Investigating Chemical Modifications in a Complex Proteome:

Author: Lisa Ann Crawford

Investigating Chemical Modifications in a Complex Proteome

Lisa Ann Crawford

A dissertation

submitted to the Faculty of

the department of Chemistry

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Boston College
Morrissey College of Arts and Sciences
Graduate School

August 2017

Investigating Chemical Modifications in a Complex Proteome

Lisa Ann Crawford

Advisor: Eranthie Weerapana, Ph.D.

Abstract

Proteins are composed of the 20 naturally occurring amino acids and are further modified by a variety of post-translational modifications (PTMS). Naturally occurring amino acids are diverse in structure and function. Catalytic amino acids, or nucleophilic amino acids, are of particular interest because of their contribution to chemical transformations in the cell.

Synthetic covalent modification is a means to further functionalize or diversify proteins. These modifications, or enhancements, allow for improved understanding of protein structure, function and activity. For instance, isotope labeling of amino acid side chains in NMR studies enable investigators to study protein dynamics upon substrate or ligand binding. Fluorescence labeling is particularly useful to investigate protein cellular localization. Covalent modification is a useful tool to investigate the relative level of activity for protein known to be regulated by PTMs. An important feature of covalent modification reactions is site specificity, as this dictates the location, number of modifications, and protein targets.

Tyrosine is of particular interest because it is both nucleophilic and aromatic. These characteristics contribute to the existence of tyrosine residues in both the protein surface and hydrophobic cores. Tyrosine is incorporated into proteins at a relatively low frequency. Unlike lysine, which is ubiquitous on protein surfaces, the low number of potential sites for general tyrosine modifications makes it an attractive site for surface bioconjugation

modifications. A low number of surface modifications is less likely to perturb native protein function. Bioconjugation reactions give access to functionalizing the surface of proteins with moieties such as fluorophores, PEG, peptides, or drugs. Tyrosine is an attractive target for modifications because it is found in the active sites of a variety of enzymes such as sialidases, glutathione-S transferases, corticosteroid 11-beta-dehydrogensase, DNA topoisomerase, and ferredoxin-NADP+ reductase. Provided here is a survey of the known non-selective and selective synthetic chemical modification reactions for tyrosine.

To investigate nucleophilic amino acids, Activity Based Protein Profiling (ABPP) may be implemented to investigate the role of these residues. ABPP utilizes small molecule covalent probes as a tool to selectively target enzymes in their active state. To investigate a protein of interest (POI) (or class of proteins) by ABPP, it is necessary to use a small molecule covalent probe that selectively reacts with the POI over other proteins within the proteome. Due to this requirement, it is necessary to expand the current ABPP probe toolbox to increase the coverage of what proteins in the proteome may be studied. Inspired by findings in the literature, our lab sought to explore the utility of various aryl halides for implementation in ABPP probes to overcome this limitation.

This study revealed dichlorotriazine as a biologically relevant and reactive electrophile. A focus was placed on a dichlorotriazine containing probe library (LAS1-LAS20). LAS17 was discovered to be a potent and selective inhibitor of human glutathione S-transferase pi (GSTP1). Further studies revealed GSTP1 as a novel therapeutic target for the treatment of triple negative breast cancer. Other studies revealed several members of the dichlorotriazine library were found to covalently modify purified recombinant human

aldolase A (ALDOA) in the presence of a complex cellular background. Additionally, LAS9 was identified as an inhibitor of ALDOA retro aldol condensation activity *in vitro*.

Lastly, the final chapter highlights two collaborations in which tandem mass spectrometry experiments aid in the characterization of experimental data. In the first collaboration, a quantitative cysteine reactivity profiling method was used to characterize the selectivity of a cysteine reactive covalent NRF2-inducing small molecule, MIND4-17. In the second collaboration, analysis of tryptic mass spectrometry data enabled high resolution characterization of peptide sequencing for superfolder green fluorescent protein (sfGFP) expressed from observed internal nonsense suppression. Identification of the misincorporated amino acid facilitated the elucidation of the cross-talk mechanism.

Acknowledgements

I would like to begin by thanking my advisor, Eranthie Weerapana. Thank you for your commitment to being an educator. I appreciate the time, effort and energy you have invested in me over the past 6 years. Thank you for giving me the space and tools to allow myself to develop into scientist I am today. I am grateful for the understanding and compassion you had for me during and after my pregnancy with Colin, always being an advocate for our wellbeing. Thank you for believing in me even during times when I did not have confidence in myself.

I would like to thank the members of my committee, Jianmin Gao and Abhishek Chatterjee for being supportive mentors during my graduate school studies and for serving on my committee. Also, I'd like to thank Abhishek and James Italia for working allowing me to work as a collaborator on the genetic code expansion project. I would also like to thank my other Collaborators, Dan Nomura and Sharon Louie for the collaboration on the GSTP1 Triple Negative Breast Cancer project. I would also like to thank Aleksey Kazantsev and co-workers for the opportunity to work on the KEAP1 project.

I would like to thank all Weerapana lab members past and present: Nick, Alex, Julie, Yani, Shalise, Tyler, Kyle, Dan, Haley, Ranjan, Masahiro, Aaron, Eleni, Julia, and Jenny. I can truthfully say that every day working with you and learning together with you has been an absolute pleasure. Nick, while it is hard to be compared to you, thank you for setting the bar high and being an exemplary scientist. Alex, thank you for your kind conversation and support, while we might not always be on the same page, "On purpose?!?," I am thankful to call you a friend. Julie, thank you for your friendship and for

being the big sister I never wanted. Thank you for hanging around Merkert waiting for me to finish teaching Gen Chem lab to drive me to The Draft for our regular Monday night wings. Yani, thank you for being a thoughtful and compassionate friend. Thank you for always having a listening ear. Tyler, thank you for being the weird one so no one else had to, but most importantly I am thankful for our friendship. Thank you for being the little brother I always wanted. Also, thank you for teaching me the true meaning of popular pop culture sayings. Dan, thank you for your commitment to keeping things afloat, lunch time team trivia, and for fanaticizing about world travel with me. Aaron, thank you for your company at odd hours in the lab. Eleni, Julia, and Jenny, while our time together was short I am thankful for the good times we had together in the Ladies Lounge including but not limited to Fro Yo Friday and lunch at Sweetgreen. Jenny, thank you for your edits!

I would like to thank my undergraduate students, Hannah O'Day and Paige Carleen. Thank you for giving me the opportunity to mentor you and help you develop your skills as scientists. I am proud of the women you are and the great places you are both going in life. I am thankful to call you both friends.

Thank you to the entire G1 2011 Chemistry class. Having such an inclusive and loving class made this experience truly memorable. I especially want to thank Amanda and Meredith for their love and support over the years. A special thanks to my dear friend, Hilan, for years of mentoring, support, and advice. I want to thank all of the other friends I made at BC of which there are too many to name.

I want to thank all of the people who logistically made this accomplishment possible for me. Than you, Dale, Lynne, Lynn, Ian, John, TJ, Howard, Bill, and Jen for all that you have done for me over the years.

I want to especially thank my fellow Left Handed Chemists Who Love food, Kevin and Carrie. Thank you for the lengthy dinners with price tags outside of a graduate student's budget. Thank you for your listening ears and comic relief.

Shalise, The "Shal" of "ShaLisa", my pseudo bridesmaid, Colin's second mom, and my best friend, thank you for all of your emotional support and encouragement over the past 6 years. Thank you for always knowing when I need a hug and having the right words to help me pick up the pieces. I am so thankful we joined Eranthie's lab at the same time, and our story will forever remind me that things happen for a reason.

I wouldn't have made it through this program without the help of the amazing women in my Tuesday night Bible study, The Foxy Ladies. Thank you, Carolyn, Shayla, Sarah, Christine, Melissa, Alice, Danielle, Karen, Evis, Gale, Amanda, and Sam, without your encouragement and advice over the years I would not have been able to finish this program. Thank you for helping me stay focused on what matters most.

Carolyn, I can't thank you enough for the years of godly advice and support you have given me. You have been my biggest cheerleader over these past 6 years. Thank you for constantly reminding me of my purpose and value. I am grateful to have met you on visiting weekend and for the wonderful memories we have shared together since.

Bailey, Marissa, and Sarah, thank you for supporting me during this journey! Thank you for always making me laugh and for being a listening ear. I am thankful for our decades of friendship and all of the memories we have shared. Not many people can say they are still best friends with people they know from high school, but I am one of the lucky few.

I'd like to thank Megan Fioto for our lifelong friendship. While we don't get to see each other as often as we'd like, I am thankful that I can pick up the phone and know you'll

answer. I am so thankful to have a constant companion in you and for the many wonderful experiences we have shared together.

I'd like to thank Dr. Virginia Parker for being the most wonderful lab partner throughout college. Thank you for being a wonderful, loyal friend. I am so proud of all you have accomplished!

I am thankful for the support from my family, Dad, Mom, Missy, Sean, Lisa, Rick, Rachel, Laura, Tim, and Patrick, and the encouragement they have relentlessly given me throughout my academic pursuits. Dad, thank you for being a role model and for always encouraging me to do my best and reach my potential. Mom and Missy, thank you for always being there when I need someone to talk to.

Colin, you are my greatest accomplishment from this season of my life. I hope to encourage you to reach for the stars one day!

To my husband, Brett, thank you for the endless encouragement and unconditional love throughout this process. There is no one else I'd rather have by myside throughout this journey. Thank you for picking up where I fell short because of late nights in the lab. Thank you for your patience. Words cannot express how thankful I am for you and all you have done for me!

Table of Contents

List of Figures

Chapter 1.

Figure 1-13. Sample preparation and mass spectrometry analysis workflow for protein target identification of LAS17.

Figure 1-14. Evaluating LAS17 reactivity towards purified recombinant GSTP1. (A) Workflow to evaluate LAS17 reactivity towards purified, recombinant GSTP1. (B) Recombinant, purified GSTP1 was treated with LAS17 and subjected to CuAAC, SDS-PAGE and in-gel fluorescence and Coomassie Blue staining.

Figure 1-15. (A) *in vitro* GSTP1 activity in the presence of increasing concentrations of LAS17. (B) *in vitro* concentration and time-dependent inhibition of GSTP1.

Figure 1-16. Intact-protein mass spectrometry of DMSO or LAS17 treated purified GSTP1. (A) Intact-protein mass spectrometry workflow. De-convoluted spectra of (B) DMSO and (C) LAS17 treated GSTP1.

Figure 1-17. Workflow to determine the site of labeling of LAS17 on purified WT GSTP1.

Figure 1-18. Fragmentation (MS/MS) spectra of the YVSLIY*NYEAGKDDYVK peptide of GSTP1 confirming that the LAS17 modification (+322.12 Da) is present on Y108, which is indicated with *.

Figure 1-19. (A) In-gel fluorescence analysis of purified Y108F mutant compared to WT GSTP1, confirms Y108 as the site of labeling (B) Proximity of Y108 to the GSH binding site of GSTP1 (PDB: 6GSS).

Figure 1-20. LAS17 labeling screening of (A) lysine and (B) cysteine mutants.

Figure 1-21. The Y108F mutant is resistant to LAS17-mediated inhibition compared to the wild-type GSTP1.

Figure 1-22. (A) Chemoproteomic profiling of a panel non-TNBC and TNBC cell lines with RB7. (B) GSTP1 reactivity in TNBC cells from each individual cell line and combined average spectral counts from non-TNBC and TNBC. (C) GSTP1 expression was measured across cell lines by Western blot. GSTP1 expression was normalized to actin loading control.

Figure 1-23. Effect of genetic inactivation of GSTP1 *in vitro* and *in vivo*. (A) Two independent shGSTP1 231MFP knockdown cell lines confirmed by qPCR and Western blotting. (B) GSTP1 231 MFP knockdown proliferation and survival studies. Assays were assessed by Hoechst stain. (C) Xenograft mouse models show impaired 231MFP tumor growth in shGSTP1-1 mice compared to shControl mice.

Figure 1-24. Effect of chemical inactivation of GSTP1 by LAS17 *in vitro* and *in vivo*. (A) LAS17 has an IC50 of 0.5 μM for the inhibition of GSTP1. (B) LAS17

fractions from WT MEFs treated with 0.5 μM MIND4- 17 show time-dependent accumulation.

Figure 3-5. Quantitative MS workflow. Light samples were pretreated with MIND4-17 (10 μM) and heavy samples were pretreated with DMSO. Both light and heavy samples were then treated with IA probe (100 μM). Figure adapted from Qian Y, et al. 2013.

Figure 3-6. Structure of isotopically labeled Azo-L and Azo-H tags. Incorporation of the heavy valine residue provides a mass difference of 6 Da between Azo-L and Azo-H. Figure adapted from Qian Y, et al. 2013.

Figure 3-7. Plot of cysteine reactivity ratios identified in cysteine proteome profiling after MIND4-17 pretreatment. The four proteins with ratios above 2 (indicated in red) are PLA2G4A Cytosolic phospholipase A2 (PLA2G4A), PRKDC DNA-dependent protein kinase catalytic subunit (PRKDC), PSMB5 Proteasome subunit beta type-5 (PSMB5), and RPS5 40S ribosomal protein S5 (RPS5). Highly reactive GAPDH active site cysteine (indicated in yellow) has a ratio of 0.86.

Figure 3-8. Strategy to create an aaRS-tRNA pair that can drive genetic code expansion in both eukaryotes and *E. coli* by liberating the endogenous *E. coli* TrpRS-$tRNA^{Trp}$.

Figure 3-9. Fluorescence microscopy images of HEK293T cells co-expressing EGFP-39-TAG reporter and $tRNAEc^{Trp}_{CUA}$ with and without co-expression of the cognate EcTrpRS. Robust EGFP-39-TAG expression is observed only when EcTrpRS is co-expressed. This confirms no cross-reactivity of $tRNAEc^{Trp}_{CUA}$.

Figure 3-10. ATMW1 exhibits comparable growth rate to progenitor EcNR1G strain containing ScTrpRS-$tRNA^{ScTrp}$ complementation plasmid.

Figure 3-11. Overlay of $tRNA^{EcTrp}$ (blue, PDB: 4YCP) and $tRNA^{EcGln}$ (red, PDB: 1QTQ) show high structural similarity despite low sequence homology.

Figure 3-12. Evaluating the nonsense-suppression efficiencies of $tRNA^{EcTrp}_{CUA}$ and $tRNA^{EcTrp}_{UCA}$ using sfGFP-151-TAG or sfGFP-151-TGA as reporters.

Figure 3-13. Deconvoluted mass of purified whole sfGFP co-expressed in ATMW1 with $tRNA^{EcTrp}_{CUA}$ in the absence of EcTrpRS.

Figure 3-14. Workflow if tryptic digest tandem protein mass spectrometry analysis of sfGFP-151-TAG.

Figure 3-15. (A) Extracted masses from LC-MS/MS chromatogram for the $[M+2]^{2+}$ tryptic peptides for the three potentially incorporated natural amino acids in sfGFP-151-TAG. (B) Isotopic envelop of the extracted average MS1 of the peak

List of Tables

List of Schemes

List of Abbreviations

Standard 3-letter and 1-letter codes are used for the 20 natural amino acids.

4-FSB	4-(fluorosulfonyl)benzoic acid
ABP	activity based protein profiling probes
ABPP	activity based protein profiling
ACC	A (CoA) carboxylase
ADC	antibody-drug conjugate
AEBSF	(2-aminoethyl)benzenesulfonyl fluoride
ALDOA	Aldolase A
ALDOB	aldolase B
ALDOC	aldolase C
AMPK	AMP kinase
APEX	ascorbate peroxidase
AREs	antioxidant response elements
BAP	alkaline phosphatase
BCN	bicyclo[6.1.0]nonyne
BDNB	1-bromo-2,4-dinitrobenzene
BITC-SG	S-(N-benzylthio-carbamoyl) glutathione
bp	base pair
BSA	bovine serum albumin
CAN	Cerium (IV) ammonium nitrate
CDNB	1-Chloro-2,4-dinitrobenzene
CHD	1,2-cyclohexanedione
COFRADIC	COmbined FRActional DIagonal Chromatography
CuAAC	copper(I)-catalyzed azide–alkyne cycloaddition
DCM	Dichloromethane
DcpS	mRNA-decapping scavenger enzyme
DFP	diisopropyl fluorophosphate
DFSA	3- fluorosialyl fluoride

DHAP	dihydroxyacetone phosphate
DIEA	Diisopropylethylamine
DMEM	Dulbecco’s modified Eagle’s medium
DMF	Dimethyl fumarate
DMSO	Dimethyl sulfoxide
DPBS	Dulbecco's phosphate-buffered saline
DPIV	dipeptidyl peptidase IV
DTT	Dithiothreitol
EcGlnRS	glutamyl-tRNA synthetase
EcTrpRS	E. coli tryptophan synthetase
EGFP	enhanced green fluorescent protein
EGFR	epidermal growth factor receptor
EM	electron multiplier
EMT	epithelial–mesenchymal transition
ER	oestrogen receptor
ESI	electrospray ionization
F1,6BP	D-fructose 1,6-bisphosphate
FBP aldolase	fructose-bisphosphate aldolase
FSBA	5’-p-fluorosulfonylbenzoyladenosine
GADP	D-glyceraldehyde 3-phosphate
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GC	gas chromatography
GCLC	glutamate-cysteine ligase catalytic
GCLM	glutamate-cysteine ligase regulator
GFP	Green fluorescent protien
GSH	glutathione
GSTP1	glutathione S-transferase Pi
GSTP1	glutathione S-transferase Pi 1
HAT	histone acetyl transferases
HD	Huntington’s disease
HDACs	Histone deacetylases

HEK	Human embryonic kidney
HER2	human epidermal growth factor receptor 2
HFI	Hereditary Fructose Intolerance
HO-1	heme oxygenase (decycling) 1
HRP	horseradish peroxidase
HTT	huntingtin
IA	Iodoacetamide
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JNK	Jun N-terminal kinase
KEAP1	Kelch-like ECH-associated protein 1
KO	Knockout
LC	liquid chromatography
LC-MS/MS	LC tandem mass spectrometry
m/z	mass to charge ratio
MALDI	matrix-assisted laser desorption ionization
MAPK	mitogen-activated protein kinases
MEF	mouse embryonic fibroblasts
MMAF	monomethyl auristatin F
MS	mass spectrometry
MS2	fragmentation spectra
mTOR	mammalian target of rapamycin
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NADH	Nicotinamide adenine dinucleotide
NQO1	NAD(P)H:quinone oxidoreductase 1
NRF2	nuclear factor-erythroid 2 p45-related factor 2
OP	Organophosphorus
PCR	Polymerase chain reaction
PLA2G4A	Cytosolic phospholipase A2
PPARgamma	peroxisome proliferator activated receptor-gamma
PR	progesterone receptor

PRKDC	DNA-dependent protein kinase catalytic subunit
PSMB5	Proteasome subunit beta type-5
PTAD	4-phenyl-3*H*-1,2,4-triazole-3,5(4*H*)-dione
PTM	Post translational modification
RPS5	40S ribosomal protein S5
SCID	severe combined immunodeficiency
SCX	Strong cation exhange
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	phenylsulfonate ester
sfGFP	superfolder GFP
SFP	sulforaphane
SIRT2	Sirtuin 2
SNAr	nucleophilic aromatic substitution
SPPS	solid phase peptide synthesis
SRM	Single-reaction monitoring
TBST	Tris-buffered saline containing Tween 20
TBTA	Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
TCA	tricarboxylic acid cycle
TCEP	tris(2-carboxyethyl)phosphine
THF	tetrahydrofuran
TMV	Tobacco Mosaic Virus
TNBC	triple-negative breast cancer
TOF	Time of flight
TOF	time-of-flight
TPI	Triosephosphate isomerase
TrpRS-tRNATrp	tryptophanyl-tRNA synthetase pair
UPS	ubiquitin–proteasome degradation system
USPSTF	US Preventive Services Task Force
VBAR	Vilmafix Blue AR
WT	wild-type

Chapter 1

The Discovery of a Tyrosine-Reactive Irreversible Inhibitor of GSTP1 and Investigating the Role of GSTP1 in Triple-Negative Breast Cancer

A significant portion of the work described in this chapter has been published in:

Crawford, L.A. and Weerapana, E. 2016. A tyrosine-reactive irreversible inhibitor for glutathione S-transferase Pi (GSTP1). *Molecular BioSystems*. 12, 6, 1768–1771.

Louie, S.M., Grossman, E.A., Crawford, L.A., Ding, L., Camarda, R., Huffman, T.R., Miyamoto, D.K., Goga, A., Weerapana, E. and Nomura, D.K. 2016. GSTP1 Is a Driver of Triple-Negative Breast Cancer Cell Metabolism and Pathogenicity. *Cell Chemical Biology*. 23, 5, 567–578.

The majority of the work investigating the role of GSTP1 in triple-negative breast cancer was performed by members of the Namura lab. Our contribution to the work focused on triple-negative breast cancer was the synthesis of LAS17.

Introduction to Tyrosine

Tyrosine residues are responsible for an array of biological activities and functions. Tyrosine is the 5th least abundant amino acid in proteins, behind methionine, histidine, cysteine, and tryptophan. Tyrosine incorporation constitutes less than 3% of total amino acids incorporated into proteins in *S. aureus*, *E. coli*, and mammalian cells. [1] Despite the low levels of incorporation, tyrosine is not an essential amino acid. In animals, *para*-Tyrosine is formed enzymatically from the hydroxylation of phenylalanine catalyzed by the liver enzyme, phenylalanine hydroxylase, with the cofactor tetrahydrobiopterin and non-heme iron. [2,3] Other organisms, such as plants, fungi, bacteria, and algae, synthesize tyrosine from shikimate acid via the shikimate pathway. [4] Like phenylalanine, tyrosine is one of only 3 aromatic amino acids. The hydrophobic nature of the phenyl ring favors the environment buried in the protein structure, away from solvent exposed surfaces. The aromatic nature of the side chain favors interaction with non-protein ligands and enhances binding through π-stacking and cation-π interactions. [5] The hydroxyl group allows the side chain to exist in two states, protonated and deprotonated. The pK_a of the phenolic hydroxyl of the side chain is about 10.0, however this value can vary significantly depending on the microenvironment of the protein. One study has shown that surface exposed, hydrophobic, and hydrogen bonded tyrosines have pK_a values of 10, 12, and above 13 respectively. [6] Despite occurring less frequently, these characteristics favor tyrosine residue placement on protein surfaces. The unique distribution of tyrosine residues leads to an assortment of functionalities it can participate in, such as catalysis[7], acid base chemistry[8,9] and function as a redox active center [4–7]. Tyrosine is an interesting target for selective site modification reactions because of the unique biological characteristics and chemistries it confers.

Chemical modifications expand upon or employ new functional capabilities in proteins through a stable covalent bond. Chemical modification of tyrosines can be both naturally occurring, by way of post translational modifications[10], or through synthetic modifications. Whether displayed on a protein surface or buried in an active site, the nucleophilic hydroxyl moiety of the tyrosine side chains makes it an ideal candidate for chemical modification. Additionally, the para-substitution of the hydroxyl on the phenyl ring of tyrosine is advantageous because modifications at this position do not lead to regio-isomers. Tyrosine is known to have many different post-translational modifications: phosphorylation, sulfation, nitration, oxidation, halogenation, glycosylation, AMPylation, cross-linking, and electron-conduit/metal binding.[10] These post translational modifications have the ability to alter the function and activity of proteins with high specificity and site selectivity. Alternatively, proteins can be covalently modified through synthetic means. To be able to study protein function through covalent modification the reaction conditions must be agreeable in aqueous solvent and have well-defined reactivity. Subsequent functional studies can increase our knowledge of the biological role of tyrosine and its reactivity.

The relatively low abundance of tyrosine residues makes it an ideal target for nonselective and selective tyrosine modifications. When applying a general nonselective modification strategy all surface exposed tyrosine residues will be modified through a bioconjugation reaction. The low abundance of tyrosine residues means the relatively low number of points of modification sites which is less likely to perturb protein function, than for example lysine. Additionally, tyrosine is a good target for selective covalent modifications because it is found in enzyme active sites. This means tyrosine is amenable

to functional studies that use covalent probes or inhibitors to interrogate the role of a specific protein or protein class. Tyrosine is located in active sites or is a site of regulation in several interesting classes of enzymes such as DNA topoisomerases[11], tyrosine kinases[12], sialidases[13], nucleotide binders, nucleic acid binders. Discovery of more bioorthogonal, tyrosine reactive probes and electrophiles can open up the possibility of targeting previously untargeted protein classes for therapeutics and enhance the current use of bioconjugates.

A bioconjugation reaction is the formation of a covalent bond linking a protein or other biomolecule with another molecule for added functionality. Typically, in protein bioconjugation, reactions modify lysine, cysteine, tyrosine, or tryptophan side chains, protein N-termini, or protein C-termini. However, bioconjugation reactions can also occur between an unnatural amino acid and its complementary bioorthogonal partner. [14] It is possible to achieve multiple degrees of bioconjugation on one substrate by employing several different orthogonal and site selective modification strategies. These reactions can be used to impart unique functionalities that facilitate our investigation and understanding of biological processes. Some of the functionalities made accessible through bioconjugation include fluorescence, PEGylation, antibody-drug conjugates, biotinylation, isotopic labeling, and halogenation. Addition of a fluorescent functionality can facilitate imaging studies that help elucidate cellular localization and distribution and protein dynamics through FRET.[15] Protein PEGylation has been found to increase protein stability. This is useful in the case of increasing protein half-life and solubility when circulating *in vivo*. Antibody-drug conjugates allow drugs to be delivered selectively to targeted cells. Conjugating a drug to an antibody results in more beneficial pharmacokinetics, lower IC_{50}

values, increased stability, and high cell target specificity than the drug alone.[16] Development and optimization of nonselective tyrosine modifications has expanded our knowledge of biology and has improved our therapeutic capacity. Regardless of the advances achieved by bioconjugation, there is a need for selective covalent modification strategies with protein or protein class specificity.

Covalent small molecule probes can target functionally active nucleophilic residues in enzymes. By targeting these sites, we can use the probe as a handle to understand the activity or functional state of the protein in biological processes. This is the motivation behind the field of Activity Based Protein Profiling (ABPP). ABPP uses small molecule ABPP probes to target functionally relevant nucleophilic sites over less reactive, more structurally important sites. We offer a unique perspective to this area given our interests lie in the pursuit of expanding the current chemical ABPP toolkit.[17,18] We achieve this by investigating electrophiles previously unexplored in the context of small molecule probe design. Tyrosine has been relatively unexplored in the fields of medicine (such as tyrosine reactive covalent drugs) and chemical biology through the use of small molecule probes.[19,20] The work in our lab and by others suggests targeting tyrosine residues, the "tyrosinome," as a means to expand current therapeutic strategies to protein targets and diseases previously unexplored. Our lab has developed two different selective small molecule tyrosine inhibitors for glutathione S transferase (GST) enzymes, which utilize two very different electrophiles: sulfonyl fluoride and dichlorotriazine.[21,22] In this account we explore these electrophiles in addition to the current landscape of known selective and non-selective chemical tyrosine modifications.

Non-selective Tyrosine Modifications

Derivatization of biomolecules such as proteins through bioconjugation provides a means to expound on our knowledge of the chemistry of biological systems. These transformations allow researchers to investigate biological interactions, perform imaging studies, and develop of diagnostic applications. Bioconjugation reactions must be both residue or site-selective and must be biocompatible. These criteria are necessary for successful functional studies and utility of the synthetically incorporated functionality. Tyrosine residues are ideal candidates for bioconjugation reactions and ligations because of the nucleophilicity of the hydroxyl on the phenyl moiety, the low abundance of natural incorporation into proteins, the presence of tyrosine on both the surface of proteins and in protein active sites, and the facile introduction of tyrosine residues through mutagenesis. While the utility and need for synthetic tyrosine modifications have been established, there are currently only a few available strategies that have been identified. Here we discuss the development of tyrosine selective bioconjugation reactions as they evolved throughout time.

The earliest example of tyrosine selective modifications was in the form of a diazonium-coupling reaction which yields an *ortho*-substituted tyrosine (Figure 1-1A). Evidence shows histidine residues react with diazonium salts, however due to the pKa (~10) of the hydroxyphenyl ring of tyrosine reactivity with tyrosine is preferential at pH 4-9.[23] There are numerous examples of the application and utility of diazonium-coupling.[24,25] In a recent example, diazonium reagents were also used to install an alkyne onto M13 bacteriophage coat protein P8 Tyr residues 21 and 24. These allowed subsequent functionalized with a fluorophore through Cu(I)-catalyzed azide-alkyne cycloaddition (CuAAC) for imaging

studies.[26] Diazonium reagents were used to install PEGylation modifications on Tyr 139 of Tobacco Mosaic Virus (TMV) coat protein to prepare organic-soluble TMV rods for application to nanoscale particle studies. This work highlighted the importance of using reactive diazonium salts, in which reactivity is enhanced by appropriately placed electron-withdrawing groups.[27] Diazonium hexafluorophosphate reagents were used to conjugate aplaviroc, small-molecule inhibitor of virus entry into host cells, to broadly neutralizing antibodies PG9 and PG16 for a new strategy for HIV therapy. It was found that PG9-aplaviroc conjugates have lower IC_{50} values than the unmodified PG9 antibody.[28] More recently, this method has been used to interrogate the "interactome" of host-pathogen interactions.[29] A photocaged diazonium species on the surface of a virus can be activated upon irradiation of 350 nm light. The free diazonium group is free to react with tyrosine residues on proteins proximal to the virus. Some drawbacks of using the diazonium reagents in solution are the balance of maintaining high diazonium reactivity in aqueous buffer while sustaining reagent stability and the ability to introduce biorthogonal functionalities. Additionally, NO_2-substituted benzene reagents are potentially immunogenic. It would be advantageous to develop a method that would allow very selective modification of tyrosine, eliminating any reactivity with histidine. Additionally, it would be beneficial to minimize the hazards of reagents used and develop a method that could modify the oxygen of the phenol moiety of the tyrosine side chain.

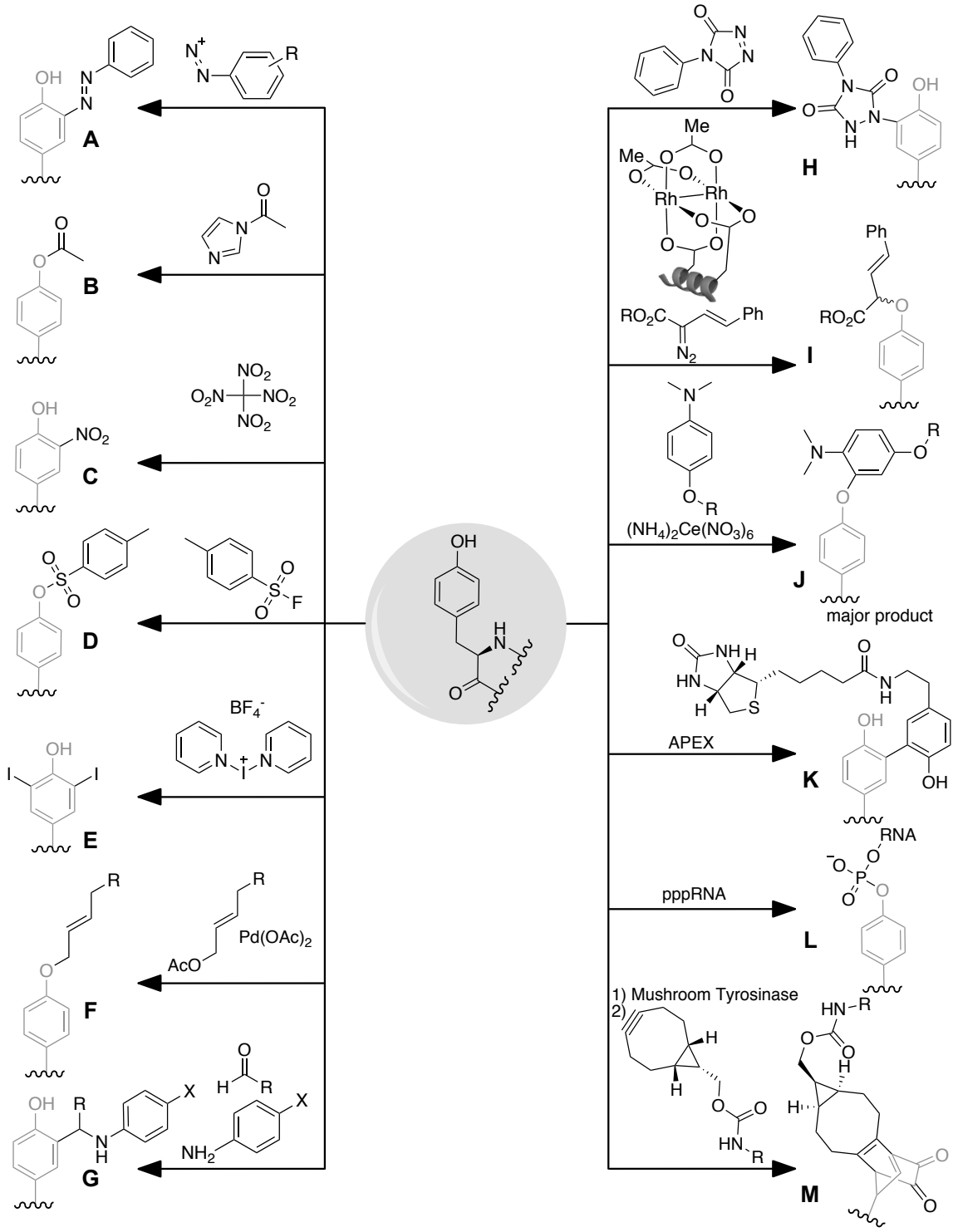


Figure 1-1. Synthetic nonselective tyrosine conjugation reactions. (A) Diazonium-coupling reaction. (B) N-acetylimidizole. (C) Tetranitromethane. (D) p-nitrobenzene sulfonyl fluoride. (E) IPy2BF4. (F) Palladium-catalyzed cross-coupling with π-allyl

complexes. (G) Metal free three component Mannich-type reaction. (H) Diazodicarboxamides. (I) Dirhodium metallopeptide catalysts. (J) Cerium (IV) ammonium nitrate (CAN) oxidation. (K) APEX. (L) Mushroom tyrosinase oxidation, strain-promoted addition.

Other early evidence of non-selective tyrosine modifications was with N-acetylimidizole (Figure 1-1B). Reactivity of a particular tyrosine towards N-acetylimidizole is a function of the degree of hydroxyl surface exposure. Proteins and peptides treated with N-acetylimidizole result in O-acetylation of the hydroxyl group on the phenyl side chain of tyrosine at neutral pH.[30,31] N-acetylimidizole has been used to investigate the role of tyrosine in Na^+-dependent transport in rabbit intestinal brush-boarders[32]. In the absence of Na^+, treatment of the transporters with N-acetylimidizole completely inhibits glucose transport. In the presence of Na^+, treatment of the transporter with N-acetylimidizole did not hinder glucose transport. This approach supports previous work that suggests tyrosine residues play a key role in Na^+ transport of the Na^+/glucose cotransporter. Furthermore, treatment with fluorescein N-acetylimidizole analogue revealed that tyrosine acetylation was Na^+ sensitive. While studies have shown the utility of this method there are some disadvantages. For instance, N-acetylimidizole has cross reactivity with amino groups. The percent of lysine residues acetylated by N-acetylimidizole appears to be protein sequence dependent.[30] Additionally, acetylation is easily reversed. Treatment with stoichiometric amounts of hydroxylamine will readily reverse the modification. While N-acetylimidizole improves on the requirement of site-selectivity, the reversibility of the modification limits the utility of the method.

Similarly, reversible nitration can be achieved with tetranitromethane (Figure 1-1C). Tetranitromethane will nitrate the hydroxyl group of tyrosine residues at pH 8. The nitrophenoxide modification can be easily observed at the strong 428 nm emission. Nitration of tyrosine residues of the Na^+/glucose cotransporter by tetranitromethane showed similar results to N-acetylimidizole treatment. Treatment of the transporter by tetranitromethane inhibited glucose transport in rabbit intestinal brush-boarders in the absence of Na^+.[32] However, one drawback of this strategy is that it is not exclusively site-selective for tyrosine. At pH 8 cysteine residues are also modified. The site-selectivity of tetranitromethane is specific for cysteine at pH 6. Similar to acetylation, nitration is also reversible, but through reduction. Tetranitromethane modification of tyrosine residues is a step in the right direction but the drawbacks of partial site-selectivity, labile modification, and inability to be further functionalized leaves much to be desired.

Around the same time as methods using N-acetylimidizole and tetranitromethane were being developed, evidence had shown *p*-nitrobenzene sulfonyl fluoride labels tyrosine residues with site specificity (Figure 1-1D).[32,33] Modification of tyrosine residues by *p*-nitrobenzene sulfonyl fluoride results in *O*-(*p*-nitrobenzesulfonyl) tyrosine. One early example of this is *p*-nitrobenzene sulfonyl fluoride treatment inactivates Ca^{2+}-DNase isolated from bovine pancreas. However, *p*-nitrobenzene sulfonyl fluoride treatment of Ca^{2+}-free DNase yields fewer *O*-(*p*-nitrobenzesulfonyl) tyrosine sites. These studies are consistent with previous studies that indicate a Ca^{2+} dependent conformational change observed in active state, Ca^{2+}-DNase, when compared to inactive state, Ca^{2+}-free DNase.[34] Despite the demonstrated utility of sulfonyl fluoride reagents, it is unclear why they were relatively untouched for the next 20 years.[35] Later, sulfonyl fluoride electrophiles emerge

as privileged warheads in chemical proteomics through ABPP probe and irreversible covalent inhibitor development.[36]

Focus on tyrosine iodination using IPy_2BF_4 was a step in the right direction towards a uniquely tyrosine selective chemical modification (Figure 1-1E). This is an improvement over previously used method that uses sodium iodide. Sodium iodide, in the presence of an oxidant, has been used to iodinate tyrosine and histidine side chains. In addition to the enhanced site-specificity, IPy_2BF_4 is compatible with the hydrophilic nature and thermal instability of amino acid based substrates, making it applicable for bioconjugation. Initially, strategies resulted in undesirable mixtures of mono-, diiodinated, and unreacted products. [37] Further optimization showed absence of an acid predictably results in the diiodinated product. Notably, even further optimization of reaction conditions revealed presence of an acid such as HBF4 or TFA resulted in monoiodinated product. However, the addition of an acid makes the reaction conditions less ideal for the application of protein modification. IPy_2BF_4 can be used to ^{125}I-iodinate proteins that can aid in the crystallographic studies. Iodination is an attractive strategy because of the high electronic density and confers information based on the accessibility and confirmation of the Tyr residue. [38] This strategy was an improvement over previous methods, such a Bolton-Hunter method, that were not residue specific and introduced additional covalent modifications that may affect the biological activity. [39] Tyrosine iodination by IPy_2BF_4 opened the possibilities for installation of allylic functionalities through Suzuki-Miyaura cross-coupling reactions.

A continuation of this work shows iodination using IPy_2BF_4 followed by same-pot Suzuki-Miyaura cross-coupling reaction conditions in water yields mono- or di-alkylated

modifications to the phenyl ring of tyrosine residues.[40] This method was demonstrated on a model peptide, Leu-enkephalin, an endogenous neurotransmitter, with morphine-like activity.[41] This application demonstrated how this method can be used to further diversify peptide libraries. Studies that investigate the physiological implications of this modification have not been explored. Additionally, this method has not been applied to whole protein substrates. Furthermore, it would be advantageous to deploy a strategy using cross-coupling that resulted in one modification per tyrosine residue.

Conveniently, palladium-catalyzed cross-coupling of biomolecules with electrophilic π-allyl complexes can be used to install allylic functionality on the hydroxyl of the tyrosine side chain (Figure 1-1F).[42,43] This method has been applied to bioconjugation of α-chymotrypsinogen A from bovine, a model protein common to the field of tyrosine bioconjugation reactions.[44] α-chymotrypsinogen A is an ideal model protein because of the presence of only four tyrosine residues, only two of which are solvent expose. A palladium-catalyzed cross-coupling reaction of α-chymotrypsinogen A with taurine carbamate was used to install molecules of hydrophobic C_{17} chains. Installation of this group demonstrated the ability to "switch" the solubility of a protein. SDS-PAGE analysis confirmed α-chymotrypsinogen A incorporation into small unilamellar vesicles. While this method is useful for installing allylic functionalities on the hydroxyl of the side chain of tyrosine it would be favorable to develop an additional method that resulted in a carbon-carbon bond formation without the use of a metal catalyst.

Around the same time Francis also developing a metal free three component Mannich-type reaction (Figure 1-1G). This strategy results in a carbon-carbon bond formation adjacent to the side chain hydroxyl on tyrosine residues on proteins

preferentially over peptides.[45] Previously, formaldehyde was used for carbon-carbon bond forming protein crosslinking. This formaldehyde crosslinking suffers from lack of control for site selectivity in the presence of high formaldehyde concentrations and heat.[46] To circumvent the need for lysine to form imine intermediates, this strategy utilizes in situ formation of imines using aldehydes and electron-rich anilines.[47] Formaldehyde was identified as the optimal aldehyde reagent. Consistent with previous tyrosine bioconjugation development efforts, the three component Mannich-type reaction was also applied to model protein, α-chymotrypsinogen A. In this study, α-chymotrypsinogen A serine protease activity was easily evaluated by monitoring its proteolytic activity on a model peptide. The results showed bioconjugation at Y146 resulted in little to no inhibition of protease activity compared to the native, active form of α-chymotrypsinogen A after exposure to conjugation conditions with formaldehyde and 4-aminophenethyl alcohol.[47] Optimization of this method led to the use of ^{13}C labeled formaldehyde which was validated as a useful tool for the characterization of bioconjugation reactions by NMR.[48] Theoretically, this method can be used to study the role of tyrosine residue dynamics in catalysis, ligand binding, and protein-protein interactions. Drawbacks of the three-component Mannich-type approach is slow reaction rate with macromolecules such as polymers and requires large excess of coupling reagents.[24]

Work in the area of tyrosine bioconjugation performed by Barbas was inspired by the metal-free Mannich-type approach first reported by Francis, et al. This strategy employs capitalization on the efficient aqueous ene-type reaction through the reactivity of diazodicarboxylate-related molecules for tyrosine orthogonal bioconjugation (Figure 1-1H).[49–52] While substituted phenols readily react with diazodicarboxylates in the presence

of an activating protic or Lewis acid additive, this strategy is limited to organic solvents due to rapid decomposition of diazodicarboxylate reagents in aqueous conditions.[53] Conveniently, cyclic diazodicarboxamides are not as sensitive to proton activation compared to acyclic diazodicarboxamides and are therefore more applicable to aqueous chemistry. A survey of diazodicarboxylate and diazodicarboxamide regents revealed 4-phenyl-3*H*-1,2,4-triazole-3,5(4*H*)-dione (PTAD) as a suitable reagent for peptide labeling. When equimolar mixture of N-acyl methyl amides of histidine, tryptophan, serine, cysteine, lysine, and tyrosine were treated with PTAD only tyrosine modification was observed by ^{1}H NMR albeit conversion was incomplete. Studies with a model acyclic peptide reaction with PTAD proceeded rapidly, and the reaction conditions were carried through experiments using model proteins: α-chymotrypsinogen A, bovine serum albumin (BSA), and myoglobin and a cyclic diazodicarboxamide rhodamine. ESI-MS analysis confirmed tyrosine modifications on α-chymotrypsinogen A and BSA but unfavorably low levels of tyrosine modification on myoglobin. Additionally, the amount of α-chymotrypsinogen A and BSA labeled by the reagents, 35-96% yield, left room for improvement.

Dirhodium metallopeptide catalysts have shown promise in the modification of tyrosine residues, through an O-H insertion resulting in ester formation (Figure 1-1I). In this method selectivity was conferred based on sequence recognition in coiled coil protein-protein interactions in contrast to residue microenvironment, which previously had been the major focus.[54] The appeal of this strategy is to be able to target and modify a single tyrosine residue previously unreactive in previously developed tyrosine bioconjugation reactions. While this strategy has been optimized and applied to model peptide substrates,

it has yet to be applied to a whole protein. The proteins and therefore tyrosine residues for which this method can be applied are limited to those that bind coiled coil domains. Furthermore, application of this method requires rational design and optimization for each target. Also, the dirhodium center, which is the site of catalysis anchored through two glutamate residue side chains, is reactive with tryptophan and phenylalanine in addition to tyrosine as confirmed by MS/MS. This strategy also requires the use of large equivalents of the diazo substrate, 50-70 eq., for complete conversion because of the reactivity of the metallocarbenoid intermediate with water.

In an effort to add to the selection of tyrosine modification reactions, focus was directed towards increasing yields and decreasing site cross reactivity. A screening of electron-rich and oxidizing compounds revealed Cerium (IV) ammonium nitrate (CAN) as an oxidant that facilitates the addition of aniline derivatives to tyrosine and tryptophan residues through oxidation (Figure 1-1J). [55] The strategy was applied to tyrosine modification of angiotensin in HEPES buffer at a pH of 4.5 for 1 h for addition of both phenylene diamine and anisidine, as confirmed by MS/MS analysis. Anilines with electron-donating substituents were found to react more amenably to reaction conditions over anilines with nitro- or halogen- substitutions which were found to have low reactivity. Alkylated anisidine shows preference for tyrosine over tryptophan which suggests the potential for side-chain reactivity specificity conferred by the aniline substrate, though the reaction resulted in a mixture of isomeric products. Despite the desire for reaction conditions compatible with biological samples, addition of the CAN reagent affectively decreases neutral buffer pH to ~4, at which point proteins begin to denature. It has been found that increasing the buffering capacity of the buffer has deleterious effects on the level

of reactivity. Consistent with previous tyrosine synthetic modification studies, viral MS2 capsid and α-chymotrypsinogen A were used as model systems to investigate PEGylation and modification with cell-targeting peptides on native and introduced surface tyrosine residues. Another drawback of this method is the formation of both a major and minor conjugation product. The major product results from bond formation on the oxygen of the tyrosine side chain. Minor production formation is the result of a carbon-carbon bond formation *ortho* to the tyrosine hydroxyl. More recently tyrosine modification by way of reactivity with enzymatically formed intermediates has gained much attention.

A unique approach to tyrosine modifications was recently demonstrated in the context of an enzymatic reaction catalyzed by ascorbate peroxidase (APEX) (Figure 1-1K). [56] An engineered mitochondrial targeting APEX, mito-APEX, was expressed in HEK cells for the purpose of identifying proteins in close proximity to mito-APEX through biotinylation. This is accomplished through the reaction of the short-lived biotin-phenol radical, generated by mito-APEX, and the interaction of these radicals with proximal proteins. Biotin labeling allowed for streptavidin enrichment and subsequent MS analysis showed modification by biotin-phenol was specific to tyrosine residues. While the goal and motivation of this work was not to develop a tyrosine specific modification strategy, it demonstrates the possibility of performing tyrosine specific modifications in live cells.

Deoxyribozymes have shown promising potential as catalysts for DNA-catalyzed tyrosine or serine selective modifications (Figure 1-1L). [57] This method has been demonstrated in both tethered and free peptide substrates, but sights are set on achieving DNA-catalyzed side chain modifications in whole proteins. From this work, Tyr1, a tyrosine specific deoxyribozyme, forms a tyrosine-RNA linkage via conjugation of 5'-

triphosphorylated RNA oligonucleotide with a tyrosine side chain.[58] This strategy can overcome limitations often encountered with small-molecule approaches such as targeting proteins in which the structure is deemed "undruggable" or have challenging or shallow binding pockets.

Another method of conjugation is protein cross-linking. Tyrosine residues in the presence of horseradish peroxidase (HRP) and H_2O_2 can undergo a one-electron oxidation that favors dimerization with other tyrosine residues.[59,60] Sequential dimerization reactions can lead to extensive cross-linking. In one example a tyrosine containing tag is genetically introduced to *E. coli* alkaline phosphatase (BAP) at the C-terminus. HRP recognition of the tyrosine containing protein tags lead to cross-linking in the presence of HRP and H_2O_2, but not for the wild type protein. BAP returned to their native monomeric state and function upon cleavage of the cross-linked tags through treatment with a protease.

Most recently progress has been made in the area of strain-promoted tyrosine modification reactions. Accessible tyrosine residues are oxidized by mushroom tyrosinase which yields a 1,2-quinone intermediate, which can be further modified by a strained quinone-alkyne, bicyclo[6.1.0]nonyne (BCN).[61] It has been demonstrated that this strategy can be applied to modification of tyrosine residues on a model protein laminarinase A at a solvent exposed tyrosine residue incorporated by site-directed mutagenesis. Additionally, the method has been applied to an antibody-drug conjugate (ADC) which involved the modification of anti-influenza AT1002 and potent peptide based mitotic inhibitor, monomethyl auristatin F (MMAF).

Another recently developed bioconjugation reaction is *O*-tyrosine modification by tetrafluoro-1,4-benzoquinone.[62] This strategy was uncovered in a study that was

investigating the reactivity of carcinogenic metabolites of polyhalogenated aromatic compounds, halogenated quinones. So far this strategy has only been applied to tyrosine and tyrosine containing peptides. This strategy can potentially be used to study the effects of tyrosine PTMs on protein function in the cell. One drawback however is the limited functionality imparted by the conjugation.

Selective Tyrosine Modifications

Ideally, covalent selective modifications target a conserved nucleophilic residue in a class of proteins or a single residue on a single protein target. These selective molecules are particularly helpful in the fields of chemoproteomics and chemical biology through the use of ABPP probes (ABP).[63] These probes can be used to interrogate the relative activity level of a particular protein or protein class. Selective covalent modulators can also be used in the identification of new therapeutic protein targets or off targets, discovery new inhibitors and potential therapeutics, and evaluation of drug-target occupancy.[19] The criteria for a ABP is that it must selectively target functional reactive residues above all other residues. Sites with heightened activity are typically concealed in enzymatic active sites or are points of regulation by PTMs. The chemical scaffold of selective covalent small molecules can be mechanism-based inhibitors, electrophilic ligand analogues, or in the simplest form merely the electrophilic species. Many selective cysteine and serine specific small molecules have been identified. A driving force for expanding the current list of known selective tyrosine modulators, is to increase the current understanding of tyrosine at the molecular level, as it relates to disease. Here is an account of the progress and developments achieved in the area of selective tyrosine synthetic modification reactions.

Organophosphorus (OP) reagents are large category of small molecules that are traditionally used as insecticides and herbicides, but are known to have an effect of neuronal transmission in humans. These agents or their metabolites function as acetylcholinesterase inhibitors that function by targeting the active site serine residue which results in a pooling of acetylcholine. More recently, further characterization of additional OP modified protein targets has revealed tyrosine as the site of modification in proteins without serine active site residues. Peptidase enzyme papain, isolated from papaya, was found to be modified by organophosphorus agent, diisopropyl fluorophosphate (DFP), at tyrosine 123 selectively over the other 17 tyrosine residues present in the protein.[64] DFP, among other FP reagents, was inspiration for the development of the mechanism-based FP-biotin ABP. FP probe analogues were found to only react with active serine hydrolases in contrast to the inactive form. Application of FP-biotin to whole proteomes has revealed over 100 distinct targets.[63] OP agents have been found to modify tyrosine residues on tubulin and albumin.[65,66] While OP based probes such as FP-biotin give access to the ability to investigate the role and activity of some tyrosine residues, it would be advantageous to have an entirely tyrosine selective strategy.

Another broadly reactive ABP with reactivity towards tyrosine residues is the carbon electrophile based phenylsulfonate ester (SE) probe (Figure 1-2A).[67] SE was found to react with aspartate, glutamate, histidine, and tyrosine residues in mouse tissue proteome. SE selectively reacts with the active site proton acceptor, tyrosine 183, of mouse corticosteroid-11 b-dehydrogenase 1. The reactivity profile of SE towards tyrosine residues is a welcome addition to the ABP toolbox. However, the promiscuity of the probe leaves room for improvement.

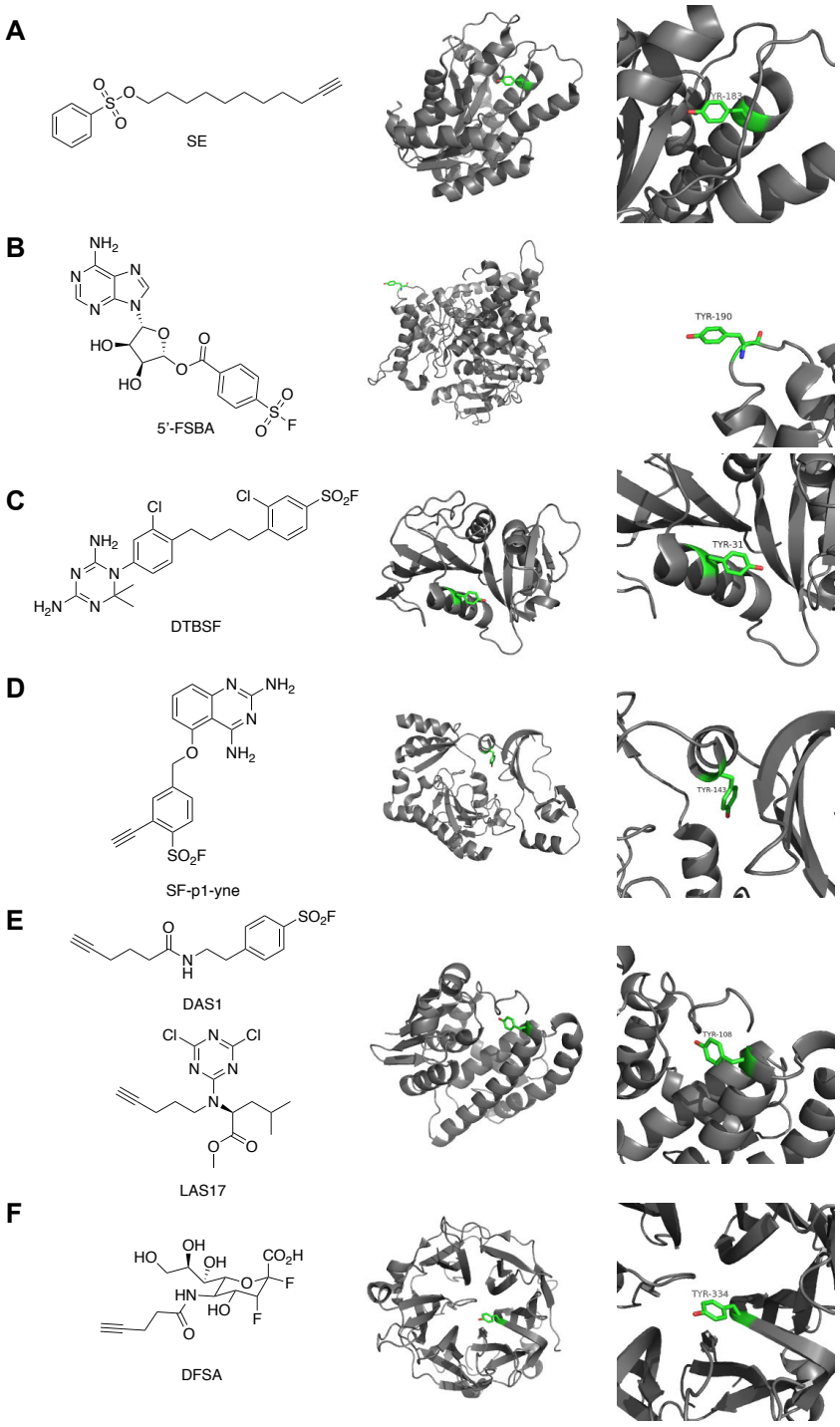


Figure 1-2. Protein and residue targets of selective tyrosine chemical modifications. (A) SE probe modifies Tyr 183 of mouse corticosteroid-11 b-dehydrogenase 1 (PDB: 1Y5R). (B) 5'-FSBA modifies Tyr 190 of bovine liver glutamate dehydrogenase (PDB: 1HWZ). (C) DTBSF (NSC 127755) modifies Tyr 31 of chicken liver dihydrofolate reductase (PDB:

1DR2). (D) SF-p1-yne and parent compound SF-p1 modify Tyr 143 of mRNA-decapping scavenger enzyme (DcpS) (PDB: 4QDV). (E) DAS1 and LAS17 target conserved tyrosine residues in the active sites of GSTP1 (PDB:6GSS). (F) Alkyne-hinged 3- fluorosialyl fluoride (DFSA) modifies Tyr 334 human cytosolic sialidase Neu2 (PDB: 1SNT).

The sulfonyl fluoride warhead has been shown to be a privileged electrophile in the field of ABPP and chemical biology.[36] Work this this area began with the development of covalent ATP and NADH analogues to identify the residues situated around the NADH binding domains. 5'-p-fluorosulfonylbenzoyladenosine (FSBA) inhibits glutamate dehydrogenase similar to the natural allosteric inhibitor, NADH (Figure 1-2B). Sequence analysis revealed that FSBA covalently modifies tyrosine 190 and lysine 420 of bovine liver glutamate dehydrogenase.[68] More recently global proteomics studies with FSBA have been performed, leading to the identification of purine-binders in Jurkat cells.[69] This study identified 185 unique FSBA-labeling sites through the technique COmbined FRActional DIagonal Chromatography (COFRADIC). Of the 185 unique FSBA-labeling sites, 124 were identified at tyrosine modifications. One example, tyrosine-protein kinase lck, proto-oncogene from the Src family, was identified as being modified by FBSA at tyrosine 393, a known autocatalytic phosphorylation site. This finding that FBSA showed reactivity towards Lck was consistent with previous reports.[70] This study demonstrated FSBA has a preference for tyrosine modifications over all other nucleophilic residues.

Another early sulfonyl fluoride, DTBSF (NSC 127755), was identified through a screening of potential triazine based dihydrofolate reductase inhibitors.[71] DTBSF was found to label chicken liver dihydrofolate reductase at tyrosine 31. The structure of DTBSF

is similar to the chemotherapeutic methotrexate, a known substrate analogue noncovalent inhibitor of dihydrofolate reductase. Inhibition of dihydrofolate reductase inhibits tetrahydrofolate synthesis causes downstream DNA and protein synthesis inhibition. DTBSF and its analogues can be used to develop more effective chemotherapeutics and give insight into potential off targets. However, global proteomics studies using DTBSF have not been explored.

More recently, a high-throughput phenotypic study revealed a sulfonyl fluoride inhibitor as an attractive potential therapeutic for spinal muscular atrophy. In this study a noncovalent small molecule, D156844, was found to upregulate the expression of SMN protein, known to be downregulated in the diseased state. Use of radiolabeled inhibitor identified mRNA-decapping scavenger enzyme (DcpS) as the protein target. Rational design of covalent DcpS inhibitors sought to specifically target tyrosine residues, proximal to the inhibitor binding area. Chemoproteomic strategies with the ABP version of SF-p1, SF-p1-yne, facilitated target occupancy studies with the parent compound (Figure 1-2D). This work demonstrates the utility of investigating the mechanism of action of covalent inhibitors with the help of ABPs.

Another useful SF containing small molecule is the serine protease inhibitor (2-aminoethyl)benzenesulfonyl fluoride (AEBSF). Despite the affinity of AEBSF for active site serine residues in serine proteases, AEBSF has been found to modify tyrosine 547 of porcine serine exopeptidase, dipeptidyl peptidase IV (DPIV).[72] The reactivity of AEBSF inspired work in our lab to develop a serine protease ABP. We synthesized DAS1, which utilizes the backbone of AEBSF and contains an alkyne handle for incorporation of a reporter group, enabling target functionalization and identification (Figure 1-2E).[73]

Analysis of DAS1 protein targets revealed reactivity towards serine protease in a whole proteome. Furthermore, we demonstrated that pretreatment with AEBSF outcompeted labeling with DAS1. A surprising result from the proteome reactivity analysis indicated DAS1 treatment enriched for 10 different glutathione s transferases (GSTs).[21] GST enzymes are most well known for their cellular detoxification activity.[74] GSTs catalyze the conjugation of electrophilic species with the sulfur of glutathione. Furthermore, site of labeling MS studies revealed DAS1 has reactivity towards a functional conserved tyrosine residue in GSTs. However, DAS1 is not the first SF containing small molecule to target tyrosine residues in GSTs.

Alkyne-hinged 3- fluorosialyl fluoride (DFSA), a probe with an electrophile specific to the sialic-acid scaffold, modifies a tyrosine within sialidase active sites (Figure 1-2F).[75] DFSA was show to be a tool for detecting influenza infections, due to the ability to react with sialidases that are essential for the pathology of invasion and escape.[76] In this study human cytosolic sialidase, Neu2, was found to be modified by DFSA at tyrosine 334. Ester-protected DFSA (PDFSA) is an improvement over previous sialidase targeting ABPs, quinone methide and photoaffinity probes, by overcoming issues of nonspecific labeling and insufficient cell permeability. Unfortunately, the mechanism-based specificity of the electrophile to the sialic-acid target-specific scaffold excludes its use from application of broad tyrosine profiling experiments.

GST modification by SF reagents has previously been demonstrated by labeling with 4-(fluorosulfonyl)benzoic acid (4-FSB). Colman et al. has used FSB to probe glutathione S-transferase (GST) enzymes in rat liver isozyme 1-1 at Tyr 8 and in isozyme 4-4 at Tyr 115, and Tyr 7 and Tyr 106 of pig lung glutathione S-transferase Pi (GSTP).[77–

[79] These studies have shown that Tyr 8 of rat liver isozyme 1-1 contributes to the catalytic activity, but is not essential for enzyme function. This study revealed the presence of a reactive tyrosine residue in the active site binding pocket of GST enzymes.

More recently our lab developed another specific tyrosine reactive probe. LAS17 is selective for glutathione S transferase pi (GSTP1) over the other GST isoforms (Figure 1-2E).[22] LAS17 utilized a dichlorotriazine electrophile and features a leucine methyl ester directing group. Treatment of human GSTP1 results in loss of catalytic GSTP1 activity. Mouse xenograft studies showed treatment with LAS17 decreased tumor growth rate and decreased the size of the tumor.[80] Furthermore, LAS17 was used to characterize GSTP1 as a driver of triple-negative breast cancer metabolism and pathogenicity. Typically, patients diagnosed with triple-negative breast cancer have a poor prognosis.[81] GSTP1 expression levels are upregulated in triple-negative breast cancer compared to other forms of breast cancer. Triple-negative breast cancer cells treated with LAS17 showed a decrease in cell survival, which was consistent with cell survival studies with GSTP1 knockdown cells. This demonstrated that chemical treatment with LAS17 has the same effect as genetically altering the expression of GSTP1. Inactivation of GSTP1 disrupts GSTP1-induced activation of GAPDH, leading to a dysregulated glycolytic metabolism, resulting in lower levels of ATP.

Breast cancer is the second leading cause of cancer related deaths for US women. The average risk of a woman in the US to develop breast cancer in her lifetime is 12.3%.[82] There have been strong initiatives to encourage early screening which have resulted in a decrease in the mortality rate. The US Preventive Services Task Force (USPSTF) recommends mammogram screening ever two years after the age of 50.[83] Breast cancer is

a broad term that encompasses a variety of disease states that are characterized by their unique histology, biology, pathology, and clinical prognosis.[84] Typically, aggressive breast cancers have a poor prognosis that can be attributed to epithelial–mesenchymal transition (EMT) which is commonly seen in cancer metastasis.[85] One particular aggressive breast cancer with poor survival rates is triple-negative breast cancer (TNBC).

TNBC is characterized by the absence of three proteins: oestrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). TNBC is the most common breast cancer subtype observed in *BRCA1* carriers, and occurs most frequently in young African-American women.[81] Currently, there are very few therapeutic options for treatment of TNBC. Due to the absence of HER2, typical chemotherapeutic strategies cannot be used to treat TNBC. Alternatively, here we report the identification of a member of the cytosolic glutathione S-transferase (GST) superfamily, glutathione S-transferase Pi 1 (GSTP1), as a metabolic driver of TNBC through an activity based chemoproteomics approach. Furthermore, we report an investigation of the effects of chemical and genetic inhibition of GSTP1 on TNBC survival *in vitro* and *in vivo*. Until this point, exclusive GSTP1 inhibition has not been proposed as a strategy for the treatment and eradication of TNBC.

The glutathione S-transferase (GST) Fold, Phase II detoxifying enzymes, consists of four distinct superfamilies (cytosolic, Kappa class mitochondrial, MAPEG, and fosfomycin resistance) which are differentiated by structure.[86] Cytosolic GSTs are the most extensively studied GST family, which consist of proteins with variable and overlapping functions. The abundance of attention has been given to the 7 distinct classes of cytosolic GSTs and is likely due to the fact that they are found in all life forms. Cytosolic GSTs have

both a glutathione (GSH) and a hydrophobic binding pocket. GSTs have been particularly helpful in the field of protein science. Researchers have taken advantage of the strong affinity of the GSH binding domain to aid in protein purification of GST tagged recombinant proteins. Incorporation of a GST GSH binding domain facilitates affinity purification tag with GSH functionalized beads.[76]

Activities of cytosolic GSTs include: GHS conjugation, tyrosine and phenylalanine degradation[87], isomerase[88], prostaglandin synthesis[89], protein glutathionylation[90], cellular signaling regulation[91], and regulation of ion channels.[92] The most well studied activity of cytosolic GSTs is the conjugation of glutathione (GSH) to electrophilic compounds (Figure 1-3). This class of molecules includes carcinogens, chemotherapeutics, xenobiotics, and pesticides or naturally occurring metabolites from oxidative stress such as aldehydes, epoxides, hydroperoxides, and quinones.[74,93] The role of GSTs in medicine, namely the neutralization of drugs and xenobiotics has fueled efforts to broaden our understanding of their role in biology.[94]

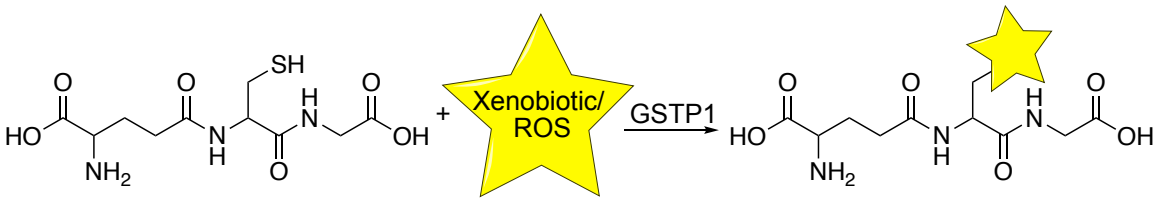


Figure 1-3. Catalytic cellular detoxification activity of GSTs.

GSTP1 is the most ubiquitous member of the soluble GST protein superfamily and is primarily responsible for catalyzing the conjugation of GSH to exogenous and endogenous electrophiles as a mechanism of cellular detoxification.[95] GSTP1 is believed

to exist and function as a homodimer (Figure 1-4). In addition to its catalytic function, GSTP1 has also been demonstrated to regulate the Jun N-terminal kinase (JNK)/mitogen-activated protein kinases (MAPK) signaling pathway through protein–protein interaction with JNK.[91]This protein-protein interaction is disrupted by self-redox regulation of post translational modification, S-glutathionylation, of Cys47 and Cys101.[90] Tyrosine phosphorylation of GSTP1 by the epidermal growth factor receptor (EGFR) kinase promotes formation of the GSTP1–JNK complex, which suppresses downstream JNK signaling.[96] A variety of human cancers, including breast, colon and ovarian cancers, express high levels of GSTP1 relative to healthy tissue.[97] Furthermore, human tumor cell lines that exhibit resistance to chemotherapeutic agents have elevated GSTP1 levels.[98] Resistance is attributed to GSTP1- catalyzed inactivation of chemotherapeutic alkylating agents through conjugation of GSH.

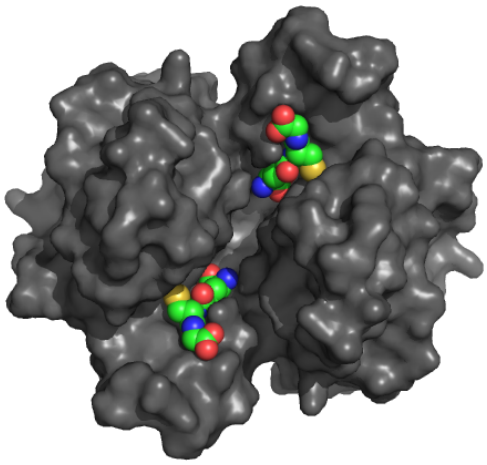


Figure 1-4. Surface representation of GSTP1 homodimer. GSH is shown in in the GSH binding domain as spheres (PDB: 6GSS).

Due to the confirmed roles of GSTP1 in promoting tumorigenesis and drug resistance, GSTP1 has emerged as a promising cancer therapeutic target. Many GSH analogs, small molecules and natural products have been identified as GSTP1 inhibitors.[91] Reversible GSTP1 inhibitors include 8-methoxypsoralen, piperlongumine, aloe emodin, and benstatin A,[91,99,100] as well as Telintra (Ezatiostat), which was in clinical trials for the treatment of myelodysplastic syndrome.[101,102] GSTP1 contains four cysteine residues, two of which (Cys47 and Cys101) have been shown to be highly reactive and targeted by irreversible GSTP1 inhibitors.[103] These cysteine-targeted covalent modifiers of GSTP1 include quercetin,[104] and S-(N-benzylthio-carbamoyl) glutathione (BITC-SG).[105] Quercetin targets Cys47, whereas BITC-SG covalently modifies Cys47 and Cys101 via an S-thiocarbamoylation reaction.[104,105]

In addition to the two characterized reactive cysteines, GSTP1 has also been shown to contain reactive tyrosine residues that are susceptible to modification by sulfonyl-fluoride electrophiles. These sulfonyl fluorides have been shown to target the equivalent of tyrosine 8 and tyrosine 108 in mouse, rat, and pig GSTP1.[21,78,79,106] Mutation of the equivalent of Y108 to phenylalanine in the Schistosoma japonicum GST resulted in loss of GSTP1 activity,[21] indicating the functional importance of this tyrosine residue. Previous sulfonyl-fluoride probes used for tyrosine labeling are not selective for GSTP1, and covalently modify various other proteins within the proteome.[21,73] Here, we report the development of a selective, irreversible GSTP1 inhibitor, LAS17, which covalently modifies the functional tyrosine, Y108, in human GSTP1. LAS17 was discovered through screening a library of dichlorotriazine-containing compounds and shown to selectively modify GSTP1 within human cancer-cell proteomes.

This work was inspired by pervious initiatives focused around investigating the proteome reactivity of a library of aryl-halides.[107] This effort was motivated by the desire to expand the currently available set of activity based protein profiling probes (ABPs). ABPs are tools implemented in the field of activity based protein profiling (ABPP). Activity based protein profiling utilizes covalent small molecule probes to selectively target reactive residues in proteins, which correlates to overall enzymatic activity.[63] Understanding the relative reactivity of a protein can give insight into the functional properties of the protein. One limitation of the field of ABPP is that in order for a protein to be studied it must be the target of an ABP. With this limitation comes the incentive to expand the so-called "chemical probe tool box," thus expanding the protein targets available for exploration.[18]

Our efforts to expand the currently available ABPs was inspired by the report of a cysteine specific inhibitor of peroxisome proliferator activated receptor-γ (PPARγ).[108] PPARγ plays a role in fat storage and PPARγ deficient mice are protected against high-fat induced obesity and insulin resistance.[109] In this study, GW9662, was found to inhibit the activity of PPARγ through selective modification of cysteine 285. The covalent adduct was achieved through a nucleophilic aromatic substitution (S_NAr) reaction between the sulfur atom of the cysteine side chain and displacement of a chloride from the *p*-chloronitrobenzene electrophile. We hypothesized that other aryl-halide compounds, such as dichlorotriazine would undergo the same reaction with nucleophilic amino acid side chains (Figure 1-5). This inspired the synthesis of an aryl-halide containing probe library.[107]Evaluation of the aryl-halide probe library for protein reactivity and site of modification revealed that dichlorotriazines predominately targeted lysine residues within

a proteome, in contrast to *p*-chloronitrobenzenes that are highly cysteine selective (Figure 1-6).

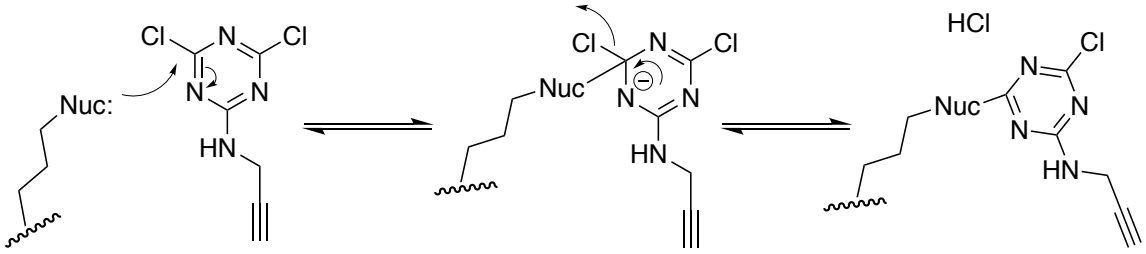


Figure 1-5. Proposed nucleophilic aromatic substitution reaction mechanism with RB7 that results in a covalent modification to protein targets.

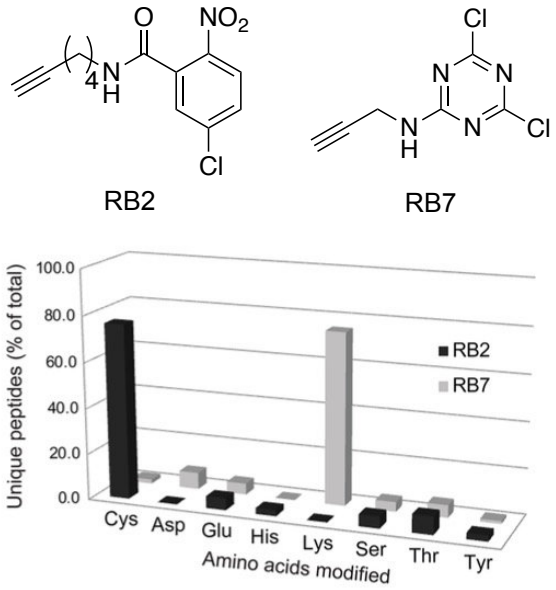


Figure 1-6. Amino acid specific reactivity profile of *p*-chloronitrobenzene (RB2) and dichlorotriazine (RB7) probes. (Figure adapted from Shannon, D.A. et al 2014.)

To further explore the unique reactivity of the dichlorotriazine electrophile and develop selective covalent modifiers to target reactive amino acids other than cysteine, we

sought to synthesize a library of probes with the following functionalities: (1) a dichlorotriazine electrophile for covalent modification of the protein target; (2) a bio-orthogonal alkyne handle for CuAAC-mediated target identification; and (3) a diversity element to direct the probes toward diverse subsets of the proteome (Figure 1-7).

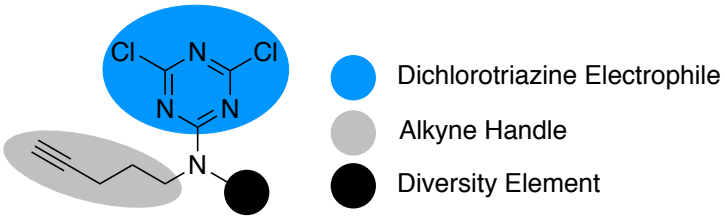


Figure 1-7. Dichlorotriazine library probe design.

Dichlorotriazines have been used by the textile industry as irreversible dyes. The covalent modification of the dye to cellulose present in cotton and flax prevents fading due to repeated wash cycles and bleaching.[110] While dichlorotriazine dyes have positively impacted the textile industry, the environmental effects of dichlorotriazine waste has been less favorable. Dichlorotriazine dyes have been found to covalently modify enzymes in biological samples in inhibit enzyme activities. One example of this is Vilmafix Blue AR (VBAR) which covalently modifies lysine 81 and/or lysine 217 of NADH-dependent bovine heart mitochondrial L-malate dehydrogenase.[111] VBAR inhibits L-malate dehydrogenase in a time and concentration manner. Inhibition of L-malate dehydrogenase prevents reduction of oxaloacetate and a dysregulation of the citric acid cycle.

Dichlorotriazine library members contains a terminal alkyne functionality that allows for evaluating target occupancy and promiscuity of the individual probes in cellular and *in vivo* systems, using established copper(I)-catalyzed azide–alkyne cycloaddition (CuAAC) methods.[112,113] CuAAC gives access to further conjugation of protein targets

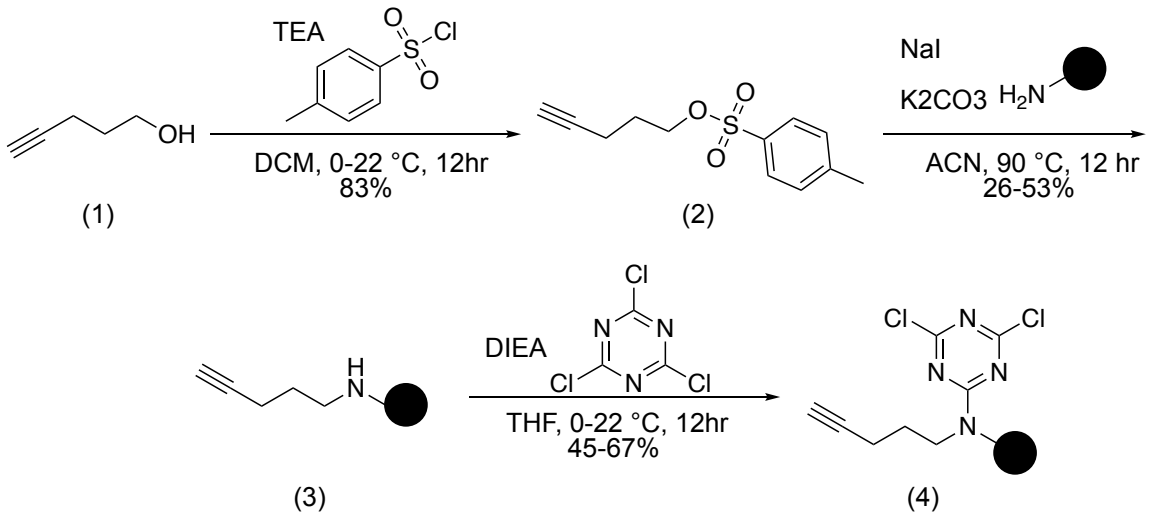
through the addition of a reporter group. Reporter groups allow for visualization or enrichment of the protein target. Examples of reporter groups are rhodamine-azide for in-gel fluorescence imaging and biotin-azide for protein target enrichment on streptavidin functionalized beads.

Dichlorotriazine library members also contain a diversity element. The incorporation of a diversity element gives both the protein target and probe addition binding contact points. Addition of these points favor more selective molecular interactions, which can fine tune the selectivity of the probe for a subset of the proteome. The selectivity of a probe library member with a diversity element will therefore be more selective for a protein or protein class above a probe containing the electrophile and alkyne alone.

Results and Discussion

Dichlorotriazine Library Synthesis and Evaluation

The dichlorotriazine probe library was synthesized via a 3-step divergent synthesis (Scheme 1-1). First, 4-pent-yn-ol (1) was tosylated with tosylchloride to generate a common intermediate (2). Then, the common intermediate (2) was then diversified through an S_N2 reaction with primary amine of 20 commercially available directing groups (3). Lastly, the resulting secondary amines were coupled to cyanuric chloride through an S_NAr reaction to afford the final library members (LAS1-LAS20) (Figure 1-8).



Scheme 1-1. Synthesis of dichlorotriazine library.

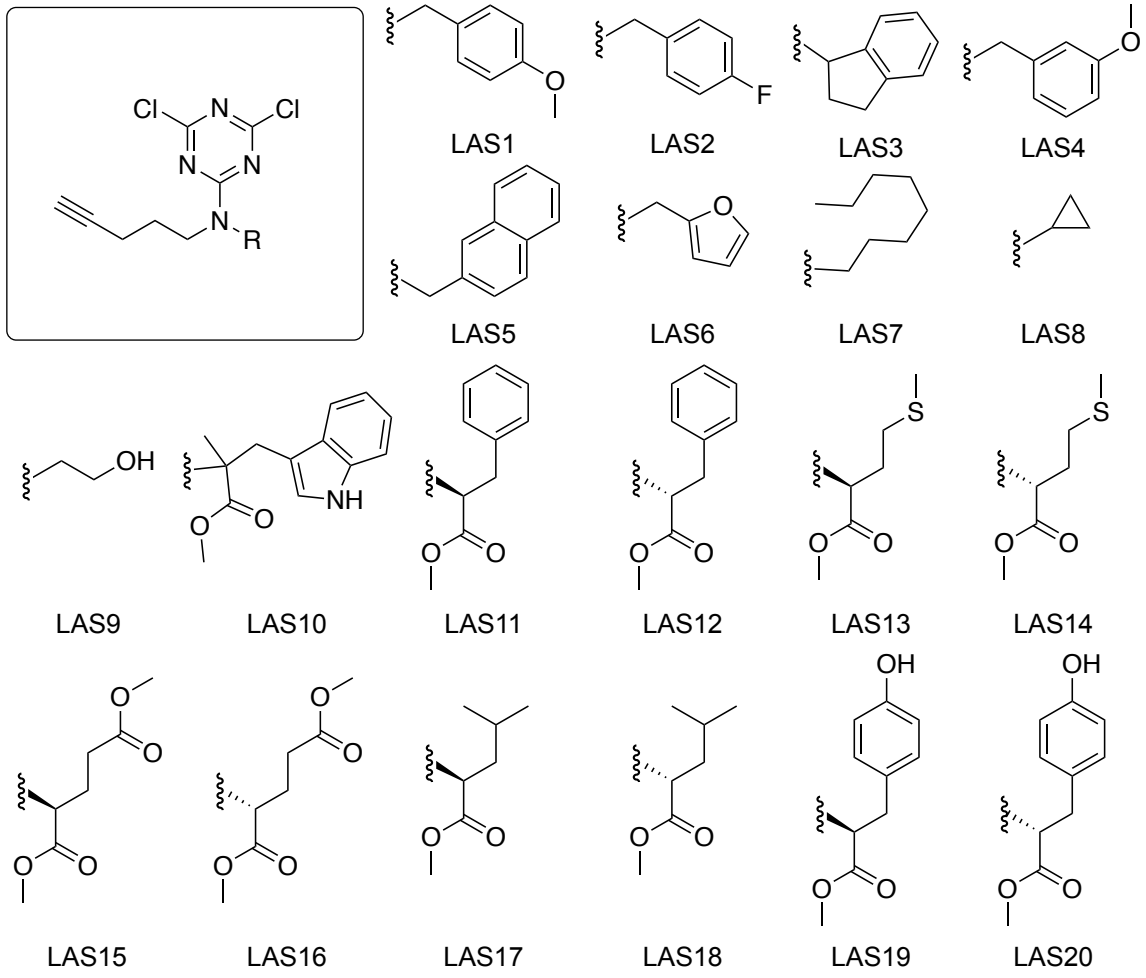


Figure 1-8. Dichlorotriazine probe design and library members, LAS1-LAS20.

LAS1-LAS20 were evaluated for covalent protein labeling in soluble HeLa cell lysates. Lysates were treated with 1 μM of each compound for 1 hour. Next, samples were subjected to CuAAC conditions with rhodamine-azide. Following the addition of rhodamine-azide, samples were resolved on SDS-PAGE and imaged for in-gel fluorescence visualization (Figure 1-9). These in-gel fluorescence screening studies demonstrated that several library members intensely labeled a low molecular weight protein band (~25 kD) (Figure 1-10). Of the 20 compounds, LAS17 (Figure 1-11A) containing the L-leucine methyl-ester directing group, was chosen for subsequent study due to the high intensity of labeling for the ~25 kD protein and minimal labeling of other proteins within the cell lysates Furthermore, treatment of live HeLa cells with 1 μM of LAS17 demonstrated the cellular permeability, potency, and selectivity of this compound for covalent modification of the ~25 kD within a cellular context (Figure 1-11B). The potency and selectivity observed in the live cell treatment was consistent with the reactivity profile observed in LAS17 treatment of lysates (Figure 1-12).

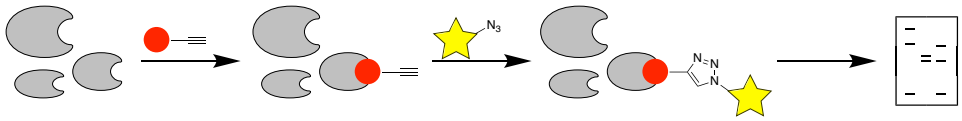


Figure 1-9. Probe library reactivity evaluation workflow in HeLa cell lysates.

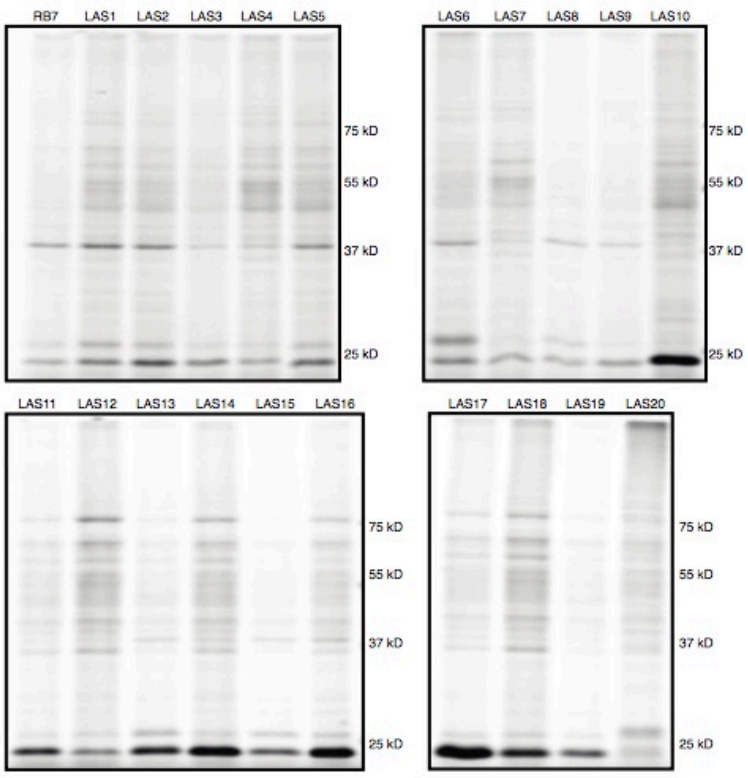


Figure 1-10. Evaluation of dichlorotriazine probe library (LAS1-LAS20) reactivity in soluble HeLa cell lysates (2.0 mg/mL).

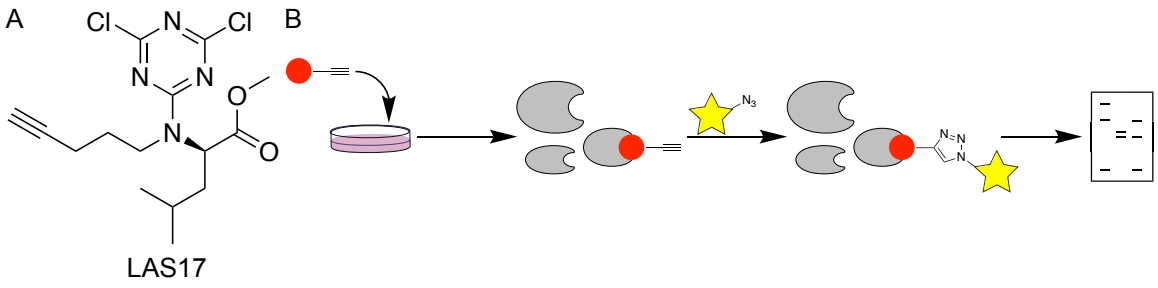


Figure 1-11. Live cell labeling. (A) Structure of LAS17. (B) LAS17 reactivity evaluation workflow in live HeLa cell labeling experiment.

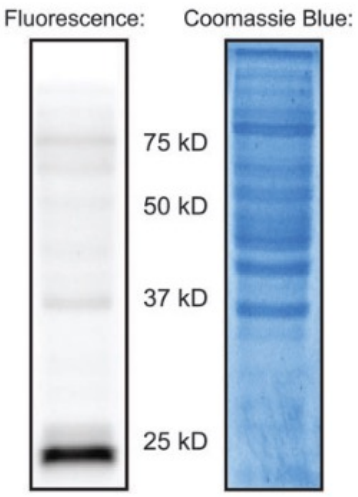


Figure 1-12. Evaluation of LAS17 reactivity in live HeLa cell treatment.

Identification of GSTP1 as the Target of LAS17

To identify the ~25 kD protein, HeLa cell lysates treated with either DMSO or 1 μM of LAS17 were subjected to CuAAC with biotin-azide (Figure 1-13). Biotinylated proteins were enriched on streptavidin-agarose beads. Unreacted proteins were removed by washing. Bead-bound proteins were subjected to on-bead trypsin digestions, and tryptic digests were subjected to LC/LC-MS/MS analysis. Spectral counts for proteins identified in three DMSO-treated samples were compared to LAS17-treated samples to identify proteins specifically labeled by LAS17 (Table 1-1, Table 1A-1). These data identified GSTP1 as the ~25 kD protein target of LAS17.

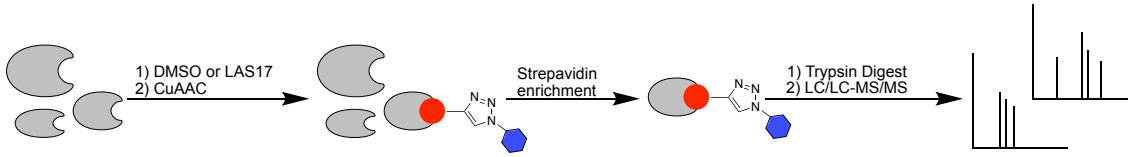


Figure 1-13. Sample preparation and mass spectrometry analysis workflow for protein target identification of LAS17.

Protein	Mol Weight (Da)	DMSO	LAS17
GSTP1 Glutathione S-Transferase	23342	0	24
HSP90AB1 Heat Shock Protein HSP 90-beta	83264	4	24
TKT Transketolase	67878	3	23
TUBA4A Tubulin alpha-4A chain	49924	4	22
PKM Pyruvate kinase isozymes M1/M2	57937	6	22
EEF1A1 Elongation factor 1-alpha 1	50141	5	20
FLNA Filamin-A	280737	6	20
TUBB Tubulin beta chain	49671	6	19
PGK1 Phosphoglycerate kinase 1	44615	1	14
HSP90AA1 Heat sock protein HSP 90-alpha	84660	0	12

Table 1-1. The top 10 proteins with the greatest difference in average spectral counts among three trial for DMSO and LAS17 treated samples.

Confirmation of GSTP1 as the Target of LAS17

To confirm the LC/LC-MS/MS findings, human GSTP1 with an N-term hexahistidine affinity tag was recombinantly expressed in *E. coli* and purified to homogeneity from the soluble fraction. The purified protein was treated with 1 μM of LAS17 and analyzed for in-gel fluorescence (Figure 1-14A). Significant protein labeling of purified GSTP1 was observed, confirming that LAS17 covalently modifies GSTP1 (Figure 1-14B).

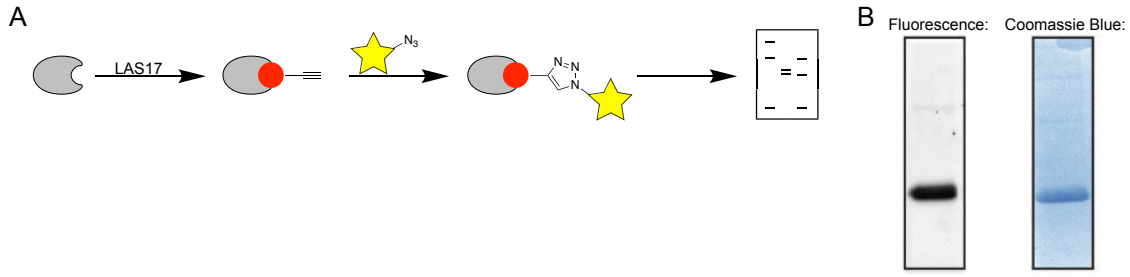


Figure 1-14. Evaluating LAS17 reactivity towards purified recombinant GSTP1. (A) Workflow to evaluate LAS17 reactivity towards purified, recombinant GSTP1. (B)

Recombinant, purified GSTP1 was treated with LAS17 and subjected to CuAAC, SDS-PAGE and in-gel fluorescence and Coomassie Blue staining.

LAS17 Treatment Inhibits GSTP1 Activity

To determine if LAS17 inhibits GSTP1 activity, an *in vitro* activity assay was performed that spectroscopically monitored the conjugation of GSH to 1-bromo-2,4-dinitrobenzene (BDNB). LAS17 inhibited *in vitro* GSTP1 activity after a 1 hr preincubation with LAS17 in a concentration dependent manner (Figure 1-15A). Given the covalent mode of action for LAS17, time-dependent inhibition of GSTP1 was monitored, affording a second-order rate constant of inactivation (k_{inact}/K_I) of 31,200 $M^{-1}s^{-1}$ (Figure 1-15B-C). To determine if the presence of other proteins affects the inhibitory activity of LAS17, inhibition studies were performed within the context of a complex background of HeLa cell lysates. The results demonstrate that the presence of other cellular proteins does not affect the inhibitory function of LAS17.

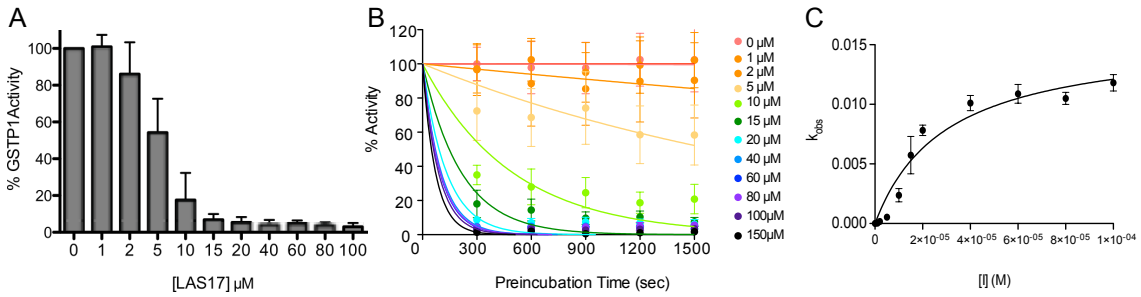


Figure 1-15. (A) *in vitro* GSTP1 activity in the presence of increasing concentrations of LAS17. (B-C) *in vitro* concentration and time-dependent inhibition of GSTP1 by LAS17.

Characterization of LAS17 Mechanism of Inhibition of GSTP1

The visualization of protein labeling after denaturing SDS-PAGE, and the time-dependent inhibition observed for LAS17, confirmed that this compound functions by covalently modifying GSTP1. However, given the high reactivity observed for the dichlorotriazine electrophile in previous proteome reactivity screening, we sought to determine in LAS17 was covalently modifying a single amino acid or multiple sites within GSTP1. To determine the stoichiometry of modification, intact-protein mass spectrometry of purified GSTP1 was performed in the presence of 3 equivalents of LAS17 (Figure 1-16A). The de-convoluted mass of the DMSO GSTP1 samples was determined to be 25779 Da (Figure 1-16B), whereas the mass of the LAS17 treated GSTP1 sample was determined to be 26102 Da (Figure 1-16C). The mass difference between the DMSO treated sample and the LAS17 treated sample is +322 Da. This mass spectrometry analysis indicates a single modified species corresponding to modification of GSPT1 with LAS17 at a single site, thus a stoichiometry of 1:1 (LAS17:GSTP1 monomer).

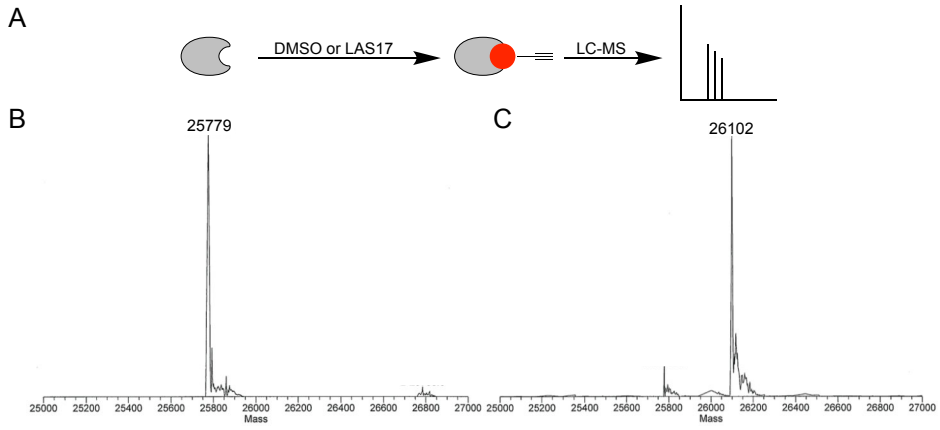


Figure 1-16. Intact-protein mass spectrometry of DMSO or LAS17 treated purified GSTP1. (A) Intact-protein mass spectrometry workflow. De-convoluted spectra of (B) DMSO and (C) LAS17 treated GSTP1.

To identify the site of GSTP1 labeling by LAS17, purified wild-type (WT) GSTP1 was incubated with LAS17 or DMSO and subjected to denaturation, reduction, cysteine alkylation, and trypsin digestion (Figure 1-17). The resulting tryptic-peptide mixtures were analyzed by LC-MS/MS. The generated fragmentation spectra were searched for a differential modification mass of +322.12 on all potentially nucleophilic amino acids. There analyses identified Y108 as the site of modification by LAS17 on GSTP1 (Figure 1-18). To confirm Y108 as the site of modification, a Y108F mutant was generated. The Y108F mutant was expressed and purified in tandem with WT GSTP1. Both the WT and Y108F-mutant GSTP1 were treated with LAS17 and analyzed by in-gel fluorescence. The results show that replacing the nucleophilic tyrosine residue with a phenylalanine eliminated the ability of LAS17 to covalently modify GSTP1, thereby confirming that modification was selectively occurring that this residue (Figure 1-19A). Y108 is located in the GSH binding site, and is conserved among the pi, mu, and theta GST classes in humans (Figure 1-19B).

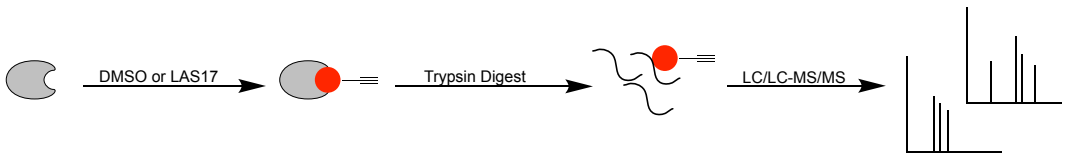


Figure 1-17. Workflow to determine the site of labeling of LAS17 on purified WT GSTP1.

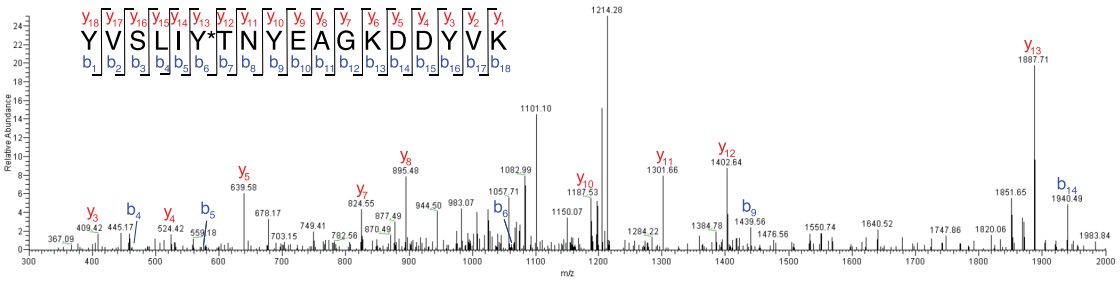


Figure 1-18. Fragmentation (MS/MS) spectra of the YVSLIY*NYEAGKDDYVK peptide of GSTP1 confirming that the LAS17 modification (+322.12 Da) is present on Y108, which is indicated with *.

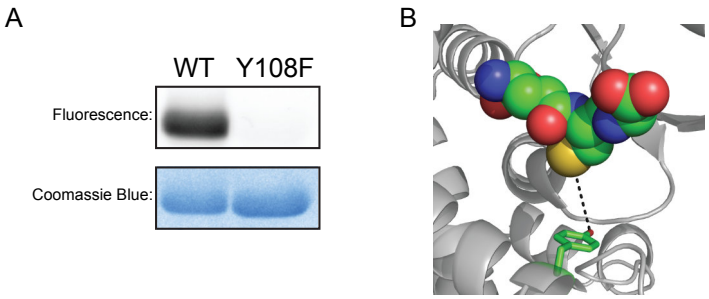


Figure 1-19. (A) In-gel fluorescence analysis of purified Y108F mutant compared to WT GSTP1, confirms Y108 as the site of labeling (B) Proximity of Y108 to the GSH binding site of GSTP1 (PDB: 6GSS).

Since the dichlorotriazine electrophile was previously shown to covalently modify lysine residues in proteomes, we sought to confirm LAS17 was not modifying a lysine residue on GSTP1. This was achieved by generating lysine to alanine mutants at each of the 12 lysine residues on GSTP1. Mutation of these lysine residues did not result in a decrease in LAS17 labeling compared to the WT control, further confirming that covalent modification was occurring at Y108 (Figure 1-20A). Additionally, a C47S mutant was

generated, which did not show loss of fluorescence compared to the WT control (Figure 1-20B).

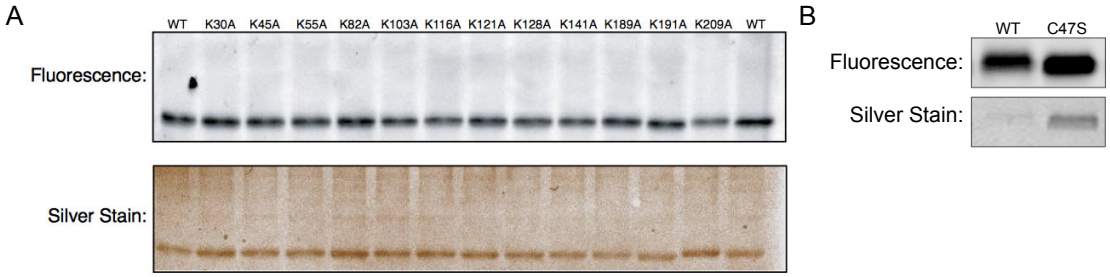


Figure 1-20. LAS17 labeling screening of (A) lysine and (B) cysteine mutants.

Lastly, activity assays with the Y108F mutant showed that LAS17 had negligible effect on the activity of the mutant protein in comparison to WT GSTP1 (Figure 1-21). These assay data thereby confirm that the inhibitor activity of LAS17 is solely due to the covalent modification of Y108. Residue This tyrosine residue has been found to be within hydrogen bond distance of the sulfur of GSH. This tyrosine residue favors formation and participates in stabilization of the thiolate.[114]

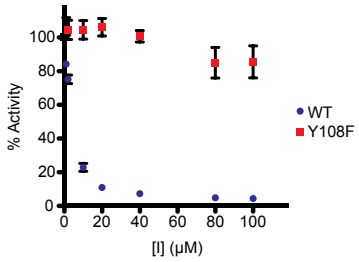


Figure 1-21. The Y108F mutant is resistant to LAS17-mediated inhibition compared to the wild-type GSTP1.

GSTP1 is a Metabolic Driver of Triple-Negative Breast Cancer

The work in the sections was performed by Louie, S.M., et al. Our contribution to the project was the synthesis of LAS17.

To identify TNBC specific metabolic drivers, an activity based chemoproteomic strategy was used to screen a panel of non-TNBC and TNBC cell lines. A lysine-reactive dichlorotriazine probe, RB7, and cysteine-reactive iodoacetamide alkyne were used to identify proteins with reactive functional residues. Lysine and cysteine residues were chosen in this study because they are nucleophilic amino acids that have been shown to contribute to enzymatic activity and function.[18,107]Probe protein targets from each of the cell lines were identified by mass spectrometry analysis of enriched tryptic digests. A filtering cut off criteria of a >5-fold increase in spectral counts following data analysis revealed unique metabolic enzymes are upregulated in TNBC cells compared to non-TNBC cells (Figure 1-22A). GSTP1 was found to be the most upregulated protein target across TNBC cell lines compared to non-TNBC cell lines (Figure 1-22B). Additionally, GSTP1 was found to have higher expression levels in precancerous cells line, MCF10A, and TNBC cells compared to non-TNBC cells by Western blotting (Figure 1-22C). This finding suggests that GSTP1 is a TNBC specific therapeutic target.

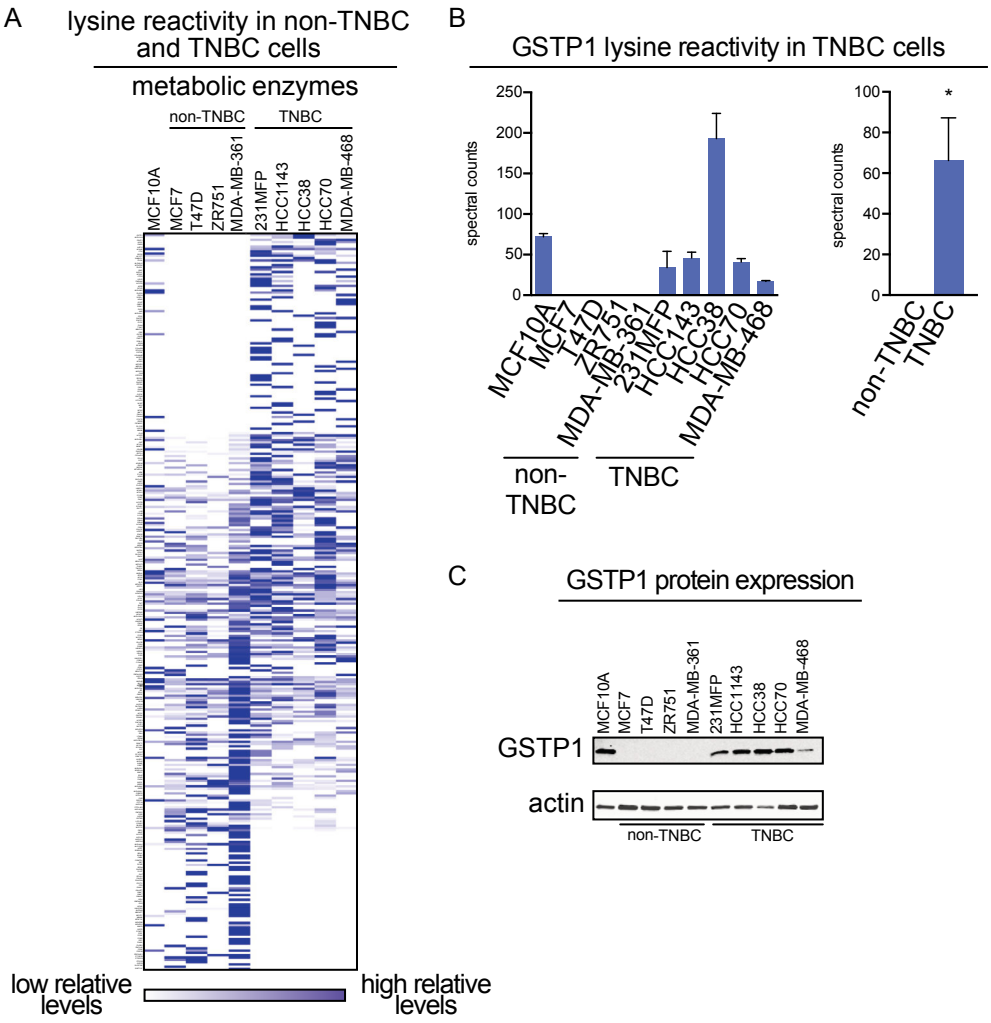


Figure 1-22. (A) Chemoproteomic profiling of a panel non-TNBC and TNBC cell lines with RB7. (B) GSTP1 reactivity in TNBC cells from each individual cell line and combined average spectral counts from non-TNBC and TNBC. (C) GSTP1 expression was measured across cell lines by Western blot. GSTP1 expression was normalized to actin loading control.

Effects of GSTP1 Inhibition on the Pathogenicity of Triple-Negative Breast Cancer

To confirm GSTP1 as a true therapeutic target for the treatment of TNBC, two GSTP1 knockdown 231MFP TNBC cell lines (shGSTP1-1 and shGSTP1-2) were generated. GSTP1 mRNA levels were assessed by qPCR and GSTP1 expression levels were assessed by Western blot (Figure 1-23A). These results show significant decreases in mRNA and protein levels, confirming successful GSTP1 knockdown. These knockdown cell lines showed impaired survival in serum-free media, but did not show decreased levels of cell proliferation compared to controls (Figure 1-23B). Furthermore, xenograft mouse studies, injected with shControl 231MFP and shGSTP1-1 231MFP knockdowns, showed decreased tumor growth rate for the shGSTP1-1 mouse compared to the shControl in immune-deficient mice (n=8) (Figure 1-23C). These results show that genetic inactivation of GSTP1 in TNBC results in decreased pathogenicity.

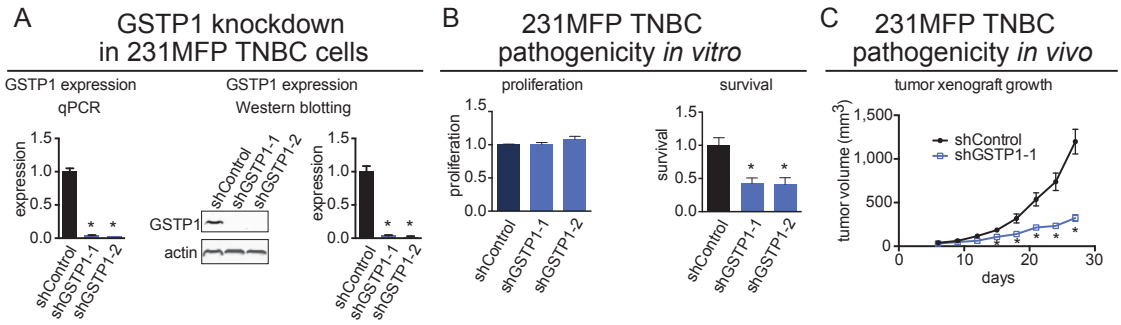


Figure 1-23. Effect of genetic inactivation of GSTP1 *in vitro* and *in vivo*. (A) Two independent shGSTP1 231MFP knockdown cell lines confirmed by qPCR and Western blotting. (B) GSTP1 231 MFP knockdown proliferation and survival studies. Assays were assessed by Hoechst stain. (C) Xenograft mouse models show impaired 231MFP tumor growth in shGSTP1-1 mice compared to shControl mice.

The results from the genetic studies motivated efforts to investigate if the same decrease in pathogenicity can be achieved by chemical inhibition of GSTP1. Thus, LAS17

was chosen to investigate the pharmacological potential of chemical inactivation of GSTP1 in TNBC. *In vitro* treatment of purified GSTP1 with LAS17 showed to have a 50% inhibitory concentration (IC_{50}) of 0.5 μM (Figure 1-24A). Consistent the results seen in shGSTP1, a survival assay shows significant decrease in cell survival in the LAS17 treated sample compared to the control (Figure 1-24B). Lastly, similar to the shGSTP1 xenograft studies, once daily treatment of immune-deficient mice with daily LAS17 (20 mg/kg) treatment starting at 2 or 16 days post subcutaneous injection with 231MFP cells shows decreased tumor growth rate in addition to decreased tumor size compared to the control (Figure 1-24C). Mice treated with LAS17 did not experience significant decreases in total body weight. LAS17 treatment of additional TNBC cell lines, HCC38, HCC70, and HCC1143 also resulted in significant decreases in cell survival (Figure 1-24).

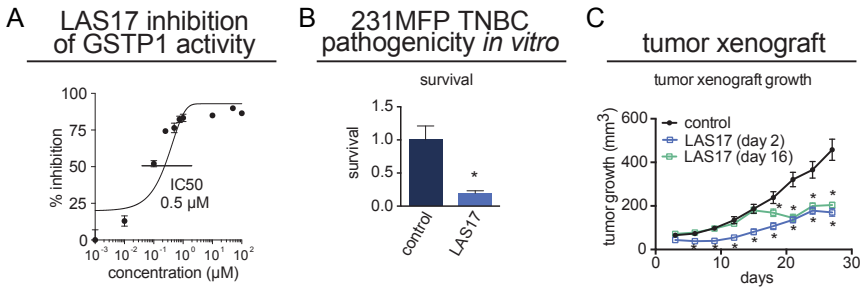


Figure 1-24. Effect of chemical inactivation of GSTP1 by LAS17 *in vitro* and *in vivo*. (A) LAS17 has an IC_{50} of 0.5 μM for the inhibition of GSTP1. (B) LAS17 treatment results in decreased cell survival. (C) Xenograft mouse models show impaired tumor growth in daily LAS17 treated mice 2 and 16 days post subcutaneous injection of 231MFP cells compared to vehicle treated mice.

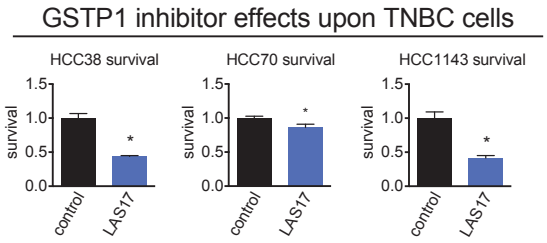


Figure 1-25. Survival assay of LAS17 treated TNBC cell lines.

Focus was then directed on elucidating the mechanism by which GSTP1 influences the pathogenicity of TNBC. GSTP1 is known to participate in protein-protein interactions as a modulator of signaling pathways. GSTP1 is known to interact with a member of the family of MAPK proteins, JNK. Additionally, GSTP1 has the ability to influence cellular levels of GSH, thus modulating the ratio of reduced and oxidized glutathione (GSH/GSSG), which is a hallmark of oxidative stress conditions. However, data suggests that neither of these are the mechanisms by which GSTP1 contributed to the pathogenicity of TNBC (data not shown).

These findings were further confirmed through metabolic profiling mapping experiments that revealed an alternative mechanism. Single-reaction monitoring (SRM)-based liquid chromatography-tandem MS (LC-MS/MS) method was used to monitor and quantify levels of ~200 metabolites spanning glycolytic, tricarboxylic acid (TCA) cycle, amino acid, nucleotide, and lipid metabolism. Additionally, an untargeted LC-MS method (m/z range set to 100-1,200) was used to identify metabolites that were altered upon GSTP1 inactivation. Through these methods the following metabolites were found to be lowered in the two shGSTP1 231MFP cell lines and with LAS17 treatment: lactic acid, ATP, nucleotides, diacylated phospholipids, and alkylacyl ether lipids. Conversely, the following metabolites were found to be increased in the two GSTP1 inactivation conditions: acyl

carnitines (ACs), ceramides, and lysophospholipids (Figure 1-26). Metabolite level differences in the two GSTP1 activation conditions were attributed to alterations that can be attributed to chronic GSTP1 inactivation (knockdown) versus acute GSTP1 inactivation (LAS17 treatment).

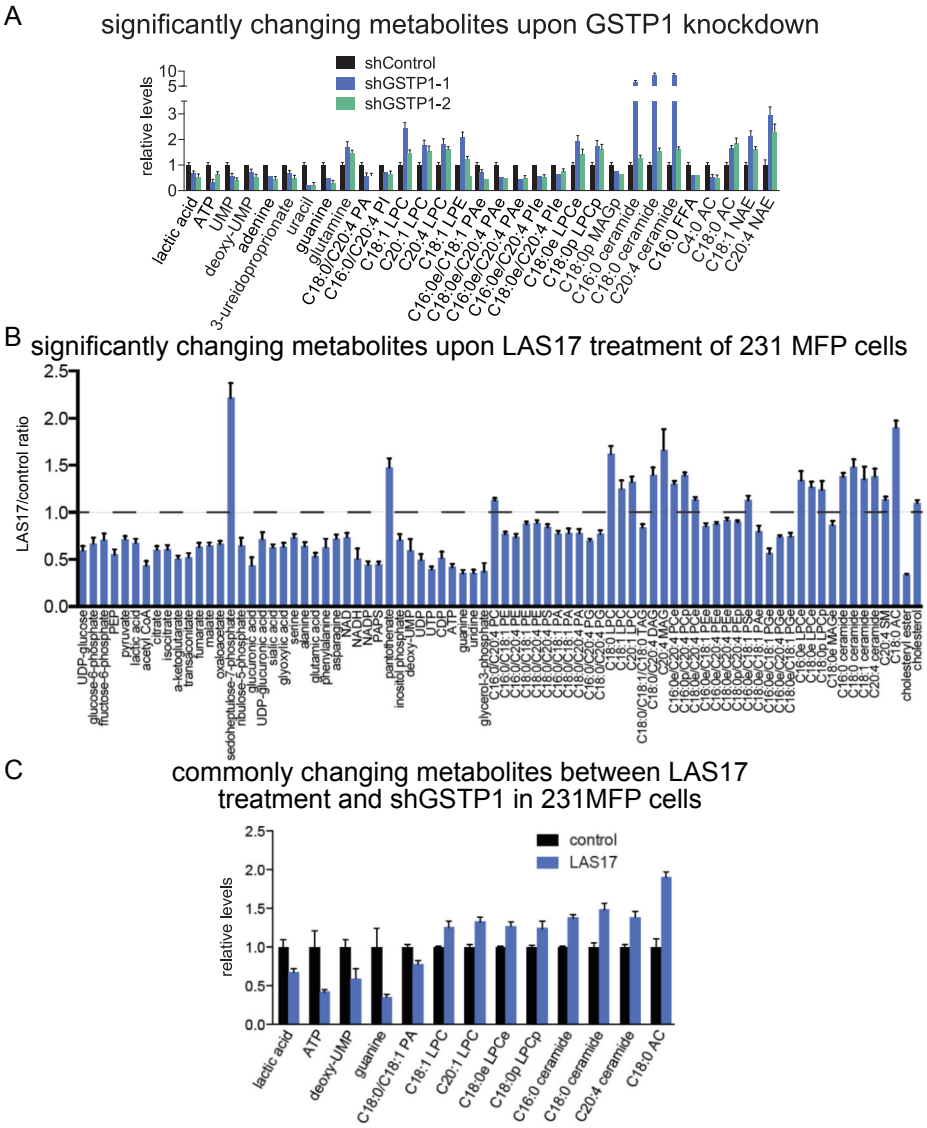


Figure 1-26. (A) Identification of metabolites that were significantly ($p < 0.05$) changed in the same direction in both shGSTP1-1 and shGSTP1-2 cells compared to shControl cells. (B) Identification of metabolites that were significantly ($p<0.05$) changed in levels in LAS17-treated (10 μM, 20 h) 231MFP cells compared to DMSO-vehicle-treated controls.

(C) Common metabolites that were significantly ($p<0.05$) changed between LAS17 treatment and shGSTP1 cells. Mapping of common observed metabolic changes to metabolic pathway maps suggested GSTP1 plays a role in the dysregulation of glycolysis, lipid metabolism, energetics, and oncogenic signaling pathways. To confirm these findings a number of the metabolite fluctuations were investigated independently.

Investigating the Influence of GSTP1 Activity in Glycolytic Metabolism

The observed decrease of lactic acid, in the metabolomics study resulting from GSTP1 inactivation, suggests a reduced glycolytic metabolism compared to controls. To confirm GSTP1 as a regulator of lactic acid levels, lactic acid secretion of 231MFP cells was monitored over time. The results show a significant decrease in the levels of lactic acid secreted by the GSTP1 inactivated, shGSTP1 and LAS17 treated, 231MFP cells compared to the controls, shControl and DMSO vehicle treated (Figure 1-27A). In an effort to identify the glycolytic enzymes influenced by GSTP1 activity, a stably overexpressing FLAG-tagged GSTP1 231MFP cell line was generated. FLAG pulldown on lysates from the FLAG-tagged-GSTP1 cell line enriched for seven proteins compared to the mock control (Figure 1-27B). Glycolytic enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was significantly enriched in the GSTP-FLAG expressing FLAG pulldown compared to the mock. However, endogenous GAPDH was not able to be detected using this method. Exogenous GADPH pulldown was observed upon the addition of pure and active GAPDH enzyme GSTP-FLAG expressing cell lysates. To investigate the effect of the protein-protein interaction between GSTP1 and GAPDH on GAPDH activity, GAPDH activity was monitored *in vitro*. GAPDH activity is greatly increased in the presence of GSTP1.

However, this enhancement is partially suppressed by LAS17 pre-treatment of GSTP1 (Figure 1-27C). Control experiments confirm the reduction in GAPDH activity is not a result of direct LAS17 inhibition of GAPDH (data not shown).

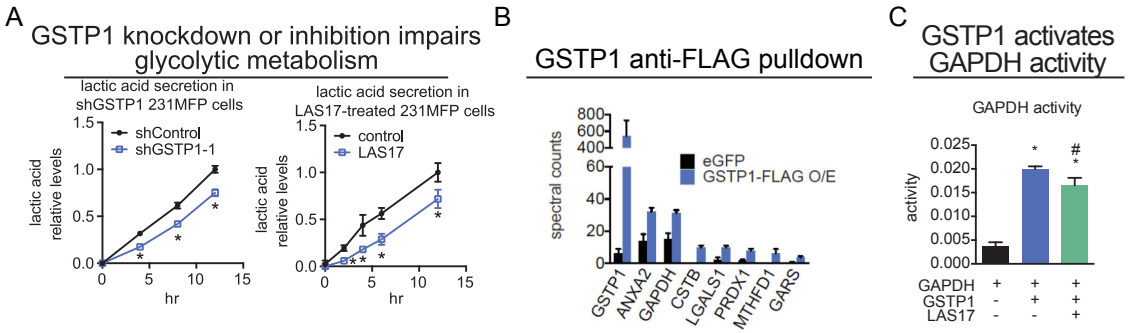


Figure 1-27. (A) Decreased levels of lactic acid secretion in media as determined by lactic assay kit. (B) Proteins enriched in FLAG pulldown proteomics experiment. (C) GSTP1 enhances the activity of GAPDH. LAS17 pretreated GSTP1 enhances GAPDH activity to a lesser degree.

GSTP1 Inhibition Impairs Oncogenic Signaling Pathways

Indication of reduced levels of ATP and increased levels of ACs (C18:0 AC) in the metabolomic profiling studies suggests enhanced phosphorylation and activity of AMP kinase (AMPK) and downstream phosphorylation and inhibition of A (CoA) carboxylase (ACC). Inhibition of mammalian target of rapamycin (mTOR) linked to the activation of AMPK has previously been demonstrated as a strategy to impair cancer pathogenicity.[115]To investigate the effect of GSTP1 inhibition on activation of AMPK and inactivation ACC, GSTP1 inhibited cells and controls were evaluated for AMPK and ACC phosphorylation in conjunction with AMPK and ACC expression levels by Western blotting (Figure 1-28A-B). These results further suggest that observed reduced levels of

ATP and increase levels of ACs can be correlated to activation of AMPK and thus impairment of oncogenic signaling. Furthermore, mTOR inactivation is further confirmed by the reduce levels of phosphorylated S6, a downstream target of mTOR, in GSTP1 inactivated cells compared to the controls (Figure 1-28C). Additionally, survival of 231MFP cells treated with mTOR inhibitor, Torin 1, showed a decrease in survival compared to the control (Figure 1-28D). Furthermore, co-treatment of LAS17 and Torin 1 did not show an additive effect on cell survival. This suggests LAS17 inhibits an essential target upstream of mTOR.

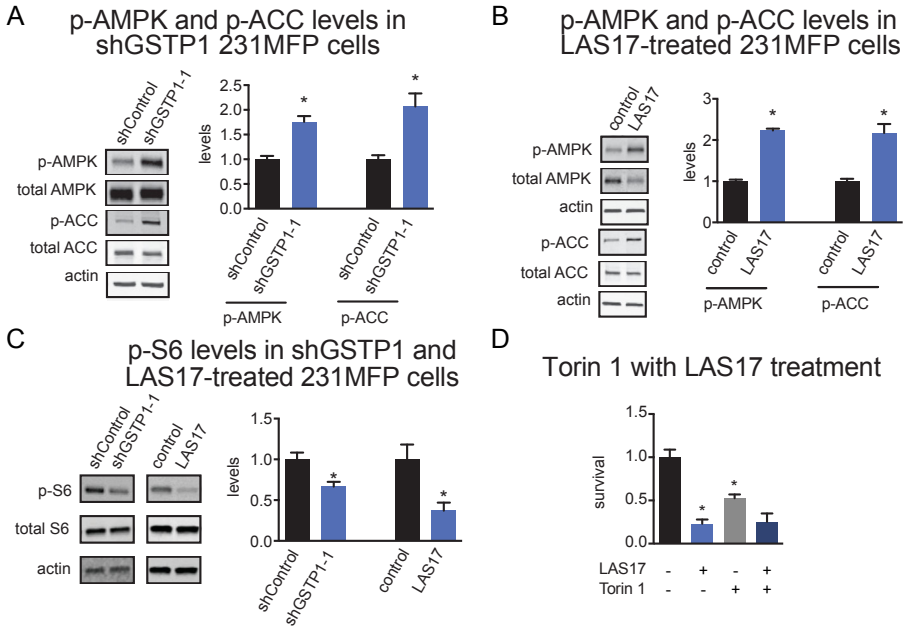


Figure 1-28. Effect of GSTP1 inhibition on oncogenic signaling pathways. (A) Genetic inhibition of GSPT1 results in increased amounts of p-AMPK and p-ACC compared to control cells. (B) Chemical inhibition of GSTP1 by LAS17 results in increased amounts of p-AMPK and p-ACC compaired to control cells. (D) mTOR inhibitor, Torin 1, and LAS17 do not have a synergistic effect on inhibiting cell survival.

Conclusions

In summary, we report the discovery and characterization of LAS17, a potent and selective tyrosine-directed irreversible inhibitor for GSTP1. Given the well-characterized role of GSTP1 in cancer pathogenesis and chemotherapeutic resistance, LAS17 provides a unique mode of irreversible GSPT1 inhibition, targeting a functional tyrosine residue (Y108), in contrast to previously reported cysteine-targeted compounds. Despite the high reactivity of the dichlorotriazine electrophile of LAS17, high selectivity is observed for GSTP1 within a complex proteome. The structural basis for this selectivity is currently unknown. Lastly, the presence of the terminal alkyne functionality for CuAAC provides the ability to determine target engagement and off-targets in cell-based and *in vivo* systems using CuAAC-mediated conjugation of reporter tags for visualization and enrichment of LAS17-bound proteins. LAS17 is a valuable addition to the currently available collection of reversible and irreversible GSTP1 inhibitors. LAS17 can be utilized as an activity-based probe (ABP) to interrogate GSTP1 activity in cancer, and as a pharmacological modulator of GSTP1 to further interrogate the therapeutic potential of this enzyme.

GSTP1 was revealed to be a unique, upregulated metabolic driver of TNBC in a global lysine reactivity profiling experiment. Genetic inactivation studies (knockdown) of GSTP1 in 231MFP, a TNBC cell line, revealed impaired glycolytic metabolism. This data was complemented by studies utilizing chemical inactivation of GSTP1 in 231MFP, made accessible, in part, by the use of the potent and selective GSTP1 inhibitor, LAS17. Metabolomics mapping experiments indicate reduced levels of macromolecular building blocks and ATP in the GSTP1 inactivated samples compared to controls, suggesting a role in the regulation of glycolysis and signaling. However, we recognize there may be other,

additional mechanism involved that are consistent with the metabolomic changes observed that remain unexplored in this study.

FLAG-GSTP1 pulldown proteomics experiments revealed glycolytic enzyme GAPDH as a protein-protein binding partner with GSTP1. In a study of GAPDH activity, we observed a dramatic increase in GAPDH activity in the presence of GSTP1. When GSTP1 is pretreated with LAS17, GAPDH activity is enhanced to a lesser degree. While we recognize the activation of GAPDH is only modestly reduced by LAS17 treatment, it is possible that in the context of a cell there are additional underlying mechanisms that are involved in GSTP1 correlated activation of GAPDH. GAPDH catalyzes the interconversion of glyceraldehyde-3-phosphate to 1,3-bisphosphoglycerate through the addition or removal of an inorganic phosphate group with one equivalent of NADH or NAD^+ in the reverse. GAPDH activity was shown to be insensitive to LAS17 treatment. Future interests lie in understanding and characterizing the mechanism by which GSTP1 interact is with GAPDH on the molecular level.

Additionally, we show inactivation of GSTP1 results in impaired oncogenic signaling as suggested by the metabolomic studies. GSTP1 inactivation leads to the increase of active p-AMPK which results in impaired breast cancer pathogenicity. Furthermore, we show GSTP1 inactivation leads to the increase of inactive p-ACC, consistent with the observed decrease in AC levels.

These data together suggest GSTP1 as a novel target, uniquely to TNBC over other forms of breast cancer. Our studies of GSTP1 inhibition, namely GSPT1 inactivation by LAS17, suggest GSTP1 inhibitors can be used as a stand-alone treatment for TNBC. These findings can be attributed to the power and usefulness of activity-based chemoproteomic

studies in conjunction with metabolomic studies to uncover underlying pathways and drivers in TNBC.

Acknowledgements

I would like to acknowledge Alex Shannon, Ranjan Banerjee, Daniel Bak, and Elizabeth Webster for their preliminary work that lead to the discovery of the dichlorotriazine as a proteome reactive electrophile. I would also like to acknowledge Alex Shannon, Daniel Bak, and Tyler Bechtel for their continued help maintaining the Orbitrap mass spectrometers.

Experimental Procedures

General

All reagents were purchased from Sigma Aldrich unless otherwise noted. All compounds were characterized by ^{1}H and ^{13}C NMR on either a Varian (Palo Alto, CA) 500 MHz or 600 MHz spectrometer. Chemical shifts (δ) are reported in parts per million (ppm) with chemical shifts referenced to internal standards: $CDCl_3$ (7.26 ppm for ^{1}H, 77.8 ppm for ^{13}C). Coupling constants (J) are reported in Hertz (Hz) and multiplicities are abbreviated as singlet (s), broad singlet (bs), doublet (d), triplet (t), pentet (p), multiplet (m), doublet of doublets (dd), and doublet of triplets (dt). High resolution mass spectra (HRMS) were obtained at the Mass Spectrometry Facility at Boston College (Chestnut Hill, MA). Analytical thin layer chromatography (TLC) was performed on Sorbent Technologies Silica G TLC Plates w/UV354 (0.25 mm). All compounds were visualized

on TLC by UV and/or $KMnO_4$ staining. Column chromatography was carried out using forced flow of indicated solvent on Sorbent Technology Standard Grade Silica Gel, 40-63 μm particle size, 60 Å pore size (Sorbent Technologies). PBS buffer, DMEM/High glucose media, and penicillin streptomycin (Pen/Strep) were purchased from Thermo Scientific (Waltham, MA). Primers were ordered from Eurofins MWG Operon (Hunstville, AL), and sequencing was performed by Genewiz (Cambridge, MA).

Synthesis of 1-methyl-4-(pent-4-yn-1-ylsulfonyl)benzene (2)

To an oven dried flask equipped with a stir bar, 4-pentyn-1-ol (11.8 mmol, 1.0 eq.), triethylamine (TEA) (23 mmol, 1.95 eq.), and dichloromethane (DCM) (33 mL) was added under N_2. The resulting solution was cooled to 0 ° C using an ice bath. Next, the 4-toluenesulfonyl chloride (12.5 mmol, 1.06 eq.) was added in one portion. The reaction vessel was purged with N_2 and allowed to slowly warm to 22 ° C. The reaction was quenched after 12 hrs with water (20 mL) and extracted with DCM (3 X 20 mL). The combined organic layers were dried over sodium sulfate (Na_2SO_4) and then concentrated *in vacuo*. The crude oil was purified by silica column chromatography (9:1, Hexanes (Hex):Ethylacetate (EtOAc)). The product was isolated as a clear oil (83 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.80 (d, J = 8.3 Hz, 2H), 7.35 (dd, J = 8.7, 0.7 Hz, 2H), 4.15 (t, J = 6.1 Hz, 2H), 2.45 (d, J = 0.7 Hz, 3H), 2.26 (td, J = 6.9, 2.7 Hz, 2H), 1.88 (t, J = 2.6 Hz, 1H), 1.88-1.83 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 144.9, 133.1, 130.0, 128.1, 82.24, 77.16, 69.56, 68.85, 27.87, 21.78, 14.84. HRMS m/z calculated for $C_{12}H_{14}O_2S$ $[M+H]^+$: 239.0664. Found: 239.0739.

Synthesis of methyl pent-4-yn-1-yl-L-leucinate (3)

To an oven dried flask equipped with a stir bar and reflux condenser, methyl L-leucinate hydrochloride (0.95 mmol 1.0 eq.), sodium iodide (0.47 mmol, 0.5 eq.), and potassium carbonate (2.37 mmol, 2.5 eq.) in acetonitrile (ACN) (1.6 mL) was added under N_2. The resulting mixture was allowed to heat to 90 °C and stirred for 1 hr before the drop-wise addition of tosylated alcohol (1) (1.04 mmol, 1.1 eq.). The reaction was allowed to mix for 12 hrs before being cooled and diluted with DCM. The suspension was filtered to remove solid precipitate. The remaining supernatant was concentrated *in vacuo*. The crude oil was purified by silica column chromatography (9:1–1:1 Hex:EtOAc) after dry loading. The product was isolated as a clear oil (36% yield). ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.72 (d, J = 1.4, 3H), 3.33 – 3.24 (m, 1H), 2.77 – 2.65 (m, 1H), 2.53 (dt, J = 12.71, 7.10, 1H), 2.26 (td, J = 7.5, 7.1, 2.4, 2H), 1.93 (dt, J = 4.2, 1.8, 1H), 1.79 – 1.61 (m, 3H), 1.46 (tt, J = 7.4, 1.7, 2H), 0.91 (ddd, J = 11.9, 6.6, 1.3 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 176.7, 84.20, 77.16, 68.57, 60.10, 51.72, 46.98, 42.97, 29.06, 25.07, 22.84, 22.46, 16.30.

Synthesis of methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-L-leucinate (LAS17) (4)

To an oven dried flask equipped with a stir bar was added, cyanuric chloride (0.211 mmol, 1 eq.) and tetrahydrofuran (THF) (6 mL), under N_2. Diisopropylethylamine (DIEA) (0.222 mmol, 1.05 eq.) and (2) (0.211 mmol, 1 eq.) were dissolved in THF (8 mL) and the resulting solution was added drop-wise to the cyanuric chloride solution. The reaction vessel was purged with N_2. The resulting mixture was allowed to cool to 0 °C in an ice bath. The reaction was allowed to warm to 22 °C over 12 hrs while stirring. The crude

product was purified by silica column chromatography (1:1 DCM:Hex–DCM). The product was isolated as a yellow-white solid (52% yield). ^{1}H NMR (600 MHz, $CDCl_3$) δ 5.16 (dd, *J* = 10.0, 4.8, 1H), 3.79 (ddd, *J* = 14.1, 10.5, 5.2, 1H), 3.73 (s, 3H), 3.41 (ddd, *J* = 14.1, 10.5, 5.4, 1H), 2.32 (dtd, *J* = 16.4, 6.7, 2.7, 1H), 2.25 (dddd, *J* = 17.0, 7.7, 6.2, 2.7, 1H), 2.01 (t, *J* = 2.7, 1H), 2.00 – 1.90 (m, 2H), 1.88 – 1.74 (m, 2H). ^{13}C NMR (151 MHz, $CDCl_3$) δ 171.4, 170.7, 170.1, 165.8, 83.28, 69.74, 58.55, 52.96, 46.45, 38.28, 30.07, 26.51, 25.21, 23.42, 22.09, 16.50. HRMS m/z calculated for $C_{15}H_{20}Cl_2N_4O_2$ $[M+H]^+$: 359.0963. Found 359.1047.

LAS1-LAS16, LAS18-LAS20 were synthesized following a similar general protocol as detailed here for LAS17.

Evaluating Protein Labeling in HeLa Cell Lysates

HeLa cells were grown in complete DMEM media with FBS, penicillin, and streptomycin on 20 cm cell culture plates. Cells were harvested and sonicated to lyse to form whole cell lysates. These lysates were separated by centrifugation at 45,000 rpm for 45 min at 4 °C to yield soluble and membrane proteins. The supernatant was collected and the pellet was discarded. Protein concentrations for soluble lysates were determined using a standard Bradford Assay (Bio-Rad DC Protein Assay). HeLa soluble protein lysates (50 μL, 2 mg/mL) were pretreated with probe (1 μM, 50X stock in DMSO) at RT for 1 hr. Samples then underwent click chemistry with TAMRA-azide (Lumiprobe, 25 μM, 50X stock in DMSO), TCEP (1 mM, 50X fresh stock in water), TBTA ligand (100 μM, 17X stock in DMSO:*t*-butanol = 1:4), and copper(II) sulfate (1 mM, 50X stock in water)

followed by incubation at RT for 1 hr. SDS-PAGE loading buffer 2X (reducing, 50 μL) was added to the samples and 25 μL of this solution was separated by SDS-PAGE at 100 volts for 130 minutes on a 10% polyacrylamide gel. Gels were visualized on a Hitachi FMBIO II multiview flatbed laser-induced fluorescent scanner. After analysis, gels underwent a typical procedure for Coomassie staining and destaining. Stained gels were visualized on a Stratagene Eagle Eye apparatus by COHU High performance CCD camera.

Protein Target Identification by LC/LC-MS/MS

HeLa soluble protein lysates in DPBS (pH 7.4) (500 μL, 2 mg/mL) were aliquoted and LAS17 (1 μM) or DMSO was added to the appropriate samples. Two aliquots were made for each inhibitor concentration or DMSO equaling 2 tubes for one final sample. All samples were then treated with biotin-azide (100 μM, 50X stock in DMSO), TCEP (1 mM, 50X fresh stock in water), TBTA ligand (100 μM, 17X stock in DMSO:*t*-butanol = 1:4), and copper(II) sulfate (1 mM, 50X stock in water) followed by incubation at 22 °C for 1 hr. Samples were combined pairwise and centrifuged (6500 g, 4 min, 4 °C) to pellet the precipitated proteins. The pellets were resuspended in cold methanol by sonication and the two samples were combined. Centrifugation was followed by a second cold methanol wash, after which the pellet was solubilized in DPBS containing 1.2% SDS via sonication and heating (90 °C, 5 min).

The SDS-solubilized proteome samples were diluted by 5 mL of DPBS for a final SDS concentration of 0.2%. The solution was incubated with 100 μL of streptavidin-agarose beads (Thermo Scientific, washed 3X with DPBS to remove storage buffer) overnight at 4 °C. Samples were rotated at 22 °C for 2 hr before being washed by 5 mL 0.2

% SDS/DPBS, 3 X 5 mL DPBS, and 3 X 5 mL water. The beads were pelleted by centrifugation (1400 X g, 3 min) between washes.

The washed beads were suspended in 500 uL of 6 M urea/DPBS and 10 mM DTT (from 20X stock in water) and placed in a 65 °C heat block for 15 min. Iodoacetamide (20 mM from 50X stock in water) was then added and the samples were allowed to react at 37 °C for 30 min while shaking. Following reduction and alkylation, the beads were pelleted by centrifugation and resuspended in 200 μL of 2 M urea/DPBS, 1 mM $CaCl_2$ (100X stock in water), and sequencing-grade trypsin (2 μg). The digestion was allowed to proceed overnight at 37 °C while shaking. The beads were pelleted by centrifugation and washed with 2 X 50 μL water. The washes were combined with the supernatant from the trypsin digestion step. Formic acid (15 μL) was added to the samples, which were stored at -20 °C until mass spectrometry analysis.

LC/LC-MS/MS analysis was performed on an LTQ-Orbitrap Discovery mass spectrometer (ThermoFisher) coupled to an Agilent 1200 series HPLC. Peptide digests were pressure loaded onto a 250 μm fused silica desalting column packed with 4 cm of Aqua C18 reverse phase resin (Phenomenex). The peptides were eluted onto a biphasic column (100 μm fused silica with a 5 μ, tip, packed with 10 cm C18 and 4 cm Partisphere strong cation exchange resin (SCX, Whatman) using a gradient 5-100% Buffer B in Buffer A (Buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; Buffer B: 20% water, 80% acetonitrile, 0.1% formic acid). The peptides were then eluted from the SCX onto the C18 resin and into the mass spectrometer using 4 salt steps previously described.[2] The flow rate through the column was set to ~0.25 μL/min and the spray voltage was set to 2.75 kV. One

full MS scan (FTMS) (400-1800 MW) was followed by 8 data dependent scans (ITMS) of the n^{th} most intense ions.

The tandem MS data were searched using the SEQUEST algorithm using a concatenated target/decoy variant of the human UniProt database. A static modification of +57.02146 on cysteine was specified to account for alkylation by iodoacetamide. SEQUEST output files were filtered using DTASelect.

Overexpression and Purification of GSTP1

The cDNA for WT-GSTP1 was subcloned into a pET-47b N-term His Tag expression vector using a forward primer (5'-CCAGGATCCGCCGCCCTACA-3') containing a BamH1 restriction site and a reverse primer (5'-AGCCTCGAGTCACTGTTTCCCG) containing a Xho1 restriction site. All constructs were verified by sequencing (Genewiz, Cambridge, MA). Constructs were transformed into BL21 competent *E. coli* (New England Biolabs). From an overnight LB culture with antibiotics at 37 °C, 5 mL were added to 500 mL LB (pH 7.0) with antibiotics and were grown to OD600 of 0.8. Protein expression was induced with IPTG (400 μM, 250X stock in water) for 5 hrs at 37 °C. Soluble cell lysates in DPBS (pH 7.4) were purified using Ni-NTA chromatography with imidazole concentrations of 25 mM and 500 mM in DPBS (pH 7.4) for the wash and elution steps, respectively. Purification fractions were analyzed for purity using SDS-PAGE. Imidazole was removed from pure protein fractions using NAP-5 desalting columns that had been buffer exchanged with DPBS.

LAS17 Labeling of Purified GSTP1

Purified GSTP1 (50 μL, 0.2 mg/mL) was pretreated with LAS17 (50 nM, 50X stock in DMSO) at 22 °C for 1 hr. Samples then underwent click chemistry with TAMRA-azide (Lumiprobe, 25 μM, 50X stock in DMSO), TCEP (1 mM, 50X fresh stock in water), TBTA ligand (100 μM, 17X stock in DMSO:*t*-butanol = 1:4), and copper(II) sulfate (1 mM, 50X stock in water) followed by incubation at 22 °C for 1 hr. SDS-PAGE loading buffer 2X (reducing, 50 μL) was added to the samples and 25 μL of this solution was separated by SDS-PAGE at 100 volts for 145 minutes on a 12% polyacrylamide gel. Gels were visualized on a Bio-Rad ChemiDoc MP Imaging System using the rhodamine setting. After analysis, gels underwent a typical procedure for Coomassie staining and destaining. Stained gels were visualized on a Bio-Rad ChemiDoc MP Imaging System.

Intact-protein MS Analysis

Purified GSTP1 (50 μL, 0.2 mg/mL) was pretreated with DMSO or probe (30 μM, 50X stock in DMSO) at 22 °C for 1 hr. Samples were centrifuged at 13,000 rpm for 5 min at RT to remove any insoluble particles. Samples were diluted by half in water and analyzed by LC/MS. 10 μL samples were injected onto a Aeris WIDEPORE 3.6u XB-C18 using an Agilent 1260 Infinity and were analyzed by an Agilent 6230 TOF Mass Spectrometer. Peak masses were extracted using Agilent MassHunter Qualitative Analysis B.06.00 software. Deconvolution software MagTran was used to determine the mass of the protein from the peak lists generated by mMass software.

Evaluating the Site of Modification

100 μg of GSTP1 in DPBS was treated with DMSO or LAS17 (10 μM, 50X stock in DMSO) for 1 hr at 22 °C. Protein was precipitated by addition of 100% trichloroacetic acid in PBS and incubated at -80 °C overnight. Thawed samples were centrifuged at 15K for 10 min at 22 °C and the supernatant was discarded. The remaining protein pellet was washed with 500 μL of cold acetone, vortexed to resuspend the pellet and centrifuged at 15K for 10 min at 22 °C. The supernatant was again discarded and the pellet was allowed to air dry until trace amounts of acetone were gone. The pellet was resuspended in 30 μL of 8M urea in PBS, then 70 μL 100 mM ammonium bicarbonate in PBS and 1.5 μL of 1 M DTT in PBS were added. Samples were incubated at 65 °C for 15 minutes. Samples were alkylated for 30 minutes at room temperature with the addition of 2.5 μL of 500 mM iodoacetamide in PBS. Sample volume was increase to 224 μL by addition of 120 μL PBS, then 2 μg of sequencing-grade trypsin (Promega) and 2.5 μL of 100 mM $CaCl_2$ was added. Samples were agitated over night at 37 °C. Then trypsin was quenched with 10 μL of formic acid (~5% of final volume) and were centrifuged at 15K for 20 minutes at room temperature to pellet undigested protein. Supernatant was transferred to a new tube and stored at -20 °C. Samples were analyzed by LC-MS/MS using a LTQ Orbitrap XL mass spectrometer (ThermoFisher) coupled to an EASY-nLC 1000 nanoLC (ThermoFisher). 10 μL of peptide digests were loaded onto 100 μm fused silica column with a 5 μm tip packed with 10 cm of Aqua C18 reverse phase resin (Phenomenex) using the EASY-nLC 1000 autosampler. The digests were eluted using a gradient 0-100% Buffer B in Buffer A (Buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; Buffer B; 20% water, 80% acetonitrile, 0.1% formic acid). The flow rate through the column was set to 400 nL/min and the spray

voltage was set to 3.5 kV. One full MS scan (FTMS) (400-1800 MW) was followed by 7 data dependent scans (ITMS) of the nth most intense ion with dynamic exclusion. The tandem MS data were searched using the SEQUEST algorithm using a concatenated target/decoy variant of the human IPI databases. A static modification of +57.02146 on cysteine was specified to account for iodoacetamide alkylation and differential modification of +322.12 (LAS17) were specified on tyrosine to account for probe modification. Modification searches on cysteine and lysine confirmed cysteine and lysine residues were not modified by LAS17. SEQUEST output files were filtered using DTASelect 2.0.

Verification of the Site of Labeling

Purified GSTP1 tyrosine mutants (50 μL, 0.2 mg/mL) or lysine mutants (50 μL, 0.5 mg/mL) were pretreated with LAS17 (50 μM, 50X stock in DMSO) at 22 °C for 1 hr. Samples then underwent click chemistry with TAMRA-azide (Lumiprobe, 25 μM, 50X stock in DMSO), TCEP (1 mM, 50X fresh stock in water), TBTA ligand (100 μM, 17X stock in DMSO:*t*-butanol = 1:4), and copper(II) sulfate (1 mM, 50X stock in water) followed by incubation at RT for 1 hr. SDS-PAGE loading buffer 2X (reducing, 50 μL) was added to the samples and 25 μL of this solution was separated by SDS-PAGE at 100 volts for 145 minutes on a 12% polyacrylamide gel. Gels were visualized on a Bio-Rad ChemiDoc MP Imaging System using the rhodamine setting. After analysis, gels underwent a typical procedure for Coomassie staining and destaining. Stained gels were visualized on a Bio-Rad ChemiDoc MP Imaging System.

GSTP1 Mutants

To verify the site of modification identified by mass spectrometry, tyrosine 108 was mutated to phenylalanine using an adopted QuikChange lightening MSDS protocol. (primer: 5'-CCCTCATCTTCACCAACTATGAGGCGGGCAAGGATGACTATGTGAAGGCACTGCC -3').

To confirm LAS17 was not modifying any of the lysine residues on GSTP1 we used the QuikChange Kit (Agilent Technologies, Santa Clara, California) to mutate each of the lysine residues to alanine.

Primers:

K30A Forward:

5'-GAGCTGGGCGGAGGAGGTGGTGACCGTGGAGACGTGGC-3'

K30A Reverse:

5'-CCTCCTCCGCCCAGCTCTGGCCCTGATCTGCCAGCAG-3'

K45A Forward:

5'-CTCACTCGCAGCCTCCTGCCTATACGGGCAGCTCCCC-3'

K45A Reverse:

5'-GAGGCTGCGAGTGAGCCCTCCTGCCACGTCTCCACG-3'

K55A Forward:

5'-GCTCCCCGCGTTCCAGGACGGAGACCTCACCCTG-3'

K55A Reverse:

5'-CCTGGAACGCGGGGAGCTGCCCGTATAGGCAGGAGG-3'

K82A Forward:

5’-CTATGGGGCGGACCAGCAGGAGGCAGCCCTGGTG-3’

K82A Reverse:

5’-GGCCGCACCCTTGGGCTCTATGGGGCGGACCAG-3’

K103A Forward:

5’-GGCGTGGAGGACCTCCGCTGCGCATACGTCTCCC-3’

K103A Reverse:

5’-GCCCGCCTCATAGTTGGTGTAGATGAGGGAGCGTATGCGCAGCG-3’

K116A Forward:

5’-GAGGCGGGCGCGCGGATGACTATGTGAAGGCACTGCCCGG-3’

K116A Reverse:

5’-

GTCATCCGCGCCCGCCTCATAGTTGGTGTAGATGAGGGAGACGTATTTGCAG

CGG-3’

K121A Forward:

5’-GACTATGTGGAGGCACTGCCCGGGCAACTGAAGCC-3’

K121A Reverse:

5’-

CAGTGCCTCCACATAGTCATCCTTGCCCGCCTCATAGTTGGTGTAGATGAGGG

-3’

K128A Forward:

5’-GCCCGGGCAACTGGCGCCTTTTGAGACCCTGC-3’

K128A Reverse:

5’-GCAGGGTCTCAAAAGGCGCCAGTTGCCCGGGC-3’

K141A Forward:

5’-CCAGGGAGGCGCGACCTTCATTGTGGGAGACCAGATCTCCTTCG-3’

K141A Reverse:

5’-GAAGGTCGCGCCTCCCTGGTTCTGGGACAGCAGGG-3’

K189A Forward:

5’-CGGCCCGCGCTCAAGGCCTTCCTGGCCTCCC-3’

K189A Reverse:

5’-CCTTGAGCGCGGGCCGGGCGCTGAGGCG-3’

K191A Forward:

5’-CCAAGCTCAAGGCCTTCCTGGCCTCCCCTGAGTACG-3’

K191A Reverse:

5’-GAAGGCCGCGAGCTTGGGCCGGGCACTGAGG-3’

K209A Forward:

5’-CCTCCCCATCAATGGCAACGGGGCACAGTGACTCGAGGC-3’

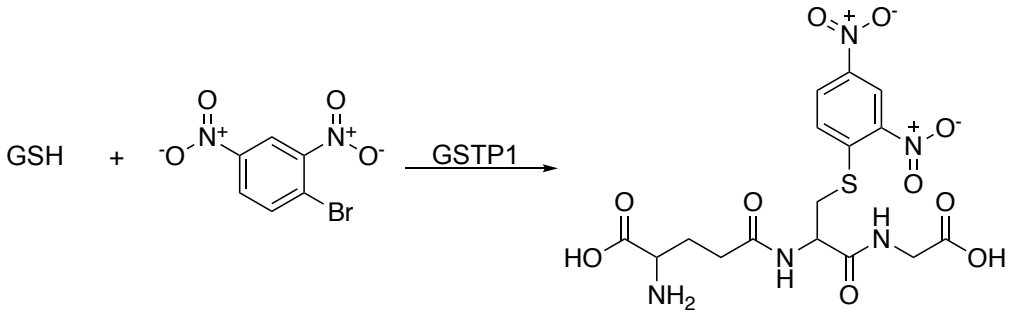
K209A Reverse:

5’-CGAGTCACTGTGCCCCGTTGCCATTGATGGGGAGGTTCACGTACTC-3’

GSTP1 Activity Assay

Purified recombinant GSTP1 was diluted to 0.02 mg/mL in Assay Buffer (100 mM NaH_2PO_4, pH 7.0). Samples were then incubated with LAS17 or DMSO for varying time points. A prepared stock of 250 mM L-Glutathione, reduced (GSH) substrate in water was diluted to 4 mM in Assay Buffer. Equal volumes of treated GSTP1 (25 μL) and GSH (25

μL, 4mM) were combined to afford final GSTP1 concentrations of 0.01 mg/mL and GSH concentrations of 2 mM . A substrate stock solution of 75 mM 1-bromo-2,4-dinitrobenzene (BDNB) in ethanol was diluted to 2 mM in Assay buffer. 50 μL of GSTP1/GSH mixture was aliquoted into wells of a 96 well plate (Costar, Catalog # 3595). The reaction (Scheme 1-2) was started by adding 50 μL of the 2 mM BDNB in Assay Buffer to each well. Reaction was monitored by kinetic mode for 5 minutes at an absorbance of 340 nm.



Scheme 1-2. Activity assay monitors the GSTP1 catalyzed conjugation of GSH to BDNB.

The following experimental procedures are from Louie, S.M., Grossman, E.A., Crawford, L.A., Ding, L., Camarda, R., Huffman, T.R., Miyamoto, D.K., Goga, A., Weerapana, E. and Nomura, D.K. 2016. GSTP1 Is a Driver of Triple-Negative Breast Cancer Cell Metabolism and Pathogenicity. *Cell Chemical Biology*. 23, 5, 567–578.

Cell Culture

The 231MFP cells were obtained from Prof. Benjamin Cravatt and were generated from explanted tumor xenografts of MDA-MB-231 cells. MCF7, MCF10A, T47D, ZR751, MDA-MB-361, HCC1143, HCC38, HCC70, MDA-MB-468, and HEK293T cells were obtained from the American Type Culture Collection. HEK293T cells were cultured in DMEM containing 10% (v/v) fetal bovine serum (FBS) and maintained at 37 C with 5%

CO_2. 231MFP, MDA-MB-361, and MDA-MB 468 cells were cultured in L15 medium containing 10% FBS and maintained at 37 C with 0% CO_2. MCF10A cells were cultured in DMEM/ F12K medium containing 5% horse serum, 20 ng/mL epidermal growth factor, 100 ng/mL cholera toxin, 10 ng/mL insulin, and 500 ng/mL hydrocortisone, and maintained at 37 C with 5% CO_2. MCF7, T47D, ZR751, HCC1143, HCC38, and HCC70 cells were cultured in RPMI medium containing 10% FBS and maintained at 37 C with 5% CO_2.

Lysine Reactivity Profiling

Cells were washed with PBS and harvested by scraping. 1 mg of protein was incubated with 10 μM dichlorotriazine-alkyne, RB7, for 30 min at room temperature. Following incubation, biotin-azide, TCEP, TBTA ligand, and Cu(II)SO_4 were added and incubated for an additional hour at room temperature. Following incubation, the mixture was subject to centrifugation at 6500 X g for 4 minutes. The supernatant was removed and 500 μL of ice cold methanol was added to each tube. Solubilized protein was centrifuged at 6500 X g for 4 minutes at 4 °C and supernatant was removed. Pellet was washed again with 500 μL of ice cold methanol and subject to centrifugation once more. The supernatant was removed and 1 mL of 1.2% SDS/PBS (w/v) was added until pellet was resolubilized. Mixture was heated at 90 °C for 5 minutes, then centrifuged for 5 minutes at 6500 X g. The 1 mL sample was then added to 5 mL of PBS. Avidin-agarose beads (Thermo Scientific) were washed with PBS and 170 μL of the bead slurry was added to each sample that was subsequently rotated overnight at 4 °C. Samples were then warmed on a rotator at room temperature the following day and subject to centrifugation at 1400 X g for 3 minutes. The

supernatant was removed and the beads were washed with 5 mL of 0.2% SDS/PBS (w/v), then subject to centrifugation at 1400 X g for 3 minutes. Beads were then transferred to Micro Bio-Spin columns (BioRad) and washed three times with PBS, followed by three times with water. The washed beads were then transferred to 500 µL of 6 M urea. 25 µL of 195 mM DTT was added to each sample, which was then incubated at 65 °C for 20 minutes. Tubes were cooled to room temperature before adding 25 µL of 400mM iodoacetamide solution, then incubated at 37 °C for 30 minutes. The reaction was diluted by adding 950 µL of PBS, then centrifuged at 1400 X g for 2 minutes, after which the supernatant was removed. 200 µL of 2 M urea and 0.5 µg/µL of Sequencing Grade Trypsin (Promega) were added and samples were incubated overnight at 37°C. Beads were centrifuged at 1400 x g for 3 minutes. Supernatant was added into a Micro Bio-Spin column to elute tryptic peptides while filtering out beads. Samples were acidified and stored at -80 °C until they were ready for proteomic analysis. Tryptic digests were analyzed using a Thermo LTQ-XL and quantified by spectral counting.

Evaluation of GSTP1 Expression in Pre-cancerous, Non-TNBC, and TNBC Cells

qPCR was performed using the manufacturer's protocol for Fisher Maxima SYBR Green with 10 mM primer concentrations or for Bio-Rad SsoAdvanced Universal Probes Supermix. Primer sequences for Fisher Maxima SYBR Green were derived from Primer Bank. Primer sequences for Bio-Rad SsoAdvanced Universal Probes Supermix were designed with Primer 3 Plus.

General Western Blotting Protocol

E-cadherin antibody was obtained from BD Biosciences. Vimentin antibody was obtained from Sigma. Antibodies to cyclophilin, GSTP1, b-actin, phospho-AMPK a (Thr172), AMPK a, phospho-ACC (Ser79), ACC, phospho-S6, total S6, phospho-JNK (Thr183/Tyr185), JNK, and GAPDH were obtained from Cell Signaling Technology. FLAG antibody was obtained from Cayman Chemicals.

Cells were lysed in lysis buffer (CST) containing both protease and phosphatase inhibitors. Proteins were resolved by electrophoresis on 4%–15% Tris-glycine precast Mini-PROTEAN TGX gel (Bio-Rad) and transferred to nitrocellulose membranes using the iBlot system (Invitrogen). Blots were blocked with 5% nonfat milk in Tris-buffered saline containing Tween 20 (TBST) solution for 1 hr at room temperature, washed in TBST, and probed with primary antibody diluted in recommended diluent per manufacturer over- night at 4 °C. Following washes with TBST, the blots were incubated in the dark with an IR-linked secondary antibody at room temperature for 1 hr. Blots were visualized using an Odyssey Li-Cor scanner after additional washes.

Constructing GSTP1 Knockdown Cell Lines

We used two independent short-hairpin oligonucleotides to knock down the expression of GSTP1 and one short-hairpin oligonucleotide to knock down the expression of CDH1 using previously described methods.[116]For generation of stable shRNA lines, lentiviral plasmids in the pLKO.1 backbone containing shRNA (Sigma) against human GSTP1 were transfected into HEK293T cells using Fugene (Roche). Lentivirus was collected from filtered cultured medium and used to infect the target cancer cell line with

Polybrene. Target cells were selected over 3 days with 1 mg/mL puromycin. The short-hairpin sequences used for generation of the GSTP1 knockdown lines were:

shGSTP1-1,

CCGGCGCTGACTACAACCTGCTGGACTCGAGTCCAGCAGGTTGTAGTCAGCG

TTTTTG;

shGSTP1-2,

CCGGCCTCACCCTGTACCAGTCCAACTCGAGTTGGACTGGTACAGGGTGAGG

TTTTTG.

The short-hairpin sequence used for generation of the CDH1 knockdown line was CCGGAGAAGGGTCTGTTCACGTATTCTCGAGAATACGTGAACAGA

CCCTTCTTTTTTG.

The control shRNA was targeted against GFP with the target sequence GCAAGCTGACCCTGAAGTTCAT. Knockdown was confirmed by qPCR or western blotting.

Cellular Phenotype Assays

Cell proliferation, serum-free cell survival, and migration assays were performed in a manner similar to previously described assays.[116,117]Migration assays were performed using Transwell chambers (Corning) with 8μm pore-sized membranes coated with collagen. 50,000 cells were seeded into the top chamber and allowed to migrate for 24 hours. Chambers were fixed with Diff-Quik (Dade Behring) solutions.

Cell survival and proliferation assays were performed using the WST-1 reagent (Roche) for the MCF7 cell line. Cells were seeded in a 96-well plate with a volume of 200

μL serum-free media for survival and serum-containing media for proliferation (20,000 cells for survival and 10,000 cells for proliferation). WST-1 reagent was added to each well and incubated at 37 °C for 30 minutes before measuring absorbance. Cell survival measures cell death under serum-free and non-proliferative conditions, whereas cell proliferation assays are performed in the presence of serum and measures proliferation of cells.

Cell survival and proliferation assays for the 231MFP cells were performed using Hoechst 33342 dye (Invitrogen) according to manufacturer's protocol. Briefly, cells were seeded into 96-well plates (40,000 for survival and 20,000 for proliferation) in a volume of 200 μL and allowed to adhere overnight. Medium was removed from each well and 100 μL of staining solution containing 10% formalin and Hoechst 33342 dye was added to each well and incubated for 15 minutes in the dark at room temperature. After incubation, staining solution was removed and wells were washed twice with PBS before imaging.

Cell survival assays were also performed using Hoechst 33342 dye for the HCC38, HCC70, and HCC1143 cell lines with the exception of HCC38 shGSTP1 cells that were performed using the WST-1 reagent. Cell survival assays using Hoechst 33342 were performed as described above. Cell survival assay for HCC38 shGSTP1 cells was performed with WST-1 reagent as described above.

Tumor Xenograft Studies

Human tumor xenografts were established by transplanting cancer cells ectopically into the flank of C.B17 severe combined immunodeficiency (SCID) mice (Taconic Farms) as previously described. [118] In brief, cells were washed twice with PBS, trypsinized, and harvested in serum-containing medium. Harvested cells were washed twice with serum-

free medium and resuspended at a concentration of 2.0 X 10^4 cells/mL, and 100 mL was injected. Tumors were measured every 2 days with calipers. Animal experiments were conducted in accordance with the guidelines of the Institutional Animal Care and Use Committees of the University of California, Berkeley.

GSTP1 Activity Assay

The IC_{50} of LAS17 was determined using a GSTP1 activity assay. 1-Chloro-2,4-dinitrobenzene (CDNB) was incubated with 200 mM reduced glutathione and 1 mg of active GSTP1 protein with 100, 10, 1, 0.1, 0.01, 0.001, and 0 mM LAS17. GSTP1 catalyzes the conjugation of glutathione to CDNB, generating the reaction product glutathione-DNB conjugate, which absorbs at 340 nm. The rate of increase in the absorption of the product is proportional to GSTP1 activity and was used to measure GSTP1 activity.

Metabolomic Profiling of Cancer Cells

Metabolomic analyses were conducted using previously described methods.[116,117,119–121]Briefly, 2 million cells were plated overnight, serum starved for 2 hours prior to harvesting, after which cells were washed twice with PBS, harvested by scraping, and flash frozen.

For nonpolar metabolomic analyses, flash frozen cell pellets were extracted in 4 mL of 2:1:1 chloroform/methanol/PBS with internal standards dodecylglycerol (10 nmoles) and pentadecanoic acid (10 nmoles). Organic and aqueous layers were separated by centrifugation, and organic layer was extracted. Aqueous layer was acidified with 0.1% formic acid followed by re-extraction with 2 mL chloroform. The second organic layer was

combined with the first extract and dried under nitrogen, after which lipids were resuspended in chloroform (120 μL). A 10 μL aliquot was then analyzed by both single-reaction monitoring (SRM)-based LC-MS/MS or untargeted LC-MS.

For polar metabolomic analyses, frozen cell pellets were extracted in 180 μL of 40:40:20 acetonitrile/methanol/water with internal standard d_3 N^{15}- serine (1 nmole). Following vortexing and bath sonication, the polar metabolite fraction (supernatant) was isolated by centrifugation. A 20 μL aliquot was then analyzed by both single-reaction monitoring (SRM)-based LC-MS/MS or untargeted LC-MS.

For the SRM transitions where we monitor the transition of parent masses to the loss of the headgroup (e.g. loss of phosphocholine from phosphatidylcholine), we have ascertained the acyl chain specificities from previously described procedures.[121]For phospholipids such as PCs and PEs, we ascertained fatty acid acyl chain composition from phospholipids using a mobile phase containing both ammonium hydroxide and formic acid and monitored the fatty acid fragmentations from [M-H+HCO_2H] m/z at 40 V collision energy in negative ionization mode. For other phospholipids such as PAs and PIs, we monitored the fatty acid fragmentations from [M-H] m/z at 40 V collision energy in negative ionization mode in mobile phase containing just ammonium hydroxide. For the lipids that we have measured in this study, the designated acyl chains represent the primary fatty acids that were on the lipid backbone. However, this method is less sensitive than monitoring the loss of headgroup from the phospholipid, and thus we used SRM transitions for many phospholipids where we monitored for loss of headgroups (e.g. PCs, PEs, PSs, PAs, PIs).

Relative levels of metabolites were quantified by integrating the area under the curve for each metabolite, normalizing to internal standard values, and then normalizing to the average values of the control groups.

Lactic Acid Secretion

Lactic acid secretion from L-15 medium was measured by collecting medium and performing a colorimetric lactic acid assay using a kit purchased from Abcam in accordance with the manufacturer's protocol.

Constructing FLAG-Tagged GSTP1 Cells

GSTP1 cDNA (Dharmacon) was subcloned into the pENTR4-FLAG vector (Addgene). This entry vector was recombined via an attL-attR (LR) reaction into a pLenti CMV puro DEST vector (Addgene). For generation of the FLAG- tagged GSTP1 line, the lentiviral plasmid containing FLAG-GSTP1 was transfected into HEK293T cells using Fugene (Roche). Lentivirus was collected from filtered cultured medium and used to infect the target cancer cell line with Polybrene. Target cells were selected over 3 days with 1 mg/mL puromycin.

Anti-FLAG Pulldown Studies

Pulldown studies were performed used Anti-FLAG M2 magnetic beads (Sigma) according to the manufacturer's protocol. FLAG-tagged GSTP1-overexpressing and GFP-overexpressing control cells were lysed in lysis buffer (CST), and 500 mg of lysate was incubated with 32 mL of anti-FLAG magnetic beads for 2 hr at 4 °C. Beads were collected

and washed three times with Tris-buffered saline before elution with 4% SDS (w/v) in 120 mM Tris-HCl. Samples were heated to 95 °C for 3 min. Eluent was subsequently prepared for proteomic profiling with a shotgun proteomic analysis protocol as described below.

Shotgun Proteomic Profiling of Anti-FLAG Pulldown

Pulldown products were precipitated in 20% trichloroacetic acid at 80 C overnight and centrifuged at 10,000 X g at 4 °C for 10 min to pellet protein. Pelleted proteins were washed three times with 8 M urea in PBS. After solubilization, 30 ml of 0.2% ProteaseMAX Surfactant (Promega) was added and the resulting mixture was vortexed followed by the addition of 40 mL of 100 mM ammonium bicarbonate and 10 mM tris(2-carboxyethyl)phosphine (TCEP). After 30 min, 12.5 mM iodoacetamide was added and allowed to react for 30 min in the dark before adding 120 mL of PBS and 1.2 mL of 1% ProteaseMAX Surfactant. The protein solution was vortexed, and 0.5 mg/mL sequencing-grade trypsin (Promega) was added and allowed to react overnight at 37 °C. The peptide solution was then centrifuged at 10,000 X g before the supernatant was subsequently analyzed by LC-MS/MS.

GAPDH Activity Assay

GAPDH activity was measured using a colorimetric kit purchased from BioVision and performed according to the manufacturer's protocol. Active human GSTP1 and active human GAPDH full-length proteins were purchased from Abcam. Proteins were co-incubated at 37 °C for 1 hr before GAPDH activity was measured using the kit.

References

(1) Okayasu, T; Ikeda, M; Akimoto, K; Sorimachi, K. The Amino Acid Composition of Mammalian and Bacterial Cells. *Amino Acids* 1997, *13* (3), 379–391.

(2) Udenfriend, S.; Cooper, J. R. The Enzymatic Conversion of Phenylalanine to Tyrosine. *The Journal of biological chemistry* 1952, *194* (2), 503–11.

(3) Fitzpatrick, P. Tetrahydropterin-dependent amino acid hydroxylases. *Annual Reviews Biochemistry* 1999, *68* (1), 355–381.

(4) Maeda, H.; Dudareva, N. The Shikimate Pathway and Aromatic Amino Acid Biosynthesis in Plants. *Annual Review Plant Biology* 2012, *63*, 73–105.

(5) McGaughey, G. B.; Gagné, M.; Rappé, A. K. Pi-Stacking Interactions. Alive and Well in Proteins. *Journal Biological Chemistry* 1998, *273* (25), 15458–63.

(6) Oktaviani, N. A.; Pool, T. J.; Kamikubo, H.; Slager, J.; Scheek, R. M.; Kataoka, M.; Mulder, F. A. A. Comprehensive Determination of Protein Tyrosine pKa Values for Photoactive Yellow Protein Using Indirect 13C NMR Spectroscopy. *Biophysical Journal.* 2012, *102* (3), 579–86.

(7) Bartlett, G. J.; Porter, C. T.; Borkakoti, N.; Thornton, J. M. Analysis of Catalytic Residues in Enzyme Active Sites. *Journal Molecular Biology* 2002, *324* (1), 105–21.

(8) Aktas, D.; Cook, P. A Lysine-Tyrosine Pair Carries Out Acid−Base Chemistry in the Metal Ion-Dependent Pyridine Dinucleotide-Linked B-Hydroxyacid Oxidative Decarboxylases†. *Biochemistry-us* 2009, *48* (16), 3565–3577.

(9) Breslow, R.; Chin, J.; Hilvert, D.; Trainor, G. Evidence for the General Base Mechanism in Carboxypeptidase A-Catalyzed Reactions: Partitioning Studies on

Nucleophiles and H2(18)O Kinetic Isotope Effects. *Proceedings of the National Academy Sciences U.S.A.* 1983, *80* (14), 4585–9.

(10) Jones, L. H.; Narayanan, A.; Hett, E. C. Understanding and Applying Tyrosine Biochemical Diversity. *Molecular Biosystems* 2014, *10* (5), 952–69.

(11) Champoux, J. J. DNA Topoisomerases: Structure, Function, and Mechanism. *Annual Review of Biochemistry* 2001, *70*, 369–413.

(12) Hunter, T. THE CROONIAN LECTURE 1997. The Phosphorylation of Proteins on Tyrosine: Its Role in Cell Growth and Disease. *Philosophical Transactions Royal Society Lond B Biological Sciences* 1998, *353* (1368), 583–605.

(13) Watts, A.; Oppezzo, P.; Withers, S.; Alzari, P.; Buschiazzo, A. Structural and Kinetic Analysis of Two Covalent Sialosyl-Enzyme Intermediates on Trypanosoma Rangeli Sialidase. *Journal of Biological Chemistry* 2006, *281* (7), 4149–4155.

(14) Kalia, J.; Raines, R. T. Advances in Bioconjugation. *Current organic chemistry* 2010, *14* (2), 138–147.

(15) Jares-Erijman, E. A.; Jovin, T. M. FRET Imaging. *Nature Biotechnology* 2003, *21* (11), 1387–95.

(16) Beck, A.; Goetsch, L.; Dumontet, C.; Corvaïa, N. Strategies and Challenges for the next Generation of Antibody-Drug Conjugates. *Nature Reviews Drug Discovery* 2017, *16* (5), 315–337.

(17) Adam. Chemical Strategies for Functional Proteomics. *Molecular & Cellular Proteomics* 2002, *1* (10), 781–790.

(18) Shannon, D. A.; Weerapana, E. Covalent Protein Modification: The Current Landscape of Residue-Specific Electrophiles. *Current opinion in chemical biology* 2015,

24, 18–26.

(19) Bunnage, M.; Chekler, E.; Jones, L. Target Validation Using Chemical Probes. *Nature Chemical Biology* 2013, *9* (4), 195–199.

(20) Hett, E. C.; Xu, H.; Geoghegan, K. F.; Gopalsamy, A.; Kyne, R. E.; Menard, C. A.; Narayanan, A.; Parikh, M. D.; Liu, S.; Roberts, L.; Robinson, R. P.; Tones, M. A.; Jones, L. H. Rational Targeting of Active-Site Tyrosine Residues Using Sulfonyl Fluoride Probes. *ACS chemical biology* 2015, *10* (4), 1094–8.

(21) Gu, C.; Shannon, D. A.; Colby, T.; Wang, Z.; Shabab, M.; Kumari, S.; Villamor, J. G.; McLaughlin, C. J.; Weerapana, E.; Kaiser, M.; Cravatt, B. F.; Hoorn, R. A. van der. Chemical Proteomics with Sulfonyl Fluoride Probes Reveals Selective Labeling of Functional Tyrosines in Glutathione Transferases. *Chemical Biology* 2013, *20* (4), 541–8.

(22) Crawford; Weerapana. A Tyrosine-Reactive Irreversible Inhibitor for Glutathione S -Transferase Pi (GSTP1). *Molecular Biosystems* 2016, *12* (6), 1768–1771.

(23) Curreli, N; Oliva, S; Rescigno, A; Rinaldi, AC; Sollai, F; Sanjust, E. Novel Diazonium-Functionalized Support for Immobilization Experiments. *Journal of applied polymer science* 1997, *66* (8), 1433–1438.

(24) Jones, M. W.; Mantovani, G.; Blindauer, C. A.; Ryan, S. M.; Wang, X.; Brayden, D. J.; Haddleton, D. M. Direct Peptide Bioconjugation/PEGylation at Tyrosine with Linear and Branched Polymeric Diazonium Salts. *Journal of the American Chemical Society* 2012, *134* (17), 7406–7413.

(25) Hooker, J. M.; Kovacs, E. W.; Francis, M. B. Interior Surface Modification of Bacteriophage MS2. *Journal of the American Chemical Society* 2004, *126* (12), 3718–9.

(26) Li, K.; Chen, Y.; Li, S.; Nguyen, H.; Niu, Z.; You, S.; Mello, C.; Lu, X.; Wang, Q.

Chemical Modification of M13 Bacteriophage and Its Application in Cancer Cell Imaging. *Bioconjugate Chemistry* 2010, *21* (7), 1369–77.

(27) Schlick, T. L.; Ding, Z.; Kovacs, E. W.; Francis, M. B. Dual-Surface Modification of the Tobacco Mosaic Virus. *Journal of the American Chemical Society* 2005, *127* (11), 3718–23.

(28) Gavrilyuk, J.; Ban, H.; Uehara, H.; Sirk, S.; Saye-Francisco, K.; Cuevas, A.; Zablowsky, E.; Oza, A.; Seaman, M.; Burton, D.; Barbas, C. Antibody Conjugation Approach Enhances Breadth and Potency of Neutralization of Anti-HIV-1 Antibodies and CD4-IgG. *Journal of Virology* 2013, *87* (9), 4985–4993.

(29) Jensen, S.; Kimani, F.; Jewett, J. Light-Activated Triazabutadienes for the Modification of a Viral Surface. *ChemBioChem* 2016, *17* (23), 2216–2219.

(30) Cohen, L. A. Group-Specific Reagents in Protein Chemistry. *Annual Reviews of Biochemistry* 1968, *37*, 695–726.

(31) Riordan, J. F.; Wacker, W. E.; Vallee, B. L. N-Acetylimidazole: A Reagent for Determination of “Free” Tyrosyl Residues of Proteins. *Biochemistry* 1965, *4* (9), 1758–1765.

(32) Peerce, B. E.; Wright, E. M. Evidence for Tyrosyl Residues at the Na+ Site on the Intestinal Na+/glucose Cotransporter. *Journal of Biological Chemistry* 1985, *260* (10), 6026–31.

(33) Liao, T.-H.; Ting, R. S.; Yeung, J. E. Reactivity of Tyrosine in Bovine Pancreatic Deoxyribonuclease with P-Nitrobenzenesulfonyl Fluoride. *Journal of Biological Chemistry* 1982, *257* (10), 5637–5644.

(34) Poulos, TL; Price, PA. Some Effects of Calcium Ions on the Structure of Bovine

Pancreatic Deoxyribonuclease A. *Journal of Biological Chemistry* 1972.

(35) Lundblad, RL. Chemical Reagents for Protein Modification. 2014.

(36) Narayanan, A.; Jones, L. Sulfonyl Fluorides as Privileged Warheads in Chemical Biology. *Chemical Science* 2015, *6* (5), 2650–2659.

(37) Espuña, G.; Arsequell, G.; Valencia, G.; Barluenga, J.; Pérez, M.; González, J. M. Control of the Iodination Reaction on Activated Aromatic Residues in Peptides. *Chemical Communications* 2000, *0* (14), 1307–1308.

(38) Espuña, G.; Andreu, D.; Barluenga, J.; Pérez, X.; Planas, A.; Arsequell, G.; Valencia, G. Iodination of Proteins by IPy2BF4, a New Tool in Protein Chemistry. *Biochemistry-us* 2006, *45* (19), 5957–63.

(39) Bolton, A. E.; Hunter, W. M. The Labelling of Proteins to High Specific Radioactivities by Conjugation to a 125I-Containing Acylating Agent. *Biochemical Journal* 1973, *133* (3), 529–39.

(40) Vilaró, M.; Arsequell, G.; Valencia, G.; Ballesteros, A.; Barluenga, J. Arylation of Phe and Tyr Side Chains of Unprotected Peptides by a Suzuki−Miyaura Reaction in Water. *Organic Letters* 2008, *10* (15), 3243–3245.

(41) Schiller, P. W.; Yam, C. F.; Prosmanne, J. Synthesis, Opiate Receptor Affinity, and Conformational Parameters of [4-Tryptophan] Enkephalin Analogs. *Journal of medicinal chemistry* 1978, *21* (11), 1110–1116.

(42) Tilley, D.; Francis, M. Tyrosine-Selective Protein Alkylation Using Π-Allylpalladium Complexes. *Journal of the American Chemical Society* 2006, *128* (4), 1080–1081.

(43) Chen, S.; Li, X.; Ma, H. New Approach for Local Structure Analysis of the

Tyrosine Domain in Proteins by Using a Site-Specific and Polarity-Sensitive Fluorescent Probe. *ChemBioChem* 2009, *10* (7), 1200–1207.

(44) Wilcox, P. E. Chymotrypsinogens—chymotrypsins. *Methods in enzymology* 1970, *19*, 64–108.

(45) Romanini, D.; Francis, M. Attachment of Peptide Building Blocks to Proteins Through Tyrosine Bioconjugation. *Bioconjugate Chem* 2008, *19* (1), 153–157.

(46) Fraenkel-Conrat, H.; Olcott, H. S. Reaction of Formaldehyde with Proteins VI. Cross-Linking of Amino Groups with Phenol, Imidazole, or Indole Groups. *Journal of Biological Chemistry* 1948, *174*, 827–843.

(47) Joshi, N. S.; Whitaker, L. R.; Francis, M. B. A Three-Component Mannich-Type Reaction for Selective Tyrosine Bioconjugation. *Journal of the American Chemical Society* 2004, *126* (49), 15942–3.

(48) McFarland, J.; Joshi, N.; Francis, M. Characterization of a Three-Component Coupling Reaction on Proteins by Isotopic Labeling and Nuclear Magnetic Resonance Spectroscopy. *Journal of the American Chemical Society* 2008, *130* (24), 7639–7644.

(49) Ban, H.; Gavrilyuk, J.; Barbas, F. Tyrosine Bioconjugation through Aqueous Ene-Type Reactions: A Click-Like Reaction for Tyrosine. *Journal of the American Chemical Society* 2010, *132* (5), 1523–1525.

(50) Ban, H.; Nagano, M.; Gavrilyuk, J.; Hakamata, W.; Inokuma, T.; Barbas, C. Facile and Stabile Linkages through Tyrosine: Bioconjugation Strategies with the Tyrosine-Click Reaction. *Bioconjugate Chemistry* 2013, *24* (4), 520–532.

(51) Sato, S.; Nakamura, K.; Nakamura, H. Tyrosine-Specific Chemical Modification with in Situ Hemin-Activated Luminol Derivatives. *ACS Chemical Biology* 2015, *10* (11),

2633–2640.

(52) Nilo, A.; Allan, M.; Brogioni, B.; Proietti, D.; Cattaneo, V.; Crotti, S.; Sokup, S.; Zhai, H.; Margarit, I.; Berti, F.; Hu, Q.-Y.; Adamo, R. Tyrosine-Directed Conjugation of Large Glycans to Proteins via Copper-Free Click Chemistry. *Bioconjugate Chemistry* 2014, *25* (12), 2105–2111.

(53) Desimoni; Faita; Righetti; Sfulcini; Tsyganov. Solvent Effect in Pericyclic Reactions. IX. The Ene Reaction. *Tetrahedron* 1994, *50* (6), 1821–1832.

(54) Popp, B.; Ball, Z. Structure-Selective Modification of Aromatic Side Chains with Dirhodium Metallopeptide Catalysts. *Journal of the American Chemical Society* 2010, *132* (19), 6660–6662.

(55) Seim, K.; Obermeyer, A.; Francis, M. Oxidative Modification of Native Protein Residues Using Cerium(IV) Ammonium Nitrate. *Journal of the American Chemical Society* 2011, *133* (42), 16970–16976.

(56) Rhee, H.-W.; Zou; Udeshi; Martell; Mootha; Carr; Ting. Proteomic Mapping of Mitochondria in Living Cells via Spatially Restricted Enzymatic Tagging. *Science* 2013, *339* (6125), 1328–1331.

(57) Wong, O.; Pradeepkumar, P.; Silverman, S. DNA-Catalyzed Covalent Modification of Amino Acid Side Chains in Tethered and Free Peptide Substrates. *Biochemistry-us* 2011, *50* (21), 4741–4749.

(58) Pradeepkumar; Höbartner, C.; Baum, D.; Silverman, S. DNA-Catalyzed Formation of Nucleopeptide Linkages. *Angewandte Chemie International Edition* 2008, *47* (9), 1753–1757.

(59) Minamihata, K.; Goto, M.; Kamiya, N. Protein Heteroconjugation by the

Peroxidase-Catalyzed Tyrosine Coupling Reaction. *Bioconjugate Chemistry* 2011, *22* (11), 2332–2338.

(60) Minamihata, K.; Goto, M.; Kamiya, N. Site-Specific Protein Cross-Linking by Peroxidase-Catalyzed Activation of a Tyrosine-Containing Peptide Tag. *Bioconjugate Chemistry* 2011, *22* (1), 74–81.

(61) Bruins, J.; Westphal, A. H.; Albada, B.; Wagner, K.; Bartels, L.; Spits, H.; Berkel, W. J. van; Delft, F. van. Inducible, Site-Specific Protein Labelling by Tyrosine Oxidation–Strain-Promoted (4+2) Cycloaddition. *Bioconjugate Chemistry* 2017, *28* (4), 1189–1193.

(62) Qu, N.; Li, F.; Shao, B.; Shao, J.; Zhai, G.; Wang, F.; Zhu, B.-Z. The Unexpected and Exceptionally Facile Chemical Modification of the Phenolic Hydroxyl Group of Tyrosine by Polyhalogenated Quinones under Physiological Conditions. *Chemical Research in Toxicology* 2016, *29* (10), 1699–1705.

(63) Evans, M. J.; Cravatt, B. F. Mechanism-Based Profiling of Enzyme Families. *Chemical Reviews* 2006, *106*, 3279–3301.

(64) Chaiken, IM; Smith, EL. Reaction of a Specific Tyrosine Residue of Papain with Diisopropylfluorophosphate. *Journal of Biological Chemistry* 1969.

(65) Schopfer, L. M.; Champion, M. M.; Tamblyn, N.; Thompson, C. M.; Lockridge, O. Characteristic Mass Spectral Fragments of the Organophosphorus Agent FP-Biotin and FP-Biotinylated Peptides from Trypsin and Bovine Albumin (Tyr410). *Analytical biochemistry* 2005, *345* (1), 122–32.

(66) Grigoryan, H.; Schopfer, L. M.; Thompson, C. M.; Terry, A. V.; Masson, P.; Lockridge, O. Mass Spectrometry Identifies Covalent Binding of Soman, Sarin, Chlorpyrifos Oxon, Diisopropyl Fluorophosphate, and FP-Biotin to Tyrosines on Tubulin:

A Potential Mechanism of Long Term Toxicity by Organophosphorus Agents *Chemico-Biological Interactions* 2008, *175* (1-3), 180–186.

(67) Weerapana, E.; Simon, G. M.; Cravatt, B. F. Disparate Proteome Reactivity Profiles of Carbon Electrophiles. *Nature Chemistry Biology* 2008, *4* (7), 405–7.

(68) Schmidt, JA; Colman, RF. Identification of the Lysine and Tyrosine Peptides Labeled by 5'-P-Fluorosulfonylbenzoyladenosine in the NADH Inhibitory Site of Glutamate Dehydrogenase. *Journal of Biological Chemistry* 1984.

(69) Hanoulle, X.; Damme, J. Van; Staes, A.; Martens, L.; Goethals, M.; Vandekerckhove, J.; Gevaert, K. A New Functional, Chemical Proteomics Technology to Identify Purine Nucleotide Binding Sites in Complex Proteomes. *Journal Proteome Research* 2006, *5* (12), 3438–45.

(70) Wang, Q.; Srinivas, P. R.; Harrison, M. L.; Geahlen, R. L. Partial Purification and Characterization of the Lck Protein-Tyrosine Kinase from Bovine Thymus. *Biochemical Journal* 1991, *279*, 567–574.

(71) Kumar, A. A.; Mangum, J. H.; Blankenship, D. T.; Freisheim, J. H. Affinity Labeling of Chicken Liver Dihydrofolate Reductase by a Substituted 4,6-Diaminodihydrotriazine Bearing a Terminal Sulfonyl Fluoride. *Journal of Biological Chemistry* 1981, *256* (17), 8970–8976.

(72) Engel, M.; Hoffmann, T.; Manhart, S.; Heiser, U.; Chambre, S.; Huber, R.; Demuth, H.-U.; Bode, W. Rigidity and Flexibility of Dipeptidyl Peptidase IV: Crystal Structures of and Docking Experiments with DPIV. *Journal of Molecular Biology* 2006.

(73) Shannon, D. A.; Gu, C.; McLaughlin, C. J.; Kaiser, M.; Hoorn, R. A. van der; Weerapana, E. Sulfonyl Fluoride Analogues as Activity-Based Probes for Serine Proteases.

Chembiochem 2012, *13* (16), 2327–30.

(74) Hayes, J. D.; Flanagan, J. U.; Jowsey, I. R. GLUTATHIONE TRANSFERASES. *Annual Review of Pharmacology and Toxicology* 2005, *45* (1), 51–88.

(75) Tsai, C.-S.; Yen, H.-Y.; Lin, M.-I.; Tsai, T.-I.; Wang, S.-Y.; Hsu, T.-L.; Cheng, Y.-S. E.; Fang, J.-M.; Wong, C.-H. Cell-Permeable Probe for Identification and Imaging of Sialidases. *Proceedings of the National Academy of Sciences* 2013.

(76) Severi, E.; Hood, D. W.; Thomas, G. H. Sialic Acid Utilization by Bacterial Pathogens. *Microbiology* 2007, *153* (9), 2817–2822.

(77) Barycki, J.; Colman, R. Affinity Labeling of Glutathione S-Transferase, Isozyme 4-4, by 4-(fluorosulfonyl)benzoic Acid Reveals Tyr115 to Be an Important Determinant of Xenobiotic Substrate Specificity. *Biochemistry* 1993, *32* (48), 13002–13011.

(78) Wang, J.; Barycki, J. J.; Colman, R. F. Tyrosine 8 Contributes to Catalysis but Is Not Required for Activity of Rat Liver Glutathione S-Transferase, 1-1. *Protein science : a publication of the Protein Society* 1996, *5* (6), 1032–42.

(79) Pettigrew, N. E.; Moyer-Myers, M.; Colman, R. F. Affinity Labeling of Pig Lung Glutathione S-Transferase Pi by 4-(fluorosulfonyl)benzoic Acid. *Archives Biochemistry Biophysics* 1999, *364* (1), 107–14.

(80) Louie, S. M.; Grossman, E. A.; Crawford, L. A.; Ding, L.; Camarda, R.; Huffman, T. R.; Miyamoto, D. K.; Goga, A.; Weerapana, E.; Nomura, D. K. GSTP1 Is a Driver of Triple-Negative Breast Cancer Cell Metabolism and Pathogenicity. *Cell Chemical Biology* 2016, *23* (5), 567–78.

(81) Dawson, S. J.; Provenzano, E.; Caldas, C. Triple Negative Breast Cancers: Clinical and Prognostic Implications. *European Journal of Cancer* 2009, *45*, 27–40.

(82) Pace, L. E.; Keating, N. L. A Systematic Assessment of Benefits and Risks to Guide Breast Cancer Screening Decisions. *Journal of the American Medical Association* 2014, *311* (13), 1327–1335.

(83) Nelson, H. D.; Tyne, K.; Naik, A.; Bougatsos, C.; Chan, B. K.; Humphrey, L. Screening for Breast Cancer: An Update for the US Preventive Services Task Force. *Annals of Internal Medicine* 2009, *151* (10), 727–737.

(84) Weigelt, B.; Geyer, F. C.; Reis-Filho, J. S. Histological Types of Breast Cancer: How Special Are They? *Molecular Oncology* 2010, *4* (3), 192–208.

(85) Huang, R.; Zong, X. Aberrant Cancer Metabolism in Epithelial–mesenchymal Transition and Cancer Metastasis: Mechanisms in Cancer Progression. *Critical Reviews in Oncology/Hematology* 2017, *115*, 13–22.

(86) Board, P. G.; Menon, D. Glutathione Transferases, Regulators of Cellular Metabolism and Physiology. *Biochimica et Biophysica Acta* 2013, *1839* (5), 3267–3288.

(87) Blackburn, A. C.; Woollatt, E.; Sutherland, G. R.; Board, P. G. Characterization and Chromosome Location of the Gene GSTZ1 Encoding the Human Zeta Class Glutathione Transferase and Maleylacetoacetate Isomerase. *Cytogenetic and Genome Research* 1998, *83* (1-2), 109–114.

(88) Benson, A. M.; Talalay, P.; Keen, J. H.; Jakoby, W. B. Relationship between the Soluble Glutathione-Dependent Delta 5-3-Ketosteroid Isomerase and the Glutathione S-Transferases of the Liver. *Proceedings of the National Academy of Sciences of the United States of America* 1977, *74* (1), 158–62.

(89) Flanagan, J. U.; Smythe, M. L. Sigma-Class Glutathione Transferases. *Drug Metabolism Reviews* 2011, *43* (2), 194–214.

(90) Townsend, D. M.; Manevich, Y.; He, L.; Hutchens, S.; Pazoles, C. J.; Tew, K. D. Novel Role for Glutathione S-Transferase Pi. Regulator of Protein S-Glutathionylation Following Oxidative and Nitrosative Stress. *The Journal of biological chemistry* 2009, *284* (1), 436–45.

(91) Laborde. Glutathione Transferases as Mediators of Signaling Pathways Involved in Cell Proliferation and Cell Death. *Cell Death & Differentiation* 2010, *17* (9), 1373–1380.

(92) Liu, D.; Hewawasam, R.; Pace, S. M.; Gallant, E. M.; Casarotto, M. G.; Dulhunty, A. F.; Board, P. G. Dissection of the Inhibition of Cardiac Ryanodine Receptors by Human Glutathione Transferase GSTM2-2. *Biochemical pharmacology* 2009, *77* (7), 1181–93.

(93) Mannervik, B.; Danielson, U. H. Glutathione Transferases--Structure and Catalytic Activity. *Critical Reviews Biochemistry* 1988, *23* (3), 283–337.

(94) Mannervik, B.; Board, P. G.; Hayes, J. D.; Listowsky, I.; Pearson, W. R. Nomenclature for Mammalian Soluble Glutathione Transferases. *Methods in Enzymology* 2005, *401*, 1–8.

(95) Nebert, D. W.; Vasiliou, V. Analysis of the Glutathione S-Transferase (GST) Gene Family. *Human Genomics* 2004, *1* (6), 460–464.

(96) Okamura, T.; Antoun, G.; Keir, S. T.; Friedman, H.; Bigner, D. D.; Ali-Osman, F. Phosphorylation of Glutathione S-Transferase P1 (GSTP1) by Epidermal Growth Factor Receptor (EGFR) Promotes Formation of the GSTP1-c-Jun N-Terminal Kinase (JNK) Complex and Suppresses JNK Downstream Signaling and Apoptosis in Brain Tumor Cells. *Journal of Biological Chemistry* 2015, *290* (52), 30866–30878.

(97) Schnekenburger, M.; Karius, T.; Diederich, M. Regulation of Epigenetic Traits of the Glutathione S-Transferase P1 Gene: From Detoxification toward Cancer Prevention

and Diagnosis. *Frontiers in pharmacology* 2014, *5*, 1–7.

(98) Townsend, D. M.; Tew, K. D. The Role of Glutathione-S-Transferase in Anti-Cancer Drug Resistance. *Oncogene* 2003, *22* (47), 7369–7375.

(99) Gong, L.-H.; Chen, X.-X.; Wang, H.; Jiang, Q.-W.; Pan, S.-S.; Qiu, J.-G.; Mei, X.-L.; Xue, Y.-Q.; Qin, W.-M.; Zheng, F.-Y.; Shi, Z.; Yan, X.-J. Piperlongumine Induces Apoptosis and Synergizes with Cisplatin or Paclitaxel in Human Ovarian Cancer Cells. *Oxidative Medicine and Cellular Longevity* 2014, *2014*, 1–10.

(100) Oliveira, D. M. M. de; Farias, M. T. de; Teles, A. L. L.; Santos Junior, M. C. Dos; Cerqueira, M. D. de; Lima, R. M.; El-Bachá, R. S. 8-Methoxypsoralen Is a Competitive Inhibitor of Glutathione S-Transferase P1-1. *Frontiers in cellular neuroscience* 2014, *8* (308), 1–11.

(101) Raza; Galili; Smith; Godwin; Lancet; Melchert; Jones; Keck; Meng; Brown; List. Phase 1 Multicenter Dose-Escalation Study of Ezatiostat Hydrochloride (TLK199 Tablets), a Novel Glutathione Analog Prodrug, in Patients with Myelodysplastic Syndrome. *Blood* 2009, *113* (26), 6533–6540.

(102) Raza, A.; Galili, N.; Mulford, D.; Smith, S.; Brown, G.; Steensma, D.; Lyons, R.; Boccia, R.; Sekeres, M.; Garcia-Manero, G.; Mesa, R. Phase 1 Dose-Ranging Study of Ezatiostat Hydrochloride in Combination with Lenalidomide in Patients with Non-Deletion (5q) Low to Intermediate-1 Risk Myelodysplastic Syndrome (MDS). *Journal of Hematology & Oncology* 2012, *5* (1), 1–8.

(103) Ricci, G.; Boccio, G. Del; Pennelli, A.; Bello, M. L.; Petruzzelli, R.; Caccuri, A. M.; Barra, D.; Federici, G. Redox Forms of Human Placenta Glutathione Transferase. *The Journal of Ciological Chemistry* 1991, *266*, 21409–21415.

(104) Zanden, J. J. van; Hamman, O. Ben; Iersel, M. L. van; Boeren, S.; Cnubben, N. H.; Bello, M. Lo; Vervoort, J.; Bladeren, P. J. van; Rietjens, I. M. C. M. Inhibition of Human Glutathione S-Transferase P1-1 by the Flavonoid Quercetin. *Chemico-Biological Interactions* 2003, *145* (2), 139–48.

(105) Quesada-Soriano, I.; Primavera, A.; Casas-Solvas, J. M.; Téllez-Sanz, R.; Barón, C.; Vargas-Berenguel, A.; Bello, M. Lo; García-Fuentes, L. Identifying and Characterizing Binding Sites on the Irreversible Inhibition of Human Glutathione S-Transferase P1-1 by S-Thiocarbamoylation. *Chembiochem* 2012, *13* (11), 1594–604.

(106) Barycki, J. J.; Colman, R. F. Affinity Labeling of Glutathione S-Transferase, Isozyme 4-4, by 4-(fluorosulfonyl)benzoic Acid Reveals Tyr115 to Be an Important Determinant of Xenobiotic Substrate Specificity. *Biochemistry* 1993, *32* (48), 13002–11.

(107) Shannon, A.; Banerjee, R.; Webster, E.; Bak, D.; Wang, C.; Weerapana, E. Investigating the Proteome Reactivity and Selectivity of Aryl Halides. *Journal of the American Chemical Society* 2014, *136* (9), 3330–3333.

(108) Leesnitzer, L. M.; Parks, D. J.; Bledsoe, R. K.; Cobb, J. E.; Collins, J. L.; Consler, T. G.; Davis, R. G.; Hull-Ryde, E. A.; Lenhard, J. M.; Patel, L.; Plunket, K. D.; Shenk, J. L.; Stimmel, J. B.; Therapontos, C.; Willson, T. M.; Blanchard, S. G. Functional Consequences of Cysteine Modification in the Ligand Binding Sites of Peroxisome Proliferator Activated Receptors by GW9662. *Biochemistry* 2002, *41* (21), 6640–6650.

(109) Jones, J. R.; Barrick, C.; Kim, K.-A.; Linder, J.; Blondeau, B.; Fulimoto, Y.; Shiota, M.; Kesterson, R. A.; Kahn, B. B.; Magnuson, M. A. Deletion of PPARy in Adipose Tissues of Mice Protects against High Fat Diet-Induced Obesity and Insulin Resistance. *Proceedings of the National Academy of Sciences* 2005, *102* (17), 6207–6212.

(110) Tam, K.; Smith, E.; Booth, J.; Compton, R.; Brennan, C.; Atherton, J. Kinetics and Mechanism of Dyeing Processes: The Dyeing of Cotton Fabrics with a Procion Blue Dichlorotriazinyl Reactive Dye. *Journal of Colloid and Interface Science* 1997, *186* (2), 387–398.

(111) Labrou, N. E.; Eliopoulos, E.; Clonis, Y. D. Dye-Affinity Labelling of Bovine Heart Mitochondrial Malate Dehydrogenase and Study of the NADH-Binding Site. *Biochemical Journal* 1996, *315* (2), 687–693.

(112) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie International Edition English* 2001, *40* (11), 2004–2021.

(113) Speers, A. E.; Adam, G. C.; Cravatt, B. F. Activity-Based Protein Profiling in Vivo Using a Copper(i)-Catalyzed Azide-Alkyne [3 + 2] Cycloaddition. *Journal of the American Chemical Society* 2003, *125* (16), 4686–4687.

(114) Wilce, M.; Parker, MW. Structure and Function of Glutathione S-Transferases. *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* 1994, *1205* (1), 1-18.

(115) Dowling, R. J. O.; Zakikhani, M.; Fantus, I. G.; Pollak, M.; Sonenberg, N. Metformin Inhibits Mammalian Target of Rapamycin–dependent Translation Initiation in Breast Cancer Cells. *Cancer Research* 2007, *67* (22), 10804–10812.

(116) Benjamin, D. I.; Cozzo, A.; Ji, X.; Roberts, L. S.; Louie, S. M.; Mulvihill, M. M.; Luo, K.; Nomura, D. K. Ether Lipid Generating Enzyme AGPS Alters the Balance of Structural and Signaling Lipids to Fuel Cancer Pathogenicity. *Proceedings of the National Academy of Sciences* 2013, *110* (37), 14912–14917.

(117) Benjamin, D. I.; Li, D. S.; Lowe, W.; Heuer, T.; Kemble, G.; Nomura, D. K. Diacylglycerol Metabolism and Signaling Is a Driving Force Underlying FASN Inhibitor Sensitivity in Cancer Cells. *ACS Chemical Biology* 2015, *10*, 1616–1623.

(118) Nomura, D. K.; Long, J. Z.; Niessen, S.; Hoover, H. S.; Ng, S.-W.; Cravatt, B. F. Monoacylglycerol Lipase Regulates a Fatty Acid Network That Promotes Cancer Pathogenesis. *Cell* 2010, *140*, 49–61.

(119) Benjamin, D. I.; Louie, S. M.; Mulvihill, M. M.; Kohnz, R. A.; Li, D. S.; Chan, L. G.; Sorrentino, A.; Bandyopadhyay, S.; Cozzo, A.; Ohiri, A.; Goga, A.; Ng, S.-W.; Nomura, D. K. Inositol Phosphate Recycling Regulates Glycolytic and Lipid Metabolism That Drives Cancer Aggressiveness. *ACS Chemical Biology* 2014, *9*, 1340–1350.

(120) Mulvihill, M. M.; Benjamin, D. I.; Ji, X.; Scolan, E. L.; Louie, S. M.; Shieh, A.; Green, M.; Narasimhalu, T.; Morris, P. J.; Luo, K.; Nomura, D. K. Metabolic Profiling Reveals PAFAH1B3 as a Critical Driver of Breast Cancer Pathogenicity. *Chemistry & biology* 2014, *21*, 831–840.

(121) Long, J. Z.; Cisar, J. S.; Milliken, D.; Niessen, S.; Wang, C.; Trauger, S. A.; Siuzdak, G.; Cravatt, B. F. Metabolomics Annotates ABHD3 as a Physiologic Regulator of Medium-Chain Phospholipids. *Nature Chemical Biology* 2011, *7*, 763–765.

Chapter 2

The Discovery of a Lysine-Reactive Covalent Probe for ALDOA

The work in this chapter is not published.

Introduction

Lysine is an essential amino acid, and it is one of most ubiquitously incorporated amino acids observed in proteins. On average, the pK_a of the amino side chain of a surface exposed lysine is ~10.4. This means at neutral pH (7.4) the vast majority of these lysine residues are in the protonated state, functioning to enhance protein structure and stability. [1] The other basic amino acid residues are arginine and histidine. Protonated lysine residues can also participate catalytically as proton donors in acid base chemistry. Lysine residues buried in a protein have been found to have a perturbed pK_a as much as ~5 units due to influences by the residue's microenvironment or ligand binding. [2–4] These deprotonated lysine residues are able to function as nucleophiles in chemical reactions.

In acid-base chemistry, lysine side chain amino groups typically act as a general base. One common example of this role is found in catalytic dyads and triads. Classically, serine proteases have been thought to contain a catalytic triad. Conversely, there is a proposed group of serine proteases that do not contain the essential histidine residue, rather function with a catalytic dyad, Ser/Lys. [5]In the catalytic dyad, the lysine residue functions as a general base to stabilize the oxygen anion, enhancing the nucleophilicity of the serine side chain. One proposed example of the serine proteases containing a catalytic dyad is *E. coli* LexA repressor. [6]Mutagenesis studies of LexA repressor have shown that lysine 156 and serine 119 are essential for activity (Figure 2-1A). A catalytic Ser/Lys dyad has been proposed for the β-lactamase TEM1, which hydrolyzes β-lactam rings. However, conflicting NMR titration experiments results in controversial roles for lysine 73 (Figure 2-1B).

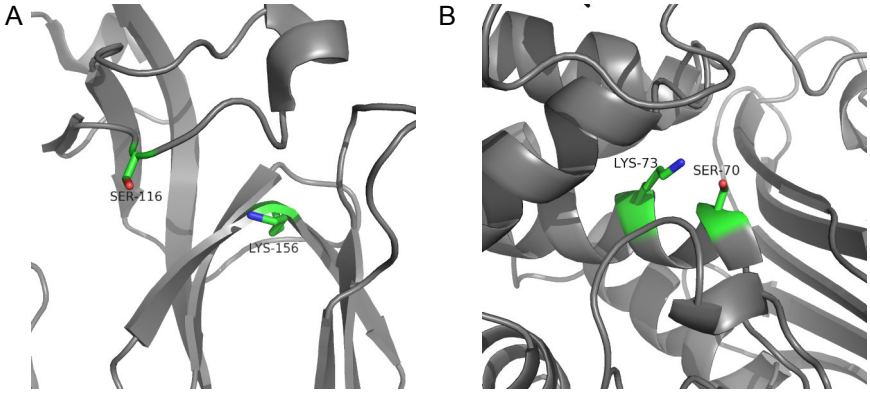
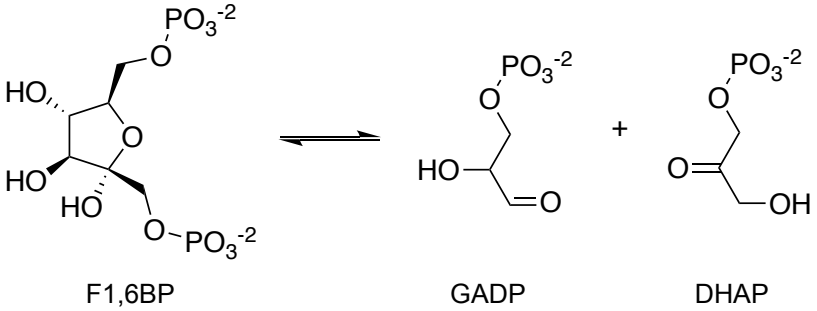


Figure 2-1. (A) LexA repressor catalytic dyad residues (PDB:1jhf). (B) TEM1 β-lactamase proposed catalytic residues (PDB: 1btl).

Protein kinases have a conserved lysine residue in their catalytic domain of subdomain II. This residue has been found to be essential for transfer for the phosphate group from ATP to an amino acid side chain.

In animals and plants class I fructose-bisphosphate aldolase (also referred to as FBP aldolase, FBA, aldolase, or ALDO) is one of 10 enzymes in glycolysis which is responsible for the conversion of glucose into pyruvate as well as the generation of ATP. Class II aldolase, found in fungi, bacteria, and yeast, is a metalloenzyme that uses cobalt(II) for direct binding to the substrate. [7] Like other lyase enzymes, in this reaction aldolase catalyzes a breakdown of its substrate, which in this case is D-fructose 1,6-bisphosphate. More specifically, aldolase catalyzes the reverse aldol condensation of D-fructose 1,6-bisphosphate (F1,6BP) to dihydroxyacetone phosphate (DHAP) + D-glyceraldehyde 3-phosphate (GADP) through a Schiff base mechanism (Scheme 2-1). Aldolase is a homotetramer and is expressed as 3 variants in vertebrates. [8,9]Aldolase A (ALDOA or GSD12) is expressed in muscle, aldolase B (ALDOB) is expressed in the liver, and aldolase C (ALDOC) is expressed in the brain (Figure 2-1). BLAST [10]analysis of the variants shows

the greatest similarity between ALDOA and ALDOAC with 82% homology. The sequence homology between ALDOA/ALDOB and ALDOB/ALDOC is 70% for both.



F1,6BP GADP DHAP

Scheme 2-1. ALDO catalyzes the interconversion of D-fructose 1,6-bisphosphate (F1,6BP) and D-glyceraldehyde 3-phosphate (GADP) + dihydroxyacetone phosphate (DHAP).

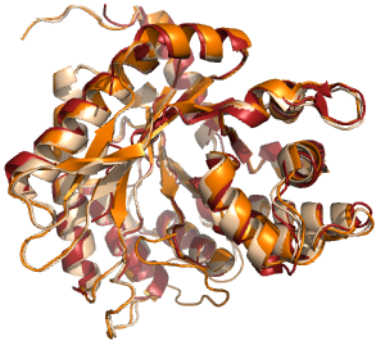


Figure 2-1. Overlay of ALDOA (Orange, PDB: 1ALD), ALDOB (Red, PDB: 1XDM), and ALDOC (Tan, PDB: 1XFB) crystal structures.

Aldolase enzymes have been found to participate in protein-protein interactions, suggesting they have both structural and catalytic related activities. A search for interactions of ALDOA in the BioGRID general repository for interactions indicates 126 unique physical interactions, of which 106 were identified in high throughput and 20 were identified through low throughput experiments. [11]One example of ALDOA protein-protein

interactions is that ALDOA is known to bind actin in the cytoskeleton, however the exact relevance and function of this interaction is unknown. [12,13]The interaction of ALDOA with actin is found to be reversed in cells with the addition of 2-deoxyglucose, a known substrate inhibitor of glycolysis. [14,15]Aldolase was also found to bind a C-terminal peptide, residues 408-451, of α-tubulin with known structural similarity to full length α-tubulin. This interaction was found to inhibit aldolase activity. [16]

There are multiple diseases associated with decreased aldolase activity and stability. Deficiency in ALDOA have been associated with Glycogen Storage Disease Type 12 (GSD12) disease. This metabolic disorder is characterized by mutations in ALDOA gene that results in decreased thermal stability [17,18], reduced catalytic ALDOA activity [19], or reduced substrate affinity [19,20]. Patients with the disease typically have hemolytic anemia and may have intellectual disabilities, short stature, muscle weakness and fatigue and other abnormal phenotypes as a result of ALDOA deficiency. [21,22]

Additionally, ALDOA activity has been linked to cancer cell metabolism. It has been established that cancer cells experience increased aerobic glycolysis, resulting in the production of large amount of ATP that fuel the progression and spread of cancer cells. [23]High levels of ATP turn off or inhibit AMPK activity, which is linked to decreased cancer survival and pathogenicity, serving as a regulatory mechanism. Since glycolysis is essential to cancer cell survival, glycolytic enzymes have been proposed as therapeutic targets for the treatment of the disease. Inhibition of ALDOA has been shown to break the forward reactions of glycolysis and shows promise in inhibiting tumor growth *in vivo*. [24]

Homozygous or heterozygous mutations or deletions in ALDOB that result in decreased ALDOB activity cause Hereditary Fructose Intolerance (HFI). Mutations or

deletions can result in decreased enzymatic catalytic activity [25–27], improper folding [28], and thermal instability [25]. Typically, fructose intolerance appears when fructose is added to an infant's diet. [29]Patients with HFI suffer from hypoglycemia and severe abdominal symptoms, and with continued exposure hepatic and renal injury. Severity of the disease symptoms are directly correlated to patient diet. [30]The disease is managed by avoiding foods that contain fructose.

While there are 3 variants of ALDO, the genes for which are all found on different chromosomes, ALDOA has received the most attention in the literature, likely due to the fact that rabbit muscle aldolase is easily isolated. [31,32]All three human aldolases have a conserved active site lysine residue (Figure 2-2). The active site, K229 makes a covalent Schiff-base intermediate through attack of the carbonyl carbon and subsequent loss of water. [33]Also, each of the aldolases have a conserved active site proton acceptor, E187, that stabilizes the Schiff base intermediate. K229 and E187 are believed to function in a catalytic dyad. [34]The exact mechanism for aldolase is not fully understood. For example, the amino acid responsible for the ring opening of D-fructose 1,6-bisphosphate is unknown, but is proposed to be initiated by K177 or K229. [35,36]There are many proposed key amino acids that have been identified through mutagenesis and crystallography studies.

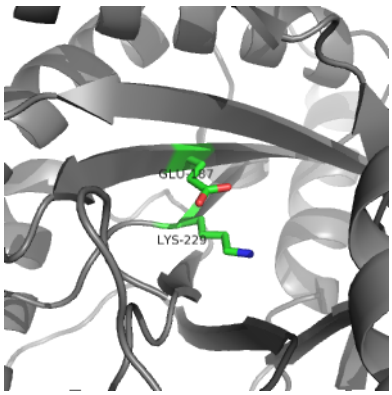


Figure 2-2. Active site residues K229 and E187 of ALDOA (PDB: 1ALD).

The active site lysine, K229, of ALDOA is an interesting target for ABP given the many functional roles of ALDOA. Lysine residues play a variety of roles that can be both functional, such as active sites, or structural, participating protein-protein interactions. These activities contribute to the demand for lysine, an essential amino acid, to be obtained from the diet. The pK_a of the amino group of the lysine side chain is 10.4 in water, however the pK_a of a lysine side chain in a protein can be depressed as far as 5.3, likely due to influence from the surrounding protein microenvironment. [2]Lysine residues can function as a nucleophile but also participate in acid/base chemistry. Protein kinases have a conserved lysine residue in the ATP-binding pocket. [37,38]Lysine residues, positively charged at physiological pH, are often located on the surface of proteins, enhancing that stability of the protein. [1]Human serum albumin uses its many surface lysine residues to bind endogenous and exogenous molecules for transport through the blood system (Figure 2-3). [39]Lysine residues are also commonly involved in post translational modifications such as acetylation, methylation, ubiquitination, and phosphorylation which function as a mechanism of cellular signaling or as a means to alter the function and activity of the protein. [40]

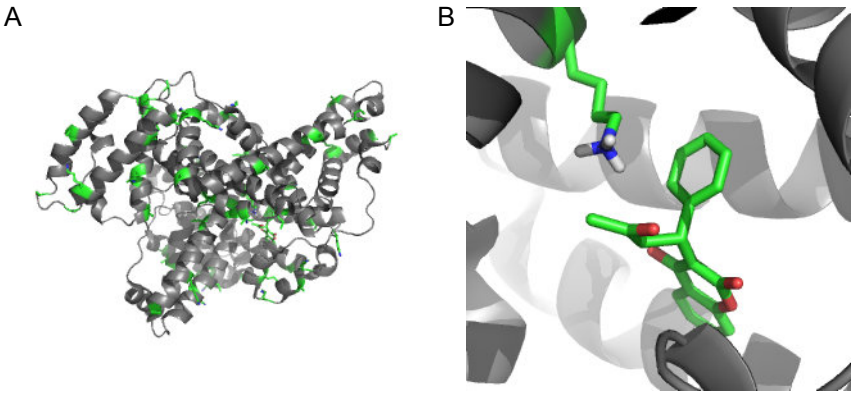


Figure 2-3. Human serum albumin. (A) Lysine residues shown in green. (B) Cation-pi interaction of K199 and ring of Warfin. (PDB: 2BXD)

One way to study reactive lysine residues is through the use of ABPs. While the biological role of lysine is vast, the number of lysine reactive electrophiles employed in ABPP is limited. One challenging aspect of targeting lysine residues is the additional presence of primary amines, protein N-terminal amines. Nonspecifically, primary amines in proteins can be modified by reductive alkylation with aldehydes, succinimidyl esters, sulfonyl chlorides, isocyanates, and thioisocyanates. [41,42]Directed lysine modification has been demonstrated with ATP and ADP analogues broadly at the conserve lysine in protein kinase active sites. [43] Most recently, efforts have been directed towards modification of functional, but not necessarily catalytic lysine residues. One study uses a pentynoic acid sulfotetrafluorophenyl ester to investigate ligandable lysine residues in a proteome. [44]The disparity between the number of nucleophilic and functional lysine residues and the number of selective lysine modifications motivates our interest to close the gap. Selective lysine ABPs will give access to studies that will increase our knowledge of the biological role of lysine in the proteome. More specifically, identification of an ABP and/or inhibitor of ALDOA will give access to understanding the underlying biological consequences of ALDOA inhibition in GSD12 disease and cancer and may give insight into new therapeutic strategies for the treatment of the diseases.

Currently, there are known inhibitors of ALDOA (Figure 2-4). 1,2-cyclohexanedione (CHD) is known to inhibit ALDOA activity through the modification of arginine residue, R55. [45]CHD has been shown to modify lysine and arginine residues on

other proteins, thus CHD is not ideal for selective ALDOA inhibition studies. [46,47]Phosphoric acid mono-(2,3-dioxo-butyl) ester is a substrate inhibitor that was designed to modify the active site lysine of ALDOA. However, Phosphoric acid mono-(2,3-dioxo-butyl) ester modifies an arginine residue in aldolase A from rabbit muscle. [48]A Chemical library screen revealed compound TDZD-8 which modifies C239 which is located distal to the active site. [24]TDZD-8 was used to show impaired glycolysis in MDA-MB-231 breast cancer cells. This list is just sampling of the known ALDOA inhibitors.

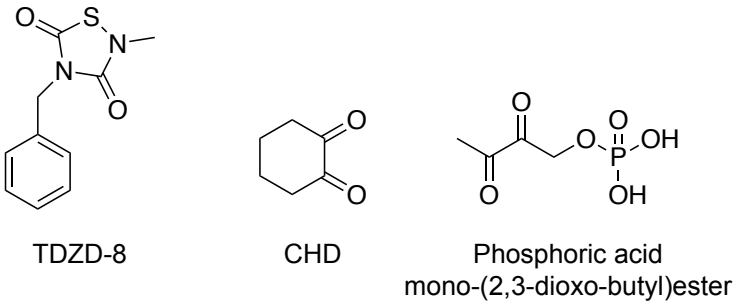


Figure 2-4. Sampling of known ALDOA inhibitors.

As mentioned in Chapter 1, our efforts to expand the current chemical tool box of ABPs started with inspiration from literature referencing a cysteine reactive arylhalide containing inhibitor of PPAR-γ. [49]Our investigation of the proteome reactivity of a dichlorotriazine electrophile revealed preferential reactivity towards lysine residues. [50]Here we report the discovery of a dichlorotriazine containing, lysine selective small molecule inhibitor of ALDOA.

Results and Discussion

Dichlorotriazine Library Synthesis and Evaluation

See Chapter 1, Results and Discussion section on page 32.

Evaluation of the dichlorotriazine library in HeLa cell soluble lysates and subsequent MS analysis revealed LAS17 as a selective ABP for the protein GSTP1. However, we desired to further characterize the protein targets of some of the additional dichlorotriazine library members.

Identification of ALDOA as a target of LAS1, LAS6, and LAS12

Given the desire to discover and characterize a lysine reactive ABP motivated by the identification of the dichlorotriazine electrophile (Chapter 1), we chose identify the protein targets of select dichlorotriazine containing probe library members. LAS1, LAS6, and LAS12 showed both individual differences and similarities in their reactivities toward the proteome in the in-gel fluorescence imaging studies (Figure 1-10).

To identify the protein targets of LAS1, LAS6, and LAS12, HeLa cell lysates were treated with the individual probe library members (1 μM) or DMSO and were subjected to CuAAC with biotin-azide as previously described in Chapter 1 (Figure 1-13). To identify common targets of LAS1, LAS6, and LAS12, the tryptic data was filtered to include proteins for which there were at least three spectral counts for the probe treated samples and zero spectral counts in the DMSO treated sample (Table 2-1, Table 2A-1). Since ALDOA has an active site lysine, it is an ideal candidate to demonstrate the utility of a

lysine reactive ABP. For this reason, we sought to investigate the potential of a dichlorotriazine probe as an inhibitor of ALDOA. Instead of choosing one member of the dichlorotriazine library for further studies, we moved forward with a panel of probe library members in an effort to identify the most selective probe members in addition to the best inhibitor.

Protein	MW(Da)	DMSO	LAS1	LAS6*	LAS12*
HSP90AB1 Heat shock protein HSP 90-β	83264	0	26	28	21
HSPD1 60 kDa heat shock protein, mito	61055	0	11	22	29
HSP90AA1 Heat shock protein HSP 90-α	84660	0	20	20	18
GSTP1 Glutathione S-transferase P	23342	0	4	24	28
ACLY ATP-citrate synthase	120839	0	11	15	21
PGK1 Phosphoglycerate kinase 1	44615	0	21	12	8
EEF2 Elongation factor 2	95338	0	6	8	25
PGAM1 Phosphoglycerate mutase 1	28804	0	9	9	14
ALDOA Fructose-bisphosphate aldolase A	39420	0	13	5	7
GANAB Neutral alpha-glucosidase AB	106874	0	7	7	8

Table 2-1. The top 10 commonly identified proteins among LAS1, LAS6, and LAS12 treated samples represented by number of observed spectral counts for each protein (* n=2).

Confirmation of ALDOA as a Target of Dichlorotriazine Probes

To confirm the LC/LC-MS/MS findings, human ALDOA with a C-term hexahistidine affinity tag was recombinantly expressed in *E. coli* and purified to homogeneity. The purified protein (0.2 mg/mL) was treated with 250 nM of individual probe library members, conjugated with fluorophore through CuAAC, and analyzed for in-

gel fluorescence (Figure 2-5). The results from the library screen showed that the fluorescence signal from LAS12, LAS15, LAS16, and LAS19 were the most intense.

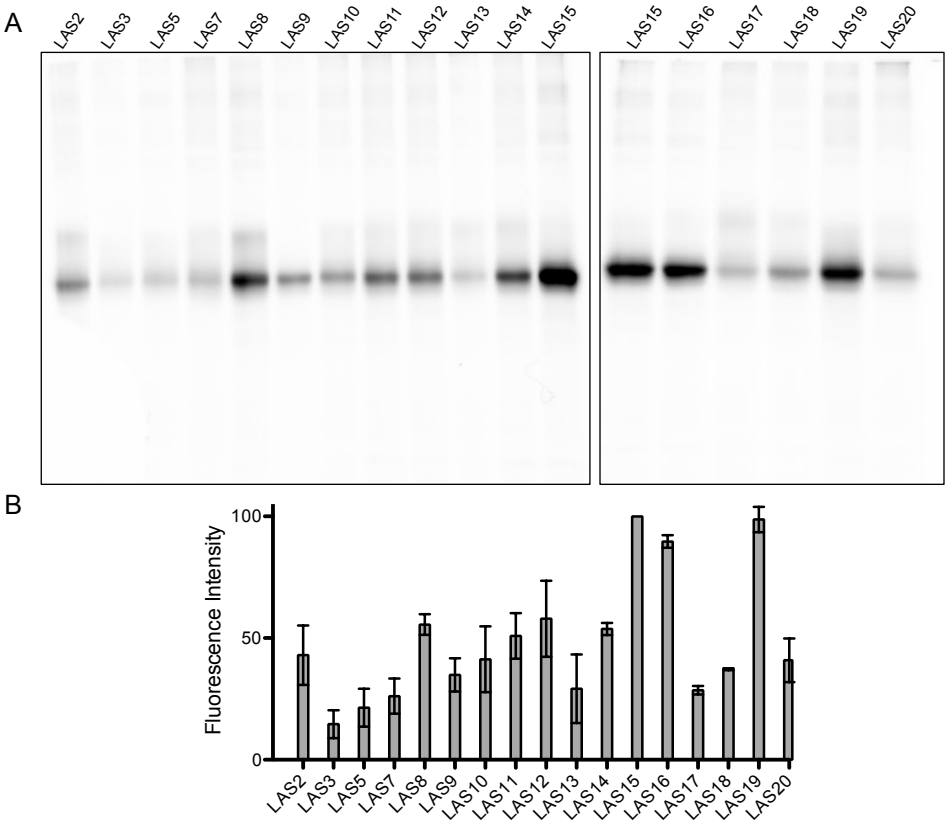


Figure 2-5. Labeling of purified ALDOA with dichlorotriazine library members. (A) Representative image of purified ALDOA (0.2 mg/mL) treated with respective probe library members (250 nM) analyzed by in-gel fluorescence. (B) Fluorescence intensities of n=3 ALDOA labeling studies. Relative fluorescence intensities were normalized to protein abundance (Commassie Blue) which were then normalized to LAS15, represented on each gel.

Next, we sought to evaluate the specificity of these four probes in a complex background to simulate a context more similar to that of a cell. LAS12, LAS15, LAS16, and LAS19 were evaluated for ALDOA labeling in the context of MCF7 lysates (Figure

2-6). LAS 15, LAS16, and LAS19 showed labeling of exogenous ALDOA. Next, we chose to investigate the potential of the dichlorotriazine probes to selectively label endogenously expressed ALDOA, to better investigate the reactivity and selectivity in a more native setting. To investigate the labeling potential of LAS 15, LAS16, and LAS19 to label overexpressed ALDOA, ALDOA-Myc was over expressed in HEK293T cells. HEK239T overexpressing ALDOA-Myc soluble cell lysates were treated with LAS 15, LAS16, and LAS19. However, we did not see a difference in labeling intensity around the anticipated molecular weight of ALDOA (Figure 2-7).

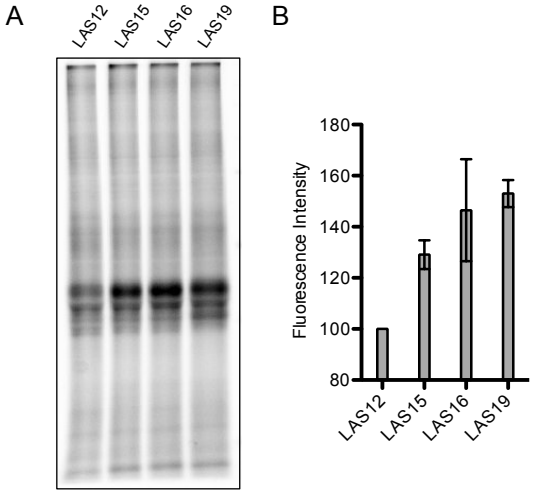


Figure 2-6. Labeling of purified ALDOA in MCF7 background lysates with dichlorotriazine library members. (A) Representative image of purified ALDOA (0.2 mg/mL) with MCF7 background (1.0 mg/mL) treated with respective probe library members (250 nm) analyzed by in-gel fluorescence. (B) Fluorescence intensities of n=3 ALDOA labeling studies. Fluorescence intensity was normalized to LAS12.

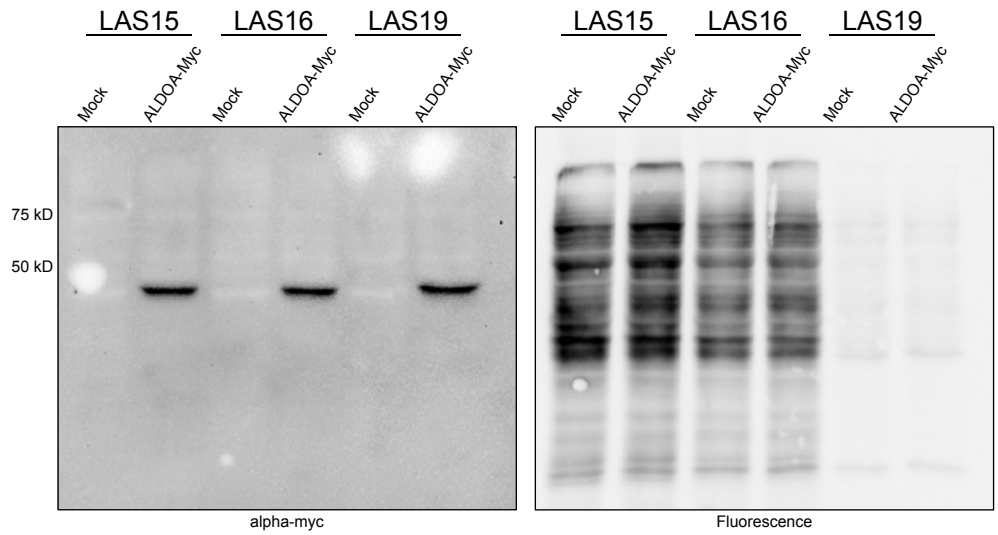


Figure 2-7. Labeling studies of overexpressed ALDOA-Myc in HEK293T lysates with LAS15, LAS16, and LAS19 (250 nM) compared to mock lysates (2.0 mg/mL).

LAS9 Treatment Inhibits ALDOA Activity

To determine if any of the dichlorotriazine probes inhibit ALDOA activity, an *in vitro* assay was performed. In this assay the activity of ALDOA was monitored through a coupled assay. ALDOA converts F1,6BP to DHAP and GADP. Triosephosphate isomerase (TPI) catalyzes the interconversion of DHAP to GADP. GADP is then irreversibly converted to 1,3BPG by glyceraldehyde phosphate dehydrogenase (GAPDH) and one equivalent of NADH. ALDOA activity can be indirectly determined by spectroscopically monitoring the disappearance of NADH at 340 nm (Figure 2-8). LAS9 showed the greatest ability to inhibit ALDOA activity *in vitro*. A concentration screen of indicates LAS9 has an IC_{50} of 8.14 μM for the inhibition of ALDOA (Figure 2-9).

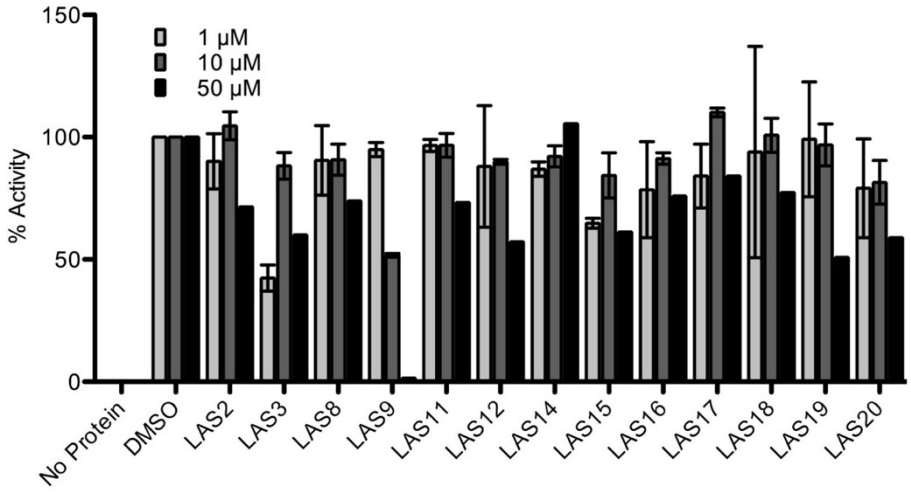


Figure 2-8. Evaluation of *in vitro* ALDOA activity in the presence of various dichlorotriazine probe library members at 1 µM, 10 µM, and 50 µM. LAS19 treatment at 50 µM showed complete inhibition of ALDOA activity.

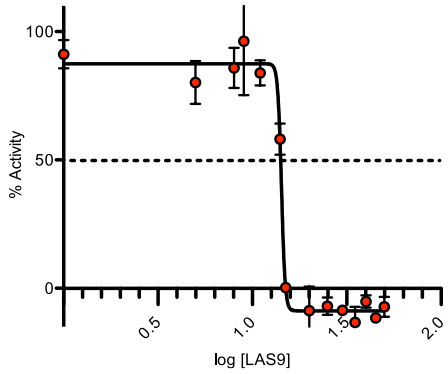


Figure 2-9. LAS9 has an IC_{50} of 8.14 µM for the inhibition of ALDOA.

Characterization of Dichlorotriazine Probe Site of Modification of ALDOA

The visualization of protein fluorescence labeling after denaturing SDS-PAGE confirmed that the dichlorotriazine probes modify ALDOA covalently. We sought to determine the site of labeling of select library members on ALDOA. This was accomplished using purified ALDOA as previously described in Chapter 1 (Figure 1-18). LAS9 was chosen for site of labeling evaluation due to the observed strong inhibitory effect

of ALDOA activity and apparent *in vitro* labeling. LAS16 was chosen because of the high fluorescence signal intensity observed in the gel based assays of purified ALDOA and ALDOA in the presence of MCF7 background lysates. However, LAS16 treatment of ALDOA was shown to possess negligible inhibitory capabilities. LAS8 serves as a control as it was shown to label ALDOA only moderately *in vitro* and also does not appear to inhibit ALDOA activity *in vitro*. Tryptic peptides were searched for differential masses on lysine, cysteine, and tyrosine residues (Table 2-2).

Probe	Peptide	Spectral Counts	Residue
LAS8 (K, +234.9777) (C, +177.0458)	R.IVAPGK*GILAADESTGSIAK.R	5	K27
	K.VDK*GVVPLAGTNGETTTQGLDGLSER.C	2	K110
	A.SIC*QQNGIVPIVEPEILPDGDHDLKR.C	1	C177
LAS9 (K, +237.9777)	H.RIVAPGK*GILAADESTGSIAK.R	1	K27
	R.IVAPGK*GILAADESTGSIAK.R	2	K27
	K.ENLK*AAQEEYVK.R	1	K321
LAS16 (K, +352.0939)	K.K*ELSDIAHR.I	1	K13
	R.IVAPGK*GILAADESTGSIAK.R	3	K27
	K.GILAADESTGSIAK*R.L	1	K41
	K.GGVVGIK*VDK.G	1	K107
	K.VDK*GVVPLAGTNGETTTQGLDGLSER.C	1	K110
	K.DGADFAK*WR.C	1	K146
	K.AAQEEYVK*R.A	1	K229

Table 2-2. Tryptic peptides identified to have a differential modification corresponding to a modification on lysine.

Site of labeling mass spectrometry studies revealed unique modification sites for each of the probes evaluated. LAS8 modifies K27, K110, and C177. K27 is located at the base of the active site beta-sheet barrel in a flexible loop that could potentially participate

in acid-base chemistry of ALDOA catalysis. K110 is located on a flexible loop at the opening of the active site binding pocket. K110 has been found to be modified as acetylated lysine. The effects of this modification on ALDOA activity are unclear. [51]C177 is located on the surface of ALDOA, and is distal to the active site binding pocket opening. LAS9 modifies K27, and K321. K321 has also been found to be acetylated. Mutagenesis studies of K to Q simulates acetylation in K321Q, resulting in a significant increase in activity for K321Q over WT ALDOA. LAS16 modifies the most resides out of all the probes analyzed for site of modification, K13, K27, K41, K107, K110, K146, and K229. However, LAS16 is the only probe library member shown to modify the active site lysine, K229 (Figure 2-10). K13 is located on an alpha-helix near the N-term region. K41 is located at the base of the active site binding pocket on an alpha-helix. Data suggests K41 interacts with the negatively charged 6-phosphate on the F1,6BP substrates. [33]K107 is located on a beta-sheet at the opening of the active site binding pocket. The presence of positively charges lysine residues (K107 and K110) at the opening of the active site may function to attract negatively charges molecules into the active site binding pocket of ALDOA. K146 is located on one of the eight beta-sheets that make up the active site binding domain, and it located very close to active site residue, K229. Data suggests K146 interacts with the negatively charges 1-phosphate of F1,6BP. [33]Conflicting reports suggest K107 is involved in phosphate binding and K149 plays an essential role in carbon-carbon bond breaking/formation, cleavage and condensation of the C3-C4 bond of F1,6BP. [52]K229 is the active site nucleophile of ALDOA and is located on a beta-sheet in the active site binding pocket proximal to E187, the catalytic acidic residue.

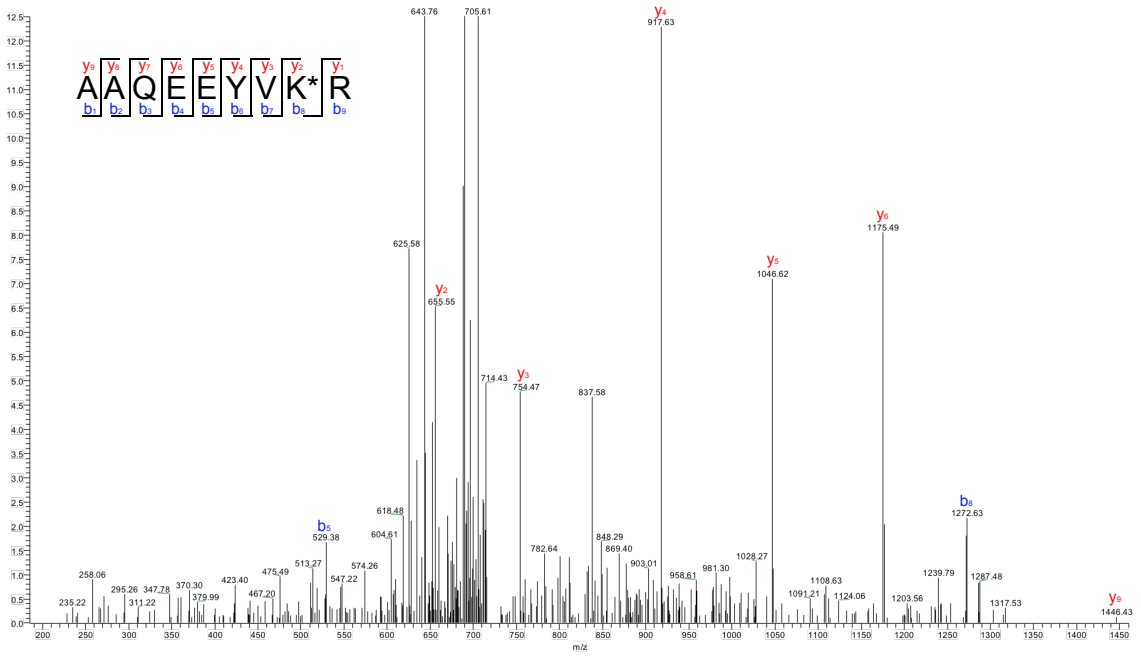


Figure 2-10. Fragmentation (MS/MS) spectra of the AAQEEYVK peptide of ALDOA confirming that the LAS16 modification (+352.0939 Da) is present on K229, which is indicated with *.

This data together suggests LAS8, LAS9, and LAS16 have different modes of action on ALDOA. First, LAS9 was the only probe library member to exhibit significant inhibition of ALDOA activity. LAS9 was the only probe evaluated that modifies K321, which is known to be a site of regulatory acetylation. The unique selectivity for K321 may explain the apparent inhibitory effect of LAS9. On the other hand, LAS16 was among the probe library members to label ALDOA with the highest intensity, as determined by in-gel fluorescence (Figure 2-5). Site of labeling studies of LAS16 treated ALDOA indicate the highest number of unique side chains modifications, compared to LAS8 and LAS9. This suggests the promiscuity of LAS16 labeling of ALDOA contributes to the intense fluorescence signal.

It is important to note the number of spectral counts from this data set are of low confidence. Additional optimization of the protocol and additional technical replicates will improve the significance of the findings mentioned above.

Future Work

While the work in this chapter leads to the discovery of a covalent irreversible inhibitor of ALDOA, there is opportunity to expand on the findings with future experiments. One such study is the confirm site of labeling studies through mutagenesis. Fluorescence-based gel experiments of ALDOA mutants corresponding to the identified sites of labeling for the probe library members will confirm the mass spectrometry based findings. Furthermore, activity assays with the mutants will give insight into the role of the residue in the catalytic mechanism.

Since the activity assay performed is an indirect assessment of ALDOA activity, this is a potential area for improvement. A more direct, and therefore ideal way to assess ALDOA activity would be to monitor the conversion of F1,3BP to DHAP and GADP. This can be accomplished using mass spectrometry with the use of an internal standard to quantify amounts of F1,3BP and chemically sequester the products of ALDOA through a non-reversible reaction to disfavor ALDOA reverse activity based on Le Chatelier's principle.

To fully understand the potential utility of the dichlorotriazine library it would be beneficial to understand the *in situ* reactivity and selectivity of the probe members. So far, the dichlorotriazine library has been screened in HeLa cell lysates. It would be beneficial to investigate the reactivity of the library in additional cell types. We anticipate the proteins

targeted by the library member to vary among cell type given the inherent variability in protein expression, activity, and functional state. Additionally, we can investigate additional protein targets of the dichlorotriazine probe library that utilized active site lysine residues.

Conclusions

In summary, we report the discover and characterization of LAS9, a lysine and cysteine reactive irreversible inhibitor of human ALDOA. Additionally, we report the discovery and characterization of LAS16, a lysine selective, yet not site selective, probe for human ALDOA, that modifies the active site lysine, K229.

The site of labeling data in conjunction with the inhibition studies suggests K321 plays a role in the mechanism or activity of ALDOA. Since K321 is located distal to the active site it is possible that K321 modification by LAS9 results in allosteric inhibition of ALDOA.

Additional significance of K321 modification is linked to the known regulatory mechanism of acetylation. Histone deacetylases (HDACs) and histone acetyl transferases (HATs) are known to regulate protein acetylation. NAD-dependent HDAC, Sirtuin 2 (SIRT2) was found to regulate acetylation of ALDOA.[51] SIRT2 knockdown cell lines show hyperacylation of K321 and K110 of ALDOA. Mutagenesis studies show a correlation between ALDOA acylation and activity. The observed site selectivity of LAS9 suggest it can potentially be used to probe the state of K321 modification in acetylation studies.

Acknowledgements

In addition to the acknowledgement mentioned previously in Chapter 1, I would like to acknowledge Alex Shannon for his work subcloning human ALDOA into pET23a and pcDNA3.1 and the optimization of the expression and purification of recombinant ALDOA from *E. Coli*.

Experimental Procedures

General

All reagents were purchased from Sigma Aldrich unless otherwise noted. All compounds were characterized by ^{1}H and ^{13}C NMR on either a Varian (Palo Alto, CA) 500 MHz or 600 MHz spectrometer. Chemical shifts (δ) are reported in parts per million (ppm) with chemical shifts referenced to internal standards: $CDCl_3$ (7.26 ppm for ^{1}H, 77.8 ppm for ^{13}C). Coupling constants (J) are reported in Hertz (Hz) and multiplicities are abbreviated as singlet (s), broad singlet (bs), doublet (d), triplet (t), pentet (p), multiplet (m), doublet of doublets (dd), and doublet of triplets (dt). High resolution mass spectra (HRMS) were obtained at the Mass Spectrometry Facility at Boston College (Chestnut Hill, MA). Analytical thin layer chromatography (TLC) was performed on Sorbent Technologies Silica G TLC Plates w/UV354 (0.25 mm). All compounds were visualized on TLC by UV and/or $KMnO_4$ staining. Column chromatography was carried out using forced flow of indicated solvent on Sorbent Technology Standard Grade Silica Gel, 40-63 μm particle size, 60 Å pore size (Sorbent Technologies). PBS buffer, DMEM/High glucose media, and penicillin streptomycin (Pen/Strep) were purchased from Thermo Scientific

(Waltham, MA). Primers were ordered from Eurofins MWG Operon (Hunstville, AL), and sequencing was performed by Genewiz (Cambridge, MA).

Protein Target Identification by LC/LC-MS/MS

HeLa soluble protein lysates in DPBS (pH 7.4) (500 μL, 2 mg/mL) were aliquoted and LAS1, LAS6, LAS12 (1 μM) or DMSO was added to the appropriate samples. Two aliquots were made for each inhibitor concentration or DMSO equaling 2 tubes for one final sample. All samples were then treated with biotin-azide (100 μM, 50X stock in DMSO), TCEP (1 mM, 50X fresh stock in water), TBTA ligand (100 μM, 17X stock in DMSO:*t*-butanol = 1:4), and copper(II) sulfate (1 mM, 50X stock in water) followed by incubation at 22 °C for 1 hr. Samples were combined pairwise and centrifuged (6500 g, 4 min, 4 °C) to pellet the precipitated proteins. The pellets were resuspended in cold methanol by sonication and the two samples were combined. Centrifugation was followed by a second cold methanol wash, after which the pellet was solubilized in DPBS containing 1.2% SDS via sonication and heating (90 °C, 5 min).

The SDS-solubilized proteome samples were diluted by 5 mL of DPBS for a final SDS concentration of 0.2%. The solution was incubated with 100 μL of streptavidin-agarose beads (Thermo Scientific, washed 3X with DPBS to remove storage buffer) overnight at 4 °C. Samples were rotated at 22 °C for 2 hr before being washed by 5 mL 0.2 % SDS/DPBS, 3 X 5 mL DPBS, and 3 X 5 mL water. The beads were pelleted by centrifugation (1400 X g, 3 min) between washes.

The washed beads were suspended in 500 uL of 6 M urea/DPBS and 10 mM DTT (from 20X stock in water) and placed in a 65 °C heat block for 15 min. Iodoacetamide (20

mM from 50X stock in water) was then added and the samples were allowed to react at 37 °C for 30 min while shaking. Following reduction and alkylation, the beads were pelleted by centrifugation and resuspended in 200 μL of 2 M urea/DPBS, 1 mM $CaCl_2$ (100X stock in water), and sequencing-grade trypsin (2 μg). The digestion was allowed to proceed overnight at 37 °C while shaking. The beads were pelleted by centrifugation and washed with 2 X 50 μL water. The washes were combined with the supernatant from the trypsin digestion step. Formic acid (15 μL) was added to the samples, which were stored at -20 °C until mass spectrometry analysis.

LC/LC-MS/MS analysis was performed on an LTQ-Orbitrap Discovery mass spectrometer (ThermoFisher) coupled to an Agilent 1200 series HPLC. Peptide digests were pressure loaded onto a 250 μm fused silica desalting column packed with 4 cm of Aqua C18 reverse phase resin (Phenomenex). The peptides were eluted onto a biphasic column (100 μm fused silica with a 5 μm, tip, packed with 10 cm C18 and 4 cm Partisphere strong cation exchange resin (SCX, Whatman) using a gradient 5-100% Buffer B in Buffer A (Buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; Buffer B: 20% water, 80% acetonitrile, 0.1% formic acid). The peptides were then eluted from the SCX onto the C18 resin and into the mass spectrometer using 4 salt steps previously described. The flow rate through the column was set to ~0.25 μL/min and the spray voltage was set to 2.75 kV. One full MS scan (FTMS) (400-1800 MW) was followed by 8 data dependent scans (ITMS) of the n^{th} most intense ions.

The tandem MS data were searched using the SEQUEST algorithm using a concatenated target/decoy variant of the human UniProt database. A static modification of

+57.02146 on cysteine was specified to account for alkylation by iodoacetamide. SEQUEST output files were filtered using DTASelect.

Overexpression and Purification of ALDOA from *E. Coli*

The cDNA for WT-ALDOA was subcloned into a pET-23a C-term His Tag expression vector. All constructs were verified by sequencing (Genewiz, Cambridge, MA). Constructs were transformed into BL21 competent *E. coli* (New England Biolabs). From an overnight LB culture with antibiotics at 37 °C, 5 mL were added to 500 mL LB (pH 7.0) with antibiotics and were grown to OD600 of 0.8. Protein expression was induced with IPTG (100 μM, 250X stock in water) for 5 hrs at 37 °C. Soluble cell lysates in DPBS (pH 7.4) were purified using Ni-NTA chromatography with imidazole concentrations of 25 mM and 500 mM in DPBS (pH 7.4) for the wash and elution steps, respectively. Purification fractions were analyzed for purity using SDS-PAGE (Figure 2-11). Imidazole was removed from pure protein fractions using NAP-5 desalting columns that had been buffer exchanged with DPBS.

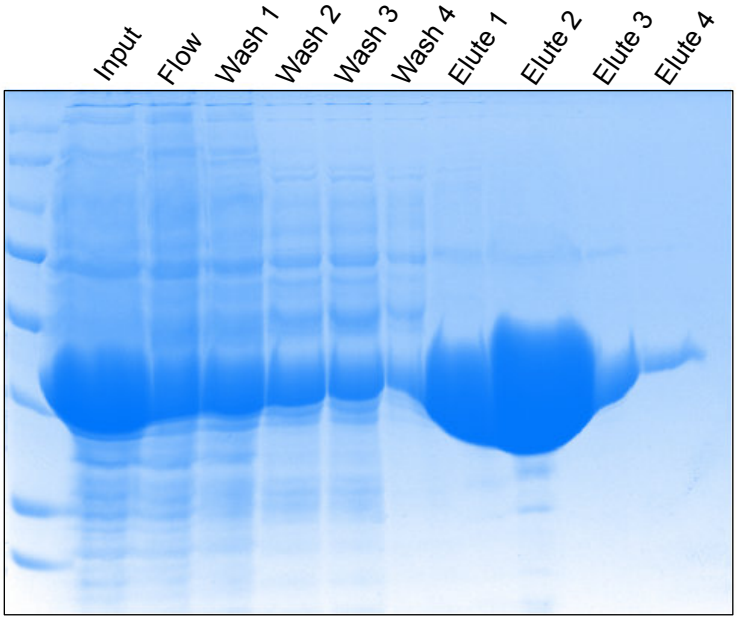


Figure 2-11. Fractions from ALDOA purification were analyzed for purity by SDS-PAGE.

Dichlorotriazine Library Labeling of Purified ALDOA

Purified ALDOA (50 µL, 0.2 mg/mL) was pretreated with probes (250 nM, 50X stock in DMSO) at 22 °C for 1 hr. Samples then underwent click chemistry with TAMRA-azide (Lumiprobe, 25 µM, 50X stock in DMSO), TCEP (1 mM, 50X fresh stock in water), TBTA ligand (100 µM, 17X stock in DMSO:*t*-butanol = 1:4), and copper(II) sulfate (1 mM, 50X stock in water) followed by incubation at 22 °C for 1 hr. SDS-PAGE loading buffer 2X (reducing, 50 µL) was added to the samples and 25 µL of this solution was separated by SDS-PAGE at 100 volts for 145 minutes on a 10% polyacrylamide gel. Gels were visualized on a Bio-Rad ChemiDoc MP Imaging System using the rhodamine setting. After analysis, gels underwent a typical procedure for Coomassie staining and destaining. Stained gels were visualized on a Bio-Rad ChemiDoc MP Imaging System.

Dichlorotriazine Library Labeling of Recombinant ALDOA in MCF7 Background

Samples of purified ALDOA (50 μL, 0.2 mg/mL) in the presence of MCF7 background lysates (1.0 mg/mL) were pretreated with probes (250 nM, 50X stock in DMSO) at 22 °C for 1 hr. Samples then underwent click chemistry with TAMRA-azide (Lumiprobe, 25 μM, 50X stock in DMSO), TCEP (1 mM, 50X fresh stock in water), TBTA ligand (100 μM, 17X stock in DMSO:*t*-butanol = 1:4), and copper(II) sulfate (1 mM, 50X stock in water) followed by incubation at 22 °C for 1 hr. SDS-PAGE loading buffer 2X (reducing, 50 μL) was added to the samples and 25 μL of this solution was separated by SDS-PAGE at 100 volts for 145 minutes on a 10% polyacrylamide gel. Gels were visualized on a Bio-Rad ChemiDoc MP Imaging System using the rhodamine setting. After analysis, gels underwent a typical procedure for Coomassie staining and destaining. Stained gels were visualized on a Bio-Rad ChemiDoc MP Imaging System.

Dichlorotriazine Library Labeling of ALDOA-Myc in HEK293T

Mock and ALDOA-Myc overexpressing soluable cell lysates (50 μL, 2.0 mg/mL) were pretreated with probes (250 nM, 50X stock in DMSO) at 22 °C for 1 hr. Samples then underwent click chemistry with TAMRA-azide (Lumiprobe, 25 μM, 50X stock in DMSO), TCEP (1 mM, 50X fresh stock in water), TBTA ligand (100 μM, 17X stock in DMSO:*t*-butanol = 1:4), and copper(II) sulfate (1 mM, 50X stock in water) followed by incubation at 22 °C for 1 hr. SDS-PAGE loading buffer 2X (reducing, 50 μL) was added to the samples and 25 μL of this solution was separated by SDS-PAGE at 100 volts for 145 minutes on a 10% polyacrylamide gel. Proteins were transferred to nitrocellulose membranes at 75 V for 2 hrs in Tris-buffered saline (TBS) buffer. Blots were blocked with 5% nonfat milk in

Tris-buffered saline containing Tween 20 (TBST) solution for 2 hrs at room temperature with aggitation, washed in TBST, and probed with primary anti-Myc antibody (Cell Signaling) diluted in TBST over- night at 4 °C. Following washes with TBST, the blots were incubated in the with an HRP-linked secondary antibody (Cell Signaling) at room temperature for 2 hrs. Blots were visualized on a Bio-Rad ChemiDoc MP Imaging System using the rhodamine setting to evaluate probe modification and chemiluminescence to detect ALDOA-Myc expression levels.

Evaluating the Site of Modification

100 μg of ALDOA in DPBS was treated with DMSO, LAS8, LAS9, or LAS16 (1 μM, 50X stock in DMSO) for 1 hr at 22 °C. Protein was precipitated by addition of 100% trichloroacetic acid in PBS and incubated at -80 °C overnight. Thawed samples were centrifuged at 15K for 10 min at 22 °C and the supernatant was discarded. The remaining protein pellet was washed with 500 μL of cold acetone, vortexed to resuspend the pellet and centrifuged at 15K for 10 min at 22 °C. The supernatant was again discarded and the pellet was allowed to air dry until trace amounts of acetone were gone. The pellet was resuspended in 30 μL of 8M urea in PBS, then 70 μL 100 mM ammonium bicarbonate in PBS and 1.5 μL of 1 M DTT in PBS were added. Samples were incubated at 65 °C for 15 minutes. Samples were alkylated for 30 minutes at room temperature with the addition of 2.5 μL of 500 mM iodoacetamide in PBS. Sample volume was increase to 224 μL by addition of 120 μL PBS, then 2 μg of sequencing-grade trypsin (Promega) and 2.5 μL of 100 mM $CaCl_2$ was added. Samples were agitated over night at 37 °C. Then trypsin was quenched with 10 μL of formic acid (~5% of final volume) and were centrifuged at 15K

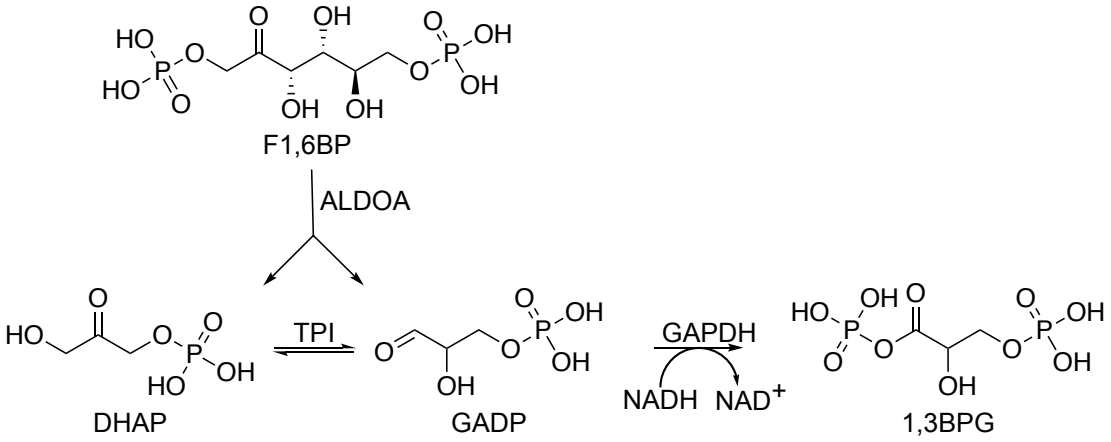
for 20 minutes at room temperature to pellet undigested protein. Supernatant was transferred to a new tube and stored at -20 °C. Samples were analyzed by LC-MS/MS using a LTQ Orbitrap XL mass spectrometer (ThermoFisher) coupled to an EASY-nLC 1000 nanoLC (ThermoFisher). 10 μL of peptide digests were loaded onto 100 μm fused silica column with a 5 μm tip packed with 10 cm of Aqua C18 reverse phase resin (Phenomenex) using the EASY-nLC 1000 autosampler. The digests were eluted using a gradient 0-100% Buffer B in Buffer A (Buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; Buffer B; 20% water, 80% acetonitrile, 0.1% formic acid). The flow rate through the column was set to 400 nL/min and the spray voltage was set to 3.5 kV. One full MS scan (FTMS) (400-1800 MW) was followed by 7 data dependent scans (ITMS) of the nth most intense ion with dynamic exclusion. The tandem MS data were searched using the SEQUEST algorithm using a concatenated target/decoy variant of the human IPI databases. A static modification of +57.02146 on cysteine was specified to account for iodoacetamide alkylation and differential modifications were specified on lysine to account for probe modification. Modification searches on cysteine and tyrosine were performed. SEQUEST output files were filtered using DTASelect 2.0.

ALDOA Activity Assay

Adapted from Bergmeyer, H.U. (1974) Methods of Enzymatic Analysis, Second Edition, Volume I, 430

Purified recombinant ALDOA was diluted in PBS. ALDOA samples were then incubated with probes at various concentrations or DMSO for 1 hr. For a final 100 μL reaction mix, F1,6BP (0.19 μM, 30X stock), Tris HCl Buffer pH 7.4 (90 mM, 100 mM

stock,), NADH (0.13 mM, 30X stock), TPI (1 unit), and GAPDH (5 units), were mixed together and allowed to pre-incubate for 10 min. The reaction was initiated with the addition of ALDOA (0.16 μM, 4X stock). Reaction was monitored by kinetic mode every 30 seconds for 5 minutes at an absorbance of 340 nm.



Scheme 2-2. Coupled activity assay monitors the GAPDH catalyzed consumption of NADH as a result of the production of GADP by ALDOA and TPI.

References

(1) Sokalingam, S.; Raghunathan, G.; Soundrarajan, N.; Lee, S.-G. A Study on the Effect of Surface Lysine to Arginine Mutagenesis on Protein Stability and Structure Using Green Fluorescent Protein. *PLoS ONE* 2012, *7* (7), e40410.

(2) Isom, D. G.; Castaneda, C. A.; Cannon, B. R.; B., B. Large Shifts in pKa Values of Lysine Residues Buried inside a Protein. *Proceedings of the National Academy of Sciences* 2011, *108* (13), 5260–5265.

(3) Czerwinski, R. M.; Harris, T. K.; Massiah, M. A.; Mildvan, A. S.; Whitman, C. P. The Structural Basis for the Perturbed pKa of the Catalytic Base in 4-Oxalocrotonate

Tautomerase: Kinetic and Structural Effects of Mutations of Phe-50†. *Biochemistry* 2001, *40* (7), 1984–1995.

(4) Ishikita, H. Origin of the pKa Shift of the Catalytic Lysine in Acetoacetate Decarboxylase. *FEBS Letters* 2010.

(5) Paetzel, M.; Dalbey, R. E. Catalytic Hydroxyl/amine Dyads within Serine Proteases. *Trends in Biochemical Sciences* 1997, *22* (1), 28–31.

(6) Slilaty, S. N.; Little, J. W. Lysine-156 and Serine-119 Are Required for LexA Repressor Cleavage: A Possible Mechanism. *Proceedings of the National Academy of Sciences USA* 1987, *84*, 3987–3991.

(7) Mildvan, A. S.; Kobes, R. D.; Rutter, W. J. Magnetic Resonance Studies of the Role of the Divalent Cation in the Mechanism of Yeast Aldolase. *Biochemistry* 1971, *10* (7), 1191–1204.

(8) Lebherz, H. G.; Rutter, W. J. Distribution of Fructose Diphosphate Aldolase Variants in Biological Systems. *Biochemistry* 1969, *8* (1), 109–121.

(9) Sygusch, J.; Beaudry, D.; Allaire, M. Molecular Architecture of Rabbit Skeletal Muscle Aldolase at 2.7-A Resolution. *Proceedings of the National Academy of Sciences U S A* 1987, *84* (22), 7846–7850.

(10) Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. Basic Local Alignment Search Tool. *Journal of Molecular Biology* 1990, *215* (3), 403–410.

(11) Stark, C.; Breitkreutz, B.-J.; Reguly, T.; Boucher, L.; Breitkreutz, A.; Tyers, M. BioGRID: A General Repository for Interaction Datasets. *Nucleic Acids Research* 2006, *34* (90001), D535–D539.

(12) Arnold, H.; Pette, D. Binding of Glycolytic Enzymes to Structure Proteins of the

Muscle. *European Journal of Biochemistry* 1968, *6* (2), 163–171.

(13) O'Reilly, G.; Clarke, F. Identification of an Actin Binding Region in Aldolase. *FEBS letters* 1993, *321* (1), 69–72.

(14) Pagliaro, L.; Taylor, D. L. 2-Deoxyglucose and Cytochalasin D Modulate Aldolase Mobility in Living 3T3 Cells. *The Journal of Cell Biology* 1992, *118* (4), 859–863.

(15) Bricknell, O. L.; Daries, P. S.; Opie, L. H. A Relationship between Adenosine Triphosphate, Glycolysis and Ischaemic Contracture in the Isolated Rat Heart. *Journal of Molecular and Cellular Cardiology* 1981, *13* (10), 941–945.

(16) Volker, K. W.; Knull, H. R. A Glycolytic Enzyme Binding Domain on Tubulin. *Archives of Biochemistry and Biophysics* 1997, *338* (2), 237–243.

(17) Kishi, H.; Mukai, T.; Hirono, A.; Fujii, H.; Miwa, S.; Hori, K. Human Aldolase A Deficiency Associated with a Hemolytic Anemia: Thermolabile Aldolase due to a Single Base Mutation. *Proceedings of the National Academy of Sciences of USA* 1987, *84* (23), 8623–8627.

(18) Takasaki, Y.; Takahashi, I.; Mukal, T.; Hori, K. Human Aldolase A of a Hemolytic Anemia Patient with Asp-128→ Substitution: Characteristics of an Enzyme Generated in E. Coli Transfected with the Expression …. *The Journal of Biochemistry* 1990, *180* (2), 153–157.

(19) Esposito, G.; Vitagliano, L.; Costanzo, P.; Borrelli, L.; Barone, R.; Pavone, L.; Izzo, P.; Zagari, A.; Saalvatore, F. Human Aldolase A Natural Mutants: Relationship between Flexibility of the C-Terminal Region and Enzyme Function. *Biochemical Journal* 2004, *380*, 51–56.

(20) Kreuder, J.; Borkhardt, A.; Repp, R.; Pekrun, A.; Göttsche, B.; Gottschalk, U.; Reichmann, H.; Schachenmayr, W.; Schlegel, K.; Lampert, F. Inherited Metabolic Myopathy and Hemolysis Due to a Mutation in Aldolase A. *New England Journal of Medicine* 1996, *334* (17), 1100–1105.

(21) Hurst, J. A.; Baraitser, M.; Winter, R. M.; Opitz, J. M.; Reynolds, J. F. A Syndrome of Mental Retardation, Short Stature, Hemolytic Anemia, Delayed Puberty, and Abnormal Facial Appearance: Similarities to a Report of Aldolase a Deficiency. *American Journal of Medical Genetics* 1987, *28* (4), 965–970.

(22) Miwa, S.; Fujii, H.; Tani, K.; Takahashi, K.; Takegawa, S.; Fujinami, N.; Sakurai, M.; Kubo, M.; Tanimoto, Y.; Kato, T.; Matsumoto, N. Two Cases of Red Cell Aldolase Deficiency Associated with Hereditary Hemolytic Anemia in a Japanese Family. *American Journal of Hematology* 1981, *11* (4), 425–437.

(23) Warburg, O.; Wind, F.; Negelein, E. The Metabolism of Tumors in the Body. *The Journal of general physiology* 1927, *8* (6), 519–30.

(24) Grandjean, G.; Jong, P. R. de; James, B. P.; Koh, M.; Lemos, R.; Kingston, J.; Aleshin, A.; Bankston, L. A.; Miller, C. P.; Cho, E.; Edupuganti, R.; Devkota, A.; Stancu, G.; Liddington, R. C.; Dalby, K. N.; Powis, G. Definition of a Novel Feed-Forward Mechanism for Glycolysis-HIF1α Signaling in Hypoxic Tumors Highlights Aldolase A as a Therapeutic Target. *Cancer Research* 2016, *76* (14), 4259–4269.

(25) Esposito, G.; Imperato, M.; Ieno, L.; Sorvillo, R.; Benigno, V.; Parenti, G.; Parini, R.; Vitagliano, L.; Zagari, A.; Salvatore, F. Hereditary Fructose Intolerance: Functional Study of Two Novel ALDOB Natural Variants and Characterization of a Partial Gene Deletion. *Human Mutation* 2010, *31* (12), 1294–1303.

(26) Santamaria, R.; Esposito, G.; Vitagliano, L.; Race, V.; Paglionico, I.; Zancan, L.; Zagari, A.; Salvatore, F. Functional and Molecular Modelling Studies of Two Hereditary Fructose Intolerance-Causing Mutations at Arginine 303 in Human Liver Aldolase. *The Biochemical journal* 2000, *350*, 823–828.

(27) Lau, J.; Tolan, D. R. Screening for Hereditary Fructose Intolerance Mutations by Reverse Dot-Blot. *Molecular and Cellular Probes* 1999, *13* (1), 35–40.

(28) Esposito, G.; Santamaria, R.; Vitagliano, L.; Ieno, L.; Viola, A.; Fiori, L.; Parenti, G.; Zancan, L.; Zagari, A.; Salvatore, F. Six Novel Alleles Identified in Italian Hereditary Fructose Intolerance Patients Enlarge the Mutation Spectrum of the Aldolase B Gene. *Human Mutation* 2004, *24* (6), 534–534.

(29) Perheentupa, J.; Pitkänen, E. Symptomless Hereditary Fructose Intolerance. *The Lancet* 1962, *279* (7243), 1358–1359.

(30) Ali, M.; Rellos, P.; Cox, T. M. Hereditary Fructose Intolerance. *J. Med. Genet.* 1998, *35* (5), 353–65.

(31) Serero, S.; Maire, P.; Nguyen, V. C.; Cohen-Haguenauer, O.; Gross, M. S.; C; Tand, M. F. de; Kahn, A.; Frezal, J. Localization of the Active Gene of Aldolase on Chromosome 16, and Two Aldolase A Pseudogenes on Chromosomes 3 and 10. *Human genetics* 1988, *78* (2), 167–174.

(32) Henry, I.; Gallano, P.; Besmond, C.; Weil, D.; Mattei, M. G.; Turleau, C.; Boué, J.; Kahn, A.; Junien, C. The Structural Gene for Aldolase B (ALDB) Maps to 9q13----32. *Annals of human genetics* 1985, *49* (Pt 3), 173–80.

(33) Gamblin, S. J.; Davies, G. J.; Grimes, J. M.; Jackson, R. M.; Littlechild, J. A.; Watson, H. C. Activity and Specificity of Human Aldolases. *Journal of Molecular*

Biology 1991, *219* (4), 573–576.

(34) Choi, K. H.; Lai, V.; Foster, C. E.; Morris, A. J.; Tolan, D. R.; Allen, K. N. New Superfamily Members Identified for Schiff-Base Enzymes Based on Verification of Catalytically Essential Residues †. *Biochemistry* 2006, *45* (28), 8546–8555.

(35) Choi, K. H.; Tolan, D. R. Presteady-State Kinetic Evidence for a Ring-Opening Activity in Fructose-1,6-(bis)phosphate Aldolase. *Journal of the American Chemical Society* 2004, *126* (11), 3402–3403.

(36) Lorentzen, E.; Siebers, B.; Hensel, R.; Pohl, E. Mechanism of the Schiff Base Forming Fructose-1, 6-Bisphosphate Aldolase: Structural Analysis of Reaction Intermediates. *Biochemistry* 2005, *44* (11), 4222–4229.

(37) Patricelli, M. P.; Szardenings, A. K.; Liyanage, M.; Nomanbhoy, T. K.; Wu, M.; Weissig, H.; Aban, A.; Chun, D.; Tanner, S.; Kozarich, J. W. Functional Interrogation of the Kinome Using Nucleotide Acyl Phosphates. *Biochemistry* 2007, *46*, 350–358.

(38) Patricelli, M. P.; Nomanbhoy, T. K.; Wu, J.; Brown, H.; Zhou, D.; Zhang, J.; Jagannathan, S.; Aban, A.; Okerberg, E.; Herring, C.; Nordin, B.; Weissig, H.; Yang, Q.; Lee, J.-D.; Gray, N. S.; Kozarich, J. W. In Situ Kinase Profiling Reveals Functionally Relevant Properties of Native Kinases. *Chemistry & Biology* 2011, *18* (6), 699–710.

(39) Kragh-Hansen, U.; Chuang, V. T. G.; Otagiri, M. Practical Aspects of the Ligand-Binding and Enzymatic Properties of Human Serum Albumin. *Biological & Pharmaceutical Bulletin* 2002, *25* (6), 695–704.

(40) Azevedo, C.; Saiardi, A. Why Always Lysine? The Ongoing Tale of One of the Most Modified Amino Acids. *Advances in Biological Regulation* 2016, *60*, 144–150.

(41) McFarland, J. M.; Francis, M. B. Reductive Alkylation of Proteins Using Iridium

Catalyzed Transfer Hydrogenation. *Journal of the American Chemical Society* 2005, *127*, 13490–13491.

(42) Shannon, D. A.; Weerapana, E. Covalent Protein Modification: The Current Landscape of Residue-Specific Electrophiles. *Current opinion in chemical biology* 2015, *24*, 18–26.

(43) Kathman, S. G.; Xu, Z.; Statsyuk, A. V. A Fragment-Based Method to Discover Irreversible Covalent Inhibitors of Cysteine Proteases. *Journal of medicinal chemistry* 2014, *57* (11), 4969–4974.

(44) Hacker, S.; Backus, K.; Lazear, M.; Forli, S.; Correia, B.; Cravatt, B. Global Profiling of Lysine Reactivity and Ligandability in the Human Proteome. *Nat Chem* 2017.

(45) Toi, K.; Bynum, E.; Norris, E.; Itano, H. A. Studies on the Chemical Modification of Arginine. *Journal of Biological Chemistry* 1967, *242* (5), 1036–1043.

(46) Suckau, D.; Mak, M.; Przybylski, M. Protein Surface Topology-Probing by Selective Chemical Modification and Mass Spectrometric Peptide Mapping. *Proceedings of the National Academy of Sciences of the United States of America* 1992, *89* (12), 5630–5634.

(47) Mendoza, V. L.; Vachet, R. W. Probing Protein Structure by Amino Acid-specific Covalent Labeling and Mass Spectrometry. *Mass Spectrometry Reviews* 2009, *28* (5), 785–815.

(48) Chabot, N.; Vinatier, V.; Gefflaut, T.; Baudoin, C.; Rodriguez, F.; Blonski, C.; Hoffmann, P. Irreversible Inhibition of Aldolase by a Phosphorylated A-Dicarbonyl Compound. *Journal of Enzyme Inhibition and Medicinal Chemistry* 2008, *23* (1), 21–27.

(49) Leesnitzer, L. M.; Parks, D. J.; Bledsoe, R. K.; Cobb, J. E.; Collins, J. L.; Consler, T. G.; Davis, R. G.; Hull-Ryde, E. A.; Lenhard, J. M.; Patel, L.; Plunket, K. D.; Shenk, J. L.; Stimmel, J. B.; Therapontos, C.; Willson, T. M.; Blanchard, S. G. Functional Consequences of Cysteine Modification in the Ligand Binding Sites of Peroxisome Proliferator Activated Receptors by GW9662. *Biochemistry* 2002, *41* (21), 6640–6650.

(50) Shannon, A.; Banerjee, R.; Webster, E.; Bak, D.; Wang, C.; Weerapana, E. Investigating the Proteome Reactivity and Selectivity of Aryl Halides. *Journal of the American Chemical Society* 2014, *136* (9), 3330–3333.

(51) Cha, Y.; Han, M.-J.; Cha, H.-J.; Zoldan, J.; Burkart, A.; Jung, J.; Jang, Y.; Kim, C.-H.; Jeong, H.-C.; Kim, B.-G.; Langer, R.; Kahn, R. C.; Guarente, L.; Kim, K.-S. Metabolic Control of Primed Human Pluripotent Stem Cell Fate and Function by the miR-200c-SIRT2 Axis. *Nature Cell Biology* 2017, *19* (5), 445–456.

(52) Morris, A. J.; Tolan, D. R. Lysine-146 of Rabbit Muscle Aldolase Is Essential for Cleavage and Condensation of the C3-C4 Bond of Fructose 1,6-Bis(phosphate). *Biochemistry* 2002, *33* (40), 12291–12297.

Chapter 3

Mass Spectrometry Strategies to Address Experimental Challenges

A significant portion of the work described in this chapter has been published in:

Quinti, L., Dayalan Naidu, S., Träger, U., Chen, X.; Kegel-Gleason, K., Llères, D., Connolly, C., Chopra, V., Low, C., Moniot, S., Sapp, E., Tousley, A. R., Vodicka, P., Kanegan, M. J. Van, Kaltenbach, L. S., Crawford, L. A., Fuszard, M., Higgins, M., Miller, J. R. C. R., Farmer, R. E., Potluri, V., Samajdar, S., Meisel, L., Zhang, N., Snyder, A., Stein, R., Hersch, S. M., Ellerby, L. M., Weerapana, E., Schwarzschild, M. A., Steegborn, C., Leavitt, B. R., Degterev, A., Tabrizi, S. J., Lo, D. C., DiFiglia, M., Thompson, L. M., Dinkova-Kostova, A. T., Kazantsev, A. G. 2017. KEAP1-Modifying Small Molecule Reveals Muted NRF2 Signaling Responses in Neural Stem Cells from Huntington's Disease Patients. Proc. Natl. Acad. Sci. U.S.A., 114, 23, E4676–E4685.

Italia, J. S., Addy, P. S., Wrobel, C. J. J., Crawford, L. A., Lajoie, M. J., Zheng, Y., Chatterjee, A. 2017. An Orthogonalized Platform for Genetic Code Expansion in Both Bacteria and Eukaryotes. Nat. Chem. Biol. 13, 4, 446–450.

Introduction

In Analytical Chemistry, mass spectrometry is used to determine the mass to charge ratio (m/z) of an ion for a chemical compound. The m/z can be used to infer the mass of the parent compound, and thus the chemical composition and identity. While the field of mass spectrometry only started in the late 19th century, the popularity and utility of the technology has made significant advances during that time. [1] Most notably, half of the 2002 Nobel Prize in Chemistry was awarded to two individuals, John B. Fenn [2] and Kiochi Tanaka [3], for their individual contributions to the mass spectrometry field which have facilitated advances in the developing field of Chemical Biology.

In the most basic form, mass spectrometers are comprised of three different components: ion source, mass analyzer, and detector. During sample analysis, the ion source converts the analyte into charged ions. Ions are then attracted into the mass spectrometer with an applied electrical current. Some common ionization sources are electrospray ionization (ESI), chemical ionization, and matrix-assisted laser desorption ionization (MALDI). Next, the mass analyzer selects and sorts the ions based on their m/z. Some common mass analyzers are time-of-flight (TOF), quadrupoles, and ion traps such as: quadrupole ion trap and orbitrap. [4] Lastly, the detector senses a change in current and interprets the signal from the ions. Detectors are usually an electron multiplier (EM) or in the case of ion traps, metal plates in the mass analyzer itself can be used to detect changes in voltage.

The development of tandem mass spectrometry increased the utility and scope of the experiments available to be applied to complex samples and molecules, particularly those of biological origin. Further expansion of utility came when chromatography, such

as gas chromatography (GC) or liquid chromatography (LC), was coupled to the ion source, providing a dimension of analytical separation to the analysis of complex samples. In LC tandem mass spectrometry (LC-MS/MS) complex samples are first separated by LC and are then typically ionized by either ESI or low flowrate nanoESI. [5] Once the ions enter the mass spectrometer the m/z is determined for the ions eluting at that particular time, known as the precursor or parent ions, which is defined as MS1. After determination of the MS1, select ions from the MS1 are targeted for fragmentation. Fragmentation can occur by different methods both chemically and physically, however a popular fragmentation technique is low-energy collision-induced dissociation, which favors cleavage of the amide bond of the peptide backbone and generates b and y ions. [6] The spectra of the m/z of the b and y ions generated by fragmentation of a singular ion from the MS1 is called an MS2. Ions targeted for fragmentation can be predetermined or determined experimentally by indicating the number of ions from MS1 desired to be targeted for generation of a fragmentation spectra (MS2).

The power of this technology is rooted in the ability to determine the amino acid sequence of the peptide precursor ion. For example, consider the peptide N-ACDEF-C, which has a mass of 583.1948. Then consider the peptide N-DEFAC-C, which also has a mass of 583.1948. While the mass of these two peptides is the same, this data does not confer information about the sequence of the peptide. In data analysis, the absence of amino acid sequence information makes correlating the protein of origin in a complex proteome impossible. However, alignment of the unique sequence dependent characteristics of the fragmentation spectra of the b and y ions of both the experimentally determined MS2 and a theoretical MS2 make data analysis feasible (Table 3-1).

m/z of B^+	B ion #	ACDEF		Y ion #	m/z of Y^+
72.04444	1	A	CDEF	4	513.16502
175.05362	2	AC	DEF	3	410.15584
290.08056	3	ACD	EF	2	295.12889
419.12316	4	ACDE	F	1	166.0863

m/z of B^+	B ion #	DEFAC		Y ion #	m/z of Y^+
116.03426	1	D	EFAC	4	469.17519
245.07686	2	DE	FAC	3	340.1326
392.14527	3	DEF	AC	2	193.06419
463.18238	4	DEFA	C	1	122.02707

Table 3-1. Calculated m/z values for b and y ions of theoretical peptides with identical precursor masses.

Here we report the use of mass tandem spectrometry based experimental strategies to solve experimental challenges in two collaborations. Implementation of tandem mass spectrometry in these studies gave access to critical data which enabled experimental validation and optimization.

The first collaboration involves the use a proteome wide profiling of quantitative cysteine reactivity to investigate the promiscuity of the cysteine reactive small molecule, MIND4-17, which was found to react with a cysteine on the protein Kelch-like ECH-associated protein 1 (KEAP1). Here, tandem mass spectrometry was used to determine the additional protein targets of MIND4-17 as an assessment of the promiscuity of the cysteine reactive small molecule. The proteome scale analysis supports the proposed mechanism of action of MIND4-17.

The second collaboration involves the determination of undesired natural amino acid incorporation into sfGFP, due to mischarged orthogonal tRNA, at an engineered internal nonsense codon. Here, tandem mass spectrometry enabled identification of key players involved in the observed cross-talk. Identification of the natural amino acid incorporated at the nonsense position enabled experimental optimization to overcome and eliminate cross-talk. This optimization enhanced the utility and validated the novel strategy for unnatural amino acid incorporation in bacterial and eukaryotic cells.

KEAP1 tightly regulates nuclear factor-erythroid 2 p45-related factor 2 (NRF2) protein levels, thus NRF2 signaling ability. In the cytoplasm KEAP1 acts as a regulator by serving as an adaptor protein between NRF2 and CUL3-based E3 ubiquitin ligase. [7–9] In the absence of stress, KEAP1 efficiently promotes the polyubiquitination of NRF2, leading to its degradation by the ubiquitin–proteasome degradation system (UPS). Polyubiquitination is mediated by the association of the KEAP1 Kelch domain with the amino-terminal Neh2 domain of NRF2 (Figure 3-1). [10–13] However, in the presence of stress, NRF2 degradation is impaired by a conformational change in the KEAP1/NRF2/CUL3 complex that occurs when oxidants or electrophilic chemicals modify the sensory cysteines of KEAP1. [14–17] Human KEAP1 has 27 cysteine residues, and the pattern of cysteine residues modified by a particular electrophilic molecule is called the cysteine code. [18] It is unclear which or if all 27 cysteines are sensory cysteine residues.

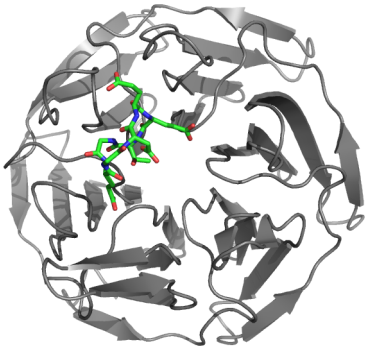


Figure 3-1. Kelch domain of KEAP1 (aas 325-609) shown in gray with the Neh2 domain of NRF2 (shown as sticks) (PBD: 3ZGC). Currently, there is no crystal structure of full length KEAP1.

Since modification of KEAP1 results in the impaired degradation of NRF2, during times of stress NRF2 begins to accumulate and eventually translocates to the nucleus. Once in the nucleus, NRF2 binds to antioxidant response elements (AREs) which activate transcription responses as a result to the stress sensed by KEAP1. [15] Phase II enzymes are transcribed as a result of NRF2 signaling, among many additional downstream proteins. Phase II enzymes, such as glutathione S-transferases (GSTs), NAD(P)H:quinone oxidoreductase 1 (NQO1), heme oxygenase (decycling) 1 (HO-1), and catalytic (GCLC) and regulator (GCLM) subunits of glutamate-cysteine ligase, are cellular protector proteins against carcinogens and oxidants through oxidation, reduction, hydrolysis, or conjugation reactions. [19,20] More recently, NRF2 is also linked to the mechanism of anti-inflammatory effects.

The pathogenicity of Huntington's disease (HD), an inherited autosomal-dominant and highly penetrant neurodegenerative disorder, commonly manifests itself in oxidative stress and neuroinflammation mechanisms. HD patients experience physical and

psychological changes including disordered movements and cognitive decline. [21] HD results from aggregate of the large protein (350 kDa) huntingtin (HTT) found highly expressed in the brain. Aggregate formation is linked to the expansion (n>37) of a polymorphic trinucleotide repeat sequence $(CAG)_n$ within the gene encoding HTT, and is not observed with wild type HTT. [22] Observed neuroinflammation in HD patients may be a result of not fully engaged or inhibited NRF2 signaling since all ARE gene products are not responsively upregulated in HD patient brain tissue. [23] Further evidence from HD mouse model suggests that protective NRF2 signaling is not fully activated and/or muted. Studies have found pharmacological stimulation of NRF2s in HD mouse model is efficacious and associated with increased expression of antioxidant proteins and reduction of ROS levels in brain. [24,25] Recent work of our collaborator identified a unique structural scaffold of triazole-containing small molecules that potently induces expression of canonical ARE genes (NQO1, GCLM, GCLC, HMOX1/HO-1) in primary mouse cells and causes a pronounced reduction of ROS. [26]

The work presented here is a continuation of the previous publication. Here, we elucidated the molecular mechanism by which the most potent lead compound, MIND4-17, developed from this scaffold induces NRF2 activation, and used this highly selective probe to examine specific activation responses and potential benefits of NRF2 signaling in mouse and human HD models. More specifically, my contribution to the work was to evaluate the selectivity of MIND4-17 towards KEAP1 through a global proteomic screening using a broadly reactive cysteine ABP, iodoacetamide alkyne.

The second collaboration highlighted in this chapter is focused on the field of unnatural amino acid incorporation. This technique allows researchers to investigate or

modify the structure and function of proteins with control over the specific site and number of modifications. This is an improvement over the use of bioconjugation reactions to incorporate new functional groups into proteins, since traditional bioconjugation methods are limited in the sites available for modification as well as the yield and stoichiometry of modifications per protein, which is dictated by protein sequence. Synthetic incorporation of unnatural amino acids in peptides through solid phase peptide synthesis (SPPS) is the limited to the size of the peptide that can be synthesized (<50 aa) through this method and the ability to ensure proper protein folding.[27] However, the use of semisynthetic methods, allow access to the synthesis of much larger macromolecules. The portion of the protein containing the unnatural amino acid is synthesized using SPPS, and the rest of the protein is synthesized recombinantly. Some of the semisynthetic strategies fall under the category of native chemical ligation: native chemical ligation with homocysteine, Staudinger ligation, intein-mediated method, and through a Click reaction.[28] The drawback of the semisynthetic method is that the protein is not synthesized in a native cell environment, hindering the ability to study the protein in a relevant context. To overcome this drawback, focus has been directed towards translational incorporation of unnatural amino acids. This method results in the synthesis of the modified protein in the context of a cell, where it is more feasible and relevant to perform studies that probe the function and activity of the protein. Enabling a cell to incorporate an unnatural amino acid, thus expanding the genetic code, has been accomplished through various strategies which include stop codon suppressor tRNAs, orthogonal aminoacyl-tRNA synthetases, codon expansion and non-natural base pairs.[28]

In the case of orthogonal aminoacyl-tRNA synthetases, an engineered aminoacyl-tRNA synthetase-tRNA pair is used to deliver an unnatural amino acid to a desired engineered position, such as a nonsense codon. It is essential for the aminoacyl-tRNA synthetase-tRNA pair to be orthogonal to the host's translational machinery. To ensure orthogonality, typically aminoacyl-tRNA synthetase-tRNA pairs are chosen from organisms of a different domain of life, thus contains evolutionarily divergent components. [29] Typically, genetic code expansion of bacterial relies on aminoacyl-tRNA synthetase-tRNA pairs derived from eukaryotes and archaea. Alternatively, genetic code expansion of eukaryotes relies on aminoacyl-tRNA synthetase-tRNA pairs derived from bacteria, since archaeal derived aminoacyl-tRNA synthetase-tRNA pairs are typically cross-reactive in eukaryotes. The use of two distinct sets of aminoacyl-tRNA synthetase-tRNA pairs for genetic code expansion in eukaryotes and bacterial leads to a significant technical disadvantage; the two individual systems must be separately genetically encoded and optimized using the *in vivo* selection systems developed using *E. coli* and *S. cerevisiae* as hosts. The platform for *E. coli* selection is more facile and more successful for genetically encoding new unnatural amino acids than the *S. cerevisiae* platform. The potential for using one system for both bacterial and eukaryote expression has been demonstrated with the use of archaea-derived pyrrolysyl-tRNA synthetase-$tRNA^{Pyl}$, which contains naturally occurring TAG codon pair. The scope of unnatural amino acids that can be incorporated by this pair is limited by the preference for unnatural amino acids with structural similarity to lysine. The focus of this study was to investigate the potential of a novel strategy that would overcome the limitations of the need to use two separate selection systems for one unnatural amino acid for incorporation into proteins in bacterial and eukaryotic cells. The

strategy requires an endogenous bacterial aminoacyl-tRNA synthetase-tRNA pair to be replaced by an aminoacyl-tRNA synthetase-tRNA pair from a eukaryote. This substitution allows the liberated endogenous bacterial aminoacyl-tRNA synthetase-tRNA pair to be available for unnatural amino acid incorporation. Additionally, this strategy allows the use of the preferred, more facile bacterial selection for directed evolution of the liberated endogenous bacterial aminoacyl-tRNA synthetase-tRNA pair.

Results and Discussion

The following results and discussion pertain to the KEAP1/NRF2 project.

Triazole-Containing Inducers Enhance NOQ1 Activity

To elucidate the mechanism of NRF2 activation, comparison of the potencies of lead compound from previous study, MIND4, new candidate MIND4-17 (Figure 3-2A), and a series of structural triazole-containing analogs were assayed in a quantitative NQO1 inducer bioassay in murine Hepa1c1c7 cells (Figure 3-2B). [30] The NQO1 bioassay is a coupled assay that indirectly measures the activity of NQO1 by monitoring the spontaneous electron transfer and colorimetric change of yellow 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) $_{(oxidized)}$ to purple formazan (MTT) $_{(reduced)}$ as a result of the electron transfer from NADPH, generated by GAPDH, by NQO1 to menadione which converts to menadiol. The concentration that doubles the specific activity (CD value) for NQO1 was used as a measure of inducer potency and compound ranking. Within this lead series, MIND4-17 was confirmed to be the inducer of highest potency (CD = 0.15

μM), comparable to the naturally occurring inducer sulforaphane (SFP) (CD = 0.18 μM) and significantly more potent than the clinically approved NRF2 activator dimethyl fumarate (DMF) (CD = 9 μM). This data indicates an increase of NQO1 activity as a result of NRF2 activation.

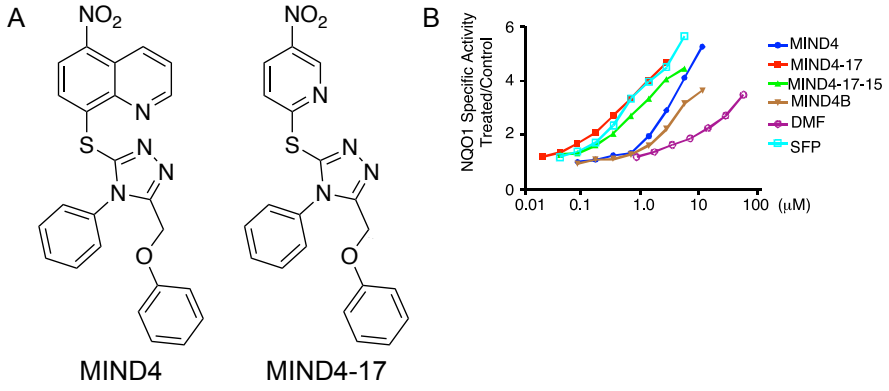


Figure 3-2. (A) Structures of triazole-containing compounds parent MIND4 and lead-inducer MIND4-17 (B) Concentration-dependent activity test (n=8) of MIND4-17, structural analogs MIND4-17-15, MIND4, MIND4B, and already known, control NRF2 inducers SFP and DMF in quantitative NQO1 inducer bioassay.

NRF2 Inducer Treatment Enhances NRF2 Induction Products

Since NQO1 transcription is a product of NRF2 activation, focus was directed to investigate whether the observed increase in NQO1 activity was correlated to increased levels of NQO1 and GCLM, another product of NRF2 activated transcription. Western blotting of NQO1 and GCLM proteins after treatment for 24 hrs with MIND4-17 and DMF in mutant HD rat embryonic striatal cells ST14A shows an increase in protein expression, indicating the compounds induce NRF2 activation (Figure 3-3A). The observed increase in products of NRF2 activated transcription in the HD mouse model further suggests the NRF2 system is not functioning at a suboptimal capacity. To further investigate the

mechanism of apparent NRF2 activation, we screened the effects of MIND4-17 and other MIND4 analogs on induction of NQO1 expression NRF2-null mouse embryonic fibroblasts (MEF) cells and which did not result in an increase of NQO1. These data together support the hypothesis that observed NOQ1 induction by the compounds and increased activity are NRF2 dependent (Figure 3-3B).

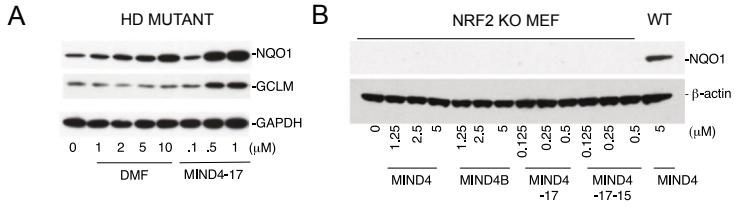


Figure 3-3. (A) Evaluation of NRF2-responsive NQO1 and GCLM protein levels in MIND4-17 and DMF-treated mutant HD ST14A cells. (B) Evaluation of NRF2-responsive protein levels indicates no observed induction of NQO1 in NRF2-KO MEFs treated with MIND4-17 and its analogs compared to control.

NRF2 is Stabilized as a Result of Treatment with NRF2 Inducing Compounds

Since stabilization of NRF2 is an essential step in the activation of the pathway and a marker of redirection from UPS, the stabilization of NRF2 in cells treated with MIND4-17 and its structural analogs preceded the induction of NQO1 was accessed. A 5 hr exposure to SFP, MIND4-17, and lower potency inducers MIND4 and MIND4B resulted in accumulation of detectable levels of NRF2 protein (Figure 3-4A). To determine the response time of NRF2 activation by the compounds, cells treated with MIND4-17 at different time points were analyzed for NRF2 protein levels. Nuclear accumulation of NRF2 was observed as early as 30 min after cell exposure to MIND4-17 (0.5 μM) and remained elevated for at least 5 hr (Figure 3-4B).

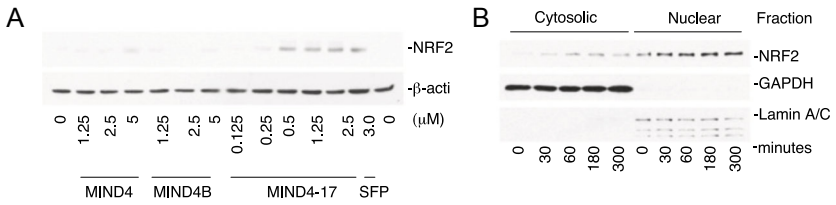


Figure 3-4. (A) NRF2 accumulation is concentration dependent. MIND4-17 showed the most NRF2 accumulation. (B) NRF2 in cytoplasmic and nuclear fractions from WT MEFs treated with 0.5 μM MIND4- 17 show time-dependent accumulation.

MIND4-17 Promiscuity Assessed by Quantitative Cysteine Reactivity Profiling

Next, focus was directed towards further conformation that the induction of antioxidant response elements are specifically due to NRF2 activation as a result of the singular cysteine modification of KEAP1 and not an unidentified off target. To determine whether MIND4-17 promiscuously alkylated other reactive cysteines within the proteome, a quantitative cysteine reactivity profiling strategy was used to globally identify cysteine residues that demonstrate significant loss of reactivity after preincubation with compound, MIND4-17 (Figure 3-5). [31] Following MIND4-17 (10 μM) treatment HeLa cell soluble lysate (2.0 mg/mL) samples were treated with a promiscuous cysteine reactive iodoacetamide-alkyne (IA) probe (100 μM). Excess IA probe was removed by methanol washes and IA probe reacted proteins were coupled to biotinylated isotopic cleavable linkers through CuAAC (Figure 3-6). Proteins modified by IA probe were enriched on streptavidin beads overnight and unreacted proteins were removed by wash steps. Bead bound proteins were subjected to on bead trypsin digestion overnight. In solution tryptic peptide digests were removed and set aside, leaving peptides containing the cysteine residue modified by IA probe remaining bound to the beads. To elute the peptides of

interest beads were treated three times with sodium dithionite solution that cleaves the diazo linkage of the cleavable linker, and the resulting washes were combined with final water wash steps to yield the final sample. [32]

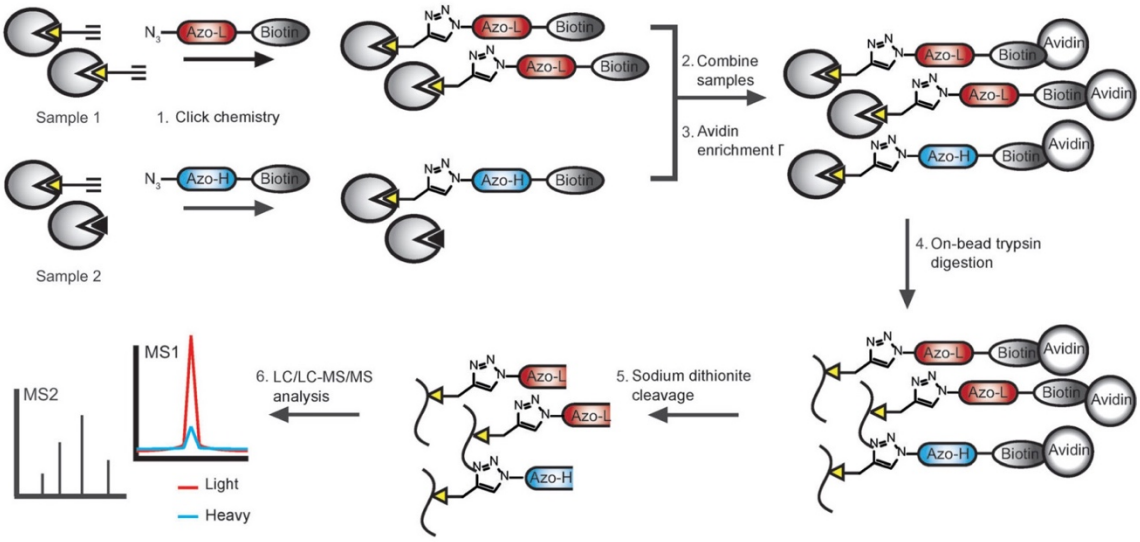


Figure 3-5. Quantitative MS workflow. Light samples were pretreated with MIND4-17 (10 μM) and heavy samples were pretreated with DMSO. Both light and heavy samples were then treated with IA probe (100 μM). Figure adapted from Qian Y, et al. 2013.

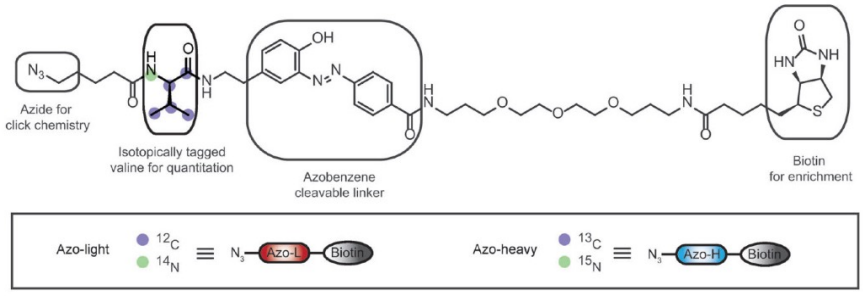


Figure 3-6. Structure of isotopically labeled Azo-L and Azo-H tags. Incorporation of the heavy valine residue provides a mass difference of 6 Da between Azo-L and Azo-H. Figure adapted from Qian Y, et al. 2013.

Of the >300 reactive cysteine-containing peptides identified in our quantitative MS studies, only four residues showed a greater than twofold change in cysteine reactivity upon treatment with MIND4-17 (Figure 3-7, Table 3-2, Table 3A-1). The small subset of sensitive cysteine residues identified in the global quantitative analysis experiment suggests MIND4-17 has a high level of selective reactivity towards KEAP1. Within the subset of cysteines with no change in reactivity were hyper-reactive cysteine residues such as the active-site nucleophile of GAPDH, attesting to the low reactivity of the MIND4-17 across even highly reactive cysteines within the proteome.

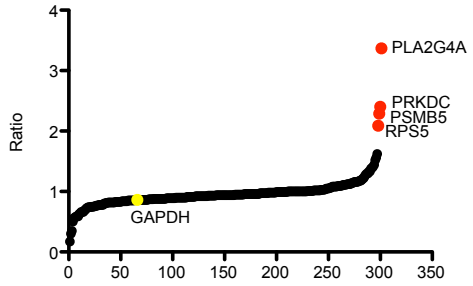


Figure 3-7. Plot of cysteine reactivity ratios identified in cysteine proteome profiling after MIND4-17 pretreatment. The four proteins with ratios above 2 (indicated in red) are Cytosolic phospholipase A2 (PLA2G4A), DNA-dependent protein kinase catalytic subunit (PRKDC), Proteasome subunit beta type-5 (PSMB5), and 40S ribosomal protein S5 (RPS5). Highly reactive GAPDH active site cysteine (indicated in yellow) has a ratio of 0.86.

IPI	Protein	Symbol	Peptide	Ratio
P47712	Cytosolic phospholipase A2	PLA2G4A	CSVSLSNVEAR	3.37
P78527	DNA-dependent protein kinase catalytic subunit	PRKDC	INQVFHGSCITEGNELTK	2.4
P28074	Proteasome subunit beta type-5	PSMB5	VIEINPYLLGTMAGGAADCSFWER	2.29
P46782	40S ribosomal protein S5	RPS5	AQCPIVER	2.09

Table 3-2. Peptides with ratios greater than 2.0 which indicate sensitivity to MIND4-17 treatment.

We note that we did not detect any cysteine containing peptides from KEAP1. Tryptic KEAP1 peptide containing C319 has been previously detected using this method (IA probe, 100 μM) in MDA-MB-231 and MCF7 cells, but not in Jurkat cells. [31] Independent studies have shown expression of KEAP1 in HeLa cells is much less than in MDA-MB-231 and MCF7 cells by Western Blot. [33] The low amounts of KEAP1 protein in the sample may account for this result. Future work to further investigate the targets of alkylation by MIND4-17 it would be beneficially to use cell lysates from a HD model cell line, as this will more closely mimic protein expression levels and reactivity characteristic of the disease.

The following results and discussion pertains to the unnatural amino acid project.

Functionally Replacing the TrpRS-tRNA^{Trp} pair of E. coli

Focus was directed to the development of the tryptophanyl-tRNA synthetase (TrpRS-tRNA^{Trp}) pair for genetic code expansion of *E. coli* and eukaryotes (Figure 3-8).

For this strategy to be applicable for unnatural amino acid incorporation into eukaryotes, it was necessary to confirm successful incorporation and orthogonality of TAG-suppressing *E. coli* $tRNA^{Trp}$ ($tRNA^{EcTrp}_{CUA}$) in eukaryotic cells. To evaluate incorporation, TAG-suppressing *E. coli* $tRNA^{Trp}$ ($tRNA^{EcTrp}_{CUA}$) and an enhanced green fluorescent protein (EGFP) reporter containing a stop codon at a permissive site (EGFP-39-TAG) was co-expressing in HEK293T, with or without the EcTrpRS synthetase. EGFP expression was only observed with EcTrpRS present confirming TAG suppression efficiency and supports no cross-reactivity in eukaryotic cells (Figure 3-9). Focus was then directed towards the substitution and subsequent deletion of the TrpRS-$tRNA^{Trp}$ pair.

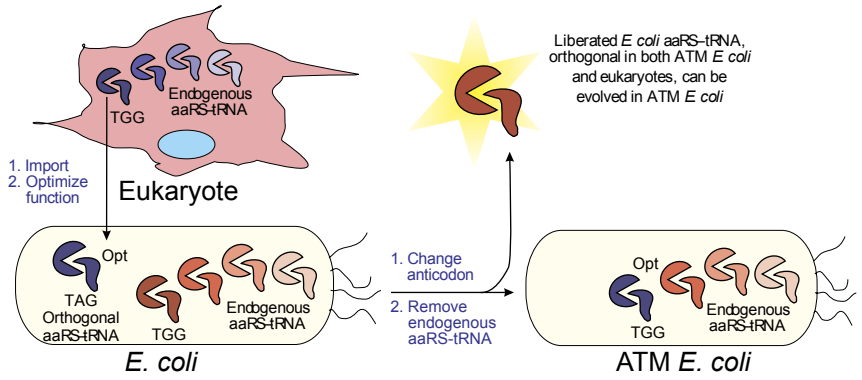


Figure 3-8. Strategy to create an aaRS-tRNA pair that can drive genetic code expansion in both eukaryotes and *E. coli* by liberating the endogenous *E. coli* TrpRS-$tRNA^{Trp}$.

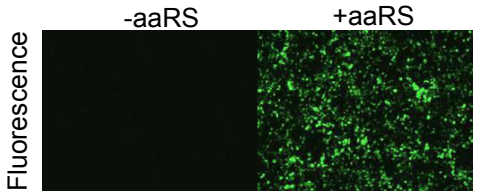


Figure 3-9. Fluorescence microscopy images of HEK293T cells co-expressing EGFP-39-TAG reporter and $tRNAEc^{Trp}_{CUA}$ with and without co-expression of the cognate EcTrpRS. Robust EGFP-39-TAG expression is observed only when EcTrpRS is co-expressed. This confirms no cross-reactivity of $tRNAEc^{Trp}_{CUA}$.

Previously, an *S. cerevisiae*-derived ScTrpRS-tRNA^{ScTrp} pair has been optimized in *E. coli* for highly efficient nonsense suppression. A plasmid containing ScTrpRS-tRNA$^{ScTrp}_{CCA}$ was engineered and incorporated which then allow for deletion of *trpS* (encoding EcTrpRS) and *trpT* (encoding tRNA$^{Trp}_{CAA}$), resulting in the ATMW1 strain. Growth rate studies of the ATMW1 strain compared to its progenitor EcNR1G strain showed no inhibition of growth, confirming efficient functional substitution by the engineered pair from *S. cerevisiae* (Figure 3-10).

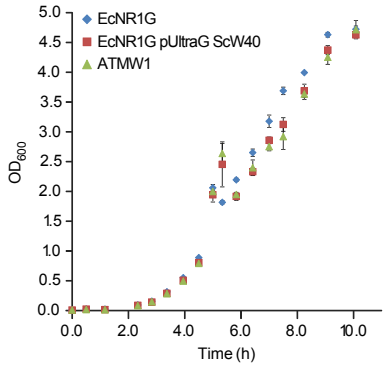


Figure 3-10. ATMW1 exhibits comparable growth rate to progenitor EcNR1G strain containing ScTrpRS-tRNA^{ScTrp} complementation plasmid.

Reintroducing the EcTrpRS-tRNA^{EcTrp} pair into ATMW1

Next, the potential to reintroduce the liberated EcTrpRS-tRNA^{EcTrp} pair was tested. The tRNA^{EcTrp} was supplied as a nonsense suppressor for TAG, however previous reports indicate TAG suppressor tRNA$^{EcTrp}_{CUA}$ is mischarged by *E. coli* glutamyl-tRNA synthetase (EcGlnRS). [34,35]Although tRNA^{EcTrp} and tRNA^{EcGln} have very little sequence homology, the observed mischarge is likely due to the overall structural similarity of the tRNA in addition to the similarity of the anticodons, CUA and CUG, for the TAG suppressor tRNA$^{EcTrp}_{CUA}$ and tRNA^{EcGln} respectively (Figure 3-11). Knowing the precedence of cross-

reactivity, we sought to evaluate the impact on protein expression of a superfolder GFP (sfGFP) reporter in the ATMW1 strain (Figure 3-12). The high levels of sfGFP expression in the ATMW1 strain with $tRNA^{EcTrp}_{CUA}$ indicate cross-talk. However, when the codon is switched to TGA (opal) from TAG (amber), and the anticodon of the $tRNA^{EcTrp}$ is changed from CUA to UCA in the absence of EcTrpRS there is no expression of sfGFP, indicating a noncross-talk system.

Figure 3-11. Overlay of $tRNA^{EcTrp}$ (blue, PDB: 4YCP) and $tRNA^{EcGln}$ (red, PDB: 1QTQ) show high structural similarity despite low sequence homology.

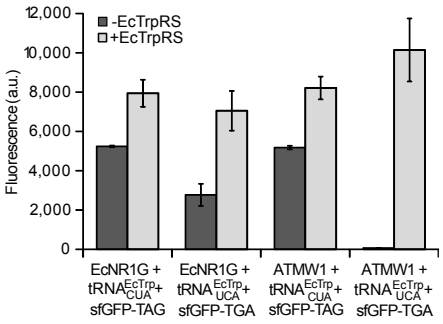


Figure 3-12. Evaluating the nonsense-suppression efficiencies of $tRNA^{EcTrp}_{CUA}$ and $tRNA^{EcTrp}_{UCA}$ using sfGFP-151-TAG or sfGFP-151-TGA as reporters.

To identify the synthetase mischarging the $tRNA^{EcTrp}_{CUA}$ sfGFP-151-TAG was purified from ATMW1 $tRNA^{EcTrp}_{CUA}$ expressed in the absence of EcTrpRS. The sfGFP-151-TAG protein was purified by Ni-NTA affinity purification. The purified sfGFP-151-TAG was analyzed by whole protein mass spectrometry and tryptic digestion mass spectrometry (LC-MS/MS). The mass wild-type sfGFP is 27617 Da, however the observed mass by whole protein mass spectrometry is 27562 Da (Figure 3-13). The observed mass indicates the misincorporation of a Q, E, or K residue into sfGFP at the nonsense codon (Table 3-3). Incorporation of W is not expected due to the absence of EcTrpRS. The expected mass of sfGFP-151-TAG with successful tryptophan incorporation is 27620 Da. The misincorporation of a natural amino acid can be traced to a mischarged $tRNA^{EcTrp}_{CUA}$ by EcGlnRS, EcGluRs, or EcLysRS.

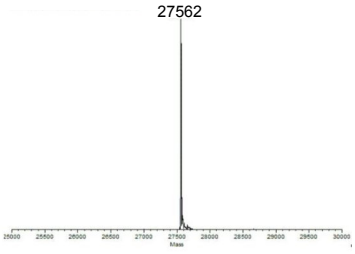


Figure 3-13. Deconvoluted mass of purified whole sfGFP co-expressed in ATMW1 with $tRNA^{EcTrp}_{CUA}$ in the absence of EcTrpRS.

Tryptic Peptide	Calculated Mass of sfGFP (Da)
LEYNFNSHNVQITADK	27562
LEYNFNSHNVEITADK	27563
LEYNFNSHNVK	27562
LEYNFNSHNVWITADK	27620

Table 3-3. Table of tryptic peptides for natural amino acid residues potentially misincorporated into sfGFP-151-TAG and the corresponding calculated mass of sfGFP with the misincorporation.

The resolution of whole protein mass spectrometry presented an experimental challenge. Whole protein mass spectrometry was able to confirm full length sfGFP was being produced, but was not able to differentiate the exact natural amino acid being incorporated into the 151 site. This presents a significant challenge for further optimization of the system, since identification of the EcRS involved in the cross talk with $tRNA^{EcTrp}_{CUA}$ is a critical player. To overcome this challenge, sfGFP-151-TAG expressed in the absence of EcTrpRS was analyzed by tryptic digestion tandem mass spectrometry (Figure 3-14).

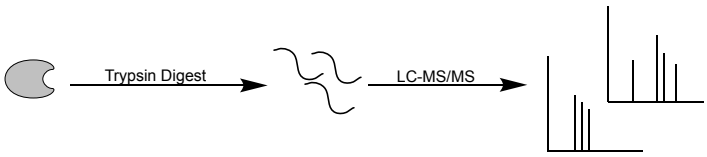


Figure 3-14. Workflow if tryptic digest tandem protein mass spectrometry analysis of sfGFP-151-TAG.

EcGlnRS Mischarges $tRNA^{EcTrp}_{CUA}$

First we searched for the m/z values, within window of 5 ppm, for the predicted tryptic peptides (Figure 3-15A). Extraction of the calculated glutamine containing peptide m/z resulted in a peak characteristic of peptide elution in the chromatogram, though there is a small signal in the same region for the expected glutamate containing peptide. To confirm the signal is indeed the glutamine containing peptide, we extracted the average MS1 spectra for the time corresponding to the elution peak. Examination of the average MS1 isotopic envelope revealed the m/z of the $[M+2]^{2+}$ peptide to be 946.9574 (Figure 3-15B). This value corresponds to the calculated m/z for the $[M+2]^{2+}$ of the glutamine containing peptide (Table 3-4). To further confirm the identity of the amino acid in position 151 of sfGFP we analyzed the fragmentation spectra (MS2) of the 946.9574 ion (Figure 3-16). The high accuracy of the orbitrap mass analyzer in conjunction with the data from the fragmentation spectra allowed us to identify the natural amino acid incorporated into the position corresponding to the nonsense TAG codon as glutamine.

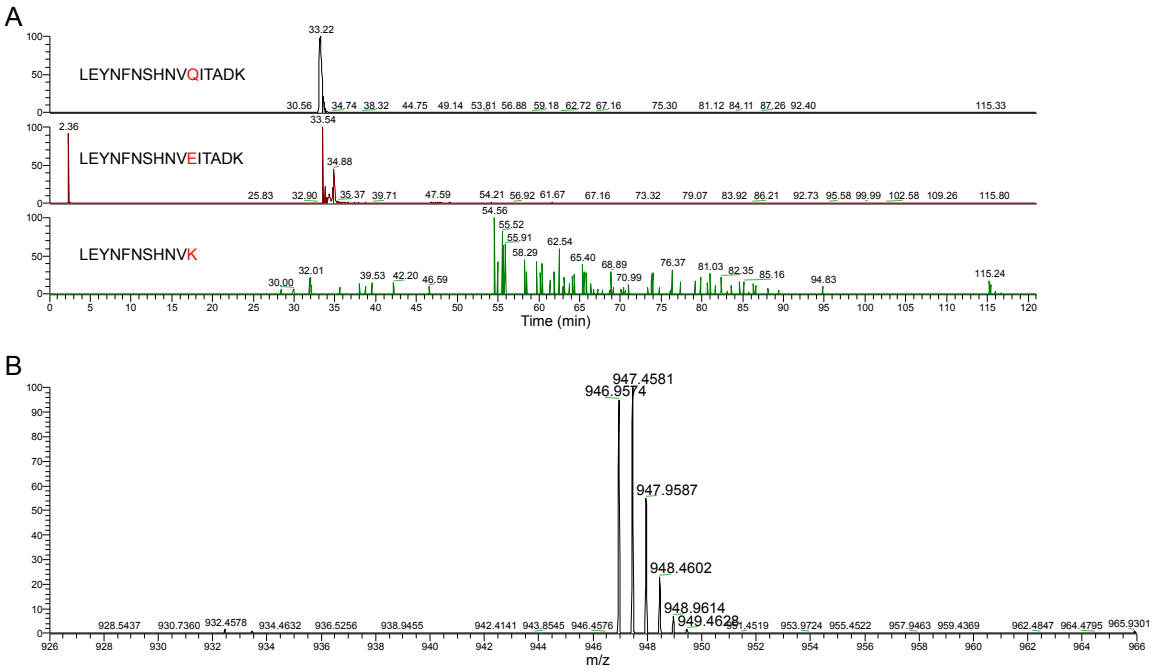


Figure 3-15. (A) Extracted masses from LC-MS/MS chromatogram for the $[M+2]^{2+}$ tryptic peptides for the three potentially incorporated natural amino acids in sfGFP-151-TAG. (B) Isotopic envelop of the extracted average MS1 of the peak indicates the m/z of the $[M+2]^{2+}$ tryptic peptide is 946.9574 which corresponds to the calculated m/z of the tryptic peptide with glutamine incorporation (946.9605).

Trypic Peptide	[M]	$[M+1]^{+}$	$[M+2]^{2+}$	$[M+3]^{3+}$
LEYNFNSHNVQITADK	1891.9065	1892.9138	946.9605	631.6428
LEYNFNSHNVEITADK	1892.8905	1893.8978	947.4526	631.9708
LEYNFNSHNVK	1363.6521	1364.6594	682.8334	455.5580

Table 3-4. Calculated m/z values for the tryptic peptides of the predicted natural amino acid incorporations.

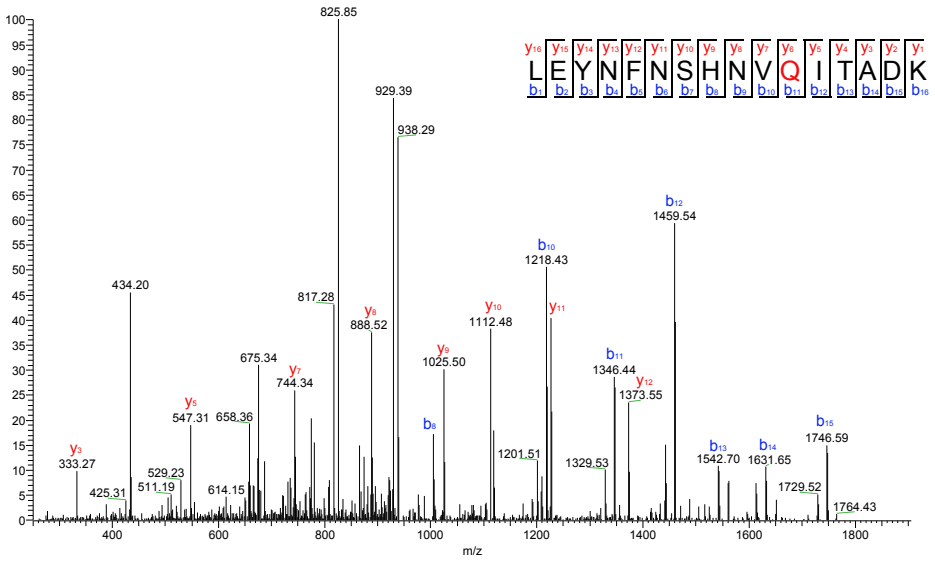


Figure 3-16. Annotation of the MS2 fragmentation spectra of the 946.9574 ion indicates glutamine is incorporated as the 151st residue in sfGFP.

Despite the homology between tRNA^{EcTrp} and tRNA^{EcGln}, EcGlnRS only recognizes tRNA^{EcGln} due to the use of the U residing in the anticodon as a major identity element. [34,36]To circumvent this issue the tRNA$^{EcTrp}_{CUA}$ was changed to tRNA$^{EcTrp}_{UCA}$ and tryptophan incorporation at TGA was observed (Figure 3-17).

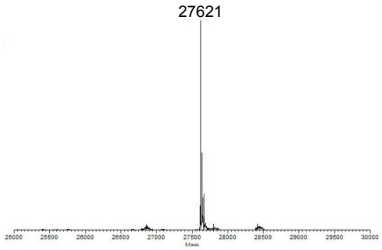


Figure 4-17. Deconvoluted mass of purified whole sfGFP-151-TGA co-expressed in ATMW1 with tRNA$^{EcTrp}_{UCA}$ and EcTrpRS indicates incorporation of tryptophan into the opal suppression 151.

Conclusion

Here we report the tandem mass spectrometry data from two collaborations. In the first collaboration, we evaluated the promiscuity of a cysteine reactive small molecule using a global quantitative cysteine reactivity strategy. We found cysteine residues on four proteins, C726 of PLA2G4A, C1904 of PRKDC, C111 of PSMB5, and C66 of RPS5, to be sensitive to alkylation of MIND4-17. We did not identify KEAP1 as a target of MIND4-17 in our dataset likely due to the low protein abundance of KEAP1 in HeLa cells. Future work using HD model cell lines would enhance the characterization of proteins susceptible to alkylation by MIND4-17. Furthermore, assessment of NRF2 signaling in KO studies of PLA2G4A, PRKDC, PSMB5, RPS5 will better characterize the role of these proteins and the effects of alkylation by MIND4-17 in the enhanced expression of antioxidant enzymes.

In the second collaboration, we take advantage of the high resolution of the orbitrap mass analyzer to accurately identify the sequence of a tryptic peptide from sfGFP. We found a glutamine residue was incorporated into the amber nonsense codon during the translation of sfGFP-151-TAG. This observation correlated with previous work that showed EcGlnRS will charge $tRNA^{EcTrp}_{CUA}$ due to improper recognition of the anticodon. By identifying the mechanism of misincorporation, through peptide sequencing, we were able to make the necessary improvements to the experimental design to eliminate RS-tRNA cross-talk. This drawback was overcome by changing the nonsense codon from amber (TAG) to opal (TGA). Reporter protein sfGFP-151-TGA is only expressed when co-expressed with the reintroduced EcTrpRS-$tRNA^{EcTrp}_{UCA}$ pair.

Acknowledgements

I would like to thank my collaborators James Italia, Abhishek Chatterjee, and Aleksey Kazantsev and the many co-authors for the opportunity to contribute to projects outside the scope of my individual projects.

Experimental Procedures

The following experimental procedures are from:

Quinti, L., Dayalan Naidu, S., Träger, U., Chen, X.; Kegel-Gleason, K., Llères, D., Connolly, C., Chopra, V., Low, C., Moniot, S., Sapp, E., Tousley, A. R., Vodicka, P., Kanegan, M. J. Van, Kaltenbach, L. S., Crawford, L. A., Fuszard, M., Higgins, M., Miller, J. R. C. R., Farmer, R. E., Potluri, V., Samajdar, S., Meisel, L., Zhang, N., Snyder, A., Stein, R., Hersch, S. M., Ellerby, L. M., Weerapana, E., Schwarzschild, M. A., Steegborn, C., Leavitt, B. R., Degterev, A., Tabrizi, S. J., Lo, D. C., DiFiglia, M., Thompson, L. M., Dinkova-Kostova, A. T., Kazantsev, A. G. 2017. KEAP1-Modifying Small Molecule Reveals Muted NRF2 Signaling Responses in Neural Stem Cells from Huntington's Disease Patients. Proc. Natl. Acad. Sci. U.S.A., 114, 23, E4676–E4685.

Compound Source and Storage

Compounds were procured from ChemBridge Corporation (purity quality control ensured by provided NMR spectroscopy data), dissolved in molecular-biology–grade DMSO to 10 mM stock concentrations, aliquoted, and stored at ⁻80 °C. MIND4-17 was resynthesized (purity, >95%) and showed essentially identical potency of NRF2 activation

to MIND4-17 in multiple batches purchased from ChemBridge. DMF was purchased from Sigma and SFP from LKT Laboratories.

NQO1 Bioassay

Inducer potency was quantified by use of the NQO1 bioassay in Hepa1c1c7 cells as described.[30] In brief, cells grown in 96-well plates were exposed to serial dilutions of each compound for 48 hr, and the NQO1 enzyme activity was determined in cell lysates. Results are shown as average values of eight replicate wells. The SD in each case was less than 5%

Compound Activity Test in the Rat Embryonic Striatal Cell Lines ST14A

ST14A cells (a generous gift from E. Cattaneo, University of Milan, Milan) stably express either a mutant expanded repeat (128Q) or WT (26Q) 546-aa HTT fragment and were treated with compounds for 24 hr. Protein levels of NQO1, GCLM, and α-tubulin (loading control) were detected in ST14A by immunoblotting with NQO1 (Sigma; 1:1,000), GCLM (Abcam; 1:800), and GAPDH (Millipore; 1:10,000) antibodies.

Compound Activity Test in Mouse Embryonic Fibroblasts

MEFs from WT, NRF2- knockout (NRF2-KO), or KEAP1-knockout (KEAP1-KO) mice (59) were cultured in plastic dishes (Invitrogen) coated for 30 min with 0.1% (wt/vol) gelatin. For experiments, cells (250,000 per well) were grown for 24 hr on six-well plates, and then treated with solvent control [0.1% DMSO (vol/vol)] or compounds for 24 h. For Western blot analysis, cells were lysed in radio-immunoprecipitation assay (RIPA) buffer

containing 1 protease inhibitor mixture tablet (Roche) per 10 mL of buffer. Proteins were resolved by SDS/ PAGE and immunoblotted with specific antibodies against NQO1 (1:1,000) or NRF2 (1:1,000), both gifts from John D. Hayes (University of Dundee, Dundee, UK), and β-actin (Sigma; mouse monoclonal; 1:10,000).

Nuclear-Cytoplasmic Extract Separation

WT MEFs (500,000 per dish) were grown for 24 hr on 6-cm plates, and then treated with solvent control [0.1% DMSO (vol/vol)] or 0.5 μM MIND4-17. At the end of each treatment time, cells were washed twice in PBS and lysed in buffer A [10 mM KCl, 5 mM $MgCl_2$, 50 mM Tris·HCl (pH 7.5), 0.5% Nonidet P-40, 1 mM DTT], containing 1 protease inhibitor and 1 phosphatase inhibitor tablets (Roche) per 10 mL of buffer. Lysates were subjected to centrifugation at 1,000 × g for 5 min at 4 °C, and supernatant (cytoplasmic fraction) was transferred to a fresh microfuge tube. The pellet (nuclear fraction) was washed three times in buffer A before being dissolved in buffer B [2% SDS, 150 mM NaCl, 65 mM Tris·HCl (pH 8.0)] and sonicated for 30 s. Finally, the lysates were subjected to Western blot analysis with NRF2 antibody and with GAPDH (rabbit polyclonal; 1:5,000) and lamin A/C (rabbit polyclonal; 1:1,000; Gene Tex) to confirm fraction purity and equal protein loading.

Proteomic Analysis by Quantitative MS of MIND4-17

Preparation of MS samples for reactive cysteine profiling. HeLa soluble protein lysates (500 μL, 2 mg/mL) were aliquoted, and MIND4-17 (10 μM), or DMSO was added to the appropriate samples at the designated concentration. Two aliquots were made for

each inhibitor concentration or DMSO equaling four tubes for one final sample. The samples were incubated at room temperature (RT) for 1 hr. IA-alkyne (100 μM) was added to the samples and incubated at RT for 1 hr. Inhibitor-treated and DMSO samples were functionalized with Azo-H tag or Azo-L tag (100 μM, 50× stock in DMSO), respectively (68). All samples were then treated with TCEP (1 mM, 50× fresh stock in water), TBTA ligand (100 μM, 17× stock in DMSO:t-butanol = 1:4), and copper(II) sulfate (1 mM, 50× stock in water) followed by incubation at RT for 1 h. Samples were combined pairwise to combine inhibitor and DMSO-treated samples and centrifuged (6,500 × g, 4 min, 4 °C) to pellet the precipitated proteins. The pellets were resuspended in cold methanol by sonication, and the two samples were combined. Centrifugation was followed by a second methanol wash, after which the pellet was solubilized in Dulbecco's PBS (DPBS) containing 1.2% SDS via sonication and heating (90 °C, 5 min). The sample resulting from inhibitor pretreatment, IA-alkyne incubation, and Azo-H or Azo-L labeling are as follows: Heavy (DMSO, 100 μM IA-Alkyne)/light (10 μM MIND4-17, 100 μM IA-Alkyne).

The SDS-solubilized proteome samples were diluted by 5 mL of DPBS for a final SDS concentration of 0.2%. The solution was incubated with 100 μL of streptavidin-agarose beads (Thermo Scientific; washed three times with DPBS to remove storage buffer) overnight at 4 °C. Samples were rotated at RT for 2 hr before washed by 5 mL of 0.2% SDS/DPBS, 3 × 5 mL of DPBS, and 3 × 5 mL of water. The beads were pelleted by centrifugation (1,400 × g, 3 min) between washes.

The washed beads were suspended in 500 μL of 6 M urea/ DPBS and 10 mM DTT (from 20× stock in water) and placed in a 65 °C heat block for 15 min. Iodoacetamide (20 mM from 50× stock in water) was then added, and the samples were allowed to react at 37

°C for 30 min. Following reduction and alkylation, the beads were pelleted by centrifugation and resuspended in 200 μL of 2 M urea/DPBS, 1 mM $CaCl_2$ (100× stock in water), and trypsin (2 μg). The digestion was allowed to proceed overnight at 37 °C. The beads were pelleted by centrifugation and washed with 3 × 500 μL of DPBS and 3 × 500 μL of water. The azo-benzene cleavage was carried out by incubating the beads with 50 μL of 25 mM sodium dithionite at RT for 1 hr. The cleavage process was then repeated twice with 75 μL of 25 and 50 mM sodium dithionite, and all of the supernatants were combined. The beads were then washed twice with 75 μL of water, and the wash was combined with the supernatant from the cleavage step to reach 350 μL final. Formic acid (17.5 μL) was added to the sample, which was stored at ¯20 °C until MS analysis.

LC/LC-MS/MS analysis was performed on an LTQ-Orbitrap Discovery mass spectrometer (Thermo Fisher) coupled to an Agilent 1200 series HPLC. Peptide digests were pressure loaded onto a 250-μm fused silica desalting column packed with 4 cm of Aqua C18 reverse phase resin (Phenom- enex). The peptides were eluted onto a biphasic column [100-μm fused silica with a 5-μm tip, packed with 10-cm C 18 and 4-cm Partisphere strong cation exchange resin (SCX, Whatman)] us- ing a gradient 5–100% Buffer B in Buffer A (Buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; Buffer B: 20% water, 80% acetonitrile, 0.1% formic acid). The peptides were then eluted from the SCX onto the C18 resin and into the mass spectrometer using four salt steps previously described. The flow rate through the column was set to ~0.25 μL/min, and the spray voltage was set to 2.75 kV. One full MS scan (FTMS) (400–1,800 M_r) was followed by eight data-dependent scans (ITMS) of the nth most intense ions.

The tandem MS data were searched using the SEQUEST algorithm using a concatenated target/decoy variant of the human IPI databases. A static modification of +57.02146 on cysteine was specified to account for alkylation by iodoacetamide, and differential modifications of +462.2987 (IA-alkyne and cleaved Azo-H) and +456.2849 (IA-alkyne and cleaved Azo-L) were specified on cysteine to account for probe modifications. SEQUEST output files were filtered using DTASelect. Quantification of heavy/light ratios ($R_{H/L}$) was performed using the CIMAGE quantification package as previously described.

Cysteine reactivity in proteomes pretreated with MIND4-17 (heavy-labeled) were compared with DMSO-treated samples (light-labeled). For every cysteine-containing peptide, a heavy/light ratio was calculated that reflects the extent of cysteine reactivity in the inhibitor-treated vs. DMSO samples. A ratio of 1 indicates that there was no change in cysteine reactivity, whereas ratio values >>1 show a loss in cysteine reactivity upon inhibitor treatment.

The following experimental procedures are from Italia, J. S., Addy, P. S., Wrobel, C. J. J., Crawford, L. A., Lajoie, M. J., Zheng, Y., Chatterjee, A. 2017. An Orthogonalized Platform for Genetic Code Expansion in Both Bacteria and Eukaryotes. *Nat. Chem. Biol*. 13, 4, 446–450.

General Methods

For cloning and plasmid propagation, the DH10b (Life Technologies) strain of *E. coli* was used. Polymerase chain reaction (PCR) was performed using the Phusion Hot Start II DNA Polymerase (Fisher Scientific) using the manufacturer's protocol. For purification

of DNA (plasmid as well as PCR products, etc.) spin columns from Epoch Life Science were used. Restriction enzymes and T4 DNA ligase were obtained from New England BioLabs (NEB). DNA oligomers for PCR were purchased from Integrated DNA Technologies (IDT). Verification of cloned DNA by Sanger sequencing was performed by Eton Biosciences. Antibiotics, isopropyl β-D-1- thiogalactopyranoside (IPTG), and L-arabinose were purchased from Sigma-Aldrich or Fisher Scientific. Components of media were obtained from Fisher Scientific. Bacteria were grown on LB or M63 agar plates and LB liquid medium with the following antibiotic concentrations unless otherwise mentioned: 95 μg/mL spectinomycin, 20 μg/mL chloramphenicol, and 100 μg/mL ampicillin, 30 μg/mL kanamycin, 15 μg/mL zeocin, 12 μg/mL tetracycline, 10 μg/mL gentamycin.

Statistical Methods

For all expression analyses (sfGFP in bacteria or EGFP in HEK293T cells), the mean of three independent experiments was reported, and error bars represent standard deviation. For the growth rate analysis, each data point represents the mean OD_{600} of three independent cultures of the same strain (error bars represent standard deviation). In our experience, a mean of three experiments provides adequate levels of accuracy for these experiments.

Strains, Cell Lines

The EcNR1 strain was a gift from G.M. Church. The HEK293T cell line was obtained from ATCC, and propagated without further confirmation. Cell lines are regularly

monitored for Mycoplasma contamination. Even though HEK293T is listed under misidentified cell lines in the ICLAC database, we used it for our expression analyses as a representative mammalian cell line. Given the widespread use of this cell line as a model mammalian expression host, and since our conclusions do not rely on its specific identity (beyond a representative mammalian cell-line), we believe that the use of this strain is justified.

Building ATMW1

EcNR1 was transformed with pUltraG_ScW40$_{CCA}$. To remove the *E. coli* tryptophanyl–tRNA synthetase (*trpS*) from this strain, the gene encoding zeocin resistance (ShBle) driven by the EM-7 promoter and the CYC1 transcription terminator was PCR amplified using primers TrpRS.Z.ab-F and TrpRS.Z.ab-R to generate the PCR product *trpS*::Zeo^R. 50 ng of the *trpS*::Zeo^R PCR cassette was transformed in the recombination following the aforementioned protocol, and the resulting strains were plated on LB–Agar plates supplemented with zeocin. The resulting colonies were screened via colony PCR using TrpRS150F + TrpRS150R, TrpRS150F + Zeo-iR, and TrpRS150F + TrpRS_sqiR, as well as sequencing these colony PCR products. This strain was named EZ4.

To replace the *E. coli* tryptophanyl-tRNA$_{CCA}$ (*trpT*), the *trpT*::Gent^R PCR cassette was amplified using trpTKO.Gent-F and trpTKO.Gent-R (965 bp). 50 ng of *trpT*::Gent^R PCR cassette was transformed into EZ4, induced as previously described. Resulting gentamycin-resistant colonies were screened for the desired recombination using colony PCR primers trpT GsqF and trpT GsqR, as well as sequencing of the PCR product. The resulting strain was named EZG4.

2 μM 90 bp oligo, galK.90 del, was used to delete *galK* from the endogenous genomic location. Following transformation, cells were recovered for 6 hrs and washed twice with M9 minimal media at 5,000 × g for 5 min. 100 μl of a 10^4 dilution was plated on M63 minimal media supplemented with glycerol and 2-deoxygalactose to select for successful galK deletion. Colony PCR was used to verify the deletion of *galK* using galK_KO_verf-F/R. The resulting strain was named G4.

galK.PCR cassette containing the endogenous promoter dlambda.GalK-F and dlambda.galK dterm-R (1,348 bp) was amplified and was used to remove the λ-RED genes from the 4 strain. Following transformation of 50 ng of this PCR product into strain 4, successful insertion of the galK.PCR cassette into the λ-RED site was selected by plating the cells on M63 plates containing galactose as the sole carbon source for 3 days. Surviving colonies were screened by colony PCR using dlambda.sqF with dlambda.sqR or dLambda.sqiR for the desired deletion. This final strain was named ATMW1 (EcNR1 *trpS*::Zeo^R *trpT*::$Gent^R$ *ΔgalK* λ-RED::*galK*).

Growth Comparison

5 mL starter cultures of EcNR1G, EcNR1G + pUltraG_ScW40$_{CCA}$, and ATMW1 strains were grown for 16 hrs in LB with all strain-dependent antibiotics. For each strain, the starter culture was diluted to an initial OD_{600} of 0.01 in three identical cultures of 80 mL LB with no antibiotics and allowed to grow in 250 mL sterile Erlenmeyer flasks at 30 °C, with shaking (250 r.p.m.). Growth was monitored every 30 min by measuring OD_{600} in a 10 mm cuvette.

Assessment of aaRS–tRNA Activity Using a Chloramphenicol Reporter

Overnight cultures of ATMW1 harboring pRepAC-EcW-TAG or pRepJI-EcW-TGA, with or without pBK–EcWRSwt, were diluted to an OD_{600} of 0.1, and 3 μL was spot plated on LB agar plates supplemented with kanamycin (+pBK plates), spectinomycin, tetracycline, and varying chloramphenicol concentrations. Growth was analyzed after 48 hrs of incubation at 37 °C.

Assessment of aaRS–tRNA Activity Using a sfGFP151 Reporter

EcNR1 or ATMW1 harboring pEvolT5-EcW sfGFP151 (TAG or TGA) with or without pBK-EcWRSwt was grown overnight in LB. The starter cultures were diluted in LB supplemented with required antibiotics to 0.05 OD_{600}. Cultures were grown at 30 °C or 37 °C (30 °C when comparing to progenitor strain EcNR1) until 0.55 OD_{600}, at which point the sfGFP expression was induced with a final concentration of 1 mM IPTG. Unnatural amino acids (UAA) were added during induction to a final concentration of 1 mM. Cultures were grown for an additional 17–20 hrs at 37 °C with shaking. To evaluate sfGFP expression, cells from 150 μL of the cultures were pelleted at 5,000 × g, resuspended in 150 μL PBS, and transferred to a 96-well clearbottom assay plate. Fluorescence was measured by using a SpectraMAX M5 (Molecular Devices) (Ex. 488 nm; Em. 534 nm). Fluorescence for each sample was normalized using its OD_{600}.

Protein Purification

The tRNAEcTrp$_{UCA}$ was expressed from the *proK* promoter. The sfGFP reporter gene (sfGFP-151-TGA or wild type sfGFP) was expressed from pET22b-T5lac plasmid

driven by the strong *t5.lac* promoter. Overnight expression cultures were centrifuged and resuspended in lysis buffer: B-PER Bacterial Protein Extraction Reagent (Thermo Scientific) + 1× Halt Protease Inhibitor Cocktail (Thermo Scientific) + 0.01% Pierce Universal Nuclease (Thermo Scientific). After 30 min incubation at room temperature, the lysate was clarified by centrifuging at 22,000 × g for 5 min. The C-terminally polyhistidine tagged soluble sfGFP in the supernatant was purified using a HisPur Ni-NTA resin (Thermo Scientific) following manufacturer's protocol. Protein purity was confirmed by SDS–PAGE and purified protein molecular weight was confirmed by ESI–MS (Agilent Technologies, 1260 Inifinity ESI–TOF).

Assessing Activity of tRNAEcTrp$_{CUA}$ and Synthetase Hits in HEK293T

Dulbecco's modified Eagle's medium (high glucose DMEM) supplemented with 10% FBS (FBS) and penicillin–streptomycin (0.5×) was used to culture HEK293T cells. Cells were incubated in a humidified incubator at 37 °C in the presence of 5% CO_2. HEK293T were seeded at a density of 600,000 cells per well for a 12-well plate 1 day before transfection and transfected at ~70% confluence. Polyethylenimine (PEI, Sigma) and DNA were mixed at a ratio of 4 μL PEI (1 mg/mL) to 1 μg DNA in DMEM. After 20 min incubation, 100 μL of this mixture was used to transfect one single well in a 12-well plate. For these transfections, 500 ng of pAcBac EGFP39* U6-EcWtR TAG was transfected in the presence or absence of pAcBac-TrpRS (wt, h14, or h9) U6-EcWtR TAG. UAAs were added to the culture medium to a final concentration of 1 mM at the time of transfection. Fluorescence images were taken at 48 h post-transfection using a Zeiss Axio Observer fluorescence microscope.

To obtain EGFP39* expression data, cells were harvested, washed once with PBS buffer (5,000 × g), and lysed with CelLytic M lysis buffer (Sigma) supplemented with 1× Halt protease inhibitor and 0.01% Pierce universal nuclease. 50 μL lysis buffer was used for each well of a 12-well plate, and was allowed to incubate for 20 min. After incubation, the lysate was clarified by centrifuging at 22,000 × g for 5 min and was transferred to a clear bottom 96-well assay plate. Fluorescence was measured using a SpectraMAX M5 (Molecular Devices) (Ex. 488 nm; Em. 534 nm).

Tryptic Digestion and LC–MS/MS Analysis of Reporter Proteins

12 μg purified reporter protein was treated with DMSO and precipitated with 5 μl 100% solution of trichloracetic acid. Sample was frozen at ¯80 °C overnight. Thawed sample was centrifuged at 15,000 r.p.m. for 10 min at room temperature. Supernatant was removed and pellets were vortexed to resuspend in 500 μL cold acetone. Samples were then centrifuged at 5,000 r.p.m. for 10 min. Acetone was then removed and pellet was allowed to air dry. Pellet was resuspended in 30 μL 8 M urea in PBS followed by 70 μL 100 mM ammonium bicarbonate, and then 1.5 μL 1 M DTT was added. Samples were incubated at 65 °C for 15 min. After incubation, 2.5 μL of 500 mM iodoacetamide in PBS was added and the sample was left at room temperature for 30 min. Following incubation, 120 μL PBS was added to each sample and vortexed rapidly. 2 μg of trypsin was added to the samples followed by 2.5 μL 100 mM $CaCl_2$. Samples were then agitated for 37 °C overnight. Trypsin was then quenched with 10 μL formic acid and pelleted at 15,000 r.p.m for 20 min. Supernatants were stored at ¯20 °C. Samples were subsequently analyzed by LC–MS/MS using a LTQ Orbitrap XL mass spectrometer (Thermo Fisher) coupled to an

EASY-nLC 1000 nanoLC (Thermo Fisher). 10 μL samples were loaded onto 100 μm fused silica column with a 5 μm tip packed with 10 cm of Aqua C18 reverse-phase resin (Phenomenex) using the EASY-nLC 1000 autosampler. The digests were eluted using a gradient 0–100% buffer B in buffer A (buffer A: 95% water, 5% acetonitrile, 0.1% formic acid; buffer B; 20% water, 80% acetonitrile, 0.1% formic acid). The flow rate through the column was set to 400 nL/min and the spray voltage was set to 3.5 kV. One full MS scan (FTMS) (400–1800 MW) was followed by seven data-dependent scans (ITMS) of the nth most intense ion from the imported mass list with dynamic exclusion. The tandem MS data were searched using the SEQUEST algorithm using a concatenated target/decoy variant of the human IPI databases with addition of sfGFP. A static modification of $^{+}$57.02146 on cysteine was specified to account for iodoacetamide alkylation. SEQUEST output files were filtered using DTASelect 2.0.

References

(1) Downard, K. *Mass Spectrometry: A Foundation Course*, 1st ed.; The Royal Society of Chemistry, 2004; p 226.

(2) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science* 1989, *246* (4926), 64–71.

(3) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T. Protein and Polymer Analyses up to M/z 100 000 by Laser Ionization Time of Flight Mass Spectrometry. *Rapid communications in mass spectrometry* 1988, *2* (8), 151–153.

(4) Perry, R. H.; Cooks, G. R.; Noll, R. J. Orbitrap Mass Spectrometry: Instrumentation, Ion Motion and Applications. *Mass Spectrometry Reviews* 2008, *27* (6), 661–699.

(5) Page, J. S.; Marginean, I.; Baker, E. S.; Kelly, R. T.; Tang, K.; Smith, R. D. Biases in Ion Transmission through an Electrospray Ionization-Mass Spectrometry Capillary Inlet. *Journal of the American Society for Mass Spectrometry* 2009, *20* (12), 2265–2272.

(6) Roepstorff, P.; Fohlman, J. Letter to the Editors. *Biological Mass Spectrometry* 1984, *11* (11), 601–601.

(7) Itoh, K.; Wakabayashi, N.; Katoh, Y.; Ishii, T.; Igarashi, K.; Engel, J. D.; Yamamoto, M. Keap1 Represses Nuclear Activation of Antioxidant Responsive Elements by Nrf2 through Binding to the Amino-Terminal Neh2 Domain. *Genes & Development* 1999, *13*, 76–86.

(8) Cullinan, S. B.; Gordan, J. D.; Jin, J.; Harper, J. W.; Diehl, J. A. The Keap1-BTB Protein Is an Adaptor That Bridges Nrf2 to a Cul3-Based E3 Ligase: Oxidative Stress Sensing by a Cul3-Keap1 Ligase. *Molecular and Cellular Biology* 2004, *24* (19), 8477–8486.

(9) Tong, K. I.; Kobayashi, A.; Katsuoka, F.; Yamamoto, M. Two-Site Substrate Recognition Model for the Keap1-Nrf2 System: A Hinge and Latch Mechanism. *Biological chemistry* 2006, *387* (10-11), 1311–1320.

(10) Zhang, D. D.; Hannink, M. Distinct Cysteine Residues in Keap1 Are Required for Keap1-Dependent Ubiquitination of Nrf2 and for Stabilization of Nrf2 by Chemopreventive Agents and Oxidative Stress. *Molecular and Cellular Biology* 2003, *23* (22), 8137–8151.

(11) McMahon, M.; Itoh, K.; Yamamoto, M.; Hayes, J. D. Keap1-Dependent Proteasomal Degradation of Transcription Factor Nrf2 Contributes to the Negative Regulation of Antioxidant Response Element-Driven Gene Expression. *Journal of Biological Chemistry* 2003, *278* (24), 21592–21600.

(12) Zhang, D. D.; Lo, S.-C.; Cross, J. V.; Templeton, D. J.; Hannink, M. Keap1 Is a Redox-Regulated Substrate Adaptor Protein for a Cul3-Dependent Ubiquitin Ligase Complex. *Molecular and Cellular Biology* 2004, *24* (24), 10941–10953.

(13) Kobayashi, A.; Kang, M.-I.; Okawa, H.; Ohtsuhji, M.; Zenke, Y.; Chiba, T.; Igarashi, K.; Yamamoto, M. Oxidative Stress Sensor Keap1 Functions as an Adaptor for Cul3-Based E3 Ligase to Regulate Proteasomal Degradation of Nrf2. *Molecular and Cellular Biology* 2004, *24* (16), 7130–7139.

(14) Kamps, M. P.; Taylor, S. S.; Sefton, B. M. Direct Evidence That Oncogenic Tyrosine Kinases and Cyclic AMP-Dependent Protein Kinase Have Homologous ATP-Binding Sites. *Nature* 1984, *310*.

(15) McMahon, M.; Lamont, D. J.; Beattie, K. A.; Hayes, J. D. Keap1 Perceives Stress via Three Sensors for the Endogenous Signaling Molecules Nitric Oxide, Zinc, and Alkenals. *Proceedings of the National Academy of Science* 2010, *107* (44), 18838–18843.

(16) Baird, L.; Llères, D.; Swift, S.; Dinkova-Kostova, A. T. Regulatory Flexibility in the Nrf2-Mediated Stress Response Is Conferred by Conformational Cycling of the Keap1-Nrf2 Protein Complex. *Proceedings of the National Academy of Sciences* 2013, *110* (38), 15259–15264.

(17) Baird, L.; Dinkova-Kostova, A. T. Diffusion Dynamics of the Keap1–Cullin3 Interaction in Single Live Cells. *Biochemical and Biophysical Research Communications*

2013, *433*, 58–65.

(18) Kansanen, E.; Kuosmanen, S. M.; Leinonen, H.; Levonen, A.-L. The Keap1-Nrf2 Pathway: Mechanisms of Activation and Dysregulation in Cancer. *Redox Biology* 2013, *1*, 45–49.

(19) Kensler, T. W.; Wakabayashi, N.; Biswal, S. Cell Survival Responses to Environmental Stresses via the Keap1-Nrf2-ARE Pathway. *Annual Review of Pharmacology and Toxicology* 2007, *47*, 89–116.

(20) Hayes, J. D.; Dinkova-Kostova, A. T. The Nrf2 Regulatory Network Provides an Interface between Redox and Intermediary Metabolism. *Trends in biochemical sciences* 2014, *39* (4), 199–218.

(21) Flier, J. S.; Underhill, L. H.; Martin, J. B.; Gusella, J. F. Huntingtons Disease. *New England Journal of Medicine* 1986, *315* (20), 1267–1276.

(22) Duyao, M.; Ambrose, C.; Myers, R.; Novelletto, A.; Persichetti, F.; Frontali, M.; Folstein, S.; Ross, C.; Franz, M.; Abbott, M.; Gray, J.; Conneally, P.; Young, A.; Penney, J.; Hollingsworth, Z.; Shoulson, I.; Lazzarini, A.; Falek, A.; Koroshetz, W.; Sax, D.; Bird, E.; Vonsattel, J.; Bonilla, E.; Alvir, J.; Bickham Conde, B.; Cha, J.; Dure, L.; Gomez, F.; Ramos, M.; Sanchez-Ramos, J.; Snodgrass, S.; Young, M. de; Wexler, N.; Moscowitz, C.; Penchaszadeh, G.; MacFarlane, H.; Anderson, M.; Jenkins, B.; Srinidhi, J.; Barnes, G.; Gusella, J.; MacDonald, M. Trinucleotide Repeat Length Instability and Age of Onset in Huntington's Disease. *Nature Genetics* 1993, *4* (4), 387–392.

(23) Sorolla, M. A.; Reverter-Branchat, G.; Tamarit, J.; Ferrer, I.; Ros, J.; Cabiscol, E. Proteomic and Oxidative Stress Analysis in Human Brain Samples of Huntington Disease. *Free Radical Biology and Medicine* 2008, *45*, 667–678.

(24) Ellrichmann, G.; Petrasch-Parwez, E.; Lee, D.-H.; Reick, C.; Arning, L.; Saft, C.; Gold, R.; Linker, R. A. Efficacy of Fumaric Acid Esters in the R6/2 and YAC128 Models of Huntington's Disease. *PLoS ONE* 2011, *6* (1), e16172.

(25) Stack, C.; Ho, D.; Wille, E.; Calingasan, N. Y.; Williams, C.; Liby, K.; Sporn, M.; Dumont, M.; Beal, M. F. Triterpenoids CDDO-Ethyl Amide and CDDO-Trifluoroethyl Amide Improve the Behavioral Phenotype and Brain Pathology in a Transgenic Mouse Model of Huntington's …. *Free Radical Biology and Medicine* 2010, *49*, 147–158.

(26) Quinti, L.; Casale, M.; Moniot, S.; Pais, T. F.; Kanegan, M. J. Van; Kaltenbach, L. S.; Pallos, J.; Lim, R. G.; Dayalan Naidu, S.; Runne, H.; Meisel, L.; Abdul Rauf, N.; Leyfer, D.; Maxwell, M. M.; Saiah, E.; Landers, J. E.; Luthi-Carter, R.; Abagyan, R.; Dinkova-Kostova, A. T.; Steegborn, C.; March, J. L.; Lo, D. C.; Thompson, L. M.; Kazantsev, A. G. SIRT2-and NRF2-Targeting Thiazole-Containing Compound with Therapeutic Activity in Huntington's Disease Models. *Cell Chemical Biology* 2016, *23*, 849–861.

(27) Raibaut, L.; Mahdi, O.; Melnyk, O. *Protein Ligation and Total Synthesis II*; 2015; Vol. 363.

(28) Graaf, A. de; Kooijman, M.; Hennink, W.; Mastrobattista, E. Nonnatural Amino Acids for Site-Specific Protein Conjugation. *Bioconjugate Chemistry* 2009, *20* (7), 1281–1295.

(29) Dumas, A.; Lercher, L.; Spicer, C. D.; Davis, B. G. Designing Logical Codon Reassignment–expanding the Chemistry in Biology. *Chemical Science* 2015, *6* (1), 50–69.

(30) Fahey, J. W.; Dinkova-Kostova, A. T.; Stephenson, K. K.; Talalay, P. The "Prochaska" Microtiter Plate Bioassay for Inducers of NQO1. *Methods in Enzymology* 2004, *384*, 243–258.

(31) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. Quantitative Reactivity Profiling Predicts Functional Cysteines in Proteomes. *Nature* 2010, *468*, 790–795.

(32) Qian, Y.; Martell, J.; Pace, N. J.; Ballard, T. E.; Johnson, D. S.; Weerapana, E. An Isotopically Tagged Azobenzene-Based Cleavable Linker for Quantitative Proteomics. *ChemBioChem* 2013, *14*, 1410–1414.

(33) Probst, B. L.; McCauley, L.; Trevino, I.; Wigley, W. W.; Ferguson, D. A. Cancer Cell Growth Is Differentially Affected by Constitutive Activation of NRF2 by KEAP1 Deletion and Pharmacological Activation of NRF2 by the Synthetic Triterpenoid, RTA 405. *PLOS ONE* 2015, *10* (8), e0135257.

(34) Jahn, M.; Rogers, M. J.; Söll, D. Anticodon and Acceptor Stem Nucleotides in tRNA(Gln) Are Major Recognition Elements for E. Coli Glutaminyl-tRNA Synthetase. *Nature* 1991, *352* (6332), 258–60.

(35) Soll, L.; Berg, P. Recessive Lethal Nonsense Suppressor in Escherichia Coli Which Inserts Glutamine. *Nature* 1969, *223* (5213), 1340–1342.

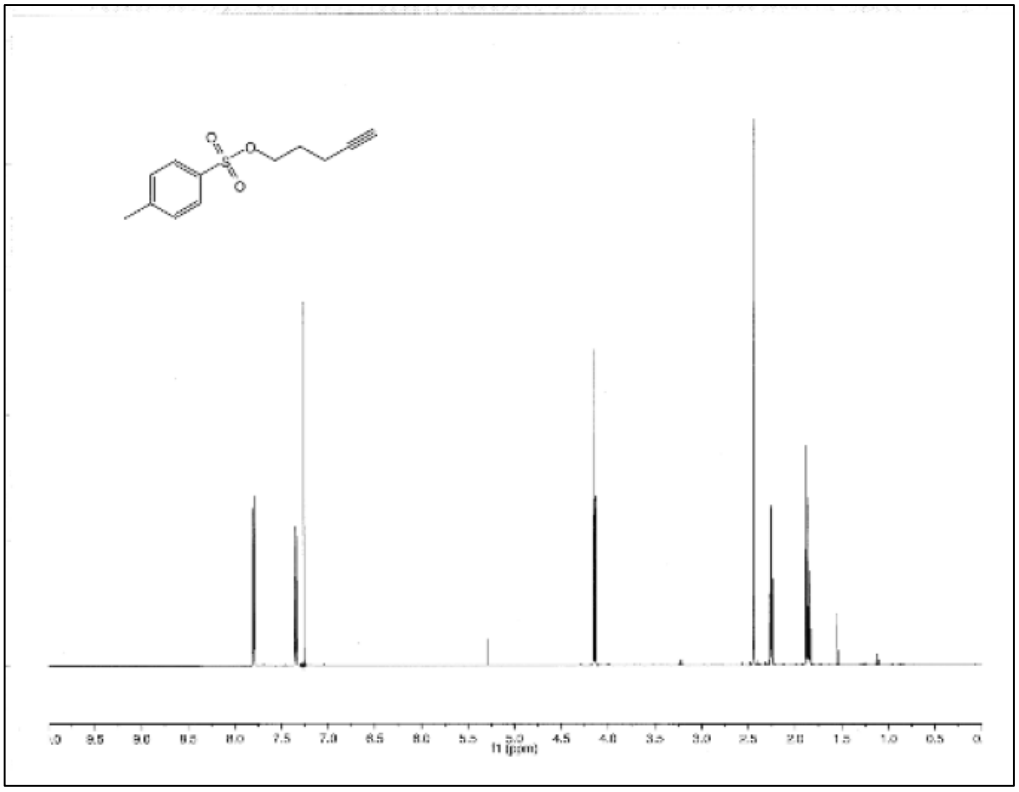
(36) Rogers, M. J.; Adachi, T.; Inokuchi, H.; Söll, D. Switching tRNA(Gln) Identity from Glutamine to Tryptophan. *Proceedings of the National Academy of Sciences of the United States of America* 1992, *89* (8), 3463–7.

Appendix I

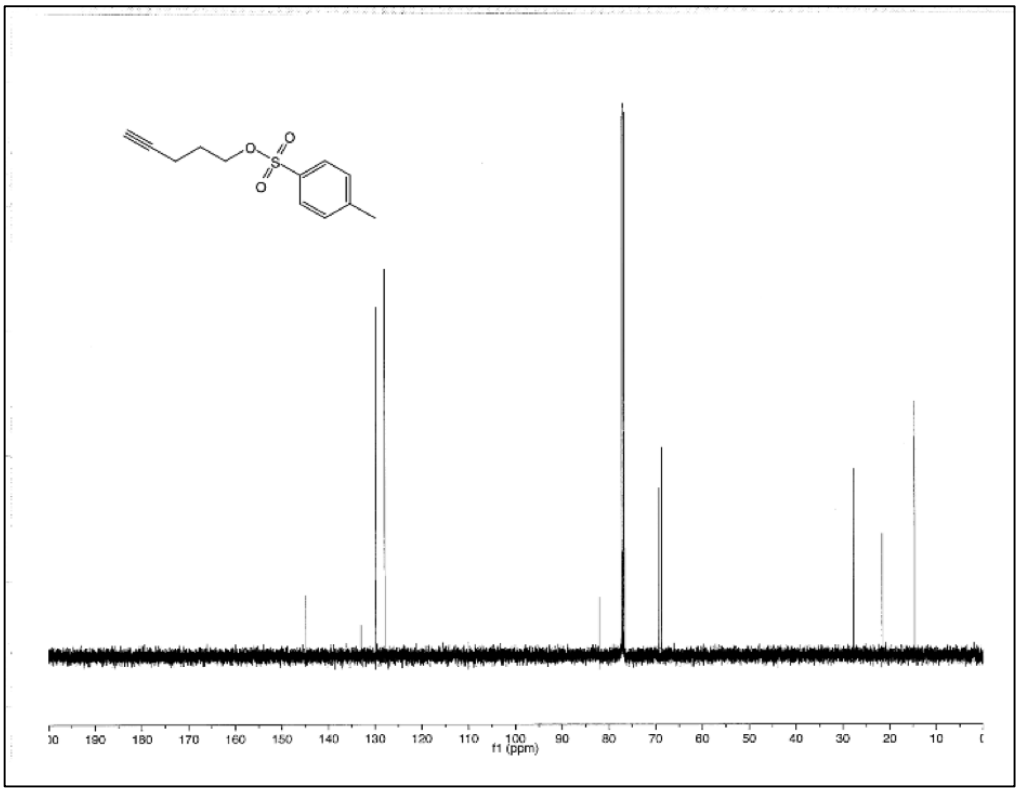
NMR Data

pent-4-yn-1-yl 4-methylbenzenesulfonate

^{1}H-NMR

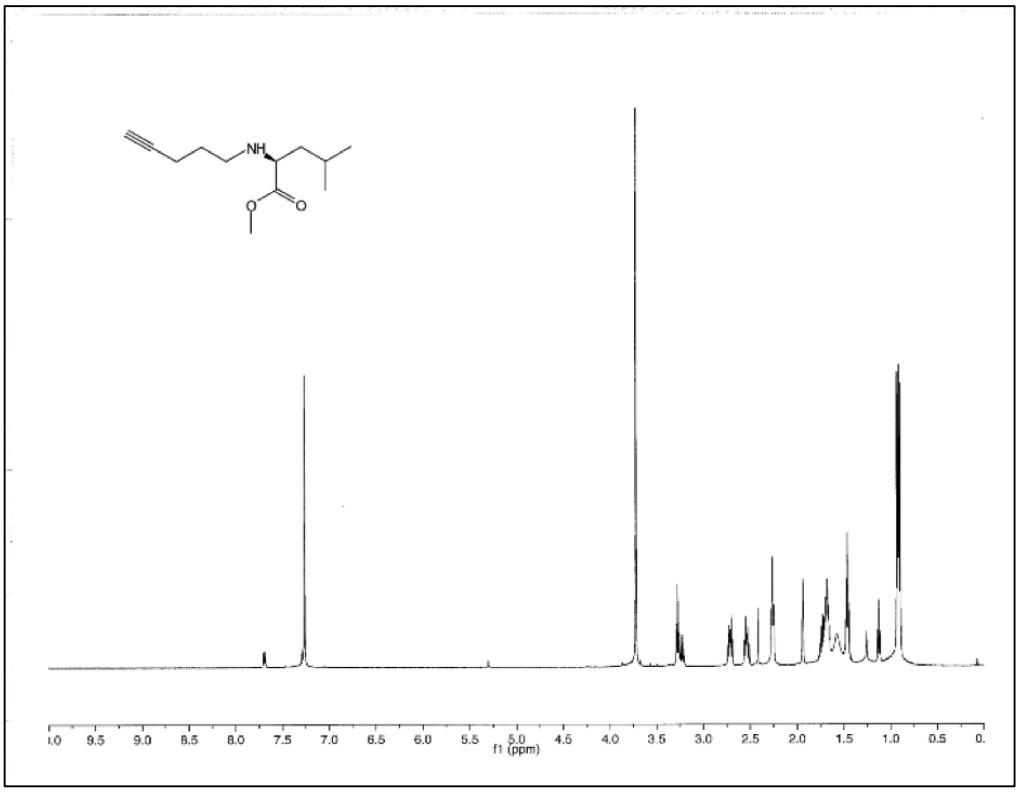


^{13}C-NMR

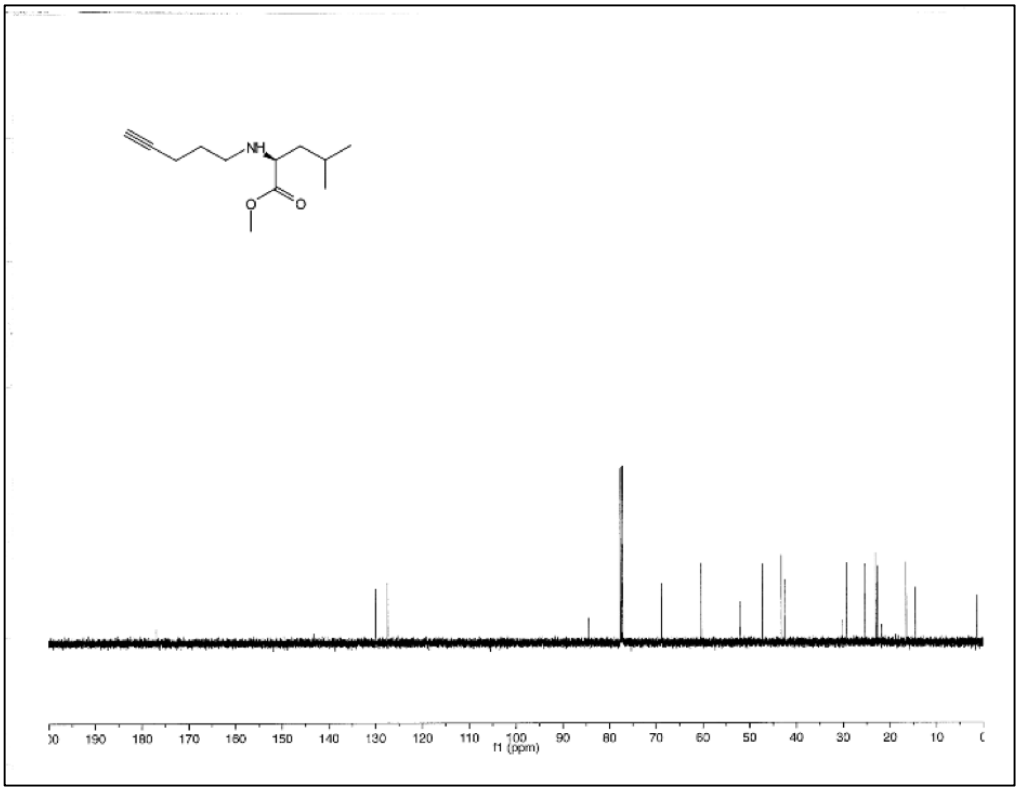


methyl pent-4-yn-1-yl-L-leucinate

^{1}H-NMR

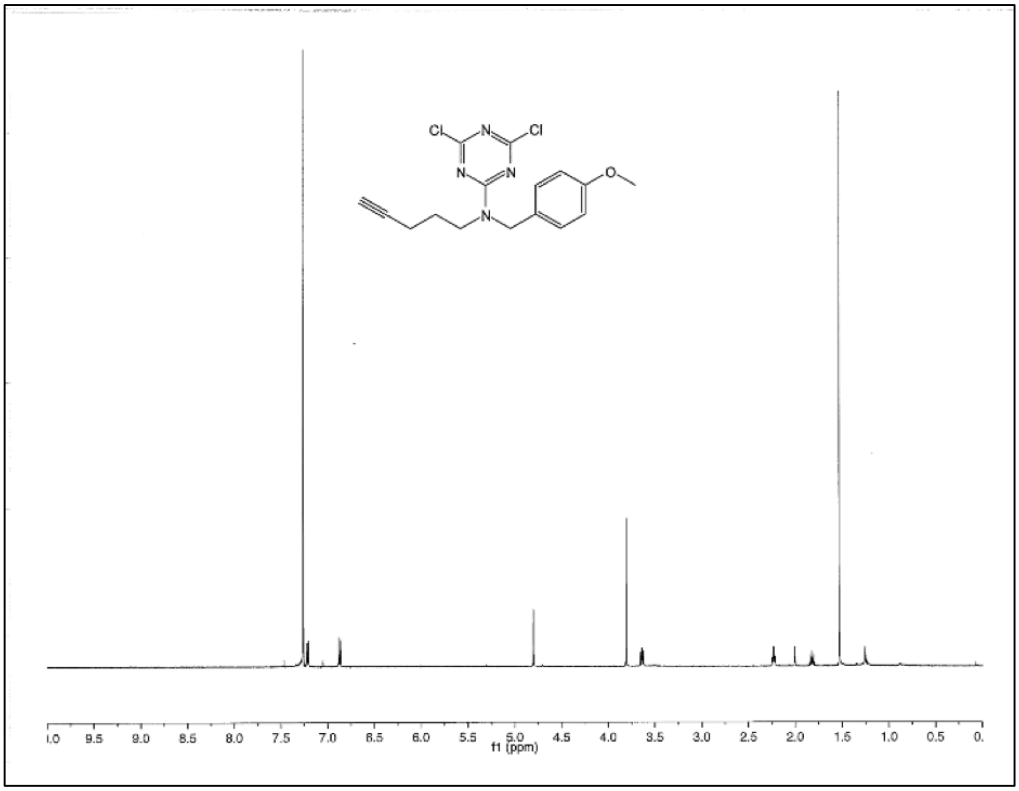


^{13}C-NMR

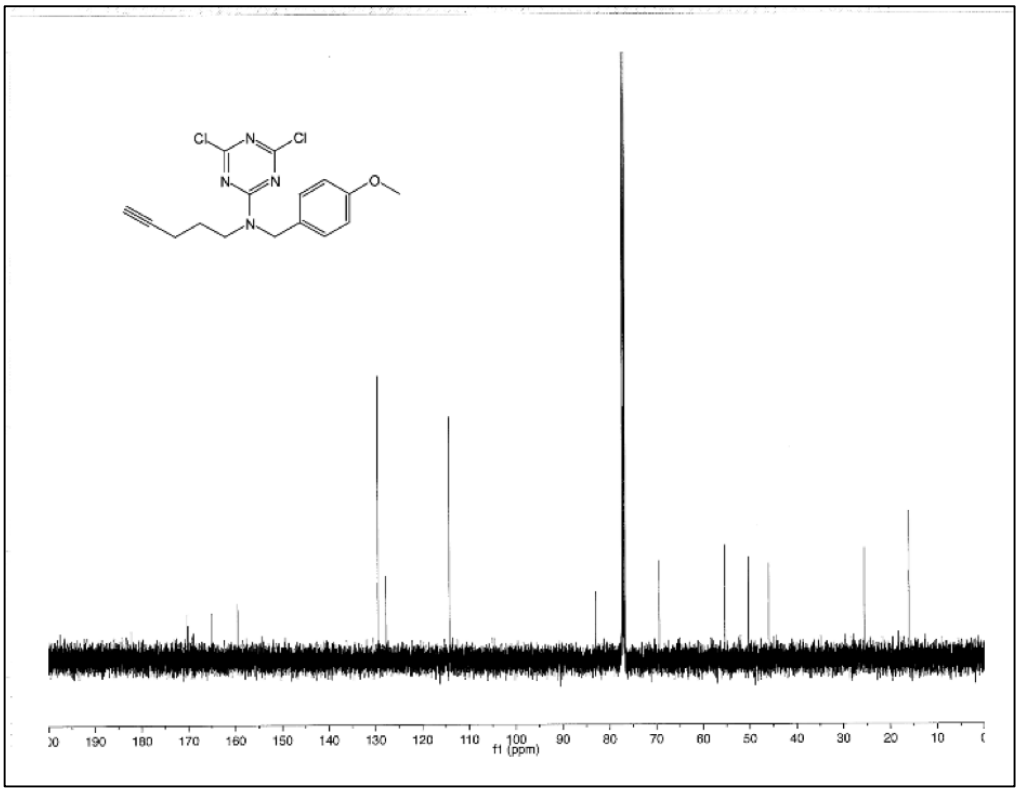


4,6-dichloro-N-(4-methoxybenzyl)-N-(pent-4-yn-1-yl)-1,3,5-triazin-2-amine (LAS1)

^{1}H-NMR

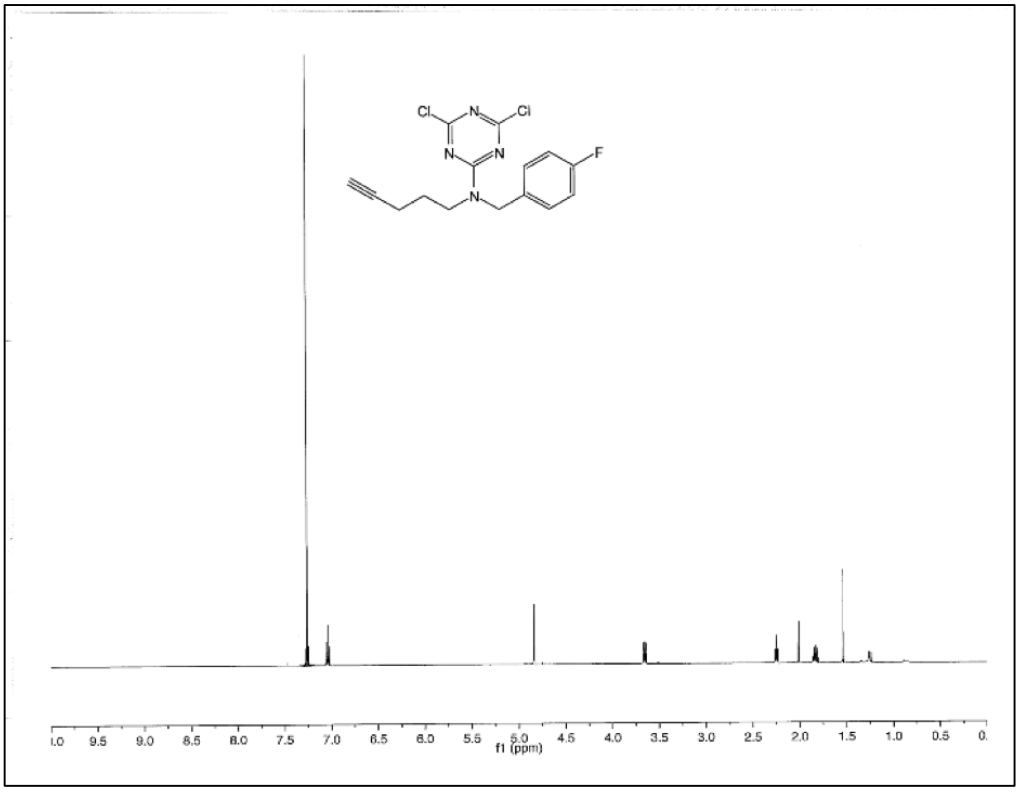


^{13}C-NMR

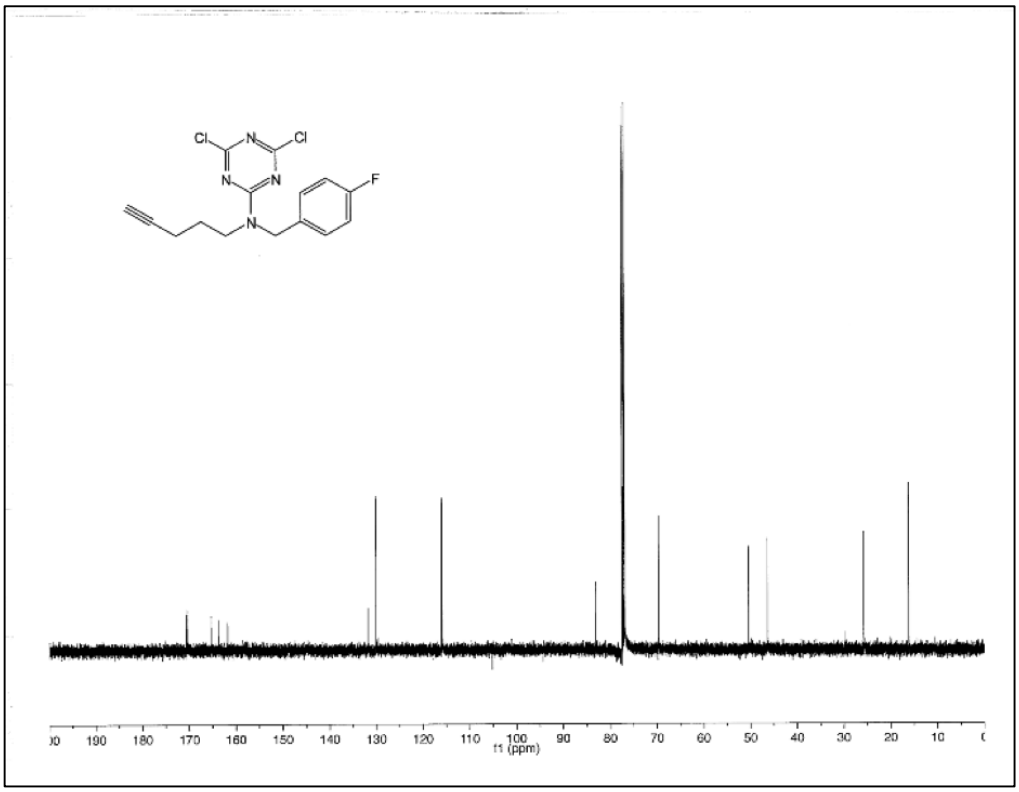


4,6-dichloro-N-(4-fluorobenzyl)-N-(pent-4-yn-1-yl)-1,3,5-triazin-2-amine (LAS2)

^{1}H-NMR

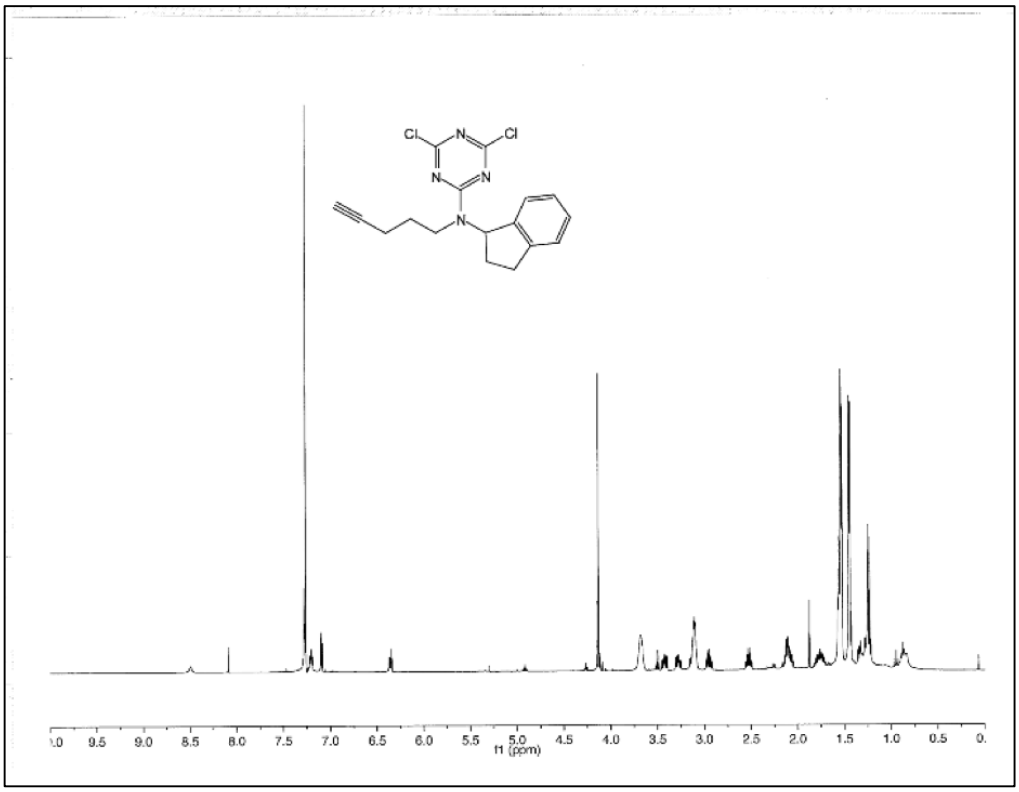


^{13}C-NMR

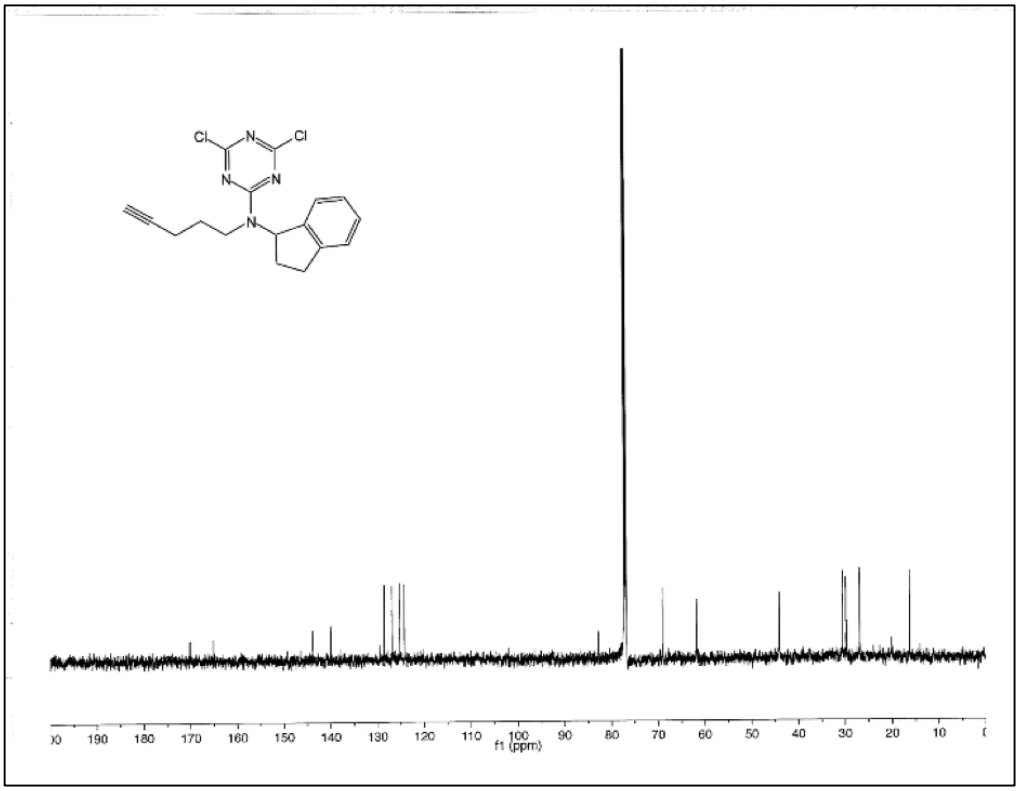


4,6-dichloro-N-(2,3-dihydro-1H-inden-1-yl)-N-(pent-4-yn-1-yl)-1,3,5-triazin-2-amine (LAS3)

^{1}H-NMR

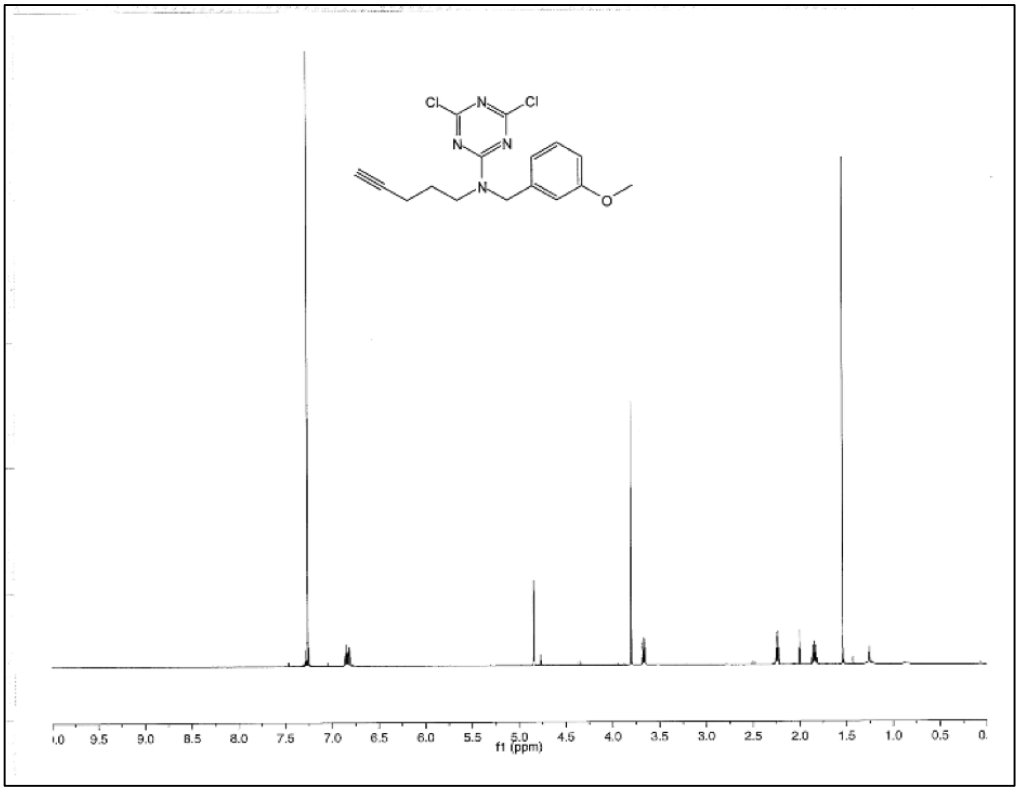


^{13}C-NMR

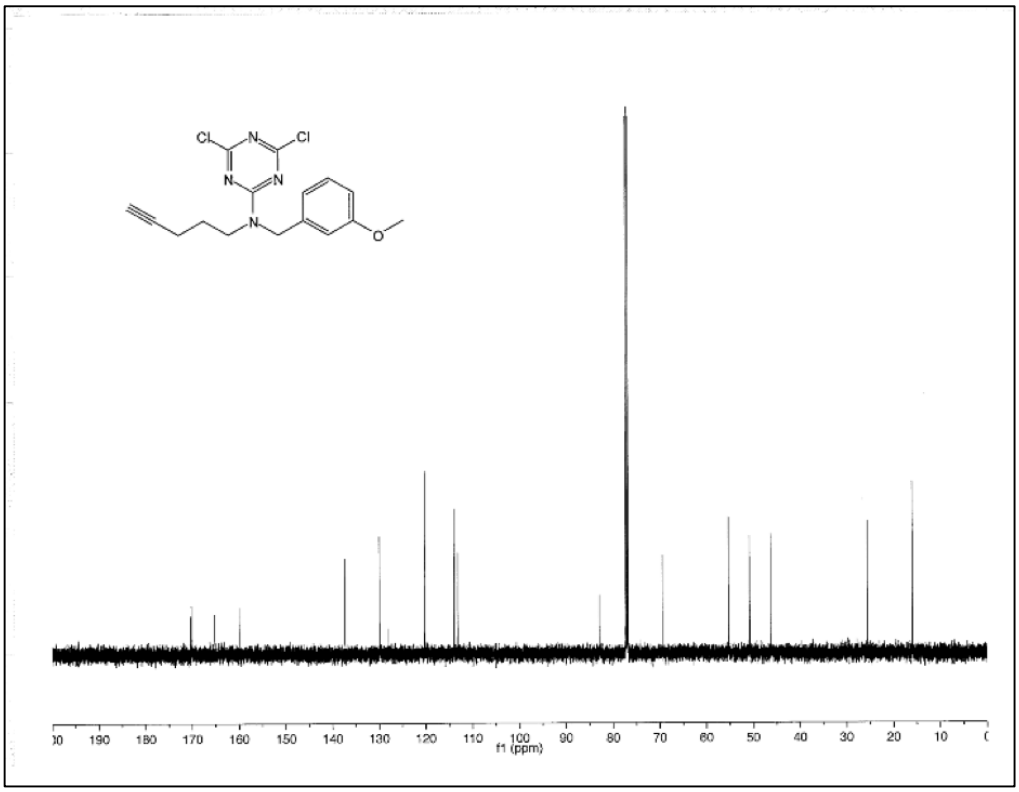


4,6-dichloro-N-(3-methoxybenzyl)-N-(pent-4-yn-1-yl)-1,3,5-triazin-2-amine (LAS4)

^{1}H-NMR

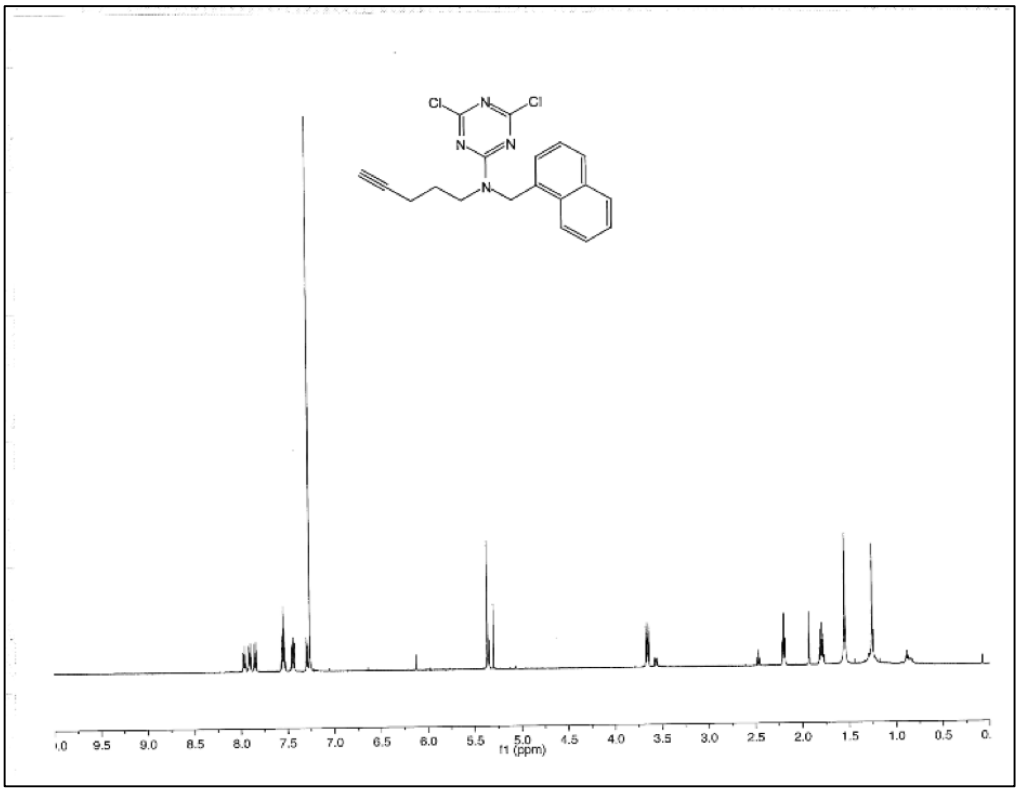


^{13}C-NMR

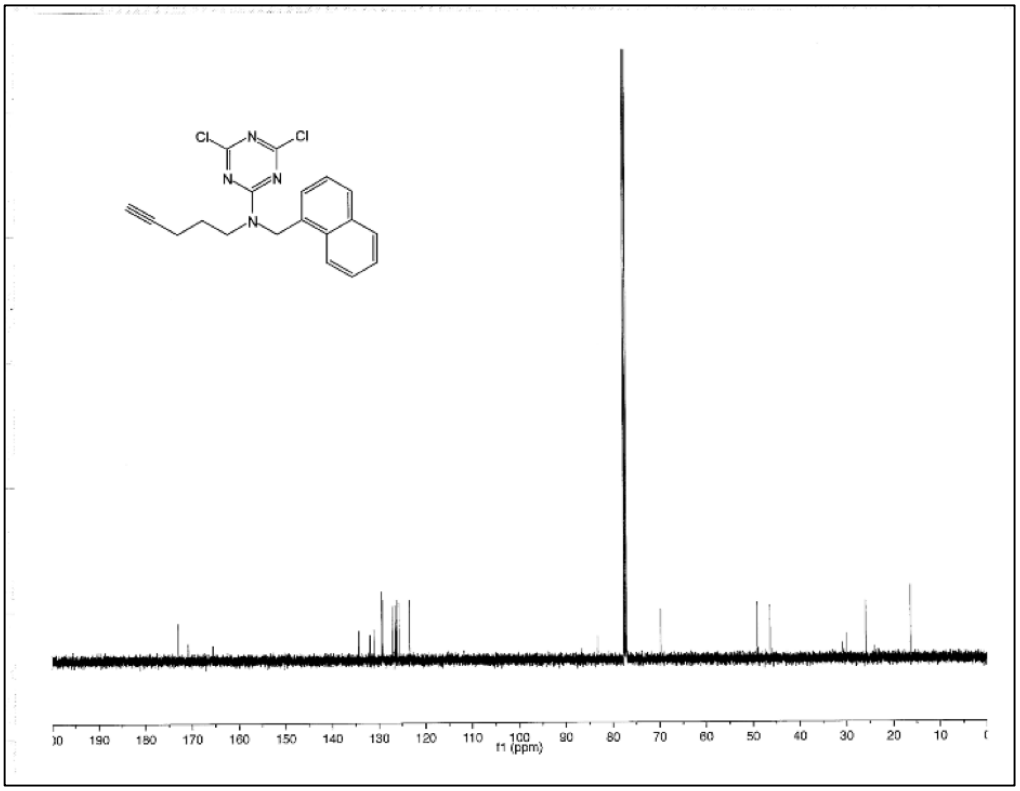


4,6-dichloro-N-(naphthalen-2-ylmethyl)-N-(pent-4-yn-1-yl)-1,3,5-triazin-2-amine (LAS5)

^{1}H-NMR

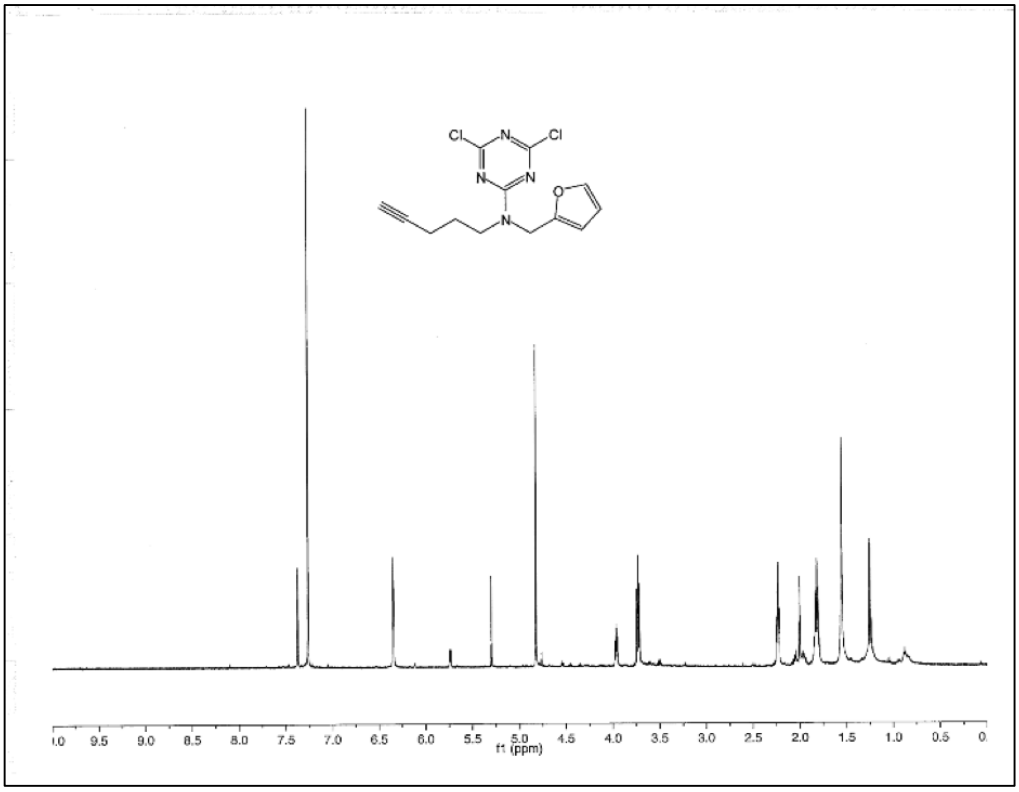


^{13}C-NMR

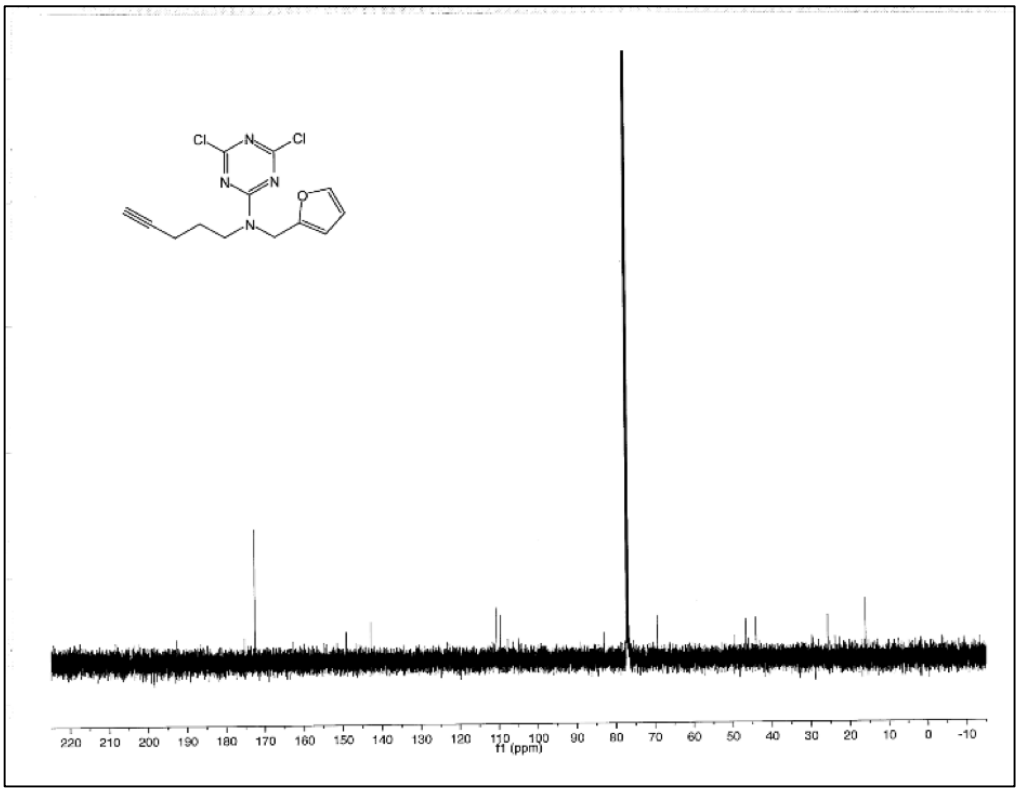


4,6-dichloro-N-(furan-2-ylmethyl)-N-(pent-4-yn-1-yl)-1,3,5-triazin-2-amine (LAS6)

^{1}H-NMR

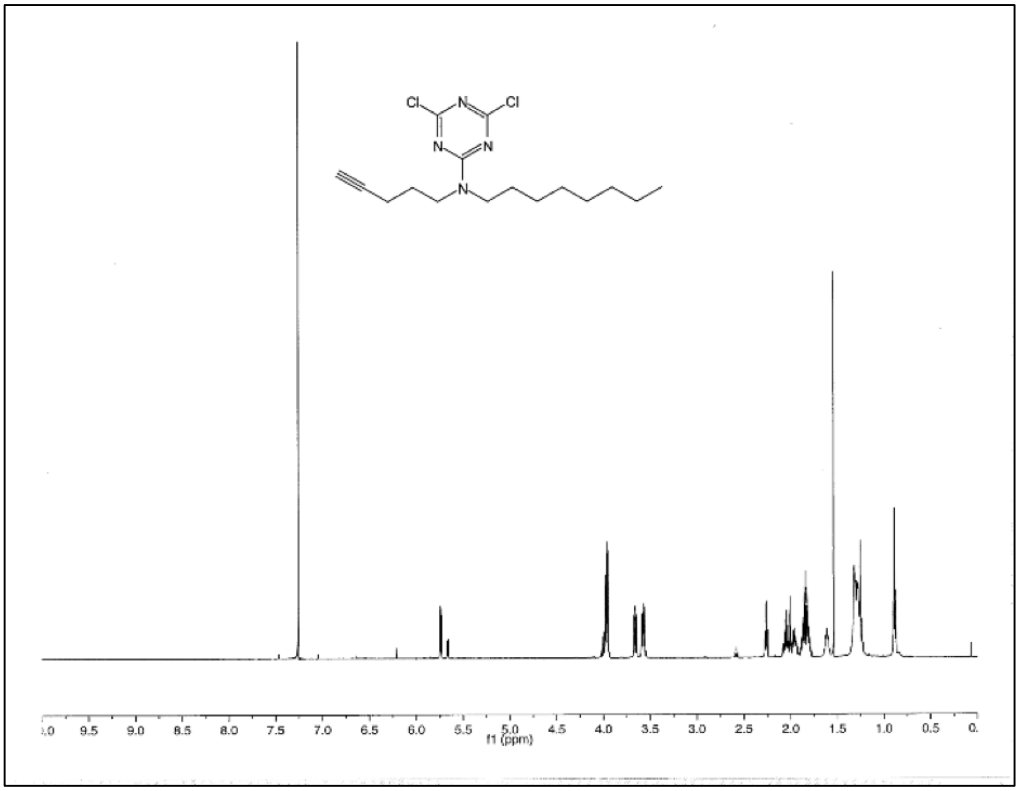


^{13}C-NMR

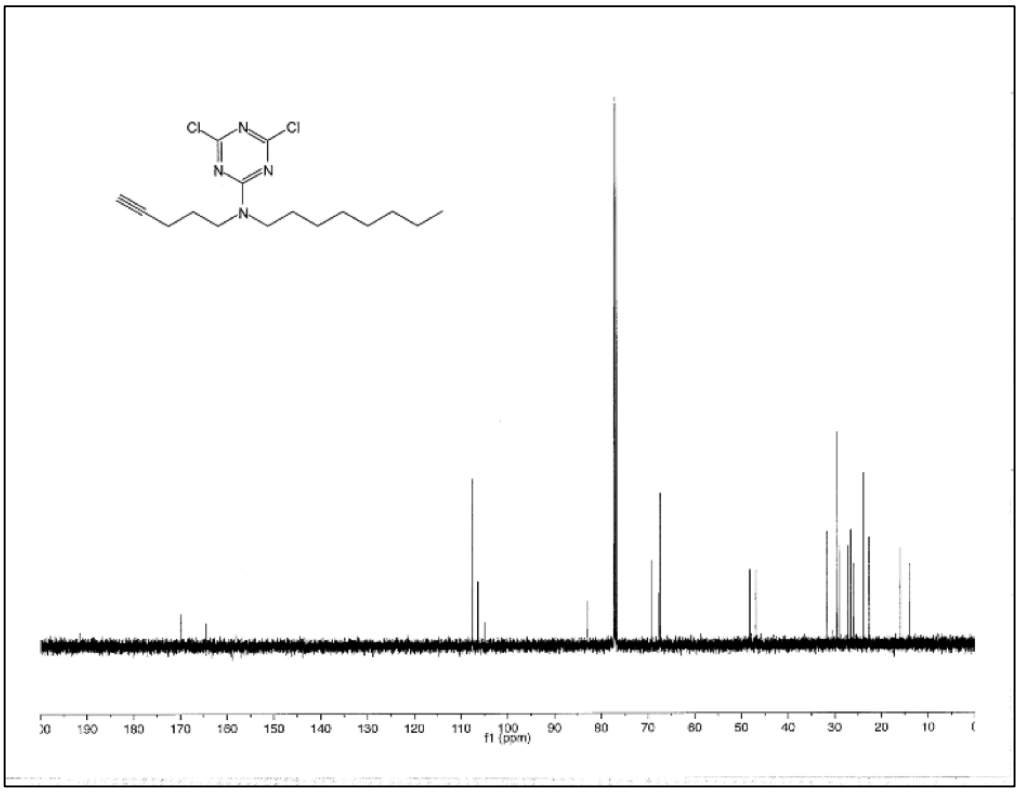


4,6-dichloro-N-octyl-N-(pent-4-yn-1-yl)-1,3,5-triazin-2-amine (LAS7)

^{1}H-NMR

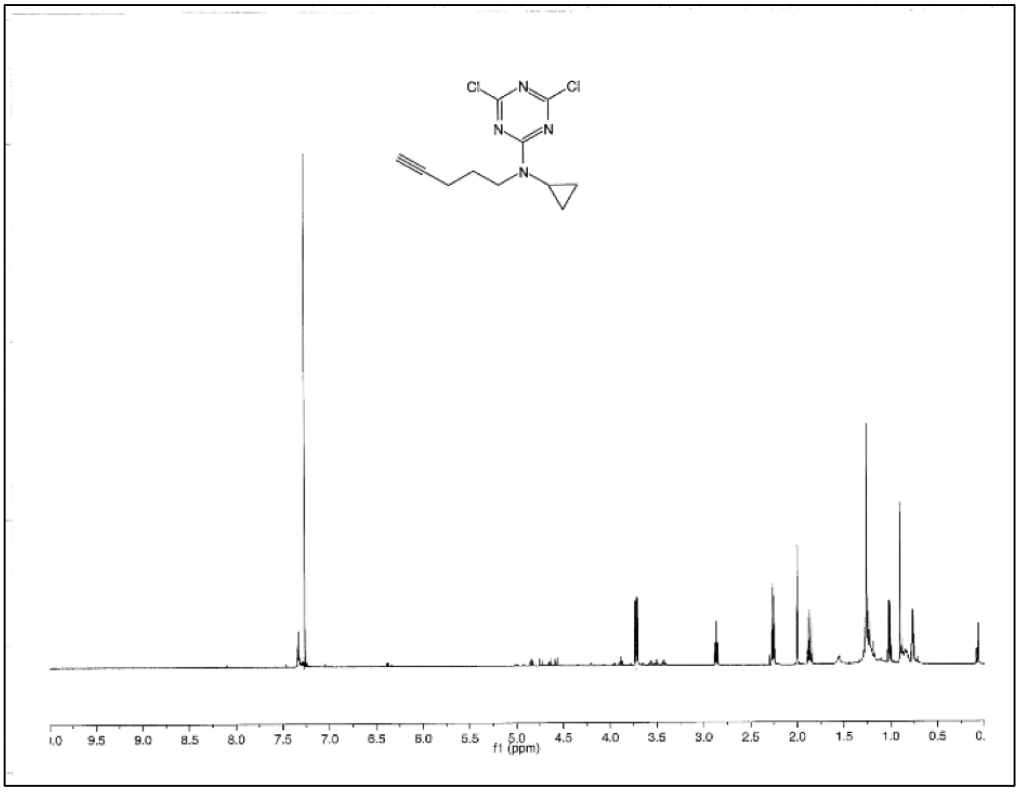


^{13}C-NMR

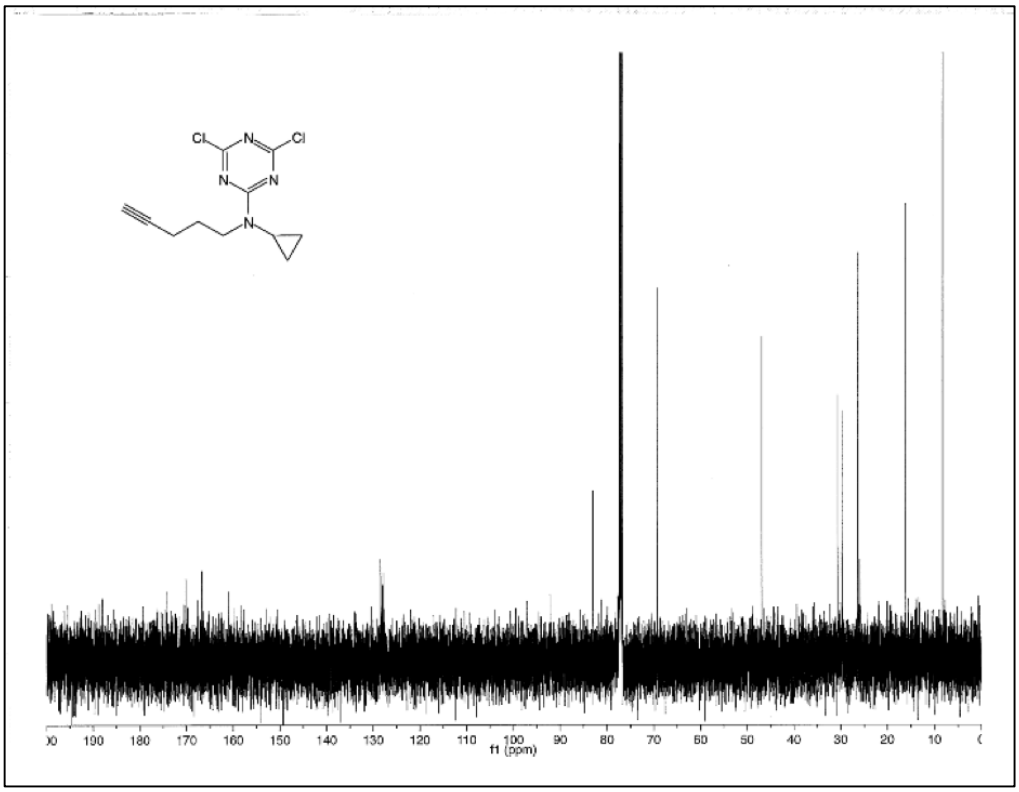


4,6-dichloro-N-cyclopropyl-N-(pent-4-yn-1-yl)-1,3,5-triazin-2-amine (LAS8)

^{1}H-NMR

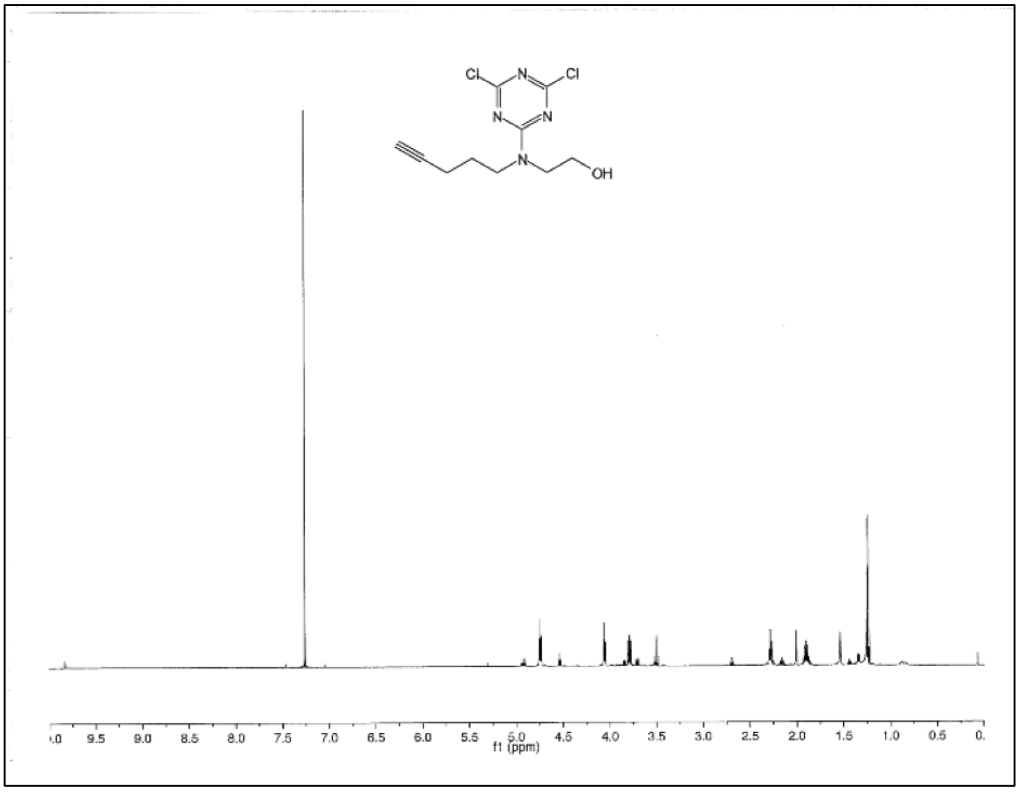


^{13}C-NMR

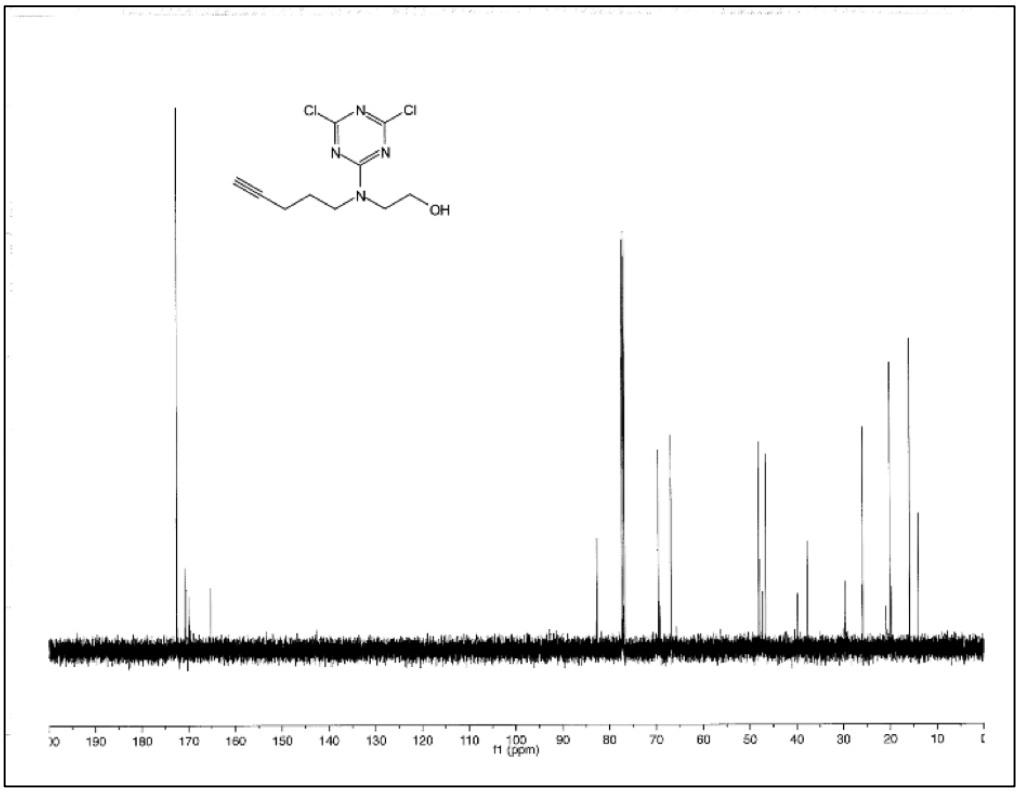


2-((4,6-dichloro-1,3,5-triazin-2-yl)(pent-4-yn-1-yl)amino)ethan-1-ol (LAS8)

^{1}H-NMR

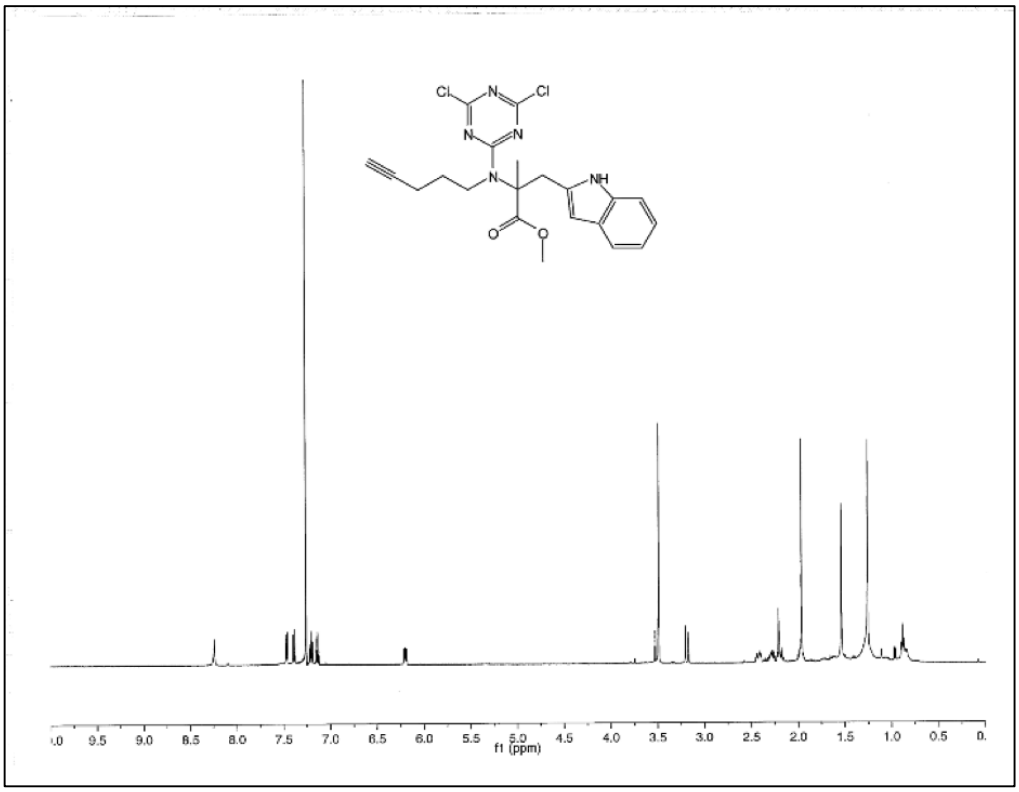


^{13}C-NMR

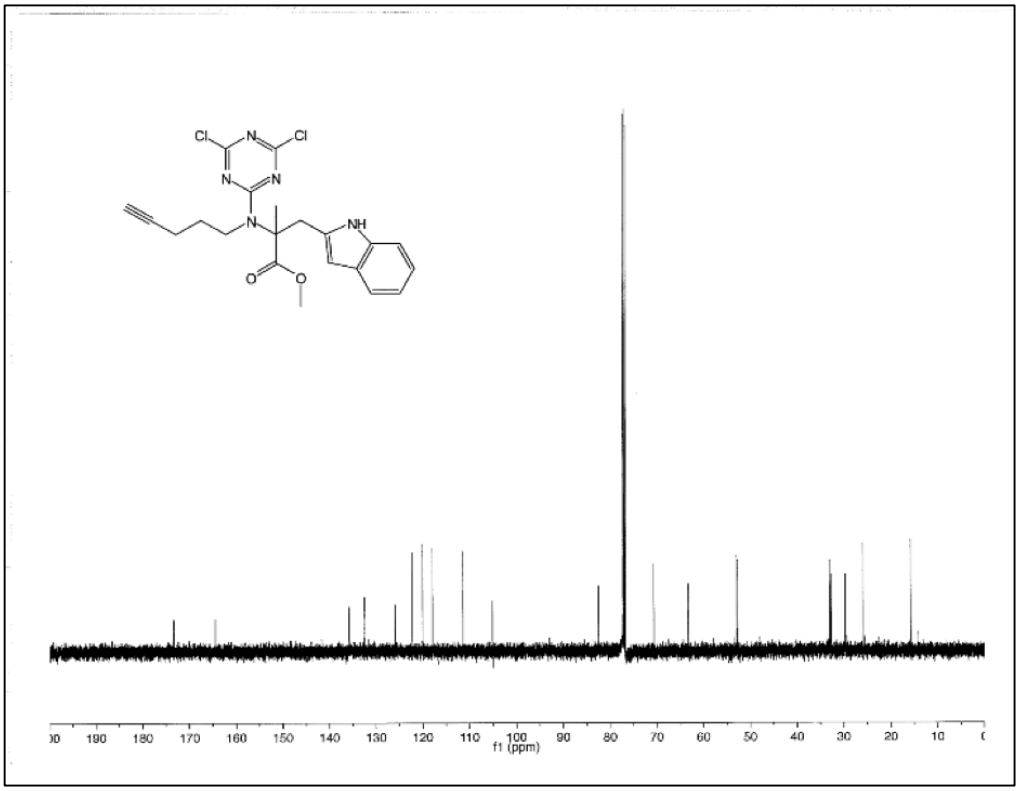


methyl 2-((4,6-dichloro-1,3,5-triazin-2-yl)(pent-4-yn-1-yl)amino)-3-(1H-indol-3-yl)-2-methylpropanoate (LAS10)

^{1}H-NMR

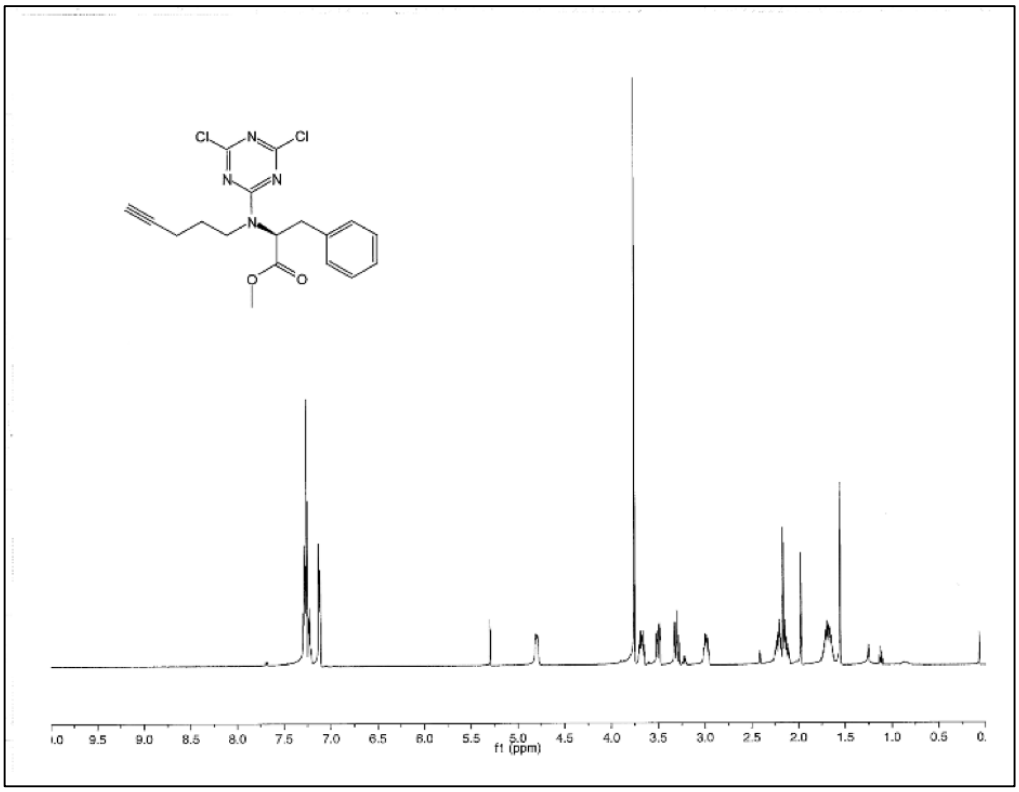


^{13}C-NMR

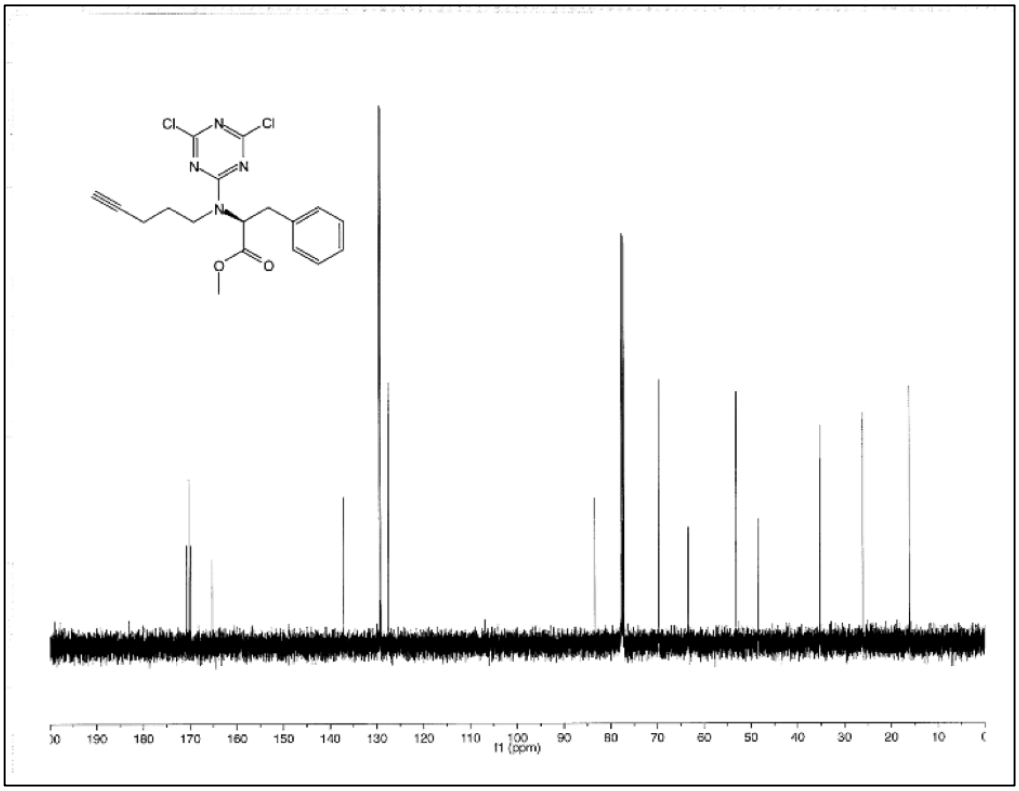


methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-L-phenylalaninate (LAS11)

^{1}H-NMR

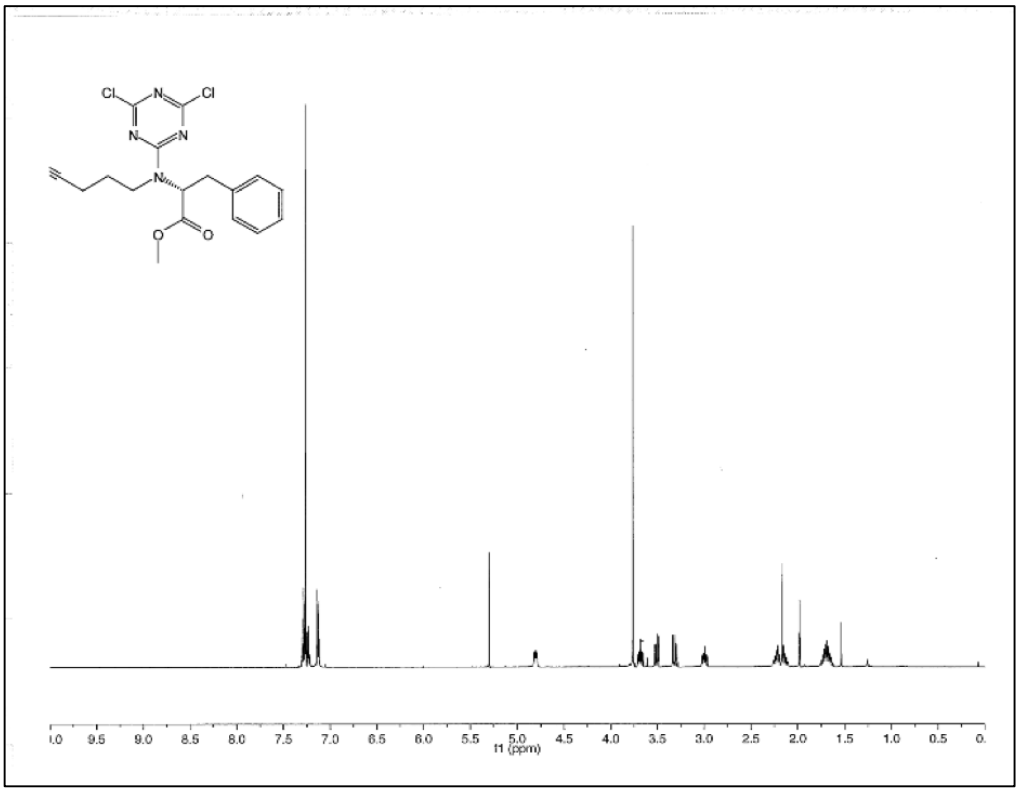


^{13}C-NMR

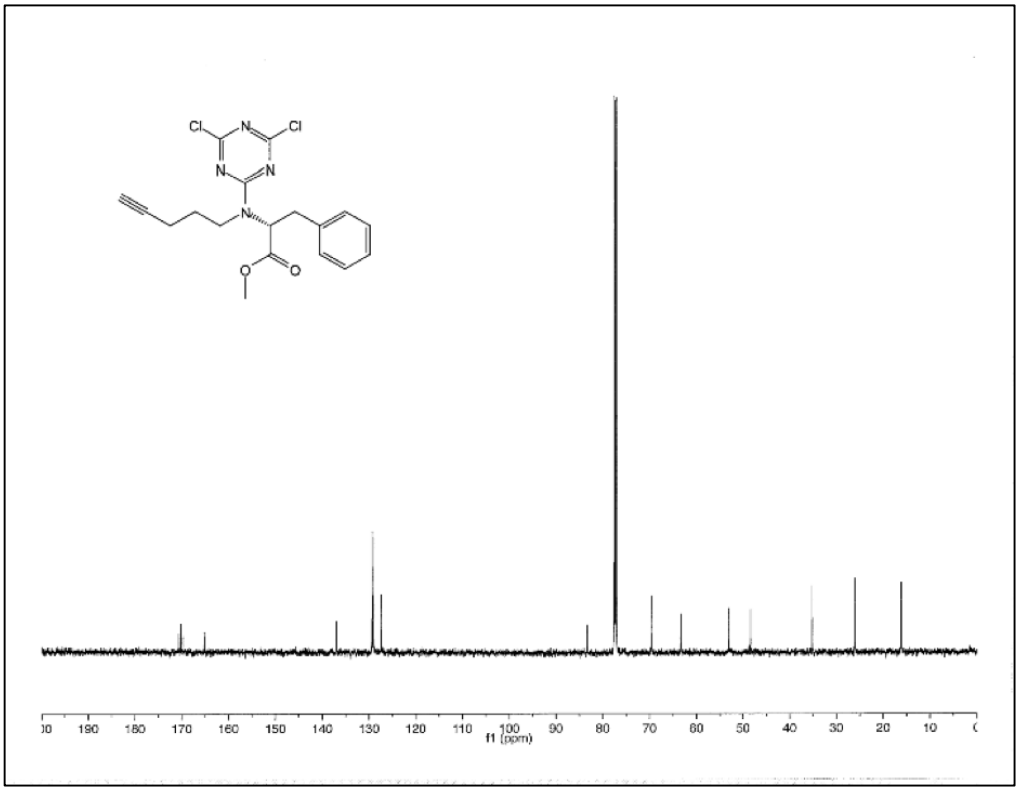


methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-D-phenylalaninate (LAS12)

^{1}H-NMR

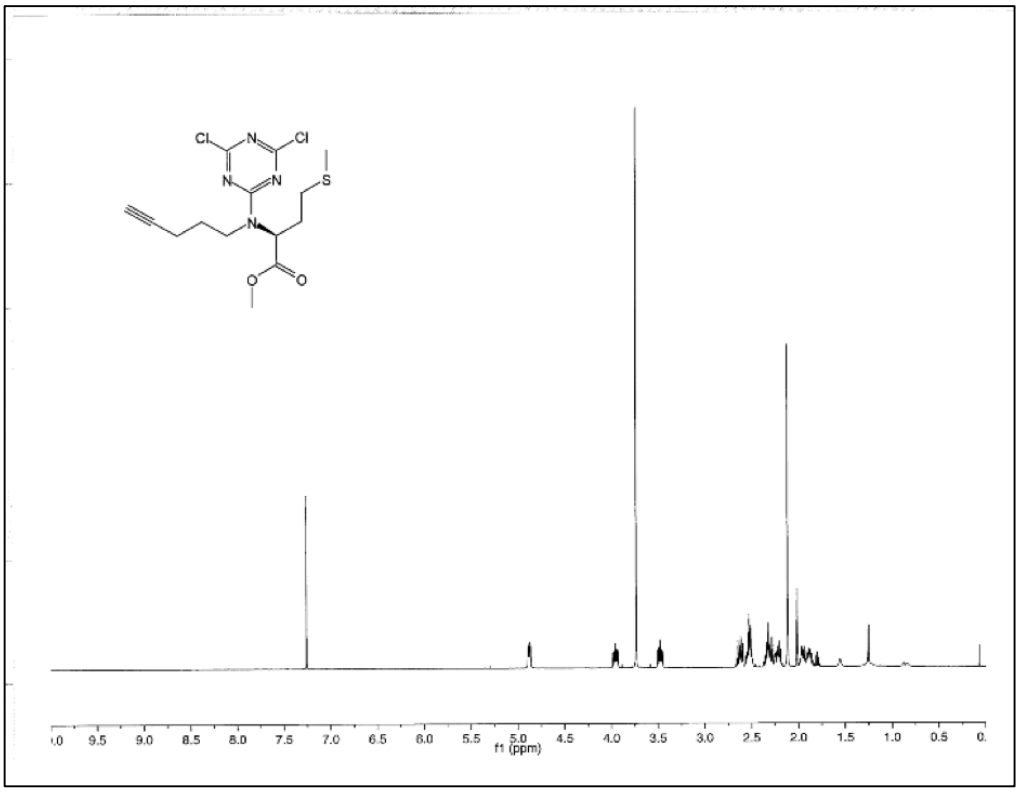


^{13}C-NMR

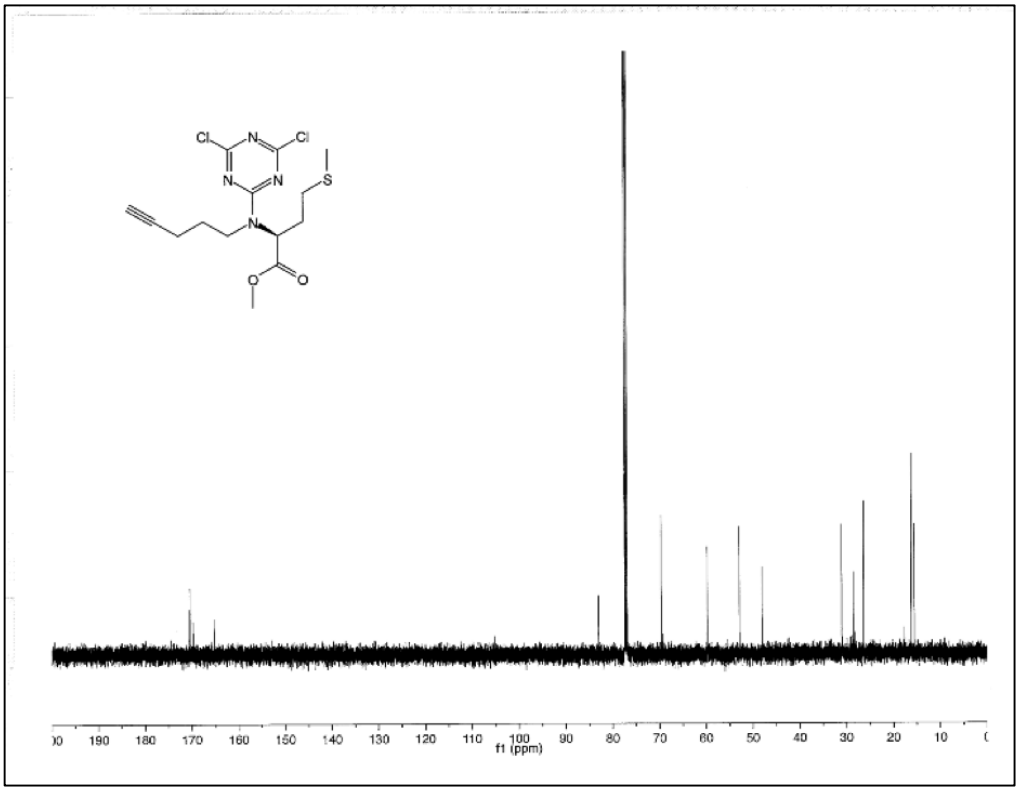


methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-L-methioninate (LAS13)

^{1}H-NMR

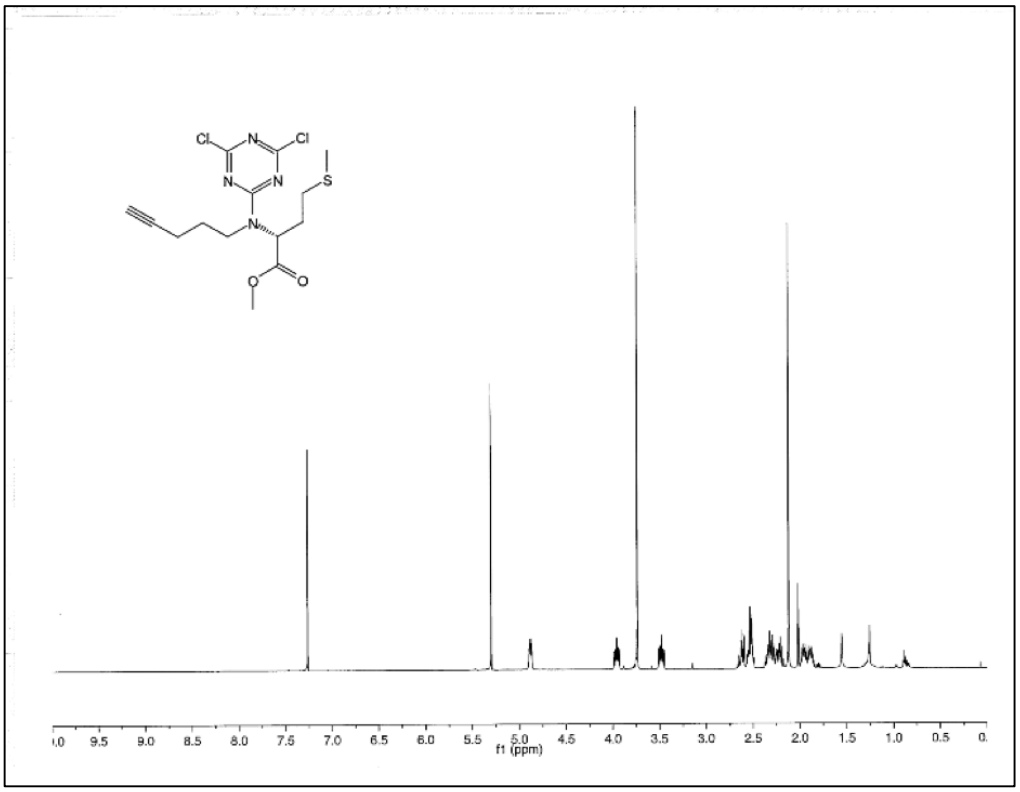


^{13}C-NMR

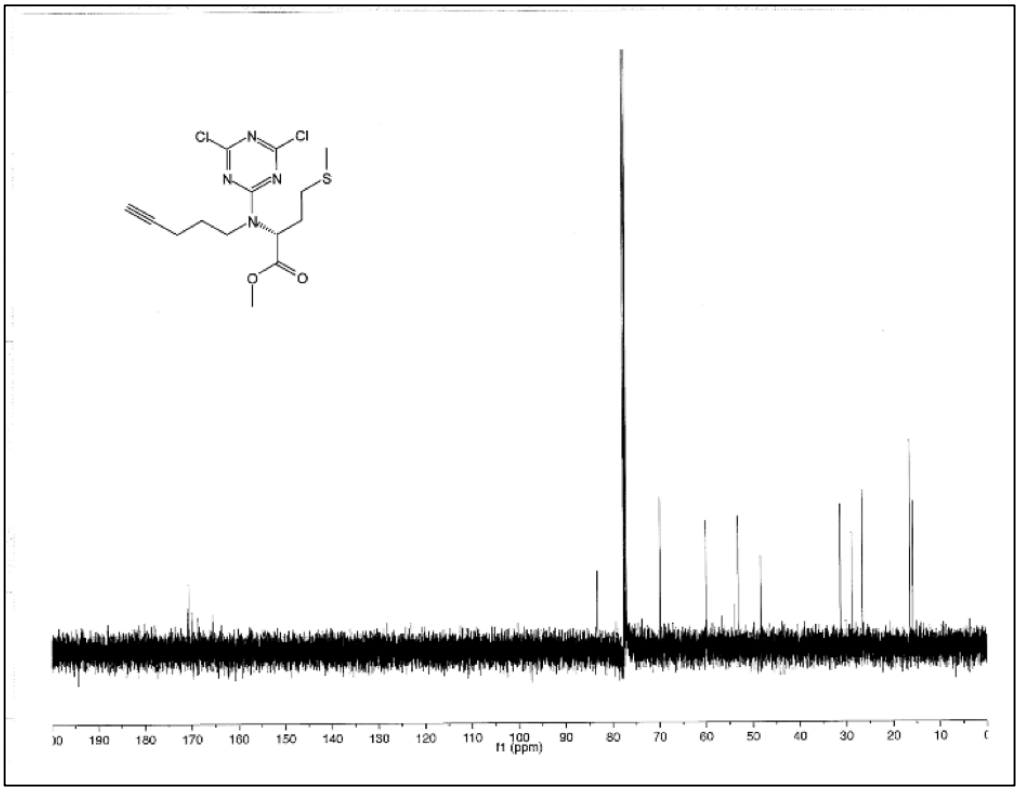


methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-D-methioninate (LAS14)

^{1}H-NMR

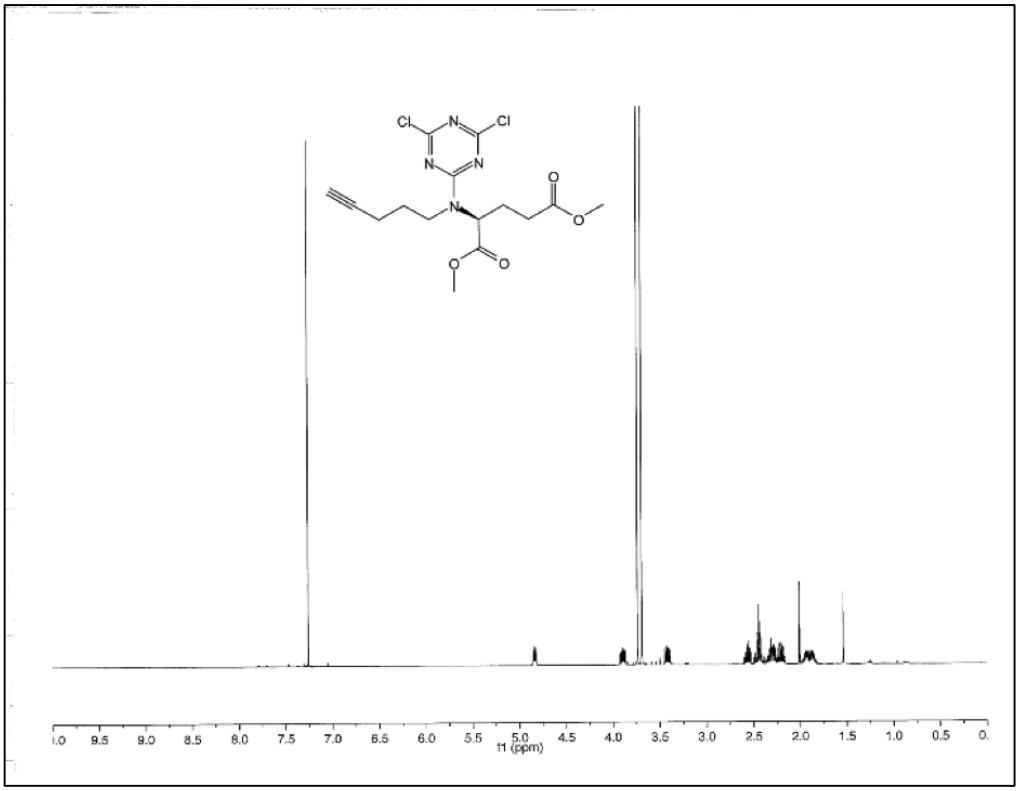


^{13}C-NMR

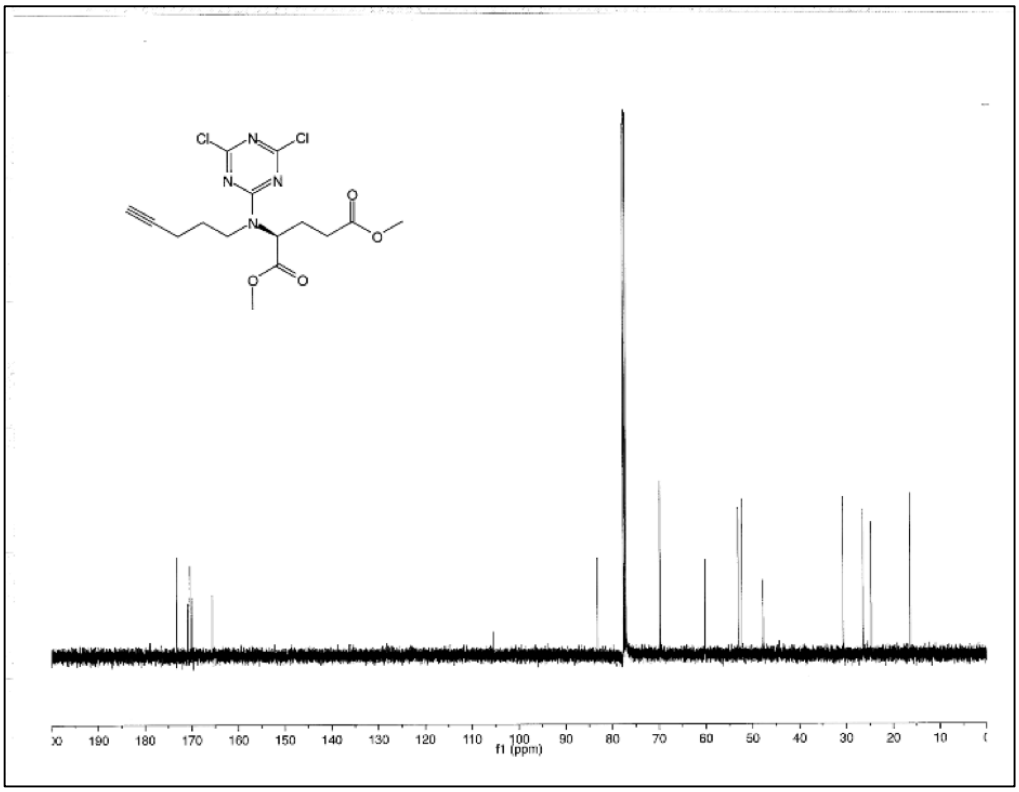


dimethyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-L-glutamate (LAS15)

^{1}H-NMR

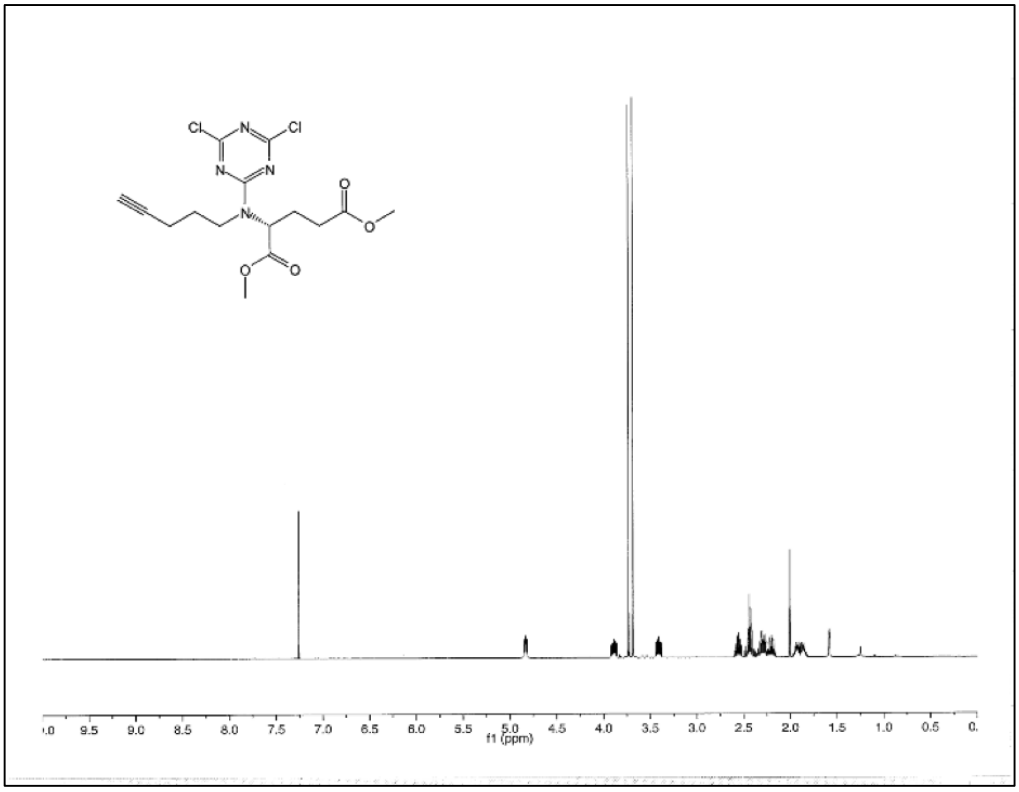


^{13}C-NMR

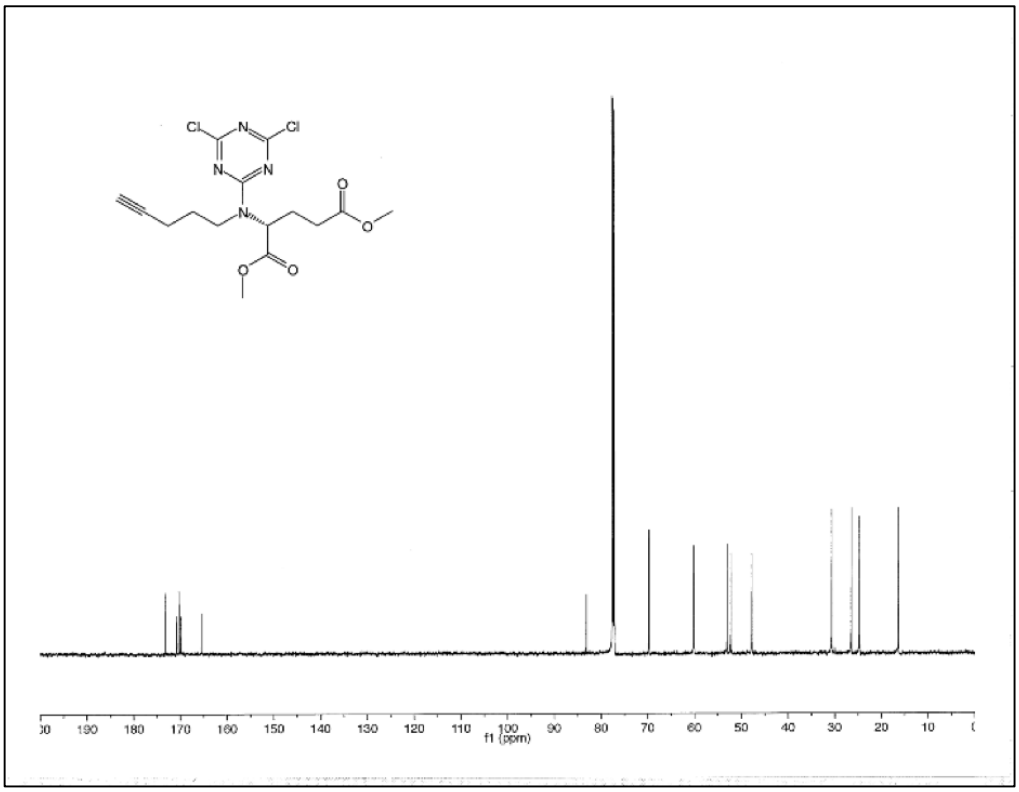


dimethyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-D-glutamate (LAS16)

^{1}H-NMR

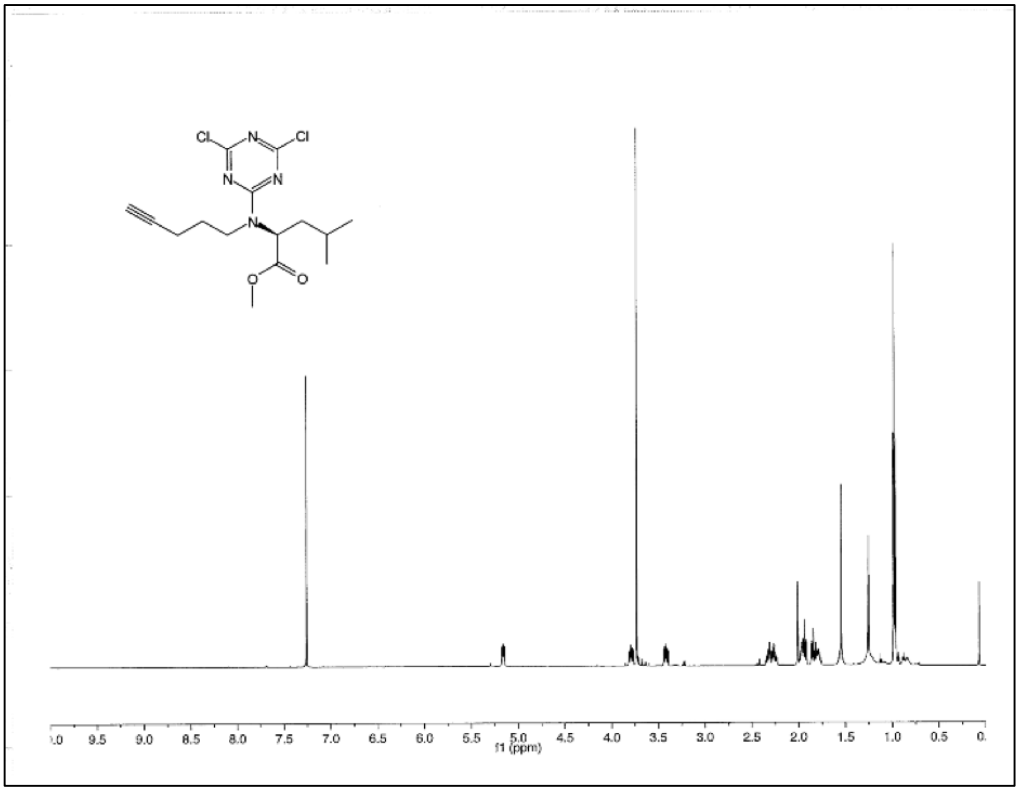


^{13}C-NMR

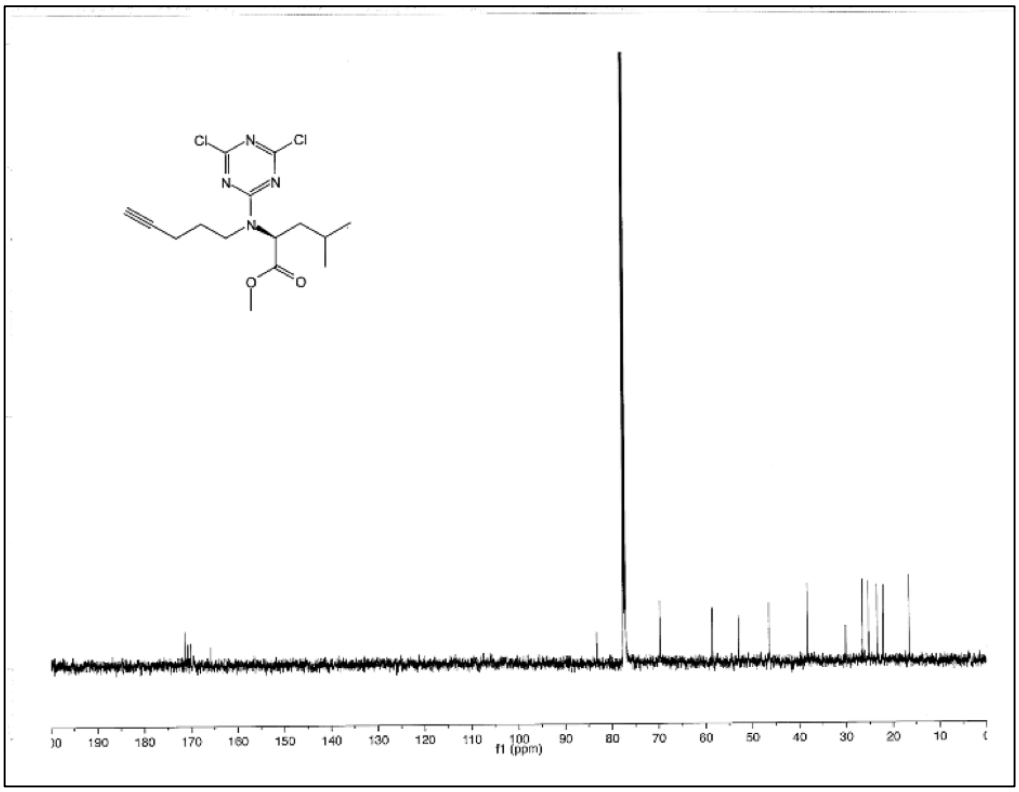


methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-L-leucinate (LAS17)

^{1}H-NMR

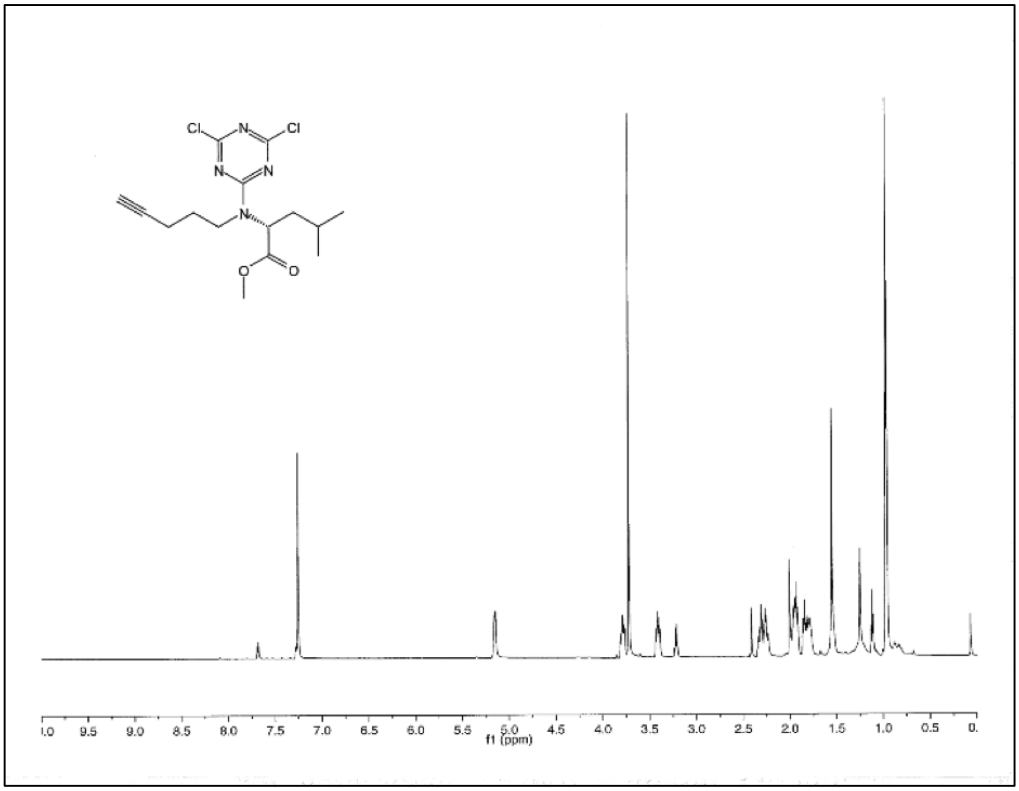


^{13}C-NMR

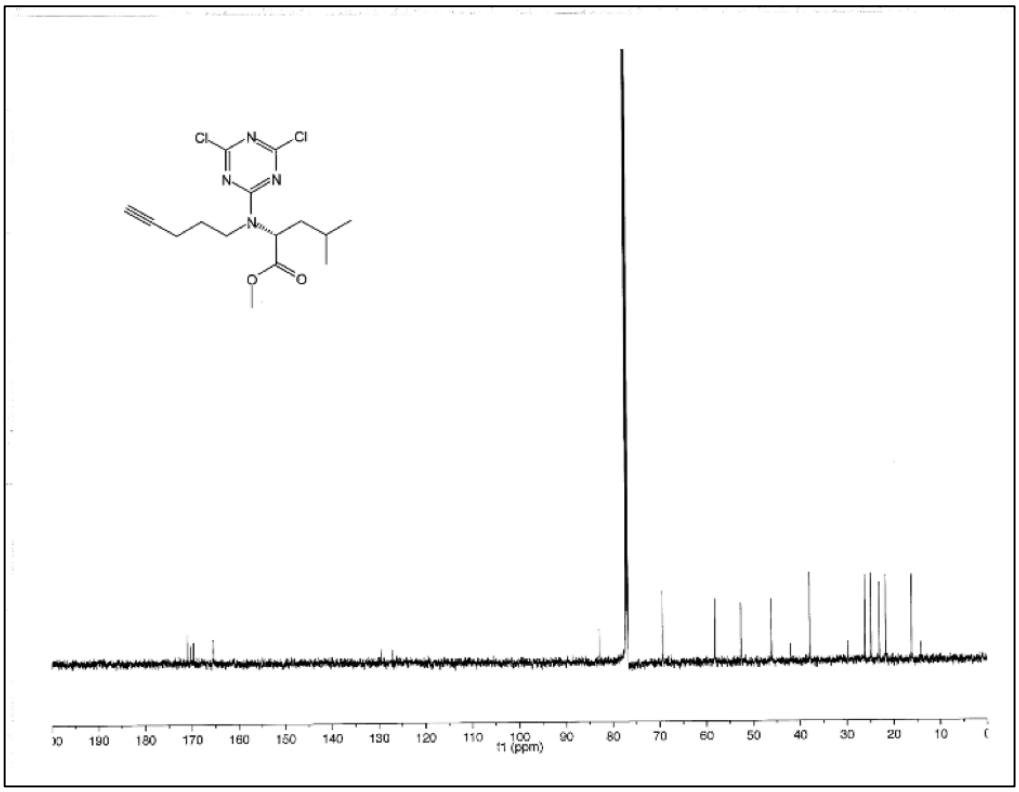


methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-D-leucinate (LAS18)

^{1}H-NMR

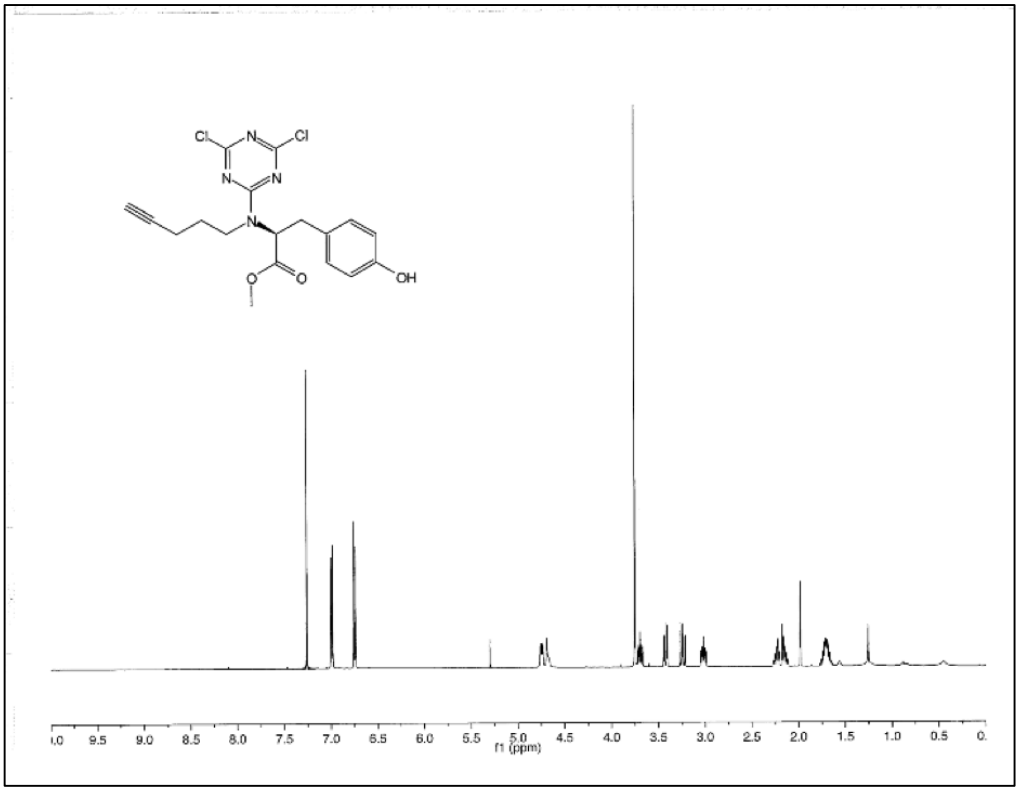


^{13}C-NMR

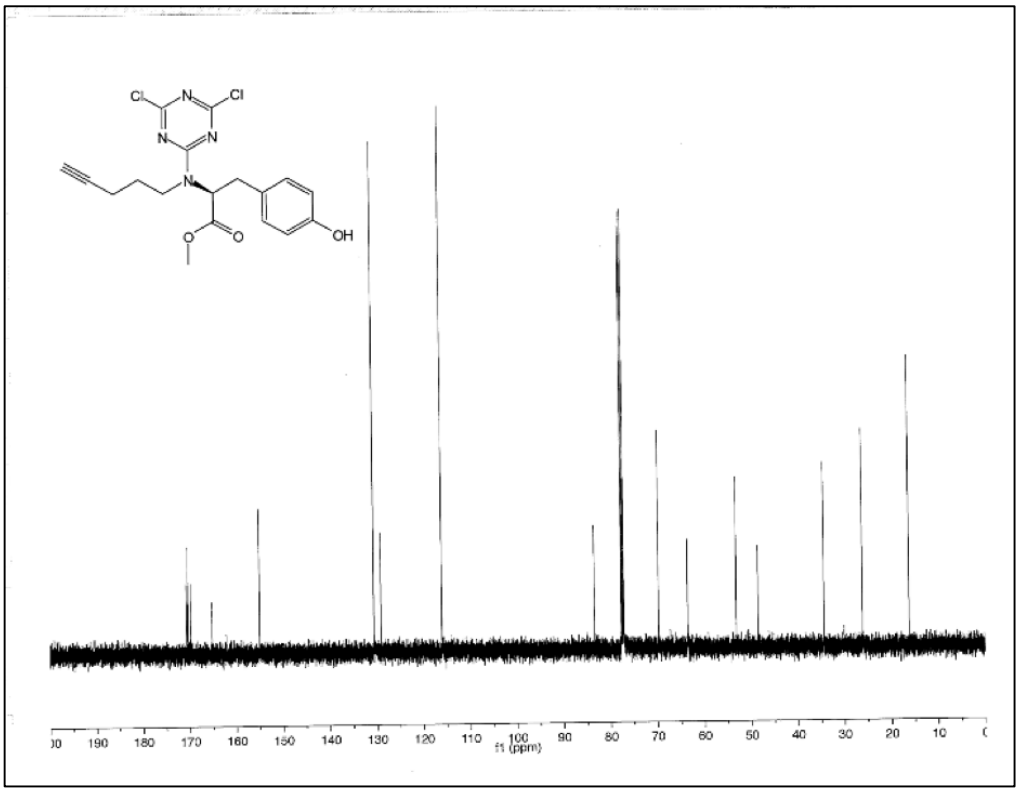


methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-L-tyrosinate (LAS19)

^{1}H-NMR

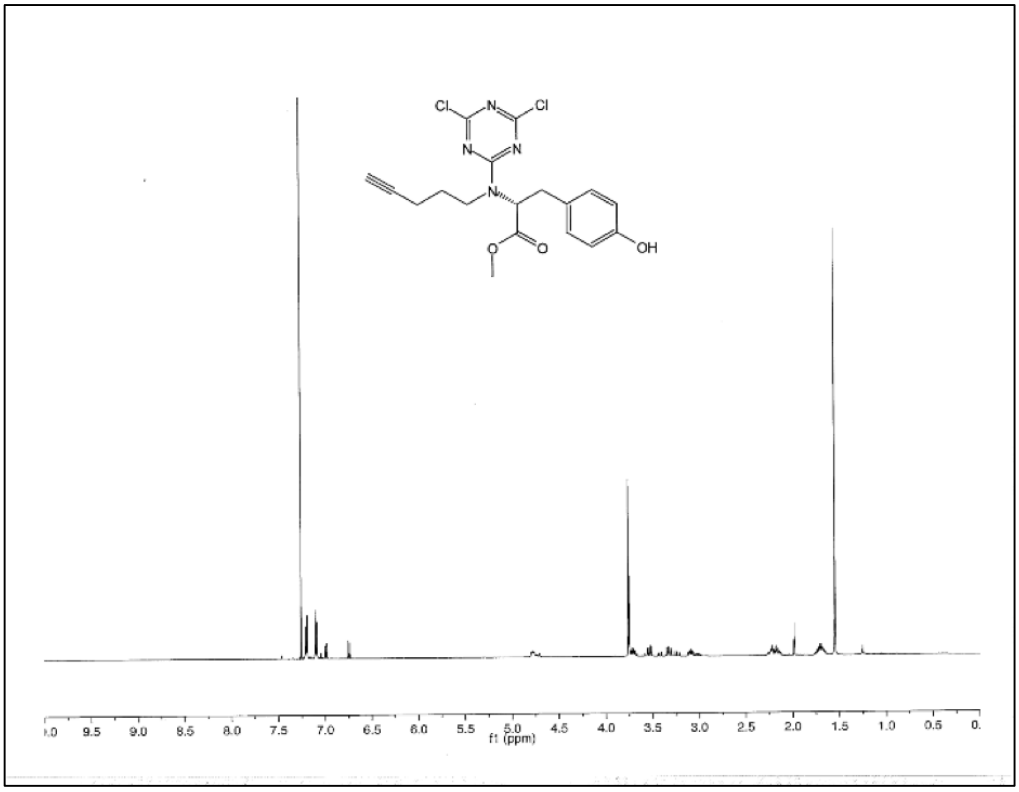


^{13}C-NMR

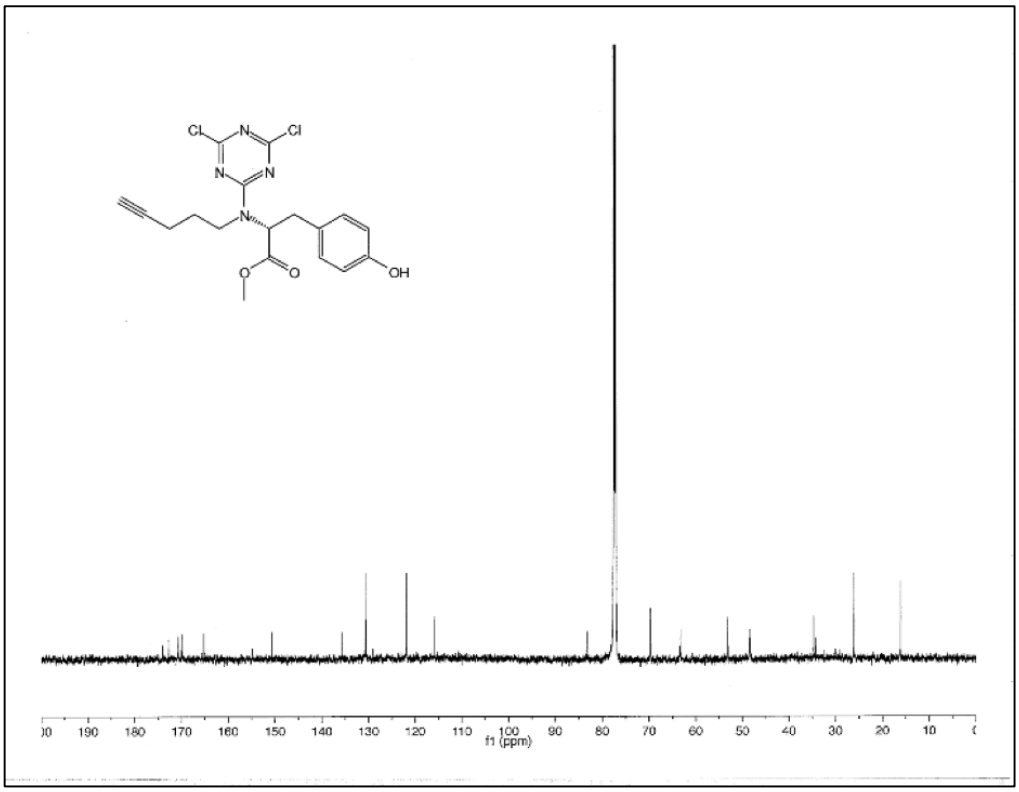


methyl N-(4,6-dichloro-1,3,5-triazin-2-yl)-N-(pent-4-yn-1-yl)-D-tyrosinate (LAS20)

^{1}H-NMR



^{13}C-NMR



Appendix II

Mass spectrometry tables

Table 1A-1. Tryptic digests of HeLa lysates +/- LAS17.

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
P09211	GSTP1_HUMAN - GSTP1 Glutathione S-transferase P	23342	0	0	0	36	23	13	0.0	24.0	24.0
P00558	PGK1_HUMAN - PGK1 Phosphoglycerate kinase 1	44615	3	0	0	3	8	31	1.0	14.0	14.0
P07900	HS90A_HUMAN - HSP90AA1 Heat shock protein HSP 90-alpha	84660	0	0	0	5	0	31	0.0	12.0	12.0
P02545	LMNA_HUMAN - LMNA Prelamin-A/C	74140	2	0	0	0	3	17	0.7	6.7	10.0
P04406	G3P_HUMAN - GAPDH Glyceraldehyde-3-phosphate dehydrogenase	36053	2	0	0	2	0	16	0.7	6.0	9.0
Q01813	K6PP_HUMAN - PFKP 6-phosphofructokinase type C	85596	2	0	0	7	5	6	0.7	6.0	9.0
P13639	EF2_HUMAN - EEF2 Elongation factor 2	95338	0	0	2	6	0	11	0.7	5.7	8.5
P26639	SYTC_HUMAN - TARS Threonine--tRNA ligase, cytoplasmic	83435	2	0	0	2	3	11	0.7	5.3	8.0
Q05639	EF1A2_HUMAN - EEF1A2 Elongation factor 1-alpha 2	50470	0	0	4	4	0	26	1.3	10.0	7.5
P62736	ACTA_HUMAN - ACTA2 Actin, aortic smooth muscle	42009	0	0	2	0	0	15	0.7	5.0	7.5
P68133	ACTS_HUMAN - ACTA1 Actin, alpha skeletal muscle	42051	0	0	2	0	0	15	0.7	5.0	7.5
P11021	GRP78_HUMAN - HSPA5 78 kDa glucose-regulated protein	72333	0	0	0	0	2	19	0.0	7.0	7.0
P29401	TKT_HUMAN - TKT Transketolase	67878	7	0	3	15	8	46	3.3	23.0	6.9
P08238	HS90B_HUMAN - HSP90AB1 Heat shock protein HSP 90-beta	83264	11	0	0	8	11	53	3.7	24.0	6.5
P04075	ALDOA_HUMAN - ALDOA Fructose-bisphosphate aldolase A	39420	2	0	0	2	0	11	0.7	4.3	6.5
P11142	HSP7C_HUMAN - HSPA8 Heat shock cognate 71 kDa protein	70898	0	0	2	5	0	8	0.7	4.3	6.5
P62258	1433E_HUMAN - YWHAE 14-3-3 protein epsilon	29174	0	0	0	0	0	19	0.0	6.3	6.3
P06744	G6PI_HUMAN - GPI Glucose-6-phosphate isomerase	63147	2	0	0	0	0	12	0.7	4.0	6.0
P18669	PGAM1_HUMAN - PGAM1 Phosphoglycerate mutase 1	28804	0	0	0	0	0	18	0.0	6.0	6.0
P14625	ENPL_HUMAN - HSP90B1 Endoplasmin	92469	2	0	0	0	0	11	0.7	3.7	5.5
P68366	TBA4A_HUMAN - TUBA4A Tubulin alpha-4A chain	49924	8	0	5	4	13	48	4.3	21.7	5.0
O43707	ACTN4_HUMAN - ACTN4 Alpha-actinin-4	104854	0	0	3	0	3	12	1.0	5.0	5.0
P10809	CH60_HUMAN - HSPD1 60 kDa heat shock protein, mitochondrial	61055	0	0	0	4	0	11	0.0	5.0	5.0
P12956	XRCC6_HUMAN - XRCC6 X-ray repair cross-complementing protein 6	69843	0	0	0	0	2	13	0.0	5.0	5.0
P07195	LDHB_HUMAN - LDHB L-lactate dehydrogenase B chain	36639	0	0	0	6	3	5	0.0	4.7	4.7
Q9Y490	TLN1_HUMAN - TLN1 Talin-1	269765	0	0	0	3	5	6	0.0	4.7	4.7
Q16658	FSCN1_HUMAN - FSCN1 Fascin	54530	0	0	0	0	0	14	0.0	4.7	4.7
Q15149	PLEC_HUMAN - PLEC Plectin	531796	2	0	0	2	0	7	0.7	3.0	4.5
Q14697	GANAB_HUMAN - GANAB Neutral alpha-glucosidase AB	106874	3	0	0	2	0	11	1.0	4.3	4.3
P55072	TERA_HUMAN - VCP Transitional endoplasmic reticulum ATPase	89322	0	0	0	0	0	13	0.0	4.3	4.3

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
P14618	KPYM_HUMAN - PKM Pyruvate kinase isozymes M1/M2	57937	7	5	5	7	5	55	5.7	22.3	3.9
P68104	EF1A1_HUMAN - EEF1A1 Elongation factor 1-alpha 1	50141	12	0	4	4	0	57	5.3	20.3	3.8
P07205	PGK2_HUMAN - PGK2 Phosphoglycerate kinase 2	44796	3	0	0	3	8	0	1.0	3.7	3.7
P53396	ACLY_HUMAN - ACLY ATP-citrate synthase	120839	0	0	0	0	0	11	0.0	3.7	3.7
P60174	TPIS_HUMAN - TPI1 Triosephosphate isomerase	30791	5	0	0	0	0	18	1.7	6.0	3.6
P21333	FLNA_HUMAN - FLNA Filamin-A	280737	12	0	5	9	9	42	5.7	20.0	3.5
P31948	STIP1_HUMAN - STIP1 Stress-induced-phosphoprotein 1	62639	0	0	2	0	0	7	0.7	2.3	3.5
P07437	TBB5_HUMAN - TUBB Tubulin beta chain	49671	5	8	4	15	10	33	5.7	19.3	3.4
P38646	GRP75_HUMAN - HSPA9 Stress-70 protein, mitochondrial	73681	0	0	0	0	2	8	0.0	3.3	3.3
P68371	TBB4B_HUMAN - TUBB4B Tubulin beta-4B chain	49831	5	8	0	0	10	33	4.3	14.3	3.3
P35579	MYH9_HUMAN - MYH9 Myosin-9	226530	5	0	4	5	3	21	3.0	9.7	3.2
P06733	ENOA_HUMAN - ENO1 Alpha-enolase	47169	4	0	2	3	0	16	2.0	6.3	3.2
P49327	FAS_HUMAN - FASN Fatty acid synthase	273424	3	0	2	0	2	13	1.7	5.0	3.0
P40926	MDHM_HUMAN - MDH2 Malate dehydrogenase, mitochondrial	35503	0	0	0	0	0	9	0.0	3.0	3.0
Q00610	CLH1_HUMAN - CLTC Clathrin heavy chain 1	191613	4	0	4	5	4	14	2.7	7.7	2.9
Q14315	FLNC_HUMAN - FLNC Filamin-C	291020	3	0	2	7	3	4	1.7	4.7	2.8
P31327	CPSM_HUMAN - CPS1 Carbamoyl-phosphate synthase	164939	0	0	3	3	0	5	1.0	2.7	2.7
O75369	FLNB_HUMAN - FLNB Filamin-B	278162	0	0	0	3	0	5	0.0	2.7	2.7
P60660	MYL6_HUMAN - MYL6 Myosin light polypeptide 6	16930	0	0	0	0	0	8	0.0	2.7	2.7
P22314	UBA1_HUMAN - UBA1 Ubiquitin-like modifier-activating enzyme 1	117849	0	0	0	0	0	8	0.0	2.7	2.7
P17812	PYRG1_HUMAN - CTPS1 CTP synthase 1	66690	0	0	0	0	0	8	0.0	2.7	2.7
P60709	ACTB_HUMAN - ACTB Actin, cytoplasmic 1	41737	14	2	10	10	3	48	8.7	20.3	2.3
P08107	HSP71_HUMAN - HSPA1B Heat shock 70 kDa protein 1A/1B	70052	0	0	0	3	0	4	0.0	2.3	2.3
P07237	PDIA1_HUMAN - P4HB Protein disulfide-isomerase	57116	0	0	0	2	0	5	0.0	2.3	2.3
P00338	LDHA_HUMAN - LDHA L-lactate dehydrogenase A chain	36689	0	0	0	0	0	7	0.0	2.3	2.3
P25786	PSA1_HUMAN - PSMA1 Proteasome subunit alpha type-1	29556	0	0	0	0	0	7	0.0	2.3	2.3
B4DEV8	B4DEV8_HUMAN - PSMA1 Proteasome subunit alpha type	18458	0	0	0	0	0	7	0.0	2.3	2.3
Q04695	K1C17_HUMAN - KRT17 Keratin, type I cytoskeletal 17	48106	2	0	0	2	0	2	0.7	1.3	2.0
P17987	TCPA_HUMAN - TCP1 T-complex protein 1 subunit alpha	60344	0	0	2	4	0	0	0.7	1.3	2.0
P07477	TRY1_HUMAN - PRSS1 Trypsin-1	26558	0	0	0	2	4	0	0.0	2.0	2.0
P07478	TRY2_HUMAN - PRSS2 Trypsin-2	26488	0	0	0	2	4	0	0.0	2.0	2.0
P52209	6PGD_HUMAN - PGD 6-phosphogluconate dehydrogenase, decarboxylating	53140	0	0	0	0	2	4	0.0	2.0	2.0

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
P48643	TCPE_HUMAN - CCT5 T-complex protein 1 subunit epsilon	59671	0	0	0	0	2	4	0.0	2.0	2.0
P09488	GSTM1_HUMAN - GSTM1 Glutathione S-transferase Mu 1	25712	0	0	0	0	6	0	0.0	2.0	2.0
Q03013	GSTM4_HUMAN - GSTM4 Glutathione S-transferase Mu 4	25561	0	0	0	0	6	0	0.0	2.0	2.0
P46439	GSTM5_HUMAN - GSTM5 Glutathione S-transferase Mu 5	25675	0	0	0	0	6	0	0.0	2.0	2.0
P31040	DHSA_HUMAN - SDHA Succinate dehydrogenase	72692	0	0	0	0	2	4	0.0	2.0	2.0
P31153	METK2_HUMAN - MAT2A S-adenosylmethionine synthase isoform type-2	43661	0	0	0	0	0	6	0.0	2.0	2.0
Q562R1	ACTBL_HUMAN - ACTBL2 Beta-actin-like protein 2	42003	0	0	0	0	0	6	0.0	2.0	2.0
P14324	FPPS_HUMAN - FDPS Farnesyl pyrophosphate synthase	48276	0	0	0	0	0	6	0.0	2.0	2.0
O00231	PSD11_HUMAN - PSMD11 26S proteasome non-ATPase regulatory subunit 11	47464	0	0	0	0	0	6	0.0	2.0	2.0
Q92769	HDAC2_HUMAN - HDAC2 Histone deacetylase 2	55364	0	0	0	0	0	6	0.0	2.0	2.0
Q13813	SPTN1_HUMAN - SPTAN1 Spectrin alpha chain, non-erythrocytic 1	284538	0	0	0	0	0	6	0.0	2.0	2.0
P50454	SERPH_HUMAN - SERPINH1 Serpin H1	46441	0	0	0	0	0	6	0.0	2.0	2.0
P28074	PSB5_HUMAN - PSMB5 Proteasome subunit beta type-5	28480	0	0	0	2	0	3	0.0	1.7	1.7
P23528	COF1_HUMAN - CFL1 Cofilin-1	18502	0	0	0	0	2	3	0.0	1.7	1.7
Q02790	FKBP4_HUMAN - FKBP4 Peptidyl-prolyl cis-trans isomerase FKBP4	51805	0	0	0	0	2	3	0.0	1.7	1.7
O00763	ACACB_HUMAN - ACACB Acetyl-CoA carboxylase 2	276539	0	0	0	0	5	0	0.0	1.7	1.7
P26447	S10A4_HUMAN - S100A4 Protein S100-A4	11728	0	0	0	0	0	5	0.0	1.7	1.7
Q13162	PRDX4_HUMAN - PRDX4 Peroxiredoxin-4	30540	0	0	0	0	0	5	0.0	1.7	1.7
P36873	PP1G_HUMAN - PPP1CC Serine/threonine-protein phosphatase PP1-gamma catalytic subunit	36984	0	0	0	0	0	5	0.0	1.7	1.7
P62136	PP1A_HUMAN - PPP1CA Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	37512	0	0	0	0	0	5	0.0	1.7	1.7
P62140	PP1B_HUMAN - PPP1CB Serine/threonine-protein phosphatase PP1-beta catalytic subunit	37187	0	0	0	0	0	5	0.0	1.7	1.7
P08670	VIME_HUMAN - VIM Vimentin	53652	0	0	0	0	0	5	0.0	1.7	1.7
P50395	GDIB_HUMAN - GDI2 Rab GDP dissociation inhibitor beta	50663	0	0	0	0	0	5	0.0	1.7	1.7
Q9H4A4	AMPB_HUMAN - RNPEP Aminopeptidase B	72596	0	0	0	0	0	5	0.0	1.7	1.7
P20618	PSB1_HUMAN - PSMB1 Proteasome subunit beta type-1	26489	0	0	0	0	0	5	0.0	1.7	1.7
P13797	PLST_HUMAN - PLS3 Plastin-3	70811	0	0	0	0	0	5	0.0	1.7	1.7
P0CG48	UBC_HUMAN - UBC Polyubiquitin-C	77039	0	0	0	0	0	5	0.0	1.7	1.7
J3QRK5	J3QRK5_HUMAN - UBBP4 Protein UBBP4	25797	0	0	0	0	0	5	0.0	1.7	1.7
P62987	RL40_HUMAN - UBA52 Ubiquitin-60S ribosomal protein L40	14728	0	0	0	0	0	5	0.0	1.7	1.7

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
P62979	RS27A_HUMAN - RPS27A Ubiquitin-40S ribosomal protein S27a	17965	0	0	0	0	0	5	0.0	1.7	1.7
P07814	SYEP_HUMAN - EPRS Bifunctional glutamate/proline--tRNA ligase	170590	0	0	0	0	0	5	0.0	1.7	1.7
O43143	DHX15_HUMAN - DHX15 Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	90933	0	0	0	0	0	5	0.0	1.7	1.7
Q71U36	TBA1A_HUMAN - TUBA1A Tubulin alpha-1A chain	50136	8	0	5	4	14	0	4.3	6.0	1.4
Q13748	TBA3C_HUMAN - TUBA3D Tubulin alpha-3C/D chain	49960	8	0	5	4	14	0	4.3	6.0	1.4
P02649	APOE_HUMAN - APOE Apolipoprotein E	36154	9	0	2	7	6	2	3.7	5.0	1.4
P49368	TCPG_HUMAN - CCT3 T-complex protein 1 subunit gamma	60534	0	0	0	4	0	0	0.0	1.3	1.3
O75874	IDHC_HUMAN - IDH1 Isocitrate dehydrogenase	46659	0	0	0	0	2	2	0.0	1.3	1.3
Q93008	USP9X_HUMAN - USP9X Probable ubiquitin carboxyl-terminal hydrolase FAF-X	292278	0	0	0	0	2	2	0.0	1.3	1.3
P60842	IF4A1_HUMAN - EIF4A1 Eukaryotic initiation factor 4A-I	46154	0	0	0	0	0	4	0.0	1.3	1.3
P42330	AK1C3_HUMAN - AKR1C3 Aldo-keto reductase family 1 member C3	36853	0	0	0	0	0	4	0.0	1.3	1.3
P61158	ARP3_HUMAN - ACTR3 Actin-related protein 3	47371	0	0	0	0	0	4	0.0	1.3	1.3
P78417	GSTO1_HUMAN - GSTO1 Glutathione S-transferase omega-1	27566	0	0	0	0	0	4	0.0	1.3	1.3
P43686	PRS6B_HUMAN - PSMC4 26S protease regulatory subunit 6B	47366	0	0	0	0	0	4	0.0	1.3	1.3
P15559	NQO1_HUMAN - NQO1 NAD(P)H dehydrogenase	30868	0	0	0	0	0	4	0.0	1.3	1.3
P54577	SYYC_HUMAN - YARS Tyrosine--tRNA ligase, cytoplasmic	59144	0	0	0	0	0	4	0.0	1.3	1.3
P61160	ARP2_HUMAN - ACTR2 Actin-related protein 2	44761	0	0	0	0	0	4	0.0	1.3	1.3
P35520	CBS_HUMAN - CBS Cystathionine beta-synthase	60587	0	0	0	0	0	4	0.0	1.3	1.3
O75828	CBR3_HUMAN - CBR3 Carbonyl reductase	30850	0	0	0	0	0	4	0.0	1.3	1.3
P16152	CBR1_HUMAN - CBR1 Carbonyl reductase	30375	0	0	0	0	0	4	0.0	1.3	1.3
P12814	ACTN1_HUMAN - ACTN1 Alpha-actinin-1	103058	0	0	0	0	0	4	0.0	1.3	1.3
Q99460	PSMD1_HUMAN - PSMD1 26S proteasome non-ATPase regulatory subunit 1	105836	0	0	0	0	0	4	0.0	1.3	1.3
P41250	SYG_HUMAN - GARS Glycine--tRNA ligase	83166	0	0	0	0	0	4	0.0	1.3	1.3
Q99832	TCPH_HUMAN - CCT7 T-complex protein 1 subunit eta	59367	0	0	0	0	0	4	0.0	1.3	1.3
B0V043	B0V043_HUMAN - VARS Valyl-tRNA synthetase	140466	0	0	0	0	0	4	0.0	1.3	1.3
P62266	RS23_HUMAN - RPS23 40S ribosomal protein S23	15808	3	0	2	0	0	6	1.7	2.0	1.2
P09429	HMGB1_HUMAN - HMGB1 High mobility group protein B1	24894	0	0	5	0	0	6	1.7	2.0	1.2
P11498	PYC_HUMAN - PC Pyruvate carboxylase, mitochondrial	129634	40	34	45	52	43	37	39.7	44.0	1.1
P04792	HSPB1_HUMAN - HSPB1 Heat shock protein beta-1	22783	2	0	0	0	0	2	0.7	0.7	1.0

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
O76009	KT33A_HUMAN - KRT33A Keratin, type I cuticular Ha3-I	45940	2	0	0	0	2	0	0.7	0.7	1.0
P02533	K1C14_HUMAN - KRT14 Keratin, type I cytoskeletal 14	51561	2	0	0	2	0	0	0.7	0.7	1.0
Q14566	MCM6_HUMAN - MCM6 DNA replication licensing factor MCM6	92889	0	0	2	0	0	2	0.7	0.7	1.0
P17066	HSP76_HUMAN - HSPA6 Heat shock 70 kDa protein 6	71028	0	0	0	3	0	0	0.0	1.0	1.0
O60361	NDK8_HUMAN - NME2P1 Putative nucleoside diphosphate kinase	15529	0	0	0	0	0	3	0.0	1.0	1.0
P22392	NDKB_HUMAN - NME2 Nucleoside diphosphate kinase B	17298	0	0	0	0	0	3	0.0	1.0	1.0
P36952	SPB5_HUMAN - SERPINB5 Serpin B5	42100	0	0	0	0	0	3	0.0	1.0	1.0
P30101	PDIA3_HUMAN - PDIA3 Protein disulfide-isomerase A3	56782	0	0	0	0	0	3	0.0	1.0	1.0
P23381	SYWC_HUMAN - WARS Tryptophan--tRNA ligase, cytoplasmic	53165	0	0	0	0	0	3	0.0	1.0	1.0
Q01105	SET_HUMAN - SET Protein SET	33489	0	0	0	0	0	3	0.0	1.0	1.0
Q15008	PSMD6_HUMAN - PSMD6 26S proteasome non-ATPase regulatory subunit 6	45531	0	0	0	0	0	3	0.0	1.0	1.0
Q96FW1	OTUB1_HUMAN - OTUB1 Ubiquitin thioesterase OTUB1	31284	0	0	0	0	0	3	0.0	1.0	1.0
Q9H832	UBE2Z_HUMAN - UBE2Z Ubiquitin-conjugating enzyme E2 Z	38210	0	0	0	0	0	3	0.0	1.0	1.0
P00390	GSHR_HUMAN - GSR Glutathione reductase, mitochondrial	56257	0	0	0	0	0	3	0.0	1.0	1.0
Q13838	DX39B_HUMAN - DDX39B Spliceosome RNA helicase DDX39B	48991	0	0	0	0	0	3	0.0	1.0	1.0
P50213	IDH3A_HUMAN - IDH3A Isocitrate dehydrogenase	39592	0	0	0	0	0	3	0.0	1.0	1.0
P31943	HNRH1_HUMAN - HNRNPH1 Heterogeneous nuclear ribonucleoprotein H	49229	0	0	0	0	0	3	0.0	1.0	1.0
P55795	HNRH2_HUMAN - HNRNPH2 Heterogeneous nuclear ribonucleoprotein H2	49264	0	0	0	0	0	3	0.0	1.0	1.0
P26641	EF1G_HUMAN - EEF1G Elongation factor 1-gamma	50119	0	0	0	0	0	3	0.0	1.0	1.0
E7EMT2	E7EMT2_HUMAN - Uncharacterized protein	24182	0	0	0	0	0	3	0.0	1.0	1.0
Q99798	ACON_HUMAN - ACO2 Aconitate hydratase, mitochondrial	85425	0	0	0	0	0	3	0.0	1.0	1.0
P11766	ADHX_HUMAN - ADH5 Alcohol dehydrogenase class-3	39724	0	0	0	0	0	3	0.0	1.0	1.0
Q9Y263	PLAP_HUMAN - PLAA Phospholipase A-2-activating protein	87157	0	0	0	0	0	3	0.0	1.0	1.0
P06576	ATPB_HUMAN - ATP5B ATP synthase subunit beta, mitochondrial	56560	0	0	0	0	0	3	0.0	1.0	1.0
P34932	HSP74_HUMAN - HSPA4 Heat shock 70 kDa protein 4	94331	0	0	0	0	0	3	0.0	1.0	1.0
Q9NSE4	SYIM_HUMAN - IARS2 Isoleucine--tRNA ligase, mitochondrial	113791	0	0	0	0	0	3	0.0	1.0	1.0
Q15029	U5S1_HUMAN - EFTUD2 116 kDa U5 small nuclear ribonucleoprotein component	109436	0	0	0	0	0	3	0.0	1.0	1.0
Q12931	TRAP1_HUMAN - TRAP1 Heat shock protein 75 kDa, mitochondrial	80110	0	0	0	0	0	3	0.0	1.0	1.0
Q14980	NUMA1_HUMAN - NUMA1 Nuclear mitotic apparatus protein 1	238257	0	0	0	0	0	3	0.0	1.0	1.0
Q14974	IMB1_HUMAN - KPNB1 Importin subunit beta-1	97170	0	0	0	0	0	3	0.0	1.0	1.0

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
Q13085	ACACA_HUMAN - ACACA Acetyl-CoA carboxylase 1	265551	50	22	39	43	41	23	37.0	35.7	1.0
P05165	PCCA_HUMAN - PCCA Propionyl-CoA carboxylase alpha chain, mitochondrial	80059	18	7	10	7	14	10	11.7	10.3	0.9
P13645	K1C10_HUMAN - KRT10 Keratin, type I cytoskeletal 10	58827	8	0	15	13	6	0	7.7	6.3	0.8
Q9BVA1	TBB2B_HUMAN - TUBB2B Tubulin beta-2B chain	49953	5	8	0	0	10	0	4.3	3.3	0.8
Q96RQ3	MCCA_HUMAN - MCCC1 Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	80473	15	0	6	4	4	8	7.0	5.3	0.8
Q9UI26	IPO11_HUMAN - IPO11 Importin-11	112535	0	0	0	2	0	0	0.0	0.7	0.7
P09622	DLDH_HUMAN - DLD Dihydrolipoyl dehydrogenase, mitochondrial	54177	0	0	0	2	0	0	0.0	0.7	0.7
P35900	K1C20_HUMAN - KRT20 Keratin, type I cytoskeletal 20	48487	0	0	0	2	0	0	0.0	0.7	0.7
P19012	K1C15_HUMAN - KRT15 Keratin, type I cytoskeletal 15	49212	0	0	0	2	0	0	0.0	0.7	0.7
P13646	K1C13_HUMAN - KRT13 Keratin, type I cytoskeletal 13	49588	0	0	0	2	0	0	0.0	0.7	0.7
P08779	K1C16_HUMAN - KRT16 Keratin, type I cytoskeletal 16	51268	0	0	0	2	0	0	0.0	0.7	0.7
Q14525	KT33B_HUMAN - KRT33B Keratin, type I cuticular Ha3-II	46214	0	0	0	0	2	0	0.0	0.7	0.7
Q8NBS9	TXND5_HUMAN - TXNDC5 Thioredoxin domain-containing protein 5	47629	0	0	0	0	2	0	0.0	0.7	0.7
Q7Z794	K2C1B_HUMAN - KRT77 Keratin, type II cytoskeletal 1b	61901	0	0	0	0	2	0	0.0	0.7	0.7
P31947	1433S_HUMAN - SFN 14-3-3 protein sigma	27774	0	0	0	0	0	2	0.0	0.7	0.7
P33316	DUT_HUMAN - DUT Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial	26563	0	0	0	0	0	2	0.0	0.7	0.7
P52907	CAZA1_HUMAN - CAPZA1 F-actin-capping protein subunit alpha-1	32923	0	0	0	0	0	2	0.0	0.7	0.7
P21266	GSTM3_HUMAN - GSTM3 Glutathione S-transferase Mu 3	26560	0	0	0	0	0	2	0.0	0.7	0.7
P09651	ROA1_HUMAN - HNRNPA1 Heterogeneous nuclear ribonucleoprotein A1	38747	0	0	0	0	0	2	0.0	0.7	0.7
P07355	ANXA2_HUMAN - ANXA2 Annexin A2	38604	0	0	0	0	0	2	0.0	0.7	0.7
P04818	TYSY_HUMAN - TYMS Thymidylate synthase	35716	0	0	0	0	0	2	0.0	0.7	0.7
P55209	NP1L1_HUMAN - NAP1L1 Nucleosome assembly protein 1-like 1	45374	0	0	0	0	0	2	0.0	0.7	0.7
P38919	IF4A3_HUMAN - EIF4A3 Eukaryotic initiation factor 4A-III	46871	0	0	0	0	0	2	0.0	0.7	0.7
O43175	SERA_HUMAN - PHGDH D-3-phosphoglycerate dehydrogenase	56651	0	0	0	0	0	2	0.0	0.7	0.7
O00148	DX39A_HUMAN - DDX39A ATP-dependent RNA helicase DDX39A	49130	0	0	0	0	0	2	0.0	0.7	0.7
P52597	HNRPF_HUMAN - HNRNPF Heterogeneous nuclear ribonucleoprotein F	45672	0	0	0	0	0	2	0.0	0.7	0.7
P62333	PRS10_HUMAN - PSMC6 26S protease regulatory subunit 10B	44173	0	0	0	0	0	2	0.0	0.7	0.7
P07954	FUMH_HUMAN - FH Fumarate hydratase, mitochondrial	54637	0	0	0	0	0	2	0.0	0.7	0.7
Q13263	TIF1B_HUMAN - TRIM28 Transcription intermediary factor 1-beta	88550	0	0	0	0	0	2	0.0	0.7	0.7
P78371	TCPB_HUMAN - CCT2 T-complex protein 1 subunit beta	57488	0	0	0	0	0	2	0.0	0.7	0.7

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
P23526	SAHH_HUMAN - AHCY Adenosylhomocysteinase	47716	0	0	0	0	0	2	0.0	0.7	0.7
P08195	4F2_HUMAN - SLC3A2 4F2 cell-surface antigen heavy chain	67994	0	0	0	0	0	2	0.0	0.7	0.7
P35237	SPB6_HUMAN - SERPINB6 Serpin B6	42622	0	0	0	0	0	2	0.0	0.7	0.7
Q15084	PDIA6_HUMAN - PDIA6 Protein disulfide-isomerase A6	48121	0	0	0	0	0	2	0.0	0.7	0.7
Q9UBB4	ATX10_HUMAN - ATXN10 Ataxin-10	53489	0	0	0	0	0	2	0.0	0.7	0.7
Q15067	ACOX1_HUMAN - ACOX1 Peroxisomal acyl-coenzyme A oxidase 1	74424	0	0	0	0	0	2	0.0	0.7	0.7
P15170	ERF3A_HUMAN - GSPT1 Eukaryotic peptide chain release factor GTP-binding subunit ERF3A	55756	0	0	0	0	0	2	0.0	0.7	0.7
P13861	KAP2_HUMAN - PRKAR2A cAMP-dependent protein kinase type II-alpha regulatory subunit	45518	0	0	0	0	0	2	0.0	0.7	0.7
Q00839	HNRPU_HUMAN - HNRNPU Heterogeneous nuclear ribonucleoprotein U	90585	0	0	0	0	0	2	0.0	0.7	0.7
Q08J23	NSUN2_HUMAN - NSUN2 tRNA (cytosine(34)-C(5))-methyltransferase	86471	0	0	0	0	0	2	0.0	0.7	0.7
P13010	XRCC5_HUMAN - XRCC5 X-ray repair cross-complementing protein 5	82705	0	0	0	0	0	2	0.0	0.7	0.7
O00154	BACH_HUMAN - ACOT7 Cytosolic acyl coenzyme A thioester hydrolase	41796	0	0	0	0	0	2	0.0	0.7	0.7
Q96QK1	VPS35_HUMAN - VPS35 Vacuolar protein sorting-associated protein 35	91707	0	0	0	0	0	2	0.0	0.7	0.7
P00491	PNPH_HUMAN - PNP Purine nucleoside phosphorylase	32118	0	0	0	0	0	2	0.0	0.7	0.7
P33991	MCM4_HUMAN - MCM4 DNA replication licensing factor MCM4	96558	0	0	0	0	0	2	0.0	0.7	0.7
P49588	SYAC_HUMAN - AARS Alanine--tRNA ligase, cytoplasmic	106810	0	0	0	0	0	2	0.0	0.7	0.7
P18206	VINC_HUMAN - VCL Vinculin	123799	0	0	0	0	0	2	0.0	0.7	0.7
P78347	GTF2I_HUMAN - GTF2I General transcription factor II-I	112416	0	0	0	0	0	2	0.0	0.7	0.7
Q15021	CND1_HUMAN - NCAPD2 Condensin complex subunit 1	157182	0	0	0	0	0	2	0.0	0.7	0.7
Q86VP6	CAND1_HUMAN - CAND1 Cullin-associated NEDD8-dissociated protein 1	136375	0	0	0	0	0	2	0.0	0.7	0.7
Q04637	IF4G1_HUMAN - EIF4G1 Eukaryotic translation initiation factor 4 gamma 1	175490	0	0	0	0	0	2	0.0	0.7	0.7
P30876	RPB2_HUMAN - POLR2B DNA-directed RNA polymerase II subunit RPB2	133896	0	0	0	0	0	2	0.0	0.7	0.7
Q92973	TNPO1_HUMAN - TNPO1 Transportin-1	102355	0	0	0	0	0	2	0.0	0.7	0.7
P46940	IQGA1_HUMAN - IQGAP1 Ras GTPase-activating-like protein IQGAP1	189251	0	0	0	0	0	2	0.0	0.7	0.7
O75116	ROCK2_HUMAN - ROCK2 Rho-associated protein kinase 2	160899	0	0	0	0	0	2	0.0	0.7	0.7
O75643	U520_HUMAN - SNRNP200 U5 small nuclear ribonucleoprotein 200 kDa helicase	244505	0	0	0	0	0	2	0.0	0.7	0.7
Q92616	GCN1L_HUMAN - GCN1L1 Translational activator GCN1	292756	0	0	0	0	0	2	0.0	0.7	0.7
Q9Y4I1	MYO5A_HUMAN - MYO5A Unconventional myosin-Va	215403	0	0	0	0	0	2	0.0	0.7	0.7
O00507	USP9Y_HUMAN - USP9Y Probable ubiquitin carboxyl-terminal hydrolase FAF-Y	291075	0	0	0	0	0	2	0.0	0.7	0.7
P12270	TPR_HUMAN - TPR Nucleoprotein TPR	267290	0	0	0	0	0	2	0.0	0.7	0.7

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
Q14204	DYHC1_HUMAN - DYNC1H1 Cytoplasmic dynein 1 heavy chain 1	532412	0	0	0	0	0	2	0.0	0.7	0.7
P04264	K2C1_HUMAN - KRT1 Keratin, type II cytoskeletal 1	66039	17	5	28	14	9	5	16.7	9.3	0.6
P06748	NPM_HUMAN - NPM1 Nucleophosmin	32575	8	0	3	0	0	6	3.7	2.0	0.5
P13647	K2C5_HUMAN - KRT5 Keratin, type II cytoskeletal 5	62378	11	2	4	9	0	0	5.7	3.0	0.5
Q9P258	RCC2_HUMAN - RCC2 Protein RCC2	56085	4	0	2	0	0	3	2.0	1.0	0.5
Q15323	K1H1_HUMAN - KRT31 Keratin, type I cuticular Ha1	47237	4	0	0	0	2	0	1.3	0.7	0.5
P35908	K22E_HUMAN - KRT2 Keratin, type II cytoskeletal 2 epidermal	65433	15	2	11	10	2	0	9.3	4.0	0.4
Q5T749	KPRP_HUMAN - KPRP Keratinocyte proline-rich protein	64136	2	0	3	2	0	0	1.7	0.7	0.4
P35527	K1C9_HUMAN - KRT9 Keratin, type I cytoskeletal 9	62064	4	2	7	3	0	0	4.3	1.0	0.2
Q86YZ3	HORN_HUMAN - HRNR Hornerin	282389	2	0	7	0	2	0	3.0	0.7	0.2
P02538	K2C6A_HUMAN - KRT6A Keratin, type II cytoskeletal 6A	60045	10	2	0	0	0	0	4.0	0.0	0.0
Q14978	NOLC1_HUMAN - NOLC1 Nucleolar and coiled-body phosphoprotein 1	73603	7	0	2	0	0	0	3.0	0.0	0.0
P19338	NUCL_HUMAN - NCL Nucleolin	76615	5	0	3	0	0	0	2.7	0.0	0.0
I3L2F9	I3L2F9_HUMAN - Uncharacterized protein	45399	0	8	0	0	0	0	2.7	0.0	0.0
Q3ZCM7	TBB8_HUMAN - TUBB8 Tubulin beta-8 chain	49776	0	8	0	0	0	0	2.7	0.0	0.0
O43790	KRT86_HUMAN - KRT86 Keratin, type II cuticular Hb6	53501	6	0	0	0	0	0	2.0	0.0	0.0
Q14533	KRT81_HUMAN - KRT81 Keratin, type II cuticular Hb1	54928	6	0	0	0	0	0	2.0	0.0	0.0
P78385	KRT83_HUMAN - KRT83 Keratin, type II cuticular Hb3	54195	6	0	0	0	0	0	2.0	0.0	0.0
P78386	KRT85_HUMAN - KRT85 Keratin, type II cuticular Hb5	55802	6	0	0	0	0	0	2.0	0.0	0.0
Q13428	TCOF_HUMAN - TCOF1 Treacle protein	152105	6	0	0	0	0	0	2.0	0.0	0.0
Q9BYQ8	KRA49_HUMAN - KRTAP4-9 Keratin-associated protein 4-9	22405	5	0	0	0	0	0	1.7	0.0	0.0
Q15185	TEBP_HUMAN - PTGES3 Prostaglandin E synthase 3	18697	4	0	0	0	0	0	1.3	0.0	0.0
Q14247	SRC8_HUMAN - CTTN Src substrate cortactin	61586	2	0	2	0	0	0	1.3	0.0	0.0
P25205	MCM3_HUMAN - MCM3 DNA replication licensing factor MCM3	90981	3	0	0	0	0	0	1.0	0.0	0.0
Q9C0C9	UBE2O_HUMAN - UBE2O Ubiquitin-conjugating enzyme E2 O	141293	3	0	0	0	0	0	1.0	0.0	0.0
Q9P2J5	SYLC_HUMAN - LARS Leucine--tRNA ligase, cytoplasmic	134466	3	0	0	0	0	0	1.0	0.0	0.0
Q9UQ35	SRRM2_HUMAN - SRRM2 Serine/arginine repetitive matrix protein 2	299616	3	0	0	0	0	0	1.0	0.0	0.0
Q09666	AHNK_HUMAN - AHNAK Neuroblast differentiation-associated protein AHNAK	629114	3	0	0	0	0	0	1.0	0.0	0.0
P15311	EZRI_HUMAN - EZR Ezrin	69413	0	0	3	0	0	0	1.0	0.0	0.0
Q92764	KRT35_HUMAN - KRT35 Keratin, type I cuticular Ha5	50361	2	0	0	0	0	0	0.7	0.0	0.0
P62269	RS18_HUMAN - RPS18 40S ribosomal protein S18	17719	2	0	0	0	0	0	0.7	0.0	0.0
Q9UNF1	MAGD2_HUMAN - MAGED2 Melanoma-associated antigen D2	64954	2	0	0	0	0	0	0.7	0.0	0.0

IPI	Protein	Mol Weight (Da)	DMSO_1	DMSO_2	DMSO_3	LAS17_1	LAS17_2	LAS17_3	DMSO Average	LAS17 Average	Fold Change
O14980	XPO1_HUMAN - XPO1 Exportin-1	123386	2	0	0	0	0	0	0.7	0.0	0.0
I3L1M0	I3L1M0_HUMAN - Uncharacterized protein	49431	2	0	0	0	0	0	0.7	0.0	0.0
Q13620	CUL4B_HUMAN - CUL4B Cullin-4B	103982	2	0	0	0	0	0	0.7	0.0	0.0
P0CG38	POTEI_HUMAN - POTEI POTE ankyrin domain family member I	121282	0	2	0	0	0	0	0.7	0.0	0.0
Q6S8J3	POTEE_HUMAN - POTEE POTE ankyrin domain family member E	121363	0	2	0	0	0	0	0.7	0.0	0.0
P0CG39	POTEJ_HUMAN - POTEJ POTE ankyrin domain family member J	117390	0	2	0	0	0	0	0.7	0.0	0.0
P0C0S8	H2A1_HUMAN - HIST1H2AM Histone H2A type 1	14091	0	0	2	0	0	0	0.7	0.0	0.0
Q9BTM1	H2AJ_HUMAN - H2AFJ Histone H2A.J	14019	0	0	2	0	0	0	0.7	0.0	0.0
Q96QV6	H2A1A_HUMAN - HIST1H2AA Histone H2A type 1-A	14233	0	0	2	0	0	0	0.7	0.0	0.0
Q8IUE6	H2A2B_HUMAN - HIST2H2AB Histone H2A type 2-B	13995	0	0	2	0	0	0	0.7	0.0	0.0
Q16777	H2A2C_HUMAN - HIST2H2AC Histone H2A type 2-C	13988	0	0	2	0	0	0	0.7	0.0	0.0
P22626	ROA2_HUMAN - HNRNPA2B1 Heterogeneous nuclear ribonucleoproteins A2/B1	37430	0	0	2	0	0	0	0.7	0.0	0.0
P54652	HSP72_HUMAN - HSPA2 Heat shock-related 70 kDa protein 2	70021	0	0	2	0	0	0	0.7	0.0	0.0
Q12873	CHD3_HUMAN - CHD3 Chromodomain-helicase-DNA-binding protein 3	226590	0	0	2	0	0	0	0.7	0.0	0.0
Q8TDI0	CHD5_HUMAN - CHD5 Chromodomain-helicase-DNA-binding protein 5	223047	0	0	2	0	0	0	0.7	0.0	0.0
Q14839	CHD4_HUMAN - CHD4 Chromodomain-helicase-DNA-binding protein 4	218003	0	0	2	0	0	0	0.7	0.0	0.0

Table 3A-1. Tryptic digests of HeLa lysates +/- LAS1, LAS6, LAS12.

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P08238	HS90B_HUMAN - HSP90AB1 Heat shock protein HSP 90-beta	23342	0	26	27.5	21
P10809	CH60_HUMAN - HSPD1 60 kDa heat shock protein, mitochondrial	61055	0	11	22	29
P07900	HS90A_HUMAN - HSP90AA1 Heat shock protein HSP 90-alpha	22110	0	20	19.5	17.5
P09211	GSTP1_HUMAN - GSTP1 Glutathione S-transferase P	28804	0	4	24	27.5
P53396	ACLY_HUMAN - ACLY ATP-citrate synthase	35077	0	11	14.5	21
P00558	PGK1_HUMAN - PGK1 Phosphoglycerate kinase 1	95338	0	21	11.5	8
P13639	EF2_HUMAN - EEF2 Elongation factor 2	44615	0	6	8	24.5
P18669	PGAM1_HUMAN - PGAM1 Phosphoglycerate mutase 1	39420	0	9	9	13.5
P04075	ALDOA_HUMAN - ALDOA Fructose-bisphosphate aldolase A	70052	0	13	4.5	6.5
Q14697	GANAB_HUMAN - GANAB Neutral alpha-glucosidase AB	83435	0	7	7	8
O75369	FLNB_HUMAN - FLNB Filamin-B	30791	0	4	7	8
P22102	PUR2_HUMAN - GART Trifunctional purine biosynthetic protein adenosine-3	120839	0	5	5	6

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P04406	G3P_HUMAN - GAPDH Glyceraldehyde-3-phosphate dehydrogenase	56651	345	144	647.5	167.5
P11498	PYC_HUMAN - PC Pyruvate carboxylase, mitochondrial	83264	131	54	238.5	95
P60709	ACTB_HUMAN - ACTB Actin, cytoplasmic 1	63147	41	88	91.5	64.5
P29401	TKT_HUMAN - TKT Transketolase	50119	8	56	70.5	101
Q13085	ACACA_HUMAN - ACACA Acetyl-CoA carboxylase 1	84660	56	17	138	40.5
P07437	TBB5_HUMAN - TUBB Tubulin beta chain	48121	35	40	57.5	52.5
P21333	FLNA_HUMAN - FLNA Filamin-A	69843	10	33	46.5	47.5
P68133	ACTS_HUMAN - ACTA1 Actin, alpha skeletal muscle	67820	36	43	51	26.5
P68104	EF1A1_HUMAN - EEF1A1 Elongation factor 1-alpha 1	97170	5	39	37.5	40.5
P07237	PDIA1_HUMAN - P4HB Protein disulfide-isomerase	67994	62	18	63	20.5
P14618	KPYM_HUMAN - PKM Pyruvate kinase isozymes M1/M2	56560	7	33	28.5	38
P06733	ENOA_HUMAN - ENO1 Alpha-enolase	269765	4	38	24.5	37
P04264	K2C1_HUMAN - KRT1 Keratin, type II cytoskeletal 1	106874	17	16	67	11
P49327	FAS_HUMAN - FASN Fatty acid synthase	191613	2	14	21	36.5
P30101	PDIA3_HUMAN - PDIA3 Protein disulfide-isomerase A3	47518	26	7	42.5	19.5
P11021	GRP78_HUMAN - HSPA5 78 kDa glucose-regulated protein	107767	9	6	43.5	14.5
P31327	CPSM_HUMAN - CPS1 Carbamoyl-phosphate synthase	24423	5	10	16.5	37.5
P35579	MYH9_HUMAN - MYH9 Myosin-9	104854	6	14	21.5	24
P13645	K1C10_HUMAN - KRT10 Keratin, type I cytoskeletal 10	60344	3	17	14.5	8.5
P35908	K22E_HUMAN - KRT2 Keratin, type II cytoskeletal 2 epidermal	101559	8	16	11	6.5
P30041	PRDX6_HUMAN - PRDX6 Peroxiredoxin-6	469093	4	6	20	5.5
P09429	HMGB1_HUMAN - HMGB1 High mobility group protein B1	59144	3	5	14	12.5
P62937	PPIA_HUMAN - PPIA Peptidyl-prolyl cis-trans isomerase A	532412	12	6	7	5.5
Q01813	K6PP_HUMAN - PFKP 6-phosphofructokinase type C	278162	0	14	5	58.5
Q15084	PDIA6_HUMAN - PDIA6 Protein disulfide-isomerase A6	36053	0	2	23	9
Q00610	CLH1_HUMAN - CLTC Clathrin heavy chain 1	129634	0	2	10	15
Q06830	PRDX1_HUMAN - PRDX1 Peroxiredoxin-1	56782	0	3	7	15
P26038	MOES_HUMAN - MSN Moesin	18012	0	3	11.5	9.5
P60174	TPIS_HUMAN - TPI1 Triosephosphate isomerase	25035	0	10	5	4.5
O43707	ACTN4_HUMAN - ACTN4 Alpha-actinin-4	41737	0	7	7	5.5
P63244	GBLP_HUMAN - GNB2L1 Guanine nucleotide-binding protein subunit beta-2-like 1	57937	0	2	8	9
Q9Y490	TLN1_HUMAN - TLN1 Talin-1	57116	0	4	7	8
P17174	AATC_HUMAN - GOT1 Aspartate aminotransferase, cytoplasmic	67878	0	4	11.5	3.5
Q14974	IMB1_HUMAN - KPNB1 Importin subunit beta-1	24894	0	5	6	7
Q99832	TCPH_HUMAN - CCT7 T-complex protein 1 subunit eta	49671	0	0	5	12.5

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P12956	XRCC6_HUMAN - XRCC6 X-ray repair cross-complementing protein 6	47169	0	5	4	6.5
Q92616	GCN1L_HUMAN - GCN1L1 Translational activator GCN1	26210	0	0	4	10.5
O43175	SERA_HUMAN - PHGDH D-3-phosphoglycerate dehydrogenase	65433	0	2	5.5	6.5
P23528	COF1_HUMAN - CFL1 Cofilin-1	265551	0	0	8.5	5.5
Q71U36	TBA1A_HUMAN - TUBA1A Tubulin alpha-1A chain	58827	28	48	16.5	83
P68366	TBA4A_HUMAN - TUBA4A Tubulin alpha-4A chain	72333	28	0	50	74.5
P68371	TBB4B_HUMAN - TUBB4B Tubulin beta-4B chain	50141	35	0	56	47.5
Q9BUF5	TBB6_HUMAN - TUBB6 Tubulin beta-6 chain	42051	12	10	22	5
Q9HB71	CYBP_HUMAN - CACYBP Calcyclin-binding protein	72692	6	6	18.5	4
Q96RQ3	MCCA_HUMAN - MCCC1 Methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial	66039	5	4	18	3
P12277	KCRB_HUMAN - CKB Creatine kinase B-type	164939	3	0	14	7
P31040	DHSA_HUMAN - SDHA Succinate dehydrogenase	273424	2	2	6.5	9.5
P78527	PRKDC_HUMAN - PRKDC DNA-dependent protein kinase catalytic subunit	226530	0	2	4.5	18.5
P38117	ETFB_HUMAN - ETFB Electron transfer flavoprotein subunit beta	280737	0	9	14.5	1.5
P08107	HSP71_HUMAN - HSPA1B Heat shock 70 kDa protein 1A/1B	18502	0	7	2.5	11.5
P62826	RAN_HUMAN - RAN GTP-binding nuclear protein Ran	85596	0	2	13.5	5.5
O76003	GLRX3_HUMAN - GLRX3 Glutaredoxin-3	46248	0	5	15.5	0
P06576	ATPB_HUMAN - ATP5B ATP synthase subunit beta, mitochondrial	53140	0	2	14	4
Q562R1	ACTBL_HUMAN - ACTBL2 Beta-actin-like protein 2	36639	0	10	9	0
P11586	C1TC_HUMAN - MTHFD1 C-1-tetrahydrofolate synthase, cytoplasmic	59367	0	10	2	6
P26639	SYTC_HUMAN - TARS Threonine--tRNA ligase, cytoplasmic	22391	0	2	3.5	12
P07195	LDHB_HUMAN - LDHB L-lactate dehydrogenase B chain	26560	0	6	1.5	8
P08195	4F2_HUMAN - SLC3A2 4F2 cell-surface antigen heavy chain	41351	0	3	4.5	7.5
P00505	AATM_HUMAN - GOT2 Aspartate aminotransferase, mitochondrial	41924	0	5	3.5	6
Q16881	TRXR1_HUMAN - TXNRD1 Thioredoxin reductase 1, cytoplasmic	54530	0	0	8	5.5
P62258	1433E_HUMAN - YWHAE 14-3-3 protein epsilon	70756	0	4	0	9
Q9P2J5	SYLC_HUMAN - LARS Leucine--tRNA ligase, cytoplasmic	49542	0	0	2.5	8.5
P11142	HSP7C_HUMAN - HSPA8 Heat shock cognate 71 kDa protein	21892	0	0	5	5.5
P55060	XPO2_HUMAN - CSE1L Exportin-2	22783	0	0	2	8.5
P37802	TAGL2_HUMAN - TAGLN2 Transgelin-2	60534	0	3	5	2
P42765	THIM_HUMAN - ACAA2 3-ketoacyl-CoA thiolase, mitochondrial	59671	0	2	2	6
P17812	PYRG1_HUMAN - CTPS1 CTP synthase 1	117849	0	0	5	5
P00338	LDHA_HUMAN - LDHA L-lactate dehydrogenase A chain	35937	0	0	2.5	6.5
P42704	LPPRC_HUMAN - LRPPRC Leucine-rich PPR motif-containing protein, mitochondrial	70898	0	0	4.5	4
P13797	PLST_HUMAN - PLS3 Plastin-3	85182	0	0	5.5	2.5

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P27708	PYR1_HUMAN - CAD CAD protein	46441	0	0	3	4
P07814	SYEP_HUMAN - EPRS Bifunctional glutamate/proline--tRNA ligase	66690	0	0	2	5
Q15149	PLEC_HUMAN - PLEC Plectin	81308	0	0	2	4
Q14315	FLNC_HUMAN - FLNC Filamin-C	57488	0	0	2	4
P62736	ACTA_HUMAN - ACTA2 Actin, aortic smooth muscle	157904	36	0	9	26.5
P05165	PCCA_HUMAN - PCCA Propionyl-CoA carboxylase alpha chain, mitochondrial	140958	14	0	18	9
P35241	RADI_HUMAN - RDX Radixin	292756	5	3	11.5	5.5
Q13509	TBB3_HUMAN - TUBB3 Tubulin beta-3 chain	144734	0	0	0	35.5
Q05639	EF1A2_HUMAN - EEF1A2 Elongation factor 1-alpha 2	90585	0	0	12	16.5
P13667	PDIA4_HUMAN - PDIA4 Protein disulfide-isomerase A4	54235	0	0	16.5	0
P06744	G6PI_HUMAN - GPI Glucose-6-phosphate isomerase	115935	0	8	3	4
P52895	AK1C2_HUMAN - AKR1C2 Aldo-keto reductase family 1 member C2	242981	0	9	1.5	3
P26641	EF1G_HUMAN - EEF1G Elongation factor 1-gamma	629114	0	7	2	4
Q14204	DYHC1_HUMAN - DYNC1H1 Cytoplasmic dynein 1 heavy chain 1	27844	0	2	2.5	7.5
P52209	6PGD_HUMAN - PGD 6-phosphogluconate dehydrogenase, decarboxylating	49924	0	4	1	7
P08237	K6PF_HUMAN - PFKM 6-phosphofructokinase, muscle type	50136	0	2	1.5	8.5
P04792	HSPB1_HUMAN - HSPB1 Heat shock protein beta-1	49831	0	4	5	2
P22314	UBA1_HUMAN - UBA1 Ubiquitin-like modifier-activating enzyme 1	42644	0	0	3	7.5
P17858	K6PL_HUMAN - PFKL 6-phosphofructokinase, liver type	68564	0	0	1.5	9
P17987	TCPA_HUMAN - TCP1 T-complex protein 1 subunit alpha	49857	0	2	3	4.5
P49411	EFTU_HUMAN - TUFM Elongation factor Tu, mitochondrial	80473	0	0	2.5	7
Q08211	DHX9_HUMAN - DHX9 ATP-dependent RNA helicase A	80059	0	2	1	6.5
Q99714	HCD2_HUMAN - HSD17B10 3-hydroxyacyl-CoA dehydrogenase type-2	29174	0	4	1.5	4
P19338	NUCL_HUMAN - NCL Nucleolin	46154	0	0	8	1.5
P46782	RS5_HUMAN - RPS5 40S ribosomal protein S5	13944	0	0	0	9.5
P40227	TCPZ_HUMAN - CCT6A T-complex protein 1 subunit zeta	13952	0	2	0	7
P48643	TCPE_HUMAN - CCT5 T-complex protein 1 subunit epsilon	13989	0	0	4	4.5
P61978	HNRPK_HUMAN - HNRNPK Heterogeneous nuclear ribonucleoprotein K	13920	0	0	1	7.5
P33992	MCM5_HUMAN - MCM5 DNA replication licensing factor MCM5	13920	0	5	0	3.5
Q9BWD1	THIC_HUMAN - ACAT2 Acetyl-CoA acetyltransferase, cytosolic	58024	0	2	2.5	3.5
P49368	TCPG_HUMAN - CCT3 T-complex protein 1 subunit gamma	15529	0	2	1.5	4.5
P78371	TCPB_HUMAN - CCT2 T-complex protein 1 subunit beta	17298	0	4	1	3
P80723	BASP1_HUMAN - BASP1 Brain acid soluble protein 1	50976	0	3	5	0
P62805	H4_HUMAN - HIST4H4 Histone H4	38604	0	6	0	2
P14625	ENPL_HUMAN - HSP90B1 Endoplasmin	64616	0	6	0	2

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P04179	SODM_HUMAN - SOD2 Superoxide dismutase	39724	0	0	7.5	0
P21266	GSTM3_HUMAN - GSTM3 Glutathione S-transferase Mu 3	36689	0	0	3.5	3.5
P53597	SUCA_HUMAN - SUCLG1 Succinyl-CoA ligase	55805	0	0	3.5	3.5
Q9BR76	COR1B_HUMAN - CORO1B Coronin-1B	56158	0	0	4	2.5
P60842	IF4A1_HUMAN - EIF4A1 Eukaryotic initiation factor 4A-I	29226	0	2	2	2.5
P04040	CATA_HUMAN - CAT Catalase	47716	0	0	5.5	1
P34932	HSP74_HUMAN - HSPA4 Heat shock 70 kDa protein 4	32575	0	0	0	6.5
P57053	H2BFS_HUMAN - H2BFS Histone H2B type F-S	31284	0	0	1.5	4.5
Q99880	H2B1L_HUMAN - HIST1H2BL Histone H2B type 1-L	25712	0	0	1.5	4.5
Q99879	H2B1M_HUMAN - HIST1H2BM Histone H2B type 1-M	110417	0	0	1.5	4.5
Q5QNW6	H2B2F_HUMAN - HIST2H2BF Histone H2B type 2-F	70811	0	0	1.5	4.5
Q16778	H2B2E_HUMAN - HIST2H2BE Histone H2B type 2-E	134466	0	0	1.5	4.5
P20810	ICAL_HUMAN - CAST Calpastatin	89322	0	0	5	1
Q86VP6	CAND1_HUMAN - CAND1 Cullin-associated NEDD8-dissociated protein 1	82705	0	0	2.5	3.5
P46940	IQGA1_HUMAN - IQGAP1 Ras GTPase-activating-like protein IQGAP1	69413	0	0	0	6
P30876	RPB2_HUMAN - POLR2B DNA-directed RNA polymerase II subunit RPB2	36250	0	0	2	3
P50395	GDIB_HUMAN - GDI2 Rab GDP dissociation inhibitor beta	123386	0	0	2	3
P35580	MYH10_HUMAN - MYH10 Myosin-10	170590	0	0	2	3
P07737	PROF1_HUMAN - PFN1 Profilin-1	76573	0	0	0	4.5
O00429	DNM1L_HUMAN - DNM1L Dynamin-1-like protein	87302	0	0	0	4.5
P25705	ATPA_HUMAN - ATP5A1 ATP synthase subunit alpha, mitochondrial	85018	0	0	0	4
P23284	PPIB_HUMAN - PPIB Peptidyl-prolyl cis-trans isomerase B	90070	0	0	4	0
P35527	K1C9_HUMAN - KRT9 Keratin, type I cytoskeletal 9	53652	3	0	31.5	0
P12821	ACE_HUMAN - ACE Angiotensin-converting enzyme	24604	0	0	102.5	0
Q6S8J3	POTEE_HUMAN - POTEE POTE ankyrin domain family member E	59257	0	48	0	0
Q09666	AHNK_HUMAN - AHNAK Neuroblast differentiation-associated protein AHNAK	102355	0	0	14.5	2
Q00839	HNRPU_HUMAN - HNRNPU Heterogeneous nuclear ribonucleoprotein U	59756	0	3	1.5	5.5
P42330	AK1C3_HUMAN - AKR1C3 Aldo-keto reductase family 1 member C3	136375	0	9	0	1
P13929	ENOB_HUMAN - ENO3 Beta-enolase	78806	0	0	0	10
P0C0S8	H2A1_HUMAN - HIST1H2AM Histone H2A type 1	103058	0	9	0	0
Q9BTM1	H2AJ_HUMAN - H2AFJ Histone H2A.J	132600	0	9	0	0
Q16777	H2A2C_HUMAN - HIST2H2AC Histone H2A type 2-C	98064	0	9	0	0
Q04828	AK1C1_HUMAN - AKR1C1 Aldo-keto reductase family 1 member C1	133896	0	9	0	0
P32119	PRDX2_HUMAN - PRDX2 Peroxiredoxin-2	531796	0	3	4	1.5
P15311	EZRI_HUMAN - EZR Ezrin	291020	0	3	0	5.5

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P54577	SYYC_HUMAN - YARS Tyrosine--tRNA ligase, cytoplasmic	36735	0	3	3	2
O15067	PUR4_HUMAN - PFAS Phosphoribosylformylglycinamidine synthase	26923	0	3	3	2
P13010	XRCC5_HUMAN - XRCC5 X-ray repair cross-complementing protein 5	26478	0	0	1	7
Q16658	FSCN1_HUMAN - FSCN1 Fascin	76615	0	2	1	4
P31939	PUR9_HUMAN - ATIC Bifunctional purine biosynthesis protein PURH	88550	0	3	1	3
P07355	ANXA2_HUMAN - ANXA2 Annexin A2	22693	0	3	1	2.5
P08758	ANXA5_HUMAN - ANXA5 Annexin A5	42003	0	2	2	2
P33993	MCM7_HUMAN - MCM7 DNA replication licensing factor MCM7	37432	0	0	2.5	3.5
P23526	SAHH_HUMAN - AHCY Adenosylhomocysteinase	42009	0	2	0	4
P54819	KAD2_HUMAN - AK2 Adenylate kinase 2, mitochondrial	15054	0	0	5	1
P54652	HSP72_HUMAN - HSPA2 Heat shock-related 70 kDa protein 2	14121	0	5	0	1
P50454	SERPH_HUMAN - SERPINH1 Serpin H1	22876	0	0	3	2.5
Q96FW1	OTUB1_HUMAN - OTUB1 Ubiquitin thioesterase OTUB1	50433	0	2	1.5	2
O14980	XPO1_HUMAN - XPO1 Exportin-1	11367	0	0	2.5	3
Q06210	GFPT1_HUMAN - GFPT1 Glucosamine--fructose-6-phosphate aminotransferase	20778	0	0	1.5	4
P54886	P5CS_HUMAN - ALDH18A1 Delta-1-pyrroline-5-carboxylate synthase	24205	0	0	3	2
P63104	1433Z_HUMAN - YWHAZ 14-3-3 protein zeta/delta	50663	0	3	0	2
P31153	METK2_HUMAN - MAT2A S-adenosylmethionine synthase isoform type-2	27764	0	3	0	2
P12268	IMDH2_HUMAN - IMPDH2 Inosine-5'-monophosphate dehydrogenase 2	50470	0	0	1	3.5
P09488	GSTM1_HUMAN - GSTM1 Glutathione S-transferase Mu 1	41943	0	0	2.5	2
P23921	RIR1_HUMAN - RRM1 Ribonucleoside-diphosphate reductase large subunit	47079	0	0	2	2.5
P27635	RL10_HUMAN - RPL10 60S ribosomal protein L10	26688	0	0	3.5	1
Q9BTW9	TBCD_HUMAN - TBCD Tubulin-specific chaperone D	27745	0	0	1.5	3
P22234	PUR6_HUMAN - PAICS Multifunctional protein ADE2	81877	0	2	0	2.5
P08559	ODPA_HUMAN - PDHA1 Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	59751	0	0	0	4.5
Q9Y678	COPG1_HUMAN - COPG1 Coatomer subunit gamma-1	94331	0	0	0	4.5
P33316	DUT_HUMAN - DUT Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial	47656	0	2	2.5	0
P39023	RL3_HUMAN - RPL3 60S ribosomal protein L3	48043	0	0	4.5	0
Q8WUM4	PDC6I_HUMAN - PDCD6IP Programmed cell death 6-interacting protein	44761	0	0	0	4.5
P24752	THIL_HUMAN - ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	92469	0	0	0	4.5
O00763	ACACB_HUMAN - ACACB Acetyl-CoA carboxylase 2	68304	0	0	4.5	0
Q9Y281	COF2_HUMAN - CFL2 Cofilin-2	43296	0	0	4.5	0
Q92820	GGH_HUMAN - GGH Gamma-glutamyl hydrolase	36853	0	0	4.5	0
P55072	TERA_HUMAN - VCP Transitional endoplasmic reticulum ATPase	51805	0	0	2	2
P08670	VIME_HUMAN - VIM Vimentin	189251	0	0	3	1

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P11413	G6PD_HUMAN - G6PD Glucose-6-phosphate 1-dehydrogenase	38747	0	0	1	3
P62241	RS8_HUMAN - RPS8 40S ribosomal protein S8	34225	0	0	0	4
P27348	1433T_HUMAN - YWHAQ 14-3-3 protein theta	90981	0	0	1	3
Q9H7Z7	PGES2_HUMAN - PTGES2 Prostaglandin E synthase 2	59621	0	0	0	4
P38606	VATA_HUMAN - ATP6V1A V-type proton ATPase catalytic subunit A	50680	0	0	0	4
Q8N257	H2B3B_HUMAN - HIST3H2BB Histone H2B type 3-B	49130	0	0	1.5	2.5
Q16576	RBBP7_HUMAN - RBBP7 Histone-binding protein RBBP7	48991	0	0	1.5	2.5
Q7L7L0	H2A3_HUMAN - HIST3H2A Histone H2A type 3	140466	0	0	0	3.5
Q09028	RBBP4_HUMAN - RBBP4 Histone-binding protein RBBP4	56085	0	0	0	3.5
P53618	COPB_HUMAN - COPB1 Coatomer subunit beta	96865	0	0	0	3.5
O00410	IPO5_HUMAN - IPO5 Importin-5	97718	0	0	0	3.5
P14868	SYDC_HUMAN - DARS Aspartate--tRNA ligase, cytoplasmic	100200	0	0	1	2.5
P47897	SYQ_HUMAN - QARS Glutamine--tRNA ligase	228997	0	0	2.5	1
P62277	RS13_HUMAN - RPS13 40S ribosomal protein S13	91707	0	0	3.5	0
P14174	MIF_HUMAN - MIF Macrophage migration inhibitory factor	107142	0	0	3.5	0
P53999	TCP4_HUMAN - SUB1 Activated RNA polymerase II transcriptional coactivator p15	123630	0	0	3.5	0
P50990	TCPQ_HUMAN - CCT8 T-complex protein 1 subunit theta	121682	0	0	0	3
Q9P258	RCC2_HUMAN - RCC2 Protein RCC2	13908	0	0	2	1
P13647	K2C5_HUMAN - KRT5 Keratin, type II cytoskeletal 5	43661	0	0	3	0
P30153	2AAA_HUMAN - PPP2R1A Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	82286	0	0	0	3
P53634	CATC_HUMAN - CTSC Dipeptidyl peptidase 1	47820	0	0	3	0
P31946	1433B_HUMAN - YWHAB 14-3-3 protein beta/alpha	57136	0	0	0	2.5
P31947	1433S_HUMAN - SFN 14-3-3 protein sigma	33361	0	0	0	2.5
P21980	TGM2_HUMAN - TGM2 Protein-glutamine gamma-glutamyltransferase 2	69285	0	0	0	2.5
Q03013	GSTM4_HUMAN - GSTM4 Glutathione S-transferase Mu 4	74140	0	0	2.5	0
P41250	SYG_HUMAN - GARS Glycine--tRNA ligase	24182	0	0	0	2
P36776	LONM_HUMAN - LONP1 Lon protease homolog, mitochondrial	55756	0	0	0	2
Q12931	TRAP1_HUMAN - TRAP1 Heat shock protein 75 kDa, mitochondrial	63945	0	0	0	2
P13804	ETFA_HUMAN - ETFA Electron transfer flavoprotein subunit alpha, mitochondrial	73461	0	0	0	2
P05091	ALDH2_HUMAN - ALDH2 Aldehyde dehydrogenase, mitochondrial	73603	0	0	0	2
Q99497	PARK7_HUMAN - PARK7 Protein DJ-1	87799	0	0	2	0
O75083	WDR1_HUMAN - WDR1 WD repeat-containing protein 1	92889	0	0	2	0
P54727	RD23B_HUMAN - RAD23B UV excision repair protein RAD23 homolog B	119516	0	0	2	0
P11940	PABP1_HUMAN - PABPC1 Polyadenylate-binding protein 1	70021	0	0	2	0
P12270	TPR_HUMAN - TPR Nucleoprotein TPR	26563	0	0	2	0

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P42166	LAP2A_HUMAN - TMPO Lamina-associated polypeptide 2, isoform alpha	35503	0	0	2	0
P42167	LAP2B_HUMAN - TMPO Lamina-associated polypeptide 2, isoforms beta/gamma	58487	0	0	2	0
P30837	AL1B1_HUMAN - ALDH1B1 Aldehyde dehydrogenase X, mitochondrial	24722	0	0	2	0
P00390	GSHR_HUMAN - GSR Glutathione reductase, mitochondrial	27402	0	0	2	0
Q9BVA1	TBB2B_HUMAN - TUBB2B Tubulin beta-2B chain	23743	35	0	0	22
P02538	K2C6A_HUMAN - KRT6A Keratin, type II cytoskeletal 6A	32193	2	0	3	0
P51858	HDGF_HUMAN - HDGF Hepatoma-derived growth factor	31362	10	0	3	0
Q04837	SSBP_HUMAN - SSBP1 Single-stranded DNA-binding protein, mitochondrial	28993	3	0	2	0
O60361	NDK8_HUMAN - NME2P1 Putative nucleoside diphosphate kinase	72933	0	3	1.5	1.5
P22392	NDKB_HUMAN - NME2 Nucleoside diphosphate kinase B	22591	0	3	1.5	1.5
Q02218	ODO1_HUMAN - OGDH 2-oxoglutarate dehydrogenase, mitochondrial	17222	0	0	2	3
Q92973	TNPO1_HUMAN - TNPO1 Transportin-1	30868	0	0	1.5	3
P12814	ACTN1_HUMAN - ACTN1 Alpha-actinin-1	62378	0	2	1	1.5
P61160	ARP2_HUMAN - ACTR2 Actin-related protein 2	46109	0	3	0	1.5
P55809	SCOT1_HUMAN - OXCT1 Succinyl-CoA:3-ketoacid coenzyme A transferase 1, mitochondrial	191480	0	0	1.5	2.5
Q13263	TIF1B_HUMAN - TRIM28 Transcription intermediary factor 1-beta	62064	0	0	3	1
P18124	RL7_HUMAN - RPL7 60S ribosomal protein L7	60045	0	0	2.5	1
P06748	NPM_HUMAN - NPM1 Nucleophosmin	34294	0	0	1	2.5
Q92598	HS105_HUMAN - HSPH1 Heat shock protein 105 kDa	65309	0	2	0	1.5
P09960	LKHA4_HUMAN - LTA4H Leukotriene A-4 hydrolase	46932	0	2	0	1.5
P11766	ADHX_HUMAN - ADH5 Alcohol dehydrogenase class-3	28082	0	0	2	1
P50570	DYN2_HUMAN - DNM2 Dynamin-2	17040	0	0	1	2
P23396	RS3_HUMAN - RPS3 40S ribosomal protein S3	27774	0	0	0	3
P25205	MCM3_HUMAN - MCM3 DNA replication licensing factor MCM3	47371	0	0	0	3
B0V043	B0V043_HUMAN - VARS Valyl-tRNA synthetase	31350	0	0	1.5	1.5
P32322	P5CR1_HUMAN - PYCR1 Pyrroline-5-carboxylate reductase 1, mitochondrial	47366	0	2	0	1
Q96HC4	PDLI5_HUMAN - PDLIM5 PDZ and LIM domain protein 5	38714	0	0	1.5	1.5
Q14978	NOLC1_HUMAN - NOLC1 Nucleolar and coiled-body phosphoprotein 1	37187	0	0	1.5	1.5
Q14566	MCM6_HUMAN - MCM6 DNA replication licensing factor MCM6	33637	0	2	0	1
P40926	MDHM_HUMAN - MDH2 Malate dehydrogenase, mitochondrial	16930	0	2	1	0
P49419	AL7A1_HUMAN - ALDH7A1 Alpha-aminoadipic semialdehyde dehydrogenase	52352	0	2	1	0
Q9UL46	PSME2_HUMAN - PSME2 Proteasome activator complex subunit 2	31387	0	0	3	0
P84085	ARF5_HUMAN - ARF5 ADP-ribosylation factor 5	27566	0	3	0	0
P15531	NDKA_HUMAN - NME1 Nucleoside diphosphate kinase A	83166	0	3	0	0
P16152	CBR1_HUMAN - CBR1 Carbonyl reductase	46659	0	3	0	0

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
Q06323	PSME1_HUMAN - PSME1 Proteasome activator complex subunit 1	53165	0	3	0	0
Q04760	LGUL_HUMAN - GLO1 Lactoylglutathione lyase	106489	0	0	1.5	1
Q7L1Q6	BZW1_HUMAN - BZW1 Basic leucine zipper and W2 domain-containing protein 1	38497	0	0	1	1.5
Q02790	FKBP4_HUMAN - FKBP4 Peptidyl-prolyl cis-trans isomerase FKBP4	43482	0	0	1.5	1
Q16181	SEPT7_HUMAN - SEPT7 Septin-7	38580	0	0	0	2.5
Q96QK1	VPS35_HUMAN - VPS35 Vacuolar protein sorting-associated protein 35	40746	0	0	0	2.5
P02545	LMNA_HUMAN - LMNA Prelamin-A/C	57924	0	0	1	1.5
O94925	GLSK_HUMAN - GLS Glutaminase kidney isoform, mitochondrial	63111	0	0	1.5	1
O95373	IPO7_HUMAN - IPO7 Importin-7	77329	0	0	1.5	1
Q07021	C1QBP_HUMAN - C1QBP Complement component 1 Q subcomponent-binding protein, mitochondrial	50909	0	0	2.5	0
P09651	ROA1_HUMAN - HNRNPA1 Heterogeneous nuclear ribonucleoprotein A1	101997	0	0	1	1
Q32P51	RA1L2_HUMAN - HNRNPA1L2 Heterogeneous nuclear ribonucleoprotein A1-like 2	42742	0	0	1	1
O00148	DX39A_HUMAN - DDX39A ATP-dependent RNA helicase DDX39A	31122	0	0	1	1
Q13838	DX39B_HUMAN - DDX39B Spliceosome RNA helicase DDX39B	106810	0	0	1	1
Q13200	PSMD2_HUMAN - PSMD2 26S proteasome non-ATPase regulatory subunit 2	48551	0	0	1	1
O00159	MYO1C_HUMAN - MYO1C Unconventional myosin-Ic	30540	0	0	1	1
E7EMT2	E7EMT2_HUMAN - Uncharacterized protein	100071	0	0	1	1
P15170	ERF3A_HUMAN - GSPT1 Eukaryotic peptide chain release factor GTP-binding subunit ERF3A	55103	0	0	1	1
Q14192	FHL2_HUMAN - FHL2 Four and a half LIM domains protein 2	55364	0	0	2	0
P30040	ERP29_HUMAN - ERP29 Endoplasmic reticulum resident protein 29	80110	0	0	2	0
P46781	RS9_HUMAN - RPS9 40S ribosomal protein S9	53489	0	0	2	0
P15559	NQO1_HUMAN - NQO1 NAD(P)H dehydrogenase	40283	0	0	2	0
Q02952	AKA12_HUMAN - AKAP12 A-kinase anchor protein 12	82432	0	0	2	0
Q9BUL8	PDC10_HUMAN - PDCD10 Programmed cell death protein 10	57221	0	2	0	0
P36952	SPB5_HUMAN - SERPINB5 Serpin B5	67568	0	2	0	0
Q14240	IF4A2_HUMAN - EIF4A2 Eukaryotic initiation factor 4A-II	69603	0	2	0	0
P61764	STXB1_HUMAN - STXBP1 Syntaxin-binding protein 1	24519	0	2	0	0
P08133	ANXA6_HUMAN - ANXA6 Annexin A6	54232	0	2	0	0
P55884	EIF3B_HUMAN - EIF3B Eukaryotic translation initiation factor 3 subunit B	61398	0	2	0	0
Q99733	NP1L4_HUMAN - NAP1L4 Nucleosome assembly protein 1-like 4	60587	0	2	0	0
P17066	HSP76_HUMAN - HSPA6 Heat shock 70 kDa protein 6	112535	0	2	0	0
P48741	HSP77_HUMAN - HSPA7 Putative heat shock 70 kDa protein 7	94512	0	2	0	0
Q16836	HCDH_HUMAN - HADH Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	113289	0	0	0	1.5
P61158	ARP3_HUMAN - ACTR3 Actin-related protein 3	96023	0	0	0	1.5

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P62140	PP1B_HUMAN - PPP1CB Serine/threonine-protein phosphatase PP1-beta catalytic subunit	96558	0	0	0	1.5
Q16719	KYNU_HUMAN - KYNU Kynureninase	83550	0	0	0	1.5
O75874	IDHC_HUMAN - IDH1 Isocitrate dehydrogenase	49229	0	0	0	1.5
P07686	HEXB_HUMAN - HEXB Beta-hexosaminidase subunit beta	45672	0	0	0	1.5
P48735	IDHP_HUMAN - IDH2 Isocitrate dehydrogenase	69367	0	0	0	1.5
Q7KZF4	SND1_HUMAN - SND1 Staphylococcal nuclease domain-containing protein 1	62943	0	0	0	1.5
P29692	EF1D_HUMAN - EEF1D Elongation factor 1-delta	76931	0	0	0	1.5
P49588	SYAC_HUMAN - AARS Alanine--tRNA ligase, cytoplasmic	90933	0	0	0	1.5
P26599	PTBP1_HUMAN - PTBP1 Polypyrimidine tract-binding protein 1	103276	0	0	0	1.5
O60506	HNRPQ_HUMAN - SYNCRIP Heterogeneous nuclear ribonucleoprotein Q	115963	0	0	0	1.5
O95757	HS74L_HUMAN - HSPA4L Heat shock 70 kDa protein 4L	103505	0	0	0	1.5
Q96T76	MMS19_HUMAN - MMS19 MMS19 nucleotide excision repair protein homolog	51712	0	0	0	1.5
P31943	HNRH1_HUMAN - HNRNPH1 Heterogeneous nuclear ribonucleoprotein H	88068	0	0	0	1.5
P52597	HNRPF_HUMAN - HNRNPF Heterogeneous nuclear ribonucleoprotein F	97746	0	0	0	1.5
O94808	GFPT2_HUMAN - GFPT2 Glucosamine--fructose-6-phosphate aminotransferase	109685	0	0	0	1.5
Q9H8V3	ECT2_HUMAN - ECT2 Protein ECT2	144498	0	0	0	1.5
Q15365	PCBP1_HUMAN - PCBP1 Poly(rC)-binding protein 1	193539	0	0	0	1.5
P26583	HMGB2_HUMAN - HMGB2 High mobility group protein B2	205556	0	0	0	1.5
P22695	QCR2_HUMAN - UQCRC2 Cytochrome b-c1 complex subunit 2, mitochondrial	171491	0	0	0	1.5
P40429	RL13A_HUMAN - RPL13A 60S ribosomal protein L13a	780624	0	0	0	1.5
O43252	PAPS1_HUMAN - PAPSS1 Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase 1	14865	0	0	0	1.5
P10155	RO60_HUMAN - TROVE2 60 kDa SS-A/Ro ribonucleoprotein	45200	0	0	0	1.5
Q92599	SEPT8_HUMAN - SEPT8 Septin-8	35080	0	0	0	1.5
Q96AY3	FKB10_HUMAN - FKBP10 Peptidyl-prolyl cis-trans isomerase FKBP10	37498	0	0	0	1.5
P05161	ISG15_HUMAN - ISG15 Ubiquitin-like protein ISG15	24034	0	0	1.5	0
P02533	K1C14_HUMAN - KRT14 Keratin, type I cytoskeletal 14	38325	0	0	1.5	0
P54725	RD23A_HUMAN - RAD23A UV excision repair protein RAD23 homolog A	48443	0	0	1.5	0
Q93052	LPP_HUMAN - LPP Lipoma-preferred partner	23577	0	0	1.5	0
Q9NTJ3	SMC4_HUMAN - SMC4 Structural maintenance of chromosomes protein 4	70833	0	0	1.5	0
Q9UQ35	SRRM2_HUMAN - SRRM2 Serine/arginine repetitive matrix protein 2	25734	0	0	1.5	0
Q9H773	DCTP1_HUMAN - DCTPP1 dCTP pyrophosphatase 1	60671	0	0	1.5	0
O00299	CLIC1_HUMAN - CLIC1 Chloride intracellular channel protein 1	56381	0	0	1.5	0
Q15942	ZYX_HUMAN - ZYX Zyxin	55756	0	0	1.5	0
Q9NVI1	FANCI_HUMAN - FANCI Fanconi anemia group I protein	64245	0	0	1.5	0

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P62841	RS15_HUMAN - RPS15 40S ribosomal protein S15	75379	0	0	0	1
P47756	CAPZB_HUMAN - CAPZB F-actin-capping protein subunit beta	53249	0	0	0	1
P43686	PRS6B_HUMAN - PSMC4 26S protease regulatory subunit 6B	79275	0	0	0	1
P04083	ANXA1_HUMAN - ANXA1 Annexin A1	87680	0	0	0	1
Q96C36	P5CR2_HUMAN - PYCR2 Pyrroline-5-carboxylate reductase 2	101272	0	0	0	1
P60660	MYL6_HUMAN - MYL6 Myosin light polypeptide 6	107546	0	0	0	1
P30084	ECHM_HUMAN - ECHS1 Enoyl-CoA hydratase, mitochondrial	244505	0	0	0	1
P78417	GSTO1_HUMAN - GSTO1 Glutathione S-transferase omega-1	152785	0	0	0	1
P23381	SYWC_HUMAN - WARS Tryptophan--tRNA ligase, cytoplasmic	14091	0	0	0	1
P40937	RFC5_HUMAN - RFC5 Replication factor C subunit 5	14019	0	0	0	1
Q6NVY1	HIBCH_HUMAN - HIBCH 3-hydroxyisobutyryl-CoA hydrolase, mitochondrial	13988	0	0	0	1
Q15366	PCBP2_HUMAN - PCBP2 Poly(rC)-binding protein 2	20530	0	0	0	1
P55039	DRG2_HUMAN - DRG2 Developmentally-regulated GTP-binding protein 2	17149	0	0	0	1
P50991	TCPD_HUMAN - CCT4 T-complex protein 1 subunit delta	30375	0	0	0	1
P30740	ILEU_HUMAN - SERPINB1 Leukocyte elastase inhibitor	24702	0	0	0	1
Q15645	PCH2_HUMAN - TRIP13 Pachytene checkpoint protein 2 homolog	121363	0	0	0	1
Q13162	PRDX4_HUMAN - PRDX4 Peroxiredoxin-4	28723	0	0	0	1
P35221	CTNA1_HUMAN - CTNNA1 Catenin alpha-1	42100	0	0	0	1
Q13547	HDAC1_HUMAN - HDAC1 Histone deacetylase 1	46402	0	0	0	1
Q92769	HDAC2_HUMAN - HDAC2 Histone deacetylase 2	67569	0	0	0	1
Q9UBB4	ATX10_HUMAN - ATXN10 Ataxin-10	36788	0	0	0	1
Q12792	TWF1_HUMAN - TWF1 Twinfilin-1	75873	0	0	0	1
Q92499	DDX1_HUMAN - DDX1 ATP-dependent RNA helicase DDX1	92482	0	0	0	1
P27824	CALX_HUMAN - CANX Calnexin	42823	0	0	0	1
Q96L21	RL10L_HUMAN - RPL10L 60S ribosomal protein L10-like	71028	0	0	0	1
Q15233	NONO_HUMAN - NONO Non-POU domain-containing octamer-binding protein	40244	0	0	0	1
P00367	DHE3_HUMAN - GLUD1 Glutamate dehydrogenase 1, mitochondrial	276539	0	0	0	1
P35520	CBS_HUMAN - CBS Cystathionine beta-synthase	19891	0	0	0	1
Q9UI26	IPO11_HUMAN - IPO11 Importin-11	11250	0	0	0	1
P33991	MCM4_HUMAN - MCM4 DNA replication licensing factor MCM4	14195	0	0	0	1
Q02809	PLOD1_HUMAN - PLOD1 Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1	12712	0	0	0	1
P02768	ALBU_HUMAN - ALB Serum albumin	12476	0	0	0	1
O43776	SYNC_HUMAN - NARS Asparagine--tRNA ligase, cytoplasmic	13833	0	0	0	1
O43143	DHX15_HUMAN - DHX15 Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	33310	0	0	0	1
P55786	PSA_HUMAN - NPEPPS Puromycin-sensitive aminopeptidase	22238	0	0	0	1

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
Q96P70	IPO9_HUMAN - IPO9 Importin-9	25561	0	0	0	1
O75390	CISY_HUMAN - CS Citrate synthase, mitochondrial	16273	0	0	0	1
P40763	STAT3_HUMAN - STAT3 Signal transducer and activator of transcription 3	20652	0	0	0	1
Q9UQ16	DYN3_HUMAN - DNM3 Dynamin-3	16773	0	0	0	1
P33176	KINH_HUMAN - KIF5B Kinesin-1 heavy chain	16793	0	0	0	1
P41252	SYIC_HUMAN - IARS Isoleucine--tRNA ligase, cytoplasmic	24328	0	0	0	1
A4D0S4	LAMB4_HUMAN - LAMB4 Laminin subunit beta-4	46588	0	0	0	1
Q4ZHG4	FNDC1_HUMAN - FNDC1 Fibronectin type III domain-containing protein 1	17887	0	0	0	1
Q9P265	DIP2B_HUMAN - DIP2B Disco-interacting protein 2 homolog B	57285	0	0	0	1
Q5CZC0	FSIP2_HUMAN - FSIP2 Fibrous sheath-interacting protein 2	66194	0	0	0	1
P62829	RL23_HUMAN - RPL23 60S ribosomal protein L23	51561	0	0	0	1
Q00796	DHSO_HUMAN - SORD Sorbitol dehydrogenase	39609	0	0	0	1
Q15102	PA1B3_HUMAN - PAFAH1B3 Platelet-activating factor acetylhydrolase IB subunit gamma	48634	0	0	0	1
P54136	SYRC_HUMAN - RARS Arginine--tRNA ligase, cytoplasmic	26558	0	0	0	1
Q9ULV4	COR1C_HUMAN - CORO1C Coronin-1C	26488	0	0	0	1
Q9Y223	GLCNE_HUMAN - GNE Bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase	38210	0	0	0	1
Q13619	CUL4A_HUMAN - CUL4A Cullin-4A	44552	0	0	0	1
Q9BXJ9	NAA15_HUMAN - NAA15 N-alpha-acetyltransferase 15, NatA auxiliary subunit	40447	0	0	0	1
O95782	AP2A1_HUMAN - AP2A1 AP-2 complex subunit alpha-1	65746	0	0	0	1
O75643	U520_HUMAN - SNRNP200 U5 small nuclear ribonucleoprotein 200 kDa helicase	43171	0	0	0	1
P52701	MSH6_HUMAN - MSH6 DNA mismatch repair protein Msh6	18506	0	0	0	1
Q9C005	DPY30_HUMAN - DPY30 Protein dpy-30 homolog	44744	0	0	1	0
A6NHG4	DDTL_HUMAN - DDTL D-dopachrome decarboxylase-like protein	74404	0	0	1	0
P30046	DOPD_HUMAN - DDT D-dopachrome decarboxylase	95786	0	0	1	0
Q9NRX4	PHP14_HUMAN - PHPT1 14 kDa phosphohistidine phosphatase	70671	0	0	1	0
Q9BPW8	NIPS1_HUMAN - NIPSNAP1 Protein NipSnap homolog 1	46871	0	0	1	0
O43399	TPD54_HUMAN - TPD52L2 Tumor protein D54	56196	0	0	1	0
P62263	RS14_HUMAN - RPS14 40S ribosomal protein S14	68330	0	0	1	0
Q9Y2B0	CNPY2_HUMAN - CNPY2 Protein canopy homolog 2	51902	0	0	1	0
Q6IS14	IF5AL_HUMAN - EIF5AL1 Eukaryotic translation initiation factor 5A-1-like	128979	0	0	1	0
Q9GZV4	IF5A2_HUMAN - EIF5A2 Eukaryotic translation initiation factor 5A-2	68935	0	0	1	0
Q9UKK9	NUDT5_HUMAN - NUDT5 ADP-sugar pyrophosphatase	103187	0	0	1	0
P11310	ACADM_HUMAN - ACADM Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	128302	0	0	1	0
P19013	K2C4_HUMAN - KRT4 Keratin, type II cytoskeletal 4	55993	0	0	1	0

IPI	Protein	Mol Weight (Da)	DMSO	LAS1	LAS6 Average	LAS12 Average
P35998	PRS7_HUMAN - PSMC2 26S protease regulatory subunit 7	147182	0	0	1	0
P07477	TRY1_HUMAN - PRSS1 Trypsin-1	73115	0	0	1	0
P07478	TRY2_HUMAN - PRSS2 Trypsin-2	113744	0	0	1	0
Q9H832	UBE2Z_HUMAN - UBE2Z Ubiquitin-conjugating enzyme E2 Z	94623	0	0	1	0
P07339	CATD_HUMAN - CTSD Cathepsin D	267290	0	0	1	0
H7C469	H7C469_HUMAN - Uncharacterized protein	284538	0	0	1	0
P60981	DEST_HUMAN - DSTN Destrin	299615	0	0	1	0
Q9NTK5	OLA1_HUMAN - OLA1 Obg-like ATPase 1	526612	0	0	1	0
Q14166	TTL12_HUMAN - TTLL12 Tubulin--tyrosine ligase-like protein 12	358201	0	0	1	0
P45974	UBP5_HUMAN - USP5 Ubiquitin carboxyl-terminal hydrolase 5	475986	0	0	1	0
P38919	IF4A3_HUMAN - EIF4A3 Eukaryotic initiation factor 4A-III	14395	0	0	1	0
Q16401	PSMD5_HUMAN - PSMD5 26S proteasome non-ATPase regulatory subunit 5	17862	0	0	1	0
Q15637	SF01_HUMAN - SF1 Splicing factor 1	18681	0	0	1	0
Q01518	CAP1_HUMAN - CAP1 Adenylyl cyclase-associated protein 1	18737	0	0	1	0
Q8WUM0	NU133_HUMAN - NUP133 Nuclear pore complex protein Nup133	75492	0	0	1	0
Q9H0B6	KLC2_HUMAN - KLC2 Kinesin light chain 2	50670	0	0	1	0
O75185	AT2C2_HUMAN - ATP2C2 Calcium-transporting ATPase type 2C member 2	35964	0	0	1	0
Q93009	UBP7_HUMAN - USP7 Ubiquitin carboxyl-terminal hydrolase 7	26923	0	0	1	0
P34897	GLYM_HUMAN - SHMT2 Serine hydroxymethyltransferase, mitochondrial	47629	0	0	1	0
Q92945	FUBP2_HUMAN - KHSRP Far upstream element-binding protein 2	57206	0	0	1	0
O00754	MA2B1_HUMAN - MAN2B1 Lysosomal alpha-mannosidase	94270	0	0	1	0
P43243	MATR3_HUMAN - MATR3 Matrin-3	56257	0	0	1	0
Q13813	SPTN1_HUMAN - SPTAN1 Spectrin alpha chain, non-erythrocytic 1	46638	0	0	1	0
P98088	MUC5A_HUMAN - MUC5AC Mucin-5AC	61277	0	0	1	0
P49792	RBP2_HUMAN - RANBP2 E3 SUMO-protein ligase RanBP2	51854	0	0	1	0
Q9C0G6	DYH6_HUMAN - DNAH6 Dynein heavy chain 6, axonemal	89988	0	0	1	0
P68036	UB2L3_HUMAN - UBE2L3 Ubiquitin-conjugating enzyme E2 L3	149323	0	0	1	0
Q8NBS9	TXND5_HUMAN - TXNDC5 Thioredoxin domain-containing protein 5	149714	0	0	1	0
Q13099	IFT88_HUMAN - IFT88 Intraflagellar transport protein 88 homolog	137499	0	0	1	0
P43034	LIS1_HUMAN - PAFAH1B1 Platelet-activating factor acetylhydrolase IB subunit alpha	381596 5	0	0	1	0
Q2QGD7	ZXDC_HUMAN - ZXDC Zinc finger protein ZXDC	17260	0	0	1	0
P55201	BRPF1_HUMAN - BRPF1 Peregrin	26788	0	0	1	0
Q8WZ42	TITIN_HUMAN - TTN Titin	49953	0	0	1	0
Q8NHL6	LIRB1_HUMAN - LILRB1 Leukocyte immunoglobulin-like receptor subfamily B member 1	70819	0	0	0	0
Q13748	TBA3C_HUMAN - TUBA3D Tubulin alpha-3C/D chain	49960	28	0	0	0

Table 4A-1. Proteome wide cysteine reactivity profiling to identify cysteine residues modified by MIND4-17.

index	ipi	Protein	Peptide sequence	MIND-17
888	P47712	PLA2G4A Cytosolic phospholipase A2	CSVSLSNVEAR	3.37
514	P78527	PRKDC DNA-dependent protein kinase catalytic subunit	INQVFHGSCITEGNELTK	2.4
201	P28074	PSMB5 Proteasome subunit beta type-5	VIEINPYLLGTMAGGAADCSFWER	2.29
118	P46782	RPS5 40S ribosomal protein S5	AQCPIVER	2.09
252	Q96CD2	PPCDC Phosphopantothenoylcysteine decarboxylase	ASCPAAAPLMER	1.62
839	Q9P258	RCC2 Protein RCC2	AVQDLCGWR	1.56
497	P12955	PEPD Xaa-Pro dipeptidase	YCTDTGVLFR	1.52
800	O14744	PRMT5 Protein arginine N-methyltransferase 5	DLNCVPEIADTLGAVAK	1.44
598	P32320	CDA Cytidine deaminase	GCNIENACYPLGICAER	1.41
440	Q9Y3F4	STRAP Serine-threonine kinase receptor-associated protein	CVLPEEDSGELAK	1.39
239	Q7RTV0	PHF5A PHD finger-like domain-containing protein 5A	PCTLVR	1.38
537	P36578	RPL4 60S ribosomal protein L4	SGQGAFGNMCR	1.33
748	Q9NUP9	LIN7C Protein lin-7 homolog C	VLQSEFCNAVR	1.33
654	P29034	S100A2 Protein S100-A2	YSCQEGDK	1.3
225	Q9UNH7	SNX6 Sorting nexin-6	IGSSLYALGTQDSTDICK	1.3
549	Q9H4A6	GOLPH3 Golgi phosphoprotein 3	LQLEACGMR	1.28
887	O00469	PLOD2 Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	LDPDMALCR	1.25
268	O14980	XPO1 Exportin-1	LDINLLDNVVNCLYHGEGAQQR	1.22
709	O43707	ACTN4 Alpha-actinin-4	EGLLLWCQR	1.21
166	Q9Y277	VDAC3 Voltage-dependent anion-selective channel protein	VCNYGLTFTQK	1.19
397	O60361	NME2P1 Putative nucleoside diphosphate kinase	GDFCIQVGR	1.18
778	Q9Y570	PPME1 Protein phosphatase methylesterase 1	GLSNLFLSCPIPK	1.18
742	P49327	FASN Fatty acid synthase	LGMLSPEGTCK	1.17
295	Q9UPQ0	LIMCH1 LIM and calponin homology domains-containing prote	AANSCTSYSGTTLNLK	1.16
619	P23921	RRM1 Ribonucleoside-diphosphate reductase large subunit	IIDINYYPVPEACLSNK	1.16
17	P45880	VDAC2 Voltage-dependent anion-selective channel protein	SCSGVEFSTSGSSNTDTGK	1.16
117	P38606	ATP6V1A V-type proton ATPase catalytic subunit A	VLDALFPCVQGGTTAIPGAFGCGK	1.16
788	P32322	PYCR1 Pyrroline-5-carboxylate reductase 1, mitochondrial	CMTNTPVVVR	1.15
104	P23396	RPS3 40S ribosomal protein S3	GLCAIAQAESLR	1.14
352	P55060	CSE1L Exportin-2	ICAVGITK	1.14
207	Q14192	FHL2 Four and a half LIM domains protein 2	DDFAYCLNCFCDLYAK	1.12

380	P00505	GOT2 Aspartate aminotransferase, mitochondrial	TCGFDFTGAVEDISK	1.12
747	P30084	ECHS1 Enoyl-CoA hydratase, mitochondrial	EMQNLSFQDCYSSK	1.12
158	P34932	HSPA4 Heat shock 70 kDa protein 4	LMSANASDLPLSIECFMNDVDVSGTMNR	1.12
328	P30084	ECHS1 Enoyl-CoA hydratase, mitochondrial	ALNALCDGLIDELNQALK	1.11
373	P36507	MAP2K2 Dual specificity mitogen-activated protein kinase	LCDFGVSGQLIDSMANSFVGTR	1.11
68	P28062	PSMB8 Proteasome subunit beta type-8	LLSNMMCQYR	1.11
705	Q16658	FSCN1 Fascin	LSCFAQTVSPAEK	1.1
61	P05388	RPLP0 60S acidic ribosomal protein P0	CFIVGADNVGSK	1.1
507	P36873	PPP1CC Serine/threonine-protein phosphatase PP1-gamma cat	GNHECASINR	1.09
342	J3KR12	Uncharacterized protein	SLLINAVEASCIR	1.09
671	P46782	RPS5 40S ribosomal protein S5	VNQAIWLLCTGAR	1.09
395	Q96QK1	VPS35 Vacuolar protein sorting-associated protein 35	TQCALAASK	1.09
249	P28838	LAP3 Cytosol aminopeptidase	SAGACTAAAFLK	1.08
70	P60174	TPI1 Triosephosphate isomerase	VPADTEVVCAPPTAYIDFAR	1.08
653	Q96FW1	OTUB1 Ubiquitin thioesterase OTUB1	QEPLGSDSEGVNCLAYDEAIMAQQDR	1.08
363	P68366	TUBA4A Tubulin alpha-4A chain	YMACCLLYR	1.08
339	Q9UL46	PSME2 Proteasome activator complex subunit 2	CGFLPGNEK	1.08
284	P15880	RPS2 40S ribosomal protein S2	GCTATLGNFAK	1.07
240	Q9BUF5	TUBB6 Tubulin beta-6 chain	EIVHIQAGQCGNQIGTK	1.06
898	O15020	SPTBN2 Spectrin beta chain, non-erythrocytic 2	IHCLENVDK	1.06
548	P31153	MAT2A S-adenosylmethionine synthase isoform type-2	TCNVLVALEQQSPDIAQGVHLDR	1.05
140	P08238	HSP90AB1 Heat shock protein HSP 90-beta	VFIMDSCDELIPEYLNFIR	1.05
95	Q14432	PDE3A cGMP-inhibited 3,5-cyclic phosphodiesterase A	ISPLSSPCSSPLQGTPASSLVSK	1.05
303	P49327	FASN Fatty acid synthase	AFDTAGNGYCR	1.04
206	P13797	PLS3 Plastin-3	VDLNSNGFICDYELHELFK	1.03
366	P68366	TUBA4A Tubulin alpha-4A chain	AVCMLSNTTAIAEAWAR	1.03
846	O75521	ECI2 Enoyl-CoA delta isomerase 2, mitochondrial	WLSDECTNAVVNFLSR	1.03
116	O95372	LYPLA2 Acyl-protein thioesterase 2	TYPGVMHSSCPQEMAAVK	1.03
536	Q86W42	THOC6 THO complex subunit 6 homolog	AQVPGSSPGLLSLSLNQQPAAPECK	1.02
341	P57721	PCBP3 Poly(rC)-binding protein 3	INISEGNCPER	1.02
48	P36952	SERPINB5 Serpin B5	ACLENLGLK	1.02
149	A6NDG6	PGP Phosphoglycolate phosphatase	FIAGTGCLVR	1.02
244	O43583	DENR Density-regulated protein	VLYCGVCSLPTEYCEYMPDVAK	1.02
765	O95456	PSMG1 Proteasome assembly chaperone 1	VFGSCPR	1.02
78	P57721	PCBP3 Poly(rC)-binding protein 3	LVVPASQCGSLIGK	1.02
343	P18085	ARF4 ADP-ribosylation factor 4	NICFTVWDVGGQDR	1.01
75	P53396	ACLY ATP-citrate synthase	FICTTSAIQNR	1.01

39	P60174	TPI1 Triosephosphate isomerase	IAVAAQNCYK	1.01
801	P35579	MYH9 Myosin-9	VEDMAELTCLNEASVLHNLK	1.01
155	P18669	PGAM1 Phosphoglycerate mutase 1	YADLTEDQLPSCESLK	1.01
837	P10768	ESD S-formylglutathione hydrolase	CPALYWLSGLTCTEQNFISK	1.01
113	P60709	ACTB Actin, cytoplasmic 1	LCYVALDFEQEMATAASSSSLEK	1
32	Q15365	PCBP1 Poly(rC)-binding protein 1	LVVPATQCGSLIGK	1
37	P00505	GOT2 Aspartate aminotransferase, mitochondrial	VGAFTMVCK	1
547	P07737	PFN1 Profilin-1	CYEMASHLR	1
574	P60174	TPI1 Triosephosphate isomerase	IIYGGSVTGATCK	1
340	P07900	HSP90AA1 Heat shock protein HSP 90-alpha	VFIMDNCEELIPEYLNFIR	1
312	P62753	RPS6 40S ribosomal protein S6	LNISFPATGCQK	1
614	P40227	CCT6A T-complex protein 1 subunit zeta	NAIDDGCVVPGAGAVEVAMAEALIK	1
144	Q9BWD1	ACAT2 Acetyl-CoA acetyltransferase, cytosolic	ATVAPEDVSEVIFGHVLAAGCGQNPVR	1
645	Q52LJ0	FAM98B Protein FAM98B	SLCNLEESITSAGR	1
251	Q9NVG8	TBC1D13 TBC1 domain family member 13	LLQDYPITDVCQILQK	1
647	Q9NQ88	TIGAR Fructose-2,6-bisphosphatase TIGAR	EQFSQGSPSNCLETSLAEIFPLGK	1
107	Q9BWD1	ACAT2 Acetyl-CoA acetyltransferase, cytosolic	QASVGAGIPYSVPAWSCQMICGSGLK	1
300	P53384	NUBP1 Cytosolic Fe-S cluster assembly factor NUBP1	GASCQGCPNQR	1
325	P61158	ACTR3 Actin-related protein 3	YSYVCPDLVK	1
19	P12532	CKMT1B Creatine kinase U-type, mitochondrial	LGYILTCPSNLGTGLR	1
368	Q7L1Q6	BZW1 Basic leucine zipper and W2 domain-containing prot	FDPTQFQDCIIQGLTETGTDLEAVAK	1
45	P16455	MGMT Methylated-DNA--protein-cysteine methyltransferase	VVCSSGAVGNYSGGLAVK	1
355	Q99832	CCT7 T-complex protein 1 subunit eta	EGTDSSQGIPQLVSNISACQVIAEAVR	1
394	P49721	PSMB2 Proteasome subunit beta type-2	NLADCLR	0.99
47	P26641	EEF1G Elongation factor 1-gamma	FPEELTQTFMSCNLITGMFQR	0.99
399	P78347	GTF2I General transcription factor II-I	SILSPGGSCGPIK	0.99
126	Q08J23	NSUN2 tRNA (cytosine(34)-C(5))-methyltransferase	MVYSTCSLNPIEDEAVIASLLEK	0.99
526	Q9UGI8	TES Testin	LPCEMDAQGPK	0.99
256	P50914	RPL14 60S ribosomal protein L14	ALVDGPCTQVR	0.99
316	P33240	CSTF2 Cleavage stimulation factor subunit 2	LCVQNSPQEAR	0.99
296	P10809	HSPD1 60 kDa heat shock protein, mitochondrial	AAVEEGIVLGGGCALLR	0.99
675	P68366	TUBA4A Tubulin alpha-4A chain	TIGGGDDSFTTFFCETGAGK	0.99
288	P29401	TKT Transketolase	MAAISESNINLCGSHCGVSIGEDGPSQMALED LAMFR	0.99
354	P62829	RPL23 60S ribosomal protein L23	ISLGLPVGAVINCADNTGAK	0.99
44	P31327	CPS1 Carbamoyl-phosphate synthase	TSACFEPSLDYMVTK	0.98
601	P07954	FH Fumarate hydratase, mitochondrial	FEALAAHDALVELSGAMNTTACSLMK	0.98
138	Q14181	POLA2 DNA polymerase alpha subunit B	VLGCPEALTGSYK	0.98

375	P62306	SNRPF Small nuclear ribonucleoprotein F	CNNVLYIR	0.98
38	P00338	LDHA L-lactate dehydrogenase A chain	VIGSGCNLDSAR	0.98
101	Q01813	PFKP 6-phosphofructokinase type C	NESCSENYTTDFIYQLYSEEGK	0.98
51	P68366	TUBA4A Tubulin alpha-4A chain	AYHEQLSVAEITNACFEPANQMVK	0.98
386	P07437	TUBB Tubulin beta chain	LTTPTYGDLNHLVSATMSGVTTCLR	0.98
40	P12004	PCNA Proliferating cell nuclear antigen	CAGNEDIITLR	0.97
102	F5H284	PPIAL4D Peptidyl-prolyl cis-trans isomerase	IIPGFMCQGGDFTR	0.97
587	P63244	GNB2L1 Guanine nucleotide-binding protein subunit beta-2-	AEPPQCTSLAWSADGQTLFAGYTDNLVR	0.97
630	Q9HB90	RRAGC Ras-related GTP-binding protein C	SCGHQTSASSLK	0.97
58	P07814	EPRS Bifunctional glutamate/proline--tRNA ligase	LGVENCYFPMFVSQSALEK	0.97
98	P29401	TKT Transketolase	QAFTDVATGSLGQGLGAACGMAYTGK	0.97
804	Q96I99	SUCLG2 Succinyl-CoA ligase	IDATQVEVNPFGETPEGQVVCFDAK	0.97
886	P13639	EEF2 Elongation factor 2	TFCQLILDPIFK	0.97
444	Q96HC4	PDLIM5 PDZ and LIM domain protein 5	GALYCELCYEK	0.96
262	A6NDG6	PGP Phosphoglycolate phosphatase	NNQESDCVSK	0.96
525	P30041	PRDX6 Peroxiredoxin-6	DFTPVCTTELGR	0.96
267	O75822	EIF3J Eukaryotic translation initiation factor 3 subunit	ITNSLTVLCSEK	0.96
314	P07203	GPX1 Glutathione peroxidase 1	FQTIDIEPDIEALLSQGPSCA	0.96
385	P21291	CSRP1 Cysteine and glycine-rich protein 1	SCFLCMVCK	0.96
629	Q14847	LASP1 LIM and SH3 domain protein 1	ACFHCETCK	0.96
310	Q99832	CCT7 T-complex protein 1 subunit eta	QLCDNAGFDATNILNK	0.96
362	P50990	CCT8 T-complex protein 1 subunit theta	IAVYSCPFDGMITETK	0.96
361	P68366	TUBA4A Tubulin alpha-4A chain	SIQFVDWCPTGFK	0.96
734	P37235	HPCAL1 Hippocalcin-like protein 1	LLQCDPSSASQF	0.96
36	I3L2F9	Uncharacterized protein	TAVCDIPPR	0.96
218	P42677	RPS27 40S ribosomal protein S27	LTEGCSFR	0.96
369	P13639	EEF2 Elongation factor 2	STLTDSLVCK	0.96
46	P53597	SUCLG1 Succinyl-CoA ligase	LIGPNCPGVINPGECK	0.96
65	O75874	IDH1 Isocitrate dehydrogenase	SEGGFIWACK	0.95
590	Q9UHD8	SEPT9 Septin-9	LTVIDTPGFGDHINNENCWQPIMK	0.95
304	P31327	CPS1 Carbamoyl-phosphate synthase	VVAVDCGIK	0.95
27	O95336	PGLS 6-phosphogluconolactonase	AACCLAGAR	0.95
246	Q13418	ILK Integrin-linked protein kinase	FSFQCPGR	0.95
76	P62280	RPS11 40S ribosomal protein S11	DVQIGDIVTVGECR	0.95
90	Q15365	PCBP1 Poly(rC)-binding protein 1	VMTIPYQPMPASSPVICAGGQDR	0.95
818	Q14192	FHL2 Four and a half LIM domains protein 2	DNQNFCVPCYEK	0.95
24	P46782	RPS5 40S ribosomal protein S5	TIAECLADELINAAK	0.95
25	Q07020	RPL18 60S ribosomal protein L18	GCGTVLLSGPR	0.94

346	P08670	VIM Vimentin	QVQSLTCEVDALK	0.94
131	Q09666	AHNAK Neuroblast differentiation-associated protein AHNA	GPFVEAEVPDVDLECPDAK	0.94
370	O43684	BUB3 Mitotic checkpoint protein BUB3	TPCNAGTFSQPEK	0.94
234	Q9NP73	ALG13 UDP-N-acetylglucosamine transferase subunit ALG13	ADLVISHAGAGSCLETLEK	0.94
605	P21333	FLNA Filamin-A	LQVEPAVDTSGVQCYGPGIEGQGVFR	0.94
277	P67936	TPM4 Tropomyosin alpha-4 chain	EENVGLHQTLDQTLNELNCI	0.94
628	Q15370	TCEB2 Transcription elongation factor B polypeptide 2	ADDTFEALCIEPFSSPPELPDVMK	0.94
1	Q9Y696	CLIC4 Chloride intracellular channel protein 4	AGSDGESIGNCPFSQR	0.94
220	P08238	HSP90AB1 Heat shock protein HSP 90-beta	LVSSPCCIVTSTYGWTANMER	0.94
390	Q9Y3F4	STRAP Serine-threonine kinase receptor-associated protei	IGFPETTEEELEEIASENSDCIFPSAPDVK	0.94
389	I3L2F9	Uncharacterized protein	NMMAACDPR	0.94
735	Q15181	PPA1 Inorganic pyrophosphatase	GISCMNTTLSESPFK	0.94
290	Q9UFW8	CGGBP1 CGG triplet repeat-binding protein 1	PLTASLQCNSTAQTEK	0.94
393	Q01813	PFKP 6-phosphofructokinase type C	LPLMECVQMTQDVQK	0.94
327	P09211	GSTP1 Glutathione S-transferase P	ASCLYGQLPK	0.94
376	P48735	IDH2 Isocitrate dehydrogenase	SSGGFVWACK	0.94
323	Q13501	SQSTM1 Sequestosome-1	FSFCCSPEPEAEAEAAAGPGPCER	0.94
106	O00244	ATOX1 Copper transport protein ATOX1	HEFSVDMTCGGCAEAVSR	0.93
274	P31947	SFN 14-3-3 protein sigma	GEELSCEER	0.93
97	P38606	ATP6V1A V-type proton ATPase catalytic subunit A	YSNSDVIIYVGCGER	0.93
572	P63208	SKP1 S-phase kinase-associated protein 1	ENQWCEEK	0.93
217	Q9NZL4	HSPBP1 Hsp70-binding protein 1	DACDTVR	0.93
721	Q9NQ88	TIGAR Fructose-2,6-bisphosphatase TIGAR	EECPVFTPPGGETLDQVK	0.93
378	H7C455	Uncharacterized protein	DCK	0.93
213	P61978	HNRNPK Heterogeneous nuclear ribonucleoprotein K	IIPTLEEGLQLPSPTATSQLPLESDAVECLNYQH YK	0.93
297	P23528	CFL1 Cofilin-1	HELQANCYEEVK	0.92
279	Q96EY8	MMAB Cob(I)yrinic acid a,c-diamide adenosyltransferase,	IQCTLQDVGSALATPCSSAR	0.92
309	P62826	RAN GTP-binding nuclear protein Ran	VCENIPIVLCGNK	0.92
356	P26641	EEF1G Elongation factor 1-gamma	AAAPAPEEEMDECEQALAAEPK	0.92
228	P55769	NHP2L1 NHP2-like protein 1	LLDLVQQSCNYK	0.92
652	P21266	GSTM3 Glutathione S-transferase Mu 3	IAAYLQSDQFCK	0.92
93	O00299	CLIC1 Chloride intracellular channel protein 1	IGNCPFSQR	0.92
706	P30084	ECHS1 Enoyl-CoA hydratase, mitochondrial	ICPVETLVEEAIQCAEK	0.92
22	P78417	GSTO1 Glutathione S-transferase omega-1	FCPFAER	0.92
329	Q15019	SEPT2 Septin-2	LTVVDTPGYGDAINCR	0.92
398	Q99714	HSD17B10 3-hydroxyacyl-CoA dehydrogenase type-2	LGNNCVFAPADVTSEK	0.92
895	Q9Y5Y2	NUBP2 Cytosolic Fe-S cluster assembly factor NUBP2	AVHQCDR	0.92

559	Q16555	DPYSL2 Dihydropyrimidinase-related protein 2	FQLTDCQIYEVLSVIR	0.91
83	P14866	HNRNPL Heterogeneous nuclear ribonucleoprotein L	LCFSTAQHAS	0.91
53	O15067	PFAS Phosphoribosylformylglycinamidine synthase	FCDNSSAIQGK	0.91
626	Q13748	TUBA3D Tubulin alpha-3C/D chain	TIQFVDWCPTGFK	0.91
253	Q13526	PIN1 Peptidyl-prolyl cis-trans isomerase NIMA-interacti	SGEEDFESLASQFSDCSSAK	0.91
41	O95861	BPNT1 3(2),5-bisphosphate nucleotidase 1	TCATDLQTK	0.91
56	P49327	FASN Fatty acid synthase	DPETLVGYSMVGCQR	0.91
306	P22695	UQCRC2 Cytochrome b-c1 complex subunit 2, mitochondrial	NALANPLYCPDYR	0.91
120	Q6XQN6	NAPRT1 Nicotinate phosphoribosyltransferase	LQALVNSLCAGQSP	0.9
326	P11802	CDK4 Cyclin-dependent kinase 4	LMDVCATSR	0.9
280	P49327	FASN Fatty acid synthase	AINCATSGVVGLVNCLR	0.9
618	P60981	DSTN Destrin	CSTPEEIK	0.9
28	P08237	PFKM 6-phosphofructokinase, muscle type	LPLMECVQVTK	0.9
211	P61289	PSME3 Proteasome activator complex subunit 3	LDECEEAFQGTK	0.9
359	P37802	TAGLN2 Transgelin-2	NMACVQR	0.9
291	P23396	RPS3 40S ribosomal protein S3	GCEVVVSGK	0.9
344	P21291	CSRP1 Cysteine and glycine-rich protein 1	DGEIYCK	0.9
54	Q14258	TRIM25 E3 ubiquitin/ISG15 ligase TRIM25	NTVLCNVVEQFLQADLAR	0.9
165	Q04695	KRT17 Keratin, type I cytoskeletal 17	LSGGLGAGSCR	0.9
293	P02795	MT2A Metallothionein-2	SCCSCCPVGCAK	0.89
94	P60981	DSTN Destrin	LGGSLIVAFEGCPV	0.89
132	P62937	PPIA Peptidyl-prolyl cis-trans isomerase A	ITIADCGQLE	0.89
388	P42330	AKR1C3 Aldo-keto reductase family 1 member C3	WVDPNSPVLLEDPVLCALAK	0.89
392	P27348	YWHAQ 14-3-3 protein theta	YLAEVACGDDR	0.89
651	P08238	HSP90AB1 Heat shock protein HSP 90-beta	GFEVVYMTEPIDEYCVQQLK	0.89
745	P23528	CFL1 Cofilin-1	DCR	0.89
787	Q9NUU7	DDX19A ATP-dependent RNA helicase DDX19A	VLVTTNVCAR	0.89
803	Q9Y3D2	MSRB2 Methionine-R-sulfoxide reductase B2, mitochondrial	GQAGGGGPGTGPGLGEAGSLATCELPLAK	0.89
324	P09110	ACAA1 3-ketoacyl-CoA thiolase, peroxisomal	DCLIPMGITSENVAER	0.88
99	O43865	AHCYL1 Putative adenosylhomocysteinase 2	LCVPAMNVNDSVTK	0.88
286	Q9NVG8	TBC1D13 TBC1 domain family member 13	SLDDSQCGITYK	0.88
381	Q7L2H7	EIF3M Eukaryotic translation initiation factor 3 subunit	VAASCGAIQYIPTELDQVR	0.88
92	P37802	TAGLN2 Transgelin-2	DGTVLCELINALYPEGQAPVK	0.88
91	P04406	GAPDH Glyceraldehyde-3-phosphate dehydrogenase	VPTANVSVVDLTCR	0.88
315	P14618	PKM Pyruvate kinase isozymes M1/M2	CCSGAIIVLTK	0.88
114	Q9NQR4	NIT2 Omega-amidase NIT2	VGLGICYDMR	0.88

311	P21291	CSRP1 Cysteine and glycine-rich protein 1	TVYFAEEVQCEGNSFHK	0.88
347	Q15417	CNN3 Calponin-3	CASQAGMTAYGTR	0.88
391	P27635	RPL10 60S ribosomal protein L10	MLSCAGADR	0.88
334	P05388	RPLP0 60S acidic ribosomal protein P0	AGAIAPCEVTVPAQNTGLGPEK	0.88
125	P04732	MT1E Metallothionein-1E	CAQGCVCK	0.87
577	P00558	PGK1 Phosphoglycerate kinase 1	GCITIIGGGDTATC*C*AK	0.87
222	Q2TAA2	IAH1 Isoamyl acetate-hydrolyzing esterase 1 homolog	VILITPTPLCETAWEEQCIIQGCK	0.87
865	P51610	HCFC1 Host cell factor 1	ACAAGTPAVIR	0.87
74	P24752	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	QAVLGAGLPISTPCTTINK	0.87
69	P12277	CKB Creatine kinase B-type	FCTGLTQIETLFK	0.86
513	Q9Y508	RNF114 RING finger protein 114	DCGGAAQLAGPAAEADPLGR	0.86
317	P04406	GAPDH Glyceraldehyde-3-phosphate dehydrogenase	IISNASCTTNCLAPLAK	0.86
209	P60709	ACTB Actin, cytoplasmic 1	CPEALFQPSFLGMESCGIHETTFNSIMK	0.86
576	P07195	LDHB L-lactate dehydrogenase B chain	GMYGIENEVFLSLPCILNAR	0.86
575	P00338	LDHA L-lactate dehydrogenase A chain	DDVFLSVPCILGQNGISDLVK	0.86
358	Q13185	CBX3 Chromobox protein homolog 3	LTWHSCPEDEAQ	0.86
604	P62879	GNB2 Guanine nucleotide-binding protein G(I)/G(S)/G(T)	TFVSGACDASIK	0.86
650	Q9NRP4	ACN9 Protein ACN9 homolog, mitochondrial	ACFGTFLPEEK	0.86
258	P49411	TUFM Elongation factor Tu, mitochondrial	GEETPVIVGSALCALEGR	0.86
115	O75663	TIPRL TIP41-like protein	VACAEEWQESR	0.86
322	P62249	RPS16 40S ribosomal protein S16	TATAVAHCK	0.86
281	Q86SX6	GLRX5 Glutaredoxin-related protein 5, mitochondrial	GTPEQPQCGFSNAVVQILR	0.86
561	Q9Y5Y2	NUBP2 Cytosolic Fe-S cluster assembly factor NUBP2	VGILDVDLCGPSIPR	0.85
351	P62888	RPL30 60S ribosomal protein L30	VCTLAIIDPGDSDIIR	0.85
43	Q9BRA2	TXNDC17 Thioredoxin domain-containing protein 17	SWCPDCVQAEPVVR	0.85
744	P04183	TK1 Thymidine kinase, cytosolic	LFAPQQILQCSPAN	0.85
305	P25398	RPS12 40S ribosomal protein S12	LGEWVGLCK	0.85
270	O14950	MYL12B Myosin regulatory light chain 12B	NAFACFDEEATGTIQEDYLR	0.85
162	Q14790	CASP8 Caspase-8	VFFIQACQGDNYQK	0.85
534	Q7L8W6	ATPBD4 ATP-binding domain-containing protein 4	CEGDEVEDLYELLK	0.84
233	P30044	PRDX5 Peroxiredoxin-5, mitochondrial	ALNVEPDGTGLTCSLAPNIISQL	0.84
603	Q9BY32	ITPA Inosine triphosphate pyrophosphatase	GCQDFGWDPCFQPDGYEQTYAEMPK	0.84
720	Q09666	AHNAK Neuroblast differentiation-associated protein AHNA	LEGDLTGPSVGVEVPDVELECPDAK	0.84
283	P10809	HSPD1 60 kDa heat shock protein, mitochondrial	CEFQDAYVLLSEK	0.84
840	P82932	MRPS6 28S ribosomal protein S6, mitochondrial	ECEGIVPVPLAEK	0.84
419	P62701	RPS4X 40S ribosomal protein S4, X isoform	ECLPLIIFLR	0.83
128	P00505	GOT2 Aspartate aminotransferase, mitochondrial	EYLPIGGLAEFCK	0.83

111	Q01518	CAP1 Adenylyl cyclase-associated protein 1	ALLVTASQCQQPAENK	0.83
332	P63244	GNB2L1 Guanine nucleotide-binding protein subunit beta-2-	FSPNSSNPIIVSCGWDK	0.83
10	P21291	CSRP1 Cysteine and glycine-rich protein 1	NLDSTTVAVHGEEIYCK	0.83
364	Q99439	CNN2 Calponin-2	AGQCVIGLQMGTNK	0.83
408	P30153	PPP2R1A Serine/threonine-protein phosphatase 2A 65 kDa reg	LNIISNLDCVNEVIGIR	0.82
88	P33316	DUT Deoxyuridine 5-triphosphate nucleotidohydrolase,	IAQLICER	0.82
34	Q96I99	SUCLG2 Succinyl-CoA ligase	SCNGPVLVGSPQGGVDIEEVAASNPELIFK	0.82
50	P04075	ALDOA Fructose-bisphosphate aldolase A	ALANSLACQGK	0.82
82	P10809	HSPD1 60 kDa heat shock protein, mitochondrial	CIPALDSLTPANEDQK	0.82
333	Q9Y5P6	GMPPB Mannose-1-phosphate guanyltransferase beta	LCSGPGIVGNVLVDPSAR	0.82
57	P40926	MDH2 Malate dehydrogenase, mitochondrial	TIIPLISQCTPK	0.82
330	P62316	SNRPD2 Small nuclear ribonucleoprotein Sm D2	NNTQVLINCR	0.81
743	P50238	CRIP1 Cysteine-rich protein 1	CNK	0.81
134	P14174	MIF Macrophage migration inhibitory factor	LLCGLLAER	0.8
157	P48047	ATP5O ATP synthase subunit O, mitochondrial	GEVPCTVTSASPLEEATLSELK	0.8
299	P30626	SRI Sorcin	DTAQQGVVNFPYDDFIQCVMSV	0.78
224	Q99873	PRMT1 Protein arginine N-methyltransferase 1	VIGIECSSISDYAVK	0.78
62	Q9UI30	TRMT112 tRNA methyltransferase 112 homolog	ICPVEFNPNFVAR	0.78
152	P24752	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	QGEYGLASICNGGGGASAMLIQK	0.78
817	P60866	RPS20 40S ribosomal protein S20	TPCGEGSK	0.78
387	P62333	PSMC6 26S protease regulatory subunit 10B	AVASQLDCNFLK	0.77
336	Q99439	CNN2 Calponin-2	CASQSGMTAYGTR	0.77
77	Q7RTV0	PHF5A PHD finger-like domain-containing protein 5A	ICDECNYGSYQGR	0.76
377	P42166	TMPO Lamina-associated polypeptide 2, isoform alpha	SGIQPLCPER	0.76
372	P63220	RPS21 40S ribosomal protein S21	TYAICGAIR	0.75
867	P35998	PSMC2 26S protease regulatory subunit 7	LCPNSTGAEIR	0.75
203	Q9H479	FN3K Fructosamine-3-kinase	AFGGPGAGCISEGR	0.75
294	Q12849	GRSF1 G-rich sequence factor 1	YIELFLNSCPK	0.74
762	O00233	PSMD9 26S proteasome non-ATPase regulatory subunit 9	GIGMNEPLVDCEGYPR	0.74
163	P24752	ACAT1 Acetyl-CoA acetyltransferase, mitochondrial	IHMGSCAENTAK	0.74
307	P49458	SRP9 Signal recognition particle 9 kDa protein	VTDDLVCLVYK	0.73
589	P38117	ETFB Electron transfer flavoprotein subunit beta	EVIAVSCGPAQCQETIR	0.72
379	P21980	TGM2 Protein-glutamine gamma-glutamyltransferase 2	VVSGMVNCNDDQGVLLGR	0.7
648	Q13200	PSMD2 26S proteasome non-ATPase regulatory subunit 2	GTLTLCPYHSDR	0.69
200	Q9BW61	DDA1 DET1- and DDB1-associated protein 1	FHADSVCK	0.66

164	Q93052	LPP Lipoma-preferred partner	TYITDPVSAPCAPPLQPK	0.66
917	P31943	HNRNPH1 Heterogeneous nuclear ribonucleoprotein H	DLNYCFSGMSDHR	0.66
212	P34896	SHMT1 Serine hydroxymethyltransferase, cytosolic	AVLEALGSCLNNK	0.64
89	P68104	EEF1A1 Elongation factor 1-alpha 1	PMCVESFSDYPPLGR	0.63
704	P23528	CFL1 Cofilin-1	AVLFCLSEDK	0.59
18	P14649	MYL6B Myosin light chain 6B	ILYSQCGDVMR	0.59
913	Q14247	CTTN Src substrate cortactin	HCSQVDSVR	0.59
60	P49189	ALDH9A1 4-trimethylaminobutyraldehyde dehydrogenase	GALMANFLTQGQVCCNGTR	0.57
319	O14929	HAT1 Histone acetyltransferase type B catalytic subunit	LCQDLPCFSR	0.56
241	Q14919	DRAP1 Dr1-associated corepressor	ACQVTQSR	0.5
761	Q9H0L4	CSTF2T Cleavage stimulation factor subunit 2 tau variant	LCVQNSHQEAR	0.35
261	P83731	RPL24 60S ribosomal protein L24	VELCSFSGYK	0.3
828	P62280	RPS11 40S ribosomal protein S11	CPFTGNVSIR	0.17