Stony Brook University

Developing novel chemical tools to study *O*-GlcNAc modifications

A Thesis Presented

by

Marian Shashika Fernando

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Master of Science

in

Chemistry

Stony Brook University

December 2008

Stony Brook University

The Graduate School

Marian Shashika Fernando

We, the thesis committee for the above candidate for the

Masters of Science degree, hereby recommend

acceptance of this thesis.

Isaac Carrico – Thesis Advisor

Assistant Professor, Department of Chemistry

Dr. Nicole Sampson – Chairperson of defense

Professor, Department of Chemistry

Dr. Daniel Raleigh – Third member

Professor, Department of Chemistry

This thesis is accepted by the Graduate School

Lawrence Martin
Dean of the Graduate School

Abstract of the thesis
Developing novel chemical tools to study *O*-GlcNAc modifications
By
Marian Shashika Fernando
Masters of Science
in
Chemistry
Stony Brook University
2008

Understanding the physiological impact of post-translational modifications (PTMs) are critical because they alter physical and chemical properties of proteins. *O*-GlcNAc modification is one such PTM discovered in early 1980's by Gerald Hart. *O*-GlcNAc modifications were thought to be a novel PTM because they seemed non-extended to a branched structure and the modified proteins were localized in the nucleocytoplasm unlike classical glycoproteins. *O*-GlcNAcylated proteins include chromatins, transcription factors, nuclear pore proteins, and certain types of cytoskeletal proteins. Changes in *O*-GlcNAc modifications are thought to play a role in some diseases including diabetes, Alzheimer's disease, and cancer. Detecting and studying *O*-GlcNAc modifications has been a great challenge due to the complex nature of these modifications. The traditional methods of detecting *O*-GlcNAc modified proteins include *O*-GlcNAc antibodies, wheat germ agglutinin (WGA), and radiolabeling using GalT. Recently developed, more advanced chemical tools include the BEMAD strategy, reporter and affinity tags, isotopic labeling, and chemical reporter stategy. Expanding the chemical toolkit to investigate the functional specificities of *O*-GlcNAc modified proteins will have a tremendous impact on the field of glycobiology.

The overall aim of this project is to develop a novel chemical tool to further study *O*-GlcNAc modifications in different cellular states. The developed novel chemical tool utilizes the bioorthogonal chemical reporter strategy and 2D Differential Gel Electrophoresis (DIGE).

The developed tool, which is a variation of traditional DIGE provides the ability to compare protein expression levels between two cellular states. The former portion of the chemical tool utilizes the chemical reporter strategy, while the latter portion of the novel chemical tool utilizes DIGE.

For my family.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS VII
LIST OF FIGURES VIII
ACKNOWLEDGMENTS V

CHAPTER 1 1

INTRODUCTION 2
PROPERTIES AND CELLULAR FUNCTIONS OF *O*-GLCNAC MODIFICATIONS 4
O-GLCNAC MODIFICATIONS AND DISEASES 7
DEVELOPMENT OF CHEMICAL TOOLS TO STUDY *O*-GLCNAC MODIFICATIONS 9
SUMMARY 13
REFERENCES 14

CHAPTER 2 18

INTRODUCTION TO THE BIOORTHOGONAL CHEMICAL REPORTER STRATEGY 19
CHEMICAL REPORTER STRATEGY: METABOLIC INCORPORATION 20
CHEMICAL REPORTER STRATEGY: CHEMICAL LABELING AND BIOORTHOGONAL REACTIONS 21
HEXOSAMINE BIOSYNTHETIC PATHWAY 23
SPECIFIC AIM OF THE PROJECT: THE APPROACH TAKEN TO DEVELOP A NOVEL CHEMICAL TOOL TO STUDY *O*-GLCNAC MODIFIED PROTEINS 27
DEVELOPMENT OF A CHEMICAL TOOL ASSOCIATED WITH DIGE 28
RESULTS AND DISCUSSION 30
CONCLUSION 49
METHODS AND MATERIALS 51
Culturing Jurkat cells 51
Metabolic labeling conditions 51
Stimulating Jurkat cells 51
Flow cytometry assay 51
Synthesis of Fluorescein alkyne probe 52
Synthesis of tetramethylrhodamine alkyne probe 52
Synthesis of Cy 3 alkyne 52
Synthesis of Cy 5 alkyne 53
Cu(I) catalyzed [3+2] cycloaddition 53
DIGE labeling 53
2D Gel electrophoresis and gel imaging 54
Post-staining and spot picking 54
Synthesis of compound (1) 55
Synthesis of Biotin N-Hydroxysuccinimide ester, compound (2) 55
Synthesis of compound (3) 56
Synthesis of Desthiobiotin- N-hydroxysuccinimide ester, compound (4) 56
Synthesis of compound (5) 57
REFERENCES 58
BIBLIOGRAPHY 61

LIST OF ABBREVIATIONS

Ac_4GlcNAz – peracetylated *N*-azidoacetylglucosamine
Ac_4GalNAz – peracetylated *N*-azidoacetylgalactosamine
AD - Alzheimer's disease
APP - amyloid precursor protein
BAP - biotin pentylamine
BEMAD - β elimination followed by Michael addition
Cy Dye – cyanine dye
DIGE - differential gel electrophoresis
DTT - dithiothreitol
eIF - eukaryotic initiation factor
eNOS - nitric oxide synthase
FITC – fluorescein isothiocyante
Gal T - galactosyltransferase
GalNAc - *N*-acetylgalactosamine
GalNAz – *N*-azidoacetylgalactosamine
GlcNAz - *N*-azidoacetylglucosamine
GlcNAz - *N*-azidoacetylglucosamine
GFAT - glutamine:fructose-6-phosphate amidotransferase
GNK - GlcNAc kinase
HBP - hexosamine biosynthetic pathway
MALDI-MS – Matrix assisted laser desorption/ionization
NMR – nuclear magnetic resonance
PMA - phorbol myristate ester
PTM - post-translational modifications
UDP - GlcNAc - Uridine diphospho-*N*-acetylglucosamine
WGA - wheat germ agglutinin

LIST OF FIGURES

FIGURE 1.1 O-GLCNACYLATION CYCLE. 4
FIGURE 1.2 BEMAD STRATEGY....10
FIGURE 1.3. GAL T REPORTER TAG....11
FIGURE 2.1. BIOORTHOGONAL CHEMICAL REPORTER STRATEGY. 19
FIGURE 2.2. BIOORTHOGONAL REACTIONS OF AZIDES....22
FIGURE 2.3. THE DE NOVO HEXOASMINE BIOSYNTHETIC PATHWAY AND THE SALVAGE PATHWAY. 24
FIGURE 2.4. THE GALACTOSAMINE SALVAGE PATHWAY. 26
FIGURE 2.5. SCHEMATIC OF TRADITIONAL DIGE EXPERIMENT 29
FIGURE 2.6. SCHEMATIC OF DIGE ASSOCIATED NOVEL CHEMICAL TOOL....30
SCHEME 1. SYNTHESIS OF AC_4GALNAZ....31
SCHEME 2. SYNTHESIS OF THE A) FLUORESCEIN ALKYNE AND B) TETRAMETHYLRHODAMINE ALKYNE PROBES. 32
FIGURE 2.7. RESULTS OF FLOW CYTOMETRY ASSAY....33
SCHEME 3. A GENERAL SCHEME FOR THE CLICK REACTION BETWEEN GALNAZ LABELED CELL LYSATE AND FLUORESCENT-ALKYNE PROBE. 34
FIGURE 2.8. RESULTS FROM THE FLUORESCEIN/ TAMRA DIGE EXPERIMENTS. 37
SCHEME 4: (A) CYANINE 3 DYE (B) SYNTHESIS OF THE CY3- ALKYNE PROBE (C) CYANINE 5 DYE (D) SYNTHESIS OF THE CY5- ALKYNE PROBE. 39
FIGURE 2.9. RESULTS FROM THE CY3/CY5 DIGE EXPERIMENTS: 41
FIGURE 2.10. SPOTS ANALYZED BY IMAGEQUANT TL SOFTWARE. 43
TABLE 1. VOLUME OF SPOTS ANALYZED AND RATIO OF ABUNDANCES FOR THE IMAGEQUANT ANALYZED SPOTS....43
FIGURE 2.10. SCHEMATIC OF AFFINITY CHROMATOGRAPHY ASSOCIATED NOVEL CHEMICAL TOOL. 45
SCHEME 5. SYNTHESIS OF COMPOUND 1....45
SCHEME 6: SYNTHESIS OF BIOTIN NHS ESTER (2). 46
SCHEME 7. SYNTHESIS OF BIOTIN-PEG-ALKYNE PROBE (3)....47
SCHEME 8. SYNTHESIS OF THE DESTHIOBIOTIN-PEG-ALKYNE PROBE (5)....47

Acknowledgments

My experience here at Stony Book University was certainly rewarding. I first want to acknowledge my advisor Dr. Isaac Carrico. Thank you for your enthusiasm and energy, as well as the tremendous encouragement and guidance you have given me. Secondly, I want to thank my committee, Dr. Nicole Sampson and Dr. Dan Raleigh, for taking time to read my thesis and for the positive experience I had during the first meeting and master's defense.

Dr. Mike White, thank you for the great advice and guidance. Thank you for being understanding and supportive. Furthermore, I want to thank Katherine Hughes for taking care of so many administrative tasks for me, for always being available and for all the advice.

It has been a pleasure working with my lab mates. Partha Banarjee, Lakshmi Rajaram, Juah Chung, Yanjie Chu, and Yoon Ohm, thank you for the intellectual conversations, troubleshooting experiments, lunch and coffee breaks, and sharing each others cultural backgrounds.

Days spent at Stony Brook would have never been the same, if not for all the friends I have made. Kellie Morgan, Paul Calderone, Dan Wilson, Matt Calder, Janine Bogaro, Kanishk Kapilashrami, Lucienne Buannic, Carl Machutta, Ed Melief, and Allison Haigney, thank you for making each day here at Stony Brook lively and cheerful. Thank you for all the great times we've had and most important keeping me sane. Carl, Ed, and Kanishk thank you for all the troubleshooting ideas and intellectual discussions.

Finally, I am grateful to my parents, for their unwavering love and support. Mom, you are truly the wind beneath my wings. Thank you dad for making sure

everything is more than fine. Thank you both for letting me know that come what may I have a loving family to fall onto. I must thank my sister for being supportive throughout the years, I miss you. Last but not least, I thank my grand parents for bringing me up with tender care, your love and commitment, and molding me with moral values. I am truly blessed to belong to such a loving and caring family, thank you.

Chapter 1

Introduction and Background to *O*-GlcNAc Modifications and Chemical Tools Used to Detect *O*-GlcNAc Modified Proteins

Introduction

In the post-genomic era, understanding the physiological impact of post-translational modifications (PTMs) is increasingly important. These modifications, which may alter physical and chemical properties of the proteins, are difficult to study as they are non-coded, often substoichiometric and transient. Glycosylation, the addition of one or more sugar moieties to proteins, represents one of the most common forms of PTMs. Most glycoproteins can be split into two categories: *N*- linked and *O*- linked. The *N*- linked glycans are attached to Asn residues. Immature *N*- linked glycans are mounted on a target protein by a single enzyme that recognize the consensus sequence Asn-Xaa-Ser/Thr (Xaa any residue but Pro) and then processed by glycosyltransferase to produce a mature glycan. In contrast, the predominant form of *O*- linked glycosylation, the mucin type, is add to the protein by an initial *N*-acetylgalactosamine (GalNAc) residue being linked to the hydroxyl group of a Ser or Thr residue; thereafter, a sugar moiety at a time is added to form a mature glycan. Unlike *N*- linked glycosylation, *O*- linked glycosylation does not have a consensus sequence (*1*). *N*- and mucin type *O*- linked glycosylations of proteins are involved in a number of physiological processes, which include signaling, protein folding, trafficking, function as structural components and ligand receptors (*2*).

In 1984, Torres and Hart discovered a new post-translational modification, *O*-GlcNAc. It was considered novel for two reasons; compared to the traditional glycoproteins, *O*-GlcNAc was not extended into a more complex structure, and it had a nucleocytoplasmic distribution (*3*). This novel PTM is not well understood; however, the studies done on *O*-GlcNAc modifications so far have shown that it is an important PTM

and could be responsible for physiological processes and many diseases. Studies have shown that changes in *O*-GlcNAc modifications are associated with diseases that include diabetes, Alzheimer's disease, and cancer. Hence, *O*-GlcNAc modifications are a vital post-translational modification to study. Certain characteristics of *O*-GlcNAcylation make this PTM inherently challenging to study. These characteristics include, *O*-GlcNAc modifications being substoichiometric, the absence of a consensus sequence, poorly understood control mechanisms, and its transient nature (*4*). Therefore, to better understand the regulation of *O*-GlcNAc modifications in cells, the development of new chemical tools is required. Developing the chemical toolbox to further investigate *O*-GlcNAc modifications will have a tremendous impact on the field of glycobiology. The overall aim of this project is to develop a novel chemical tool to further study *O*-GlcNAc modifications.

Properties and cellular functions of *O*-GlcNAc modifications

To better understand the biological function of *O*-GlcNAc, it is worthwhile to look into the biochemical characteristics of *O*-GlcNAc. Unlike some post translational modifications *O*-GlcNAc modifications have no consensus sequences, which make these modifications a challenge to study. Proteins bearing this modification are mostly localized to the cytoplasm and nucleus, but *O*-GlcNAc modified proteins are also found in mitochondrial and membrane associated proteins (*3*). Examples of known *O*-GlcNAc modified proteins include RNA polymerases, cytoskeletal proteins, nuclear oncogenes, tumor suppressors, steroid receptors, phosphatases, kinases, and kinase regulatory proteins (*5*).

O-GlcNAcylation is known to be highly dynamic and inducible. Two enzymes regulate the addition and the removal of *O*-GlcNAc from protein: *O*-GlcNAc transferase (OGT: uridine diphospho-*N*-acetylglucosamine: poly-peptide β-acetylglucosaminyltransferase) and *O*-GlcNAcase (β-*N*-acetylglucosaminidase) (*3, 4, 6*). OGT is responsible for the addition of GlcNAc to naked proteins, while *O*-GlcNAcase is responsible for the GlcNAc removal (figure 1.1) (*4, 6*).

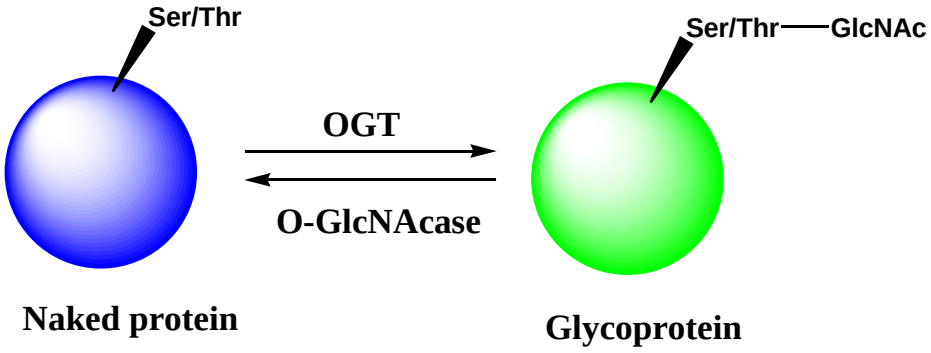


Figure 1.1 O-GlcNAcylation cycle.
OGT and *O*-GlcNAcase enzymes are involved in the addition and the removal of GlcNAc in the modification of a protein.

A recent study carried out by Shafi *et al.* demonstrated that *O*-GlcNAc modified proteins are crucial for organism viability by performing knockout of the OGT gene that resulted in embryonic stem cell lethality in mice (*7*). Such investigations suggested that addition of *O*-GlcNAc might be a dynamic regulatory modification; in other words, *O*-GlcNAc is in a constant state of flux compared to other types of glycosylations.

Different mechanisms are affected by the dynamic addition of *O*-GlcNAc to proteins. These mechanisms include mediating protein phosphorylation, altering protein degradation, and regulating the localization of protein (*3*). A limited number of cases show that *O*-GlcNAc and *O*-phosphate are mapped on the same residue. Such examples include nitric oxide synthase (eNOS) (*8*), c-myc (*9*), RNA polymerase II (*10*), and estrogen receptor β (*11*). By occupying the Ser and Thr residues *O*-GlcNAc would change the pattern of available sites for phosphorylation, resulting in changes in protein activity (*3*). The insulin resistance models studied by Burnett *et al.* demonstrated that the levels of *O*-GlcNAc increased on eNOS while the phosphorylation of Ser-1177 were reduced (*3, 8*). Hence, these studies provide circumstantial evidence that *O*-GlcNAc modifications are dynamic and can be reciprocal to phosphorylation.

Protein degradation has a critical role in regulating cellular homeostasis. This is regulated through destruction of cell-cycle regulators, anti-apoptotic proteins, transcription factors, and tumor suppressors. *O*-GlcNAcylation alters protein degradation by modifying the targeting of some proteins to the proteasome or by modifying the activity of the proteasome (*3*). *O*-GlcNAc modifications on eukaryotic initiation factor (eIF) 2α-p67 inhibit protein degradation. Once degraded, eIF2α becomes phosphorylated, which results in transitional inhibition (*12*).

The dynamic nature of *O*-GlcNAc is involved in mediating protein localization (*3*). Studies done in *Aplyisa* neurons have shown that numerous *O*-GlcNAc modified proteins rapidly move from the cytoplasm to the nucleus (*13*). Many studies have shown that *O*-GlcNAc plays a role in nuclear localization.

The above mentioned examples illustrate that *O*-GlcNAc is an important modification and exemplify its significance in cellular functions. *O*-GlcNAc may play a role in signaling cellular events akin to phosphorylation. For a protein modification to be involved in signal cascades it should be dynamic and it should be inducible by specific stimuli. *O*-GlcNAc satisfies these considerations and in combination with phosphorylation can provide control of protein function and activity.

O-GlcNAc modifications and diseases

Previous studies have shown that changes in *O*-GlcNAc levels are associated with various diseases including diabetes, Alzheimer's disease, and cancer (*14*). Two major hallmarks of diabetes are insulin resistance and hyperglycemia (*15*). Increased levels in UDP-GlcNAc due to hyperglycemia, results in an increase in *O*-GlcNAc levels that are thought to lead to insulin resistance (*16*). Insulin resistance by high glucose levels in adipocytes can be blocked by inhibiting glutamine:fructose-6-phosphate amidotransferase (GFAT) in the hexosamine biosynthetic pathway (*15*). It has also been observed that a reciprocal interaction between *O*-GlcNAcylation and phosphorylation affecting the insulin pathway proteins. The studies indicate that increased glucose flux results in *O*-GlcNAcylation of proteins linked to insulin resistance. Some other evidence also suggest that hyperglycemia is associated with global *O*-GlcNAcylation modification (*15*).

O-GlcNAc modifications, OGT, and *O*-GlcNAcase are abundant in the brain. The impairment of the intake of glucose by the brain is a known metabolic defect in Alzheimer's disease (AD). A majority of the proteins are thought to be involved in AD are *O*-GlcNAc modified, including tau and beta- amyloid precursor protein (APP). APP is thought to be *O*-GlcNAcylated in a region that may cause degradation (*15*). In AD tau protein is hyperphosphorylated which causes tau to aggregate and form neurofibrillary tangles. *O*-GlcNAc modifications of tau negatively regulate its *O*-phosphorylation in a site specific manner. Comparison studies done between normal and AD brain tissue have shown that there is a global decrease in *O*-GlcNAc modifications in AD (*4*).

Recently, several research groups have begun to entertain the idea that *O*-GlcNAcylation plays a role in cell division and cancer. Amongst the proteins that are *O*-

GlcNAcylated, there are a number of transcription factors whose activity and degradation may be altered in response to changes in the *O*-GlcNAc level. c-Myc is a proto-oncogene product that is *O*-GlcNAc modified in its transactivation domain, a region that is mutated in many lymphomas (*17*). Fang and Miller studied the function of *O*-GlcNAc in cell division by microinjecting galactosyltransferase into stage VI Xenopus laevis oocytes. The galactosyltransferase caps the *O*-GlcNAc residues with galactose and prevents deglycosylation. This interrupts the affinity for *O*-GlcNAc specific lectins. The investigators saw a toxic effect on stage VI Xenopus laevis oocytes due to galactose capping when these oocytes were stimulated with progesterone. Fang and Miller explained that the capping appeared to disrupt aster formation during M phase of the cell cycle (*17*).

Development of chemical tools to study *O*-GlcNAc modifications

Although *O*-GlcNAc modifications were discovered 20 years ago, it is not until recently that investigators began to learn about its functional significance. Detecting and studying *O*-GlcNAc modifications has been a great challenge for multiple reasons. *O*-GlcNAc modifications are dynamic, targeted to specific sub-cellular compartments, and low abundant proteins (*18*). Furthermore, *O*-GlcNAcylations are often substoichiometric, which makes it a challenge to detect by mass spectroscopy (*4*). *O*-GlcNAc is labile; therefore, when ionized in the mass spectrometer *O*-GlcNAc is susceptible for cleavage (*4, 18*). As a result, GlcNAc specific detection methods are needed to probe the importance of this PTM.

Classical methods for detecting *O*-GlcNAc modifications include *O*-GlcNAc antibodies, wheat germ agglutinin (WGA), and radiolabeling using β-1,4-galactosyltransferase (GalT) (*18*). Although, these techniques have greatly contributed to the study of *O*-GlcNAc modifications, there are limitations, such as detection sensitivity, limited binding affinity and specificity (*18*) that demand for enhanced techniques to study *O*-GlcNAc modified proteins. Recently, investigators have developed methods that involve reporter and affinity tags, isotopic labeling, and bioorthogonal chemistry (*18*) to further explore *O*-GlcNAc modified proteins.

Chemical tools have opened up new doors for visualizing and profiling *O*-GlcNAc modified proteins. Development of chemical tools involving the use of mass spectroscopy has enabled the determination of exact glycosylation sites. BEMAD (β elimination followed by Michael addition) is one such technique (figure 2.1) (*19*). The BEMAD strategy was developed to enrich low abundance modified proteins, map the sites of modification, and identify the modified proteins. The BEMAD approach is developed so that Ser or Thr linked to *O*-GlcNAc will be replaced by the stable affinity tags dithiothreitol (DTT) or biotin pentylamine (BAP) after β elimination (*19*). These affinity tags are more stable than the *O*-GlcNAc they replaced; hence, peptides with the affinity tags will be more stable to be analyzed by mass spectrometry than the peptides capped with *O*-GlcNAc (*20*). However, BEMAD introduces a few limitations that include, non-specific mapping between *O*-GlcNAc and *O*- phosphate, resistance of some modified sites to β elimination, and potential non-specific Michael addition to unmodified sites because of relatively harsh conditions used during β elimination (*20*).

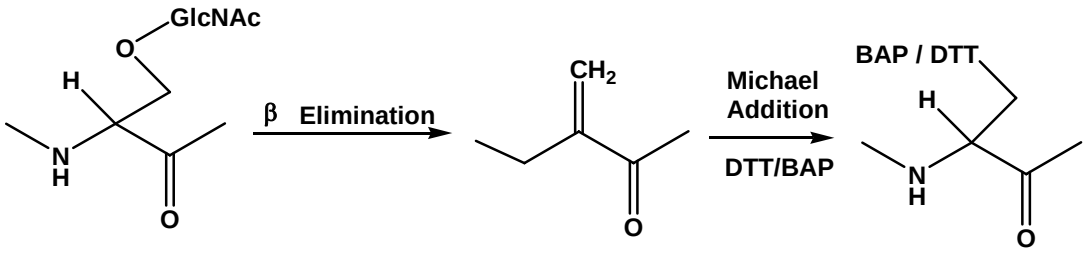


Figure 1.2 BEMAD strategy.
Scheme for DTT/BAP tagging of Post-translationally modified Ser/Thr residues using BEMAD.

Another approach that has been developed in the recent past is tagging *O*-GlcNAc modified proteins with reporter groups such as biotin (figure 3.1) (*21*). This reporter tag strategy utilizes GalT to label the C4 hydroxyl of GlcNAc with a ketone analog. Thereafter, the ketone analog reacts with the aminooxy biotin derivative (*21*). This strategy provides improvements in sensitivity and detection of proteins more efficiently than the traditional methods and eliminates false positives that may arise from methods such as antibody affinity chromatography (*18*). However, limitations of this strategy include the ketone probe is difficult to access synthetically, use of the recombinant enzyme, and complications when incorporating GalT and ketone probe to cells.

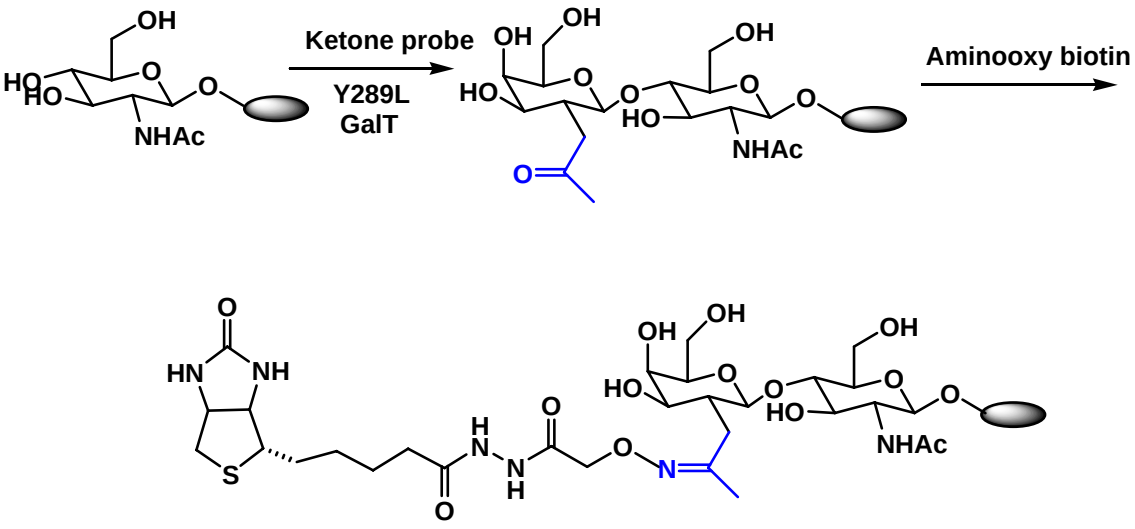


Figure 1.3. Gal T reporter tag.
O-GlcNAc-glycosylated proteins are chemoenzymatically tagged with a ketogalactose sugar, which allows selective biotinylation of the proteins.

Finally, bioorthogonal chemistry has opened new avenues in labeling and visualizing biomolecules *in vivo*. The functional group, a chemical reporter, is incorporated into the target biomolecule using cell's metabolic machinery. Subsequently, the metabolic labeling is detected with exogenously delivered probes (*22*). The bioorthogonal chemistry strategy will be discussed in detail in the subsequent chapter.

Summary

O-GlcNAc modifications were discovered in the early 1980's. It was thought to be a novel modification because *O*-GlcNAc appeared to be a non-extended modification, and was localized in the nucleocytoplasm unlike classical glycoproteins. Subsequently, it was found that *O*-GlcNAc is a dynamic, inducible, and an abundant post translational modification with no consensus sequence. It is thought to play a reciprocal role to *O*-phosphorylation in some examples. *O*-GlcNAc modified proteins are largely found in the nucleus and the cytoplasm, and may be responsible for altering cellular functions. Studies have shown that *O*-GlcNAc is responsible for up or down regulation in the transcription of several genes, mediating localization in the cell and mediating protein-protein interactions. By altering cellular functions, *O*-GlcNAc modifications can impact various diseases such as diabetes, Alzheimer's disease, and cancer. *O*-GlcNAc modifications require advanced techniques to study its specific functions. The traditional methods of detecting *O*-GlcNAc modified proteins include *O*-GlcNAc antibodies, wheat germ agglutinin (WGA), and radiolabeling using GalT. The recently developed and more advanced chemical tools include the BEMAD strategy, reporter and affinity tags, isotopic labeling, and bioorthogonal chemistry. These novel methods improved sensitivity and included steps to enrich low abundance proteins. Additionally, these techniques allow mapping modified sites and the identification of modified proteins through mass spectroscopy. Ultimatly, *O*-GlcNAc is a ubiquitous and abundant modification that needs to be explored in more depth.

References

1. Hang, H. C., Yu, C., Kato, D. L., and Bertozzi, C. R. (2003) A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation, *Proc Natl Acad Sci U S A 100*, 14846-14851.
2. Pratt, M. R., and Bertozzi, C. R. (2005) Synthetic glycopeptides and glycoproteins as tools for biology, *Chem Soc Rev 34*, 58-68.
3. Zachara, N. E., and Hart, G. W. (2006) Cell signaling, the essential role of O-GlcNAc!, *Biochim Biophys Acta 1761*, 599-617.
4. Hart, G. W., Housley, M. P., and Slawson, C. (2007) Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins, *Nature 446*, 1017-1022.
5. Hart, G. W., Kreppel, L. K., Comer, F. I., Arnold, C. S., Snow, D. M., Ye, Z., Cheng, X., DellaManna, D., Caine, D. S., Earles, B. J., Akimoto, Y., Cole, R. N., and Hayes, B. K. (1996) O-GlcNAcylation of key nuclear and cytoskeletal proteins: reciprocity with O-phosphorylation and putative roles in protein multimerization, *Glycobiology 6*, 711-716.
6. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., and Hart, G. W. (2002) Dynamic O-glycosylation of nuclear and cytosolic proteins: further characterization of the nucleocytoplasmic beta-N-acetylglucosaminidase, O-GlcNAcase, *J Biol Chem 277*, 1755-1761.
7. Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny, *Proc Natl Acad Sci U S A 97*, 5735-5739.

8. Federici, M., Menghini, R., Mauriello, A., Hribal, M. L., Ferrelli, F., Lauro, D., Sbraccia, P., Spagnoli, L. G., Sesti, G., and Lauro, R. (2002) Insulin-dependent activation of endothelial nitric oxide synthase is impaired by O-linked glycosylation modification of signaling proteins in human coronary endothelial cells, *Circulation 106*, 466-472.

9. Kamemura, K., Hayes, B. K., Comer, F. I., and Hart, G. W. (2002) Dynamic interplay between O-glycosylation and O-phosphorylation of nucleocytoplasmic proteins: alternative glycosylation/phosphorylation of THR-58, a known mutational hot spot of c-Myc in lymphomas, is regulated by mitogens, *J Biol Chem 277*, 19229-19235.

10. Comer, F. I., and Hart, G. W. (2001) Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II, *Biochemistry 40*, 7845-7852.

11. Cheng, X., Cole, R. N., Zaia, J., and Hart, G. W. (2000) Alternative O-glycosylation/O-phosphorylation of the murine estrogen receptor beta, *Biochemistry 39*, 11609-11620.

12. Datta, R., Choudhury, P., Bhattacharya, M., Soto Leon, F., Zhou, Y., and Datta, B. (2001) Protection of translation initiation factor eIF2 phosphorylation correlates with eIF2-associated glycoprotein p67 levels and requires the lysine-rich domain I of p67, *Biochimie 83*, 919-931.

13. Elliot, S. P., Schmied, R., Gabel, C. A., and Ambron, R. T. (1993) An 83 kDa O-GlcNAc-glycoprotein is found in the axoplasm and nucleus of Aplysia neurons, *J Neurosci 13*, 2424-2429.

14. Khidekel, N., Ficarro, S. B., Peters, E. C., and Hsieh-Wilson, L. C. (2004) Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain, *Proc Natl Acad Sci U S A 101*, 13132-13137.

15. Dias, W. B., and Hart, G. W. (2007) O-GlcNAc modification in diabetes and Alzheimer's disease, *Mol Biosyst 3*, 766-772.

16. Zachara, N. E., O'Donnell, N., Cheung, W. D., Mercer, J. J., Marth, J. D., and Hart, G. W. (2004) Dynamic O-GlcNAc modification of nucleocytoplasmic proteins in response to stress. A survival response of mammalian cells, *J Biol Chem 279*, 30133-30142.

17. Slawson, C., and Hart, G. W. (2003) Dynamic interplay between O-GlcNAc and O-phosphate: the sweet side of protein regulation, *Curr Opin Struct Biol 13*, 631-636.

18. Rexach, J. E., Clark, P. M., and Hsieh-Wilson, L. C. (2008) Chemical approaches to understanding O-GlcNAc glycosylation in the brain, *Nat Chem Biol 4*, 97-106.

19. Wells, L., Vosseller, K., Cole, R. N., Cronshaw, J. M., Matunis, M. J., and Hart, G. W. (2002) Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications, *Mol Cell Proteomics 1*, 791-804.

20. Vosseller, K., Hansen, K. C., Chalkley, R. J., Trinidad, J. C., Wells, L., Hart, G. W., and Burlingame, A. L. (2005) Quantitative analysis of both protein expression and serine / threonine post-translational modifications through stable isotope labeling with dithiothreitol, *Proteomics 5*, 388-398.

21. Khidekel, N., Ficarro, S. B., Clark, P. M., Bryan, M. C., Swaney, D. L., Rexach, J. E., Sun, Y. E., Coon, J. J., Peters, E. C., and Hsieh-Wilson, L. C. (2007) Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics, *Nat Chem Biol 3*, 339-348.

22. Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A., and Bertozzi, C. R. (2006) A comparative study of bioorthogonal reactions with azides, *ACS Chem Biol 1*, 644-648.

Chapter 2

Development of a Novel Chemical Tool Utilizing Bioorthogonal Chemical Reporter Strategy and 2D Differential Gel Electrophoresis

Introduction to the bioorthogonal chemical reporter strategy

As mentioned at the end of the previous chapter, bioorthogonal chemistry has opened new avenues in labeling and visualizing biomolecules *in vivo.* The bioorthogonal chemical reporter strategy consists of two steps: metabolic and chemical labeling (figure 2.1). The functional group, a chemical reporter, is incorporated into the target biomolecule using the cell's metabolic machinery (*1*). The first step, metabolic labeling, requires unnatural analogs (amino acids, monosaccharides, lipids, nucleotides, etc.) that are tolerated by the cell's natural biosynthetic machinery. These unnatural analogs must contain chemoselective chemical "handles", which will allow specific chemical labeling within the context of complex biological samples. Chemoselective chemical reactions must tolerate ambient thiols, ketone, aldehydes, and redox active metals. Finally, the stability of the introduced unnatural analogs, chemical "handles" and the resultant conjugates is critical. Inherently difficult to achieve, this combination of speed, stability, small size and limited toxicity has limited the number of accessible chemoselective ligation reactions.

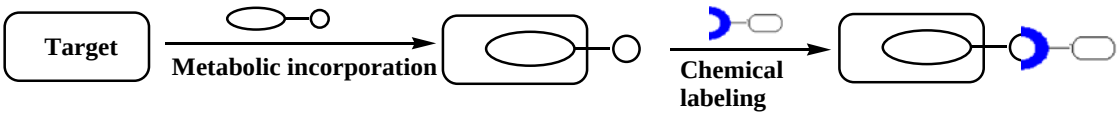


Figure 2.1. Bioorthogonal chemical reporter strategy.
The two step chemical tool utilizes metabolic incorporation followed by chemical labeling.

Chemical reporter strategy: metabolic incorporation

Metabolic incorporation of an unnatural moiety is the initial step of the bioorthogonal chemical reporter strategy. A chemical reporter must be stable in the context of ambient thiols, ketones, aldehydes, and redox active metals in the biological systems (*2*). Regardless of the biosynthetic pathway employed, the enzymes involved in the process must be able to tolerate the unnatural tag. An ideal chemical reporter would have to be small so that it would not cause any structural perturbation in the biological system. Therefore, an ideal chemical reporter must be stable, tolerant to other physiological context, and small in size.

Ketones, aldehydes, and azides are commonly used metabolic labeling tags (*2*). Ketones and aldehydes have restricted function as chemical reporters. This is a result of slow kinetics achieved at physiological pH (optimum pH is < 6). Furthermore, ketones and aldehydes are not genuinely bioorthogonal because they can be found in free sugars, pyruvate, oxaloacetate and various cofactors within the cell (*2*).

Azides on the other hand, may represent the ideal chemical reporters compared to ketones and aldehydes. Azide is the most widely used chemical reporter. This is a result of its small size, kinetic stability, lack of reactivity with natural biofunctional moieties, significant accessible potential energy, biosynthetic tolarence of azide (*2, 3*).

Chemical reporter strategy: chemical labeling and bioorthogonal reactions

The second step of the bioorthogonal chemical reporter strategy is to identify the metabolic incorporation by exogenously delivered probes (*1*). These include fluorescent, peptide, and biotinylated probes. The stability of the introduced chemical reporter tag, chemical handle and the resultant conjugate is critical. This difficult to achieve combination of speed, stability, and limited toxicity has limited the number of accessible chemoselective ligations. These features are utilized in bioorthogonal reactions, which include Staudinger ligation of azides with functionalized phosphines (*4*), strain promoted [3+2] cycloaddition using cyclooctynes that are activated by ring strain (*5*), and copper catalyzed azide-alkyne [3+2] cycloaddition reactions (*6*) (figure 2.2). The Staudinger ligation is used to covalently attach phosphine probes to the azide, which is useful for labeling whole cells and living models (*7*). One disadvantage of Staudinger ligation is that oxidation of the phosphine which may reduce the amount of available probe. Another minor drawback is the slow kinetics of the Staudinger reaction (*8*).

An alternative bioorthogonal reaction is the Cu(I) catalyzed azide-alkyne [3+2] cycloaddition, which is also known as the 'click' reaction. Notably, both reactants are stable under physiological conditions and are rarely found naturally. These reactions proceed readily in water and tolerate a wide range of functionality. The click reactions have higher kinetic rates and are fairly insensitive to pH compared to Staudinger ligation (*2*). The reaction can proceed at room temperature to provide a triazole product with 1,4-regioselectivity, and can be used to label proteins, glycans, lipids, virus particles and

nucleic acids. However, Cu(I) catalyzed [3+2] cycloaddition is unsuitable for labeling biomolecules in living organisms due to its cellular toxicity.

Another bioorthogonal reaction that can utilize azides is the strain promoted cycloaddition reactions. It is a catalyst free [3+2] cycloaddition reaction with azides that involves the use of ring strain. Strain promoted reactions are useful in labeling biomolecules in *in vivo* studies (*2*). However, the cyclooctynes are challenging to synthesize.

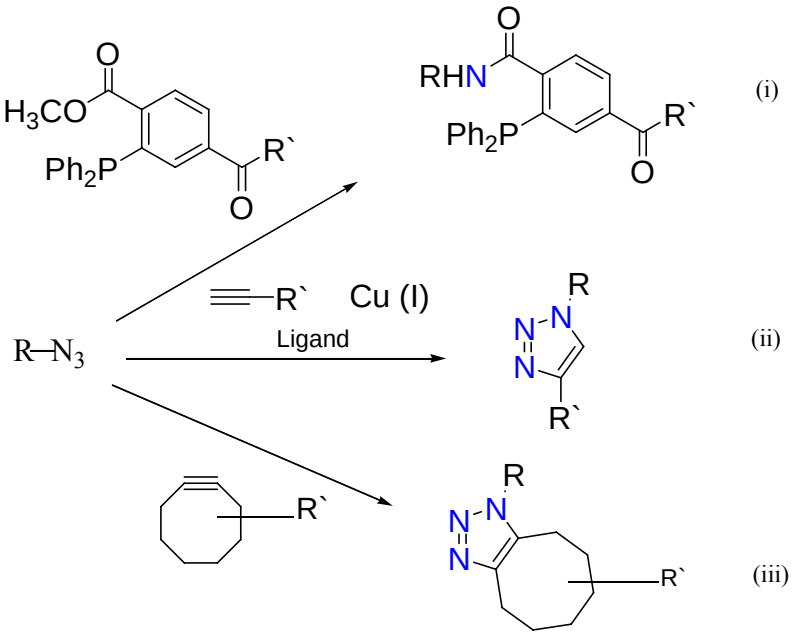


Figure 2.2. Bioorthogonal reactions of azides.
(i) Staudinger ligation (ii) click Chemistry (iii) Strain-promoted cycloaddition.

Although, the bioorthogonal chemical reporter strategy can target lipids, glycans, proteins, amino acids, etc., this project exploits monosaccharides. In the initial step, the metabolic labeling, unnatural sugars such as GalNAz are used for metabolic incorporation. The chemical labeling step is carried out utilizing "click"chemistry

reactions. GalNAz is introduced into the cell in the peracetylated form in order to enhance cell permeability.

Hexosamine biosynthetic pathway

The hexosamine biosynthetic pathway (HBP) is a branch of the glucose metabolic pathway (*9*). Figure 2.3 demonstrates the *de novo* HBP and the salvage pathway. GlcNAz can be introduced to the hexosamine salvage pathway by incorporating peracetylated *N*-azidoacetylglucosamine (Ac_4GlcNAz) into the cell. Ac_4GlcNAz is deacetylated by intracellular esterases and then enters the salvage pathway. The salvage pathway bypasses the *de novo* pathway through the action of GlcNAc kinase (GNK) to produce GlcNAc-6-P (*10*). Uridine diphospho-*N*-acetylglucosamine (UDP-GlcNAc) is the end product of both the *de novo* HBP and the salvage pathway.

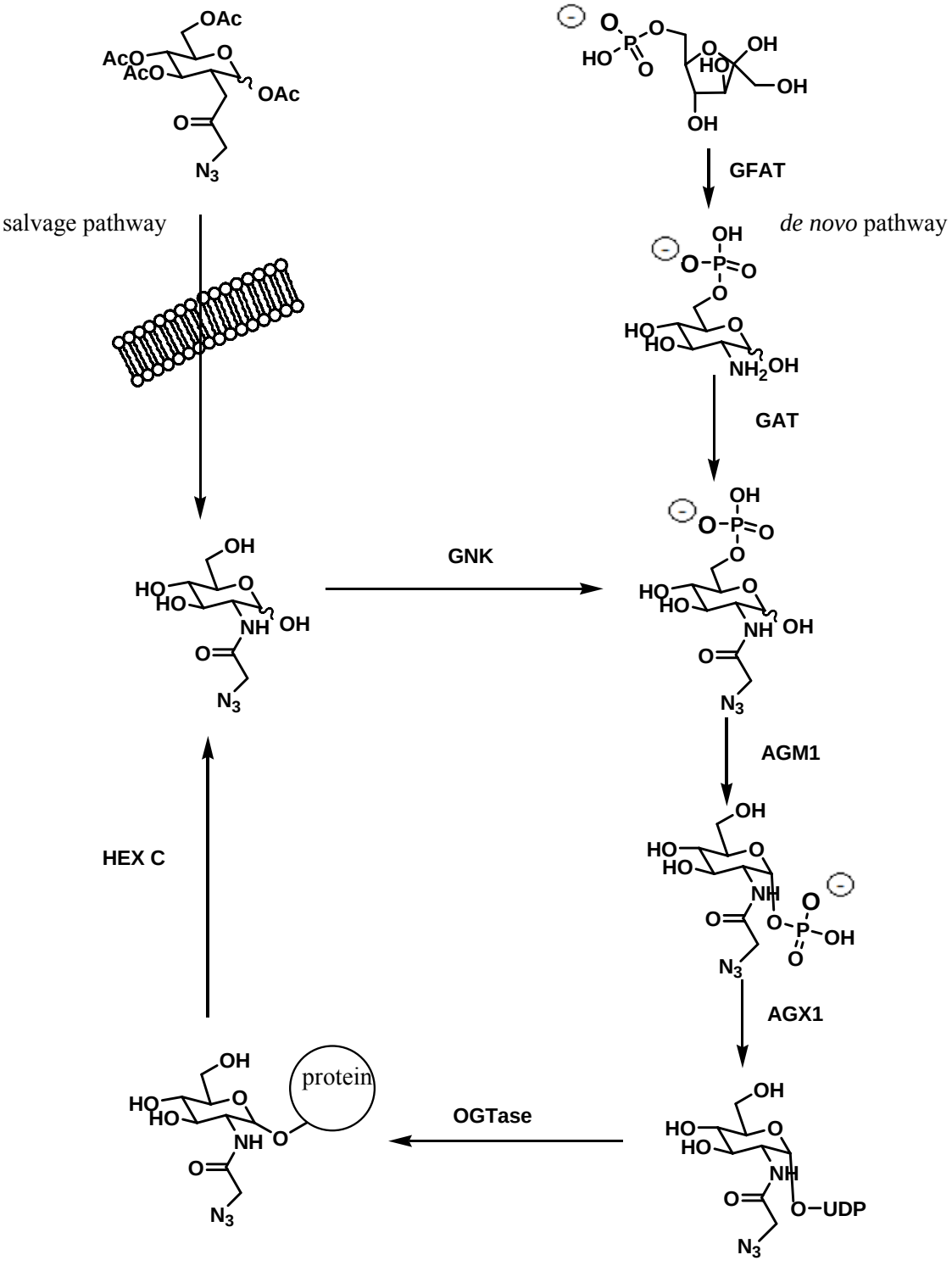


Figure 2.3. The de novo hexoasmine biosynthetic pathway and the salvage pathway.
O-GlcNAc modified proteins are biosynthesized in a series of enzymatic steps. The salvage pathway bypasses the *de novo* pathway through the action of *O*-GlcNAc Kinase and generates GlcNAc-6-Phosphate.

It was observed that GlcNAz is not efficiently accommodated by the enzymes of the salvage pathway (*11*). This drawback was overcome by exploiting the GalNAc salvage pathway (figure 2.4) (Carrico personal communication). Importantly, GalNAz was not found to be converted to any other cell surface azido sugars, such as GlcNAz or ManNAz in cells. UDP-GalNAc can be generated from GalNAc by the action of GalNAc 1-kinase and UDP-GalNAc pyrophosphorylase enzymes of the salvage pathway. UDP-GalNAc is then converted to UDP-GlcNAc through the action of UDP-GlcNAc/GalNAc C_4-epimerase (*12, 13*). The GalNAc salvage pathway has been estimated to be 400 fold more efficient than the GlcNAc salvage pathway for the production of GlcNAc modified proteins (Carrico personal communication). Hence, the GalNAc pathway is the more efficient route to monitor the *O*-GlcNAc modifications in the cell.

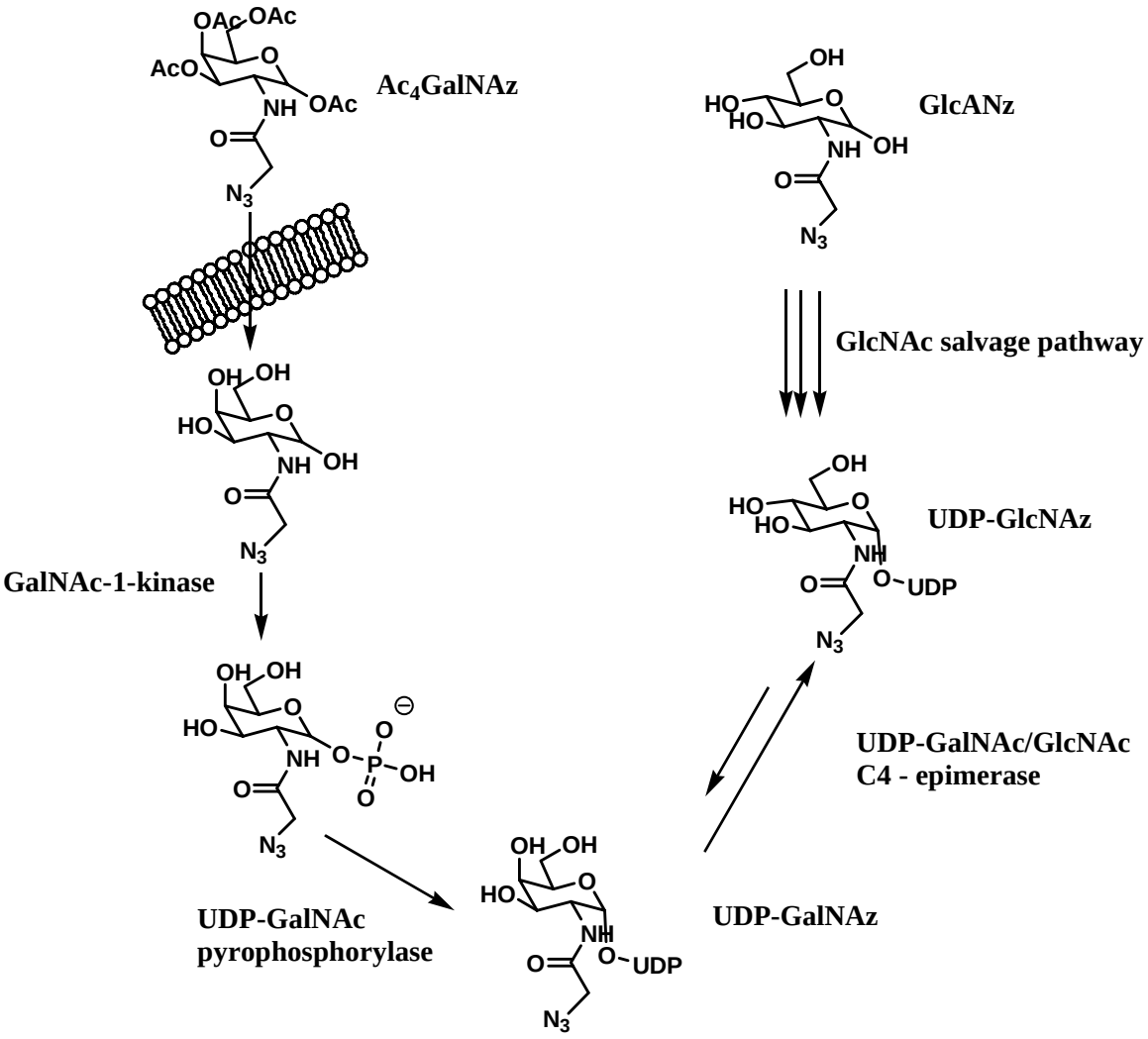


Figure 2.4. The galactosamine salvage pathway.
UDP-GalNAc can be generated from GalNAc by the action of GalNAc 1-kinase and UDP-GalNAc pyrophosphorylase enzymes of the salvage pathway. UDP-GalNAc is then generated from UDP-GlcNAc through the action of UDP-GlcNAc/GalNAc C_4-epimerase.

Specific aim of the project: the approach taken to develop a novel chemical tool to study *O*-GlcNAc modified proteins

Recently, it has been shown that the up-regulation of OGT is a key factor for T cell stimulation (*14*). Activation of transcription factors such as NFAT and NFκB occurs due to various "activities" that occur downstream of the T cell receptor. Moreover, NFAT and NFκB have been discovered to be post translationally modified with *O*-GlcNAc. An increase of *O*-GlcNAc modified NFAT and NFκB were observed shortly after T cell activation, which correlated with its translocation to the nucleus. In activated cells, NFAT is a critical factor for transcriptional induction of IL-2 and some other cytokines. Jurkat cells are immature T cells, which can be used to study the mechanism of T cell stimulation.

Development of new chemical tools to study *O*-GlcNAc modified proteins and the physiology of those proteins is essential. Expanding the existing chemical toolbox by utilizing the bioorthogonal chemical reporter strategy, which employs the GalNAc salvage pathway, is the main goal of the project. We have focused on developing a novel chemical tool exploiting the chemical reporter strategy and differential gel electrophoresis (DIGE). By introducing this novel technique, we hope to compare *O*-GlcNAc changes in proteins between different cellular states.

Development of a chemical tool associated with DIGE

2D DIGE ((Differential Gel Electrophoresis) was developed by Musfata Unlu in 1997 (*15*). It is a modification of 2D gel electrophoresis and allows a comparison of the relative abundance of proteins between two samples on a 2D gel. In the past DIGE has been useful in studying differential protein expression in breast, lung, and pancreatic cancer cells with respect to normal cells (*16*). DIGE has also been associated in developing biomarkers to examine abundance in serum protein levels (*17*).

2D DIGE is a method preferred over conventional 2D- PAGE because of various limitations of the latter. In 2D PAGE, to observe a differential protein expression a comparison of at least two gels is necessary. However, variations between gels preclude superimposability. This limitation is critical when comparing gels with significantly different spot patterns (*15*).

Classical DIGE experiments are designed as follows. Protein samples of two cellular states are pre-labeled with different fluorescent dyes that have identical charge and almost identical mass. Lysine residues of proteins are labeled with fluorescent dyes conjugated with N- hydroxysuccinimidyl esters. The variation between gels is reduced because the two samples are subjected to the same environment during 2D separation. An equal combination of both samples are labeled with a third dye as an internal standard. Normalization of internal standard across gels allows the ratio of relative expression of the same protein on a gel to be compared to another. The samples are then combined and resolved on a 2D SDS-page gel. This results in co-migration of proteins originating from separate samples. Proteins that are expressed differentially will show different levels of protein abundance that can be visualized using fluorescence imaging techniques.

Thereafter, the proteins spots are analyzed by mass spectrometry. The schematic diagram of traditional DIGE is illustrated in figure 2.5.

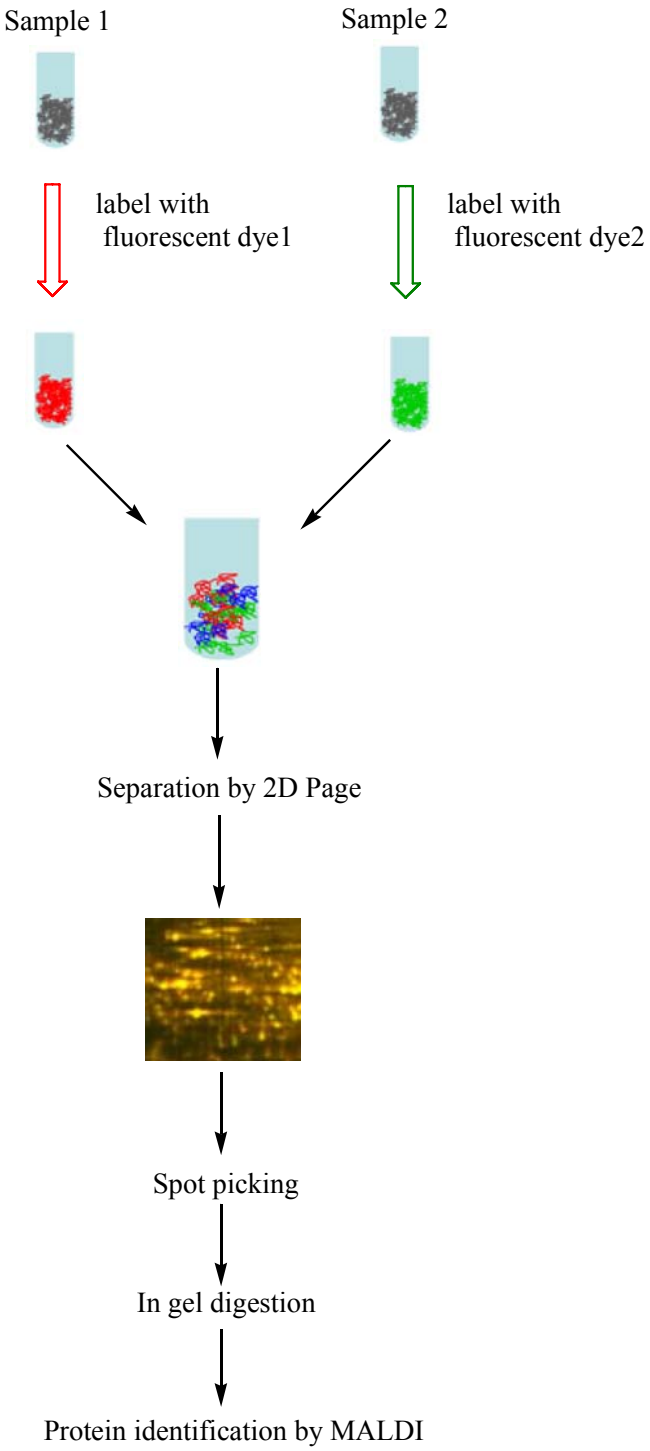


Figure 2.5. Schematic of traditional DIGE experiment

In this project the DIGE experiments were designed with variations to the classical DIGE method. Stimulated and unstimulated protein samples are first metabolically labeled with peracetylated GalNAz; thereafer, chemically labeled with two different fluorescent alkyne tags through "click" reactions. The two samples are then combined and resolved on a single 2D gel. Thereafter, the labeled *O*-GlcNAcylated proteins are visualized using a fluorescent detecting Typhoon scanner. Once the spots of interest are identified, they are excised using an automatic spot cutter. Later, these excised spots are trypsin digested and analyzed by MALDI-MS. Figure 2.6 depicts a detailed mechanism of the developed DIGE technique.

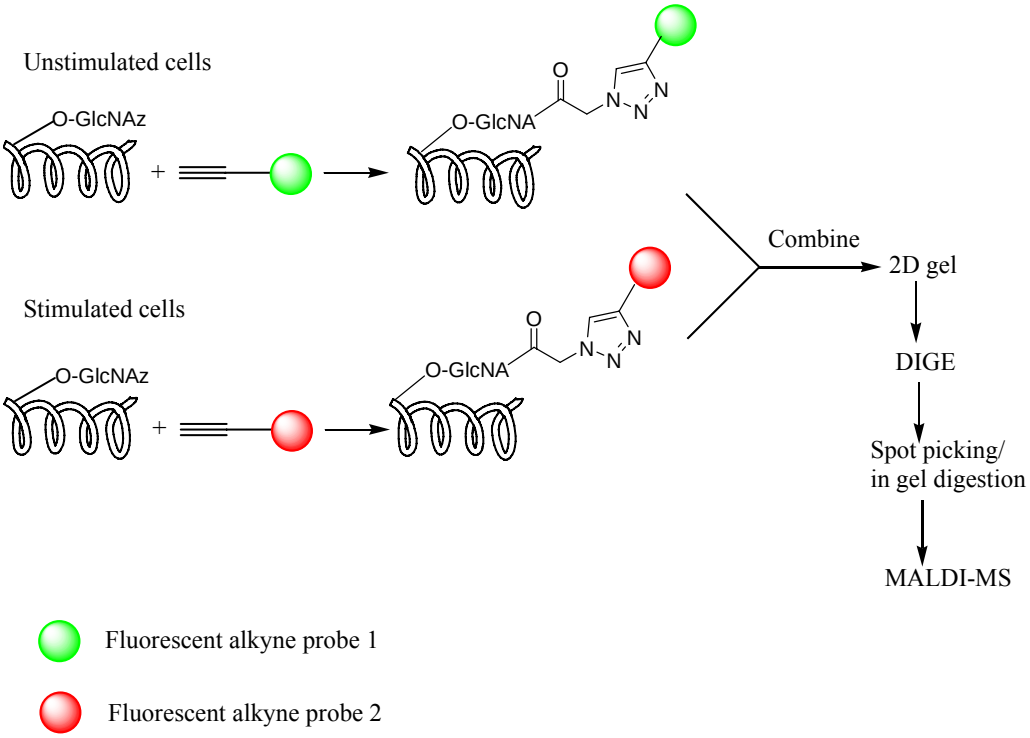
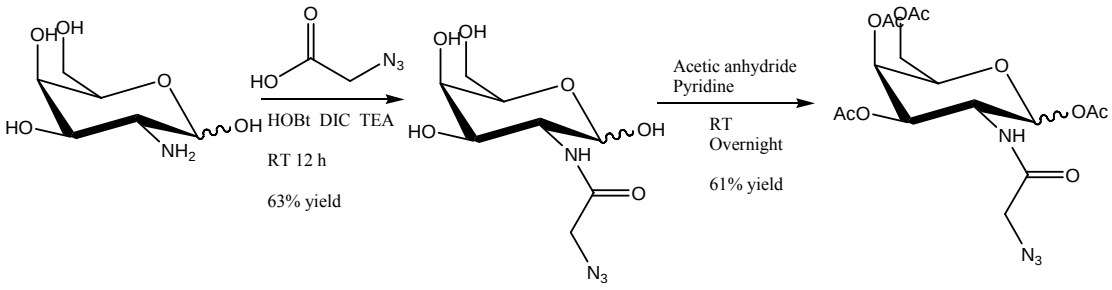


Figure 2.6. Schematic of DIGE associated novel chemical tool

Results and discussion

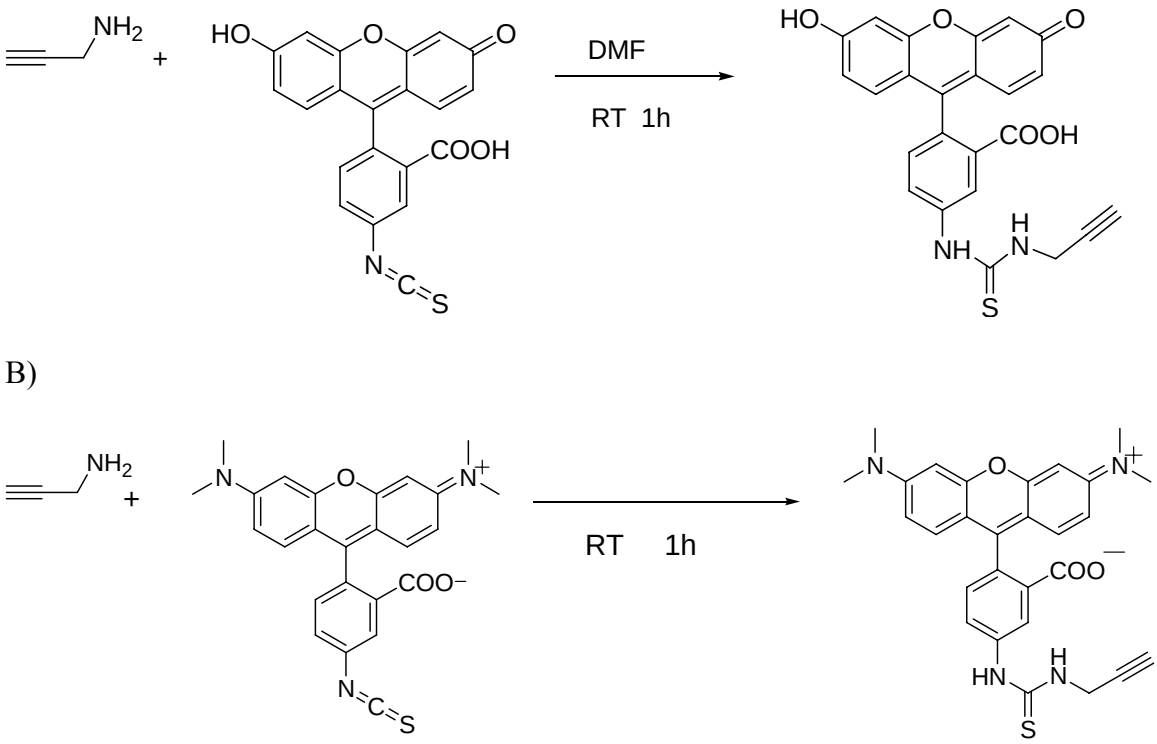
Ac_4GalNAz was synthesized as shown in scheme 1 (generously synthesized by Lakshmi Rajaram, Carrico group). Ac_4GalNAz was synthesized using galactosamine hydrochloride as starting material, which was reacted with azido acetic acid along with TEA, HOBt and DIC to produce GalNAz. The reaction was allowed to run overnight and the reaction mixture was concentrated and purified by silica gel chromatography eluting with a gradient of 10-20% methanol:dichloromethane. Thereafter, acetic anhydride was added to a solution of GalNAz in pyridine and the reaction mixture was stirred overnight at room temperature. The crude material was purified by silica gel chromatography eluting with 33% ethyl acetate: hexanes.



Scheme 1. Synthesis of Ac_4GalNAz.

The subsequent step in building up the novel chemical tool using DIGE is to synthesize fluorescent alkyne probes. Two fluorescent alkyne probes, fluorescein alkyne and tetramethylrhodamine alkyne, were synthesized as shown in scheme 2.

A)



Scheme 2. Synthesis of the A) Fluorescein alkyne and B) Tetramethylrhodamine alkyne probes.

Jurkat cells can be activated by many stimulants, which include phytohaemagglutinin, phorbol myristate ester (PMA), ionomycin, and CD28 and CD3 antibodies (*14*). A combination of PMA (10ng/ml) and ionomycin (500ng/ml) was used to activate the T cells. A PMA/ionomycin cocktail was used to activate the T cells because it is known to be a better stimulant for Jurkat cells as compared to other stimulants (*18*) and is comparatively cost effective. The cells were first labeled with 50 mM $Ac_4GalNAz$ for 15 hours. Subsequently, PMA/ionomycin was added and stimulation was allowed for 2 hours. The cocktail induces a robust activation in protease kinase C and the activation of the Jurkat cells were monitored by flow cytometry assay using fluorescein isothiocyante labeled anti CD 69. The results are depicted in figure 2.7. The activation of T lymphocytes, both *in vivo* and *in vitro*, induces expression of CD69.

This molecule, which appears to be the earliest inducible cell surface glycoprotein acquired during lymphoid activation, is involved in lymphocyte proliferation.

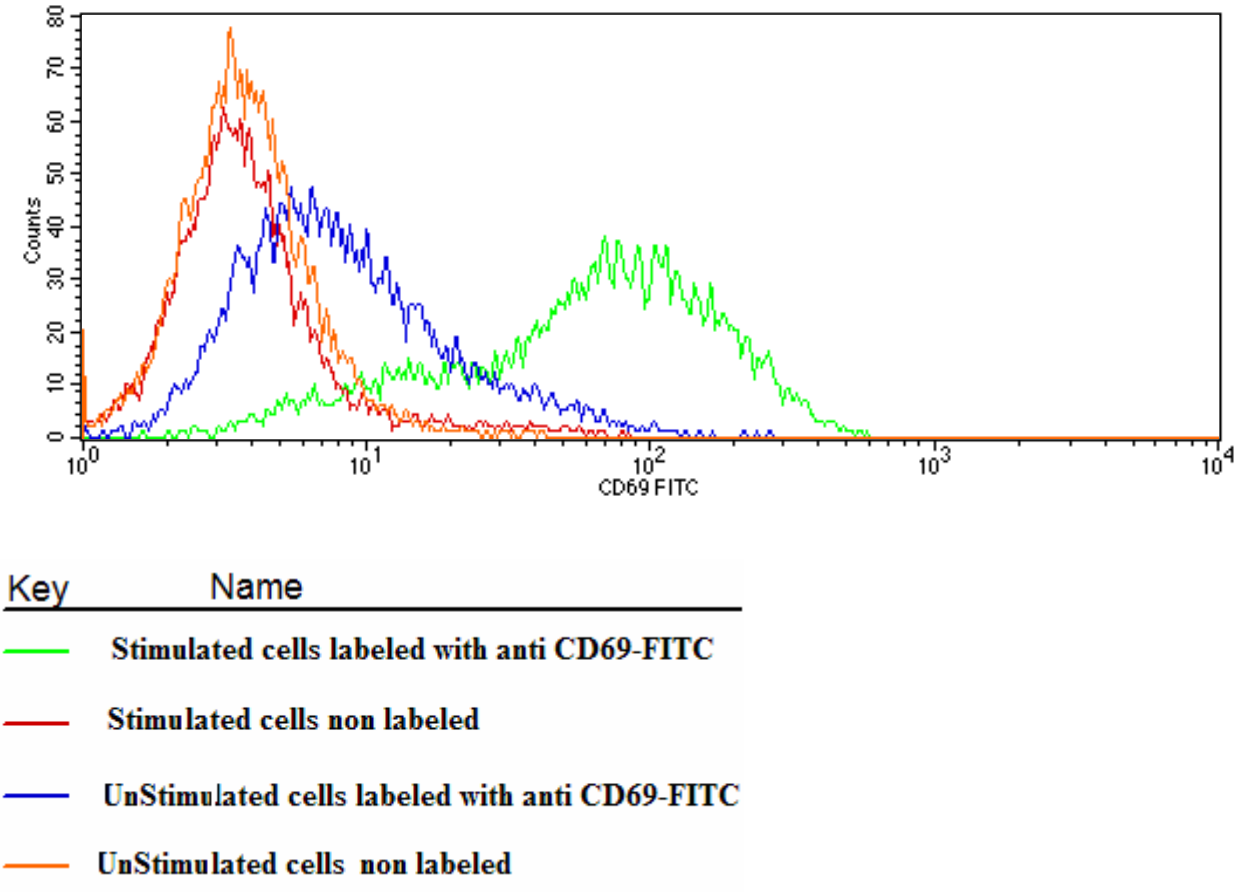
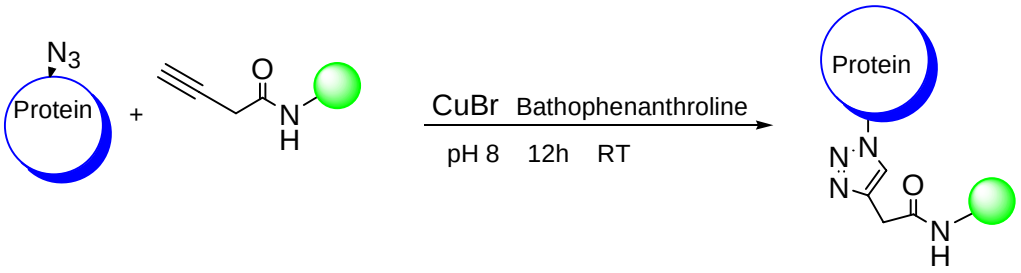


Figure 2.7. Results of flow cytometry assay

As expected the stimulated cell sample, which was labeled with anti CD69-FITC displayed the highest fluorescence intensity compared to stimulated cell sample that was not labeled with anti CD69-FITC, and the two non-stimulated cell samples. The non-stimulated cells labeled with anti CD69-FITC showed higher fluorescence intensity compared to the two unlabeled samples due to the natural induction of CD69 on the cell surface.

Each cell passed through the laser beam of the flow cytometer is known as an event. The stimulated cell sample, which is conjugated with anti CD 69- FITC showed a total count of 8254 events, an event being a single cell passing through the flow cytometer laser. Out of the total events, 88.66% were recognized to be positive events. Even though, above 80% positive events are statistically acceptable, according to previous studies the statistical data can be optimized by stimulating the cells for a shorter time (up to 1 hr) using identical PMA/ionomycin concentrations.

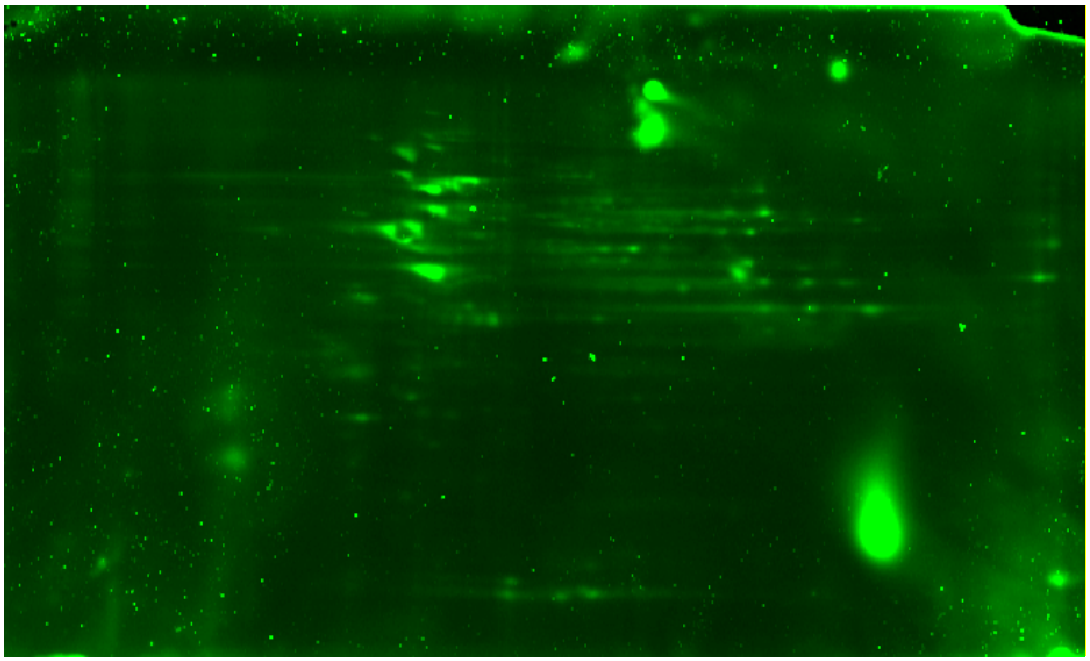
Metabolically labeled unstimulated and stimulated cell lysates were chemically labeled exploiting click chemistry with the synthesized fluorescent probes (scheme 3). Click reactions were allowed to proceed for 12 hours. Thereafter, the chemically labeled proteins are precipitated using methanol and chloroform (4: 1) to remove the excess fluorescent probe and other reagents. The metabolically and chemically labeled lysates were then combined and resolved on a single 2D gel. The gel was imaged using a Typhoon 9400 scanner exciting the fluorescent dyes at relevant excitation wavelength.



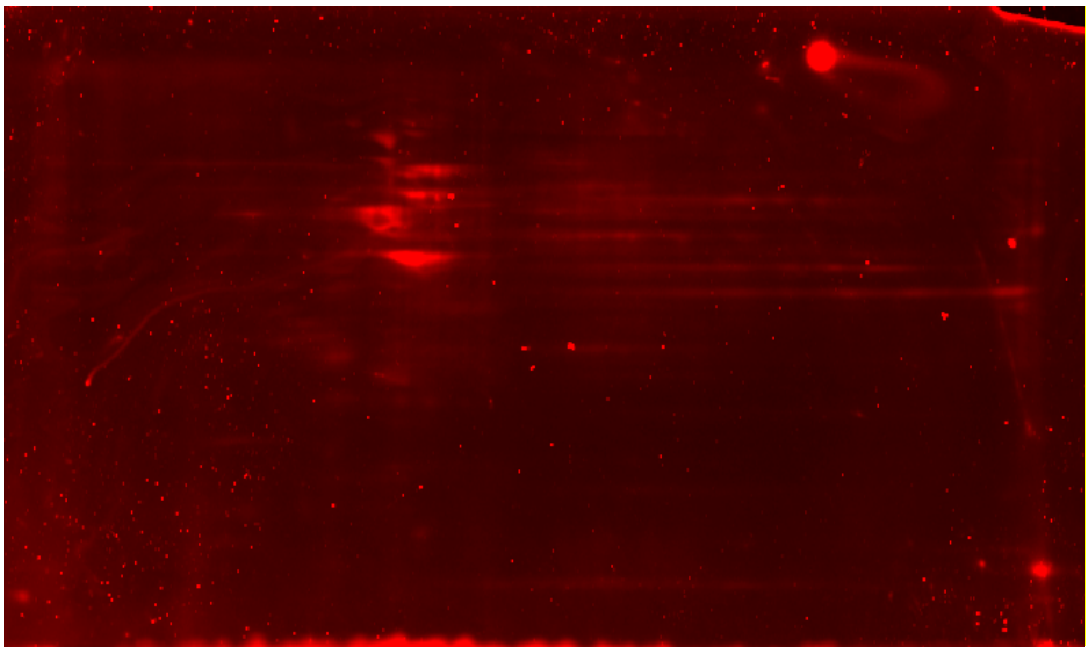
Scheme 3. A general scheme for the Click reaction between GalNAz labeled cell lysate and fluorescent-alkyne probe.

Figure 2.8 shows the results of 2D SDS- Page gel electrophoresis for the lysates labeled with the two fluorescent alkyne probes and the DIGE image. The non-stimulated cells were chemically labeled with Fluorescein-alkyne and the stimulated cells were labeled with Tetramethylrhodamine-alkyne. First, the two images were scanned independently on the Typhoon scanner. Thereafter, the two images are overlaid to produce the DIGE image. The 2D images of Fluorescein, Tetramethylrhodamine, the DIGE and the Sypro ruby stain image are shown respectively in figure 2.7.

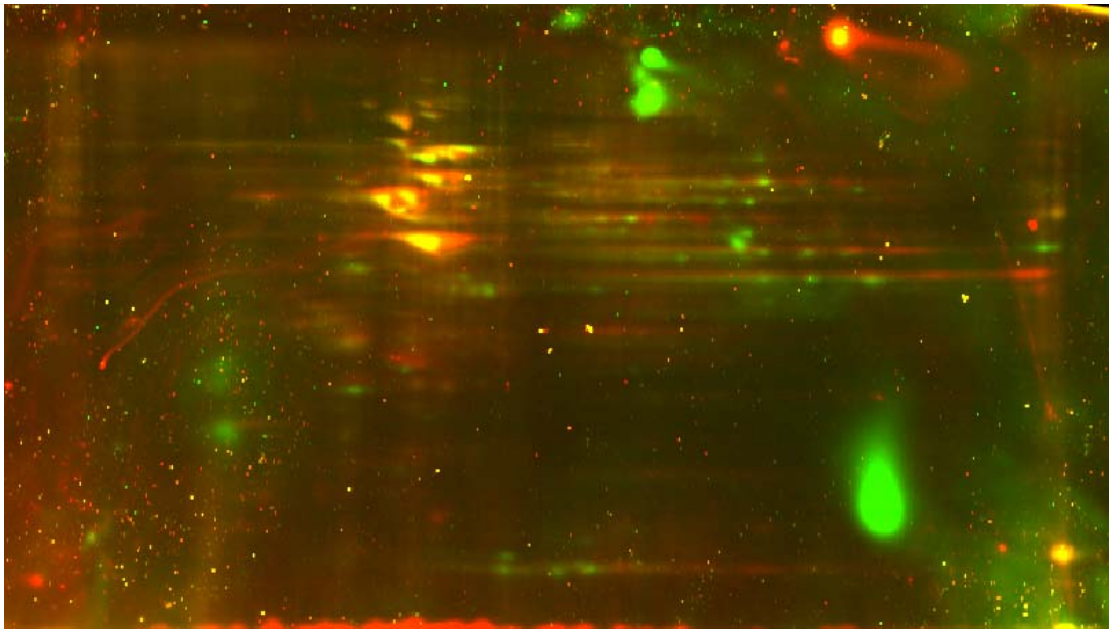
(A)



(B)



(C)



(D)

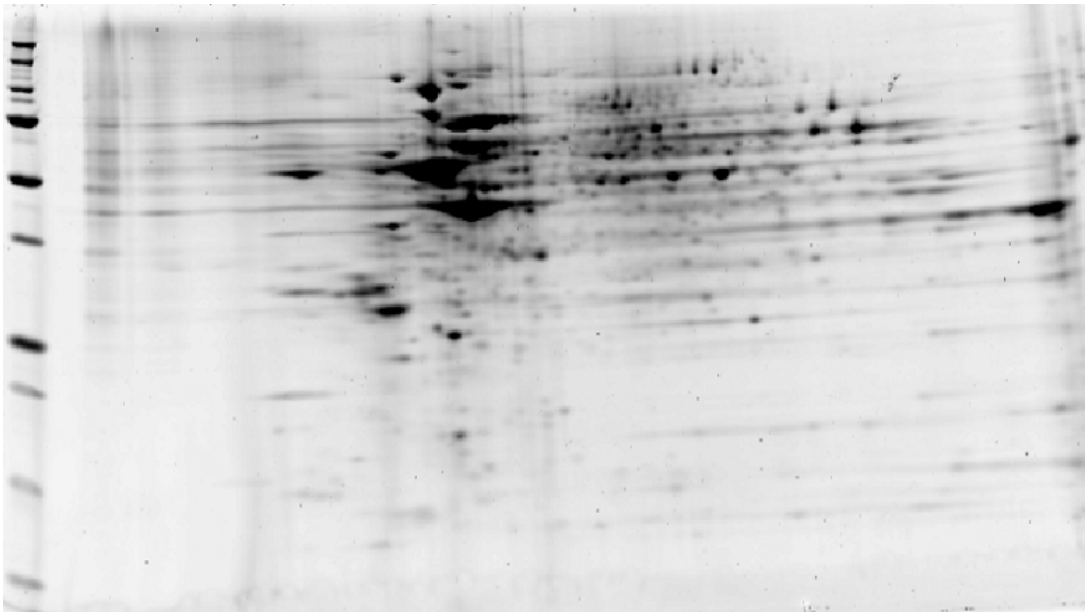


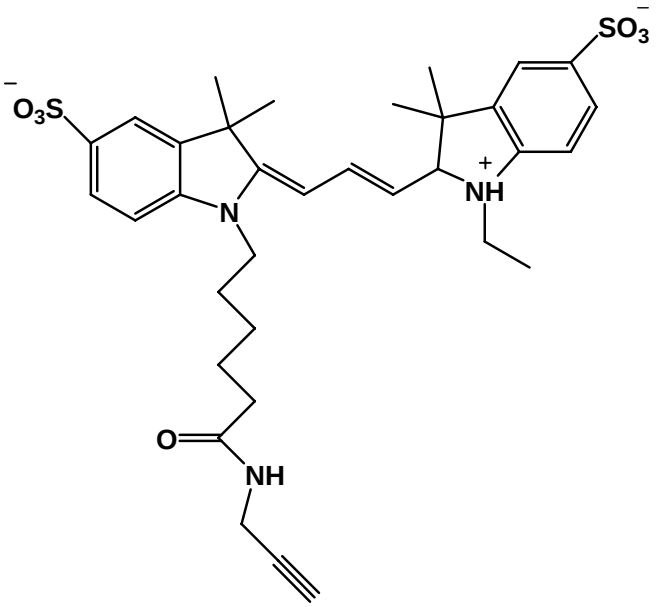
Figure 2.8. Results from the Fluorescein/ TAMRA DIGE experiments.
(A) Non stimulated Jurkat cell lysate chemically labeled with Fluorescein -alkyne (false colored in green) excitation/emission 488/526 nm (B) Stimulated Jurkat cell lysate chemically labeled with TAMRA-alkyne (false colored in red) excitation /emission 532/580 nm (C) DIGE image when images A and B are overlaid (D) The Sypro ruby image. Proteins were sseparated on pH 3-10 IEF strips in the first dimension and 4-12 % SDS page gel in the second dimension. The fluorescent label proteins were imaged on Typhoon 9400 scanner at appropriate wavelengths.

The preliminary study was conducted using Fluorescein and Tetramethylrhodamine fluorescent dye pair in order to develop and optimize experimental conditions. However, the DIGE image revealed that many spots did not exactly align. This was particularly noticeable in the low abundance proteins. It was due to the discrepancy in matched charge on the fluorescent dyes. In addition, a high level of background labeling was observed.

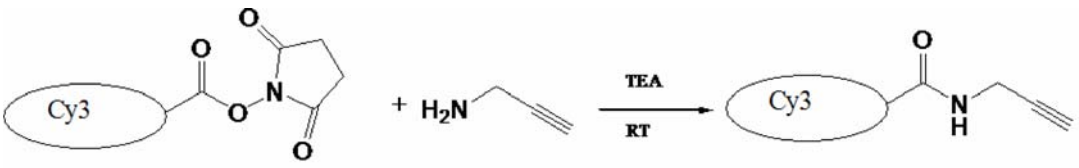
After developing and optimizing the conditions with the above fluorescent dye pair, the experiment was repeated using a novel pair of fluorescent dyes, Cyanine dyes. Cyanine dyes are ideal for DIGE experiments because of the matched charge and the

mass, water solubility, the relatively higher quantum yield and extinction coefficients, and distinguishably different excitation wavelengths. The Cy 3-alkyne and Cy 5-alkyne probes were synthesized as is shown in scheme 4.

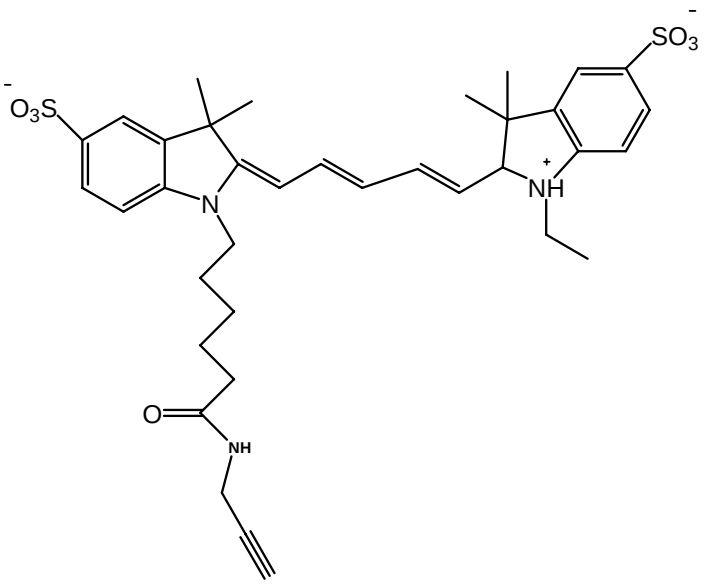
(A)



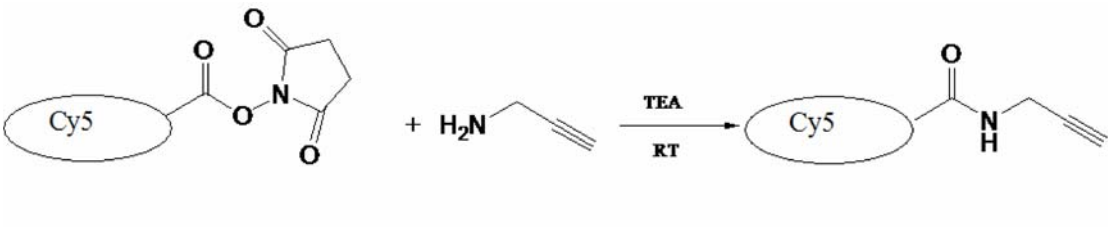
(B)



(C)



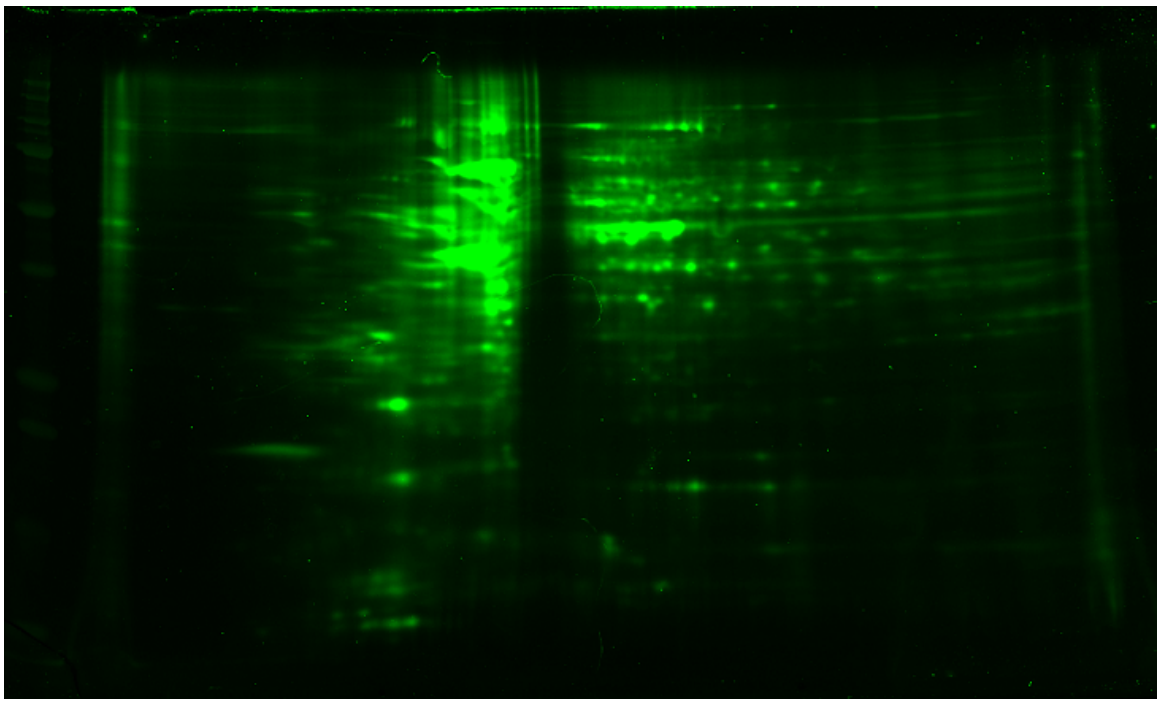
(D)



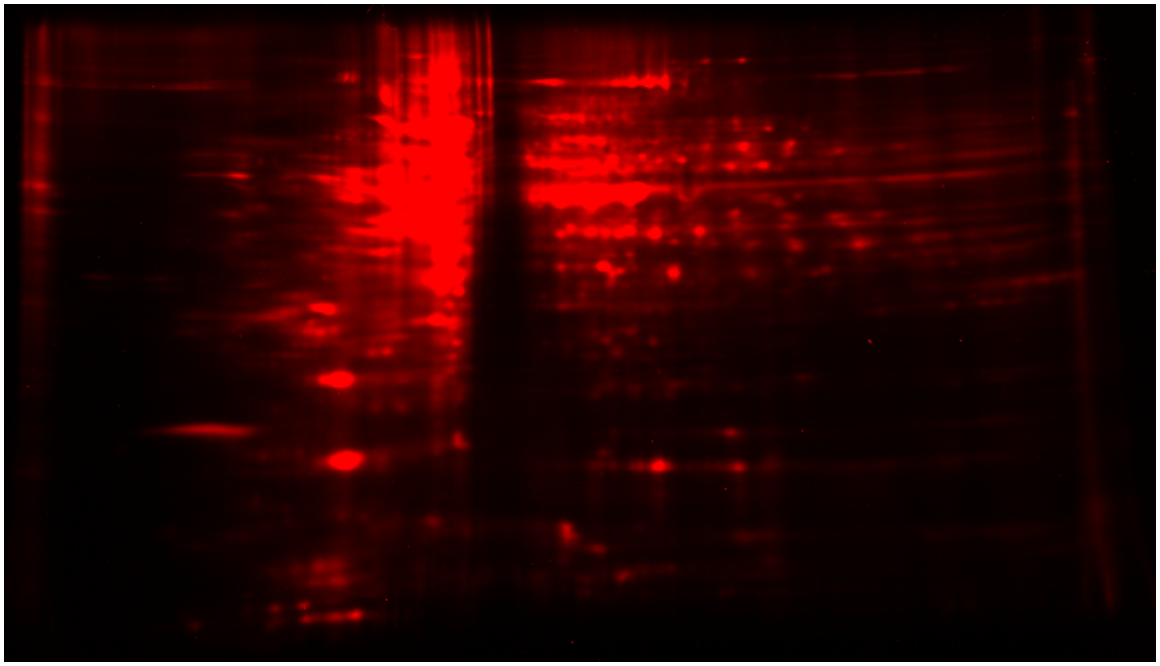
Scheme 4: (A) Cyanine 3 dye (B) Synthesis of the Cy3- alkyne probe (C) Cyanine 5 dye (D) Synthesis of the Cy5- alkyne probe.

Figure 2.8 illustrate the results of 2D SDS- Page gel electrophoresis for the samples labeled with the two Cy dye probes and the DIGE image. The non stimulated cells were chemically labeled with Cy3–alkyne (false colored in green) and the stimulated cells are labeled with Cy5–alkyne (false colored in red). The 2D images of Cy3, Cy5, the DIGE and the Sypro Ruby image are shown respectively in figure 2.8.

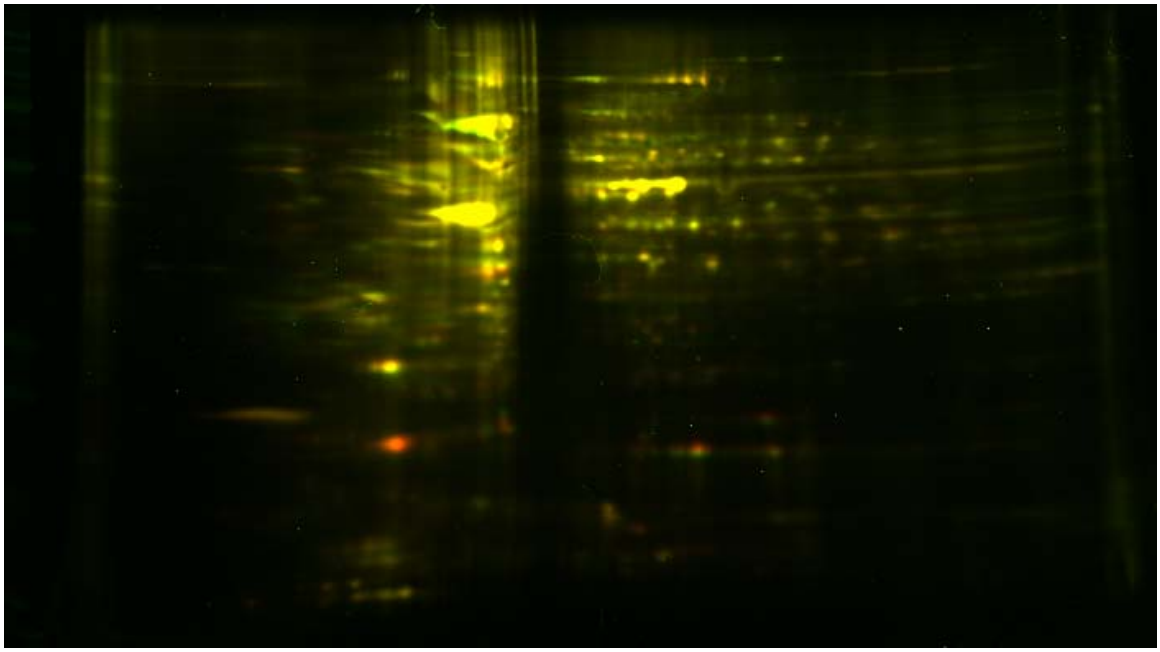
(A)



(B)



(C)



(D)

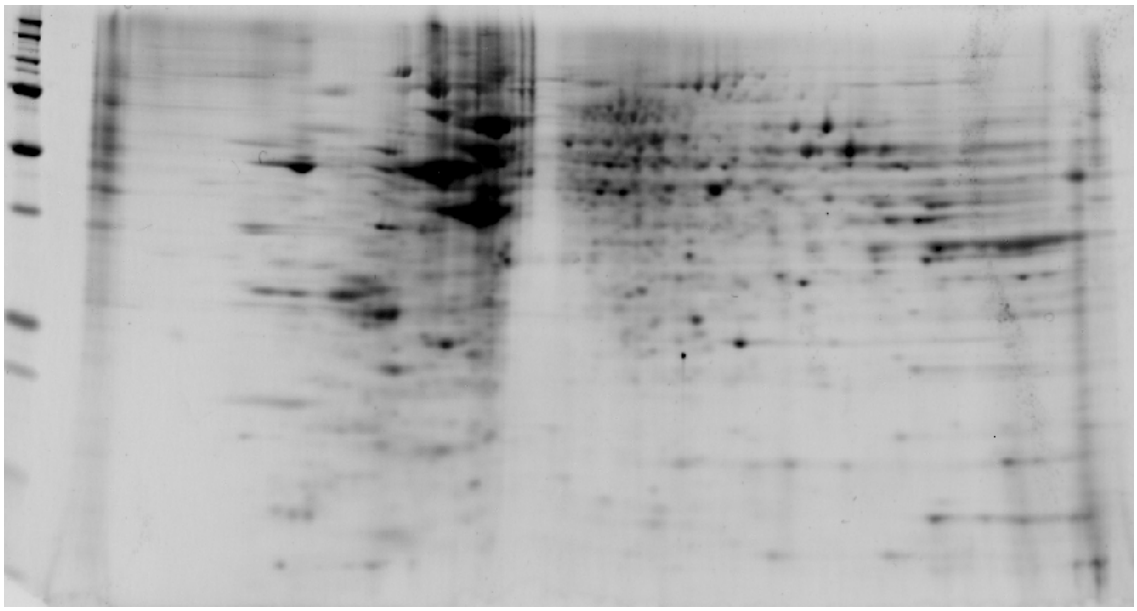


Figure 2.9. Results from the Cy3/Cy5 DIGE experiments:
(A) Non stimulated Jurkat cell lysate chemically labeled with Cy3 -alkyne (false colored in green) excitation/emission 532/580 nm (B) Stimulated Jurkat cell lysate chemically labeled with Cy5-alkyne (false colored in red) excitation /emission 633/670 nm (C) DIGE image when images A and B are overlaid (D) The Sypro ruby image.

As expected, *O*-GlcNAc proteins are expressed differentially. The highly abundant *O*-GlcNAcylated proteins common to both samples appeared yellow (false colored). A couple of low abundant proteins appeared to be either red or green (false colored), indicating a change in *O*-GlcNAcylation between the samples. The sypro ruby image demonstrates the total protein expression for both samples. Certain protein spots that appear in the syrpo ruby image does not appear in the DIGE, which indicates that the metabolic and the chemical labeling is very specific for *O*-GlcNAc modified proteins. Furthermore, the horizontal streaks in higher pI range might indicate phosphorylated proteins. Even though, Typhoon laser scanner is capable of detecting proteins at low levels, the signal from abundant proteins masked the fluorescence signal from some of the low abundance proteins.

It is known that the volume of the spots from the Cy dye images correlate with the abundance of protein (*16*). The volumes of 20 spots were obtained by using the Image quant TL software. Table 1 demonstrates the volumes of the protein spots labeled with Cy3 alkyne, Cy 5 alkyne, and the ratio of abundances. Spot A is ~ 4.7 folds more abundant in Cy 5 labeled sample, which is the stimulated sample. On the other hand, spot D is ~4.7 folds more abundant in Cy 3 labeled sample, which is the non-stimulated sample. The ratio of abundance for most analyzed spots is shown to be ~ 1:1.

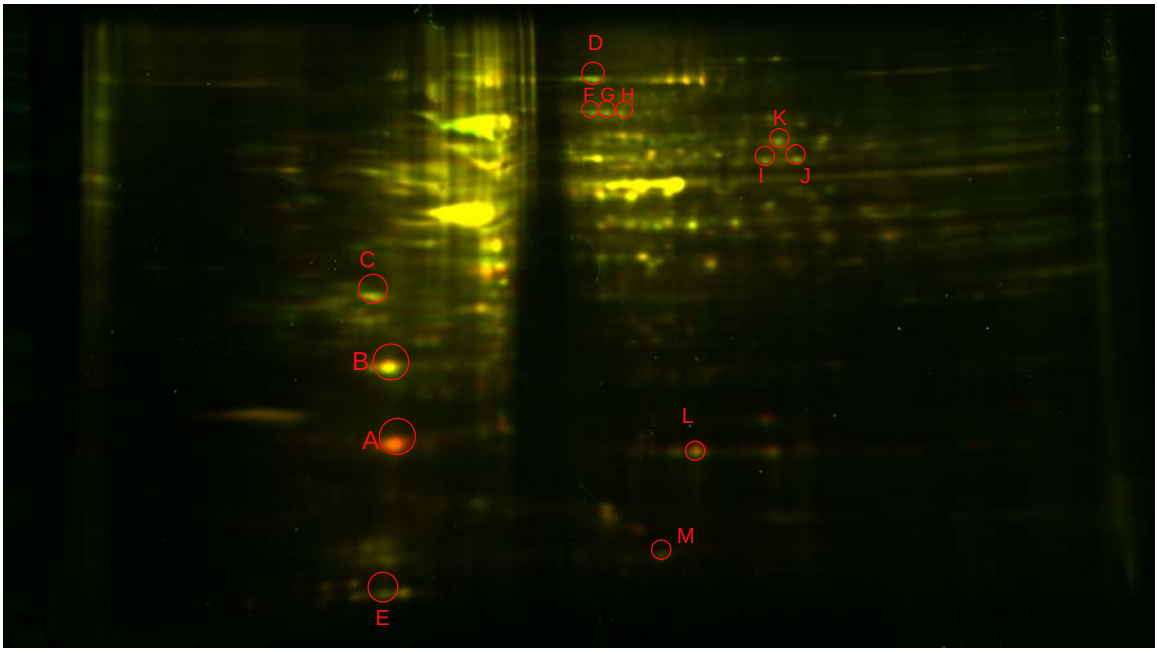


Figure 2.10. Spots analyzed by Imagequant TL software.

Label	vol. of spots (Cy3)	vol. of spots (Cy5)	ratio of abundance
A	118235.6	552907.1	4.68
B	366981.5	510997.5	1.39
C	123581.4	91083.37	0.74
D	81392.19	50563.72	0.62
E	46427.6	29780.04	0.64
F	46936.03	21695.95	0.46
G	48189.07	29241.41	0.61
H	31432.89	17268.75	0.55
I	62001.44	37505.23	0.61
J	60801.82	49929.35	0.82
K	58266.55	38823.04	0.66
L	40610.98	82392.58	2.03
M	6548.18	14546.43	2.22

Table 1. volume of spots analyzed and ratio of abundances for the Imagequant analyzed spots

Currently, MALDI-MS studies for distinctive protein spots that showed a significant difference in expression levels are underway.

As an alternative strategy for the DIGE associated chemical tool, another novel chemical tool utilizing the technique of affinity chromatography is currently under development. This technique would help in enriching the low abundance *O*-GlcNAzylated proteins, which is critical to study changes involving *O*-GlcNAc modifications. The enrichment can be achieved by immunoprecipitating the proteins of interest using avidin or streptavidin resin column chromatography. The avidin-biotin complex is a strong non-covalent interaction (Ka = 10^{15} M^{-1}) between a protein and ligand which is utilized for immunoprecipitation. Cells will be metabolically labeled and thereafter chemically labeled with the desthiobiotinylated-PEG-alkyne probe. The labeled cell lysates will be loaded onto an avidin or streptavidin column and eluted with biotin which has a higher affinity to avidin compared to desthiobiotin-PEG-alkyne probe. Once pure fractions of *O*-GlcNAc modified proteins are collected, two-dimensional gel electrophoresis can be performed to identify the proteins that are being modified. The above-mentioned technique is performed separately on both stimulated and unstimulated states. Thereafter, the two 2D gels can be compared to recognize the changes in the *O*-GlcNAc modified levels in the two different cellular stages. Figure 2.10 shows a detailed mechanism of the developed affinity technique.

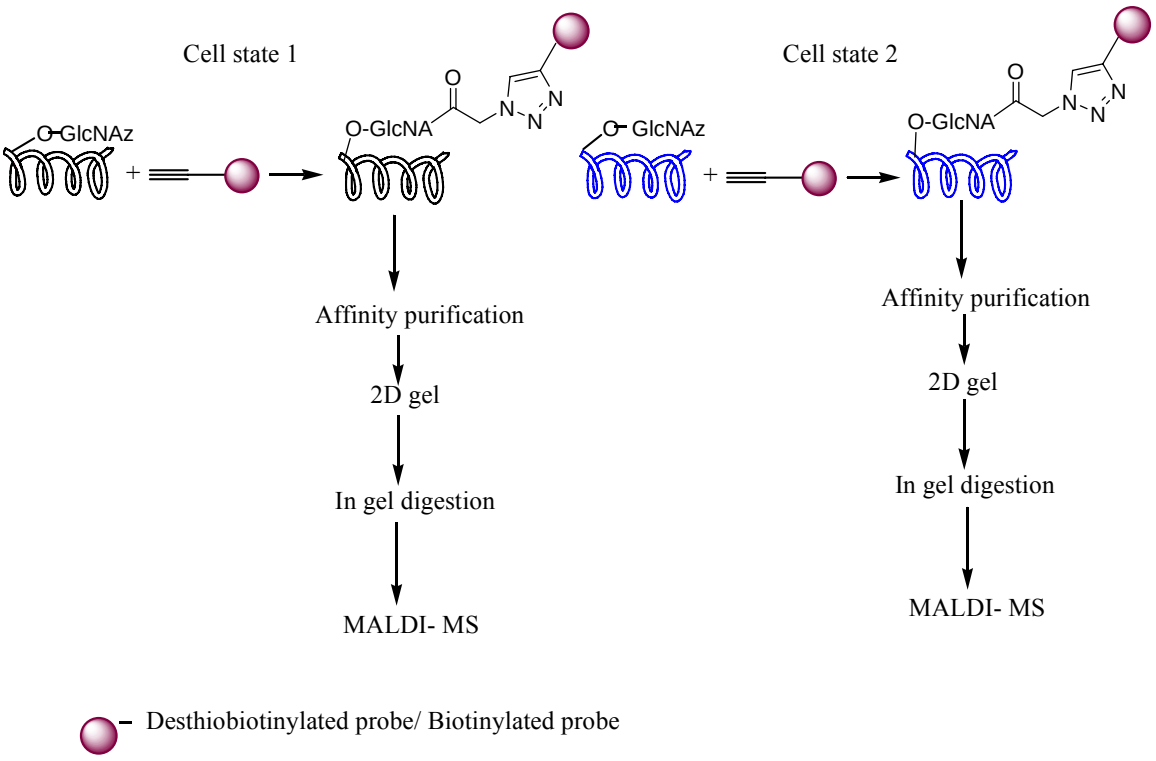
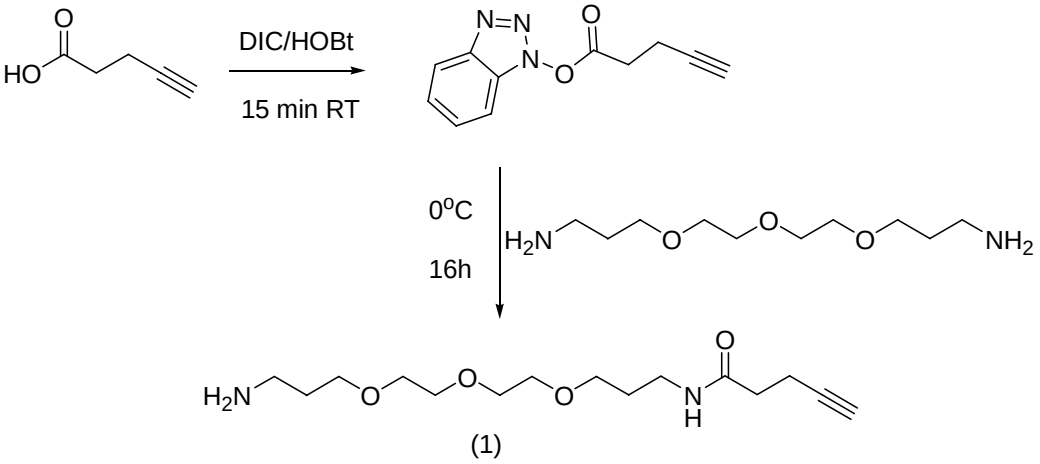


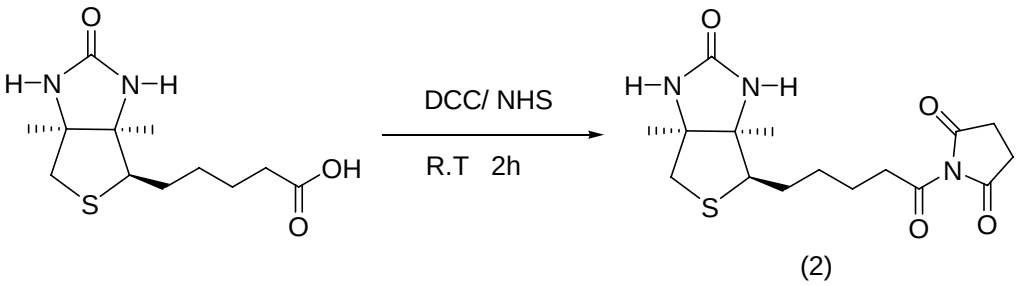
Figure 2.10. Schematic of affinity chromatography associated novel chemical tool.

As the initial step in developing second chemical tool, the necessary probes, biotin-PEG-alkyne and desthiobiotin-PEG- alkyne, were synthesized. The reaction schemes are shown below.



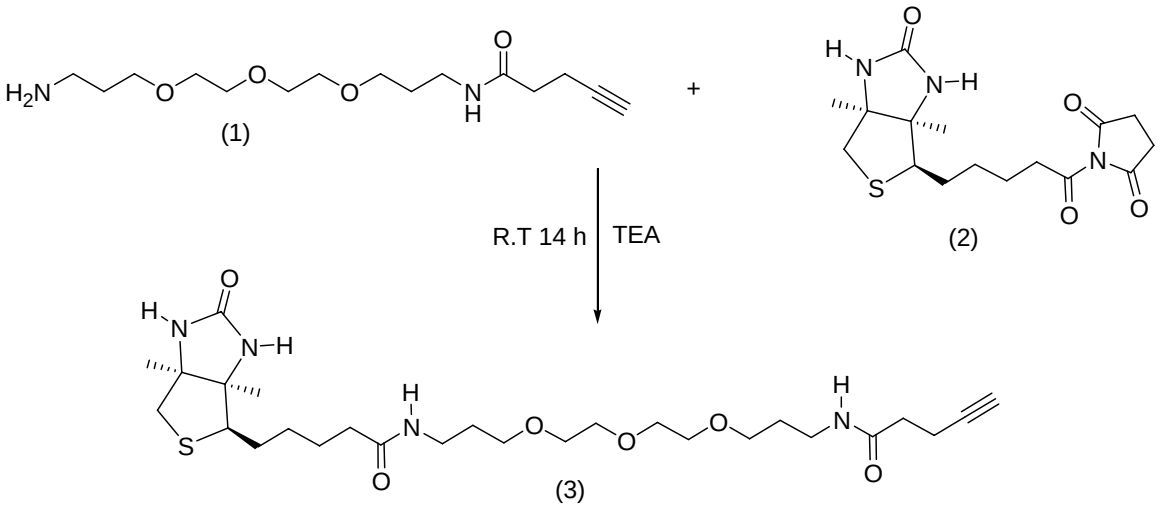
Scheme 5. Synthesis of compound 1.

The initial step in the synthesis of compound **1** involves esterification of the pentynoic acid using N,N`-diisopropylcarbodiimide (DIC) as a strong dehydrating agent. During the reaction DIC hydrates and forms N,N`-diisopropylurea (DIU). Along with DIC, Hydroxybenzotriazole (HOBt) is used for esterification reaction as a nucleophilic catalyst. Subsequently, the activated pentynoic acid ester is added to 4,7,10-trioxa-1,13-tridecanediamine. The activated ester of pentynoic acid should be added dropwise in 0^0C over an hour to minimize the forming of the bis-coupled byproduct. The reaction was stirred at room temperature for an hour. The resulting oil was dissolved in dichloromethane and loaded onto a silica column and eluted with dichloromethane:methanol (90:10 to 80:20).

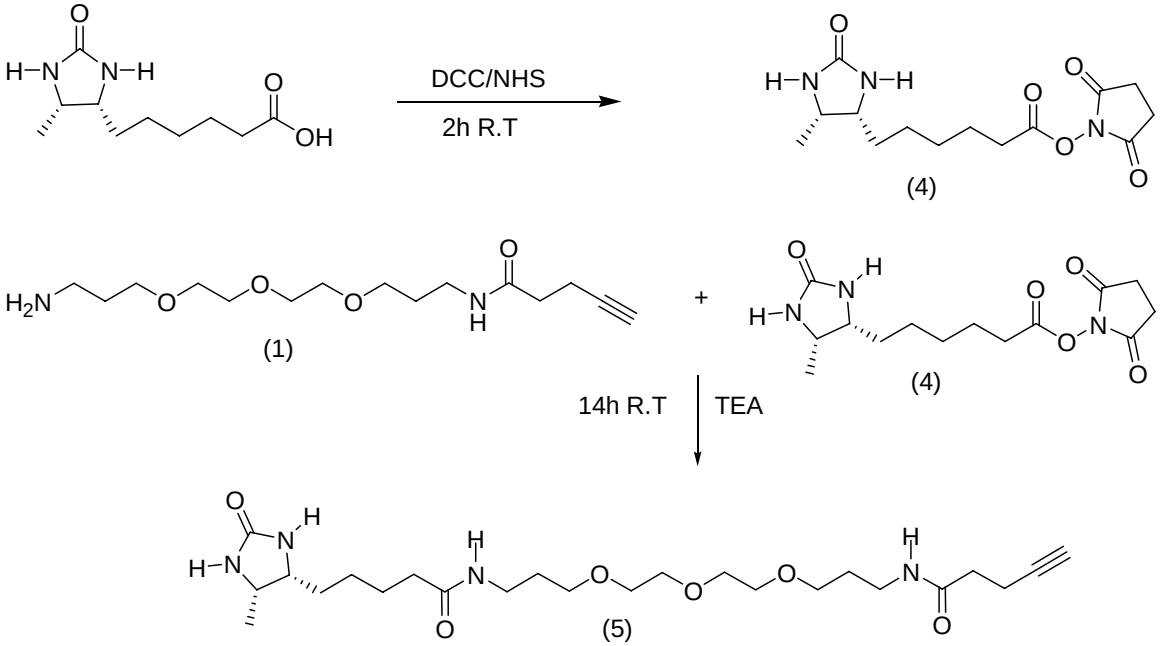


Scheme 6: Synthesis of biotin NHS ester (2).

Synthesis of Compound **2** follows the same reaction principles as described previously for pentynoic acid. Dicyclohexylcarbadiimide (DCC) was used in synthesizing biotin-NHS ester. Biotin was reacted with DCC to form a highly unstable activated acid intermediate. The NHS reacts to form a less labile activated acid, biotin - OSU. After the mixture was stirred for 2 hours, the precipitate was removed by filtration and the final product was precipitated with diethylether. The isolated product was dried under vacuum to yield (81%) of Biotin-NHS ester as a white solid.



Scheme 7. Synthesis of Biotin-PEG-alkyne probe (3)



Scheme 8. Synthesis of the desthiobiotin-PEG-alkyne probe (5)

In order to synthesize compound **3**, biotinylated-PEG-alkyne probe, compound **1** was added to biotin-OSU, in the presence of triethylamine (TEA). A polyethylene glycol (PEG) chain, which is highly water soluble, is attached between the biotin and pentynoic ester in order to allow incorporation of water insoluble biotin into cells. PEGylation enables easy incorporation of the probe into eukaryotic cells.

The desthiobiotin-PEG-alkyne probe was synthesized in a similar manner (Scheme 8). Activated desthiobiotin was synthesized under the same conditions used in synthesizing biotin-OSU. The reaction yielded a 30% pure product, which is a similar to previously reported yield (37%). In order to synthesize compound **5**, compound **1** was added to desthiobiotin-OSU in the presence of TEA.

Conclusion

Post-translational modifications are important for the structure and function of various proteins. Studying post-translational modifications (PTM) has been a challenge in the post-genomic era since they are non-coded, often substoichiometric and transient. *O*-GlcNAcylation of proteins is an interesting post-translational modification because it is thought to play a role in disease states such as, diabetes, Alzheimer's disease and cancer. It was found that *O*-GlcNAc is a dynamic, inducible, and an abundant PTM lacking any apparent consensus sequence. The dynamic character of *O*-GlcNAc is significant for cell regulation, cytoskeletal structure and modulation of protein function. This PTM is thought to play a reciprocal role to *O*-phosphorylation. Chemical tools including isotopic labeling, BEMAD, and the bioorthogonal chemical reporter strategy are utilized to study *O*-GlcNAcylated proteins. Although, these chemical tools have been useful to study *O*-GlcNAc modifications, there is a need for more sophisticated chemical tools due to the complex nature of these modifications.

This project focused on developing a novel chemical tool, which is associated with the chemical reporter strategy and DIGE. The chemical reporter strategy consists of two steps: metabolic labeling and chemical labeling. The biosynthetic machinery native to the cell is exploited to carry out the metabolic labeling, while the chemical labeling is done in a bioorthogonal fashion with click reactions. DIGE provides the ability to compare differential protein expression between two cellular states. Compared to traditional 2 D electrophoresis, DIGE provides increased reproducibility and eliminates gel to gel variations when comparing the protein expression of two samples.

The novel chemical tool is a variation of traditional DIGE was developed using GalNAz for metabolic labeling and fluorescent alkyne probes for chemical labeling. Thereafter, proteins from the two samples are combined and resolved on a 2D gel. The *O*-GlcNAcylated proteins are then imaged using a Typhoon laser scanner.

The preliminary results using Cy3 and Cy5 dyes shows that the metabolic and the chemical labeling are specific to *O*-GlcNAcylated proteins. Proteins that are expressed differentially in the two samples showed different levels of abundance and migration patterns on the DIGE image. These results were also confirmed by the statistical analysis done by using the Imagequant software. Presently, MALDI-MS studies are underway to study proteins that showed a significant difference in expression levels. In future, we plan to extrapolate our results and utilize the said chemical tool to study *O*-GlcNAc modifications in diseased cell lines.

Methods and materials

Culturing Jurkat cells

Jurkat cells (Cell Culture/Hybridoma Facility, SUNY Stony Brook) were maintained in a 5% CO_2, water saturated atmosphere at 37^0C, and grown in RPMI 1640 supplemented with 10% Fetal Bovine Serum and 2% penicillin/streptomycin. Cell densities were maintained between 1 x 10^5 and 2.5 x 10^5 cells/ml.

Metabolic labeling conditions

50µl Aliquots of 50 mM peracetylated *N*-azidoacetylgalactosamine (Ac_4GalNAz) (generously synthesized by Lakshmi Rajaram, Carrico group) in ethanol were delivered into 200 ml culture flasks. Once the ethanol was evaporated, media and Jurkat cells were seeded in the flasks at the density of 2.5 x 10^5 cells/ml. The cells were incubated at 37^0C for 2-3 days till the cell density reach 1 x 10^6 cells/ml.

Stimulating Jurkat cells

Jurkat cells at a density of 1 x 10^6 cells/ml were metabolically labeled with Ac_4GalNAz for 15 hours. Thereafter, 10µl of PMA (1mg/ml) and 25 µl of ionomycin (1mg/ml) were added to the culture flasks. The stimulations were allowed for 2 hours.

Flow cytometry assay

The cell concentrations of stimulated and unstimulated samples were adjusted to 1 million cells/ ml with culture media. 1 ml of cell suspension was placed into each tube and centrifuge at 250g for 5 minutes. The supernatant was removed and precipitate the cells 2-3 times with FACS buffer. The pellet from the final wash was suspended in 50 µl of FACS buffer. Add 20 µl of anti CD69-FITC antibody and mixed gently. Incubate the

reaction mixture on ice for 30 minutes. The cells were washed 2-3 times with FACS buffer and suspend in 500 µl of FACS bufffer for analysis.

Synthesis of Fluorescein alkyne probe

0.1g of fluorescein-5-isothiocyanate was mixed with 10 ml of methanol. 20 µl of propargylamine was added to this mixture and let react for 2 hours with protecting from the light. Once the reaction reached completion, the solvent was evaporated using rotary evaporation *in vacuo.* The isolated product yielded 91% as a yellow powder. ESI calculated for $C_{24}H_{16}N_2O_5S$ $[M+H]^+$ 444.1, found 445.1

Synthesis of tetramethylrhodamine alkyne probe

500µl of tetramethylrhodamine -5-isothiocyanate (10mM) was mixed with 10 ml of methanol. 0.79 µl of propargylamine was added to this mixture and let react for 2 hours with protecting from the light. Once the reaction reached completion, the solvent was evaporated using rotary evaporation *in vacuo.* The isolated product yielded 95% as a magenta colored powder. ESI calculated for $C_{28}H_{26}N_4O_3S$ $[M+H]^+$ 498.2, found 499.2

Synthesis of Cy 3 alkyne

Cy3-NHS (1mg) was dissolved in DMF. Propargyl amine (20.13 µl) was added to Cy3 in DMF and triethylamine (22.05 µl) was added to the reaction mixture and the reaction was stirred for 14 hours. Once the reaction reached completion, the solvent was evaporated using rotary evaporation *in vacuo.* The isolated product yielded 90% as a pink colored film. Rf = 0.48 (acetonitrile:acetic acid:water 50:5:45) ESI calculated for $C_{34}H_{42}N_3O_7S_2$ $[M+H]^+$ 667.1, found 668.1.

Synthesis of Cy 5 alkyne

Cy5-NHS (1mg) was dissolved in DMF. Propargyl amine (20.13 µl) was added to Cy5 in DMF and triethylamine (22.05 µl) was added to the reaction mixture and the reaction was stirred for 14 hours. Once the reaction reached completion, the solvent was evaporated using rotary evaporation *in vacuo.* The isolated product yielded 90% as a blue colored film. Rf = 0.54 (acetonitrile:acetic acid:water 50:5:45) ESI calculated for $C_{36}H_{44}N_3O_7S_2$ $[M+H]^+$ 693.1, found 694.1.

Cu(I) catalyzed [3+2] cycloaddition

The click reactions were set up as follows. GalNAz labeled Cell lysates reacted either with fluorescein alkyne (10nM), tetramethylrhodamine alkyne (10nM), or biotin-PEG-alkyne probe. The reactions were performed under anaerobic conditions, with Tris buffer at pH 8, bathophenanthroline (3mM) and CuBr (1mM). The reaction was allowed to proceed for 12- 15 hours in room temperature (*19*).

DIGE labeling

After the metabolic labeling the cells were lysed following a standard protocol [19]. Thereafter, the samples were chemically labeled with Cy 3 alkyne and Cy 5 alkyne utilizing click reactions which as outlined above. After the click reactions, 600 µl of methanol, 150 µl of chloroform and 400 µl of water were added to the reaction mixture and vortexed briefly followed by centrifugation for 5 minutes at 13,000g. The upper aqueous phase was carefully removed and discarded while leaving the interface layer containing the protein precipitate intact. 450 µl of methanol was added to the tube and the process was repeated. The tube was covered with lint free tissue and tube cap kept open

for 15 minutes. Thereafter, the cap was replaced and stored in -20 ^{0}C until ready for 2D electrophoresis.

2D Gel electrophoresis and gel imaging

An aliquot of 75 µg of protein from each labeled sample was precipitated and resolublized in SDS-page loading buffer. Immobilized non linear pH gradient (IPG) strips, pH 3-10 were rehydrated with Cy dye labeled samples in the dark at room temperature overnight. Isoelectric focusing was performed using a Multiphor 11 apparatus for a total of 80kV-h at 20 ^{o}C, 10mA. Strips were equilibrated for 15 min in 50 mM Tris-HCl, pH 8.8, 6M urea, 30% glycerol, 1% SDS containing 65 mM dithiothreitol and then for 15 min in the same buffer containing 240mM iodoacetamide. Equilibrated IPG strips were transferred onto 4-12 % gradient polyacrylamide gels. Gels were run in Protean 11 gel tanks at 30mA per gel at 10 ^{o}C until the dye front run off the bottom of the gels. 2D gels were scanned directly on the Typhoon 9400 laser scanner. The laser scanner has 4 lasers at different wavelengths enabling scanning at the different wavelengths specific for each of the Cy dyes.

Post-staining and spot picking

The gels were fixed in 30 % methanol, 7.5% acetic acid overnight and washed in water and total protein was detected by spot staining with Sypro Ruby dye for 3 hours at room temperature. Excess dye was removed by washing twice in water and the gels were imaged at the appropriate excitation and emission wavelengths for the stain. Spots of interest were excised from 2D gels using an automated spot picker. Spots were collected in 200 µl of water in 96- well plates and kept frozen at -20 ^{o}C for protein identification by MALDI-MS.

Synthesis of compound (1)

Pentynoic acid (0.5g, 5.1mmol) was added to a dry round-bottom flask containing anhydrous N,N Dimethylformamine (DMF) (5ml). Hydroxybenzotriazole (HOBt) (0.78g, 5.1mmol) and N,N`-diisopropylcarbodiimide (DIC) (0.8ml, 5.1mmol) were added subsequently. 4,7,10-trioxa-1,13-tridecanediamine (3.37ml, 15.3mmol) was added to another dry flask with acetonitrile (10 ml). Both flasks were cooled to 0-5 ^{0}C in ice water baths. After 30 minutes, the Pentynoic ester was added to the tridecanediamine drop wise over an hour. The reaction was stirred at room temperature for an hour. The precipitate was removed by filtration. The filtrate was rotary evaporated *in vacuo*. The resulting oil was dissolved in dichloromethane and loaded onto a silica column. The column was eluted with dichloromethane:methanol (90:10 to 80:20). The isolated product was dried under vacuum to yield 61% of product as a white solid. Rf = 0.39 (dichloromethane:methanol 70:30). ^{1}H NMR (300 MHz, CD_3OD)δ 3.64 (s, 8 H), 3.60 (t, 4 H), 3.50 (t, 2 H), 3.10 (t, 2 H), 2.82 (t,1H), 2.50-2.31 (m, 4 H), 1.98 -1.88(m, 2 H), 1.80-1.70 (m, 2 H). ESI calculated for $C_{15}H_{28}N_{14}O_{16}S$ $[M+H]^{+}$ 300.2, found 302.0

Synthesis of Biotin N-Hydroxysuccinimide ester, compound (2)

D-Biotin (0.5g, 2.04mmol) was dissolved in 25 ml of hot (70^{0}C) DMF under argon. The solution was cooled down to room temperature. Subsequently, N-Hydroxysuccinimide (NHS) and Dicyclohexylcarbadiimide (DCC) were added to the solution. After the mixture was stirred for 2 hours, the precipitate was removed by filtration. The volume was reduced to about 10 ml by applying vacuum at 50^{0}C. Diethylether was added into the residual solution to precipitate the product. The filtered

product was washed successively with diethylether and 2-proponol. The isolated product was dried under vacuum to yield 0.403 g (81%) of Biotin-NHS ester as a white solid. ^{1}H NMR (300 MHz, DMSO) δ 5.60 (s, 1 H), 5.50 (s, 1H), 4.46(m, 1 H), 4.27 (m, 2 H), 3.14 (m, 1 H), 2.8 (dd, 1 H), 2.75 (d, 1 H),2.5 (s,4H), 2.29 (t, 2 H), 1.73 (m, 1 H), 1.64 (m, 3 H), 1.46 (m, 2 H).

Synthesis of compound (3)

Compound **1** was dissolved in anhydrous DMF. The biotin-NHS ester was dissolved in DMF. Compound **1** was added to the biotin NHS ester. Triethylamine was added to the solution and the reaction was stirred for 14 hours. Purification was done by silica gel chromatography (dichloromethane:methanol 90:10).The product, biotinylated-PEG-alkyl probe, was a white powder (62% yield). R_f = 0.34 (dichloromethane:methanol 80:20)

^{1}H NMR (300 MHz, CD_3CN) δ 6.77 (br s, 2 H), 5.74 (s, 1 H), 5.38 (s, 1 H), 4.44 (t, 1 H), 4.27 (m, 1 H), 3.61-3.54 (m, 8 H), 3.51 (t, 4 H), 3.25- 3.16 (m, 5 H), 2.91 (dd, 1 H), 2.68 (d, 1 H), 2.47 (m, 2 H), 2.35-2.30 (m, 6 H), 2.22 (t, 1 H), 2.17 (t, 2 H), 1.75-1.53 (m, 8 H), 1.44-1.37 (m, 2 H). ESI calculated for $C_{25}H_{43}N_4O_6S$ $[M+H]^+$ 527.4, found 527.2.

Synthesis of Desthiobiotin- N-hydroxysuccinimide ester, compound (4)

Desthiobiotin (0.2g) NHS (0.107g) and DCC (0.117g) were dissolved in DMF. The solution was stirred for 10 hours and room temperature and then was stored at 4 ^{0}C for 2 hours. The Dicyclohexylurea (DCU) was removed by filtration and the filtrate was evaporated on dryness in vacuum. The residue was dissolved in 2-propanol, and the insoluble yellow material was removed by filtration, and the filtrate was stored in 4 ^{0}C for

crystallization. The crystals were collected washed with ice-cold 2-propanol and dried. The clear crystaline product was isolated with a yield of 40%.

^{1}H NMR δ (300 MHz, DMSO) 6.60 (s,1H) 6.20 (s, 1 H), 5.50 (s, 1H), 3.60 (m, 1 H), 3.50 (m, 1 H), 2.80 (s, 4 H),2.65 (s,2H), 2.40 (t, 2 H), 1.63 (t, 2 H), 1.50-1.10 (m, 4H), 0.99 (d, 3 H).

Synthesis of compound (5)

Compound **1** was dissolved in anhydrous DMF. The Desthiobiotin-NHS ester was dissolved in DMF. Compound **1** was added to the Desthiobiotin-NHS ester. Triethylamine was added to the solution and the reaction was stirred for 14 hours. Purification was done by silica gel chromatography (dichloromethane:methanol 90:10).The product was a white powder with a yield of 60%.

^{1}H NMR (300 MHz, CD_3OD) δ 6.60 (s,1H) 6.20 (s, 1 H), 4.30 (s, 1H), 4.20 (m, 1 H), 3.50 (m, 8 H), 3.4 (m, 4 H), 3.20 (t, 4 H), 2.40 (t, 4 H), 2.00 (m, 2H), 1.70- 1.50 (m ,8 H) 1.30 (m,3H). ESI calculated for $C_{25}H_{44}N_4O_3$ $[M+H]^+$ 496.4 ,found 497.4

References

1. Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A., and Bertozzi, C. R. (2006) A comparative study of bioorthogonal reactions with azides, *ACS Chem Biol 1*, 644-648.
2. Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems, *Nat Chem Biol 1*, 13-21.
3. Griffin, R. J. (1994) The medicinal chemistry of the azido group, *Prog Med Chem 31*, 121-232.
4. Saxon, E., and Bertozzi, C. R. (2000) Cell surface engineering by a modified Staudinger reaction, *Science 287*, 2007-2010.
5. Agard, N. J., Prescher, J. A., and Bertozzi, C. R. (2004) A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems, *J Am Chem Soc 126*, 15046-15047.
6. Rostovtsev, V. V., Green, L. G., Fokin, V. V., and Sharpless, K. B. (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes, *Angew Chem Int Ed Engl 41*, 2596-2599.
7. Prescher, J. A., Dube, D. H., and Bertozzi, C. R. (2004) Chemical remodelling of cell surfaces in living animals, *Nature 430*, 873-877.
8. Speers, A. E., and Cravatt, B. F. (2004) Profiling enzyme activities in vivo using click chemistry methods, *Chem Biol 11*, 535-546.
9. Dias, W. B., and Hart, G. W. (2007) O-GlcNAc modification in diabetes and Alzheimer's disease, *Mol Biosyst 3*, 766-772.

10. Vocadlo, D. J., Hang, H. C., Kim, E. J., Hanover, J. A., and Bertozzi, C. R. (2003) A chemical approach for identifying O-GlcNAc-modified proteins in cells, *Proc Natl Acad Sci U S A 100*, 9116-9121.

11. Dube, D. H., and Bertozzi, C. R. (2003) Metabolic oligosaccharide engineering as a tool for glycobiology, *Curr Opin Chem Biol 7*, 616-625.

12. Dube, D. H., Prescher, J. A., Quang, C. N., and Bertozzi, C. R. (2006) Probing mucin-type O-linked glycosylation in living animals, *Proc Natl Acad Sci U S A 103*, 4819-4824.

13. Hang, H. C., Yu, C., Kato, D. L., and Bertozzi, C. R. (2003) A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation, *Proc Natl Acad Sci U S A 100*, 14846-14851.

14. Golks, A., Tran, T. T., Goetschy, J. F., and Guerini, D. (2007) Requirement for O-linked N-acetylglucosaminyltransferase in lymphocytes activation, *EMBO J 26*, 4368-4379.

15. Unlu, M., Morgan, M. E., and Minden, J. S. (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts, *Electrophoresis 18*, 2071-2077.

16. Gharbi, S., Gaffney, P., Yang, A., Zvelebil, M. J., Cramer, R., Waterfield, M. D., and Timms, J. F. (2002) Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system, *Mol Cell Proteomics 1*, 91-98.

17. Dowling, P., O'Driscoll, L., Meleady, P., Henry, M., Roy, S., Ballot, J., Moriarty, M., Crown, J., and Clynes, M. (2007) 2-D difference gel electrophoresis of the

lung squamous cell carcinoma versus normal sera demonstrates consistent alterations in the levels of ten specific proteins, *Electrophoresis 28*, 4302-4310.

18. Rebeaud, F., Hailfinger, S., Posevitz-Fejfar, A., Tapernoux, M., Moser, R., Rueda, D., Gaide, O., Guzzardi, M., Iancu, E. M., Rufer, N., Fasel, N., and Thome, M. (2008) The proteolytic activity of the paracaspase MALT1 is key in T cell activation, *Nat Immunol 9*, 272-281.

19. Sen Gupta, S., Kuzelka, J., Singh, P., Lewis, W. G., Manchester, M., and Finn, M. G. (2005) Accelerated bioorthogonal conjugation: a practical method for the ligation of diverse functional molecules to a polyvalent virus scaffold, *Bioconjug Chem 16*, 1572-1579.

20. Proteins and Proteomics: A laboratory manual. R.J Simpson. Cold Spring Habour Laboratory Press (2002).

Bibliography

1. Hang, H. C., Yu, C., Kato, D. L., and Bertozzi, C. R. (2003) A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation, *Proc Natl Acad Sci U S A 100*, 14846-14851.

2. Pratt, M. R., and Bertozzi, C. R. (2005) Synthetic glycopeptides and glycoproteins as tools for biology, *Chem Soc Rev 34*, 58-68.

3. Zachara, N. E., and Hart, G. W. (2006) Cell signaling, the essential role of O-GlcNAc!, *Biochim Biophys Acta 1761*, 599-617.

4. Hart, G. W., Housley, M. P., and Slawson, C. (2007) Cycling of O-linked beta-N-acetylglucosamine on nucleocytoplasmic proteins, *Nature 446*, 1017-1022.

5. Hart, G. W., Kreppel, L. K., Comer, F. I., Arnold, C. S., Snow, D. M., Ye, Z., Cheng, X., DellaManna, D., Caine, D. S., Earles, B. J., Akimoto, Y., Cole, R. N., and Hayes, B. K. (1996) O-GlcNAcylation of key nuclear and cytoskeletal proteins: reciprocity with O-phosphorylation and putative roles in protein multimerization, *Glycobiology 6*, 711-716.

6. Wells, L., Gao, Y., Mahoney, J. A., Vosseller, K., Chen, C., Rosen, A., and Hart, G. W. (2002) Dynamic O-glycosylation of nuclear and cytosolic proteins: further characterization of the nucleocytoplasmic beta-N-acetylglucosaminidase, O-GlcNAcase, *J Biol Chem 277*, 1755-1761.

7. Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., and Marth, J. D. (2000) The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny, *Proc Natl Acad Sci U S A 97*, 5735-5739.

8. Federici, M., Menghini, R., Mauriello, A., Hribal, M. L., Ferrelli, F., Lauro, D., Sbraccia, P., Spagnoli, L. G., Sesti, G., and Lauro, R. (2002) Insulin-dependent activation of endothelial nitric oxide synthase is impaired by O-linked glycosylation modification of signaling proteins in human coronary endothelial cells, *Circulation 106*, 466-472.

9. Kamemura, K., Hayes, B. K., Comer, F. I., and Hart, G. W. (2002) Dynamic interplay between O-glycosylation and O-phosphorylation of nucleocytoplasmic proteins: alternative glycosylation/phosphorylation of THR-58, a known mutational hot spot of c-Myc in lymphomas, is regulated by mitogens, *J Biol Chem 277*, 19229-19235.

10. Comer, F. I., and Hart, G. W. (2001) Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II, *Biochemistry 40*, 7845-7852.

11. Cheng, X., Cole, R. N., Zaia, J., and Hart, G. W. (2000) Alternative O-glycosylation/O-phosphorylation of the murine estrogen receptor beta, *Biochemistry 39*, 11609-11620.

12. Datta, R., Choudhury, P., Bhattacharya, M., Soto Leon, F., Zhou, Y., and Datta, B. (2001) Protection of translation initiation factor eIF2 phosphorylation correlates with eIF2-associated glycoprotein p67 levels and requires the lysine-rich domain I of p67, *Biochimie 83*, 919-931.

13. Elliot, S. P., Schmied, R., Gabel, C. A., and Ambron, R. T. (1993) An 83 kDa O-GlcNAc-glycoprotein is found in the axoplasm and nucleus of Aplysia neurons, *J Neurosci 13*, 2424-2429.

14. Khidekel, N., Ficarro, S. B., Peters, E. C., and Hsieh-Wilson, L. C. (2004) Exploring the O-GlcNAc proteome: direct identification of O-GlcNAc-modified proteins from the brain, *Proc Natl Acad Sci U S A 101*, 13132-13137.

15. Dias, W. B., and Hart, G. W. (2007) O-GlcNAc modification in diabetes and Alzheimer's disease, *Mol Biosyst 3*, 766-772.

16. Zachara, N. E., O'Donnell, N., Cheung, W. D., Mercer, J. J., Marth, J. D., and Hart, G. W. (2004) Dynamic O-GlcNAc modification of nucleocytoplasmic proteins in response to stress. A survival response of mammalian cells, *J Biol Chem 279*, 30133-30142.

17. Slawson, C., and Hart, G. W. (2003) Dynamic interplay between O-GlcNAc and O-phosphate: the sweet side of protein regulation, *Curr Opin Struct Biol 13*, 631-636.

18. Rexach, J. E., Clark, P. M., and Hsieh-Wilson, L. C. (2008) Chemical approaches to understanding O-GlcNAc glycosylation in the brain, *Nat Chem Biol 4*, 97-106.

19. Wells, L., Vosseller, K., Cole, R. N., Cronshaw, J. M., Matunis, M. J., and Hart, G. W. (2002) Mapping sites of O-GlcNAc modification using affinity tags for serine and threonine post-translational modifications, *Mol Cell Proteomics 1*, 791-804.

20. Vosseller, K., Hansen, K. C., Chalkley, R. J., Trinidad, J. C., Wells, L., Hart, G. W., and Burlingame, A. L. (2005) Quantitative analysis of both protein expression and serine / threonine post-translational modifications through stable isotope labeling with dithiothreitol, *Proteomics 5*, 388-398.

21. Khidekel, N., Ficarro, S. B., Clark, P. M., Bryan, M. C., Swaney, D. L., Rexach, J. E., Sun, Y. E., Coon, J. J., Peters, E. C., and Hsieh-Wilson, L. C. (2007) Probing the dynamics of O-GlcNAc glycosylation in the brain using quantitative proteomics, *Nat Chem Biol 3*, 339-348.

22. Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A., and Bertozzi, C. R. (2006) A comparative study of bioorthogonal reactions with azides, *ACS Chem Biol 1*, 644-648.

23. Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A., and Bertozzi, C. R. (2006) A comparative study of bioorthogonal reactions with azides, *ACS Chem Biol 1*, 644-648.

24. Prescher, J. A., and Bertozzi, C. R. (2005) Chemistry in living systems, *Nat Chem Biol 1*, 13-21.

25. Griffin, R. J. (1994) The medicinal chemistry of the azido group, *Prog Med Chem 31*, 121-232.

26. Saxon, E., and Bertozzi, C. R. (2000) Cell surface engineering by a modified Staudinger reaction, *Science 287*, 2007-2010.

27. Agard, N. J., Prescher, J. A., and Bertozzi, C. R. (2004) A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems, *J Am Chem Soc 126*, 15046-15047.

28. Rostovtsev, V. V., Green, L. G., Fokin, V. V., and Sharpless, K. B. (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes, *Angew Chem Int Ed Engl 41*, 2596-2599.

29. Prescher, J. A., Dube, D. H., and Bertozzi, C. R. (2004) Chemical remodelling of cell surfaces in living animals, *Nature 430*, 873-877.

30. Speers, A. E., and Cravatt, B. F. (2004) Profiling enzyme activities in vivo using click chemistry methods, *Chem Biol 11*, 535-546.

31. Dias, W. B., and Hart, G. W. (2007) O-GlcNAc modification in diabetes and Alzheimer's disease, *Mol Biosyst 3*, 766-772.

32. Vocadlo, D. J., Hang, H. C., Kim, E. J., Hanover, J. A., and Bertozzi, C. R. (2003) A chemical approach for identifying O-GlcNAc-modified proteins in cells, *Proc Natl Acad Sci U S A 100*, 9116-9121.

33. Dube, D. H., and Bertozzi, C. R. (2003) Metabolic oligosaccharide engineering as a tool for glycobiology, *Curr Opin Chem Biol 7*, 616-625.

34. Dube, D. H., Prescher, J. A., Quang, C. N., and Bertozzi, C. R. (2006) Probing mucin-type O-linked glycosylation in living animals, *Proc Natl Acad Sci U S A 103*, 4819-4824.

35. Hang, H. C., Yu, C., Kato, D. L., and Bertozzi, C. R. (2003) A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation, *Proc Natl Acad Sci U S A 100*, 14846-14851.

36. Golks, A., Tran, T. T., Goetschy, J. F., and Guerini, D. (2007) Requirement for O-linked N-acetylglucosaminyltransferase in lymphocytes activation, *EMBO J 26*, 4368-4379.

37. Unlu, M., Morgan, M. E., and Minden, J. S. (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts, *Electrophoresis 18*, 2071-2077.

38. Gharbi, S., Gaffney, P., Yang, A., Zvelebil, M. J., Cramer, R., Waterfield, M. D., and Timms, J. F. (2002) Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system, *Mol Cell Proteomics 1*, 91-98.

39. Dowling, P., O'Driscoll, L., Meleady, P., Henry, M., Roy, S., Ballot, J., Moriarty, M., Crown, J., and Clynes, M. (2007) 2-D difference gel electrophoresis of the lung squamous cell carcinoma versus normal sera demonstrates consistent alterations in the levels of ten specific proteins, *Electrophoresis 28*, 4302-4310.

40. Rebeaud, F., Hailfinger, S., Posevitz-Fejfar, A., Tapernoux, M., Moser, R., Rueda, D., Gaide, O., Guzzardi, M., Iancu, E. M., Rufer, N., Fasel, N., and Thome, M. (2008) The proteolytic activity of the paracaspase MALT1 is key in T cell activation, *Nat Immunol 9*, 272-281.

41. Sen Gupta, S., Kuzelka, J., Singh, P., Lewis, W. G., Manchester, M., and Finn, M. G. (2005) Accelerated bioorthogonal conjugation: a practical method for the ligation of diverse functional molecules to a polyvalent virus scaffold, *Bioconjug Chem 16*, 1572-1579.

42. Proteins and Proteomics: A laboratory manual. R.J Simpson. Cold Spring Habour Laboratory Press (2002).