 2.5 mM sodium ascorbate in 1mM HEPES buffer. Reactions were incubated at RT for 16 hours. The lysate samples were then dialyzed in 1X PBS buffer for 4 hours to get rid of excess biotin-alkyne, and incubated with excess streptavidin (C_f =10uM) for 1 hour at RT. Lysates were loaded onto12% SDS-PAGE gel without boiling samples in order to preserve mCherry fluorescence.

Exploring strain-promoted azide-alkyne cycloaddition reaction inside live cells using alternative source of azide

To proceed systematically to develop the two-step fluorophore labeling technique, we wished to decouple the intracellular strain-promoted cycloaddition second step from the azide ligation step in order to have a clear assay of the cycloaddition specificity and yield. We have to analyze the specificity and efficiency of the cycloaddition step independent of the possible background and non-specific reaction from the azide ligation step. In order to achieve this, we needed an isolated population of azide inside live cells that was not from LplA azide ligation. We obtained peracetylated azido-GalNAz to use as nuclear localized pools of azides inside live cells.

In this approach to independently study the strain-promoted cycloaddition in live cells, we used an azide-derivatized sugar analog, peracetylated *N*-azidoacetylgalactosamine, also called azido-GalNAc or GalNAz as a source of nuclear localized azide. We acquired azido-GalNAc from the Bertozzi lab at UC Berkeley, who was using the sugar analog to study proteins bearing O-linked glycans (14,22–24). Azido-GalNAc is an azide-derivative of GalNAc, whose addition and removal are controlled by two enzymes glycosyltransferase (OGT) and O-GlcNAcase (OGA), respectively (25). GalNAc glycosylation is the most prevalent form of O-linked glycosylation, which puts the sugar α-linked to Ser or Thr residue. There are two main pathways of GalNAc incorporation. In the first pathway, GalNAc-1-kinase and UDP-GalNAc pyrophosphorylase convert GalNAc to UDP-GalNAc, which is then recognized by the enzyme polypeptide N-acetyl-α-galactosaminyltransferase (ppGalNAcT) to put on the hydroxyl group of Ser or Thr (22,25). The second avenue of GalNAc addition is through the GlcNAc salvage pathway, which converts GlcNAc to UDP-GalNAc using an epimerase (22,25). O-GlcNAc glycosylation has been found in numerous proteins involved in transcription, protein trafficking, cell survival, and neurodegenerative diseases (25). Since there is no defined sequence consensus for the ppGalNAcT, predictions of O-linked glycosylated sites have been difficult. However, studies have shown that O-GlcNAcylation, unlike traditional glycosylation, is nearly exclusively on cytoplasmic and nuclear proteins (26). It is possible that GalNAz could be converted to GlcNAc by the epimerase and end up on mainly cytosolic and nuclear proteins. The

Bertozzi group observed that metabolic incorporation of GalNAz followed by detection using Staudinger ligation with phosphine showed predominant presence of GalNAz labeled proteins in the nucleus as well as weaker signal from the cytosol (personal communication). Having evidence that GalNAz gets incorporated onto nuclear proteins, we decided to try it as a nuclear-specific, non-enzymatic source of azide to test our strain-promoted azide-alkyne cycloaddition reaction in live cells.

We chose CHO cells for this experiment because of its higher rate of metabolic incorporation compare to other cell lines. CHO cells were plated on glass coverslips coated with GalNAz or no GalNAz as a negative control. After undergoing metabolic incorporation for three days, the cells were fixed, permeabilized, and reacted with cy3-alkyne via CuAAC. We can observe in Figure 2-11 that cells incorporated with azido-GalNAc gave strong nuclear cy3 signal. Both 100 μM (data not shown) and 200 μM azido-GalNAc incubation gave similar level of signal. The same experiment was also demonstrated using HeLa cells to show that azido-GalNAc incorporation followed by CuAAC detection is generalizable to other cells lines.

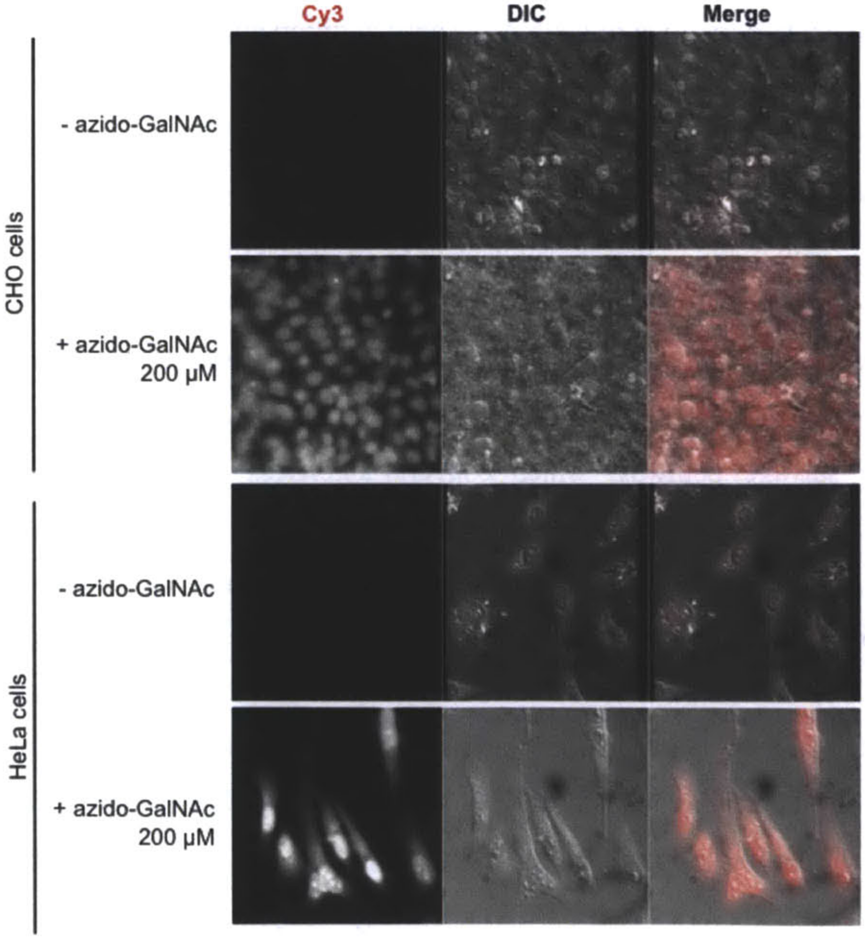


Figure 2-11. Fixed cell detection of metabolically incorporated azido sugar analog azido-GalNAc with cy3-alkyne via CuAAC. Metabolic incorporation of azido-GalNAc was performed in both CHO (top) and HeLa (bottom) cells. Cells were incubated with azido-GalNAc for three days followed by fixation and reaction with cy3-alkyne. Strong cy3 signal was observed in the nucleus as well as weaker signal throughout the cytosol because azido-GalNAc also weakly labels O-linked glycoproteins in the cytosol. Negative control with no azido-GalNAc added to cells shows no cy3 retention.

The success of live cell azido-GalNAc incorporation and fixed cell cy3-alkyne labeling via CuAAC gave us confidence that cells can successfully take up the azido sugar analog to a detectable level and the majority of the signal localized to the nucleus. We tried live cell strain-promoted reaction with a cyclooctyne-fluorophore conjugate after azido-GalNAc metabolic incorporation (Figure 2-12, top scheme). After azido-GalNAc incubation, we replaced cell culture media with media containing a cyclooctyne derivatized fluorophore, monofluorinated cyclooctyne-carboxyfluorescein diacetate (MOFO-CFDA, Figure 2-12). At physiological pH, fluorescein interchanges between the monoanion and the dianion form, rendering the probe membrane impermeant. We protected the phenolic oxygens in fluorescein with acetate groups to make MOFO-fluorescein diacetate (MOFO-CFDA), which is nonfluorescent and can readily across the cell membrane. Once inside the cells, the acetate esters can be cleaved by endogenous cellular esterases to restore fluorescence. Gradually, the excess MOFO-fluorescein is removed from cells by the action of nonspecific anion transporters. We washed out the unreacted probe by replacing MOFO-CFDA containing media with fresh media three times and further incubated the cells in fresh media for an hour to rely on the cell's nonspecific anionic transporters to pump out excess probe.

Figure 2-12 shows MOFO-CFDA nuclear signal intensity and distribution at various time points after dye incubation. Clear nuclear signal can be seen in most cells, but signal distribution has a large variation from cell to cell, perhaps due to the different levels of metabolic incorporation of azido-GalNAc. Metabolic incorporation of azido-GalNAc followed by detection using cyclooctyne-fluorescein conjugate demonstrated that strain-promoted cyclooctyne is able to react specifically with a localized population of azide and the excess probe can be washed out of cells without significant background reaction with small molecules such as thiols in the cytosol.

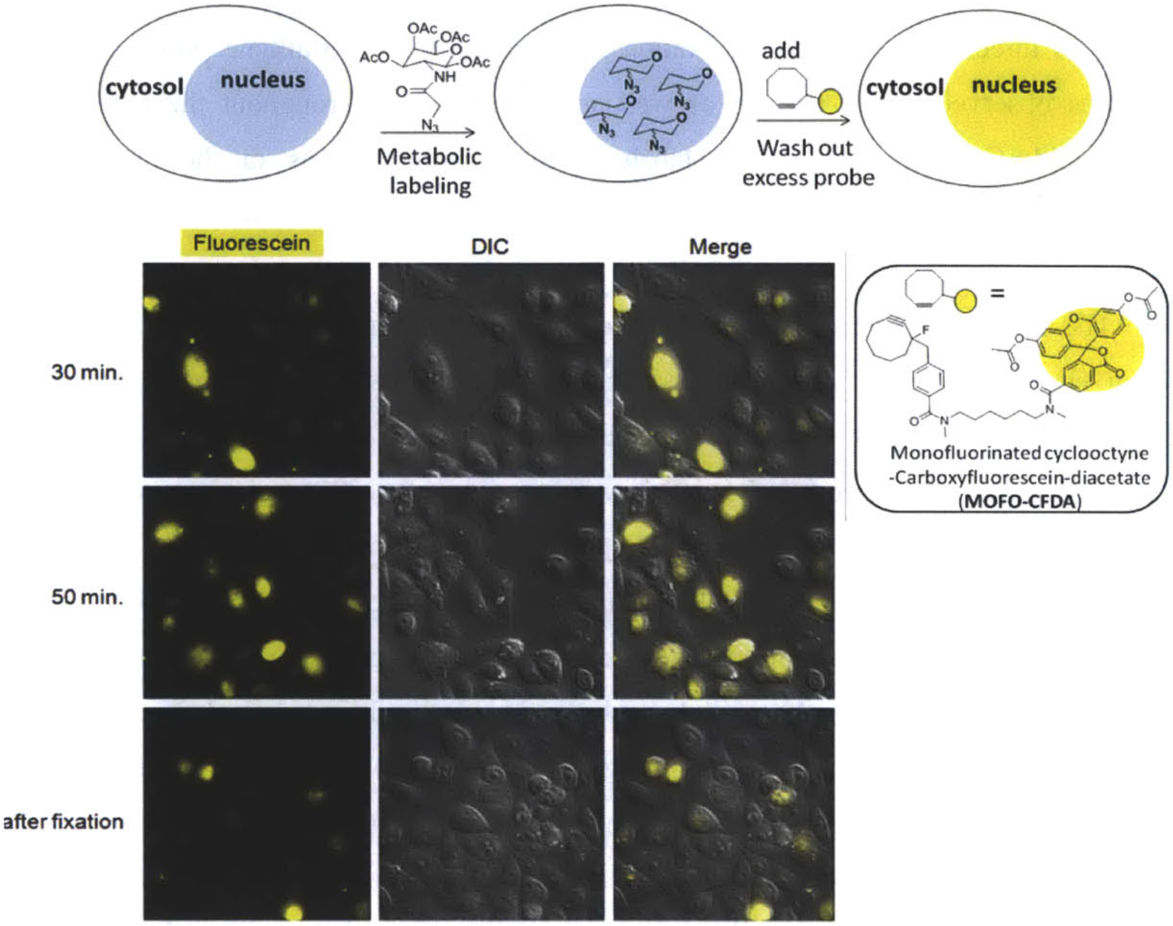

cytosol
nucleus
AcO
OAc
OAc
OAc
NH
N3
Metabolic labeling
add
Wash out excess probe
cytosol
nucleus
Fluorescein
DIC
Merge
30 min.
50 min.
after fixation
F
Monofluorinated cyclooctyne
-Carboxyfluorescein-diacetate
(MOFO-CFDA)

Figure 2-12. Live cell detection of metabolically incorporated azido sugar analog azido-GalNAc with fluorescein conjugated to cyclooctyne via strain-promoted azide-alkyne cycloaddition. (*Top*) Labeling scheme. Metabolic incorporation of azido-GalNAz was performed in CHO cells following protocol previously described. Afterwards, cells were incubated with 10 μM MOFO-CFDA (structure shown below) for 10 min. (*Bottom*) Images of live cells were taken at time points (indicated on the left side of images) post MOFO-CFDA addition. Strongest nuclear fluorescein signal was observed within the first hour after MOFO-CFDA labeling. Cells were also imaged after fixation and permeabilization to demonstrate the retention of nuclear signal and also to completely wash out any remaining unreacted MOFO-CFDA.

Conclusion

In this chapter, we described the screening and characterization of a new LplA ligase and azide pair, evaluation of azide ligation specificity, and performed strain-promoted azide-alkyne cycloaddition using an azido-sugar analog inside live mammalian cells. We developed assays to examine the two-step fluorophore targeting scheme systematically by isolating the azide ligation step from the strain-promoted azide-alkyne cycloaddition step so we can troubleshoot easily and identify potential problems without many confounding and unknown factors inside the cell.

The *in vitro* characterization shows the LplA(W37I) and azide 9 pair has an improved k_{cat}/K_m of two-fold over the wild-type LplA and azide 7 pair. Inside live mammalian cells, LplA(W37I)/azide 9 is able to achieve quantitative azide ligation yield over an incubation time of a brief 30 minutes, while wild-type LplA/azide 7 only had 73% yield over an hour. Furthermore, LplA(W37I)/azide 9 is also highly specific to the LAP-tagged protein. We did not observe cytosolic background signal when labeling nuclear-localized LAP-YFP. The strain-promoted cycloaddition reaction was tested using an LplA-independent, nuclear-localized source of azide, azido-GalNAc, and monofluorinated cyclooctyne conjugated to fluorescein (MOFO-CFDA). The reaction proceeded successfully inside live cells, giving strong nuclear fluorescein signal. The excess MOFO-CFDA probe was washed out of cells within an hour.

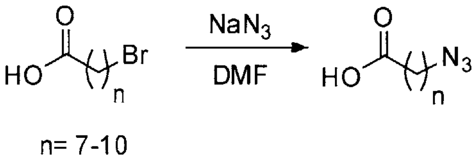
In the next chapter, we discuss using this optimized platform to evaluate the cycloaddition efficiencies of different cyclooctynes and demonstrating various large fluorophore targeting to different proteins across multiple cell lines. We also discuss some of the challenges and problems we faced in labeling with certain fluorophores inside live mammalian cells and ways we tried to overcome these challenges.

Experimental methods

General synthetic methods

All reagents were the highest grade available and purchased from Sigma-Aldrich, Anaspec, Thermal Scientific, TCI America, Alfa Aesar, or Life Technologies and used without further purification. Anhydrous solvents were drawn from Sigma-Aldrich SureSeal bottles. Analytical thin layer chromatography was performed on 0.25 mm silica gel 60 F254 plates and visualized under short or long wavelength UV light, or after staining with bromocresol green or ninhydrin. Flash column chromatography was carried out using silica gel (ICN SiliTech 32-63D). Mass spectrometric analysis was performed on an Applied Biosystems 200 QTRAP mass spectrometer using electrospray ionization. HPLC analysis and purification were performed on a Varian Prostar Instrument equipped with a photo-diode-array detector. A reverse-phase Microsorb-MV 300 C18 column (250 × 4.6 mm dimension) was used for analytical HPLC. NMR spectra were recorded on a Bruker AVANCE 400 MHz instrument.

Synthesis of alkyl azide probes



To a solution of the corresponding bromoalkanoic acid (~1 g, 5 mmol) in 10 mL N,N-dimethylformamide (DMF) was added sodium azide (~0.5 g, 7.5 mmol). The mixture was allowed to stir at room temperature overnight. The progress of the reaction was monitored by thin layer chromatography (1:2 hexanes:ethyl acetate) followed by bromocresol green stain. Upon completion, DMF was removed under reduced pressure. The resulting residue was re-dissolved in 15 mL of 1 M HCl and extracted with ethyl acetate (3 x 15 mL). The organic layer was dried over magnesium sulfate, then filtered. After removal of ethyl acetate *in vacuo*, the crude product was purified by silica gel chromatography (solvent gradient 0-15% ethyl acetate in hexanes) to afford the corresponding azidoalkanoic acid as clear or pale yellow oil. Yields ranged from 50-70%.

Characterization of n=7 azide (8-azidooctanoic acid). ^{1}H NMR ($CDCl_3$): 11.87 (s, 1H) 3.20 (t, 2H, J = 6.9), 2.28 (t, 2H, J = 7.5), 1.56 (m, 5H), 1.33 (m, 5H). ESI-MS calculated for $[M\text{-}H]^-$: 184.11; observed 183.66.

Characterization of n=8 azide (9-azidononanoic acid). ^{1}H NMR ($CDCl_3$) 3.22 (t, 2H, J = 6.9), 2.30, (t, 2H, J = 7.5), 1.60 (m, 5H), 1.29 (m, 7H). ESI-MS calculated for $[M{-}H]^-$: 198.12; observed 198.65.

Characterization of n=9 azide (10-azidodecanoic acid). ^{1}H NMR ($CDCl_3$): 3.23 (t, 2H, J = 6.9), 2.28 (t, 2H, J = 7.5), 1.53 (m, 5H), 1.31 (m, 9H). ESI-MS calculated for $[M\text{-}H]^-$: 212.14; observed 212.28.

Characterization of n=10 azide (11-azidoundecanoic acid) ^{1}H NMR ($CDCl_3$) 3.27 (t, 2H, J = 7.1), 2.39, (t, 2H, J = 7.5), 1.65 (m, 5H), 1.20 (m, 11H). ESI-MS calculated for $[M\text{-}H]^-$: 226.16; observed 226.12.

Protein expression and purification

Wild-type LplA and its mutant variants, and E2p were expressed in BL21-DE3 bacteria with IPTG induction as previously described (2). After Ni-NTA purification, LplA enzymes were dialyzed and stored in 20 mM tri pH 7.5, 1 mM DTT, and 10% glycerol in aliquots at -80°C. Protein concentrations were determined using the bicinchoninic acid assay (Pierce).

Kinetic analysis of azide 9 ligation (Figure 2-6 C)

Reactions were set up as described in the figure legend. Aliquots were taken and quenched before product conversion exceeded 5%. To calculate initial rates, we determined the amount of product at each time point by generating a calibration curve using purified LAP and LAP-azide 9 mixed at different ratios. This curve correlated the measured ratio of integrated HPLC peak areas to the actual ratio, i.e. adjusted for any differences in extinction coefficient of LAP vs. LAP-azide 9. Initial rates (V_o) were determined at each azide 9 concentration by plotting the amount of LAP-azide 9 product against time. The slope of the line gives V_o. V_o values were then plotted against azide 9 concentration in Figure 2-6 C, and Origin 8.5.1 was used to fit the curve to the Michaelis-Menten equation $V_o = V_{max}[\text{azide } 9] / (K_m + [\text{azide } 9])$. From the V_{max}, k_{cat} was calculated using $V_{max} = k_{cat}[E]_{total}$. Measurements of V_o values at each azide 9 concentration were performed in triplicate.

Detection of live cell azide ligation using cy5-alkyne via CuAAC (Figure 2-7)

HEK 293T cells plated on fibronectin-coated glass coverslips in wells of a 48-well cell culture plate (0.95 cm^2 per well) were transfected with 400 ng pcDNA3-LAP-YFP-NLS (nuclear-localized) and either 20 ng pcDNA3-wild-type LplA or 200 ng pcDNA3-LplA(W37I). 18 hours after transfection, cells were incubated in growth media containing 200 μM azide 7 or azide 9 for 1 hour at 37C. Excess azide probe was washed out by replacing azide-containing media with fresh media 3x over 1 hour. After azide washout, cells were washed 3x with DPBS and fixed with 3.7% formaldehyde in DPBS pH 7.4 for 10 min at room temperature followed by cold precipitation with methanol for 5 min at -20 °C. Cells were then blocked with 3% (w/v) BSA in DPBS for 1 hour at room temperature. To react the fixed cells with cy5-alkyne, cells were incubated in reaction mixture containing 50 μM cy5-alkyne, 1mM $CuSO_4$, 100 μM TBTA ligand, and 2.5 mM sodium ascorbate in DPBS overnight at room temperature. Cells were then washed 3x with DPBS and imaged.

Native gel-shift analysis of azide 7 and azide 9 ligation yields in cells (Figures 2-8)

HEK 293T cells were plated into wells of a 12-well culture plate (4 cm^2 per well) 18 hours prior to transfection and grown to 60% confluency. For azide 7 ligation, cells were transfected with 50 ng wild-type LplA and 1000 ng pcDNA3-LAP-YFP. For azide 9 ligation, cells were transfected with 500 ng LplA(W37I) and either 1000 ng pcDNA3-LAP-YFP or pcDNA3-LAP(K→A)-YFP using Lipofectamine 2000 (Life Technologies). The LplA:LAP plasmid ratios are identical to the conditions used for imaging in Figures 3-2 and 3-4. 18 hr after transfection, cells were incubated in growth media (MEM supplemented with 10% FBS) containing 200 μM azide 7 or azide 9 for 30 minutes or 1 hour at 37 °C. Excess azide probe was washed out over 1 hour. Cells were then harvested and lysed in 500 μL hypotonic lysis buffer (1 mM HEPES pH 7.5, 5 mM $MgCl_2$, 1 mM PMSF (Thermal Scientific, phenylmethanesulfonyl fluoride), 1 mM protease inhibitor cocktail (Sigma-Aldrich)), frozen at -20 °C, thawed at room temperature, then mixed by vortexing for 2 min. This freeze-thaw-vortex cycle was repeated three times. Cells were then centrifuged at 13,000 rpm for 2 min, and the supernatant was separated and analyzed on a 12% polyacrylamide native gel without SDS (5 μL lysate per lane) at constant 200 V. Prior to Coomassie staining, in-gel fluorescence of YFP was visualized on a FUJIFILM FLA-9000 instrument using LD473 laser and Long Pass Blue (LPB) filter. A repeat of the experiment in Figure 2-8 gave ligation yields of 67% for ^{WT}LplA (50 ng plasmid) + azide 7, and 89% for ^{W37I}LplA (500 ng plasmid) + azide 9 (data not shown).

Streptavidin gel-shift analysis of azide 9 ligation yield in cells (Figure 2-10)

HEK 293T cells were plated into wells of a 12-well culture plate (4 cm^2 per well) 18 hours prior to transfection and grown to 60% confluency. For azide 9 ligation, cells were transfected with 500 ng LplA(W37I) and 1000 ng pcDNA3-LAP-mCherry using Lipofectamine 2000 (Life Technologies). 18 hours after transfection, cells were incubated in growth media (MEM supplemented with 10% FBS) containing 200 μM azide 9 for 1 hour at 37 °C. Excess azide probe was washed out over 1 hour. Cells were then harvested and lysed in 500 μL hypotonic lysis buffer (1 mM HEPES pH 7.5, 5 mM $MgCl_2$, 1 mM PMSF (Thermal Scientific, phenylmethanesulfonyl fluoride), 1 mM protease inhibitor cocktail (Sigma-Aldrich)), frozen at -20 °C, thawed at room temperature, then mixed by vortexing for 2 min. This freeze-thaw-vortex cycle was repeated three times. Cells were then centrifuged at 13,000 rpm for 2 min, and the supernatant was collected. Click reaction was performed in cell lysate supernatant with condition as follows: Biotin-alkyne 100 μM, $CuSO_4$ 1mM, TBTA ligand 100 μM, 2.5 mM sodium ascorbate 2.5 mM. Reactions were incubated at RT for 16 hours. The lysate samples were then dialyzed in 1X PBS buffer for 4 hours to remove excess biotin-alkyne, and incubated with excess streptavidin (C_f =10uM) for 1 hour at RT. Lysates were loaded onto 12% SDS-PAGE gel without boiling samples in order to preserve mCherry fluorescence.

Detection of metabolic incorporation of azido-GalNAc with cy3-alkyne in fixed cells (Figure 2-11)

Glass coverslips were first coated with 100 μM or 200 μM azido-GalNAc in ethanol by letting the ethanol evaporate inside the cell culture fume hood. CHO cells or HeLa cells were then plated on azido-GalNAc-coated glass coverslips at 40% confluency in wells of a 48-well cell culture plate (0.95 cm^2 per well) and let grown for 2-3 days. After metabolic incorporation, cells were washed 3x with DPBS and fixed with 3.7% formaldehyde in DPBS pH 7.4 for 10 min at room temperature followed by cold precipitation with methanol for 5 min at -20 °C. Cells were then blocked with 3% (w/v) BSA in DPBS for 1 hour at room temperature. To react the fixed cells with cy3-alkyne, cells were incubated in reaction mixture containing 50 μM cy3-alkyne, 1mM $CuSO_4$, 100 μM TBTA ligand, and 2.5 mM sodium ascorbate in DPBS overnight at room temperature. Cells were then washed 3x with DPBS and imaged.

Detection of metabolic incorporation of azido-GalNAc with MOFO-CFDA in live cells (Figure 2-12)

Glass coverslips were first coated with 200 μM azido-GalNAc in ethanol by letting the ethanol evaporate inside the cell culture fume hood. CHO cells or HeLa cells were then plated on azido-GalNAc-coated glass coverslips at 40% confluency in wells of a 48-well cell culture plate (0.95 cm^2 per well) and let grown for 2-3 days. After metabolic incorporation, cells were incubated in growth media containing 10 μM MOFO-CFDA for 10 minutes at 37C. Excess MOFO-CFDA was washed out of cells by replacing media with fresh media 3x over 1 hour. Cells were imaged live at 30 minutes and 50 minutes post MOFO-CFDA addition. Afterwards, cells were washed 3x with DPBS, fixed with 3.7% formaldehyde in DPBS pH 7.4 for 10 min at room temperature followed by cold precipitation with methanol for 5 min at -20 °C. Fixed cells were imaged again to illustrated that the nuclear fluorescence signal can be retained, indicating the reaction was probably covalent.

References

1. Ali ST, Guest JR. Isolation and characterization of lipoylated and unlipoylated domains of the E2p subunit of the pyruvate dehydrogenase complex of Escherichia coli. Biochem J. 1990 Oct 1;271(1):139–45.

2. Green DE, Morris TW, Green J, Cronan JE, Guest JR. Purification and properties of the lipoate protein ligase of Escherichia coli. Biochem J. 1995 Aug 1;309(Pt 3):853–62.

3. Fernández-Suárez M, Baruah H, Martínez-Hernández L, Xie KT, Baskin JM, Bertozzi CR, et al. Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. Nature Biotechnology. 2007;25(12):1483–7.

4. Puthenveetil S, Liu DS, White KA, Thompson S, Ting AY. Yeast Display Evolution of a Kinetically Efficient 13-Amino Acid Substrate for Lipoic Acid Ligase. J. Am. Chem. Soc. 2009 Nov 18;131(45):16430–8.

5. Uttamapinant C, White KA, Baruah H, Thompson S, Fernández-Suárez M, Puthenveetil S, et al. A fluorophore ligase for site-specific protein labeling inside living cells. PNAS. 2010 Jun 15;107(24):10914–9.

6. Slavoff SA, Liu DS, Cohen JD, Ting AY. Imaging Protein–Protein Interactions inside Living Cells via Interaction-Dependent Fluorophore Ligation. J. Am. Chem. Soc. 2011 Dec 14;133(49):19769–76.

7. Jin X, Uttamapinant C, Ting AY. Synthesis of 7-Aminocoumarin by Buchwald–Hartwig Cross Coupling for Specific Protein Labeling in Living Cells. ChemBioChem. 2011;12(1):65–70.

8. Cohen JD, Thompson S, Ting AY. Structure-Guided Engineering of a Pacific Blue Fluorophore Ligase for Specific Protein Imaging in Living Cells. Biochemistry. 2011 Sep 27;50(38):8221–5.

9. Baruah H, Puthenveetil S, Choi Y-A, Shah S, Ting AY. An Engineered Aryl Azide Ligase for Site-Specific Mapping of Protein–Protein Interactions through Photo-Cross-Linking. Angewandte Chemie International Edition. 2008;47(37):7018–21.

10. Liu DS, Tangpeerachaikul A, Selvaraj R, Taylor MT, Fox JM, Ting AY. Diels–Alder Cycloaddition for Fluorophore Targeting to Specific Proteins inside Living Cells. J. Am. Chem. Soc. 2012 Jan 18;134(2):792–5.

11. Saxon E, Bertozzi CR. Cell surface engineering by a modified Staudinger reaction. Science. 2000 Mar 17;287(5460):2007–10.

12. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. Proc. Natl. Acad. Sci. U.S.A. 2002 Jan 8;99(1):19–24.

13. Prescher JA, Dube DH, Bertozzi CR. Chemical remodelling of cell surfaces in living animals. Nature. 2004 Aug 19;430(7002):873–7.

14. Dube DH, Prescher JA, Quang CN, Bertozzi CR. Probing mucin-type O-linked glycosylation in living animals. Proc. Natl. Acad. Sci. U.S.A. 2006 Mar 28;103(13):4819–24.

15. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. A Comparative Study of Bioorthogonal Reactions with Azides. ACS Chem. Biol. 2006 Nov 1;1(10):644–8.

16. Baskin JM, Prescher JA, Laughlin ST, Agard NJ, Chang PV, Miller IA, et al. Copper-free click chemistry for dynamic in vivo imaging. PNAS. 2007 Oct 23;104(43):16793–7.

17. Agard NJ, Prescher JA, Bertozzi CR. A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. J. Am. Chem. Soc. 2004 Nov 1;126(46):15046–7.

18. Chang PV, Prescher JA, Hangauer MJ, Bertozzi CR. Imaging cell surface glycans with bioorthogonal chemical reporters. J. Am. Chem. Soc. 2007 Jul 11;129(27):8400–1.

19. Lin FL, Hoyt HM, Van Halbeek H, Bergman RG, Bertozzi CR. Mechanistic investigation of the staudinger ligation. J. Am. Chem. Soc. 2005 Mar 2;127(8):2686–95.

20. Beatty KE, Fisk JD, Smart BP, Lu YY, Szychowski J, Hangauer MJ, et al. Live-Cell Imaging of Cellular Proteins by a Strain-Promoted Azide–Alkyne Cycloaddition. ChemBioChem. 2010;11(15):2092–5.

21. Chang PV, Prescher JA, Sletten EM, Baskin JM, Miller IA, Agard NJ, et al. Copper-free click chemistry in living animals. PNAS. 2010 Feb 2;107(5):1821–6.

22. Hang HC, Yu C, Kato DL, Bertozzi CR. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. Proc. Natl. Acad. Sci. U.S.A. 2003 Dec 9;100(25):14846–51.

23. Hang HC, Yu C, Pratt MR, Bertozzi CR. Probing glycosyltransferase activities with the Staudinger ligation. J. Am. Chem. Soc. 2004 Jan 14;126(1):6–7.

24. Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR. In vivo imaging of membrane-associated glycans in developing zebrafish. Science. 2008 May 2;320(5876):664–7.

25. Freeze, H. H., Ernst, B., Hart, G. W. & Sinay, P. Carbohydrates in Chemistry and Biology. Carbohydrates in Chemistry and Biology. Wiley, New York; p. 3–14.

26. Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu. Rev. Biochem. 2011 Jun 7;80:825–58.

Chapter 3. Optimization and demonstration of diverse fluorophore targeting inside live cells

The work discussed in this chapter has been published in: J. Yao, C. Uttamapinant, A. Poloukhtine, J.M. Baskin, J.A. Codelli, E.M. Sletten, C.R. Bertozzi, V.V. Popik, A.Y. Ting, "Fluorophore targeting to cellular proteins via enzyme-mediated azide ligation and strain-promoted cycloaddition", *J Am Chem Soc* **2012** (134), 3720-3728.

Introduction

In the previous chapter, we describe the *in vitro* screening and characterization of the LplA mutant enzyme and azide probe, as well as assays developed to assess the live-cell azide ligation and strain-promoted cycloaddition reaction. We have established working conditions providing quantitative azide ligation yield, with optimized azide concentration and incubation time. Live-cell labeling of cyclooctyne-fluorescein conjugate with metabolically incorporated azido-GalNAc gave nuclear specific labeling signal after excess probe washout. In this chapter, we describe further screening, quantitative comparisons, and optimization to the final two-step labeling protocol carried out inside live cells.

First, we performed additional intracellular labeling assays to compare the labeling efficiency of LplA(W37I) and azide 9 with three other top LplA mutants and azide pairs from the HPLC screen (Chapter 2). We ensured for equal expression of all enzymes by immunostaining. Second, we give a brief overview of the scientific literature regarding cyclooctynes and their rich evolution and engineering history over the years to improve their reaction kinetics and hydrophilicity for applications in live cells and animals. We chose five cyclooctynes and compared their reactivity, labeling efficiency, and background using our two-step fluorophore targeting platform. Third, we discuss the challenges in intracellular labeling with large, hydrophobic, and positively charged fluorophores, such as ATTO 647N, and our attempts to overcome the challenges. Finally, we demonstrate using the optimized site-specific two-step fluorophore targeting method to label many intracellular proteins with various fluorophores across multiple cell lines.

In summary, this chapter encloses our intracellular efforts to optimize the two-step fluorophore labeling technique. Several of the fluorophores targeted using LplA and strain-promoted cycloaddition are attractive alternatives to fluorescent proteins. The methodology will be most useful as a nontoxic labeling method for abundant proteins, whose fusion to large tags perturbs function.

Comparison of LplA mutant enzymes and azide pairs using a live-cell labeling assay

Knowing that the azide ligation can achieve quantitative ligation and that the strain-promoted azide-alkyne cycloaddition can give specific labeling inside live cells, we were ready to combine the two steps. We have shown previously that LplA(W37I) and azide 9 exhibited better *in vitro* ligation kinetics with 2-fold improvement in k_{cat}/K_m over wild-type LplA and azide 7 pair. Also, native gel-shift assay showed LplA(W37I) ligation of azide 9 achieved almost complete live cell azide ligation yield in just 30 minutes, while wild-type LplA ligation of azide 7 achieved only about 73% yield in 1 hour (Figure 2-8). We wanted to further differentiate among the LplA mutant/azide pairs by comparing LplA(W37I)/azide 9 with three other top enzyme/azide pairs (starred in Figure 2-5) from the HPLC screen under the more stringent context of live cell labeling.

We first empirically optimized the azide washout time required to fully remove excess unligated azide, using cyclooctyne-fluorescein retention as readout. We incubated untransfected HEK 293T cells with azide 9 for 1 hour, followed by azide washout for either 10 minutes or 1 hour. We then added the cyclooctyne-fluorophore conjugate, MOFO-CFDA to the cells for 10 minutes, followed by 2 hours washout. Figure 3-1 shows that 1 hour of azide washout was enough to prevent MOFO-fluorescein retention inside cells, as the background was comparable to cells not treated with azide 9.

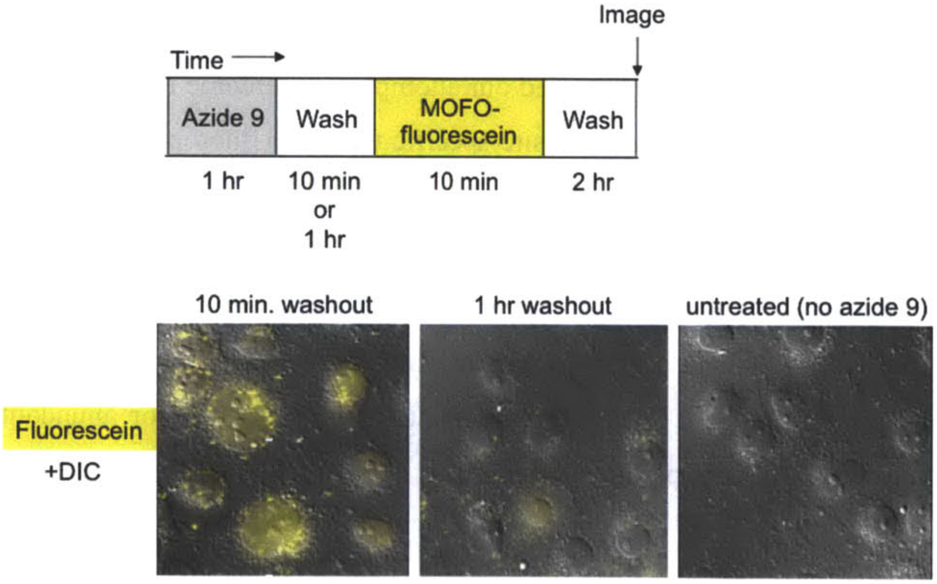


Figure 3-1. Determination of azide 9 washout time in live cells. Following the protocol shown at the top, untransfected HEK 293T cells were incubated with 200 μM azide 9 for

1 hour, then washed for either 10 min or 1 hour, before adding 10 μM MOFO-fluorescein diacetate for 10 min. Cells were imaged after a 2 hours wash. The images show that fluorescein background is almost as low with 1 hour azide washout, as in cells not treated with azide at all.

Since azide 9 is two methylene units longer than azide 7 and thus more hydrophobic, we assumed that the 1 hour washout should be sufficient for azide 7 as well. To compare the top four LplA mutant/azide pairs using the two-step labeling method inside live cells, we transfected HEK 293T cells with the substrate protein LAP-tagged blue fluorescent protein (BFP) and LplA plasmids corresponding to each of the top four ligases. After transfection, following the labeling protocol shown in Figure 3-2, we incubated the cells with azide 7 or azide 9 for 1 hour and washed out excess azide for 1 hour. MOFO-fluorescein diacetate was added to the cells for 10 minutes to derivatize the azide-LAPs. After 2 hours of washing to remove excess fluorophore, cells were imaged live. Figure 3-2 shows specific labeling of LAP-BFP for all four ligase and azide combinations, but the highest signal-to-background ratio was obtained for the LplA(W37I) and azide 9 pair. Compared to the wild-type LplA and azide 7 pair previously used for cell-surface protein labeling, we can see that LplA(W37I) and azide 9 labeling offered substantial improvement in signal intensity, about 4-fold greater on average.

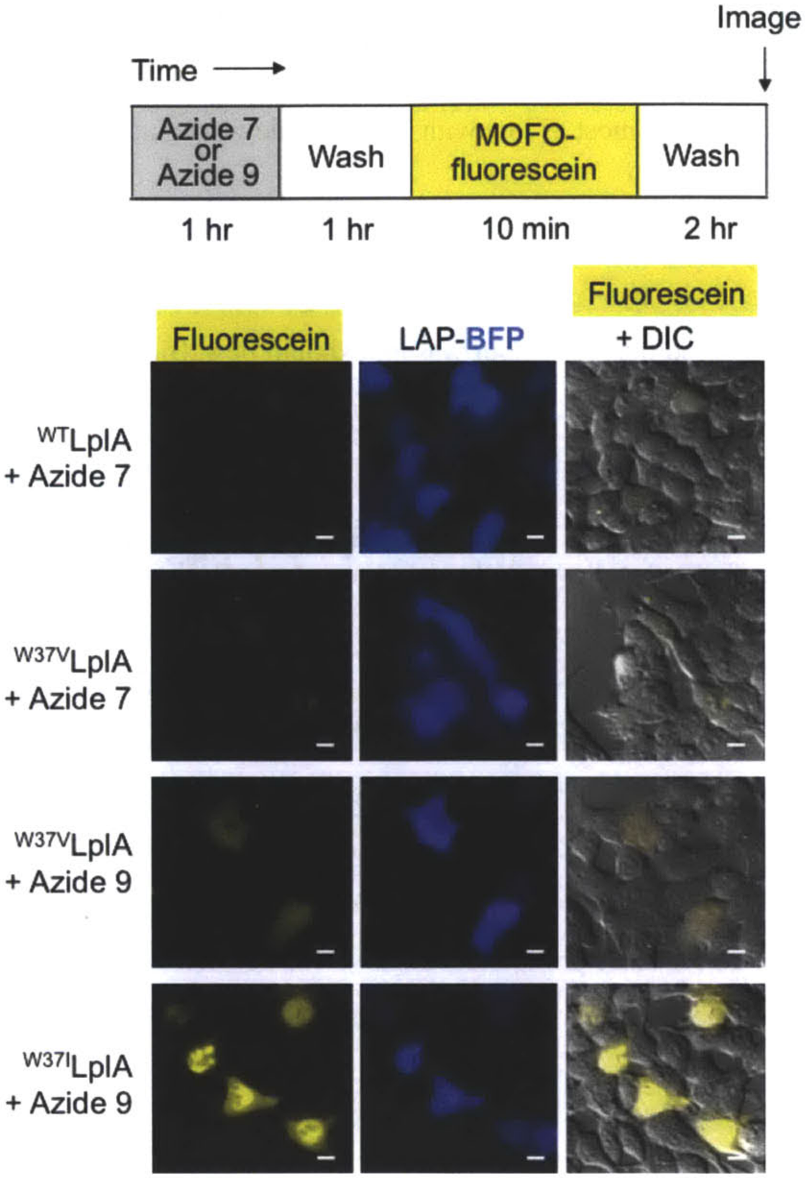


Image
Time
Azide 7 or Azide 9
Wash
MOFO-fluorescein
Wash
1 hr
1 hr
10 min
2 hr
Fluorescein
LAP-BFP
Fluorescein + DIC
WTLplA + Azide 7
W37VLplA + Azide 7
W37VLplA + Azide 9
W37ILplA + Azide 9

Figure 3-2. Identification of the best LplA mutant/azide substrate pair for intracellular protein labeling. (top) Labeling protocol: HEK cells co-expressing indicated LplA mutant enzyme and LAP-tagged BFP were labeled with azide 7 or azide 9 for 1 hour, washed for 1 hour, then labeled with monofluorinated cyclooctyne conjugated to fluorescein diacetate (MOFO-CFDA, structure in Figure 3-5) for 10 minutes. Thereafter, cells were washed again for 2 hours to remove excess unconjugated fluorophore. (bottom) Images of HEK cells labeled as following to top protocol, with different LplA mutant/azide probe pairs (starred combinations in Figure 2-5). Scale bars, 10 μm.

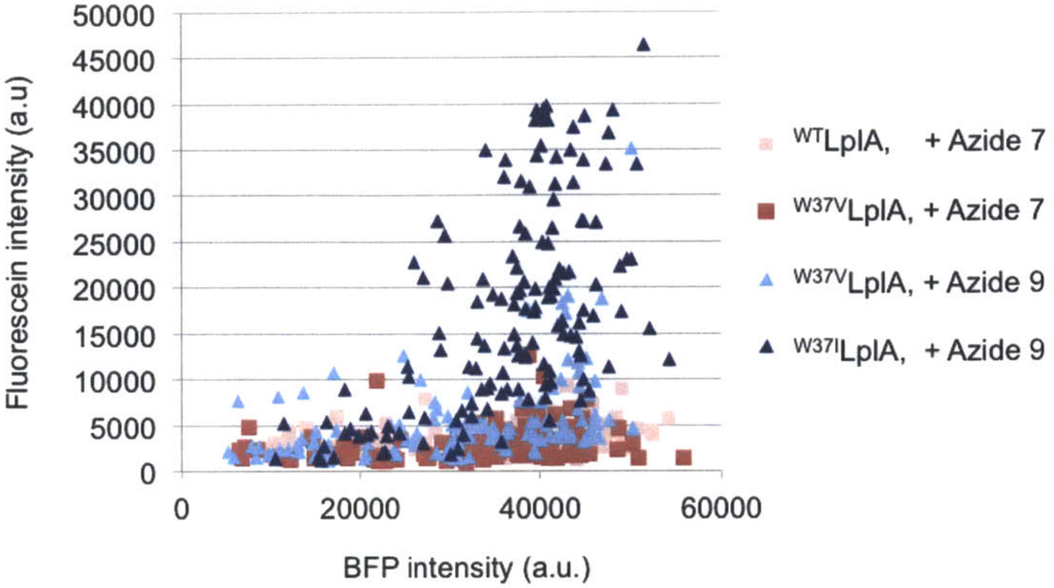


Figure 3-3. Quantitation of data shown in Figure 3-2. For each condition, the mean fluorescein intensity was plotted against the mean BFP intensity, for >100 single cells. Fluorescein ligation yield is highest for the LplA(W37I)/azide 9 combination.

We further quantified the labeling signal differences in Figure 3-3, in which fluorescein intensity is plotted against LAP-BFP expression level for >100 single cells for each condition. Two-step labeling with LplA(W37I) and azide 9 gave about 5-fold higher fluorescein signal than the other three ligase and azide pairs. After live cell labeling and imaging, we fixed and permeabilized the cells to perform anti-FLAG immunofluorescence staining to detect the enzyme expression level of FLAG-tagged LplA in cells (Figure 3-4). We needed to ensure that the perceived higher labeling signal was not caused simply by higher enzyme expression level of LplA(W37I). Anti-FLAG immunostaining showed that ligase expression levels are all comparable under our experimental conditions.

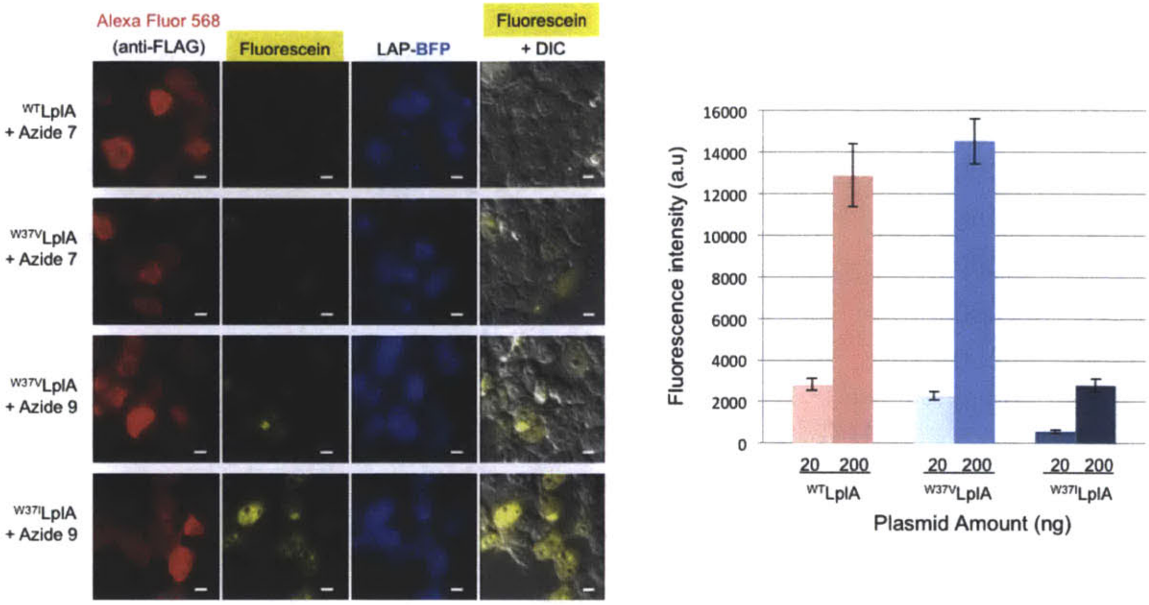


Figure 3-4. Immunofluorescence with anti-FLAG to compare labeling against expression of all enzymes. (left) A repeat of labeling in Figure 3-2 following the same labeling protocol, except that after live cell labeling, cells were fixed and stained with anti-FLAG antibody to visualize FLAG-tagged LplA expression levels. To obtain comparable LplA expression levels of mutants here and in Figure 3-2, ten-fold more LplA(W37I) plasmid was introduced compared to wild-type LplA and LplA(W37V) plasmids. All scale bars, 10 μm. (right) Graph showing quantitation of LplA mutant expression levels in cells. Mean Alexa Fluor 568 intensities (anti-FLAG staining) for > 30 single cells from each condition were averaged and graphed here.

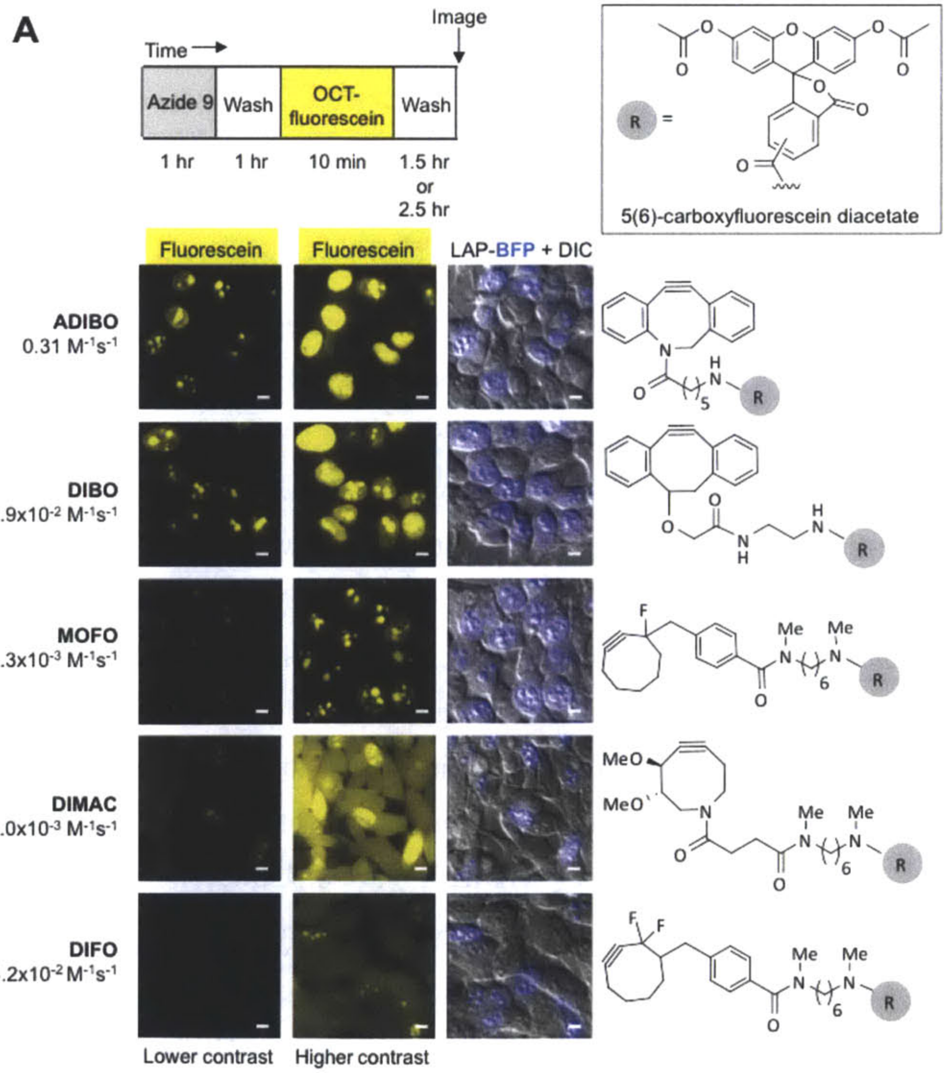
Comparison of cyclooctyne structures

After optimizing the azide ligation of the two-step fluorophore labeling method, we proceeded to the second step of strain-promoted azide-alkyne cycloaddition. Numerous cyclooctynes have been developed by labs all around the world. The toolbox of cyclooctynes has been expanding rapidly over the past decade owing to the wide application of strain-promoted [3+2] azide-alkyne cycloaddition from *in vitro* small molecules (1–3) and protein functionalization (4–8), proteome profiling (9), cell surface glycan modification (10), tagging and imaging proteins inside cells (11), to *in vivo* ligation in mice (12). With a rich panel of cyclooctynes available, we needed to choose a cyclooctyne best suited for our goal of fluorophore targeting to proteins inside live cells.

Since the first report of using cyclooctynes to label cell surface glycoproteins (13), many research groups have engineered numerous variations of subsequent generations of cyclooctynes (4,14–19). These structures vary in terms of ring strain, electron deficiency and hydrophilicity, which in turn influence reactivity toward azide and other intracellular molecules, such as thiols. The design and synthesis of these cyclooctynes revolved around two main goals: to increase cyclooctyne's reactivity and to improve its hydrophilicity. A cyclooctyne with fast reaction kinetics with azide can largely improve the sensitivity of labeling which is necessary especially within an environment as complex as inside the cell or even in animals. Two routes have been explored to promote reactivity: first, introducing electron-withdrawing groups, like fluorine, adjacent to the alkyne to lower the LUMO of the cyclooctyne so it has better energy orbital overlap with the HOMO of the azide (14,15); second, imposing ring strain around the alkyne by fusing two benzene rings on each side of cyclooctyne (4,17,18). Using these two strategies, we have seen a second order rate constant of nearly 800-fold increase from simple cyclooctyne ($1.2\text{x}10^{-3}\ M^{-1}S^{-1}$) (14) to biarylazacyclooctynone (BARAC, $0.96\ M^{-1}S^{-1}$) (17). However, increased reactivity also comes at the cost of increased hydrophobicity especially when dibenzo groups are attached to raise ring strain (4,17,18). Although the dibenzo rings impart a large enhancement in reactivity, the hydrophobic character of these cyclooctynes leads to high chance of nonspecific binding to cellular membranes and sequestration by hydrophobic pockets. Placement of di-methoxy groups and installing a nitrogen atom within the cyclooctyne ring to break its hydrophobic surface area and to reduce the extent of nonspecific hydrophobic binding to cellular hydrophobic areas have been tried to improve the hydrophilicity of these probes (4,16). With respect to our LplA-mediated technique of intracellular fluorophore targeting, in addition to having efficient strain-promoted azide-alkyne cycloaddition and stability inside live cells, the ideal cyclooctyne would strike a balance between high reactivity and low hydrophobicity.

We selected a panel of five cyclooctyne structures to test inside live cells because it was not clear which cyclooctyne structure(s) would be the best for our purpose. In order to compare the reactivity of all five cyclooctynes in parallel without the compounding effect caused by the cell entry and washout properties of the fluorophore, we derivatized each cyclooctyne with fluorescein diacetate via a stable amide bond (Figure 3-5). We selected fluorescein for this purpose because it has excitation (λ_{ex} 490 nm) and emission wavelengths (λ_{em} 514 nm) suitable for live-cell imaging and high quantum yield (0.95). Most importantly, it is the least hydrophobic among the fluorophores we have tested and can be washed out of cells during a relatively short period of time (less than 2 hours) which we had seen during previous experiments using MOFO-CFDA to label metabolically incorporated azido-GalNAc.

With the optimized protocol of LplA(W37I) mediated azide 9 ligation and a panel of cyclooctyne-fluorescein diacetate conjugates, we performed two-step intracellular labeling on nuclear localized LAP-BFP in live HEK 293T cells following the scheme in Figure 3-5 A (Top). After azide 9 ligation and washout, cells were treated with one of the five cyclooctyne-fluorescein diacetate conjugates. Confocal images were taken at multiple time points after probe incubation. It is immediately clear that the probes ADIBO, DIBO, and MOFO showed very specific nuclear labeling with low background (Figure 3-5 A). MOFO-fluorescein washed out the fastest among the three and specific labeling was observed at only 1.5 hours after probe addition. However, after normalization, it is clear that the nuclear signal of ADIBO and DIBO appeared to be much higher than signal from MOFO. This is most likely due to the superior second order rate constants of ADIBO (0.31 $M^{-1}s^{-1}$ in methanol) (19) and DIBO (5.92×10^{-2} $M^{-1}s^{-1}$ in methanol) (18) compared to MOFO (4.3×10^{-3} $M^{-1}s^{-1}$) (14). We quantitatively analyzed the signal to noise ratio of ADIBO-, DIBO-, and MOFO-fluorescein by plotting the signal from the nucleus versus background in the cytosol of > 50 single cells for each cyclooctyne-fluorescein (Figure 3-5 B). ADIBO- and DIBO-fluorescein both gave about 4-5 times higher nuclear signal than MOFO-fluorescein, however, their cytosolic backgrounds are also higher. The overall signal to noise ratios (defined as nuclear signal over cytosolic background) of all three cyclooctynes are comparable. The higher absolute signal of ADIBO and DIBO makes them superior cyclooctynes than MOFO. However, we note that the faster washout time of MOFO-fluorescein can be advantageous in certain applications.


A
Image
Time
Azide 9
Wash
OCT-fluorescein
Wash
1 hr
1 hr
10 min
1.5 hr or 2.5 hr
R =
5(6)-carboxyfluorescein diacetate
Fluorescein
Fluorescein
LAP-BFP + DIC
ADIBO
$0.31\ M^{-1}s^{-1}$
DIBO
$5.9x10^{-2}\ M^{-1}s^{-1}$
MOFO
$4.3x10^{-3}\ M^{-1}s^{-1}$
DIMAC
$3.0x10^{-3}\ M^{-1}s^{-1}$
DIFO
$4.2x10^{-2}\ M^{-1}s^{-1}$
Lower contrast
Higher contrast

B

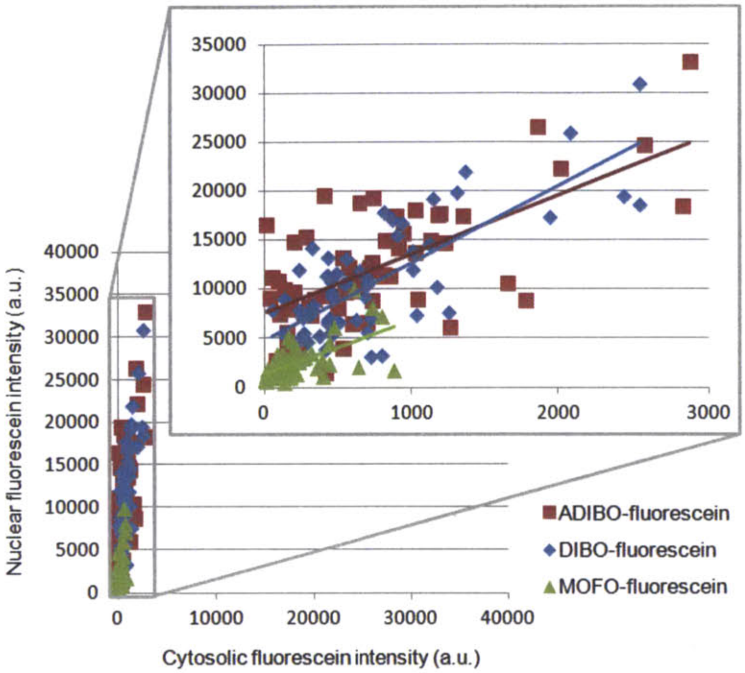


Figure 3-5. Evaluation of various cyclooctyne structures for site-specific intracellular protein labeling. (A) Top: labeling protocol for HEK 293T cells co-expressing LplA(W37I) and nuclear-localized LAP−BFP (LAP−BFP−NLS). After labeling with azide 9 for 1 hour and washing for 1 hour, cells were treated with the indicated cyclooctyne-fluorescein conjugate for 10 minutes. Cells were washed again for 2.5 hours to remove excess unconjugated fluorophore, except for MOFO, which was washed for only 1.5 hours. Bottom: images of labeled HEK 293T cells. The LAP−BFP−NLS image is overlaid on the DIC image. Fluorescein signal intensities can be compared in the first two columns, which show the fluorescein images at lower contrast (left) and higher contrast (middle). Cyclooctyne structures are shown at right, and second-order rate constants are shown at left. ADIBO, aza-dibenzocyclooctyne; DIBO, 4-dibenzocyclooctynol; MOFO, monofluorinated cyclooctyne; DIMAC, 6,7-dimethoxyazacyclooct-4-yne; DIFO, difluorinated cyclooctyne. All scale bars, 10 μm. (B) Quantitation of data in (A). For the three cyclooctynes (ADIBO, DIBO, and MOFO), the mean nuclear fluorescein intensity (representing specific labeling) was plotted against the mean cytosolic fluorescein intensity (representing nonspecific labeling), for the same cell. More than 50 single cells were analyzed for each cyclooctyne.

Nuclear labeling using DIMAC-fluorescein showed very weak signal, perhaps due to the slow second order rate constant of DIMAC ($3.0x10^{-3}$ $M^{-1}s^{-1}$ in CD_3CN) (16). We tried to increase signal by incubating cells with up to 30 μM of DIMAC-fluorescein without significant improvement (data not shown). Aside from low signal, we also observed retention of the probe in certain non-transfected cells, which was unexpected because DIMAC was noted for the hydrophilicity conferred by the di-methoxy groups. The background fluorescence was much reduced after cell fixation and further washout (Figure 3-6), indicating it was perhaps not caused by a covalent reaction.

For DIFO-fluorescein, we observed nuclear signal specific to transfected cells as soon as 10 min after probe addition, however, this was accompanied by the high cytosolic background that persisted even after cells were fixed and permeabilized (Figure 3-6), at which point the small molecules inside cells should be thoroughly washed out. We also tried to reduce or prevent this seemingly permanent cytosolic background by lowering the concentration of DIFO-FDA to 1 μM and incubation time as short as 40 seconds, but the cytosolic background still persisted even though at lower intensity (data not shown). Despite the fast reactivity of DIFO ($7.6x10^{-2}$ in CD_3CN) (15), our observation indicates that DIFO was retained covalently inside cells perhaps through reaction with nonspecific nucleophiles. This hypothesis is strengthened by the result from Beatty and coworkers that showed DIFO reacting with cysteine and glutathione at the strained-alkyne moiety *in vitro* (11). With cellular glutathione concentration in the millimolar range (20), it is very likely that the irreversible high background in intracellular DIFO labeling could be caused by reaction with glutathione.

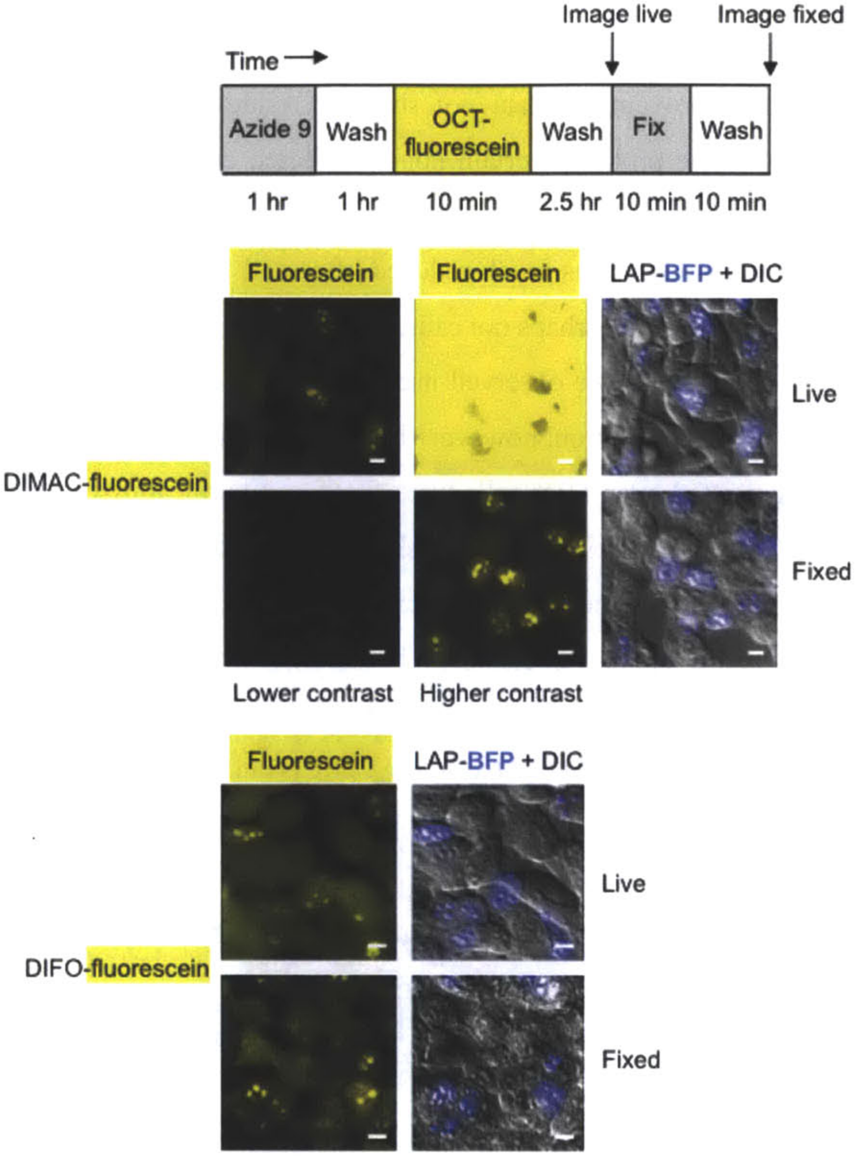


Image live
Image fixed
Time
Azide 9
Wash
OCT-fluorescein
Wash
Fix
Wash
1 hr
1 hr
10 min
2.5 hr
10 min
10 min
Fluorescein
Fluorescein
LAP-BFP + DIC
Live
DIMAC-fluorescein
Fixed
Lower contrast
Higher contrast
Fluorescein
LAP-BFP + DIC
Live
DIFO-fluorescein
Fixed

Figure 3-6. Analysis of background labeling by cyclooctyne-fluorescein conjugates. Samples from Figure 3-5 A labeled with DIMAC-fluorescein and DIFO-fluorescein were re-imaged after cell fixation and permeabilization. Labeling and imaging protocol is given at the top. Non-specific fluorescence is much reduced after fixation for DIMAC-fluorescein but not as much for DIFO-fluorescein.

Furthermore, work from Chang *et al* using DIMAC and DIFO, each conjugated to a FLAG epitope tag to detect cell surface sialic acids labeled with azides in live mouse also showed background non-specific labeling (12). In the negative control experiment, DIMAC and DIFO seemed to bind to mouse serum albumin (MSA) in heart lysates in an azide-independent manner. Further investigation with MSA immunoprecipitation and probing with anti-MSA and anti-FLAG antibodies showed that the nonspecific signals still persisted even after the samples were treated with 1 M urea, β-mercaptoethanol, and boiled, suggesting that it might be caused by a combination of hydrophobic and covalent interactions, perhaps with the two free cysteine residues in MSA. Even though the high reactivity of DIFO offers the potential of high sensitivity and fast labeling, we need to be concerned about nonspecific reactions when using this cyclooctyne.

On the basis of these results, we selected ADIBO for most of our intracellular protein labeling experiments. However, as shown later, due to ADIBO's hydrophobicity, MOFO is a better option when working with very hydrophobic fluorophores such as ATTO 647N.

Demonstration of intracellular protein labeling with ADIBO-fluorescein diacetate

Having optimized both the azide ligase and the cyclooctyne, we proceeded to characterize two-step fluorophore labeling inside cells, and explore its generality. We first empirically optimized the ADIBO-fluorophore loading concentration and washout time by testing five concentrations of the probe and imaging cells at various times post probe incubation. As shown in Figure 3-7, washout time increased to more than 2.5 hours at 40 μM of ADIBO-fluorescein diacetate. We determined that 10 μM probe, followed by 2.5 hours of washout was effective.

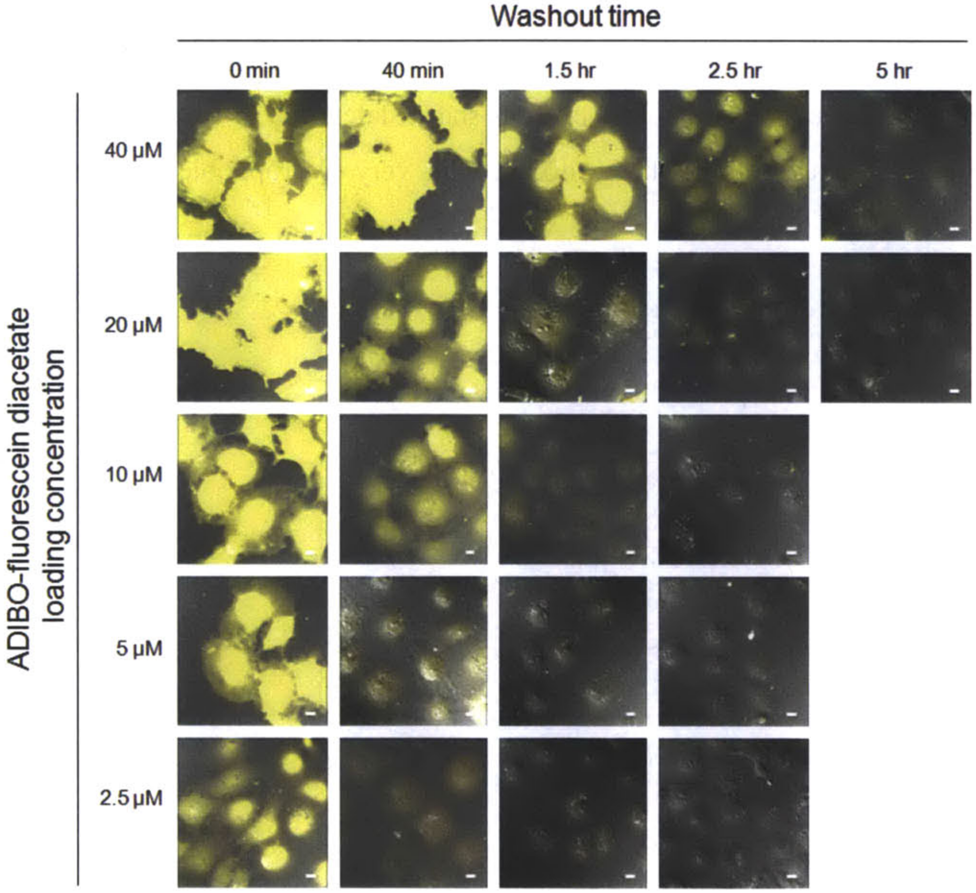


Figure 3-7. Two-dimensional optimization of ADIBO-fluorescein diacetate loading concentration and washout time. Various amounts of ADIBO-fluorescein were loaded into untransfected COS-7 cells for 10 minutes at 37 °C. Various washout times were tested, ranging from 0 to 5 hours. Fluorescein images are shown with DIC overlay. All scale bars, 10 μm. Based on this data, we selected 10 μM ADIBO-fluorescein and 2 - 2.5 hour washout time for subsequent experiments.

Following the scheme shown in Figure 3-8, we labeled LAP-tagged BFP protein localized to various compartments inside the cell. Figure 3-8 shows that only cells expressing LAP-BFP retained fluorescein, while neighboring untransfected cells did not. Negative controls performed in parallel showed labeling is completely abolished when the catalytic lysine residue in LAP is mutated to alanine, when azide 9 is omitted, and when using a catalytically inactive LplA enzyme which has residue K133 that is involved in both lipoic acid adenylation and lipoate transfer mutated to R.

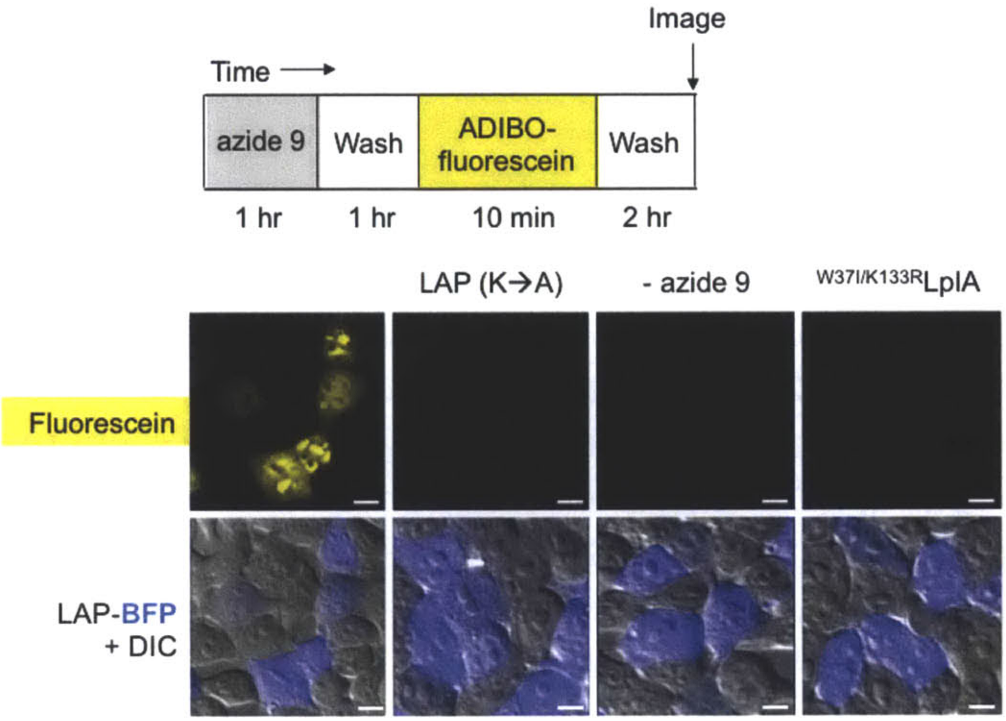


Figure 3-8. Two-step fluorophore targeting via strain-promoted azide-alkyne cycloaddition. Following the labeling scheme shown on top, HEK 293T cells expressing LplA(W37I) and LAP-BFP were labeled with azide 9 and ADIBO-fluorescein diacetate. Only transfected cells retained fluorescein signal. Negative controls with non-catalytic LAP(K→A), omitting azide 9, and dead mutant LplA(W37I, K133R) did not show labeling.

We also demonstrated labeling of different LAP fusion proteins (Figure 3-9) and in different cell lines. Using the two-step protocol shown in the scheme in Figure 3-8, we successfully labeled LAP in the nucleus, cytosol, and plasma membrane, as well as LAP fusions to actin and MAP2 (microtubule associated protein 2) with robust signal. These experiments were performed in multiple mammalian cell lines, including HEK 293T, HeLa, and COS-7, demonstrating the versatility of the method.

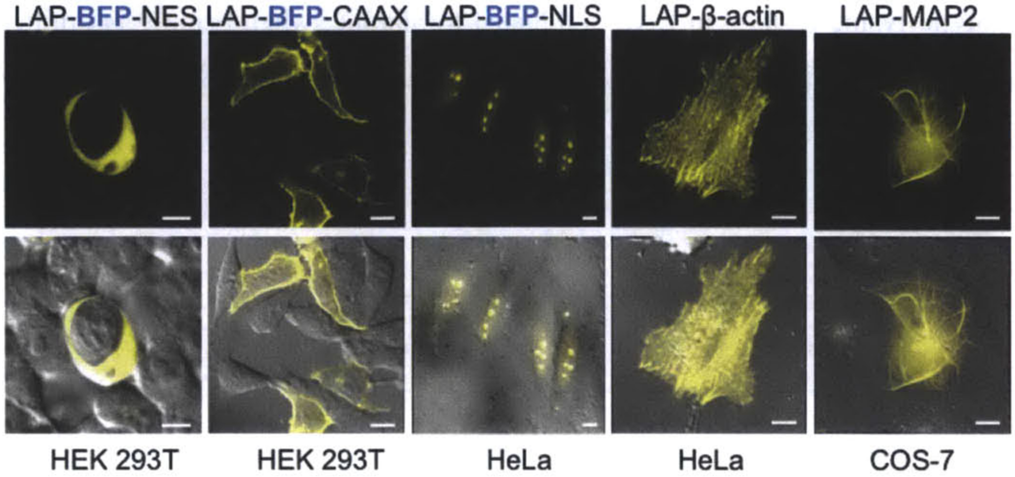

LAP-BFP-NES
LAP-BFP-CAAX
LAP-BFP-NLS
LAP-β-actin
LAP-MAP2
HEK 293T
HEK 293T
HeLa
HeLa
COS-7

Figure 3-9. ADIBO−fluorescein labeling of three localized LAP−BFP fusions, LAP−β-actin, and LAP−MAP2 (microtubule-associated protein 2). Labeling in the cell type indicated beneath each image was performed as in Figure 3-8, except that for LAP−β-actin and LAP−MAP2, azide 9 was incubated for 2 hours, and washed for 1.5 hours before fluorophore addition. NES = nuclear export sequence; CAAX = prenylation tag; NLS = nuclear localization sequence. All scale bars, 10 μm.

Challenges and optimization for targeting large fluorophores inside live cells

Following the success of ADIBO-fluorescein diacetate labeling on many LAP-fused proteins inside live cells, we wished to extend the method to other large fluorophores such as rhodamine and ATTO dyes. Initial labeling experiments using cyclooctyne conjugated to these fluorophores were challenging on two major fronts. First, the hydrophobic nature of these fluorophores made the probes stick to hydrophobic surfaces of proteins inside cells, causing very high background and prolonged washout time. Second, in addition to being hydrophobic, these fluorophores are often positively charged. The positive charge caused them to be attracted and eventually sequestered inside the mitochondrial matrix, which has a more negative electrical potential (21). Even by itself, without any additional conjugation, we found that ATTO 647N gave a high level of nonspecific cell staining, primarily in the mitochondria (data not shown). In each case, the time required to remove excess probe depended on the different physical properties of the fluorophore. Comparing to labeling with the relatively hydrophilic probe ADIBO-fluorescein diacetate, where excess probe in the background can be washed out quickly to see specific labeling, we observed that labeling performed with ADIBO conjugated to the red or far-red emitting fluorophores generally took from 3-9 hours to remove background in order to see specific labeling signal.

To alleviate the nonspecific hydrophobic sticking background, aside from frequent media change over a prolonged period of time, we found that using a less hydrophobic cyclooctyne such as MOFO rather than ADIBO, especially in conjugation with a very hydrophobic fluorophore such as ATTO 647N helped in lowering the background. Figure 3-10 shows a comparison of LAP-BFP-NLS labeling with ADIBO and MOFO, each conjugated to ATTO 647N. The graph on the bottom plots the specific labeling (nuclear ATTO signal) to the nonspecific labeling (cytosolic ATTO signal) for >50 cells for each probe. It can be seen that MOFO-ATTO 647N gives much more specific labeling than ADIBO-ATTO 647N, likely because the total hydrophobicity of the conjugate is reduced. This ultimately permitted us to perform ATTO 647N labeling of LAP-β-actin in live COS-7 cells as shown in a later section.

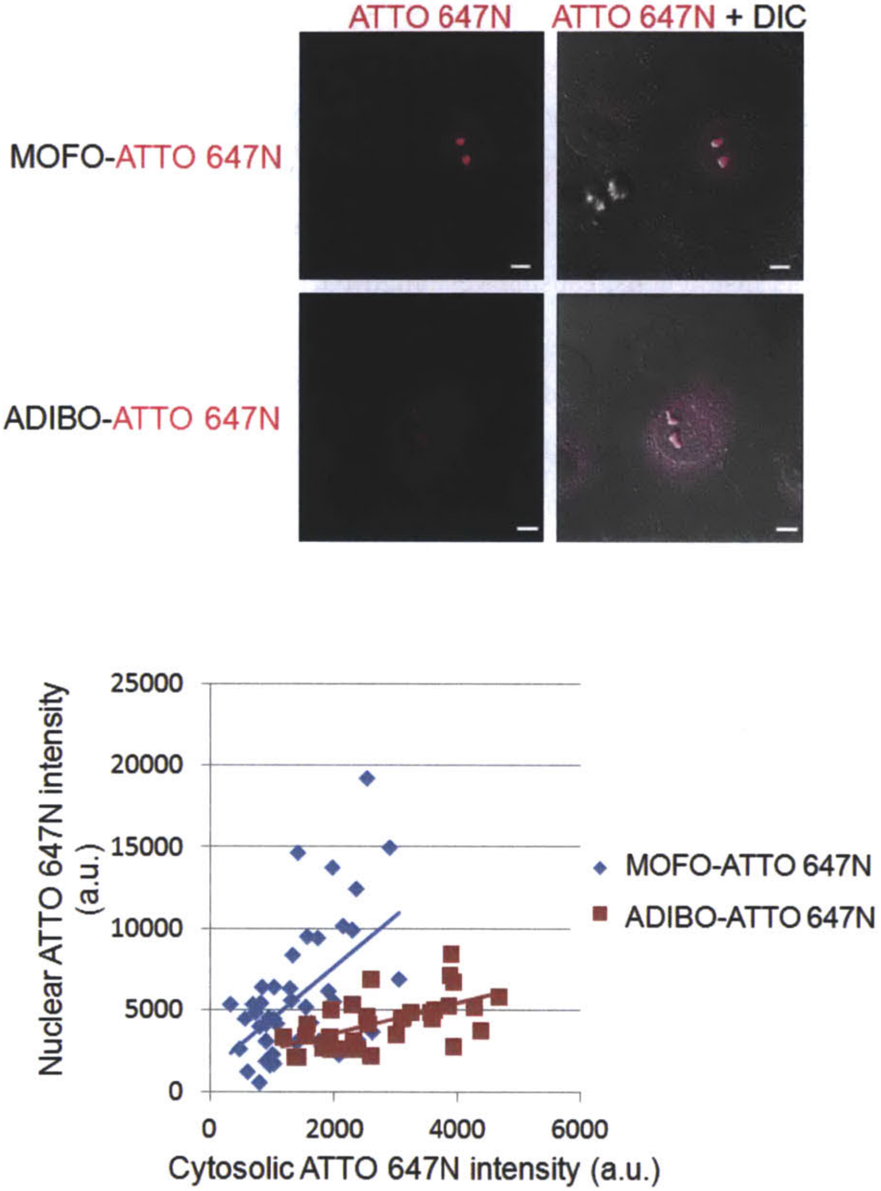


Figure 3-10. Comparison of MOFO-ATTO 647N and ADIBO-ATTO 647N nuclear labeling specificities. (Top) Nuclear labeling of LAP-BFP-NLS with MOFO- and ADIBO-ATTO 647N in COS-7 cells. Images are shown after 8 hours of ATTO 647N conjugates washout. (Bottom) The mean nuclear ATTO 647N intensity (representing specific signal) was plotted against the mean cytosolic ATTO 647N intensity (representing nonspecific signal), for the same cell, for >50 single cells for each condition.

Aside from pairing especially hydrophobic and positively charged fluorophores with less hydrophobic cyclooctynes, we also tested the effect of varying linker structures between MOFO and ATTO 647N in an attempt to further reduce the nonspecific labeling background. The conventional *N, N'*-dimethyl-1,6-hexanediamine (HDDA) linker that we used for most fluorophore conjugates was replaced by the more hydrophilic PEG linker (Figure 3-11). However, we did not observe any significant improvement in reducing cytosolic background or shortening washout time while using MOFO-PEG-ATTO 647N (Figure 3-11), suggesting that the degree of hydrophobicity of the cyclooctyne and fluorophore moieties dominated probe distribution inside cells.

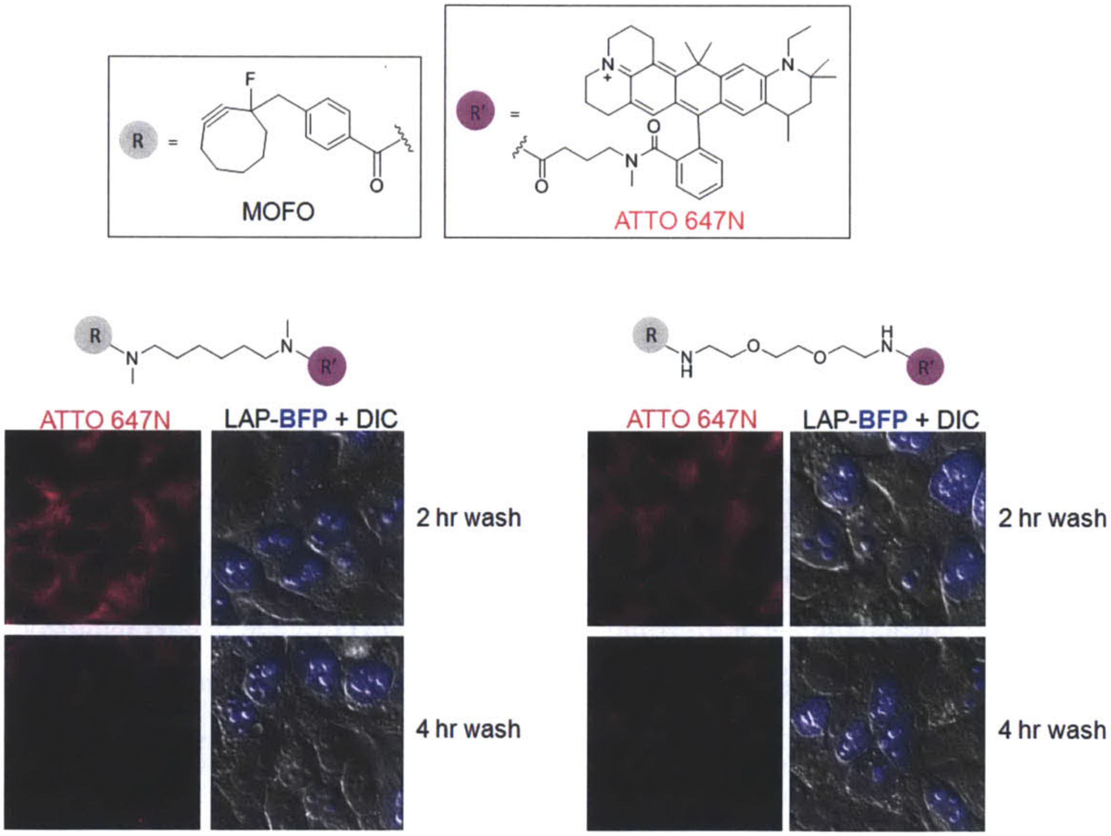


Figure 3-11. Comparison of two different linker structures for MOFO-ATTO 647N. Conjugates were synthesized with either an N,N'-dimethyl-1,6-hexanediamine linker (left) or a polyethylene glycol PEG linker (right). Labeling was performed on HEK 293T cells expressing LplA(W37I) and nuclear-localized LAP-BFP as described previously. Images are shown after the indicated fluorophore washout times. Specific labeling is

detectable, particularly in the nucleoli of transfected cells, but the images are dominated by the non-specific signal in these experiments.

The second major source of background fluorescence was from dyes trapped inside the mitochondrial matrix. Since the organelle has a negative inner-electrical potential established by the electron transport chain (21,22), positively charged fluorophores such as tetramethylrhodamine and ATTO 647N are especially attracted to the mitochondria. We incubated HEK 293T cells with ADIBO-ATTO 647N together with Mitotracker Green that confirmed the mitochondrial localization of the probe (data not shown). Compared to the hazy nonspecific background in the cytosol, the mitochondrial background appeared as very bright dots or tubular structures depending on the cell type. It was even more difficult to wash out because the dyes are sequestered in the organelle as long as the inner mitochondrial potential is maintained. Low levels of mitochondrial background fluorescence still remained after cells were fixed and permeabilized (data not shown). Moreover, mitochondrial trapping of the probe can reduce the effective concentration of the probe that is available elsewhere to react with azide-ligated LAP. To improve the signal to noise ratio of two-step fluorophore targeting using these hydrophobic and positively charged dyes, we explored two strategies to reduce mitochondrial sequestering of the dye. The first strategy is to apply a mitochondrial uncoupling agent that can temporarily disrupt the electrical potential across the mitochondrial inner membrane, so trapped dye can escape and be washed out. One routinely used uncoupling agent, *p*-trifluoromethoxy carbonyl cyanide phenyl hydrazone (FCCP), is a lipid-soluble weak acid that freely diffuses in and out of the mitochondrial matrix. It acts like a proton shuttle to disrupt the chemiosmotic gradient across the inner membrane without interrupting the electron transport chain (23).

To test whether using FCCP to dissipate the proton gradient could help in reducing mitochondrial trapping of positively charged fluorophore, we performed labeling on nuclear-localized LAP-BFP with either ADIBO-X-rhodamine or ADIBO-ATTO 647N as described in Figure 3-12. The ionophore FCCP was added 6 hours after probe incubation. The mitochondrial background fluorescence, which initially appeared as dots in HEK 293T cells, gradually lost its brightness and started to appear as blurred nonspecific background fluorescence in the cytosol. Cells were maintained in growth media containing FCCP for 2 hours to allow the fluorophore to escape from the mitochondrial matrix. During this time, we relied on the nonspecific ion transporters of the cell to pump out excess probe. Even though disrupting the mitochondrial potential aided the removal of some of the trapped probe, further dependence on the action of nonspecific transporters and prolonged washout time were still not sufficient to obtain a clean background for such hydrophobic and positively charged dyes (Figure 3-12).

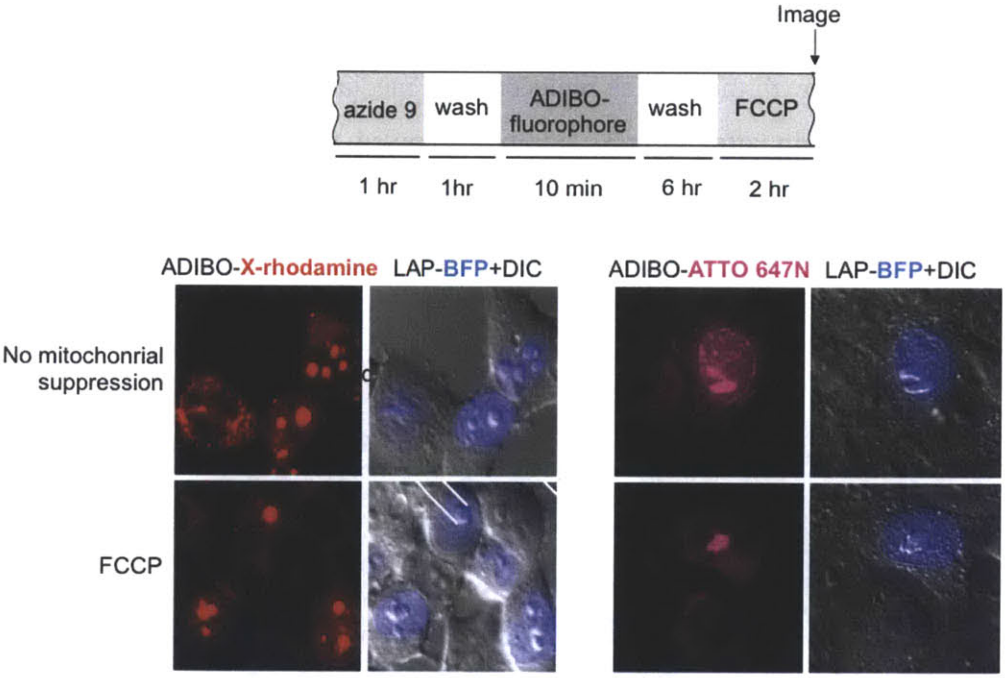


Image
azide 9
wash
ADIBO-fluorophore
wash
FCCP
1 hr
1hr
10 min
6 hr
2 hr
ADIBO-X-rhodamine
LAP-BFP+DIC
ADIBO-ATTO 647N
LAP-BFP+DIC
No mitochonrial suppression
FCCP

Figure 3-12. Suppression of mitochondrial background labeling with FCCP. Following labeling scheme shown on top, after azide 9 incubation and washout, HEK 293T cells expressing LplA(W37I) and LAP-BFP-NLS were labeled with either ADIBO-X-rhodamine (left) or ADIBO-ATTO 647N (right) for 10 minutes. Thereafter, cells were washed and incubated in fresh media. After 6 hours of probe washout, media was replaced with media containing FCCP (*p*-trifluoromethoxy carbonyl cyanide phenyl hydrazone, a mitochondrial uncoupler). Live cells were imaged at 2 hours after FCCP addition. Labeled cells with no FCCP addition were imaged at 8 hours after cyclooctyne-dye addition for comparison.

The second strategy to reduce mitochondrial sequestering of the probe is to apply small molecule quenchers. The two major factors that lead to fluorophores entering the mitochondria are hydrophobicity and positive charge, which are also properties of some small molecule fluorescence quencher dyes. We imagined that if the same mechanism applies, these quenchers would also be attracted to the mitochondria, co-localizing with the fluorophores trapped there, and efficiently quenching their fluorescence. We chose the quencher QSY21 because its absorption wavelength overlaps well with the emission wavelength of ATTO 647N at 660 nm. Labeling of nuclear-localized LAP-BFP in HEK 293T cells was carried out as described in Figure 3-13. At 6 hours after ADIBO-ATTO 647N washout, cell growth media was replaced with media containing QSY21. Live cell images taken at 10 min after QSY21 addition showed much reduced mitochondrial background fluorescence in cells incubated with QSY21.

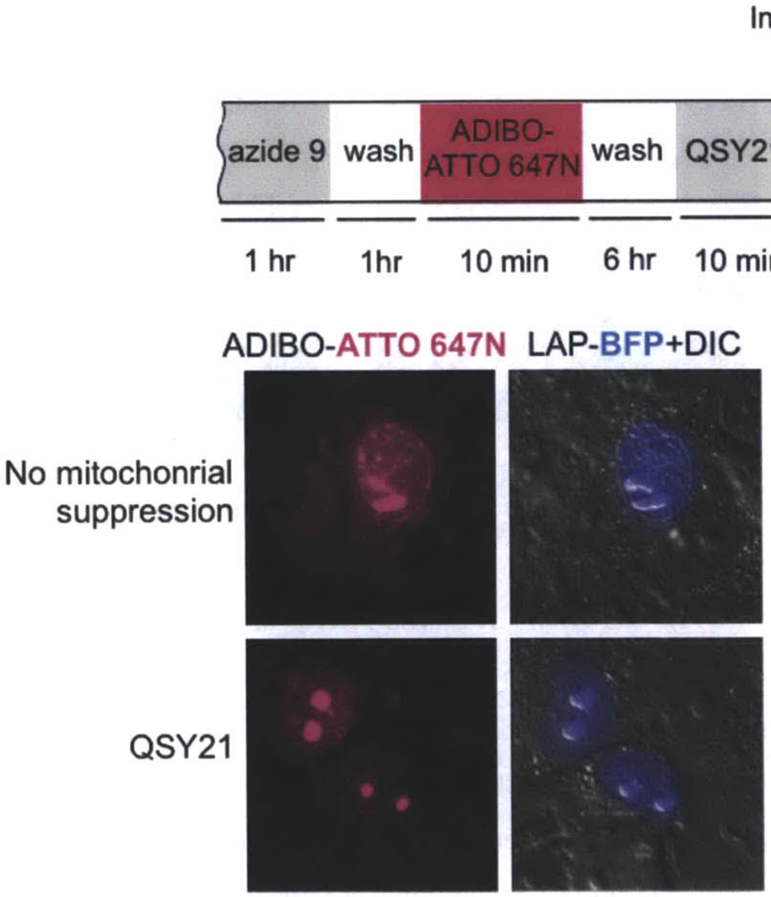


Image
azide 9
wash
ADIBO-ATTO 647N
wash
QSY21
1 hr
1hr
10 min
6 hr
10 min
ADIBO-ATTO 647N
LAP-BFP+DIC
No mitochonrial suppression
QSY21

Figure 3-13. Suppression of mitochondrial background labeling with QSY21. Following labeling scheme shown on top, after azide 9 incubation and washout, HEK 293T cells expressing LplA(W37I) and LAP-BFP-NLS were labeled with ADIBO-ATTO 647N for 10 minutes. Thereafter, ADIBO-ATTO 647N was removed from the cells, which were washed and incubated in fresh media. After 6 hours of probe washout, media was replaced with media containing 10 μM QSY21 for 10 minutes. Live cells were imaged after QSY21 incubation. Labeled cells with no QSY21 addition were imaged at 6 hours after ADIBO-ATTO 647N addition for comparison.

We also tested the combined effect of mitochondrial uncoupler FCCP and QSY21 quencher by first adding FCCP after ADIBO-ATTO 647N incubation to free some of the fluorophore trapped there and allow it to wash out over 2 hours in the presence of FCCP (Figure 3-14, top scheme). Afterwards, cell growth media containing FCCP was replaced with fresh growth media without FCCP to allow the electrical potential to recover over 1 hour. After the mitochondrial electrical potential has recovered, QSY21 was added to further quench the remaining fluorophore in the mitochondrial matrix. Using both FCCP and QSY21, we were able to increase the nuclear signal to cytosolic background ratio about two-fold (Figure 3-14 bottom).

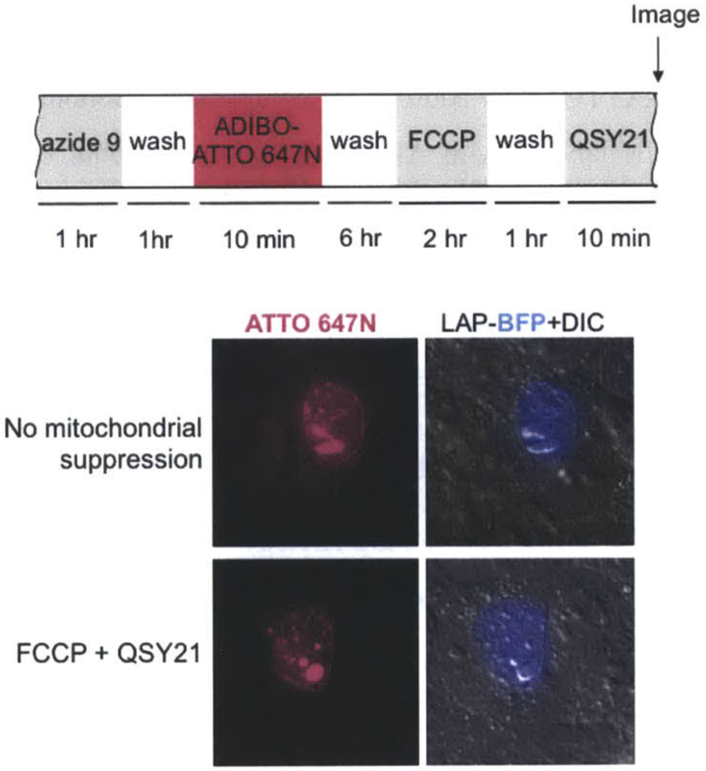


Image
azide 9
wash
ADIBO-ATTO 647N
wash
FCCP
wash
QSY21
1 hr
1hr
10 min
6 hr
2 hr
1 hr
10 min
ATTO 647N
LAP-BFP+DIC
No mitochondrial suppression
FCCP + QSY21

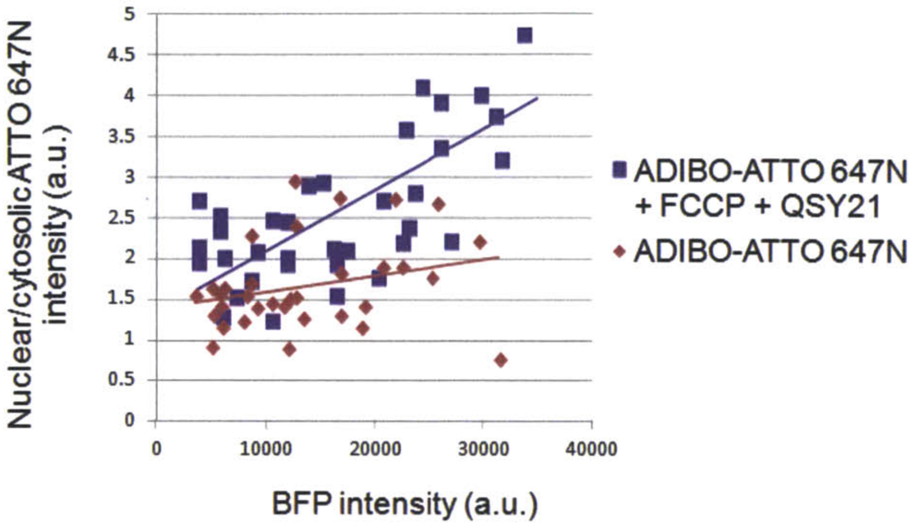


Figure 3-14. Suppression of mitochondrial background labeling with FCCP and QSY21. (Top) Following labeling scheme shown to test the combined effect of FCCP and QSY21, after azide 9 incubation and washout, HEK 293T cells expressing LplA(W37I) and LAP-BFP-NLS were labeled with ADIBO-ATTO 647N for 10 minutes. Thereafter, ADIBO-ATTO 647N was removed from the cells and replaced with complete media. After 6 hours of probe washout, cells were treated with mitochondrial uncoupler FCCP and quencher QSY21 sequentially as shown in the labeling scheme. Live cells were imaged after QSY21 incubation. Labeled cells with no QSY21 addition were imaged at 9 hours after ADIBO-ATTO 647N addition for comparison. (Bottom) Quantitation of data on top. For the cells labeled with ADIBO-ATTO 647N, with or without FCCP and QSY21 application to reduce background, the ratio of the mean nuclear fluorescein intensity (representing specific labeling) and the mean cytosolic fluorescein intensity (representing nonspecific labeling) was plotted against the mean nuclear BFP intensity (representing LAP-BFP-NLS expression level), for the same cell. More than 50 single cells were analyzed for each condition.

We applied the QSY21 quenching strategy in ADIBO-ATTO 647N labeling of LAP-tagged β–actin in live cells with the hope that we would acquire images with more clear and distinctive actin filaments without the background fluorescence from dyes in the cytosol. COS-7 cells transfected with LAP-β-actin and LplA(W37I) were labeled with azide 9 and ADIBO-ATTO 647N as described previously. Following 6 hours of ADIBO-ATTO 647N washout, cells were incubated with media containing 10 μM of QSY21 for 10 minutes and imaged live. However, we found that even though QSY21 was supposed to be trapped inside the mitochondria and only quench the background fluorescence there, some amount of QSY21 in the cytosol also reduced cytosolic LAP-β-actin specific labeling signal by about 2-fold comparing to conventional labeling with no quencher addition (Figure 3-15). If all the images were normalized with respect to the labeling performed without QSY21, ADIBO-ATTO 647N labeled LAP-β-actin pattern was not visible when QSY21 was added (Figure 3-15, first column). Even though QSY21 helped to reduce the mitochondrial background in ADIBO-ATTO 647N labeling of nuclear LAP-BFP without decreasing nuclear signal, in the more challenging case of LAP-β-actin labeling, the signal was very low and can be reduced by QSY21 quencher remaining in the cytosol. Since QSY21, like other similarly large, hydrophobic, and positively charged fluorophores, is not completely trapped in the mitochondria, trace amounts of it may still nonspecifically stick to the cytosol to quench cytosolic ADIBO-ATTO 647N signal. We decided this could not become a widely applicable strategy in reducing mitochondrial background especially when it also reduced LAP-specific signal. However, if the protein of interest were localized to the nucleus or if the LAP-specific signal were very robust, the strategy of quenching mitochondrial background using a hydrophobic and positively charged quencher whose absorption spectrum overlaps well with the emission spectrum of the fluorophore in use could be applied.

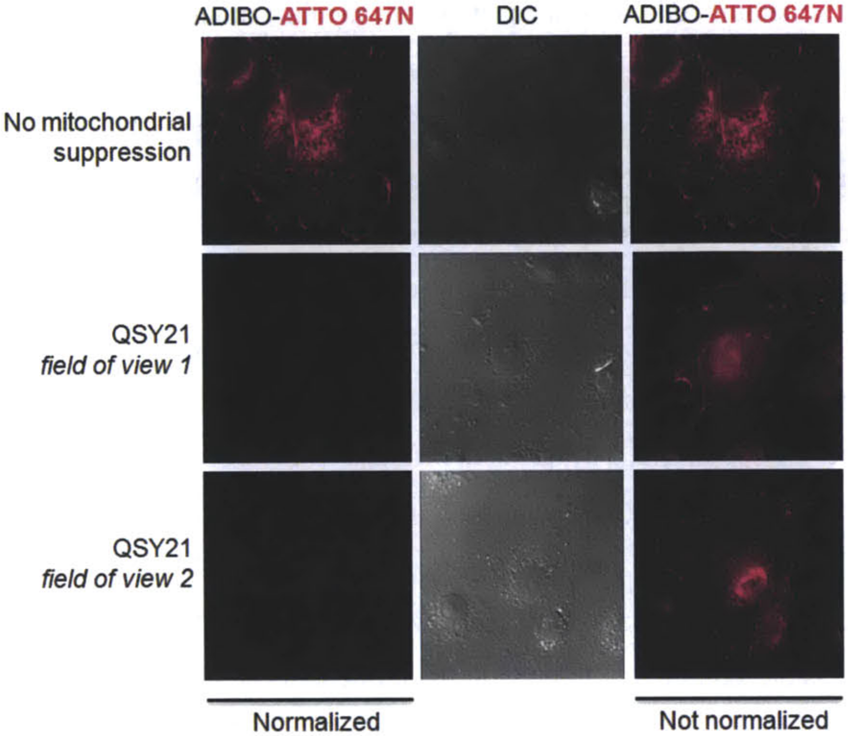

ADIBO-ATTO 647N
DIC
ADIBO-ATTO 647N
No mitochondrial suppression
QSY21 *field of view 1*
QSY21 *field of view 2*
Normalized
Not normalized

Figure 3-15. Comparison of ADIBO-ATTO 647N labeling of LAP-β-actin with and without suppression of mitochondrial background using QSY21 quencher. COS-7 cells expressing LplA(W37I) and LAP-β-actin were labeled with azide 9 and ADIBO-ATTO 647N. After 6 hours of ADIBO-ATTO 647N washout, cells were incubated with QSY21 quencher for 10 minutes. Live cells were imaged 10 min after QSY21 incubation. Labeled cells with no QSY21 addition were imaged at 6 hours after ADIBO-ATTO 647N incubation for comparison.

Extension of site-specific two-step labeling method to diverse fluorophore structures with labeling inside cells and on the cell surface

One of the advantages of the intracellular two-step fluorophore targeting technique is its potential to be generalizable towards any type of fluorophore as long as the fluorophore can be conjugated to a cyclooctyne and the conjugate is cell permeable. As a way to showcase the versatility of the method, in addition to fluorescein, we prepared ADIBO conjugates to tetramethylrhodamine (TMR), ATTO 647N, and ATTO 655. ADIBO-TMR and ADIBO-ATTO 655 both gave specific labeling (Figure 3-16), but ADIBO-ATTO 647N produced a high level of nonspecific binding. As shown in the previous section, we demonstrated that conjugating the more hydrophobic fluorophore ATTO 647N with the less hydrophobic cyclooctyne MOFO instead of ADIBO was able to provide more specific labeling. Therefore, we conjugated the hydrophobic fluorophores like X-rhodamine and ATTO 647N to MOFO.

Figure 3-16 shows live cell labeling of multiple LAP fusion proteins with a diverse palette of fluorophores spanning from fluorescein to ATTO 647N. ADIBO is used for the more hydrophilic dyes such as fluorescein, TMR and ATTO 655. DIBO, which is structurally similar to ADIBO, is used for Oregon Green. MOFO is used for the more hydrophobic dyes, X-rhodamine and ATTO 647N.

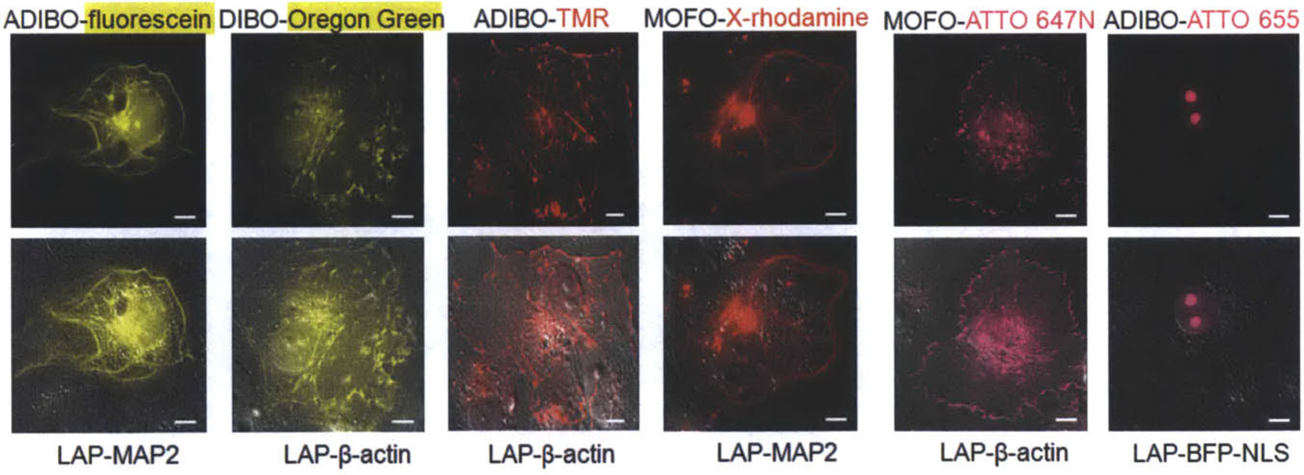
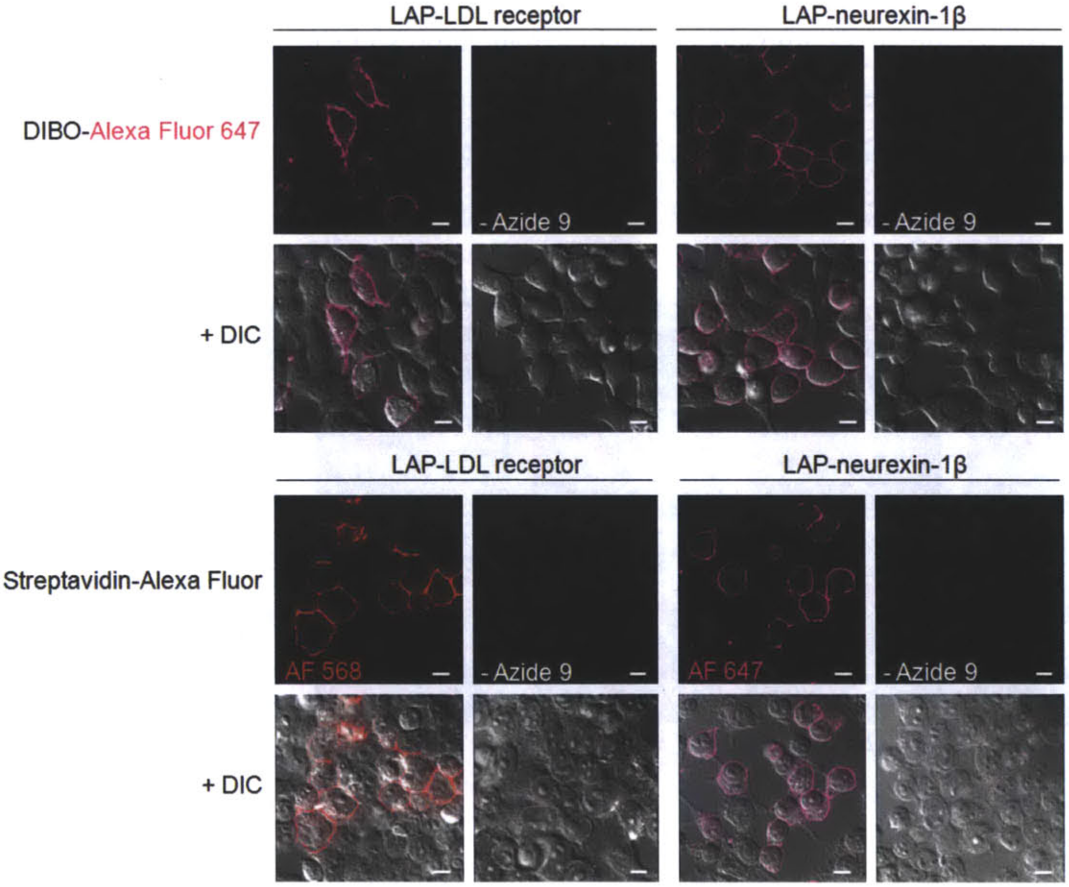


Figure 3-16. Intracellular protein labeling with diverse fluorophore structures. COS-7 cells coexpressing LplA(W37I) and the indicated LAP fusion protein (labeled across bottom) were labeled with azide 9, followed by the indicated cyclooctyne–fluorophore

conjugate (labeled across top). MOFO was used for the more hydrophobic fluorophores (X-rhodamine and ATTO 647N); ADIBO and DIBO were used for the others. Chemical structures are shown in Experimental Methods. TMR = tetramethylrhodamine.

After demonstrating two-step fluorophore ligation on various intracellular proteins, we also wished to perform labeling on cell surface proteins. We chose two cell surface proteins, LAP-tagged LDL receptor and neuronal adhesion protein neurexin -1β. After transfecting HEK 293T cells with each plasmid, we added purified LplA(W37I) protein, azide 9, and ATP to the cell medium for 20 minutes. Following azide ligation, azide ligated LAP-tagged proteins were derivatized with DIBO-Alexa Fluor 647, which is membrane-impermeant, or DIBO-biotin. Cells labeled with DIBO-biotin were visualized by staining with streptavidin-Alexa Fluor conjugates. Figure 3-17 (top and middle) shows both DIBO-Alexa Fluor 568 and DIBO-biotin label cell surface proteins with specific and high signal to background ratio.


LAP-LDL receptor
LAP-neurexin-1β
DIBO-Alexa Fluor 647
- Azide 9
- Azide 9
+ DIC
LAP-LDL receptor
LAP-neurexin-1β
Streptavidin-Alexa Fluor
AF 568
- Azide 9
AF 647
- Azide 9
+ DIC

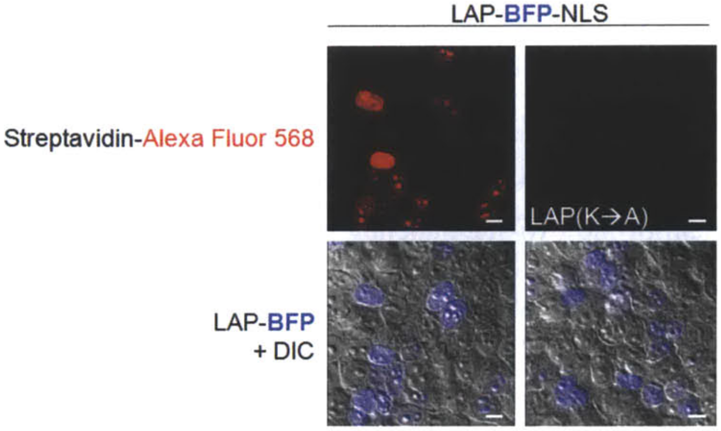


Figure 3-17. Cell surface and intracellular protein labeling with DIBO-Alexa Fluor 647 and DIBO-biotin (both commercially available from Life Technologies). (Top and middle) HEK 293T cells expressing LAP-LDL receptor (low density lipoprotein receptor) or LAP-neurexin-1β were labeled by first adding purified LplA(W37I) enzyme and azide 9 to the cell medium for 20 min. After 5 min of washing, DIBO-Alexa Fluor 647 or DIBO-biotin was added for 10 min. DIBO-biotin was detected by staining live cells with streptavidin-Alexa Fluor conjugate. Negative controls are shown with azide 9 omitted. (Bottom) DIBO-biotin is cell permeable and its labeling could also be detected on LAP-fused intracellular proteins. HEK 293T cells expressing LplA(W37I) and nuclear-targeted LAP-BFP-NLS were labeled live with azide 9 and then DIBO-biotin. Thereafter, cells were fixed and permeabilized and labeled proteins were detected by staining with streptavidin-Alexa Fluor 568 conjugate. A negative control is shown with an alanine mutation in LAP. All scale bars, 10 μm.

Since DIBO-biotin is membrane-permeant, it can also be applied in performing this labeling inside cells. However, unlike cell surface proteins that can be directly visualized using streptavidin-fluorophore conjugate, biotinylated intracellular LAP-tagged proteins can only be detected after membrane permeabilization. Figure 3-17 (bottom) shows intracellular labeling of azide-ligated LAP-BFP-NLS in HEK 293T cells using DIBO-biotin. After DIBO-biotin addition for 10 minutes, the cells were washed, fixed and permeabilized, then biotinylated LAP was detected with streptavidin-Alexa Fluor 568.

Measurement of overall two-step ligation yield in cells

We have demonstrated previously that azide 9 ligation by the LplA(W37I) enzyme inside live cells is close to quantitative. We were curious what percentage of azide 9-ligated LAP protein reacts with cyclooctyne inside cells. We used DIBO-biotin labeling to measure the overall yield of the two-step labeling method in live cells. After performing azide 9 and DIBO-biotin labeling in HEK 293T cells expressing LplA(W37I) and LAP-mCherry fusion proteins, cells were lysed, incubated with excess streptavidin protein to bind biotinylated LAP-mCherry. The lysate was analyzed by SDS-PAGE. In-gel mCherry fluorescence imaging in Figure 3-18 shows that LAP-mCherry runs at the expected molecular weight of 27 kDa in negative control samples in which azide 9 or streptavidin is omitted. However, lane 1 shows about 21% of LAP-mCherry shifted up to approximately 80 kDa, corresponding the combined molecular weight of LAP-mCherry (27 kDa) and streptavidin protein (56 kDa). This higher molecular weight band represents the population of LAP-mCherry that was ligated with azide 9, which subsequently reacted with DIBO-biotin. From the streptavidin gel-shift assay using DIBO-biotin to label intracellular LAP-tagged protein, we conclude that under our current labeling protocol, with azide 9 incubation for 1 hour, excess azide 9 removal for 1 hour, and cyclooctyne-fluorophore conjugate incubation for 10 minutes, the two-step labeling yield is approximately 20%.

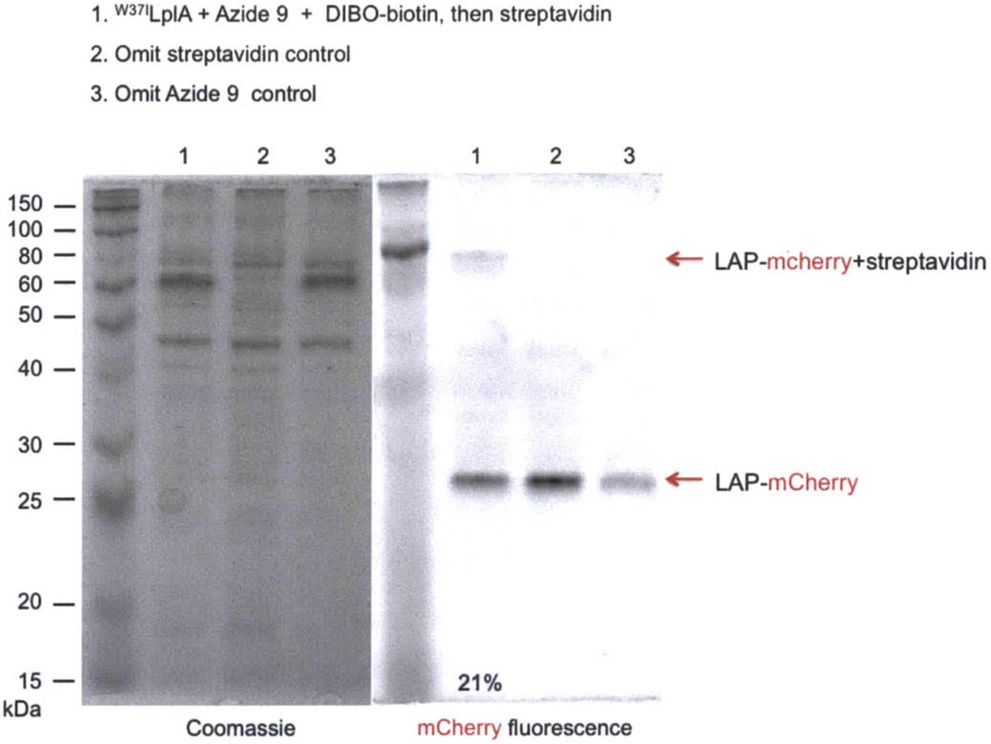

1. W37ILplA + Azide 9 + DIBO-biotin, then streptavidin
2. Omit streptavidin control
3. Omit Azide 9 control
1
2
3
1
2
3
150
100
80
60
50
40
30
25
20
15
kDa
LAP-mcherry+streptavidin
LAP-mCherry
21%
Coomassie
mCherry fluorescence

Figure 3-18. Streptavidin gel-shift analysis of two-step labeling yield in live cells. HEK 293T cells co-expressing LplA(W37I) and LAP-mCherry were labeled with azide 9 for 1 hour. Excess azide was washed out for 1 hour. Cell were then incubated with DIBO-biotin for 10 minutes. Excess DIBO-biotin was washed out for 2 hours. After cell lysis, excess streptavidin protein (56 kDa) was added to shift the molecular weight of biotinylated LAP-mCherry (27 kDa). Coomassie and mCherry fluorescence visualization are shown for lysates run on a 12% SDS-polyacrylamide gel. Lane 2 shows a negative control with streptavidin omitted. Lane 3 shows omission of azide 9. Estimated percent conversion to biotin-alkyne ligated LAP-mCherry product from LAP-mCherry was approximated using densitometry and given at the bottom of the mCherry fluorescence gel image.

Conclusion

We have developed an enzyme based labeling method that can target a diverse panel of fluorophores to a small 13-amino acid peptide tag LAP fused to proteins of interest inside living cells. The targeting is accomplished first by enzyme-mediated alkyl azide ligation and then by strain-promoted cycloaddition with a fluorophore-conjugated cyclooctyne. To develop the method, we systematically optimized the azide ligation reaction through screening of lipoic acid ligase mutants and alkyl azide variants. We then evaluated five different cyclooctynes structures differing in their reactivity, selectivity, hydrophobicity, and extent of nonspecific binding to cells, using a live-cell fluorescein targeting assay. Our final, optimized two-step labeling scheme was used to target a diverse panel of fluorophores with emission wavelengths spanning the visible spectrum, to a variety of LAP fusion proteins in multiple mammalian cell lines.

A few aspects that govern the choice of cyclooctyne are hydrophobicity, reactivity, and stability. Our comparison of cyclooctynes in live cells yielded observations that should prove useful even beyond the context of enzyme-mediated fluorophore targeting, due to the numerous and diverse applications to which cyclooctynes are being applied. In choosing a cyclooctyne for intracellular ligation, it is important that the probe has high reactivity, yet remains hydrophilic and inert towards endogenous cellular nucleophiles. One of the earliest cyclooctynes, MOFO (monofluorinated cyclooctyne) (14), performed well inside cells, giving signal to background ratios consistently >5:1 in the context of fluorescein targeting to nuclear LAP. However, its slow second-order rate constant (4.3×10^{-3} $M^{-1}s^{-1}$) (14) was a severe drawback. DIFO (di-fluorinated cyclooctyne) (15) was developed in attempt to increase the kinetics of cyclooctynes. However, we have seen that in the case of DIFO, although the probe is quite reactive, its nonspecific reactions inside cells led to a permanent background. The next generation cyclooctynes, fused to two benzene rings, increased ring strain and hence the second-order rate constants. Not surprisingly, we found both cyclooctynes, ADIBO (19) and DIBO (18), gave about 4-fold higher absolute signal in cells, compared to MOFO, probably due to increased yield of the cycloaddition product. ADIBO and DIBO also labeled the LAP-tagged cytoskeletal proteins actin and MAP2 efficiently. However, the dibenzo rings that confer the enhanced reactivity for such cyclooctynes also made them more hydrophobic.

The increase in signal was accompanied by an increase in background, likely due to nonspecific binding of these hydrophobic probes. Consequently, the signal to background *ratios* were comparable for ADIBO-, DIBO-, and MOFO-fluorescein conjugates.

When we extended the cyclooctyne comparison to other fluorophores, we found that ADIBO and DIBO conjugates to relatively well-behaved hydrophilic fluorophores such as fluorescein and Oregon Green gave satisfactory labeling, but when we tried to target very hydrophobic fluorophores such as ATTO 647N, the combined hydrophobicity of the dye and the cyclooctyne (ADIBO) precluded successful labeling, due to high nonspecific binding. This was alleviated by using the less hydrophobic MOFO instead. We were able to label LAP-tagged β-actin with ATTO 647N conjugated to MOFO. To reduce the nonspecific and mitochondrial background from large, hydrophobic, and positively charged fluorophores such as ATTO 647N, we explored strategies utilizing the mitochondrial uncoupling agent FCCP that disrupts the chemiosmotic gradient across the mitochondrial inner membrane, as well as the fluorescence quencher QSY21 to reduce mitochondrial-specific background, but these approaches only led to minimal reduction of background. In the case of LAP-β-actin labeling with ATTO 647N, QSY21 quencher remaining in the cytosol also quenched labeling signal. Our study illustrates the need for new cell-permeable, bright, and hydrophilic fluorophore, as well as the design and synthesis of cyclooctyne probes that combine high reactivity and stability with low hydrophobicity.

For some applications, however, prolonged washout of unbound probe or using a fluorescence quencher to reduce background is sometimes impossible, or the remaining probe can further frustrate detection. To minimize the amount of excess probe, new and improved bioorthogonal ligation chemistries that have higher reactivity than strain-promoted azide-alkyne cycloaddition open up alternative possibilities that offer high labeling sensitivity and yield, even at low probe concentrations. ADIBO and DIBO have been reported to react with nitrones at a rate ranging from 3 to 5 magnitudes faster than with azide (24,25). Cycloaddition reaction with nitrones have been applied in protein functionalization *in vitro* (25), attaching antibodies to surface of hybrid multifunctional nanoparticles (26), and labeling EGF receptors on the cell surface of human breast cancer cells (27). However, the intracellular behavior of nitrone and this fast cycloaddition reaction have yet to be tested and characterized.

Another reaction involving the Diels-Alder cycloaddition between tetrazine and *trans*-cyclootene has reported an explosive second-order rate constant of 2000 $M^{-1}s^{-1}$ (28), several magnitudes faster than strain-promoted azide-alkyne cycloaddition. Daniel Liu and Anupong Tangpeerachaikul from our lab have demonstrated in a separate work using LplA-based probe ligation in combination with this inverse Diels-Alder cycloaddition to target fluorophores to intracellular and cell surface proteins in live cells (29). The very fast cycloaddition kinetics yields substantial improvements in signal to background ratio for protein labeling. Furthermore, the ability of tetrazine to quench fluorescence of conjugated fluorophore until reaction with *trans*-cyclooctene enables the reaction to be fluorogenic, a very desirable feature of bioorthogonal reactions. A fluorogenic bioorthogonal reaction is essentially a background-free labeling system that would yield a detectable signal only after a chemoselective reaction occurs. A fluorogenic system would eliminate the time-consuming washout steps to remove excess reagents and would allow high concentration of reagent to be used if needed without worrying about increasing background. The inverse Diels-Alder cycloaddition has been applied in orthogonal labeling of cell surface glycans with the click reaction between dibenzocyclooctyne and azide (10), *in vitro* labeling of purified proteins that are genetically modified through the UAA method with trans-cyclooctene (30), as well as on the cell surface of cancer cells (31). It would be exciting to see demonstration of fluorogenic labeling inside live cells, where the background signal caused by hydrophobic fluorophore sticking would be quenched by tetrazine.

Another interesting advance is in copper-catalyzed click chemistry. Previously discounted for cellular applications due to copper toxicity, new improvements in copper ligand design and reactive oxygen species scavenging have made it possible to perform click chemistry on live cell surfaces and even animals (32). If the toxicity can be further reduced, while preserving the fast kinetics of ligation (currently 104–107 fold greater than strain-promoted cycloaddition), then copper-catalyzed click chemistry will be competitive with other methods for bioorthogonal derivatization on the cell surface.

Furthermore, aside from utilizing bioorthogonal chemistries to target large fluorophores, postdoc Wenjing Wang in our lab is working towards evolving LplA to accept and ligate various large fluorophores directly in a single step. This challenging approach involves engineering an "exit tunnel" in LplA for large fluorophore structures so they do not have to directly interact with LplA or fit into its binding pocket. This method could lead to an efficient labeling technique that shortens labeling time and possibly gives higher ligation yield than what we achieved using strain-promoted cycloaddition. To make the enzyme recognize and ligate not a single large fluorophore but applicable towards, for example, all rhodamine based fluorophores, yet maintain the specificity towards LAP substrate, Wenjing is currently using yeast-evolution to evolve a novel LplA mutant with an exit tunnel to accommodate large fluorophores.

Several fluorophores targeted using LplA and strain-promoted cycloaddition in this study have exemplary properties that make them attractive alternatives to fluorescent proteins. X-rhodamine is a bright and photostable fluorophore commonly used for speckle imaging of actin (33). ATTO dyes are some of the best fluorophores in terms of their superior photophysical properties that make them suitable for both STED (stimulated emission depletion) (34,35) and STORM-type (36) super-resolution microscopies. On the cell surface, we targeted Alexa Fluor 647, an excellent fluorophore that has been used for countless ensemble and single-molecule imaging experiments. If methods can be developed to deliver sulfonated fluorophores - which include the cyanine dyes and Alexa Fluors - across cell membranes (37), then these too should be targetable to specific intracellular proteins using PRIME.

In conclusion, we have extended the PRIME technology to target larger fluorophores with superior photophysical properties than coumarin, via strain-promoted azide-alkyne cycloaddition. Considered in the context of other protein labeling methods (38), the disadvantages of the approach presented here are the requirement for co-expression of the LplA labeling enzyme, the unavoidable background caused by nonspecific binding of cyclooctyne–fluorophore conjugates (albeit low in the case of hydrophilic fluorophores such as fluorescein and Oregon Green), and the signal which is fundamentally limited by the kinetics of strain-promoted cycloaddition chemistry. Given these factors, the methodology will be most useful as a nontoxic (in contrast to FlAsH

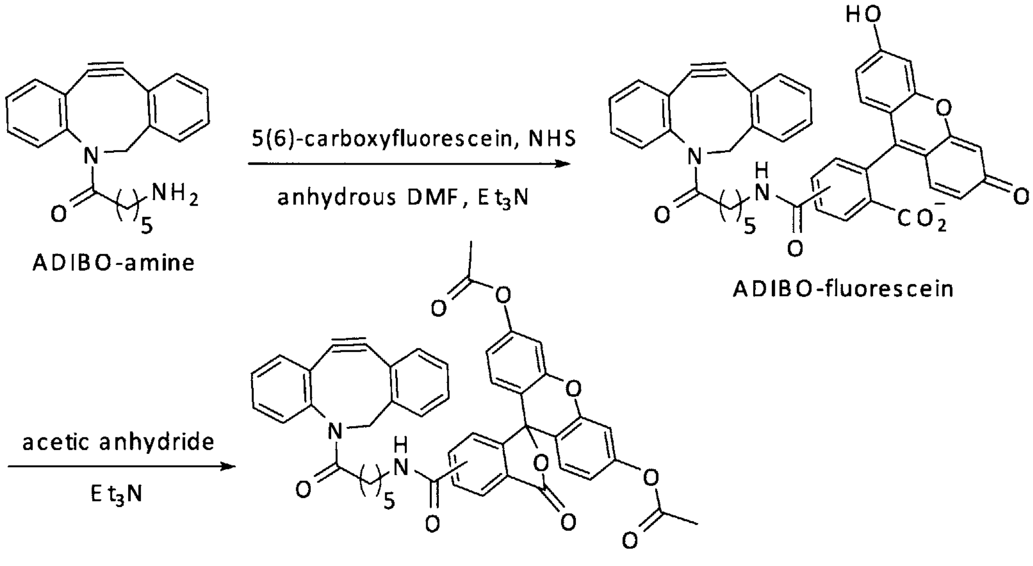
(39)) labeling method for abundant proteins, whose fusion to large tags (such as fluorescent proteins, HaloTag (40), or SNAP tag (41) perturb function, such as actin. We hope to offer insights in choosing the appropriate cyclooctynes and fluorophores for specific applications. With the development of new ligation methods and superior fluorophores, we anticipate the routine labeling and imaging of proteins with small-molecule fluorescent probes inside living cells.

Experimental methods

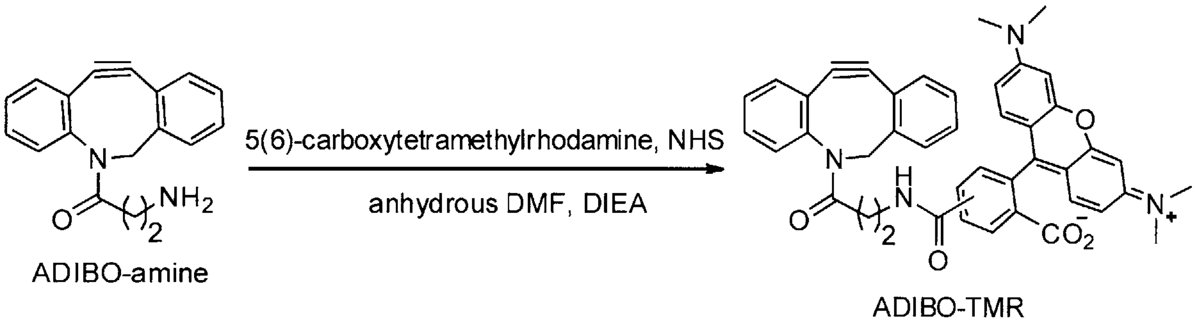
General synthetic methods

All reagents were the highest grade available and purchased from Sigma-Aldrich, Anaspec, Thermal Scientific, TCI America, Alfa Aesar, or Life Technologies and used without further purification. Anhydrous solvents were drawn from Sigma-Aldrich SureSeal bottles. Analytical thin layer chromatography was performed on 0.25 mm silica gel 60 F254 plates and visualized under short or long wavelength UV light, or after staining with bromocresol green or ninhydrin. Flash column chromatography was carried out using silica gel (ICN SiliTech 32-63D). Mass spectrometric analysis was performed on an Applied Biosystems 200 QTRAP mass spectrometer using electrospray ionization. HPLC analysis and purification were performed on a Varian Prostar Instrument equipped with a photo-diode-array detector. A reverse-phase Microsorb-MV 300 C18 column (250 × 4.6 mm dimension) was used for analytical HPLC. NMR spectra were recorded on a Bruker AVANCE 400 MHz instrument.

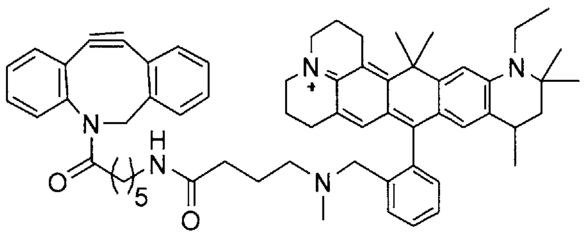
Synthesis of ADIBO- and DIBO-fluorophore conjugates



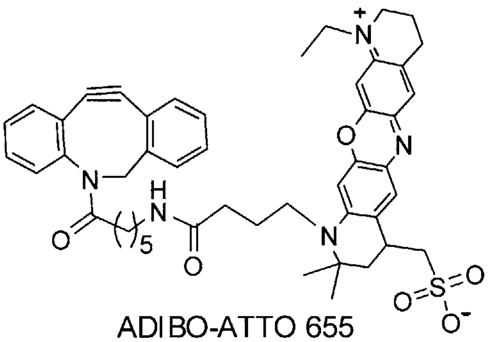
ADIBO-fluorescein diacetate The synthesis of aza-dibenzocyclooctyne-amine (ADIBO-amine) has been previously described.[1] To a solution of ADIBO-amine (3 mg, 9 μmol) in anhydrous DMF (500 μL) was added triethylamine (Et_3N, 3.8 μL, 27 μmol) and 5,6-carboxyfluorescein succinimidyl ester (NHS) (9.9 μmol, AnaSpec). The reaction was allowed to proceed for 10 hr at room temperature. Solvent was then removed under reduced pressure. The residue was subsequently dissolved in acetic anhydride (200 μL, 2.1 mmol) and allowed to stir for 30 min at room temperature. The color of the solution changed from bright yellow to colorless during the course of the reaction. Excess acetic anhydride was removed under reduced pressure. The resulting residue was purified by silica gel chromatography using 0-5% methanol in dichloromethane to afford ADIBO-fluorescein diacetate (R_f = 0.5 in 10% methanol in dichloromethane). Purified product was analyzed on HPLC which showed single peak with absorbance at 210 nm. Estimated yield for two steps is ~60%. ESI-MS for ADIBO-fluorescein diacetate: calculated for $[M+H]^+$: 761.24; observed 760.74. We note that the 5- and 6-regioisomers of fluorescein derivatives could have different cell entry and washout properties, which were not investigated in this study.



ADIBO-TMR To a solution of ADIBO-amine (70 mg, 0.46 mmol) in anhydrous DMF (2 mL) was added N,N-diisopropylethylamine (DIEA, 0.12 mL, 0.69 mmol) and 5,6-carboxytetramethylrhodamine (TMR) succinimidyl ester (NHS) (100 mg, 0.23 mmol, Sigma-Aldrich). The reaction was allowed to proceed for 12 hr at room temperature. Solvent was then removed under reduced pressure. Conjugate was purified by silica gel chromatography using 0-2% methanol in chloroform to provide dark red crystalline. The purity of the product was checked by HPLC. ESI-MS for ADIBO-TMR: calculated $[M+H]^+$: 688.27; observed 688.8. We note that the 5- and 6-regioisomers of TMR derivatives could have different cell entry and washout properties, which were not investigated in this study.

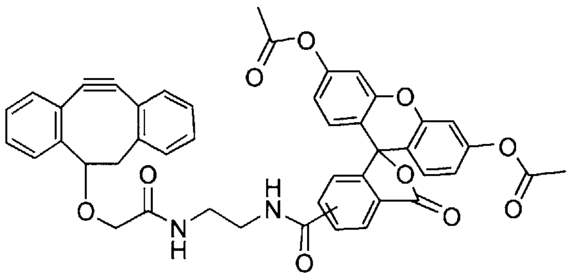


ADIBO-ATTO 647N



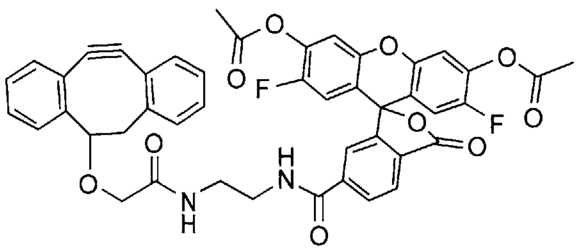
ADIBO-ATTO 655

ADIBO-ATTO 647N, ADIBO-ATTO 655 ADIBO conjugates to ATTO 647N and ATTO 655 were synthesized in a similar manner from ADIBO-amine.[1] ATTO 647 NHS ester (Sigma-Aldrich) and ATTO 655 NHS ester (Sigma Aldrich) were used. Conjugates were purified by silica gel chromatography using 0-2% methanol in dichloromethane. ESI-MS for ADIBO-ATTO 647N: calculated $[M+H]^+$: 946.56; observed 946.29. ESI-MS for ADIBO-ATTO 655: calculated $[M+H]^+$: 827.34; observed 827.51.



DIBO-fluorescein diacetate

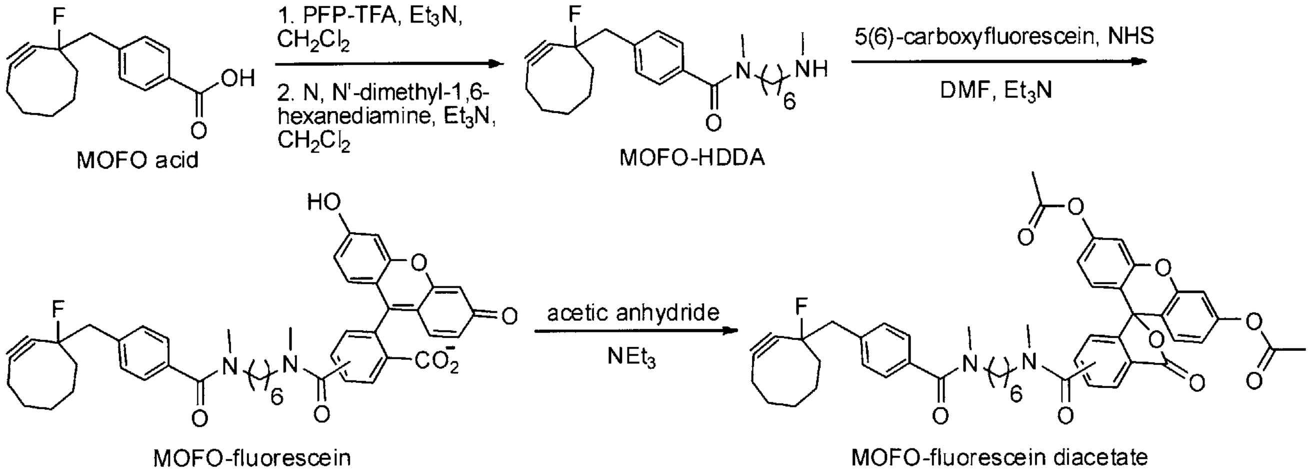
DIBO-fluorescein diacetate DIBO-fluorescein diacetate was synthesized in an analogous manner to ADIBO-fluorescein diacetate, from commercial DIBO-amine (Invitrogen) and fluorescein NHS ester (AnaSpec). The conjugate was purified by silica gel chromatography using 0-5% methanol in dichloromethane. ESI-MS for DIBO-fluorescein diacetate: calculated $[M+H]^+$: 763.22; observed 763.86.



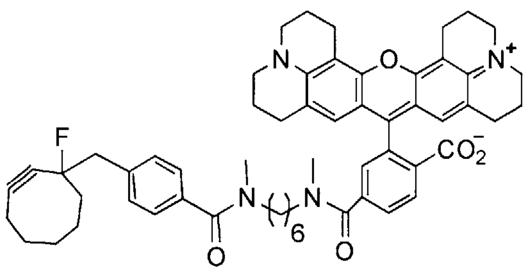
DIBO-Oregon Green 488 diacetate

DIBO-Oregon Green diacetate DIBO-Oregon Green 488 diacetate was a gift from Kyle Gee (Life Technologies).

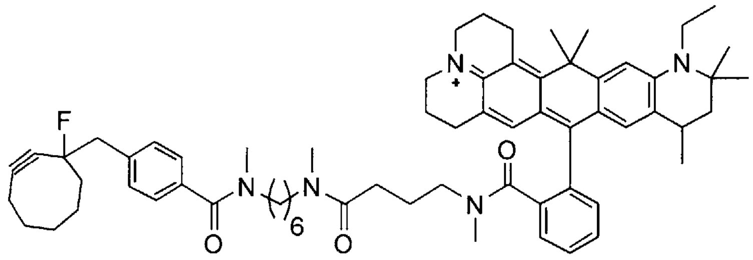
Synthesis of MOFO-, DIMAC-, and DIFO-fluorophore conjugates


1. PFP-TFA, Et3N, CH2Cl2
2. N, N'-dimethyl-1,6-hexanediamine, Et3N, CH2Cl2
MOFO acid
MOFO-HDDA
5(6)-carboxyfluorescein, NHS
DMF, Et3N
acetic anhydride
NEt3
MOFO-fluorescein
MOFO-fluorescein diacetate

MOFO-fluorescein diacetate To a solution of MOFO cyclooctyne acid[2] (5 mg, 19 μmol) in 500 μL anhydrous dichloromethane was added pentafluorophenyl trifluoroacetate (PFP-TFA, 9.8 μL, 57 μmol) and Et_3N (8 μL, 57 μmol). The reaction was allowed to proceed for 2 hr at room temperature. N, N'-dimethyl-1,6-hexanediamine (HDDA, 114 μmol) was then added to the reaction mixture, which was allowed to stir for 5 hr at room temperature. Solvent was removed under reduced pressure. The reaction mixture was purified by silica gel chromatography (10-15% methanol in dichloromethane) to afford MOFO-N,N'-dimethyl-1,6-hexanediamine (MOFO-HDDA). MOFO-HDDA was dissolved in anhydrous DMF (300 μL), and 5(6)-carboxyfluorescein, succinimidyl ester (9.8 mg, 20.9 μmol) and Et_3N (8 μL, 57 μmol) were added to the mixture, which was allowed to stir for 10 hr at room temperature. Solvent was removed under reduced pressure. The residue was dissolved in a small amount of acetic anhydride (<200 μL) and allowed to stir for 30 min at room temperature. After removal of acetic anhydride under reduced pressure, the reaction mixture was purified by silica gel chromatography (solvent gradient 0-5% methanol in dichloromethane) to afford MOFO-fluorescein diacetate (R_f = 0.4 in 10% methanol in dichloromethane). Estimated overall yield for four steps, 30-40%. ESI-MS for MOFO-fluorescein diacetate: calculated $[M+H]^+$: 829.34; observed 829.44.

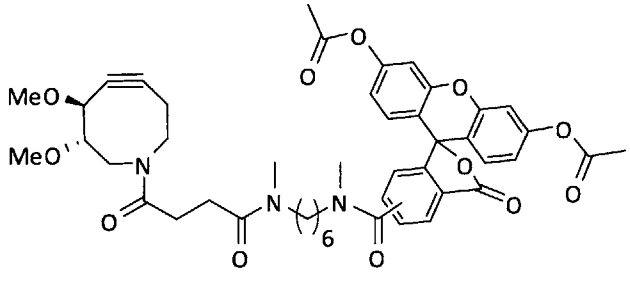


MOFO-X-rhodamine

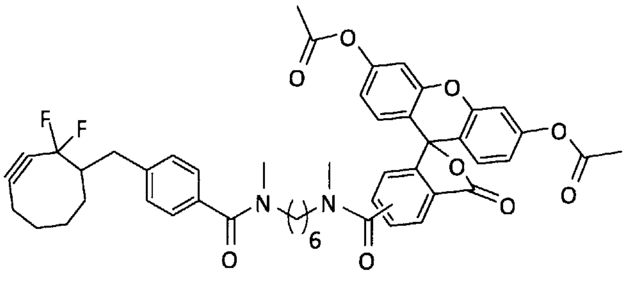


MOFO-ATTO 647N

MOFO-X-rhodamine, MOFO-ATTO 647N MOFO-HDDA was synthesized as described above, then conjugated to 5(6)-X-rhodamine NHS ester (Anaspec, 5(6)-ROX, SE) or ATTO 647N NHS ester (Sigma-Aldrich). Conjugates were purified by silica gel chromatography using 0-5% methanol in dichloromethane for MOFO-X-rhodamine and 0-2% methanol in dichloromethane for MOFO-ATTO 647N. ESI-MS for MOFO-X-rhodamine: calculated $[M+H]^+$: 903.49; observed 903.72. ESI-MS for MOFO-ATTO 647N: calculated $[M+H]^+$: 1014.66; observed 1014.42.



DIMAC-fluorescein diacetate



DIFO-fluorescein diacetate

DIMAC-fluorescein diacetate, DIFO-fluorescein diacetate Fluorescein diacetate conjugates to DIMAC[3] and DIFO[4] were synthesized in a similar manner from their respective acids. Conjugates were purified by silica gel chromatography using 0-5% methanol in dichloromethane. ESI-MS for DIMAC-fluorescein diacetate: calculated $[M-H]^-$: 752.33; observed 752.40. ESI-MS for DIFO-fluorescein diacetate: calculated $[M+H]^+$: 847.33; observed 847.26.

Plasmids

For bacterial expression of LplA, we used His$_6$-LplA in pYFJ16.[5] For mammalian expression of LplA, we used His$_6$-FLAG-LplA in pcDNA3.[5] For mammalian expression of LAP fusion proteins, we used LAP-β-actin and LAP-MAP2 in Clontech vector,[5] LAP-LDL receptor in pcDNA4, and LAP-neurexin-1β in pNICE. LAP-BFP expression constructs (LAP-BFP, LAP-BFP-NLS, LAP-BFP-CAAX, and LAP-BFP-NES) in pcDNA3 and LAP-mCherry in pcDNA3 were generated from corresponding pcDNA3-LAP-YFP plasmids[5] by replacing YFP with BFP or mCherry, using the BamHI and EcoRI restriction sites. All LplA and LAP point mutants were prepared via QuikChange site-directed mutagenesis. Complete sequences of plasmids used in this study are available at http://stellar.mit.edu/S/project/tinglabreagents/r02/materials.html (Ting lab website).

Immunofluorescence staining of LplA (Figure 3-4)

After live cell imaging, cells were fixed with 3.7% formaldehyde in Dulbecco's Phosphate Buffered Saline (DPBS) pH 7.4 for 10 min at room temperature followed by cold precipitation with methanol for 5 min at -20 ºC, then blocked with 3% BSA in DPBS for 1 hr at room temperature. To visualize FLAG-tagged LplA, cells were incubated with 4 μg/mL mouse monoclonal anti-FLAG antibody (Sigma-Aldrich) in 1% BSA in DPBS for 1 hr at room temperature. Cells were further washed and incubated with 4 μg/mL goat anti-mouse IgG antibody conjugated to Alexa Fluor 568 (Life Technologies) in 1% BSA in DPBS for 1 hr at room temperature, then washed and imaged.

Cell fixation after live cell DIMAC-fluorescein and DIFO-fluorescein labeling (Figure 3-6)

After live cell imaging, cells were fixed with 3.7% formaldehyde in DPBS pH 7.4 for 10 min at room temperature followed by cold precipitation with methanol for 5 min at -20 °C. Cells were then washed with DPBS several times over 10 min, before imaging.

Cell surface and intracellular labeling with commercial DIBO conjugates (Figure 3-17)

DIBO-Alexa Fluor 647 cell surface labeling (Figure 3-17 top)

HEK cells plated on glass coverslips in wells of a 48-well cell culture plate (0.95 cm^2 per well) were transfected with 100 ng pcDNA4-LAP-LDL receptor or 400 ng pNICE-LAP-neurexin-1β using Lipofectamine 2000. At 18 hr after transfection, cells were washed three times with MEM. Enzymatic ligation of azide 9 on the cell surface was performed in MEM with 5 μM ^{W37I}LplA, 500 μM azide 9, 2 mM ATP and 2 mM magnesium acetate for 20 min at room temperature (to minimize internalization of cell-surface proteins). After washing three times with MEM, cells were incubated with 10 μM DIBO-Alexa Fluor 647 in MEM for 10 min at room temperature. Cells were then washed three times with MEM and imaged live.

DIBO-biotin cell surface and intracellular labeling (Figure 3-17 middle and bottom)

DIBO-biotin cell surface labeling was performed in the same manner as DIBO-Alexa Fluor 647 cell surface labeling, described above. After DIBO-biotin incubation, cells were washed three times with DPBS and fixed with 3.7% formaldehyde in DPBS pH 7.4 for 10 min at room temperature, followed by cold precipitation with methanol for 5 min at -20 °C. Fixed cells were then blocked with 1% casein in DPBS for 1 hr at room temperature. To visualize specific labeling, cells were stained with streptavidin conjugated to Alexa Fluor 568 or Alexa Fluor 647 in 0.5% casein in DPBS for 5 min at room temperature, followed by washing three times with DPBS and imaging.

For DIBO-biotin intracellular labeling, HEK cells plated on glass coverslips in wells of a 48-well cell culture plate (0.95 cm^2 per well) were transfected with 400 ng pcDNA3-LAP-BFP-NLS and 200 ng pcDNA3-^{W37I}LplA. Azide 9 labeling/washout and DIBO-biotin labeling/washout were performed in the same manner as in Figure 3-18. After DIBO-biotin washout, cells were fixed and stained with streptavidin-Alexa Fluor 568, as described above.

Analysis of overall two-step ligation yield after strain-promoted cycloaddition in live cells (Figure 3-18)

HEK cells plated into wells of a 12-well culture plate (4 cm^2 per well) were transfected with 500 ng pcDNA3-^{W37I}LplA and 1000 ng pcDNA3-LAP-mCherry using Lipofectamine 2000 (Life Technologies). Azide 9 labeling and washout were performed in the same manner as in Figure 2-8. After excess azide 9 washout, cells were incubated in MEM containing 10 μM DIBO-biotin (Life Technologies) for 10 min at 37 °C. Thereafter, cells were further washed for 2.5 hr to remove excess DIBO-biotin. Cells were then harvested and lysed in the same manner as in Figure 2-8. The cell lysate was incubated with 5 μM of streptavidin for 1 hr at 4 °C, then analyzed on a 12% SDS-polyacrylamide gel at constant 200 V, under conditions known to preserve biotin-streptavidin binding as well as streptavidin's subunit association.[6] Prior to Coomassie staining, in-gel fluorescence of mCherry was visualized on FUJIFILM FLA-9000 instrument using SHG532 laser and Long Pass Green (LPG) filter.

Quantitative analysis of MOFO-fluorescein labeling of LAP-BFP using four LplA mutant/azide substrate pairs (Figure 3-2)

Cells with fluorescein signal at least 2-fold greater than autofluorescence from untransfected cells, and BFP signal at least 5-fold greater than autofluorescence were selected by hand for analysis. For each of these cells, the entire area of the cell representing signal was circled. SlideBook was used to calculate the mean intensities in both channels. The background-corrected mean fluorescein intensity was plotted against the background-corrected mean BFP intensity using Excel.

Quantitative analysis of LplA mutant expression levels in cells (Figure 3-4)

Cells with Alexa Fluor 568 signal at least 1.5-fold great than background (area without any cell) were selected by hand for analysis. For each of these cells, the entire area of the cell representing signal was circled. SlideBook was used to calculate the mean intensities in the channel. The background-corrected mean Alexa Fluor 568 intensity was plotted using Excel.

Quantitative analysis of fluorophore-cyclooctyne labeling specificity (Figures 3-5 B and 3-10)

Cells with signal at least 3-fold greater than autofluorescence from untransfected cells in the cyclooctyne channel were selected by hand for analysis. For each of these cells, one region in the cytosol (representing background) and one region in the nucleus (representing specific signal) were manually circled. The background-corrected mean fluorescence intensity was determined for both regions using SlideBook. Excel was used to plot the nuclear versus cytosolic fluorescence intensity for each cell. Since ATTO 647N labeling signal was low, we selected for analysis cells with signal at least 2-fold greater than autofluoresence from untransfected cells in the ATTO 647N channel.

Other protocols

LplA and mutants were expressed and purified as previously described.[5] The 13-amino acid LAP peptide (H_2N-GFEID$\underline{K}$VWYDLDA-CO_2H)[7] was synthesized by the Tufts University Peptide Synthesis Core Facility and purified to >96% homogeneity.

References

1. Manova RK, Pujari SP, Weijers CAGM, Zuilhof H, Van Beek TA. Copper-Free Click Biofunctionalization of Silicon Nitride Surfaces via Strain-Promoted Alkyne–Azide Cycloaddition Reactions. Langmuir. 2012 Jun 12;28(23):8651–63.

2. Singh I, Freeman C, Madder A, Vyle JS, Heaney F. Fast RNA conjugations on solid phase by strain-promoted cycloadditions. Org. Biomol. Chem. 2012 Aug 1;10(33):6633–9.

3. Chakrabarty R, Stang PJ. Post-assembly Functionalization of Organoplatinum(II) Metallacycles via Copper-free Click Chemistry. J. Am. Chem. Soc. 2012 Sep 12;134(36):14738–41.

4. Debets MF, Van Berkel SS, Schoffelen S, Rutjes FPJT, Van Hest JCM, Van Delft FL. Aza-dibenzocyclooctynes for fast and efficient enzyme PEGylation via copper-free (3+2) cycloaddition. Chemical Communications. 2010;46(1):97.

5. Khatwani SL, Kang JS, Mullen DG, Hast MA, Beese LS, Distefano MD, et al. Covalent protein–oligonucleotide conjugates by copper-free click reaction. Bioorganic & Medicinal Chemistry. 2012 Jul 15;20(14):4532–9.

6. Jang S, Sachin K, Lee H, Kim DW, Lee HS. Development of a Simple Method for Protein Conjugation by Copper-Free Click Reaction and Its Application to Antibody-Free Western Blot Analysis. Bioconjugate Chem. 2012 Nov 21;23(11):2256–61.

7. german group, UAA into GFP OCT, and bicyclo[6.1.0]nonyne, and terminal alkyne, 1.

8. Temming RP, Eggermont L, Eldijk MB van, Hest JCM van, Delft FL van. N-terminal dual protein functionalization by strain-promoted alkyne–nitrone cycloaddition. Org. Biomol. Chem. 2013 Apr 3;11(17):2772–9.

9. Hahne H, Sobotzki N, Nyberg T, Helm D, Borodkin VS, Van Aalten DMF, et al. Proteome Wide Purification and Identification of O-GlcNAc-Modified Proteins Using Click Chemistry and Mass Spectrometry. J. Proteome Res. 2013 Feb 1;12(2):927–36.

10. Niederwieser A, Späte A-K, Nguyen LD, Jüngst C, Reutter W, Wittmann V. Two-Color Glycan Labeling of Live Cells by a Combination of Diels–Alder and Click Chemistry. Angewandte Chemie International Edition. 2013;52(15):4265–8.

11. Beatty KE, Fisk JD, Smart BP, Lu YY, Szychowski J, Hangauer MJ, et al. Live-Cell Imaging of Cellular Proteins by a Strain-Promoted Azide–Alkyne Cycloaddition. ChemBioChem. 2010;11(15):2092–5.

12. Chang PV, Prescher JA, Sletten EM, Baskin JM, Miller IA, Agard NJ, et al. Copper-free click chemistry in living animals. PNAS. 2010 Feb 2;107(5):1821–6.

13. Agard NJ, Prescher JA, Bertozzi CR. A Strain-Promoted [3 + 2] Azide–Alkyne Cycloaddition for Covalent Modification of Biomolecules in Living Systems. J. Am. Chem. Soc. 2004 Nov 1;126(46):15046–7.

14. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR. A Comparative Study of Bioorthogonal Reactions with Azides. ACS Chem. Biol. 2006 Nov 1;1(10):644–8.

15. Codelli JA, Baskin JM, Agard NJ, Bertozzi CR. Second-Generation Difluorinated Cyclooctynes for Copper-Free Click Chemistry. J. Am. Chem. Soc. 2008 Aug 1;130(34):11486–93.

16. Sletten EM, Bertozzi CR. A Hydrophilic Azacyclooctyne for Cu-Free Click Chemistry. Org. Lett. 2008 Jul 1;10(14):3097–9.

17. Jewett JC, Sletten EM, Bertozzi CR. Rapid Cu-Free Click Chemistry with Readily Synthesized Biarylazacyclooctynones. J Am Chem Soc. 2010 Mar 24;132(11):3688–90.

18. Ning X, Guo J, Wolfert MA, Boons G-J. Visualizing Metabolically Labeled Glycoconjugates of Living Cells by Copper-Free and Fast Huisgen Cycloadditions. Angewandte Chemie International Edition. 2008;47(12):2253–5.

19. Kuzmin A, Poloukhtine A, Wolfert MA, Popik VV. Surface Functionalization Using Catalyst-Free Azide–Alkyne Cycloaddition. Bioconjugate Chem. 2010 Nov 17;21(11):2076–85.

20. Griffith OW. Biologic and pharmacologic regulation of mammalian glutathione synthesis. Free Radical Biology and Medicine. 1999 Nov;27(9–10):922–35.

21. Vafai SB, Mootha VK. Mitochondrial disorders as windows into an ancient organelle. Nature. 2012 Nov 15;491(7424):374–83.

22. Poburko D, Demaurex N. Regulation of the mitochondrial proton gradient by cytosolic Ca2+ signals. Pflugers Arch - Eur J Physiol. 2012 Jul 1;464(1):19–26.

23. Blerkom JV, Davis P, Alexander S. Inner mitochondrial membrane potential (ΔΨm), cytoplasmic ATP content and free Ca2+ levels in metaphase II mouse oocytes. Hum. Reprod. 2003 Nov 1;18(11):2429–40.

24. McKay CS, Chigrinova M, Blake JA, Pezacki JP. Kinetics studies of rapid strain-promoted [3 + 2]-cycloadditions of nitrones with biaryl-aza-cyclooctynone. Organic & Biomolecular Chemistry. 2012;10(15):3066.

25. Ning X, Temming RP, Dommerholt J, Guo J, Ania DB, Debets MF, et al. Protein Modification by Strain-Promoted Alkyne–Nitrone Cycloaddition. Angewandte Chemie International Edition. 2010;49(17):3065–8.

26. Colombo M, Sommaruga S, Mazzucchelli S, Polito L, Verderio P, Galeffi P, et al. Site-Specific Conjugation of ScFvs Antibodies to Nanoparticles by Bioorthogonal Strain-Promoted Alkyne–Nitrone Cycloaddition. Angewandte Chemie International Edition. 2012;51(2):496–9.

27. McKay CS, Blake JA, Cheng J, Danielson DC, Pezacki JP. Strain-promoted cycloadditions of cyclic nitrones with cyclooctynes for labeling human cancer cells. Chem. Commun. 2011 Aug 30;47(36):10040–2.

28. Blackman ML, Royzen M, Fox JM. The Tetrazine Ligation: Fast Bioconjugation based on Inverse-electron-demand Diels-Alder Reactivity. J Am Chem Soc. 2008 Oct 15;130(41):13518–9.

29. Liu DS, Tangpeerachaikul A, Selvaraj R, Taylor MT, Fox JM, Ting AY. Diels–Alder Cycloaddition for Fluorophore Targeting to Specific Proteins inside Living Cells. J. Am. Chem. Soc. 2012 Jan 18;134(2):792–5.

30. Lang K, Davis L, Wallace S, Mahesh M, Cox DJ, Blackman ML, et al. Genetic Encoding of Bicyclononynes and trans-Cyclooctenes for Site-Specific Protein Labeling in Vitro and in Live Mammalian Cells via Rapid Fluorogenic Diels–Alder Reactions. J. Am. Chem. Soc. 2012 Jun 27;134(25):10317–20.

31. Devaraj NK, Upadhyay R, Haun JB, Hilderbrand SA, Weissleder R. Fast and Sensitive Pretargeted Labeling of Cancer Cells via Tetrazine/Trans-Cyclooctene Cycloaddition. Angew Chem Int Ed Engl. 2009;48(38):7013–6.

32. Soriano del Amo D, Wang W, Jiang H, Besanceney C, Yan AC, Levy M, et al. Biocompatible Copper(I) Catalysts for in Vivo Imaging of Glycans. J. Am. Chem. Soc. 2010 Dec 1;132(47):16893–9.

33. Lim JI, Sabouri-Ghomi M, Machacek M, Waterman CM, Danuser G. Protrusion and Actin Assembly are Coupled to the Organization of Lamellar Contractile Structures. Exp Cell Res. 2010 Aug 1;316(13):2027–41.

34. Mueller V, Ringemann C, Honigmann A, Schwarzmann G, Medda R, Leutenegger M, et al. STED Nanoscopy Reveals Molecular Details of Cholesterol- and Cytoskeleton-Modulated Lipid Interactions in Living Cells. Biophys J. 2011 Oct 5;101(7):1651–60.

35. Westphal V, Rizzoli SO, Lauterbach MA, Kamin D, Jahn R, Hell SW. Video-Rate Far-Field Optical Nanoscopy Dissects Synaptic Vesicle Movement. Science. 2008 Apr 11;320(5873):246–9.

36. Dempsey GT, Vaughan JC, Chen KH, Bates M, Zhuang X. Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging. Nat Meth. 2011 Dec;8(12):1027–36.

37. Pauff SM, Miller SC. Synthesis of Near-IR Fluorescent Oxazine Dyes with Esterase-Labile Sulfonate Esters. Org. Lett. 2011 Dec 2;13(23):6196–9.

38. Wombacher R, Cornish VW. Chemical tags: Applications in live cell fluorescence imaging. Journal of Biophotonics. 2011;4(6):391–402.

39. Adams SR, Campbell RE, Gross LA, Martin BR, Walkup GK, Yao Y, et al. New Biarsenical Ligands and Tetracysteine Motifs for Protein Labeling in Vitro and in Vivo: Synthesis and Biological Applications. J. Am. Chem. Soc. 2002 May 1;124(21):6063–76.

40. Los GV, Encell LP, McDougall MG, Hartzell DD, Karassina N, Zimprich C, et al. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. ACS Chem. Biol. 2008 Jun 1;3(6):373–82.

41. Gautier A, Juillerat A, Heinis C, Corrêa Jr. IR, Kindermann M, Beaufils F, et al. An Engineered Protein Tag for Multiprotein Labeling in Living Cells. Chemistry & Biology. 2008 Feb 22;15(2):128–36.

Chapter 4. Development of a photocrosslinker ligase for protein-protein interaction detection

The work discussed in this chapter is unpublished. We would like to thank Laura Gerber (Professor Dick Schrock lab, MIT Chemistry Department) for lending us the Porta-Ray 400R photocrosslinking instrument and the Lippard Lab (MIT Chemistry Department) for lending us Stratalinker photocrosslinking instrument for initial photocrosslinking experiments. We would like to thank Arturo Pizano (Professor Dan Nocera lab, MIT Chemistry Department) for his help in obtaining transient absorption spectra of benzophenone compounds. We would also like to thank Professor Jason Chin (Molecular Research Council, Laboratory of Molecular Biology) for helpful discussions.

Introduction to methods of protein-protein interaction detection

Protein-protein interactions, or the "interactome", form the basis of almost every cellular event, from growth, migration, division, to cell-cell communication. It has been estimated that 650,000 protein-protein interactions exist in the human interactome (1). The interactome not only dictates the normal functioning of the organism, but also underlies its diseased state. Mapping protein-protein interactions using biophysical, biochemical, or genetic means have provided a wealth of information on macromolecular protein complexes. Several major methods for mapping protein-protein interactions include co-immunoprecipitation, the yeast two-hybrid assay, protein-fragment complementation assay (PCA), and photoaffinity labeling (PAL).

Co-immunoprecipitation is one of the earliest methods to identify members of interacting protein complexes. In co-immunoprecipitation, a protein of interest is identified through affinity capture and subjected to a series of biochemical purification methods that rely on typically conventional chromatography steps. Even though it has undoubtedly led to new protein-protein interaction discoveries, co-immunoprecipitation has many limitations (2). First, the technique does not allow any spatial resolution because interacting proteins are identified in the cell lysates when compartmentalization of various organelles has been broken down. When taken out of endogenous environment inside the cell, the stoichiometry and concentration of interacting proteins are severely disturbed which may lead to false loss of interacting partners. Furthermore, in a protein-protein interaction complex, the antigen may be buried and inaccessible by the antibody. Finally, the efficiency of co-immunoprecipitation can depend largely on the availability and affinity of the antibodies.

One of the most widely used methods for protein-protein interaction detections is the yeast two-hybrid system (3,4). The technology splits the two functional domains, often the DNA binding domain and the activation domain, of a transcriptional activator and fuses each domain to one of a pair of potentially interacting proteins, called the bait and the prey protein. The physical interaction of the bait and prey protein brings near the two activator domains and reconstitutes the transcriptional activator's ability to turn on a reporter gene. The yeast two-hybrid method has not only been used to assess individual protein-protein interactions, but also to map the interactions in the entire proteome (5).

Array-based automated yeast two-hybrid approaches have been instrumental for interaction studies on a global proteomic scale. In the seminal work of Stelzl *et al.* by fusing 4456 baits and 532 preys to activator domains, the authors identified more than 3000 novel protein-protein interactions including interactions proposed to form disease-related proteins. Furthermore, yeast two-hybrid screens that focus on disease pathways and functions have uncovered among many other remarkable studies, proteins involved in inherited ataxias (6) and neurodegenerative Huntington's disease (7).

However, despite the high success of large-scale yeast two-hybrid approaches, the technique is not without some serious drawbacks. First, the frequent occurrence of false positives and false negatives imposes a challenge to meaningful data interpretation. False positives can be caused by overexpression of the heterologous proteins as well as self-activation of the reporter gene by the bait protein in the absence of bait-prey interaction. Some false negatives can be caused by improper folding of the bait and prey proteins once they are fused to their respective domains, or mislocation and the inability of the protein partners to translocate to the nucleus, such as interactions of membrane proteins (4). Estimations for false positives range from 25 to 45 % (8). Individual confirmation of protein-protein interactions with alternative methods is not always possible for large-scale yeast two-hybrid approaches. Another serious drawback of yeast two-hybrid method is the poor time resolution. The amplification of the reporter gene signal requires first the formation of the protein interaction under study, reconstitution of the transcriptional activator, and finally the translocation of the transcriptional activator to the nucleus. The entire process can take hours after the protein-protein interaction has occurred (9). Moreover, protein-protein interactions requiring relevant cofactors that are not present in the yeast host organism may be lost. Because of these limitations, individual confirmation of protein-protein interactions with alternative methods, such as co-immunoprecipitation, is often required.

Variants of the yeast two-hybrid method have been developed to address the disadvantages and limitations of original yeast two-hybrid technique. Regarding the cellular compartment-dependent protein-protein interactions, such as the interaction between membrane proteins or between a membrane protein and a cytosolic protein, systems based on split-ubiquitin and split-TEV have been created so that the reconstituted protein can act to release the reporter protein, which can be either a transcription activator or light emitting enzyme like luciferase. In the case of split-ubiquitin, reconstituted ubiquitin is recognized and cut by ubiquitin-specific proteases, which in turn release the transcription factor that is attached to ubiquitin. Similarly, reconstituted TEV enzyme can directly cleave and free a reporter protein. However, the strategy is still very limited temporally.

One of the methods derived from the idea of the yeast two-hybrid technology is the protein-fragment complementation assay (PCA) (10). In PCA, protein-protein interactions are measured by fusing each interacting protein of interest to a split fragment of a reporter protein that has been rationally dissected into two fragments using protein-engineering strategies. An improvement over yeast two-hybrid, PCA offers the direct reconstitution of a reporter protein and avoids the long time lag from transcription activation of the reporter gene to obtaining enough expressed protein to have a measurable signal, thus allowing for faster detection of the dynamics of protein-protein interaction. Split-ubiquitin is one of the first and the smallest split-protein reporters. Split-ubiquitin offers multiple readout options, such as change in molecular weight, subcellular localization, and protein stability. Basically any protein that changes a measureable quality after being released from the reassembled ubiquitin can be utilized as a reporter for the system. Other notable split-protein pairs often used include those derived from GFP and its variants (11), dihydrofolate reductase (DHFR) (12), β-lactamase (13), and luciferases (14). However, for the GFP-derived split-protein pairs, the irreversible nature of the reconstituted fluorescent proteins limits their ability to report the temporal information of the specific protein interacting pair under study.

Imaging based microscopy methods have also been applied to study protein-protein interaction. In fluorescence resonance energy transfer (FRET), classical energy donor and acceptor proteins, such as CFP and YFP, are fused to interacting proteins, respectively. However, FRET experiments are only effective in assessing protein-protein proximity due to the resolution limit and are unable to provide concrete evidence of physical interaction. Moreover, design of protein fusions often require structural knowledge of the protein interacting pairs under study, thus it is rather limited to well-known, or highly proposed interacting partners. Without structural knowledge, one could unknowingly block the interaction sites with protein fusions, leading to false negatives. Additionally, the energy donor and acceptor proteins must be free of intrinsic dimerization tendency, which would cause false positive results. The method is also intrinsically restricted to studying only two proteins at a time.

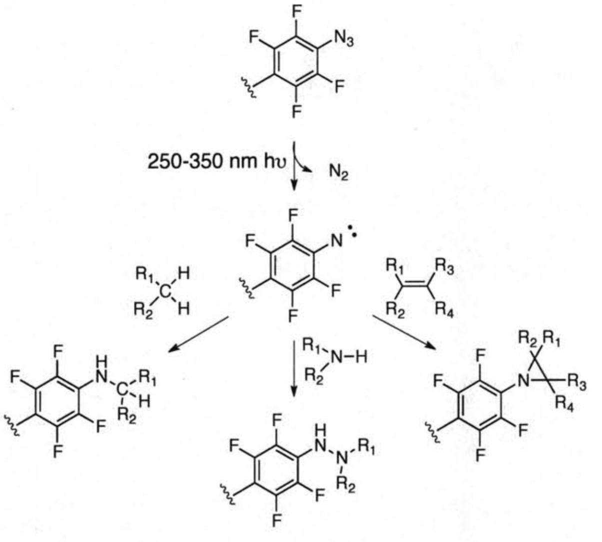
Finally, photoaffinity labeling (PAL) is yet another method that has been used to determine the position of a protein relative to those of other proteins within a macromolecular complex (15). PAL requires the use of a photoactivatable but otherwise chemically inert analogue, also called a photocrosslinker that can produce a highly reactive species once activated by light. The site-specific incorporation of a photocrosslinker to a protein of interest can pinpoint the site, or even the amino acid in the vicinity of the interacting proteins. A tremendous advantage of PAL over previously mentioned yeast two-hybrid, PCA and microscopy-based protein-protein interaction detection methods is that it can capture interactions with endogenous proteins rather than recombinant proteins. The proteins do not have to be modified with large split protein or fluorescent protein tags, and the crosslinked protein complexes are stable due to the formation of a covalent bond. Photocrosslinkers can be site-specifically introduced to proteins through labeling of cysteine residues or using the unnatural amino acid (UAA) mutagenesis technique (16) that has been developed to include several photocrosslinkers, such as benzophenone, aryl azide, and diazirine as unnatural amino acids (17,18). The high temporal resolution due to the reactive radical species of light-activated photocrosslinkers is especially advantageous in capturing transient interactions between proteins; some of these may be lost or not detectable by traditional co-immunoprecipitation or split-protein techniques.

However, even though *in vivo* crosslinking based on the genetic code expansion of UAA mutagenesis to site-specifically incorporate a photocrosslinker has been widely used, examples of applying it to map protein interactions inside live mammalian cells are still scarce (19,20). The prevalence of natural amber codons (18), low suppression efficiency (21), and generation of truncated byproducts (an average efficiency of 20% of full-length protein with one unnatural amino acid to truncated protein can be achieved) (22) make UAA mutagenesis challenging to carry out. We wanted to develop a robust, simple, and minimally invasive method to site-specifically label proteins with photocrosslinkers inside live cells, so they can become a routine tool for protein-protein interaction detection.

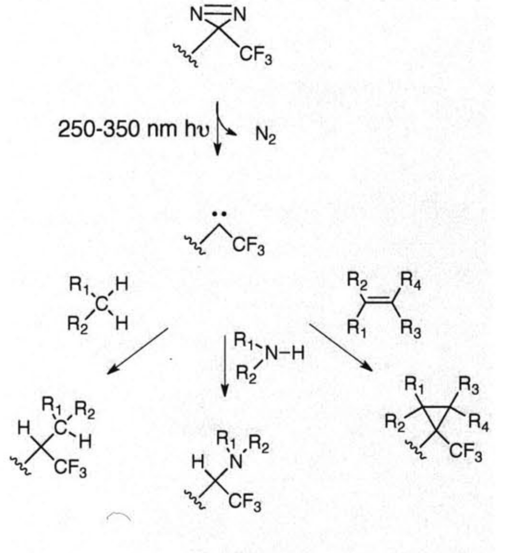
In this project, we describe the work taken towards the development and application of a photocrosslinker ligase to be used for protein-protein interaction detection inside live cells. The photocrosslinker ligase can specifically recognize and ligate an organic small molecule photocrosslinker onto its protein substrate. We hoped the method to be minimally invasion to endogenous proteins, has high spatial and temporal resolution, and high efficiency of photocrosslinking. Starting from the enzyme lipoic acid ligase (LplA) that has previously been applied to ligate an aryl azide photocrosslinker, we first describe our enzyme engineering of LplA to make it accept the more efficient photocrosslinker benzophenone. We then describe the effort to demonstrate LplA ligation of benzophenone on purified proteins, in cell lysate and in live cells. Finally, knowing that benzophenone ligation works inside live cells, we worked toward applying benzophenone photocrosslinking to detect protein-protein interactions in live cells.

Classes of photocrosslinkers

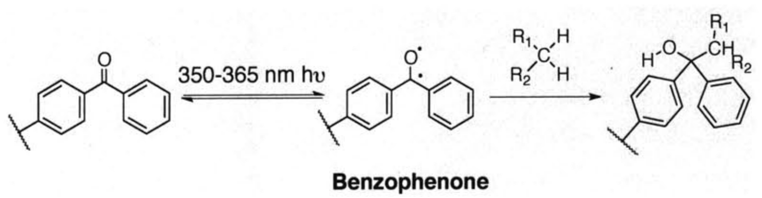
Detection of protein-protein interaction using a crosslinking probe has been used to map protein networks over the years. The light-activated crosslinking probe, often referred to as a photocrosslinker or a photoaffinity probe, must satisfy several requirements for it to function ideally in a live cell environment. First of all, the inactivated form of the photocrosslinker must be bioorthogonal to all small molecules and functional groups inside the cell. It must remain stable in the reducing environment of the cytosol and the oxidizing environment in cellular organelles such as the ER. The photocrosslinker must not be metabolized or degraded by cellular enzymes. Furthermore, the UV light activated radical of the photocrosslinker should not react or be quenched by the cellular medium but has high specificity towards its nearby proteins. The desirable half life of the reactive species should be long enough to give it sufficient time to crosslink with other proteins, but not too long that it crosslinks promiscuously. Long excitation wavelength is especially preferred for cellular applications to minimize the damage caused by high energy UV light. Three main types of photocrosslinkers are commonly used for photoaffinity labeling experiments: aryl azide, diazirines, and benzophenone.



Aryl Azide



Diazirine



Benzophenone

Figure 4-1. Three classes of photocrosslinkers. The structures shown at the top are the photochemical precursors, followed by the reactive species after UV irradiation. The nitrene radical and the carbene radical can insert into C-H bonds, heteroatom-H bonds (i.e. N-H bond), and alkenes. Benzophenone is very selective towards C-H bond. Note that formation of the benzophenone ketyl diradical is reversible.

Aryl azides are most efficiently activated by UV light within 250-350 nm wavelengths depending on the substitution on the phenyl ring (Figure 4-1). The reactive form of aryl azide is produced through a concomitant loss of N_2 and nitrene formation (23–25). Nitrene can insert into neighboring C-H bonds, heteroatom-H bonds, and form covalent bonds through adding to alkene functional groups. However, if the reactive nitrene did not find a suitable partner during its typical 100 μs life time, it will convert to the less reactive but more stable 7-membered ring ketenimine, which cannot react with proteins but only with nucleophiles (25,26). The rearrangement of the unstable nitrene to ketenimine significantly decreases photocrosslinking yields. For this reason, perfluoro aryl azides are often favored over nonfluorinated aryl azides because fluorine substitution slows down the rate of ring expansion of the perfluoro aryl nitrene to ketenimine. Aryl azide has also been utilized to study the interaction between RNA-dependent protein kinase (PKR) and RNA molecules by site-specifically introducing the photocrosslinker to various positions of PKR via cysteine-maleimide chemistry (27). Despite its low protein-protein crosslinking efficiency, aryl azide can serve as a functional handle to react with triphenyl phosphine in Staudinger ligation or with an alkyne in copper-catalyzed or strain-promoted click chemistry (28). This chemoselective reaction serves as an ancillary benefit of aryl azide that allows its incorporation into cellular systems to be monitored. In our previous publication using perfluoro aryl azide to photocrosslink the FK506 binding protein (FKBP) and the FKBP-rapamycin binding domain (FRB) (29), Dr. Hemanta Baruah used the reaction between aryl azide and phosphine as readout of the site-specific incorporation of the probe.

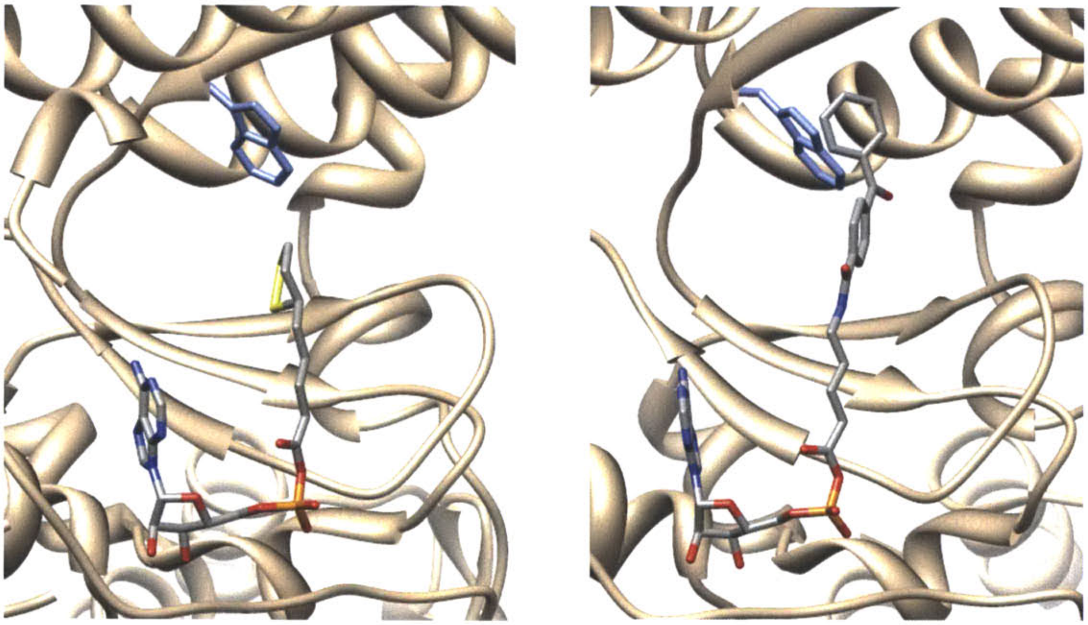
Another class of photocrosslinker is the diazirine (Figure 4-1). Diazirines form a reactive carbene intermediate with concomitant loss of N_2 upon 350 nm irradiation (25,30). The small size of diazirine groups introduces minimal perturbation to the protein under study. Its longer wavelength of activation light reduces the damage to the targeted biological system. Carbenes generated from diazirines are highly reactive and can react with nearby C-H or heteroatom-H bonds (28) (Figure 4-1). However, the striking reactive carbene intermediate has a very short half-life in the nanoseconds range (31) and can react rapidly with water, which decreases its photocrosslinking efficiency (23). Another significant drawback of the diazirine based photocrosslinkers is that they can undergo internal rearrangement to form diazo isomers that are long lived electrophiles and can cause unwanted background reactions with nucleophiles (30). For this reason, trifluoromethyl diazirines, whose diazo isomers are much less reactive, are preferred over alkyl diazirines. Photoaffinity labeling using diazirines ranges from probing receptor and ligand interaction, identification of drug targets, and protein-protein interaction detection (32). In studying calcium signaling in muscles, diazirine-based dantrolene, which is a muscle relaxant, crosslinked to a protein proposed to be an inhibitor of calcium release (33). Diazirines have also been utilized in studying glycoproteins and carbohydrate-mediated protein interactions, which are often poorly understood due to their heterogenous nature (34). Diazirines are of particular interest in this application over the other photocrosslinking probes because of their small size. Specifically, it has been introduced into O-linked β-N-acetylglucosamine (O-GlcNAc) by posttranslational modification (35). This unnatural sugar was incorporated by a mutant UDP-GlcNAc pyrophosphorylase onto O-GlcNAcylated FG-repeat nucleoporins, which was shown to crosslink with nuclear transport factors (35).

Benzophenone is yet another class of photocrosslinker. Benzophenone is activated by long wavelength excitation light in the range of 350 –365 nm (Figure 4-1) (36). The activated form of benzophenone, a ketyl diradical, is stable in the biological environment and can persist for as long as 120 μs. The excitation in the n-π* band of the α, β–unsaturated ketone produces a diradical triplet state which makes the radical an efficient hydrogen abstractor. Benzophenone reacts primarily and selectively with nearby C-H bonds in preference to the stronger O-H bonds of the solvent (28,36). If an interaction partner is not present, the ketyl diradical relaxes back to the ground state and can be activated again. The ability to be repeatedly excited, together with long life time of the radical and low background reaction with solvent, leads to benzophenone's relatively high photocrosslinking efficiency compare to those of aryl azide and diazirine, which rearranges to ketenimine and diazo isomer, respectively, or get quenched by water. However, benzophenone exhibits a high geometric constraint with regard to the angle and distance between the ketyl diradical and the nearby C-H bond. Recent work by Pavlova and coworkers (37) showed benzophenone (Bpa) introduced by UAA mutagenesis to various positions of an *E. coli* serine protease autotransporter, EspP, crosslinked to amino acids within 4 Å of the polypeptide backbone. The Yokoyama group introduced Bpa at Arg85 in gankyrin, a liver oncoprotein, through UAA mutagenesis approach (38). Crystal structure of photocrosslinked complex Bpa85 gankyrin and its interacting partner C-terminal domain of S6 proteasomal protein (S6C) revealed that the carbonyl group of Bpa85 crosslinked exclusively to the Cγ atom of Glu356 in S6C. In the 16 Bpa mutants of gankyrin studied, the photocrosslinked products showed benzophenone exhibited preferences for narrow range of distance from 2.36 Å to 4.11 Å and an angle between 104.10° and 128.29°. In addition, several groups (39–43) have discovered that benzophenone appeared to preferentially crosslink with nearby methionine residues, termed the "methionine magnet effect" (42). Despite its strict geometry requirement, benzophenone has been incorporated through UAA mutagenesis (16,18) into many proteins such as adaptor protein Grb2 (19) in live mammalian cells and glutathione *S*-transferase in *E. coli* (44).

Our lab has previously demonstrated LplA mutant-based aryl azide labeling and crosslinking in mammalian cell lysate and on the cell surface of live cells (29). Based on the same method, we wanted to develop an LplA ligase that can bind and ligate benzophenone, which has longer radical lifetime and lower nonspecific reaction with solvent than aryl azide. More importantly, benzophenone can be repeatedly activated to increase photocrosslinking efficiency. We hoped that the better photophysical properties of the benzophenone photocrosslinker would allow us to detect endogenous protein interacting partners inside live mammalian cells that we were not able to achieved previously with aryl azide. We wished to add the LplA-mediated photocrosslinker ligation methodology to the toolbox of protein-protein interaction mapping techniques and extend this photocrosslinking method to inside live mammalian cells.

Screening for a benzophenone ligase

As we discussed in Chapter 2, our lab has engineered the enzyme lipoic acid ligase (LplA) (45–48) to incorporate many unnatural small molecule fluorophores (49–52) and a photocrosslinker (29). Dr. Hemanta Baruah first laid the groundwork for engineering LplA towards accepting other organic small molecule substrates (29). Based on the crystal structure of LplA bound with lipoic acid (47), Dr. Baruah performed alanine scanning mutagenesis of amino acid residues that are within 7.5 Å of the dithiolane ring of lipoic acid and identified the Trp37 position to be critical for the LplA mutant to accept a small aryl azide molecule. Further mutagenesis studies discovered that LplA(W37V) was able to accept an aryl azide carboxylic acid with a four methylene linker. To see if an unnatural probe as large as benzophenone could possibly fit in the lipoic acid binding pocket of LplA, we compared crystal structures of LplA with lipoic acid bound (Figure 4-2, left, PDB 3A7R) and LplA with a benzophenone probe docked into the active site of the enzyme (Figure 4-2, right). The Trp37 residue is shown in blue. We can see that one of the phenyl rings of benzophenone clashes into the Trp37 residue.


E. coli lipoic acid ligase with bound lipoic acid

E. coli lipoic acid ligase with docked benzophenone probe

Figure 4-2. Lipoic acid ligase (LplA) with (left) lipoic acid bound (PDB 3A7R) and (right) a benzophenone unnatural small molecule docked into the enzyme active site. The "gatekeeper" residue Trp37 is rendered in blue.

Even though we have seen engineered LplA being able to recognize small molecule substrates that bear little resemblance to its natural substrate, lipoic acid, we were not certain if simply mutating the Trp37 "gatekeeper" residue in the binding pocket of LplA would be sufficient to make LplA bind to benzophenone, which is larger than previous small molecule substrates of LplA. Nonetheless we decided to begin with a panel of Trp37 LplA mutant enzymes to test their ability to ligate one of two benzophenone carboxylic acid probes varying in linker lengths onto LAP peptide in an *in vitro* HPLC assay. We picked six amino acid residues with smaller side chains than Trp – Gly, Ala, Val, Ile, Leu, Ser. We chose mainly hydrophobic amino acids to maintain a hydrophobic binding pocket of LplA, which is suitable for benzophenone.

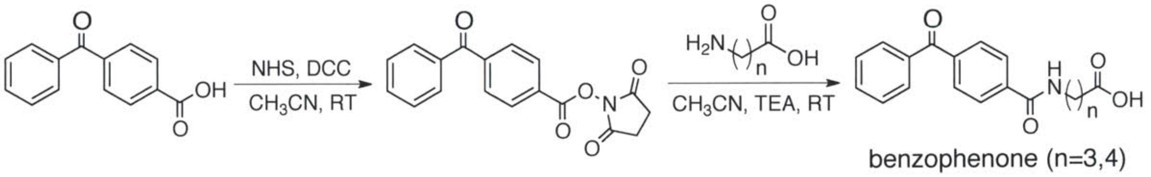
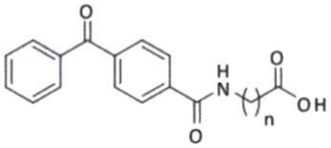


Figure 4-3. Synthetic scheme of benzophenone photocrosslinkers. NHS, *N*-hydroxysuccinimide; DCC, *N,N'*-dicyclohexylcarbodiimide; RT, room temperature; TEA, triethylamine.

Benzophenone probes were synthesized by first activating 4-benzoylbenzoic acid to the corresponding *N*-hydroxysuccinimide (NHS) ester via *N,N'*-dicyclohexylcarbodiimide (DCC) activation. 4-Benzoylbenzoic-NHS ester was then reacted with either 4-aminobutyric acid or 5-aminovaleric acid, through formation of an amide bond (Figure 4-3, Experimental Methods). Each probe was characterized by ^{1}H NMR and mass spectrometry. We screened six Trp37 LplA mutants against these two benzophenone substrates using an HPLC assay based on a shift in peak retention time to determine the conversion of the free 13-amino acid LplA Acceptor Peptide (LAP) (53) into LAP-benzophenone conjugate (Figure 4-4). Interestingly, only benzophenone with the four methylene linker (n=4) was able to get incorporated onto LAP by three LplA mutants. LplA(W37G) appeared to be the most efficient benzophenone ligase.

Subsequent benzophenone crosslinking experiments were performed with benzophenone with n=4 alkyl linker and LplA(W37G).



Benzophenone	LplA variant						
	WT	W37G	W37A	W37V	W37I	W37L	W37S
n=3	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.	N.D.
n=4	N.D.	100%	N.D.	65%	N.D.	73%	N.D.

Figure 4-4. Screening to identify LplA mutant/benzophenone substrate pair. The table shows percent conversions of LAP to LAP-benzophenone product conjugate. Wild-type LplA and six W37 point mutants were screened against two benzophenone substrates of different lengths (structures shown on left). N.D. indicates that product was not detected. Screening was performed with 500 nM LplA enzyme, 100 μM LAP peptide, and 300 μM benzophenone substrate for 40 min at 30 °C.

Detection of ligated benzophenone via reaction with fluorophore-hydroxylamine

Unlike the aryl azide photocrosslinker that has the azide moiety to act like a functional group handle to react with triphenyl phosphine in the Staudinger ligation or alkyne in click chemistry to monitor its incorporation, benzophenone does not have this ancillary benefit. For live cell application of benzophenone ligation, we needed an assay to assess the probe's ligation efficiency and specificity. The only available functional handle on benzophenone is the carbonyl group. Methods to detect aldehyde and ketone functional groups in proteins were first explored as a way to monitor protein oxidation by reactive oxygen species (54). Additionally, lipid oxidation and non-enzymatic glycation can produce protein-bound aldehydes and carbonyl functionalities. Hydrazide was explored as an aldehyde- and ketone-reactive probe (55,56). The reaction relies upon the condensation between the carbonyl group and the hydrazide to yield a stable hydrazone product. The hydrazide can be functionalized with either biotin or a fluorophore as readout. Following the introduction of keto amino acid via the UAA mutagenesis approach, the incorporation and site-specificity of *p*-acetylphenylalanine into mammalian GPCRs (57) and bacteriophage T4 lysozyme (58) have been monitored using hydrazide-fluorescein conjugate under mild acidic condition. The *p*-acetylphenylalanine incorporated protein can be separated by SDS-PAGE and visualized under a fluorescence imager.

Even though the condensation reaction with hydrazide has been successfully implemented to detect aldehyde and acetophenone, it was a challenge in the case of benzophenone because of its very deactivated and electron-rich carbonyl group. Regardless, we decided to try reacting benzophenone-ligated E2p (9 kDa) with biotin-hydrazide under acidic condition (Figure 4-5). Following *in vitro* benzophenone ligation onto E2p by LplA(W37G) (see Experimental Methods), the samples were dialyzed in phosphate buffered saline (PBS) to remove excess benzophenone probe. Biotin-hydrazide was added and the pH of the reaction was adjusted to 4.5. After 12 hours incubation with biotin-hydrazide, reactions were separated on an SDS-PAGE gel followed by Western blot transfer and detection with streptavidin-HRP. However, streptavidin-HRP stained both LplA(W37G) and E2p proteins and no benzophenone-specific signal was observed (Figure 4-5). Several explanations could be possible for the lack of benzophenone-specific signal. Biotin-hydrazide could nonspecifically stick to proteins. It is also possible that the reaction product hydrazone was hydrolyzed during subsequent analysis.

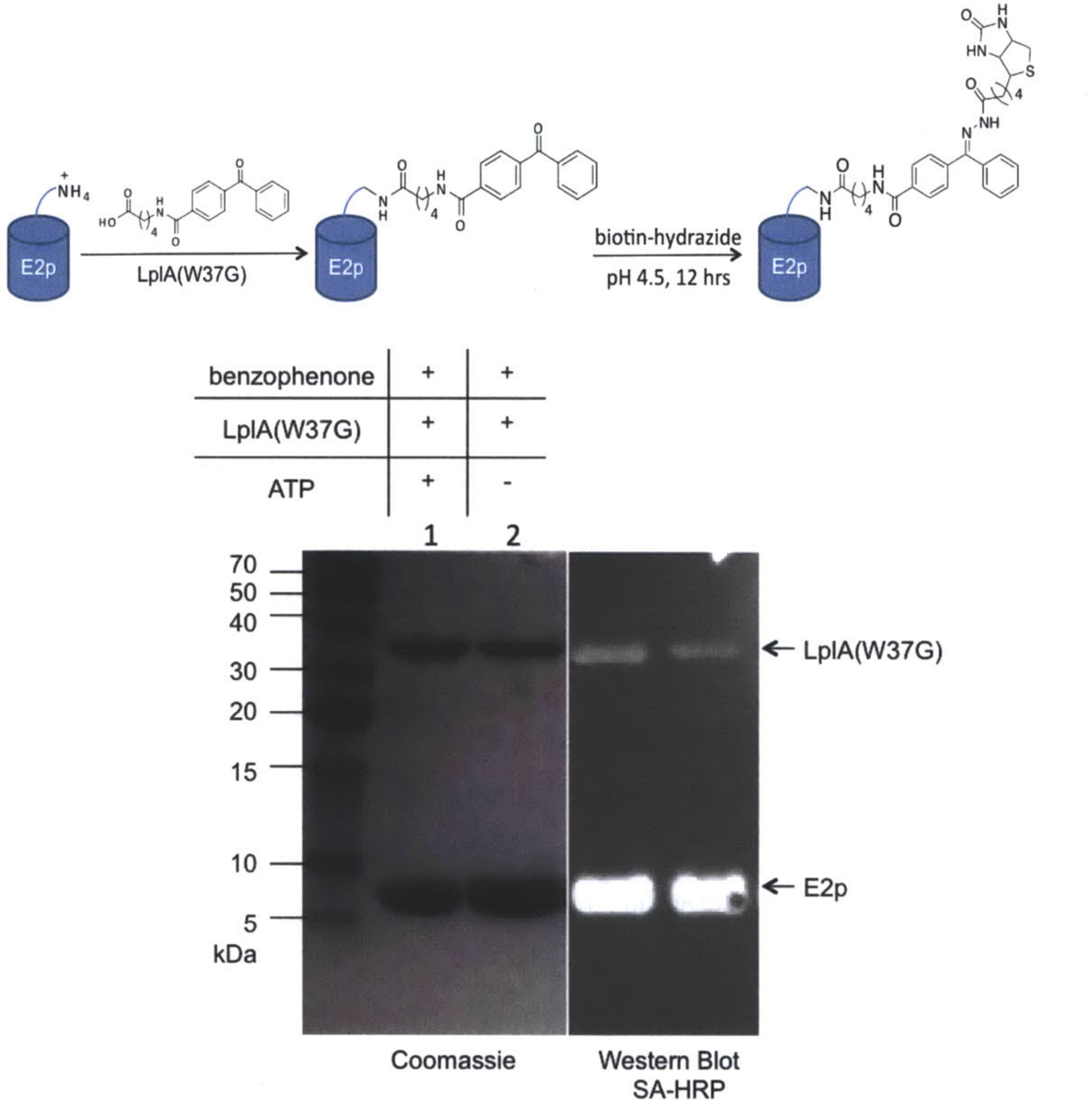

E2p
LplA(W37G)
biotin-hydrazide
pH 4.5, 12 hrs
benzophenone + +
LplA(W37G) + +
ATP + -
1
2
70
50
40
30
20
15
10
5
kDa
← LplA(W37G)
← E2p
Coomassie
Western Blot
SA-HRP

Figure 4-5. Unsuccessful attempt to detect *in vitro* ligation of benzophenone to E2p protein using biotin-hydrazide. Benzophenone ligation was performed using LplA(W37G) enzyme on purified E2p protein. Subsequently the mixture was combined with 200 μM of biotin-hydrazide and the pH of the reaction was adjusted to 4.5 to detect the benzophenone-E2p adduct via hydrazone formation. Blotting with streptavidin conjugated HRP (SA-HRP) is shown to the right of the Coomassie-stained gel. Negative control is shown with ATP omitted in the *in vitro* reaction (lane 2). Bands of equal

intensity are seen in lanes 1 and 2 in the SA-HRP blot, suggestive of non-specific binding rather than specific signal due to successful ligation and detection of benzophenone.

In fact, studies comparing the hydrolytic stabilities of hydrazones and oximes have shown that hydrazones are more prone to hydrolysis because of the higher likelihood of protonation on the nitrogen atom attached to the carbon, whose electrophilicity is enhanced as a result (59). Oximes, on the other hand, are more stable to hydrolysis than hydrazones because of the inductive effect and the higher electronegativity of the oxygen atom attached to the reactive nitrogen atom, which is in turn made less likely to be protonated. Thus, the carbonyl carbon atom in oximes is less electrophilic than the carbonyl carbon in hydrazones (Figure 4-6). Comparison of first-order rate of hydrolysis between hydrazone and oxime shows that the half-life of methyl oxime is 25 days (600 hours) while for methyl hydrazone, it is only 2 hours (59). Because the ultimate application is to ligate benzophenone in live mammalian cells, where the cell medium is hydrophilic and contains nucleophiles that could possibly react with the hydrazone product, we decided to react benzophenone with hydroxylamine instead of hydrazide to form a more stable oxime.

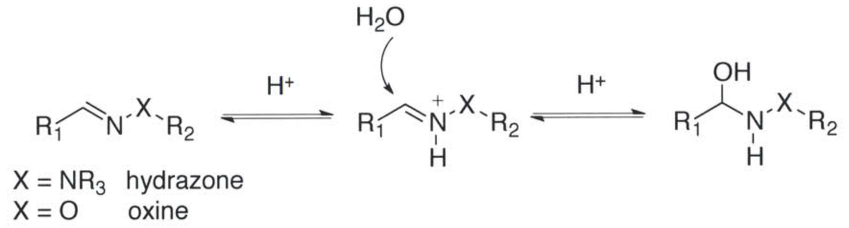


Figure 4-6. Hydrolysis of hydrazones and oximes. Hydrazones are more likely to be hydrolyzed than oximes because its carbon atom is made more electrophilic by the higher likelihood of protonation on the attached nitrogen (59).

Chemoselective reaction between keto groups and hydroxylamines has been demonstrated in detecting a benzaldehyde-containing sialic acid analog that is incorporated into cell-surface glycans through metabolic labeling (60). Furthermore, this reaction has also been directly compared with the reaction between keto groups and hydrazide in monitoring site-specific incorporation of *p*-acetylphenylalanine into bacteriophage T4 lysozyme (58). The Deniz group showed that hydroxylamine-Alexa Fluor 488 conjugate reacted with the acetophenone amino acid more effectively than hydrazide-Alexa Fluor 488 conjugate, providing on average greater than 90% labeled protein under optimized conditions. Hydroxylamine-Alexa Fluor 488 demonstrated more pH flexibility, being able to detect T4 lysozyme from pH range 4 to 9, while hydrazide-Alexa Fluor 488 only showed activity at pH 4 and 5. Reaction with hydroxylamine also appeared to be more sensitive, detecting as low as 1.5 μM of protein, while hydrazide only showed reaction at 100 μM protein concentration (58).

To test whether hydroxylamine derivatives can be used as a reagent to detect benzophenone, we first tested the reaction between benzophenone and biotin-hydroxylamine to see if they can form the oxime product (Figure 4-7). We dissolved 5 μM benzophenone and 5 μM biotin-hydroxylamine in 200 μl phosphate buffered saline (PBS) and adjusted the pH of the reaction to 5. We chose PBS as the reaction medium because it is compatible with later protein ligation and photocrosslinking experiments. The reaction was incubated at room temperature for 12 hours. We were able to observe the correct molecular weight of the oxime product by mass spectrometry at $[M+H]^+$ 639.12 (expected molecular weight: 638.25), which indicated that the reaction between benzophenone and biotin-hydroxylamine can successfully occur.

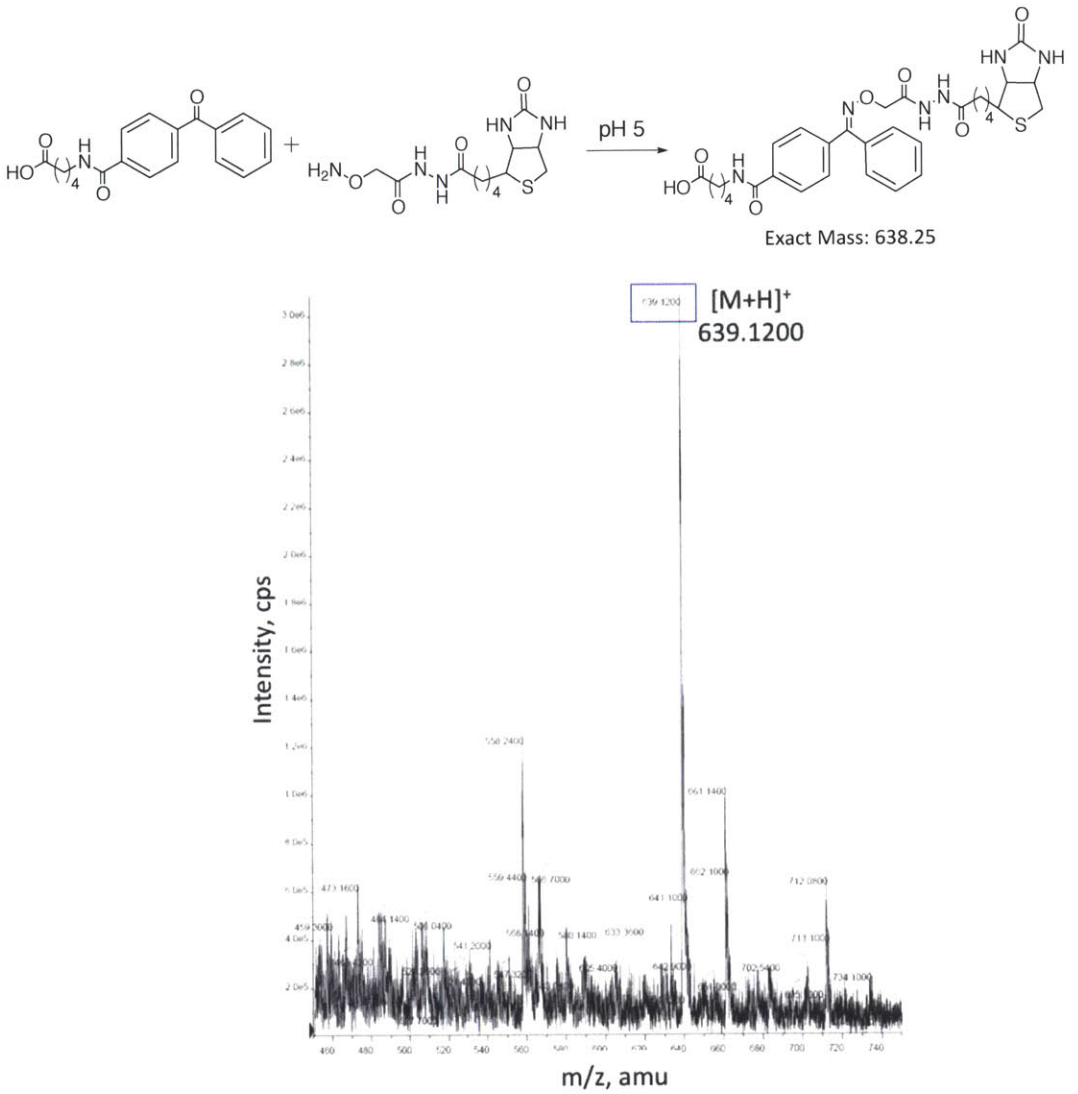
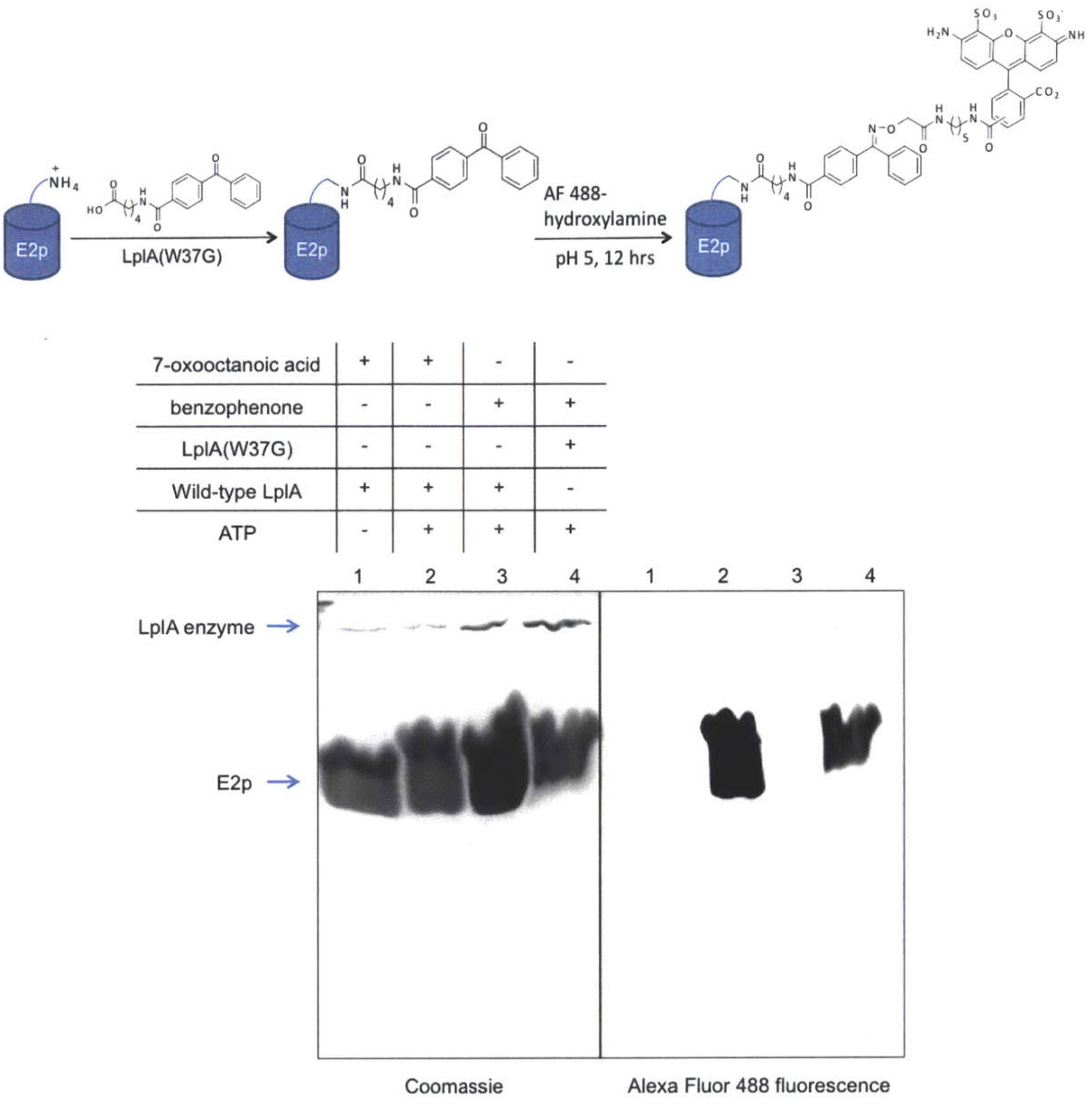


Figure 4-7. Reaction between benzophenone and biotin-hydroxylamine to form oxime product. (Top) Scheme of chemical reaction. Each reactant (5 μM) was combined in phosphate buffered saline (PBS) at pH 5. Reaction was incubated at room temperature for 12 hours. (Bottom) Mass spectrometry analysis of oxime product, which was observed at molecular weight $[M+H]^+$ 629.1200.

Subsequently, we tried to detect *in vitro* ligation of benzophenone to E2p protein using hydroxylamine-Alexa Fluor 488 conjugate. After *in vitro* benzophenone ligation onto E2p protein using purified enzyme LplA(W37G) (see Experimental Methods), excess benzophenone probe was removed by dialyzing the reaction in PBS. Hydroxylamine-Alexa Fluor 488 was added to the reaction whose pH was adjusted to 5.

After incubating at room temperature for 12 hours, the samples were separated on an SDS-PAGE gel and analyzed on a fluorescence gel imager (Figure 4-8). A positive control reaction using a ketone small molecule, 7-oxooctanoic acid that can be incorporated by wild-type LplA onto E2p was performed side by side with benzophenone ligation and reaction with hydroxylamine. 7-Oxooctanoic acid was chosen as a positive control because it is a more conventional alkyl ketone whose carbonyl group is electrophilic and not deactivated as in benzophenone. Therefore, 7-oxooctanoic acid should react readily with hydroxylamine. As shown in Figure 4-8, strong in-gel fluorescence of Alexa Fluor 488 was seen specifically at the position of E2p only when benzophenone was ligated (lane 4). The negative control reaction using wild-type LplA, which does not accept benzophenone did not show any signal (lane 3). Compared to the intensity of the Alexa Fluor 488 in-gel fluorescence on E2p ligated with 7-oxooctanoic acid, E2p-benzophenone conjugate showed weaker fluorescence retention (Figure 4-8, lane 2 and 4). The comparison of the reactivity of hydroxylamine towards the two ketone structures indicated that the benzophenone did react with hydroxylamine to the same extent; the reactivity was much weaker compared to a conventional alkyl ketone without a deactivated carbonyl center.


E2p
LplA(W37G)
E2p
AF 488-
hydroxylamine
pH 5, 12 hrs
E2p
1
2
3
4
1
2
3
4
LplA enzyme
E2p
Coomassie
Alexa Fluor 488 fluorescence

7-oxooctanoic acid	+	+	-	-
benzophenone	-	-	+	+
LplA(W37G)	-	-	-	+
Wild-type LplA	+	+	+	-
ATP	-	+	+	+

Figure 4-8. Detection of *in vitro* ligated benzophenone using AF488-hydroxylamine. As shown in the top scheme, benzophenone ligation performed using LplA(W37G) enzyme on purified E2p protein was reacted with AF488-hydroxylamine at pH 5 to detect ligated benzophenone. AF488 in-gel fluorescence is shown to the right of the Coomassie-stained gel. Positive control reaction with wild-type LplA ligated 7-oxooctanoic acid onto E2p was performed in parallel (lane 2). Negative controls are shown with ATP omitted (lane 1) and wild-type LplA in place of LplA(W37G) (lane 3).

Following the success of detecting benzophenone ligation onto E2p with hydroxylamine, we tested the reaction in the more complex environment of HEK 293T cell lysate expressing LAP-tagged blue fluorescent protein (BFP). We performed *in vitro* benzophenone ligation with purified enzyme LplA(W37G) in cell lysate, followed by dialysis to remove excess benzophenone probe and incubated the cell lysate with Alexa Fluor 488-hydroxylamine at pH 5. However, Alexa Fluor 488-hydroxylamine stuck heavily to material in the lysate, causing high background. We were not able to see any specific reaction with benzophenone-ligated LAP-BFP above this high background. There are two possible causes of backgrounds: simply hydrophobic sticking of the dye, or non-specific reactions with endogenous ketones and aldehydes in the cell lysate. To test the second hypothesis, we first treated cell lysate with biotin-hydroxylamine to block endogenous ketones and aldehydes, so they would not react with Alexa Fluor 488-hydroxylamine to cause background signal. We then treated the cell lysate with benzophenone and LplA(W37G), followed by reaction with Alexa Fluor 488-hydroxylamine in a similar fashion as before. However, there was no significant reduction in background (data not shown).

In order to reduce nonspecific background, we tried a similar experiment using biotin-hydroxylamine to react with benzophenone with the hope that the more hydrophilic biotin-hydroxylamine would stick less to proteins compared to Alexa Fluor 488-hydroxylamine. In this experiment, we first activated 7-oxooctanoic acid or benzophenone acid to its corresponding *N*-hydroxysuccinimide (NHS) ester, then chemically conjugated each probe to a purified fusion protein, FRB-YFP, to maximize signal. 7-Oxooctanoic acid or benzophenone conjugated FRB-YFP protein was incubated with biotin-hydroxylamine at pH 5 for 12 hours. Samples were separated on a 12% SDS-PAGE gel and analyzed by Western blot with streptavidin (SA)-HRP staining to detect proteins derivatized with biotin-hydroxylamine. Figure 4-9 shows strong SA-HRP signal on 7-oxooctanoic conjugated FRB-YFP (lane 3), indicating efficient reaction between 7-oxooctanoic acid and biotin-hydroxylamine. Benzophenone conjugated FRB-YFP shows very minimal signal in comparison (lane 5). We subsequently tried to react biotin-hydroxylamine in HEK 293T cell lysate containing benzophenone-ligated LAP protein. However, we were unable to detect any specific signal (data not shown).

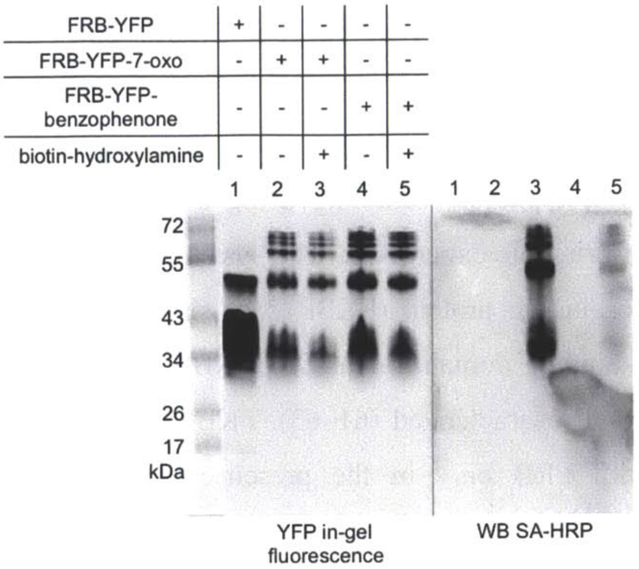


Figure 4-9. Detection of benzophenone chemically conjugated to FRB-YFP using biotin-hydroxylamine. 7-Oxooctanoic acid or benzophenone conjugated FRB-YFP protein was incubated with biotin-hydroxylamine at pH 5 to detect 7-oxooctanoic acid or benzophenone. YFP in-gel fluorescence detects FRB-YFP. Immunoblot with streptavidin (SA)-HRP detects biotinylated proteins. 7-Oxooctanoic acid conjugated FRB-YFP (lane 3) and benzophenone conjugated FRB-YFP (lane 5) both reacted with biotin-hydroxylamine. Negative controls are shown with only FRB-YFP protein (lane 1), or biotin-hydroxylamine omitted (lanes 2 and 4).

Although hydroxylamine was able to react with benzophenone as shown above (Figure 4-8, lane 4; Figure 4-9, lane 5) using *in vitro* purified E2p ligated with benzophenone, or benzophenone conjugated FRB-YFP protein, this detection method is not sensitive enough to react with benzophenone-ligated protein in cell lysates. Furthermore, reaction with hydroxylamine would only allow us to assay for whether benzophenone ligation has occurred, but is sunable to provide any information regarding the ligation yield. It is critical for us to have a method that enables quantitative measurement of the benzophenone ligation yield inside live cells to ensure the ligation is complete before proceeding to the photocrosslinking reaction. We later describe the development of a native gel shift assay to determine benzophenone ligation yield in section titled "Quantify benzophenone ligation yield inside live mammalian cells".

In vitro photocrosslinking of LAP-tagged FKBP and FRB proteins using benzophenone probe

Knowing that LplA(W37G) can ligate benzophenone probe onto LAP substrate *in vitro*, we wanted to test photocrosslinking using a protein-protein interaction model system as a proof-of-principle experiment. The system we chose is the ternary complex consisting of FK506-binding protein (FKBP12 or simply FKBP), rapamycin, and the FKBP12-rapamycin binding domain of mTOR (FRB). The interaction between FKBP and FRB has been well characterized (61–63). FKBP is known to form a heteromeric protein complex with FRB only in the presence of the immunosuppressive drug rapamycin. Rapamycin binds with high affinity ($K_d = 0.2$ nM) to FKBP (61). The FKBP-rapamycin complex in turn binds tightly to FRB. Previously, Dr. Hemanta Baruah demonstrated LplA-mediated aryl azide ligation onto LAP-tagged FKBP in mammalian cell lysate (29). Aryl azide-ligated FKBP-LAP was then photocrosslinked to its interaction partner FRB in the presence of rapamycin in cell lysate. We would like to test benzophenone photocrosslinking with the same interaction system as a proof-of-principle experiment and hopefully obtain higher photocrosslinking efficiency.

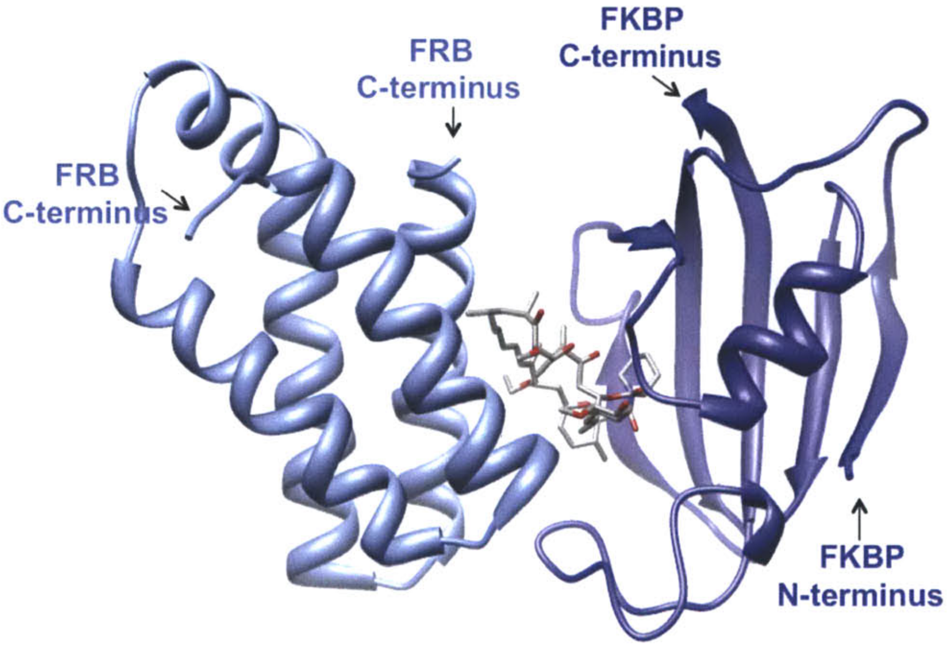


Figure 4-10. Crystal structure of FKBP-rapamycin-FRB ternary complex (PDB 2AFP) with termini of each protein indicated.

Upon examination of the crystal structure of FKBP-rapamycin-FRB ternary complex (Figure 4-10), we saw that both C-terminus of FRB and the C-terminus of FKBP are located relatively close to the interface of the complex. We decided to try FKBP with LAP fused to its C-terminus first because it has been demonstrated previously in photocrosslinking using aryl azide (29). We purified myc-tagged FKBP-LAP and HA-tagged FRB-YFP fusion proteins from *E.coli*. FKBP-LAP was then incubated with purified enzyme LplA(W37G), ATP, Mg^{2+}, and benzophenone probe to ligate benzophenone onto FKBP-LAP. Reactions were run on a reverse-phase HPLC column to confirm complete conversion of FKBP-LAP to benzophenone-ligated FKBP-LAP by the shift in peak retention time (data not shown). Excess benzophenone probe was removed through extensive dialysis. Purified FRB-YFP protein and rapamycin were then added to the reaction mixture. The complex was allowed to incubate with rotation at room temperature for 1 hour. Afterwards, the reaction was exposed to 365 nm UV light.

In our preliminary experiments, we experimented with three photocrosslinking instruments to optimize photocrosslinking conditions such as UV exposure time, distance from UV lamp, temperature, and ease of use. First, Porta-Ray 400R from UVitron is equipped with 400-Watt UVA enhanced arc lamp with intensity output concentrated from 350 to 400 nm wavelength. Its curing area is about 5 X 3 inches, which is ideal for the size of a typical mammalian cell culture plate. The cells can be placed anywhere from 1 inch to 3 inches from the UV lamp. Second, handheld long wavelength UV lamp from UVP has 4–Watt power and can be placed directly on top of cell culture plate. Even though the power of the handheld UV lamp is low, the distance between the UV light and cells can be as short as 1 centimeter. The third instrument is the Stratalinker from Stratagene, which has 5 lamps with 8-Watt output per lamp. The sample can be placed anywhere from 1 inch to 10 inches from the lamp. Since all the instruments differ in UV lamp power, distance between cells and UV light, and the effective curing area, we tried all three instruments with multiple UV irradiation conditions for each instrument. The most optimized condition for each instrument is listed in Table 4-1. All three instruments performed similarly under their optimized conditions.

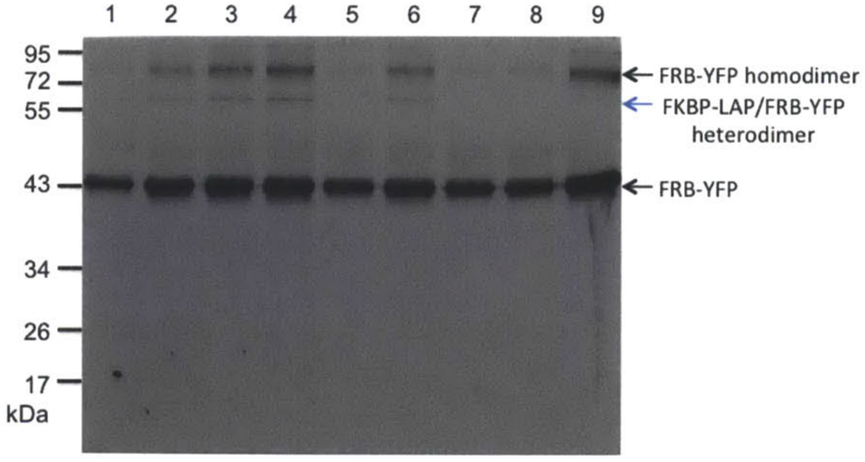
Instrument	**Wavelength**	**Power output**	**Distance from samples**	**Time of UV irradiation**
Porta-Ray 400R	350 – 400 nm	400-Watt	3 inches	2 minutes
Handheld long wavelength UV lamp	365 nm	4-Watt	1 cm	15 minutes
Stratalinker	365 nm	5 lamps, 8-Watt/lamp	5 inches	30 minutes

Table 4-1. Optimized UV photocrosslinking conditions for each instrument tested.

We first ligated benzophenone onto FKBP-LAP using purified LplA(W37G) protein and ATP (see Experimental Methods). After ligation was complete, we dialyzed the samples to remove excess unligated benzophenone. Dialyzed samples were incubated with purified FRB-YFP protein and rapamycin for 1 hour at room temperature. Using the Stratalinker instrument, a 96-well cell culture plate containing the *in vitro* ligated samples was placed on top of an ice tray that was put at 5 inches below the UV lamps. After UV irradiation for the indicated duration from 2 to 60 minutes (Figure 4-11, lanes 1-4), samples were separated on a 12% SDS-PAGE gel. To detect photocrosslinked products, we utilized YFP in-gel fluorescence to detect the YFP-tagged FRB. Many similar experiments were performed with varying UV irradiation time from 2 minutes to 15 minutes depending on the power of the instrument, and the distance between sample and UV lamps. Figure 4-11 shows a representative result from *in vitro* photocrosslinking using purified FKBP-LAP and FRB-YFP proteins. We detected an YFP-positive band (blue arrow) at 60 kDa, corresponding to the combined molecular weights of the two interaction partners, HA-FRB-YFP (41.2 kDa) and myc-FKBP-LAP (17.1 kDa). The intensity of the band increased with UV irradiation time (Figure 4-11, lanes 1-4). Negative control reactions with either UV irradiation omitted (Figure 4-11, lanes 5), benzophenone probe omitted (lane 7), ATP omitted (lane 8), or FKBP-LAP omitted (lane 9) did not show the photocrosslinked heterodimer product.

However, the crosslinked product did not depend on the presence of rapamycin (Figure 4-11, lanes 6). One explanation for this could be that the *in vitro* concentrations of the two interacting proteins and the small molecule rapamycin can be difficult to control precisely, especially when the initial concentration of benzophenone-ligated FKBP-LAP can be slightly changed after dialysis. During *in vitro* photocrosslinking, the chance of benzophenone-ligated FKBP-LAP encountering FRB in the absence of rapamycin was high because they were the most concentrated proteins present in the test tube. Later on, we tested lower concentrations of each binding partner and observed that at lower concentrations, the chance of getting rapamycin-independent background photocrosslinking was much lower (see Figure 4-20 in later section). We also observed another YFP-positive band whose molecular weight corresponds to a homodimer product (82.4 kDa) of FRB-YFP (Figure 4-11). We speculated that this could be caused by the intrinsic dimerization tendency of YFP (39) in combination with UV exposure. We later tested photocrosslinking using HA-tagged FRB without YFP fusion to eliminate this problem.

UV irradiation	+	+	+	+	-	+	+	+	+
rapamycin	+	+	+	+	+	-	+	+	+
benzophenone	+	+	+	+	+	+	-	+	+
ATP	+	+	+	+	+	+	+	-	+
FKBP-LAP	+	+	+	+	+	+	+	+	-
UV irradiation time (min.)	2	15	30	60	NA	30	30	30	30



YFP fluorescence

Figure 4-11. *In vitro* photocrosslinking with purified proteins FKBP-LAP and FRB-YFP. Benzophenone-ligated FKBP-LAP protein was incubated with FRB-YFP and rapamycin followed by UV irradiation in Stratalinker instrument. Samples were separated on an SDS-PAGE gel. YFP in-gel fluorescence was detected using a chemiluminescence imager. FRB-YFP (41.2 kDa), FKBP-LAP (17.1 kDa). Photocrosslinked product of the two proteins is indicated by blue arrow.

Optimizations to improve *in vitro* photocrosslinking yield

Since the photocrosslinking yield of FKBP-LAP and FRB-YFP was very low despite trying different UV instruments, varying UV irradiation times, and optimizing the distance between sample and UV light, we thought perhaps something else instead of photocrosslinking condition was causing low yield. We tried several things to troubleshoot the low yield of photocrosslinked product.

1. *Methionine to Alanine mutation on FKBP-LAP*

Benzophenone is highly sensitive to its surrounding amino acids and requires specific geometric configurations of the nearby C-H bonds to react (36,38). Furthermore, several studies have shown that photocrosslinking experiments with benzophenone continued to identify methionine as the crosslinked residue at the site of interaction (39–43). Benzophenone could perhaps photocrosslink and insert into one of the C-H bonds on either of the methylene units of methionine. This preference of benzophenone towards methionine has been termed the "magnet effect" of methionine (42). We suspected that in our LplA-mediated ligation system, perhaps the methionine residues near the ligation site of benzophenone in FKBP-LAP protein were reacting with the ligated benzophenone before it had a chance to photocrosslink to residues of the interacting protein FRB. The possibility of benzophenone ketyl diradical being "quenched" by nearby methionine residues could severely handicap or completely prevent the ability of the radical to interact with other C-H bonds. Crystal structure of the ternary complex of FRB-rapamycin-FKBP (PDB 2FAP) (61,62) (Figure 4-12) shows a total of three methionines on FKBP and six methionines on FRB, highlighted as ball-and-sticks. Since LAP peptide was fused to the C-terminus of FKBP, the two methionines on FKBP that are closest to benzophenone are Met49 and Met66.

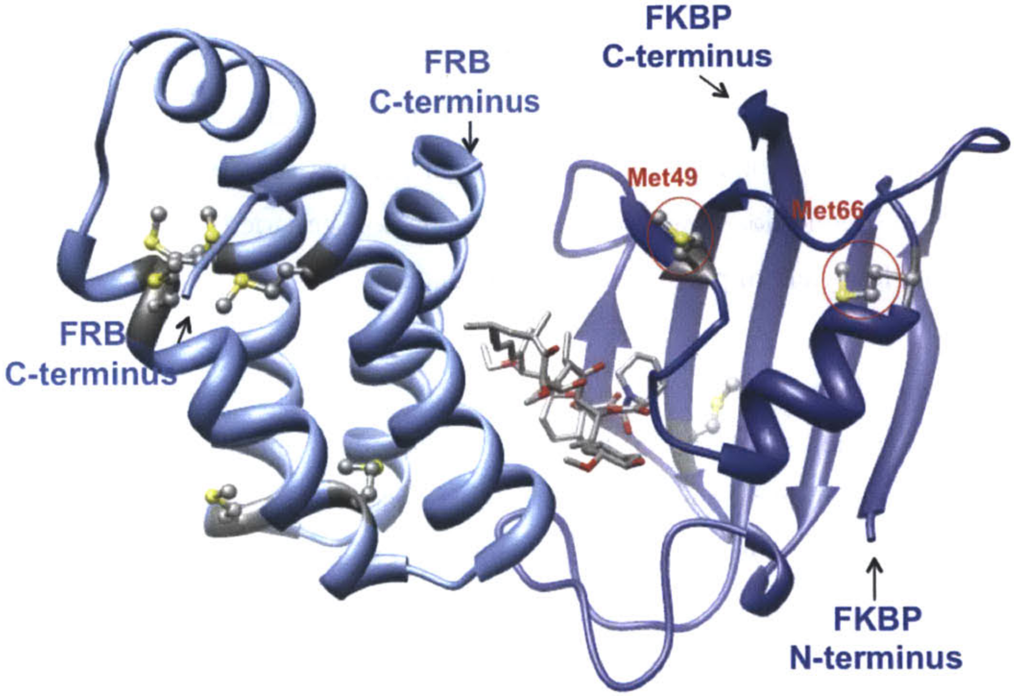


Figure 4-12. Crystal structure of FRB-rapamycin-FKBP ternary complex at 2.2 Å resolution (PDB 2FAP) with methionine residues in ball-and-stick. The termini of FRB and FKBP are indicated by black arrows. Two methionines (Met49, Met66) most proximal to the C-terminus of FKBP are circled in red. The LAP peptide is fused to the C-terminus of FKBP.

To test the hypothesis of the "magnet effect" of methionine, we cloned single and double alanine mutants of Met49 and Met66 in FKBP-LAP. Crystal structures as well as biophysical analysis of the FKBP-rapamycin-FRB ternary complex show that these two methionine residues are not involved in the binding of FKBP to rapamycin and there is little interaction between FKBP and FRB (61–63). Thus these mutations should not affect ternary complex formation. Three new FKBP-LAP constructs, FKBP(M49A)-LAP, FKBP(M66A)-LAP, and double mutant FKBP(M49A, M66A)-LAP were cloned, expressed and purified from *E.coli*. These proteins were *in vitro* ligated with benzophenone probe using purified LplA(W37G) enzyme (see Experimental Methods), and incubated with FRB-mApple (YFP was changed to mApple to reduce the amount of homodimer formation caused by YFP as seen previously in Figure 4-11) and rapamycin to form the ternary complex. However, photocrosslinking experiments using three methionine to alanine mutant proteins of FKBP-LAP did not increase heterodimer formation as shown in Figure 4-13. The band corresponding to FBKP-FRB heterodimer product is indicated by a red arrow. Photocrosslinking of FKBP(M66A)-LAP and FRB-mApple gave the highest heterodimer product signal (Figure 4-13, lane 3).

UV irradiation	-	+	+	+	+	+	+
FKBP(M49A)-LAP	-	+	-	-	-	-	-
FKBP(M66A)-LAP	-	-	+	-	-	-	-
FKBP(M49A, M66A)-LAP	-	-	-	+	-	-	-
FKBP-LAP	+	-	-	-	-	+	+
benzophenone	+	+	+	+	+	-	+

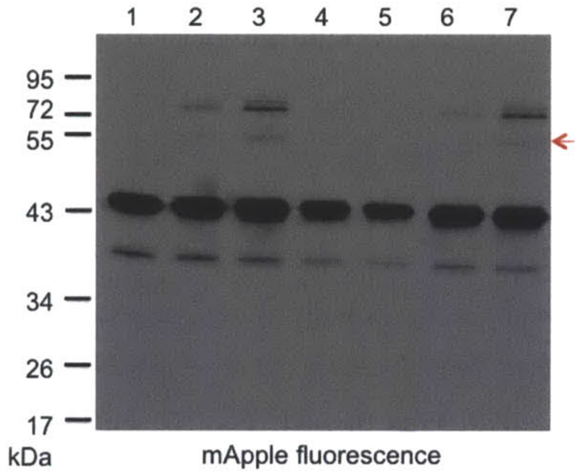


Figure 4-13. *In vitro* photocrosslinking with methionine to alanine mutant proteins of FKBP-LAP to avoid methionine quenching of benzophenone ketyl diradical. Lanes 2, 3 and 4 contain three methionine to alanine mutants of FKBP-LAP photocrosslinked to FRB-mApple. Photocrosslinked FKBP/FRB heterodimer product is indicated by a red arrow. Compared to the wild-type FKBP-LAP, methionine mutants of FKBP-LAP did not help to significantly increase crosslinking efficiency with FRB.

2. *Removal of fluorescent proteins fused to FRB and FKBP*

It is conceivable that the YFP or mApple fused to the C-terminus of FRB could hamper movement of the proteins as they interact with each other upon rapamycin stimulation, causing inefficient ternary complex formation. Furthermore, YFP caused significant amount of homodimer formation of FRB-YFP, which might have reduced the effective concentration of FRB-YFP that could interact with rapamycin and FKBP-LAP. A new FRB construct with HA epitope tag fused to the N-terminus of FRB was cloned, expressed, and purified from *E.coli*. Similarly, benzophenone-ligated myc-FKBP-LAP (see Experimental Methods) was incubated with HA-FRB and rapamycin, and subsequently irradiated with UV light. Samples were separated on a 12% SDS-PAGE gel and transferred to a nitrocellulose membrane for Western blot analysis. FKBP-LAP was detected on Western blot with antibody against the myc tag fused to the N–terminus of FKBP.

In Figure 4-14, photocrosslinked product of benzophenone-ligated myc-FKBP-LAP and HA-FRB is boxed in red. The crosslinked complex is dependent on UV irradiation, but shows a weak background under both rapamycin omitted (lane 2) and benzophenone omitted (lane 4) conditions. We suspected that the rapamycin-independent background was caused by the high concentration (10 μM) of proteins in the *in vitro* reaction mixture. We repeated the same experiment with lower concentrations of FKBP and FRB. Rapamycin-dependent formation of UV photocrosslinked product was achieved if concentration of the proteins were decreased to 2 μM (see Figure 4-20 in later section). The weak background seen in the absence of benzophenone (Figure 4-14, lane 4) may be caused by crosslinking through Tyr2105 on FRB (PDB 2FAP), which points directly into the interface of the ternary complex. Furthermore, even though *E. coli* purified myc-FKBP-LAP appeared to be pure by Coomassie staining (data not shown), Western blot staining with anti-myc antibody shows a nonspecific antibody staining background (Figure 4-14, green arrow) of myc-FKBP-LAP. In addition, myc-FKBP-LAP appeared to form higher molecular weight aggregates (Figure 4-14, blue arrow) in the presence of the benzophenone photocrosslinker. To try to remove the anti-myc nonspecific staining of myc-FKBP-LAP, we performed a second Ni-NTA column purification of the His6-tagged myc-FKBP-LAP protein, but the background band still persisted (Figure 4-14, green arrow). For the benzophenone-dependent higher molecular weight aggregates of myc-FKBP-LAP protein (Figure 4-14, blue arrow), we tried to remove excess benzophenone through extensive buffer exchange using minispin columns instead of dialysis, but the background still persisted.

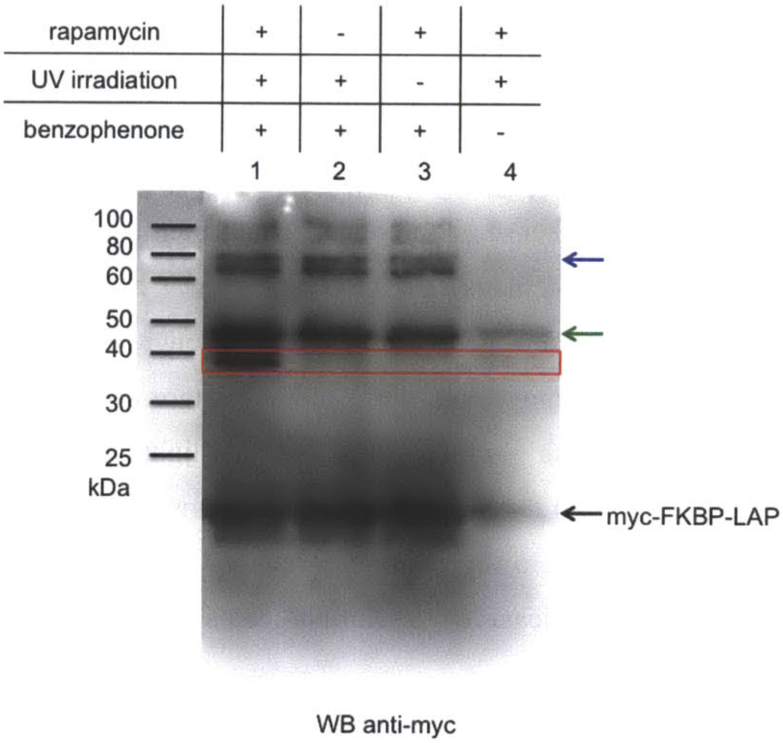


Figure 4-14. *In vitro* photocrosslinking of myc-FKBP-LAP (17.1 kDa) and HA-FRB (14.2 kDa). Fluorescent protein fusions on FRB were replaced by the smaller epitope tag, human influenza hemagglutinin epitope (HA). Western blot analysis with antibody against myc tag on FKBP detect photocrosslinked product (boxed in red), but also weak background crosslinking (green arrow) in the absence of rapamycin (lane 2) and benzophenone (lane 4). The myc-FKBP-LAP protein also formed higher molecular weight complexes (blue arrow) in the presence of benzophenone (lanes 1-3).

3. *LAP placement on FRB instead of FKBP*

Since *in vitro* photocrosslinking of FRB and LAP-tagged FKBP gave very minimal photocrosslinked product, we thought perhaps the position of the LAP peptide, and therefore the placement of benzophenone probe, was not ideal for the photocrosslinker to react with nearby C-H bonds on FRB. Looking at the crystal structure of the ternary complex in Figure 4-12, we observed that fusing LAP to the C-terminus of FRB might bring it closer to the interface of the interacting proteins, therefore providing benzophenone with more accessibility to crosslink to C-H bonds on FKBP. We switched the fusion site of LAP from the C-terminus of FKBP to the C-terminus of FRB to test this hypothesis. We further added a 10-amino acid linker (GSGSTSGSGK) between LAP and FRB to increase the spatial reachability of benzophenone. The new protein pairs we tried to photocrosslink are FKBP/FRB-LAP and FKBP/FRB-(10 a.a. linker)-LAP. Following the same protocol as previously described, we first *in vitro* ligated benzophenone onto FRB-LAP or FRB-(10 a.a. linker)-LAP protein using purified LplA(W37G) protein, ATP and Mg^{2+}(see Experimental Methods). Ligation product was confirmed using the HPLC assay. We removed unligated benzophenone probe through dialysis. Afterwards, we added FKBP protein and rapamycin to the reaction mixture to induce ternary complex formation, then irradiated the samples with 365 nm light. After photocrosslinking, samples were separated on a 12% SDS-PAGE gel and transferred to a nitrocellulose blot for immunostaining analysis. However, immunostaining against the myc epitope tag on FKBP did not show any significant increase of photocrosslinked complex as compared to previous results (data not shown).

At this stage of the project, we have tried many photocrosslinking conditions to try to increase the *in vitro* photocrosslinking yield, but none of the attempts was able to provide high yield of the FKBP and FRB photocrosslinked product without further increasing background. We decided to take a closer look at the design of our benzophenone photocrosslinking probe. In the next section, we describe the synthesis of a new benzophenone probe and the reason behind changing to a new photocrosslinker.

Synthesis and LplA(W37G) incorporation of alkyl benzophenone

Since none of the previously described attempts to increase the photocrosslinking yield of the FKBP/FRB protein-protein interaction system was successful, we thought perhaps the problem lay with the photochemistry of the benzophenone probe. We went back to examine the synthetic design of the benzophenone photocrosslinker. The chemistry of benzophenone UV activation involves first absorption of a photon at 300-365 nm, resulting in the promotion of one electron from the nonbonding n-orbital on oxygen to an antibonding π^* orbital of the carbonyl group. In this ketyl diradical state, the half-empty n-orbital on oxygen causes oxygen to be electron-deficient and abstract a nearby hydrogen atom from a C-H bond, leaving a carbon or alkyl radical. Subsequently, the benzophenone ketyl radical readily combines with the alkyl radical near it to form a C-C bond (64). A scheme of benzophenone insertion into nearby C-H bond is shown in Figure 4-15.

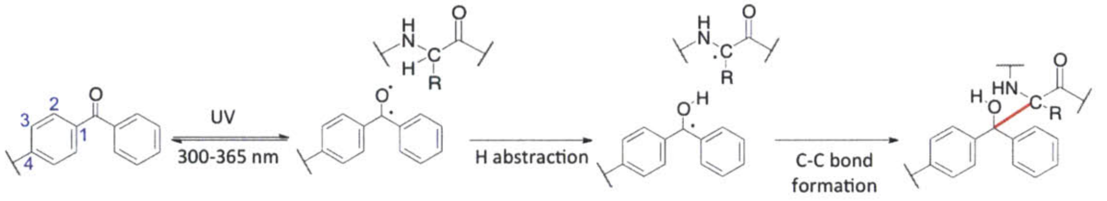


Figure 4-15. Scheme of benzophenone insertion into nearby C-H bond after UV irradiation. Numbering of carbon positions in benzophenone is indicated in blue. Benzophenone can be repeatedly activated to ketyl diradical via absorption of a photon at 300-365 nm. After UV activation, if a nearby C-H bond were at the optimal geometry and distance from the ketyl diradical, the ketyl diradical would abstract a hydrogen atom from it, leaving behind a carbon radical. If the ketyl radical and the carbon radical still remain at the optimal interactive distance, they would combine to form a new C-C bond (indicated in red).

Although detailed mechanisms and kinetics relating to benzophenone radical formation and photocrosslinking remain largely unknown, we examined benzophenone-based crosslinking reagents described in literature (64). Two main groups of crosslinkers were used in biological experiments. The first group consists of modification of small molecules and natural ligands as photolabile analogs used to identify their interaction sites on protein receptors and macromolecular complexes. The second group includes modifying specific amino acids of a protein of interest into photocrosslinking moieties, such as the introduction of benzophenone through unnatural amino acid mutagenesis and site-specific chemical conjugation of a crosslinker via cysteine-maleimide or cysteine-iodoacetamide chemistry (65). We noticed that many benzophenone-based crosslinkers have the benzophenone reactive core structure directly connected to a simple methylene unit (64), which is different from our benzophenone probe that has a carbonyl functional group directly attached to the fourth position of the benzophenone ring (Figure 4-15). We suspected that perhaps the carbonyl group would interrupt the electronics of radical generation by participating in the triplet excited state, or act as an electron-withdrawing group that could cause the ketyl radical to delocalize.

In order to test this theory and confirm radical generation of the benzophenone probe, we took transient absorption (TA) spectra of four different benzophenone compounds (Figure 4-16) with the help of graduate student Arturo Pizano in the lab of Professor Dan Nocera. Compound A has no attachment at the benzophenone ring; Compounds B, the unnatural amino acid benzophenone, has a methyl group attached to the fourth position of the benzophenone core; and Compounds C and D both have carbonyl groups directly attached to the benzophenone ring, Compound D is our benzophenone probe used in previous experiments. All four compounds produced TA spectra with a sharp peak at 330 nm and a broad peak at 550 nm with a slight shoulder to the blue region, which are all characteristics of a reduced benzophenone ketyl radical (66). It is reassuring to know that ketyl radical can be generated. However, Compounds C and D both showed much higher absorbance at 330 nm than Compounds A and B (data not shown). Previous work from the Tanaka group demonstrated that having a higher 330 nm to 550 nm absorbance ratio is an indication of greater ketyl radical character (66). Based on our result, we can state that having a carbonyl group directly attached to the benzophenone core made Compounds C and D have more ketyl radical character than Compounds A and B. But further experiments would be needed to investigate the exact effect of greater ketyl radical character on the photocrosslinking efficiency of benzophenone crosslinkers. We decided to design and synthesize a new benzophenone probe that does not have a functional group immediately linked to the benzophenone reactive core.

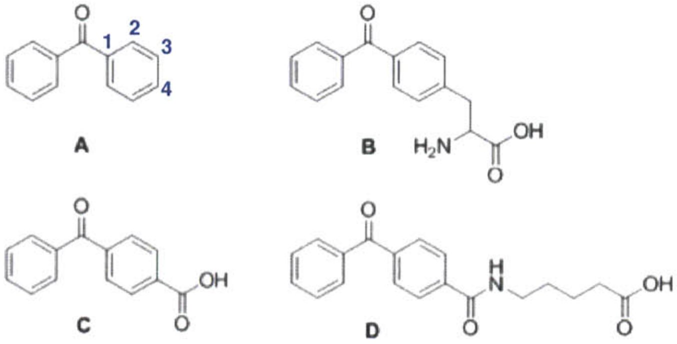


Figure 4-16. Benzophenone (A) and benzophenone derivatives (B-D) with different functional groups attached to the fourth position of the benzophenone core. Numbering of carbon positions in benzophenone is indicated in blue.

Following the synthetic scheme shown in Figure 4-17, we synthesized a new benzophenone probe with a methylene group directly attached to the benzophenone ring through Friedel-Craft acylation of benzoyl chloride and phenyl valeric methyl ester, followed by hydrolysis of the methyl ester to generate the corresponding carboxylic acid (see Experimental Methods). We obtained both para and ortho products in approximately 5:1 ratio. The products were characterized by ^{1}H NMR and mass spectrometry.

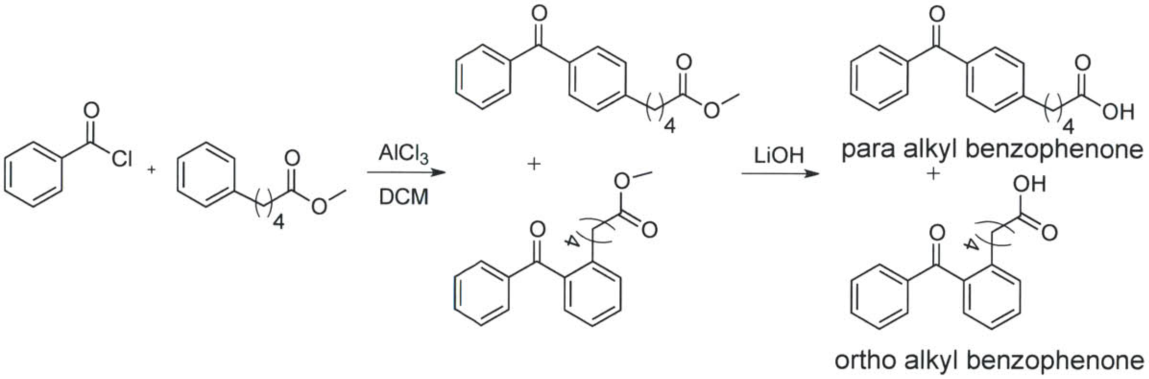


Figure 4-17. Synthetic scheme of alkyl benzophenone probes with methylene unit attached to the benzophenone core. The para and ortho products were obtained in a 5:1 ratio.

Aside from changing the design of the benzophenone photocrosslinker, we also modified the LAP acceptor peptide by changing the Trp residue located two amino acids downstream of the catalytic Lys residue to Phe residue, thus creating LAP(F) (GFEIDKV**F**YDLDA). The aromatic Trp residue is well known as a mediator of electron transfer between distant electron acceptor and donor centers in proteins (67). So we suspected the Trp residue near the ligated benzophenone might quench the formation of ketyl radicals by acting as an electron transfer site.

To ensure that the new alkyl benzophenone probes can be incorporated by LplA enzyme, both para and ortho alkyl benzophenone probes were tested in *in vitro* ligation reactions using purified LplA(W37G) enzyme and LAP or LAP(F) peptide. Either para or ortho alkyl benzophenone probe was incubated with purified LplA(W37G) protein, LAP or LAP(F) peptide, ATP and Mg^{2+} for 3 hours at 30 °C. Only para alkyl benzophenone was incorporated and ligated by LplA(W37G) (Figure 4-18). Perhaps the orientation of the benzophenone core with respect to the alkyl chain in ortho alkyl benzophenone prevented it from fitting into the rather straight binding tunnel of LplA (47).

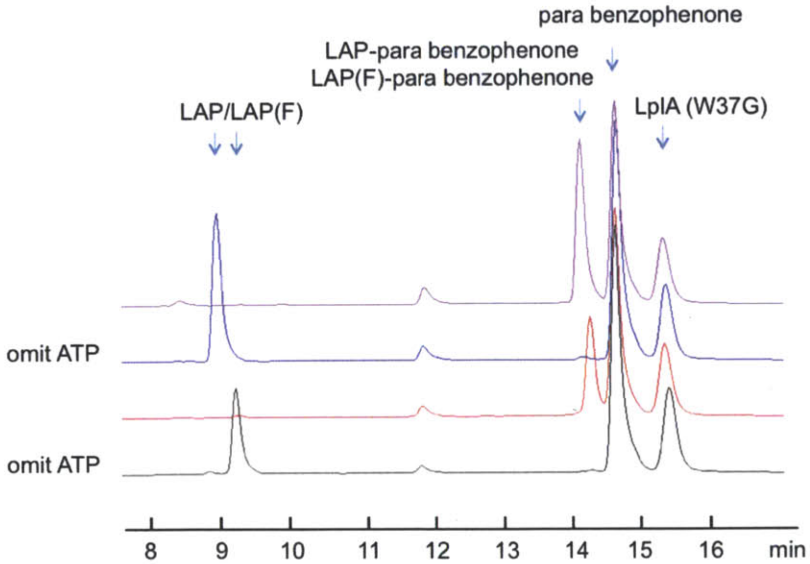
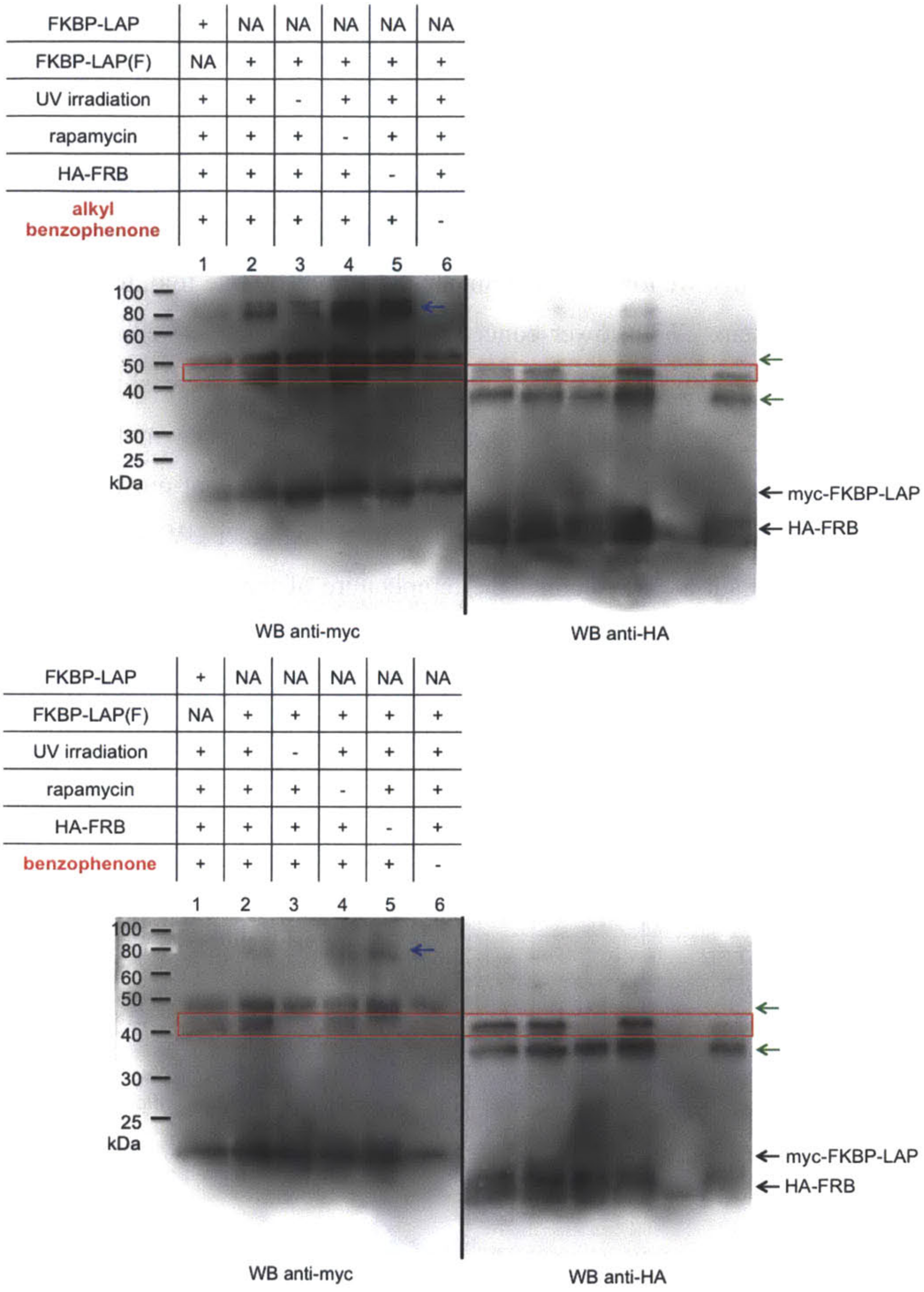


Figure 4-18. *In vitro* characterization of para alkyl benzophenone ligation by LplA(W37G). HPLC trace showing LAP (pink trace) or LAP(F) (red trace) peptide conversion to the benzophenone conjugate. Negative controls are shown with ATP omitted (black and blue trace). Reactions were performed for 3 hours with 1 μM LplA(W37G), 100 μM LAP or LAP(F), and 300 μM para alkyl benzophenone.

In vitro photocrosslinking with new alkyl benzophenone probe

We wished to compare the new alkyl benzophenone photocrosslinker with our previous benzophenone probe. *In vitro* we ligated either 10 μM FKBP-LAP or FKBP-LAP(F) protein with one of the photocrosslinkers using purified LplA(W37G) protein, ATP, and Mg^{2+} (see Experimental Methods). After ligation was complete, unligated photocrosslinker probe was removed through dialysis. We then added 10 μM FRB protein and 500 nM rapamycin to the reaction mixture to form the ternary complex. Samples were exposed to UV light using the Porta-Ray 400R instrument from UVitron for various lengths of time from 1 to 5 minutes. Samples were kept on ice during UV irradiation time to prevent heat-induced protein aggregation. Samples were then analyzed by Western blot using antibodies against the myc epitope tag on FKBP-LAP and the HA epitope tag on FRB. Coomassie stain of UV-irradiated samples showed proteins started to smear and aggregate if UV irradiation time is longer than 2 minutes using the Porta-Ray 400R which has very powerful UV lamps (data not shown). Thus, UV irradiation time was set to 2 minutes for subsequent experiments. Figure 4-19 (Top) shows *in vitro* photocrosslinking using the new benzophenone probe, alkyl benzophenone, while Figure 4-19 (Bottom) shows same experiment as on the top but using the original benzophenone probe.

First of all, even though both proteins appeared to be pure by Coomassie staining, myc-FKBP-LAP and HA-FRB both showed nonspecific antibody staining (indicated by green arrows) by mouse anti-myc and rabbit anti-HA antibodies, respectively. In addition, myc-FKBP-LAP appeared to form higher molecular weight aggregates (indicated by blue arrow) in the presence of the photocrosslinker. The backgrounds are the same as we have seen before in Figure 4-14. Photocrosslinking was tested using either benzophenone-ligated FKBP-LAP or benzophenone-ligated FKBP-LAP(F). The photocrosslinked product stained positive for both the myc tag on FKBP and the HA tag on FRB (lanes 1-2, boxed in red). However, we also saw background crosslinking in the absence of rapamycin or the photocrosslinker (Figure 4-19, lanes 4 and 6), which we have also observed before. These observations were the same for photocrosslinking carried out with either the new alkyl benzophenone probe (Figure 4-19, top) or the original benzophenone probe (Figure 4-19, bottom). The benzophenone-independent background (Figure 4-19, lane 6) could be caused by photocrosslinking via the Tyr2105 residue on FRB (PDB 2FAP), located at the interface of the FKBP/FRB complex. We suspected that the rapamycin-independent background (Figure 4-19, lane 4) could be caused by the high concentration (10 μM) of proteins in the *in vitro* reaction mixture. We repeated the same experiment with lower concentrations of FKBP and FRB. Rapamycin-dependent formation of UV photocrosslinked product was achieved if concentration of the proteins were decreased to 2 μM (Figure 4-20). In summary, alkyl benzophenone and benzophenone, appeared to perform similarly during *in vitro* crosslinking of purified proteins.



FKBP-LAP	+	NA	NA	NA	NA	NA
FKBP-LAP(F)	NA	+	+	+	+	+
UV irradiation	+	+	-	+	+	+
rapamycin	+	+	+	-	+	+
HA-FRB	+	+	+	+	-	+
alkyl benzophenone	+	+	+	+	+	-

FKBP-LAP	+	NA	NA	NA	NA	NA
FKBP-LAP(F)	NA	+	+	+	+	+
UV irradiation	+	+	-	+	+	+
rapamycin	+	+	+	-	+	+
HA-FRB	+	+	+	+	-	+
benzophenone	+	+	+	+	+	-

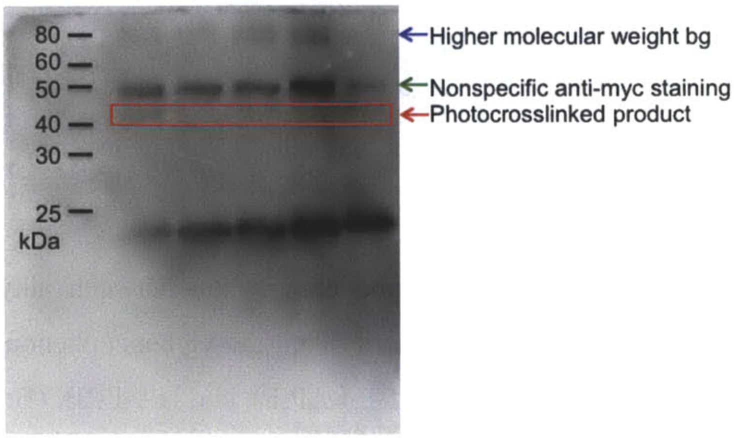
1
2
3
4
5
6
100
80
60
50
40
30
25
kDa
myc-FKBP-LAP
HA-FRB
WB anti-myc
WB anti-HA

Figure 4-19. *In vitro* photocrosslinking comparison between alkyl benzophenone and benzophenone. (Top) Photocrosslinking using alkyl benzophenone probe and purified proteins myc-FKBP-LAP or myc-FKBP-LAP(F), and HA-FRB. (Bottom) Same as (Top), except the original benzophenone probe was used. Each LAP or LAP(F)-tagged FKBP protein (10 μM), and FRB protein (10 μM) was used in this experiment. Photocrosslinked

complex is boxed in red; nonspecific immunostaining background is indicated by green arrows; higher molecular weight complex of FKBP-LAP or FKBP-LAP(F) is indicated by blue arrow.

We further tried to optimize the system by lowering the concentrations of interacting proteins to reduce background photocrosslinking. Instead of using 10 μM of FKBP and FRB proteins, we lowered protein concentrations by five fold, to 2 μM of each FKBP and FRB protein. With lower concentration of proteins, we were able to obtain rapamycin-dependent and alkyl benzophenone-dependent photocrosslinking product (Figure 4-20, lanes 3 and 5). However, we still saw a higher molecular weight background (Figure 4-20, blue arrow) that appears to be dependent on the presence of benzophenone. We tried to reduce this background by trying other UV lamps, decreasing the time of UV irradiation, and keeping the temperature of the samples cool by storing the samples on ice during UV irradiation, but the background still persisted. We were also unable to remove the nonspecific anti-myc staining (Figure 4-20, green arrow) even though myc-FKBP-LAP(F) appeared to be pure by Coomassie staining.

FKBP-LAP(F)	+	+	+	+	+
UV irradiation	+	-	+	+	+
rapamycin	+	+	-	+	+
HA-FRB	+	+	+	-	+
alkyl benzophenone	+	+	+	+	-
	1	2	3	4	5


80
60
50
40
30
25
kDa
←Higher molecular weight bg
←Nonspecific anti-myc staining
←Photocrosslinked product

WB anti-myc

Figure 4-20. Reducing concentrations of FKBP-LAP(F) and FRB helps in eliminating background photocrosslinking. Alkyl benzophenone-ligated FKBP-LAP(F) (2 μM) was incubated with 2 μM FRB and 500 nM rapamycin before UV irradiation for 1 minute using the Porta-Ray 400R crosslinking instrument. Western blot immunostaining with anti-myc to detect the myc tag on FKBP-LAP(F) showed crosslinked product (boxed in red) which is dependent on UV irradiation, rapamycin, FRB, and alkyl benzophenone.

After extensive optimization of photocrosslinking conditions, we were able to obtain photocrosslinking product of FKBP-LAP and FRB that was dependent on UV irradiation, rapamycin, FRB and alkyl benzophenone. However, we were not able to increase the yield of the crosslinking product without increasing background at the same time. It is possible that the concentrations of the interacting protein partners and the small molecule rapamycin were difficult to fine-tune to the most optimized state, where most proteins were in the ternary complex form. Since our ultimate goal was to perform photocrosslinking inside live mammalian cells, which may require a completely different set of photocrosslinking conditions from the conditions for *in vitro* experiments, we decided to move the FKBP/FRB photocrosslinking system to test and further optimize inside live cells.

Quantify benzophenone ligation yield inside live mammalian cells

The first step in photocrosslinking inside live cells was to ensure benzophenone ligation onto LAP-tagged protein substrate is quantitative. To determine the yield of benzophenone ligation by LplA(W37G) inside live cells, we applied the same native gel-shift assay first described in Chapter 2 for estimating azide ligation yield. The same principle applies that after a small molecule probe has been ligated onto the catalytic lysine residue in the LAP substrate, changing the positively charged lysine to a neutral amide, the small molecule-conjugated LAP now migrates faster than the unligated LAP on native gel. Therefore, the population of ligated LAP can effectively separate itself from the free LAP. The assay was tested on benzophenone ligated LAP to ensure that the same method can be applied to other small molecules aside from azide described in Chapter 2. Different LAP substrates were also tested to make certain that the native gel migration pattern was generalizable to all LAP conjugated fluorescent proteins and other proteins of interest.

Three LAP substrates were tested as shown in Figure 4-21 – LAP-tagged YFP (left), LAP-tagged mCherry (middle), and LAP-tagged mCherry-FKBP (right). In each experiment, HEK 293T cells were transfected with the indicated LplA enzyme and LAP substrate, incubated with benzophenone probe for 3 hours, followed by washing to remove unligated probe for 1 hour. Cells were then lysed with hypotonic lysis buffer and separated on 12% polyacrylamide native gel. In-gel fluorescence imaging of LAP-tagged fluorescent proteins was used to analyze and detect the position of ligated and unligated LAP substrates. Figure 4-21 (left, lane 1) shows benzophenone ligated LAP-YFP migrates faster than unlabeled LAP-YFP (Figure 4-21, left, lane 3 and 4) due to removal of a positive charge. In addition, negative control with the lysine in LAP mutated to alanine, LAP(Ala) (Figure 4-21, left, lane 2), also ran faster on the native gel because the positive charge has been removed. The assay is also generalizable across various LAP fusion proteins, such as LAP-tagged mCherry (Figure 4-21, middle) and LAP-tagged FKBP (Figure 4-21, right).

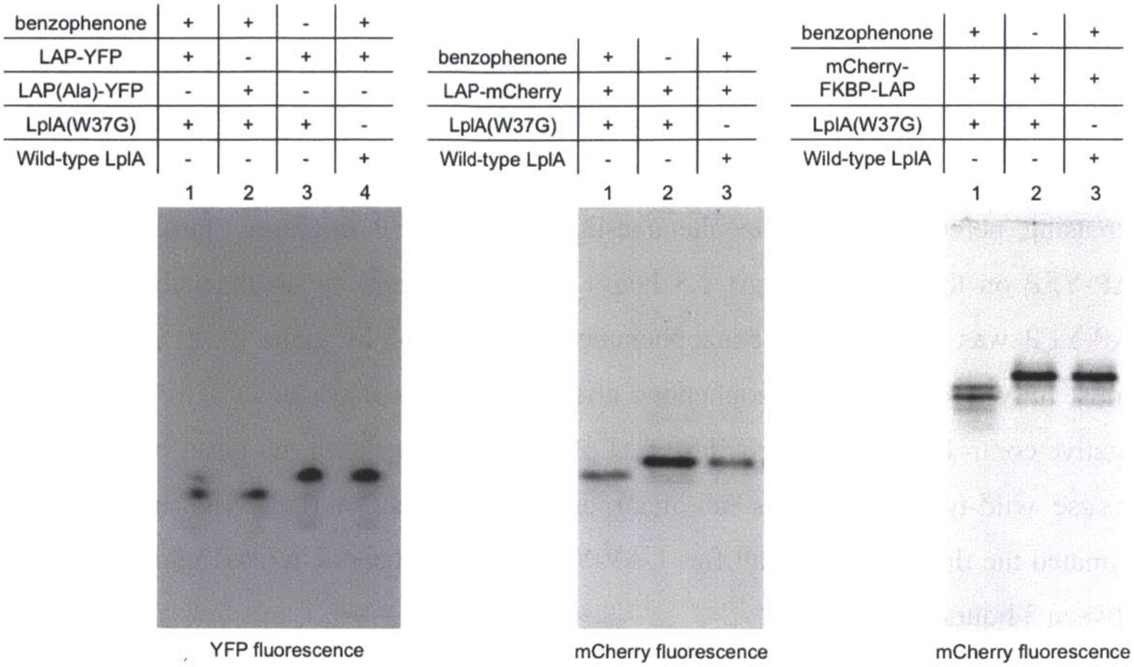

benzophenone + + - +
LAP-YFP + - + +
LAP(Ala)-YFP - + - -
LplA(W37G) + + + -
Wild-type LplA - - - +
1 2 3 4
YFP fluorescence
benzophenone + - +
LAP-mCherry + + +
LplA(W37G) + + -
Wild-type LplA - - +
1 2 3
mCherry fluorescence
benzophenone + - +
mCherry-FKBP-LAP + + +
LplA(W37G) + + -
Wild-type LplA - - +
1 2 3
mCherry fluorescence

Figure 4-21. Gel-shift analysis of benzophenone ligation yield in cells. (left) Benzophenone ligation on LAP-YFP (lane 1). Negative controls are shown with catalytically inactive LAP (K→A)-YFP (lane 2), benzophenone omitted (lane 3), and wild-type LplA which does not incorporate benzophenone (lane 4). (middle) Benzophenone ligation on LAP-mCherry (lane 1). Negative controls are shown with benzophenone omitted (lane 2), and wild-type LplA, which does not incorporate benzophenone (lane 3). (right) Benzophenone ligation on mCherry-FKBP-LAP (lane 1). Negative controls are shown with benzophenone omitted (lane 2), and wild-type LplA, which does not incorporate benzophenone (lane 3).

Having established an assay to estimate benzophenone ligation yield, we performed a time course labeling experiment to determine the time it would take for LAP to achieve quantitative benzophenone ligation yield, which was necessary before proceeding to live cell photocrosslinking. We hoped to shorten the benzophenone incubation time to reduce the chance of probe nonspecifically sticking to hydrophobic areas inside the cell, causing high background crosslinking. HEK 293T cells expressing LAP-YFP and LplA(W37G) were incubated with benzophenone probe from 1 hour to 3.5 hours. At the end of each ligation time point, cells were washed three times with benzophenone-free serum-supplemented media and incubated in fresh media for an additional hour before cells were pelleted and lysed. After benzophenone washout, cells were lysed and separated on a 12% native gel with DTT-free and SDS-free protein loading buffer to preserved native structure of the protein. Figure 4-22 shows that as benzophenone ligation time progressed from 1 hour to 3.5 hours (lanes 1-3), we see an increasing percentage of benzophenone-ligated LAP-YFP migrating faster than free LAP-YFP on the native gel. At 2.5 hours of benzophenone incubation, about 52% of LAP-YFP was converted to benzophenone-ligated LAP-YFP (lane 2). LAP-YFP was almost fully ligated with benzophenone after about 3.5 hours (lane 4). LAP-YFP in the negative control ligation using wild-type LplA does not appear to run faster on native gel because wild-type LplA does not incorporate benzophenone (lane 4). Therefore, we estimated the time to convert all free LAP-YFP to benzophenone-ligated LAP-YFP to be between 3 hours to 3.5 hours.

benzophenone	+	+	+	+
LAP-YFP	+	+	+	+
LplA(W37G)	+	+	+	-
Wild-type LplA	-	-	-	+
benzophenone incubation time (hr)	1	2	3.5	3.5

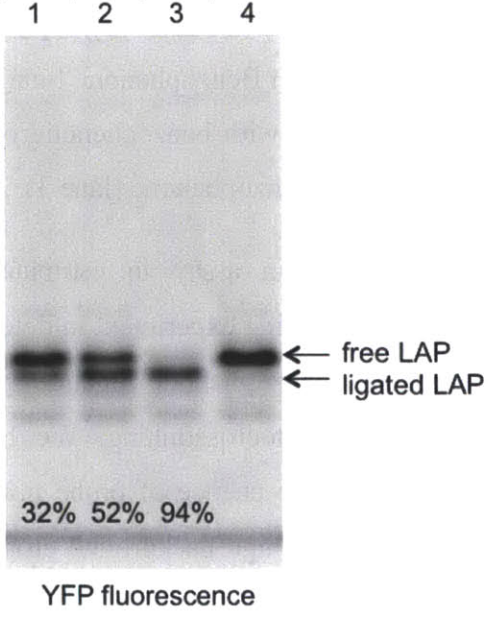


Figure 4-22. Native gel-shift assay to measure benzophenone ligation yield over time. HEK 293T cells expressing the indicated LplA enzyme and LAP-YFP were incubated

with 300 μM benzophenone for 1 hour, 2 hours, or 3.5 hours. Cell lysates were analyzed on a 12% polyacrylamide native gel. Using YFP in-gel fluorescence as readout, lanes 1-3 shows an increasing percentage of benzophenone-ligated LAP-YFP. Estimated percent conversions to product are given at the bottom of the YFP fluorescence gel image. LAP-YFP in the negative control reaction (lane 4) with HEK 293T cells expressing wild-type LplA does not run faster because wild-type LplA does not incorporate benzophenone.

Photocrosslinking inside live mammalian cells

Having established that benzophenone ligation onto LAP-tagged protein can be quantitative inside live cells, we moved forward with performing photocrosslinking inside live mammalian cells using both alkyl benzophenone and benzophenone probes. In live cell ligation and photocrosslinking, we needed to ensure optimal expression of the LplA(W37G) enzyme and the LAP-tagged protein substrate. UV irradiation time must be ideal for the radical to form covalent bonds with nearby amino acids but not too long to cause cell aggregation or cell death. In addition, live cells should be carefully kept on ice during UV irradiation to prevent heat-induced protein aggregation, which could cause false positives.

We first tried live-cell photocrosslinking on the FKBP-LAP and FRB protein-protein interaction system. HEK 293T cells were transfected with three plasmids – FLAG-tagged LplA(W37G), myc-tagged FKBP-LAP, and HA-tagged FRB. We checked for the expression of each protein via immunostaining against the epitope tag on each protein (see Experimental Methods). Figure 4-23 shows the expressions of myc-FKBP-LAP and HA-FRB proteins. We can see that both proteins exhibited strong fluorescence signals through the entire cell, with and without LplA(W37G) co-expression. Since the FLAG tag on LplA(W37G) and the myc tag on FKBP-LAP share the same secondary antibody, we can only immunostain for one of the two proteins with HA-FRB at the same time. Separate immunostaining against the FLAG tag on LplA(W37G) confirmed that the enzyme was also expressing well in the presence of FKBP-LAP and FRB (data not shown).

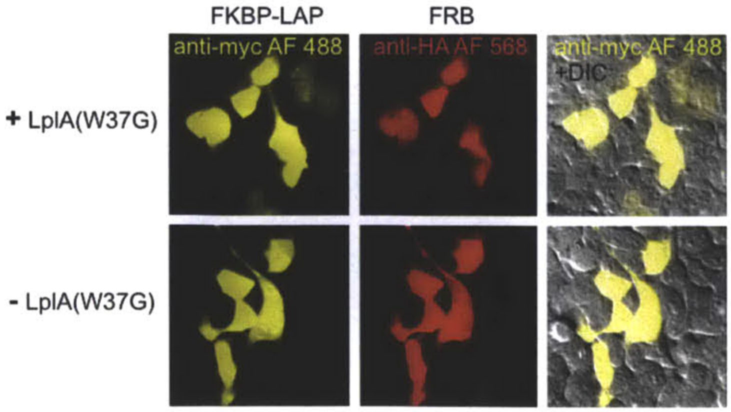


Figure 4-23. Expression of FKBP-LAP and FRB in HEK 293T cells. HEK 293T cells were transfected with FKBP-LAP, FRB and LplA(W37G) plasmids. At 18 hours after transfection, cells were fixed and immunostained with antibodies against the myc epitope tag on FKBP-LAP and the HA epitope tag on FRB. Both proteins showed strong expression in the presence and absence of LplA(W37G) protein.

Following transfection, live cells were incubated in cell culture media containing either 300 μM alkyl benzophenone or benzophenone probe for up to 4 hours. Cells were then washed with fresh media for 1 hour, so excess unligated benzophenone can be removed from cells via the action of organic anionic transporters. Cells were further incubated in fresh media containing 500 nM rapamycin for an additional hour to stimulate ternary complex FKBP-rapamycin-FRB formation. Finally, cell culture media was replaced with ice-cold DPBS buffer and the cells were placed on ice before exposure to 365 nm UV light for either 2 minutes using the Porta-Ray 400R instrument at 2 inches below the lamp or 12 minutes using long-wavelength handheld UV lamp with the lamp placed directly on top of the cell culture plate. After UV irradiation, we lysed the cells using radio-immunoprecipitation assay (RIPA) buffer. The lysates were analyzed by Western blot with anti-myc antibody staining to detect myc-FKBP-LAP and its crosslinked complexes (Figure 4-24). Lanes 1 and 5 in Figure 4-24 are the experimental lanes. The negative control reactions were performed with either photocrosslinker omitted (lanes 2 and 6), LplA(W37G) omitted (lane 3 and 7), or UV irradiation omitted (lane 4 and 8).

alkyl benzophenone	+	-	+	+	-	-	-	-
benzophenone	-	-	-	-	+	-	+	+
FKBP-LAP	+	+	+	+	+	+	+	+
LplA(W37G)	+	+	-	+	+	+	-	+
UV irradiation	+	+	+	-	+	+	+	-
	1	2	3	4	5	6	7	8

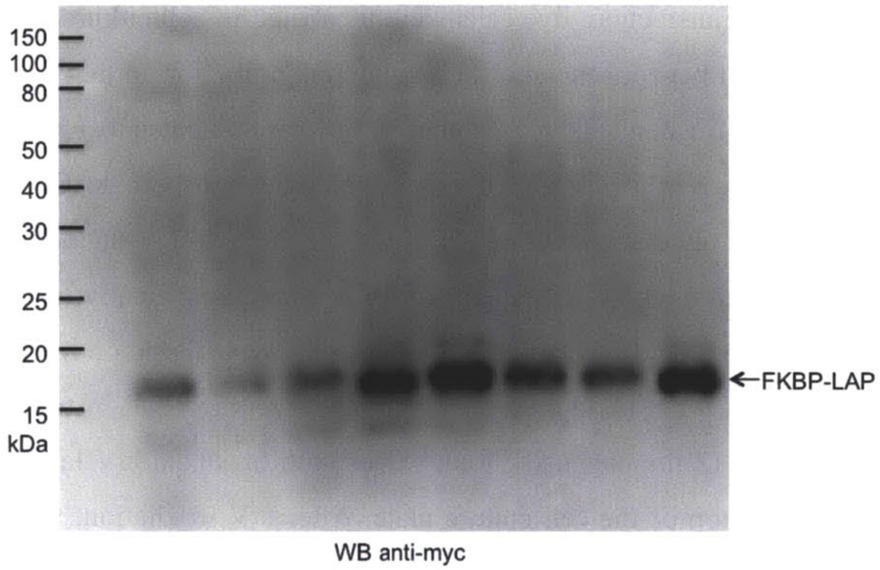


Figure 4-24. Live cell benzophenone ligation and photocrosslinking of myc-FKBP-LAP and HA-FRB. HEK 293T cells expressing myc-FKBP-LAP, HA-FRB, and LplA(W37G) were labeled with alkyl benzophenone or benzophenone and incubated with rapamycin to stimulate complex formation. Live cells were irradiated with 365 nm UV light for 2 minutes. Cells were then lysed and analyzed with anti-myc staining to detect the myc tag on FKBP-LAP.

Live cell photocrosslinking with myc-FKBP-LAP and HA-FRB did not show any crosslinked product, which was expected at around 40 kDa for the heterodimer. We only detected myc-FKBP-LAP protein at 17.1 kDa on the Western blot and did not observe any other myc-positive bands. Since we confirmed previously that benzophenone ligation onto LAP protein substrate in live cells was quantitative under the current condition, it was possible that ligated benzophenone did not have an ideal interaction partner nearby. We can see from the crystal structure in Figure 4-12 that LAP tag fused to the C-terminus of FKBP placed ligated benzophenone in a position outside of the direct interacting interface between FKBP and FRB (61). Since benzophenone ketyl diradical prefers to have nearby C-H bonds at an optimal distance of about 2 to 4 Å (38) to have efficient crosslinking, the placement of the LAP tag put the photocrosslinker at a disadvantaged location. Another explanation is that since analysis of crystal structures of FKBP-rapamycin-FRB shows extensive interactions between rapamycin and its two protein partners, but relatively limited interactions between the proteins themselves, perhaps this system is not the most ideal for testing crosslinking with benzophenone, whose diradical species can only attack geometrically accessible C-H bonds (38,64).

Next, we decided to try a different LAP-tagged protein for live cell photocrosslinking experiment. We chose LAP-tagged actin because of actin's abundant interacting protein partners. In addition, actin dynamics are tightly regulated by numerous proteins, which associate with monomeric actin and/or filamentous actin. We were hopeful that benzophenone-labeled LAP-actin would photocrosslink to some of its interacting proteins, such as profilin which interacts preferentially with actin monomers; capping proteins which interact with barbed ends of actin filament assembly; and Arp2/3 complex which serve as nucleation sites for new actin filaments.

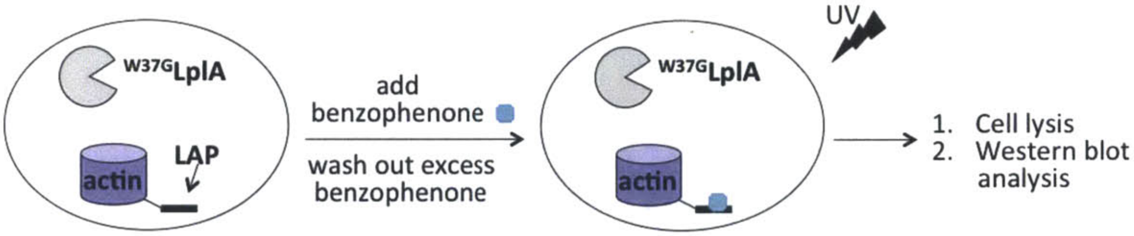


Figure 4-25. Benzophenone photocrosslinker labeling scheme of LAP-actin in live cells. HEK 293T cells expressing LplA(W37G) and LAP-tagged actin were incubated with benzophenone probe. After excess unligated benzophenone is removed, cells were irradiated with 365 nm UV light. Afterwards, cells were lysed and the lysate was analyzed by Western blot to detect crosslinked complexes of LAP-actin.

Following the scheme in Figure 4-25, we incubated HEK 293T cells expressing HA-LAP-actin and LplA(W37G) in cell culture media containing either 300 μM alkyl benzophenone probe or benzophenone probe for 3 hours. After washing the cells with fresh media to remove excess probe via the action of nonspecific anionic transporters, we replaced the culture media with ice-cold DPBS buffer and put the cells on ice before exposure to 365 nm UV light for 12 minutes using long-wavelength handheld UV lamp that was placed directly on top of the cell culture plate, at approximately 1 centimeter from the live cells. After UV irradiation, cells were lysed and the lysate was analyzed with anti-HA staining to detect HA-LAP-actin and its crosslinked complexes on Western blot (Figure 4-26).

alkyl benzophenone	+	+	-	+	-	-	-	+
benzophenone	-	-	-	-	+	+	-	-
HA-LAP-actin	+	-	+	+	+	-	+	+
HA-LAP(F)-actin	-	+	-	-	-	+	-	-
LplA(W37G)	+	+	+	-	+	+	+	+
UV irradiation	+	+	+	+	+	+	+	-

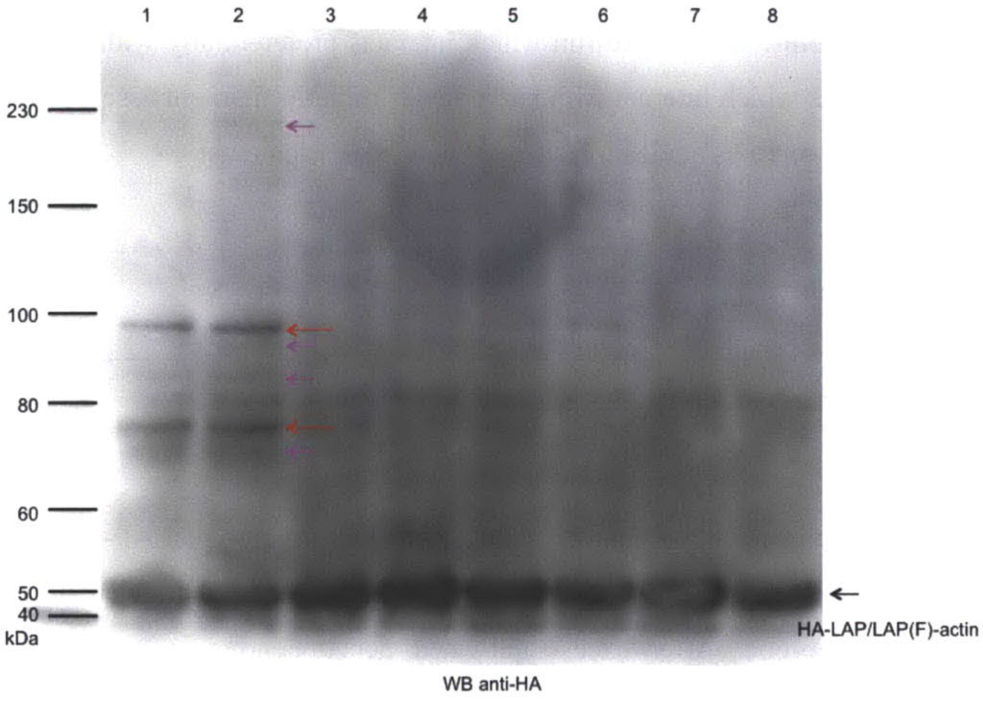
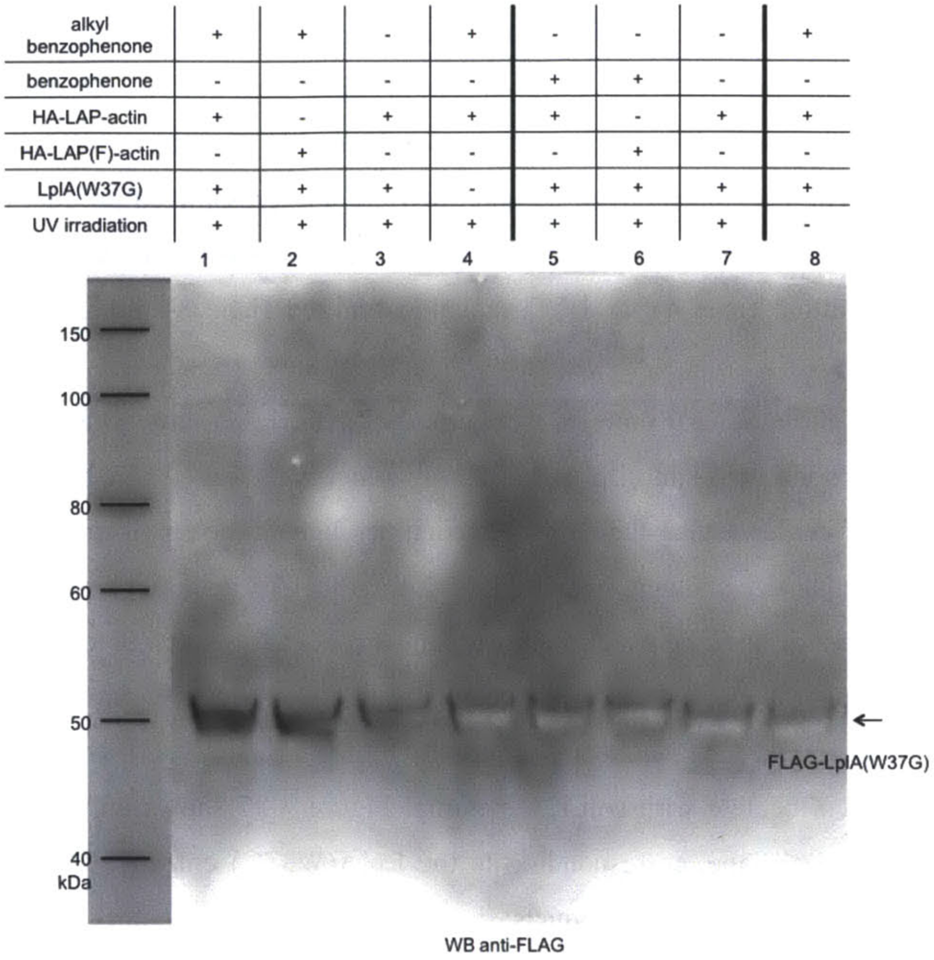


Figure 4-26. Benzophenone ligation and photocrosslinking on HA-LAP/LAP(F)-actin in live mammalian cells. After live cell ligation with either alkyl benzophenone or benzophenone probe and UV irradiation, cells were lysed and analyzed on Western blot with anti-HA antibody staining. Negative controls are shown with either benzophenone probe, LplA(W37G), or UV irradiation omitted.

We were glad to see that photocrosslinking of alkyl benzophenone ligated HA-LAP/LAP(F)-β-actin gave two major crosslinked products at about 75 kDa and 95 kDa indicated by red arrows and four minor products indicated by pink arrows (Figure 4-26, lanes 1 and 2). The benzophenone probe did not yield any crosslinked product (lanes 5 and 6). The negative control experiments with either photocrosslinker omitted (lane 3), LplA(W37G) omitted (lane 4), or UV irradiation omitted (lane 8) did not show any crosslinked bands.

We were curious what proteins were crosslinked to LAP-actin. One of the most obvious speculations was the LplA(W37G) enzyme. Previous LplA-based protein labeling gave us evidence that the enzyme sometimes ligates the small molecule probe onto itself. We stained another Western blot from the live cell alkyl benzophenone labeling and photocrosslinking experiment with anti-FLAG antibody to detect the FLAG-epitope tag on LplA(W37G). If any of the photocrosslinked product (Figure 4-26, red and pink arrows) were HA-LAP-actin crosslinked to FLAG-LplA(W37G), the anti-FLAG staining should co-localize with anti-HA staining. Figure 4-27 shows only one FLAG-positive band in each lane, corresponding to the LplA(W37G) enzyme. Therefore, the HA-positive, photocrosslinked products in Figure 4-26 were not simply HA-LAP/LAP(F)-actin crosslinked to the enzyme.



	1	2	3	4	5	6	7	8
alkyl benzophenone	+	+	-	+	-	-	-	+
benzophenone	-	-	-	-	+	+	-	-
HA-LAP-actin	+	-	+	+	+	-	+	+
HA-LAP(F)-actin	-	+	-	-	-	+	-	-
LplA(W37G)	+	+	+	-	+	+	+	+
UV irradiation	+	+	+	+	+	+	+	-

Figure 4-27. Western blot stained with anti-FLAG antibody to detect FLAG-tagged LplA(W37G) after live cell photocrosslinking of benzophenone-ligated HA-LAP/LAP(F)-actin. Experiment in Figure 4-26 was repeated to stain for FLAG-tagged LplA(W37G). No additional FLAG-positive band was detected aside from LplA(W37G) enzyme.

Knowing that we were able to obtain photocrosslinked complex of HA-LAP-actin, we wanted to see if we could reduce the benzophenone incubation time without sacrificing crosslinking signal. Lowering the photocrosslinker incubation time could decrease the chance of photocrosslinker sticking inside cells, which may cause unwanted background. We performed an alkyl benzophenone ligation time course experiment to determine the optimal probe incubation time to achieve highest amount of photocrosslinked products. HEK 293T cells expressing LplA(W37G) and HA-LAP-actin were incubated with alkyl benzophenone probe from 1 hour to 4 hours. At the end of each ligation time point, cells were washed and incubated in fresh cell culture media for 1 hour to remove excess probe. We prepared cell lysate from each time point and immunostained with anti-HA antibody to detect HA-LAP-actin. Figure 4-28 shows that 1 hour of alkyl benzophenone live-cell incubation (lane 1) was able to give comparable signal of photocrosslinked product as 4 hours of probe incubation (lane 4). However, we observed that negative control with alkyl benzophenone probe omitted shows a background photocrosslinked band at around 75 kDa (Figure 4-28, lane 6, green arrow). The intensity of the background band is weaker than that of the experiment (lanes 1-4). However, this background band is only present in the negative control with alkyl benzophenone omitted (lane 6, green arrow) and not in other negative controls with either UV irradiation omitted (lane 5), LplA(W37G) omitted (lane 7), with wild-type LplA (lane 8), or with the substrate LAP-actin omitted (lane 9). We speculated that the background could be caused by surface-exposed Tyr residues in actin that were also activated by the 365 nm UV light and crosslinked to nearby protein. Nonetheless, we were glad to see that HA-LAP-actin photocrosslinked products can be obtained with only 1 hour of live cell alkyl benzophenone incubation.

UV irradiation	+	+	+	+	-	+	+	+	+
alkyl benzophenone	+	+	+	+	+	-	+	+	+
LplA(W37G)	+	+	+	+	+	+	-	-	+
Wild-type LplA	-	-	-	-	-	-	-	+	-
LAP-actin	+	+	+	+	+	+	+	+	-
alkyl benzophenone incubation time (hr)	1	2	3	4	4	4	4	4	4

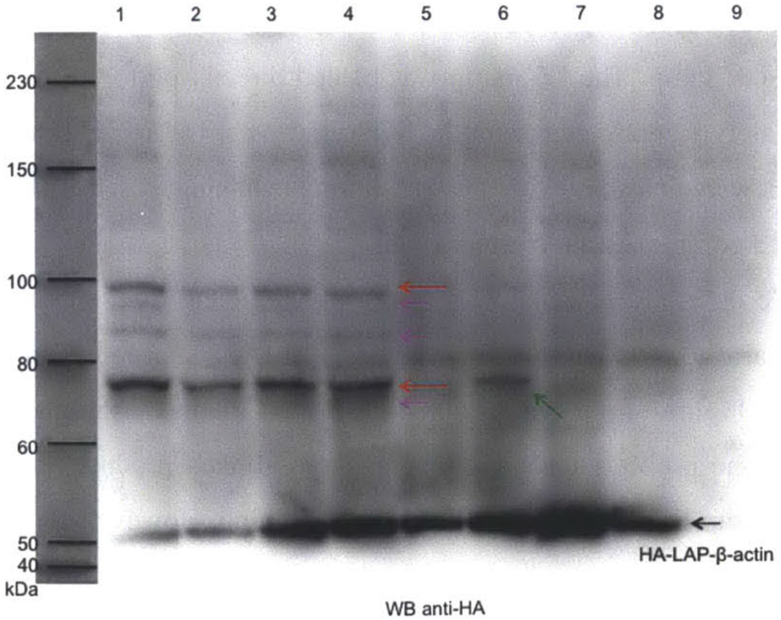


Figure 4-28. Live cell photocrosslinking of HA-LAP-actin with alkyl benzophenone incubation time course. HEK 293T cells expressing LplA(W37G) and HA-LAP-actin were incubated with 300 μM of alkyl benzophenone probe from 1 hour to 4 hours (lanes 1-4). Cell lysates were analyzed with anti-HA staining on Western blot. Negative controls are shown with either UV irradiation omitted (lane 5), alkyl benzophenone omitted (lane 6), LplA(W37G) omitted (lane 7), HA-LAP-actin omitted (lane 9), or with wild-type LplA (lane 8), which does not incorporate alkyl benzophenone probe.

In another live cell photocrosslinking experiment, we performed benzophenone ligation and photocrosslinking on the microtubule-associated protein 2 (MAP2). MAP2 is predominantly found in neurons. Its principal function is to regulate tubulin concentration and to maintain cell morphology by regulating the spacing of microtubules (68). We previously demonstrated LplA-mediated small molecule labeling on MAP2 with coumarin (49) as well as larger fluorophores like fluorescein and X-rhodamine using the LplA-mediated two-step labeling methodology (69). Benzophenone probe ligation followed by photocrosslinking in live HEK 293T cells also yielded photocrosslinked products. Following a similar protocol used in LAP-actin crosslinking described previously, we incubated HEK 293T cells expressing LplA(W37G) and HA-LAP-MAP2 in cell culture media containing either alkyl benzophenone or benzophenone photocrosslinker probe for 3 hours. After washing the cells with fresh media to remove excess probe, we replaced the culture media with ice-cold DPBS buffer. Cells were placed on ice before exposure to 365 nm UV light for 12 minutes using long-wavelength handheld UV lamp. The UV lamp was placed directly on top of the cell culture plate, at approximately 1 centimeter from the live cells. After UV irradiation, cells were lysed and the lysates were analyzed with anti-HA immunostaining to detect HA-LAP-MAP2 and its crosslinked complexes on Western blot.

We observed one major photocrosslinked band that is at 120 kDa (Figure 4-29, dark blue arrow), 40 kDa more than HA-LAP-MAP2 substrate; and one minor product at 90 kDa (Figure 4-29, light blue arrow). Photocrosslinking with the original benzophenone probe did not produce any crosslinked product (lane 6), consistent with our result from live cell HA-LAP-actin crosslinking. However, a weak background band at 120 kDa also appeared in the omitting alkyl benzophenone negative control reaction (Figure 4-29, lane 3, green arrow), similar to the background band we observed while performing LAP-actin crosslinking in live cells (Figure 4-28, lane 6, green arrow). Compared to the photocrosslinking signal (Figure 4-29, dark blue arrow), this background signal is a lot weaker. Other negative control reactions with either UV irradiate omitted (lane 2), LplA(W37G) omitted (lane 4), or with wild-type LplA (lane 5) do not show any crosslinking signal.

UV irradiation	+	-	+	+	+	+		+	+
alkyl benzophenone	+	+	-	+	+	-		-	-
benzophenone	-	-	-	-	-	+		-	+
LplA(W37G)	+	+	+	-	-	+		+	-
Wild-type LplA	-	-	-	-	+	-		-	+
	1	2	3	4	5	6		7	8

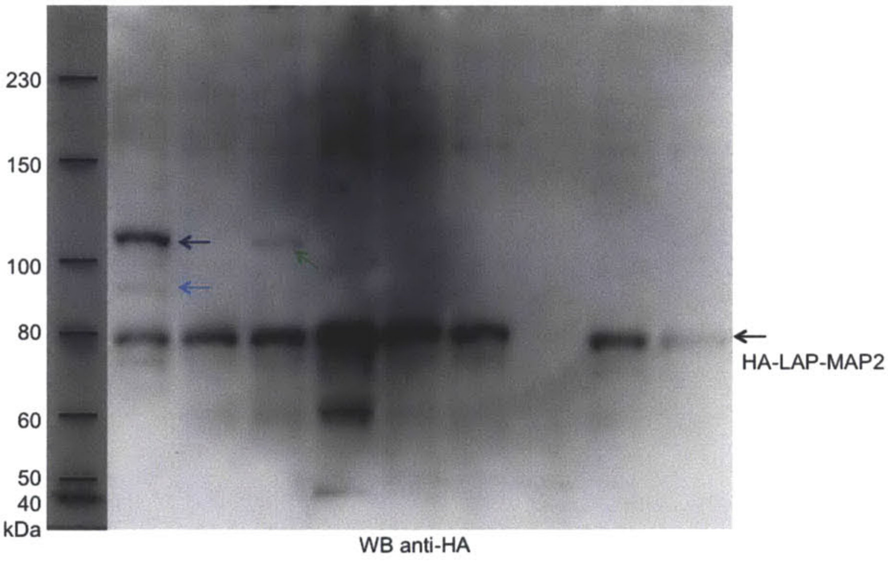


Figure 4-29. Benzophenone ligation and photocrosslinking performed on HA-LAP-MAP2 in live mammalian cells. After live cell ligation with either alkyl benzophenone or benzophenone probe and UV irradiation, cells were analyzed on Western blot with anti-HA antibody staining. Negative controls are shown with either UV irradiation, benzophenone probe, or LplA(W37G) omitted (lanes 2-4, lane 7) or with wild-type LplA (lane 5), which does not incorporate the benzophenone probe. Dark blue and light blue arrows indicate photocrosslinked products; green arrow indicates background crosslinked product in the absence of alkyl benzophenone.

In summary, live cell photocrosslinker probe ligation and crosslinking using alkyl benzophenone on HA-LAP-actin and HA-LAP-MAP2 proteins showed crosslinked products that were detected by immunostaining with antibody against the HA tag. Compared to the original benzophenone probe that showed no photocrosslinked products in either actin or MAP2 context, the alkyl benzophenone probe appeared to work better in live cell experiments. We decided to pursue subsequent live cell photocrosslinking experiments with the alkyl benzophenone probe.

LAP-actin photocrosslinking with drug treatment inside live mammalian cells

Following the success of LAP-actin and LAP-MAP2 photocrosslinking in live mammalian cells, we wanted to use different drug stimulants to induce or prevent different proteins from interacting with actin to change the protein-protein interaction pattern. This experiment also served to help us in checking the validity of our intracellular benzophenone crosslinking method. We should be able to obtain different photocrosslinked products since different proteins are known to associate with actin whether it is in its monomeric or filamentous state (Figure 4-30).

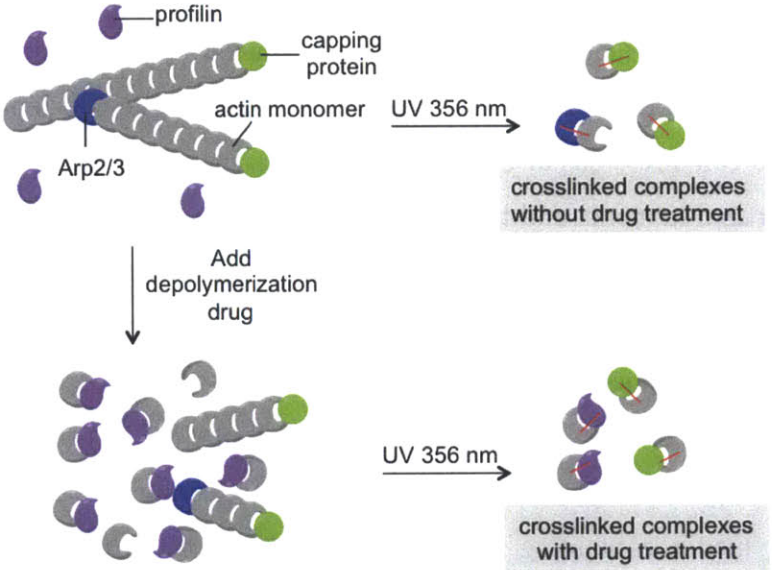

profilin
capping protein
actin monomer
Arp2/3
UV 356 nm
crosslinked complexes without drug treatment
Add depolymerization drug
UV 356 nm
crosslinked complexes with drug treatment

Figure 4-30. Schematic representations of different LAP-actin crosslinked complexes with and without depolymerization drug treatment. At 365 nm UV light irradiation of F-actin (a polymer of benzophenone-ligated LAP-actin monomers) causes LAP-actin to crosslink to capping protein and also Arp2/3 complex. Addition of an actin depolymerization drug depolymerizes a majority of F-actin into G-actin, which is then bound by profilin. As a result, UV irradiation of G-actin leads to a different pool of crosslinked complexes. Red bars represent covalent bond formation.

We first used various drug treatments to change the polymerization state of actin as well as its localization. We examined a total of five stimulants. Three stimulants changed the polymerization state of actin. First, latrunculin A is a potent macrolide that binds to monomeric G-actin, forming a 1:1 complex to disrupt microfilament polymerization (70). Second, cytochalasin B is a drug that disassembles stress fibers, causing aggregation of actin fibers within the cytosol. Cytochalasin B is about 10- to 100-fold less potent than latrunculin A (71). Third, CK-666 inhibits nucleation of new actin filaments by binding at the interface of Arp2 and Arp3 proteins, locking Arp2/3 complex in an inactive conformation (72). In addition, two stimulants increased the population of actin in the nucleus. Leptomycin B is a fatty acid that inhibits the nuclear export of proteins containing nuclear export signal. Previously in Dr. Chayasith Uttamapinant's work (49), leptomycin B was used to block actin export from the nucleus, thus increasing the amount of nuclear actin pool, so it can be ligated by coumarin ligase to provide further evidence for the presence of nuclear actin. In addition to leptomycin B, heat-shock induced cell stress has also been shown to increase the abundance of nuclear actin (73). Since drug treatments and heat shock either alter actin polymerization states or its location inside the cell dramatically, we expected to see different protein-protein interaction complexes being crosslinked whether actin is being inhibited, disassembled, or translocated to the nucleus.

We incubated HEK 293T cells expressing LplA(W37G) and HA-LAP-actin in media containing the alkyl benzophenone probe. Following probe ligation, cells were incubated in media containing each drug stimulant for 40 minutes. For heat-shock induced cell stress, the cells was moved form the 37 °C incubator to 43 °C incubator for 1 hour. At the end of drug or heat treatment, the cell culture media was replaced with DPBS and the cells were irradiated with 365 nm UV light for 12 minutes using the handheld UV lamp. The cells were then lysed and analyzed on Western blot with anti-HA immunostaining to detect HA-LAP-actin crosslinked complexes.

However, as shown in Figure 4-31, none of the stimulants was able to induce new HA-LAP-actin crosslinked product that is different from the crosslinked complexes already detected in the absence of drug and heat treatments. In Figure 4-31, lane 1 represents crosslinked complexes already identified without stimulant, and each lane in lanes 4, 6, 7, 8, and 10 represents one of the five stimulants tested in this experiment. The crosslinked products were not observed if alkyl benzophenone or UV irradiation was omitted (lanes 2, 3, 5, 9 11). We were puzzled regarding the lack of different crosslinked products. We were confident that the drug treatments were successful because we observed drastic cell morphology changes such as retraction of the filopodia and cell rounding, similar to literature reports (74,75).

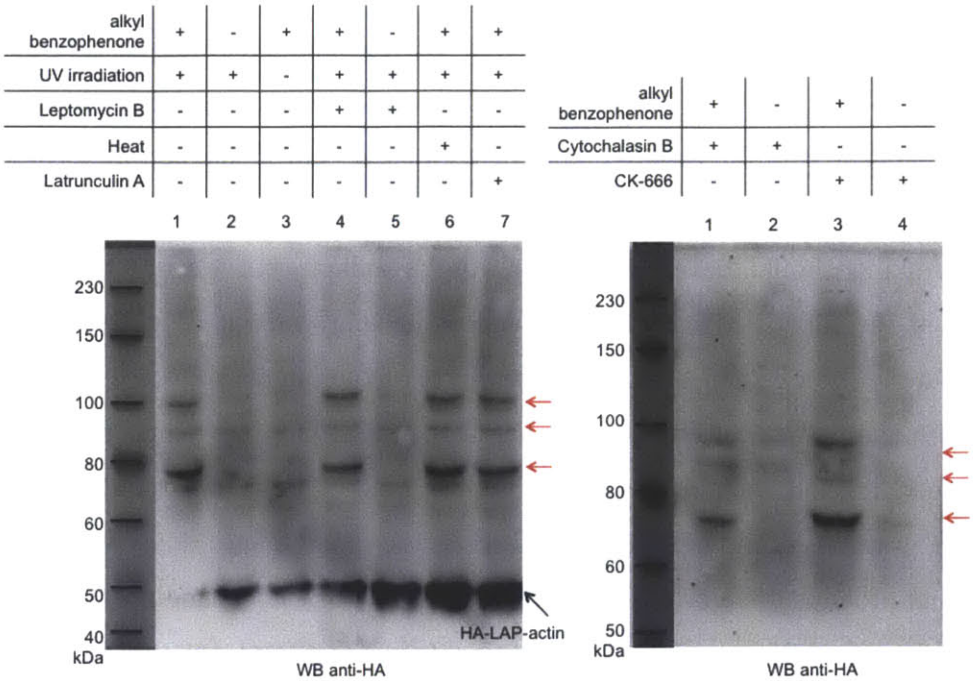

alkyl benzophenone + - + + - + +
UV irradiation + + - + + + +
Leptomycin B - - - + + - -
Heat - - - - - + -
Latrunculin A - - - - - - +
1 2 3 4 5 6 7
230
150
100
80
60
50
40
kDa
HA-LAP-actin
WB anti-HA
alkyl benzophenone + - + -
Cytochalasin B + + - -
CK-666 - - + +
1 2 3 4
230
150
100
80
60
50
kDa
WB anti-HA

Figure 4-31. Live cell photocrosslinking of LAP-actin with various drug treatments. HEK 293T cells expressing LplA(W37G) and HA-LAP-actin were incubated with 300 μM alkyl benzophenone for 3 hours. Afterward alkyl benzophenone ligation, cells were incubated for 40 minutes at 37 °C with 4 nM of leptomycin B, 500 nM of latrunculin A, 10 μM cytochalasin B, or 10 μM CK-666. For heat-induced stress, cells were moved from 37 °C incubator to 43 °C incubator for 60 minutes after alkyl benzophenone ligation. HA-LAP-actin is indicated by black arrow. Crosslinked complex of HA-LAP-actin are indicated by red arrow.

These results led us to question the validity and the efficiency of our photocrosslinking method. It should be obvious that different proteins interact with actin whether actin is in its monomeric state or in its fibrous state. The fact that the photocrosslinking pattern stayed the same under many potent drug treatments alerted us to reexamine our system more closely. The photocrosslinked complexes of HA-LAP-actin we identified in live cell photocrosslinking were reproducible under many experiments and dependent on UV irradiation, LplA(W37G) enzyme, and alkyl benzophenone probe. However, there are still some unresolved issues that should be addressed in future experiments. First of all, we continued to see a weaker background crosslinked band in the alkyl benzophenone negative control experiments in both live-cell HA-LAP-actin photocrosslinking and HA-LAP-MAP2 photocrosslinking (Figure 4-28, lane 6; Figure 4-29, lane 3). We do not yet understand the source of this background and why it only appeared in the negative control with alkyl benzophenone omitted. Second, the photocrosslinked complexes we identified in HA-LAP-actin live-cell photocrosslinking should be altered under drug treatment, but we did not observe a difference in the identified crosslinked bands with and without drug treatment. It is possible that the proteins crosslinked to actin in our experiments were not changed with different stimulants, but we find this unlikely because actin's interacting partners differ greatly depending on its polymerization states. In the future, we could isolate the crosslinked proteins from the gel to perform mass spectrometry analysis in order to identify what proteins were crosslinked to HA-LAP-actin. Understanding the identity of the crosslinked proteins would offer additional evidence and support regarding the validity of our photocrosslinking method.

Conclusion

We have demonstrated *in vitro* photocrosslinking with the FKBP-rapamycin-FRB interaction system using two benzophenone probes, benzophenone and alkyl benzophenone. The two probes differ at the point of attachment to the fourth position of the benzophenone core – the first version benzophenone has a carbonyl group attached which we speculated might interfere with the ketyl diradical generation through electron delocalization, therefore, alkyl benzophenone with simply a methylene group attached at the fourth position was synthesized to address this concern. Both probes behaved similarly under *in vitro* crosslinking environment with purified proteins. However, only alkyl benzophenone produced crosslinked products once we tried photocrosslinking inside live cells with either an alkyl benzophenone-ligated or benzophenone-ligated LAP-actin or LAP-MAP2 protein. To further explore the nature of the crosslinked complexes, we used various drug stimulants as well as heat-induced cellular stress in an attempt to change the polymerization state or localization of actin, which would cause different proteins to interact with actin. We were not able to obtain any additional or different crosslinking signal from the ones already observed under normal state without drug treatment.

Several aspects of the project made us reevaluate our goal of developing a robust and easy to implement protein-protein interaction detection method. First, the placement of the photocrosslinker often has to be optimized by screening many regions in the protein of interest. This can be challenging in the LplA-mediated labeling technique because it requires many different genetic fusions of the LAP peptide to the protein under study. So far, LplA-mediated labeling technique has been demonstrated mostly on proteins with LAP fusions at the termini (29,49,51,52,69,76,77), where LAP is likely to be easily accessible. It is conceivable that certain regions, especially internal or buried, would have different degrees of accessibility and probe ligation yield by LplA. Second, prior to photocrosslinking, one would need to quantify the ligation yield for each LAP-tagged protein using native gel-shift assay, which can be labor-intensive when many fusion positions are tested. Third, for the protein under study, one would need to choose the region of LAP fusion carefully in order to avoid disrupting the native structure or the binding ability of the protein to other interacting proteins. A protein-specific functional assay comparing the wild-type and LAP-tagged protein should be accomplished. Information regarding the protein's interacting regions or a crystal structure of the protein can help in deciding possible regions for LAP fusions. For example, we did not fuse LAP to the internal region of FKBP because it contains many residues that form hydrogen bonds or hydrophobic interactions with rapamycin. Ideally, we need to achieve quantitative ligation yield of the photocrosslinker on LAP-tagged protein in order to maximize crosslinking signal, but this requirement cannot always be guaranteed especially if LAP is inaccessible.

Finally, we also need synthetic chemists to design novel photocrosslinkers that could work both specifically and efficiently with high sensitivity and less background reaction or quenching by media or solvent.

Experimental methods

Genetic constructs

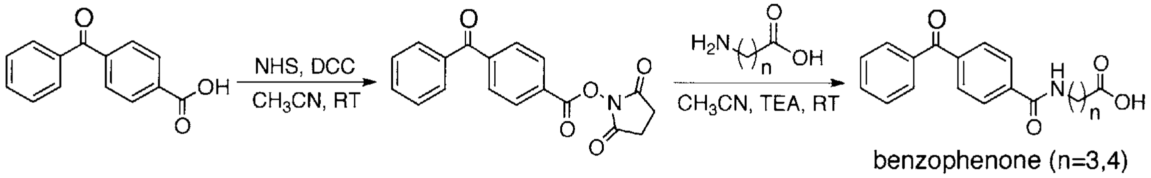
Construct name	Vector	Features and variants	References
FLAG-LplA	pet21a	FLAG: DYKDDDDK Other LplA mutants: W37G, W37A, W37V, W37I, W37L, W37S	(69)
E2p	pet21a		(77)
myc-FKBP-LAP	pet21a	myc: EQKLISEEDL LAP: GFEIDKVWYDLDA Other FKBP mutants: M49A, M66A, M49A and M66A	(50)
myc-FKBP-LAP(F)	pet21a	LAP(F): GFEIDKVFYDLDA	
myc-FKBP	pet21a		(50)
HA-FRB-YFP	pet21a	HA: YPYDVPDYA YFP: yellow fluorescent protein	
HA-FRB-mApple	pet21a	mApple: a red fluorescent protein	
HA-FRB-LAP	pet21a		
HA-FRB-(10 a.a. linker)-LAP	pet21a	10 a.a. linker: GSGSTSGSGK	
HA-FRB	pet21a		
FLAG-LplA	pcDNA3	Other LplA mutant: W37G	(69)
FLAG-LplA	pCAG	Other LplA mutant: W37G	
LAP-YFP	pcDNA3		(69)
LAP (Ala)-YFP	pcDNA3	LAP(Ala): GFEIDAVWYDLDA	(69)
LAP-mCherry	pcDNA3	mCherry: a red fluorescent protein	(69)
mCherry-FKBP-LAP	pcDNA3		
myc-FKBP-LAP	pcDNA3		(50)
HA-FRB	pcDNA3		
HA-LAP-actin	Clontech vector	Other LAP variant: LAP(F)	(69)
HA-LAP-MAP2	Clontech vector		(69)

General synthetic methods

All reagents were the highest grade available and purchased from Sigma-Aldrich, Anaspec, Thermal Scientific, TCI America, Alfa Aesar, or Life Technologies and used without further purification. Anhydrous solvents were drawn from Sigma-Aldrich SureSeal bottles. Analytical thin layer chromatography was performed on 0.25 mm silica gel 60 F254 plates and visualized under short or long wavelength UV light, or after staining with 2,4-dinitrophenylhydrazine (DNPH) or bromocresol green. Flash column chromatography was carried out using silica gel (ICN SiliTech 32-63D). Mass spectrometric analysis was performed on an Applied Biosystems 200 QTRAP mass spectrometer using electrospray ionization. HPLC analyses were performed on a Varian Prostar Instrument equipped with a photo-diode-array detector. A reverse-phase Microsorb-MV 300 C18 column (250 × 4.6 mm dimension) was used for analytical HPLC. NMR spectra were recorded on a Bruker AVANCE 400 MHz instrument.

Synthesis of benzophenone probes

Synthesis of benzophenone

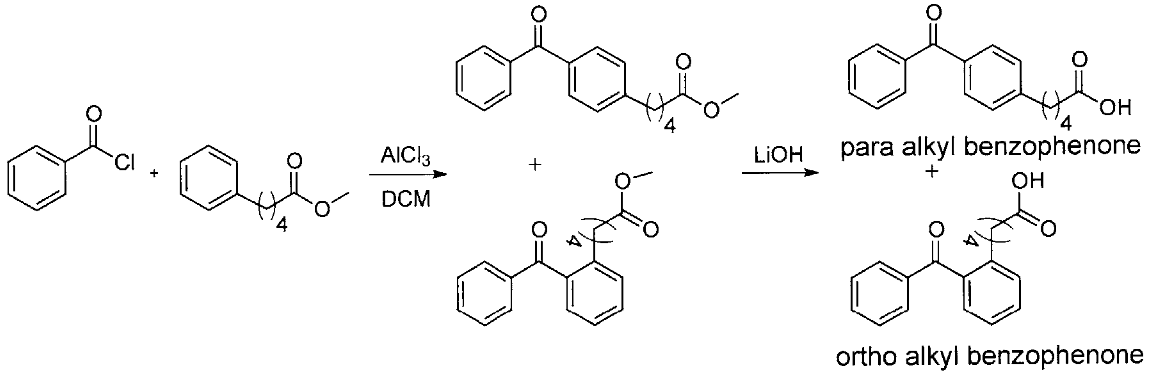


To a solution of 40 mg of 4-benzoylbenzoic acid (0.18 mmole, Sigma Aldrich) dissolved in 10 ml of anhydrous acetonitrile (MeCN) was added 22 mg of *N*-hydroxysuccinimide (NHS) (0.19 mmol, Sigma Aldrich) and 43 mg of *N,N*'-dicyclohexylcarbodiimide (DCC) (21 mmol, TCI America). The reaction was stirred at room temperature for 10 hours. Afterwards, MeCN was removed under reduced pressure and the crude material was purified by silica gel column chromatography (column size: 4 inches tall, 0.8 inch in diameter; solvent gradient from 25 to 50% ethyl acetate in hexane). Purified 4-benzoylbenzoic-NHS ester (48 mg, 0.15 mmol) was dissolved in 4 ml of anhydrous MeCN and reacted with either 17 mg of 4-aminobutyric acid (0.17 mmol, Sigma Aldrich) or 20 mg of 5-aminovaleric acid (0.17 mmol, Sigma Aldrich), and 14 μl of anhydrous triethylamine (TEA) (0.10 mmol, Sigma Aldrich) for 12 hours at room temperature. Solvent was removed under reduced pressure and the crude mixture was purified by silica gel column chromatography (column size: 4 inches tall, 0.8 inch in diameter; solvent gradient from 5 to 10% methanol in dichloromethane (DCM)). Both benzophenone products were confirmed by ^{1}H NMR and mass spectrometry.

Characterization of benzophenone (n=3): ^{1}H NMR ($CDCl_3$): 7.93 (d, 2H, *J* = 4.2), 7.83 (t, 1H, *J* = 4.8), 7.45 (m, 4H), 7.57 (m, 1H), 7.45 (m, 2H), 3.47 (m, 1H), 3.06 (q, 4H, *J* = 7.2), 2.41 (t, 2H). ESI-MS calculated for $[M+H]^+$: 311.12; observed 312.03.

Characterization of benzophenone (n=4): ^{1}H NMR ($CDCl_3$): 7.92 (d, 2H, J = 4.2), 7.80 (m, 4H), 7.60 (m, 1H), 7.47 (m, 2H), 6.92 (m, 1H), 3.49 (m, 2H), 3.05 (m, 4H), 2.36 (m, 3H). ESI-MS calculated for $[M+H]^+$: 325.13; observed 326.41.

Synthesis of alkyl benzophenone



To a solution of 500 mg of benzoyl chloride (3.6 mmol, Sigma Aldrich) dissolved in 10 ml of anhydrous DCM was added 142 mg of anhydrous $AlCl_3$ (1.1 mmol, TCI America). The reaction was stirred at room temperature for 30 minutes, during which time the color of the reaction changed gradually from clear yellow to red, indicating the formation of the acylium ion. After the reaction turned red, 750 mg of 5-phenyl valeric methyl ester (3.9 mmol, Sigma Aldrich) was added. The reaction was stirred at room temperature for 2 hours and monitored by thin layer chromatography (TLC). TLC followed by 2,4-dinitrophenylhydrazine (DNPH) stain indicated the formation of the ketone functional group. After 2 hours, DCM was removed under reduced pressure. The resulting residue of the crude reaction mixture was purified by silica gel chromatography (column size: 6 inches tall, 0.8 inch in diameter; solvent gradient from 5 to 20% ethyl acetate in hexane). Benzophenone alkyl methyl ester was hydrolyzed to the corresponding carboxylic acid using LiOH in 5 ml of 50% tetrahydrofuran (THF) in water for 1 hour at room temperature. Hydrolysis was monitored by TLC and bromocresol green stain to identify the carboxylic acid. The reaction mixture was diluted into 10 ml of 1M HCL and extracted with ethyl acetate (3 X 10 ml). The organic layer was dried over magnesium sulfate, then filtered. Ethyl acetate was removed under reduced pressure, Even though the *para* benzophenone alkyl methyl ester was favored, both *para* and *ortho* products were obtained in approximately 5 to 1 ratio. Product yield for the *para* product was approximately 30% and the yield for the *ortho* product was approximately 6%. Both products were confirmed by ^{1}H NMR and mass spectrometry

Characterization of para alkyl benzophenone: ^{1}H NMR (CD_3OD): 7.79 (m, 2H), 7.74 (m, 2H), 7.58 (m, 1H), 7.48 (m, 2H), 7.28 (m, 2H), 2.72 (t, 2H, $J = 5.6$), 2.40 (t, 2H, $J = 5.6$), 1.712 (m, 4H). ESI-MS calculated for $[M+H]^+$: 282.13; observed 283.4.

Characterization of ortho alkyl benzophenone: ^{1}H NMR (CD_3OD): 7.83 (m, 2H), 7.61 (m, 1H), 7.46 (m, 2H), 7.30 (m, 3H), 2.70 (m, 2H), 2.31 (t, 2H, $J = 5.6$), 1.63 (m, 1H), 1.28 (m, 4H). ESI-MS calculated for $[M+H]^+$: 282.13; observed 283.5.

Protein expression and purification: see Chapter 2 Experimental Methods.

HPLC screening to identify LplA enzyme and benzophenone probe pair

For the screen in Figure 4-4, reactions containing 500 nM LplA enzyme, 300 μM benzophenone probe, 100 μM LAP peptide (sequence: H_2N-GFEID$\underline{K}$VWYDLDA-CO_2H), 2 mM ATP, and 2 mM magnesium acetate in 25 mM Na_2HPO_4 pH 7.2 were incubated at 30 °C for 40 minutes. Reactions were quenched with 40 mM EDTA (ethylenediaminetetraacetic acid, final concentration). Percent conversion to LAP–benzophenone adduct was quantified by comparing area under the peaks on HPLC with a C18 reverse phase column, recording absorbance at 210 nm. Elution conditions were 30–60% acetonitrile in water with 0.1% trifluoroacetic acid over 20 min (linear gradient) at 1.0 mL/min flow rate. The percent conversion was calculated from the ratio of LAP–benzophenone to the sum of unmodified LAP + LAP– benzophenone.

In vitro benzophenone ligation onto E2p or purified LAP-tagged protein

In vitro benzophenone ligation reactions contained 1 μM LplA(W37G) enzyme, 500 μM benzophenone or alkyl benzophenone probe, 50 μM E2p or 10 μM LAP-tagged protein (except in Figure 4-20, where 2 μM of FKBP-LAP was used), 2 mM ATP, and 2 mM magnesium acetate in phosphate-buffered saline (PBS) were incubated at 30 °C for 3 hours. Reactions were analyzed on HPLC using a linear gradient from 30–60% acetonitrile in water with 0.1% trifluoroacetic acid over 20 minutes at 1.0 mL/min flow rate to confirm quantitative benzophenone ligation on E2p or LAP-tagged proteins. After ligation was complete, reactions were transferred into dialysis membrane tubing (SnakeSkin, Cat. # 88242, Pierce) with 3,500 MWCO or dialysis cups (Cat. # 69570, Pierce) with 10,000 MWCO to removed excess unligated benzophenone probe through dialysis in phosphate buffered saline (PBS) buffer for a total of 10 hours. Dialysis buffer was changed to fresh PBS buffer after 5 hours.

Detecting benzophenone-ligated E2p with biotin-hydrazide (Figure 4-5) or hydroxylamine-Alexa Fluor 488 (Figure 4-8)

Benzophenone-ligated E2p (50 μM) was incubated with biotin-hydrazide (200 μM) or hydroxylamine-Alexa Fluor 488 (200 μM) at pH 4.5 or 5, respectively, for 12 hours at room temperature. Afterwards, the samples were mixed with protein loading buffer and separated on a 12 % SDS-PAGE gel.

For the samples reacted with biotin-hydrazide, the proteins were transferred onto a nitrocellulose membrane at 400 mA (constant mA) for 2.5 hours. Complete transfer of the pre-stained protein ladder was used as an indication of successful protein transfer. The nitrocellulose membrane was blocked with 3% w/v BSA (Pierce), in TBS-T (0.1% Tween-20 in Tris-buffered saline) buffer at room temperature for 1 hour. To detect biotinylated proteins, membrane was immersed in a solution of 0.3 mg/mL streptavidin-HRP (Thermo Scientific) at room temperature for 1 hour before rinsing with TBS-T 4× 5 minutes. Blots were developed with the SuperSignal West Pico reagent (Thermo Scientific), and imaged using an Alpha Innotech digital camera.

For the samples reacted with hydroxylamine-Alexa Fluor 488, after separating the proteins on 12% SDS-PAGE gel, in-gel fluorescence of Alex Fluor 488 was visualized on a FUJIFILM FLA-9000 instrument.

In vitro photocrosslinking with purified FKBP and FRB proteins

Benzophenone ligation on LAP-tagged protein is described above. After ligation is complete, reactions were transferred into dialysis membrane tubing (SnakeSkin, Cat. # 88242, Pierce) with 3,500 MWCO to removed excess unligated benzophenone probe through dialysis in phosphate buffered saline (PBS) buffer for a total of 10 hours. Dialysis buffer was changed to fresh PBS buffer after 5 hours. Following dialysis, 10 μM of either purified FKBP or FRB protein (except in Figure 4-20, where 2 μM of FRB protein was added) and 500 nM rapamycin (1 mM stock concentration in DMSO) were added to the reactions. After 1 hour of incubation with rotation at room temperature for ternary complex formation, reactions were transferred into wells in a 96-well plate, which was placed on ice and exposed to UV light using the specified instrument for the indicated amount of time. All samples were kept on ice during UV irradiation to prevent heat-induced protein aggregation.

Optimized conditions for each photolysis instrument

Instrument	wavelength	Power output	Distance from samples	Time of UV irradiation
Porta-Ray 400R	350 – 400 nm	400-Watt	3 inches	2 minutes
Handheld long wavelength UV lamp	365 nm	4-Watt	1 cm	15 minutes
Stratalinker	365 nm	5 lamps, 8-Watt/lamp	5 inches	30 minutes

Western blot analysis of *in vitro* FKBP and FRB photocrosslinking

After UV irradiation, samples were combined with protein loading buffer and separated on a 12% SDS-PAGE gel at constant 200 V. For blotting analysis, proteins in the SDS-PAGE gel were transferred to nitrocellulose membrane. Protein transfer was run for 2.5 hours at constant 400 mA. Complete transfer of pre-stained protein ladder was used as an indication of successful protein transfer. Afterwards, the nitrocellulose membrane was blocked with 3% w/v BSA (Pierce), in TBS-T (0.1% Tween-20 in Tris-buffered saline) buffer at room temperature for 1 hour. To stain with mouse anti-myc (Millipore) or rabbit anti-HA (Invitrogen) primary antibody, the membranes were incubated with antibody (1:1000 dilution) in TBS-T buffer at room temperature for 1 hour, then rinsed with TBS-T buffer 4 x 5 minutes. The membranes were then incubated with goat anti-mouse or anti-rabbit horseradish peroxidase conjugate (BioRad, 1:5000 dilution) in TBS-T buffer at room temperature for 1 hour. Membranes were washed 4 x 5 minutes with TBS-T, then maintained in TBS-T at room temperature for imaging. The blots were developed using Supersignal West Pico substrate (Pierce) and visualized on an Alpha Innotech ChemiImager 5500 instrument.

Native gel-shift analysis of benzophenone ligation yields in cells (Figures 4-21, 4-22)

HEK 293T cells were plated into wells of a 12-well culture plate (4 cm^2 per well) 18 hours prior to transfection and grown to 60% confluency. For benzophenone ligation, cells were transfected with 400 ng LplA(W37G) and either 800 ng LAP-YFP, LAP(Ala)-YFP, LAP-mCherry, or mCherry-FKBP-LAP using Lipofectamine 2000 (Life Technologies) following manufacture's protocol. At 18 hours after transfection, cells were incubated in growth media (MEM supplemented with 10% FBS) containing 300 μM benzophenone probe for up to 3.5 hours at 37 °C. For Figure 4-22, cells were incubated with benzophenone for either 1 hour, 2 hours, or 3.5 hours. At the end of benzophenone incubation time, cells were then washed with fresh media for 1 hour, so excess unligated benzophenone can be removed from cells via the action of organic anionic transporters. Cells were then harvested using a cell scrapper and lysed in 100 μl hypotonic lysis buffer (1 mM HEPES pH 7.5, 5 mM $MgCl_2$, 1 mM PMSF (phenylmethanesulfonyl fluoride, Thermal Scientific), 1 mM protease inhibitor cocktail (Sigma-Aldrich)) using freeze-thaw lysis method. Cell lysates were frozen at -20 °C, thawed at room temperature, then mixed by vortexing for 2 minutes. This freeze-thaw-vortex cycle was repeated three times. Cells were then centrifuged at 13,000 rpm for 2 minutes to pellet insoluble cellular fractions. The lysate supernatant (5 μl lysate per lane) was combined with DTT-free and SDS-free protein loading buffer and separated and analyzed on a 12% polyacrylamide native gel without SDS at constant 200 V. Lysate samples were not boiled to preserve the native structure and the fluorescence of the LAP-tagged fluorescent proteins. Prior to Coomassie staining, in-gel fluorescence of YFP or mCherry was visualized on a FUJIFILM FLA-9000 instrument.

Immunostaining to check expressions of FKBP-LAP and FRB (Figure 4-23)

HEK 293T cells were plated on 5 mm × 5 mm glass coverslips placed within wells of a 48-well cell culture plate (0.95 cm^2 per well) 12–16 hours prior to transfection. Glass coverslips were pre-coated with 50 μg/mL fibronectin (Millipore) to increase HEK 293T adherence. Plated cells were transfected with 100 ng FKBP-LAP in pcDNA3, 100 ng HA-FRB in pcDNA3, and 100 ng LplA(W37G) in pCAG using lipofectamine. At 18 hours after transfection, cells were washed 3x with Dulbecco's phosphate buffered saline (DPBS) and fixed with 3.7% paraformaldehyde in DPBS for 15 minutes at room temperature. Fixed cells were washed 3x with DPBS followed by protein precipitation with cold methanol for 6 minutes at -20 °C. After cold methanol precipitation, cells were washed with DPBS 3x and blocked with 0.5% casein in DPBS for 1 hour at room temperature.

To visualize myc-FKBP-LAP and HA-FRB proteins, fixed cells were immunostained by first incubating with primary antibodies, mouse monoclonal anti-myc antibody (Millipore, 1:1000 dilution) and rabbit monoclonal anti-HA antibody (Invitrogen, 1:1000 dilution) in 1% (w/v) BSA in DPBS for 1 hour at room temperature. Cells were further washed with DPBS 3x and incubated with secondary antibody, goat anti-mouse IgG antibody conjugated to Alexa Fluor 488 (Life Technologies) and goat anti-rabbit IgG antibody conjugated to Alexa Fluor 568 (Life Technologies) in 1% (w/v) BSA in DPBS for 1 hour at room temperature. Cells were then washed 3x and imaged on a confocal microscope.

Benzophenone probe ligation and photocrosslinking on HA-LAP-actin or HA-LAP-MAP2 in live cells (Figures 4-26, 4-27, 4-28, 4-29)

HEK 293T cells were plated into wells of a 12-well culture plate (4 cm^2 per well) 18 hours prior to transfection and grown to 60% confluency. For benzophenone ligation, cells were transfected with 400 ng LplA(W37G) and 800 ng HA-LAP-actin or HA-LAP-MAP2 using Lipofectamine 2000 (Life Technologies) following manufacture's protocol. 18 hours after transfection, cells were incubated in growth media (MEM supplemented with 10% FBS) containing 300 μM benzophenone or alkyl benzophenone probe 3 hours (except in Figure 4-28, where alkyl benzophenone incubation was from 1 hour to 4 hours as indicated in the figure) at 37 °C. Excess benzophenone probe was removed by rinsing cells with fresh cell culture media 3 times over 1 hour and relying on the action of organic anionic transporters. Media was replaced with cold DPBS and the cells were placed on ice. Handheld, long wavelength UV lamp was placed directly on top of cell culture plate, 1 centimeter from the live cells. Cells were UV irradiated for 12 minutes and then lysed with 50 μl RIPA lysis buffer (50 mM Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100) containing 1 mM PMSF (Thermal Scientific, phenylmethanesulfonyl fluoride) and 1 mM protease inhibitor cocktail. Cell lysate was separated on a 9% SDS-PAGE gel. For blotting analysis, proteins in the SDS-PAGE gel were transferred to nitrocellulose membrane. Protein transfer was run for 2.5 hours at constant 400 mA. Complete transfer of pre-stained protein ladder was used as an indication of successful protein transfer. Afterwards, the nitrocellulose membrane was blocked with 3% w/v BSA (Pierce), in TBS-T (0.1% Tween-20 in Tris-buffered saline) buffer at room temperature for 1 hour.

To immunostain with rabbit anti-HA primary antibody (Invitrogen) or mouse anti-FLAG (Sigma) primary antibody, the membranes were incubated with antibody (1:1000 dilution) in TBS-T buffer at room temperature for 1 hour, then rinsed with TBS-T buffer 4 x 5 minutes. The membranes were then incubated with anti-rabbit horseradish peroxidase conjugate or goat anti-mouse (BioRad, 1:5000 dilution) in TBS-T buffer at room temperature for 1 hour. Membranes were washed 4 x 5 minutes with TBS-T, then maintained in TBS-T at room temperature for imaging. The membranes were developed using Supersignal West Pico substrate (Pierce) and visualized on an Alpha Innotech ChemiImager 5500 instrument.

References

1. Stumpf MPH, Thorne T, Silva E de, Stewart R, An HJ, Lappe M, et al. Estimating the size of the human interactome. PNAS. 2008 May 13;105(19):6959–64.

2. Velasco-García R, Vargas-Martínez R. The study of protein-protein interactions in bacteria. Can. J. Microbiol. 2012 Nov;58(11):1241–57.

3. Fields S, Song O. A novel genetic system to detect protein–protein interactions. , Published online: 20 July 1989; | doi:10.1038/340245a0. 1989 Jul 20;340(6230):245–6.

4. Fields S. Interactive learning: Lessons from two hybrids over two decades. PROTEOMICS. 2009;9(23):5209–13.

5. Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, et al. A Human Protein-Protein Interaction Network: A Resource for Annotating the Proteome. Cell. 2005 Sep 23;122(6):957–68.

6. Lim J, Hao T, Shaw C, Patel AJ, Szabó G, Rual J-F, et al. A Protein–Protein Interaction Network for Human Inherited Ataxias and Disorders of Purkinje Cell Degeneration. Cell. 2006 May 19;125(4):801–14.

7. Kaltenbach LS, Romero E, Becklin RR, Chettier R, Bell R, Phansalkar A, et al. Huntingtin Interacting Proteins Are Genetic Modifiers of Neurodegeneration. PLoS Genet. 2007 May 3(5).

8. Caufield JH, Sakhawalkar N, Uetz P. A comparison and optimization of yeast two-hybrid systems. Methods. 2012 Dec;58(4):317–24.

9. Suter B, Kittanakom S, Stagljar I. Two-hybrid technologies in proteomics research. Current Opinion in Biotechnology. 2008 Aug;19(4):316–23.

10. Shekhawat SS, Ghosh I. Split-protein systems: beyond binary protein–protein interactions. Current Opinion in Chemical Biology. 2011 Dec;15(6):789–97.

11. Hu C-D, Kerppola TK. Simultaneous visualization of multiple protein interactions in living cells using multicolor fluorescence complementation analysis. Nat Biotech. 2003 May;21(5):539–45.

12. Pelletier JN, Campbell-Valois F-X, Michnick SW. Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. PNAS. 1998 Oct 13;95(21):12141–6.

13. Galarneau A, Primeau M, Trudeau L-E, Michnick SW. β-Lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein–protein interactions. Nat Biotech. 2002 Jun;20(6):619–22.

14. Remy I, Michnick SW. A highly sensitive protein-protein interaction assay based on Gaussia luciferase. Nat Meth. 2006 Dec;3(12):977–9.

15. Hatanaka Y, Sadakane Y. Photoaffinity Labeling in Drug Discovery and Developments: Chemical Gateway for Entering Proteomic Frontier. Current Topics in Medicinal Chemistry. 2002 Mar 1;2(3):271–88.

16. Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang Z, Schultz PG. An Expanded Eukaryotic Genetic Code. Science. 2003 Aug 15;301(5635):964–7.

17. Chou C, Uprety R, Davis L, Chin JW, Deiters A. Genetically encoding an aliphatic diazirine for protein photocrosslinking. Chem. Sci. 2011 Feb 14;2(3):480–3.

18. Liu W, Brock A, Chen S, Chen S, Schultz PG. Genetic incorporation of unnatural amino acids into proteins in mammalian cells. Nat Meth. 2007 Mar;4(3):239–44.

19. Hino N, Okazaki Y, Kobayashi T, Hayashi A, Sakamoto K, Yokoyama S. Protein photo-cross-linking in mammalian cells by site-specific incorporation of a photoreactive amino acid. Nat Meth. 2005 Mar;2(3):201–6.

20. Hino N, Oyama M, Sato A, Mukai T, Iraha F, Hayashi A, et al. Genetic Incorporation of a Photo-Crosslinkable Amino Acid Reveals Novel Protein Complexes with GRB2 in Mammalian Cells. Journal of Molecular Biology. 2011 Feb 18;406(2):343–53.

21. Wang L, Xie J, Schultz PG. Expanding the Genetic Code. Annual Review of Biophysics and Biomolecular Structure. 2006;35(1):225–49.

22. De Graaf AJ, Kooijman M, Hennink WE, Mastrobattista E. Nonnatural amino acids for site-specific protein conjugation. Bioconjug. Chem. 2009 Jul;20(7):1281–95.

23. Knowles JR. Photogenerated reagents for biological receptor-site labeling. Acc. Chem. Res. 1972 Apr 1;5(4):155–60.

24. Voskresenska V, Wilson RM, Panov M, Tarnovsky AN, Krause JA, Vyas S, et al. Photoaffinity Labeling via Nitrenium Ion Chemistry: Protonation of the Nitrene Derived from 4-Amino-3-nitrophenyl Azide to Afford Reactive Nitrenium Ion Pairs. J. Am. Chem. Soc. 2009 Aug 19;131(32):11535–47.

25. Brunner J. New photolabeling and crosslinking methods. Annu. Rev. Biochem. 1993;62:483–514.

26. Doering W vo. E, Odum RA. Ring enlargement in the photolysis of phenyl azide. Tetrahedron. 1966;22(1):81–93.

27. Spanggord RJ, Beal PA. Site-specific modification and RNA crosslinking of the RNA-binding domain of PKR. Nucleic Acids Res. 2000 May 1;28(9):1899–905.

28. Tanaka Y, Bond MR, Kohler JJ. Photocrosslinkers illuminate interactions in living cells. Mol. BioSyst. 2008 May 20;4(6):473–80.

29. Baruah H, Puthenveetil S, Choi Y-A, Shah S, Ting AY. An Engineered Aryl Azide Ligase for Site-Specific Mapping of Protein–Protein Interactions through Photo-Cross-Linking. Angewandte Chemie International Edition. 2008;47(37):7018–21.

30. Brunner J, Senn H, Richards FM. 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. J. Biol. Chem. 1980 Apr 25;255(8):3313–8.

31. Bayley H. Photogenerated Reagents in Biochemistry and Molecular Biology. Elsevier; 2000.

32. Dubinsky L, Krom BP, Meijler MM. Diazirine based photoaffinity labeling. Bioorganic & Medicinal Chemistry. 2012 Jan 15;20(2):554–70.

33. Hosoya T, Hiramatsu T, Ikemoto T, Aoyama H, Ohmae T, Endo M, et al. Design of dantrolene-derived probes for radioisotope-free photoaffinity labeling of proteins involved in the physiological Ca2+ release from sarcoplasmic reticulum of skeletal muscle. Bioorganic & Medicinal Chemistry Letters. 2005 Mar 1;15(5):1289–94.

34. Yu S-H, Wands AM, Kohler JJ. Photoaffinity probes for studying carbohydrate biology. J Carbohydr Chem. 2012;31(4-6):325–52.

35. Yu S-H, Boyce M, Wands AM, Bond MR, Bertozzi CR, Kohler JJ. Metabolic labeling enables selective photocrosslinking of O-GlcNAc-modified proteins to their binding partners. PNAS. 2012 Mar 27;109(13):4834–9.

36. Dorman G, Prestwich GD. Benzophenone Photophores in Biochemistry. Biochemistry. 1994 May 1;33(19):5661–73.

37. Pavlova O, Peterson JH, Ieva R, Bernstein HD. Mechanistic link between β barrel assembly and the initiation of autotransporter secretion. Proc. Natl. Acad. Sci. U.S.A. 2013 Mar 5;110(10):E938–947.

38. Sato S, Mimasu S, Sato A, Hino N, Sakamoto K, Umehara T, et al. Crystallographic Study of a Site-Specifically Cross-Linked Protein Complex with a Genetically Incorporated Photoreactive Amino Acid,. Biochemistry. 2011 Jan 18;50(2):250–7.

39. Bremer AA, Leeman SE, Boyd ND. Evidence for Spatial Proximity of Two Distinct Receptor Regions in the Substance P (SP)·Neurokinin-1 Receptor (NK-1R) Complex Obtained by Photolabeling the NK-1R withp-Benzoylphenylalanine3-SP. J. Biol. Chem. 2001 Jun 22;276(25):22857–61.

40. Lequin O, Bolbach G, Frank F, Convert O, Girault-Lagrange S, Chassaing G, et al. Involvement of the Second Extracellular Loop (E2) of the Neurokinin-1 Receptor in

the Binding of Substance P PHOTOAFFINITY LABELING AND MODELING STUDIES ,. J. Biol. Chem. 2002 Jun 21;277(25):22386–94.

41. Pham V, Dong M, Wade JD, Miller LJ, Morton CJ, Ng H, et al. Insights into Interactions between the α-Helical Region of the Salmon Calcitonin Antagonists and the Human Calcitonin Receptor using Photoaffinity Labeling. J. Biol. Chem. 2005 Aug 5;280(31):28610–22.

42. Wittelsberger A, Thomas BE, Mierke DF, Rosenblatt M. Methionine acts as a "magnet" in photoaffinity crosslinking experiments. FEBS Letters. 2006 Mar 20;580(7):1872–6.

43. Bisello A, Adams AE, Mierke DF, Pellegrini M, Rosenblatt M, Suva LJ, et al. Parathyroid Hormone-Receptor Interactions Identified Directly by Photocross-linking and Molecular Modeling Studies. J. Biol. Chem. 1998 Aug 28;273(35):22498–505.

44. Farrell IS, Toroney R, Hazen JL, Mehl RA, Chin JW. Photo-cross-linking interacting proteins with a genetically encoded benzophenone. Nat Meth. 2005 May;2(5):377–84.

45. Ali ST, Guest JR. Isolation and characterization of lipoylated and unlipoylated domains of the E2p subunit of the pyruvate dehydrogenase complex of Escherichia coli. Biochem J. 1990 Oct 1;271(1):139–45.

46. Green DE, Morris TW, Green J, Cronan JE, Guest JR. Purification and properties of the lipoate protein ligase of Escherichia coli. Biochem J. 1995 Aug 1;309(Pt 3):853–62.

47. Fujiwara K, Toma S, Okamura-Ikeda K, Motokawa Y, Nakagawa A, Taniguchi H. Crystal Structure of Lipoate-Protein Ligase A from Escherichia coli DETERMINATION OF THE LIPOIC ACID-BINDING SITE. J. Biol. Chem. 2005 Sep 30;280(39):33645–51.

48. Fujiwara K, Maita N, Hosaka H, Okamura-Ikeda K, Nakagawa A, Taniguchi H. Global Conformational Change Associated with the Two-step Reaction Catalyzed by Escherichia coli Lipoate-Protein Ligase A. J. Biol. Chem. 2010 Mar 26;285(13):9971–80.

49. Uttamapinant C, White KA, Baruah H, Thompson S, Fernández-Suárez M, Puthenveetil S, et al. A fluorophore ligase for site-specific protein labeling inside living cells. PNAS. 2010 Jun 15;107(24):10914–9.

50. Slavoff SA, Liu DS, Cohen JD, Ting AY. Imaging Protein–Protein Interactions inside Living Cells via Interaction-Dependent Fluorophore Ligation. J. Am. Chem. Soc. 2011 Dec 14;133(49):19769–76.

51. Jin X, Uttamapinant C, Ting AY. Synthesis of 7-Aminocoumarin by Buchwald–Hartwig Cross Coupling for Specific Protein Labeling in Living Cells. ChemBioChem. 2011;12(1):65–70.

52. Cohen JD, Thompson S, Ting AY. Structure-Guided Engineering of a Pacific Blue Fluorophore Ligase for Specific Protein Imaging in Living Cells. Biochemistry. 2011 Sep 27;50(38):8221–5.

53. Puthenveetil S, Liu DS, White KA, Thompson S, Ting AY. Yeast Display Evolution of a Kinetically Efficient 13-Amino Acid Substrate for Lipoic Acid Ligase. J. Am. Chem. Soc. 2009 Nov 18;131(45):16430–8.

54. Stadtman ER, Levine RL. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. Amino Acids. 2003 Dec 1;25(3-4):207–18.

55. Hensley K. Detection of Protein Carbonyls by Means of Biotin Hydrazide-Streptavidin Affinity Methods. In: Kurien BT, Scofield RH, editors. Protein Blotting and Detection [Internet]. Humana Press; 2009 [cited 2013 Apr 15]. p. 457–62. Available from: http://link.springer.com/protocol/10.1007/978-1-59745-542-8_46

56. Hensley K, Williamson KS. Protein Carbonyl Determination Using Biotin Hydrazide. Methods in Biological Oxidative Stress [Internet]. New Jersey: Humana Press; [cited 2013 Apr 15]. p. 195–200. Available from: http://www.springerprotocols.com/Abstract/doi/10.1385/1-59259-424-7:195

57. Ye S, Köhrer C, Huber T, Kazmi M, Sachdev P, Yan ECY, et al. Site-specific Incorporation of Keto Amino Acids into Functional G Protein-coupled Receptors Using Unnatural Amino Acid Mutagenesis. J. Biol. Chem. 2008 Jan 18;283(3):1525–33.

58. Brustad EM, Lemke EA, Schultz PG, Deniz AA. A General and Efficient Method for the Site-Specific Dual-Labeling of Proteins for Single Molecule Fluorescence Resonance Energy Transfer. J. Am. Chem. Soc. 2008 Dec 31;130(52):17664–5.

59. Kalia J, Raines RT. Hydrolytic Stability of Hydrazones and Oximes. Angewandte Chemie International Edition. 2008;47(39):7523–6.

60. Zeng Y, Ramya TNC, Dirksen A, Dawson PE, Paulson JC. High-efficiency labeling of sialylated glycoproteins on living cells. Nat Meth. 2009 Mar;6(3):207–9.

61. Liang J, Choi J, Clardy J. Refined structure of the FKBP12–rapamycin–FRB ternary complex at 2.2 Å resolution. Acta Crystallographica Section D Biological Crystallography. 1999 Apr 1;55(4):736–44.

62. Choi J, Chen J, Schreiber SL, Clardy J. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. Science. 1996 Jul 12;273(5272):239–42.

63. Banaszynski LA, Liu CW, Wandless TJ. Characterization of the FKBP.rapamycin.FRB ternary complex. J. Am. Chem. Soc. 2005 Apr 6;127(13):4715–21.

64. Dormán G, Prestwich GD. Benzophenone photophores in biochemistry. Biochemistry. 1994 May 17;33(19):5661–73.

65. Hassan AQ, Stubbe J. Mapping the subunit interface of ribonucleotide reductase (RNR) using photo cross-linking. Bioorganic & Medicinal Chemistry Letters. 2008 Nov 15;18(22):5923–5.

66. Tsubomura H, Yamamoto N, Tanaka S. Transient absorption spectra of benzophenone studied by the flash excitation. Chemical Physics Letters. 1967 Oct;1(8):309–10.

67. Wierzchowski KL. Intramolecular electron transfer between tryptophan radical and tyrosine in oligoproline-bridged model peptides and hen egg-white lysozyme. Acta Biochim. Pol. 1997;44(4):627–44.

68. Kalcheva N, Albala J, O'Guin K, Rubino H, Garner C, Shafit-Zagardo B. Genomic structure of human microtubule-associated protein 2 (MAP-2) and characterization of additional MAP-2 isoforms. Proc Natl Acad Sci U S A. 1995 Nov 21;92(24):10894–8.

69. Yao JZ, Uttamapinant C, Poloukhtine A, Baskin JM, Codelli JA, Sletten EM, et al. Fluorophore Targeting to Cellular Proteins via Enzyme-Mediated Azide Ligation and Strain-Promoted Cycloaddition. J. Am. Chem. Soc. 2012 Feb 29;134(8):3720–8.

70. Ayscough KR. Coupling actin dynamics to the endocytic process in Saccharomyces cerevisiae. Protoplasma. 2005 Oct;226(1-2):81–8.

71. Pedersen S., Hoffmann E., Mills J. The cytoskeleton and cell volume regulation. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology. 2001 Oct;130(3):385–99.

72. Nolen BJ, Tomasevic N, Russell A, Pierce DW, Jia Z, McCormick CD, et al. Characterization of two classes of small molecule inhibitors of Arp2/3 complex. Nature. 2009 Aug 20;460(7258):1031–4.

73. Wada A, Fukuda M, Mishima M, Nishida E. Nuclear export of actin: a novel mechanism regulating the subcellular localization of a major cytoskeletal protein. EMBO J. 1998 Mar 16;17(6):1635–41.

74. Effects of cytochalasin and phalloidin on actin. J Cell Biol. 1987 Oct 1;105(4):1473–8.

75. Yarmola EG, Somasundaram T, Boring TA, Spector I, Bubb MR. Actin-latrunculin A structure and function. Differential modulation of actin-binding protein function by latrunculin A. J. Biol. Chem. 2000 Sep 8;275(36):28120–7.

76. Liu DS, Tangpeerachaikul A, Selvaraj R, Taylor MT, Fox JM, Ting AY. Diels–Alder Cycloaddition for Fluorophore Targeting to Specific Proteins inside Living Cells. J. Am. Chem. Soc. 2012 Jan 18;134(2):792–5.

77. Fernández-Suárez M, Baruah H, Martínez-Hernández L, Xie KT, Baskin JM, Bertozzi CR, et al. Redirecting lipoic acid ligase for cell surface protein labeling with small-molecule probes. Nature Biotechnology. 2007;25(12):1483–7.

Chapter 5. Effort towards membrane protein topology mapping with an ascorbate peroxidase enzyme

We would like to thank Dawei Jiang of the Clapham lab for providing us with LETM1-mCherry plasmid in pEGFP-N_3 vector.

Introduction

Membrane proteins are important players in the cell and carry out crucial functions from acting as small molecule transporters to taking roles in complex signaling pathways. It has been estimated that more than half of the drugs on the market are directed against membrane protein receptors (1). However, membrane proteins are still very underrepresented among solved atomic protein structures, making up less than 1% of the structures available in the Protein Data Bank (PDB) (1). The reason for the lack of membrane protein structures lies in the technical challenges associated with expressing them in large enough quantities for structural analysis, solubilizing them in appropriate detergents, and crystallizing them for structure determination. Aside from solving an entire three-dimensional structure of a membrane protein, knowing the *topology* - the transmembrane regions and the in-out orientation across the membrane - of a membrane protein can also provide important clues as to the protein's function, potential interaction partners, and trafficking. Here, we describe our preliminary effort towards developing a method that can be used to study the topology of membrane proteins.

Membrane proteins are difficult to mass-produce using the standard protein production workhorse, *E. coli.* Because the bacterial machinery is quite different from that of mammalian cells, researchers need to identify suitable bacterial growth conditions that can induce proper folding of membrane proteins. Recent advances in cell-free systems of protein translation have enabled the production of tens of milligrams of membrane proteins (2), but the challenge is still to keep them soluble for subsequent purification and analysis. Even when expressed at high levels, membrane proteins are inherently difficult to purify to extremely high purity. In addition, because of the large hydrophobic regions in membrane proteins, they are often prone to aggregation, making it necessary to screen and test numerous stabilizing buffers and detergent solutions to keep these proteins soluble. Solid-state NMR spectroscopy methods, which avoid the requirement for crystallization, are rapidly advancing, but sensitivity and resolution enhancements are still needed to make this a robust technology (3,4).

Since there are no general and reliable methods to obtain topological information of membranes proteins, computational prediction techniques and biochemical methods are also used to acquire structural information. Computational methods to predict membrane protein topology or putative transmembrane region have been very helpful, but the precision of these methods is limited by the amount of available structures (5,6). Various biochemical techniques have been instrumental in elucidating membrane protein topology and orientation. One method introduces enzymes such as alkaline phosphatase, β-galactosidase and β-lactamase as reporters by fusing them to specific regions of the membrane protein (7). These enzymes often have different stabilities and activities depending on which compartment of the cell they are exposed. Alkaline phosphatase is active upon the formation of four disulfide bonds and dimer formation in an oxidizing environment (8). The presence of an active alkaline phosphatase, as measured by its activity towards a chromogenic substrate, can provide evidence that a region of the membrane protein is localized to an oxidizing environment. The enzymes β-galactosidase and β-lactamase work in a similar and complementary fashion. The β-galactosidase located in the cytoplasm is properly folded and active, while β-galactosidase fused downstream of an export signal is not folded properly and therefore is inactive (9). For β-lactamase, if the enzyme is fused to the portion of the protein that is exposed to the outside of a bacterial cell, the enzyme has the ability to protect the cells against antibiotic such as ampicillin, which otherwise would inactivate essential proteins anchored on the outside of the cytoplasmic membrane that are crucial for cell wall biosynthesis. All three enzymes are applicable to heterologously expressed eukaryotic membrane proteins. However, they are not generalizable to all compartmental membranes in the cell. For example, alkaline phosphatase is only applicable if the membrane separates between an oxidizing and a reducing environment (7).

Aside from fusing enzymes to various regions of membrane proteins, researchers have often used exogenous proteases to assay for membrane protein topology. Integral membrane proteins cross the membrane either in a single-pass or a multi-pass fashion that exposes their hydrophilic loops to one of the two compartments separated by the membrane. The impermeability of the membrane allows membrane-impermeant small molecules and proteases to be utilized to verify putative transmembrane structures. (10). The most common practice is to externally supply promiscuous proteolytic enzymes that cannot cross the membrane, then determine which regions of the membrane protein are cleaved or protected from proteolysis. Regions of cleavage can be analyzed by immunoblot using antibodies directed against the protein of interest. However, protease accessibility assays may fail to identify exposed regions of the membrane protein, especially if the region is small (11).

The goal of this project is to develop a microscopy-based imaging methodology to study membrane protein topology. The method uses an enzymatic labeling approach in which the enzyme is fused to different regions of a membrane protein of interest. The enzyme is an engineered plant ascorbate peroxidase enzyme named APEX developed in our lab that has been utilized as a genetic tag for EM imaging (12) as well as in proteomic mapping inside live cells (13). In EM imaging, APEX oxidizes the substrate 3, 3'-diaminobenzidine (DAB) in an H_2O_2-dependent manner to form DAB polymers that can serve as an EM contrast agent after treatment with OsO_4. Martell *et al.* utilized the localized and membrane restricted feature of DAB staining to generate EM images of various cellular proteins fused to APEX. Based on the location of the EM contrast, Martell and coworkers were able to demonstrate that both termini of the mitochondrial calcium uniporter (MCU), an inner mitochondrial membrane (IMM) protein, face the mitochondrial matrix. Even though EM imaging provides nanometer resolution, the instrumentation and specific skills required to perform EM imaging are not always available or accessible to all research labs. Therefore, we wanted to develop an alternative method that is able to use microscopy imaging of the APEX labeling patterns to provide clues to the topology of membrane proteins.

The imaging method is based on APEX labeling of a phenol small molecule, biotin-tyramide. In addition to using DAB as a substrate, APEX can also oxidize phenol to phenoxyl radicals in an H_2O_2-dependent manner. When live mammalian cells expressing APEX targeted to a specific cellular compartment, such as the mitochondrial matrix, were incubated with biotin-tyramide and H_2O_2, endogenous proteins were covalently modified with biotin only in the compartment where APEX was targeted (13). In proteomic mapping, Rhee *et al.* targeted APEX to the mitochondrial matrix or intermembrane space (IMS) and performed biotin-tyramide labeling by adding H_2O_2 for 1 minute (13). The short lifetime of phenoxyl radicals allow biotin-tyramide radicals to only react with APEX-proximal proteins. After live cell labeling, biotinylated proteins in the APEX-specific cellular compartment were enriched and analyzed by tandem mass spectrometry.

Dr. Hyun-woo Rhee subsequently discovered that not only did DAB labeling by APEX produce a localized signal that delineated cellular ultrastructures, APEX labeling using biotin-tyramide followed by staining with streptavidin-fluorophore conjugate also showed different labeling patterns depending on the location of the APEX enzyme. Knowing from previous experiments that the biotin-tyramide radical does not cross the membrane (13), Dr. Rhee observed that if APEX is localized in a strictly membrane-bound compartment, such as the mitochondrial matrix, the pattern of biotin-tyramide labeling appeared to be tight and restricted, delineating the mitochondria of the cell. However, if APEX is in the cytosol or in places where the biotin-tyramide radical can easily diffuse during the time of labeling, such as the mitochondrial intermembrane space (IMS) because of the porous outer mitochondrial membrane (OMM), biotin-tyramide labeling pattern appeared diffusive and throughout the cell. We wanted to use the location-specific APEX biotinylation pattern as a tool to infer membrane protein topology (Figure 5-1). APEX fused to the matrix side of an inner mitochondrial membrane (IMM) protein would show a restricted and mitochondrial-specific labeling pattern because matrix is enclosed by the tight IMM (Figure 5-1, left). In contrast, APEX fused to the IMS side of an IMM protein would show a diffusive biotinylation pattern because the outer mitochondrial membrane (OMM) is porous (Figure 5-1, right), allowing biotin-tyramide radical to diffuse. Depending on the enzymatic fusion site of APEX to the membrane protein of interest, the labeling pattern can appear either restricted inside certain cellular organelle or diffusive throughout the entire cell. The restricted or diffusive labeling patterns can provide topological clues to the location of certain regions of a membrane protein.

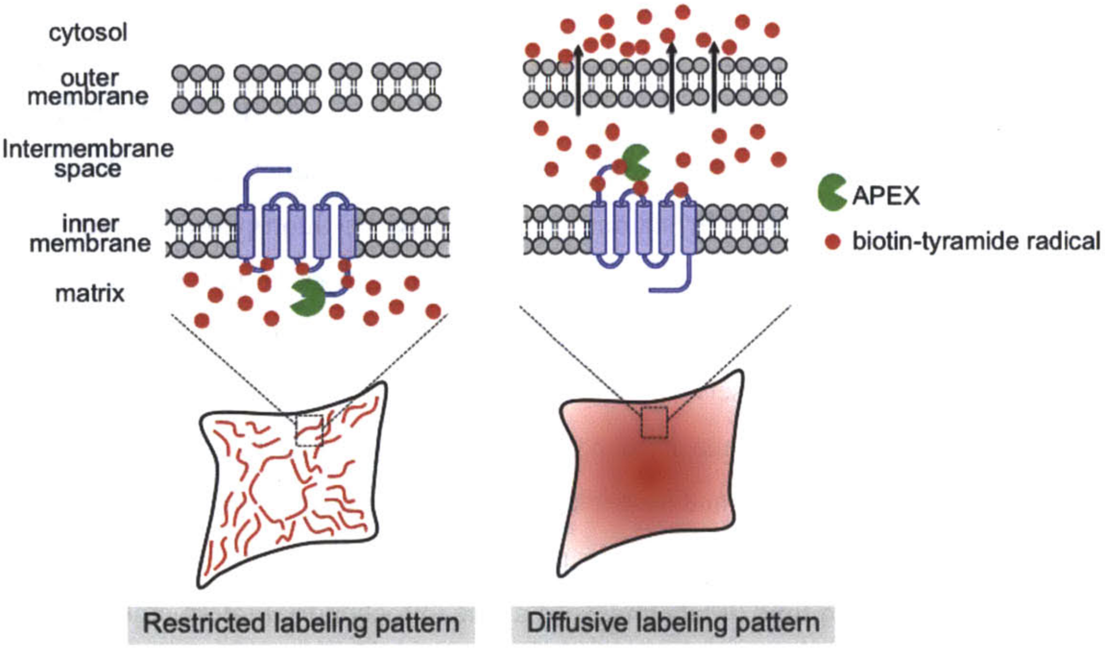

cytosol
outer membrane
Intermembrane space
inner membrane
matrix
APEX
● biotin-tyramide radical
Restricted labeling pattern
Diffusive labeling pattern

Figure 5-1. Scheme depicting labeling with APEX (green pacman) fused to either matrix or IMS side of an IMM protein. Biotin-tyramide radical (red dot) generated in the matrix would remain localized there because the IMM is very tight and impassable, therefore generating what appeared to be mitochondria-localized biotinylation signal on a conventional microscope. In contrast, biotin-tyramide radical generated in the IMS would have the opportunity to diffuse across the porous OMM and escape into the cytosol of the cell, leading to a diffusive or whole-cell labeling signal.

We wished to fuse APEX to different regions of a membrane protein and apply APEX labeling of biotin-tyramide to develop a technique for membrane protein topology prediction. The method is easy to implement and requires only a basic microscope to observe tight or diffusive labeling pattern. In this chapter, we describe the development and effort towards applying APEX labeling to investigate the membrane topology of a mitochondrial inner membrane protein LETM1.

EF-hand domain-containing protein 1 - LETM1

In the work of Rhee and coworkers, mitochondrial matrix-targeted APEX was used to label and enrich matrix localized proteins (13). By analyzing the 495 proteins identified in the matrix proteome, we focused on the 83 kDa mitochondrial inner membrane protein EF-hand domain-containing protein 1 (LETM1) because it was highly enriched in the mitochondrial matrix proteome with two biotinylated peptides identified (13). The first biotinylated peptide, KLEEGGPVYSPPAEVVVK (biotin-tyramide modified tyrosine residue is in red) in LETM1 mapped to amino acid residues 133-150, which corresponds to a region upstream from the putative transmembrane domain (amino acids 206-229) (Figure 5-2). The second biotinylated peptide in LETM1 (TGEEKYVEESK) is located near the C-terminus of the protein, from amino acid residues 593-603 (Figure 5-2). Because the biotin-tyramide radical does not cross the membrane, both biotinylated peptides of LETM1 must be located in the mitochondrial matrix. This result is interesting because under the assumption that the N-terminus of LETM1 is located in the matrix, there would have to be at least a second or possibly more transmembrane domains in order for the second biotinylated peptide near the C-terminus to be located in the matrix. The topology of LETM1 has significant implications on the protein's proposed functions (11,14–16). Therefore, we decided to study the topology of LETM1 in the inner mitochondrial membrane (IMM).

In 1999, Endele *et al.* first cloned and characterize the LETM1 gene, which was identified as a chromosomally deleted gene in almost all patients with Wolf-Hirschhorn syndrome (WHS), a genetic disease characterized by growth delay, seizures, mental retardation, and characteristic facial features (17). Computational sequence analysis of LETM1 shows its sequence identity ranges from 31% in *S. cerevisiae* to 83% in mouse (18). Putative structural features of human LETM1 include a 24-amino acid transmembrane region at the N-terminus; a highly conserved PKC phosphorylation site; two EF-binding motifs near the C-terminus; and several coiled-coil regions (15,18) (Figure 5-2). Studies by Schlickum *et al.* further demonstrated that a mitochondrial targeting signal sequence is located at the N-terminus, without which LETM1 appeared to be mis-localized to the cytosol (18).

MASILLRSCRGRAPARLPPP	PRYTVPRGSPGDPAHLSCAS	TLGLRNCLNVPFGCCTPIHP	60
VYTSSRGDHLGCWALRPECL	RIVSRAPWTSTSVGFVAVGP	QCLPVRGWHSSRPVRDDSVV	120
EKSLKSLKDKNK**KLEEGGPV**	**YSPPAEVVVK**KSLGQRVLDE	LKHYYHGFRLLWIDTKIAAR	180
MLWRILNGHSL**TRRE**RRQFL	**RICADLFRLVPFLVFVVVPF**	**MEFLLPVA**VKLFPNMLPSTF	240
ETQSLKEERLKKELRVKLEL	AKFLQDTIEEMALKNKAAKG	SATKDFSVFFQKIRETGERP	300
SNEEIMRFSKLFEDELTLDN	LTRPQLVALCKLLELQSIGT	NNFLRFQLTMRLRSIKADDK	360
LIAEEGVDSLNVKELQAACR	ARGMRALGVTEDRLRGQLKQ	WLDLHLHQEIPTSLLILSRA	420
MYLPDTLSPADQLKSTLQTL	PEIVAKEAQVKVAEVEGEQV	DNKAKLEATLQEEAAIQQEH	480
REKELQKRSEVAKDFEPERV	VAAPQRPGTEPQPEMPDTVL	QSETLKDTAPVLEGLKEEEI	540
TKEEIDILSDACSKLQEQKK	SLTKEKEELELLKEDVQ**DYS**	**EDLQEIKKEL**SK**TGEEKYVE**	600
ESKASKRLTKRVQQMIGQID	GLISQLEMDQQAGKLAPANG	MPTGENVISVAELINAMKQV	660
KHIPESKLTSLAAAL**DENKD**	**GKVNIDDL**VKVIELVDKEDV	HISTSQVAEIVATLEKEEKV	720
EEKEKAKEKAEKEVAEVKS			739

Figure 5-2. Amino acid sequence of human LETM1. Orange indicates conserved PKC phosphorylation site; light blue indicates putative transmembrane region; underlining represents two coiled-coil regions; green indicates two 13-amino acids EF-hand binding motifs. Two biotinylated peptides labeled by matrix-targeted APEX and identified by tandem MS in matrix proteome (13) were in dark blue with biotinylated tyrosine residues highlighted in red.

Two yeast homologs of LETM1, Mdm38 and Ylh47, have been studied in depth to gain insight into the role of LETM1 (15,16,18,19). They do not contain the EF-binding motifs present in human LETM1 but otherwise share high protein sequence identity especially near the N-terminus of the proteins (15). Both yeast homolog proteins are localized to the mitochondrial inner membrane as shown by subcellular fractionation (19,20), protease accessibility assays (19), and alkaline extraction of membrane proteins (15,19). Functional analysis of Mdm38 further revealed that deletion of this gene in yeast caused elevated intramitochondrial potassium level, swollen mitochondria, and decreased membrane potential (19). Numerous functional roles have been proposed for LETM1, from maintaining mitochondrial volume (20) and shape (21), acting as a Ca^{2+}/H^{+} antiporter (14), regulating mitochondrial fragmentation (22), to aiding mitochondrial protein export (15) and regulating mitochondrial ATP production (23). However, despite the many functional implications of this protein, the exact correlation between LETM1 and WHS, as well as how it is transported into the mitochondria and its topology in the IMM are still unclear.

Experiments investigating the topology of LETM1 in the IMM revolved around subcellular fractionation of the mitochondria and protease accessibility assay. A number of works agree that LETM1 is an IMM protein with no detectable part exposed to the IMS (15,19,24). However, Hasegawa *et al.* suggest that LETM1 is partially exposed to the IMS because when the IMM was compromised by the addition of digitonin, matrix proteins like Hsp60 and GDH were protected from trypsin digestion, but LETM1 was partially digested (20). The work of Dimmer *et al.* shows a proposed membrane orientation of LETM1 as a single transmembrane IMM protein with its N-terminus in the IMS (24). Yet, they were unable to detect the IMS-exposed region using a proteinase K accessibility assay. In addition to the insensitivity of most protease accessibility or protection assays, subcellular fractionations are often prone to contamination (24), further leading to misguided or contradictory conclusions. We hope to provide an alternative method based on our APEX/biotin-tyramide labeling assay to infer the topology of LETM1.

Restricted and diffusive labeling patterns of localized APEX

To first demonstrate the restricted and diffusive labeling patterns, we performed biotin-tyramide labeling in live HEK 293T and COS-7 cells expressing either APEX localized to the mitochondrial matrix or the IMS. We expected to see a tight labeling pattern when APEX is in the mitochondrial matrix because the matrix is a tightly membrane-bound compartment. In contrast, we expected to see a diffusive labeling pattern with signal spread out through the cell when APEX is in the IMS, which is not tightly bound region because the OMM is very porous. We incubated the cells in cell culture media containing biotin-tyramide for 1 hour at 37 °C. We then added 1 mM H_2O_2 to the cells to initiate labeling. Following labeling for 1 minute, the cells were fixed with paraformaldehyde, stained with neutravidin-Alexa Fluor 647 conjugate, and imaged on a confocal microscope.

In HEK 293T cells, we observed tight and restricted Alexa Fluor 647 signal co-localizing with mitochondrial green fluorescent protein (GFP) marker when labeling with APEX in the matrix (Figure 5-3). The labeling pattern highlights the mitochondrial network, which appears as small dots in HEK 293T cells. In contrast, labeling with APEX localized to the mitochondrial IMS shows Alexa Fluor 647 signal spread out throughout the cell. In COS-7 cells, we also saw tight Alexa Fluor 647 signal delineating the mitochondria and co-localizing with mitochondrial maker. When labeling with APEX in the IMS of COS-7 cells, Alexa Fluor 647 signal appeared to be fuzzy and spread out in COS-7 cells (Figure 5-3).

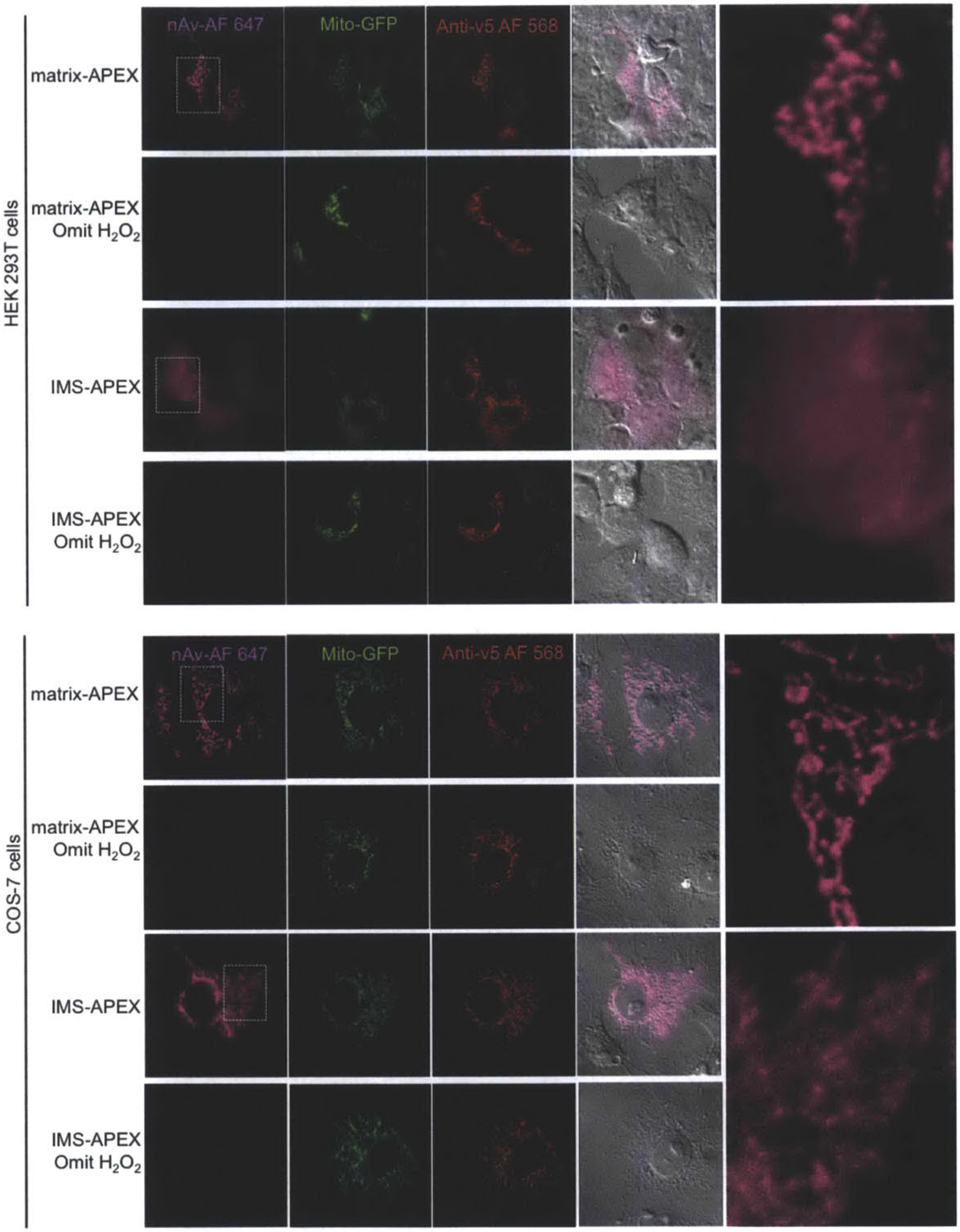

HEK 293T cells
nAv-AF 647
Mito-GFP
Anti-v5 AF 568
matrix-APEX
matrix-APEX Omit H_2O_2
IMS-APEX
IMS-APEX Omit H_2O_2
COS-7 cells
nAv-AF 647
Mito-GFP
Anti-v5 AF 568
matrix-APEX
matrix-APEX Omit H_2O_2
IMS-APEX
IMS-APEX Omit H_2O_2

Figure 5-3. Restricted and diffusive biotin-tyramide labeling patterns in HEK 293T and COS-7 cells. HEK 293T and COS-7 cells expressing matrix localized APEX or IMS localized APEX were labeled with biotin-tyramide and H_2O_2. Cells were then fixed and

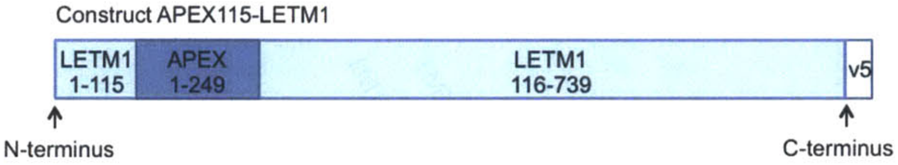
stained with neutravidin (nAv)-Alexa Fluor 647 to visualize biotinylated proteins, and antibody to visualize v5-epitope tag on the APEX enzyme. Mito-GFP is a mitochondrial marker. Negative control was performed with H_2O_2 omitted. Region of the labeled image boxed in white, dashed box is zoomed in as a larger image on the right.

Having established that APEX localized to either the mitochondrial matrix or the IMS gave different biotinylation patterns, we moved forward to fuse APEX to various regions of the inner mitochondrial membrane protein LETM1 to investigate which regions of LETM1 are in the matrix or in the IMS.

Biotin-tyramide labeling patterns of APEX-fused LETM1

Patterns and levels of expression of APEX-LETM1 fusion proteins

To apply the difference in labeling patterns in deducing LETM1 topology in the mitochondrial inner membrane, we fused APEX to various regions of LETM1 and performed biotin-tyramide labeling in HEK 293T cells. Based on the amino acid sequence of human LETM1 in Figure 5-2, we chose regions of LETM1 that do not have previous functional designation to fuse the APEX enzyme. We cloned a total of 16 LETM1-APEX fusion constructs (Figure 5-4), all of which have v5 epitope tag at the end of LETM1. For example, APEX115-LETM1 indicates that APEX is fused after amino acid 115 of LETM1 (Figure 5-4, top). Most of the constructs expressed well with clear mitochondrial pattern as confirmed by anti-v5 immunostaining and co-localization with mitochondrial GFP marker (Figure 5-6).



Construct name	APEX insertion after	Expression pattern/level
APEX115-LETM1	aa 115	mitochondrial / unstable
APEX156-LETM1	aa 156	mitochondrial / strong
APEX180-LETM1	aa 180	mitochondrial / strong
APEX187-LETM1	aa 187	mitochondrial / strong
APEX201-LETM1	aa 201	mitochondrial / strong
APEX235-LETM1	aa 235	mitochondrial / strong
APEX285-LETM1	aa 285	mitochondrial / strong
APEX300-LETM1	aa 300	mitochondrial / strong
APEX365-LETM1	aa 365	mitochondrial / strong
APEX385-LETM1	aa 385	mitochondrial / strong
APEX444-LETM1	aa 444	mitochondrial / weak
APEX524-LETM1	aa 524	mitochondrial / strong
APEX591-LETM1	aa 591	whole cell / weak
APEX609-LETM1	aa 609	whole cell / strong
APEX657-LETM1	aa 657	whole cell / strong
LETM1-v5-APEX	aa 739	mitochondrial / strong
LETM1-v5	-	mitochondrial / strong

Figure 5-4. Constructs with APEX fused to different regions of LETM1. (Top) Schematic representation of APEX115-LETM1 fusion construct. APEX (total 249 amino acids) is inserted after amino acid 115 of LETM1. An epitope tag v5 is fused at the C-terminus of LETM1. (Bottom) List of APEX-LETM1 fusion constructs, their expression patterns and levels of expression. Microscopy images of each fusion construct are shown in Figure 5-6.

A few constructs did not express stably or appeared to be mistargeted. Construct APEX115-LETM1 sometimes was undetectable by immunostaining on fixed HEK 293T cells and showed proteolytic products on Western blot (Figure 5-5, see Experimental methods for transfection and Western blot procedures). This is likely caused by the APEX fusion site being immediately after the N-terminal mitochondrial targeting region from amino acids 1 - 115 of LETM1. Schlickum *et al.* (18) showed that fusion of GFP before the targeting region or deletion of the targeting region both resulted in diffusive staining in the cytoplasm, not restricted to the mitochondria of the cell.

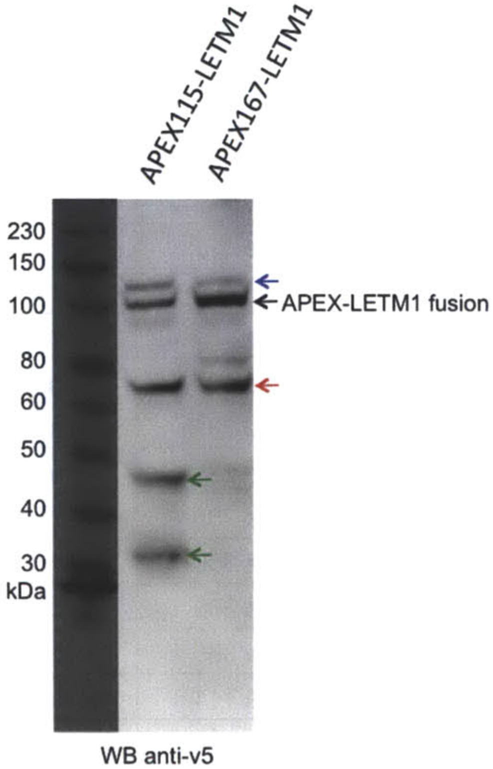


Figure 5-5. APEX115-LETM1 fusion protein shows cleavage products on Western blot. Immunostaining with anti-v5 antibody detected possible cleavage products (green arrows) of APEX115-LETM1. APEX167-LETM1 fusion protein does not show the same

cleavage products. APEX-LET1 fusion protein (~114 kDa) is indicated by black arrow. Red and blue arrows indicate artifacts caused by protein overexpression, which is discussed in the following section.

Three constructs with APEX fused near the C-terminus of LETM1, APEX591-LETM1, APEX609-LETM1, and APEX657-LETM1, did not target to the mitochondria, but instead appeared to have whole cell expression pattern (see Figure 5-6, anti-v5 AF 568 channel). Interestingly, in all three constructs, APEX is placed between the two EF-hand binding motifs. The first EF-hand binding motif is from amino acids 577 – 590; the second EF-hand binding motif is from amino acids 676 – 688. This is an indication that the region between the two EF-hand binding motifs could be critical to protein folding and therefore does not tolerate other protein tags. It is also possible that protein fusion in this region disrupts LETM1 folding, making the mitochondrial targeting region at the N-terminus inaccessible or cannot be recognized by transport proteins or chaperones to bring LETM1 to the mitochondria, leading to the whole cell expression pattern we observed with all three constructs.

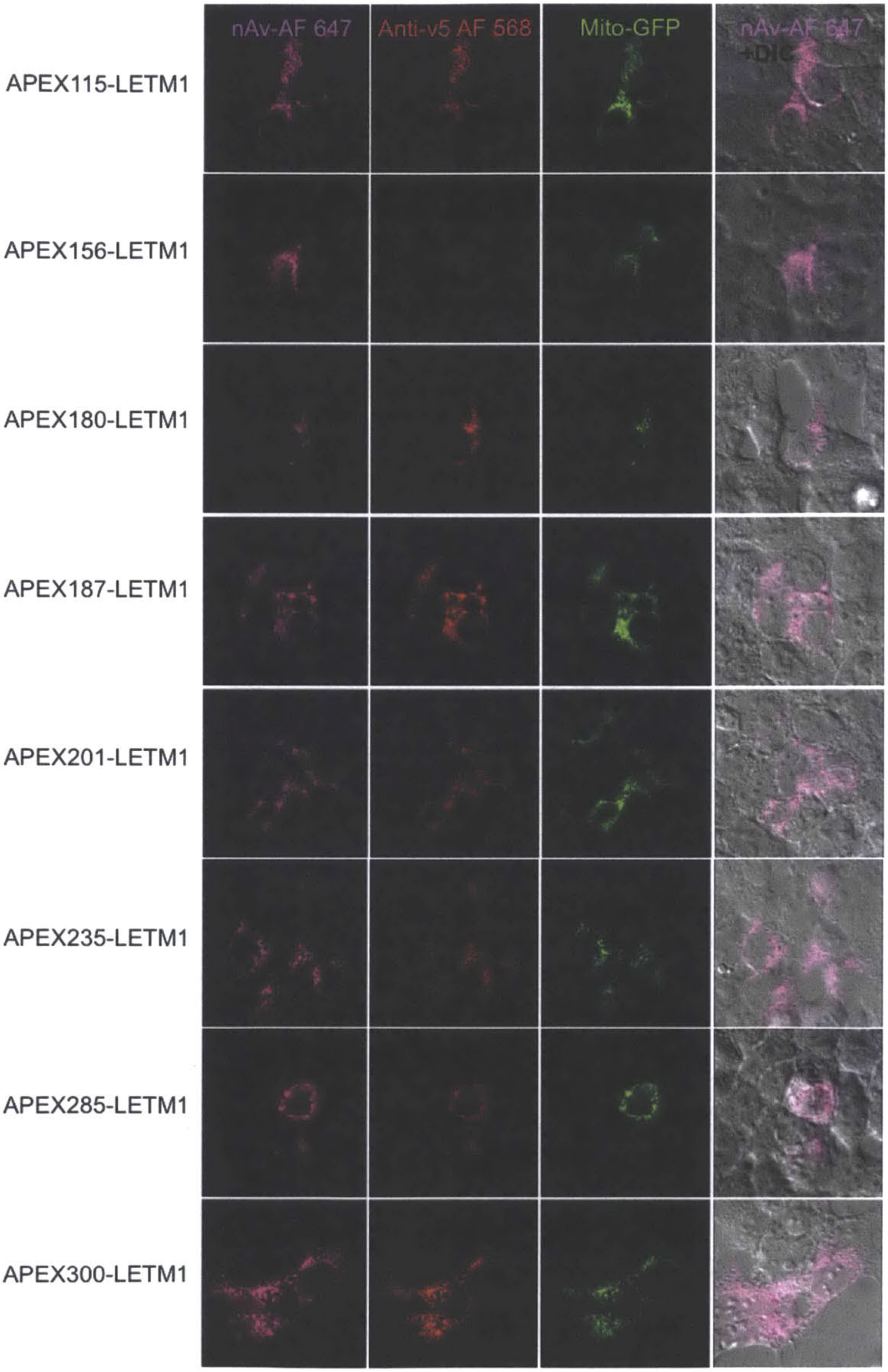
Biotin-tyramide labeling patterns of APEX-LETM1 fusion proteins

We performed biotin-tyramide labeling in HEK 293T cells expressing the APEX-LETM1 fusion constructs. Since LETM1 is a mitochondrial inner membrane protein (IMM), depending on the fusion sites of the APEX, certain APEX could be in the mitochondrial matrix, while other APEX could be in the IMS. We expected to see the fusion proteins give either restricted or diffusive labeling patterns. Based on the labeling patterns, we can infer the topology of LETM1 in the IMM.

We transfected HEK 293T cells with each of the APEX-LETM1 fusion constructs and a mitochondrial GFP marker using lipofectamine. At 18 hours after transfection, we incubated the cells in cell culture media containing biotin-tyramide for 1 hour at 37 °C. We then added 1 mM H_2O_2 to the media to initiate labeling. After 1 minute of labeling, the cells were washed with Dulbecco's phosphate buffered saline (DPBS) and fixed using 3.7% paraformaldehyde in DPBS. We immunostained the fixed cells with anti-v5 antibody to visualize the expression of each APEX-LETM1 protein and with neutravidin-

Alexa Fluor 647 conjugate (nAv-AF 647) to visualize the biotin-tyramide labeled proteins.

Figure 5-6 shows the biotin-tyramide labeling pattern (nAv-AF 647 channel) and the expression (mouse anti-v5-AF 568 channel) of each APEX-LETM1 fusion construct. We can also visualize the mitochondrial network via the expression of a mitochondrial marker, mito-GFP. The first nine APEX-LETM1 constructs all showed tight and restricted labeling pattern, with nAv-AF 647 signal co-localizing with both APEX-LETM1 expression (Anti-v5 AF 568 channel) and the mitochondrial GFP marker. Three constructs with APEX fused after amino acids 385, 444, and 524 of LETM1 all showed diffusive labeling pattern, with nAv-AF 647 signal spread out throughout the cell.


nAv-AF 647
Anti-v5 AF 568
Mito-GFP
nAv-AF 647
+DIC
APEX115-LETM1
APEX156-LETM1
APEX180-LETM1
APEX187-LETM1
APEX201-LETM1
APEX235-LETM1
APEX285-LETM1
APEX300-LETM1

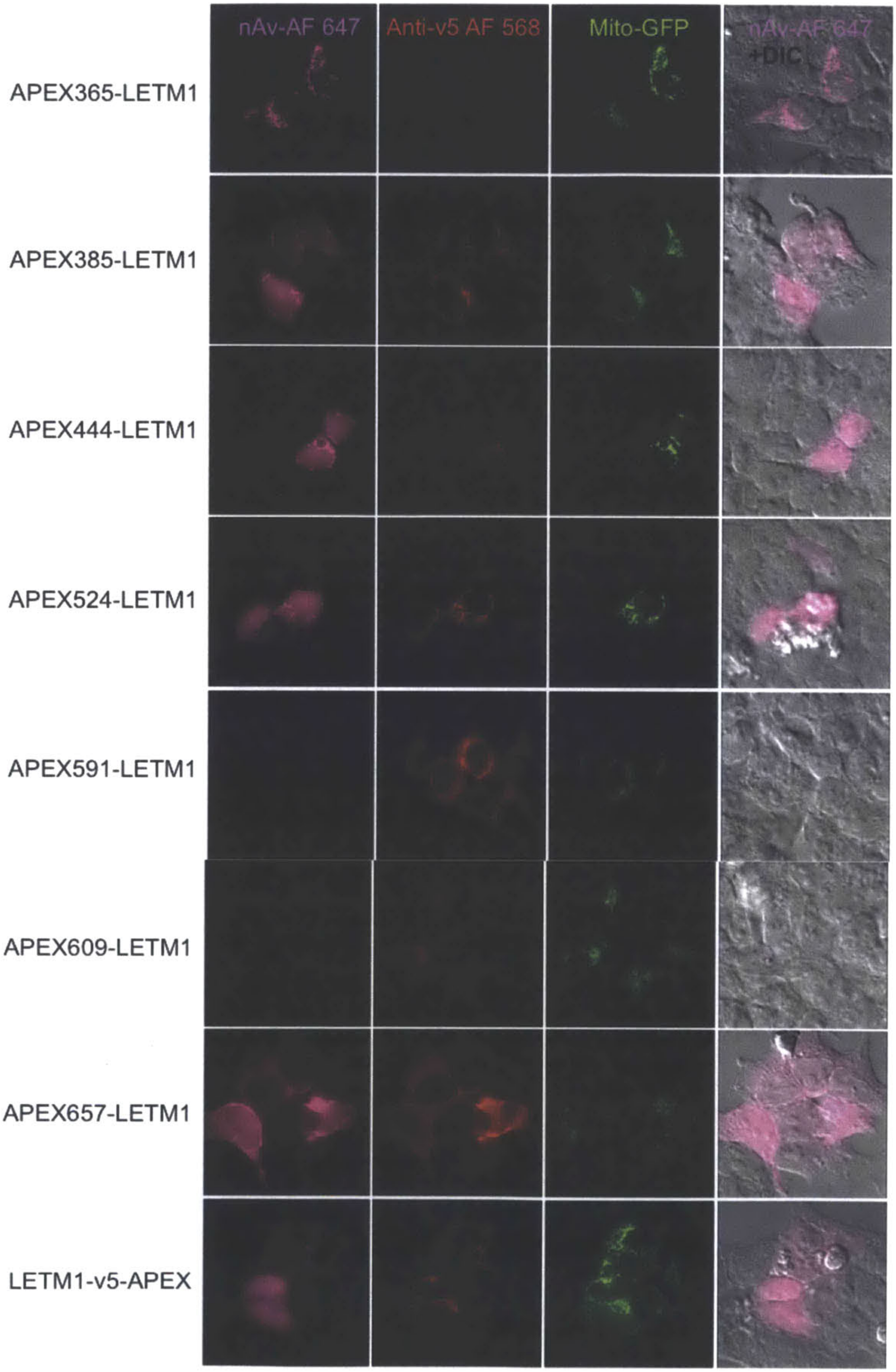

nAv-AF 647
Anti-v5 AF 568
Mito-GFP
nAv-AF 647 +DIC
APEX365-LETM1
APEX385-LETM1
APEX444-LETM1
APEX524-LETM1
APEX591-LETM1
APEX609-LETM1
APEX657-LETM1
LETM1-v5-APEX

Figure 5-6. Biotin-tyramide labeling with APEX-LETM1 fusion constructs in HEK 293T cells. HEK 293T cells expressing the indicated APEX-LETM1 fusion construct and a mitochondrial marker, mito-GFP were labeled with biotin-tyramide and H_2O_2. Cells were then fixed and stained with neutravidin-Alexa Fluor 647 conjugate (nAv-AF 647) to visualize biotinylated proteins, and with anti-v5 antibody to visualize APEX-LETM1 expression.

Our results from APEX-LETM1 fusion protein labeling provided hints to the membrane topology of LETM1. First, if we assume the direction of the putative transmembrane region from amino acids 206 - 229 is going from the matrix to the IMS, then we would expect APEX235-LETM1 to be located in the IMS since APEX is fused only 6 amino acids downstream of the transmembrane segment. However, we observed tight and restricted biotinylation labeling pattern of APEX235-LETM1, co-localizing with the mitochondrial GFP marker, which indicates that APEX is in the matrix (Figure 5-6). On the other hand, if we assume the direction of the transmembrane region were going from the IMS to the matrix, then we would expect construct APEX201-LETM1 with APEX fused after amino acid 201, just 5 amino acids upstream of the transmembrane segment, to be localized in the IMS. However, APEX201-LETM1 also showed tight labeling pattern (Figure 5-6). This result led us to speculate that perhaps the putative transmembrane region of LETM1 does not go through the entire IMM. Indeed, work by Frazier *et al.* using subcellular fractionation of mitochondria and carbonate extraction of membrane proteins suggested that the yeast homologs of LETM1, Mdm38 and Ylh47, could be only partially embedded in the lipid environment and may not fully penetrate the IMM (15). This is in agreement with our imaging result showing APEX located very near the transmembrane region (amino acids 206 - 229), either upstream in the case of APEX201-LETM1 or downstream in the case of APEX235-LETM1, both generated a restricted biotinylation labeling pattern. Our hypothesis is that the putative transmembrane segment of LETM1 might be embedded in IMM, perhaps passing through only one phospholipid layer of the IMM.

Controlling expression levels of LETM1 fusion proteins

In order to draw any conclusion regarding LETM1 membrane topology, in addition to having correct mitochondrial localization of the APEX-LETM1 fusion construct, we must also ensure that the constructs expressed at the expected molecular weight without artifacts such as protease cleavage. Either overexpression of the protein or proteolysis could cause false positives or negatives of the labeling signal. Furthermore, if APEX is cleaved from LETM1, it could diffuse away from the IMM causing false labeling pattern that does not reflect the localization or topology of LETM1.

We performed Western blot analysis of cell lysate from HEK 293T cells expressing different APEX fusion constructs of LETM1 (Figure 5-7). We observed with anti-LETM1 immunostaining that in addition to the band at the expected molecular weight (114 kDa for APEX-LETM1 fusion protein, 83 kDa for LETM1 or LETM1-v5 protein) (Figure 5-7, green and black arrows), each transfected construct also had a smaller protein species that appeared at either 65 kDa, or 90 kDa if APEX were fused near the C-terminus of LETM1 (Figure 5-8, red arrows). Furthermore, there is also a larger protein species appearing at about 5 kDa heavier than the expected molecular weight of each fusion protein (Figure 5-8, blue arrows). These artifacts were also observed in a parallel experiment performed in COS-7 cells (data not shown).

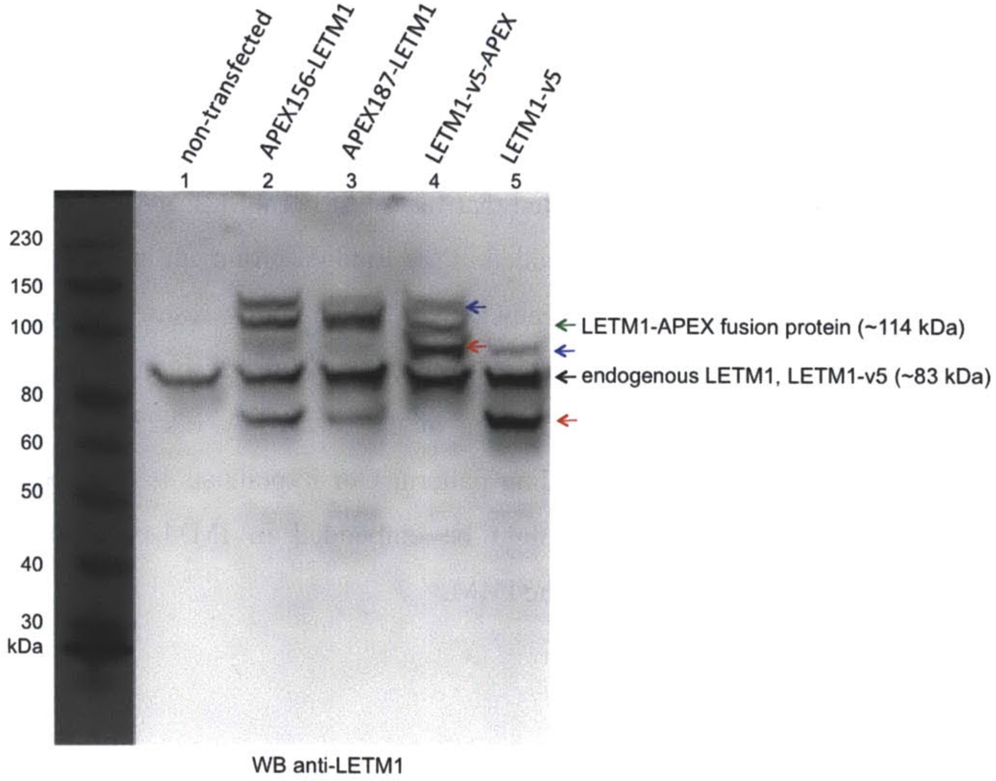

non-transfected
APEX156-LETM1
APEX187-LETM1
LETM1-v5-APEX
LETM1-v5
1
2
3
4
5
230
150
100
80
60
50
40
30
kDa
← LETM1-APEX fusion protein (~114 kDa)
← endogenous LETM1, LETM1-v5 (~83 kDa)
WB anti-LETM1

Figure 5-7. Western blot analysis of APEX-LETM1 fusion proteins expressed in HEK 293T cells. HEK 293T cells were transfected with the indicated APEX-LETM1 fusion construct using lipofectamine. At 18 hours after transfection, cells were lysed and the lysates were separated on a 9% SDS-PAGE gel and transferred onto nitrocellulose membrane for immunostaining with anti-LETM1 antibody (see Experimental methods for details). Black arrow indicates endogenous LETM1 protein (in all lanes) and transfected LETM1-v5 protein (lane 5) at 83 kDa (LETM1 and LETM1-v5 proteins cannot be separated on a 9% SDS-PAGE gel). Green arrow indicates APEX-LETM1 fusion proteins (lanes 2, 3, and 4) at 114 kDa. Red arrows indicate possible protease cleavage products. Blue arrows indicate unknown complex about 5 kDa heavier than expected molecular weights of the expressed proteins.

Other labs have also observed larger and smaller protein species expressed in addition to LETM1 (18,25,26). Schlickum and coworkers detected in HEK 293T cells transfected with LETM1 (83 kDa) two extra bands at 95 kDa and 60 kDa (18). Dimmer and coworkers transfected HeLa cells with LETM1 and saw a lower molecular weight species (25). Authors speculated that the higher molecular weight protein could be a result of posttranslational modification, while the lower protein species may represent protein cleavage product by an unknown protease or degradation product possibly caused by overexpression of LETM1 (18,25). We further conducted Western blot assays of cell lysates harvested at either 24 or 72 hours after lipofectamine transfection of APEX-LETM1 fusion constructs. We hoped that perhaps protein expression level would decrease over time, leading to more controlled expression of LETM1. However, we observed the same cleavage pattern as shown previously (data not shown).

In order to reduce the expression of transfected LETM1 or LETM1 fusion protein, we decided to try a different transfection method. We made lentivirus encoding either LETM1 or LETM1-mCherry fusion construct. We infected COS-7 cells with low amounts of viruses. Thereafter, we lysed the cells and analyzed the cell lysates on Western blot with anti-LETM1 immunostaining. Figure 5-8 shows that cells transfected with LETM1 (lane 2) or LETM1-mCherry (lane 6) using lipofectamine transfection method appeared to have cleavage products (Figure 5-8, red arrows) as well as the unknown larger complexes (blue arrows) as we have seen previously in HEK 293T cells. In the case of LETM1-mCherry, many more cleavage products were observed when the construct was transfected with lipofectamine (lane 6, green arrows). In contrast, cells with viral infected LETM1 (lanes 3-5) or LETM1-mCherry (lanes 7-9) appeared to generate LETM1 or LETM1-mCherry protein at the expected molecular weight of either 83 kDa or 111 kDa, respectively, without the cleavage products. Viral infection of LETM1 gave the cleanest expression of LETM1 using 10 μl of the virus (lane 3). The upper band of the two bands we observed in Figure 5-8, lanes 4 and 5 are likely the cytosolic precursor protein of LETM1, before the mitochondrial targeting sequence of LETM1 was cleaved. We plan to perform subsequent experiments with lentiviral infection of LETM1 fusion constructs to avoid overexpression artifacts.

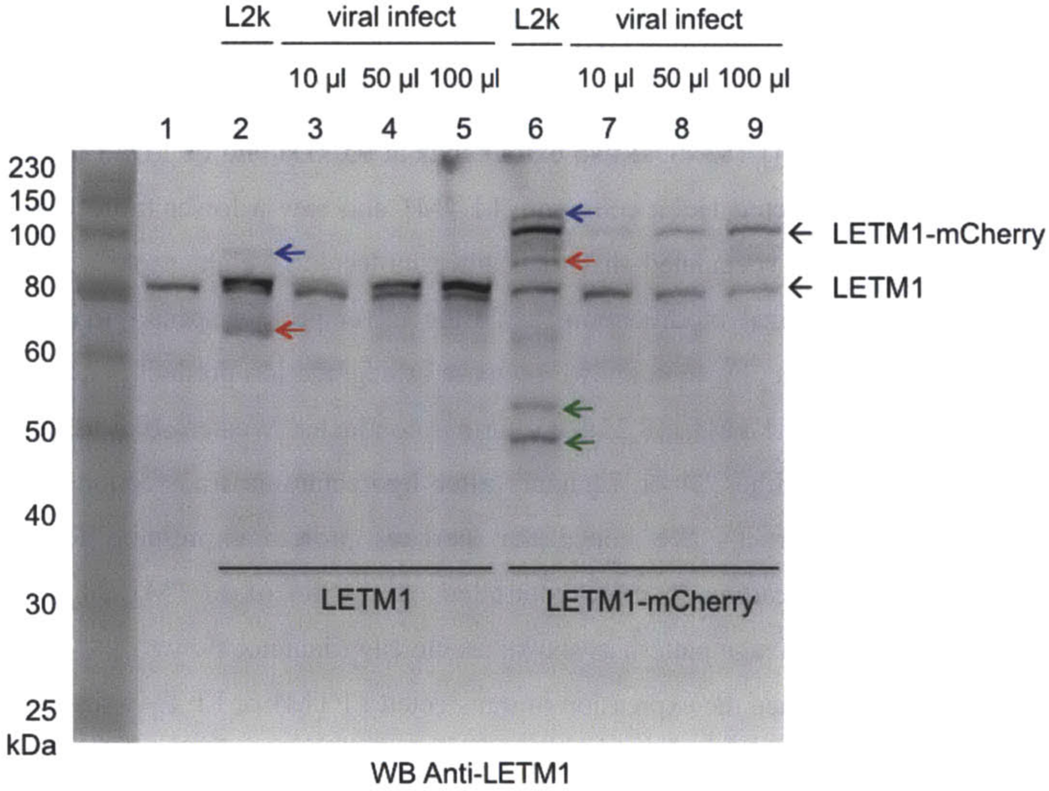

L2k
viral infect
L2k
viral infect
10 μl
50 μl
100 μl
10 μl
50 μl
100 μl
1
2
3
4
5
6
7
8
9
230
150
100
80
60
50
40
30
25
kDa
← LETM1-mCherry
← LETM1
LETM1
LETM1-mCherry
WB Anti-LETM1

Figure 5-8. Western blot analysis of COS-7 cells lipofected (L2k) or viral infected with LETM1 or LETM1-mCherry construct. COS-7 cells were either transfected using lipofectamine (L2k) (lanes 2 and 6) or viral infected using increasing amounts of virus - 10 μl, 50 μl, or 100 μl. Cells were then lysed and analyzed on Western blot. Viral

infected cell lysates appeared to have much cleaner expression of LETM1 (lanes 3-5) or LETM1-mCherry protein (lanes 7-9) without cleavage products or unknown larger protein complex. Lane 1 represents endogenous LETM1 protein in COS-7 cells.

Conclusion and future direction

In this chapter, we described our effort towards developing an enzymatic labeling and imaging-based method for assaying membrane protein topology. The method uses the engineered ascorbate peroxidase enzyme APEX that generates biotin tyramide radicals in the presence of H_2O_2. APEX can be fused to different regions of the membrane protein of interest. Depending on whether APEX is located in a membrane-bound compartment, the biotinylation pattern can appear tight and restricted to the location of APEX or diffusive throughout the cell. The biotinylation patterns can be used to infer the topology of a membrane protein of interest.

We fused APEX to different regions of the mitochondrial inner membrane protein, EF-hand domain-containing protein 1, LETM1. Biotin-tyramide labeling in HEK 293T cells expressing different APEX-LETM1 constructs showed some APEX-LETM1 fusion proteins gave restricted biotinylation pattern, co-localizing with the mitochondrial GFP marker. The tight labeling pattern is indicative of those specific regions of LETM1 being located in the mitochondrial matrix. Three APEX-LETM1 proteins gave diffusive biotinylation patterns, suggesting that those regions could be in the mitochondrial intermembrane space (IMS). Our initial results indicated that the proposed transmembrane region of LETM1, from amino acids 206 to 229, may not traverse the entire phospholipid bilayer of the IMM from matrix to IMS. Instead the proposed transmembrane region may only be embedded in one leaflet of the phospholipid bilayer, so the regions upstream (before amino acid 206) and downstream (after amino acid 229) of the transmembrane region of LETM1 could be both in the matrix.

However, imaging results with APEX-LETM1 fusion proteins would need to be repeated for fear of overexpression artifacts. Each LETM1-APEX fusion protein must be analyzed by Western blot to ensure the expression is free of proteolytic product or other unknown protein species. We performed lentiviral induction and lipofection transfection of APEX-LETM1 constructs and found viral induction expressed the fusion proteins at an optimal level free of proteolytic products. Subsequent experiments will be performed with viral infected plasmids to strictly control for fusion protein expression level.

Even though readout from streptavidin-fluorophore staining and imaging using a conventional microscope would not provide the super-resolution of EM, we believed the different labeling patterns of biotin-tyramide could still be exploited as a useful and simple tool to provide hints to the localization of certain regions of a membrane protein. In conclusion, we hope to develop a robust technique based on biotin-tyramide and APEX labeling that can be routinely used to infer membrane protein topology.

Experimental methods

Genetic constructs

Construct name	Vector	Features and variants	References
matrix-APEX	pcDNA3	mitochondrial matrix targeting sequence: MLATRVFSLVGKRAISTSVCVRAH v5: GKPIPNPLLGLDST (between matrix targeting sequence and APEX)	(13)
IMS-APEX	pcDNA3	Targeted to the IMS via the LACTB$_{(1-68)}$ leader sequence: MYRLLSSVTARAAATAGPAWDGGRRGAHRRPGLPVLGLGWAGGLGLGLGLALGAKLVVGLRGAVPIQS v5: GKPIPNPLLGLDST (between LACTB targeting sequence and APEX)	(13,27)
APEX115-LETM1	pcDNA3	APEX located after amino acid 115 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX156-LETM1	pcDNA3	APEX located after amino acid 156 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX180-LETM1	pcDNA3	APEX located after amino acid 180 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX187-LETM1	pcDNA3	APEX located after amino acid 187 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX201-LETM1	pcDNA3	APEX located after amino acid 201 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX235-LETM1	pcDNA3	APEX located after amino acid 235 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX285-LETM1	pcDNA3	APEX located after amino acid 285 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX300-LETM1	pcDNA3	APEX located after amino acid 300 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX365-LETM1	pcDNA3	APEX located after amino acid 365 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX385-LETM1	pcDNA3	APEX located after amino acid 385 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX444-LETM1	pcDNA3	APEX located after amino acid 444 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX524-LETM1	pcDNA3	APEX located after amino acid 524 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX591-LETM1	pcDNA3	APEX located after amino acid 591 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX609-LETM1	pcDNA3	APEX located after amino acid 609 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
APEX657-LETM1	pcDNA3	APEX located after amino acid 657 of LETM1 v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
LETM1-v5-APEX	pcDNA3	APEX located after v5 tag v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
LETM1-v5	pcDNA3	v5: GKPIPNPLLGLDST (at C-terminus of LETM1)	
LETM1-mCherry	pEGFP-N3	red fluorescent protein mCherry (at C-terminus of LETM1)	(14)

LETM1	pcDNA3		
LETM1	FCPGW		
LETM1-mCherry	FCPGW	red fluorescent protein mCherry (at C-terminus of LETM1)	

Labeling and imaging of mitochondrial matrix and IMS proteins (Figure 5-3, Figure 5-6)

HEK 293T or COS-7 cells were plated on 5 mm × 5 mm glass coverslips placed within wells of a 48-well cell culture plate (0.95 cm^2 per well) 12–16 hours prior to transfection. HEK 293T cells were plated on glass coverslips pre-coated with 50 μg/mL fibronectin (Millipore) to increase adherence. In general, plated cells were transfected with 100 ng indicated APEX construct at 50-70% confluency using lipofectamine following manufacture's protocol. At 18 hours after transfection, cells were incubated in cell culture media (minimum essential media (MEM) + 10% FBS) containing 500 μM biotin-tyramide for 1 hour at 37 °C. Afterwards, 1 mM H_2O_2 was added directly to the cell media and the reaction was allowed to proceed for 1 minute at room temperature. Cells were then washed 3x with Dulbecco's phosphate buffered saline (DPBS) and fixed with 3.7% paraformaldehyde in DPBS for 15 minutes at room temperature. Fixed cells were washed 3x with DPBS followed by protein precipitation with cold methanol for 6 minutes at -20 °C. Fixed cells were washed with DPBS 3x and blocked with 0.5% casein in DPBS for 1 hour at room temperature.

To visualize biotinylated proteins and the expression of APEX or APEX fusion proteins, fixed cells were immunostained by first incubating with primary antibody, 4 μg/mL mouse monoclonal anti-v5 antibody (Sigma-Aldrich) in 1% (w/v) BSA in DPBS for 1 hour at room temperature. Cells were further washed with DPBS and incubated with secondary antibody, 4 μg/mL goat anti-mouse IgG antibody conjugated to Alexa Fluor 568 (Life Technologies) and also neutravidin-Alexa Fluor 647 conjugate (nAv-AF 647) in 1% (w/v) BSA in DPBS for 1 hour at room temperature. Cells were then washed 3x and imaged on a confocal microscope.

Western blot with anti-LETM1 antibody to detect LETM1 and APEX-LETM1 fusion proteins (Figure 5-5, Figure 5-7, Figure 5-8)

HEK 293T or COS-7 cells were plated in wells of a 6-well cell culture plate (9.6 cm^2 /well) and transfected with 200 ng of the indicated APEX-LETM1 fusion construct (Figure 5-5, 5-7), LETM1-v5 (Figure 5-7), LETM1 (Figure 5-8) or LETM1-mCherry (Figure 5-8) using lipofectamine, following manufacture's protocol.

For viral infection of LETM1 and LETM1-mCherry constructs (Figure 5-8), 10 µl, 50 µl, or 100 µl of the LETM1 or LETM1-mCherry virus (see below for protocol of lentivirus generation) was added to 1 well (in 6-well plate) of HEK 293T cells in 1 ml of cell culture media.

At 18 hours after lipofectamine transfection or 48 hours after viral infection, cells were washed 3x with DPBS and collected in DPBS by gentle pipetting followed by centrifugation at 200 rpm for 5 minutes at room temperature. Cell pellets were lysed with 100 µl RIPA lysis buffer (50 mM Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100) supplemented with 1x protease cocktail (Sigma Aldrich), 1 mM PMSF. Lysates (5 µl out of 100 µl lysate per well of a 6-well plate) were mixed with protein loading buffer and separated on 9% SDS-PAGE gel. Proteins were transferred to nitrocellulose membrane using transfer buffer (0.25 M Tris, 1.8 M glycine) and transferred at constant 400 mM for 2 hours. The complete transfer of the colored protein ladder was used as an indication of successful transfer. Thereafter, the nitrocellulose membrane was blocked with 3% (w/v) BSA in TBS-T (0.1% Tween-20 in Tris-buffered saline) buffer at room temperature for 1 hour. To immunostain with mouse monoclonal anti-LETM1 antibody (Abcam, Cat. ab55434), nitrocellulose membrane was immersed in a solution of 1 ng/mL anti-LETM1 antibody in TBS-T at room temperature for 60 minutes before rinsing with TBS-T 4× 5 minutes. The membrane was then incubated with goat anti-mouse horseradish peroxidase conjugate (BioRad, 1:5000 dilution) in TBS-T at room temperature for 60 minutes. Membrane was wash 4x 5 minutes with TBS-T, then developed with the SuperSignal West Pico reagent (Thermo Scientific), and imaged using an Alpha Innotech digital camera.

Lentivirus generation

HEK 293T cells were plated in wells of a 6-well cell culture plate (9.6 cm^2 /well) and transfected at 50 – 70% confluency with either 1.25 µg LETM1 or LETM1-mCherry in

FCPGW plasmid (lentiviral transfer vector), 0.125 μg pVSVG (lentiviral packaging vector), and 1.125 μg Δ8.9 (lentiviral packaging vector) using lipofectamine (Invitrogen), following manufacture's protocol. Transfection MEM solution was changed to 1 ml of fresh cell culture media (MEM+10% FBS) at 4 hours after transfection. When the three plasmids are transfected into HEK 293T cells, viral particles accumulate in the supernatant. Cells were allowed to grow for 72 hours during which time lentivirus was being made. To harvest the virus, cell culture media was collected and filtered through a 0.45 μm filter and stored at 4 °C. Viruses used (Figure 5-8) were no more than 5 days old.

References

1. Baker M. Making membrane proteins for structures: a trillion tiny tweaks. Nat Meth. 2010 Jun;7(6):429–34.

2. Park N, Um SH, Funabashi H, Xu J, Luo D. A cell-free protein-producing gel. Nat Mater. 2009 May;8(5):432–7.

3. Verardi R, Traaseth NJ, Masterson LR, Vostrikov VV, Veglia G. Isotope Labeling for Solution and Solid-State NMR Spectroscopy of Membrane Proteins. In: Atreya HS, editor. Isotope labeling in Biomolecular NMR [Internet]. Springer Netherlands; 2012 [cited 2013 Apr 22]. p. 35–62. Available from: http://link.springer.com/chapter/10.1007/978-94-007-4954-2_3

4. Ding X, Zhao X, Watts A. G-protein-coupled receptor structure, ligand binding and activation as studied by solid-state NMR spectroscopy. Biochemical Journal. 2013 Mar 15;450(3):443–57.

5. Shimizu T, Mitsuke H, Noto K, Arai M. Internal Gene Duplication in the Evolution of Prokaryotic Transmembrane Proteins. Journal of Molecular Biology. 2004 May 21;339(1):1–15.

6. Rapp M, Granseth E, Seppälä S, Von Heijne G. Identification and evolution of dual-topology membrane proteins. Nat Struct Mol Biol. 2006 Feb;13(2):112–6.

7. Geest M van, Lolkema JS. Membrane Topology and Insertion of Membrane Proteins: Search for Topogenic Signals. Microbiol. Mol. Biol. Rev. 2000 Mar 1;64(1):13–33.

8. Akiyama Y, Ito K. Folding and assembly of bacterial alkaline phosphatase in vitro and in vivo. J. Biol. Chem. 1993 Apr 15;268(11):8146–50.

9. Silhavy TJ, Shuman HA, Beckwith J, Schwartz M. Use of gene fusions to study outer membrane protein localization in Escherichia coli. Proc. Natl. Acad. Sci. U.S.A. 1977 Dec;74(12):5411–5.

10. Yost-Daljev MK, Cornelissen CN. Determination of Surface-Exposed, Functional Domains of Gonococcal Transferrin-Binding Protein A. Infect Immun. 2004 Mar;72(3):1775–85.

11. Tamai S, Iida H, Yokota S, Sayano T, Kiguchiya S, Ishihara N, et al. Characterization of the mitochondrial protein LETM1, which maintains the mitochondrial tubular shapes and interacts with the AAA-ATPase BCS1L. J. Cell. Sci. 2008 Aug 1;121(Pt 15):2588–600.

12. Martell JD, Deerinck TJ, Sancak Y, Poulos TL, Mootha VK, Sosinsky GE, et al. Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. Nat Biotech. 2012 Nov;30(11):1143–8.

13. Rhee H-W, Zou P, Udeshi ND, Martell JD, Mootha VK, Carr SA, et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science. 2013 Mar 15;339(6125):1328–31.

14. Jiang D, Zhao L, Clapham DE. Genome-wide RNAi screen identifies Letm1 as a mitochondrial Ca2+/H+ antiporter. Science. 2009 Oct 2;326(5949):144–7.

15. Frazier AE, Taylor RD, Mick DU, Warscheid B, Stoepel N, Meyer HE, et al. Mdm38 interacts with ribosomes and is a component of the mitochondrial protein export machinery. J Cell Biol. 2006 Feb 13;172(4):553–64.

16. Dimmer KS, Fritz S, Fuchs F, Messerschmitt M, Weinbach N, Neupert W, et al. Genetic basis of mitochondrial function and morphology in Saccharomyces cerevisiae. Mol. Biol. Cell. 2002 Mar;13(3):847–53.

17. Endele S, Fuhry M, Pak S-J, Zabel BU, Winterpacht A. LETM1, A Novel Gene Encoding a Putative EF-Hand Ca2+-Binding Protein, Flanks the Wolf–Hirschhorn Syndrome (WHS) Critical Region and Is Deleted in Most WHS Patients. Genomics. 1999 Sep 1;60(2):218–25.

18. Schlickum S, Moghekar A, Simpson JC, Steglich C, O'Brien RJ, Winterpacht A, et al. LETM1, a gene deleted in Wolf–Hirschhorn syndrome, encodes an evolutionarily conserved mitochondrial protein. Genomics. 2004 Feb;83(2):254–61.

19. Nowikovsky K, Froschauer EM, Zsurka G, Samaj J, Reipert S, Kolisek M, et al. The LETM1/YOL027 Gene Family Encodes a Factor of the Mitochondrial K+ Homeostasis with a Potential Role in the Wolf-Hirschhorn Syndrome. J. Biol. Chem. 2004 Jul 16;279(29):30307–15.

20. Hasegawa A, Van der Bliek AM. Inverse correlation between expression of the Wolfs Hirschhorn candidate gene Letm1 and mitochondrial volume in C. elegans and in mammalian cells. Hum. Mol. Genet. 2007 Sep 1;16(17):2061–71.

21. Tamai S, Iida H, Yokota S, Sayano T, Kiguchiya S, Ishihara N, et al. Characterization of the mitochondrial protein LETM1, which maintains the mitochondrial tubular shapes and interacts with the AAA-ATPase BCS1L. J Cell Sci. 2008 Aug 1;121(15):2588–600.

22. Piao L, Li Y, Kim SJ, Sohn K-C, Yang K-J, Park KA, et al. Regulation of OPA1-mediated mitochondrial fusion by leucine zipper/EF-hand-containing transmembrane protein-1 plays a role in apoptosis. Cell. Signal. 2009 May;21(5):767–77.

23. Piao L, Li Y, Kim SJ, Byun HS, Huang SM, Hwang S-K, et al. Association of LETM1 and MRPL36 contributes to the regulation of mitochondrial ATP production and necrotic cell death. Cancer Res. 2009 Apr 15;69(8):3397 404.

24. Dimmer KS, Navoni F, Casarin A, Trevisson E, Endele S, Winterpacht A, et al. LETM1, deleted in Wolf–Hirschhorn syndrome is required for normal mitochondrial morphology and cellular viability. Hum. Mol. Genet. 2008 Jan 15;17(2):201–14.

25. Dimmer KS, Navoni F, Casarin A, Trevisson E, Endele S, Winterpacht A, et al. LETM1, deleted in Wolf–Hirschhorn syndrome is required for normal mitochondrial morphology and cellular viability. Hum. Mol. Genet. 2008 Jan 15;17(2):201–14.

26. Tamai S, Iida H, Yokota S, Sayano T, Kiguchiya S, Ishihara N, et al. Characterization of the mitochondrial protein LETM1, which maintains the mitochondrial tubular shapes and interacts with the AAA-ATPase BCS1L. J Cell Sci. 2008 Aug 1;121(15):2588–600.

27. Brown TA, Tkachuk AN, Shtengel G, Kopek BG, Bogenhagen DF, Hess HF, et al. Superresolution fluorescence imaging of mitochondrial nucleoids reveals their spatial range, limits, and membrane interaction. Mol. Cell. Biol. 2011 Dec;31(24):4994–5010.

Jennifer (Zhengzheng) Yao

Tel: (626)758-2370
Email: jzyao@mit.edu

70 Pacific Street, Apt. 548C
Cambridge, MA 02139

EDUCATION

Ph.D. in Chemistry
Massachusetts Institute of Technology (MIT) 9/2008-expected 6/2013
B.S. in Chemical Biology
University of California at Berkeley (UC Berkeley) 9/2004-5/2008

RESEARCH EXPERIENCE

Department of Chemistry, MIT 9/2008-expected 6/2013
Advisor: **Alice Y. Ting**
Developed of new methods for targeting diverse fluorophores to specific proteins inside live cells. Engineered benzophenone photocrosslinker ligase to study protein-protein interactions.

Department of Chemistry, UC Berkeley 8/2006-5/2008
Advisor: **Dirk Trauner**
Designed, synthesized, and purified small molecule photoswitchable agonists of the nicotinic acetylcholine receptor and two-photon sensitive chromophores for light-activated ion channels.

Department of Chemical Engineering, UC Berkeley 5/2005-8/2006
Advisor: **David Schaffer**
Studied of regulation of proneural Mash1 and Ngn1 genes in the Wnt signaling pathway.

Department of Biological Science, California State University, LA 6/2003-8/2003
Center for Environmental Analysis-Centers of Research Excellence in Science and Technology (CEA-CREST)
Advisor: **John Gamon**
Applied of optical tools to study plant ecophysiology and ecosystem gas exchange.

PUBLICATIONS

Yao, J.Z., Uttamapinant, C., Poloukhtine, A., Baskin, J.M., Codelli, J.A., Sletten, E.M., Bertozzi, C.R., Popik, V.V., Ting, A.Y. "Fluorophore targeting to cellular proteins via enzyme-mediated azide ligation and strain-promoted cycloaddition", *J Am Chem Soc.* 2012 (134), 3720-8.

Tochitsky, I., Banghart, M.R., Mourot, A., Yao, J.Z., Gaub, B., Kramer, R.H., Trauner, D. "Optochemical control of genetically engineered neuronal nicotinic acetylcholine receptors", *Nat Chem.* 2012 (4), 105-11.

Banghart, M.R., Mourot, A., Fortin, D.L., Yao, J.Z., Kramer, R.H., Trauner, D. "Photochromic blockers of voltage-gated potassium channels", *Angew Chem Int Ed Engl.* 2009 (48), 9097-101.

LABORATORY SKILLS

- Biology-related lab skills: mammalian cell and neuron culture, protein purification from mammalian and bacterial cells, chromatography (normal and reverse phase, ion exchange, affinity binding), molecular cloning, gel-electrophoresis, immunoprecipitation, Western blot, microscopy, LC-MS
- Organic synthesis-related lab skills: organic synthesis and purification of small molecules on large (>1 g) and small scales (<0.5 mg), distillation, extraction, recrystallization, SiO_2 and C-18 column chromatography, LC-MS, NMR

TEACHING AND MENTORING EXPERIENCE

Mentor, Women In Chemistry, MIT 9/2009-Present

- Guided first-year female graduate students in Chemistry in course selections and choosing Ph.D. research advisors.
- Introduced new students to campus resources and met with mentees monthly to offer effective mentorship.

Teaching Assistant, Biological Chemistry II (5.08), MIT 2/2009-5/2009

- Taught two recitation classes with 30 students each throughout the semester.
- Developed effective recitation material working in conjunction with lecture material.
- Selected recent seminal research publications to discuss in class.
- Helped professors design problem sets and exams.

Teaching Assistant, Principles of Chemical Science (5.111), MIT 9/2008-12/2008

- HHMI teaching fellow
- Developed and implemented innovative general chemistry teaching material that demonstrates the chemical principles behind inspiring applications in biology and medicine.
- Taught two recitation classes with 30 students each throughout the semester.

Laboratory Instructor, General Chemistry Laboratory, UC Berkeley 1/2008-5/2008

- Conducted two general chemistry laboratory classes with 35 students each throughout the semester.
- Guided students in developing safe and proper experimental skills and writing lab reports.

Laboratory Instructor, Organic Chemistry Laboratory, UC Berkeley 1/2007-5/2007

- Conducted two organic chemistry laboratory classes with 35 students each throughout the semester.
- Taught organic chemistry lab skills including handling moisture and shock-sensitive chemicals, TLC, extraction, column purification, recrystallization, and distillation.

LEADERSHIP AND COMMUNITY SERVICE

MIT Chemistry Outreach 6/2009-Present

- Conducted engaging demonstrations illustrating a broad range of chemical principles to classes of thirty to ninety students at local schools ranging from elementary to high schools.
- Illustrated concepts relating to polymer formation, natural products, temperature, pressure, and light through demonstrations and PowerPoint slides.

Chemistry Education and Outreach Career Panel Organizer 1/2013

- Organized and networked with college lecturers, high school teachers, and leaders in education-related nonprofit organizations to establish the first departmental Education and Outreach Career Panel.

HONORS AND AWARDS

MIT Presidential Fellowship	8/2008-5/2009
HHMI Teaching Assistant Fellowship	8/2008-12/2008
Berkeley Molecular Cell Biology Undergraduate Symposium Honorary Poster	3/2008
Melvin J. Heger-Horst Scholarship (College of Chemistry, UC Berkeley)	2007
Merck Undergraduate Research Scholarship (David Schaffer Lab, UC Berkeley)	2006