DEVELOPING CHEMICAL TOOLS TO MAP MOLECULAR MECHANISMS THAT DRIVE DISEASE

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Alexander Kenneth Hurben

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Dr. Natalia Y. Tretyakova

April 2023

© Alexander K. Hurben 2023

Acknowledgements

Throughout my scientific journey thus far, I have realized that having a strong support system and luck is absolutely critical for individual success. I am extremely fortunate to have benefited from both, and thus I have many people I need to acknowledge here.

I would like to thank my advisor, Professor Natalia Tretyakova, for all her support and mentorship over the years. Her passion for science has been truly inspiring and I have grown tremendously during my tenure in her group. I would also like to thank my committee members, Professors Carrie Haskell-Luevano, Daniel Harki, and Steven Graves, for their guidance and support. I also would like to acknowledge Professor Mark Distefano in providing me with the opportunity to perform undergraduate research. This experience and training were truly formative in cementing my decision to pursue graduate school and prepared me well for the trials ahead.

The entire Department of Medicinal Chemistry has been instrumental in my success. I owe a great deal to the outstanding faculty and staff whose ongoing support and advocacy have been critical during my time in graduate school.

Additionally, I have been most fortunate to have been surrounded by an excellent cohort of graduate students and lab mates. Through them, I have forged lifelong friendships, received invaluable mentorship, and garnered a support system that was essential in navigating my graduate career.

Completion of this thesis is a testament to the amazing people I have been surrounded by. Without them, I would not be where I am today.

Dedication

This thesis is dedicated to my family along with my partner Samantha. Their unwavering love and support were foundational in making my achievements possible.

Abstract

Human health is impacted by molecular level events. Within this context, DNA encodes for cellular instructions and proteins execute these genetic protocols to ensure the cell's components are functioning properly. Thus, the underlying biochemistry within the cell is tightly regulated to ensure vitality. However, when these processes become compromised or damaged, there is increased susceptibility towards developing cancer and neurodegenerative diseases. Often, small chemical changes to our DNA and proteins can be the culprits of such dysregulation. These seemingly minuscule modifications can impair the cell's proper functions, which can lead to cell death or uncontrolled growth, and ultimately manifest as degenerative disease or cancer. Electrophilic small molecules can be responsible for chemically altering cellular machinery. Additionally, enzymes can make chemical changes such as post-translational modifications on proteins and epigenetic DNA modifications, which can have pronounced effects on cell function and can be equally damaging if not properly regulated.

Understanding how molecular events influence health is of paramount importance in designing therapeutics that effectively prevent and treat disease, as well as discovering biomarkers which enable early detection. Despite immense research efforts from the scientific community, there is much remaining to learn about these microscopic processes. This is due to their inherent complexity and lack of technologies to study them. This thesis aims to contribute to addressing this problem by developing new chemical tools which advance our knowledge of the molecular mechanisms that drive

disease. This work is composed of seven chapters which explore reactive dopamine metabolites linked to Parkinson's disease, tools to study the biological implications of elevated intracellular methylglyoxal concentrations, and the development of small molecule epigenetic modulators to regulate aberrant DNA methylation.

Chapters I, **II**, and **III** of this thesis explores how dysregulated dopamine may contribute to Parkinson's disease initiation. **Chapter I** commences with a review of dopamine metabolism and the subsequent generation of reactive dopamine derived metabolites in neurons. This is followed by an overview of protein damaged induced by these metabolites and a review of chemical tools and techniques implemented to study dysregulated dopamine in various experimental systems. Next, **Chapter II** describes our efforts in designing and implementing a dopamine derived chemical probe to profile dopamine modified proteins which found that dopamine metabolites disrupt protein-folding pathways critical for maintaining healthy neurons. We also detail our development of photoactivatable dopamine probes in **Chapter III**. Collectively, this work improves the understanding of dopamine protein modification and by extension, molecular events that may contribute to Parkinson's, which may inform future Parkinson's therapeutic development.

Chapter IV and **Chapter V** of this thesis focuses on the reactive metabolite methylglyoxal. Methylglyoxal is a sugar-derived metabolite produced naturally in all cells. This reactive compound forms adducts with DNA and proteins, thereby altering their function and influencing cell signalling. Consequently, methylglyoxal protein

adducts are implicated in numerous diseases such as cancer, neurodegeneration, diabetes, and cardiovascular disease. In many of these diseases, the cellular processes that break down methylglyoxal become compromised, leading to elevated levels of this reactive molecule within cells. Existing chemical tools to investigate methylglyoxal biology are limited, leading to an incomplete understanding of its physiological and disease-causing roles. Here, we disclose a chemical tool that confers light-mediated release of a methylglyoxal probe within cell models. We use this chemical to identity of the resulting protein adducts. This work enables studying protein adducts induced by methylglyoxal in a controlled fashion to illuminate how this reactive compound impacts various disease states. We also detail our efforts in profiling proteins which undergo covalent DNA crosslinking in the presence of methylglyoxal in **Chapter V**. This effort is the first study to identify this type of methylglyoxal adduct at a proteome wide scale, which provides a list of candidate methylglyoxal derived DNA-protein cross to investigate in future work. Collectively, these efforts further our understanding of basic methylglyoxal biology and its role in disease progression.

Finally, **Chapters VI** and **VII** of this thesis describe our efforts towards developing small molecule epigenetic modulators. Regulating gene expression is critical for keeping cells healthy. Over- or under-expressed genes can lead to cancer and other diseases. Accordingly, cells have many methods to control when specific genes are turned on or off in order to function properly; DNA methylation being an example. There are many proteins which control the addition and removal of DNA methyl marks across the

genome to ensure appropriate gene expression. Mutations in, or dysfunction of, these proteins can initiate certain cancers. One essential group of proteins involved in removing DNA methylation marks is ten-eleven translocation (TET) methylcytosine dioxygenases (TET). Given TET's central role in cancer development, theses protein represent a potential drug target. However, there is a paucity of small molecules which selectively inhibit their function without affecting other cellular processes. Thus, there is a need to develop potent and selective compounds which block TET-meditated DNA demethylation. Within **Chapter VI**, we show our efforts towards the development of novel small molecule TET inhibitors which led us to uncover that copper contamination is responsible for the activity of a reported TET inhibitor. In **Chapter VII**, we present work on a novel TET inhibitor scaffold which features a bifunctional cofactor-substrate mimetic design. This work has the potential to generate new anticancer therapeutics and improve our understanding of how TET and DNA methylation is linked to cancer development.

Ultimately, the work in this thesis provides a novel suite of chemical tools for studying dopamine dysregulation, methylglyoxal metabolism, and TET function. These tools provide insights into cellular damage caused by dopamine and methylglyoxal adducts as well as probes for altering DNA methylation status. Such tools are critical for mapping molecular mechanisms that drive disease.

Table of Contents

List of Tables

List of Figures

1 Chapter I: Role of Protein Damage Inflicted by Dopamine Metabolites in Parkinson's Disease: Evidence, Tools, and Outlook

Adapted with permission by the American Chemical Society from:

Hurben A.K., Tretyakova N.Y. The Role of Protein Damage Inflicted by Dopamine Metabolites in Parkinson's Disease: Evidence, Tools, and Outlook. *Chemical Research in Toxicology*. **2022**, *35* (10), 1789–1804.

This chapter is adopted from a manuscript published in *Chemical Research in Toxicology* drafted by Alexander K. Hurben under the supervision of Prof. Tretyakova.

1.1 Introduction

Parkinson's disease (PD) is a progressive neurodegenerative disorder that afflicts over five million individuals globally.[1] PD symptomology manifests as a continual decline in motor function, resulting in rigidity, bradykinesia, tremor at rest, and postural instability.[2] The cause of motor dysfunction in PD is largely attributed to the depletion of striatal dopamine (DA) due to the death of dopaminergic neurons, particularly in the substantia nigra pars compacta (SNc) (**Figure 1.1**).[3] However, the events that initiate dopaminergic collapse in the SNc are thought to take place up to a decade prior to diagnosis, as it is estimated that 50−90% of dopaminergic neurons have already been lost by that time.[4] Currently, there are no approved therapeutics that halt or slow PD. Thus, improving understanding of the molecular underpinnings that initiate PD remains an important area in developing novel treatments. Herein, we discuss the role that dysregulated DA plays in driving PD pathology,[5] highlighting the contribution of covalent protein damage induced by reactive DA metabolites. In addition, novel methodologies that enable studying protein damage inflicted by DA metabolites are presented. We conclude with an outlook on this area of PD research and a perspective on how further work investigating DA−protein damage in the brain will aid in elucidating PD etiology.

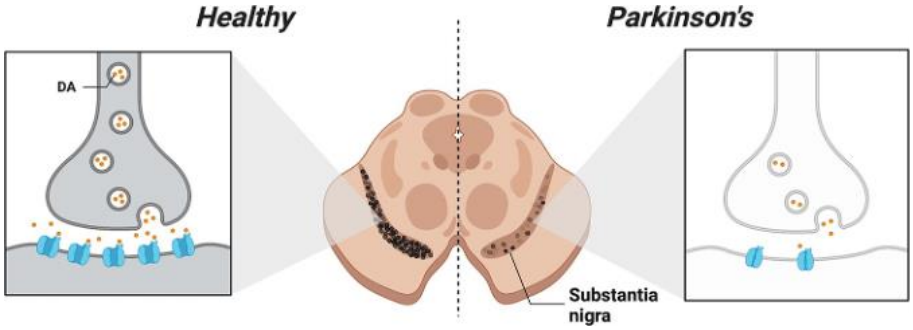


Figure 1.1. Loss of dopaminergic neurons in Parkinson's disease

Representation of autopsied midbrain sections from healthy and PD diagnosed individuals (right and left, respectively). Impairment and death of dopaminergic neurons is a hallmark of PD.

1.2 Dopamine Metabolism and Formation of Reactive Dopamine Metabolites

DA was first reported to be a neurotransmitter by Arvid Carlsson in 1957.[6] Since then, DA has been found to modulate a multitude of physiological processes, ranging from motivational salience to motor coordination, and roles outside the central nervous system.[7, 8] The link between DA and PD was shown by Oleh Hornykiewicz, who reported that DA concentration is drastically decreased in the midbrains of PD patients compared to healthy controls.[9] This landmark study launched the use of the DA metabolic precursor, L-3,4dihydroxyphenylalanine (L-DOPA), as a treatment for PD.[10] Despite the symptomatic relief L-DOPA can provide in PD, the underlying neurodegenerative processes involved in the disease are not affected, which has sparked some debate over its continued use.[11-14] Regardless, DA remains a central player in PD etiology, and emerging evidence points to reactive DA metabolites being contributors to the pathology.[15]

DA metabolism is a carefully orchestrated process overseen by various enzymatic and sequestering systems.[16] It has been hypothesized that disruption of these systems can lead to production of reactive DA metabolites and buildup of reactive oxygen species (ROS) and impair carbonyl metabolism, all of which can impart neurotoxicity.[17] A summary of DA metabolism is shown in **Figure 1.2**. DA generation commences with the hydroxylation of L-Tyr by tyrosine hydroxylase (TH) to give L-DOPA.[18] Next, L-DOPA is decarboxylated via aromatic acid decarboxylase (AADC) through a tetrahydrobiopterin (H4-biopterin) and iron-dependent manner to furnish DA.[19] Once DA is created, it can be converted to norepinephrine (NE) by a copper-containing monooxygenase, dopamine β-

hydroxylase (DβH), or catabolized by one of two major and minor pathways.[20] In the major pathway, DA is deaminated to 3,4-dihydroxyphenylacetaldehyde (DOPAL) by monoamine oxidase (MAO). MAO is a flavin adenine dinucleotide (FAD) oxidase that produces hydrogen peroxide and ammonia as additional byproducts in this reaction.[21] MOA is also anchored to the outer mitochondrial membrane and can shuttle electrons from DA metabolism to the mitochondrial electron transport chain.[22] DOPAL is further oxidized to its corresponding acid, 3,4-dihydroxyphenylacetic acid (DOPAC), by an aldehyde dehydrogenase (ALDH).[23] DOPAC is subsequently methylated to form homovanillic acid (HVA) through catechol-O-methyltransferase (COMT).[24, 25] HVA is an inert byproduct capable of being excreted. In the minor pathway, one of the hydroxyls of DA is first capped by a methyl group via COMT to afford 3-methoxytyramine (MTY), which is consequently deaminated by MAO to give 3-methoxy-4hydroxyphenyl acetaldehyde (MOPAL) and oxidized by ADH to HVA.[26] In addition, the aldehyde functionality within DOPAL can be reduced, primarily by cytosolic aldehyde reductases (AR) or by alcohol dehydrogenases (ADH), to produce 3,4-dihydroxyphenylethanol (DOPET) as a minor metabolite.[27-29] Together, these enzymatic pathways regulate the biosynthesis and metabolism of DA in presynaptic dopaminergic neurons.

Beyond enzymatic regulation, DA concentrations in neurons are controlled by subcellular compartmentalization. An overview of these processes is depicted in **Figure 1.3**. Within DA presynaptic neurons, DA is primarily housed within synaptic vesicles. Vesicular monoamine transporter 2 (VMAT2) is the transporter responsible for this shuttling of DA.[30] The two enzymes that biosynthesize DA from L-Tyr, TH and AACD,

are also physically associated with VMAT2 to assist in the confinement of DA.[31, 32] The H+ ATPase localized in the vesicular membrane pumps protons inside the vesicle to maintain an acidic environment in order to stabilize the confined DA.[33] DA concentrations are reported to reach up to 1 M in vesicles, whereas cytosolic concentrations are measured in the low micromolar range.[34, 35] DA released into the synaptic cleft undergoes reuptake by dopamine transporter (DAT) for vesicular repacking or is broken down by astrocytes in close proximity.[36] Residual cytosolic DA in presynaptic neurons is catabolized to HVA or sequestered into granules of neuromelanin.[37] Proper regulation of DA handling systems is critical for dopaminergic health.

Because of its catechol structure, DA has a propensity to spontaneously form reactive metabolites (**Figure 1.4A**). When not confined to the acidic environment of synaptic vesicles, DA is susceptible to oxidation by a variety of processes to give dopamine semiquinone (DASQ) or dopaquinone (DAQ).[38] DASQ can be generated *in vivo* enzymatically through H_2O_2 activation by peroxidases and prostaglandin H synthase.[39, 40] DAQ is formed through auto-oxidation at cytosolic pH; this process can be accelerated by metals like Cu and Fe.[41] DAQ rapidly undergoes intramolecular cyclization to afford leukodopaminochrome (LDAC).[42] However, in some cases, DAQ can be converted to 6-OHDA (6-hydroxydopamine) and 6-OHDQ (6-hydroxydopaquinone) by ROS and Fe.[43, 44] LDAC is also prone to oxidation, akin to DA, to form dopaminochrome (DAC).[45] DAC undergoes a tautomeric rearrangement to 5,6-dihydroxyindole (DHI), which further oxidizes to 5,6-dihydroxyindolequinone (DHIQ)

and ultimately polymerizes to neuromelanin.[46] As shown in **Figure 1.4B**, the catechol in DOPAL can also be oxidized to a quinone to give DOPAL quinone (DPQAL), another reactive DA derived metabolite.[47] Collectively, DA oxidation *in vivo* leads to a complex array of potentially toxic electrophilic metabolites that can react with proteins, necessitating tight regulation over neuronal DA metabolism.

There is accumulating evidence that DA handling systems are dysregulated in the context of PD, contributing to disease progression.[48, 49] Specifically, vesicular uptake of DA by VMAT2 is reduced by 87−90% in autopsied striatum sections of PD patients as compared to healthy controls.[50] Impaired vesicular DA uptake leads to accumulation of cytosolic DA, which could have deleterious consequences on neuronal health.[51, 52] Accordingly, genetically enhanced vesicular function of DA provides neuroprotection against the dopaminergic toxin 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) in murine models of PD.[53] Furthermore, several promoter haplotypes that increase the expression of VMAT2 are associated with a decreased incidence of PD.[54] In other investigations of PD etiology, accumulation of oxidized DA has been longitudinally linked to a toxic cascade of neurological dysfunction correlated with idiopathic and familial PD cases.[55] Together, these studies point to the importance of DA regulation in PD.

Beyond improper DA storage in PD, other pathological factors of PD can exacerbate DA-inflicted stress.[56] One of these factors is higher Fe and other metal content in the brains of PD patients as compared to healthy controls.[57-59] The presence of redox-active metals can induce oxidative stress and promote DA oxidation.[60, 61] In addition, the

concentrations of ROS scavenging molecules like glutathione (GSH) are decreased in PD, perpetuating the toxicity of cytosolic DA.[62, 63] Collectively, there are multiple pathways that can lead to DA dysregulation and faulty DA toxification which can contribute to PD progression. Next, we will explore specific reactions of DA metabolites along with the ensuing DA damage protein and its relevance to PD.

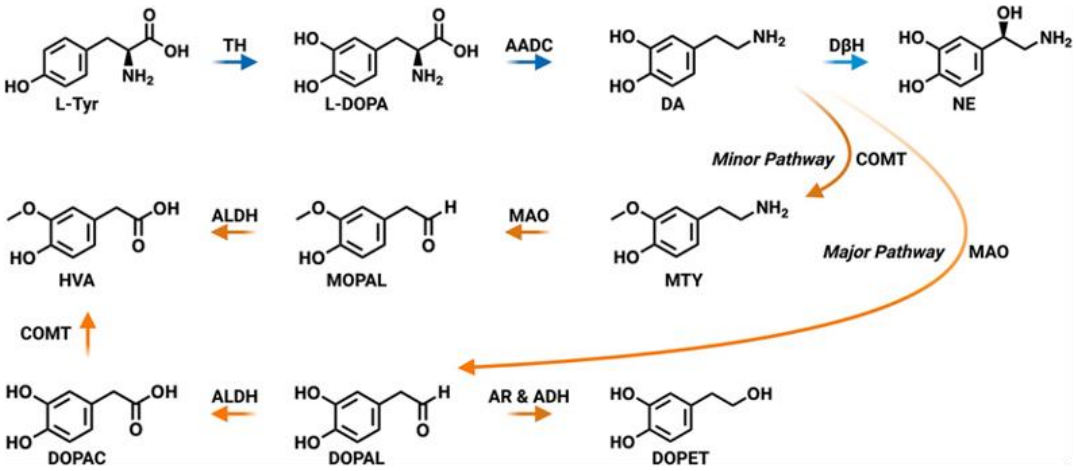


Figure 1.2. Primary pathways for dopamine biosynthesis and metabolism

DA is generated from L-Tyr by TH and AADC. DA can be converted by DβH to NE or broken down by two major and minor pathways consisting of transformations catalyzed by MAO, ADH, ALDH, AR, and COMT.

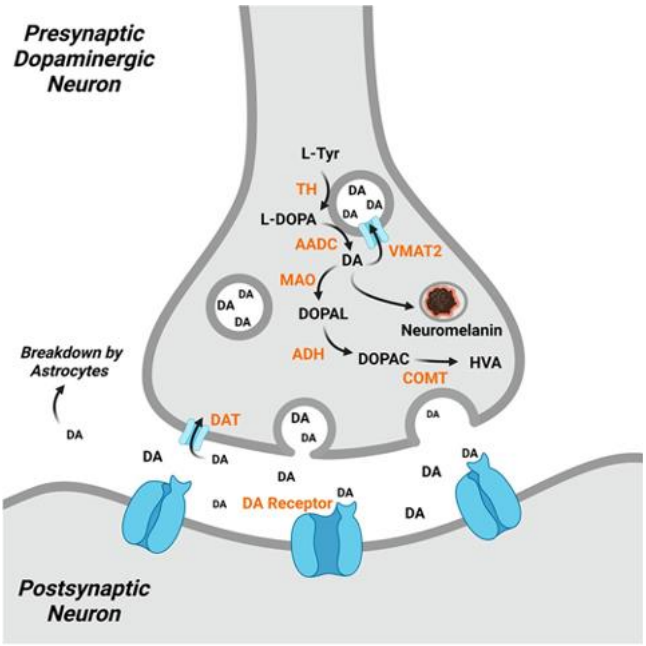


Figure 1.3. Dopamine handling systems in neurons

DA is created by TH and AADC, which is then sequestered into synaptic vesicles by VMAT2. Excess cytosolic DA is converted to HVA by MAO, ADH, and COMT or enveloped into neuromelanin. Residual DA in the synaptic cleft following release by the presynaptic neuron undergoes reuptake by DAT and is repackaged in synaptic vesicles or is broken down by neighboring astrocytes.

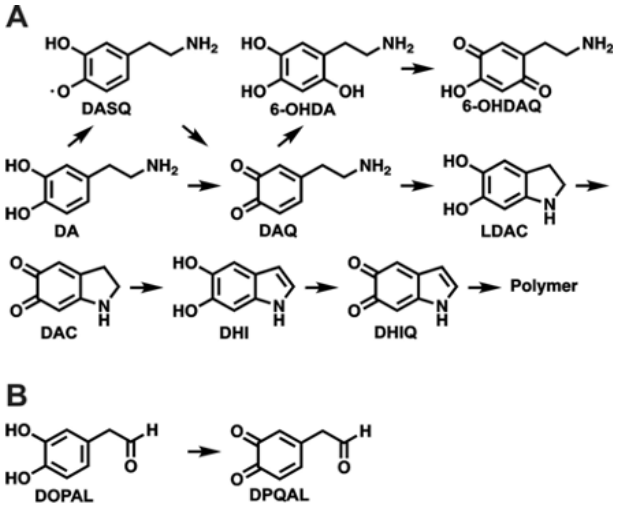


Figure 1.4. Dopamine-derived reactive metabolites

A) Oxidation products of DA. DA can undergo a 1 e− oxidation to DASQ or a 2 e− oxidation to DAQ. DAQ cyclizes to LDAC or reacts with ROS to form 6-OHDA. 6-OHDA can further oxidize to 6-OHDQ. LDAC can oxidize to DAC, which can rearrange to DHI. DHI can oxidize to DHIQ, which polymerizes due to instability. B) Oxidization of DOPAL to DPQAL.

1.3 Protein Damage Caused by Dopamine Metabolites

DA has long been recognized as a driver of oxidative stress in the brain.[64, 65] Intrastriatal injections of DA in rats caused neuronal degeneration, providing an early demonstration of DA toxicity in the SNc.[66] Importantly, DA toxicity could be alleviated by coadministration of a reductant such as ascorbic acid or GSH, suggesting that oxidized DA metabolites were responsible for inducing damage. Understanding the molecular mechanisms responsible for DA toxicity in the brain remains an active area of research.[67]

DA metabolites can undergo a variety of chemical reactions that can impart cellular damage.[68] An overview of these reactions is shown in **Figure 1.5**. Some DA metabolites are less stable than others, DASQ being an example. As shown in **Figure 1.5A**, DASQ can couple with other radicals, scavenge cellular thiols like GSH, undergo a one-electron oxidation by O_2 to generate DAQ and O_2^-, or disproportionate to DA and DAQ, all of which can inflict toxicity.[69, 70] DAQ and other quinone species like DAC, DHIQ, and 6-OHDQ can react with various nucleophilic amino acid residues, which is depicted in **Figure 1.5B**.[71] Among the amino acids side chains, Cys has the highest reactivity toward DA-derived quinones, as thiols have a rate constant 10 000-fold higher than amines for nucleophilic attack at the quinone core.[72-74] It has been postulated that the reaction between thiols and ortho-quinones may proceed through a radical mechanism which accounts for the rapid reaction rate and high regioselectivity of thiol addition to the 5 position in the ring.[75] In addition to Cys, Lys and His side chains can react with DA metabolites.[76-78] The metabolites DOPAL and DPQAL also harbor reactive groups susceptible to modification and thus have been reported as capable of inducing toxicity.[17,

[79, 80] As shown in **Figure 1.5C**, the aldehyde groups of DOPAL and DPQAL can form a Schiff base with the amine groups of proteins and peptides.[81] Following Schiff base formation DOPAL can oxidize to its DPQAL form and undergo addition to give rise to cross-linked products.[28, 82, 83] It is likely that free DPQAL alone is a minor contributor to protein cross-linking due to its inherit instability. However, the chemistry of DOPAL is complex; therefore, studies investigating the reactivity of this metabolite need to carefully consider experimental variables such as pH and use of deoxygenated media. Such experimental conditions may bias observed DOPAL reactivity profiles as these factors influence the rates of Schiff base formation and auto-oxidation processes and thus may not fully recapitulate *in vivo* conditions. In some cases, two molecules of DOPAL can react with a Lys residue to generate a dicatechol pyrrole adduct.[84] In addition, once formed, DA protein adducts can redox cycle and produce additional ROS.[85] Collectively, the reactivity of DA metabolites is the basis for the damage they can impart on proteins.

DA metabolites have been reported to react with a diverse array of proteins. The resulting adducts can impair protein function or cause structural alterations to the native conformation (**Figure 1.6**). These protein modifications can have deleterious effects on many biochemical pathways relevant to PD. Below, we discuss various reported DA adducted proteins which are also summarized in **Table 1.1**.

DA metabolites have been shown to damage numerous regulatory proteins involved in DA homeostasis. One of the first proteins reported to react with DA metabolites was tyrosine hydroxylase (TH).[86, 87] Through the use of radiolabeled DA, TH was shown to be covalently modified by oxidized DA metabolites and to become

functionally inhibited as a result. DAT, another DA handling enzyme, was also found to be adducted by DA metabolites at Cys residues, which resulted in blocked DA uptake.[88] In addition, dihydropteridine reductase (DHPR) activity is inhibited by DA-derived quinones; however, it is not known if this occurs through covalent modification.[89] DHPR regenerates tetrahydrobiopterin, which is a cofactor required for hydroxylation reactions and biosynthesis of DA precursors L-Tyr and LDOPA.[90] The metabolite DHI has been shown to covalently bind a regulatory Cys in the nuclear receptor related-1 protein (Nurr1) which stimulates Nurr1 activity in cell culture and zebrafish.[91] Nurr1 is a transcription factor that regulates expression of critical maintenance and survival genes of dopaminergic neurons, including DA handling systems like TH, VMAT2, and DAT.[92-97] In fact, all of these aforementioned proteins are important players in DA regulation; thus, their inhibition by aberrant DA metabolites fuels a vicious cycle of DA dysregulation relevant to PD.

Oxidized DA metabolites have been linked to mitochondrial dysfunction and disruption of mitochondrial electron transport.[98, 99] Consequently, various mitochondrial proteins are reported to be impaired by reactive DA species. For example, superoxide dismutase 2 (SOD2) aggregates in the presence of DAQ due to covalent Cys adducts *in vitro*, which results in diminished catalytic activity.[100] SOD2 is responsible for converting superoxide into H_2O_2 and O_2.[101] Inhibition of SOD2 activity can result in mitochondrial dysfunction *in vivo*.[102] The selenoprotein glutathione peroxidase 4 is also functionally inhibited by DAQ, forming aggregates *in vitro* and in rat neuronal PC12 cells.[103] Glutathione peroxidase 4 is a mitochondrial antioxidant enzyme that reduces a variety of

cellular peroxides which is combative against oxidative stress and thus hypothesized to be protective against neurodegeneration.[104] In isolated rat brain mitochondria treated with radiolabeled DA, ATP synthase was found to be a target of DA metabolites,[105] particularly at the γ-subunit of the F 1 catalytic domain as determined by mass spectrometry. This modification resulted in impaired ATP production. In the same study, DA-derived quinones also affected the mitochondrial morphology. As a whole, mounting evidence suggests that oxidized DA metabolites can elicit mitochondrial dysfunction through protein adduct formation and damage.

Excess reactive DA species have been shown to damage antioxidant pathways. DOPAL inhibits glutathione-S-transferase (GST) *in vitro* through covalent modification of this protein.[106] GST catalyzes the addition of GSH to reactive metabolites of DA, which serves as a general cellular detoxification mechanism.[107] Overexpression of GST was found to be protective in a Drosophila model of PD.[108] In addition, high levels of DA administered to astroglial cells resulted in GSH depletion.[109] Thus, inhibition of GST and GSH reduction by dysregulated DA may promote oxidative damage in PD.

DA metabolites have been found to modify endoplasmic reticulum (ER) proteins. The ER plays a critical role in maintaining protein homeostasis. ER stress results in accumulation of unfolded proteins and activation of the unfolded protein response (UPR).[110] ER stress and UPR are drivers of neurodegenerative processes and PD progression.[111-114] Protein disulfide isomerase (PDI) is a prominent ER chaperon protein responsible for properly folding proteins by catalyzing disulfide bond formation.[115] Disruption of PDI activity can lead to ER stress and UPR activation.[116] Oxidized DA

metabolites have been reported to modify PDI in cell models and inhibit PDI activity *in vitro*.[117, 118] In addition, ER stress is induced by 6-OHDA in cell models.[114] Thus, it is possible that dysregulated DA metabolism may influence ER stress experienced in PD.

Various proteins with genetic ties to PD are known to be damaged by DA metabolites. Among these is the vesicular trafficking protein alpha synuclein (AS). AS is closely linked with PD as mutations in this protein result in PD development.[119, 120] AS is also a main component of Lewy bodies, which are insoluble intracellular protein aggregates observed in the brains of PD patients.[121, 122] Numerous studies, ranging from *in vitro* analyses to animal models of PD, have shown that DA metabolites interact with AS, promote AS aggregation, disrupt AS-linked synaptic vesicle function, impede chaperone mediated autophagy processes, and induce degeneration.[82, 123-130] In addition, DA-modified AS has been found in the plasma of PD patients, further strengthening the importance of this adduct.[131] Parkin is another protein where mutations are linked to early onset PD.[132] Parkin is a E3 ubiquitin ligase that helps clear damaged mitochondria via autophagy and proteasomal mechanisms.[133] DA metabolites have been shown to inhibit parkin activity in cell culture, and parkin−DA adducts have been found in autopsied PD brains.[134, 135] DJ-1 is a redox sensor protein associated with familial PD.[136] The physiological function of DJ-1 is still being elucidated, but it is well established that DJ-1 confers neuroprotection against oxidative stress.[137, 138] DJ-1 has been found to be covalently modified by DA-derived quinones *in vitro*, leading to structural alteration of DJ-1.[139] Mutations in glucocerebrosidase (GCase) are correlated with PD development.[140] GCase hydrolyzes the glycosidic bond in glucocerebroside glycolipids and is a key

component in proper lysosome function.[141] Reactive DA metabolites have been found to adduct a Cys residue in the active site of GCase in neurons obtained from patients with idiopathic and familial PD, resulting in impaired lysosomal function.[55] Given that DA metabolites target and inhibit proteins whose dysfunction is already linked to PD, a compelling case can be made for mishandled DA being involved in idiopathic occurrences of PD.

Histones, the proteins involved in DNA packaging within the nucleus, have been reported to be modified by DA via transglutaminase 2 (TGM2).[142, 143] TGM2 catalyzes amide bond formation between the carboxylic acid side chain of Glu and the amine of DA to yield a DA−histone linkage. The levels of this modification in rat brains were modulated by cocaine and heroin administration. Consequently, the degree of histone DA modification changed levels of gene expression and drug seeking behavior. Although this histone modification does not involve reactive metabolites of DA, dysregulated DA concentrations in PD may influence levels of this histone mark and lead to pathological alterations in gene expression.[144] Whether histone modifications are relevant in PD pathology remains to be seen; however, it is an intriguing area to explore in the future.

Beyond protein dysfunction induced by DA modification, antibodies reactive to DA-adducted ovalbumin have been found in PD patient serum. The presence of such DA damage recognizing antibodies creates a possible link to the immune system being involved in PD as DA damage recognizing antibodies could generate or amplify inflammatory responses in PD.[145] In addition, DA-derived neuromelanin is capable of activating innate immune cells, microglia, which induce degeneration of DA neurons.[146]

Growing evidence points to immune system dysfunction playing a role in PD progression.[147-149] The studies mentioned above suggest that DA and DA-induced protein damage may contribute to aberrant immune responses in PD.

Collectively, covalent protein modifications by reactive DA metabolites can drive neuronal toxicity.[150, 151] The studies described above provide evidence that dysregulation of DA is a critical factor in PD progression. Further elucidation of the molecular underpinnings of DA-dependent damage is key in fully understanding PD etiology. Many efforts in this area utilize chemical tools to drive this research forward. Below, various strategies for studying DA dysregulation using chemical biology and analytical chemistry tools are presented.

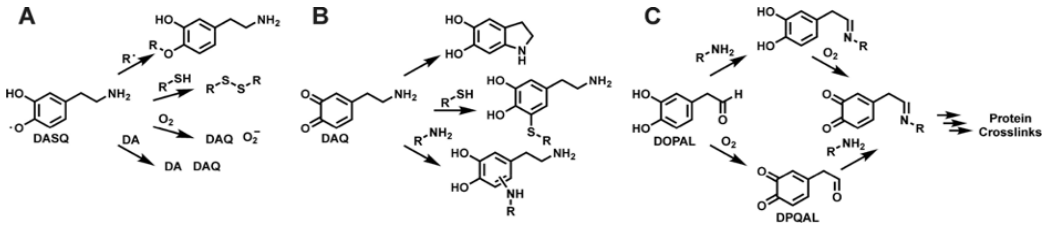


Figure 1.5. Chemical reactions of dopamine metabolites

A) Reactions by DASQ: radical cross coupling, thiol oxidation, superoxide generation, and disproportionation (shown from top to bottom). B) Reactions by DAQ: intramolecular cyclization, thiol addition, amine addition (shown from top to bottom). Note that other DA-derived quinones can undergo similar additions. C) Reactions by DOPAL: imine condensation, oxidation to DPQAL, and protein cross-link formation to due to nucleophile addition into the ring.

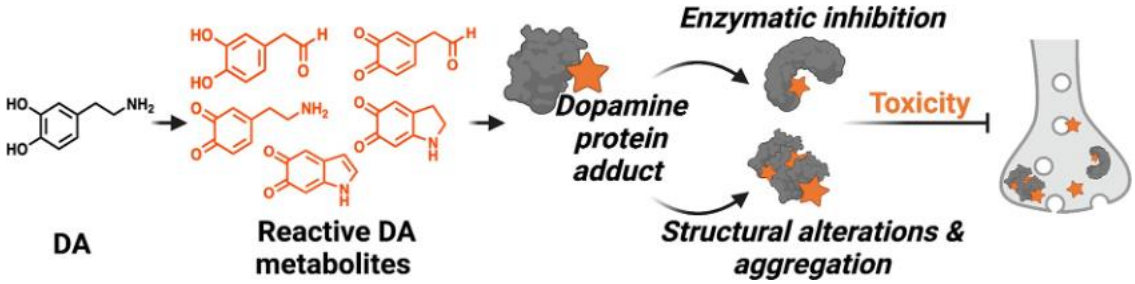


Figure 1.6. Protein damage inflicted by reactive DA metabolites can drive neurotoxicity through enzymatic inhibition and conformational changes

Table 1.1. Reported dopamine-adducted proteins, functional ramifications, and methods of detection

protein	functional consequences of DA modification	method of adduct detection	refs
tyrosine hydroxylase	inhibition of enzymatic activity	radiolabel, NBT staining	86, 87
dopamine transporter	inhibition of transporter activity	mutagenesis	88
nuclear receptor related-1 protein	stimulation of Nurr1 transcriptional activity	X-ray	89
superoxide dismutase 2	inhibition of enzymatic activity	radiolabel, NBT staining, mutagenesis	90
glutathione peroxidase 4	inhibition of enzymatic activity	radiolabel	91
ATP synthase	impairment of mitochondrial ATP synthesis	radiolabel, MS	92

glutathione-*S*-transferase	inhibition of enzymatic activity	NBT staining	93
alpha synuclein	alteration of oligomer and fibril formation	radiolabel, boronate enrichment	94-101
parkin	inhibition of E3 ligase function	radiolabel, boronate enrichment, MS	102, 103
glucocerebrosidase	inhibition of enzymatic activity	nIRF, MS	55
DJ-1	alteration of protein structure	radiolabel, NBT staining, mutagenesis	104
protein disulfide isomerase	inhibition of enzymatic activity	bioorthogonal probe, MS	105, 106

1.4 Tools to Study Dopamine Dysfunction

The chemical toolbox used to interrogate the consequences of dysregulated DA is growing. Some of the existing approaches utilize DA mimetic probes, whereas others rely on the chemical structure and reactivity of DA itself. Analytical methodologies are also critical for studying DA dynamics and dysfunction. In the case of identifying DA-modified proteins, certain methods provide a higher degree of unambiguous evidence for DA adduction than others. For instance, MS/MS analysis of an adducted peptide provides concrete verification that a particular protein is indeed susceptible to DA modification, whereas other approaches (e.g., radiolabeling, immunoprecipitation, redox detection, etc.) provide only presumptive evidence for a DA adduct within a particular protein. Below, we will discuss examples of these classes, highlight how different tools provide valuable insights into DA's role in PD, and underscore the advantages and disadvantages of the given technologies.

DA derivatives come in many modalities (**Figure 1.7**). For example, radiolabeled DA is an effective probe for studying DA-driven protein damage. In pioneering work, ^{14}C-dopamine in combination with autoradiography was used to identify proteins modified by DA metabolites.[152, 153] These studies provided an initial set of DA-adducted proteins to be further investigated. While introduction of ^{14}C in DA is advantageous in terms of imposing minimal structural changes to the DA scaffold, the toxicity and regulatory hurdles associated with ^{14}C hamper the use of radioactive isotopes.[154] Alternative probe based strategies to identify proteins damaged by DA-derived species have leveraged bioorthogonal chemistries to enrich DA-adducted proteins prior to

identification by mass spectrometry.[155] Enrichment of structurally modified protein species is important for their identification due to signal suppression by unmodified peptides during mass spectrometry analyses. For example, alkynylated derivatives of 6-OHDA and DA, dubbed 6-OHDA-PEG3-yne and DA^{yne}, respectively, have been reported (**Figure 1.7**).[117, 118] These researchers have successfully demonstrated the utility of bioorthogonal functionalization of DA as it enabled detection and proteomic identification of proteins modified by DA metabolites when 6OHDA-PEG3-yne was administered to cellular lysates or when DA^{yne} was incubated in cell culture. It should be noted that care must be taken in the design of structurally altered DA mimetic probes as to not drastically alter their physicochemical properties from those of DA. As a testament to this, 6-OHDAPEG3-yne is unable to traverse membranes in living cells whereas DA^{yne} can. As a whole, functionalized DA derivatives make powerful tools for broad-scale proteomic studies of DA adducted proteins.

Additional DA-based probes have been designed for tracking DA dynamics and signaling pathways. These efforts have been recently reviewed.[156] Briefly, ^{18}F fluorodopamine and ^{19}F-LDOPA probes were developed to be used for positron emission tomography (PET) and magnetic resonance imaging (MRI) (**Figure 1.7**).[157, 158] These tools can report on DA localization *in vivo*. Another class of tools are fluorescent false neurotransmitters (FFNs) like FFN102, FFN200, and FFN206 which are substrates for DA handling enzymes and used in confocal microscopy applications (**Figure 1.7**).[159-161] While these molecules do not directly report on protein damage caused by dysregulated DA, they are critical for deciphering DA dynamics and trafficking *in vitro* and *in vivo*.

The second category of tools for studying DA disposition in the brain take advantage of the molecule's redox reactivity (**Figure 1.8**). Early efforts in detecting proteins covalently modified by DA metabolites relied on the redox-sensitive catechol functionality possessed by DA metabolites.[162] Following immobilization of quinone-modified proteins on a membrane, a basic nitroblue tetrazolium (NBT) solution, which redox cycles with the membrane-bound adducted proteins to produce nitroblue formazan (NBF), can be applied. NBF in turn deposits on and stains the membrane purple, enabling visual detection of protein-bound quinones. This technique was utilized in a comparative proteomic study aimed at profiling DA-modified proteins in 2- and 15-month rats.[163] These researchers found that levels of DA-adducted proteins increase with age. This methodology is useful for comparing relative amounts of quinone-modified proteins between samples; however, it is unable to distinguish different species of quinones from one another. Other redox-cycling quinone protein detection strategies have been reported using a dithiothreitol (DTT) and luminol system to generate quantifiable chemiluminescence.[164] Such techniques may have utility for DA metabolites as well. A similar approach for detecting immobilized protein-bound ortho-quinones or free ortho-quinones alone utilizes the spectroscopic properties of these molecules.[165] Here, a near-infrared fluorescence (nIRF) imager is used to excite bound quinones at 685 nm, and the resultant emission from the ortho-quinone is measured at 700 nm. This is selective for ortho-quinones/catechols over meta- and para-substituted analogs. This method enables the detection of oxidized DA at nanomole levels. However, similar to detection via redox cycling, this procedure is unable to differentiate species of protein-bound quinones.

Despite the limitations, these tools are robust methodologies for quantifying global levels of protein damage induced by DA metabolites.

Other tools that utilize the chemical structure of DA to study DA−protein adducts enable enrichment and capture of modified proteins (**Figure 1.9**). One example is the use of boronate agarose resin, which chelates the cis-diols present on DA at alkaline pH and releases them at acidic pH.[166] Experimental workflows based on boronate resins have enabled the capture of DA-modified parkin in human brain samples and DA metabolites modified by AS *in vitro*.[123, 134] It should be noted that this approach is useful for enriching endogenously DA-damaged proteins by any metabolites that retain catechol functionality upon adduction but is prone to capturing other biological cis-diols such as glycosylated proteins as well. Enrichment strategies and other applications of boronate chemistry can be found in a previous review.[167] Another reported strategy to detect proteins susceptible to DA modification used DA-functionalized quantum dots (QDs).[168] Covalent attachment of a protein to the DA anchored on the CdTe/ZnS QDs causes a measurable change in the fluorescence intensity of the QDs, which can report levels of adducted protein captured by the nanoparticles. These DA-functionalized QDs are cell penetrant and thus can provide a relative measure of the amount of proteins vulnerable to DA modification in a particular cell. However, this approach cannot determine the identity of the adducted proteins.

Beyond DA protein adduct detection, there are additional tools to monitor endogenous DA to study DA dynamics in a variety of capacities. Biosensor and chemical sensor tools for DA detection have been recently reviewed.[169, 170] These methods rely on

the intrinsic chemical reactivity of DA for signal generation. Genetically encoded DA sensors that produce fluorescence upon DA binding have enabled recording DA synaptic dynamics in mice.[171, 172] Small molecule chemical sensors have been designed to covalently trap DA, which results in measurable fluorescence changes in the sensor molecule which can be used to track DA location in cells. However, distinguishing between DA and NE by this approach remains challenging.[173] Similar reactivity-based DA-sensing strategies have also been applied to nanoparticle approaches.[174-177] These tools are valuable constituents of the toolbox needed to probe DA dysfunction in PD as they can be applied in animal and cell models of PD.

Analytical chemistry techniques are also invaluable additions to the chemical toolbox focused on studying DA's involvement in PD. Fast scan cyclic voltammetry is a powerful way to quantify DA concentration and dynamics.[178] This electrochemical approach detects the amounts of DA adsorbed onto an electrode by increasing and decreasing the voltage applied and simultaneously measuring the resulting current generated by the oxidation and reduction of the catechol in DA; further details can be found in ref [179]. This methodology has enabled measurement of subsecond DA release *in vivo* and provided valuable insights in how ROS, synaptic vesicle dysfunction, and aging impact PD progression in animal models but has a low throughput and requires a skilled operator.[180-186]

Liquid chromatography-mass spectrometry (LC-MS) is a powerful analytical tool for studying dysregulated DA in PD. LC-MS is the foundation for any proteomic-based study aimed at identifying protein targets of reactive DA metabolites.[187] Typically,

protocols will enrich DA-adducted proteins from cellular lysates through a variety of the methods discussed above or generate DA−protein adducts *in vitro* by incubating a particular protein of interest with reactive DA metabolites prior to LC-MS analysis. Enrichment of adducted proteins is a key component of such analysis. Often adducts are in low abundance in comparison to unmodified proteins, which necessitates enrichment to enable their detection by MS. The enriched pool of DA−protein adducts is commonly digested with a trypsin protease to afford peptide fragments.[188] These peptides are separated by a LC system with an in-line orbitrap mass analyzer. Inside this type of mass spectrometer, the peptides are fragmented to give a series of Y and B ions, which are deterministic of the parent peptide sequences.[189] Sequenced peptides are then searched in various databases to identify their protein of origin.[190] An overview of this process is shown in **Figure 1.10**. Our laboratory has successfully identified 38 DA−protein adducts in SH-SY5Y cells through such an approach, which led us to discover that reactive DA metabolites modify and functionally inhibit PDI.[117] Studies like this improve the understanding of how reactive DA metabolites inflict damage across the proteome.

The specific sites of DA−protein modification can also be elucidated by LC-MS. A notable example of this was highlighted in a study by Burbulla et al., where a specific DA modification was located on a catalytic Cys residue within the active site of GCase.[55] Mapping the sites of modification provides valuable information into the structural motifs that are susceptible to DA adduction and gives insight into the molecular mechanisms by which reactive DA metabolites impair protein function. However, there are still challenges in comprehensively mapping the specific sites of DA adducts across the

proteome, namely, ensuring the peptide containing the DA adduct ionizes well and has a m/z within the analysis range of the spectrometer. In such cases, the locations of DA modification can be inferred through mutagenesis experiments. In these workflows, the hypothesized adducted amino acid is mutated to a nonreactive side chain and the mutant protein is assayed to determine if DA adduction still occurs.

In addition, LC-MS methodology allows for detection of potential PD biomarkers linked to DA dysfunction.[191, 192] 5-SCysteinyl-dopamine, a byproduct of reactive DA metabolites coupling with Cys or GSH, has been reported to be elevated with age and in the brains of PD patients.[193-195] Thus, these DA-derived metabolic adducts may have utility as potential PD biomarkers.[196]

Collectively, there is a large array of chemical tools to study DA. These tools can be leveraged to better understand the roles DA plays in PD progression. Each tool possesses unique advantages and disadvantages that must be kept in mind when designing experiments. The continued expansion and innovation of the DA chemical toolbox will diversify the methodologies researchers have at their disposal to interrogate DA dynamics and will further advance the understanding of functional consequences of DA-induced protein damage.

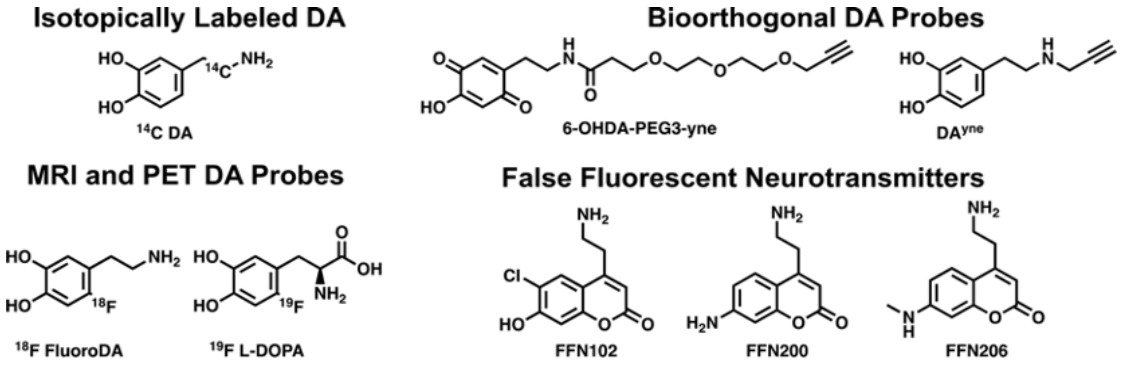


Figure 1.7. Classes and structures of dopamine mimetic probes

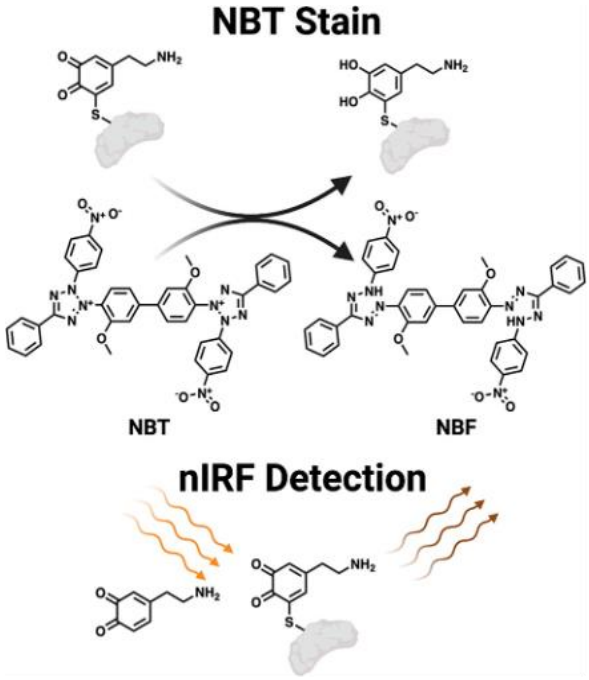


Figure 1.8. Methods to detect dopamine protein adducts

Schematic of the NBT redox cycle with a DA protein adduct to yield NBF, enabling visualization of DA protein adducts (Top). Depiction of the nIRF properties of ortho-quinones which enable the detection of DA protein adducts (Bottom).

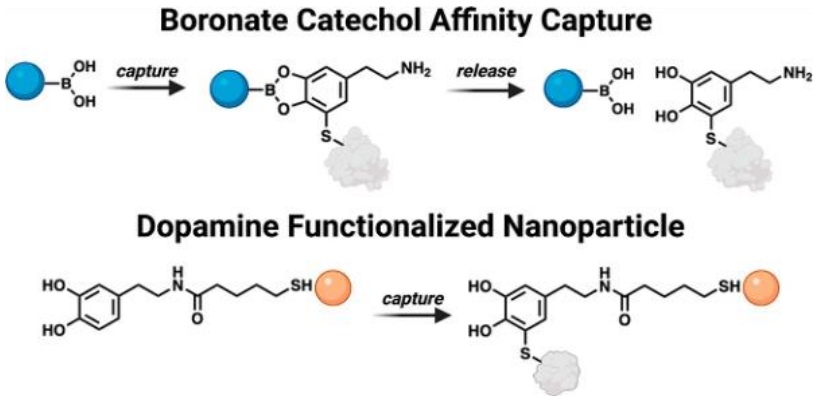


Figure 1.9. Technologies to capture dopamine protein adducts

Schematic of boronate-functionalized resin utilized to immobilize and release proteins containing a DA adduct (Top). Visualization of DA-decorated nanoparticles enabling detection of proteins susceptible to DA modification (Bottom).

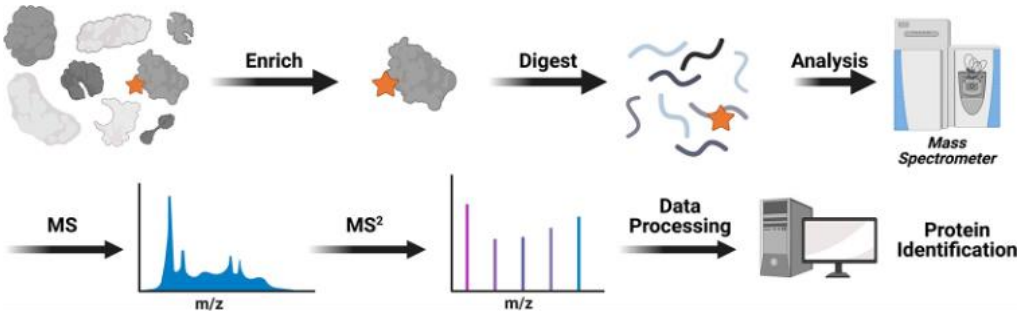


Figure 1.10. Identification of dopamine modified proteins by mass spectrometry

Low-abundance DA protein adducts are enriched via various strategies (e.g., bioorthogonal chemistry and boronate affinity). This enriched set of proteins is digested to peptides and subjected to LC-MS/MS analysis. MS^2 spectra can then be searched against a database to enable protein identification.

1.5 Outlook

PD is a complex and heterogeneous disease. The exact cause of PD remains unknown, and there are many competing hypotheses concerning the origins and molecular events that drive PD pathology.[197] It is hypothesized that intracellular accumulation of misfolded AS is the main driver of neuronal toxicity and PD development.[198-200] There is evidence that misfolded AS originating in the gut can propagate through the gut−brain axis and induce neurodegeneration.[201, 202] Other postulated origins for PD are the age-dependent buildup of oxidative stress, mitochondrial dysfunction, and lysosomal impairment.[203-206] There are also genetic risk factors that predispose individuals to PD.[207, 208] However, a majority of PD cases are idiopathic in origin, and thus, a combination of genetic and environmental factors likely work in concert to influence PD development.[202, 209] In addition, there are particular features that SNc dopaminergic neurons possess which confer vulnerability and predispose them to degradation in PD, such as their long length, low levels of myelination, and high energy consumption.[210, 211] Another risk factor for degeneration of these dopaminergic neurons is the presence of DA itself.

The propensity for DA to degrade into reactive metabolites when improperly displaced from vesicular confinement is why DA dysregulation is a potential player in PD. Rogue DA derived electrophilic metabolites react with a host of proteins, which in turn can inhibit the activity of the adducted proteins and lead to neurotoxicity. Understanding the nuances of these molecular events is important as it provides insights

to the multifaceted etiology of PD, which can aid biomarker discovery for early detection and drive therapeutic development.[212]

Tremendous progress has been made to characterize protein damage caused by reactive DA metabolites. However, there are still many unanswered questions pertaining to DA's contribution to PD. For instance, what percentage of a particular protein within a cell needs to be modified by DA metabolites to drive a phenotypic response? Answering questions like this are especially important in cases like Nurr1 DA adducts, which can influence gene expression. Implementing activity-based protein-profiling (ABPP) strategies capable of reporting percent occupancy of an electrophile bound to a particular protein along with the specific site of modification would be useful toward this end.[213-218] In addition, controlled and targeted release of DA metabolites to particular proteins and locations in cellular systems would assist in clarifying signaling pathways impacted by DA−protein adducts.[219, 220] These types of precise chemical tools have been developed to study the consequences of protein damage caused by lipid-derived electrophiles like 4-hydroxynonenal (HNE) and reactive dicarbonyl metabolites like methylglyoxal (MGO).[221-225] Similar DA tools could be built from reported photoactivatable DA precursors.[226]

Other remaining questions include how does the abundance of DA−protein adducts change over time and what DA metabolites are most influential in driving neuronal dysfunction. Longitudinal studies tracking the formation and identity of proteins modified by DA metabolites over an organism's life correlated to PD incidence would provide insight to how DA induced damage contributes to PD. Continued development of

tools enabling the enrichment and characterization of DA−protein adducts would be valuable assets in this area. Further optimization of the LC-MS methodology that comprehensively identifies DA−protein adducts would aid this effort. Harnessing the unique reactivity of ortho-quinones may prove useful in enriching DA−protein adducts in biological samples through cycloadditions with strained cyclic alkenes.[227] In addition, antibodies raised against particular DA-modified proteins are currently lacking and would also be a useful enrichment and detection tool. To determine the pathological contribution of various DA-derived metabolites, probes akin to the reported DA^{yne} and 6-OHDA-PEG3-yne could be designed to mimic specific oxidation states of DA metabolites.[117, 118] A comprehensive suite of DA-derived probe molecules would enable comparative studies between the complex array of different DA metabolites and illuminate their contribution to DA-induced protein damage.

To conclude, multiple lines of evidence point to DA dysregulation being a contributor to PD. Reactive DA metabolites can elicit neurotoxicity when they are not properly broken down or when their levels overwhelm the natural detoxification machinery. Chemical tools have greatly improved our understanding of DA-induced dysfunction. Continued research effort in the area of DA-induced protein damage in PD and further expansion of the chemical toolbox needed to study these processes will further elucidate the complex disease etiology of PD.

2 Chapter II: Proteome-Wide Profiling of Cellular Targets Modified by Dopamine Metabolites Using a Bio-Orthogonally Functionalized Catecholamine

Adapted with permission by the American Chemical Society from:

Hurben A.K. Erber L.N., Tretyakova N.Y., Doran T.M. Proteome-Wide Profiling of Cellular Targets Modified by Dopamine Metabolites Using a Bio-Orthogonally Functionalized Catecholamine. *ACS Chemical Biology,* **2021**, *16* (11), 2581–2594.

This chapter is adopted from a manuscript published in *ACS Chemical Biology* with contributions from Alexander K. Hurben (synthesis, *in vitro* and cellular experiments, and proteomics) and Dr. Luke Erber (assistance in proteomics), under the supervision of Prof. Tretyakova and Prof. Doran.

2.1 Introduction

As established in **Chapter I** of this thesis, Parkinson's disease (PD) is the second most common neurodegenerative disease, which currently affects five million individuals worldwide.[228] PD is characterized by a progressive decline in motor function due to the selective death of midbrain dopaminergic neurons in the striatum, and in the substantia nigra pars compacta (SNc).[3] The molecular events that initiate dopaminergic collapse in the SNc at the onset of PD occur up to a decade before clinical presentation. Other features of PD that precede symptom onsets include reductions in synaptic density, decreased synaptic dopamine (DA), and loss of DA-derived neuromelanin, suggesting that DA dysregulation is a key factor in PD pathophysiology.[229] Indeed, DA loss correlates with motor dysfunction severity and DA metabolites have recently garnered attention as potential diagnostic biomarkers.[230] Therefore, unraveling the mechanisms by which DA influences PD- related clinical and pathological changes will be critical for developing effective drugs and pre-symptomatic PD diagnostic assays.[15]

Clinical studies have associated PD incidence with impairments in cellular regulation of cytosolic DA levels. A large-scale genome-wide association study (GWAS) meta- analysis identified PD-linked mutations in genes that are involved in DA regulation, such as synaptic vesicle endocytosis.[231] Genetic mutations or pharmacological agents that perturb DA homeostasis are also known to drive neuropathic phenotypes *in vivo*.[232] Conversely, lowering cytosolic DA by increasing vesicular monoamine transporter (VMAT) activity enhances DA release and opposes PD- associated neurodegeneration in mice.[53] This evidence suggests that DA dysregulation plays a

significant role in PD etiology. Functional measures of DA transfer out of the cytosol by VMAT are 87−90% lower in autopsied brains of PD patients as compared to control donors.[50] While the specific mechanisms underlying this observation have not been characterized, they implicate DA toxicity as an early and key component of PD-related neurodegeneration. Time-resolved mechanistic analyses of dopaminergic neurons derived from human induced pluripotent stem cells found that DQs accumulate in the neurons of PD donors, but not healthy donors.[55] Additionally, these oxidized DA metabolites induced mitochondrial and lysosomal dysfunction, two cellular processes involved in early mechanisms of PD-associated cell death. Collectively, these studies support a causal role for DA in neuronal dysfunction and degeneration.

In addition to its receptor-binding activity, DA is also chemically reactive toward biomolecules under conditions that mirror cytosolic environments. The chemical reactivity of DA derives from its electron-rich catechol moiety, which renders DA prone to auto-oxidation outside of acidic vesicles.[71] These oxidative processes produce a series of electrophilic dopaquinone (DQ) metabolites through sequential oxidative transformations (**Figure 2.1**). To prevent oxidative damage resulting from a surge in cellular DQ metabolites, DA levels are maintained through metabolic clearance, repackaging into synaptic vesicles and polymerization into neuromelanin.[233]

Dysfunction in any of these regulatory processes could result in high levels of DQs. DQs induce oxidative damage by increasing reactive oxygen species (ROS), a byproduct of DA-DQ redox cycling. Additionally, DQs nonenzymatically modify protein nucleophiles to form DA-protein adducts (DPAs).[26] These modifications can occur at

low-occupancy particularly at the active site of enzymes, resulting in alterations to the native protein structure and functions that affect critical cellular processes. Many PD-implicated proteins that are susceptible to DA modification have been identified using target-based approaches.[86, 88, 89, 100, 103, 124, 125, 128, 129, 134, 139, 234] For example, interactions between DQ and α-synuclein (αsyn) promote the self-assembly of toxic oligomers, leading to neurodegeneration in mouse models.[124, 125, 129] DQ modification of parkin found in human brain tissue samples reportedly inhibits its E3 ubiquitin ligase function, which can lead to proteotoxic stress underlying PD- related neurodegeneration. Finally, covalent dimerization of DJ-1, a protein linked genetically to PD susceptibility, occurs in the presence of DQs, although its biological significance *in vivo* remains unclarified.[134] These cases exemplify target-based approaches to identify DPAs, which rely on foreknowledge of well-established PD targets. Conversely, few studies have successfully applied unbiased approaches to characterize and validate novel DPAs at the proteome-wide level.[152, 153, 163] This is largely because DA reactivity vitiates the sensitivity of analytical platforms capable of comprehensively profiling DPAs. Developing sensitive analytical characterization methods for reactive metabolites of DA guarantees to enhance our understanding of how DPAs contribute to aging and early stages of PD progression.

To access caches of candidate DPAs involved during early stages of PD pathology, we developed a chemo-proteomic probe that enables DPA profiling at low micromolar exposures of DA in human cells. In this proof-of-concept study, we synthesized a novel dopamine-mimetic probe, which we call DA^{yne}, containing a bio-orthogonal alkyne handle for proteomic enrichment of DPAs. Herein, we show that

DA^{yne} is a suitable biochemical mimetic of DA and demonstrate the utility of DA^{yne} as a visualization and enrichment probe enabled by the versatility of click chemistry modification. We highlight the generality and impact of our newly developed chemical probe by identifying DPAs and extracting cellular processes likely perturbed in the early stages of PD pathogenesis using chemoproteomics. Further, we provide validation of our findings by demonstrating that one of the proteins identified as a target of DA^{yne} in our proteomic analyses, protein disulfide isomerase (PDI) involved in endoplasmic reticulum (ER) stress, is functionally inactivated by DA. These studies demonstrate that DA^{yne} is a useful reactivity-based probe molecule that enables rapid discovery and characterization of cellular DPAs using a suite of qualitative and quantitative assays.

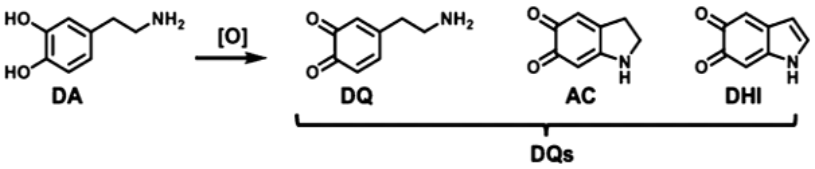


Figure 2.1. Simplified schematic of dopamine oxidation to electrophilic DQs metabolites Shown are the 2 e−, 4 e−, and 6 e− oxidation products of DA; dopa-o-quinone (DQ), aminochrome (AC), and oxidized dihydroxyindole (DHI), respectively.

2.2 Materials and Methods

All chemicals required for the synthesis of DA^{yne}, preparation of buffers, mushroom tyrosinase (Lot#: SLBZ0022), NBT, Amicon ultra 0.5 mL centrifugal filters (3k) and click reactions were purchased from Sigma-Aldrich unless otherwise stated. NMR solvents were obtained from Cambridge Isotope Laboratories Inc. SH-SY5Y neuroblastoma cells were acquired from ATCC. Cell culture media, additives, and consumables were purchased from Corning. High-capacity streptavidin agarose resin, Zeba Spin Desalting Columns (7K MWCO, 0.5 mL), PDIA3 (ERp57) antibody (Cat. no. CL2444), and DAPI were obtained from Thermo Scientific. alamarBlue HS reagent and ProLong Glass were purchased from Invitrogen. BCA kits and C-18 spin columns were bought from Pierce. Mini-PROTEAN TGX Precast Gels, 0.2 μm nitrocellulose, and filter paper transfer stacks were obtained from BioRad. Biotin azide, THPTA, Alexa Fluor 647 azide, and azide agarose beads were purchased from Click Chemistry Tools. REVERT total protein stain kit, Odyssey TBS blocking buffer, and IRDye800CW streptavidin were purchased from LI-COR. Recombinant PDIA3 and PDI inhibitor screening kit (fluorometric) were obtained from BioVision (Cat. no. 7601-100 and K840-100, respectively).

Synthetic Procedures:

General Procedures:

Chemical reactions were conducted in a fumehood with oven-dried glassware under an argon atmosphere unless otherwise stated. Thin-layer chromatography (TLC) was performed with analtech silica uniplates and visualized under 254 nm UV light.

Column chromatography was conducted with 60 mesh silica gel. A Buchi R-200 rotary evaporator and V-100 vacuum pump were used to remove solvent. NMR spectra were taken on a 400 MHz Varian spectrometer. HRMS spectra were taken on a Burker ESI-BioTOF II instrument and internally calibrated with PEG 300. The synthetic route to DA^{yne} is shown in **Figure 2.2**.

N-(3,4-dihydroxyphenethyl)-2,2,2-trifluoroacetamide **2**.

Dopamine HCl (3.42 g, 18 mmol, 1 eq) was dissolved into 100 mL of MeOH. The resultant solution was purged with argon for 20 minutes, then methyl trifluoroacetate (3.63 mL, 36 mmol, 2 eq) was added dropwise, followed by Et_3N (10 mL, 72 mmol, 4 eq). The reaction was allowed to proceed for 1 hour at RT. Solvent was removed *in vacuo* and the residue was dissolved in EtOAc, washed with 0.5 M HCl, water, brine, and dried with Mg_2SO_4. EtOAc was removed by rotary evaporation and the product was crystalized in CH_2Cl_2 to afford **7** (4.22 g, 93% yield) as fine white crystals. ^{1}H NMR (400 MHz, METHANOL-d_4) δ ppm 2.67 (2 H, t, *J*=7.63 Hz) 3.40 (2 H, t, *J*=7.43 Hz) 6.50 (1 H, dd, *J*=7.83, 1.96 Hz) 6.62 (1 H, d, *J*=1.96 Hz) 6.66 (1 H, d, *J*=7.83 Hz). ^{13}C NMR (101 MHz, METHANOL-d_4) δ ppm 33.83 (1 C, s) 41.17 (1 C, s) 114.96 (1 C, s) 115.40 (1 C, s) 117.51 (1 C, s) 119.60 (1 C, s) 129.82 (1 C, s) 143.52 (1 C, s) 144.90 (1 C, s) 157.64 (1 C, s). HRMS $C_{10}H_{10}F_3NNaO_3^+$ calc 272.0505 found 272.0481

N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-2,2,2-trifluoroacetamide **3**.

Compound **2** (2 g, 8 mmol, 1 eq) and 2,2-dimethoxypropane (3.9 mL, 32 mmol, 4 eq) was add to 100 mL of toluene and heated to reflux to fully dissolve. *p*-Toluenesulfonic acid (68.8 mg, 0.4 mmol, 0.05 eq) was added and the solution was

refluxed for 3 hours. Solvent was removed *in vacuo* and recrystallization in hexanes afforded **3** (1.82 g, 79% yield) as dark crystalline solid. ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.65 (6 H, s) 2.76 (2 H, t, *J*=6.85 Hz) 3.54 (2 H, q, *J*=6.65 Hz) 6.23 - 6.38 (1 H, m) 6.53 - 6.58 (2 H, m) 6.63 - 6.67 (1 H, m). ^{13}C NMR (101 MHz, CHLOROFORM-*d*) δ ppm 25.80 (1 C, s) 34.67 (1 C, s) 41.14 (1 C, s) 108.31 (1 C, s) 108.60 (1 C, s) 118.05 (1 C, s) 120.97 (1 C, s) 130.41 (1 C, s) 146.44 (1 C, s) 147.91 (1 C, s) 157.26 (1 C, s). HRMS $C_{13}H_{15}F_3NO_3^+$ calc 290.0999 found 290.1028

2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethan-1-amine **4**.

Compound **3** (2.00 g, 6.92 mmol, 1 eq) was dissolved into MeOH (50 mL). LiOH (0.66 g, 27.68 mmol, 4 eq) was dissolved into 25 mL of H_2O and added dropwise. The resultant solution was stirred for 10 minutes at RT. MeOH was removed *in vacuo* and the residue was dissolved in brine and extracted with EtOAc, dried with $MgSO_4$, and concentrated to give **4** (1.30 g, 98%) as a dark brown viscous oil. ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.64 (6 H, d, *J*=1.17 Hz) 2.63 (2 H, t, *J*=6.65 Hz) 2.89 (2 H, t, *J*=6.65 Hz) 6.55 - 6.66 (3 H, m). ^{13}C NMR (101 MHz, CHLOROFORM-*d*) δ ppm 25.80 (2 C, s) 39.58 (1 C, s) 43.60 (1 C, s) 107.95 (1 C, s) 108.86 (1 C, s) 117.57 (1 C, s) 120.98 (1 C, s) 132.73 (1 C, s) 145.76 (1 C, s) 147.48 (1 C, s). HRMS $C_{11}H_{16}NO_2^+$ calc 194.1176 found 194.1186

N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-4-nitrobenzenesulfonamide **5**.

Compound **4** (730 mg, 3.7 mmol, 1.1 eq) and 4-nitrobenzenesulfonyl chloride (760 mg, 3.4 mmol, 1 eq) were dissolved in 10 mL CH_2Cl_2, placed under N_2, and cooled to 0 °C. Et_3N (0.514 mL, 3.7 mmol, 1.1 eq) was added dropwise and the reaction was

stirred at 0 °C for 5 minutes and allowed to warm to RT and stirred for an additional 15 minutes. 15 mL of sat. NH_4Cl was added and the reaction mixture was extracted with CH_2Cl_2, washed with brine, dried with $MgSO_4$, and concentrated *in vacuo* to afford **5** (1.25 g, 97%) as a yellow solid. ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.63 (6 H, s) 2.67 (2 H, t, *J*=6.65 Hz) 3.17 - 3.27 (2 H, m) 4.61 (1 H, br. s.) 6.37 - 6.49 (2 H, m) 6.59 (1 H, d, *J*=7.83 Hz) 7.92 - 8.00 (2 H, m) 8.27 - 8.35 (2 H, m).^{13}C NMR (101 MHz, CHLOROFORM-*d*) δ ppm 25.79 (2 C, s) 35.47 (1 C, s) 44.41 (1 C, s) 108.31 (1 C, s) 108.53 (1 C, s) 118.13 (1 C, s) 121.07 (1 C, s) 124.32 (2 C, s) 128.22 (2 C, s) 129.91 (1 C, s) 145.83 (1 C, s) 146.50 (1 C, s) 147.86 (1 C, s) 149.99 (1 C, s). HRMS $C_{17}H_{17}N_2O_6S^-$ calc 377.0813 found 377.1297

N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-4-nitro-*N*-(prop-2-yn-1-yl)benzenesulfonamide **6**.

Compound **5** (1.2 g, 3.17 mmol, 1 eq) and potassium carbonate (876 mg, 6.35 mmol, 2 eq) were dissolved in 50 mL of acetone under N_2. Propargyl bromide solution 80 wt. % in toluene (.374 mL, 3.8 mmol, 1.2 eq) was added dropwise and the resultant solution was refluxed for 12 hours. Solvent was removed *in vacuo* and the residue was dissolved in H_2O and extracted with CH_2Cl_2, washed with brine, dried with $MgSO_4$, and subjected to column chromatography hexanes:EtOAc v/v gradient of 4:1 to 1:1) to afford **6** (1.27 g, 96%) as a dark amber oil. ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.62 - 1.66 (6 H, m) 2.05 (1 H, d, *J*=2.74 Hz) 2.81 (2 H, dd, *J*=8.41, 6.85 Hz) 3.39 (2 H, dd, *J*=8.22, 7.04 Hz) 4.14 (2 H, d, *J*=2.35 Hz) 6.53 - 6.66 (3 H, m) 7.95 - 8.04 (2 H, m) 8.26 - 8.35 (2 H, m) ^{13}C NMR (101 MHz, CHLOROFORM-*d*) δ ppm 25.78 (2 C, s) 34.41 (1 C,

s) 36.86 (1 C, s) 48.30 (1 C, s) 74.41 (1 C, s) 75.89 (1 C, s) 108.25 (1 C, s) 108.85 (1 C, s) 117.94 (1 C, s) 121.06 (1 C, s) 124.06 (1 C, s) 128.87 (1 C, s) 130.49 (1 C, s) 144.70 (1 C, s) 146.29 (1 C, s) 147.64 (1 C, s) 150.07 (1 C, s). HRMS $C_{20}H_{19}N_2O_6S^-$ calc 415.0969 found 415.2538

N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)prop-2-yn-1-amine **7**.

Potassium hydroxide (328 mg, 5.8 mmol, 2 eq) predissolved into a 10 M aq. solution and thiophenol (0.654 mL, 5.8 mmol, 2 eq) were combined in 25 mL of CH_3CN and cooled to 0 °C. **6** (1.22 g, 2.9 mmol, 1 eq) dissolved in 25 mL of CH_3CN was added. The solution was heated to 50 °C for 40 minutes. The reaction was quenched with H_2O, extracted with CH_2Cl_2, washed with brine, dried with $MgSO_4$, and subjected to column chromatography (hexanes:EtOAc v/v gradient of 4:1 to 1:1) to afford **7** (338 mg, 50%) as a yellow oil. ^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.62 - 1.66 (6 H, m) 2.16 - 2.21 (1 H, m) 2.71 (2 H, t, *J*=7.04 Hz) 2.90 (2 H, t, *J*=7.04 Hz) 3.39 - 3.45 (2 H, m) 6.57 - 6.66 (3 H, m). ^{13}C NMR (101 MHz, CHLOROFORM-*d*) δ ppm 25.83 (2 C, s) 35.72 (1 C, s) 38.04 (1 C, s) 49.87 (1 C, s) 71.44 (1 C, s) 81.84 (1 C, s) 108.01 (1 C, s) 108.78 (1 C, s) 117.61 (1 C, s) 120.88 (1 C, s) 132.66 (1 C, s) 145.81 (1 C, s) 147.50 (1 C, s). HRMS $C_{14}H_{18}NO_2^+$ calc 232.1332 found 232.1296

4-(2-(prop-2-yn-1-ylamino)ethyl)benzene-1,2-diol **8a** (DA^{yne}).

Compound **7** (19.5 mg, 0.11 mmol, 1 eq) was dissolved in $CHCl_3$ (2 mL) and H_2O (50 µL). TFA (400 µL) was added and the resultant solution was stirred under Ag at RT overnight. Solvent was removed *in vacuo* followed by the addition of 1 mL of 0.1 N HCl in THF which was also removed by rotatory evaporation. The reside was then dissolved

aqueous 0.1 N HCl and lyophilizatied to afford **8a** (DA^{yne}) (21 mg, quant.) as white solid. ^{1}H NMR (400 MHz, METHANOL-d_4) δ ppm 2.84 (2 H, t, *J*=7.63 Hz) 3.16 - 3.22 (1 H, m) 3.27 (2 H, t, *J*=7.83 Hz) 3.92 (2 H, d, *J*=2.74 Hz) 6.55 - 6.62 (1 H, m) 6.70 (1 H, d, *J*=1.96 Hz) 6.74 (1 H, d, *J*=7.83 Hz) ^{13}C NMR (101 MHz, METHANOL-d_4) δ ppm 32.89 (1 C, s) 37.72 (1 C, s) 74.81 (1 C, s) 79.51 (1 C, s) 117.09 (1 C, s) 117.13 (1 C, s) 121.35 (1 C, s) 129.06 (1 C, s) 146.02 (1 C, s) 147.12 (1 C, s). HRMS $C_{11}H_{14}NO_2^+$ calc 192.1019 found 192.1078

alamarBlue Cell Viability Assay:

Undifferentiated SH-SY5Y cells were cultured in 1:1 MEM:F12 media containing 10% FBS and Pen/Strep 100 U/mL under a 5% CO2 atmosphere at 37 °C. Cells were seeded into 96 well plates at a density of 20,000 cells per well. 24 h after seeding, the media were removed and replaced with the indicated treatment condition in 1:1 MEM:F12 media absent of FBS, 0.2% DMSO. Following the allotted incubation time, the plates were spun at 1000 g for 5 min and the media were removed and replaced with 100 μL of a 10% alamarBlue solution in FBS free 1:1 MEM:F12. The plates were incubated for 4 h in the dark at 37 °C and then fluorescence was measured at Ex/Em 560/590 nm on a BioTek Synergy H1 plate reader. Cell viability for a given treatment condition was calculated by normalizing the fluorescence signal to the DMSO treated control.

UV−vis Spectroscopy:

Fresh stock solutions of DA, DA^{yne}, $NaIO_4$, and TY were prepared in Millipore Milli-Q water or PBS (pH 7.4) that had been purged for 15 min under argon. The

oxidation reactions were initiated by combining the reactants in a 2 mL Beckman quartz cuvette prior to reading. The spectra were acquired every 30 s for 30 min on a Thermo Scientific NanoDrop One.

NMR Spectroscopy:

Spectra were obtained on a Varian 400 MHz spectrometer. Fresh stock solutions of DA, DA^{yne}, and $NaIO_4$ were prepared in deuterated PBS (pH 7.4, pD uncorrected) that had been purged for 15 min with argon. The oxidation reactions were initiated by combining the reactants in an NMR tube. Reactions were allowed to proceed for 4 h at 25 °C and monitored using a data accumulation of eight scans for each spectrum.

Click Reaction and Western Blot:

Undifferentiated SH-SY5Y cells were cultured in 6 well plates to 90% confluency with 1:1 MEM:F12 media containing 10% FBS and Pen/Strep 100 U/mL under a 5% CO_2 atmosphere at 37 °C. The media was removed and replaced with the indicated treatment condition in 1:1 MEM:F12 media absent of FBS, 0.2% DMSO. Following the allotted incubation time, the cells were washed three times with PBS and lysed in 200 μL of MPER containing HALT protease and phosphatase inhibitor cocktail. The lysates were centrifuged at 17,000 g for 12 min and the supernatant was incubated with 10 μL of high capacity streptavidin agarose resin (Cat: 20359; Lot; TK27535) at 4 °C for 16 h to remove naturally biotinylated proteins. Samples were then centrifuged at 1000 g for 5 min and the supernatant was carefully aspirated. Protein concentrations were determined through a BCA assay per manufacturer's protocol.

Click reactions were performed in 200 μL of MPER at a protein concentration of 100 μg/mL with 25 μM biotin azide, 1.5 mM THPTA, 1 mM CuSO4, and 3 mM Na ascorbate. The reactions were agitated at 25 °C for 1 h in the dark. Proteins were precipitated with 800 μL cold acetone overnight at −20 °C. Following 20 min of centrifugation at 17,000g, acetone was carefully aspirated, and the protein pellets were dissolved into 15 μL of 1× LI-COR loading buffer with 10% BME and denatured for 5 min at 95 °C. Each sample was loaded using 13 μL of sample into 4−20% gradient Mini-PROTEAN TGX Precast Gels for SDS-PAGE. The gels were then blotted onto 0.2 μm nitrocellulose with a BioRad Trans-Blot Turbo apparatus. Blots were stained with REVERT total protein stain per manufacturer's protocol and imaged by a LI-COR Odyssey CLX. This was followed by a 1 h block at RT in Odyssey TBS blocking buffer and incubated with IRDye800CW streptavidin (Lot#: C80918-11) at a 1:10,000 dilution for 1 h. The membranes were then washed and imaged by a LI-COR Odyssey CLX. Western blot quantification was performed with Image Studio Software following the manufacturer's protocol for normalization with REVERT total protein stain.

NBT Staining: [162]

Nitrocellulose blots were prepared previously described. Immediately following transfer, blots were washed twice with water and incubated with a 0.75 mM nitro blue tetrazolium (NBT), 2 M potassium glycinate buffer (pH 10) in the dark for 45 mins at room temp. The NBT solution was removed and the blot was washed with 160 mM sodium borate solution two times. The blot was further incubated in the 160 mM sodium

borate overnight. The blot was then washed with water and imaged with an optical camera.

nIR Detection: [165]

Nitrocellulose blots were prepared previously described. Immediately following transfer, blots were washed twice with water and imaged with a Licor Odyssey CLX.

Confocal Microscopy:

Undifferentiated SH-SY5Y cells were cultured on Neuvitro GC-25-1.5-Laminin coverslips to 90% confluency with 1:1 MEM:F12 media containing 10% FBS and Pen/Strep 100 U/mL under a 5% CO_2 atmosphere at 37 °C. The media were removed and replaced with the indicated treatment condition in 1:1 MEM:F12 media absent of FBS, 0.2% DMSO. Following the allotted incubation time, the cells were washed three times with 2 mL PBS and fixed with 2 mL 4% PFA (molecular biology grade) in PBS and washed an additional three times with PBS. In-cell click reactions were performed by incubating the cells in 0.1% Triton X-100 PBS (PBT) solution containing 10 μM Alexa Fluor 647 azide, 1.5 mM THPTA, 1 mM $CuSO_4$ and 50 mM Na ascorbate for 1 h. The cells were washed three times with PBT, stained with 300 nM of DAPI for 30 min, and washed an additional three times with PBT. A drop of ProLong Glass anti-photobleaching solution was placed on the coverslips and was allowed to cure for 48 h. Coverslips were mounted to glass slides, sealed with nail polish and stored at 4 °C until imaging. Images were acquired on an Olympus FluoView FV1000 BX2 upright confocal microscope using a PLAPON 60 × O NA:1.42 objective. Laser Ex/Em were set at 405:461 nm and 635:668 nm for DAPI and Alexa Fluor 647, respectively. Acquisition

parameters for example voltage, gain, offset, ect were optimized for 10 μM DA^{yne} treatment and kept constant when imaging other conditions.

Proteomics:

Probe Treatment. SH-SY5Y cells in 1:1 MEM:F12 media containing 10% FBS and Pen/Strep 100 U/mL were grown to 90% confluency in T75 flasks. The cells were then washed with PBS and incubated with 1:1 MEM:F12 media (no FBS) containing 10 μM DA^{yne} or vehicle for 12 h at 37 °C. The cells were then washed with PBS, suspended in a trypsin-containing buffer, pelleted at 500 g for 5 min, and further washed three times in PBS. The cells were lysed with 200 μL of MPER containing HALT protease and phosphatase inhibitors for 10 min at RT per the manufacture's protocol. The samples were subjected to centrifugation at 17,000 g for 12 min to remove cellular debris and the protein concentration was determined by BCA.

On-Bead Click Reaction. Click Chemistry Tools azide agarose beads (100 μL of Cat: 31038-2; Lot: 2333) were washed twice with 1 mL of water. The click reaction was conducted by adding a 1 mL solution of 100 μg of protein lysate, 1.5 mM THPTA, 1 mM $CuSO_4$, 3 mM Na ascorbate in MPER to the beads. The reaction was allowed to proceed for 16 h before the beads were pelleted by centrifugation and the supernatant was aspirated.

Washing of Agarose Beads. The beads were washed once with water and the proteins bound to the beads were reduced by addition of 100 mM dithiothreitol (DTT) in 1 mL of 1% SDS, 5 mM ethylenediaminetetraacetic acid (EDTA), 250 mM NaCl, 100 mM Tris pH 8.0 and incubation at 70 °C for 15 min. After pelleting the beads and

removing the supernatant, alkylation was performed with 40 mM iodoacetamide in the dark for 30 min in the same buffer as the previous step. Following centrifugation and removal of the super- natant, the beads were washed five times with 1 mL of 1% SDS, 5 mM EDTA, and 100 mM NaCl in PBS; then, 10 times with 8 M urea in 100 mM Tris pH 8.0; and finally, 10 times with 20% CH_3CN.

On Bead Digestion. Digestion was performed by suspending the beads in 200 μL of 100 mM Tris, 2 mM $CaCl_2$ pH 8.0 and the addition 1 μg of trypsin. The reaction was rotated at 37 °C overnight. The supernatant was transferred to a new tube and combined with two additional 300 μL water 0.1% formic acid washes. Samples were desalted with Pierce C-18 spin columns (Cat: 69725; Lot: TJ276015) per the manufacturer's instruction with MS grade water, CH_3CN, and formic acid. Samples were concentrated by vacuum evaporation and stored at −20 °C until LC−MS/MS analysis

LC−MS/MS Analysis. Tryptic peptides from each biological replicate were analyzed by HPLC-electrospray ionization + -MS/ MS using a Thermo Scientific Q Exactive Hybrid Quadrupole- Orbitrap Mass Spectrometer in line with an Eksigent NanoLC-Ultra 2D HPLC system, a nanospray source, and Xcalibur 4.1.31.9 software for instrument control. Dried peptide samples were reconstituted in 20 μL of 2% CH_3CN; 0.1% formic acid and 4 μL were separated on an in-house packed C18 column (15 cm × 75 μm, Luna C18) over a 90 min gradient with degassed HPLC buffer; buffer A (0.1% formic acid); buffer B (CH_3CN; 0.1% formic acid). Source parameters detailed a spray voltage of 4.0 kV, a capillary temperature of 320 °C, and an S-lens RF level of 55%. MS analysis was performed in a top 12 data-dependent mode. The survey scan was performed

using a resolution setting of 70,000 at 200 m/z with an AGC target of 1 × 106. Dynamic exclusion was enabled with an exclusion duration for 30 s. The 12 most intense precursor ions (excluding singly charged species) were selected for higher-energy C-trap dissociation fragmentation using a collision energy of 30%. The fragmentation scan was performed using a resolution of 17,500 with a AGC target of 5 × 104 or maximum injection time of 50 ms and an isolation window of 2.0 m/z.

Data Analysis. MaxQuant (version 1.6.5.0) was used to search the raw mass spectrometry data. Default parameters were applied and the human Uniprot database (downloaded on 2019/05/15, with 74,349 sequences) was searched against.[190] Trypsin was selected as the protease with allowance for a maximum of two missing cleavages. Variable modifications of methionine oxidation, protein N-terminal acetylation and the fixed modification of cysteine carbamidomethylation were included. The precursor ion and fragmentation mass tolerance of 4.5 ppm and 0.5 Da were selected, respectively. The false discovery rate (FDR) was set to 0.01 at the protein, peptide, and site levels with a minimum peptide length of 6 and a minimum Andromeda score of 40. A minimum of two peptides were required for protein identification. To obtain label-free protein quantification, default LFQ parameters were selected with normalized protein intensities as the output. The Perseus software suite (version 1.6.6.0) was used to process the resultant LFQ intensities.[235] The data were Log2 transformed and filtered following previously described methods.[236] A two-tailed, two-sample t-test was performed to compare protein abundance between groups. Statistically significant enrichment was

determined with a Benjamini-Hochberg corrected FDR of 0.05 and a minimal coefficient of variation (S0) of 2.0.

PDIA3 Modification by DA^{yne} and DA:

Click Reaction with Recombinant PDIA3 Protein. Aliquots of PDIA3 (10 μL of a 1 mg/ mL solution in 50 mM Tris buffer, pH 8) were treated with the indicated concentration of preoxidized DA^{yne} for 1 h at 37 °C, the final reaction volume was 50 μL. For the DA competition experiments, PDIA3 was incubated with preoxidized DA for 30 min prior the addition of DA^{yne}. DA^{yne} and DA oxidation was conducted in 500 μL Tris buffer with 600 units of TY and 0.5 mM of DA or DA^{yne} where TY was removed after a 30 min incubation at RT with an amicon ultra 0.5 mL centrifugal filter (3k). Following PDIA3 treatment with DA and DA^{yne}, excess unreacted small molecules were removed with Zeba spin desalting columns (7K MWCO, 0.5 mL), and the samples were subjected to a CuAAC click reaction with Alexa Fluor 647 azide under the same conditions described above. Excess click reaction reagents were removed with another Zeba desalting column and samples were analyzed through SDS-PAGE.

DA^{yne} Meditated Capture of PDIA3 in Lysates:

Aliquots of SH- SY5Y lysate (50 μL of a 2 mg/mL solution) were treated with the indicated concentration of preoxidized DA^{yne} for 1 h at 37 °C, the final reaction volume was 100 μL. For the DA competition experiments, lysate was incubated with preoxidized DA for 30 min prior to the addition of DA^{yne}. The oxidation procedure for DA and DA^{yne} is described above. Following incubation, excess DA and DA^{yne} were removed with Zeba spin desalting columns, a CuAAC click reaction with biotin azide was performed, and the

excess click reaction reagents were removed with another Zeba desalting column. The samples were then incubated with 50 μL of high capacity streptavidin agarose resin for 15 min at RT to capture DPAs. The beads were then washed five times with PBS and then boiled in 1× gel loading buffer to release DPAs. Samples were separated by SDS-PAGE, transferred to nitrocellulose membranes and blotted for PDIA3.

PDI Activity Assay:

The PDI inhibitor screening kit obtained from BioVision was used following the manufacturer's instructions. Briefly, 2× solutions of the indicated treatments were prepared in the provided assay buffer. 50 μL of these solutions were placed into a black 96 well plate. 45 μL of additional assay buffer and 5 μL of PDI enzyme stock solution was then added and the plate was incubated at 37 °C for 30 min. Next, a 45 μL PDI substrate solution containing oxidized insulin, DTT, and fluorescence turn on probe molecule that binds to reduced insulin was added. The plate was then placed in a plate reader at 37 °C where the fluorescence was measured at Ex/Em 440/490 (indicative of reduced insulin formation) every minute for 30 min. PDI activity was calculated by normalizing the slope of a given treatment to DMSO control.

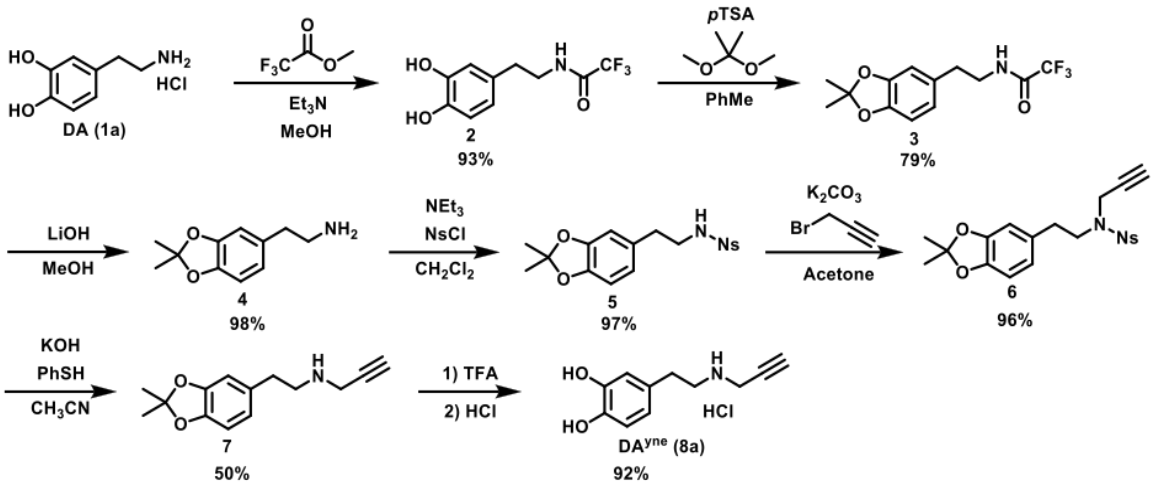


Figure 2.2. Synthesis of DA^{yne}

2.3 Results

DA^{yne} was designed to functionally mimic DA while presenting a bio-orthogonal alkyne handle distal to the o-quinone for click chemistry modification to enable fluorescence visualization and adduct enrichment preceding analytical characterization (**Figure 2.3A**). The DA mimetic probe was synthesized in seven steps as shown in **Figure 2.3B**. Briefly, the catechol was protected with an acid- labile acetonide to enable elaboration of the amine.[237] A pendant propargyl moiety was installed at the amine terminus through a Fukuyama-inspired route.[238] Global acidic deprotection furnished DA^{yne} as a HCl salt. The overall reaction sequence gave DA^{yne} in a 32% yield.

The oxidation pathways of DA^{yne} and DA were monitored spectroscopically to characterize the reaction intermediates (**Figure 2.4A**) and kinetics upon oxidation. We used UV−vis spectroscopy to monitor DQ generation over time. Initially, we monitored the oxidation of DA and DA^{yne} in the presence of equimolar $NaIO_4$ (**Figure 2.4B,C**) in argon- purged water. As shown in **Figure 2.4B**, an increase in absorbance at 395 nm was observed at the reaction onset, indicating the formation of the o-quinone 1b. This gave rise to the appearance of two new maxima at 300 and 475 nm, which are characteristic absorbance bands for aminochrome 1d.[41] However, no o-quinone 8b was observed in the oxidation of DA^{yne} as the reaction proceeded rapidly to the aminochrome 8d (**Figure 2.4C**). Kinetically, DA^{yne} cyclized to the aminochrome 8d five times faster than DA cyclized to 1d in the presence of $NaIO_4$ (**Figure 2.4D**), which we attribute to the increased nucleophilicity of the amine due to alkyl substitution. $NaIO_4$ oxidation in phosphate-buffered saline (PBS) (pH 7.4) 4 led to DA and DA^{yne} oxidation products that

were unstable under these conditions, which was evident by the rapid decay of the respective aminochrome species (**Figure 2.4E**). In the absence of $NaIO_4$, the rates of DA and DA^{yne} auto-oxidation at pH 7.4 were comparable and auto-oxidation could be inhibited in the presence of an antioxidant, N-acetylcysteince (NAC), (**Figure 2.5**) when aminochrome absorbance was monitored at 475 nm.

Given the instability of DA metabolites observed in PBS using $NaIO_4$ as an oxidant, we confirmed that DA^{yne} is a substrate for the enzyme tyrosinase (TY), which is used routinely in DA oxidation studies.[239, 240] TY oxidizes DA to DQs and can be conveniently separated from resultant DQs by molecular weight centrifugal filters. Therefore, TY is an attractive oxidant as it obviates the use of chemical oxidizing agents such as $NaIO_4$ which, due to its strong and promiscuous oxidation potential and ability to chelate vicinal oxygens, may create artifacts in downstream cellular experiments. The absorbance of DA and DA^{yne} monitored at 300 and 475 nm over a 1 h incubation (**Figure 2.6**) shows that DA and DA^{yne} oxidize at similar rates under treatment with TY. Collectively, these UV−vis experiments indicate that DA and DA^{yne} are oxidized at comparable rates in the presence of TY and during auto-oxidation in PBS. The ability of TY to bind and oxidize DA^{yne} also demonstrates the suitability of DA^{yne} as a DA-mimetic as a protein binding substrate.

As an additional confirmation that DA and DA^{yne} feature comparable reaction directionalities, oxidation intermediates of both were characterized structurally using time-resolved nuclear magnetic resonance (NMR) spectroscopy (**Figure 2.4F,G**). The NMR spectral fingerprints of intermediates 1d and 1e were observed upon DA oxidation

using $NaIO_4$ in deuterated PBS as the oxidant. Proton NMR shifts were consistent with literature values.[45] Oxidation of DA^{yne} produced the metabolites 8d and 8e, mirroring those generated by DA (**Figure 2.7**). Taken together, our NMR and UV−vis spectroscopic data show that DA and DA^{yne} are oxidized to chemically analogous electrophilic species, namely, 1d and after long exposures to 1e, supporting the application of DA^{yne} as a DA surrogate in a cellular context.

We compared the biological activity of DA and DA^{yne} in undifferentiated neuronal SH-SY5Y cells to determine whether similarities in their chemical reactivity parallel toxicity.[241] Cells were dosed with various concentrations of DA or DA^{yne}, and their viability was measured with an alamarBlue assay, which indirectly reports oxidative stress.[242] As shown in **Figure 2.8A,B**, DA^{yne} and DA toxicity were dependent on both dose and exposure time. At physiologically relevant concentrations of 0.1−1 μM, no statistical differences were observed between DA and DA^{yne} toxicity and after 48 h, with most cells remaining viable. At 100 μM, a concentration that exceeds physiological levels of DA by 100−1000×, a significant impact on cell viability was observed for DA^{yne} when compared to DA. At this concentration, 24 h exposure to DA^{yne} reduced the cell viability to 6%, whereas equivalent treatment with DA left 41% of cells viable. While DA^{yne} is slightly more toxic at high concentrations, the inherent toxicity of native DA at high doses and relatively short exposure times underscores the necessity to develop probes that enable sub-micromolar detection limits for use in standard analytical platforms.

To ensure that the observed DA and DA^{yne} toxicity was caused by oxidation of the catechol to a quinone, we co-treated cells with DA or DA^{yne} and an antioxidant, NAC.

Addition of 2.5 mM NAC significantly rescued viability in cells treated with 100 μM of DA and DA^{yne} (**Figure 2.8C**). Co-treatment with NAC for 24 h improved cell viability to 95 and 76% for DA and DA^{yne}, respectively, compared with 41% for DA and 6% for DA^{yne} when NAC is absent. Rescue due to other processes influenced by NAC cannot be ruled out by our experiments.[243] Nevertheless, our *in vitro* data reported here support DA and DA^{yne} quinone electrophiles as drivers of cellular toxicity.

Having demonstrated conditions wherein cells tolerate the DA^{yne} probe, we focused on characterizing DPAs in SH-SY5Y cells. We first quantified the relative number of DPAs generated in undifferentiated cells treated with DA^{yne} under a variety of conditions that we knew to be favorable for generating DPAs *in vivo*. Notably, the cells were treated with the reduced probe, which auto-oxidized *in situ*. Following DA^{yne} administration and washing off excess probe, the cells were lysed, and a biotin handle was conjugated to DPAs using copper-catalyzed azide−alkyne cycloaddition (CuAAC) click chemistry.[244] The resultant lysates were separated using sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) and transferred to a nitrocellulose membrane, where DPAs were visualized with streptavidin (SA) conjugated to an 800 nm fluorescent dye, denoted as SA800 (**Figure 2.9A**). Protein concentrations were normalized following staining with the REVERT total protein stain, and the relative amounts of DPAs per treatment condition were quantified (**Figure 2.9B**). DA^{yne} concentration and duration of treatment influenced the quantity of DPAs observed per treatment group. Incubation for 1 h proved to be insufficient to produce detectable levels of DPAs by Western blotting. Treatment for 12 h with 10 μM DA^{yne} provided optimal

results, as DPAs were detectable at 2.5 times above background (**Figure 2.9B**), and relatively little toxicity was observed. Treatment with any concentration of DA^{yne} for 24 h resulted in decreased levels of detectable DPAs compared to a 12 h incubation. In the case of 10 μM DA^{yne} treatment, the normalized DPA signal for 24 h exposure was 3.3-fold lower than the 12 h treatment group. Co-incubation of 100 μM DA^{yne} and 2.5 mM NAC for 12 h resulted in a 25-fold reduction of DPA formation compared to DA^{yne} alone, which is consistent with results in **Figure 2.8C** that show that NAC mediates rescue cell viability upon DQ exposure.

Reduction in visualized DPAs at extended exposure times was expected when considering the toxicity of DQs. For example, only 75% of the cells remained viable following 10 μM treatment with DA for 24 h. The presence of melanin-like aggregates in cellular debris pellets at higher DA^{yne} treatments suggests that DPAs polymerize and aggregate, which likely sequester DPAs, reducing their levels in western blots (**Figure 2.10**). This activation of DA regulatory or DPA stress responses may account for time-dependent reductions in DPA concentrations.

Additionally, we demonstrated that fluorescently modified DPAs generated using DA^{yne} could be mapped cellularly using confocal microscopy. Following treatment with DA^{yne}, cells were fixed with paraformaldehyde, and DPAs were conjugated with Alexa Fluor 647 azide using an in-cell click reaction. The cell's nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI) and imaged on a confocal microscope. In this manner, we verified that DA^{yne} permeates the cell membrane and forms adducts homogeneously

throughout the cytosol (**Figure 2.9C**). The use of DA^{yne} to visualize DPAs subcellularly demonstrates the versatility of a DA mimetic with alkyne functionality.

Next, we wished to determine whether the rate of DA^{yne} oxidation was causing low detection limits at 1 μM doses. We treated SH-SY5Y cells with DA^{yne} in the presence of TY to generate DQs more rapidly than auto-oxidation. Using TY, 1 h exposure of DA^{yne} was sufficient to produce detectable DPAs at doses as low as 10 nM (**Figure 2.11**). Thus, DA^{yne} can be used to improve the sensitivity in DPA visualization assays requiring physiologically relevant concentrations of the DA-mimetic.

To verify that DA^{yne} forms adducts with the same proteins that bind DA, we performed a competition binding experiment in which SH-SY5Y cells were pre-incubated with DA and TY for 30 min prior to the addition of DA^{yne} (**Figure 2.12**). A dose-dependent decrease in DA^{yne} labeling upon increasing concentration of DA pre-treatment was observed. Incomplete competition at the highest concentration of DA competitor (50 μM) suggests that DQs binds to target proteins at low occupancy, which is consistent with previous reports of similarly functionalized quinone methides.[245] Since only a fraction of a particular protein population is modified in the initial DA treatment, the subsequent DA^{yne} exposure results in additional partial labeling. Nevertheless, low levels of DQ protein modification can still drive robust cellular responses,[219] supporting the need for biological probes that enhance analytical sensitivity so that DPAs can be characterized in cell models treated with physiological levels of DA- mimetic probes.

To confirm that the levels of DPA formation were similar between DA^{yne} and DA treatment, two previously described methods of detecting quinone-protein adducts, nitro

blue tetrazolium (NBT) staining and near infrared (nIR) fluorescence, were used to test the levels of DPAs produced by DA^{yne} and DA.[162, 165] Qualitative comparison of stained nitrocellulose membranes between DA and DA^{yne} generated DPAs revealed similar amounts of adduct formation (**Figure 2.13**), providing confidence that our probe is a relevant mimetic of DA. However, in our hands, these methods proved less sensitive at detecting DPAs compared to our bio-orthogonal visualization strategy using DA^{yne}, which may be due to the lower concentration of DQs used in the current study. Despite increased protein loading to 50 μg per lane to improve signal, quantitative analysis between DA and DA^{yne} produced DPAs could not be performed using these adduct detection methods.

After demonstrating that cellular DPAs formed using DA^{yne} were similar to those generated with DA, we performed a chemoproteomic experiment to enrich and characterize the proteins that are modified by oxidized metabolites of DA^{yne}. A 12 h incubation period with 10 μM DA^{yne} was selected based on our earlier labeling results. After exposing SH-SY5Y cells to DA^{yne}, DPAs were immobilized onto azide-functionalized agarose beads using CuAAC click chemistry. This approach is advantageous compared to traditional biotin-streptavidin enrichment as it eliminates the need to remove excess biotin and click reagents prior to immobilization. It also precludes interference from naturally biotinylated proteins, avoids streptavidin contamination, and enables stringent washing steps to remove proteins that nonspecifically bind the Sepharose column without removing DPAs. Using this approach, we verified that DPAs were captured by the azide beads through Western blot (**Figure 2.14**). Following capture,

the DPA bound beads were washed extensively and subjected to on-bead trypsin digestion. The resultant peptides were desalted and analyzed mass spectrometrically and quantified using the MaxQuant software package.

SH-SY5Y cells were treated with DA^{yne} or dimethyl sulfoxide (DMSO) across three biological replicates. 38 proteins were significantly enriched by DA^{yne} using the methods outlined above (**Figure 2.15A**, and **Table 2.1** for a full listing of proteins). We ensured that enrichment with DA^{yne} was consistent between treatment groups by correlating the label-free quantitation (LFQ) intensity values of identified proteins between different treatment conditions and replicates (**Figure 2.16**). We observed a correlation of 0.91 ± 0.02 in our DA^{yne}-enriched samples, indicating reproducible enrichment. Additionally, the correlation within the DMSO treatment group was also high, 0.79 ± 0.06; whereas the correlation between DA^{yne} and DMSO groups was low, 0.38 ± 0.06, verifying that DA^{yne} enrichment was robust.

Gene ontology (GO) analysis of these enriched proteins (http://www.pantherdb.org/) revealed a variety of over- represented molecular functions, such as RNA binding, disulfide oxidoreductase activity, and cytoskeleton regulation (**Figure 2.15B**).[246] Other notable DPAs found in our proteomic results were pyruvate kinase and fatty acid synthase, which play important roles in cellular energy homeostasis. The DPA interactome network of enriched proteins was analyzed using the STRING (Search Tool for the Retrieve of Interacting Genes) database (https://string-db.org/) to visualize the pathways and functions that might be impacted by DPA formation (**Figure 2.15C**).[247] Notably, cell redox homeostasis and ubiquitin/proteosome pathways were

enriched, suggesting DPAs may interfere with these processes. Enriched proteins featured a high prevalence of sulfhydryl-containing active sites, such as those within the oxidoreductase pathway. Pathways such as cellular energy homeostasis and clathrin-mediated processes identified in our STRING analysis were unexpected as proteins involved in these processes have not yet been characterized as DPAs. Such protein adducts, once validated, would represent novel targets as potential contributors to neurotoxicity during PD.

Our next objective was to validate candidate DPAs from our proteomic analysis. We focused on the PDIs because three isoforms of PDIs (PDIA3, PDIA6 and P4BH) were significantly enriched by DA^{yne}. PDIs are critical in a variety of cell stress responses, including five molecular functions: peptide disulfide oxidoreductase activity, PDI activity, transposing S−S bonds, disulfide oxidoreductase activity, and intramolecular oxidoreductase activity (shown in **Figure 2.15B**) that were significantly altered upon treatment with DA^{yne}. Notably, there is evidence that PDIs play an important role in preventing neurodegeneration.[248]

We initially verified that DA^{yne} covalently modifies recombinant PDIA3 using a simple binding assay by exposing expressed PDIA3 to DA^{yne} and monitoring the appearance of adduct formation by SDS-PAGE analysis. To do so, DA^{yne} was oxidized with TY and incubated with recombinant PDIA3 for 1 h, then Alexa Fluor 647 azide was conjugated to DA^{yne} in the mixture, and conjugate formation was visualized as a band at 58 kDa corresponding to the DPA. As shown in **Figure 2.17A**, DA^{yne} labeled PDIA3, and binding was competed upon exposure of PDIA3 to oxidized DA for 30 min prior to

DA^{yne} treatment. These data support labeling of PDIA3 by DA^{yne} at sites also occupied natively by DA.

Native PDIA3 in non-transfected SH-SY5Y cells was also captured by DA^{yne} following 1 h exposures of the lysates to oxidized DA^{yne}. In these experiments, DPAs were modified with biotin azide using click chemistry and captured on streptavidin agarose beads. After DPA capture, non-specifically bound proteins were washed off and DPAs were eluted at 95 °C in gel-loading buffer. Immunoblots shown in **Figure 2.17B** revealed that DA^{yne}-modified PDIA3 was successfully captured in our pulldown. DA also competed with DA^{yne} binding to PDIA3 (**Figure 2.17B**). Attempts to identify the specific site(s) of DA modification on PDIA3 by liquid chromatography (LC)−mass spectrometry (MS)/MS) analysis of tryptic peptides were unsuccessful. These results are consistent with previous studies that have reported difficulties detecting site specific DA adducted peptides.[87, 91, 249]

Finally, we explored the influence of DA on PDI enzymatic activity using a functional assay that measures the protein folding activity of PDI. We incubated PDI with DA and pre- oxidized DA (DQ), both in the presence or absence of NAC, and measured enzymatic activity through a fluorescence-based assay as shown in **Figure 2.17C**. We found that treatment with 150 μM DA and 150 μM DQ significantly lowered PDI activity to 67 and 29%, respectively, and that NAC and TY treatment alone had no effect on activity. However, cotreatment of DA and DQ with 1 mM NAC to intercept DQs, significantly rescued PDI activity to 90 and 84%, respectively, suggesting that covalent modification of PDI by oxidized DA metabolites inhibits the function of PDI.

Collectively, our results validate PDIA3 as an extravesicular target of oxidized DA and that DPA formation results in comprised enzymatic activity.

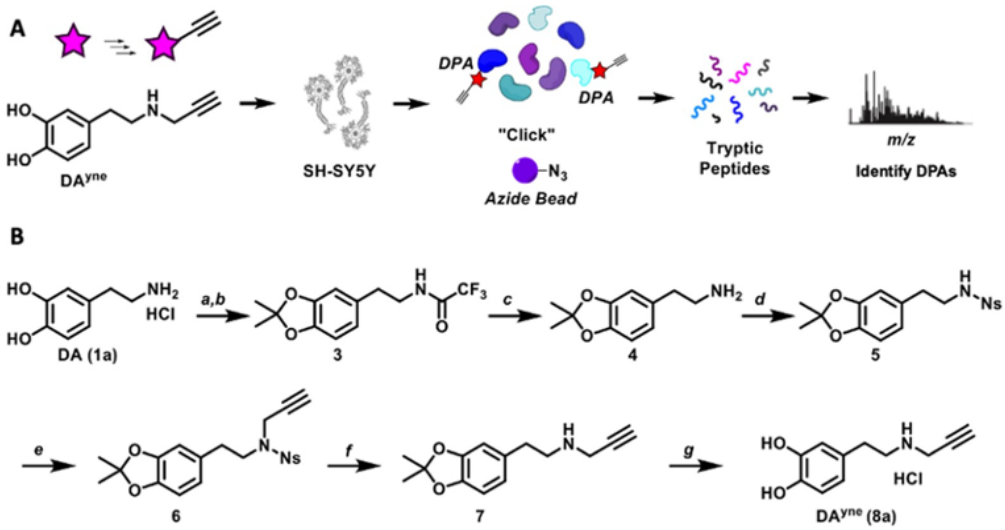


Figure 2.3. Chemical biology strategy to visualize and identify DPAs

A) Workflow of DPA enrichment via DA^{yne}. B) Synthesis of DA^{yne}. (a) Methyl trifluoroacetate, triethylamine, CH_3OH; 93%. (b) pTSA, 2,2-dimethoxypropane, toluene; 79%. (c) LiOH, CH_3OH; 98%. (d) Nosyl chloride, triethylamine, CH_2Cl_2; 97%. (e) Propargyl bromide, K_2CO_3, acetone; 96%. (f) Thiophenol, KOH, CH_3CN; 50%. (g) Trifluoroacetic acid, $CHCl_3$ then HCl, tetrahydrofuran; 96%.

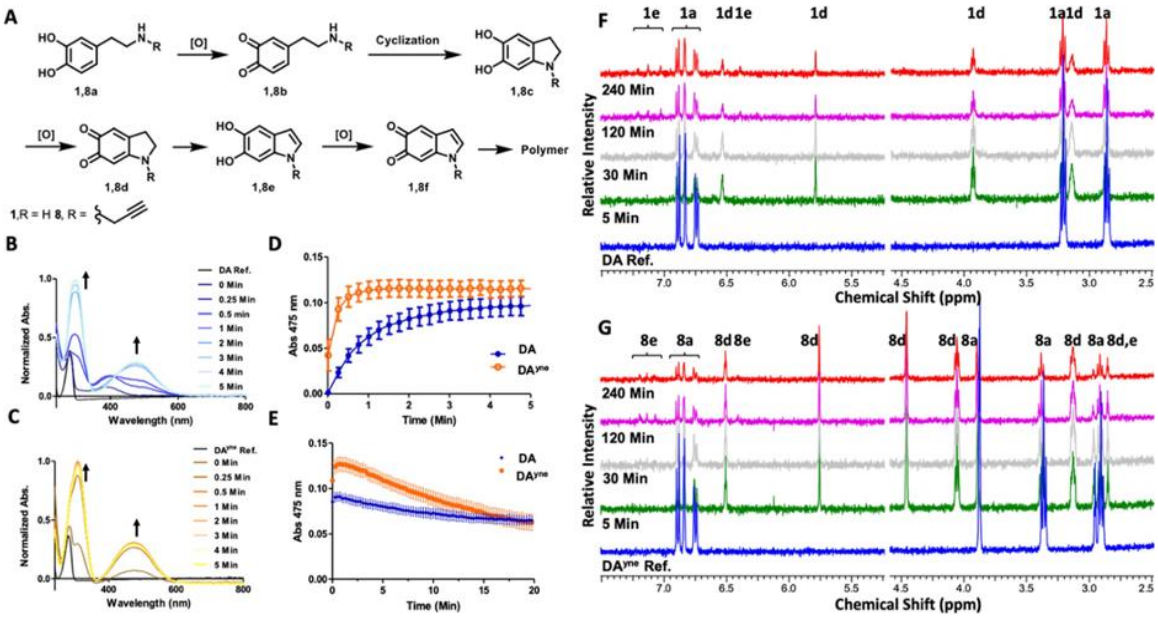


Figure 2.4. Spectroscopic characterization of DA and DA^{yne} oxidation

A) Oxidation products of DA and DA^{yne}. B/C) Absorbance spectra time course of 50 μM DA/ DA^{yne} respectively and 50 μM $NaIO_4$ in H_2O. D/E) Change in absorbance at 475 nm of 50 μM DA/ DA^{yne} and 50 μM $NaIO_4$ in H_2O or PBS pH 7.4 respectively. 475 nm absorption is characteristic of aminochrome formation. Data are represented as mean ± SEM from three replicates. F/G) NMR spectra time course of 1 mM DA/ DA^{yne}, respectively, and 1 mM $NaIO_4$ in PBS pH 7.4. Labels above proton resonances denote corresponding compounds. Water solvent peak was removed for clarity.

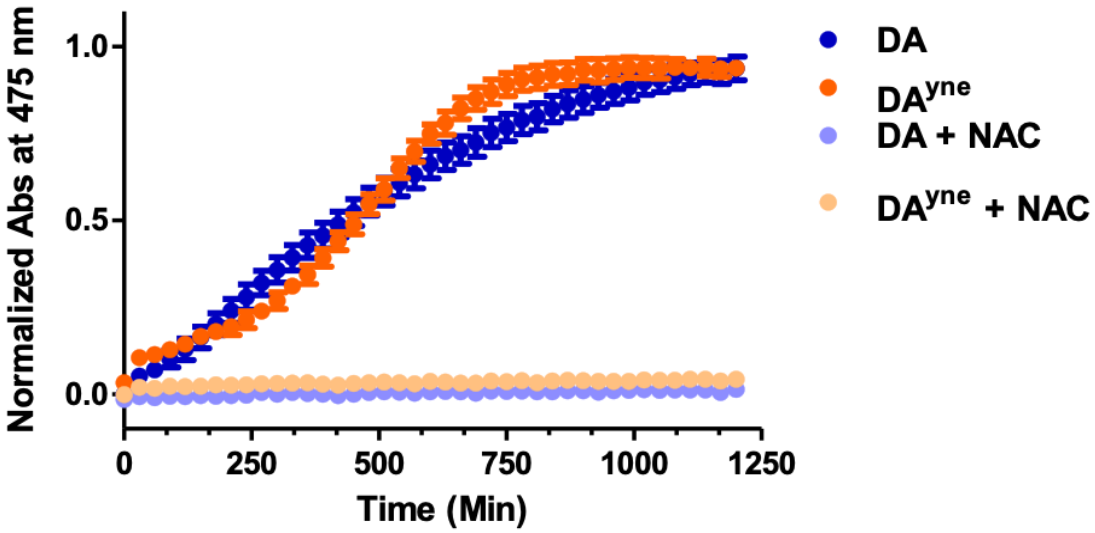


Figure 2.5. Auto-oxidation of DA and DA^{yne}

DA and DA^{yne} (50 µM) were incubated at 37 °C in PBS, pH 7.4 in the presence or absence of 1 mM NAC. Absorbance at 475 nm, indicative of aminochrome formation, was measured at 30 min intervals over the course of 20 hrs. Data are represented as mean ± SEM from three replicates.

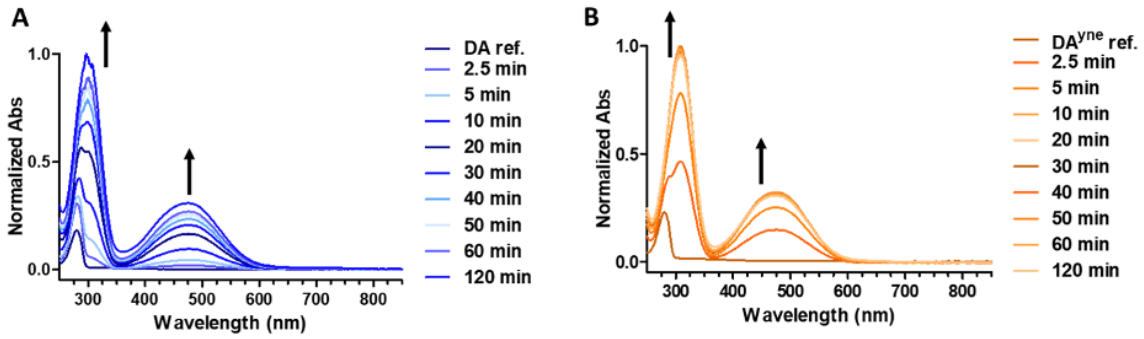


Figure 2.6. Tyrosinase oxidation of DA and DA^{yne}

TY (300 units) was incubated with 150 µM DA A) or DA^{yne} B) and the reaction was monitored via UV-Vis. The increase at 475 nm indicates successful oxidation of the substrate by TY to its corresponding aminochrome.

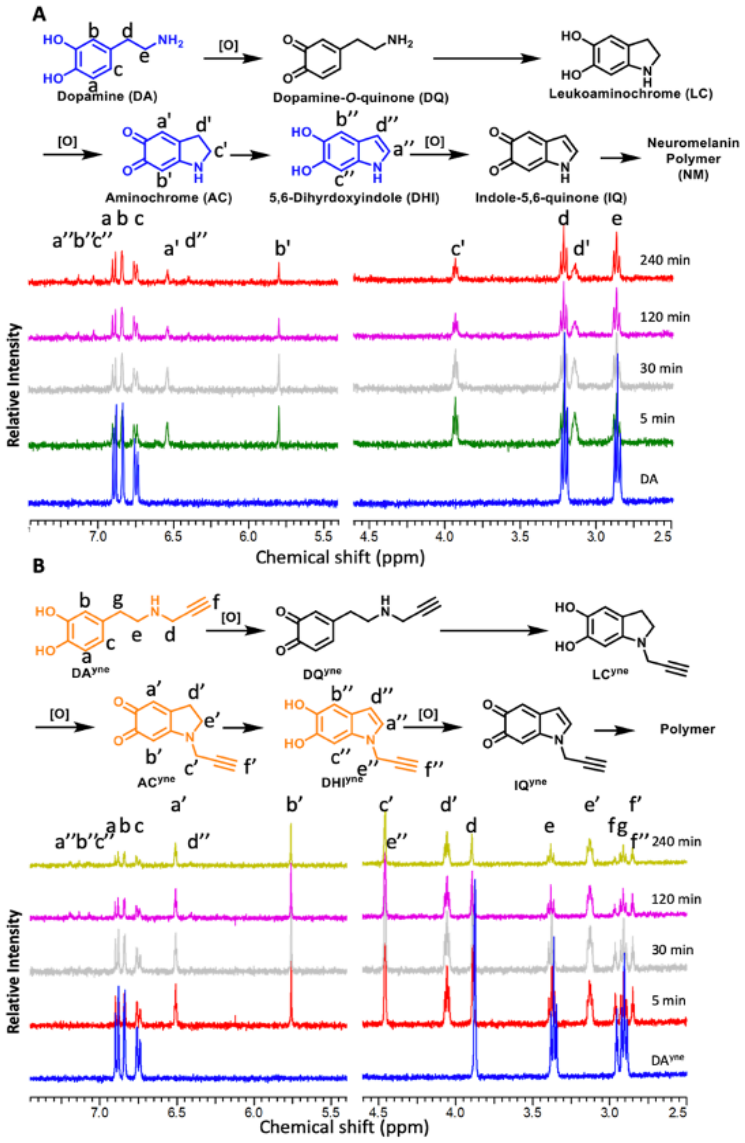


Figure 2.7. NMR spectra of reaction time course between 1 mM DA A) or DA^{yne} B) and 1 mM $NaIO_4$ in PBS pH 7.4

Labels above proton resonances denote corresponding molecular labels. Water solvent peak was removed for clarity.

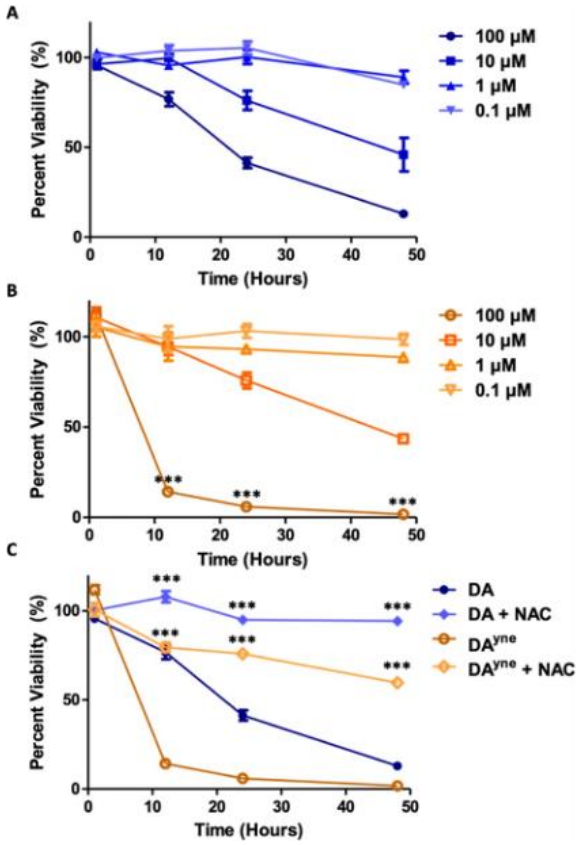


Figure 2.8. DA and DA^{yne} reduce SH-SY5Y cell viability

Cell viability was determined through an alamarBlue assay. Data are represented as mean ± SEM from four replicates and normalized to DMSO control. All treatments of DA and DA^{yne} at analogous times and concentrations were analyzed via a one-way ANVOA and an unpaired t-test (***$p < 0.001$). Percent viability of cells was determined following treatment with A) DA or B) DA^{yne} alone. C) Cells were treated with 100 µM DA or DA^{yne} in the presence or absence of 2.5 mM NAC, which showed that cell viability was rescued under these conditions.

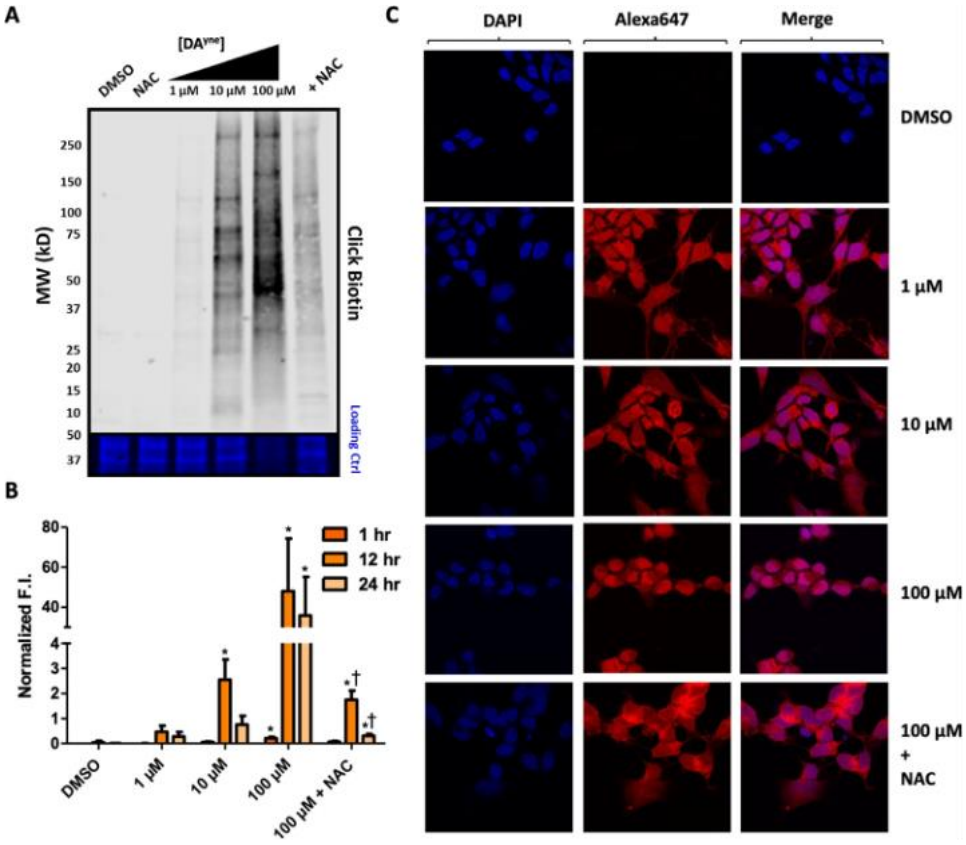


Figure 2.9. Visualization of DPAs by DA^{yne}

A) Representative nitrocellulose blot following incubation of SH-SY5Y cells with DMSO, 2.5 mM NAC, 1 µM, 10 µM, 100 µM, or 100 µM DA^{yne} and 2.5 mM NAC for 12 h, lysis and ensuing CuAAC click reaction with biotin azide. Top: 800 nm channel following incubation with SA800. Bottom: 700 nm channel following REVERT total protein stain. B) Quantification of DPAs. Fluorescence intensity (F.I.) of SA800 for each treatment was normalized by total protein through the REVERT stain. Data are represented as mean ± SEM from three replicates. Data were analyzed via a two-way ANOVA and an unpaired t-test. (Comparisons between: DMSO *$p < 0.05$; 100 µM DA^{yne} †$p < 0.05$ for given treatment duration). C) Confocal imaging of SH-SY5Y cells

following treatment with DMSO, 1 μM, 10 μM, 100 μM, or 100 μM DA^{yne} and 2.5 mM NAC, respectively for 12 h, fixation with paraformaldehyde, ensuing CuAAC click reaction with Alexa Fluor 647 azide and staining with DAPI. Images were acquired at 60× magnification.

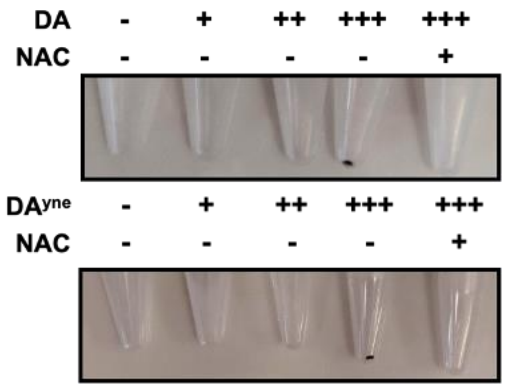


Figure 2.10. DA and DA^{yne} treatment induce melanin formation in cellular debris

Photographed tubes contain cellular debris pellet following lysis. SH-SY5Y cells were treated with DMSO, 2.5 mM NAC, 1 µM, 10 µM, 100 µM, or 100 µM DA/DA^{yne} and 2.5 mM NAC as indicated. Black debris was observed in 100 µM treatments, which presumably are melanin polymers and aggregates.

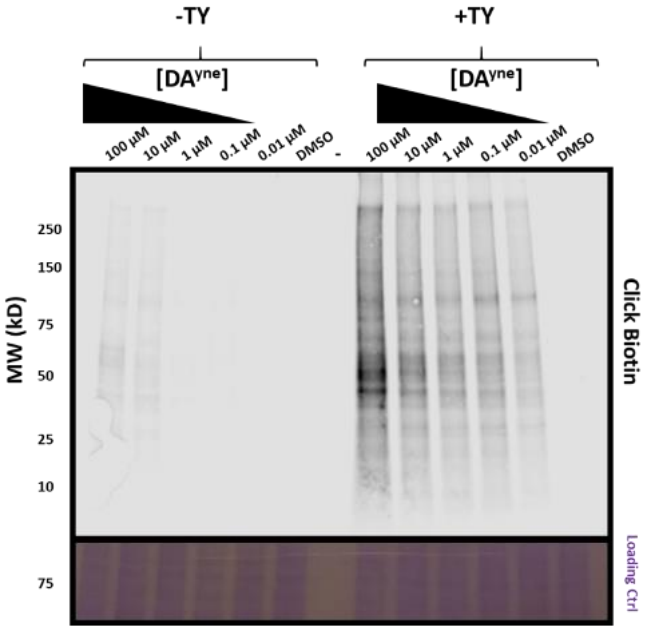


Figure 2.11. Visualization of DPAs by pre-oxidized DA^{yne}

Representative nitrocellulose blot following 1-hour incubation of SH-SY5Y cells with respective DA^{yne} concentration in the presence or absence of 600 units of TY, lysis and ensuing CuAAC click reaction with biotin azide. Top: 800 nm channel following incubation with SA800. Bottom: Coomassie blue stain to confirm equal protein loading.

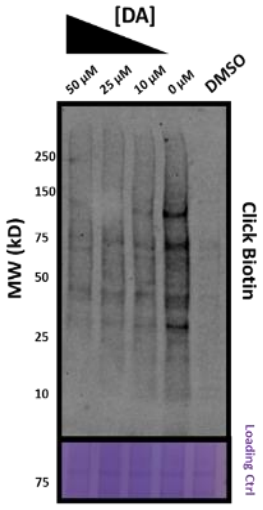


Figure 2.12. Competition of DA^{yne} derived DPAs with DA

Representative nitrocellulose blot of three independent experiments following 1-hour incubation of SH-SY5Y cells with 10 µM DA^{yne} after a 30-min pre-incubation at the specified DA concentration in the presence of 600 units of TY, lysis and ensuing CuAAC click reaction with biotin azide. Top: 800 nm channel following incubation with SA800. Bottom: Coomassie blue stain to confirm equal protein loading.

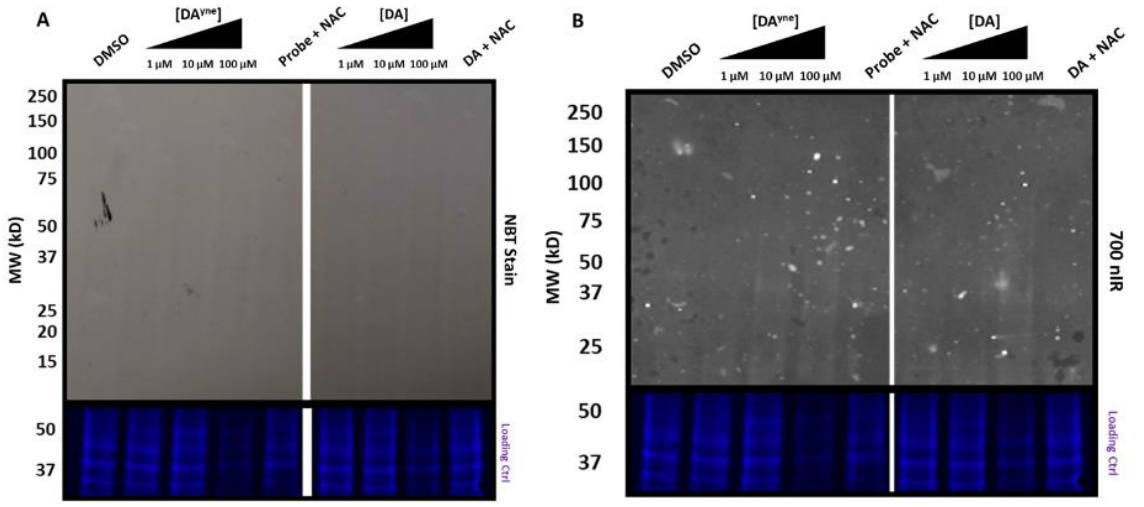


Figure 2.13. Comparison of DA and DA^{yne} derived DPAs

A) DPAs detection through NBT staining. Representative nitrocellulose blot following 1-hour incubation of SH-SY5Y cells with respective concentration of DA or DA^{yne} in the presence of 600 units of TY, lysis and NBT staining. Top: NBT stain; purple bands mark DPAs. Bottom: 700 nm channel following revert total protein stain. B) DPAs detection via nIR florescence. Representative nitrocellulose blot following 1-hour incubation of SH-SY5Y cells with respective concentration of DA or DA^{yne} in the presence of 600 units of TY, lysis, and nIR imaging. Top: 700 nm channel imaging immediately after membrane transfer; white bands mark DPAs. Bottom: 700 nm channel following revert total protein stain.

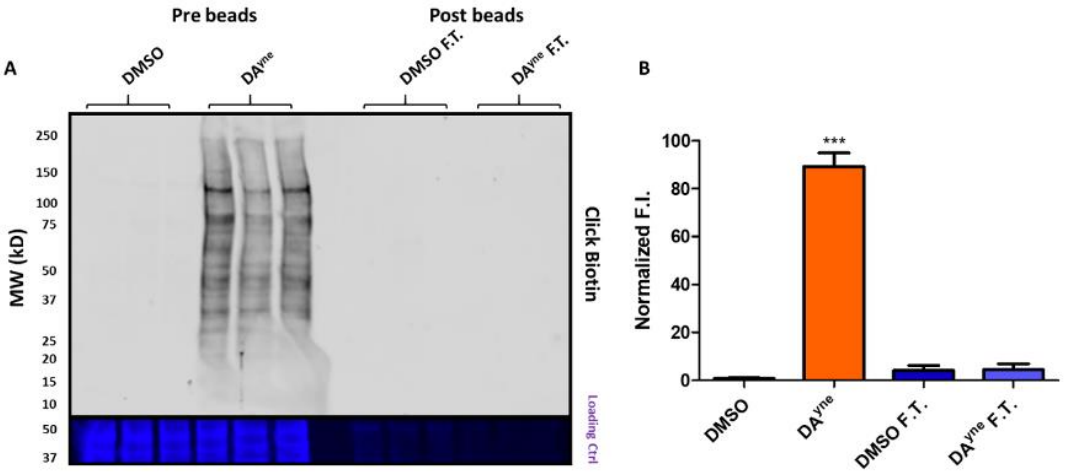


Figure 2.14. Azide functionalized agarose beads capture DA^{yne} induced DPAs

A) Nitrocellulose blot following incubation of SH-SY5Y cells with DMSO, 10 µM DA^{yne} for 12 hrs, lysis and ensuing Cu click reaction with biotin azide pre or post reaction with azide beads. F.T. denotes flowthrough. Top: 800 nm channel following incubation with SA800. Bottom: 700 nm channel following revert total protein stain. B) Quantification of DPAs. Fluorescence intensity (F.I.) of SA800 for each treatment was normalized by total protein through the revert stain. Data are represented as mean ± SEM from three replicates. Data were analyzed via a one-way ANVOA and an unpaired t-test to DMSO (***$p < 0.001$).

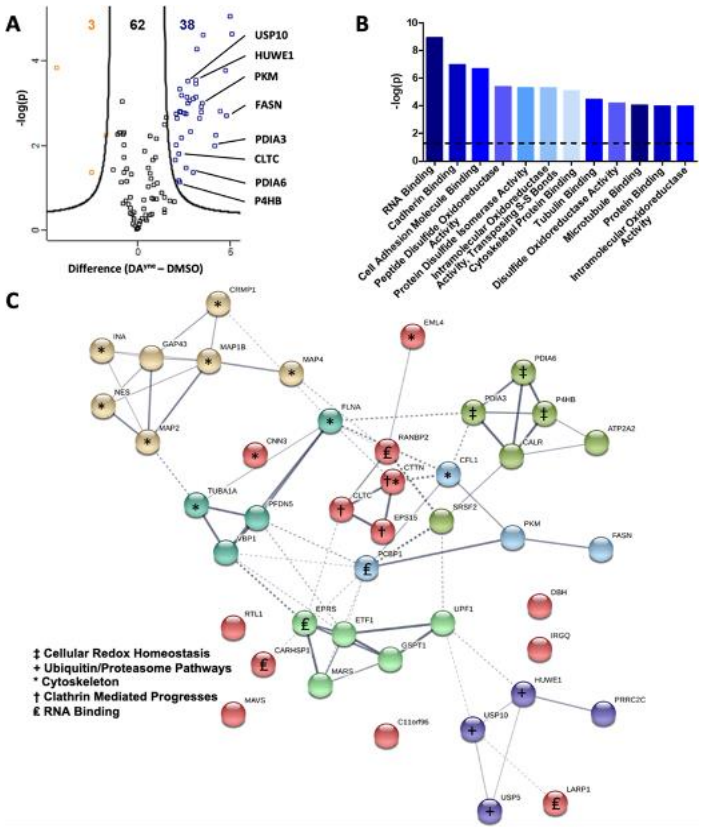


Figure 2.15. Proteomic enrichment and bioinformatic analysis of DPAs

A) Volcano plot for quantitative analysis of proteins identified following 12 h treatment of SH-SY5Y cells with DMSO or 10 μM DA^{yne} and enrichment at a FDR of 0.05 and a minimal coefficient of variation (S0) of 2.0. Proteins highlighted in orange or blue were significantly enriched in the DMSO or DA^{yne} treatment, respectively. Data are representative of 3 biological replicates from each treatment condition. B) Molecular function of DA^{yne} enriched proteins. Analysis performed with PANTHER overrepresentation test (released 20190711); annotation version: GO Ontology database (released 20190703), using Fisher's exact test with FDR correction. The dashed line represents a statistical cutoff of p = 0.05. C) Pathway analysis of DA^{yne} enriched proteins

was performed with the STRING (Search Tool for the Retrieve of Interacting Genes, version 11) database. Confidence of interaction represented by line thickness. Proteins involved in enriched processes are labeled.

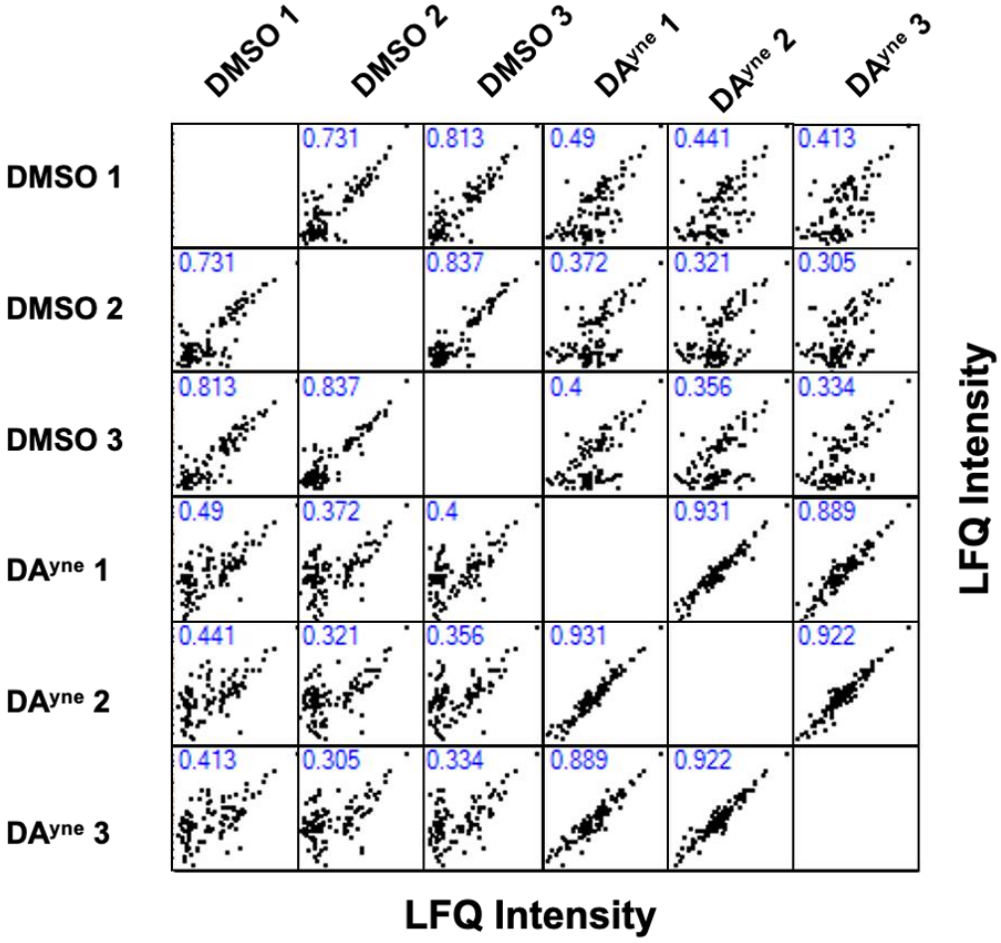


Figure 2.16. Proteomic enrichment with is DA^{yne} reproducible

Correlation plots show the intensity of label free quantitation (LFQ) values for a given treatment condition on the X and Y axis. Identified proteins are represented as black squares. The R^2 value for each comparison in written in blue. The Perseus software suite (version 1.6.6.0) was used generate this plot.

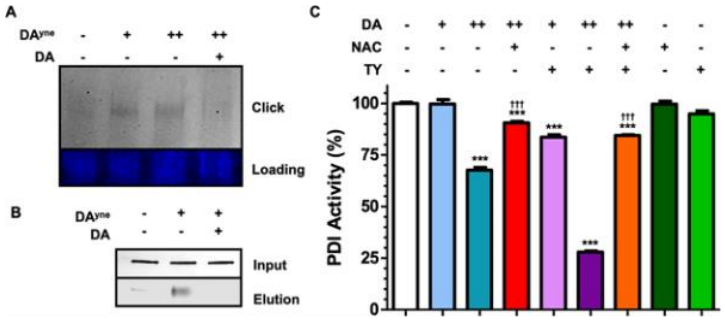


Figure 2.17. PDIA3 is covalently modified and functionally inhibited by DQs

A) Representative SDS-PAGE gel after incubating PDIA3 for 1 h with DMSO, 10 μM, 150 μM DA^{yne} preoxidized with TY, or with 150 μM DA^{yne} after a 30 min 150 μM DA pre-treatment (both preoxidized with TY), and ensuing CuAAC click reaction with Alexa Fluor 647 azide. Top: Alexa Fluor 647 fluorescence intensity. Bottom: Imperial protein stain. B) Representative nitrocellulose blot probed for PDIA3 following capture with DA^{yne}. SH-SY5Y lysate was incubated with DMSO, 150 μM DA^{yne}, or with 150 μM DA^{yne} after a 30 min 150 μM DA pre-treatment lysis and subjected to a CuAAC click reaction with biotin azide. Samples were then incubated with streptavidin agarose beads to capture DPAs. Top: sample lysate prior to streptavidin enrichment. Bottom: elution of DPAs following streptavidin agarose enrichment. C) Percent PDI activity following a 30 min incubation with the indicated treatment. Final concentrations of DA used were 10 and 150 μM. TY was used to oxidize DA to DQs. The final NAC concentration was 1 mM. Data are represented as mean ± SEM from four replicates and normalized to DMSO control. Data were analyzed via a two-way ANOVA and a tukey post hoc test. (Comparisons between: DMSO ***P < 0.0001; 150 μM DA †††P < 0.0001) for given treatment duration.

Table 2.1. List of enriched DPAs

List of enriched DPAs. The third column reports the T-test pValue to assess enrichment of DPAs. The fourth column reports the difference between the mean log2 transformed LFQ intensity values between DA^{yne} and DMSO treated groups, which is indicative of fold change difference in protein abundance. Data is representative of 3 biological replicates from each treatment condition. An input of 100 µg of protein lysate was used to prepare each sample.

Protein	Gene	Negative Log p	Difference (DA^{yne} - DMSO)	Description
Microtubule-associated protein 4	MAP4	**5.054**	**5.001**	microtubule binding
Poly(rC)-binding protein 1	PCBP1	**4.645**	**5.096**	Fe binding, mRNA splicing factor, ribonucleoprotein, serine protease
Bifunctional glutamate/proline--tRNA ligase	EPRS	**4.608**	**3.543**	ligase
Dopamine beta-hydroxylase	DBH	**4.288**	**3.2**	conversion of dopamine to noradrenaline
Cofilin-1	CFL1	**3.785**	**4.738**	actin filament binding
E3 SUMO-protein ligase RanBP2	RANBP2	**3.552**	**3.144**	G-protein modulator
Ubiquitin carboxyl-terminal hydrolase 10	USP10	**3.533**	**2.734**	cysteine protease
E3 ubiquitin-protein ligase	HUWE1	**3.463**	**3.146**	ubiquitin-protein ligase
Neuromodulin	GAP43	**3.339**	**2.286**	nerve growth
Mitochondrial antiviral-signaling protein	MAVS	**3.197**	**2.385**	innate immune defense against viruses
Serine/arginine-rich splicing factor 2	SRSF2	**3.149**	**3.149**	mRNA splicing factor
Retrotransposon-like protein 1	RTL1	**3.147**	**2.656**	developmental protein
Src substrate cortactin	CTTN	**3.112**	**3.004**	basic helix-loop-helix transcription factor, non-motor actin binding protein

Microtubule-associated protein 2	MAP2	**3.074**	**2.793**	microtubule binding
Protein PRRC2C	PRRC2C	**3.002**	**2.729**	nuclease, transcription factor
Pyruvate kinase	PKM	**2.997**	**3.488**	key role in glycolysis
Nestin	NES	**2.913**	**3.431**	brain development
Dihydropyrimidinase-related protein 1	CRMP1	**2.821**	**4.442**	metalloprotease
Prefoldin subunit 3	VBP1	**2.812**	**2.279**	chaperone
Eukaryotic peptide chain release factor GTP-binding subunit	GSPT1	**2.799**	**3.423**	G-protein, hydrolase, translation elongation factor, translation initiation factor, translation release factor
Uncharacterized protein C11orf96	C11orf96	**2.785**	**2.331**	-
Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	ATP2A2	**2.774**	**2.549**	cation transporter, hydrolase, ion channel
Immunity-related GTPase family Q protein	IRGQ	**2.757**	**2.087**	GTP binding
Calponin-3	CNN3	**2.751**	**2.6**	non-motor actin binding protein
Fatty acid synthase	FASN	**2.714**	**4.811**	acyltransferase, dehydrogenase, esterase, ligase, methyltransferase
Calcium-regulated heat-stable protein 1	CARHSP1	**2.65**	**3.354**	RNA binding protein
La-related protein 1	LARP1	**2.639**	**2.12**	ribonucleoprotein
Tubulin alpha-1A chain	TUBA1A	**2.334**	**2.857**	tubulin
Microtubule-associated protein 1B	MAP1B	**2.248**	**4.199**	non-motor microtubule binding protein
Epidermal growth factor receptor substrate 15	EPS15	**2.19**	**2.388**	G-protein modulator, calcium-binding protein, membrane traffic protein
Alpha-internexin	INA	**2.014**	**2.102**	morphogenesis of neurons
Protein disulfide-isomerase A3	PDIA3	**2.006**	**4.148**	catalyzes the formation, breakage and rearrangement of disulfide bonds
Clathrin heavy chain 1	CLTC	**1.81**	**2.258**	vesicle coat protein
Echinoderm microtubule-	EML4	**1.669**	**2.021**	microtubule

associated protein-like 4				cytoskeleton
Filamin-A	FLNA	**1.466**	**2.613**	actin filament binding
Protein disulfide-isomerase A6	PDIA6	**1.361**	**3.028**	catalyzes the formation, breakage and rearrangement of disulfide bonds
Protein disulfide-isomerase	P4HB	**1.175**	**2.238**	catalyzes the formation, breakage and rearrangement of disulfide bonds

2.4 Discussion

Therapeutic interventions capable of halting or reversing PD progression will require a complete understanding of the initial molecular events that predispose dopaminergic neurons to degenerate and the subsequent biochemistry that results in clinical symptoms. There is substantial evidence indicating that oxidized dopamine metabolites, DQs, are involved in early mechanisms of neurodegeneration during PD, including their spatiotemporal tissue distribution and their increased abundance observed in tissue samples from patients suffering from more severe forms of PD.[15] These studies are consistent with a model of PD progression that is highly dependent on or exacerbated by DA metabolites in early stages of cellular dysfunction. As these reactive metabolites that are measurable before PD symptoms are detected clinically through movement tests, they are enticing biomarkers that predict PD.[15] DPAs represent surrogates of DQs that could be measured in patients in a clinical setting and form the basis for early biochemical tests that predict both the probability of developing PD symptoms and their severity. Moreover, DQ-modification leads to generation of DPAs that result in neurotoxic outcomes and many DPAs are expected to represent novel targets for therapeutic design and development. Given the heterogeneity and complexity of PD pathogenesis, unbiased screening for DPAs on a proteome-wide scale using mass spectrometry technologies is ideal for comprehensively enumerating the DA-modified proteome without foreknowledge of disease etiology. Such DPAs can later be validated for disease-relevance using standard biochemical assays, which we demonstrated for PDIA3. However, owing to the complex reactivity profile of DA, few studies have

reported the development of such unbiased techniques for DA.[118] Several limitations have prevented such technologies from reaching fruition, including the poor sensitivity of analytical methods used for the detection of incipient DPAs in cell models. These include the low DPA levels expected natively in cell and *in vivo* models due to low occupancy binding of DA and regulatory mechanisms that maintain low DPA and cytosolic DA levels, such as neuromelanin formation.

Given that DPA levels and detection sensitivities in the presence of such regulatory mechanisms must be low, we reasoned that DPA detection sensitives might be increased by enriching adducts. Thus, a DA-mimetic proteomic probe (DA^{yne}) was designed with an alkynyl enrichment handle included to extract DPAs and improve proteomic and fluorometric sensitivity. Reduction of ostensible off-targets in these enrichment studies could then be reduced by modifying DA exposure to cells without compromising analytical sensitivity. These principles led to the design of DA^{yne}.

Our DA-mimetic reactivity-based probe DA^{yne}, includes an amino-modified propargyl enrichment tag on the DA scaffold, which makes it unique among DA-mimetic probes. This modification does not prevent our probe from cyclizing into an aminochrome-mimetic upon oxidation in a reaction trajectory that mirrors DA oxidation both kinetically and thermodynamically. Developing a probe that shares a similar kinetic and thermodynamic profile to DA is of paramount importance to ensure DQ metabolite lifetimes and protein bindings mirror native DQ metabolite binding when profiling representative DPAs in cell models. Indeed, we show this to be the case spectroscopically, which demonstrates that the oxidation of DA and DA^{yne} leads to the

formation of DQ (1b) and aminochrome species (1d) at similar apparent rates. Similar to previous studies, we observed the aminochrome species to be the most thermodynamically stable under our *in vitro* conditions for both DA and DA^{yne} and likely the primary DQ metabolites involved in DPA generation. Alkylating the amine did not drastically alter DA-induced cytotoxicity at low doses *in vitro* and enabled visualization of DPAs by fluorescent microscopy. These latter studies confirm, at least qualitatively, that the functional profile of DA is represented faithfully by DA^{yne} due to kinetic and thermodynamic properties among their chief metabolites. Therefore, amino-tagged DA offers the potential to comprehensively profile DPAs generated from all DA oxidative metabolites.

Probe molecules designed previously to mimic DA have taken advantage of nonreactive functional groups to prevent cyclization of DQ, increasing the abundance of DQ-modified DPAs but preventing aminochrome (1d) formation and their DPA counterparts. For example, functionalized hydroxydopamine (6-OHDA) quinone-based reactivity probes have incorporated enrichment handles via an amide bond at nitrogen.[118] However, these probes rely on aminoacyl installation chemistries for propargyl incorporation. Such probes bear polarized amide carbonyls that significantly alter the electrostatic potential of the DA scaffold, preventing it from traversing the membrane of live cells, limiting its broad utility. Additionally, 6-OHDA based probes can only generate DQ-like oxidation products and not cyclic intermediates like 1d, 1e, and 1f, limiting the scope of potential DPAs detected to those generated with DA. Given the short lifetime of DQ-like metabolites (1b) and the long lifetime of aminochrome (1d)

metabolites observed for DA, the biological pathways of these product metabolites may be different, and DPAs identified using 6-OHDA probes may not represent PD-specific DPAs. The failure of 6-OHDA to cyclize into 1d and 1e might account for differences in pathologies observed between 6-OHDA rat models and clinical PD.[3] However, DQ-based probes should complement the use of probes like DA^{yne} as they can aid in characterization efforts that distinguish the relative importance of each class of DPA modification in disease etiology. DPA enrichment should serve as a complement to other powerful characterization methods that use native DA as a probe, such as ^{14}C autoradiography and isotopic tandem orthogonal proteolysis-activity based protein profiling (iso-TOP-ABPP) methodology.[152, 153, 213]

There are several advantages to using DA^{yne} as a DPA characterization that cannot be achieved using state-of-the-art analytical platforms discussed above. For example, the use of a DA mimic to identify native DPAs eliminates the use of toxic radioactive ^{14}C detection isotopes to visualize low-abundant DPAs. Alkyne−azide pull-downs also provide higher selectivity than boronic acid resins used previously to enrich catechol- containing DPAs, but that are unable to distinguish catechols from other vicinal alcohols. Additionally, enrichment chemistries for DA^{yne} are robust, allowing exposures at physiological concentrations. Distinct protein bands were observed in our immunoblots after exposing SH-SY5Y cells to concentrations as low as 10 nM after oxidation with TY. DA^{yne} enrichment also proved advantageous for carrying out studies requiring long-term exposures. The amount of external DA or, as in the case here, DA^{yne} added to cells for the detection of DPAs are more than ten times lower and fall within physiologically

relevant concentrations than many previous studies.[152, 153] These lower concentrations enabled long time- resolved *in vitro* experiments to be conducted without significant cell death or compromising detection of DPAs. Parallel multidimensional analyses, such as fluorescence microscopy and proteomic analyses also enhance the utility of this probe molecule, enabling reproducible discovery and validation of DPA targets. Collectively, DA^{yne} proved to be a versatile probe molecule to generate reproducible adducts that may begin to shape our understanding of DA's contribution to neurodegeneration during the entire aging spectrum.

One of the useful features of DA^{yne} highlighted here is the high sensitivity realized in chemoproteomic enrichment experiments, which revealed that DQs form adducts with proteins contained within several diverse families that influence myriad functions. Proteins that are inactivated or structurally altered within these families and pathways have the potential to significantly compromise neuronal integrity and influence age-related neurodegeneration or PD progression. Within our collection of candidate DPAs are various classes of proteins that have been identified in previous studies, including cytoskeletal proteins, ubiquitin proteosome machinery and protein folding.[152, 153, 163] PDIs were among the major classes of enriched DPAs identified in our GO analysis. PDIs assist in protein folding and rely on thiol disulfide oxidoreductase activity to create S−S bonds. Impairment of that function by DQ modification is expected to activate the unfolded protein response, elicit ER stress and perturb cell redox homeostasis, which were identified as major pathways affected by DA^{yne} exposure in a STRING analysis. We confirmed that both DA and DA^{yne} modify PDIA3 through

competitive experiments with recombinant protein and in cellular lysates. We also showed that DQs functionally inhibit PDI activity. Collectively, the data presented here are consistent with reports from other laboratories that demonstrate the initiation of ER stress due to quinone exposure and, specifically, inactivation of PDI family proteins upon quinone modification.[118, 250] For example, Farzam et al. identified that PDIA1 oxidoreductase activity was inhibited by 6-OHDA.[118] They also show that this inhibition is rescued in the presence of cysteine, but not methionine, suggesting that inhibition is a direct result of PDI modification and not the generation of ROS. The results from our study, therefore, provide conclusive evidence that PDI proteins are modified and functionally inactivated by cytosolic DA metabolites.

PD is a multifactorial and heterogeneous degenerative disease and several possible pathways involved in protein folding (e.g., ER stress), oxidative stress, and cellular organization (e.g., cytoskeletal and endosomal dysfunction) have been proposed as primary mechanisms leading to neurodegeneration.[111] Therefore, multiple pathways are likely involved in PD pathogenesis and an unbiased proteomics approach for identifying protein targets perturbed by the presence of DQs may help elucidate precise mechanisms by which DA contributes to neurodegeneration. PDI is upregulated during ER stress, for example, which is highly associated with protein misfolding and oxidative stress responses. Members of the PDI family are highly expressed during PD and are found in Lewy bodies from PD donor SNc.[251] Thus, our data suggest that PDI inactivation by DA contributes to ER dysfunction that may contribute to protein misfolding and Lewy body pathology, however more detailed experiments are needed to

confirm this. Interrogating the role of DA and PDI within these specific pathways may uncover molecular-level targets for future therapeutic discovery.

Modification of cytoskeletal proteins by DA has been proposed as another molecular mechanism for DPA-induced toxicity,[42] and several proteins involved in these pathways were identified as DA targets in our study. The microtubule associated proteins (MAPs) were identified here as proteomic hits are particularly interesting due to their interactions with the tau protein that forms neurofibrillary tangles during Alzheimer's disease pathogenesis. MAPs organize subcellular components, organelles, and synaptic vesicles. We also found proteins impacted by DPAs that are involved in proteasomal degradation pathways, such as UPS10 and HUWE3. Other studies have postulated that DA modification of these proteolytic enzymes contribute to DA derived toxicity, as cells exposed to DA lose the ability to break down damaged proteins.[134, 234] The identification of DPAs here that have been reported elsewhere further validates the utility of our probe technology to characterize DPAs. [152, 153, 163]

In addition to the identification of DPAs featured in PD- related pathogeneses, our chemo-proteomic experiments also identified novel DPAs that have not been reported previously. Notably, we found a significant enrichment of clathrin heavy chain 1 (CLTC), which affects synaptic vesicle formation.[252] This result is notable given recent GWAS data associating mutations in CLTC with impaired synaptic vesicle endocytosis and PD incidence.[253] Our discovery that DA modifies CLTC requires more thorough investigation to validate how DA modification specifically influences synaptic vesicle endocytosis.

It is important to note significant differences between our study and prior works, which likely account for the absence of PD-related DPAs enriched in our proteomics data set, including the redox sensor protein DJ-1 and the E3 ubiquitin ligase parkin. We attribute this difference to our use of undifferentiated SH-SY5Y cells and our dosing of cells at lower concentrations for longer exposures. Our cell-based experiments were conducted at 10 μM DA^{yne} to reduce the level of nonspecific or off-target adducts, while other studies have employed up to 150 μM DA.[152, 153] Low probe concentrations also enabled us to expose cells for 24 h. We hypothesize that these longer treatment conditions allow for study of time- resolved PD models, which should enable interrogation of the earliest mechanisms of synaptic density decline during aging and PD.

Controlling the complex oxidation chemistry of DA has been a major hurdle in understanding the consequences of dysregulated DA during normal aging and PD. While the development of a chemoproteomic DA-mimetic represents a significant advance in characterizing DPAs generated from oxidized DA, the rapid kinetics of AC formation under our conditions may limit enriched adducts to those generated with AC and, potentially, DHI. It is possible that cellular conditions favoring DA dysregulation may enhance DQ levels, in which case probes designed with amido-alkyne linkages or 6-substituted catechols may be more suitable to characterize these quinoproteins. The rate constant for ortho-quinones modification by thiols is approximately 10,000-fold higher than for amines and several studies have demonstrated that cysteine sidechains are the most susceptible protein residues towards DQ adduction, so we assume that the proteins identified in this current study are generated through cysteine residues.[71, 73] We foresee

utility in developing conditions that permit DA^{yne} to map specific sites of DQ modification on DPAs by enriching DQ-adducted peptides for mass spectrometric analyses to validate this and aid in determining motifs with high propensities for DQ-modification.[254-257] Until this problem is solved, future studies could explore a suite of DA mimetic probes that each represent a specific DA oxidation state (DQ, AC, DHI) to access any differential reactivity between these metabolites.

3 Chapter III: Development of Photocaged Dopamine Probes for Controlled Release Applications

This chapter is comprised of unpublished results collected by Alexander K. Hurben under the supervision of Prof. Tretyakova.

3.1 Introduction

Dopamine (DA) is an important neurotransmitter that plays numerous signaling roles. DA also possesses a catechol moiety which is prone to auto oxidation at cellular pH. As stated in **Chapters I** and **II**, the inherent reactivity of DA complicates analyzing its signaling functions in cellular systems. Thus, the development of caged DA delivery systems, which block DA signaling and metabolism until desired, would be valuable tools to study DA dependent signaling events, both in terms of neurotransmission processes as well as protein damage caused by reactive DA metabolites.

Photocaging represents an attractive strategy for controlling DA release. Light is a biocompatible trigger and a non-intrusive stimulus for DA release in cellular models. Development and application of photocaged biomolecules and therapeutics is a growing field of research.[258] Thus far, there are only a few reports of photocaged DA systems, which necessitates further exploration in this area.[226, 259, 260] Additionally, there are no current chemical tools capable of localizing DA release to specific subcellular locations. Such tools would enable uncaging DA molecules in specific subcellular locations in order to elucidate downstream signaling events in a controlled manner.

Here, we set out to expand the suite of photoactivatable DA compounds and to develop subcellular targeting systems to localize the site of DA release. We found that nitro benzene derived photocages enable rapid DA release under 350 nm light irradiation and that incorporation of a HaloTag targeting element, paired with the expression of the Halo protein in cells allows for subcellular localization of DA photocages in live cells.

These new chemical tools should prove useful in the continued study of DA mediated biology.

3.2 Materials and Methods

All chemicals and reagents required for the synthesis of DA photocages and preparation of buffers were purchased from Sigma Aldrich unless otherwise stated. NMR solvents were obtained from Cambridge Isotope Laboratories Inc. HEK 293T cells were acquired from ATCC. Cell culture media, additives, and consumables were purchased from Corning. DAPI and anti-actin antibody were obtained from ThermoScientific. AlamarBlueHS reagent and ProlongGlass were purchased from Invitrogen. The BCA kit was bought from Pierce. Mini-protean TGX precast gels, 0.2 µm nitrocellulose and filter paper transfer stacks were obtained from BioRad. THPTA and AlexaFlour647 azide were purchased from Click Chemistry Tools. Odyssey TBS blocking buffer and secondary antibodies were purchased from Licor. TMR Halo-ligand was obtained from Promega. The transfection reagent TransIT-2020 was obtained from Mirus Bio LLC. A 0.2-amp Spectroline (Model ENF-280C) UV hand lamp was used as the 365 nm light source in all experiments. Plasmids were obtained from Addgene. Specifically, Halo-Cyto, was a gift from Jin Wang (Addgene plasmid #124314). Licor Image Studio software was used to process acquired gel images and statistical analysis was performed with GraphPad.

Synthetic Procedures:

General Procedures:

Chemical reactions were conducted in a fumehood with oven-dried glassware under an argon atmosphere unless otherwise stated. Thin-layer chromatography (TLC) was performed with analtech silica uniplates and visualized under 254 nm UV light. Column chromatography was conducted with 60 mesh silica gel. A Buchi R-200 rotary

evaporator and V-100 vacuum pump were used to remove solvent. NMR spectra were taken on a 400 MHz Varian spectrometer.

tert-butyl (3,4-dihydroxyphenethyl)carbamate (**1**).

To an Ar purged solution of DA HCl (1.53 g, 10 mmol, 1.0 equiv) in MeOH (100 mL) was added Boc_2O (4.6 mL, 20 mmol, 2.0 equiv) followed by DIPEA (3.4 mL, 20 mmol, 2.0 equiv). The reaction was stirred for 6 h at 25 °C, and then concentrated *in vacuo*. The resultant mixture was diluted in DCM (100 mL) and washed successively with 0.1 N aqueous HCl (2 × 100 mL), brine (100 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo* to afford the title compound **1** (2.5 g, 99%) as an white solid: R_f = 0.5 (1:1 hexanes:EtOAc). ^{1}H NMR (400 MHz, METHANOL-d_4) δ ppm 1.40 (9 H, s) 2.57 (2 H, t, *J*=7.43 Hz) 3.16 (2 H, t, *J*=7.43 Hz) 6.49 (1 H, d, *J*=8.22 Hz) 6.57 - 6.63 (1 H, m) 6.65 (1 H, d, *J*=8.22 Hz) ^{13}C NMR (101 MHz, METHANOL-d_4) δ ppm 27.3 (3 C, s) 35.2 (1 C, s) 41.9 (1 C, s) 78.5 (1 C, s) 114.9 (1 C, s) 115.4 (1 C, s) 119.6 (1 C, s) 130.7 (1 C, s) 143.3 (1 C, s) 144.8 (1 C, s) 157.0 (1 C, s) LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{13}H_{20}NO_4^+$ 245.1, found 245.1.

General method for photocage installation:

To a flask charged with **1** (253 mg, 1.0 mmol, 1.0 equiv), K_2CO_3 (270 mg, 2.0 mmol, 2.0 equiv), and KI (166 mg, 1.0 mmol, 1.0 equiv) and precursor bromide a-c (1.0 equiv) was added acetone (10 mL). The reaction was stirred for 48 h at 25 °C, and then concentrated *in vacuo*. The resultant mixture was diluted with EtOAc (20 mL) and washed successively with 0.1 N aqueous HCl (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. Purification by silica gel flash

chromatography (gradient of 0 to 50% EtOAc in hexanes) afforded the title compounds **2a-c** (80%, 77%, 50%) as a 1:1 mixture of regioisomers and as orange solids.

tert-butyl (4-hydroxy-3-((2-nitrobenzyl)oxy)phenethyl)carbamate (**2a**)

^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.41 (8 H, s) 2.58 - 2.80 (2 H, m) 3.16 - 3.40 (2 H, m) 4.50 (1 H, br. s.) 5.49 - 5.58 (2 H, m) 6.57 - 6.85 (2 H, m) 6.86 - 6.94 (1 H, m) 7.42 - 7.57 (1 H, m) 7.61 - 7.79 (2 H, m) 7.91 - 8.01 (1 H, m) 8.10 - 8.20 (1 H, m) LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{20}H_{25}N_2O_6^+$ 389.2, found 389.2.

tert-butyl (3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-4-hydroxyphenethyl)carbamate (**2b**)

^{1}H NMR (400 MHz, CHLOROFORM-*d*) δ ppm 1.24 (2 H, br. s.) 1.41 (8 H, br. s.) 2.56 - 2.78 (2 H, m) 3.31 (2 H, br. s.) 3.73 - 3.85 (1 H, m) 3.85 - 4.08 (5 H, m) 4.53 (1 H, br. s.) 5.43 - 5.59 (2 H, m) 6.65 - 6.82 (2 H, m) 6.84 - 6.97 (1 H, m) 7.13 (1 H, br. s.) 7.72 (1 H, dd, *J*=15.26, 3.52 Hz) LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{22}H_{29}N_2O_8^+$ 449.2, found 449.2.

tert-butyl (4-hydroxy-3-(1-(3-nitrodibenzo[*b*,*d*]furan-2-yl)ethoxy)phenethyl)carbamate (**2c)**

^{1}H NMR (METHANOL-d_4,400MHz): δ = 8.54 (s, 1 H), 8.22 (s, 1 H), 8.09 (d, *J*=7.8 Hz, 1 H), 7.53 - 7.66 (m, 2 H), 7.42 (t, *J*=7.4 Hz, 1 H), 6.65 (s, 1 H), 6.58 (d, *J*=8.2 Hz, 1 H), 6.43 (d, *J*=7.0 Hz, 1 H), 6.03 (d, *J*=6.3 Hz, 1 H), 3.10 (t, *J*=7.2 Hz, 2 H), 2.52 (t, *J*=7.2 Hz, 2 H), 2.13 (s, 12 H), 1.78 ppm (d, *J*=6.3 Hz, 3 H). LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{27}H_{29}N_2O_7^+$ 493.2, found 493.2.

General method for Boc deprotection:

To a flask charged with **2a-c** (0.05 mmol, 1.0 equiv) was added a mixture of 1:1 DCM:TFA (2 mL). The reaction was stirred for 48 h at 25 °C, and then concentrated *in vacuo*. Purification by HPLC using a Sunfire C18 column (5 μm, 150 × 2.0 mm, 10.0 mL/min flow rate, buffer A: H_2O), buffer B: ACN) involved a linear gradient of 0-95% B (0-25 min), followed by an isocratic hold at 95% B (25-30 min) afforded the title compounds **3a-c** (95%, 97%, 90%) as a 1:1 mixture of regioisomers and as orange solids.

4-(2-aminoethyl)-2-((2-nitrobenzyl)oxy)phenol **3a**

^{1}H NMR (400 MHz, METHANOL-d_4) δ ppm 2.77 - 2.89 (2 H, m) 3.05 - 3.16 (2 H, m) 5.47 - 5.56 (2 H, m) 6.78 (1 H, m) 6.82 - 6.91 (2 H, m) 7.57 (1 H, t, *J*=7.43 Hz) 7.73 (1 H, q, *J*=7.17 Hz) 7.85 - 7.97 (1 H, m) 8.09 - 8.19 (1 H, m). LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{15}H_{17}N_2O_4^+$ 289.1, found 289.1.

4-(2-aminoethyl)-2-((4,5-dimethoxy-2-nitrobenzyl)oxy)phenol (**3b**)

^{1}H NMR (400 MHz, METHANOL-d_4) δ ppm 2.81 (2 H, td, *J*=7.63, 3.91 Hz) 3.04 - 3.13 (2 H, m) 3.84 - 3.94 (6 H, m) 5.51 (2 H, d, *J*=10.56 Hz) 6.66 (1 H, dd, *J*=8.22, 1.96 Hz) 6.73 - 6.80 (1 H, m) 6.81 - 6.91 (2 H, m) 7.42 (1 H, s) 7.46 (1 H, s) 7.76 (1 H, d, *J*=1.57 Hz) LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{17}H_{21}N_2O_6^+$ 349.1, found 349.1.

4-(2-aminoethyl)-2-(1-(3-nitrodibenzo[*b*,*d*]furan-2-yl)ethoxy)phenol (**3c)**

^{1}H NMR (700 MHz, METHANOL-d_4) δ ppm 1.85 (3 H, d, *J*=6.19 Hz) 2.78 (2 H, t, *J*=7.30 Hz) 3.07 (2 H, t, *J*=7.30 Hz) 6.12 (1 H, d, *J*=6.19 Hz) 6.55 (1 H, d, *J*=8.40 Hz) 6.70 (1 H, d, *J*=8.40 Hz) 6.77 (1 H, br. s.) 7.46 - 7.50 (1 H, m) 7.62 - 7.67 (1 H, m) 7.70

(1 H, d, *J*=7.96 Hz) 8.14 (1 H, d, *J*=7.52 Hz) 8.29 (1 H, s) 8.58 (1 H, s). LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{22}H_{21}N_2O_5^+$ 393.1, found 393.1.

ethyl 2-(4-formyl-2-methoxyphenoxy)acetate (**5**)

To a solution of ethyl 2-bromoacetate (2.0 g, 13 mmol, 1.0 equiv) in ACN (50 mL) was added ethyl 2-bromoacetate (1.2 mL, 14 mmol, 1.1 equiv) and K_2CO_3 (3.6 g, 26 mmol, 2.0 equiv). The reaction was stirred for 14 h at 25 °C, and then concentrated *in vacuo*. The resultant mixture was diluted in DCM (100 mL) and washed successively with H_2O (2 × 100 mL), brine (100 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo* to afford the title compound **5** (3.7 g, 99%) as a white solid. ^{1}H NMR (CHLOROFORM-*d*,400MHz): δ = 9.84 (s, 1 H), 7.34 - 7.48 (m, 2 H), 6.86 (d, *J*=8.2 Hz, 1 H), 4.76 (s, 2 H), 4.18 - 4.32 (m, 2 H), 3.94 (s, 3 H), 1.20 - 1.33 ppm (m, 3 H). LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{12}H_{15}O_5^+$ 239.0, found 239.0.

ethyl 2-(4-formyl-2-methoxy-5-nitrophenoxy)acetate (**6**)

Compound **5** (574 mg, 2.4 mmol, 1.0 equiv) was slowly added portion wise to ice cooled soln of HNO_3 (5 mL). The reaction was stirred for 30 min and allowed to warm to 25 °C. The reaction was quenched with the addition of 50 mL ice cold water, which yielded the final product **6** (479 mg, 60%) as a pale yellow precipitate. ^{1}H NMR (CHLOROFORM-*d*, 400MHz): δ = 10.43 (s, 1 H), 7.52 (s, 1 H), 7.35 - 7.47 (m, 1 H), 4.81 (s, 2 H), 4.28 (q, *J*=7.0 Hz, 2 H), 4.02 (s, 3 H), 1.30 ppm (t, *J*=7.2 Hz, 3 H). LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{12}H_{13}NO_7^+$ 284.0, found 284.0.

2-(4-formyl-2-methoxy-5-nitrophenoxy)acetic acid (**7**)

To a solution of **6** (1 g, 3.8 mmol, 1.0 equiv) in MeOH (10 mL) was added a soln of LiOH (632 mg, 15.4, 4 equiv) in 2 mL of H_2O. The reaction was stirred for 30 min at 25 °C, and then concentrated *in vacuo*. The resultant mixture was diluted in 1:1 of 1 N HCl:DCM (100 mL) and washed successively with H_2O (2 × 100 mL), brine (100 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo* to afford the title compound **7** (811 mg, 84%) as an white solid. ^{1}H NMR (CHLOROFORM-*d*,400MHz): δ= 10.45 (s, 1 H), 7.59 (s, 1 H), 7.44 (s, 1 H), 4.88 (s, 2 H), 4.03 ppm (s, 3 H) LRMS (ESI) *m/z* [M + H]$^+$ calculated for $C_{10}H_{10}NO_7^+$ 256.0, found 256.0.

N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-(4-formyl-2-methoxy-5-nitrophenoxy)acetamide (**8**)

To a solution of **7** (50 mg, 0.16 mmol, 1.0 equiv), DIPEA (272 uL, 0.64 mmol, 4.0 equiv), and halo ligand 2-(2-((6-chlorohexyl)oxy)ethoxy)ethan-1-amine[261] (70 mg, 0.17 mmol, 1.05 equiv) in DMF (3 mL) was added EDC HCl (41 mg, 0.17, 1.1 equiv). The reaction was stirred for 18 h at 25 °C, and then concentrated *in vacuo*. The resultant mixture was diluted with DCM (20 mL) and washed successively with 0.1 N aqueous HCl (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. Purification by silica gel flash chromatography (gradient of 0 to 100% EtOAc in hexanes) afforded the title compound **8** (15 mg, 17%) as yellow oil. ^{1}H NMR (CHLOROFORM-*d*, 400MHz): δ = 10.45 (s, 1 H), 7.62 (s, 1 H), 7.43 (s, 1 H), 4.63 (s, 2 H), 4.02 (s, 3 H), 3.53 - 3.64 (m, 8 H), 3.50 (t, *J*=6.7 Hz, 2 H), 3.43 (t, *J*=6.7 Hz, 2 H), 1.73 (d, *J*=7.4 Hz, 3 H), 1.53 - 1.61 ppm (m, 3 H) LRMS (ESI) *m/z* [M + H]$^+$ calculated for $C_{20}H_{30}ClN_2O_8^+$ 461.2, found 461.2.

N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-(4-(hydroxymethyl)-2-methoxy-5-nitrophenoxy)acetamide (**9**)

To a solution of **8** (15 mg, 0.03 mmol, 1.0 equiv) in THF (2 mL) was added $NaBH_4$ (15 mg, 0.13, 4 equiv). The reaction was stirred for 15 min at 25 °C, and then quenched with 5 mL of sat. NH_4Cl. The resultant mixture was diluted with EtOAc (10 mL) and washed successively with H_2O (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. Purification by silica gel flash chromatography (gradient of 0 to 100% EtOAc in hexanes) afforded the title compound **9** (12 mg, 80%) as yellow oil. ^{1}H NMR (CHLOROFORM-*d*,400MHz): δ = 7.73 (s, 1 H), 7.13 (br. s., 1 H), 4.98 (s, 2 H), 4.56 (s, 2 H), 4.00 (s, 3 H), 3.52 - 3.63 (m, 9 H), 3.50 (t, *J*=6.7 Hz, 2 H), 3.43 (t, *J*=6.7 Hz, 2 H), 1.69 - 1.79 (m, 3 H), 1.50 - 1.62 (m, 7 H), 1.40 (br. s., 4 H), 1.34 ppm (d, *J*=7.0 Hz, 3 H) LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{20}H_{32}ClN_2O_8^+$ 463.2, found 463.2.

N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-(4-((2-hydroxy-4-(2-(prop-2-yn-1-ylamino)ethyl)phenoxy)methyl)-2-methoxy-5-nitrophenoxy)acetamide (**10**, TRDP)

To a solution of **9** (12 mg, 0.025 mmol, 1 equiv) in 1 ml of THF was added PBr_3 (3.5 uL, 0.038, 1.5 equiv) dropwise. The reaction was stirred for 18 h at 25 °C and then $BocDA^{yne}$ (5 mg, 0.025 mmol, 1 equiv) and K_2CO_3 (6 mg, 0.05 mmol, 2 equiv) were added. The reaction was stirred for 3 days at 25 °C, and then concentrated *in vacuo*. Next, 2 mL of a 1:1 soln of TFA and DCM were added. The reaction was stirred for 30 min at 25 °C, and then concentrated *in vacuo*. Purification by HPLC using a Sunfire C18 column (5 μm, 150 × 2.0 mm, 10.0 mL/min flow rate, buffer A: H_2O), buffer B: ACN)

involved a linear gradient of 0-95% B (0-25 min), followed by an isocratic hold at 95% B (25-30 min) afforded the title TRDP **10** (0.7 mg, 5%) as a 1:1 mixture of regioisomers and an orange solid. ^{1}H NMR (METHANOL-*d*4, 400MHz): δ = 8.10 (s, 1 H), 7.83 (s, 1 H), 7.56 (s, 1 H), 7.51 (br. s., 1 H), 6.81 - 6.93 (m, 2 H), 6.73 - 6.81 (m, 2 H), 5.48 - 5.56 (m, 2 H), 4.62 (s, 2 H), 3.94 - 4.01 (m, 2 H), 3.91 (dd, *J*=5.7, 2.5 Hz, 2 H), 3.52 - 3.62 (m, 4 H), 3.40 - 3.52 (m, 4 H), 3.11 (br. s., 1 H), 2.81 - 2.90 (m, 2 H), 1.91 (s, 2 H), 1.63 - 1.75 (m, 3 H), 1.49 - 1.58 (m, 2 H), 1.30 - 1.46 (m, 4 H), 1.27 ppm (s, 1 H). LRMS (ESI) *m/z* $[M + H]^+$ calculated for $C_{31}H_{43}ClN_3O_9^+$ 636.3, found 636.3.

Photolysis of Caged DA Probes Monitored by HPLC:

Caged DA samples designated for photolysis were prepared by adding 50 μL of a solution containing 100 μM in PBS (pH 7.4, 0.15% DMSO) to a 0.6 mL PCR tubes. The samples were irradiated at 365 nm via a Spectroline UV hand lamp positioned 1 inch above the midpoint of the tube. At 30 s timepoints, 20 μL of the reaction were removed and analyzed by HPLC using a Luna C18 column (5 μm, 150 × 2.0 mm, 1.0 mL/min flow rate, buffer A: ammonium acetate (pH 4), buffer B: MeOH) involved a linear gradient of 0-95% B (0-15 min), followed by an isocratic hold at 95% B (15-20 min)). Absorbance was monitored at 254 nm and peak areas corresponding to caged compounds were integrated and normalized to the peak area at time zero to determine the half-life of under 365 nm light exposure over three independent replicates.

Photolysis of Caged DA Probes Monitored by NMR:

Solutions of caged DA compounds (2 mM, deuterated PBS) were placed into quartz NMR and were irradiated at 365 nm via a Spectroline UV hand lamp positioned 1

inch above the midpoint of the tube. Samples were analyzed via ^{1}H NMR at the indicated timepoints.

Alamar Blue Cell Viability Assay:

HEK 293T cells were seeded into a 96 well plate at a density of 20,000 cells per well in 100 µL of FBS medium and incubated for 24 h. The FBS medium was removed and replaced with the indicated treatment condition in FBS-free MEM media. Specifically, following a 3 h TRDP treatment to mimic control released assay conditions, the medium was refreshed, and the cells were further incubated for 18 h. After incubation, the medium was removed and replaced with 100 µL of a 10% alamar blue solution in FBS-free MEM medium. The plates were incubated for 4 h in the dark at 37 °C and then fluorescence was measured at Ex/Em 560/590 nm on a BioTek Synergy H1 plate reader. Cell viability for a given treatment condition was calculated by normalizing the fluorescence signal to the DMSO treated control.

Confocal Microscopy:

HEK 293T cells were cultured on Neuvitro GC-25-1.5-Laminin coverslips to 70% confluency with MEM medium containing 10% FBS and Pen/Strep 100 U/mL under a 5% CO_2 atmosphere at 37 °C. Cells were transfected with 2.5 µg Halo plasmids via the serum-compatible Mirus TransIT-2020 transfection reagent following the manufacturer's instructions. After 24 h, the medium was removed and replaced with the indicated treatment condition in FBS-free medium. For competitive binding experiments, cells were treated with DMSO vehicle or 12.5 µM TRDP for 2 h. Then the cells were washed with serum-free medium and then incubated with 5 µM HaloTag TMR for 1 h. Following

the allotted incubation time, the cells were washed three times with 2 mL PBS, fixed with 2 mL of 4% PFA (molecular biology grade) in PBS for 10 min, and washed an additional three times with PBS. Next, three washes with PBT were performed. All cells' nuclei were stained with 300 nM of DAPI for 30 min and were further washed an additional three times with PBS. A drop of prolong glass anti-photobleaching solution was placed on the coverslips, which were mounted to glass slides and allowed to cure for 48 h. The slides were then sealed with clear nail polish and stored at 4 °C until they were imaged. Images were acquired on an Olympus FluoView FV1000 BX2 upright confocal microscope using a PLAPON 60X O NA:1.42 objective. Laser Ex/Em were set at 405:461, and 557:576 for DAPI, and TMR, and AlexaFlour647, respectively.

General Protocol for the Controlled Release of DA^{yne} with TRDP in Live Cells:

HEK 293T cells were cultured in MEM supplemented with 10% v/v FBS in the presence of 100 U/mL Pen/Strep antibiotics under a 5% CO_2 atmosphere at 37 °C. Cells were seeded into 6-well tissue treated plates at 400,000 cells per well. Cells were transfected with 2.5 μg Cyto-Halo, via the serum-compatible Mirus TransIT-2020 transfection reagent (7.5 μL) following the manufacturer's instructions. After 24 h following transfection, cells were washed with PBS and subjected to the desired treatment conditions in FBS free medium. Cells treated with TRDP were incubated with 12.5 μM TRDP for 2 h, then the medium was removed and replaced with fresh FBS-free medium and was further incubated for 30 min. Two additional 30-min washes were performed to ensure unbound TRDP was removed. For metal treated samples, the cells were incubated in PBS containing $FeCl_3$ or $CuCl_2$ (50 μM) in FBS-free medium 10 min

prior to irradiation. Samples were irradiated at 365 nm with a handheld UV lamp for 20 min which was positioned 1 inch above the cell monolayer. Cells were incubated at 37 °C for the specified amount of time prior to lysis and downstream analysis.

Following the allotted incubation time, the cells were collected through trypsinization, washed twice with PBS, and lysed in 200 µL of MPER lysis buffer supplemented with HALT protease inhibitors. The lysates were centrifuged at 14,000 g for 8.5 min to remove debris. The supernatant was collected, and the protein concentrations were determined through a BCA assay per manufacturer's protocol.

Click reactions were performed in a total volume of 100 µL at a protein concentration of 50 µg/mL with 10 µM AlexaFlour647 azide, 1.5 mM THPTA, 1 mM $CuSO_4$ and 3 mM sodium ascorbate. The reactions were agitated at 25 °C for 1 h in the dark. Proteins were precipitated with 400 µL cold acetone overnight at -20 °C. Following a 20 min centrifugation at 21,000 g, the acetone was carefully aspirated, and the protein pellets were dissolved into 30 µL of 1× Licor loading buffer with 10% β-mercaptoethanol and denatured for 5 min at 95 °C, then 13.5 µL of each sample was loaded into 4–20% gradient mini-protean TGX precast gels for SDS-PAGE. Gels were fixed overnight in an aqueous solution of 30% EtOH and 10% acetic acid to wash out excess fluorophore and imaged with a Licor odyssey CLX instrument to visualize protein adducts. The gels were then stained with coomassie blue to confirm equivalent loading.

3.3 Results

Our first objective was to synthesize a set of photocaged DA compounds. We selected nitro benzene derived photocage scaffolds based on their reported ability to release phenols and their synthetic tractability.[262] As shown in **Figure 3.1** photocaged DA 3a-c were obtained in three steps which commenced with the Boc protection of DA's free amine. Next the photoprotecting groups were installed via reaction with the corresponding brominated precursors of 2-nitrobenzene (Nitro), 4,5-dimethoxy-2-nitrobenzene (DMN), and nitrodibenzofuran (NDBF). The Nitro and DMN bromides were commercially available and the NDBF bromide was synthesized according to previous reports.[263] Reaction with the bromides yielded a mixture of regioisomers at both hydroxyl positions on the catechol. The final compounds were furnished upon acidic removal of the Boc group.

Next, we examined the UV-vis spectra of Nitro, DMN, and NDBF DA. The photocaged DA compounds were dissolved in PBS and analyzed in a spectrometer to afford the spectra shown in **Figure 3.2** Both the DMN and NDBF derivatives produced red-shifted local maxima at ~350 nm compared to Nitro DA which had a local maxima at ~280 nm. This result is expected due to the increased electronic density within the ring system. These data also indicate that DMN and NDBF DA photocaged should photolyze efficiently at 360 nm irradiation.

With our DA photocaged compounds in hand, we moved to evaluate their photolysis efficiency. DA photocaged compounds were dissolved in phosphate-buffered saline (PBS) pH 7.4 and irradiated with 365 nm. Photolysis was tracked via HPLC-UV as

shown in **Figure 3.3**. NDBF DA underwent rapid photolysis with a half-life of ~0.5 min. DMA DA had a half-life of ~ 3 min and Nitro DA had a haft-life of ~5 min. These results revealed that increased electron density within the nitro benzene based photocage led to an increased photolysis rate. Such data is important for the continued development of DA and other phenol based photocages as rapid uncaging is desirable in biological applications.

We also confirmed DA release via NMR. As shown in **Figure 3.4A**, the proton resonances corresponding to NDBF DA dimmish in intensity as a function of 365 nm irradiation time and give way to peaks corresponding to DA and 1-(3-nitrosodibenzofuran-2-yl)ethan-1-one. Given the NMR evidence and reported nitro benzene photolysis mechanisms, we propose that our photoprotected DA compounds uncage through mechanism shown in **Figure 3.4B**. This mechanism begins with the light mediated excitation of the nitro group, which rearranges into a hemiacetal species and undergoes hydrolysis to afford DA and the nitroso byproduct.[264]

Our next objective was to design and synthesize a caged DA probe which could be targeted to different subcellular locations to release DA in a regionally defined area of a cell on demand. To do so, we selected the HaloTag technology,[265] comprised of a Halo protein and chloroalkane targeting ligand, which react to generate a 1 : 1 ligand to protein covalent complex. This Halo system provides synthetic tractability, rapid target engagement, and documented success in similar applications.[224] We initially attempted to install the HaloTag targeting element on to the NDBF photocage, but encountered synthetic obstacles. Thus we opted to modify the DMN scaffold to yield the targetable

releasable dopamine probe, TRDP, (**Figure 3.5A**), taking inspiration from reports of similar photocages immobilized to solid supports.[266]

The overall synthetic route to obtain TRPD is shown in **Figure 3.5B**. Synthesis commenced with the alkylation and nitration of 4-hydroxy-3-methoxybenzaldehyde. The installed ester was then deprotected to the corresponding acid to enable installation of the halotag ligand through EDC mediated amide coupling. Next, the aldehyde was reduced and converted to a bromide to allow for DA, or in our case here, the alkynylated DA probe, DA^{yne}, to be incorporated. As discussed in **Chapter II**, the DA^{yne} probe can be used to report on DA adducted proteins. Finally, acidic deprotection furnished TRDP.

With TRDP in hand, we sought to examine its effect on cell viability and ability to engage transiently expressed Halo protein within the cell. We found that TRDP had no significant effect on HEK 293T cell viability as measured by an alamar blue assay at concentrations as high as 12.5 μM (**Figure 3.6A**). Thus, TRDP is tolerated at low doses. We next assessed the cell penetrability and target engagement of TRDP with Halo protein in live cells. To do so, we performed a competitive binding experiment with TRDP and a tetramethyl rhodamine (TMR) HaloTag ligand. The TMR HaloTag ligand contains a chloroalkane moiety which covalently binds Halo protein and can be visualized through confocal microscopy. However, this interaction can be blocked by another chloroalkane containing compound, in our case TRDP, resulting in the loss of TMR signal. Cells transfected with Halo protein and treated with 5 μM TMR HaloTag ligand alone displayed a high degree of fluorescence in their cytosol, whereas pre-treatment with 12.5 μM TRDP prior to TMR HaloTag ligand administration ablated fluorescent signal as

shown in **Figure 3.6B**. These results demonstrate that TRDP penetrates cell membranes and can engage cytosolic Halo protein.

Finally, we explored if our TRDP platform could be used to confer spatiotemporal control over DA protein adduct formation in live cells. We implemented a gel-based assay to visualize proteins adducted by the released DA^{yne}. Halo transfected HEK 293T cells were treated with 12.5 μM TRDP for two hours and then subjected to three washes, once every 30 min, to remove any unbound TRDP. The cells were then irradiated with 365 nm light for 20 min, to release DA^{yne}. After light exposure, the cells were incubated for either 1, 16 or 24 hours. Afterwards the cells were collected, washed, lysed, and subjected to a CuAAC reaction, which appended the Alexa647 fluorophore to any DA^{yne} protein adducts. The lysates were resolved through SDS-PAGE, allowing proteins modified by DA^{yne} and remaining Halo–TRDP complex to be visualized via fluorescent gel imaging. As shown in **Figure 3.7A**, 20 min of 365 nm irradiation reduced the fluorescent intensity of the Halo–TRDP complex by 80%, indicating that TRDP was successfully photolyzed in a cellular setting. However, despite the successful release of DA^{yne}, we did not observe any prominent fluorescent bands corresponding to DA protein adducts at any of the timepoints the cells were incubated post light exposure. We hypothesize that the lack of DA protein adducts generated might be because of the relatively low concentration of DA^{yne} released by the TRDP platform. The amount of DA^{yne} released is estimated to be ~ 10 μM (see **Chapter IV** for Halo protein intracellular concentration calculations). This amount may be too low to generate a substantial amount of reactive DA metabolites responsible for driving adduct formation. As at low

concentrations the released DA^{yne} is likely rapidly detoxified into nonreactive species, and thus abundant DA proteins adducts are not observed. This is unlike when the cells were treated with 500 μM DA^{yne} for 4 hours, which yielded numerous fluorescently labeled DA protein adducts (**Figure 3.7A**).

In an effort to improve the yield of intracellular DA protein adducts generated under our controlled release platform, we exposed cells to 50 μM Cu^{2+} or Fe^{3+} prior to light exposure. These treatments were to simulate a pro-oxidative cellular environment by increasing levels of redox active metals. Previous reports suggest that altered neuronal metal concentrations may be linked to PD progression and induce intracellular DA oxidation.[267, 268] As shown in **Figure 3.7B**, pretreatment with Cu^{2+} prior to DA release produced a higher level of high molecular weight florescent protein bands in comparison to Fe^{3+} or non metal treated cells. This suggests that elevated levels of intracellular Cu^{2+} could induce DA oxidation towards reactive adduct forming metabolites. Collectively, these results show that the TRDP platform is capable of releasing DA^{yne} upon light exposure, however the levels of DA^{yne} released do not generate abundant DA protein adducts.

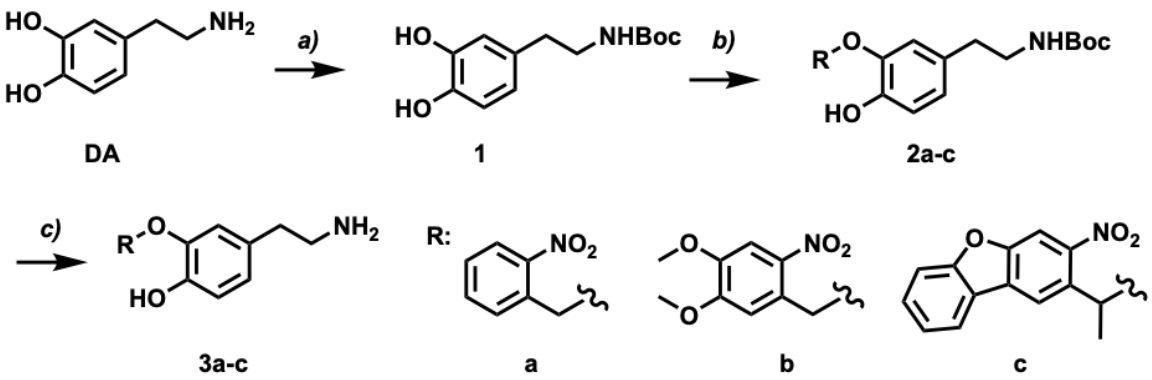


Figure 3.1. Synthesis of photocaged DA compounds

(a) Boc_2O, DIPEA, MeOH; 99%. (b) bromide precursor a-c, K_2CO_3, KI, acetone; 80-50%. (c) TFA, DCM; 98-95%.

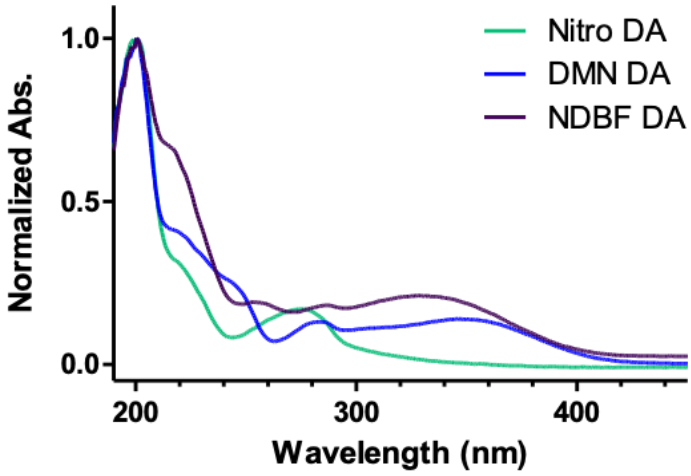


Figure 3.2. UV-vis spectra of photocaged DA compounds

Compounds were analyzed at 0.1 mM in PBS. Spectra were obtained an a nanodrop spectrometer.

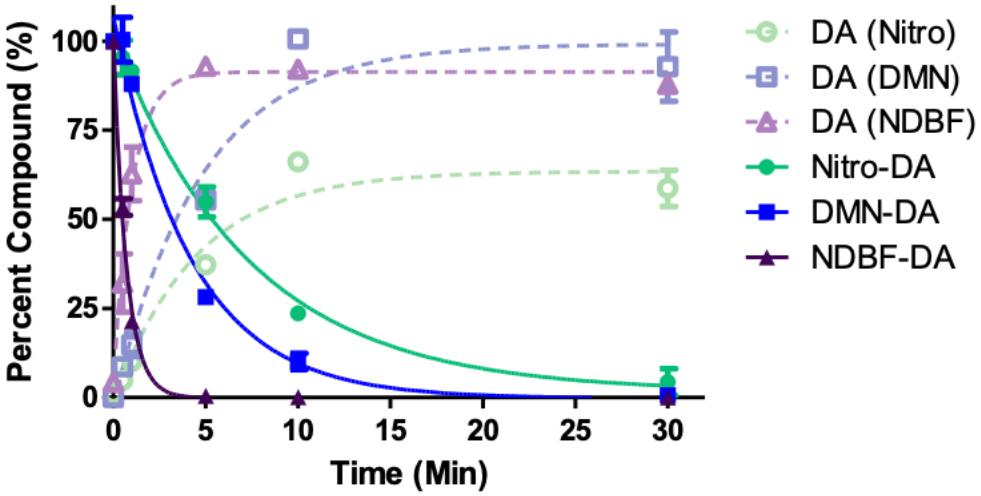


Figure 3.3. Photolysis of DA photocages

Solutions of DA photocages (0.1 mM) in PBS (pH 7.4) were irradiated for 0.5, 1-, 5-, 10- and 30-min time intervals at which aliquots were removed and analyzed via HPLC. Integrated peak areas of indicated compounds were normalized to time zero plotted against 365 nm light exposure time. Data are represented as mean ± SEM of 3 independent replicates and fit to a one phase exponential decay function.

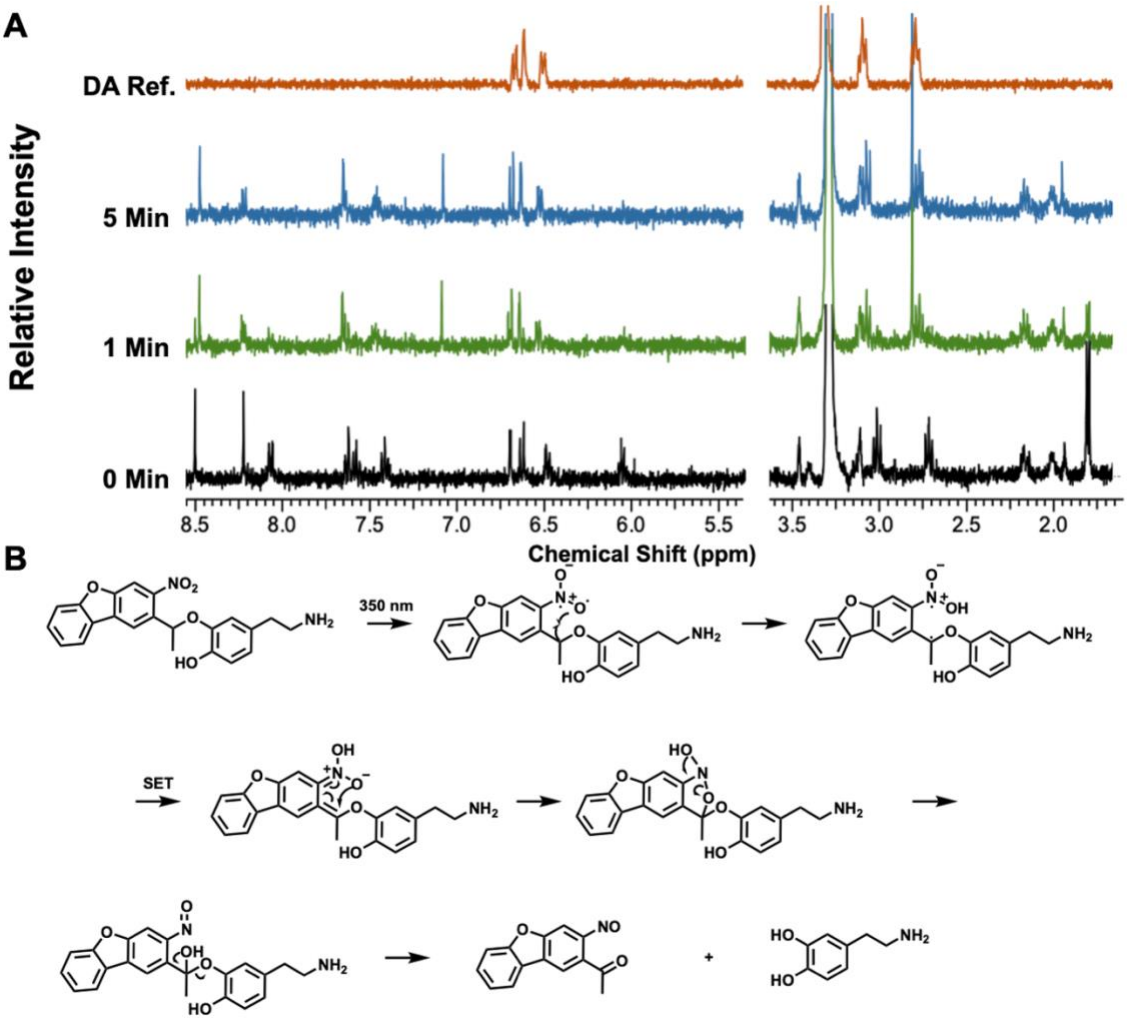


Figure 3.4. NMR characterization and proposed mechanism of NDBF DA photolysis

A) NMR spectra time course of 1 mM NDBF DA in a solution of MeOD irradiated with 360 nm light for the specified times. Water solvent peak was removed for clarity. B) Postulated mechanism of NDBF DA photolysis.

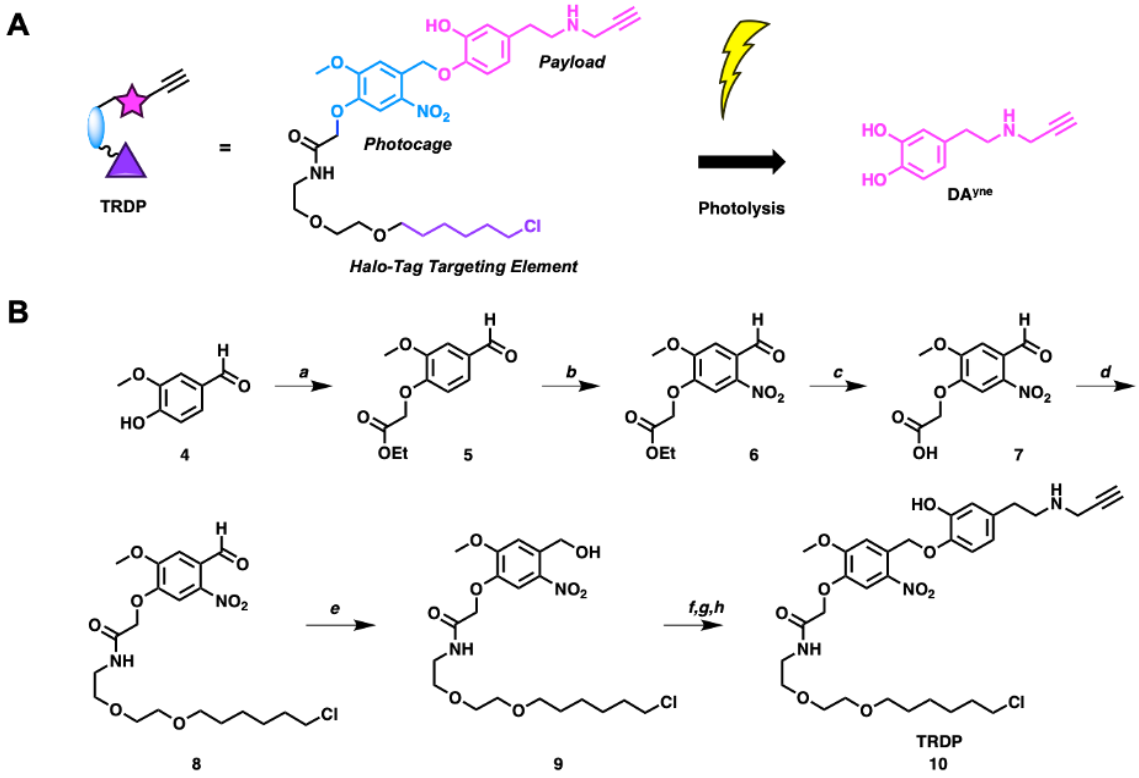


Figure 3.5. Design and synthesis of TRDP

A) Schematic of TRDP and photo release of DA^{yne}. B) Synthesis of TRDP: (a) ethyl 2-bromoacetate and K_2CO_3; 99% (b) HNO_3; 60% (c) LiOH; 84% (d) EDC, Halo ligand, and DIPEA; 17%, (e) $NaBH_4$; 60% (f) PBr_3 (g) Boc DA^{yne} and K_2CO_3 60% (h) TFA; 5% (over three steps).

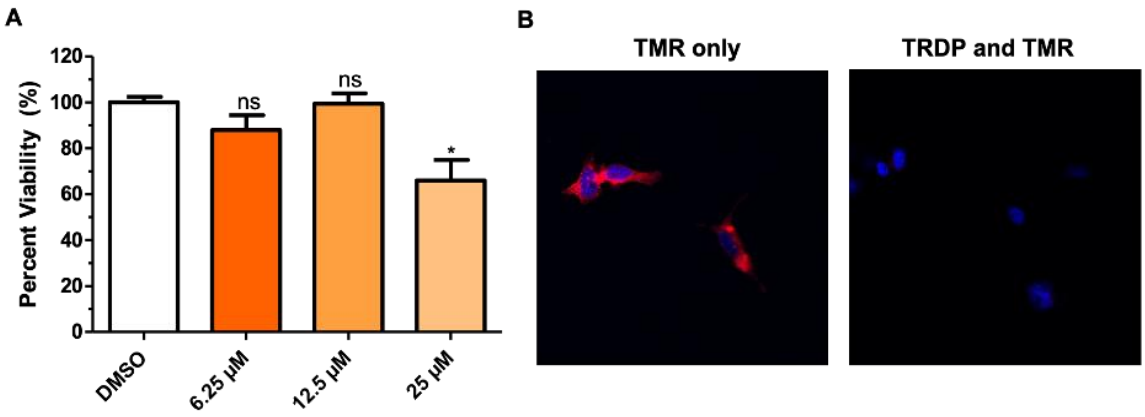


Figure 3.6. TRDP is nontoxic and engages halo protein

A) Cell viability of HEK 293T cells as measured by alamar blue following 18 h incubation with TRDP. Data are represented as mean ± SEM from 4 technical replicates and normalized to DMSO control. Statistical analysis was performed by a one-way ANOVA (top bar) and Tukey post hoc test (* $P < 0.05$). B) Confocal imaging of HEK 293T cells transfected with Halo protein treated with either 5 µM TMR HaloTag only (left), or with 12.5 µM TRDP for 1 h prior to treatment with 5 µM TMR HaloTag (right). DAPI is shown in blue and TMR is shown in red.

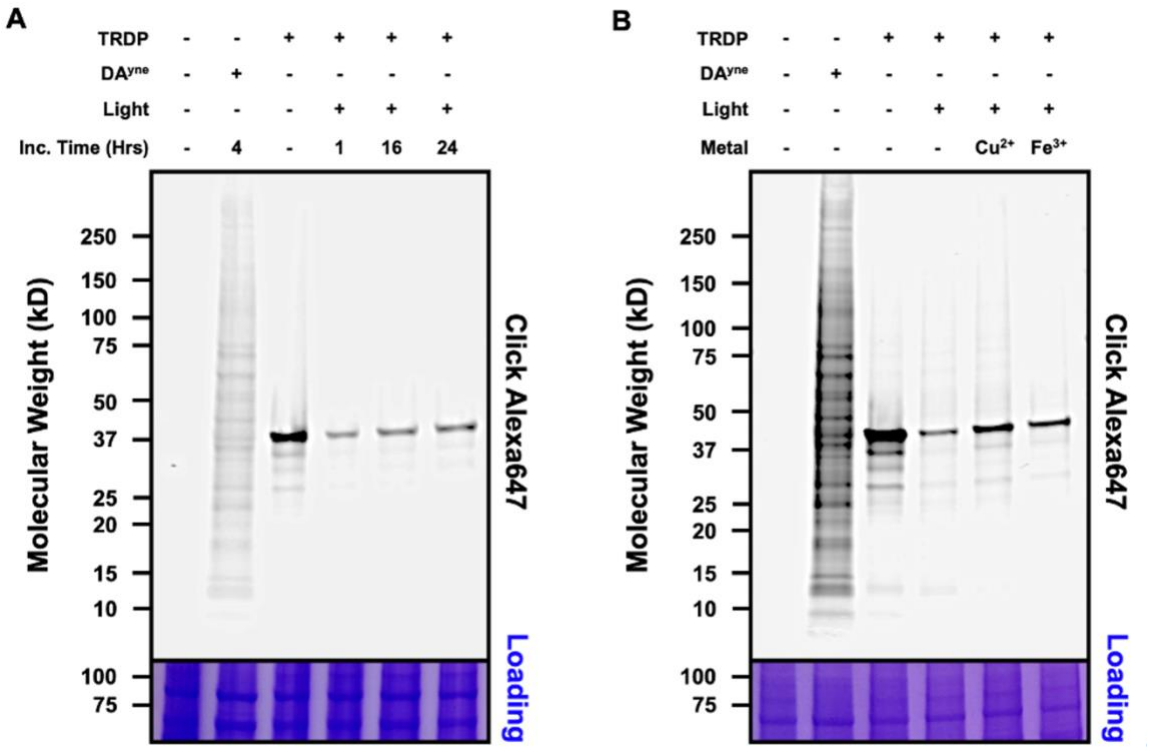


Figure 3.7. Controlled release of DA^{yne} by TRDP platform in live HEK 293T cells

A,B) Representative gels of Halo- transfected cells treated with either the indicated conditions: 100 μM DA^{yne} for 4 h, pretreatment with X μM of Cu^{2+} or Fe^{3+}, 12.5 μM of TRDP with or without a 20 min 365 nm light. Following incubation at the specified time, the cells were lysed, and lysates were subjected to a CuAAC reaction with Alexa647 azide to visualize DA protein adducts. Top: Alexa Fluor 647 fluorescence intensity. Bottom: Imperial protein stain to confirm equivalent protein loading.

3.4 Discussion

Controlling DA dynamics is important for studying DA signaling pathways and DA metabolism. Here, we showed how DA can be photocaged effectively with nitro based photoprotecting groups. We also designed a photocage functionalized with a Haloligand for targeting DA release to subcellular locations with the Halotag system. While releasing DA^{yne} with this targeted release system failed to produce abundant DA protein adducts, this tool is expected to have utility in studying other DA dependent biological processes that respond to small, localized bursts of released DA.

In the case of generating DA protein adducts on demand, TRDP could be redesigned to release a more reactive quinone species instead of a catechol moiety. Controlled release approaches to generated protein adducts on demand have been successful when releasing lipid derived electrophiles and dicarbonyls.[221, 223] Thus, direct release of a more reactive DA quinone metabolite is expected to yield higher levels of protein adducts. However, controlled release of DA is still useful in studying DA signaling as shown by work with reported photocaged DA molecules.[226, 259, 260] We envision that the nitro based photocages and the TRDP platform will aid in elucidating the precise mechanisms by which DA influences cellular phenotype and signaling events.

4 Chapter IV: Photocaged Dicarbonyl Probe Provides Spatiotemporal Control Over Protein Glycation

Adapted in part with permission by the Royal Chemistry Society from:

Hurben A.K., Ge P., Bouchard J.L., Doran T.M., Tretyakova N.Y. Controlling Protein Glycation in Cells with a Light Activatable Dicarbonyl Probe. *Chemical Communications*, **2022**, *58* (6), 855–858.

This chapter is adopted from a manuscript published in *Chemical Communications* with contributions from Alexander K. Hurben (synthesis, photolysis experiments, cellular experiments), Peng Ge (transfection protocol optimization) and Jacob Bouchard (expression and purification of halo protein) under the supervision of Prof. Tretyakova and Prof. Doran, which was supplemented with additional unpublished results obtained by Alexander K. Hurben under the supervision of Prof. Tretyakova.

4.1 Introduction

Posttranslational modification (PTM) of proteins greatly expands the number of possible proteoforms beyond the number of those explicitly genetically encoded.[269-272] PMTs can alter the structure and function of proteins and add a layer of chemical complexity to the proteome.[273-277] To date, there are over 700 reported PTMs in the UniProt's PTM Knowledgebase.[278] PMTs can be installed and removed through enzymatic mechanisms,[279] or caused by reactive metabolites which form covalent adducts with amino acid side chains.[280] The abundance of non-enzymatically derived PMTs suggest a link between metabolism and regulation of protein function and link between health and disease.[273, 281-283]

Protein glycation is an important non-enzymatic PTM that influences cell signaling, alters chromatin structure, and is implicated in numerous age- related diseases such as cancer, neurodegeneration, diabetes, and cardiovascular disease.[284-286] The principle metabolic driver of protein glycation is methylglyoxal (MGO), an endogenous dicarbonyl reactive electrophile species (RES) generated by the spontaneous degradation of glycolytic intermediates.[287] The α-oxoaldehyde moiety within MGO readily forms covalent adducts with proteins and DNA, resulting in advanced glycation end products (AGEs) (**Figure. 4.1A**).[288] Arginine and lysine are the primary amino acid residues modified by MGO, but cysteine is also susceptible to reversible modifications.[289] Of the DNA bases, guanine and adenine are prone to glycation by MGO.[290] It should be noted that MGO can also generate crosslinks between proteins, proteins and DNA and proteins and metabolites, this will be further explored in **Chapter V** of this thesis.[291] As shown in

Figure 4.1B, glycation modifications are structurally diverse. Currently there are unanswered questions about the biological significance of these non-enzymatic PTMs in terms of how they affect protein function and influence phenotype. The development of technologies capable of identifying proteins susceptible to glycation and understanding the functional ramifications of such adducts is critical to elucidating the role dicarbonyl metabolites like MGO play in ageing and disease.[292] As highlighted in **Chapters I** and **II** of this thesis, protein modifications caused by reactive electrophile species are important events which can drive toxicity and influence phenotypic outcomes.

Various tools have been developed to study AGEs generated by MGO, including MGO specific antibodies and alkyne functionalized MGO probes.[293-297] While these approaches allow for detection and identification of glycated proteins within cells, they often require non-native conditions and employ MGO concentrations of > 100 μM, which is orders of magnitude higher than physiological levels (1–10 μM).[298] Such dosing regimens, coupled with the innate reactivity of dicarbonyl compounds, exerts extraneous influence on the cellular responses being investigated (**Figure 4.2A**). Thus, novel strategies enabling precise release of biologically relevant concentrations of dicarbonyl compounds intracellularly are needed in order to comprehensively understand the processes that regulate glycated proteins and to elucidate their influence on cellular phenotype.

To address the current gap in methodology, we sought to develop a platform capable of releasing an intracellular α-oxoaldehyde probe in living cells (**Figure 4.2B**). We drew our inspiration from pioneering work on photocaged lipid derived electrophiles

by the Aye laboratory.[222, 299, 300] We envisioned that rendering a dicarbonyl precursor inert with a photolabile protecting group (photocaging) would be an effective strategy as light is a precise and bio-compatible stimulus, allowing for spatiotemporal control over cargo release. Additionally, we sought to incorporate a HaloTag-targeting element in our design for subcellular localization of our caged electrophilic precursor (akin to **Chapter III** for DA). We hypothesized that this design would allow us to identify proteins that are highly susceptible to modification by dicarbonyl species and may act as sensors of metabolites like MGO.[301] Identification of such proteins would greatly improve our understanding of the cellular processes that regulate dicarbonyl stress. Here, we present the design and synthesis of T-DiP (targetable dicarbonyl precursor) and demonstrate how T-DiP can be utilized to release a dicarbonyl probe with spatiotemporal control within live cells.

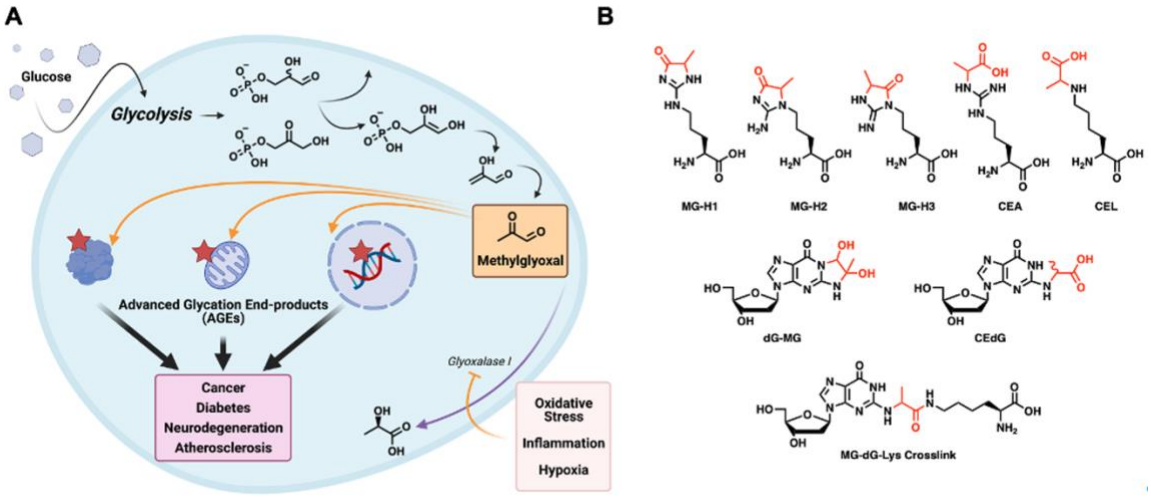


Figure 4.1. Methylglyoxal reacts with biomolecules to form advanced glycation end products

A) Generation of MGO and downstream advanced glycation end products (AGE) formation or detoxification. B) Structures of representative MGO adducts.

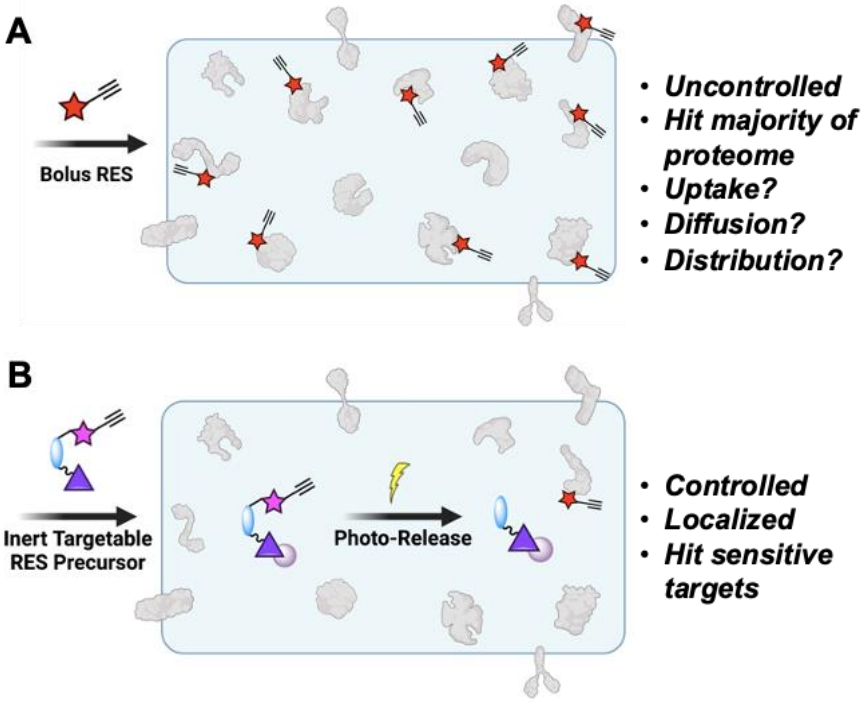


Figure 4.2. Strategies in studying proteins adducted by reactive electrophile species (RES) in cellular systems

A) Bolus dosing of a RES probe, here shown with bio-orthogonal alkyne functionality for target enrichment and identification. Treatment modifies numerous targets in a relatively uncontrolled fashion, complicating the analysis and elucidation of cellular responses. B) Controlled release of RES probe. An inert, targetable RES precursor is initially dosed to cells; following incubation and removal of excess unbound probe, cells are irradiated to release RES payload with spatiotemporal control.

4.2 Materials and Methods

All chemicals and reagents required for the synthesis of targetable dicarbonyl precursor (T-DiP) and preparation of buffers were purchased from Sigma Aldrich unless otherwise stated. NMR solvents were obtained from Cambridge Isotope Laboratories Inc. HEK 293T cells were acquired from ATCC. Cell culture media, additives, and consumables were purchased from Corning. DAPI and anti-actin antibody were obtained from ThermoScientific. AlamarBlueHS reagent and ProlongGlass were purchased from Invitrogen. The BCA kit was bought from Pierce. Mini-protean TGX precast gels, 0.2 µm nitrocellulose and filter paper transfer stacks were obtained from BioRad. THPTA and AlexaFlour647 azide were purchased from Click Chemistry Tools. Odyssey TBS blocking buffer and secondary antibodies were purchased from Licor. TEV protease was purchased from Genscript. Anti-Halo antibody and TMR Halo-ligand were obtained from Promega. The transfection reagent TransIT-2020 was obtained from Mirus Bio LLC. A 0.2-amp Spectroline (Model ENF-280C) UV hand lamp was used as the 365 nm light source in all experiments. Plasmids were obtained from Addgene. Specifically, Halo-Cyto, Halo-Mito, and Halo-ER were gifts from Jin Wang (Addgene plasmid #124314, #124315, #124316) and pet28a-His6-Halo-Tev-Keap1 was a gift from Yimon Aye (Addgene plasmid #62456). UBA1 was purchased from ABCAM. Licor Image Studio software was used to process acquired gel images and statistical analysis was performed with GraphPad.

General Reaction Conditions:

Chemical reactions were conducted with oven-dried glassware under argon atmosphere unless otherwise stated. Anhydrous solvents were dried with activated 4 Å molecular sieves. Thin-layer chromatography (TLC) was performed with Analtech silica uniplates and visualized under 254 nm UV light or with $KMnO_4$ staining. Column chromatography was conducted with 60 mesh silica gel. NMR spectra were acquired on a 400 MHz Bruker spectrometer at room temperature. HRMS spectra were obtained on a Bruker ESI-BioTOF II instrument and internally calibrated with PEG 300. Schemes of reactions are shown in **Figures 4.3-4.7**.

Synthetic Procedures:

2-Allyl-1-[(2-oxohex-5-yn-1-yl)oxy]anthracene-9,10-dione (**1**, caged DiC^{yne}):

To a solution of 2-Allyl-1-hydroxyanthra-9,10-quinone (75 mg, 0.33 mmol, 3.0 equiv), K_2CO_3 (45 mg, 0.33 mmol, 3.0 equiv), and TBAI (20 mg, 0.050 mmol, 0.50 equiv) in THF:DMF (1:1 v/v, 2 mL) was added a solution of **6** (30 mg, 0.11 mmol, 1.0 equiv) in THF (1 mL). The reaction was stirred for 18 h at 25 °C, then diluted with EtOAc (20 mL) and washed successively with 0.1 N aqueous HCl (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. Purification by silica gel flash chromatography (gradient of 0 to 10% EtOAc in hexanes) afforded the title compound **1** (30 mg, 77%) as an orange solid: R_f = 0.5 (9:1 hexanes:EtOAc). ^{1}H NMR (400 MHz, $CDCl_3$) δ 2.00 (t, *J* = 2.4 Hz, 1H), 2.64 (td, *J* = 7.6 Hz, *J* = 2.8 Hz, 2H), 3.11 (t, *J* = 7.6 Hz, 2H), 3.60 (d, *J* = 6.4 Hz, 2H), 4.58 (s, 2H), 5.09 (dd, *J* = 16.8 Hz, *J* = 1.6 Hz, 1H), 5.17 (dd, *J* = 10.2 Hz, *J* = 1.2 Hz, 1H) 5.94–6.04 (m, 1H), 7.66 (d, *J* = 8.0 Hz, 1H), 7.76–7.79 (m, 2H), 8.15 (d, *J* = 8.0 Hz, 1H), 8.22–8.26 (m, 2H). ^{13}C NMR (101

MHz, $CDCl_3$) δ 12.5, 34.3, 38.2, 68.8, 78.0, 83.1, 117.4, 124.2, 125.3, 126.8, 127.3, 132.6, 133.8, 134.2, 134.3, 134.5, 135.5, 136.2, 142.3, 156.3, 182.6, 182.8, 205.4. HRMS (ESI-TOF) *m/z* $[M + Na]^+$ calculated for $C_{23}H_{18}O_4Na^+$ 381.1103, found 381.1117.

2-Oxohex-5-ynal (**2**, DiC^{yne}):

Compound **2** was prepared with slight modification of a previously reported protocol.[297] Briefly, compound **5** (47 mg, 0.41 mmol, 1.00 equiv) was dissolved into 10 mL of anhydrous CH_2Cl_2. Dess-Martin periodinane (181 mg, 0.42 mmol, 1.02 equiv) was added and the reaction mixture was stirred for 18 h at 25 °C. The reaction was then cooled to 0 °C for 10 min and filtered over celite to remove precipitate. Next, 10 mL of aqueous 2.5% $Na_2S_2O_3$ in saturated $NaHCO_3$ was added to the filtrate and the resultant solution was for vigorously stirred for 1 h. The reaction mixture was extracted with CH_2Cl_2 (3 × 20 mL), washed with brine (60 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. The title compound was purified via silica gel flash chromatography (gradient of 0 to 100% CH_2Cl_2 in hexanes) to afford DiC^{yne} **2** (11 mg, 24%) as a pale oil: R_f = 0.8 (3:2 hexanes:EtOAc). Spectral characterization of this compound was in agreement with previous reports.[296, 297]

2-Oxohex-5-yn-1-yl 4-methylbenzenesulfonate (**6**):

Intermediates **4** and **5** were synthesized as previously reported.[302] Compound **5** (170 mg, 1.5 mmol, 1.0 equiv) was dissolved into 3 mL of anhydrous CH_2Cl_2 and cooled to 0 °C. Pyridine (180 μL, 2.3 mmol, 1.5 equiv) and TsCl (430 mg, 2.3 mmol, 1.5 equiv) were added and the reaction was allowed to warm to room temperature and stirred for 1 h. The reaction was diluted with CH_2Cl_2 (20 mL) and washed with 1 N aqueous HCl (20

mL), brine (mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. Purification by silica gel flash chromatography (gradient of 0 to 50% EtOAc in hexanes) afforded the title compound (172 mg, 43%) as a clear oil: R_f = 0.4 (4:1 hexanes:EtOAc). ^{1}H NMR (400 MHz, $CDCl_3$) δ ppm 1.94 (t, *J* = 2.8 Hz, 2H), 2.39–2.52 (m, 4 H, m), 2.78 (t, *J* = 7.6 Hz, 1H), 4.52 (s, 2H), 7.37 (d, *J* = 8.0 Hz, 2H), 7.82 (d, *J* = 8.0 Hz, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 12.3, 21.7, 38.0, 69.2, 71.8, 82.1, 128.1, 130.1, 132.2, 145.6, 201.4. HRMS (ESI-TOF) *m/z* $[M + Na]^+$ calculated for $C_{13}H_{14}O_4SNa^+$ 289.0510, found 289.0523.

N-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethyl)-2-(9,10-dioxo-1-((2-oxohex-5-yn-1-yl)oxy)-9,10-dihydroanthracen-2-yl)acetamide (**8**, T-DiP):

Compound **7** was synthesized according to previous reports.[299] To a solution of **7** (61 mg, 0.125 mmol, 1 equiv), K_2CO_3 (51 mg, 0.375 mmol, 3 equiv), and TBAI (14 mg, 0.038 mmol, 0.3 equiv) in THF:DMF (1:1 v/v, 2 mL) was added a solution of **6** (100 mg, 0.375 mmol, 3 equiv) in THF (1 mL). The reaction was stirred for 18 h at 25 °C, then diluted with EtOAc (20 mL) and washed successively with 0.1 N aqueous HCl (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. The crude material was purified by silica gel flash chromatography (gradient of 50 to 100 % EtOAc in hexanes) to afford T-DiP **8** (37 mg, 51%) as an orange solid: R_f = 0.5 (9:1 hexanes:EtOAc). T-DiP **8** was subjected to an additional RP-HPLC purification step on a Sunfire C18 column (5 μm, 19 × 250 mm) prior to cellular studies (buffer A: H_2O, buffer B: CH_3CN, flow rate; 10 mL/min, gradient: 5% B to 70% B from 0 to 12.5 min, 70% B to 95% B from 12.5 to 27 min, 95% B until 35 min; **8** eluted at ~ 90% B). ^{1}H NMR (400

MHz, $CDCl_3$) δ ppm 1.31–1.39 (m, 3H), 1.39–1.48 (m, 3H), 1.53–1.63 (m, 5H), 1.71–1.81 (m, 3H), 2.00 (t, *J* = 2.4 Hz, 1H), 2.60 (td, *J* = 7.2 Hz, *J* = 2.4 Hz, 2H), 2.90 (t, *J* = 7.2 Hz, 2H), 3.41–3.47 (m, 4H), 3.49–3.63 (m, 10H), 3.74 (s, 2H), 4.79 (s, 2H), 6.73 (m, 1H), 7.74–7.84 (m, 3H), 8.16 (d, *J* = 7.6 Hz, 1H), 8.25 (m, 2H). ^{13}C NMR (101 MHz, $CDCl_3$) δ ppm 12.7, 25.4, 26.7, 29.4, 32.5, 37.7, 38.6, 39.5, 45.0, 69.12, 69.6, 70.0, 70.3, 71.2, 82.7, 124.4, 126.8, 127.3, 133.9, 134.3, 134.4, 134.8, 137.1, 156.7, 169.4, 182.55, 204.6. HRMS (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{32}H_{37}NO_7Cl^+$ 582.2253, found 582.2265.

2-allyl-1-(2-oxopropoxy)anthracene-9,10-dione (**9**, caged MGO):

To a solution of -Allyl-1-hydroxyanthra-9,10-quinone (75 mg, 0.33 mmol, 3.0 equiv), K_2CO_3 (45 mg, 0.33 mmol, 3.0 equiv), and TBAI (20 mg, 0.050 mmol, 0.50 equiv) in THF:DMF (1:1 v/v, 2 mL) was added a solution of 2-oxopropyl tosylate[303] (30 mg, 0.11 mmol, 1.0 equiv) in THF (1 mL). The reaction was stirred for 18 h at 25 °C, then diluted with EtOAc (20 mL) and washed successively with 0.1 N aqueous HCl (2 × 20 mL), brine (20 mL), dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. Purification by silica gel flash chromatography (gradient of 0 to 10% EtOAc in hexanes) afforded the title compound **1** (27 mg, 26%) as an orange solid: R_f = 0.5 (9:1 hexanes:EtOAc). ^1H NMR (500 MHz, $CDCl_3$) δ 8.22 - 8.28 (m, 2 H), 8.14 (d, *J*=7.9 Hz, 1 H), 7.74 - 7.82 (m, 2 H), 7.66 (d, *J*=7.9 Hz, 1 H), 5.99 (ddt, *J*=16.9, 10.3, 6.4 Hz, 1 H), 5.13 - 5.19 (m, 1 H), 5.09 (dq, *J*=17.1, 1.6 Hz, 1 H), 4.57 (s, 2 H), 3.60 (d, *J*=6.4 Hz, 2 H), 2.44 ppm (s, 3 H) ^{13}C NMR (126 MHz, $CDCl_3$) δ 205.3, 182.8, 182.6, 156.4, 142.3,

136.1, 135.6, 134.6, 134.3, 134.2, 133.7, 132.6, 127.3, 126.7, 125.4, 124.2, 117.3, 78.2, 34.3, 26.7 LRMS (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{20}H_{17}O_4^+$ 321.1, found 321.1.

Photolysis of Caged DiC^{yne} Monitored by HPLC:

Caged DiC^{yne} 1 samples designated for photolysis were prepared by adding 50 µL of a solution containing 150 µM 1 in PBS (pH 7.4, 0.15% DMSO) to a 0.6 mL PCR tubes. The samples were irradiated at 365 nm via a Spectroline UV hand lamp positioned 1 inch above the midpoint of the tube. At 30 s timepoints, 20 µL of the reaction were removed and analyzed by HPLC using a Luna C18 column (5 µm, 150 × 2.0 mm). The analysis method (1.0 mL/min flow rate, buffer A: H_2O, buffer B: CH_3CN) involved a linear gradient of 5-95% B (0-25 min), followed by an isocratic hold at 95% B (25-30 min). Absorbance was monitored at 254 nm and peak areas corresponding to caged DiC^{yne} were integrated and normalized to the peak area at time zero to determine the half-life of 1 under 365 nm light exposure over three independent replicates.

Light Dependent DiC^{yne}-Peptide Adduct Formation:

A solution of the peptide SGFRY or Ac-LESRHYAG (100 µM) in PBS (pH 7.4) was incubated with either DMSO, 150 µM DiC^{yne}, MGO, 150 µM caged DiC^{yne}, or caged MGO in a total volume of 30 µL containing 1.6% DMSO. Samples designated for light exposure were irradiated with 365 nm light for 5 min. All samples were subsequently incubated for 18 h in the dark at 37 °C. The solutions were then prepared for MALDI-MS by desalting with Millipore C18 ziptips following the manufacturer's instructions. Desalted samples were either spotted onto a MALDI plate with a saturated solution of α-cyano-4-hydroxycinnamic acid as the matrix or prepared for LC-MS/MS. MALDI

spectra were acquired on an AB-Sciex 5800 MALDI/TOF-MS to monitor DiC^{yne}-peptide adduct formation over two independent replicates. LC-MS/MS analysis was performed on a Eksigent NanoLC-Ultra 2D HPLC system coupled to a nanospray ESI source and the Thermo Scientific Q Exactive Mass Spectrometer. Peptides were resuspended in 16 uL 0.1% formic acid. 4 µL sample were separated using an in-house packed C-18 column (15 cm x 75 um, Luna C18) across a 60 min gradient. Mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1% formic acid in acetonitrile. Source parameters were set as 3.0 kV spray voltage, capillary temperature of 300C and S-lens RF level of 50%. MS analysis was performed using a top 12 data-dependent mode. The full scan was run using the resolution setting of 70,000 and an AGC target of 1 x 10^6. The exclusion duration was set for 30 s for dynamic exclusion. The 12 most intense precursor ions were chosen for HCD fragmentation which was set at 30%. The fragmentation scan was acquired with an isolation window of 2.0 *m/z*, a resolution of 17,500 at an AGC target of 5 x 10^4 or a maximum injection time of 50 ms.

General Protocol for DiC^{yne} Light Mediated Release in Lysates:

HEK 293T cell lysates (1 mg/mL) in MPER lysis buffer with 1.0% DMSO were incubated with the indicated concentration of caged DiC^{yne} for a given experiment in a total reaction volume of 50 µL and irradiated with 365 nm light via a Spectroline UV hand lamp for the specified amount of time. Following further incubation at 37 °C, a click reaction was performed with 10 µM AlexaFlour647 azide, 1.5 mM tris((1-hydroxy-propyl-1H-1,2,3-triazol-4-yl)methyl)amine (THTPA), 1 mM $CuSO_4$ and 3 mM sodium ascorbate. The reactions were agitated with an orbital shaker at 25 °C for 1 h in the dark.

Proteins were then precipitated with 400 µL of cold acetone overnight at -20 °C. Following centrifugation at 21,000 g for 20 min, the acetone was carefully aspirated, and the protein pellets were dissolved into 30 µL of 1× Licor loading buffer (125 mM Tris-HCl, pH 6.8, 50% glycerol, 4% SDS, 0.2% Orange G dye) containing 10% β-mercaptoethanol and denatured for 5 min at 95 °C. Next, 13.5 µL of each sample was loaded into 4–20% gradient mini-protean TGX precast gels for SDS-PAGE. Gels were fixed overnight in an aqueous solution of 30% EtOH and 10% acetic acid to wash out excess fluorophore and imaged with a Licor odyssey CLX instrument to quantify DiC^{yne} derived AGEs. The gels were then stained with coomassie blue to confirm equivalent loading.

Expression of His6 Halo Protein:

Recombinant His6-HaloTag protein was expressed in *E. coli* BL21 cells as a His6-Halo-TEV-Keap1 fusion protein as previously described with slight modification.[224] Briefly, expression of the fusion protein was carried out in LB Broth (Miller, 25 g/L) containing 50 µg/mL of carbenicillin (LB media). Agar plates containing 50 µg/mL of carbenicillin were streaked and allowed to grow overnight. The following day, a single colony was selected and used to inoculate 50 mL of LB media and was grown overnight at 37 °C. Then 5 mL of the overnight culture was diluted in 500 mL of LB media. The flask was placed on a shaker at 37 °C at 225 RPM and allowed to grow for ~4 h (OD = ~0.45 – 0.6). Expression was induced with 250 µM IPTG over 18 h at 19 °C. The resulting suspension was pelleted at 4,500 g for 12 min at 4 °C. The pellet was suspended in lysis buffer (50 mM NaH_2PO_4 (pH 8.0), 10 mM imidazole, 5 mM β-mercaptoethanol,

0.5 mM phenylmethylsulfonyl fluoride, and 0.01% Triton X-100) and lysed by sonication at 40 W in 30 s intervals for T-DiP min (4 min total sonication). The cellular debris was pelleted at 38,000 g for 25 min at 4 °C and the supernatant was collected. The supernatant was purified using TALON affinity chromatography and fractions containing the fusion protein were confirmed through western blot. The desired fractions were subsequently purified via size exclusion chromatography (SEC) by injecting them on a Bio-RAD NGC SEC system equipped with a GE Superdex HiLoad 16/60 column equilibrated with SEC buffer (50 mM Tris (pH 8.0), 10 mM dithiothreitol, 100 mM NaCl and 5% glycerol) at 4 °C. The fractions containing the desired fusion protein were characterized by western blot. Cleavage of the fusion protein was achieved using TEV protease. TEV (10,000 U/mL) was incubated with the fusion protein (1.5 mg/mL) in SEC buffer for 18 h to generate His6-Halo protein. The resulting cleavage product was purified using TALON affinity chromatography and characterized through SDS-PAGE and western blot. The protein concentration was determined by A280 (calculated extinction coefficient of 59930 $M^{-1}cm^{-1}$). The protein was aliquoted and stored in SEC buffer at -80 °C until needed.

Photolysis of T-DiP-Halo Complex in vitro:

For photolysis experiments, 1 µg of His6-Halo protein, (66 µM in PBS, pH 7.4, 3% DMSO) was incubated with 66 µM T-DiP in a total volume of 30 µL per condition. The resultant solutions were incubated in the dark at 37 °C for 30 min to allow the covalent complex to form between the Halo protein and chloroalkane ligand. Samples were then subjected to 365 nm light exposure for 0, 1, 5, or 15 min followed by a click

reaction through the addition of 0.1 mM AlexaFlour647 azide, 1.5 mM THPTA, 1 mM $CuSO_4$ and 3 mM sodium ascorbate. The reaction was allowed to proceed for 1 h and quenched with 7.5 µL 4× gel loading buffer. Samples were boiled at 95 °C for 5 min, then 30 µL of each sample was loaded into 4–20% gradient mini-protean TGX precast gels for SDS-PAGE. Gels were fixed overnight in an aqueous solution of 30% EtOH and 10% acetic acid to wash out excess fluorophore and imaged with a Licor odyssey CLX instrument to quantify the T-DiP-Halo protein complex. The gels were then stained with coomassie blue to confirm equivalent loading.

Alamar Blue Cell Viability Assay:

HEK 293T cells were seeded into a 96 well plate at a density of 20,000 cells per well in 100 µL of FBS medium and incubated for 24 h. The FBS medium was removed and replaced with the indicated treatment condition in FBS-free MEM media. Specifically, following a 3 h T-DiP treatment to mimic control released assay conditions, the medium was refreshed, and the cells were further incubated for 18 h, or cells were subjected to a 20 min light exposure and further incubated for 18 h. After incubation, the medium was removed and replaced with 100 µL of a 10% alamar blue solution in FBS-free MEM medium. The plates were incubated for 4 h in the dark at 37 °C and then fluorescence was measured at Ex/Em 560/590 nm on a BioTek Synergy H1 plate reader. Cell viability for a given treatment condition was calculated by normalizing the fluorescence signal to the DMSO treated control.

Confocal Microscopy:

HEK 293T cells were cultured on Neuvitro GC-25-1.5-Laminin coverslips to 70% confluency with MEM medium containing 10% FBS and Pen/Strep 100 U/mL under a 5% CO_2 atmosphere at 37 °C. Cells were transfected with 2.5 μg Halo plasmids via the serum-compatible Mirus TransIT-2020 transfection reagent following the manufacturer's instructions. After 24 h, the medium was removed and replaced with the indicated treatment condition in FBS-free medium. For competitive binding experiments, cells were treated with DMSO vehicle or 12.5 μM T-DiP for 2 h. Then the cells were washed with serum-free medium and then incubated with 5 μM HaloTag TMR for 1 h. Following the allotted incubation time, the cells were washed three times with 2 mL PBS, fixed with 2 mL of 4% PFA (molecular biology grade) in PBS for 10 min, and washed an additional three times with PBS. For in-cell click imaging experiments, the cells were treated with either DMSO, 500 μM 2 for 4 h, or 12.5 μM T-DiP with a 20-min 365 nm light exposure followed by a 15-min incubation prior to fixation as descried above. Following fixation, the cells were incubated in a 0.1% Triton X-100 PBS (PBT) solution containing 10 μM AlexaFlour647 azide, 1.5 mM THPTA, 1 mM $CuSO_4$, and 50 mM sodium ascorbate for 1 h to label proteins modified by DiC^{yne}. Next, three washes with PBT were performed. All cells' nuclei were stained with 300 nM of DAPI for 30 min and were further washed an additional three times with PBS. A drop of prolong glass anti-photobleaching solution was placed on the coverslips, which were mounted to glass slides and allowed to cure for 48 h. The slides were then sealed with clear nail polish and stored at 4 °C until they were imaged. Images were acquired on an Olympus FluoView FV1000 BX2 upright confocal

microscope using a PLAPON 60X O NA:1.42 objective. Laser Ex/Em were set at 405:461, 557:576, and 635:668 nm for DAPI, TMR, and AlexaFlour647, respectively.

General Protocol for the Controlled Release of DiC^{yne} with T-DiP in Live Cells:

HEK 293T cells were cultured in MEM supplemented with 10% v/v FBS in the presence of 100 U/mL Pen/Strep antibiotics under a 5% CO_2 atmosphere at 37 °C. Cells were seeded into 6-well tissue treated plates at 400,000 cells per well. Cells were transfected with 2.5 μg Cyto-Halo, Mito-Halo, or ER-Halo plasmid via the serum-compatible Mirus TransIT-2020 transfection reagent (7.5 μL) following the manufacturer's instructions. After 24 h following transfection, cells were washed with PBS and subjected to the desired treatment conditions in FBS free medium. Cells treated with T-DiP were incubated with 12.5 μM for 2 h, then the medium was removed and replaced with fresh FBS-free medium and the cells were further incubated for 30 min. Two additional 30-min washes were performed to ensure any unbound T-DiP was removed. Samples were irradiated at 365 nm with a handheld UV lamp for 20 min which was positioned 1 inch above the cell monolayer. Cells were incubated at 37 °C for the specified amount of time prior to lysis and downstream analysis.

Following the allotted incubation time, the cells were collected through trypsinization, washed twice with PBS, and lysed in 200 μL of MPER lysis buffer supplemented with HALT protease inhibitors. The lysates were centrifuged at 14,000 g for 8.5 min to remove debris. The supernatant was collected, and the protein concentrations were determined through a BCA assay per manufacturer's protocol.

Click reactions were performed in a total volume of 100 µL at a protein concentration of 50 µg/mL with 10 µM AlexaFlour647 azide, 1.5 mM THPTA, 1 mM $CuSO_4$ and 3 mM sodium ascorbate. The reactions were agitated at 25 °C for 1 h in the dark. Proteins were precipitated with 400 µL of cold acetone overnight at -20 °C. Following a 20 min centrifugation at 21,000 g, the acetone was carefully aspirated, and the protein pellets were dissolved into 30 µL of 1× Licor loading buffer with 10% β-mercaptoethanol and denatured for 5 min at 95 °C, then 13.5 µL of each sample was loaded into 4–20% gradient mini-protean TGX precast gels for SDS-PAGE. Gels were fixed overnight in an aqueous solution of 30% EtOH and 10% acetic acid to wash out excess fluorophore and imaged with a Licor odyssey CLX instrument to visualize and quantify DiC^{yne} derived protein adducts. The gels were then stained with coomassie blue to confirm equivalent loading.

Proteomic Identification of DiC^{yne} *Modified Proteins Under Controlled Release Conditions:*

HEK 293T cells were cultured in MEM media supplemented with 10% v/v FBS in the presence of 100 U/mL Pen/Strep antibiotics under a 5% CO_2 atmosphere at 37 °C. Cells were seeded into 10 cm dishes tissue treated plates with 2,000,000. Cells were transfected with Cyto-Halo plasmid via the serum-compatible Mirus TransIT-2020 transfection reagent following the manufacturer's instructions. After 24 h following transfection, cells were washed with PBS and subjected to the desired treatment conditions in FBS free medium. Cells treated with T-DiP were incubated with 12.5 µM T-DiP for 2 h, then the medium was removed and replaced with fresh FBS-free medium

and was further incubated for 30 min. Two additional 30-min washes were performed to ensure unbound T-DiP was removed. Samples were irradiated at 365 nm with a handheld UV lamp for 20 min which was positioned 1 inch above the cell monolayer. Cells were incubated at 37 °C for the 30 mins prior to lysis and downstream analysis.

Following the allotted incubation time, the cells were collected through trypsinization, washed twice with PBS, and lysed in 500 µL of MPER lysis buffer supplemented with HALT protease inhibitors. The lysates were centrifuged at 14,000 g for 8.5 min to remove debris. The supernatant was collected, and the protein concentrations were determined through a BCA assay per manufacturer's protocol.

On-Bead Click Reaction: Click Chemistry Tools azide agarose beads (100 µL of Cat: 31038-2; Lot: 2333) were washed twice with 1 mL of water. The click reaction was conducted by adding a 1 mL solution of 750 µg of protein lysate, 1.5 mM THPTA, 1 mM $CuSO_4$, 3 mM Na ascorbate in MPER to the beads. The reaction was allowed to proceed for 16 h before the beads were pelleted by centrifugation and the supernatant was aspirated.

Washing of Agarose Beads: The beads were washed once with water and the proteins bound to the beads were reduced by addition of 100 mM dithiothreitol (DTT) in 1 mL of 1% SDS, 5 mM ethylenediaminetetraacetic acid (EDTA), 250 mM NaCl, 100 mM Tris pH 8.0 and incubation at 70 °C for 15 min. After pelleting the beads and removing the supernatant, alkylation was performed with 40 mM iodoacetamide in the dark for 30 min in the same buffer as the previous step. Following centrifugation and removal of the supernatant, the beads were washed five times with 1 mL of 1% SDS, 5

mM EDTA, and 100 mM NaCl in PBS; then, 10 times with 8 M urea in 100 mM Tris pH 8.0; and finally, 10 times with 20% CH_3CN.

On Bead Digestion: Digestion was performed by suspending the beads in 200 μL of 100 mM Tris, 2 mM $CaCl_2$ pH 8.0 and the addition 1 μg of trypsin. The reaction was rotated at 37 °C overnight. The supernatant was transferred to a new tube and combined with two additional 300 μL water 0.1% formic acid washes. Samples were desalted with Pierce C-18 spin columns (Cat: 69725; Lot: TJ276015) per the manufacturer's instruction with MS grade water, CH_3CN, and formic acid. The digests were concentrated by vacuum evaporation and stored at −20 °C until HPLC-MS/MS analysis.

HPLC-MS/MS Analysis: Tryptic peptides from each biological replicate were analyzed by HPLC-ESI -MS/ MS using a Thermo Scientific Q Exactive Hybrid Quadrupole- Orbitrap Mass Spectrometer in line with an Eksigent NanoLC-Ultra 2D HPLC system, a nanospray source, and Xcalibur 4.1.31.9 software for instrument control. Dried peptide samples were reconstituted in 20 μL of 2% CH_3CN; 0.1% formic acid Samples (4 μL) were injected onto an in-house packed C18 column (15 cm × 75 μm, Luna C18) eluted over 90 min gradient with buffer A (0.1% formic acid) and buffer B (CH_3CN; 0.1% formic acid). Typical ion source parameters included spray voltage of 4.0 kV, a capillary temperature of 320 °C, and an S-lens RF level of 55%. MS/MS analysis was performed in data-dependent mode using top 12 most abundant ions. The survey scan was performed using a resolution setting of 70,000 at 200 m/z with an AGC target of 1×10^6. Dynamic exclusion was enabled with an exclusion duration for 30 s. 12 most intense precursor ions (excluding singly charged species) were selected for higher-energy

C-trap dissociation fragmentation using a collision energy of 30%. The fragmentation scan was performed using a resolution of 17,500 with a AGC target of 5×10^4 or maximum injection time of 50 ms and an isolation window of 2.0 *m/z*.

Data Analysis. MaxQuant (version 1.6.5.0) was used to search the raw mass spectrometry data. Default parameters were applied and the human Uniprot database (downloaded on 2019/05/15, with 74,349 sequences) was searched against.[190] Trypsin was selected as the protease with allowance for the maximum of two missing cleavages. Variable modifications of methionine oxidation, protein N-terminal acetylation and the fixed modification of cysteine carbamidomethylation were included. The precursor ion and fragmentation mass tolerance of 4.5 ppm and 0.5 Da were selected, respectively. The false discovery rate (FDR) was set to 0.01 at the protein, peptide, and site levels with a minimum peptide length of 6 and a minimum Andromeda score of 40. A minimum of two peptides were required for protein identification. To obtain label-free protein quantification, default LFQ parameters were selected with normalized protein intensities as the output. The Perseus software suite (version 1.6.6.0) was used to process the resultant LFQ intensities.[235] The data were Log2 transformed and filtered following previously described methods.[236] A two-tailed, two-sample t-test was performed to compare protein abundance between groups. Statistically significant enrichment was determined with a Benjamini-Hochberg corrected FDR of 0.05 and a minimal coefficient of variation (S0) of 2.0.

MS Analysis of UBA1 MGO Adducts:

Solutions of UBA1 (6.7 μM, PBS, pH 7.4) were incubated with MGO (10 mM) in 25 μL of PBS (pH 7.4) at 37 °C. After the allotted reaction time, the reactions were quenched upon addition of 6.25 μL of 4x NuPAGE gel loading buffer and 25 μL of reaction mixtures were subjected to SDS-PAGE on NuPAGE 4-12% bis tris 1 mm gels. Protien bands were excised from the gel. These bands were reduced with TCEP and alkylated with iodoacetamide following the instructions of the Thermo In-Gel Tryptic Digestion Kit (89871). The gel pieces were dehydrated with acetonitrile for 15 min. The solvent was removed and the gel pieces were air dried and then subjected to tryptic digestion (1 μg), followed by StageTip desalting, and the resultant peptides were stored at -20 °C prior to MS analysis. Prior to MS analysis, samples were resuspended in 12 μL 0.1% formic acid and 4 μL sample injections were characterized by LC-MS/MS using the same liquid chromatography and mass spectrometry settings as described above. The raw MS data was searched with proteome discoverer 2.2 against the UBA1 FASTA with the additional variable modifications: CEL (+72 Da), CEA (+72 Da), and MG-H1 (+54 Da).

General Notes:

For all experiments involving photocaged compounds, care was taken to shield samples from stray light to prevent undesired photolysis. Experiments were performed in a dimly lit environment and samples were wrapped in foil whenever possible.

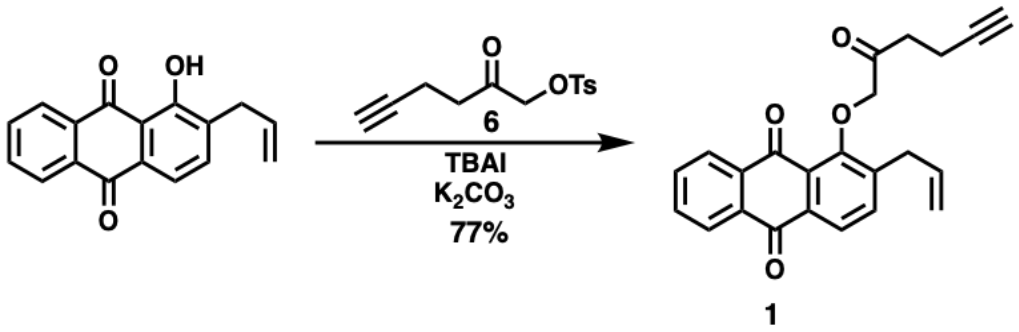


Figure 4.3. Synthesis of caged DiC^{yne}

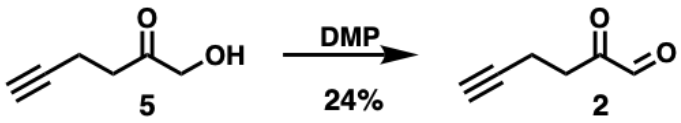


Figure 4.4. Synthesis of DiC^{yne}

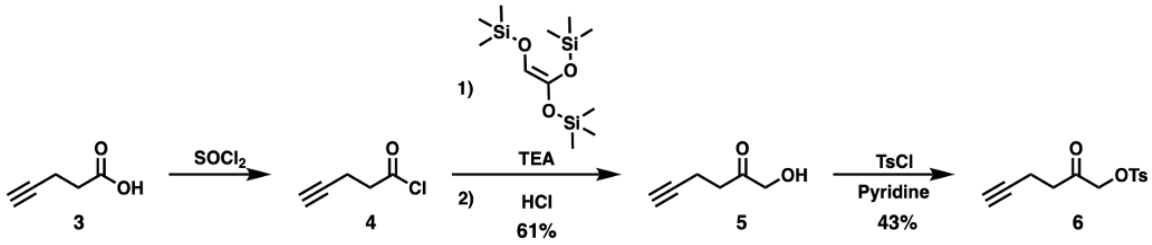


Figure 4.5. Synthesis of T-DiP intermediates

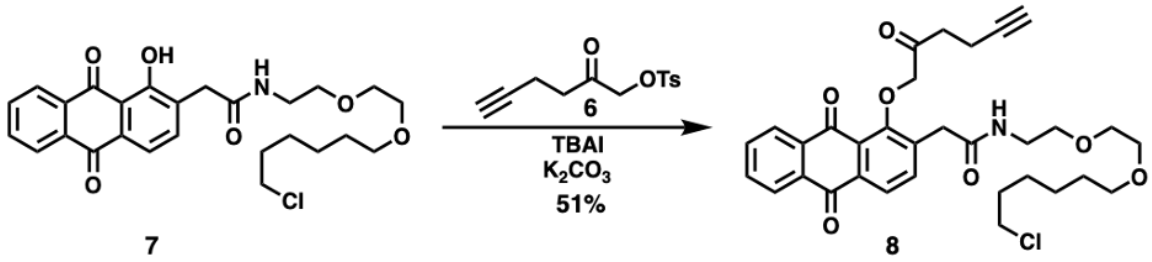


Figure 4.6. Synthesis of T-DiP

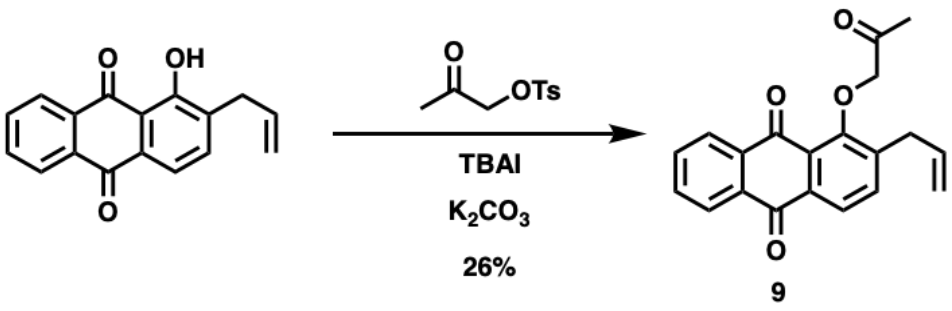


Figure 4.7. Synthesis of caged MGO

4.3 Results

Our first objective was to design a photocage for α-oxoaldehydes. We selected the hydroxy anthraquinone scaffold due to its ability to release aldehydes under 365 nm irradiation.[304] We synthesized caged MGO and caged DiC^{yne} (dicarbonyl alkyne) to release the bioorthogonally functionalized 2-oxohex-5-ynal, DiC^{yne} upon light exposure as shown in **Figure. 4.8A**. The alkyne handle featured in DiC^{yne} facilitates visualization and identification of resultant protein adducts through the attachment of an azide linked tag via Cu catalyzed azide–alkyne cycloaddition (CuAAC) chemistry. DiC^{yne} has been previously reported to be a suitable MGO mimetic.[296, 297] However, it must be noted DiC^{yne} is significantly larger and more hydrophobic than MGO, which may result in differential biological activity.

Initially, caged DiC^{yne} was subjected to photolysis experiments. Using HPLC-UV, we found caged DiC^{yne} to have a half- life of 8 ± 2 s when irradiated with 365 nm light in phosphate- buffered saline (PBS) pH 7.4 (**Figure 4.8B**). We also photolyzed caged DiC^{yne} in the presence of an arbitrary arginine containing peptide (SGFRY), which is susceptible to modification by the released DiC^{yne}. Through MALDI-MS, we verified the formation of the dicarbonyl peptide adduct by observing a 92 Da mass increase in irradiated samples, indicative of hydroimidazolone formation at the arginine side chain (**Figure 4.8C**). These results confirm successful light-induced release of DiC^{yne} by our photocage. We also performed complementary experiments with caged MGO to release MGO in the presence of the peptide Ac-LESRHYAG. Here we monitored adduct formation via LC-MS/MS. As shown in **Figure 4.9**, MGO modification was dependent

on light exposure and the site of adduct formation was observed on the Arg residue, further indicting release of MGO via the anthraquinone scaffold.

Next, we investigated if caged DiC^{yne} allowed for on demand protein glycation in HEK 293T lysates. Cell lysates were treated with compounds DiC^{yne} or caged DiC^{yne}, irradiated for 5 min with 365 nm light, and further incubated for an hour at 37 °C to allow for protein adducts to form. We then utilized the alkyne functionality in DiC^{yne} to visualize the resultant modified proteins by appending a fluorophore to the alkynylated proteins with CuAAC chemistry. This allowed the adducts to be visualized with a fluorescent gel imager following SDS-PAGE separation. Caged DiC^{yne} successfully glycated proteins following light exposure with a similar banding pattern to DiC^{yne} alone (**Figure 4.8D**).

We also monitored the stability of our dicarbonyl photocage in lysates. Samples treated with caged DiC^{yne} were either kept in the dark or exposed to 365 nm light for 5 min prior to 18 h incubation at 37 °C (**Figure 4.10).** The non-irradiated samples generated 95 ± 4% less fluorescent signal attributed to DiC^{yne} - protein adducts than their illuminated counterparts. This indicates negligible degradation of the caged DiC^{yne} precursor and minimal inherit dark reactivity. Collectively, these data show that the anthraquinone scaffold is an effective α-oxoaldehyde photocage under physiological conditions.

Our subsequent goal was to incorporate a targeting element into our dicarbonyl photocage scaffold, so that DiC^{yne} release could be directed intracellularly. We selected the HaloTag technology[265] comprised of a Halo protein and chloroalkane targeting

ligand, which react to generate a 1 : 1 ligand to protein covalent complex. This system provides synthetic tractability, rapid target engagement, and documented success in similar applications.[224] The chloroalkane functionalized T-DiP was prepared in five steps (**Figure 4.11A**).

We verified that T-DiP covalently engaged recombinant Halo protein and released DiC^{yne} upon irradiation with 365 nm light (**Figure 4.11B**). This was accomplished by first incubating recombinant Halo protein and T-DiP in equimolar amounts together to form a 1:1 covalent complex. This complex was irradiated for various lengths of time, 0–15 min, to liberate DiC^{yne}. Next a CuAAC reaction with Alexa647 azide was performed, followed by SDS PAGE separation to visualize the remaining Halo : T-DiP complex. Fluorescence intensity corresponding to the complex was quantified and normalized to the amount of protein loaded to calculate the half-life of the complex under 365 nm light exposure, which was found to be 0.9 ± 0.3 min (**Figure 4.11C**), indicating that DiC^{yne} was rapidly released from the Halo bound anthraquinone photocage.

Upon transitioning to cellular experiments, we ensured that our inert precursor molecule was nontoxic. We found that T-DiP had no significant effect on HEK 293T cell viability as measured by an alamar blue assay at concentrations as high as 75 μM (**Figure 4.12A**). Thus, T-DiP is well tolerated at low doses.

We next generated HEK 293T cells which transiently express Halo protein containing targeting sequences towards the cytosol, mitochondria, and endoplasmic reticulum (ER) to traffic Halo protein to these areas of the cell.[305] We chose these locations for targeted release of T-DiP because glycolysis, and thus MGO generation,

primarily occurs within the cytosol.[306] It is also known that mitochondrial dysfunction, along with ER stress, is associated with glycated proteins.[307] We estimated the concentration of Halo protein expressed within each cell to be ~12, 4, and 8 μM, for the respective cytosol, mitochondria, and ER targeted Halo variants based off a standard curve western blot generated with recombinant Halo protein (**Figure 4.13**). These concentrations are consistent with previous studies.[224]

To assess cell penetrability and target engagement of T-DiP with Halo protein in live cells, we performed a competitive binding experiment with T-DiP and a tetramethyl rhodamine (TMR) HaloTag ligand. The TMR HaloTag ligand contains a chloroalkane moiety which covalently binds Halo protein and can be visualized through confocal microscopy. However, this interaction can be blocked by another chloroalkane containing compound, in our case T-DiP, resulting in the loss of TMR signal. Cells transfected with Halo protein and treated with 5 μM TMR HaloTag ligand alone displayed a high degree of fluorescence in their cytosol, mitochondria, or ER, respectively, whereas pre-treatment with 12.5 μM T-DiP prior to TMR HaloTag ligand administration ablated fluorescent signal as shown in **Figure 4.12B**. These results demonstrate that compound T-DiP penetrates cell membranes and engages proteins at different subcellular locations.

In order to optimize intracellular DiC^{yne} release with our T-DiP platform in live cells, we implemented a gel-based assay to visualize proteins adducted by DiC^{yne}. Halo transfected HEK 293T cells were treated with 12.5 μM T-DiP for two hours and then subjected to three washes, once every 30 min, to remove any unbound T-DiP. The cells were then irradiated with 365 nm light for varying amounts of time, 0–20 min, to release

DiC^{yne}. After light exposure, the cells were collected, washed, lysed, and subjected to a CuAAC reaction, which appended the Alexa647 fluorophore to any DiC^{yne} - protein adducts. The lysates were resolved through SDS-PAGE, allowing proteins modified by DiC^{yne} and remaining Halo–T-DiP complex to be visualized via fluorescent gel imaging. We found the half-life of the Halo : T-DiP complex to be 1.7 ± 0.9 min upon 365 nm light exposure in live cells (**Figure 4.14**) We postulate that interference from the cell medium may account for this slower rate of uncaging as compared to *in vitro* experiments. The photolysis of the complex plateaued at 83 ± 9% completion after 20 min irradiation. This completion percentage was consistent across cytosolic, mitochondrial and ER targeted Halo proteins.

We ensured that 20 min of light exposure did not significantly affect cell viability as measured by an alamar blue assay (**Figure 4.15**). Additionally, this cellular light dose falls within reported tolerable ranges.[308] Based on these results; we selected a 20 min irradiation time for our subsequent experiments.

We sought to estimate the relative amount of intracellular DiC^{yne} released by T-DiP following precedent set from similar calculations with caged lipid-derived electrophiles.[224] Taking the concentration of cytosolic Halo protein per cell to be approximately 12 μM, a photolysis efficiency of 83%, and assuming complete saturation of Halo protein by T-DiP, we estimate that ~10 μM DiC^{yne} is liberated in the cytosol under our controlled release conditions. Analogously, we surmise ~3 and ~7 μM are locally released to the mitochondria and ER, respectively. These are physiologically

relevant concentrations of dicarbonyl probe as intracellular levels of the metabolite MGO have been reported in the range of 1–10 μM.[298, 309]

Finally, we utilized our T-DiP platform to confer spatiotemporal control over protein glycation in live cells. Through our gel-based fluorescence assay, we observed that targeting T-DiP to the cytosol, mitochondria, or ER, followed by 20 min light exposure to release DiC^{yne} produced unique sets of modified proteins as indicated by the differential banding patterns seen in **Figure 4.16A**. Under dark conditions, T-DiP displayed minimal off-target reactivity shown by the selective binding to Halo protein; the additional lower molecular weight fluorescent bands below the principal Halo band correspond to truncated Halo protein found in western blots (**Figure 4.17**). We further verified the selectivity of T-DiP for Halo protein by treating non-transfected cells with our inert T-DiP probe. The fluorescent signal attributed to DiC^{yne} modified proteins was 93 ± 8% lower in non-transfected cells in comparison to those expressing cytosolic Halo protein (**Figure 4.16B**).

Additionally, we mapped the cellular localization of proteins modified under controlled DiC^{yne} release by an in-cell CuAAC reaction. As shown in **Figure 4.16C**, cytosolic, mitochondrial, and ER targeted DiC^{yne} liberation generated distinctly localized areas of protein adducts within the cell. These results highlight the utility of targeted electrophile release for studies of reactive metabolites and suggest that there may be unique proteins which are particularly susceptible to dicarbonyl modification housed in different subcellular locations. These results correlate with the subcellular differences seen in glycated proteins via our gel-based assay.

Next, we performed proteomic analysis of proteins modified via DiC^{yne} under our controlled release conditions. HEK 293T cells were transfected with cytosolic halo protein and subjected to the controlled release protocol with T-DiP as described above. Three experimental conditions were used, a DMSO control group, cells treated with T-DiP and exposed to 365 nm irradiation (light), and cells treated with T-DiP with no light irradiation (dark). Following cell lysis, DiC^{yne} adducted proteins were captured onto azide functionalized beads via CuAAC chemistry akin to the procedure detailed in **Chapter I** of this thesis. After covalent immobilization, the beads were washed to remove nonspecific proteins and subjected to tryptic digestion. The resultant peptides were desalted and analyzed through LC-MS/MS paired with spectral data search to identify proteins within a given sample.

Label free analysis of enriched proteins through the Maxquant software revealed light dependent enrichment of DiC^{yne} modified proteins through our controlled release platform. As shown in **Figure 4.17A-C**, statistically significant enriched proteins were only observed in light treated samples. Full list of enriched proteins are provided in **Table 4.1** and **Table 4.2**. Additionally, the identity of the enriched proteins were consistent between the two control groups with 19 proteins overlapping (**Figure 4.17D**), indicating reliable enrichment.

Subsequent bioinformatic analysis of the light enriched proteins under controlled release conditions was performed through the DAVID biological process over enrichment test (**Figure 4.17E**). Processes such as ubiquitin protein ligase binding, ubiquitin-like protein ligase binding, RNA binding, and processes relating to protein folding were the

most significantly enriched in our analysis. These results indicate that dicarbonyl metabolites could modify and disrupt the activity of proteins involved in these processes (eg UBA1, HSP60, HSP70, and EIF4A1). Thus, interference with ubiquitin protein ligase binding and protein folding processes be a mechanism of MGO induced stress. However, subsequent experiments are needed to confirm this.

Next, we sought to validate and identify the specific sites of MGO modification for the protein UBA1. This protein catalyzes the first step in ubiquitin conjugation to mark cellular proteins for degradation.[310] Disruption of this process by MGO modification may impact ubiquitin dependent process which we found in our bioinformatic analysis. We incubated UBA1 with an excess of MGO (10 mM) for 2 or 4 hours and subjected the reaction mixtures to SDS-PAGE separation to remove excess MGO. As shown in **Figure 4.18A**, prolonged MGO treatment decreased the abundance of the UBA1 protein band at the expected MW (117 Da) and produced high molecular weight aggregates. These high molecular weight bands were excessed from the gel and subjected to tryptic digestion and LC-MS/MS analysis to locate the specific sites of MGO modification. We obtained a 48.2% sequence coverage and observed MGO modifications in primarily solvent accessible regions of UBA1, namely on Lys89, Lys304, Arg384, Lys385, Lys443, Lys465, Lys470, Lys593, Lys671, Agr678, and Lys830 (**Figure 4.18B**) which could be the targets for MGO modification in a cellular context.[311] Additionally, MS/MS spectra of modified peptides were inspected to ensure high quality MS/MS data. A representative MS/MS spectrum the MGO modification at Agr678 for is shown in

Figure 4.18C which provide spectral evidence that this sidechain in modified by MGO. Collectively, these results suggest that UBA1 is a target for MGO modification.

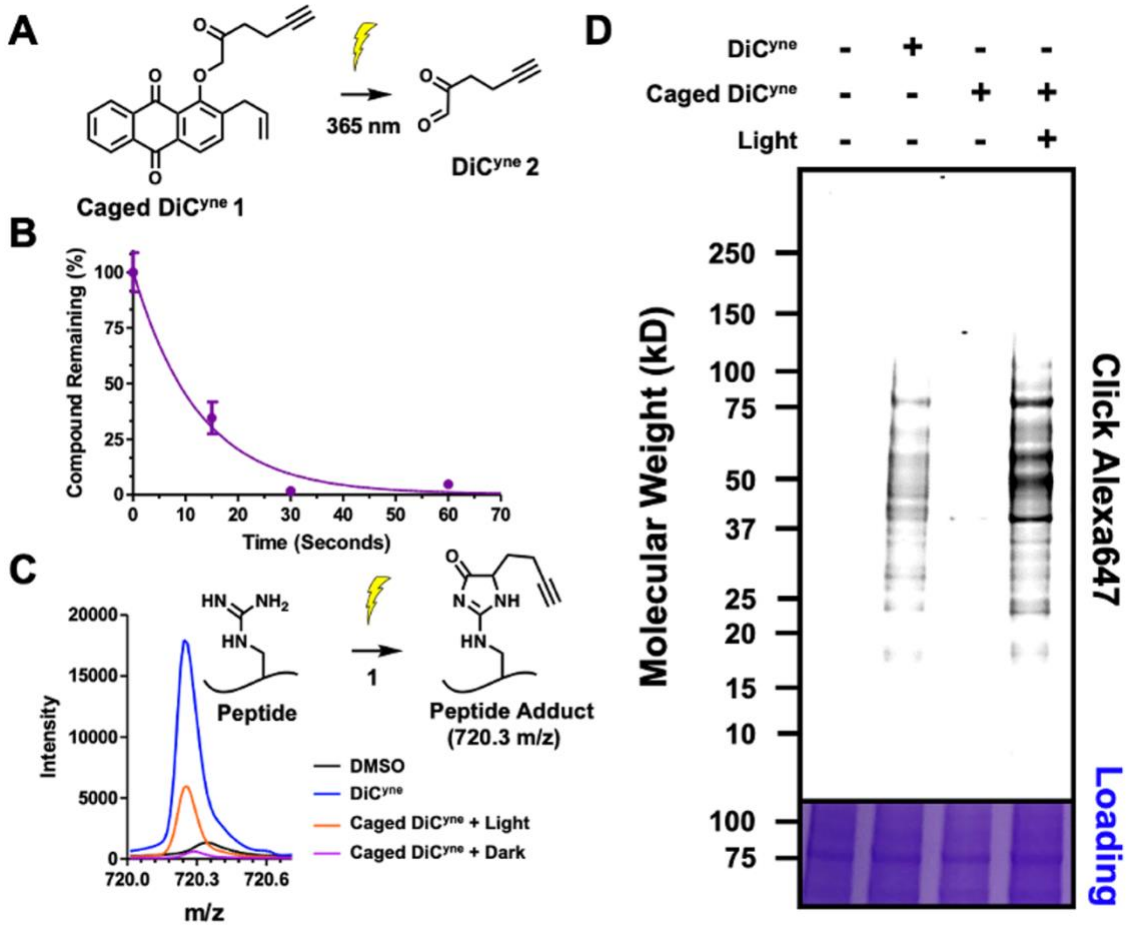


Figure 4.8. Development of an effective dicarbonyl photocage

A) Schematic of caged DiC^{yne} 1 photolysis to release DiC^{yne} 2. B) Kinetics of caged DiC^{yne} 1 photolysis monitored through HPLC-UV. Aliquots of 1 (150 μM) were irradiated with 365 nm light and subjected to HPLC analysis. Percent compound remaining was calculated by integrating the peak area corresponding to 1 and normalizing to time zero. Data are represented as mean ± SEM of 3 independent replicates and fit to a one phase exponential decay function, $t_{1/2}$:8 ± 2 s. (C) Light induced peptide glycation. Overlayed MALDI spectra of peptide (SGFRY) only, peptide

incubated with 2, or peptide incubated with 1 with or without 5 min of 365 nm irradiation. D) Light controlled protein glycation in lysate. Representative gel of HEK 293T lysates dosed with 10 μM of 2 or 1 and irradiated for 5 min then further incubated for 1 h at 37 °C prior to CuAAC reaction with Alexa647 azide.

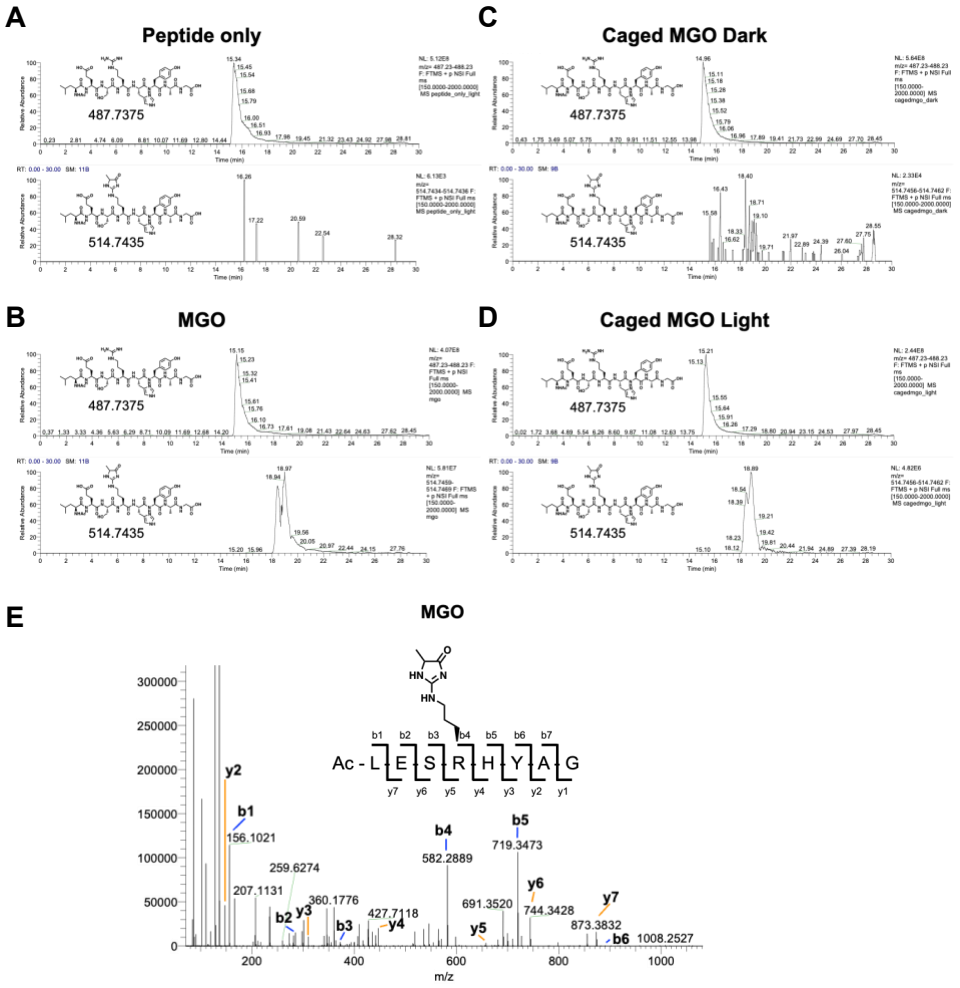


Figure 4.9. Caged MGO releases MGO and generates MGO peptide adducts upon light exposure

A,B,C and D) Extracted ion chromatogram of for Ac-LESRHYAG (top) and MGO adducted Ac-LESRHYAG (bottom) masses for peptide only, treatment with MGO, treatment with caged MGO with no exposure to 365 nm light, and treatment with treatment with caged MGO with exposure to 365 nm light. E) Respective MS/MS spectra of MGO adducted Ac-LESRHYAG.

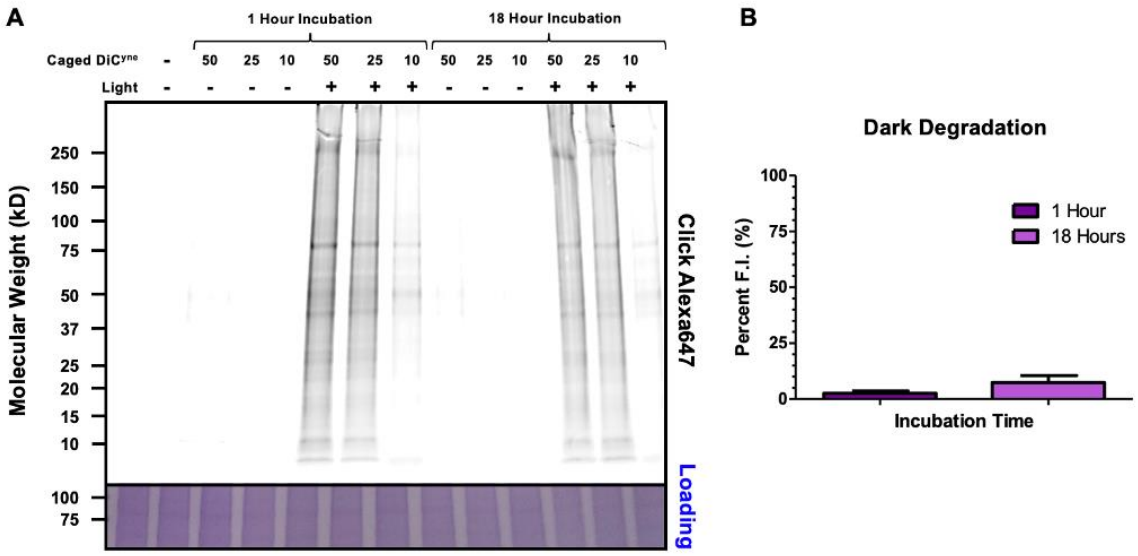


Figure 4.10. Stability of DiC^{yne} photocage

A) Representative agarose gel of HEK293T lysates (1 mg/mL) incubated with 50 μM, 25 μM, or 10 μM caged DiC^{yne} **1,** which were kept in the dark or irradiated for 10 min with 365 nm light and further incubated for the allotted time at 37 °C prior to CuAAC reaction with Alexa647 azide and separation by SDS-PAGE. Top: Fluorescence imaging of Alexa647. Bottom: Coomassie stain loading control. B) Integrated fluorescent intensity (F.I.) of Alexa647 signal in samples kept in the dark normalized to their irradiated counterparts. Data are represented as mean ± SEM from 3 technical replicates.

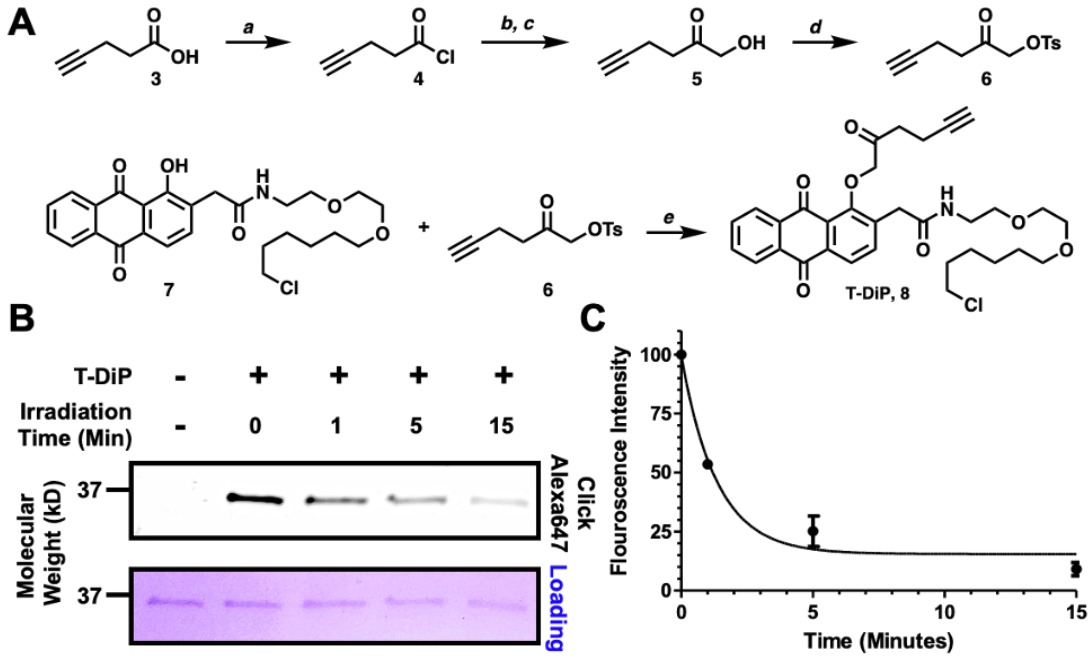


Figure 4.11. T-DiP engages halo protein and releases DiC^{yne} following exposure to light A) Synthesis of T-DiP 8: (a) $SOCl_2$; (b) tris(trimethylsiloxy)ethylene and TEA; (c) HCl 61% (over three steps); (d) TsCl and pyridine 43%, (e) TBAI and K_2CO_3 60%. B) Representative gel of Halo protein incubated with equimolar T-DiP to from a 1:1 complex followed by 365 nm irradiation and subsequent CuAAC reaction with Alexa647 azide. C) Normalized fluorescence intensity of Halo : T-DiP complex following light exposure. Data are represented as mean ± SEM of 3 independent replicates and fit to a one phase exponential decay function, $t_{1/2}$:0.9 ± 0.3 min.

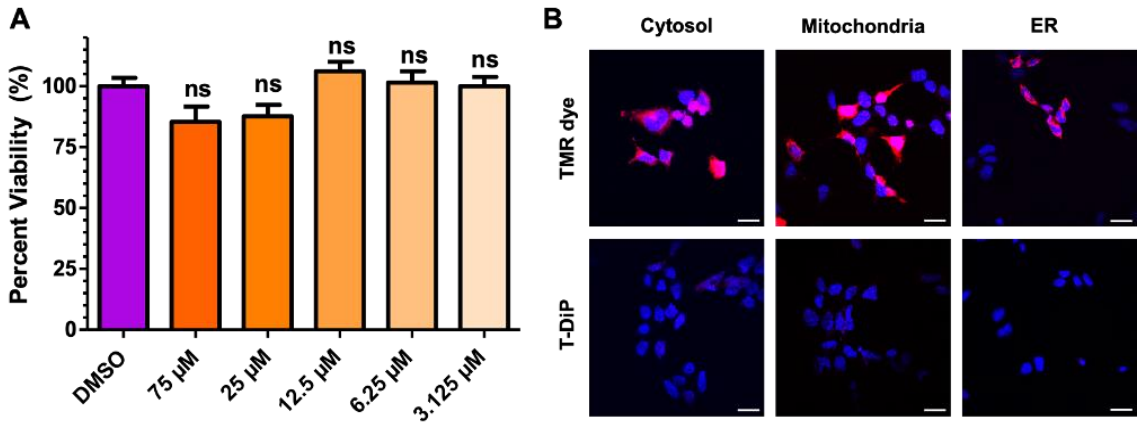


Figure 4.12. T-DiP is nontoxic and targetable to different subcellular areas

A) Cell viability of HEK 293T cells as measured by alamar blue following 18 h incubation with T-DiP. Data are represented as mean ± SEM from 4 technical replicates and normalized to DMSO control. Statistical analysis was performed by a one-way ANOVA (top bar) and Tukey post hoc test (ns: $P > 0.05$). B) Confocal imaging of HEK 293T cells transfected with Halo proteins targeted to the cytosol, mitochondria, and ER treated with either 5 μM TMR HaloTag only (top), or with 12.5 μM T-DiP for 1 h prior to treatment with 5 μM TMR HaloTag (bottom). DAPI is shown in blue and TMR is shown in red. Scale bar in bottom left denotes 20 mm.

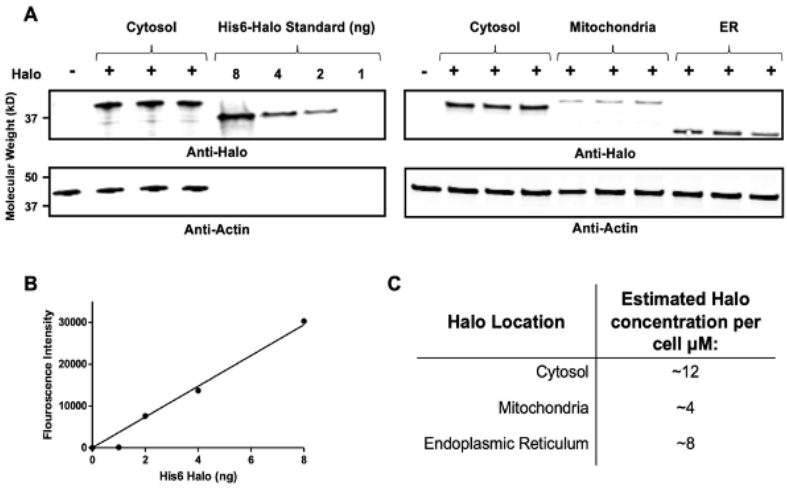


Figure 4.13. Estimation of intracellular halo protein concentration

A) Representative western blots of HEK 293T cells 24 h post transfection with targeted Halo plasmid alongside a standard curve of titrated His6-Halo protein. Top: Anti-Halo. Bottom: Anti-Actin B) Standard curve of recombinant His6-Halo protein fluorescence intensity as measured through western blot. Data are represented as mean ± SEM from 2 technical replicates. C) Estimated cellular Halo protein concentration. Halo protein concentration was estimated based off the following calculations and assumptions:[224, 312] Following cell lysis, 1% of the sample was loaded into the gel and the cytosolic Halo protein concentration in each lane was calculated to be 6 ± 2 ng by the standard curve of recombinant protein. The average number of cells collected per well across 6 biological replicates was $1 \times 10^6 \pm 2 \times 10^4$. Thus, ~$6 \times 10^{-5}$ ng (1.7×10^{-8} nmol) of Halo protein is expressed in per cell. Using data from Bionumbers (http://bionumbers.hms.harvard.edu/bionumber.aspx?id=108893) and the assumption the cells are spherical, the radius of an HEK cell is calculated to be 6.5 µm. The approximate

volume of each cell (¾ π r^3) is estimated to be 1.44 x 10^{-12} L. Therefore, the average concentration of cytosolic Halo protein expressed in each cell is predicted to be ~12 μM. Similar calculations were performed for mitochondria and ER targeted Halo protein.

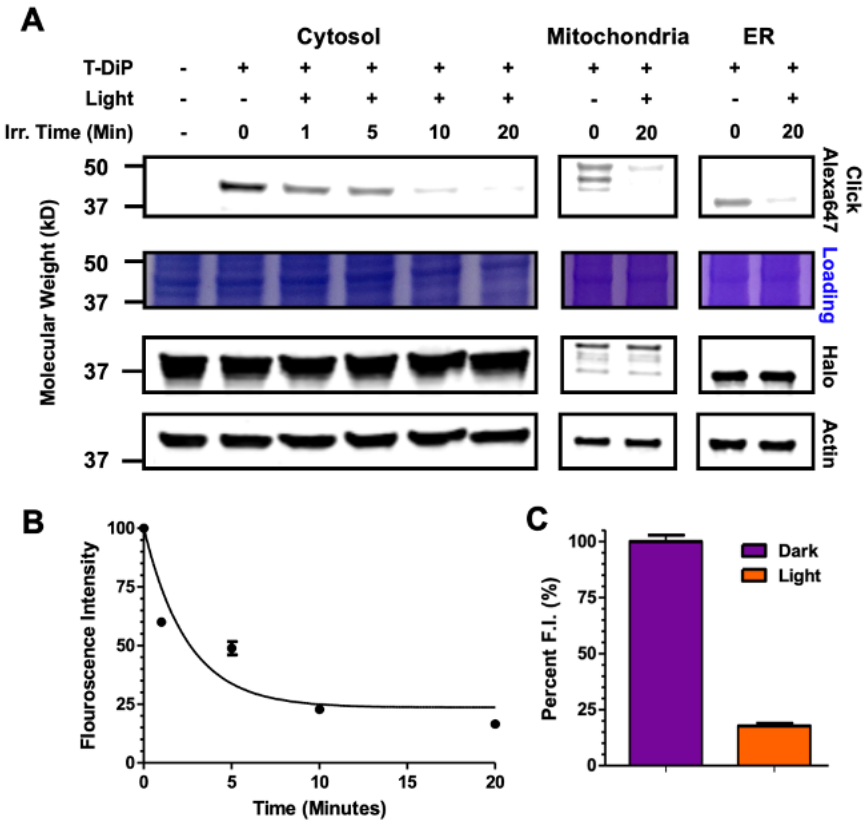


Figure 4.14. T-DiP controlled release system rapidly uncages DiC^{yne} in live cells

Representative agarose gels and western blots of Halo transfected HEK 293T cells treated with 12.5 μM T-DiP and irradiated with 365 nm light for the specified time. Lysates were subjected to a subsequent CuAAC reaction with Alexa647 azide and separated by SDS-PAGE. Top: Fluorescence imaging of Alexa647. Second from top: Coomassie stain loading control. Second from bottom: anti-Halo. Bottom: anti-Actin. B) Normalized fluorescence intensity of Halo:T-DiP complex following light exposure. Data are represented as mean ± SEM of two biological replicates and fit to a one phase exponential decay function, $t_{1/2}$: 1.7 ± 0.9 min. C) Normalized fluorescence intensity of Halo:T-DiP complex following 20 min light exposure. Data are represented as mean ±

SEM of two biological replicates for each cytosol, mitochondria, and ER targeted Halo protein.

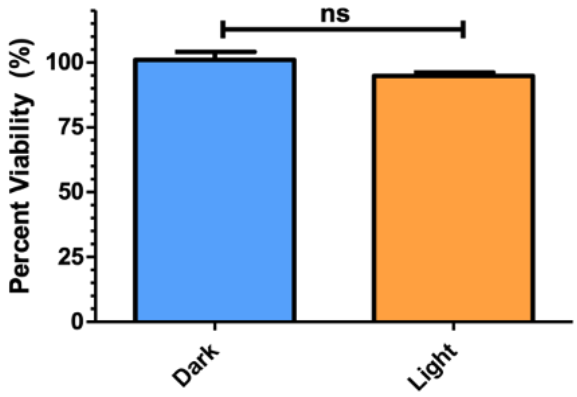


Figure 4.15. 20 minutes of 365 nm light exposure is nontoxic

Cell viability of HEK 293T cells following 20 min irradiation of 365 nm light as measured by alamar blue. Data are represented as mean ± SEM from 12 technical replicates and normalized to dark control. Statistical analysis was performed by a t-test.

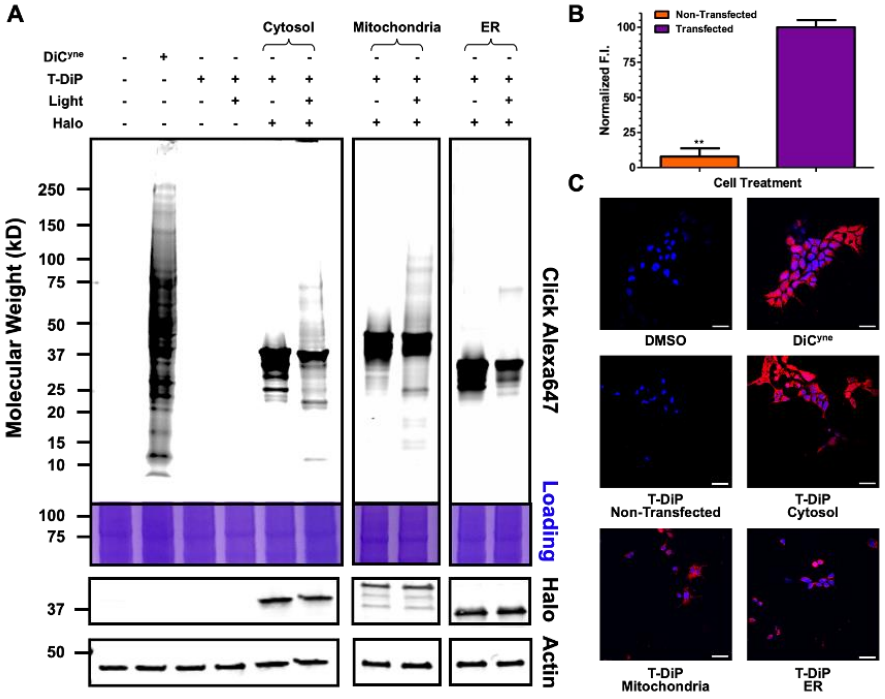


Figure 4.16. Controlled release of DiC^{yne} by T-DiP platform in live HEK 293T cells

A) Representative gels and western blots of non- or Halo- transfected cells treated with either DMSO, 500 μM 2 for 4 h, or 12.5 μM T-DiP with or without a 20 min 365 nm light exposure followed by a 15 min incubation prior to lysis. Lysates were subjected to a CuAAC reaction with Alexa647 azide. B) Normalized fluorescent intensity (F.I.) of Alexa647 signal in non- transfected cells vs. cells expressing cytosolic Halo protein following T-DiP treatment and light exposure. Data are represented as mean ± SEM from 3 biological replicates. Statistical analysis was performed by a t-test (**P < 0.01). C) Confocal imaging of non- or Halo- transfected cells and treatment with either DMSO, 500 μM 2 for 4 h, or 12.5 μM 8 with a 20 min 365 nm light exposure followed by a 15

min incubation prior to fixation. DAPI is shown in blue and Alexa647 is shown in red. Scale bar in bottom left denotes 20 mm.

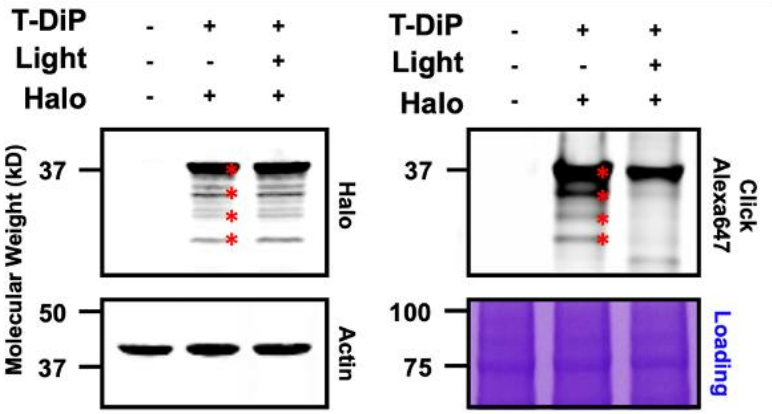


Figure 4.17. T-DiP binds to truncated halo protein

Representative agarose gels and western blots of HEK 293T with or without transfection of cytosolic Halo treated with DMSO or 12.5 μM T-DiP, with or without 20 min exposure to 365 nm light. Lysates were subjected to a subsequent CuAAC reaction with Alexa647 azide and separated by SDS-PAGE. Left: western blots probed for Halo protein (top) and actin (bottom). Right: Fluorescence imaging of Alexa647 (top) and Coomassie stain loading control (bottom). Red stars (*) denote overlap of Halo protein western bands and fluorescent bands observed at the same molecular weight, signifying binding of T-DiP to these truncation products.

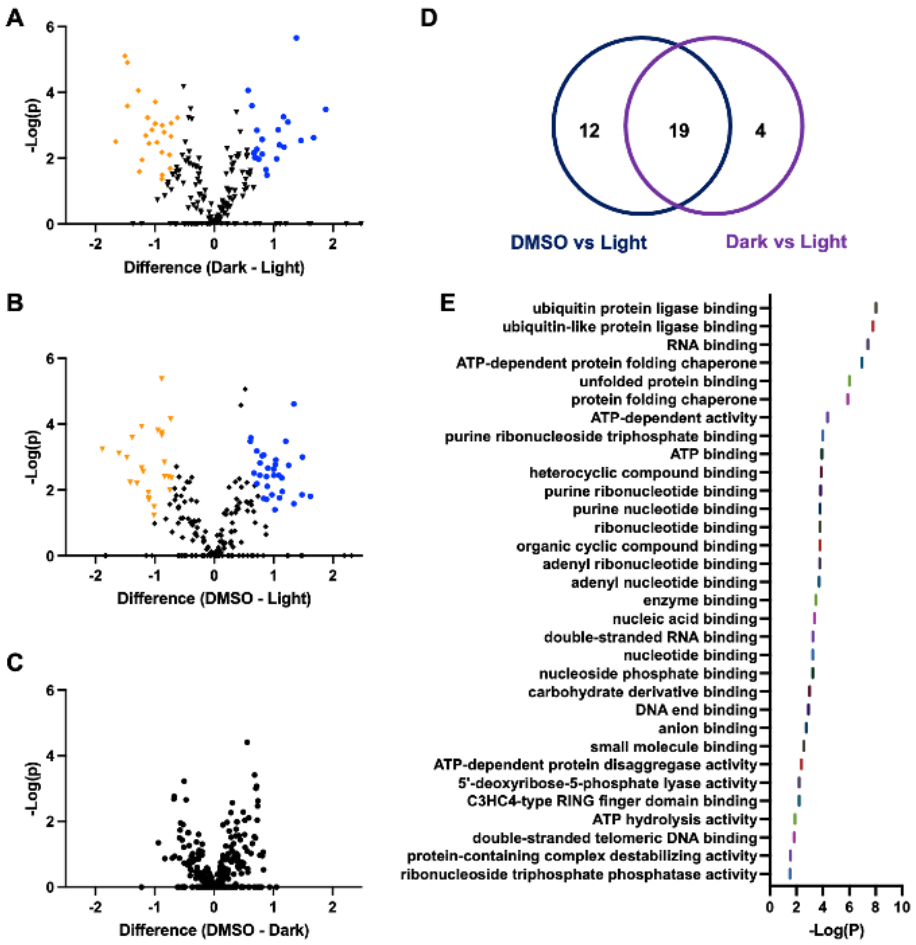


Figure 4.18. Identification and bioinformatic analysis of proteins enriched via T-DiP controlled release

A-C) Volcano plot of proteins identified as DiC^{yne} adducts comparing against Dark vs Light, DMSO vs Light, and DMSO vs Dark treatment conditions, respectively. Proteins were identified via LC-MS/MS and statistically analyzed for enrichment at an FDR of 0.01 and a minimal coefficient of variation (S0) of 0.5. Proteins highlighted in blue were significantly enriched in the Light conidiations. Data are representative of 3 biological replicates from each treatment condition. D) Venn diagram showing overlap in the

proteins enriched in the light treatment group between DMSO vs Light and Dark vs Light comparisons. E) Biological process analysis of proteins enriched by T-DiP controlled release. Analysis was performed with the DAVID overrepresentation test from the protein enriched in the DMSO vs Light dataset and results are shown as a Benjamini-Hochberg adjusted p-value.

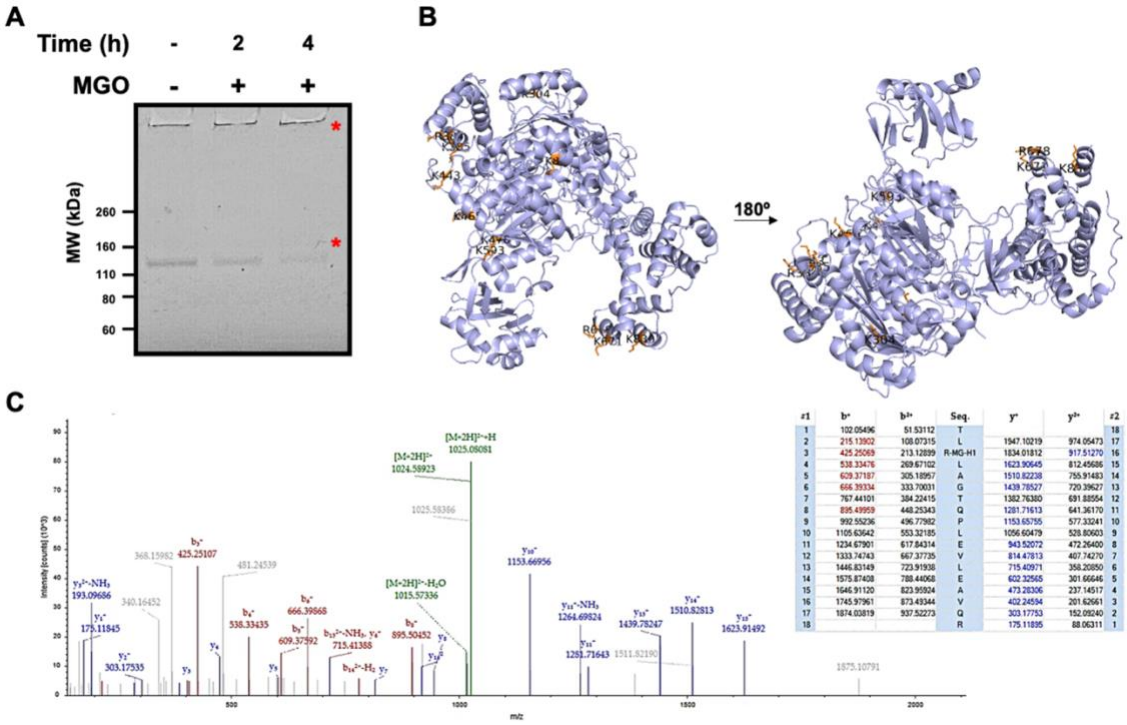


Figure 4.19. UBA1 is covalently modified by MGO

A) Representative SDS-PAGE gel after incubating UBA1 for with MGO for 2 or 4 hours. The * denotes higher molecular weight aggerates induced by MGO treatment. B) Visualization of UBA residues modified by MGO. Modified sidechains are highlighted in orange. PDB 6DC6. C) Representative MS/MS spectrum for the peptide containing the MGO modification at Agr678. The tabled on the right denotes the expected Y and B ion series with the observed ions colored in red and blue.

Table 4.1. List of enriched proteins between dark and light treatment groups

The third column reports the T-test pValue to assess enrichment. The fourth column reports the difference between the mean log2 transformed LFQ intensity values between Dark and Light treated groups, which is indicative of fold change difference in protein abundance. Data is representative of 3 biological replicates from each treatment condition.

Protein Name	Gene name	Negative Log(P)	Difference (Dark - Light)
nucleosome assembly protein 1 like 1	NAP1L1	**3.482**	**1.876**
peroxiredoxin 6	PRDX6	**2.625**	**1.673**
heterogeneous nuclear ribonucleoprotein H1	HNRNPH1	**2.539**	**1.46**
ubiquitin like modifier activating enzyme 1	UBA1	**5.657**	**1.382**
heat shock protein family A (Hsp70) member 1B	HSPA1B	**3.097**	**1.239**
X-ray repair cross complementing 5	XRCC5	**2.339**	**1.174**
calnexin	CANX	**3.257**	**1.164**
eukaryotic translation initiation factor 4A1	EIF4A1	**2.403**	**1.097**
phosphoglycerate mutase 1	PGAM1	**2.863**	**1.075**
karyopherin subunit beta 1	KPNB1	**1.977**	**1.047**
protein disulfide isomerase family A member 6	PDIA6	**1.482**	**0.887**
leucine rich pentatricopeptide repeat containing	LRPPRC	**1.652**	**0.872**
t-complex 1	TCP1	**2.12**	**0.808**
heat shock protein family D (Hsp60) member 1	HSPD1	**2.566**	**0.807**
tripartite motif containing 28	TRIM28	**2.572**	**0.804**
tubulin beta 4B class IVb	TUBB4B	**1.966**	**0.748**
phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole	GART	**2.847**	**0.719**

synthetase			
heat shock protein family A (Hsp70) member 4	HSPA4	**2.271**	**0.717**
synaptotagmin binding cytoplasmic RNA interacting protein	SYNCRIP	**2.013**	**0.708**
poly(U) binding splicing factor 60	PUF60	**2.018**	**0.683**
X-ray repair cross complementing 6	XRCC6	**2.169**	**0.676**
clathrin heavy chain	CLTC	**2.157**	**0.663**
heat shock protein family A (Hsp70) member 8	HSPA8	**3.594**	**0.634**
creatine kinase B	CKB	**4.064**	**0.567**

Table 4.2. List of enriched proteins between DMSO and light treatment groups

The third column reports the T-test pValue to assess enrichment. The fourth column reports the difference between the mean log2 transformed LFQ intensity values between DMSO and Light treated groups, which is indicative of fold change difference in protein abundance. Data is representative of 3 biological replicates from each treatment condition.

Protein Name	Gene name	Negative Log(P)	Difference (DMSO - Light)
peroxiredoxin 6	PRDX6	**1.804**	**1.618**
heterogeneous nuclear ribonucleoprotein H1	HNRNPH1	**2.999**	**1.486**
nucleosome assembly protein 1 like 1	NAP1L1	**1.852**	**1.477**
eukaryotic translation initiation factor 4A1	EIF4A1	**1.578**	**1.34**
ubiquitin like modifier activating enzyme 1	UBA1	**4.61**	**1.34**
phosphoglycerate mutase 1	PGAM1	**2.748**	**1.25**
poly(U) binding splicing factor 60	PUF60	**3.475**	**1.204**
X-ray repair cross complementing 6	XRCC6	**1.948**	**1.145**
t-complex 1	TCP1	**2.362**	**1.138**
clathrin heavy chain	CLTC	**1.756**	**1.098**
tubulin beta 4B class IVb	TUBB4B	**2.442**	**1.087**
nucleolin	NCL	**2.91**	**1.037**
heat shock protein family A (Hsp70) member 4	HSPA4	**2.766**	**1.026**
leucine rich pentatricopeptide repeat containing	LRPPRC	**1.4**	**1.02**
karyopherin subunit beta 1	KPNB1	**2.442**	**1.001**
tripartite motif containing 28	TRIM28	**2.626**	**0.997**
calnexin	CANX	**1.852**	**0.972**

enolase 1	ENO1	**2.664**	**0.903**
nuclear distribution C, dynein complex regulator	NUDC	**2.109**	**0.894**
nucleophosmin 1	NPM1	**2.404**	**0.88**
protein disulfide isomerase family A member 6	PDIA6	**1.714**	**0.88**
importin 9	IPO9	**1.74**	**0.832**
tubulin beta class I	TUBB	**3.059**	**0.831**
lysyl-tRNA synthetase 1	KARS1	**3.039**	**0.807**
ribosomal protein SA	RPSA	**2.82**	**0.77**
TNF receptor associated protein 1	TRAP1	**2.441**	**0.769**
tubulin alpha 1b	TUBA1B	**2.188**	**0.718**
synaptotagmin binding cytoplasmic RNA interacting protein	SYNCRIP	**3.184**	**0.715**
ribosomal protein L4	RPL4	**2.514**	**0.669**
heat shock protein family A (Hsp70) member 8	HSPA8	**3.58**	**0.616**
phosphoribosylaminoimidazole carboxylase and phosphoribosylaminoimidazolesuccinocarboxamide synthase	PAICS	**3.479**	**0.608**

4.4 Discussion

Dicarbonyl protein adducts are an important class of non-enzymatic posttranslational modification. Continued development of methods for studying the proteins prone to modification by dicarbonyl metabolites like MGO is an important area to improve understanding of the biological role such adducts play in normal cell function and under disease conditions. Our T-DiP system described here enables controlled release of dicarbonyls upon light exposure, making it a useful chemical biology tool to add the collection of other caged electrophile systems.[225, 313] However, despite the lack of protein adducts observed in the absence of light, we cannot rule out that the exposed redox sensitive ketone in T-DiP could potentially react reversibly with cellular components and lead to artifacts. Second generation dicarbonyl photocages could explore masking the ketone moiety in addition to the reactive aldehyde.[258]

Our initial MS based proteomic results suggest that low levels of increased intracellular MGO may influence ubiquitin dependent cellular processes. Validating that UBA1 is susceptible MGO adduction in an *in vitro* setting provides additional evidence for this. We hypothesize that MGO modification of UBA1 may impede UBA1 activity or block UBA1 protein-protein interactions and thereby induce toxicity in a cellular context. It has been documented that posttranslational modification to UBA1 via acylation and phosphorylation affects activity its activity.[314] Interestingly, diminished UBA1 expression and activity has been linked to neurodegenerative disorders that MGO is also implicated in.[315] However, the functional consequences of MGO modification to UBA1 activity

remain to be elucidated. Future work exploring how intracellular MGO levels are associated with protein ubiquitination is warranted to validate these fundings.

Collectively, we have developed an effective photocage for α-oxoaldehydes in biological systems. By installing a HaloTag targeting element in our photolabile dicarbonyl precursor, we demonstrated that T-DiP technology can be utilized to rapidly release physiologically relevant concentrations of dicarbonyls intracellularly and generate glycated proteins under spatiotemporal control in live cells. We envision that this chemical tool will aid in elucidating the precise mechanisms by which dicarbonyl stress influences cellular fitness and contributes to ageing-related diseases.

5 Chapter V: Proteomic Profiling of DNA-Protein Crosslinks Generated by Methylglyoxal

Adapted in part from:

Erber L.,* Hurben A.K.,* Tretyakova N.Y. Proteomic Profiling of DNA-Protein Cross-Links Generated by Methylglyoxal. *manuscript in preparation*.

* denotes equal contribution.

This chapter details the work performed by Alexander K. Hurben (conducted cell culture and *in-vitro* generation of DPCs) and Dr. Luke Erber (performed DPC isolation and characterization) under the supervision of Prof. Tretyakova.

5.1 Introduction

DNA-protein crosslinks (DPCs) are formed upon covalent trapping of cellular proteins to DNA. DPCs are formed upon exposure to both physical and chemical agents, including endogenous agents like formaldehyde[316] and reactive metabolic intermediates,[317] exogenous agents such as 1,3-butadiene,[318-320] ionizing radiation,[321] UV light,[322] transition metals,[323] as well as therapeutic agents; platinum drugs,[324] nitrogen mustards,[325-327] and haloethylnitrosoureas.[328] These bulky DNA lesions distort the DNA helix, and thereby obstruct essential DNA-protein interactions necessary for DNA recombination, transcription, replication, repair, and remodeling of chromatin. [322, 329-331] If not repaired, DPCs lead to permanent DNA alterations and toxicity.[322, 332-334] Such DNA damage is observed in individuals suffering from Ruijs-Aalfs syndrome, a genetic disorder defined by polymorphisms in the SPRTN gene which encodes a protease involved in DPC repair. Patients with Ruijs-Aalfs syndrome exhibit elevated levels of DPCs, genomic instability, premature aging and are at high risk for hepatocellular carcinomas.[335-337] Despite the ubiquitous formation of DPCs, the structural identities and repair pathways are not completely understood due to the heterogeneity and challenge of studying bulky biomolecular conjugates containing characteristics of proteins and DNA.

Methylglyoxal (MGO) is an electrophilic α-oxoaldehyde byproduct of cellular metabolism and is capable of reacting with nucleophilic biomolecules to form covalent adducts in a process defined as glycation. MGO is primarily derived through the spontaneous hydrolysis of glycolysis intermediates and is present at low micromolar levels in eukaryotic cells.[338] To maintain low micromolar levels, MGO reacts with

glutathione and is further processed by the glyoxalase activity of GLO1 and GLO2 to form D-lactate.[339] MGO levels are elevated in diabetes, cancer, renal failure, and neurodegenerative diseases.[339-341] Increased levels of MGO leads to formation of protein adducts and elevated inflammation signaling and thereby contributes to disease phenotype.[342, 343]

As stated in **Chapter IV**, MGO covalently modifies DNA and proteins.[344] MGO primarily modifies the nucleophilic sites of arginine and lysine amino acids and guanine and adenine bases.[290, 345] These adducts are collectively referred to as advanced glycation end-products (AGEs). However, MGO has also been shown to generate covalent DNA-protein crosslinks. *In vitro* assays have revealed MGO crosslinking of DNA polymerase 1 with deoxyguanosine.[317] More recently, MGO was shown to crosslink deoxyguanosine with N-acetyl-lysine and cross link DNA to histones in nucleosome core particles.[346, 347] Additionally, SPRTN deficient cells and Ruijs-Aslfs patient lymphoblastoid cell lines with monogenic and biallelic mutations in SPRTN exhibit increased sensitivity to MGO exposure.[348] Thus, amounting evidence suggests that MGO derived DPCs may play a role in disease etiology, warranting deeper exploration into this class of DPCs.

There is an unmet need to comprehensively profile and identify MGO-derived DPCs formed in cells. Earlier *in vitro* experiments have been limited to synthetic DNA duplexes which do not reflect normal DNA–protein interactions observed in cells. The purpose of the present study was to characterize DNA–protein cross-linking in human cells treated with MGO. We coupled cell culture treatments and DPC purification assays to characterize MGO-dependent DPC formation and utilized mass spectrometry-based

proteomics to identify proteins trapped to DNA upon MGO exposure. Our efforts produced a list of MGO derived DPCs which provides insight into potential DPCs that may be relevant in diseases linked to elevated MGO levels. Subsequent bioinformatic analysis revealed that these DPCs may play a role in telomere DNA organization, prompting us to validate DPC formation between GAPDH and telomere DNA in the presence of MGO. Ultimately, this work provides a foundational set of MGO derived DPCs to further investigate.

5.2 Materials and Methods

MGO was obtained from Millipore sigma as a 40% solution in water. SimplyBlue™ SafeStain was purchased from Invitrogen. Pierce™ color silver stain kit, NuPAGE 4-12% bis tris SDS-PAGE gels, cell culture media, reagents, recombinant BSA, and other consumables were obtained from Thermo Fisher Scientific. Antibodies were purchased from Cell Signaling. Recombinant GAPDH was obtained from Cayman Chemical. The FAM labeled DNA template sequence for generation of DNA-protein crosslinks FAM-Tel FAM - AGG GTT AGG GTT AGG GTT was purchased from IDT.

Cell Culture:

The human fibrosarcoma (HT1080) cells were collected from the American Type Cell Culture Collection.[349] SPRTN-deficient mouse embryonic fibroblast (MEF7, SPRTN f/-) and wild type (MEF5, SPRTN +/+) cells were collected from Dr. Yuichi Machida's laboratory (Mayo Clinic, MN).[350] GLO1 and GLO2 knockout HEK293T cells were obtained from Dr. James Galligan's laboratory.[309] Wild type HEK293T cells were obtained from ATCC. Cells were grown using Dulbecco's Modified Eagle's Medium containing 10% fetal bovine serum and 1% penicillin/streptomycin. Cells were maintained in a humidified incubator with 5% CO_2 at 37 °C.

AlamarBlue Cell Viability Assay:

HEK293T cells were cultured in MEM media containing 10% FBS and Pen/Strep 100 U/mL under a 5% CO2 atmosphere at 37 °C. Cells were seeded into 96 well plates at a density of 20,000 cells per well. 24 h after seeding, the media was removed and replaced with the indicated MGO concentration in complete MEM media. Following the

allotted incubation time, the media was removed and replaced with 100 μL of a 10% alamarBlue solution in complete MEM media. The plates were incubated for 4 h in the dark at 37 °C and then fluorescence was measured at Ex/Em 560/590 nm on a BioTek Synergy H1 plate reader. Cell viability for a given treatment condition was calculated by normalizing the fluorescence signal to the vehicle treated control.

Isolation of DPC-associated DNA by K-SDS Assay:

This assay was performed according to the following reference.[351] The specified cell lines were seeded into 6 well plates at a concentration of 600,000 cells per well. The cells were further incubated for 24 h after which the media was removed and replaced with the indicated treatment conditions. Following MGO treatment for the allotted time, the media was removed, and the cells were washed with PBS. The cells were then lysed in 500 μL 20 mM Tris-HCl (pH 7.4), 2% SDS. Lysate was sheared and solubilized using a syringe (25G). Cells were treated with 200 μg/mL RNAse for 15 min at 37 °C. 10% of the sample was removed and set aside as the “input fraction". The input fraction was treated with 250 μL 200 M KCl, 10 mM EDTA, 20 mM Tris-HCl (pH 7.5) and proteinase K overnight at 37 °C. To the remaining sample, an equal volume of precipitation buffer was added (200 mM KCl in 20 mM Tris-HCl (pH 7.5). This sample represented the “DPC pellet fraction”. Samples were vortex-mixed vigorously. Samples were incubated on ice for 5 min and then centrifuged for 5 min at 18,000 x g at 4 °C. The supernatant was discarded. The pellet was resuspended in 1 mL of wash buffer, 100 mM KCl in 20 mM Tris-HCl (pH 7.5) and heated for 10 min at 65 °C. Samples were removed from the heat block, incubated on ice for 5 min and centrifuged for 5 min at 18,000 x g at

4 °C. The supernatant was removed, and the sample washing was repeated twice for a total of three washes. When the wash steps were completed, the final pellet was resuspended in 1 mL digestion buffer, 100 mM KCl, 10 mM EDTA, proteinase K in 20 mM Tris-HCl (pH 7.5). Samples were incubated overnight at 37 °C. Following proteinase K digestion, the samples were centrifuged for 10 min at 20,000 x g and the supernatant for the input fraction and DPC pellet fraction were collected for DNA quantitation.

Isolation of DPC-associated Proteins by Phenol-chloroform Assay:

This protocol was adapted from previous reports with slight modification.[320, 326, 333] Briefly, HT1080 cells were seeded into 15 cm plates and grown to 90% confluency. The cells were then treated with MGO at the indicated concentration for 2 hours. Following treatment, the media was removed, and the cells were collected via trypsinization. Cell pellets were washed, resuspended in 2 mL PBS and lysed in 2 mL 2X cell lysis buffer. Lysis buffer contained 20 mM Tris-HCl (pH 7.4), 10 mM $MgCl_2$, 2% Triton X-100 and 20% sucrose in water. Lysates were incubated on ice for 5 min and centrifuged at 2,000 g for 10 min at 4 °C. The supernatant was discarded. Nuclear pellets were resuspended in 2 mL saline-EDTA solution and then treated with 10 μL RNAseA (10 μg/mL) and 300 μL protease inhibitor cocktail. SDS was added to a final concentration of 2%. Samples were mixed well and incubated for 2 h at 37 °C with gentle shaking in a water bath. 2 mL phenol (Sigma, UltraPure) was added to each sample and mixed for 5 min. Samples were centrifuged at 1500 rpm for 10 min at room temperature. The aqueous and interface layers were collected. 2 mL phenol:chloroform 1:1 was added to pooled aqueous and interface fractions. Samples were mixed for 5 min and centrifuged

at 1500 rpm for 10 min at room temperature. The aqueous and interface layers were collected and 4 mL 100% cold EtOH was added and mixed with each sample. Samples were incubated at -20 °C overnight. Samples were centrifuged at 3500 rpm for 15 min. The supernatant was discarded, and the DNA pellet was transferred to a new Eppendorf tube. The pellets were washed twice with 1 mL 70% ice-cold ethanol with centrifugation at 14000 g for 5 min and the supernatant was removed after each wash. After the second wash, the pellets were air dried. The DNA pellets were resuspended in 400 μL TE (Tris-EDTA buffer) and sheared into solution using 19 and 25 gauge needles. The DNA concentration was measured with the PicoGreen assay.

DPC Quantitation by PicoGreen Assay:

DNA quantitation was performed according to the manufacturer's instructions as summarized here. A 96 well plate was prepared wherein each well contained 100 μL TE buffer with 0.5% Picogreen reagent. For each biological sample, two wells were prepared. 1 μL DNA from the input fraction was mixed into well one and 1 μL DNA from the DPC pellet fraction was mixed into well two. A control sample from a mock sample was used as a blank. All samples were mixed and incubated in the dark for two minutes. A DNA standard curve was generated using 1 μL from 50, 10, 5, 1, 0.5, and 0.1 ng/μL DNA stocks. Fluorescence detection was completed using a Synergy2 plate reader (BioTek). The fluorescence excitation and emission wavelengths were 485 nm and 528 nm respectively. The DNA amount for each sample was calculated by using the standard curve generated using the fluorescence comparison with the known DNA stocks. Protein-

crosslinked DNA μg amounts from DPC pellet fractions are divided by total μg DNA amount calculated from the input fractions to obtain percent of protein-crosslinked DNA.

DPC Quantitation by Dot-blotting:

For immunodetection, the DPC fraction from the samples isolated by the Phenol-chloroform assay were diluted in tris buffered saline (TBS), 10 mM Tris-HCl (pH 7.5) in 150 mM NaCl. Diluted samples were loaded onto a dot blot apparatus containing a nitrocellulose membrane. Following sample loading, the membrane was blocked using 2% BSA in TBS for 1 hour at room temperature and incubated overnight with primary antibodies at 4 °C. the nitrocellulose membrane was washed with TBS buffer containing 0.1% tween20 three times, 5 min per wash on a shaker. The membrane was further incubated with secondary antibody (LiCor, 1:10,000) for 1 h at room temperature on a shaker. The membrane was further washed again in TBS buffer containing 0.1% tween20 three times, 5 min per wash on a shaker. The membrane was imaged and quantified using the LiCor OdysseyFc instrument on the 800 nm channel.

Mass Spectrometry-based Identification of DPC-Associated Proteins:

Chromosomal DNA containing DPCs were extracted using the phenol-chloroform method described above. 50 μg DNA was adjusted to pH 8 using 20 mM triethylammonium bicarbonate (TEAB) and was treated with 5 mM dithiothreitol for 1 h at 56 °C. Samples were returned to room temperature for 5 min and treated with 10 mM iodoacetamide for 30 min at room temperature in the dark. Samples were treated with 2 μg trypsin overnight at 37 °C in a water bath. Samples were desalted with Pierce C-18 spin columns according to the manufacturer's instructions. Samples were dried by

vacuum evaporation and stored at -20 °C. Tryptic peptides from each sample was analyzed by liquid chromatography coupled to tandem mass spectrometry using a Eksigent NanoLC-Ultra 2D HPLC system coupled to a nanospray ESI source and the Thermo Scientific Q Exactive Mass Spectrometer. Peptides were resuspended in 16 μL 0.1% formic acid. 4 μL sample were separated using an in-house packed C-18 column (15 cm x 75 μm, Luna C18) across a 60 min gradient. Mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1% formic acid in acetonitrile. Source parameters were set as 3.0 kV spray voltage, capillary temperature of 300 °C and S-lens RF level of 50%. MS analysis was performed using a top 12 data-dependent mode. The full scan was run using the resolution setting of 70,000 and an AGC target of 1×10^6. The exclusion duration was set for 30 s for dynamic exclusion. The 12 most intense precursor ions were chosen for HCD fragmentation which was set at 30%. The fragmentation scan was acquired with an isolation window of 2.0 *m/z*, a resolution of 17,500 at an AGC target of 5×10^4 or a maximum injection time of 50 ms. Raw mass spectrometry data was processed using MaxQuant (version 1.6.5.0).[190] Data was processed using default parameters using the human Uniprot database (downloaded on 2019/05/15, with 74,349 sequences). Label-free quantification was enabled in MaxQuant using default LFQ parameters.

Bioinformatic Analysis of Mass Spectrometry Results:

The Perseus software (version 1.6.6.0) was used to process the protein LFQ intensities. The LFQ intensities were log2 transformed and processed according to previously described methods.[235, 236] A two-tailed, two-sample t-test was performed to

compare protein abundance between experimental groups. Statistically significant enriched proteins were identified with a Benjamini-Hochberg corrected FDR of 5% and a minimal coefficient of variation (S_0) of 0.5. Functional gene ontology enrichment was performed using the tool DAVID Bioinformatics Resources 2021.[352]

Characterization of DNA-Protein Crosslinks Generated with Recombinant Proteins:

In vitro DPCs were prepared by incubating 1 μg of the specified protein with FAM labeled ss-telomeric DNA $(TTAGGG)_3$ (25 μM) in the presence of MGO (10 mM) in 25 μL of PBS (pH 7.4) at 37 °C. After the allotted reaction time, the reactions were quenched upon addition of 6.25 μL of 4x NuPAGE gel loading buffer and 25 μL of reaction mixtures were subjected to SDS-PAGE on NuPAGE 4-12% bis tris 1 mm gels. Following separation, gels were imaged on a Typhoon7000 fluorescence imager to visualize FAM labeled DPCs and subsequently stained with by simply blue or silver stain to visualize protein loading.

5.3 Results

Our first objective was to verify that MGO treatment generated DPCs in cell models. We assessed MGO toxicity to identify appropriate MGO treatment conditions for generating MGO induced DPCs. We treated HEK293T cells across a 50 μM – 10 mM MGO treatment range for 2-, 4-, and 24-hour durations and measured their viability using an almarblue assay (**Figure 5.1A**). We observed that escalation of MGO concentration led to a dose dependent decrease in cell viability. Additionally, we observed that treatment duration significantly impacted cell viability. Treatment with 10 mM MGO for 2, 4, and 24 hours reduced cell viability to 75 ± 20 %, 35 ± 18 %, and 0.4 ± 1 %, respectively. To limit excessive cell death caused by MGO, we selected a 2-hour MGO treatment window prior to DPC isolation. Similar studies which generated cellular DPCs via treatment of crosslinking agents have also reported using 1 – 3 hour long incubation times.[326, 353-355]

Next, we performed a K-SDS assay to quantify the relative amounts of DPCs generated upon MGO treatment. In the K-SDS assay, treatment with sodium dodecyl sulfate (SDS) enables protein-specific precipitation with KCl, permitting separation of protein-bound DNA from free DNA. Following multiple rounds of protein precipitation and washing, the free DNA is removed, and the remaining DNA is quantified to reveal the global DPC level in the sample as a function of percent input.[351] As shown in **Figure 5.1B**, DPC levels rose in response to increased MGO treatment in HEK293T cells, indicating MGO induces DPC formation.

MGO is detoxified in the cell by the glycosylase activity of GLO1 and GLO2. We obtained cells that contained a knockout of either GLO1 or GLO2, two of the principal enzymes responsible for detoxifying MGO.[356-360] MGO treatment of these GLO1 or GLO2 deficient cells led to significantly increased formation of DPCs by 2.4-fold and 2.7-fold, respectively, as measured by the K-SDS assay (**Figure 5.1C**). These results suggest that reduced MGO detoxification can lead to increased levels of MGO derived DPCs.

To test whether MGO-dependent DPCs are recognized for proteolytic repair, MEF cells exhibiting a deficiency in SPRTN expression, a critical gene required for DPC repair were obtained.[350, 361] Knockdown of the DPC repair enzyme SPRTN resulted in an elevated DPC signal in comparison to the WT cells upon exposure to 2.5 and 5 mM MGO when measured by the K-SDS assay (**Figure 5.1D**). These data indicate that DPCs generated by MGO are recognized by the SPRTN repair pathway.

After verifying that MGO forms DPCs, we sought to isolate these crosslinked adducts. To isolate proteins covalently linked to DNA, we performed a phenol:chloroform extraction. This extraction methodology enables DPCs to be isolated via phase partitioning.[326, 362, 363] As shown in **Figure 5.1E**, HT1080 cells were treated with 0.5, 1, 2.5, and 5 mM of MGO and DPCs were successfully isolated using the phenol:chloroform method. The samples were normalized by total DNA content, separated by SDS-PAGE and visualized via total protein staining. These results verified our ability to isolate MGO derived DPCs and enabled us to move forward onto identifying these crosslinks through mass spectrometry.

Upon establishing that MGO exposure led to the formation of DPCs, we sought to identify the proteins prone to crosslinking. We treated HT1080 cells with 0 or 5 mM MGO for 2 hours and isolated the DPCs via phenol:chloroform extraction (**Figure 5.2**). The samples were normalized by DNA content and processed for mass spectrometry by reducing and alkylating exposed Cys residues, digesting the proteins with trypsin and desalting the tryptic peptides. LC-MS/MS was performed and proteins within the samples were identified and quantified using the MaxQuant software package. Proteomic analysis revealed 265 statistically enriched DPC proteins with pValues lower than 0.01 within the MGO treatment group (**Figure 5.3A** and **Table 5.1** for a full listing of proteins). These enriched DPC proteins included known notable DNA binders previously identified as DPCs like GAPDH, PARP1, histones, XRCC1, and TOP1. Additionally, we verified that DPC enrichment was consistent between biological replicates by correlating the label-free quantitation (LFQ) intensity values of identified proteins between all samples (**Figure 5.4**). We observed a correlation of 0.94 ± 0.02 in our MGO treated samples, indicating reproducible DPC isolation. Additionally, the correlation within the vehicle treatment group was moderate, 0.81 ± 0.07; whereas the correlation between MGO and vehicle groups was low, 0.47 ± 0.03, indicating that DPC enrichment was robust.

Following proteomic identification of DPCs enriched upon MGO treatment, gene ontology analysis was performed.[364, 365] As shown in **Figure 5.3B**, a cellular compartment overrepresentation test with the DAVID database found nuclear proteins associated with the nucleoplasm, nucleus, nucleolus, and chromatin to be significantly enriched. Of the proteins identified, 69.2% of them were annotated with nucleus

affiliation. These results are expected as nuclear proteins are in close proximity to DNA and thus make likely targets for MGO crosslinking. Enrichment for ribosomal and membrane proteins was also observed, which were found in 10.5% and 53.8% of the proteins, however it should be noted that many of these proteins participate in multiple biological process and can be found in multiple subcellular locations.[366] Previous work in identifying DPCs induced by the chemical crosslinker cisplatin has also found ribosomal proteins to be enriched by phenol:chloroform extraction following treatment.[354] Subsequent bioinformatic analysis of the identified MGO-derived DPCs was performed using the DAVID biological process enrichment test (**Figure 5.3C**). Processes such as telomere organization, and nucleosome assembly were found to be significantly enriched. These results indicate that MGO derived DPCs could affect nucleosome architecture. Interesting RNA dependent processes were also enriched. Given the RNAase digestion step was included in the phenol:chloroform extraction, we hypothesis these proteins also possess DNA binding activity.

We also compared the enriched MGO-derived DPC dataset to previously identified DPCs discovered in our laboratory. This repository of DPCs was obtained from treating HT1080 cells with either diepoxybutane (DEB), cisplatin (Cpt) or phosphoramide mustard (PM) and isolating DPCs via phenol:chloroform extraction.[320, 354, 363] As shown in **Figure 5.3D**, there was a 19.2%, 33.0% and 22.7% overlap in DPC identified from DEB, Cpt, and PM treatment, respectively, when compared to DPC enriched upon MGO treatment. Notable overlapping proteins with between the dataset include DExD-Box Helicase 21, eukaryotic translation termination factor 1, and histone

H4. A full list of identified DPCs can be found in **Table 5.2**. Additionally, MGO derived DPCs were compared to those generated by mechlorethamine and no overlap between the dataset was found.[326] We also compared our MGO DPC data to another study which profiled at identified 36 DPCs formed via UV 254 nm irradiation in HeLa cells.[367] We observed 5 overlapping proteins between the two studies, namely desmoyokin, histone H3, histone H4, polypyrimidine tract binding protein 1, and topoisomerase 2A. Comparison of these datasets suggest that there are proteins which are highly susceptible to promiscuous DPC formation and other covalent protein-DNA interactions that are dependent on the species of electrophile or irradiation used.

To further confirm our proteomic results, isolated MGO induced DPCs were subjected to dotblot analysis with commercial antibodies against suspected DPCs. HT1080 cells were treated with 0.5, 1, 2.5, and 5 mM of MGO for 2 hours and DPCs were extracted via phenol:chloroform as described above. Extracted DPCs were normalized based on DNA content and analyzed by dotblot for proteins found in our proteomic data (**Figure 5.4A**). Quantification of the fluorescent intensity signals for antibodies against PARP1, histone H3A, XRCC1, TOP1, VINC, and GAPDH, revealed a dose dependent increase of protein abundance upon MGO exposure when normalized to dsDNA input (**Figure 5.4B**). Additionally, the protein GSTP1, which was not found in our enriched DPC dataset, served as a negative control, and displayed no dose dependent increase in signal upon MGO treatment. We also performed a dotblot analysis probing for ubiquitin (UBB) and observed an MGO dependent increase in signal (**Figure 5.5**). While we do not expect UBB to be covalently adducted to DNA via MGO, these results suggest

that UBB marks are being installed on MGO generated DPCs and may be participating in a DNA repair response. Collectively, these dotblot data further confirm our proteomic results and indicate that PARP1, histone H3A, XRCC1, TOP1, VINC, and GAPDH are susceptible to form DNA protein crosslinks through reaction with MGO.

Finally, we sought to confirm MGO dependent DPC formation using an *in vitro* system with GAPDH. GAPDH has been reported to bind telomeric DNA sequences with a K_d of 45 nM.[368] Given bioinformatic analysis of the identified MGO DPCs found telomeric organization to be significantly enriched, we sought to validate that GAPDH was susceptible to crosslinking to DNA upon MGO treatment. To assess MGO-GAPDH DPC formation, recombinant GAPDH was incubated with a 3` FAM labeled single-stranded oligonucleotide $(TTAGGG)_3$ in the presence or absence of MGO. The reaction mixture was then subjected to SDS-PAGE to separate unmodified protein from resultant DPCs. Following separation, the DPCs could be visualized via fluorescent imaging and protein staining. As shown in **Figure 5.6A,** a new band visualized by both fluorescent imaging and protein staining is observed upon MGO treatment of GAPDH and the FAM labeled oligonucleotide, indicating MGO dependent DPC formation. Furthermore, the yield of this crosslink is dependent of the duration of MGO treatment (**Figure 5.7)**. To test the protein specificity of MGO dependent DPC formation, we preformed the same experiment as described above, except with BSA, a non-DNA binding protein. As shown in **Figure 5.6B**, no new DPC band was observed, signifying that DNA-protein pre-complexation is likely an important factor in DPC formation by MGO. Based off studies which have modeled the location of GAPDH DNA interactions, we hypothesize that DPC

linkages could be formed at Lys residues 84, 107, 136, 186, 191, 194, or 219 as these residues reside in the DNA binding pocket of tetrameric GAPDH (**Figure 5.6C**).[369, 370] Collectively, these data show that GAPDH is susceptible to DNA protein crosslinking in the presence of MGO.

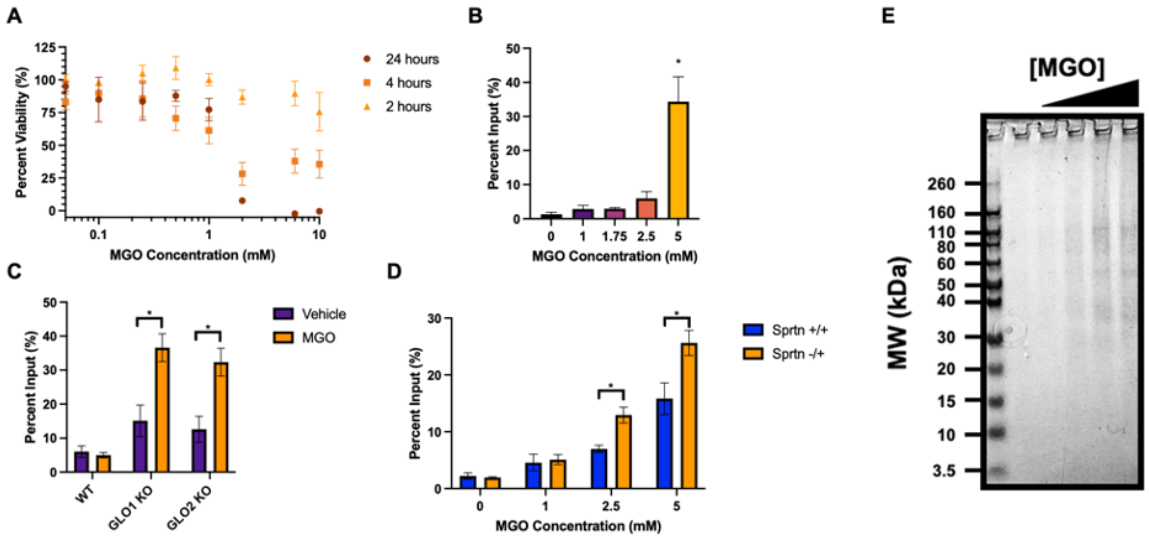


Figure 5.1. MGO treatment generates DPCs in cell models

A) MGO reduces HEK293T cell viability as a function of time and concentration. HEK293T cells were treated with MGO at multiple timepoints and concentrations. Viability was determined through an alamarBlue assay. Data is represented as mean ± SEM from four replicates and normalized to vehicle control. B) MGO treatment generates DPCs in HEK293T cells. HEK293T cells were treated with 0, 1, 1.75, 2.5, or 5 mM MGO for 2 hours and were processed by the K-SDS assay to quantify DPC levels. Data are represented as mean ± SEM from three replicates and normalized to DNA input and were analyzed via a one-way ANVOA and an unpaired t-test (*p < 0.05). C) Knockout of GLO1 or GLO2 leads to increased levels of DPC formation following MGO treatment. HEK293T cells with or without GLO1 or GLO2 were treated with 0 or 1 mM MGO for 2 hours and were processed by the K-SDS assay to quantify DPC levels. Data are represented as mean ± SEM from three replicates and normalized to DNA input and were analyzed via a one-way ANVOA and an unpaired t-test (*p < 0.05). D) Knockdown

of SPRTN leads to increased levels of DPC formation following MGO treatment. MEF7 cells with or without SPRTN deficiency were treated with 0, 1, 2.5, or 5 mM MGO for 2 hours and then processed by the K-SDS assay to quantify DPC levels. Data are represented as mean ± SEM from three replicates and normalized to DNA input and were analyzed via a one-way ANVOA and an unpaired t-test (*$p < 0.05$). E) Visualization of MGO induced DPCs isolated via phenol-$CHCl_3$ extraction. HT1080 cells were treated with 0, 0.5, 1, 2.5, or 5 mM of MGO for 2 hours and were processed by the phenol-$CHCl_3$ extraction to isolate DPCs. Isolated DPCs were resolved by 4 - 12% SDS−PAGE and visualized by SimplyBlue™ SafeStain.

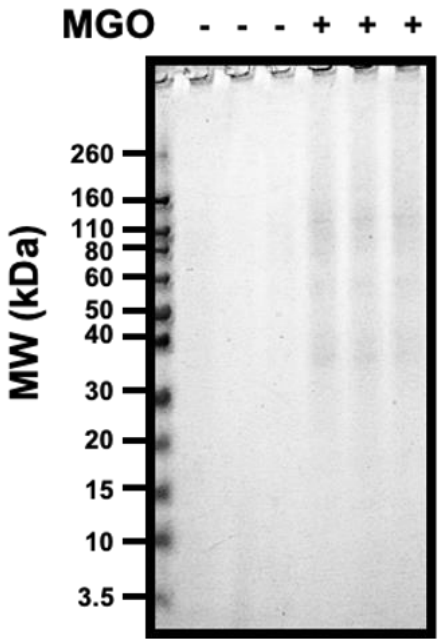


Figure 5.2. MGO induced DPCs isolated via phenol-$CHCl_3$ extraction for proteomic identification

HT1080 cells were treated with 0 or 5 mM of MGO for 2 hours and then subjected to a phenol-$CHCl_3$ extraction to isolate DPCs. Isolated DPCs were resolved by 4 - 12% SDS−PAGE and visualized by SimplyBlue™ SafeStain.

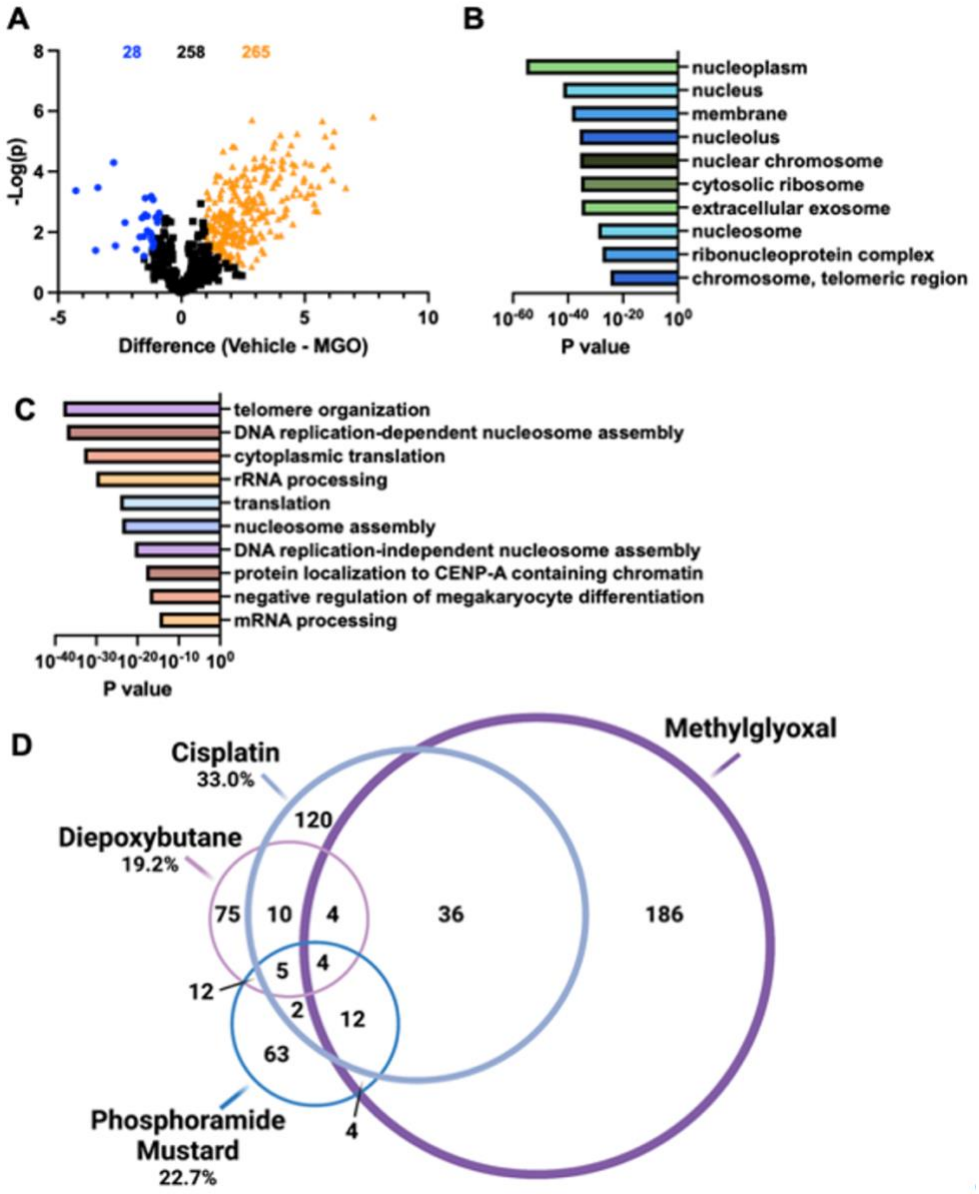


Figure 5.3. Identification and bioinformatic analysis of MGO derived DPCs

A) Volcano plot of proteins identified as potential MGO DPCs. HT1080 cells were treated with 0 or 5 mM of MGO for 2 hours and then subjected to a phenol-$CHCl_3$ extraction to isolate DPCs. DPCs were identified via LC-MS/MS and statistically analyzed for enrichment at an FDR of 0.01 and a minimal coefficient of variation (S0) of

0.5. Proteins highlighted in orange or blue were significantly enriched in the MGO or vehicle treatment, respectively. Data are representative of 3 biological replicates from each treatment condition. B) Cellular compartment analysis of DPCs enriched by MGO treatment. Analysis was performed with the DAVID overrepresentation test and results are shown as a benjamini-hochberg adjusted p-value. C) Biological process analysis of DPCs enriched by MGO treatment. Analysis was performed with the DAVID overrepresentation test and results are shown as a benjamini-hochberg adjusted p-value. D) Depiction of DPCs identified in HT1080 cells upon exposure to diepoxybutane, cisplatin, phosphoramide mustard, and MGO. Venn diagram shows overlap of proteins between each condition. Circle size is representative of the total number of DPCs identified. The precent listed below each treatment condition is indicative of precent similarity to DPCs identified following MGO exposure.

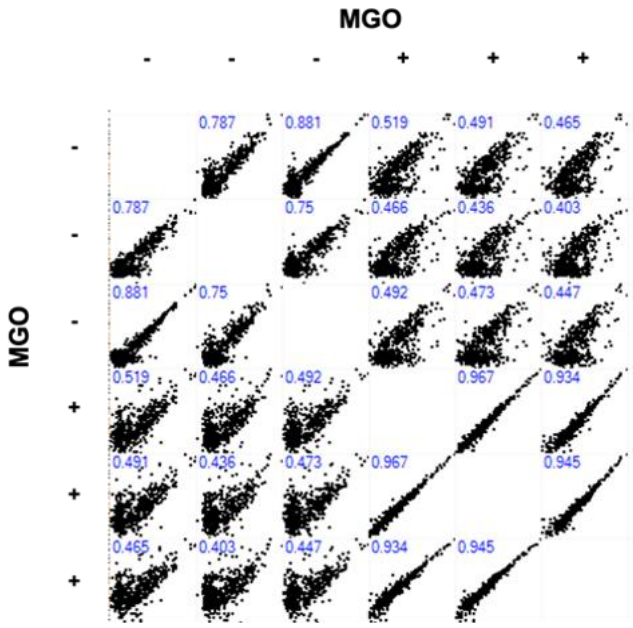


Figure 5.4. Proteomic enrichment of MGO generated DPCs is reproducible

HT1080 cells were treated with 0 or 5 mM of MGO for 2 hours and then subjected to a phenol-$CHCl_3$ extraction to isolate DPCs. DPCs were identified via LC-MS/MS. Correlation plots show the intensity of label free quantitation (LFQ) values for a given treatment condition on the X and Y axis. Identified proteins are represented as black squares. The R^2 value for each comparison is written in blue. The Perseus software suite (version 1.6.6.0) was used generate this plot.

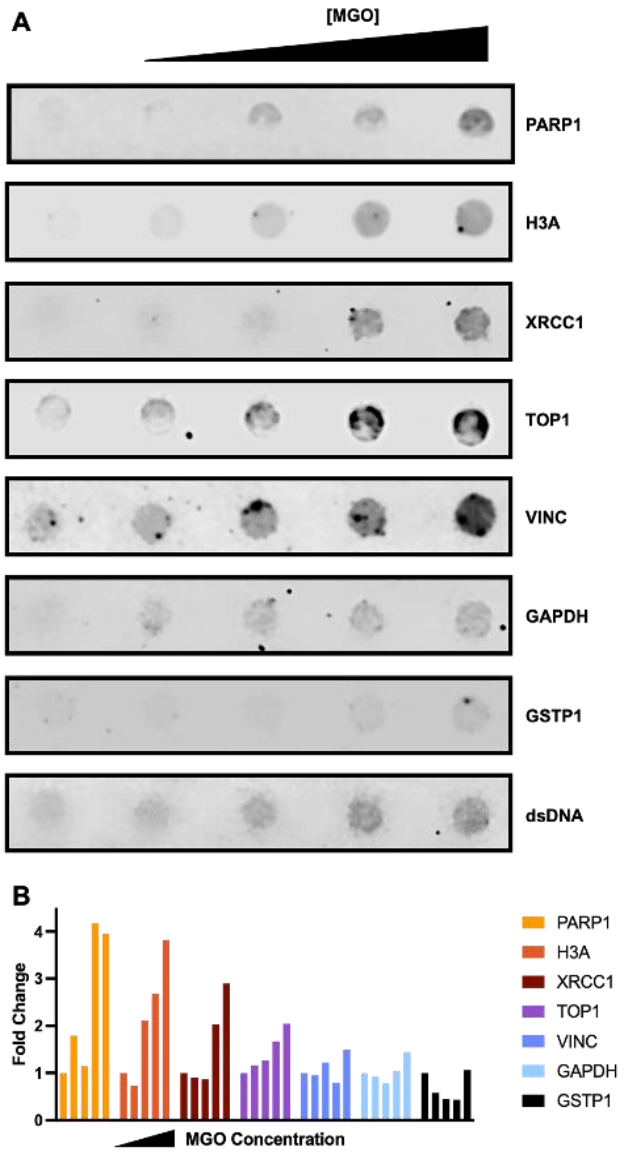


Figure 5.5. Dotblot analysis of MGO induced DPCs

A) Representative dotblots of DPCs isolated via phenol-$CHCl_3$ extraction from HT1080 cells treated with 0, 0.5, 1, 2.5, or 5 mM of MGO for 2 hours. Samples were normalized for DNA content, immobilized on nitrocellulose membranes, and probed with primary antibodies specific for PARP1, histone H3A, XRCC1, TOP1, VINC, GAPDH, GSTP1,

and dsDNA. B) Fold change in measured dotblot fluorescence intensity for PARP1, histone H3A, XRCC1, TOP1, VINC, and GAPDH signal normalized to input dsDNA for 0, 0.5, 1, 2.5, or 5 mM MGO treatments, shown left to right for each protein respectively.

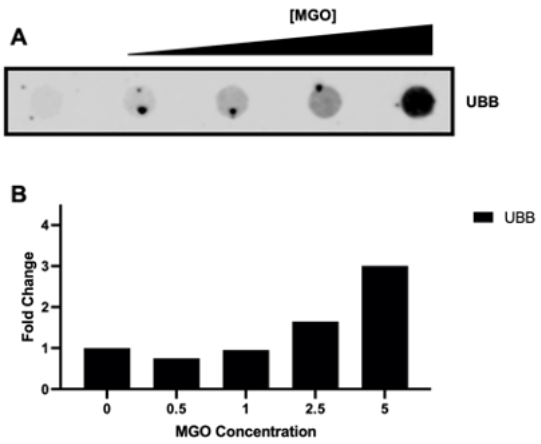


Figure 5.6. Dotblot analysis for ubiquitin in MGO induced DPCs

A) Representative dotblot of DPCs isolated via phenol-$CHCl_3$ extraction from HT1080 cells treated with 0, 0.5, 1, 2.5, or 5 mM of MGO for 2 hours. Samples were normalized for DNA content, immobilized on nitrocellulose membranes, probed with primary antibody specific for UBB. B) Fold change in measured dotblot fluorescence intensity for UBB signal normalized to input dsDNA for the given MGO concentration.

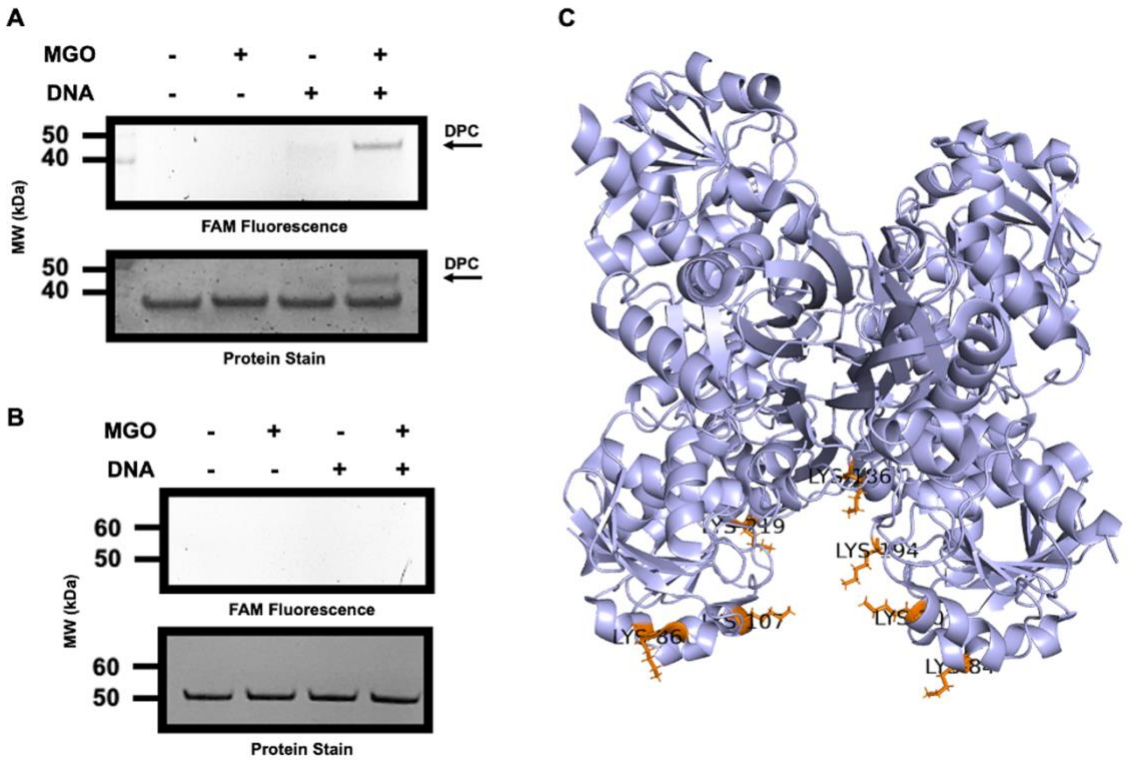


Figure 5.7. GAPDH forms DPCs in the presence of MGO

A) Representative gel of GAPDH (1 μg) incubated with or without FAM labeled ss-telomeric DNA $(TTAGGG)_3$ (25 μM) in the presence or absence of MGO (10 mM) in 25 μL of PBS (pH 7.4) at 37 °C for 1 h; top: FAM fluorescence; bottom: silver stain. B) Representative gel of BSA (1 μg) incubated with or without FAM labeled ss-telomeric DNA $(TTAGGG)_3$ (25 μM) in the presence or absence of MGO (10 mM) in 25 μL of PBS (pH 7.4) at 37 °C for 1 h; top: FAM fluorescence; bottom: simply blue stain. C) Crystal structure of tetrameric GAPDH with possible sites of MGO DPC formation in orange. PBD: 4WNC.

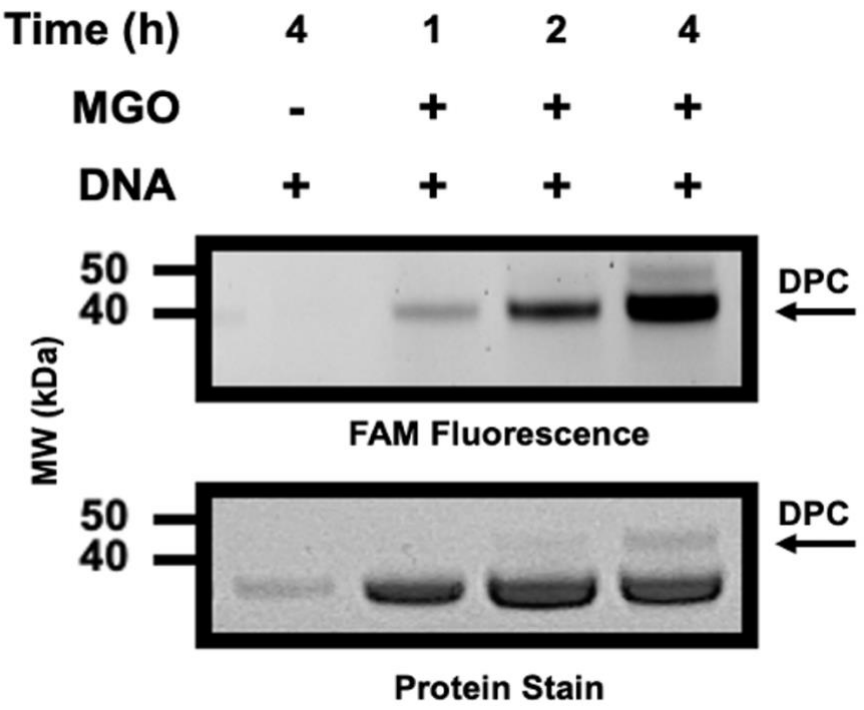


Figure 5.8. Time dependence of MGO induced GAPDH DPC formation

Representative gel of GAPDH (1 μg) incubated with or without FAM labeled ss-telomeric DNA $(TTAGGG)_3$ (25 μM) in the presence or absence of MGO (10 mM) in 25 μL of PBS (pH 7.4) at 37 °C for the specified times; top: FAM fluorescence; bottom: simply blue stain.

Table 5.1. List of enriched MGO induced DPCs

The third column reports the T-test pValue to assess enrichment of DPCs. The fourth column reports the difference between the mean log2 transformed LFQ intensity values between MGO and vehicle treated groups, which is indicative of fold change difference in protein abundance. Data is representative of 3 biological replicates from each treatment condition.

Protein Name	Gene name	Negative Log(P)	Difference (Vehicle - MGO)
heterogeneous nuclear ribonucleoprotein M	HNRNPM	**5.818**	**7.767**
H4 clustered histone 9	H4C9	**3.473**	**6.664**
H2B clustered histone 14	H2BC14	**5.34**	**6.188**
H2A clustered histone 20	H2AC20	**3.733**	**6.13**
STIP1 homology and U-box containing protein 1	STUB1	**4.848**	**6.123**
H2A.Z variant histone 2	H2AZ2	**4.205**	**5.935**
glutaredoxin 3	GLRX3	**5.176**	**5.833**
small ubiquitin like modifier 2	SUMO2	**4.095**	**5.771**
tripartite motif containing 28	TRIM28	**5.678**	**5.711**
flap structure-specific endonuclease 1	FEN1	**3.516**	**5.563**
programmed cell death 11	PDCD11	**2.694**	**5.525**
cofilin 1	CFL1	**2.733**	**5.41**
tubulin beta 6 class V	TUBB6	**2.947**	**5.384**
peptidylprolyl isomerase A	PPIA	**3.29**	**5.338**
AHNAK nucleoprotein	AHNAK	**3.545**	**5.267**
importin 4	IPO4	**4.236**	**5.265**
poly(ADP-ribose) polymerase 1	PARP1	**3.278**	**5.142**
myosin light chain 12A	MYL12A	**3.159**	**5.052**
RAN binding protein 2	RANBP2	**4.306**	**5.018**
GCN1 activator of EIF2AK4	GCN1	**3.791**	**4.981**
general transcription factor IIi	GTF2I	**3.678**	**4.955**
eukaryotic translation initiation factor 5A	EIF5A	**3.252**	**4.831**
ribosomal protein L27	RPL27	**5.255**	**4.683**

DEAD-box helicase 5	DDX5	**3.436**	**4.64**
marker of proliferation Ki-67	MKI67	**3.445**	**4.632**
ribosomal protein L26	RPL26	**4.206**	**4.618**
non-POU domain containing octamer binding	NONO	**3.976**	**4.54**
ribosomal protein L23a	RPL23A	**3.238**	**4.52**
H2A clustered histone 14	H2AC14	**2.28**	**4.514**
peroxiredoxin 1	PRDX1	**4.27**	**4.495**
ribosomal protein L27a	RPL27A	**4.881**	**4.443**
ribosome production factor 2 homolog	RPF2	**4.596**	**4.325**
tubulin alpha 1b	TUBA1B	**2.774**	**4.272**
poly(rC) binding protein 2	PCBP2	**3.152**	**4.27**
heterogeneous nuclear ribonucleoprotein L	HNRNPL	**2.787**	**4.226**
NOP2/Sun RNA methyltransferase 2	NSUN2	**2.612**	**4.22**
proteasome 26S subunit ubiquitin receptor, non-ATPase 2	PSMD2	**4.351**	**4.125**
ribosomal protein S26	RPS26	**2.53**	**4.119**
tubulin beta 4B class IVb	TUBB4B	**4.104**	**4.108**
chromosome segregation 1 like	CSE1L	**4.408**	**4.081**
Ly1 antibody reactive	LYAR	**3.058**	**4.069**
TAR DNA binding protein	TARDBP	**4.219**	**4.056**
pre-mRNA processing factor 8	PRPF8	**1.904**	**4.026**
DnaJ heat shock protein family (Hsp40) member A1	DNAJA1	**5.212**	**4.012**
Tubulin beta-3 chain		**3.85**	**3.988**
poly(rC) binding protein 1	PCBP1	**3.8**	**3.957**
ribosomal protein S7	RPS7	**2.445**	**3.925**
eukaryotic translation initiation factor 4A1	EIF4A1	**3.621**	**3.899**
annexin A2	ANXA2	**1.88**	**3.857**
H3 clustered histone 4	H3C4	**3.634**	**3.788**
ribonuclease/angiogenin inhibitor 1	RNH1	**4.968**	**3.779**
profilin 1	PFN1	**4.804**	**3.712**
tubulin beta class I	TUBB	**4.372**	**3.71**
BRCA2 and CDKN1A interacting protein	BCCIP	**3.478**	**3.707**
EBNA1 binding protein 2	EBNA1BP2	**2.703**	**3.696**

ribosomal protein L32	RPL32	**4.117**	**3.642**
DExD-box helicase 50	DDX50	**2.041**	**3.593**
exportin 1	XPO1	**3.906**	**3.475**
eukaryotic translation elongation factor 2	EEF2	**4.114**	**3.469**
X-ray repair cross complementing 5	XRCC5	**2.357**	**3.462**
RAN, member RAS oncogene family	RAN	**2.739**	**3.457**
adenosine deaminase RNA specific	ADAR	**3.933**	**3.43**
macroH2A.1 histone	MACROH2A1	**2.654**	**3.428**
chloride intracellular channel 1	CLIC1	**3.889**	**3.4**
tubulin alpha 1c	TUBA1C	**3.751**	**3.377**
tubulin beta 2A class IIa	TUBB2A	**4.477**	**3.354**
Heterogeneous nuclear ribonucleoprotein M		**2.528**	**3.329**
importin 7	IPO7	**1.316**	**3.328**
karyopherin subunit beta 1	KPNB1	**2.988**	**3.324**
ribosomal protein S24	RPS24	**4.25**	**3.315**
cell cycle and apoptosis regulator 2	CCAR2	**4.226**	**3.305**
ribosomal protein L35	RPL35	**3.675**	**3.303**
acidic nuclear phosphoprotein 32 family member E	ANP32E	**1.833**	**3.294**
polypyrimidine tract binding protein 1	PTBP1	**1.432**	**3.255**
eukaryotic translation elongation factor 1 gamma	EEF1G	**2.616**	**3.243**
ribosomal protein L10a	RPL10A	**3.212**	**3.239**
heterogeneous nuclear ribonucleoprotein A0	HNRNPA0	**2.519**	**3.229**
DEAD-box helicase 24	DDX24	**2.995**	**3.226**
fatty acid synthase	FASN	**1.962**	**3.189**
PDZ and LIM domain 7	PDLIM7	**4.174**	**3.17**
RBM14-RBM4 readthrough	RBM14-RBM4	**3.29**	**3.133**
filamin A	FLNA	**2.764**	**3.108**
sequestosome 1	SQSTM1	**2.455**	**3.093**
ribosomal protein L10	RPL10	**1.116**	**3.047**
leucyl-tRNA synthetase 1	LARS1	**1.726**	**3.028**
prostaglandin E synthase 3	PTGES3	**1.79**	**2.994**
ribosomal protein S5	RPS5	**2.967**	**2.985**
HEAT repeat containing 1	HEATR1	**1.71**	**2.981**
ribosomal RNA processing 8	RRP8	**4.176**	**2.973**

DNA topoisomerase II beta	TOP2B	**2.303**	**2.957**
H1.10 linker histone	H1-10	**4.445**	**2.954**
ribosomal protein S13	RPS13	**3.78**	**2.95**
GTP binding protein 4	GTPBP4	**3.114**	**2.944**
mutS homolog 6	MSH6	**4.109**	**2.934**
Histone H3-7		**1.741**	**2.932**
RNA binding motif protein X-linked	RBMX	**2.84**	**2.897**
RNA helicase		**1.899**	**2.868**
ribonucleoprotein, PTB binding 1	RAVER1	**2.479**	**2.862**
protein kinase, DNA-activated, catalytic subunit	PRKDC	**5.71**	**2.852**
DEAD-box helicase 47	DDX47	**3.026**	**2.849**
actin alpha 1, skeletal muscle	ACTA1	**0.872**	**2.849**
DExD-box helicase 21	DDX21	**2.098**	**2.81**
splicing factor 3b subunit 1	SF3B1	**1.677**	**2.8**
nucleolar protein interacting with the FHA domain of MKI67	NIFK	**3.026**	**2.795**
heterogeneous nuclear ribonucleoprotein H2	HNRNPH2	**2.776**	**2.792**
minichromosome maintenance complex component 7	MCM7	**3.548**	**2.775**
protein phosphatase, Mg2+/Mn2+ dependent 1G	PPM1G	**1.708**	**2.769**
H3.3 histone B	H3-3B	**3.167**	**2.741**
U2 small nuclear RNA auxiliary factor 2	U2AF2	**2.671**	**2.74**
spliceosome associated factor 3, U4/U6 recycling protein	SART3	**2.678**	**2.709**
ribosomal protein S18	RPS18	**2.751**	**2.703**
peter pan homolog	PPAN	**2.834**	**2.681**
high density lipoprotein binding protein	HDLBP	**3.744**	**2.678**
thymopoietin	TMPO	**2.085**	**2.674**
RecQ like helicase	RECQL	**3.396**	**2.656**
DEAD-box helicase 46	DDX46	**3.038**	**2.613**
DEAH-box helicase 15	DHX15	**4.121**	**2.59**
SDA1 domain containing 1	SDAD1	**3.549**	**2.583**
ribosomal protein S3	RPS3	**2.904**	**2.553**
karyopherin subunit alpha 1	KPNA1	**2.238**	**2.546**
ribosomal protein S4 X-linked	RPS4X	**3.918**	**2.536**
eukaryotic translation initiation factor 4	EIF4G1	**2.664**	**2.518**

gamma 1			
vinculin	VCL	**1.769**	**2.515**
chaperonin containing TCP1 subunit 6A	CCT6A	**2.006**	**2.495**
RuvB like AAA ATPase 1	RUVBL1	**1.536**	**2.493**
ribosomal protein S11	RPS11	**2.399**	**2.484**
heterogeneous nuclear ribonucleoprotein H1	HNRNPH1	**0.946**	**2.482**
nucleolar protein 9	NOL9	**1.855**	**2.479**
ribosomal protein S2	RPS2	**2.368**	**2.439**
DIS3 homolog, exosome endoribonuclease and 3'-5' exoribonuclease	DIS3	**2.134**	**2.432**
small nuclear ribonucleoprotein U1 subunit 70	SNRNP70	**1.855**	**2.431**
transglutaminase 2	TGM2	**2.956**	**2.413**
ubiquitin C	UBC	**3.851**	**2.412**
FA complementation group I	FANCI	**1.711**	**2.407**
interferon gamma inducible protein 16	IFI16	**2.108**	**2.387**
ribosomal protein S25	RPS25	**1.053**	**2.372**
Mtr4 exosome RNA helicase	MTREX	**3.058**	**2.329**
Rho/Rac guanine nucleotide exchange factor 2	ARHGEF2	**1.557**	**2.32**
tubulin alpha 1a	TUBA1A	**1.103**	**2.307**
N-acetyltransferase 10	NAT10	**1.629**	**2.286**
Heterogeneous nuclear ribonucleoprotein C-like 4		**1.285**	**2.278**
DEAD-box helicase 54	DDX54	**1.932**	**2.258**
Putative heat shock protein HSP 90-beta 2		**2.279**	**2.252**
dynein axonemal assembly factor 5	DNAAF5	**3.578**	**2.23**
talin 1	TLN1	**3.751**	**2.23**
proteasome 26S subunit, ATPase 1	PSMC1	**3.223**	**2.227**
NOP56 ribonucleoprotein	NOP56	**2.28**	**2.199**
cell division cycle 73	CDC73	**1.846**	**2.191**
myosin light chain 6	MYL6	**3.194**	**2.19**
U2 small nuclear RNA auxiliary factor 1	U2AF1	**1.996**	**2.168**
tripartite motif containing 25	TRIM25	**4.814**	**2.167**
Putative 60S ribosomal protein L39-like 5		**3.331**	**2.127**
leucine rich repeat containing 59	LRRC59	**1.611**	**2.114**
protein phosphatase 1 catalytic subunit beta	PPP1CB	**3.463**	**2.097**
ribosomal protein L18a	RPL18A	**1.753**	**2.088**

small nuclear ribonucleoprotein D1 polypeptide	SNRPD1	**2.392**	**2.087**
karyopherin subunit alpha 2	KPNA2	**1.968**	**2.074**
polybromo 1	PBRM1	**2.759**	**2.072**
translocase of outer mitochondrial membrane 34	TOMM34	**4.173**	**2.065**
importin 5	IPO5	**3.609**	**2.063**
heat shock protein family A (Hsp70) member 8	HSPA8	**4.845**	**2.062**
glutamine-fructose-6-phosphate transaminase 2	GFPT2	**1.184**	**2.05**
tubulin alpha 4a	TUBA4A	**1.674**	**2.038**
Heterogeneous nuclear ribonucleoprotein A1		**3.069**	**2.032**
symplekin scaffold protein	SYMPK	**2.218**	**2.012**
triosephosphate isomerase 1	TPI1	**3.528**	**2**
testis expressed 10	TEX10	**1.889**	**1.994**
unc-45 myosin chaperone A	UNC45A	**2.465**	**1.988**
DNA topoisomerase I	TOP1	**2.711**	**1.984**
ribosomal protein L13	RPL13	**1.275**	**1.976**
DnaJ heat shock protein family (Hsp40) member A2	DNAJA2	**2.63**	**1.973**
UTP18 small subunit processome component	UTP18	**2.287**	**1.97**
coiled-coil domain containing 47	CCDC47	**0.974**	**1.963**
ubiquitin protein ligase E3 component n-recognin 5	UBR5	**2.39**	**1.96**
coilin	COIL	**2.548**	**1.96**
SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	SMARCA5	**1.868**	**1.956**
Putative elongation factor 1-alpha-like 3		**3.715**	**1.954**
RNA binding motif protein 25	RBM25	**1.087**	**1.935**
A-kinase anchoring protein 8	AKAP8	**1.003**	**1.929**
heterogeneous nuclear ribonucleoprotein A3	HNRNPA3	**1.177**	**1.921**
ribosomal protein S15	RPS15	**1.154**	**1.918**
ribosomal RNA processing 1B	RRP1B	**2.146**	**1.91**
ribosomal protein L7 like 1	RPL7L1	**2.007**	**1.875**
exosome component 10	EXOSC10	**2.509**	**1.868**
biogenesis of ribosomes BRX1	BRIX1	**2.346**	**1.864**

X-ray repair cross complementing 1	XRCC1	**2.435**	**1.841**
DR1 associated protein 1	DRAP1	**1.235**	**1.839**
nuclear cap binding protein subunit 1	NCBP1	**1.724**	**1.804**
glyceraldehyde-3-phosphate dehydrogenase	GAPDH	**1.037**	**1.793**
DEAD-box helicase 10	DDX10	**1.93**	**1.788**
fibrillarin	FBL	**1.132**	**1.787**
ribosomal L1 domain containing 1	RSL1D1	**2.596**	**1.783**
PDS5 cohesin associated factor A	PDS5A	**1.672**	**1.775**
heat shock protein family A (Hsp70) member 1A	HSPA1A	**2.093**	**1.767**
oxysterol binding protein like 8	OSBPL8	**3.657**	**1.767**
chromodomain helicase DNA binding protein 4	CHD4	**1.892**	**1.759**
tropomyosin 4	TPM4	**1.069**	**1.735**
basic leucine zipper and W2 domains 1	BZW1	**2.207**	**1.731**
DNA topoisomerase II alpha	TOP2A	**1.341**	**1.728**
THO complex subunit 2	THOC2	**2.625**	**1.714**
RNA binding motif protein 34	RBM34	**1.425**	**1.708**
dyskerin pseudouridine synthase 1	DKC1	**1.731**	**1.706**
copine 1	CPNE1	**4.681**	**1.69**
uveal autoantigen with coiled-coil domains and ankyrin repeats	UACA	**1.991**	**1.689**
chaperonin containing TCP1 subunit 4	CCT4	**1.989**	**1.681**
DExH-box helicase 9	DHX9	**1.496**	**1.681**
DExD-box helicase 52	DDX52	**2.474**	**1.663**
muscleblind like splicing regulator 1	MBNL1	**3.409**	**1.64**
splicing factor proline and glutamine rich	SFPQ	**2.558**	**1.624**
pre-mRNA processing factor 3	PRPF3	**1.772**	**1.623**
nuclear autoantigenic sperm protein	NASP	**2.53**	**1.62**
WD repeat domain 1	WDR1	**1.454**	**1.602**
cullin associated and neddylation dissociated 1	CAND1	**2.452**	**1.601**
5'-3' exoribonuclease 2	XRN2	**1.448**	**1.588**
WD repeat domain 3	WDR3	**1.293**	**1.582**
ribosomal RNA processing 12 homolog	RRP12	**2.708**	**1.579**
ribosomal protein L12	RPL12	**3.118**	**1.575**
serum amyloid A like 1	SAAL1	**2.266**	**1.571**

tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein theta	YWHAQ	**3.251**	**1.57**
H1.0 linker histone	H1-0	**1.532**	**1.569**
ribosomal protein L24	RPL24	**3.864**	**1.56**
KRR1 small subunit processome component homolog	KRR1	**1.571**	**1.539**
peptidylprolyl isomerase D	PPID	**1.404**	**1.539**
Aly/REF export factor	ALYREF	**1.661**	**1.535**
histone deacetylase 2	HDAC2	**1.177**	**1.53**
A-kinase anchoring protein 8 like	AKAP8L	**1.153**	**1.503**
dynein cytoplasmic 1 heavy chain 1	DYNC1H1	**2.031**	**1.497**
dpy-30 histone methyltransferase complex regulatory subunit	DPY30	**2.56**	**1.494**
replication protein A1	RPA1	**1.234**	**1.484**
nucleolar protein 10	NOL10	**1.516**	**1.45**
small nuclear ribonucleoprotein polypeptide N	SNRPN	**1.573**	**1.447**
nucleolar complex associated 4 homolog	NOC4L	**1.615**	**1.443**
eukaryotic translation initiation factor 5	EIF5	**1.209**	**1.44**
four and a half LIM domains 2	FHL2	**3.891**	**1.422**
jumonji domain containing 6, arginine demethylase and lysine hydroxylase	JMJD6	**2.386**	**1.41**
myosin heavy chain 9	MYH9	**3.69**	**1.376**
lactate dehydrogenase A	LDHA	**2.077**	**1.375**
SUMO specific peptidase 3	SENP3	**2.999**	**1.352**
small ubiquitin like modifier 3	SUMO3	**1.918**	**1.336**
ribosomal protein L7a	RPL7A	**2.583**	**1.333**
heterogeneous nuclear ribonucleoprotein A2/B1	HNRNPA2B1	**2.24**	**1.293**
heterogeneous nuclear ribonucleoprotein H3	HNRNPH3	**1.704**	**1.289**
nucleolar protein 11	NOL11	**1.467**	**1.286**
ribosomal protein L7	RPL7	**2.48**	**1.282**
mitogen-activated protein kinase kinase kinase kinase 4	MAP4K4	**1.667**	**1.276**
transportin 1	TNPO1	**1.442**	**1.271**
guanine monophosphate synthase	GMPS	**2.079**	**1.264**
DEAD-box helicase 18	DDX18	**1.491**	**1.258**
DExH-box helicase 30	DHX30	**1.569**	**1.217**

general transcription factor IIIC subunit 1	GTF3C1	**2.072**	**1.206**
ribosomal protein L18	RPL18	**2.21**	**1.186**
FMR1 autosomal homolog 1	FXR1	**2.073**	**1.148**
ribosomal protein S17	RPS17	**3.132**	**1.142**
ribosome biogenesis regulator 1 homolog	RRS1	**2.271**	**1.138**
CCAAT enhancer binding protein zeta	CEBPZ	**2.797**	**1.115**
ribosomal protein S14	RPS14	**2.189**	**1.105**
ribosomal protein L21	RPL21	**2.194**	**1.101**
heat shock protein 90 alpha family class B member 1	HSP90AB1	**3.129**	**1.078**
URB1 ribosome biogenesis homolog	URB1	**1.742**	**1.072**
ribosomal protein L8	RPL8	**3.334**	**1.034**
NOP58 ribonucleoprotein	NOP58	**2.588**	**1.027**
ribosomal protein L14	RPL14	**1.85**	**1.012**
myosin phosphatase Rho interacting protein	MPRIP	**1.98**	**1.01**
RNA binding motif protein 26	RBM26	**2.215**	**1.005**
ribosomal protein L28	RPL28	**2.299**	**0.918**

Table 5.2. Comparison of DPCs identified in HT1080 cells upon MGO, cisplatin, diepoxybutane (DEB), or phosphoramide mustard (PM) treatment

Data compiled from the following refs: [320, 354, 363]

Uniprot accession number	Protein	Crosslinker
B2RPK0	Putative high mobility group protein B1-like 1	Cisplatin
O00148	ATP-dependent RNA helicase DDX39	Cisplatin
O00571	ATP-dependent RNA helicase DDX3X	Cisplatin
O00622	Protein CYR61	Cisplatin
O14979	Heterogeneous nuclear ribonucleoprotein D-like	Cisplatin
O43707	α-Actinin-4	Cisplatin
O60506	Heterogeneous nuclear ribonucleoprotein Q	Cisplatin
O60812	Heterogeneous nuclear ribonucleoprotein C-like 1	Cisplatin
O94880	PHD finger protein 14	Cisplatin
O95602	DNA-directed RNA polymerase I subunit RPA1	Cisplatin
P02545	Prelamin-A/C	Cisplatin
P02751	Fibronectin	Cisplatin
P07910	Heterogeneous nuclear ribonucleoproteins C1/C2	Cisplatin
P62888	60S ribosomal protein L30	Cisplatin
P05783	Keratin, type I cytoskeletal 18	Cisplatin
P05787	Keratin, type II cytoskeletal 8	Cisplatin
P06576	ATP synthase subunit β, mitochondrial	Cisplatin
P07237	Protein disulfide-isomerase	Cisplatin
P07437	Tubulin β chain	Cisplatin
P07900	Heat shock protein HSP 90-α	Cisplatin
Q96A08	Histone H2B type 1-A	Cisplatin
P09382	Galectin-1	Cisplatin

P09429	High mobility group protein B1 (HMG B1)	Cisplatin
P09661	U2 small nuclear ribonucleoprotein A′	Cisplatin
P10809	60 kDa heat shock protein, mitochondrial	Cisplatin
P11021	78 kDa glucose-regulated protein OS=Homo	Cisplatin
P12814	α-Actinin-1	Cisplatin
P13667	Protein disulfide-isomerase A4	Cisplatin
P13796	Plastin-2	Cisplatin
P14618	Pyruvate kinase isozymes M1/M2	Cisplatin
P15311	Ezrin	Cisplatin
P17844	Probable ATP-dependent RNA helicase DDX5	Cisplatin
P18085	ADP-ribosylation factor 4	Cisplatin
P18621	60S ribosomal protein L17	Cisplatin
P18669	Phosphoglycerate mutase 1	Cisplatin
P20700	Lamin-B1	Cisplatin
P21333	Filamin-A	Cisplatin
P27797	Calreticulin	Cisplatin
P27824	Calnexin	Cisplatin
P31943	Heterogeneous nuclear ribonucleoprotein H	Cisplatin
P35580	Myosin-10	Cisplatin
P36578	60S ribosomal protein L4	Cisplatin
P37802	Transgelin-2	Cisplatin
P39023	60S ribosomal protein L3	Cisplatin
P40429	60S ribosomal protein L13a	Cisplatin
P43307	Translocon-associated protein subunit α	Cisplatin
P46777	60S ribosomal protein L5	Cisplatin
P46782	40S ribosomal protein S5	Cisplatin
P49207	60S ribosomal protein L34	Cisplatin
P51149	Ras-related protein Rab-7a	Cisplatin
P52272	Heterogeneous nuclear ribonucleoprotein M	Cisplatin
P55209	Nucleosome assembly protein 1-like 1	Cisplatin
P61026	Ras-related protein Rab-10	Cisplatin
P61247	40S ribosomal protein S3a	Cisplatin
P61313	60S ribosomal protein L15	Cisplatin
P61513	60S ribosomal protein L37a	Cisplatin

P61927	60S ribosomal protein L37	Cisplatin
P62136	Serine/threonine-protein phosphatase PP1-α catalytic subunit	Cisplatin
P62495	Eukaryotic peptide chain release factor subunit 1	Cisplatin
P62736	Actin, aortic smooth muscle	Cisplatin
P62910	60S ribosomal protein L32	Cisplatin
P62913	60S ribosomal protein L11	Cisplatin
P68104	Elongation factor 1-α 1 (EF-1α1)	Cisplatin
P78371	T-complex protein 1 subunit β	Cisplatin
P84077	ADP-ribosylation factor 1	Cisplatin
Q00839	Heterogeneous nuclear ribonucleoprotein U	Cisplatin
Q01082	Spectrin β chain, brain 1	Cisplatin
Q01518	Adenylyl cyclase-associated protein 1	Cisplatin
Q02878	60S ribosomal protein L6	Cisplatin
Q08170	Splicing factor, arginine/serine-rich 4	Cisplatin
Q08945	FACT complex subunit SSRP1	Cisplatin
Q12905	Interleukin enhancer-binding factor 2	Cisplatin
Q12906	Interleukin enhancer-binding factor 3	Cisplatin
Q13427	Peptidyl-prolyl cis-trans isomerase G	Cisplatin
Q13813	Spectrin α chain, brain	Cisplatin
Q13823	Nucleolar GTP-binding protein 2	Cisplatin
Q14137	Ribosome biogenesis protein BOP1	Cisplatin
Q15185	Prostaglandin E synthase 3	Cisplatin
Q15287	RNA-binding protein with serine-rich domain 1	Cisplatin
Q15424	Scaffold attachment factor B1	Cisplatin
Q58FF7	Putative heat shock protein HSP 90-β-3	Cisplatin
Q5SSJ5	Heterochromatin protein 1-binding protein 3	Cisplatin
Q68CQ4	Digestive organ expansion factor homolog	Cisplatin
Q8IY81	Putative rRNA methyltransferase 3	Cisplatin
Q8N7H5	RNA polymerase II-associated factor 1 homolog	Cisplatin
Q8NHW5	60S acidic ribosomal protein P0-like	Cisplatin
Q92841	Probable ATP-dependent RNA helicase DDX17	Cisplatin
Q92922	SWI/SNF complex subunit SMARCC1	Cisplatin
Q96GQ7	Probable ATP-dependent RNA helicase DDX27	Cisplatin

Q96L21	60S ribosomal protein L10-like	Cisplatin
Q96MU7	YTH domain-containing protein 1	Cisplatin
Q99613	Eukaryotic translation initiation factor 3 subunit C	Cisplatin
Q99988	Growth/differentiation factor 15	Cisplatin
Q9BV38	WD repeat-containing protein 18	Cisplatin
Q9H307	Pinin	Cisplatin
Q9HCG8	Pre-mRNA-splicing factor CWC22 homolog	Cisplatin
Q9NQZ2	Something about silencing protein 10	Cisplatin
Q9NVI7	ATPase family AAA domain-containing protein 3A	Cisplatin
Q9NW13	RNA-binding protein 28	Cisplatin
Q9UH99	SUN domain-containing protein 2	Cisplatin
Q9UIG0	Tyrosine-protein kinase BAZ1B	Cisplatin
Q9UKV3	Apoptotic chromatin condensation inducer in the nucleus	Cisplatin
Q9Y3U8	60S ribosomal protein L36	Cisplatin
Q9Y4L1	Hypoxia up-regulated protein 1	Cisplatin
Q9Y4W2	Protein LAS1 homolog	Cisplatin
P07196	Neurofilament light polypeptide	Cisplatin
P78316	Nucleolar protein 14	Cisplatin
Q9UQ88	Cell division protein kinase 11A	Cisplatin
P15144	Aminopeptidase N	Cisplatin
O60841	Eukaryotic translation initiation factor 5B	Cisplatin, DEB
O75494	Splicing factor, arginine/serine-rich 10	Cisplatin, DEB
P09493	Tropomyosin α-1 chain	Cisplatin, DEB
Q14103	Heterogeneous nuclear ribonucleoprotein D0	Cisplatin, DEB
P14314	Glucosidase 2 subunit β	Cisplatin, DEB
P14625	Endoplasmin	Cisplatin, DEB
P16403	Histone H1D	Cisplatin, DEB
Q96QV6	Histone H2A type 1-A	Cisplatin, DEB
P19338	Nucleolin (C-23)	Cisplatin, DEB
P26583	High mobility group protein B2 (HMG B2)	Cisplatin, DEB
Q9UKM9	RNA-binding protein Raly	Cisplatin, DEB
P51531	Probable global transcription activator SNF2L2	Cisplatin, DEB
P61978	Heterogeneous nuclear ribonucleoprotein K	Cisplatin, DEB
P62241	40S ribosomal protein S8	Cisplatin, DEB

P84098	60S ribosomal protein L19	Cisplatin, DEB
Q03252	Lamin-B2	Cisplatin, DEB
Q13185	Chromobox protein homolog 3	Cisplatin, DEB
Q13243	Splicing factor, arginine/serine-rich 5	Cisplatin, DEB
Q13435	Splicing factor 3B subunit 2	Cisplatin, DEB
Q14978	Nucleolar and coiled-body phosphoprotein 1	Cisplatin, DEB
Q15061	WD repeat-containing protein 43	Cisplatin, DEB
Q58FF3	Putative endoplasmin-like protein	Cisplatin, DEB
Q8IYB3	Serine/arginine repetitive matrix protein 1	Cisplatin, DEB
Q8IZL8	Proline-, glutamic acid- and leucine-rich protein 1	Cisplatin, DEB
Q99729	Heterogeneous nuclear ribonucleoprotein A/B	Cisplatin, DEB
Q99733	Nucleosome assembly protein 1-like 4	Cisplatin, DEB
Q9Y2W1	Thyroid hormone receptor-associated protein 3	Cisplatin, DEB
P06748	Nucleophosmin	Cisplatin, DEB
P12270	Nucleoprotein TPR	Cisplatin, DEB
P17096	High mobility group protein HMG-I/HMG-Y	Cisplatin, DEB
P17480	Nucleolar transcription factor 1	Cisplatin, DEB
P35659	Protein DEK	Cisplatin, DEB
P52926	High mobility group protein HMGA2	Cisplatin, DEB
P62995	Transformer-2 protein homolog beta	Cisplatin, DEB
Q16629	Splicing factor, arginine/serine-rich 7	Cisplatin, DEB
Q86VM9	Zinc finger CCCH domain-containing protein 18	Cisplatin, DEB
Q9Y3B9	RRP15-like protein	Cisplatin, DEB
O00566	U3 small nucleolar ribonucleoprotein protein MPP10	Cisplatin, DEB, PM
O00541	Pescadillo homolog	Cisplatin, PM
O75643	U5 small nuclear ribonucleoprotein 200 kDa helicase	Cisplatin, PM
P05198	Eukaryotic translation initiation factor 2 subunit 1	Cisplatin, PM
P12956	X-ray repair cross-complementing protein 6 (XRCC-6 or Ku70)	Cisplatin, PM
P04083	Annexin A1	Cisplatin, PM
Q9BQG0	Myb-binding protein 1A	Cisplatin, PM
Q9BXP5	Serrate RNA effector molecule homolog	Cisplatin, PM

Q9BXY0	Protein MAK16 homolog	Cisplatin, PM
Q9Y3T9	Nucleolar complex protein 2 homolog	Cisplatin, PM
Q9Y5B9	FACT complex subunit SPT16	Cisplatin, PM
P23396	40S ribosomal protein S3	Cisplatin, PM
P25705	ATP synthase subunit α, mitochondrial	Cisplatin, PM
P31946	14-3-3 protein β/α	Cisplatin, PM
P32969	60S ribosomal protein L9	Cisplatin, PM
P33778	Histone H2B type 1-B	Cisplatin, PM
P46087	Putative ribosomal RNA methyltransferase NOP2	Cisplatin, PM
P51532	Transcription activator BRG1	Cisplatin, PM
P60709	Actin, cytoplasmic 1 (β-actin)	Cisplatin, PM
P62244	40S ribosomal protein S15a	Cisplatin, PM
P62249	40S ribosomal protein S16	Cisplatin, PM
P62258	14-3-3 Protein epsilon	Cisplatin, PM
P62266	40S ribosomal protein S23	Cisplatin, PM
P62847	40S ribosomal protein S24	Cisplatin, PM
Q02543	60S ribosomal protein L18a	Cisplatin, PM
Q14692	Ribosome biogenesis protein BMS1 homolog	Cisplatin, PM
Q15029	116 kDa U5 small nuclear ribonucleoprotein component	Cisplatin, PM
Q1KMD3	Heterogeneous nuclear ribonucleoprotein U-like protein 2	Cisplatin, PM
Q32P51	Heterogeneous nuclear ribonucleoprotein A1-like 2	Cisplatin, PM
Q5BKZ1	Zinc finger protein 326	Cisplatin, PM
P08670	Vimentin	Cisplatin, PM
Q969G3	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily E member 1	Cisplatin, PM
P08865	40S ribosomal protein SA	Cisplatin, PM, DEB
Q9NY61	Protein AATF	Cisplatin, PM, DEB
Q9UKD2	mRNA turnover protein 4 homolog	Cisplatin, PM, DEB
P46781	40S ribosomal protein S9	Cisplatin, PM, DEB
P55081	Microfibrillar-associated protein 1	Cisplatin, PM,

		DEB
P62753	40S ribosomal protein S6	Cisplatin, PM, DEB
P62829	60S ribosomal protein L23	Cisplatin, PM, DEB
P84103	Splicing factor, arginine/serine-rich 3	Cisplatin, PM, DEB
Q07065	Cytoskeleton-associated protein 4	Cisplatin, PM, DEB
Q13765	Nascent polypeptide-associated complex subunit α	Cisplatin, PM, DEB
Q16643	Drebrin	Cisplatin, PM, DEB
P05556	Integrin β-1	Cisplatin, PM, DEB
P43243	Matrin-3	Cisplatin, PM, DEB
Q07955	Splicing factor, arginine/serine-rich 1	Cisplatin, PM, DEB
O00592	Podocalyxin	DEB
O43290	U4/U6.U5 tri-snRNP-associated protein 1	DEB
O75475	PC4 and SFRS1-interacting protein	DEB
O75526	RNA-binding motif protein, X-linked-like-2	DEB
O75683	Surfeit locus protein 6	DEB
P05387	60S acidic ribosomal protein P2	DEB
P06753	Isoform TM30nm of Tropomyosin alpha-3 chain	DEB
P07951	Tropomyosin beta chain	DEB
P07996	Thrombospondin-1	DEB
P12273	Prolactin-inducible protein	DEB
P14923	Junction plakoglobin	DEB
P16070	CD44 antigen	DEB
P17096-2	Isoform HMGA1b of High mobility group protein HMG-I/HMG-Y	DEB
P23497	Nuclear autoantigen Sp-100	DEB
P25440	Bromodomain-containing protein 2	DEB
P29692	Elongation factor 1-delta	DEB
P35269	General transcription factor IIF subunit 1	DEB
P35613	Basigin	DEB
P35637	RNA-binding protein FUS	DEB

P45973	Chromobox protein homolog 5	DEB
P47914	60S ribosomal protein L29	DEB
P50502	Hsc70-interacting protein	DEB
P61244	Protein max	DEB
P62807	Histone H2B type 1-C/E/F/G/I	DEB
P62988	Ubiquitin	DEB
P67809	Nuclease-sensitive element-binding protein 1	DEB
P67936	Tropomyosin alpha-4 chain	DEB
Q13442	28 kDa heat- and acid-stable phosphoprotein	DEB
P83916	Chromobox protein homolog 1	DEB
Q01105	Protein SET	DEB
Q01130	Splicing factor, arginine/serine-rich 2	DEB
Q01469	Fatty acid-binding protein, epidermal	DEB
Q02413	Desmoglein-1	DEB
Q02539	Histone H1.1	DEB
Q08554	Desmocollin-1	DEB
Q13242	Splicing factor, arginine/serine-rich 9	DEB
Q13428	Treacle protein	DEB
Q14011	Cold-inducible RNA-binding protein	DEB
Q14444	Caprin-1	DEB
Q14919	Dr1-associated corepressor	DEB
Q15059	Bromodomain-containing protein 3	DEB
Q15149	Plectin-1	DEB
Q16695	Histone H3.1t	DEB
Q3SYE8	Putative high mobility group protein B3-like-1	DEB
Q5QJE6	Deoxynucleotidyltransferase terminal-interacting protein 2	DEB
Q6NZI2	Polymerase I and transcript release factor	DEB
Q7L014	Probable ATP-dependent RNA helicase DDX46	DEB
Q86U42	Polyadenylate-binding protein 2	DEB
Q8IZQ5	Selenoprotein H	DEB
Q8NC51	Plasminogen activator inhibitor 1 RNA-binding protein	DEB
Q8WXX5	DnaJ homolog subfamily C member 9	DEB
Q92688	Acidic leucine-rich nuclear phosphoprotein 32 family member B	DEB
Q96DR8	Mucin-like protein 1	DEB

Q96ST2	Protein IWS1 homolog	DEB
Q9BRL6	Splicing factor, arginine/serine-rich 2B	DEB
Q9H1E3	Nuclear ubiquitous casein and cyclin-dependent kinases substrate	DEB
Q9H930	Nuclear body protein SP140-like protein	DEB
Q9UMY1	Nucleolar protein 7	DEB
Q9UQ35	Serine/arginine repetitive matrix protein 2	DEB
Q9Y5S9	RNA-binding protein 8A	DEB
O95218	Zinc finger Ran-binding domain-containing protein 2, Isoform 1	DEB
P46821	Microtubule-associated protein 1B	DEB
Q7Z4V5	Hepatoma-derived growth factor-related protein 2, Isoform 1	DEB
Q8TF01	Splicing factor, arginine/serine-rich 18, Isoform 1	DEB
Q9NYF8	Bcl-2-associated transcription factor 1, Isoform 1	DEB
A0A087WTP3	Heterogeneous nuclear ribonucleoprotein M	MGO
A0A087WYT3	prostaglandin E synthase 3	MGO
A0A087WZ13	ribonucleoprotein, PTB binding 1	MGO
A0A087X0H9	RNA binding motif protein 26	MGO
A0A087X0X3	heterogeneous nuclear ribonucleoprotein M	MGO
A0A0B4J269	Tubulin beta-3 chain	MGO
A0A0C4DG89	DEAD-box helicase 46	MGO
A0A0C4DGG9	chromodomain helicase DNA binding protein 4	MGO
A0A0G2JPF8	Heterogeneous nuclear ribonucleoprotein C-like 4	MGO
A0A1X7SBZ2	RNA helicase	MGO
A0A2R8YGX3	tropomyosin 4	MGO
A8MV53	peter pan homolog	MGO
B2WTI3	jumonji domain containing 6, arginine demethylase and lysine hydroxylase	MGO
C9JP00	muscleblind like splicing regulator 1	MGO
D3YTB1	ribosomal protein L32	MGO
E7EMW7	ubiquitin protein ligase E3 component n-recognin 5	MGO
E7ENQ1	mitogen-activated protein kinase kinase kinase kinase 4	MGO
E7ETK0	ribosomal protein S24	MGO

E7EU85	FMR1 autosomal homolog 1	MGO
E9PL09	ribosomal protein S3	MGO
E9PQX9	DR1 associated protein 1	MGO
F5H5D3	tubulin alpha 1c	MGO
F5H8D7	X-ray repair cross complementing 1	MGO
F8VZX2	poly(rC) binding protein 2	MGO
F8W1R7	myosin light chain 6	MGO
F8W6I7	Heterogeneous nuclear ribonucleoprotein A1	MGO
F8W7R3	FA complementation group I	MGO
G3V529	DEAD-box helicase 24	MGO
G8JLB6	heterogeneous nuclear ribonucleoprotein H1	MGO
H0Y5B5	polybromo 1	MGO
H0YNH8	uveal autoantigen with coiled-coil domains and ankyrin repeats	MGO
H7BXY3	DExH-box helicase 30	MGO
J3KTA4	DEAD-box helicase 5	MGO
J3QLE5	small nuclear ribonucleoprotein polypeptide N	MGO
J3QLI9	small nuclear ribonucleoprotein D1 polypeptide	MGO
J3QRI7	ribosomal protein L26	MGO
K7EK07	H3.3 histone B	MGO
M0R0F0	ribosomal protein S5	MGO
M0R3D6	ribosomal protein L18a	MGO
O00299	chloride intracellular channel 1	MGO
O00410	importin 5	MGO
O14980	exportin 1	MGO
O15355	protein phosphatase, Mg2+/Mn2+ dependent 1G	MGO
O43143	DEAH-box helicase 15	MGO
O43159	ribosomal RNA processing 8	MGO
O43395	pre-mRNA processing factor 3	MGO
O43823	A-kinase anchoring protein 8	MGO
O60264	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 5	MGO
O60287	URB1 ribosome biogenesis homolog	MGO
O60832	dyskerin pseudouridine synthase 1	MGO
O60884	DnaJ heat shock protein family (Hsp40)	MGO

	member A2	
O75083	WD repeat domain 1	MGO
O75533	splicing factor 3b subunit 1	MGO
O76003	glutaredoxin 3	MGO
O94808	glutamine-fructose-6-phosphate transaminase 2	MGO
P07737	profilin 1	MGO
P0CG48	ubiquitin C	MGO
P0DMV9	heat shock protein family A (Hsp70) member 1A	MGO
P11388	DNA topoisomerase II alpha	MGO
P13489	ribonuclease/angiogenin inhibitor 1	MGO
P18206	vinculin	MGO
P21980	transglutaminase 2	MGO
P22087	fibrillarin	MGO
P26368	U2 small nuclear RNA auxiliary factor 2	MGO
P26641	eukaryotic translation elongation factor 1 gamma	MGO
P27348	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein theta	MGO
P27694	replication protein A1	MGO
P31689	DnaJ heat shock protein family (Hsp40) member A1	MGO
P33993	minichromosome maintenance complex component 7	MGO
P38432	coilin	MGO
P39748	flap structure-specific endonuclease 1	MGO
P42166	thymopoietin	MGO
P42285	Mtr4 exosome RNA helicase	MGO
P42696	RNA binding motif protein 34	MGO
P42766	ribosomal protein L35	MGO
P46013	marker of proliferation Ki-67	MGO
P46063	RecQ like helicase	MGO
P49321	nuclear autoantigenic sperm protein	MGO
P49327	fatty acid synthase	MGO
P49792	RAN binding protein 2	MGO
P49915	guanine monophosphate synthase	MGO
P50991	chaperonin containing TCP1 subunit 4	MGO
P52292	karyopherin subunit alpha 2	MGO

P52294	karyopherin subunit alpha 1	MGO
P52701	mutS homolog 6	MGO
P55010	eukaryotic translation initiation factor 5	MGO
P55060	chromosome segregation 1 like	MGO
P55265	adenosine deaminase RNA specific	MGO
P55795	heterogeneous nuclear ribonucleoprotein H2	MGO
P55854	small ubiquitin like modifier 3	MGO
P60174	triosephosphate isomerase 1	MGO
P61956	small ubiquitin like modifier 2	MGO
P62140	protein phosphatase 1 catalytic subunit beta	MGO
P62191	proteasome 26S subunit, ATPase 1	MGO
P62854	ribosomal protein S26	MGO
P62906	ribosomal protein L10a	MGO
P68133	actin alpha 1, skeletal muscle	MGO
P68366	tubulin alpha 4a	MGO
P68431	H3 clustered histone 4	MGO
P78347	general transcription factor IIi	MGO
P78527	protein kinase, DNA-activated, catalytic subunit	MGO
Q00341	high density lipoprotein binding protein	MGO
Q01081	U2 small nuclear RNA auxiliary factor 1	MGO
Q01780	exosome component 10	MGO
Q02880	DNA topoisomerase II beta	MGO
Q08752	peptidylprolyl isomerase D	MGO
Q08J23	NOP2/Sun RNA methyltransferase 2	MGO
Q09161	nuclear cap binding protein subunit 1	MGO
Q12789	general transcription factor IIIC subunit 1	MGO
Q13148	TAR DNA binding protein	MGO
Q13151	heterogeneous nuclear ribonucleoprotein A0	MGO
Q13200	proteasome 26S subunit ubiquitin receptor, non-ATPase 2	MGO
Q13206	DEAD-box helicase 10	MGO
Q13263	tripartite motif containing 28	MGO
Q13501	sequestosome 1	MGO
Q13601	KRR1 small subunit processome component homolog	MGO
Q14192	four and a half LIM domains 2	MGO
Q14204	dynein cytoplasmic 1 heavy chain 1	MGO

Q14258	tripartite motif containing 25	MGO
Q14684	ribosomal RNA processing 1B	MGO
Q14690	programmed cell death 11	MGO
Q14974	karyopherin subunit beta 1	MGO
Q15020	spliceosome associated factor 3, U4/U6 recycling protein	MGO
Q15233	non-POU domain containing octamer binding	MGO
Q15785	translocase of outer mitochondrial membrane 34	MGO
Q16666	interferon gamma inducible protein 16	MGO
Q16777	H2A clustered histone 20	MGO
Q29RF7	PDS5 cohesin associated factor A	MGO
Q59GN2	Putative 60S ribosomal protein L39-like 5	MGO
Q5JP53	tubulin beta class I	MGO
Q5SY16	nucleolar protein 9	MGO
Q5TEC6	Histone H3-7	MGO
Q5VTE0	Putative elongation factor 1-alpha-like 3	MGO
Q60FE5	filamin A	MGO
Q6DKI1	ribosomal protein L7 like 1	MGO
Q6P1J9	cell division cycle 73	MGO
Q6P2Q9	pre-mRNA processing factor 8	MGO
Q6WCQ1	myosin phosphatase Rho interacting protein	MGO
Q71UI9	H2A.Z variant histone 2	MGO
Q7L1Q6	basic leucine zipper and W2 domains 1	MGO
Q86V81	Aly/REF export factor	MGO
Q86VP6	cullin associated and neddylation dissociated 1	MGO
Q86Y56	dynein axonemal assembly factor 5	MGO
Q8N163	cell cycle and apoptosis regulator 2	MGO
Q8NI27	THO complex subunit 2	MGO
Q8TDN6	biogenesis of ribosomes BRX1	MGO
Q8TEX9	importin 4	MGO
Q92616	GCN1 activator of EIF2AK4	MGO
Q92769	histone deacetylase 2	MGO
Q92797	symplekin scaffold protein	MGO
Q92973	transportin 1	MGO
Q92974	Rho/Rac guanine nucleotide exchange factor 2	MGO

Q96AG4	leucine rich repeat containing 59	MGO
Q96ER3	serum amyloid A like 1	MGO
Q96PK6	RBM14-RBM4 readthrough	MGO
Q99829	copine 1	MGO
Q99878	H2A clustered histone 14	MGO
Q99879	H2B clustered histone 14	MGO
Q9BQ39	DExD-box helicase 50	MGO
Q9BSC4	nucleolar protein 10	MGO
Q9BTT0	acidic nuclear phosphoprotein 32 family member E	MGO
Q9BVI4	nucleolar complex associated 4 homolog	MGO
Q9BYG3	nucleolar protein interacting with the FHA domain of MKI67	MGO
Q9BZE4	GTP binding protein 4	MGO
Q9BZF1	oxysterol binding protein like 8	MGO
Q9C005	dpy-30 histone methyltransferase complex regulatory subunit	MGO
Q9H0A0	N-acetyltransferase 10	MGO
Q9H0D6	5'-3' exoribonuclease 2	MGO
Q9H3U1	unc-45 myosin chaperone A	MGO
Q9H4L4	SUMO specific peptidase 3	MGO
Q9H7B2	ribosome production factor 2 homolog	MGO
Q9H8H0	nucleolar protein 11	MGO
Q9NR12	PDZ and LIM domain 7	MGO
Q9NVU7	SDA1 domain containing 1	MGO
Q9NX58	Ly1 antibody reactive	MGO
Q9NXF1	testis expressed 10	MGO
Q9P287	BRCA2 and CDKN1A interacting protein	MGO
Q9P2J5	leucyl-tRNA synthetase 1	MGO
Q9ULX6	A-kinase anchoring protein 8 like	MGO
Q9UNE7	STIP1 homology and U-box containing protein 1	MGO
Q9UNX4	WD repeat domain 3	MGO
Q9Y2L1	DIS3 homolog, exosome endoribonuclease and 3'-5' exoribonuclease	MGO
Q9Y2R4	DExD-box helicase 52	MGO
Q9Y490	talin 1	MGO
Q9Y5J1	UTP18 small subunit processome component	MGO
X1WI28	ribosomal protein L10	MGO

O75367	macroH2A.1 histone	MGO, Cisplatin
O95373	importin 7	MGO, Cisplatin
P07355	annexin A2	MGO, Cisplatin
P08238	heat shock protein 90 alpha family class B member 1	MGO, Cisplatin
P09874	poly(ADP-ribose) polymerase 1	MGO, Cisplatin
P11387	DNA topoisomerase I	MGO, Cisplatin
P13010	X-ray repair cross complementing 5	MGO, Cisplatin
P13639	eukaryotic translation elongation factor 2	MGO, Cisplatin
P14866	heterogeneous nuclear ribonucleoprotein L	MGO, Cisplatin
P23528	cofilin 1	MGO, Cisplatin
P26373	ribosomal protein L13	MGO, Cisplatin
P26599	polypyrimidine tract binding protein 1	MGO, Cisplatin
P31942	heterogeneous nuclear ribonucleoprotein H3	MGO, Cisplatin
P38159	RNA binding motif protein X-linked	MGO, Cisplatin
P40227	chaperonin containing TCP1 subunit 6A	MGO, Cisplatin
P46779	ribosomal protein L28	MGO, Cisplatin
P50914	ribosomal protein L14	MGO, Cisplatin
P51991	heterogeneous nuclear ribonucleoprotein A3	MGO, Cisplatin
P61353	ribosomal protein L27	MGO, Cisplatin
P62269	ribosomal protein S18	MGO, Cisplatin
P62424	ribosomal protein L7a	MGO, Cisplatin
P62701	ribosomal protein S4 X-linked	MGO, Cisplatin
P62826	RAN, member RAS oncogene family	MGO, Cisplatin
P62917	ribosomal protein L8	MGO, Cisplatin
P62937	peptidylprolyl isomerase A	MGO, Cisplatin
P63241	eukaryotic translation initiation factor 5A	MGO, Cisplatin
P68371	tubulin beta 4B class IVb	MGO, Cisplatin
Q08211	DExH-box helicase 9	MGO, Cisplatin
Q15365	poly(rC) binding protein 1	MGO, Cisplatin
Q58FF8	Putative heat shock protein HSP 90-beta 2	MGO, Cisplatin
Q5JTH9	ribosomal RNA processing 12 homolog	MGO, Cisplatin
Q9BUF5	tubulin beta 6 class V	MGO, Cisplatin
Q9H0S4	DEAD-box helicase 47	MGO, Cisplatin
Q9H583	HEAT repeat containing 1	MGO, Cisplatin
Q9NVP1	DEAD-box helicase 18	MGO, Cisplatin
Q9Y265	RuvB like AAA ATPase 1	MGO, Cisplatin

O00567	NOP56 ribonucleoprotein	MGO, Cisplatin, DEB
P04406	glyceraldehyde-3-phosphate dehydrogenase	MGO, Cisplatin, DEB
P23246	splicing factor proline and glutamine rich	MGO, Cisplatin, DEB
P35579	myosin heavy chain 9	MGO, Cisplatin, DEB
P62750	ribosomal protein L23a	MGO, Cisplatin, DEB
P62805	H4 clustered histone 9	MGO, Cisplatin, DEB
P83731	ribosomal protein L24	MGO, Cisplatin, DEB
Q06830	peroxiredoxin 1	MGO, Cisplatin, DEB
Q07020	ribosomal protein L18	MGO, Cisplatin, DEB
Q13885	tubulin beta 2A class IIa	MGO, Cisplatin, DEB
Q71U36	tubulin alpha 1a	MGO, Cisplatin, DEB
Q9Y2X3	NOP58 ribonucleoprotein	MGO, Cisplatin, DEB
O76021	ribosomal L1 domain containing 1	MGO, Cisplatin, PM
P08708	ribosomal protein S17	MGO, Cisplatin, PM
P11142	heat shock protein family A (Hsp70) member 8	MGO, Cisplatin, PM
P15880	ribosomal protein S2	MGO, Cisplatin, PM
P30050	ribosomal protein L12	MGO, Cisplatin, PM
P46776	ribosomal protein L27a	MGO, Cisplatin, PM
P46778	ribosomal protein L21	MGO, Cisplatin, PM
P60842	eukaryotic translation initiation factor 4A1	MGO, Cisplatin, PM
P62081	ribosomal protein S7	MGO, Cisplatin, PM
P62277	ribosomal protein S13	MGO, Cisplatin,

		PM
P62280	ribosomal protein S11	MGO, Cisplatin, PM
Q03701	CCAAT enhancer binding protein zeta	MGO, Cisplatin, PM
P18124	ribosomal protein L7	MGO, Cisplatin, PM, DEB
P22626	heterogeneous nuclear ribonucleoprotein A2/B1	MGO, Cisplatin, PM, DEB
P62263	ribosomal protein S14	MGO, Cisplatin, PM, DEB
Q9NR30	DExD-box helicase 21	MGO, Cisplatin, PM, DEB
P07305	H1.0 linker histone	MGO, DEB
P19105	myosin light chain 12A	MGO, DEB
P49756	RNA binding motif protein 25	MGO, DEB
P62851	ribosomal protein S25	MGO, DEB
Q09666	AHNAK nucleoprotein	MGO, DEB
P00338	lactate dehydrogenase A	MGO, PM
P62841	ribosomal protein S15	MGO, PM
P68363	tubulin alpha 1b	MGO, PM
Q04637	eukaryotic translation initiation factor 4 gamma 1	MGO, PM
Q15050	ribosome biogenesis regulator 1 homolog	MGO, PM
Q8TDD1	DEAD-box helicase 54	MGO, PM
Q92522	H1.10 linker histone	MGO, PM
Q96A33	coiled-coil domain containing 47	MGO, PM
P08621	small nuclear ribonucleoprotein U1 subunit 70	MGO, PM, DEB
Q99848	EBNA1 binding protein 2	MGO, PM, DEB
A0LTM5	Cell division protein 11B	PM
O00203	AP-3 complex subunit beta-1	PM
O14786	Neuropilin-1	PM
O15371	Eukaryotic translation initiation factor 3 subunit D	PM
O60216	Double-strand-break repair protein rad21 homolog	PM
O75400	Pre-mRNA-processing factor 40 homolog A	PM
P11201	78 kDa glucose-regulated protein	PM
P17661	Desmin	PM

P20042	Eukaryotic translation initiation factor 2 subunit 2-like protein	PM
P23284	Peptidyl-prolyl cis-trans isomerase B	PM
P25205	DNA replication licensing factor MCM3	PM
P36776	Lon protease homolog, mitochondrial	PM
P46783	40S ribosomal protein S10	PM
P51665	26S proteasome non-ATPase regulatory subunit 7	PM
P54652	Heat shock-related 70 kDa protein 2	PM
P55884	Eukaryotic translation initiation factor 3 subunit B	PM
P63104	14-3-3 protein zeta/delta	PM
P63173	60S ribosomal protein L38	PM
P66905	Hemoglobin subunit alpha	PM
P83881	60S ribosomal protein L36a	PM
Q00610	Clathrin heavy chain 1	PM
Q08378	Golgin subfamily A member 3	PM
Q12873	Chromodomain-helicase-DNA-binding protein 3	PM
Q13268	Dehydrogenase/reductase SDR family member 2, mitochondrial	PM
Q14152	Eukaryotic translation initiation factor 3 subunit A	PM
Q14257	Reticulocalbin-2	PM
Q14320	Protein FAM50A	PM
Q14566	DNA replication licensing factor MCM6	PM
Q14839	Chromodomain-helicase-DNA-binding protein 4	PM
Q14980	Nuclear mitotic apparatus protein 1	PM
Q15006	Tetratricopeptide repeat protein 35	PM
Q15393	Splicing factor 3B subunit 3	PM
Q16352	Alpha-internexin	PM
Q3TIV5	Zinc finger CCCH domain-containing protein 15	PM
Q5JNZ5	Putative 40S ribosomal protein S26-like 1	PM
Q5TAP6	U3 small nucleolar RNA-associated protein 14 homolog A	PM
Q6NW1	Putative 60S ribosomal protein L13a-like MGC87657	PM
Q8WTT2	Nucleolar complex protein 3 homolog	PM

Q96FZ7	Charged multivesicular body protein 6	PM
Q99623	Prohibitin-2	PM
P04083	26S proteasome non-ATPase regulatory subunit 3	PM
Q9NQ39	Putative 40S ribosomal protein S10-like	PM
Q9NZM5	Glioma tumor suppressor candidate region gene 2 protein	PM
Q9P0M6	Core histone macro-H2A.2	PM
Q9P2E9	Ribosome-binding protein 1	PM
Q9UBS4	DnaJ homolog subfamily B member 11	PM
T1W3K1	60S ribosomal protein L17	PM
P62979	40S ribosomal protein S27a	PM
O15347	High mobility group protein B3	PM, DEB
P05109	Protein S100 A8	PM, DEB
P06702	Protein S100 A9	PM, DEB
P35268	60S ribosomal protein L22	PM, DEB
P39687	Acidic leucine-rich nuclear phosphoprotein 32 family member A	PM, DEB
P41219	Peripherin	PM, DEB
P53999	Activated RNA polymerase II transcriptional coactivator p15	PM, DEB
P60866	40S ribosomal protein S20	PM, DEB
P62899	60S ribosomal protein L31	PM, DEB
Q13283	Ras GTPase-activating protein-binding protein 1	PM, DEB
Q8WVC0	RNA polymerase-associated protein LEO1	PM, DEB
P16989	DNA-binding protein A	PM, DEB

5.4 Discussion

Studying the molecular mechanisms of how reactive small molecules effect cellular phenotype is critical for understanding how they contribute to overall health and disease susceptibility.[371] In this study, we investigated the formation of DNA-protein crosslinks formed upon exposure of human cell culture models to MGO. MGO is an important electrophile linked to many diseases and can participate in a variety of chemical reactions with biomolecules due to its electrophilic oxo-aldehyde moiety.[372] MGO forms both protein and DNA adducts which can lead to inhibition of enzymatic activity, confirmational alterations, cell signaling events, depurination, and formation of promutagenic abasic sites.[290, 373, 374]

Recently, there has been an effort to characterize the complex array of AGE adducts elicited by MGO.[375] For example, an emerging suite of chemical tools for enriching and identifying MGO adducts are being developed, along with strategies to profile MGO induced protein-protein crosslinks, reversible MGO-protein modifications, and MGO protein-metabolite crosslinks.[221, 296, 297, 309, 376-379] Additionally, previous studies have shown MGO exposure can form DNA protein crosslink conjugates within *in-vitro* models. These models were constructed using synthetic or recombinant systems to model MGO crosslinking potential.[285, 317, 346] However, no system-wide investigation of the proteins participating in formation of MGO DPCs has been completed in cell culture models. We took advantage of the K-SDS assay and the phenol:chloroform methodology to quantify and isolate DPCs from MGO treated cells. By coupling DPC isolation methods with mass spectrometry, proteins participating in DPC formation could be

readily identified and the crosslink sites could be characterized. While it should be noted that high 5 mM concentrations of MGO were used in the study, this approach reveals candidate DPCs for more detailed functional analysis to assess MGO cellular activity.

Through the K-SDS assay, we established that MGO produced DPCs in a dose dependent manner. It was also observed that knockout of either GLO1 or GLO2, genes which detoxify MGO, also significantly increased DPC levels upon MGO exposure. These results suggest that MGO derived DPCs may be relevant adducts in aging and diseases associated with elevated MGO levels like diabetes, chronic renal disease, cancer, and Alzheimer's disease.[380-383] Intriguingly, studies investigating differences in cisplatin derived DPCs formed between ER- and ER+ breast cancer cells found that the protein PRDX1 only formed DPCs in ER- cells, suggesting DPCs may be dependent of cell line.[384] Notably, we also found PRDX1 to be a MGO DPC. Thus, further investigation aimed at quantifying MGO derived DPCs across various disease states and cell types is warranted.

We observed that human fibrosarcoma HT1080 cells exposed to 5 mM MGO led to formation of DNA-protein crosslinks that were heterogeneous in size and structure. A majority of these 265 cellular proteins that exhibited crosslinking characteristics were known to be localized in the nucleus, nuclear chromosomes, telomere DNA regions, cellular membranes, and ribosomes. These identified proteins participate in nucleosome assembly (eg. Histones, SMARCA5, NASP, START3), telomere organization (eg. XRRCC5, PARP1, Histones, RPA1), mRNA processing (eg. Helicase proteins, heterogeneous nuclear ribonucleoproteins, pre-mRNA processing factors) and DNA

repair. DNA repair proteins over-represented upon MGO exposure include BCCIP, FANCI, PDS5A, RECQL, RUVBL1, STUB1, SMARCA5, FEN1, MSH6, NONO, PARP1, RPA1, RPS3, TRIM28, UBR5, and ubiquitin. These DNA repair proteins may be either directly crosslinked to the DNA or represent the DNA damage response upon exposure to elevated levels of MGO.

Profiling of proteins participating in DNA protein crosslinking have revealed differences in activity of crosslinking agents. Previous mass spectrometry-based studies of proteins participating in DPC formation upon exposure to cisplatin, diepoxybutane, and phosphoramide mustard have revealed many DNA-binding proteins including GAPDH, histones, PARP1, TOP1, and XRCC1 to exhibit DNA crosslinking characteristics in the presence of electrophilic species. 30% of proteins participating in DPC formation in response to MGO treatment were previously observed in HT1080 cells exposed to cisplatin, diepoxybutane and phosphoramide mustard. Observed differences in proteins involved in DPCs may be due to differences in crosslinking mechanisms and site selectivity. This points to the continued need to profile DPCs formed by various crosslinking agents. A side by side comparison between DPCs generated by other endogenously produced electrophiles such as glyoxal and formaldehyde would provide further insight to this area.[385, 386]

Although selected proteins such as histone H3A, PARP1, XRCC1, GAPDH, VINC and TOP1 exhibit a MGO dose dependent association with DNA via dotblot, this approach is limited by a lack of structural characterization of the covalent crosslink site

and calls for additional analysis to provide direct evidence for crosslinking, which we performed with GAPDH.

It is well documented that GAPDH plays numerous roles beyond catalyzing the conversion of glyceraldehyde-3-phosphate (G3P) to 1,3-biphosphoglycerate in the glycolytic pathway.[387-389] One of these roles includes GAPDH's recognition and binding of telomeric DNA sequences.[368] This GAPDH – DNA interaction protects telomeres against rapid degradation in response to chemotherapeutic agents in human lung cancer cells, suggesting an important biological role for GAPDH – DNA interactions.[370] Our *in vitro* validation of MGO induced GAPDH DNA protein crosslinking provides confidence in our DPC extraction methodology and however MS characterization would provide additional structural information on the location the crosslinking site. These findings suggests that GAPDH MGO derived DPCs effect telomere organization, which was a process also found in our bioinformatic analysis. Thus, this crosslinking may play an additional role in the association between MGO and ageing processes linked to telomeres, however more detailed studies are needed to clarify this hypothesis.[344, 390, 391] It should be noted that MGO also covalently modifies GAPDH, leading to enzymatic inhibition and alteration in isoelectric point.[392] In neural precursor cells (NPCs), it was found that MGO modification of GAPDH impacted Notch signaling which effected NPC homeostasis.[393] These results, coupled with our discovery that GAPDH is a target for MGO mediated DNA-protein crosslinking, points to the complexity of GAPDH biology.

Elucidating the repair mechanisms of DPCs is an active area of research.[394] Our results showed that knockdown of the SPRTN gene, which repairs DPCs, leads to

elevated DPC levels upon 2.5 and 5 mM MGO treatment. Thus it is likely that SPRTN plays a role in recognizing and removing MGO derived DPCs, which is in line with previous reports finding SPRTN to be involved in repairing DPCs generated through formaldehyde.[395] Additionally, the observation that MGO treatment induces a dose dependent increase in UBB signal in isolated DPCs, coupled with finding the small ubiquitin-like modifier (SUMO) protein to be enriched in our proteomic dataset suggests that MGO derived DPCs are activating repair pathways.[353] Based on previous studies, some of the possible repair mechanisms that could be involved in removing these MGO crosslinks include nucleotide excision repair (NER),[396] homologous recombination (HR),[397] and other proteolytic processes.[398] Future work will be needed to fully elucidate the predominant mechanisms for how MGO derived DPCs are repaired.

In conclusion, this system-wide study demonstrates DNA protein crosslinks are formed in human fibrosarcoma HT1080 cells following exposure to methylglyoxal. In these experiments, MGO exposure led to crosslinking of over 260 proteins to chromosomal DNA. These proteins were identified using mass spectrometry-based proteomics approaches and many crosslinked proteins were confirmed by immunoblotting. We identified that GAPDH crosslinks to telomere DNA. Ongoing studies in our laboratory include using analytical standards to study the formation and repair of methylglyoxal DPCs in the context of diabetes, neurological disorders, and cancer as well as interrogate the functional ramifications of MGO-GAPDH DPCs and any mutagenetic potential these bulky DNA lesions possess. Ultimately, the work here

provides a basis for studying MGO derived DPCs to better understand mechanisms of MGO toxicity and MGO cell signaling events relevant in human disease.

6 Chapter VI: Small Molecule Inhibitors of TET Dioxygenases: Bobcat339 Activity is Mediated by Contaminating Copper (II)

Adapted in part with permission by the American Chemical Society from:

Weirath N.A., Hurben A.K., Chao C., Pujari S.S., Cheng T., Liu S., Tretyakova N.Y.* Small Molecule Inhibitors of TET Dioxygenases: Bobcat339 Activity is Mediated by Contaminating Copper (II). *ACS Medicinal Chemistry Letters*. 2022, *13* (5), 792–798.

This chapter is adopted form a manuscript published in *ACS Medicinal Chemistry Letters* with contributions from Alexander K. Hurben (*in vitro* studies and LC-MS experiments), Nick A. Weirath (*in vitro* studies and LC-MS experiments), Chris Chao (synthesis), Suresh S. Pujari (synthesis), Tao Cheng (cell assays and dot blotting), and Shujun Liu (cell assays and dot blotting), under the supervision of Prof. Tretyakova.

6.1 Introduction

The DNA epigenetic mark 5-methylcytosine (5mC) is required for regulating gene expression levels in mammalian cells.[399] DNA methylation establishes stable gene expression patterns, allows for cell differentiation, and controls cell identity. Methylation of CpG islands within gene promotor regions is typically associated with gene silencing, while the demethylated state correlates with active transcription.[400] DNA methyltransferases (DNMTs) are responsible for installing and maintaining cytosine methylation marks, while Ten-eleven translocation (TET) methyl-cytosine dioxygenases (TET1, TET2, and TET3) are indirectly responsible for their removal through the iterative oxidation of the methyl group. TET enzymes are Fe(II) and α-ketoglutarate dependent enzymes that convert 5mC to 5-hydroxymethyl-C (hmC) and further to 5-formyl-C (fC) and 5-carboxy-C (caC) (**Figure 6.1**).[401] The oxidized forms of 5mC block maintenance methylation by DNMT1, contributing to passive DNA demethylation.[402] Additionally, these oxidized forms of 5mC are recognized by the DNA repair machinery to be replaced with unmodified cytosine, leading to active DNA demethylation.[403] Given their role in the maintenance of DNA methylation states, TET enzymes are central players in transcriptional regulation.

TET dioxygenases are non-heme Fe(II) enzymes that employ the α-ketoglutarate cofactor to split molecular oxygen and insert it between the 5-methyl carbon and hydrogen proximal to the Fe center via a concerted two-step reaction wherein the hydrogen is abstracted from the target carbon followed by hydroxyl rebound to the substrate (**Figure 6.2**).[404, 405] TET enzymes are capable of further oxidizing 5hmC to 5fC

and 5caC.[7] The TET family of proteins consists of three closely related isoforms: TET1, TET2, and TET3. While each of the three TET enzymes has varying substrate preferences, expression levels, and catalytic activity, the catalytic domain of all three TET family members is highly conserved.[406]

While the essential function of TET proteins in cell development and differentiation is well established, there is limited consensus regarding their role in human disease.[407] TET enzymes are often downregulated or mutated in cancer, leading to decreased levels of hmC and fC.[408, 409] For example, Lemonnier and colleagues demonstrated that TET2 mutations are responsible for significantly decreased levels of 5hmC in Tcell lymphoma.[410] Ko and colleagues obtained similar results for TET2 mutations in myeloid leukemia.[411] TET2 knockout mice are more susceptible toward myeloid malignancies.[412, 413] It was therefore hypothesized that TET enzymes act as tumor suppressor genes, and their dysfunction prevents the removal of aberrant DNA methylation marks.[409, 414-416]

While the above evidence supports a protective role of TET proteins in human cancer, increased levels of hmC are found in hepatoblastomas due to elevated TET activity.[417] TET1 is often overexpressed in cases of mixed lineage leukemias (MLLs) and some breast cancers.[418, 419] Moreover, clinical studies revealed that TET2 mutations are predictive of increased MLL patient survival.[420, 421] Therefore, it is possible that in some cases, TET enzymes are essential for creating the epigenetic landscape necessary for tumor development and growth. In these instances, inhibiting TET could be a viable therapeutic strategy. Indeed, TET2 deficient myeloid neoplasms have been selectively

targeted with TET inhibitors, leading to synthetic lethality. [422] Such promising results call for further exploration in this area.

The availability of small molecule inhibitors of TET is key to elucidating their role in human cell function and may lead to novel therapeutic strategies for cancer. However, potent and selective inhibitors remain elusive. Current methods of modulating TET activity broadly rely on mimicking the α-ketoglutarate cofactor.[423] For example, N-oxalylglycine, which is capable of displacing α-ketoglutarate from the cofactor site, inhibits TET enzymes at low mM concentrations.[424] These classes of inhibitors, however, are not specific for TET but are instead broad-spectrum inhibitors of many α-ketoglutarate dependent enzymes. Significant work by C. Schofield and colleagues has demonstrated the utility of these class of molecules as inhibitors of hypoxia-inducible factor (HIF) prolyl hydroxylases (PHDs), another class of α-ketoglutarate dependent enzymes.[425-427] These biological probes provide utility for structural investigations and *in vitro* interrogations of TET enzymes, but have limited utility for *in vivo* studies where upward of 60 of these enzymes could be modulated, making it difficult to discern physiological changes directly associated with TET modulation.[424]

A few structurally distinct TET inhibitors have also been explored. In silico screening of natural product libraries against TET identified C35, a catechol containing small molecule, as a potential inhibitor.[428] However, based on its structure, C35 likely has limited utility for future drug design due to its structural classification as a pan-assay interfering compound (PAIN) and its metabolic liability.[429, 430]

More recently, additional analogues of α-ketoglutarate with greater complexity have shown promising selectivity and potency for TET isoforms.[422] Other scaffolds, however, aim to incorporate motifs that access the substrate binding pocket within the conserved TET catalytic domain. Following this approach, Chua and colleagues reported a series of cytosine based inhibitors, so named the Bobcat series, with a lead compound exhibiting activity against TET1 and TET2 at 33 and 73 μM, respectively.[431] Their series of inhibitors were prepared via copper-catalyzed Chan-Lam coupling of various aromatic systems to 5-chloro cytosine.

With the goal of probing the biological role of TET proteins, we synthesized and characterized the lead compound (Bobcat339, **Figure 6.3**) in our laboratory. Using the Bobcat339 preparation from our laboratory that was carefully purified by HPLC to remove any residual Cu(II), we failed to observe any inhibition of recombinant human TET1 or TET2 using a quantitative mass spectrometry-based assay. Further exploration of in-house-synthesized (referred to throughout as house-Bobcat339) and commercially available (referred to throughout as "manufacturer"-Bobcat339) sources of Bobcat339 was done to elucidate the reasons underpinning this discrepancy in activity. Herein we report that trace Cu(II) greatly enhances Bobcat339's inhibitory activity, which tells a cautionary tale of ensuring that potential small molecule inhibitors are rigorously purified and characterized prior to biological evaluation.

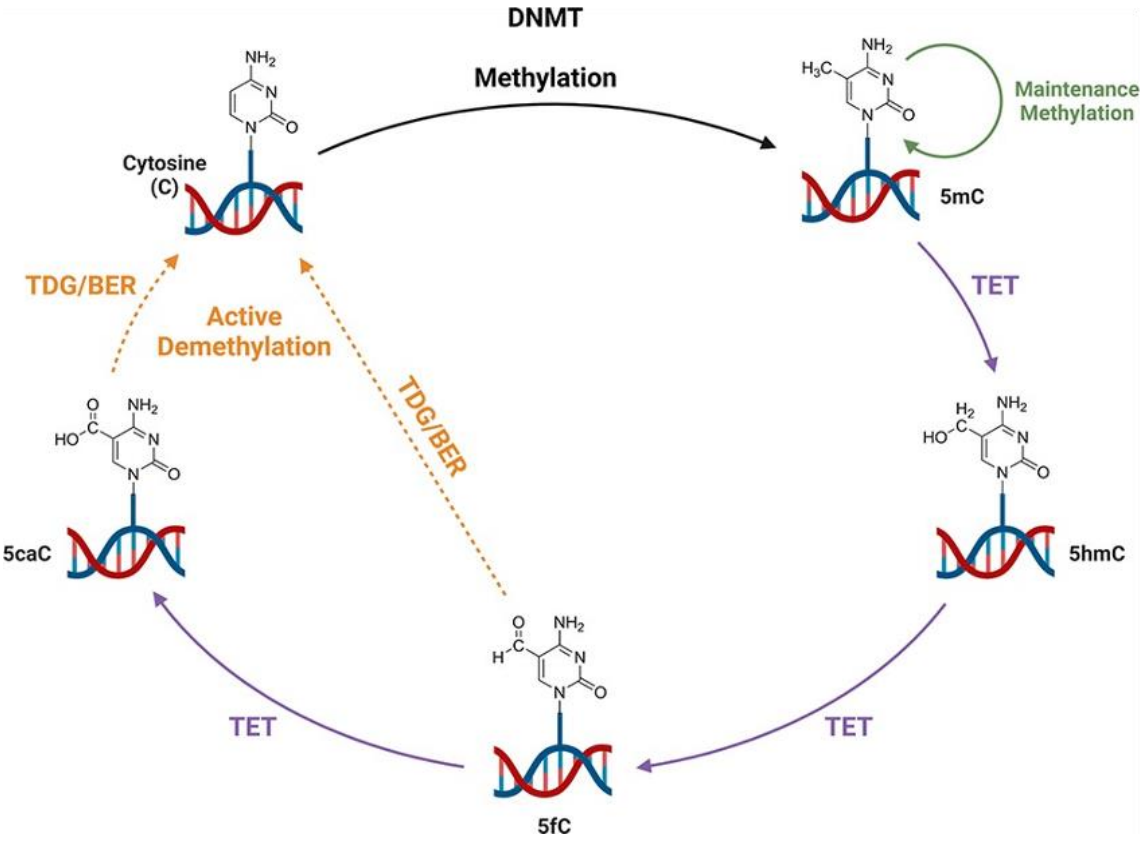


Figure 6.1. TET-mediated conversion of 5mC to 5hmC, required for regulating gene expression levels in 5fC, and 5caCa mammalian cells

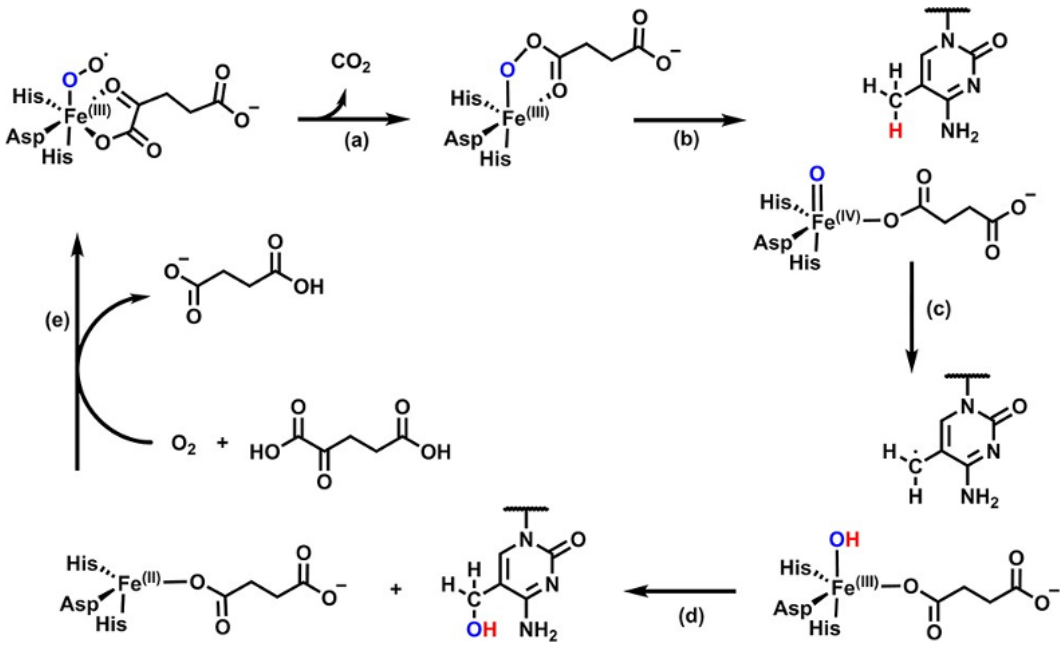


Figure 6.2. Catalytic cycle of TET dioxygenases

(a) Decarboxylation of the α-ketoglutarate cofactor, (b) heterolysis of the dioxygen bridge, (c) hydrogen abstraction, (d) OH- rebound, and (e) expulsion of succinate and regeneration of the active site.

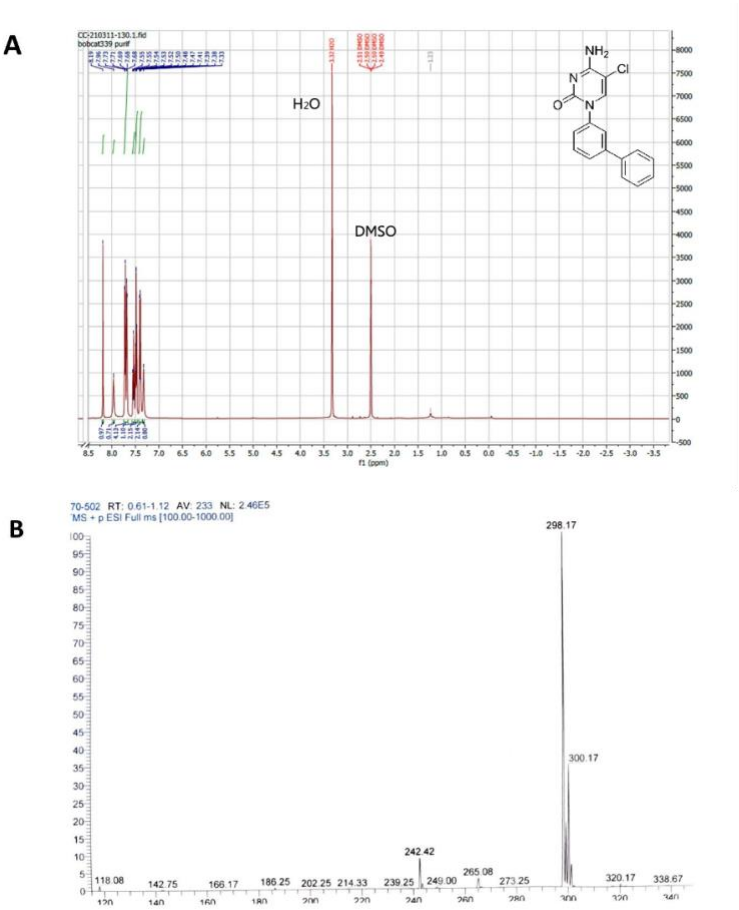


Figure 6.3. Characterization of house-synthesized Bobcat339

A) ^{1}H-NMR of Bobcat339 in DMSO. B) LCMS of Bobcat339 in positive mode.

6.2 Materials and Methods

All chemicals and reagents required for the synthesis, purification, and assaying of Bobcat339 were purchased from Sigma Aldrich unless otherwise stated. NMR and LCMS solvents were obtained from Fisher Scientific. Hep3B cells used for dot blotting were obtained from ATCC. All materials for tissue culture and related experiments were obtained from Corning and Gibco. DNA extraction kits used for dot blotting were obtained from Qiagen. TET1 (#31417) and TET2 (#31418) enzymes were purchased from Active Motif. Synthetic DNA oligomer substrates were ordered from Integrated DNA Technologies. Internal standards for LCMS quantitation of 5mC and 5hmC were obtained from Toronto Research Chemicals. Nucleotide digestion buffer was supplied by New England Biolabs (#M0649S). Filters used for purification of nucleotides were acquired from Pall Corporation and the aldehyde reactive probe used in 5fC derivatization was sourced from Cayman Chemical. Licor Image Studio was used to generate process acquired dot blotting images and all statistical analyses herein were performed with GraphPad Prism v9.3.1.

Synthesis of Bobcat339:

To a 5-chlorocytosine (200 mg, 1.37 mmol) dissolved in MeOH: H_2O (5 mL, 3:1, v/v), 3biphenylboronic acid (410 mg, 2.06 mmol), $Cu(OAc)_2.H_2O$ (275 mg, 1.37 mmol) were added. TEMED (320 mg, 413 µL, 2.57 mmol) was added dropwise to the reaction mixture and stirred at room temperature for 4 hours. The reaction mixture turned blue in color. The solvent was evaporated under high vacuum and applied for flash chromatography (FC). The compound was eluted at 20% MeOH in DCM as a dark blue

solid. This compound was further washed with water repeatedly to remove the $Cu(OAc)_2$ from the compound. A pale blue color was still observable in the compound. This compound was again purified twice by FC to obtain the colorless compound (70 mg) and further purified by HPLC (Synergi Hydro RP 80A, 250x10 mm; Flow: 2.5 mL/min.; Buffer A: Water; Buffer B: ACN; 5% – 95% buffer B over 50 minutes). ^{1}H-NMR: (DMSO-*d*6, 500MHz) δ 8.19 (s, 1H), 7.96 (br s, 1H), 7.71-7.73 (d, 2H), 7.68-7.69 (m, 2H), 7.52-7.55 (m, 1H), 7.47-7.50 (m, 2H), 7.38-7.41 (m, 2H), 7.32 (br s, 1H). LRMS (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{16}H_{13}ClN_3O$ Calc. 298.1. Obtained. 298.1.

ICP-OES Quantitation of Copper:

A Thermo Scientific iCAP 7600 ICP-OES Duo in conjunction with a CETAC ASX-560 Autosampler was used for this analysis. Copper was measured at 324.754 nm in the Axial plasma view. Samples were submitted as 100 µM solutions of Bobcat339 in 1 mL aliquots and diluted 1:8 prior to analysis to remain within the linear range of the instrument. Calibration standards were prepared from a multi-element custom standard stock (SCP Science, Quebec, Canada). Reported sample concentrations are the result of three technical replicate readings.

In-vitro TET Assay:

For *in vitro* TET assays, 10 pmol of 16-mer hemi-methylated synthetic dsDNA oligomer designed after the *Rassf1a* tumor suppressor gene promotor (5'-AATTAGAA[5meC]GCTCCTT -3') were incubated with 2xTET reaction buffer (50 mM HEPES [pH 7.9], 100 mM NaCl, 75 µM $Fe(NH_4)_2(SO_4)_2$, 2 mM ascorbate, 1 mM α-ketoglutarate, 1 mM DTT), 0.2 µg (2.3 pmol, 115 nM) recombinant human TET1 or

TET2 catalytic domain, and 5% DMSO (or inhibitor in DMSO) on a 20 µL scale. Timing was achieved by withholding TET1 protein until the 'start' of the incubation window. Reactions were completed on a dry block at 37 °C for 30 minutes unless otherwise indicated. Reactions were terminated at the 'end' of the incubation window by transferring reaction tubes to dry ice to snap-freeze the reaction until all replicates were completed. Upon completion, all samples were heat-inactivated at 90 °C for 10 minutes and then dried entirely in SpeedVac concentrator.

HPLC-ESI-MS/MS Quantitation of 5hmC:

Quantitation of 5hmC is based on methodology previously described by our group in Seiler et al. with modification.[432] Dried enzymatic reaction samples were reconstituted with mQ-water and spiked with 1 pmol each of 5-Methyl-2'-deoxycytidine-d_3 and 5-(Hydroxymethyl)-2'deoxycytidine-d_3. Nucleotide digestion enzyme mix and buffer were added per product guidelines on a 20 µL scale. Samples were digested for 60 min at 37 °C on a dry block. Following digestion, samples were filtered on nanosep 10K Omega filters. For samples being assessed for 5fC, fC was derivatized with 4 mM O-(biotinylcarbazoylmethyl) hydroxylamine in ammonium acetate buffer (pH 4.5) with 100 mM aniline for 24 hours at room temperature as previously reported. Hydrolysates were transferred to MS vials, dried entirely, and resuspended in 100 µL mQ-water for offline HPLC cleanup to enrich 5hmC. An Atlantis T3 column (*Waters*, 4.6 x 150 mm, 3 µm) was eluted at a flow of 0.9 mL/min with a linear gradient previously described of 5 mM NH_4CHO_2 buffer, pH 4.0 (A) and MeOH (B). Fractions containing dC, 5meC, and 5hmC were collected and pooled before analysis by isotope dilution HPLC-ESI-MS/MS. Pooled

samples were dried and reconstituted in 2 mM ammonium formate buffer (NH_4CHO_2) with a Dionex Ultimate 3000UHPLC (*Thermo Fisher*) interfaced with a Thermo TSQ Quantiva mass spectrometer (*Thermo Fisher*) flowing 2 mM NH_4CHO_2 (A) and ACN (B) at 15 µL/min through a Zorbax SB-C18 column (*Agilent*, 0.5 x 150 mm, 3 µm) with the gradients as previously described. 5meC and 5-d_3-meC eluted at 5.1 minutes while 5hmC and 5-d_3-hmC eluted at 3.5 minutes. Quantitation was completed monitoring MS/MS transitions *m/z* 242.1 [M + H^+] → *m/z* 126.1 [M – deoxyribose + H^+] for 5meC, *m/z* 245.2 [M + H^+] → *m/z* 129.0 [M – deoxyribose + H^+] for 5-d_3-meC, *m/z* 258.1 [M + H^+] → *m/z* 142.1 [M – deoxyribose + H^+] for 5hmC, *m/z* 261.1 [M + H^+] → *m/z* 145.1 [M – deoxyribose + H^+] for 5-d_3-hmC, m/z 569.1 [M + H+] → m/z 453.3 [M – deoxyribose + H+] for biotinyl-5fC, and m/z 581.2 [M + H+] → m/z 460.4 [M – deoxyribose + H+] for biotinyl-$^{13}C_{10}{}^{15}N_2$-5fC . All mass spectrometry parameters were optimized through direct infusion of authentic standards with settings as follows: spray voltage of 2700 V, sheath gas of 15 au, source fragmentation voltage of 5 V, and ion transfer tube temperature of 350°C. Fragmentation was induced with an Argon gas flow of 1 mTorr.

Cell Culture and Treatment for Immunostaining:

The Hep3b cell line was obtained from American Type Culture Collection with no further authentication or testing for mycoplasma. The Hep3b cells were grown in MEM media supplemented with 10% FBS (Gibco by Life Technologies TM, #16140-071) and AntibioticAntimycotic (Gibco by Life Technologies TM, #15240062) at 37 °C under 5% CO_2. TET2 inhibitor Bobcat339 was obtained from MedKoo Biosciences (#408006) and Sigma-Aldrich (SML2611). Solutions of Bobcat339 were prepared in

dimethyl sulfoxide (DMSO) at 100 mM. Bobcat339 from either vendor was incubated with cells at either 50 µM for 48 hours prior to collection and extraction of genomic DNA.

DNA Dot Blotting:

Genomic DNA was extracted using DNA Blood/Tissue Kit, denatured and subjected to dot blotting analysis using antibody against 5hmC (Active Motif, #39769) as previously described.[433] Briefly, 2 µg of purified DNA was denatured at 100 °C and mixed 1:1 with 2M NaOAc. DNA was loaded to the membrane, washed, and blocked with 5% NFM. Primary antibody for 5-hmC was incubated overnight and detected by chemiluminescence. The DNA spotted membrane was then stained with 0.02% methylene blue (Sigma) in 0.5 M sodium acetate (pH 5.0) for DNA loading control.

6.3 Results

Bobcat339 was first tested *in vitro* against the enzymatically active catalytic domain of human TET1 using a hemimethylated synthetic oligomer substrate (**Figure 6.4A**). Using isotope dilution liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS), we accurately quantified the conversion of 5mC to 5hmC over time. When houseBobcat339 was tested at 125 μM over three time points, we noted no detectable effect of the inhibitor on TET-catalyzed production of 5hmC and 5fC when compared to dimethylsulfoxide (DMSO) controls (**Figure 6.4B**). Upon obtaining these results after multiple replicates, we turned our attention to the cited publication in an effort to determine why we were observing no inhibition in our assay. We made note that during their synthesis of Bobcat339, the authors reported obtaining a "green powder" after purification of their product.[431] Their synthetic route employed a Chan−Lam coupling reaction to join the 5-chlorocytosine and biphenylboronic acid using a copper acetate catalyst. This immediately raised a question whether residual copper was present in the formulation of Bobcat339, which could be responsible for the observed inhibitory effects.

To examine the potential role of residual Cu(II) in biological activity attributed to Bobcat339, we obtained the compound from Sigma-Aldrich for use in comparative tests. Upon arrival, the physical appearances of the compounds were strikingly different. The sample from MedKoo was colored seafoam-green, while both the compounds prepared in-house and from Sigma-Aldrich were off-white in appearance (**Figure 6.5**). Each of these compounds were characterized by ^{1}H NMR and their molecular weights were

confirmed by high-performance liquid chromatography/electrospray ionization mass spectrometry (HPLC-ESI$^+$ MS), confirming identical chemical structures. Aliquots of Bobcat339 obtained from each of the three sources were analyzed for copper content by inductively coupled plasma−optical emission spectrometry (ICP-OES). ICP-OES analyses confirmed that the sample obtained from MedKoo, which had the green coloration, did contain copper at roughly 3 mol % (**Table 6.1**) while the house- and Sigma-Bobcat339 samples had no detectable levels of copper.

We next tested each of these compounds using the quantitative HPLC-ESI-MS/MS assay in triplicate using recombinant human TET1 and TET2 catalytic domains with a hemimethylated synthetic oligonucleotide substrate. In addition, we included incubations with 125 μM copper introduced by way of $CuSO_4$, as well as our house-Bobcat339 that was spiked with an equimolar concentration of $CuSO_4$ to mimic the conditions of copper contamination. All of the Bobcat339 formulations and copper solutions were tested at 125 μM and were incubated with the catalytic domain of TET1 or TET2 for 30 or 60 min respectively at 37 °C. The levels of 5hmC produced during that time were compared to those generated in the presence of DMSO or heat-inactivated aliquots of protein (**Figure 6.6**). We observed no significant reduction of 5hmC formation as compared to the DMSO control for our house-synthesized compound or the SigmaAldrich compound in either tested isoform of the TET enzyme. There was, however, nearly total inhibition of TET1 and TET2 activity in samples incubated with MedKooBobcat339, $CuSO_4$, and house-Bobcat339 spiked 1:1 with $CuSO_4$. These data strongly suggest that the Cu(II) present in commercial and spiked samples is responsible

for the observed inhibition of TET1 and TET2, and that the Bobcat339 molecule alone is not capable of inhibiting TET at the concentrations previously reported.

We next asked whether the inhibitory activity of Cu contaminated Bobcat339 was solely due to the presence of copper. We constructed a six-point dose response curve for our house-Bobcat339, MedKoo-Bobcat339, and $CuSO_4$ from 0 to 125 μM with 30 min incubations and analysis of 5hmC production by HPLC-ESI-MS/MS. We observed no inhibition of 5hmC production at any point with our in-house prepared compound. For MedKoo sourced Bobcat339 and $CuSO_4$ alone, we noticed an immediate loss of activity at the first concentration point of 25 μM, indicating a much stronger inhibitory effect for Cu(II) than Bobcat339 (**Figure 6.7A**). To better understand the relationship between Bobcat339 and Cu(II), we assessed the inhibitory effect of Cu(II) on TET1 with and without the addition of Bobcat339 at 125 μM. Cu(II) was titrated in from 0.01 to 125 μM in both conditions to closer mimic the 3 mol % Cu(II) contamination we observed in the MedKoo-Bobcat339 (**Figure 6.7B**). A two-way ANOVA revealed a significant effect on inhibitory activity with respect to the presence of Bobcat339 with two points on the curve, 1 and 4 μM Cu(II), showing significant differences in inhibition. These results indicate that the Bobcat339−Cu mixture is a more potent inhibitor of TET1 than Cu(II) alone. This increased potency could be attributed to an increased number of favorable interactions between the TET catalytic domain and the Bobcat339−copper complex as opposed to just Bobcat339 or Cu(II) alone, though further studies would be needed to elucidate the exact mechanism of this interaction.

To confirm our *in vitro* results for Bobcat339 activity against recombinant TET enzymes, the experiments were repeated using a cell-based assay. Hep3B cells in culture were treated with Bobcat339 obtained from either MedKoo or SigmaAldrich for 48 h at 50 μM. DNA was extracted, and immunoblotting was performed to detect the formation of 5hmC compared to DMSO controls. In line with our *in vitro* results, we observed a significant decrease in 5hmC levels within genomic DNA after dosing with Cu(II) contaminated MedKoo-Bobcat339 but saw no reduction in genomic 5hmC levels when dosed with Sigma-Bobcat339 via a dot blot assay (**Figure 6.8A,B**). These results were further verified via isotope dilution HPLC-ESI-MS/MS (**Figure 6.8C**). Taken together, these data further suggest that the reported TET inhibitory activity of Bobcat339 is due to the presence of contaminating copper remaining from the synthesis of the compound and that Bobcat339 itself is less active against TET than was initially reported.

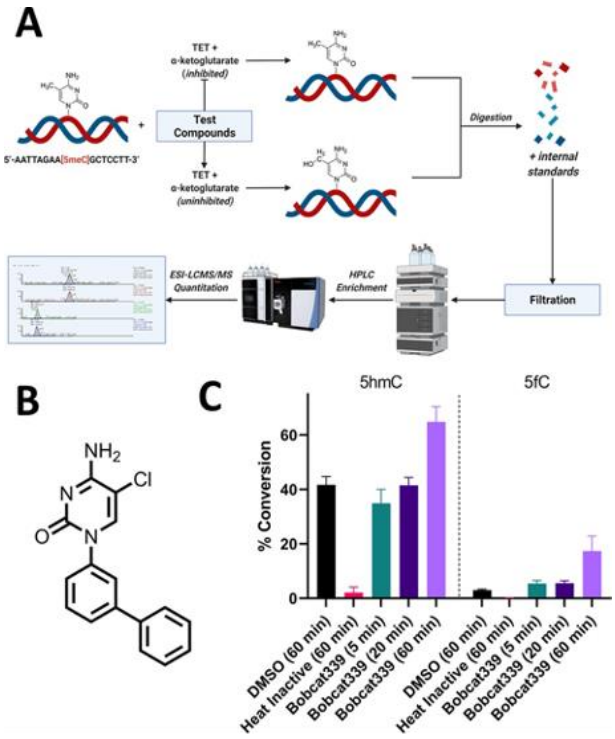


Figure 6.4. Activity of in-house synthesized Bobcat339 against TET1

A) Schematic overview of the quantitative HPLC-ESI-MS/MS assay of TET activity. B) Structure of Bobcat339. C) Percent conversion of 5mC-containing DNA duplexes to 5hmC and 5fC by TET1 in the presence or absence of Bobcat339. 5mC containing DNA oligomers were incubated with recombinant human TET1 catalytic domain and α-ketoglutarate cofactor. The formation of 5hmC and 5fC was quantified by isotope dilution HPLC-ESI-MS/MS at multiple time points (5, 20, 60 min.). Data shown represent experimental triplicates ± SD.

Figure 6.5. Visual Comparison of Bobcat339. 50 mM Bobcat339 in DMSO from MedKoo (left) and house-synthesized (right)

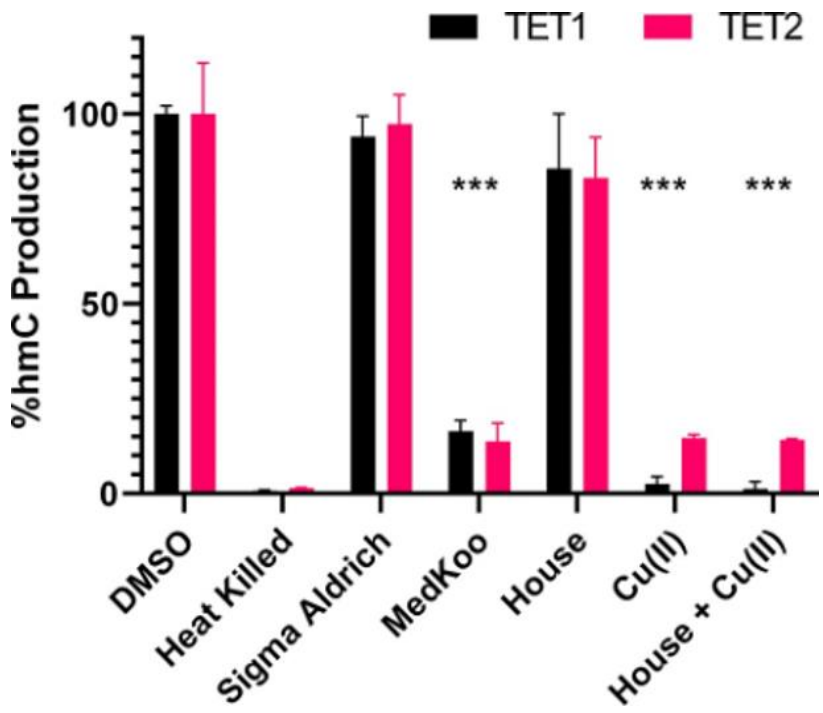


Figure 6.6. Normalized activity of TET1 and TET2 in the presence of Bobcat339 preparations synthesized in our laboratory or purchased from different vendors

5mC containing DNA was incubated with the TET1 or TET2 catalytic domain for 30 min in the presence or in the absence of Bobcat339 (125 μM) in DMSO. The formation of hmC was detected by a quantitative LC-MS assay and is displayed as % activity relative to DMSO controls All data are averaged experimental triplicates normalized to DMSO ± SD (exception to House + Cu(II) for TET2 where n = 2). Statistical analysis was performed by 2-way ANOVA with multiple comparisons and significance reported in relation to the DMSO control (***$P < 0.001$).

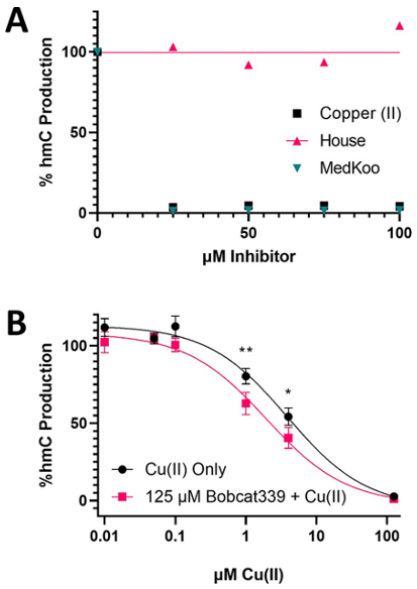


Figure 6.7. Titration of Cu(II) reveals limited effect of Bobcat339 on Cu(II) inhibition of TET1

A) Inhibition of hmC production by TET1 catalytic domain in the presence of of Bobcat339 from MedKoo, Bobcat339 made in house, and Cu(II) alone. Inhibitory activity was assessed by ESI-LCMS/MS quantitation of 5hmC formed upon incubation of 5meC containing DNA duplex with TET1. Experimental duplicates were averaged, normalized to DMSO control, and fit with a non-linear curve. B) Dose-dependence of Cu(II) inhibition of TET1 in the presence or absence of 125 µM Bobcat339. Activity was assessed by HPLC-ESI-MS/MS quantitation of 5hmC formed. Data are normalized to DMSO control and shown as average ± SD. Statistical analysis was performed by 2-way ANOVA with multiple comparisons and significance reported between the Cu(II) doses with and without Bobcat339 (*$P < 0.05$, **$P < 0.01$).

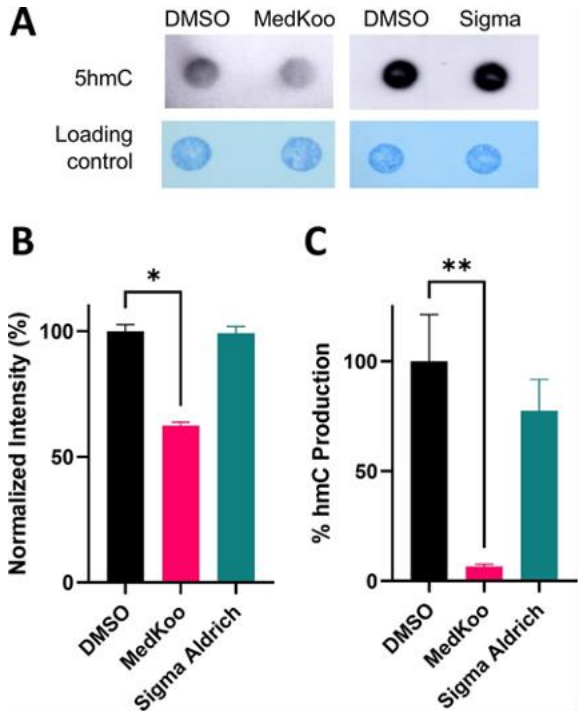


Figure 6.8. Effects of Bobcat339 on global hmC levels in DNA of cultured human cells A) 5hmC quantification in genomic DNA of Hep3B cells via dot blotting after treatment with Bobcat339 sourced from MedKoo or Sigma Aldrich (48 hours at 50 μM). B) Quantified hmC blotting intensity via densitometry is normalized to maximal intensity. Statistical comparison was performed with one-way ANOVA (*$P < 0.05$). C) HPLC-ESI-MS/MS quantitation of 5hmC from the same DNA samples confirms the trends observed via dot blotting. Data are averaged experimental triplicates normalized to DMSO ± SD. Statistical analysis was performed by one-way ANOVA (**$P < 0.01$).

Table 6.1. Copper content of Bobcat339 from different sources

Source	[Cu] (ppm)	Mol % Cu
House	<0.001 (n.d.)	0
Sigma-Aldrich	<0.001 (n.d.)	0
MedKoo	0.198	3.12

6.4 Discussion

In summary, our data suggest that Bobcat339 as a pure compound is not capable of inhibiting TET1 or TET2 at the concentrations reported by Chua and colleagues.[431] Using a combination of enzymatic and cellular assays with two independent methodologies, we have been able to demonstrate that the inhibitory activity of this molecule is a result of Cu(II) contamination. Further, we have been able to show that the effect of copper contamination is greater than that of free copper in solution. There is significant literature that suggests Cu(II) is capable of displacing Fe within the active site of nonheme Fe-dependent enzymes, which would explain why inhibition is observed in the presence of copper and copper contaminated Bobcat339.[434, 435] Our results demonstrate that, depending on the manufacturing process, researchers could expect to see significantly different results from their experiments depending on the levels of copper contamination, and thus how they source their materials should be a particular source of consideration. It appears that the Bobcat339 sample from Sigma-Aldrich was purified to the same extent as the compound synthesized in our laboratory, with no observable copper contamination or discoloration, but this was not the case for the material obtained from MedKoo.

It should be noted that in our synthetic procedure, Bobcat339 underwent three rounds of flash chromatography followed by HPLC purification to fully eliminate the copper contamination, which is a more stringent purification process than most compounds require. While there are additional methods available to strip transition metals from crude synthetic mixtures, the level of quality control of different chemical

vendors in their synthesis of compounds will undoubtedly play a large role in the perceived performance of Bobcat339. While the capability of Bobcat339 as a standalone inhibitor of TET1 and TET2 has been diminished, this now raises questions about the biological activity of other compounds within the Bobcat series reported by Chua and colleagues.[431] An extensive assessment should be done to determine the inhibitory activity of those compounds' ultrapure formulations. Additional studies should also be considered assessing the copurification of copper salts with the Bobcat molecules that could explain the ease of contamination observed in the MedKoo formulations.

7 Chapter VII: Progress Towards the Development of Bifunctional Nucleoside-based Inhibitors of Ten-eleven Translocation Enzymes to Modulate DNA Demethylation

Adapted in part from:

Hurben A.K., Le S., Chao C., Kumar V., Sclanger S., Jha B., Tretyakova N.Y. Development of Bifunctional Nucleoside-Based TET Enzyme Inhibitors as Epigenetic Modulators. *manuscript in preparation*

This chapter features contributions from Alexander K. Hurben (synthesis, *in vitro* studies, and LC-MS experiments) Sang Le (docking and synthesis), Vijay Kumar (synthesis), Simon Schlanger (cell assays and dot blotting), Babal Jha (cell assays and dot blotting), under the supervision of Prof. Tretyakova.

7.1 Introduction

As described in **Chapter VI**, 5-methylcytosine (mC, **Figure 6.1**) is an essential epigenetic DNA mark which controls the levels of gene expression.[400] Proper regulation of DNA methylation across the genome is critical for early development, cell differentiation, and the maintenance of a healthy phenotype.[436] Aberrant DNA methylation is linked to a variety of pathologies, including neurological disorders and cancer.[437] DNA methylation patterns are maintained by the opposing action of two classes of enzymes. DNA methylation marks are introduced and maintained by DNA methyltransferases (DNMT) and are removed by ten-eleven translocation (TET) enzymes, which initiate DNA demethylation via a series of iterative oxidative transformations to the 5-methyl group of mC in DNA.[438, 439] TET enzymes were discovered in 2009 and since then have been shown to be instrumental in cell development and differentiation.[406, 440]

TET enzymes are Fe^{II} dependent dioxygenases which rely on an alpha-ketoglutarate (αKG) cofactor to catalyze the oxidation of methylcytosine (mC) to hydroxymethylcytosine (hmC), formylcytosine (fC), and carboxylcytosine (caC).[441] There are three TET encoding genes, TET1, TET2, and TET3, which share a highly conserved active site (**Figure 7.1**). The active site is comprised of a double-stranded β-helix (DSBH) domain, a cysteine-rich domain, and a Fe^{II} and αKG binding pocket.[439] TET1 and TET3 both possess a N-terminal CXXC zinc finger DNA binding domain.[442] TET2 lacks this DNA binding domain and relies on the protein IDAX to direct its recruitment to DNA.[443]

Given the importance of TET enzymes in normal epigenetic processes and their deregulation in disease, selective and potent small molecule inhibitors of TET are urgently needed. Such inhibitors would aid our basic understanding of TET-mediated biology and also catalyze the development of novel therapeutic strategies for cancer and other diseases. Recent studies have suggested that TET enzymes may be viable therapeutic targets. For example, TET2 mutations are predictive of increased mixed-lineage leukemia patient survival and TET2 deficient myeloid neoplasms have been selectively targeted with TET inhibitors, leading to synthetic lethality and selective tumor cell targeting. [422] [420, 421] Therefore, in certain cancer indications, TET enzymes are essential for creating the epigenetic landscape necessary for tumor development and inhibiting TET mediated DNA demethylation could be a viable therapeutic strategy.

However, it should be noted that in certain cancers TET enzymes are downregulated or mutated, leading to decreased global levels of hmC and fC.[408, 409] For example, TET2 mutations result in decreased levels of hmC in T cell lymphoma and acute myeloid leukemia.[410] [411] Additionally, TET2 knockout mice are more susceptible toward developing myeloid malignancies than their WT conterparts.[412, 413] Thus, in some cases, TET enzymes may act as tumor suppressor genes, and their dysfunction could prevent the removal of aberrant DNA methylation marks, leading to cancer initiation.[409, 414-416] In these instances, TET inhibitors would make valuable tools to probe epigenetic mechanisms of cancer etiology.

Another potential therapeutically relevant area where TET inhibitors could be applied is chimeric antigen receptor (CAR) T cell immunotherapy. TET2 knockdown in

CAR T cells has been shown to enhance their efficacy and cause leukemia remission.[444] Therefore, *ex vivo* administration of TET inhibitors to CAR-T cells prior to CAR-T patient infusion may improve treatment outcomes. Thus, further development of TET inhibitors is warranted to explore this possibility.

To date, the majority of reported small molecule TET inhibitors have been αKG cofactor mimics (**Figure 7.2**). R and S 2-hydroxyglutarate (R-2-HG and S-2-HG) as well as itaconic (IVA), a metabolite produced by activated macrophages, have been reported to inhibit TET1 and 2 with low mM activity.[445, 446] N-oxalylglycine (NOG) has been used as a enzymatically inactive cofactor analog that was substituted for αKG in TET crystallographic studies and NOG has a IC_{50} of 80 μM against TET2.[447] Derivatives of NOG have been utilized in "bump and hole strategies" aimed at modulating the activity of carefully designed mutant TETs.[448]

The use of αKG inhibitors to inhibit TET is problematic due to the redundant use of this co-factor across many biological pathways. In addition to TET, human genome codes for 40 other αKG dependent enzymes .[449] Many of the reported αKG mimetics aimed to modulate TET activity are likely to have unintended effects on other enzymatic processes which use this co-factor.

Efforts have been made to develop TET inhibitors targeting other parts of the protein besides the αKG binding site (**Figure 7.2**). For example, Bobcat339 is mC mimic developed to target the substrate binding pocket and reported to have a double digit μM IC_{50} against TET1 and TET2.[431] However, we recently showed that residual Cu salts left

over as an artifact of synthesis of this molecule via an Ulman coupling was responsible for the observed TET inhibition (**Chapter VI)**.[450] Virtual screening efforts in have identified C35, a dicatechol containing natural product, to be a signal digit μM TET inhibitor.[428] However, the catechol functionality is a metabolic lability due to its propensity to oxidize into reactive quinone species.[451] Eltrombopag, an existing drug used to treat thrombocytopenia, was identified in focused library screen and found to inhibit TET with an IC_{50} of 1.3 μM.[452] Lastly, TETi, a 5-carbon long diacid, containing alcohol and olefin functionality (**Figure 7.2**), has been reported to inhibit all 3 TETs with an IC_{50} of 10 μM.[453] As mentioned above, this compound displayed efficacy in selectively targeting cancer cells in a hematological cancer model that harbored a mutant TET2 and thus is an exciting development in the field of TET inhibitor development.[422]

Despite the recent progress made thus far in the development of small molecule inhibitors of TET, there is still a need for selective and potent small molecule inhibitors as such inhibitors could hold therapeutic value. To overcome obstacles impeding the development of selective TET inhibitors, we sought to implement bifunctionality in our molecular design by designing molecules targeting both the substrate binding site and the catalytic site (**Figure 7.3**). Bifunctional inhibitors have garnered success against other epigenetic enzymes, however such a linked cofactor and substrate approach has yet to be applied to TET inhibition.[454-456] Herein, we disclose a series of first generation bifunctional nucleoside-based inhibitors and present preliminary biological activity data in both *in vitro* and human cell models. Given the need for TET inhibitors for use as chemical probes to investigate DNA methylation processes and their therapeutic

potential, this design provides a novel class of TET inhibitors to be expanded upon in future work.

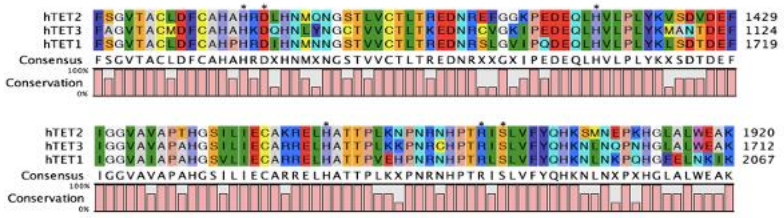


Figure 7.1. Sequence alignment of two anti-parallel beta-sheets in the DSBH catalytic domains of TET1, TET2, and TET3

The starred residues indicate residues involved in Fe(II) coordination and cofactor binding.[457]

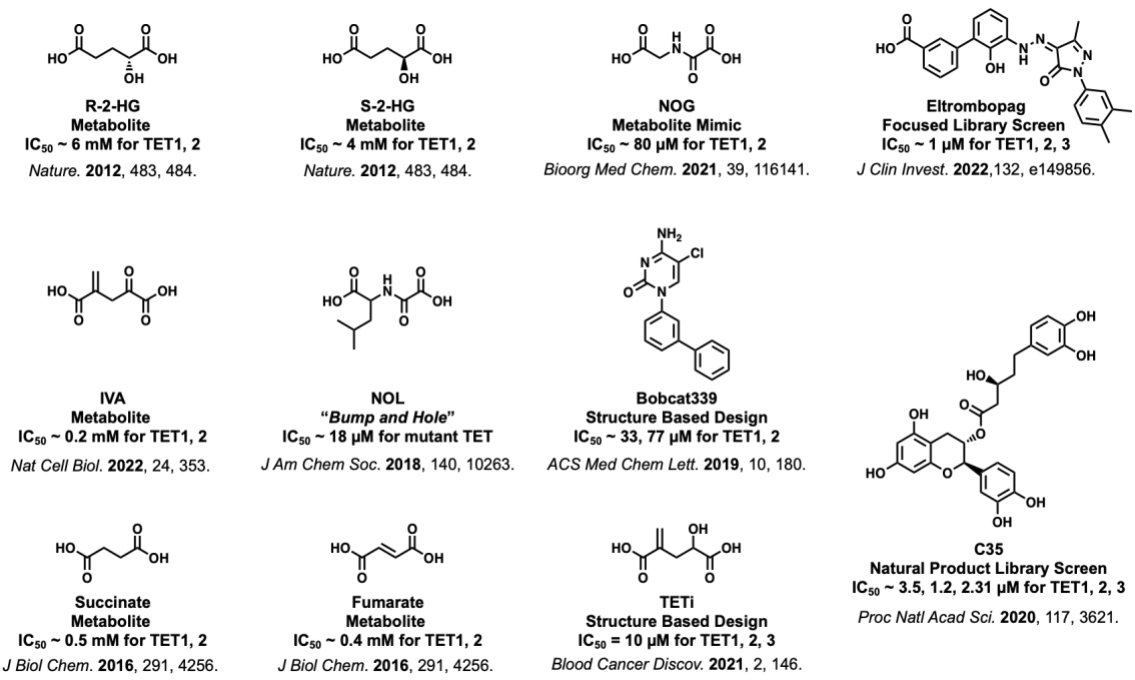


Figure 7.2. Structures of reported TET inhibitors

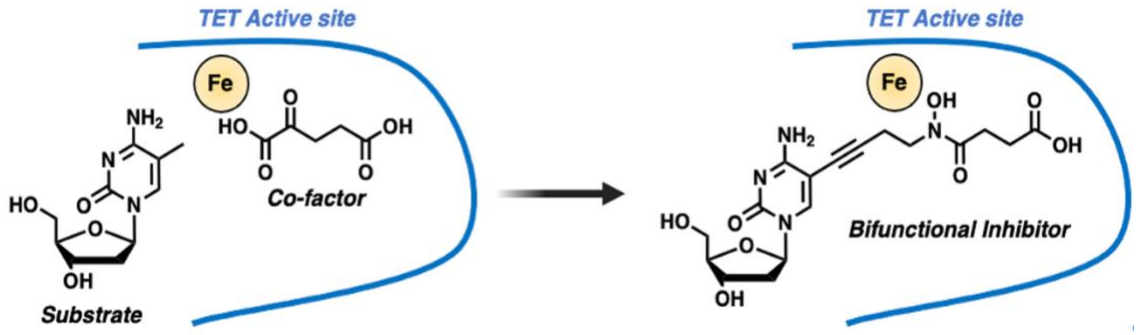


Figure 7.3. Pictorial representation of the design for bifunctional TET inhibitors through the linkage of the alpha-ketoglutarate cofactor and mC substrate

7.2 Materials and Methods

All chemicals and reagents required for the synthesis, purification, and assaying of TET inhibitors were purchased from Sigma Aldrich unless otherwise stated. NMR and LCMS solvents were obtained from Fisher Scientific. K562 cells used for cell-based assays were obtained from the Cleveland Clinic. All materials for tissue culture and related experiments were obtained from Corning and Gibco. DNA extraction kits were obtained from Qiagen. Recombinant TET1 (#31417) enzyme was purchased from Active Motif. Synthetic DNA oligomer substrates were ordered from Integrated DNA Technologies. Internal standards for LCMS quantitation of 5mC and 5hmC were obtained from Toronto Research Chemicals. Nucleotide digestion buffer was supplied by New England Biolabs (#M0649S). Filters used for purification of nucleotides were acquired from Pall Corporation. Licor Image Studio was used to generate process acquired dot blotting images and all statistical analyses herein were performed with GraphPad Prism v9.3.1.

Cell culture and Treatment:

Cells were grown in MEM media supplemented with 10% FBS (Gibco by Life Technologies TM, #16140-071) and AntibioticAntimycotic (Gibco by Life Technologies TM, #15240062) at 37°C under 5% CO_2. Solutions of inhibitors were prepared in dimethyl sulfoxide (DMSO) were incubated with cells at the specific concentration and duration prior to collection and extraction of genomic DNA.

In-vitro TET Assays:

For *in vitro* TET assays, 10 pmol of 16-mer hemi-methylated synthetic dsDNA oligomer designed after the *Rassf1a* tumor suppressor gene promotor (5'-AATTAGAA[5meC]GCTCCTT3') were incubated with 2xTET reaction buffer (50 mM HEPES [pH 7.9], 100 mM NaCl, 75 µM $Fe(NH_4)_2(SO_4)_2$, 2 mM ascorbate, 20 µM α-ketoglutarate, 1 mM DTT), recombinant human TET1 (0.2 µg, 2.3 pmol, 115 nM) and 5% DMSO (or inhibitor in DMSO) on a 20 µL scale. The reaction was initiated by adding the protein to the reaction mixture and placing on a dry block at 37 °C for 30 minutes. Reactions were terminated at by transferring reaction tubes to dry ice to snap-freeze the reaction until all replicates were completed. Upon completion, all samples were heat-inactivated at 90°C for 10 minutes and then dried entirely in SpeedVac concentrator. Sample were then prepared for LCMS or dot blot analysis.

HPLC-ESI-MS/MS quantitation of 5mC and 5hmC:

Quantitation of 5hmC is based on methodology previously described by our group with slight modification.[432] Dried enzymatic reaction samples were reconstituted with mQ-water and spiked with 1 pmol each of 5-Methyl-2'-deoxycytidine-d_3 and 5-(Hydroxymethyl)-2'deoxycytidine-d_3. Nucleotide digestion enzyme mix and buffer were added per product guidelines on a 20 µL scale. Samples were digested for 60 min at 37 °C on a dry block. Following digestion, samples were filtered on nanosep 10K Omega filters. The samples were then dried entirely and resuspended in 100 µL mQ-water for offline HPLC cleanup to enrich 5hmC. An Atlantis T3 column (*Waters*, 4.6 x 150 mm, 3 µm) was eluted at a flow of 0.9 mL/min with a linear gradient previously described of 5 mM NH_4CHO_2 buffer, pH 4.0 (A) and MeOH (B). Fractions containing 5meC and 5hmC

were collected and pooled before analysis by isotope dilution HPLC-ESI-MS/MS. Pooled samples were dried and reconstituted in 2 mM ammonium formate buffer (NH_4CHO_2) with a Dionex Ultimate 3000UHPLC (*Thermo Fisher*) interfaced with a Thermo TSQ Quantis mass spectrometer (*Thermo Fisher*) flowing 2 mM NH_4CHO_2 (A) and ACN (B) at 15 µL/min through a Zorbax SB-C18 column (*Agilent*, 0.5 x 150 mm, 3 µm) with the gradients as previously described. 5meC and 5-d_3-meC eluted at 5.1 minutes while 5hmC and 5-d_3-hmC eluted at 3.5 minutes. Quantitation was completed monitoring MS/MS transitions *m/z* 242.1 [M + H^+] → *m/z* 126.1 [M – deoxyribose + H^+] for 5meC, *m/z* 245.2 [M + H^+] → *m/z* 129.0 [M – deoxyribose + H^+] for 5-d_3-meC, *m/z* 258.1 [M + H^+] → *m/z* 142.1 [M – deoxyribose + H^+] for 5hmC, *m/z* 261.1 [M + H^+] → *m/z* 145.1 [M – deoxyribose + H^+] for 5-d_3-hmC. All mass spectrometry parameters were optimized through direct infusion of authentic standards with settings as follows: spray voltage of 2700 V, sheath gas of 15 au, source fragmentation voltage of 5 V, and ion transfer tube temperature of 350 °C. Fragmentation was induced with an Argon gas flow of 1 mTorr.

DNA Dot Blotting:

For cellular assays, genomic DNA was extracted using DNA Blood/Tissue Kit per the manufactures instructions, denatured and subjected to dot blotting analysis using antibody against 5hmC (Active Motif, #39769) as previously described.[433] Briefly, 2 µg DNA was diluted in 8 µL water, and was denatured by adding 2 µL of 2 N NaOH/50 mM EDTA, and incubated for 10 min at 95 °C. Samples were quickly moved to ice and were neutralized with 10 µL of ice cold 2 M ammonium acetate. Denatured DNA was spotted on a nitrocellulose membrane with a Bio-Dot Apparatus Assembly (Bio-Rad). The

blotted membranes were air-dried and the DNA was hybridized to the membrane by UV cross linking with a Spectrolinker XL-1000 (1200 µJ/cm2). The membranes were blocked in 5% milk/TBST for 1 h followed by primary antibody blotting at 4 °C overnight. The next day, the blots were washed with TBST followed by incubation with secondary antibody for 1 h, washed, and imaged using a Licor imager. The DNA spotted membrane was then stained with 0.02% methylene blue in 0.5 M sodium acetate (pH 5.0) for DNA loading control. The same protocol was followed for *in vitro* assays.

General Reaction Conditions.

Chemical reactions were conducted with oven-dried glassware under argon atmosphere unless otherwise stated. Thin- layer chromatography (TLC) was performed with Analtech silica uniplates and visualized under 254 nm UV light or with $KMnO_4$ staining. Column chromatography was conducted with 60 mesh silica gel. NMR spectra were acquired on a 500 MHz Bruker spectrometer at room temperature. MS spectra were acquired on a Thermo Scientific LTQ XL Ion Trap. Detailed reaction procedures and **Figures 7.4 – 10** are found below.

General method for O-benzoylation reactions:

To a round-bottomed flask were sequentially added DCM (120 mL), starting alkyne **1a** or **1b** (60 mmol, 1 equiv), and $NaHCO_3$/NaOH aqueous solution (pH 10.5, 60 mL). After five minutes of vigorous stirring, benzoyl peroxide was added (72 mmol, 1.2 equiv) and the reaction mixture was allowed to stir for four hours at room temperature. The organic layer was collected, then the aqueous layer was extracted with DCM (6 x 20 mL). The organic layers were pooled, washed with brine, dried with magnesium sulfate,

and concentrated *in vacuo*. Purification via silica gel flash column chromatography (2 column volumes of 0%, 5 columns volumes from 0% to 10%, 6 column volumes of 10%, 5 column volumes from 10% to 20%, and 8 column volumes of 20% EtOAc in hexanes) afforded compound **2a-b**.

O-benzoyl-N-(prop-2-yn-1-yl)hydroxylamine (**2a**): Orange oil (32% yield). R_f = 0.32 (20% EtOAc in hexanes). ^{1}H NMR (500 MHz, $CDCl_3$) δ 2.26 (t, *J*= 2.4 Hz, 1H), 3.92 (d, *J*= 2.5 Hz, 2H), 7.46 (t, *J*= 7.8 Hz, 2H), 7.56 – 7.62 (m, 1H), 8.03 – 8.07 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 41.8, 73.3, 78.6, 128.1, 128.7, 129.6, 133.6, 166.7. MS (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{10}H_{10}NO_2^+$ 176.1, found 176.1.

O-benzoyl-N-(but-3-yn-1-yl)hydroxylamine (**2b**): Clear, colorless oil (47% yield). R_f = 0.32 (20% EtOAc in hexanes). ^{1}H NMR (500 MHz, $CDCl_3$) δ 2.05 (t, *J*= 2.7 Hz, 1H), 2.53 (td, *J*= 6.7, 2.7 Hz, 2H), 3.30 (t, *J*= 6.7 Hz, 2H), 7.42 (dd, *J*= 8.7, 7.0 Hz, 1H), 7.51 – 7.59 (m, 2H), 7.99 (dd, *J*= 8.4, 1.5 Hz, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 17.5, 50.8, 70.4, 81.0, 128.3, 128.6, 129.4, 133.4, 166.5. MS (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{11}H_{12}NO_2^+$ 190.1, found 190.1.

General method for N-acylation reactions:

A round-bottomed flask was purged with argon and sequentially added were monomethyl succinate or monomethyl fumarate (1-1.8 equiv) and anhydrous DCM. After five minutes of stirring, oxalyl chloride (1.05-1.85 equiv) was added dropwise, followed by anhydrous DMF (3-8 drops). The acid chloride reaction mixture was allowed to stir

for two hours at room temperature. After two hours, a second round-bottomed flask was placed in an ice bath, purged with argon and were sequentially added compound **2a-b** or **1a-b** (1 equiv), anhydrous THF, and triethylamine (1.85 equiv). After five minutes of stirring, the acid chloride reaction mixture was transferred to the second reaction mixture *via* anhydrous transfer and was added dropwise. The reaction mixture was allowed to stir for 18 hours at room temperature. Afterwards, DCM (10 mL) and water (10 mL) was added to the reaction mixture. The organic layer was collected, then the aqueous layer was extracted with DCM (4 x 10 mL). The organic layers were pooled, washed with brine, dried with magnesium sulfate, and concentrated *in vacuo*. Purification *via* silica gel flash column chromatography (2 column volumes of 0%, 5 columns volumes from 0% to 15%, 4 column volumes of 15%, 5 column volumes from 15% to 20%, and 8 column volumes of 20% EtOAc in hexanes) afforded compound **3a-b**, **4a-b**, and **5a-b**.

methyl 4-((benzoyloxy)(prop-2-yn-1-yl)amino)-4-oxobutanoate (**3a**): Cloudy, colorless oil (57% yield). R_f = 0.33 (40% EtOAc in hexanes). ^{1}H NMR (500 MHz, $CDCl_3$) δ 2.25 (t, *J* = 2.5 Hz, 1H), 2.64 – 2.72 (m, 4H), 3.68 (s, 3H), 4.61 (d, *J* = 2.5 Hz, 2H), 7.50 – 7.56 (m, 2H), 7.67 – 7.72 (m, 1H), 8.13 (dd, *J* = 8.4, 1.4 Hz, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 27.68, 28.32, 38.14, 52.02, 73.06, 76.59, 126.60, 129.09, 130.34, 134.79, 164.36, 172.99. MS (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{15}H_{16}NO_5^+$ 290.1, found 290.1.

methyl 4-((benzoyloxy)(but-3-yn-1-yl)amino)-4-oxobutanoate (**3b**): Clear, colorless oil (88% yield). R_f = 0.45 (40% EtOAc in hexanes). ^{1}H NMR (500 MHz, $CDCl_3$) δ 1.96 (s, 1H), 2.58 (td, *J* = 7.1, 2.7 Hz, 2H), 2.65 (s, 4H), 3.67 (s, 3H), 4.00 (t, *J* = 7.1 Hz, 2H), 7.52 (t, *J* = 7.7 Hz, 2H), 7.68 (t, *J* = 7.5 Hz, 1H), 8.06 – 8.14 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 17.5, 27.4, 28.4, 47.3, 52.0, 70.3, 80.8, 126.7, 129.1, 130.6, 134.7, 164.5, 172.2, 173.1. (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{16}H_{17}NO_5^+$ 304.1, found 304.1.

methyl (*E*)-4-(but-3-yn-1-ylamino)-4-oxobut-2-enoate (**4a**): White solid (93% yield). R_f = 0.42 (40% EtOAc in hexanes). ^{1}H NMR (CHLOROFORM-*d* ,400MHz): δ = 8.13 (dd, *J*=8.2, 1.2 Hz, 2 H), 7.71 (s, 1 H), 7.50 - 7.63 (m, 2 H), 7.16 (d, *J*=15.7 Hz, 1 H), 6.98 (d, *J*=15.3 Hz, 1 H), 4.69 (d, *J*=2.3 Hz, 2 H), 3.76 ppm (s, 3 H) ^{13}C NMR (126 MHz, $CDCl_3$) δ 38.3, 52.4, 73.6, 76.0, 126.1, 129.2, 130.4, 130.7, 133.7, 135.0, 164.4, 164.8, 165.6. MS (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{15}H_{14}NO_5^+$ 288.1, found 288.1.

methyl (*E*)-4-((benzoyloxy)(but-3-yn-1-yl)amino)-4-oxobut-2-enoate (**4b**): White solid (80% yield). R_f = 0.35 (40% EtOAc in hexanes). ^{1}H NMR (METHANOL-d_4 ,400MHz): δ = 8.02 - 8.19 (m, 2 H), 7.75 (t, *J*=7.4 Hz, 1 H), 7.59 (t, *J*=7.6 Hz, 2 H), 7.19 (br. s., 1 H), 6.84 (d, *J*=15.7 Hz, 1 H), 4.06 (br. s., 2 H), 3.74 (br. s., 3 H), 2.62 ppm (td, *J*=6.6, 2.5 Hz, 2 H).). ^{13}C NMR (126 MHz, $CDCl_3$) δ 17.4, 47.4, 52.3, 70.5, 80.4, 126.1, 129.0, 130.3, 130.9, 133.1, 1349, 164.4, 165.5. MS (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{16}H_{15}NO_5^+$ 302.1, found 302.1.

methyl (*E*)-4-oxo-4-(prop-2-yn-1-ylamino)but-2-enoate (**5a**): White solid (97% yield). R_f = 0.20 (40% EtOAc in hexanes). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.94 (d, *J* = 15.4 Hz, 1H), 6.65 (d, *J* = 15.4 Hz, 1H), 4.17 (s, 2H), 3.80 (s, 3H), 2.27 (s, 1H). (ESI-TOF) *m/z* [M + H]$^+$ calculated for $C_8H_{10}NO_3^+$ 168.1, found 168.1.

methyl (*E*)-4-(but-3-yn-1-ylamino)-4-oxobut-2-enoate (**5b**): White solid (95% yield). R_f = 0.15 (40% EtOAc in hexanes). ^{1}H NMR (500 MHz, $CDCl_3$) δ 6.96 (d, *J* = 15.4 Hz, 1H), 6.81 (d, *J* = 15.4 Hz, 1H), 6.66 (d, *J* = 7.7 Hz, 1H), 3.77 (s, 3H), 3.49 (q, *J* = 6.3 Hz, 2H), 2.44 (td, *J* = 6.4, 2.6 Hz, 2H), 2.01 (t, *J* = 2.6 Hz, 1H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 19.3, 38.5, 52.3, 70.4, 81.3, 130.2, 136.5, 163.8, 166.2. (ESI-TOF) *m/z* [M + H]$^+$ calculated for $C_{10}H_{12}NO_3^+$ 182.1, found 182.1.

General method for consecutive debenzoylation-acetylation reactions:

A round-bottomed flask was purged with argon and sequentially added were compound **3a-b** (1 equiv), anhydrous methanol, and potassium carbonate (1.2 equiv). The reaction mixture was allowed to stir for 30 minutes at room temperature and was monitored by TLC ($KMnO_4$ stain). Afterwards, the methanol was removed *in vacuo* and to the reaction mixture was added ethyl acetate (20 mL) and water (20 mL). The reaction mixture was brought to a pH of 12 by adding NaOH (1 M) in which the aqueous layer was collected. Subsequently, the aqueous layer was acidified to a pH of 4 by adding HCl (1 M). The aqueous phase was extracted with ethyl acetate (3 x 15 mL). The organic layers were pooled, washed with brine, dried with magnesium sulfate, and concentrated *in vacuo* to afford a crude oil which was used subsequently without additional

purification. To another round-bottomed flask were sequentially added the crude oil (1 equiv), triethylamine (10 equiv), acetic anhydride (10 equiv), and ethyl acetate (minimal amount required to dissolve the crude oil into the reaction mixture). The reaction mixture was allowed to stir for 30 minutes at room temperature and was monitored by TLC ($KMnO_4$ stain). Afterwards, water (10 mL) was added to the reaction mixture. The organic layer was collected, then the aqueous layer was extracted with ethyl acetate (3 x 10 mL). The organic layers were pooled, neutralized with saturated $NaHCO_3$ solution, washed with brine, dried with magnesium sulfate, and concentrated *in vacuo* to afford compound **6a-b**.

methyl 4-(acetoxy(prop-2-yn-1-yl)amino)-4-oxobutanoate (**6a**): Orange oil (52% yield). R_f = 0.42 (50% EtOAc in hexanes, $KMnO_4$ stain). ^1H NMR (500 MHz, $CDCl_3$) δ 2.22 (s, 3H), 2.25 (t, *J* = 2.5 Hz, 1H), 2.60 (dq, *J* = 10.6, 5.9 Hz, 4H), 3.65 (s, 3H), 4.45 (d, *J* = 2.6 Hz, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 18.4, 27.4, 28.2, 37.9, 51.9, 73.0, 76.4, 168.2, 172.2, 172.8. (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{10}H_{14}NO_5^+$ 228.1, found 228.1.

methyl 4-(acetoxy(but-3-yn-1-yl)amino)-4-oxobutanoate (**6b**): Orange oil (36% yield). R_f = 0.37 (50% EtOAc in hexanes, $KMnO_4$ stain). ^1H NMR (500 MHz, $CDCl_3$) δ 1.98 (t, *J* = 2.8 Hz, 1H), 2.19 (s, 3H), 2.41 – 2.50 (m, 2H), 2.50 – 2.64 (m, 4H), 3.64 (s, 3H), 3.83 (t, *J* = 7.1 Hz, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 17.3, 18.5, 27.2, 28.3, 47.0, 51.9, 70.2,

80.7, 168.4, 171.8, 173.0. (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{11}H_{16}NO_5^+$ 242.1, found 242.1.

General Method for Sonogashira Cross Coupling Reaction:

A Schlenk flask was purged with argon and sequentially added were 5-iodo-2'-deoxycytidine or 5-iodo-2′-deoxyuridine (1 equiv), $Pd(PPh_3)_2Cl_2$ (0.1 equiv), CuI (0.05 equiv), half of the anhydrous DMF (1 mL), and anhydrous *N, N*-diisopropylethylamine (2 equiv). A separate round-bottomed flask was purged with argon and sequentially added compound **4b**, **5b**, or **6a-b**. (1.5 equiv) and half of the anhydrous DMF (1 mL). After allowing each reaction mixture to stir and purge for thirty minutes, the reaction mixture in the round-bottomed flask was transferred to the reaction mixture in the Schlenk flask *via* anhydrous transfer and was added dropwise. Afterwards, the reaction was heated to 50 C and was covered in aluminum foil to achieve dark conditions, and the reaction mixture was allowed to stir for 24-91 hours, as tracked by TLC. The reaction mixture was then concentrated *in vacuo* to afford a dark crude oil. Purification *via* silica gel flash column chromatography (2 column volumes of 0%, 5 column volumes from 0% to 10%, and 20 column volumes of 10% methanol in DCM) afforded compounds **7a-b**, **8a-b**, and **9a-b**.

methyl (*E*)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(benzoyloxy)amino)-4-oxobut-2-enoate (**7a**): Light yellow foam (35% yield). ^{1}H NMR (500 MHz, MeOD) δ = 8.07 (s, 1 H), 7.97 - 8.05 (m, 2 H), 7.61 - 7.67 (m, 1 H), 7.46 (t, J=7.8 Hz, 2 H), 6.75 (d, J=15.3 Hz,

1 H), 6.07 (t, J=6.3 Hz, 1 H), 4.22 - 4.28 (m, 1 H), 4.09 (br. s., 2 H), 3.84 (q, J=3.4 Hz, 1 H), 3.66 - 3.75 (m, 2 H), 3.57 - 3.66 (m, 3 H), 2.76 (t, J=6.1 Hz, 1 H), 2.24 - 2.31 (m, 1 H), 2.00 ppm (dt, J=13.4, 6.6 Hz, 1 H). (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{25}H_{27}N_4O_9^+$ 527.2, found 527.2.

methyl (*E*)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoate (**7b**): This compound was taken forward to the following step crude (crude yield 56%).

methyl 4-(acetoxy(3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)amino)-4-oxobutanoate (**8a**): Light orange solid (10% yield). R_f = 0.23 (15% methanol in DCM). ^{1}H NMR (500 MHz, MeOD) δ 2.14 (dt, *J* = 13.2, 6.4 Hz, 1H), 2.27 (s, 3H), 2.39 (ddd, *J* = 13.6, 6.2, 4.1 Hz, 1H), 2.59 – 2.70 (m, 4H), 3.67 (s, 3H), 3.74 (dd, *J* = 12.1, 3.7 Hz, 1H), 3.82 (dd, *J* = 12.1, 3.2 Hz, 1H), 3.95 (q, *J* = 3.5 Hz, 1H), 4.36 (dt, *J* = 6.3, 4.0 Hz, 1H), 4.72 (s, 2H), 6.19 (t, *J* = 6.3 Hz, 1H), 8.35 (s, 1H). ^{13}C NMR (126 MHz, MeOD) δ 18.2, 28.3, 29.0, 42.4, 52.3, 62.4, 71.7, 88.0, 89.1, 89.6, 91.8, 146.4, 156.6, 166.3, 170.1, 174.7. (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{19}H_{25}N_4O_9^+$ 453.2, found 453.2.

methyl 4-(acetoxy(4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-

yl)amino)-4-oxobutanoate (**8b**): Light yellow foam (42% yield). R_f = 0.37 (10% methanol in DCM). ^{1}H NMR (500 MHz, MeOD) δ 2.14 (td, *J* = 13.5, 6.9 Hz, 1H), 2.25 (s, 3H), 2.38 (ddd, *J* = 13.5, 6.2, 4.0 Hz, 1H), 2.55 – 2.70 (m, 4H), 2.75 (t, *J* = 6.4 Hz, 2H), 3.64 (s, 3H), 3.73 (dd, *J* = 12.0, 3.6 Hz, 1H), 3.82 (dd, *J* = 12.0, 3.2 Hz, 1H), 3.92 – 3.98 (m, 3H), 4.36 (dt, *J* = 6.3, 3.9 Hz, 1H), 6.20 (t, *J* = 6.3 Hz, 1H), 8.26 (s, 1H). ^{13}C NMR (126 MHz, MeOD) δ 18.2, 19.0, 28.1, 29.1, 42.4, 52.3, 54.8, 62.5, 71.8, 88.0, 89.0, 93.8, 145.4, 156.7, 166.6, 174.6. (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{20}H_{27}N_4O_9^+$ 467.2, found 467.2.

methyl (*E*)-4-((benzoyloxy)(4-(1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoate (**9a**): ^{1}H NMR (METHANOL-d_4 ,400MHz): δ = 8.05 - 8.20 (m, 3 H), 7.69 - 7.75 (m, 1 H), 7.55 (t, *J*=7.8 Hz, 2 H), 6.83 (d, *J*=15.7 Hz, 1 H), 6.20 (t, *J*=6.5 Hz, 1 H), 4.38 (dt, *J*=6.2, 3.4 Hz, 1 H), 4.15 (br. s., 2 H), 3.93 (q, *J*=3.1 Hz, 1 H), 3.68 - 3.83 (m, 5 H), 2.82 (t, *J*=6.3 Hz, 2 H), 2.24 - 2.33 (m, 1 H), 2.12 - 2.23 ppm (m, 1 H). (ESI-TOF) *m/z* $[M + H]^+$ calculated for $C_{25}H_{26}N_3O_{10}^+$ 528.2, found 528.2.

methyl (*E*)-4-((4-(1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoate (**9b**). This compound was taken forward to the following step crude (crude yield 43%).

General Method for Hydroxyamide Deprotection:

A round-bottomed flask was purged with argon and sequentially added were compound **7a**, **8a-b**, or **9a** (1 equiv), anhydrous methanol, and potassium carbonate (1.2 equiv). The reaction mixture was allowed to stir for 30 minutes at room temperature and was monitored by TLC ($KMnO_4$ stain). The reaction mixture was then concentrated *in vacuo* to afford a crude oil. Purification *via* silica gel flash column chromatography (1 column volume of 0%, 3 column volumes from 0% to 15%, and 30 column volumes of 15% methanol in DCM) and HPLC purification using a Sunfire C18 column (5 μm, 150 × 2.0 mm, 10.0 mL/min flow rate, buffer A: H_2O), buffer B: ACN) involved a gradient of 0-30% B (0-30 min), of 30-95% B (30-40 min), followed by an isocratic hold at 95% B (40-45 min) afforded compounds **10a-b**, **11**, or **12.**

methyl 4-((3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)(hydroxy)amino)-4-oxobutanoate (**10a**): Light yellow solid (61% yield). R_f = 0.24 (15% methanol in DCM, $KMnO_4$ stain). ^{1}H NMR (500 MHz, MeOD) δ 2.14 (dt, *J* = 13.3, 6.4 Hz, 1H), 2.39 (ddd, *J* = 13.6, 6.2, 4.0 Hz, 1H), 2.61 (t, *J* = 6.7 Hz, 2H), 2.82 (t, *J* = 6.6 Hz, 2H), 3.67 (s, 3H), 3.74 (dd, *J* = 12.0, 3.7 Hz, 1H), 3.82 (dd, *J* = 12.1, 3.1 Hz, 1H), 3.95 (q, *J* = 3.6 Hz, 1H), 4.36 (dt, *J* = 7.2, 3.9 Hz, 1H), 4.60 (s, 2H), 6.20 (t, *J* = 6.3 Hz, 1H), 8.33 (s, 1H). ^{13}C NMR (126 MHz, MeOD) δ 28.4, 29.2, 40.2, 42.3, 49.9, 52.2, 62.5, 71.7, 75.4, 88.0, 89.0, 90.8, 92.1, 146.2, 156.7, 166.4, 175.1. (ESI-TOF) *m/z* $[M - H]^-$ calculated for $C_{17}H_{21}N_4O_8^-$ 409.1, found 409.1.

methyl 4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobutanoate (**10b**): White foam (59% yield). R_f = 0.28 (15% methanol in DCM, $KMnO_4$ stain). ^1H NMR (500 MHz, MeOD) δ 2.13 (dt, *J* = 13.3, 6.5 Hz, 1H), 2.37 (ddd, *J* = 13.6, 6.2, 4.0 Hz, 1H), 2.59 (t, *J* = 6.7 Hz, 2H), 2.78 (dt, *J* = 34.3, 6.5 Hz, 4H), 3.64 (s, 3H), 3.70 – 3.88 (m, 4H), 3.94 (q, *J* = 3.5 Hz, 1H), 4.36 (dt, *J* = 6.3, 3.8 Hz, 1H), 6.20 (t, *J* = 6.4 Hz, 1H), 8.24 (s, 1H). ^{13}C NMR (126 MHz, MeOD) δ 18.6, 28.3, 29.3, 42.35, 48.0, 52.2, 62.5, 71.8, 73.3, 87.9, 89.0, 93.1, 94.0, 145.2, 166.6, 175.0, 175.1. (ESI-TOF) *m/z* $[M - H]^-$ calculated for $C_{18}H_{23}N_4O_8^-$ 423.1, found 423.1.

methyl (*E*)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobut-2-enoate (**11**): Yellow solid (66% yield). R_f = 0.48 (15% methanol in DCM, $KMnO_4$ stain). ^1H NMR (METHANOL-d_4 ,500MHz): δ = 8.13 (s, 1 H), 7.57 (d, *J*=15.6 Hz, 1 H), 6.68 (d, *J*=15.6 Hz, 1 H), 6.09 (t, *J*=6.3 Hz, 1 H), 4.22 - 4.30 (m, 1 H), 3.81 - 3.93 (m, 3 H), 3.66 - 3.74 (m, 4 H), 3.57 - 3.66 (m, 2 H), 3.21 (dt, *J*=3.3, 1.6 Hz, 2 H), 2.71 (t, *J*=6.4 Hz, 2 H), 2.23 - 2.32 (m, 1 H), 1.98 - 2.06 ppm (m, 1 H). (ESI-TOF) m/z $[M - H]^-$ calculated for $C_{18}H_{21}N_4O_8^-$ 421.1, found 421.1.

methyl (E)-4-(hydroxy(4-(1-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoate (**12**): Yellow solid (34% yield). R_f = 0.32 (15% methanol in DCM, $KMnO_4$ stain).

^{1}H NMR (METHANOL-d_4 ,400MHz): δ = 8.22 (s, 1 H), 7.67 (d, *J*=15.3 Hz, 1 H), 6.73 (d, *J*=16.0 Hz, 1 H), 6.22 (t, *J*=6.7 Hz, 1 H), 4.36 - 4.42 (m, 1 H), 3.86 - 3.96 (m, 2 H), 3.68 - 3.85 (m, 3 H), 2.75 (t, *J*=6.8 Hz, 2 H), 2.28 (dd, *J*=6.1, 3.7 Hz, 1 H), 2.14 - 2.24 ppm (m, 1 H). (ESI-TOF) m/z $[M - H]^-$ calculated for $C_{18}H_{20}N_3O_9^-$ 422.1, found 422.1.

General Method for Ester Hydrolysis:

To a round-bottomed flask were sequentially added compounds **7b**, **8b**, **10a-b**, **11**, or **12** (1 equiv), THF, and water. After five minutes of stirring, lithium hydroxide (3 equiv) was added and the reaction mixture was allowed to stir for five minutes at room temperature. The reaction mixture was then concentrated *in vacuo* to afford a crude oil. Purification *via* silica gel flash column chromatography (1 column volume of 0%, 3 column volumes from 0% to 20%, 3 column volumes of 20%, 3 column volumes from 20% to 40%, and 20 column volumes of 40% methanol in DCM) and HPLC purification using a Sunfire C18 column (5 μm, 150 × 2.0 mm, 10.0 mL/min flow rate, buffer A: H_2O), buffer B: ACN) involved a gradient of 0-30% B (0-30 min), of 30-95% B (30-40 min), followed by an isocratic hold at 95% B (40-45 min) afforded compounds afforded compounds **13a-b**, **14**, **15**, and **16**.

4-((3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)(hydroxy)amino)-4-oxobutanoic acid (**13a**): Light yellow solid (56% yield). R_f = 0.05 (40% methanol in DCM, $KMnO_4$ stain). ^{1}H NMR (500 MHz, MeOD) δ 2.13 (dt, *J* = 13.3, 6.5 Hz, 1H), 2.38 (ddd, *J* = 13.6, 6.1, 4.0 Hz, 1H), 2.58 (t, *J* = 6.8 Hz, 2H), 2.77 (t, *J* = 6.7 Hz, 2H), 3.73 (dd, *J* = 12.1, 3.7 Hz, 1H),

3.82 (dd, *J* = 12.0, 3.1 Hz, 1H), 3.94 (q, *J* = 3.5 Hz, 1H), 4.36 (dt, *J* = 6.4, 3.9 Hz, 1H), 4.60 (s, 2H), 6.20 (t, *J* = 6.3 Hz, 1H), 8.34 (s, 1H). ^{13}C NMR (126 MHz, MeOD) δ 28.7, 30.3, 31.7, 37.0, 40.1, 42.4, 62.5, 71.7, 75.3, 88.0, 89.0, 90.8, 92.2, 143.8, 146.2, 156.7, 166.4, 175.5. (ESI-TOF) m/z [M - H]$^-$ calculated for $C_{16}H_{19}N_4O_8^-$ 395.1, found 395.1.

4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobutanoic acid (**13b**): White solid (58% yield). R_f = 0.05 (40% methanol in DCM, $KMnO_4$ stain). ^{1}H NMR (500 MHz, MeOD) δ 2.13 (dt, *J* = 13.3, 6.5 Hz, 1H), 2.37 (ddd, *J* = 13.5, 6.1, 4.0 Hz, 1H), 2.55 (t, *J* = 7.0 Hz, 2H), 2.75 (q, *J* = 6.6 Hz, 4H), 3.70 – 3.90 (m, 4H), 3.92 – 3.96 (m, 1H), 4.36 (dt, *J* = 6.4, 3.9 Hz, 1H), 6.20 (t, *J* = 6.3 Hz, 1H), 8.24 (s, 1H). ^{13}C NMR (126 MHz, MeOD) δ 18.6, 29.2, 32.0, 42.4, 47.9, 49.9, 62.5, 71.8, 73.2, 87.9, 89.0, 93.19, 94.2, 145.2, 156.8, 166.6, 175.7. (ESI-TOF) m/z [M - H]$^-$ calculated for $C_{17}H_{21}N_4O_8^-$ 409.1, found 409.1.

(*E*)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoic acid (**14**): White solid (48% yield). R_f = 0.04 (40% methanol in DCM, $KMnO_4$ stain). ^{1}H NMR (METHANOL-d_4 ,500MHz): δ = 8.15 (s, 1 H), 6.77 (d, *J*=15.6 Hz, 1 H), 6.64 (d, *J*=15.6 Hz, 1 H), 6.14 (t, *J*=6.6 Hz, 1 H), 4.28 - 4.32 (m, 1 H), 3.83 (q, *J*=3.2 Hz, 1 H), 3.71 (dd, *J*=12.2, 3.1 Hz, 1 H), 3.64 (dd, *J*=12.1, 3.5 Hz, 1 H), 3.37 (td, *J*=6.6, 2.6 Hz, 2 H), 2.53 (t, *J*=6.7 Hz, 2 H), 2.19 (dd, *J*=6.1, 3.7 Hz, 1 H), 2.08 - 2.15 ppm (m, 1 H). ^{13}C NMR (126

MHz, MeOD) δ 20.8, 39.7, 41.7, 54.8, 62.6, 72.1, 74.0, 86.9, 89.2, 92.1, 100.7, 135.3, 144.9, 151.2, 164.8, 167.2. MS (ESI-TOF) m/z [M - H]$^-$ calculated for $C_{17}H_{20}N_4O_7^-$ 391.1, found 391.1.

(*E*)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobut-2-enoic acid (**15**): White solid (41% yield). R_f = 0.04 (40% methanol in DCM, $KMnO_4$ stain). ^{1}H NMR (METHANOL-d_4 ,500MHz): δ = 8.13 (s, 1 H), 7.34 (d, *J*=15.9 Hz, 1 H), 6.74 (d, *J*=15.6 Hz, 1 H), 6.10 (t, *J*=6.3 Hz, 1 H), 4.23 - 4.28 (m, 1 H), 3.78 - 3.94 (m, 4 H), 3.71 (dd, *J*=12.2, 2.7 Hz, 1 H), 3.63 (dd, *J*=12.1, 3.5 Hz, 1 H), 2.70 (t, *J*=6.3 Hz, 2 H), 2.22 - 2.30 (m, 1 H), 2.02 ppm (dt, *J*=13.4, 6.6 Hz, 1 H). ^{13}C NMR (METHANOL-d_4 ,126MHz): δ = 165.8, 163.4, 149.8, 143.5, 133.9, 99.3, 90.7, 87.7, 85.5, 72.5, 70.7, 61.2, 40.3, 38.2, 19.4 ppm. (ESI-TOF) m/z [M - H]$^-$ calculated for $C_{17}H_{19}N_4O_8^-$ 407.1, found 407.1.

(*E*)-4-((4-(1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoic acid (**16**): White solid (40% yield). R_f = 0.04 (40% methanol in DCM, $KMnO_4$ stain). ^{1}H NMR (METHANOL-d_4 ,500MHz): δ = 8.15 (s, 1 H), 6.70 (d, *J*=15.6 Hz, 1 H), 6.58 (d, *J*=15.6 Hz, 1 H), 6.15 (t, *J*=6.7 Hz, 1 H), 4.28 - 4.33 (m, 1 H), 3.83 (q, *J*=3.2 Hz, 1 H), 3.71 (dd, *J*=12.2, 3.1 Hz, 1 H), 3.65 (dd, *J*=12.1, 3.5 Hz, 1 H), 3.30 - 3.40 (m, 2 H), 2.52 (t, *J*=6.7

Hz, 2 H), 2.18 (dd, *J*=6.3, 3.5 Hz, 1 H), 2.07 - 2.15 ppm (m, 1 H). (ESI-TOF) m/z [M - H]$^-$ calculated for $C_{17}H_{18}N_3O_8^-$ 392.1, found 392.1.

(*E*)-4-(hydroxy(4-(1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoic acid (**17**). ^{1}H NMR (METHANOL-d_4 ,400MHz): δ = 8.21 (s, 1 H), 7.34 (d, *J*=14.1 Hz, 1 H), 6.85 (d, *J*=15.7 Hz, 1 H), 6.23 (t, *J*=6.7 Hz, 1 H), 4.40 (dt, *J*=6.1, 3.2 Hz, 1 H), 3.87 - 3.97 (m, 2 H), 3.70 - 3.86 (m, 3 H), 2.74 (br. s., 2 H), 2.25 - 2.32 (m, 1 H), 2.15 - 2.25 ppm (m, 1 H). (ESI-TOF) m/z [M - H]$^-$ calculated for $C_{17}H_{18}N_3O_9^-$ 408.1, found 408.1.

General Method for linker reduction:

To a hydrogenation flask were added **13a** (1 equiv), Pd/C (0.01 equiv) and MeOH (5 mL). The flask was purged with N_2 and was then placed under H_2 at a pressure of 45 bar on a hydrogenator. The reaction was shaken overnight. The reaction was then filtered over celite and concentrated *in vacuo* and subjected to HPLC purification using a Sunfire C18 column (5 μm, 150 × 2.0 mm, 10.0 mL/min flow rate, buffer A: H_2O), buffer B: ACN) involved a gradient of 0-30% B (0-30 min), of 30-95% B (30-40 min), followed by an isocratic hold at 95% B (40-45 min) to afford compound **18**.

4-((3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)propyl)(hydroxy)amino)-4-oxobutanoic acid (**18**): White solid (51% yield). ^{1}H NMR (METHANOL-d_4 ,500MHz): δ = 7.82 (s, 1 H), 6.18 (t, *J*=6.6 Hz, 1 H), 4.27 - 4.35 (m, 1 H), 3.81 (q, *J*=3.7 Hz, 1 H), 3.75 (dd, *J*=12.1, 3.2 Hz, 1 H), 3.65

(dd, *J*=12.1, 3.8 Hz, 1 H), 3.55 (t, *J*=6.3 Hz, 2 H), 2.58 - 2.63 (m, 2 H), 2.44 - 2.50 (m, 2 H), 2.19 - 2.26 (m, 3 H), 2.05 - 2.12 (m, 1 H), 1.69 - 1.78 ppm (m, 2 H). (ESI-TOF) m/z $[M - H]^-$ calculated for $C_{16}H_{23}N_4O_8^-$ 399.2, found 399.2.

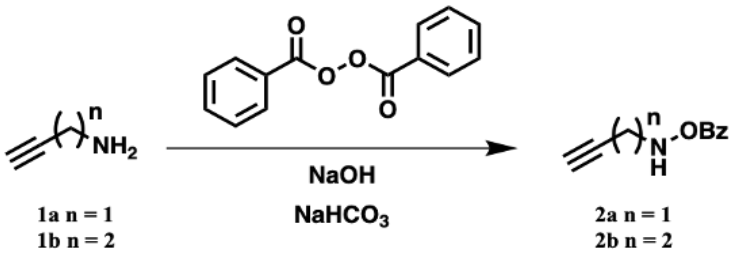


Figure 7.4. O-benzoylation reaction on linker region

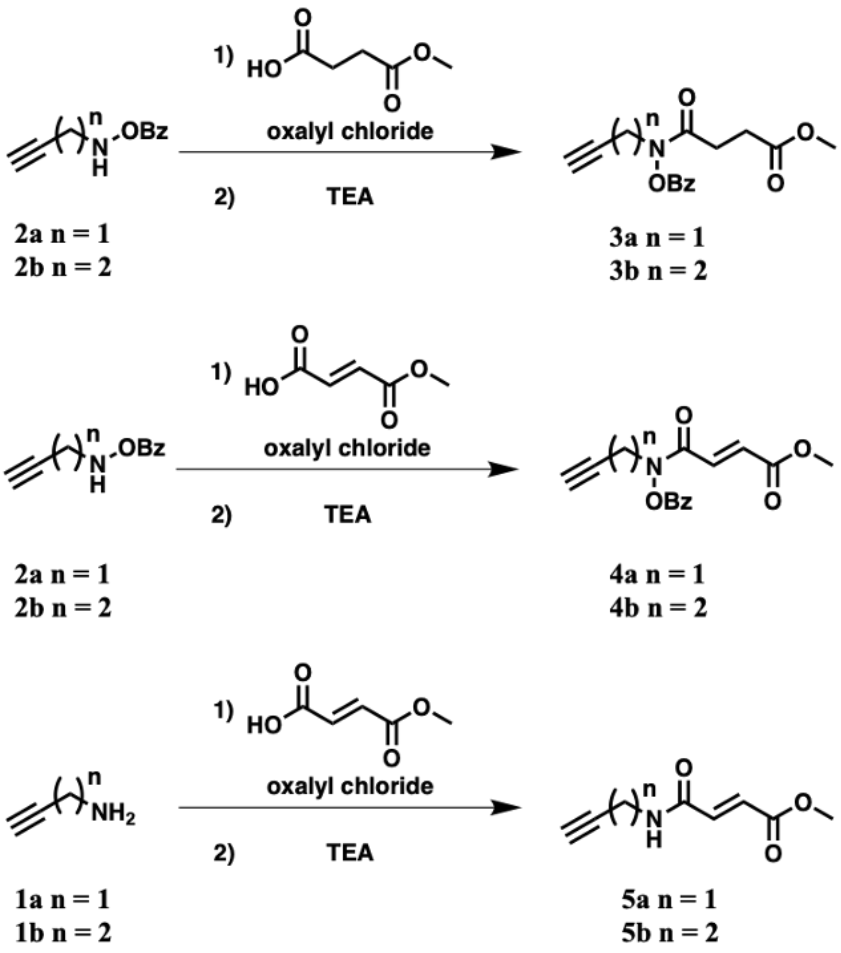


Figure 7.5. Linker cofactor synthesis

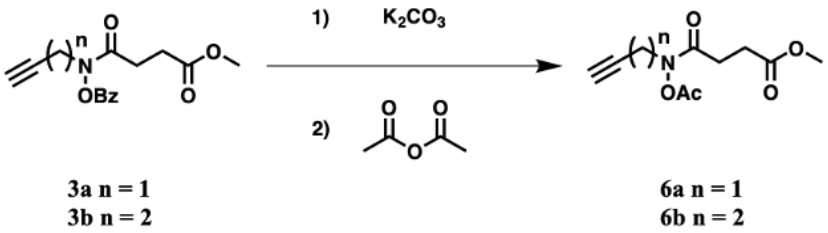


Figure 7.6. Linker cofactor O-Ac reprotection synthesis

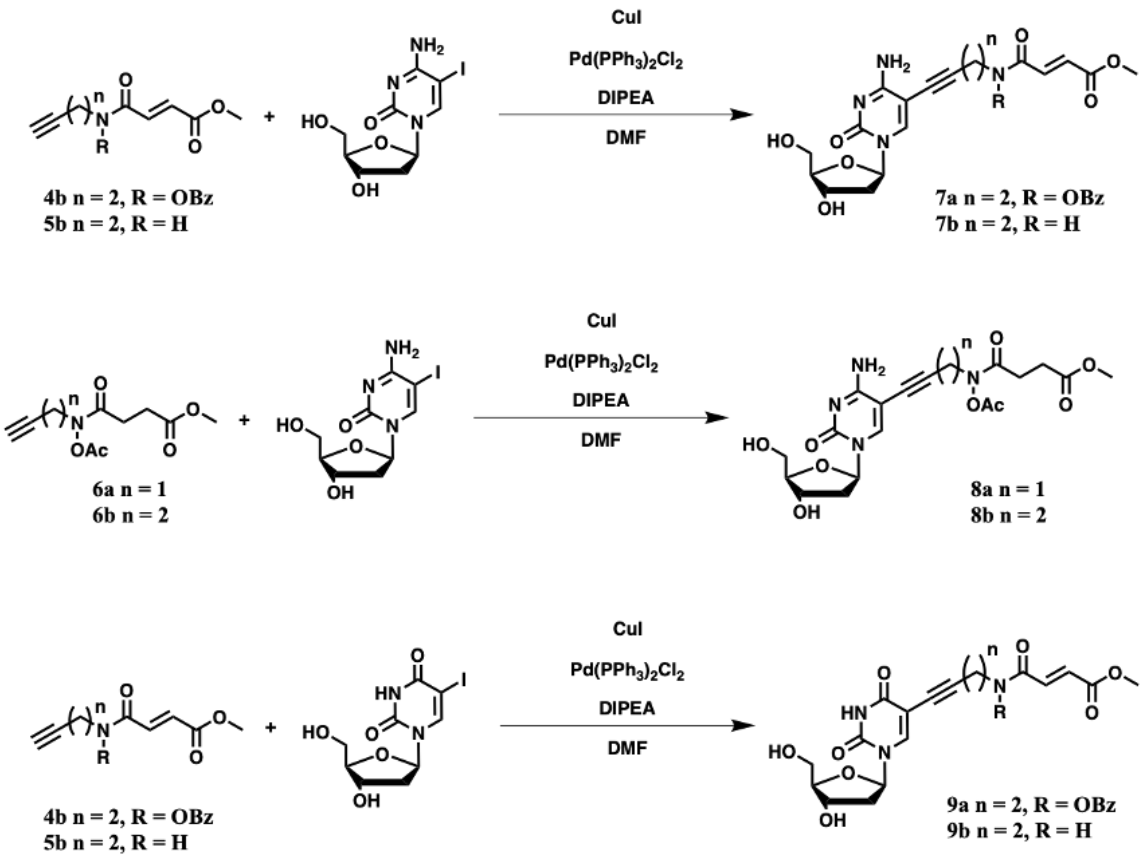


Figure 7.7. Conjunction of substrate and linker-cofactor synthesis

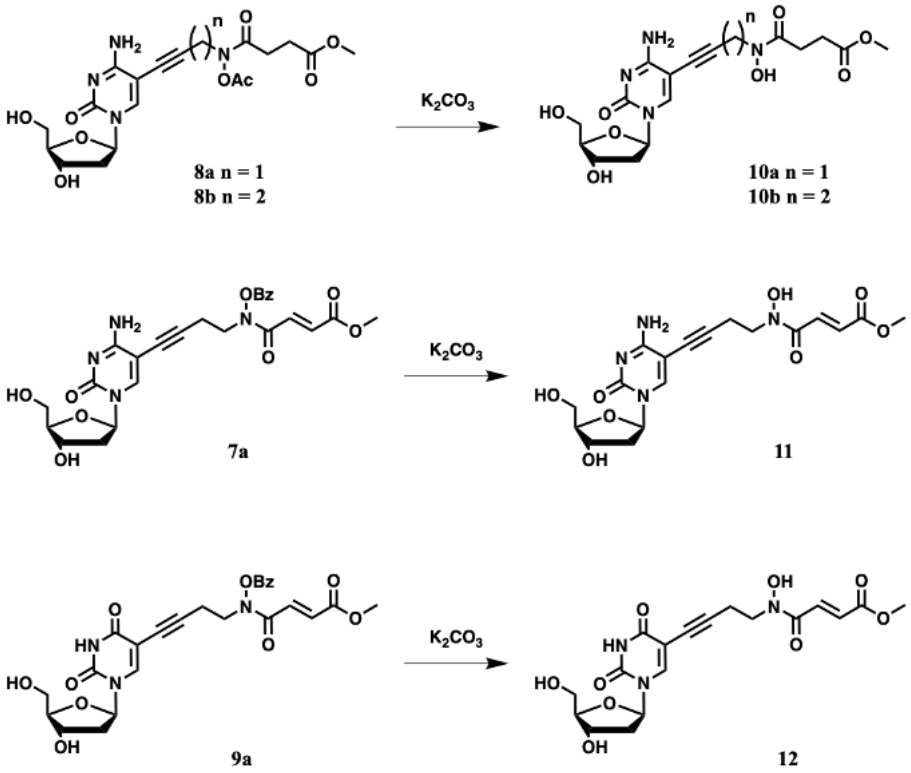


Figure 7.8. Selective deprotection of hydroxyamide moiety

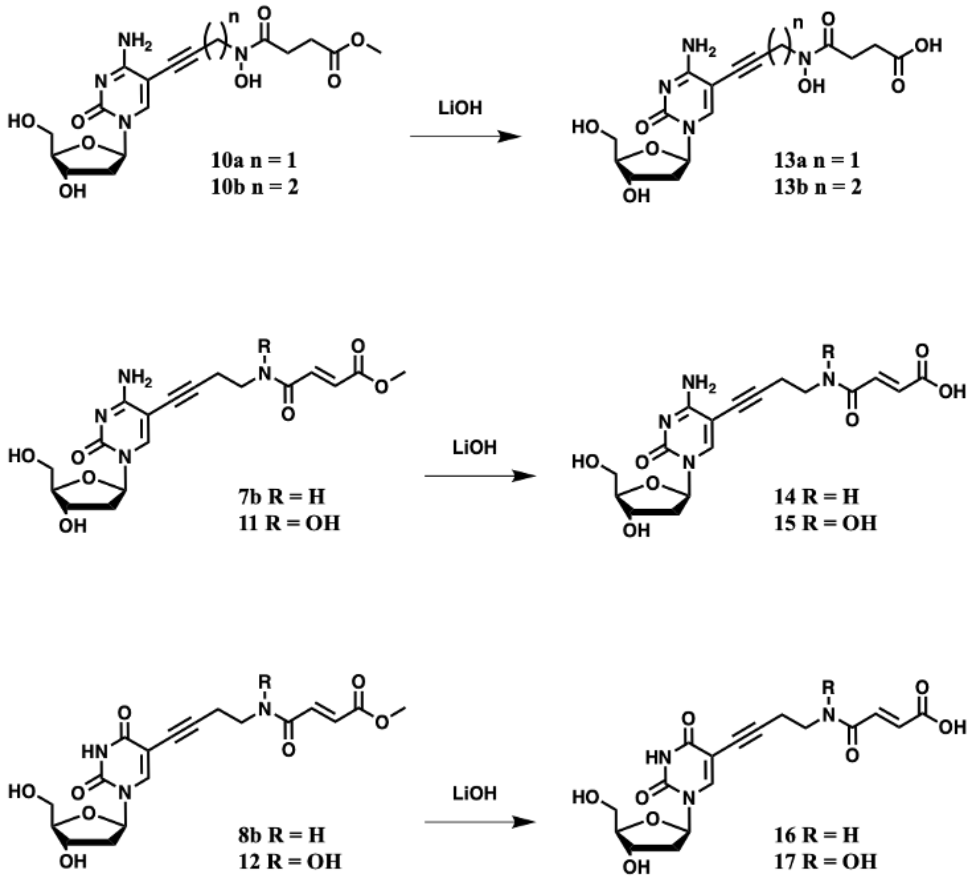


Figure 7.9. Desertification of bifunctional inhibitors

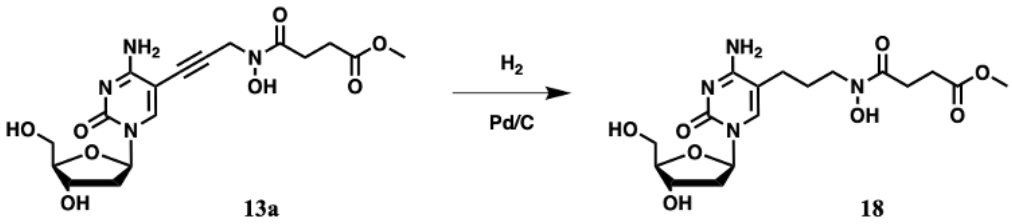


Figure 7.10. Reduction of linker

7.3 Results

Our initial objective was to design a bifunctional pan TET inhibitor by incorporating elements of both the substrate mC and alpha-ketoglutarate cofactor into the structure. We computationally docked a small library of rationally designed bifunctional molecules into the crystal structure of TET2 bound to a mC containing oligonucleotide (PDB: 4NM6) with the Schrodinger Maestro Glide program (**Figure 7.11**).[447] We explored varying the nucleobase, changing linker length, and installing a metal binding hydroxyamide moiety in our molecules. A tabulated list of representative docking scores is found in **Table 7.1**. Our results found that dC-based compounds were preferred over dU- and C-based substrate mimics, that incorporation of a hydroxyamide moiety improved molecular docking scores, and that a linker length of 4 carbons was optimal. We also observed that hydrogen bond interactions with Arg1896 and Arg1261 were important for the cofactor portion of the molecule and that pi-pi stacking with Tyr1902 and hydrogen bond interactions with Asn1397 were important for the substrate portion. We used this information to guide our synthetic efforts towards an initial series of inhibitors for TET activity screening.

The synthetic approach to access our bifunctional inhibitors is shown in **Figure 7.12A**. We envisioned key retrosynthetic disconnections at the nucleoside-linker and linker-cofactor junctions. Our route drew inspiration from bifunctional inhibitors targeting the Jumonji C domain-containing histone demethylases, which also use an alpha-ketoglutarate cofactor.[458] Detailed synthetic information can be found in the methodology section of this chapter. Briefly, for analogs containing an hydroxyamide

moiety, an O-Bz was installed in the linker section via benzoyl peroxide under basic conditions.[459] This alkyne linker was then acylated with the cofactor mimetic. Following construction of the protected cofactor portion of the molecule, the cofactor was linked to the desired iodinated nucleoside through a Sonogashira cross coupling reaction.[460] Through the course of our synthesis, we found that converting the O-Bz group to O-Ac in the linker cofactor precursors improved cross coupling yields and minimized side product formation when synthesizing compounds with 3 carbon linkers.[461, 462] Once the substrate and cofactor regions were linked, treatment with the appropriate base enabled isolation of the respective acid and methyl ester versions of the inhibitors. Our efforts thus far have produced an initial series of candidate TET inhibitors as shown in **Figure 7.12B**.

With inhibitors in hand, we moved to screen our molecules' TET inhibitory activity in a cellular assay. As an initial experiment, we treated K562 cells, a TET proficient leukemia model, with 5 μM of respective compound for 24 hours. Next, genomic DNA was extracted and 5hmC levels were quantified via dot blot (**Figure 7.13A** and **B**). The most active compounds were, 11, 13b 14, and 15 which lowered the relative levels of 5hmC by 30 ± 8%, 55 ± 20%, 26 ± 21%, and 47 ± 45%, respectively. These data revealed interesting SAR trends, namely that the substrate mimetic portion of the molecule required a deoxycytidine scaffold opposed to deoxyuridine and that a pentyne based linker was optimal, which correlated with our docking studies. However, given the large variability observed in these experiments we moved to confirm these results by mass spectrometry-based measurements.

The compounds 13b and 15 were taken forward and evaluated in an *in vitro* TET activity assay with the enzymatically active catalytic domain of human TET1 using a hemimethylated synthetic oligomer substrate. Following a 30 minute enzymatic incubation time in the presence or absence of inhibitor, the reactions were halted through denaturation and levels of 5hmC were quantified via LCMS/MS via direct isotope dilution.[402] As shown in **Figure 7.13**, both compounds 13b and 15 displayed dose-dependent TET inhibition. Of the two, the more potent compound 13b, reduced hmC production by 85 ± 3%, 30 ± 4%, 22 ± 5%, at 500 µM, 100 µM, 10 µM, respectively, in comparison to DMSO control. These data indicate that this class of bifunctional substrate cofactor mimic compounds possess mid µM TET1 inhibitory activity and can serve a starting point for continued SAR investigations for the development of more potent TET inhibitors.

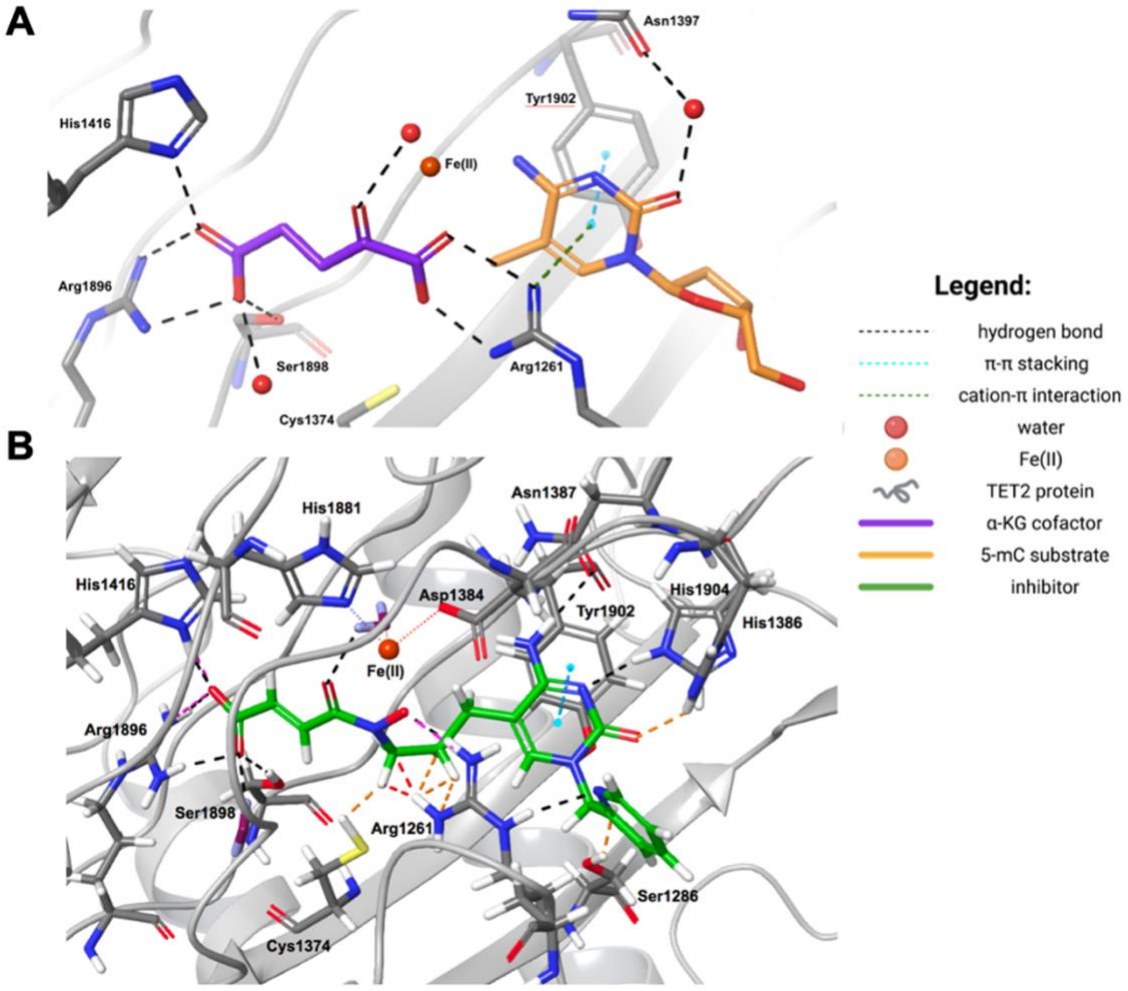


Figure 7.11. Computational docking of potential TET inhibitors

Visualization of mC and alpha-ketoglutarate docked into the active site of TET2 (PDB: 4NM6). B) Visualization of compound dC3RHAS (18) docked into the active site of TET2 (PDB: 4NM6). Legend defines observed interactions.

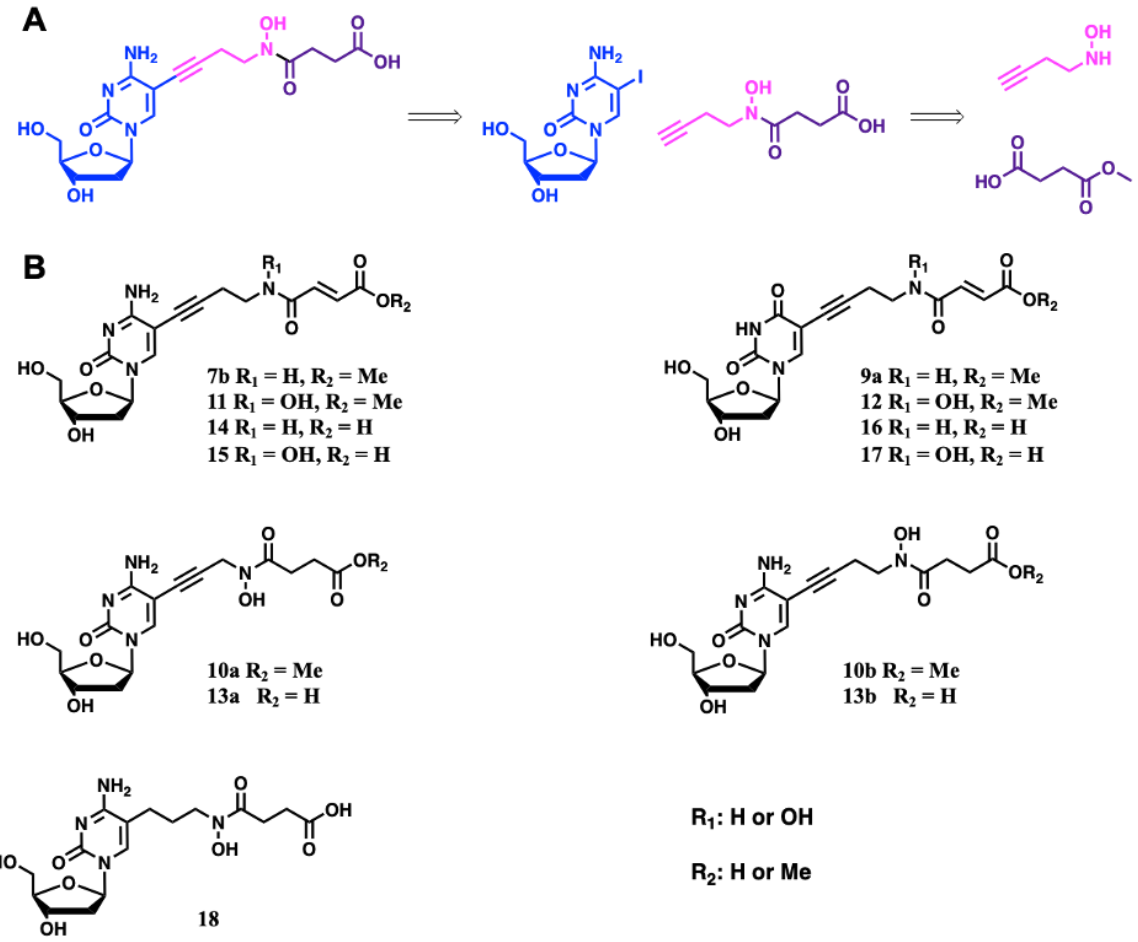


Figure 7.12. Synthetic approach to bifunctional TET inhibitors

A) Retrosynthetic disconnections for the substrate, linker, and cofactor portion of the molecule. B) Structure of synthesized inhibitors.

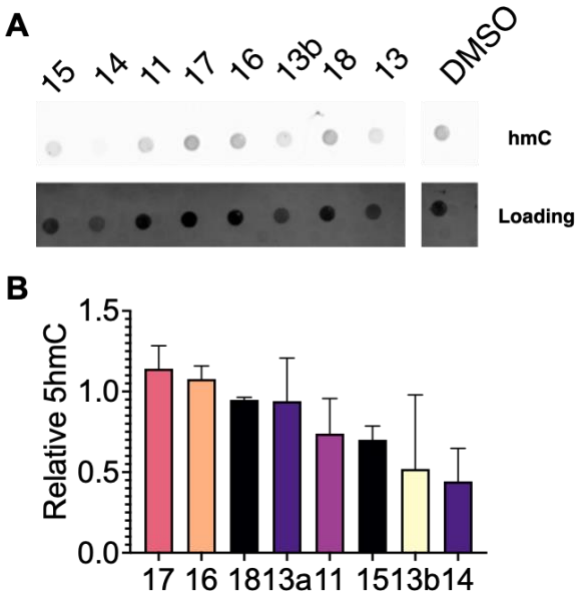


Figure 7.13. Preliminary screen of TET inhibitors in leukemia cells to assess effects on global hmC levels

A) Representative dot blot for 5hmC following treatment of K562 with 5 μM of specified inhibitor for 24 hours and extraction of cells genomic DNA; top: anti 5hmC antibody signal; bottom: methylene blue staining for loading control B) Quantified hmC blotting intensity via densitometry normalized to DNA loading and DMSO treatment. Data represent mean ± SD of two biological replicates.

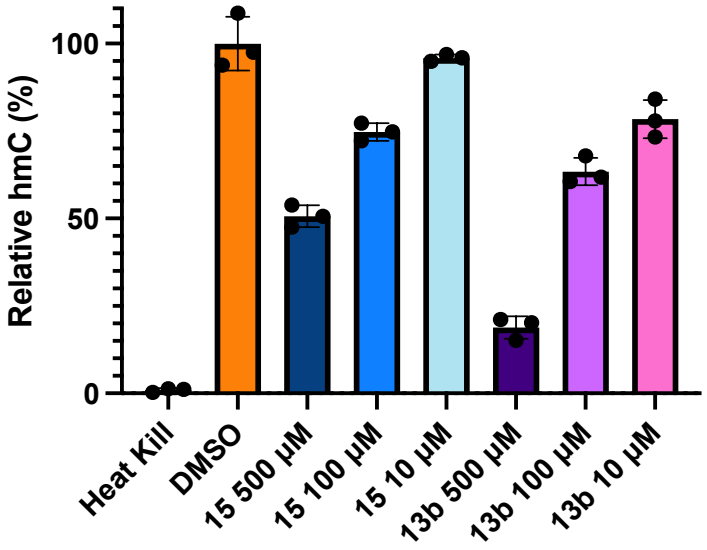
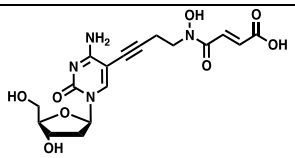
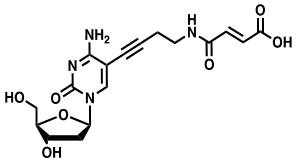
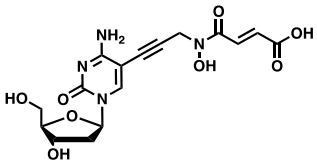
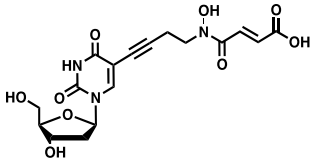
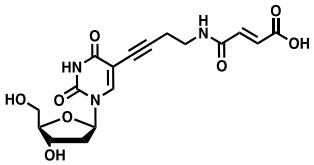
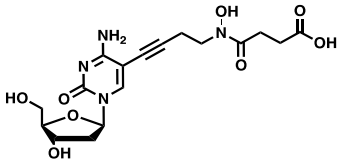
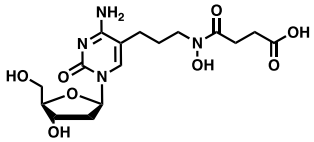


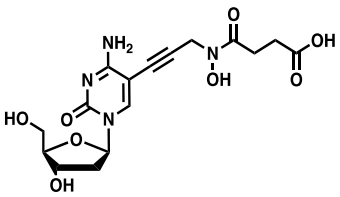
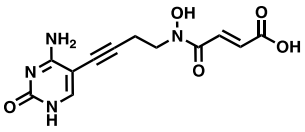
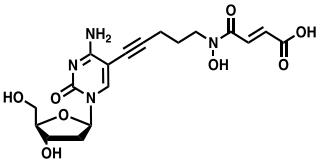
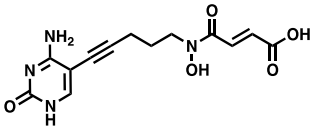
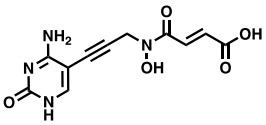
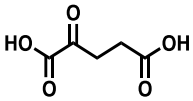
Figure 7.14. Effects of inhibitor treatment on recombinant TET1 activity

5mC containing DNA was incubated with the TET1 catalytic domain for 30 min in the presence or in the absence of inhibitor in DMSO. The formation of hmC was detected by a quantitative LC-MS assay and is displayed as % activity relative to DMSO controls. All data are averaged experimental triplicates normalized to DMSO ± SD.

Table 7.1. Tabulation of docking scores

Docking score magnitude suggests more favorable ligand interactions.

Compound	Structure	Docking Score
dC4HAF		**-12.387**
dC4AF		**-8.277**
dC3HAF		**-8.827**
dU4HAF		**-5.913**
dU4AF		**-7.201**
dC4HAS		**-6.251**
dC3RHAS		**-11.625**

dC3HAS		- **10.591**
C4HAF		- **10.563**
dC5HAS		- **6.872**
C5HAS		- **8.597**
C3HAS		- **7.065**
alpha-ketoglutarate		- **12.046**

7.4 Discussion

DNA methylation is an essential epigenetic process.[463] In turn, proper regulation of methyl marks is critical for maintaining a healthy phenotype. TET enzymes play key roles in controlling DNA demethylation and dictating phenotype, however a complete understanding of how TET activity is associated with cancer and other diseases is lacking.[438, 464] Thus potent and selective small molecule inhibitors of TET are important tools for fully elucidating its function. Our bifunctional design strategy outlined here provides new chemical space for TET inhibitors and such scaffolds will be the focus of future biological evaluation and SAR campaigns in our laboratory.

Our preliminary screening efforts have shown that our most active molecules have promising activity at 5 μM in cell-based assays. Interestingly, one of our most active compounds, 13b, possessed the most structural similarity to the native substrate and alpha-ketoglutarate cofactor within our inhibitor series, suggesting that the bifunctional approach taken here is a viable strategy for this class of enzyme. We also found that this compound inhibited TET1 function in an *in vitro* assay. However, further validation of these compounds is needed to confirm they are truly selective TET inhibitors and do not inhibit other αKG dependent enzymes.

In terms of *in vitro* experiments, full TET IC_{50} values need to be obtained against TET 1, 2, and 3 for our set of inhibitors. These experiments will assess any isoform selectivity our compounds may possess. We hypothesize that this series of compounds will be pan selective due to the high degree of active site homology across all three TETs.[407] However, the development of isoform selective compounds would provide a

valuable set of chemical probes to study the function of different TET enzymes. Competition based *in vitro* experiments with this class of bifunctional inhibitors and varying amounts of alpha-ketoglutarate and Cu would also be useful to assess potency. The cellular concentration of alpha-ketoglutarate is estimated to be in the range of 10 – 50 μM, however its levels can fluctuate in tumors and influence phenotype.[465, 466] Thus, determining the concentrations at which alpha-ketoglutarate can outcompete this class of inhibitors may be useful in assessing their efficacy. Additionally, our compounds should be screened against a panel of alpha-ketoglutarate enzymes like AlkB and other Jumonji C domain-containing demethylases to ensure these molecules are indeed selective for TET.[467] Furthermore, biophysical assays like isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), or microscale thermophoresis (MST) should be conducted.[468-470] These assays can be used to measure the K_d of these inhibitors, which will be critical in understanding and characterizing their affinity for TET. Obtaining an X-ray co-crystal structure with our inhibitors bound to TET would also provide additional binding site information and insight into the key molecular interactions these compounds have with the TET active site. Such information would be valuable for future SAR studies aimed at making more potent TET inhibitors.

Expanding the panel of cell lines used in cellular treatment experiments is also a key area to explore in future work. Previous studies focused on hematological cancer models have found that TET inhibitor treatment in cell lines harboring a TET2 deficiency resulted in selective lethality, suggesting that a particular cell's basal level of TET expression is critical in determining its phenotypic response to inhibitor treatment.[471] In

lung cancer, aberrant DNA methylation and TET expression has been reported, thus assessing the effect of this class of bifunctional inhibitors across a variety of lung cancer cell models would provide deeper mechanistic insight into the role that TET plays in cancer development and progression.[472-477] Quantifying cell viability, assessing transcriptional changes via RNA seq, and profiling changes in protein expression through global proteomics upon inhibitor treatment would provide useful data on this front. Additionally, mapping the locational changes of mC and hmC marks across the genome through reduced representation bisulfite sequencing (RRBS) following inhibitor treatment would give additional insight into how inhibiting TET function influences gene expression beyond measuring global levels of these epigenetic marks.[478]

In conclusion, we have detailed the design and synthesis of a novel class of TET inhibitors which utilize a bifunctional cofactor substrate mimetic design. This class of molecules shows activity in preliminary assays and lays the foundation for future evaluation and SAR studies. Ultimately, small molecule TET inhibitors are important chemical tools for studying the complex biological roles TET enzymes play and further investigation into this area is expected to provide mechanistic insights into cancer development and influence the design of potential epigenetic-based therapeutics.

8 Chapter VIII: Conclusion

In **Chapter II** of this thesis, we designed and a novel DA-mimetic probe (DA^{yne}) containing a bio-orthogonal alkyne handle that exhibits a reactivity profile similar to DA in aqueous buffers. By linking DPAs formed with DA^{yne} to a fluorescent reporter molecule, DPAs were visualized in fixed cells and within lysates. DA^{yne} enabled global mapping of cellular proteins affected by DQ modification and their bioactive pathways through enrichment. Our proteomic profiling of DPAs in neuronal SH-SY5Y cells indicates that proteins susceptible to DPA formation are extant throughout the proteome, potentially influencing several diverse biological pathways involved in PD such as endoplasmic reticulum stress, cytoskeletal instability, proteotoxicity, and clathrin function. We validated that a protein involved in the ER stress pathway, protein disulfide isomerase 3 (PDIA3), which was enriched in our chemoproteomic analysis, is functionally inhibited by DA, providing evidence that dysregulated cellular DA may induce or exacerbate ER stress. Thus, DA^{yne} provided new mechanistic insights into DA toxicity that may be observed during PD by enabling characterization of DPAs generated reproducibly at physiologically relevant quinone exposures. We anticipate our design and application of this reactivity- based probe will be generally applicable for clarifying mechanisms of metabolic quinone toxicity.

In **Chapter III**, we designed photo releasable DA probes which can be targeted to desired subcellular locations to release DA on demand in defined cellular regions. We envision that such a controlled release platform will aid in elucidating the precise mechanisms by which DA influences cellular phenotype and signaling events.

Chapter IV and **V** of this thesis focuses on studying the reactive metabolite MGO. Currently, there is a lack of chemical tools capable of studying protein adducts caused by this class of reactive electrophile species. In **Chapter IV** of this thesis, we report a chemical biology platform, termed T-DiP (targetable-dicarbonyl precursor), that releases a physiologically relevant dose of bio-orthogonally functionalized dicarbonyl probe upon irradiation with 365 nm light. This approach enables protein glycation to be controlled with spatiotemporal precision within live cells. Profiling and bioinformatic analysis of proteins adducted by the bio-orthogonally functionalized dicarbonyl probe under controlled release conditions suggested that dicarbonyl metabolites like MGO may affect ubiquitin ligase protein binding pathways. This led us to find that UBA1 is indeed susceptible to MGO modification *in vitro*, signifying that MGO may exert toxicity through inhibiting protein ubiquitin pathways. Collectively, this work expands the chemical toolbox needed to elucidate the roles of glycated proteins across various pathologies.

Recently, MGO was shown to crosslink deoxyguanosine with acetyl-lysine, thereby covalently crosslinking DNA to DNA-binding proteins. Therefore, in **Chapter V** of this thesis, we characterized DNA–protein cross-linking in human cells exposed to elevated levels of MGO. We coupled cell culture treatments and DPC purification assays to characterize MGO-dependent DPC formation and utilized mass spectrometry-based proteomics to identify and characterize proteins trapped to DNA upon MGO exposure. This study revealed 265 proteins participating in MGO derived DPC formation. Subsequent bioinformatic analysis revealed that these DPCs may play a role in telomeric

organization, prompting us to validate DPC formation between GAPDH and telomeric DNA in the presence of MGO. Ultimately, this work provides a foundational set of MGO derived DPCs to be followed up on in future studies and provides insight into potential DPCs that may be relevant in diseases linked to elevated MGO levels.

Chapter VI and **VII** of this thesis details our work studying the TET family of enzymes. These dioxygenases catalyze the oxidation of 5-methylcytosine (5mC) in DNA to hydroxymethyl-C, formyl-C, and carboxy-C, leading to gene activation and epigenetic remodeling. Aberrant TET function is linked to many different cancers, necessitating the development of small molecule inhibitors to study TET biology and for use as potential therapeutics. Recently, a novel cytosine-based inhibitor of TET, Bobcat339, was reported to have mid μM inhibitory activity against TET1 and TET2. The molecule is now sold as a TET inhibitor by several vendors. We independently prepared Bobcat339 in our laboratory and observed that it had minimal inhibitory activity against human TET1 and TET2 via a quantitative LC-ESI-MS/MS assay. Furthermore, the inhibitory activity of commercial Bobcat339 preparations was directly correlated with Cu(II) content. We therefore conclude that Bobcat339 alone is not capable of inhibiting TET enzymes at the reported concentrations, and that its activity is enhanced by contaminating Cu(II) which is an important discovery for researchers studying TET activating using this compound.

In **Chapter VII** of this thesis, we also disclose our initial efforts towards developing potent and selective TET inhibitors. Here we utilized structure-based design to construct a set of novel TET-specific inhibitors. We implemented a bifunctional strategy to simultaneously engage the enzyme's substrate and cofactor binding sites. We

utilized *in silico* molecular docking results to guide our synthetic effort, where we successfully accessed an initial series of potential inhibitors. These compounds were evaluated as TET inhibitors through a cellular based TET activity assay in a leukemia model and our most promising compounds lowered 5-hydroxymethylcytosine (5-hmC) levels upon 5 μM treatment, indicating that this class of molecules has the potential to inhibit TET proteins in cells. Collectively, this work represents a critical first step in developing potent and selective TET inhibitors to serve as epigenetic modulators to elucidate the role of DNA methylation in human disease.

9 Chapter IX: Future Directions

As described in **Chapters I**, **II**, and **III**, DA dysregulation may play an important role in driving PD pathology. Our work presented in **Chapter II** provides a first generation chemoproteomic tool for studying protein adducts caused by reactive DA metabolites. Future work could utilize this probe to enrich proteins susceptible to DA damage in primary cells or neuronal tissue cross sections. Such systems could provide a more representative model of PD pathology and be used in comparative studies which assess differences in DA modified proteins between cells derived from healthy and PD patients. Additionally, future work could explore the design and synthesis of chemoproteomic probes which mimic specific DA oxidation states to determine if there are differences in protein adducts derived from specific metabolites. These aforementioned studies could also be paired with SILAC or isobaric peptide tag proteomic strategies to improve protein quantification between control and treatment groups.[479, 480]

Chapter III describes our efforts in developing photoactivatable DA probes. Future work on this front could include continued design and synthesis of photocaged DA molecules with rapid release kinetics and application of our TRDP platform in neuronal cell models to study DA signaling. Continued effort in DA probe development is expected to yield deeper insights into DA dependent biology.

As detailed in **Chapters IV** and **V**, MGO is an electrophilic α-oxoaldehyde byproduct of cellular metabolism and can react with nucleophilic biomolecules to form a variety of covalent adducts in a process defined as glycation. Increased levels of MGO

leads to formation of protein and DNA adducts, which is associated with elevated inflammation signaling and the progression of diabetes, cancer, renal failure, and neurodegenerative diseases. Our work presented in **Chapter IV** yielded a novel chemical biology platform capable of directing intracellular MGO release with spatiotemporal control which we used to identify resultant protein adducts when release was targeted to the cytosol. Future work could utilize this platform by expanding profiling efforts to other areas of the cell like the nucleus, mitochondria, and ER. Additionally, the photocage in our first-generation T-DiP platform could be optimized to improve photolysis efficiency. Furthermore, the alkyne portion of the probe could be removed and replaced with an isotopically labeled MGO precursor to enable identification by mass spectrometry while mimicking the physiochemical properties of MGO more faithfully.

Chapter V of this thesis presents results gathered from profiling MGO induced DPCs in HT1080 cells. Subsequent experiments that could be pursued include, the profiling of MGO derived DPCs across various cell lines and disease models at physiologically relevant MGO concentrations to determine if cell type impacts MGO DPC formation and studying the repair pathways involved for removing MGO induced DPCs. Additionally, synthesis of an isotopically labeled MGO-Lys-dG standard for accurate quantification of MGO induced DPCs would be a valuable tool to enable the measurement of this adduct in human samples by mass spectrometry to establish the relevance of this adduct as a potential biomarker. Continued study of the biological impact MGO exerts on cells through chemical tools and investigations into MGO derived

DPCs will expand our understanding of the molecular mechanisms that underlie the abundant disease associations that MGO possesses.

As discussed in **Chapters VI** and **VII**, TET enzymes are important regulators of DNA methylation status and are associated with various cancers. Thus, development of selective and potent inhibitors against TET enzymes would assist in clarifying TETs' role in disease. One area of future work with our series of TET inhibitors could be continued SAR studies. Specifically, designing new molecules which replace the nucleoside portion and carboxylic acid region in the current structures through scaffold hopping and bioisosteric replacement may prove advantageous in terms of improving any metabolic liabilities this initial series may have.[481, 482] Additionally, continued *in vitro* and cellular analysis of these inhibitors should be conducted to confirm the selectivity of these compounds. One interesting experiment that could assess the selectivity and target engagement of these inhibitors in a cellular context is thermal protein profiling, which would allow for any proteins that interact with these inhibitors to be identified and paneled as off-targets.[483, 484]

Also, other approaches beyond occupancy-based-pharmacology methods to inhibit cellular TET DNA demethylation could be explored. Such alternative strategies could utilize proteolysis-targeting chimera (PROTAC) based technologies to selectively degrade TET enzymes and thereby halt any TET mediated DNA demethylation.[485, 486] In this case, PROTAC ligands could be designed to target unique allosteric sites on TET1 TET2 and TET3 and thereby gain selectivity for each isoform.[487]

Long term, TET inhibitors could be utilized to explore TETs' biological role in cancer progression through a battery of phenotypic cell based assays across multiple cell lines. TET inhibitors could also be applied in xenograft models to assess any potential therapeutic value that TET inhibitors may have against particular cancer indications. Continued development of TET inhibitors is expected to improve our mechanistic insight of how epigenetic dysregulation is linked to cancer initiation and progression as well as offer a new class of potential therapeutics.

10 Bibliography

(1) Dorsey, E. R.; Elbaz, A.; Nichols, E.; Abbasi, N.; Abd-Allah, F.; Abdelalim, A.; Adsuar, J. C.; Ansha, M. G.; Brayne, C.; Choi, J.-Y. J.; et al. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Neurology* **2018**, *17* (11), 939-953.

(2) Dickson, D. W. Parkinson's Disease and Parkinsonism: Neuropathology. *Cold Spring Harbor Perspectives in Medicine* **2012**, *2* (8), a009258-a009258.

(3) Poewe, W.; Seppi, K.; Tanner, C. M.; Halliday, G. M.; Brundin, P.; Volkmann, J.; Schrag, A. E.; Lang, A. E. Parkinson disease. *Nat Rev Dis Primers* **2017**, *3*, 17013.

(4) Kordower, J. H.; Olanow, C. W.; Dodiya, H. B.; Chu, Y.; Beach, T. G.; Adler, C. H.; Halliday, G. M.; Bartus, R. T. Disease duration and the integrity of the nigrostriatal system in Parkinson's disease. *Brain* **2013**, *136* (8), 2419-2431.

(5) Stokes, A. H.; Hastings, T. G.; Vrana, K. E. Cytotoxic and genotoxic potential of dopamine. *Journal of neuroscience research* **1999**, *55* (6), 659-665.

(6) Carlsson, A.; Lindqvist, M.; Magnusson, T.; Waldeck, B. On the Presence of 3-Hydroxytyramine in Brain. *Science (American Association for the Advancement of Science)* **1958**, *127* (3296), 471-471.

(7) Iversen, S. D.; Iversen, L. L. Dopamine: 50 years in perspective. *Trends Neurosci* **2007**, *30* (5), 188-193.

(8) Bucolo, C.; Leggio, G. M.; Drago, F.; Salomone, S. Dopamine outside the brain: The eye, cardiovascular system and endocrine pancreas. *Pharmacol Ther* **2019**, *203*, 107392.

(9) Ehringer, H.; Hornykiewicz, O. [Distribution of noradrenaline and dopamine (3-hydroxytyramine) in the human brain and their behavior in diseases of the extrapyramidal system]. *Klin Wochenschr* **1960**, *38*, 1236-1239.

(10) Cotzias, G. C.; Van Woert, M. H.; Schiffer, L. M. Aromatic amino acids and modification of parkinsonism. *N Engl J Med* **1967**, *276* (7), 374-379.

(11) Fahn, S.; Oakes, D.; Shoulson, I.; Kieburtz, K.; Rudolph, A.; Lang, A.; Olanow, C. W.; Tanner, C.; Marek, K. Levodopa and the progression of Parkinson's disease. *N Engl J Med* **2004**, *351* (24), 2498-2508.

(12) Mosharov, E. V.; Borgkvist, A.; Sulzer, D. Presynaptic effects of levodopa and their possible role in dyskinesia. *Mov Disord* **2015**, *30* (1), 45-53.

(13) Cilia, R.; Akpalu, A.; Sarfo, F. S.; Cham, M.; Amboni, M.; Cereda, E.; Fabbri, M.; Adjei, P.; Akassi, J.; Bonetti, A.; et al. The modern pre-levodopa era of Parkinson's disease: insights into motor complications from sub-Saharan Africa. *Brain* **2014**, *137* (Pt 10), 2731-2742.

(14) Olanow, C. W. Levodopa: effect on cell death and the natural history of Parkinson's disease. *Mov Disord* **2015**, *30* (1), 37-44.

(15) Monzani, E.; Nicolis, S.; Dell'Acqua, S.; Capucciati, A.; Bacchella, C.; Zucca, F. A.; Mosharov, E. V.; Sulzer, D.; Zecca, L.; Casella, L. Dopamine, Oxidative Stress and Protein–Quinone Modifications in Parkinson's and Other Neurodegenerative Diseases. *Angewandte Chemie International Edition* **2019**, *58* (20), 6512-6527.

(16) Eisenhofer, G.; Kopin, I. J.; Goldstein, D. S. Catecholamine Metabolism: A Contemporary View with Implications for Physiology and Medicine. *Pharmacological Reviews* **2004**, *56* (3), 331-349.

(17) Casida, J. E.; Ford, B.; Jinsmaa, Y.; Sullivan, P.; Cooney, A.; Goldstein, D. S. Benomyl, Aldehyde Dehydrogenase, DOPAL, and the Catecholaldehyde Hypothesis for the Pathogenesis of Parkinson's Disease. *Chemical Research in Toxicology* **2014**, *27* (8), 1359-1361.

(18) Daubner, S. C.; Le, T.; Wang, S. Tyrosine hydroxylase and regulation of dopamine synthesis. *Archives of Biochemistry and Biophysics* **2011**, *508* (1), 1-12.

(19) Montioli, R.; Cellini, B.; Dindo, M.; Oppici, E.; Voltattorni, C. B. Interaction of Human Dopa Decarboxylase with L-Dopa: Spectroscopic and Kinetic Studies as a Function of pH. *BioMed Research International* **2013**, *2013*, 1-10.

(20) Vendelboe, T. V.; Harris, P.; Zhao, Y.; Walter, T. S.; Harlos, K.; El Omari, K.; Christensen, H. E. The crystal structure of human dopamine β-hydroxylase at 2.9 Å resolution. *Science advances* **2016**, *2* (4), e1500980.

(21) Edmondson, D. E.; Mattevi, A.; Binda, C.; Li, M.; Hubalek, F. Structure and mechanism of monoamine oxidase. *Current medicinal chemistry* **2004**, *11* (15), 1983-1993.

(22) Graves, S. M.; Xie, Z.; Stout, K. A.; Zampese, E.; Burbulla, L. F.; Shih, J. C.; Kondapalli, J.; Patriarchi, T.; Tian, L.; Brichta, L.; et al. Dopamine metabolism by a monoamine oxidase mitochondrial shuttle activates the electron transport chain. *Nat Neurosci* **2020**, *23* (1), 15-20.

(23) Grünblatt, E.; Riederer, P. Aldehyde dehydrogenase (ALDH) in Alzheimer's and Parkinson's disease. *Journal of Neural Transmission* **2016**, *123* (2), 83-90.

(24) Vidgren, J.; Svensson, L. A.; Liljas, A. Crystal structure of catechol O-methyltransferase. *Nature* **1994**, *368* (6469), 354-358.

(25) Lambert, G.; Eisenhofer, G.; Jennings, G.; Esler, M. Regional homovanillic acid production in humans. *Life sciences* **1993**, *53* (1), 63-75.

(26) Meiser, J.; Weindl, D.; Hiller, K. Complexity of dopamine metabolism. *Cell Communication and Signaling* **2013**, *11* (1), 1-18.

(27) Elsworth, J. D.; Roth, R. H. Dopamine synthesis, uptake, metabolism, and receptors: relevance to gene therapy of Parkinson's disease. *Exp Neurol* **1997**, *144* (1), 4-9.

(28) Jinsmaa, Y.; Florang, V. R.; Rees, J. N.; Anderson, D. G.; Strack, S.; Doorn, J. A. Products of Oxidative Stress Inhibit Aldehyde Oxidation and Reduction Pathways in Dopamine Catabolism Yielding Elevated Levels of a Reactive Intermediate. *Chemical Research in Toxicology* **2009**, *22* (5), 835-841.

(29) Maardh, G.; Vallee, B. L. Human class I alcohol dehydrogenases catalyze the interconversion of alcohols and aldehydes in the metabolism of dopamine. *Biochemistry* **1986**, *25* (23), 7279-7282.

(30) Eiden, L. E.; Schäfer, M. K.-H.; Weihe, E.; Schütz, B. The vesicular amine transporter family (SLC18): amine/proton antiporters required for vesicular accumulation and regulated exocytotic secretion of monoamines and acetylcholine. *Pflügers Archiv* **2004**, *447* (5), 636-640.

(31) Mulvihill, K. G. Presynaptic regulation of dopamine release: Role of the DAT and VMAT2 transporters. *Neurochemistry international* **2019**, *122*, 94-105.

(32) Lohr, K. M.; Masoud, S. T.; Salahpour, A.; Miller, G. W. Membrane transporters as mediators of synaptic dopamine dynamics: implications for disease. *European Journal of Neuroscience* **2017**, *45* (1), 20-33.

(33) Blakely, R. D.; Edwards, R. H. Vesicular and plasma membrane transporters for neurotransmitters. *Cold Spring Harbor perspectives in biology* **2012**, *4* (2), a005595.

(34) Staal, R. G.; Mosharov, E. V.; Sulzer, D. Dopamine neurons release transmitter via a flickering fusion pore. *Nature neuroscience* **2004**, *7* (4), 341-346.

(35) Omiatek, D. M.; Bressler, A. J.; Cans, A.-S.; Andrews, A. M.; Heien, M. L.; Ewing, A. G. The real catecholamine content of secretory vesicles in the CNS revealed by electrochemical cytometry. *Scientific reports* **2013**, *3* (1), 1-6.

(36) Hovde, M. J.; Larson, G. H.; Vaughan, R. A.; Foster, J. D. Model systems for analysis of dopamine transporter function and regulation. *Neurochemistry international* **2019**, *123*, 13-21.

(37) Zecca, L.; Tampellini, D.; Gerlach, M.; Riederer, P.; Fariello, R.; Sulzer, D. Substantia nigra neuromelanin: structure, synthesis, and molecular behaviour. *Molecular Pathology* **2001**, *54* (6), 414.

(38) Zhang, S.; Wang, R.; Wang, G. Impact of dopamine oxidation on dopaminergic neurodegeneration. *ACS Chemical Neuroscience* **2018**, *10* (2), 945-953.

(39) Galzigna, L.; De Iuliis, A.; Zanatta, L. Enzymatic dopamine peroxidation in substantia nigra of human brain. *Clinica chimica acta* **2000**, *300* (1-2), 131-138.

(40) Hastings, T. G. Enzymatic oxidation of dopamine: the role of prostaglandin H synthase. *Journal of neurochemistry* **1995**, *64* (2), 919-924.

(41) Senoh, S.; Witkop, B. Formation and Rearrangements of Aminochromes from a New Metabolite of Dopamine and Some of its Derivatives1. *Journal of the American Chemical Society* **1959**, *81* (23), 6231-6235.

(42) Segura-Aguilar, J.; Paris, I.; Muñoz, P.; Ferrari, E.; Zecca, L.; Zucca, F. A. Protective and toxic roles of dopamine in Parkinson's disease. *Journal of Neurochemistry* **2014**, *129* (6), 898-915.

(43) Napolitano, A.; Crescenzi, O.; Pezzella, A.; Prota, G. Generation of the neurotoxin 6-hydroxydopamine by peroxidase/H2O2 oxidation of dopamine. *Journal of medicinal chemistry* **1995**, *38* (6), 917-922.

(44) Pezzella, A.; d'Ischia, M.; Napolitano, A.; Misuraca, G.; Prota, G. Iron-mediated generation of the neurotoxin 6-hydroxydopamine quinone by reaction of fatty acid hydroperoxides with dopamine: a possible contributory mechanism for neuronal degeneration in Parkinson's disease. *Journal of medicinal chemistry* **1997**, *40* (14), 2211-2216.

(45) Bisaglia, M.; Mammi, S.; Bubacco, L. Kinetic and structural analysis of the early oxidation products of dopamine: analysis of the interactions with α-synuclein. *Journal of Biological Chemistry* **2007**, *282* (21), 15597-15605.

(46) Sulzer, D.; Bogulavsky, J.; Larsen, K. E.; Behr, G.; Karatekin, E.; Kleinman, M. H.; Turro, N.; Krantz, D.; Edwards, R. H.; Greene, L. A. Neuromelanin biosynthesis is

driven by excess cytosolic catecholamines not accumulated by synaptic vesicles. *Proceedings of the National Academy of Sciences* **2000**, *97* (22), 11869-11874.

(47) Anderson, D. G.; Mariappan, S. S.; Buettner, G. R.; Doorn, J. A. Oxidation of 3, 4-dihydroxyphenylacetaldehyde, a toxic dopaminergic metabolite, to a semiquinone radical and an ortho-quinone. *Journal of Biological Chemistry* **2011**, *286* (30), 26978-26986.

(48) Bisaglia, M.; Greggio, E.; Beltramini, M.; Bubacco, L. Dysfunction of dopamine homeostasis: clues in the hunt for novel Parkinson's disease therapies. *The FASEB Journal* **2013**, *27* (6), 2101-2110.

(49) Caudle, W. M.; Colebrooke, R. E.; Emson, P. C.; Miller, G. W. Altered vesicular dopamine storage in Parkinson's disease: a premature demise. *Trends in neurosciences* **2008**, *31* (6), 303-308.

(50) Pifl, C.; Rajput, A.; Reither, H.; Blesa, J.; Cavada, C.; Obeso, J. A.; Rajput, A. H.; Hornykiewicz, O. Is Parkinson's disease a vesicular dopamine storage disorder? Evidence from a study in isolated synaptic vesicles of human and nonhuman primate striatum. *Journal of Neuroscience* **2014**, *34* (24), 8210-8218.

(51) Asanuma, M.; Miyazaki, I.; Ogawa, N. Dopamine- or L-DOPA-induced neurotoxicity: The role of dopamine quinone formation and tyrosinase in a model of Parkinson’s disease. *Neurotoxicity Research* **2003**, *5* (3), 165-176.

(52) Miyazaki, I.; Asanuma, M. Dopaminergic neuron-specific oxidative stress caused by dopamine itself. *Acta Medica Okayama* **2008**, *62* (3), 141-150.

(53) Lohr, K. M.; Bernstein, A. I.; Stout, K. A.; Dunn, A. R.; Lazo, C. R.; Alter, S. P.; Wang, M.; Li, Y.; Fan, X.; Hess, E. J. Increased vesicular monoamine transporter

enhances dopamine release and opposes Parkinson disease-related neurodegeneration in vivo. *Proceedings of the National Academy of Sciences* **2014**, *111* (27), 9977-9982.

(54) Glatt, C. E.; Wahner, A. D.; White, D. J.; Ruiz-Linares, A.; Ritz, B. Gain-of-function haplotypes in the vesicular monoamine transporter promoter are protective for Parkinson disease in women. *Human molecular genetics* **2006**, *15* (2), 299-305.

(55) Burbulla, L. F.; Song, P.; Mazzulli, J. R.; Zampese, E.; Wong, Y. C.; Jeon, S.; Santos, D. P.; Blanz, J.; Obermaier, C. D.; Strojny, C. Dopamine oxidation mediates mitochondrial and lysosomal dysfunction in Parkinson's disease. *Science* **2017**, *357* (6357), 1255-1261.

(56) Sulzer, D. Multiple hit hypotheses for dopamine neuron loss in Parkinson's disease. *Trends in neurosciences* **2007**, *30* (5), 244-250.

(57) Hirsch, E.; Brandel, J. P.; Galle, P.; Javoy-Agid, F.; Agid, Y. Iron and aluminum increase in the substantia nigra of patients with Parkinson's disease: an X-ray microanalysis. *Journal of neurochemistry* **1991**, *56* (2), 446-451.

(58) Dexter, D.; Wells, F.; Lee, A.; Agid, F.; Agid, Y.; Jenner, P.; Marsden, C. Increased nigral iron content and alterations in other metal ions occurring in brain in Parkinson's disease. *Journal of neurochemistry* **1989**, *52* (6), 1830-1836.

(59) Sofic, E.; Riederer, P.; Heinsen, H.; Beckmann, H.; Reynolds, G.; Hebenstreit, G.; Youdim, M. Increased iron (III) and total iron content in post mortem substantia nigra of parkinsonian brain. *Journal of neural transmission* **1988**, *74* (3), 199-205.

(60) Jenner, P.; Dexter, D.; Sian, J.; Schapira, A.; Marsden, C. Oxidative stress as a cause of nigral cell death in Parkinson's disease and incidental Lewy body disease. *Annals of*

Neurology: Official Journal of the American Neurological Association and the Child Neurology Society **1992**, *32* (S1), S82-S87.

(61) Hare, D. J.; Double, K. L. Iron and dopamine: a toxic couple. *Brain* **2016**, *139* (4), 1026-1035.

(62) Smeyne, M.; Smeyne, R. J. Glutathione metabolism and Parkinson's disease. *Free Radical Biology and Medicine* **2013**, *62*, 13-25.

(63) Ballatori, N.; Krance, S. M.; Notenboom, S.; Shi, S.; Tieu, K.; Hammond, C. L. Glutathione dysregulation and the etiology and progression of human diseases. *Biological Chemistry* **2009**, *390* (3).

(64) Cobley, J. N.; Fiorello, M. L.; Bailey, D. M. 13 reasons why the brain is susceptible to oxidative stress. *Redox Biol* **2018**, *15*, 490-503.

(65) Goldstein, D. S.; Kopin, I. J.; Sharabi, Y. Catecholamine autotoxicity. Implications for pharmacology and therapeutics of Parkinson disease and related disorders. *Pharmacol Ther* **2014**, *144* (3), 268-282.

(66) Hastings, T. G.; Lewis, D. A.; Zigmond, M. J. Role of oxidation in the neurotoxic effects of intrastriatal dopamine injections. *Proc Natl Acad Sci U S A* **1996**, *93* (5), 1956-1961.

(67) Bisaglia, M.; Filograna, R.; Beltramini, M.; Bubacco, L. Are dopamine derivatives implicated in the pathogenesis of Parkinson's disease? *Ageing Res Rev* **2014**, *13*, 107-114.

(68) Graham, D. G. Oxidative pathways for catecholamines in the genesis of neuromelanin and cytotoxic quinones. *Mol Pharmacol* **1978**, *14* (4), 633-643.

(69) Terland, O.; Almås, B.; Flatmark, T.; Andersson, K. K.; Sørlie, M. One-electron oxidation of catecholamines generates free radicals with an in vitro toxicity correlating with their lifetime. *Free Radic Biol Med* **2006**, *41* (8), 1266-1271.

(70) Kalyanaraman, B.; Premovic, P. I.; Sealy, R. C. Semiquinone anion radicals from addition of amino acids, peptides, and proteins to quinones derived from oxidation of catechols and catecholamines. An ESR spin stabilization study. *J Biol Chem* **1987**, *262* (23), 11080-11087.

(71) Tse, D. C.; McCreery, R. L.; Adams, R. N. Potential oxidative pathways of brain catecholamines. *J Med Chem* **1976**, *19* (1), 37-40.

(72) Jameson, G. N.; Zhang, J.; Jameson, R. F.; Linert, W. Kinetic evidence that cysteine reacts with dopaminoquinone via reversible adduct formation to yield 5-cysteinyl-dopamine: an important precursor of neuromelanin. *Org Biomol Chem* **2004**, *2* (5), 777-782.

(73) Li, Y.; Jongberg, S.; Andersen, M. L.; Davies, M. J.; Lund, M. N. Quinone-induced protein modifications: Kinetic preference for reaction of 1,2-benzoquinones with thiol groups in proteins. *Free Radic Biol Med* **2016**, *97*, 148-157.

(74) Xu, R.; Huang, X.; Kramer, K. J.; Hawley, M. D. Characterization of Products from the Reactions of Dopamine Quinone withN-Acetylcysteine. *Bioorganic Chemistry* **1996**, *24* (1), 110-126.

(75) Alfieri, M. L.; Cariola, A.; Panzella, L.; Napolitano, A.; d'Ischia, M.; Valgimigli, L.; Crescenzi, O. Disentangling the Puzzling Regiochemistry of Thiol Addition to o-Quinones. *The Journal of organic chemistry* **2022**, *87* (7), 4580-4589.

(76) Labenski, M. T.; Fisher, A. A.; Lo, H.-H.; Monks, T. J.; Lau, S. S. Protein Electrophile-Binding Motifs: Lysine-Rich Proteins Are Preferential Targets of Quinones. *Drug Metabolism and Disposition* **2009**, *37* (6), 1211-1218.

(77) Xu, R.; Huang, X.; Morgan, T. D.; Prakash, O.; Kramer, K. J.; Hawley, M. D. Characterization of products from the reactions of N-acetyldopamine quinone with N-acetylhistidine. *Arch Biochem Biophys* **1996**, *329* (1), 56-64.

(78) Nicolis, S.; Zucchelli, M.; Monzani, E.; Casella, L. Myoglobin Modification by Enzyme-Generated Dopamine Reactive Species. *Chemistry - A European Journal* **2008**, *14* (28), 8661-8673.

(79) Marchitti, S. A.; Deitrich, R. A.; Vasiliou, V. Neurotoxicity and Metabolism of the Catecholamine-Derived 3,4-Dihydroxyphenylacetaldehyde and 3,4-Dihydroxyphenylglycolaldehyde: The Role of Aldehyde Dehydrogenase. *Pharmacological Reviews* **2007**, *59* (2), 125-150.

(80) Goldstein, D. S.; Sullivan, P.; Holmes, C.; Miller, G. W.; Alter, S.; Strong, R.; Mash, D. C.; Kopin, I. J.; Sharabi, Y. Determinants of buildup of the toxic dopamine metabolite DOPAL in Parkinson's disease. *Journal of Neurochemistry* **2013**, *126* (5), 591-603.

(81) Ramanujam, V.; Charlier, C.; Bax, A. Observation and Kinetic Characterization of Transient Schiff Base Intermediates by CEST NMR Spectroscopy. *Angewandte Chemie International Edition* **2019**, *58* (43), 15309-15312.

(82) Follmer, C.; Coelho-Cerqueira, E.; Yatabe-Franco, D. Y.; Araujo, G. D. T.; Pinheiro, A. S.; Domont, G. B.; Eliezer, D. Oligomerization and Membrane-binding Properties of Covalent Adducts Formed by the Interaction of α-Synuclein with the Toxic Dopamine

Metabolite 3,4-Dihydroxyphenylacetaldehyde (DOPAL). *Journal of Biological Chemistry* **2015**, *290* (46), 27660-27679.

(83) Rees, J. N.; Florang, V. R.; Eckert, L. L.; Doorn, J. A. Protein Reactivity of 3,4-Dihydroxyphenylacetaldehyde, a Toxic Dopamine Metabolite, Is Dependent on Both the Aldehyde and the Catechol. *Chemical Research in Toxicology* **2009**, *22* (7), 1256-1263.

(84) Werner-Allen, J. W.; Dumond, J. F.; Levine, R. L.; Bax, A. Toxic Dopamine Metabolite DOPAL Forms an Unexpected Dicatechol Pyrrole Adduct with Lysines of α-Synuclein. *Angewandte Chemie International Edition* **2016**, *55* (26), 7374-7378.

(85) Kuhn, D. M.; Arthur, R., Jr. Dopamine inactivates tryptophan hydroxylase and forms a redox-cycling quinoprotein: possible endogenous toxin to serotonin neurons. *J Neurosci* **1998**, *18* (18), 7111-7117.

(86) Xu, Y.; Stokes, A. H.; Roskoski, R., Jr.; Vrana, K. E. Dopamine, in the presence of tyrosinase, covalently modifies and inactivates tyrosine hydroxylase. *J Neurosci Res* **1998**, *54* (5), 691-697.

(87) Kuhn, D. M.; Arthur, R. E.; Thomas, D. M.; Elferink, L. A. Tyrosine Hydroxylase Is Inactivated by Catechol-Quinones and Converted to a Redox-Cycling Quinoprotein. *Journal of Neurochemistry* **2001**, *73* (3), 1309-1317.

(88) Whitehead, R. E.; Ferrer, J. V.; Javitch, J. A.; Justice, J. B. Reaction of oxidized dopamine with endogenous cysteine residues in the human dopamine transporter. *Journal of Neurochemistry* **2001**, *76* (4), 1242-1251.

(89) Armarego, W. L.; Waring, P. Inhibition of human brain dihydropteridine reductase [EC 1.6. 99.10] by the oxidation products of catecholamines, the aminochromes. *Biochemical and Biophysical Research Communications* **1983**, *113* (3), 895-899.

(90) Lockyer, J.; Cook, R. G.; Milstien, S.; Kaufman, S.; Woo, S. L.; Ledley, F. D. Structure and expression of human dihydropteridine reductase. *Proceedings of the National Academy of Sciences* **1987**, *84* (10), 3329-3333.

(91) Bruning, J. M.; Wang, Y.; Oltrabella, F.; Tian, B.; Kholodar, S. A.; Liu, H.; Bhattacharya, P.; Guo, S.; Holton, J. M.; Fletterick, R. J.; et al. Covalent Modification and Regulation of the Nuclear Receptor Nurr1 by a Dopamine Metabolite. *Cell Chemical Biology* **2019**, *26* (5), 674-685.e676.

(92) Zetterstrom, R. H.; Solomin, L.; Jansson, L.; Hoffer, B. J.; Olson, L.; Perlmann, T. Dopamine neuron agenesis in Nurr1-deficient mice. *Science* **1997**, *276* (5310), 248-250.

(93) Kadkhodaei, B.; Ito, T.; Joodmardi, E.; Mattsson, B.; Rouillard, C.; Carta, M.; Muramatsu, S. I.; Sumi-Ichinose, C.; Nomura, T.; Metzger, D.; et al. Nurr1 Is Required for Maintenance of Maturing and Adult Midbrain Dopamine Neurons. *Journal of Neuroscience* **2009**, *29* (50), 15923-15932.

(94) Decressac, M.; Volakakis, N.; Björklund, A.; Perlmann, T. NURR1 in Parkinson disease—from pathogenesis to therapeutic potential. *Nature Reviews Neurology* **2013**, *9* (11), 629-636.

(95) Iwawaki, T.; Kohno, K.; Kobayashi, K. Identification of a potential nurr1 response element that activates the tyrosine hydroxylase gene promoter in cultured cells. *Biochemical and biophysical research communications* **2000**, *274* (3), 590-595.

(96) Hermanson, E.; Joseph, B.; Castro, D.; Lindqvist, E.; Aarnisalo, P.; Wallén, Å.; Benoit, G.; Hengerer, B.; Olson, L.; Perlmann, T. Nurr1 regulates dopamine synthesis and storage in MN9D dopamine cells. *Experimental cell research* **2003**, *288* (2), 324-334.

(97) Sacchetti, P.; Mitchell, T. R.; Granneman, J. G.; Bannon, M. J. Nurr1 enhances transcription of the human dopamine transporter gene through a novel mechanism. *Journal of Neurochemistry* **2001**, *76* (5), 1565-1572.

(98) Berman, S. B.; Hastings, T. G. Dopamine Oxidation Alters Mitochondrial Respiration and Induces Permeability Transition in Brain Mitochondria. *Journal of Neurochemistry* **2001**, *73* (3), 1127-1137.

(99) Khan, F. H.; Sen, T.; Maiti, A. K.; Jana, S.; Chatterjee, U.; Chakrabarti, S. Inhibition of rat brain mitochondrial electron transport chain activity by dopamine oxidation products during extended in vitro incubation: implications for Parkinson's disease. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* **2005**, *1741* (1-2), 65-74.

(100) Belluzzi, E.; Bisaglia, M.; Lazzarini, E.; Tabares, L. C.; Beltramini, M.; Bubacco, L. Human SOD2 Modification by Dopamine Quinones Affects Enzymatic Activity by Promoting Its Aggregation: Possible Implications for Parkinson's Disease. *PLoS ONE* **2012**, *7* (6), e38026.

(101) Johnson, F.; Giulivi, C. Superoxide dismutases and their impact upon human health. *Molecular aspects of medicine* **2005**, *26* (4-5), 340-352.

(102) Melov, S.; Coskun, P.; Patel, M.; Tuinstra, R.; Cottrell, B.; Jun, A. S.; Zastawny, T. H.; Dizdaroglu, M.; Goodman, S. I.; Huang, T. T.; et al. Mitochondrial disease in superoxide dismutase 2 mutant mice. *Proceedings of the National Academy of Sciences* **1999**, *96* (3), 846-851.

(103) Hauser, D. N.; Dukes, A. A.; Mortimer, A. D.; Hastings, T. G. Dopamine quinone modifies and decreases the abundance of the mitochondrial selenoprotein glutathione peroxidase 4. *Free Radical Biology and Medicine* **2013**, *65*, 419-427.

(104) Cardoso, B. R.; Hare, D. J.; Bush, A. I.; Roberts, B. R. Glutathione peroxidase 4: a new player in neurodegeneration? *Molecular Psychiatry* **2017**, *22* (3), 328-335.

(105) Biosa, A.; Arduini, I.; Soriano, M. E.; Giorgio, V.; Bernardi, P.; Bisaglia, M.; Bubacco, L. Dopamine Oxidation Products as Mitochondrial Endotoxins, a Potential Molecular Mechanism for Preferential Neurodegeneration in Parkinson's Disease. *ACS Chemical Neuroscience* **2018**, *9* (11), 2849-2858.

(106) Crawford, R. A.; Bowman, K. R.; Cagle, B. S.; Doorn, J. A. In vitro inhibition of glutathione-S-transferase by dopamine and its metabolites, 3, 4-dihydroxyphenylacetaldehyde and 3, 4-dihydroxyphenylacetic acid. *NeuroToxicology* **2021**, *86*, 85-93.

(107) Strange, R. C.; Spiteri, M. A.; Ramachandran, S.; Fryer, A. A. Glutathione-S-transferase family of enzymes. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2001**, *482* (1-2), 21-26.

(108) Whitworth, A. J.; Theodore, D. A.; Greene, J. C.; Benes, H.; Wes, P. D.; Pallanck, L. J. Increased glutathione S-transferase activity rescues dopaminergic neuron loss in a

Drosophila model of Parkinson's disease. *Proceedings of the National Academy of Sciences* **2005**, *102* (22), 8024-8029.

(109) Hirrlinger, J.; Schulz, J. B.; Dringen, R. Effects of dopamine on the glutathione metabolism of cultured astroglial cells: implications for Parkinson's disease. *Journal of Neurochemistry* **2002**, *82* (3), 458-467.

(110) Walter, P.; Ron, D. The unfolded protein response: from stress pathway to homeostatic regulation. *science* **2011**, *334* (6059), 1081-1086.

(111) Perri, E. R.; Thomas, C. J.; Parakh, S.; Spencer, D. M.; Atkin, J. D. The unfolded protein response and the role of protein disulfide isomerase in neurodegeneration. *Frontiers in cell and developmental biology* **2016**, *3*, 80.

(112) Ryu, E. J.; Harding, H. P.; Angelastro, J. M.; Vitolo, O. V.; Ron, D.; Greene, L. A. Endoplasmic reticulum stress and the unfolded protein response in cellular models of Parkinson's disease. *Journal of Neuroscience* **2002**, *22* (24), 10690-10698.

(113) Yoshida, H. ER stress and diseases. *FEBS Journal* **2007**, *274* (3), 630-658.

(114) Yamamuro, A.; Yoshioka, Y.; Ogita, K.; Maeda, S. Involvement of Endoplasmic Reticulum Stress on the Cell Death Induced by 6-Hydroxydopamine in Human Neuroblastoma SH-SY5Y Cells. *Neurochemical Research* **2006**, *31* (5), 657-664.

(115) Wilkinson, B.; Gilbert, H. F. Protein disulfide isomerase. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* **2004**, *1699* (1-2), 35-44.

(116) Muller, C.; Bandemer, J.; Vindis, C.; Camaré, C.; Mucher, E.; Guéraud, F.; Larroque-Cardoso, P.; Bernis, C.; Auge, N.; Salvayre, R. Protein disulfide isomerase

modification and inhibition contribute to ER stress and apoptosis induced by oxidized low density lipoproteins. *Antioxidants & redox signaling* **2013**, *18* (7), 731-742.

(117) Hurben, A. K.; Erber, L. N.; Tretyakova, N. Y.; Doran, T. M. Proteome-Wide Profiling of Cellular Targets Modified by Dopamine Metabolites Using a Bio-Orthogonally Functionalized Catecholamine. *ACS Chemical Biology* **2021**, *16* (11), 2581-2594.

(118) Farzam, A.; Chohan, K.; Strmiskova, M.; Hewitt, S. J.; Park, D. S.; Pezacki, J. P.; Özcelik, D. A functionalized hydroxydopamine quinone links thiol modification to neuronal cell death. *Redox biology* **2020**, *28*, 101377.

(119) Stefanis, L. -Synuclein in Parkinson's Disease. *Cold Spring Harbor Perspectives in Medicine* **2012**, *2* (2), a009399-a009399.

(120) Krüger, R.; Kuhn, W.; Müller, T.; Woitalla, D.; Graeber, M.; Kösel, S.; Przuntek, H.; Epplen, J. T.; Schols, L.; Riess, O. AlaSOPro mutation in the gene encoding α-synuclein in Parkinson's disease. *Nature genetics* **1998**, *18* (2), 106-108.

(121) Spillantini, M. G.; Schmidt, M. L.; Lee, V. M. Y.; Trojanowski, J. Q.; Jakes, R.; Goedert, M. α-Synuclein in Lewy bodies. *Nature* **1997**, *388* (6645), 839-840.

(122) Spillantini, M. G.; Crowther, R. A.; Jakes, R.; Hasegawa, M.; Goedert, M. α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson’s disease and dementia with Lewy bodies. *Proceedings of the National Academy of Sciences* **1998**, *95* (11), 6469-6473.

(123) Bisaglia, M.; Tosatto, L.; Munari, F.; Tessari, I.; de Laureto, P. P.; Mammi, S.; Bubacco, L. Dopamine quinones interact with α-synuclein to form unstructured adducts. *Biochemical and biophysical research communications* **2010**, *394* (2), 424-428.

(124) Lee, H.-J.; Baek, S. M.; Ho, D.-H.; Suk, J.-E.; Cho, E.-D.; Lee, S.-J. Dopamine promotes formation and secretion of non-fibrillar alpha-synuclein oligomers. *Experimental and Molecular Medicine* **2011**, *43* (4), 216.

(125) Norris, E. H.; Giasson, B. I.; Hodara, R.; Xu, S.; Trojanowski, J. Q.; Ischiropoulos, H.; Lee, V. M. Y. Reversible Inhibition of α-Synuclein Fibrillization by Dopaminochrome-mediated Conformational Alterations. *Journal of Biological Chemistry* **2005**, *280* (22), 21212-21219.

(126) Burke, W. J.; Kumar, V. B.; Pandey, N.; Panneton, W. M.; Gan, Q.; Franko, M. W.; O'Dell, M.; Li, S. W.; Pan, Y.; Chung, H. D.; et al. Aggregation of α-synuclein by DOPAL, the monoamine oxidase metabolite of dopamine. *Acta Neuropathologica* **2008**, *115* (2), 193-203.

(127) Plotegher, N.; Berti, G.; Ferrari, E.; Tessari, I.; Zanetti, M.; Lunelli, L.; Greggio, E.; Bisaglia, M.; Veronesi, M.; Girotto, S.; et al. DOPAL derived alpha-synuclein oligomers impair synaptic vesicles physiological function. *Scientific Reports* **2017**, *7* (1), 40699.

(128) Martinez-Vicente, M.; Talloczy, Z.; Kaushik, S.; Massey, A. C.; Mazzulli, J.; Mosharov, E. V.; Hodara, R.; Fredenburg, R.; Wu, D.-C.; Follenzi, A.; et al. Dopamine-modified α-synuclein blocks chaperone-mediated autophagy. *Journal of Clinical Investigation* **2008**, 10.1172/jci32806.

(129) Mor, D. E.; Tsika, E.; Mazzulli, J. R.; Gould, N. S.; Kim, H.; Daniels, M. J.; Doshi, S.; Gupta, P.; Grossman, J. L.; Tan, V. X.; et al. Dopamine induces soluble α-synuclein oligomers and nigrostriatal degeneration. *Nature Neuroscience* **2017**, *20* (11), 1560-1568.

(130) Conway, K. A.; Rochet, J. C.; Bieganski, R. M.; Lansbury, P. T., Jr. Kinetic stabilization of the alpha-synuclein protofibril by a dopamine-alpha-synuclein adduct. *Science* **2001**, *294* (5545), 1346-1349.

(131) Zhao, H.; Huang, S.; Palanisamy, S.; Wang, C.; Rainer, G.; Zhang, X. Alpha-Synuclein Dopaminylation Presented in Plasma of Both Healthy Subjects and Parkinson's Disease Patients. *PROTEOMICS – Clinical Applications* **2020**, *14* (5), 1900117.

(132) Lücking, C. B.; Dürr, A.; Bonifati, V.; Vaughan, J.; De Michele, G.; Gasser, T.; Harhangi, B. S.; Meco, G.; Denèfle, P.; Wood, N. W.; et al. Association between early-onset Parkinson's disease and mutations in the *Parkin* gene. *N. Engl. J. Med.* **2000**, *342* (21), 1560-1567.

(133) Alicia; Richard. The Roles of PINK1, Parkin, and Mitochondrial Fidelity in Parkinson's Disease. *Neuron* **2015**, *85* (2), 257-273.

(134) Lavoie, M. J.; Ostaszewski, B. L.; Weihofen, A.; Schlossmacher, M. G.; Selkoe, D. J. Dopamine covalently modifies and functionally inactivates parkin. *Nature Medicine* **2005**, *11* (11), 1214-1221.

(135) Tokarew, J. M.; El-Kodsi, D. N.; Lengacher, N. A.; Fehr, T. K.; Nguyen, A. P.; Shutinoski, B.; O'Nuallain, B.; Jin, M.; Khan, J. M.; Ng, A. C. H.; et al. Age-associated

insolubility of parkin in human midbrain is linked to redox balance and sequestration of reactive dopamine metabolites. *Acta Neuropathologica* **2021**, *141* (5), 725-754.

(136) Bonifati, V.; Rizzu, P.; Van Baren, M. J.; Schaap, O.; Breedveld, G. J.; Krieger, E.; Dekker, M. C.; Squitieri, F.; Ibanez, P.; Joosse, M. Mutations in the DJ-1 gene associated with autosomal recessive early-onset parkinsonism. *science* **2003**, *299* (5604), 256-259.

(137) Wilson, M. A. The Role of Cysteine Oxidation in DJ-1 Function and Dysfunction. *Antioxidants & Redox Signaling* **2011**, *15* (1), 111-122.

(138) Canet-Avilés, R. M.; Wilson, M. A.; Miller, D. W.; Ahmad, R.; McLendon, C.; Bandyopadhyay, S.; Baptista, M. J.; Ringe, D.; Petsko, G. A.; Cookson, M. R. The Parkinson's disease protein DJ-1 is neuroprotective due to cysteine-sulfinic acid-driven mitochondrial localization. *Proceedings of the National Academy of Sciences* **2004**, *101* (24), 9103-9108.

(139) Girotto, S.; Sturlese, M.; Bellanda, M.; Tessari, I.; Cappellini, R.; Bisaglia, M.; Bubacco, L.; Mammi, S. Dopamine-derived Quinones Affect the Structure of the Redox Sensor DJ-1 through Modifications at Cys-106 and Cys-53. *Journal of Biological Chemistry* **2012**, *287* (22), 18738-18749.

(140) Sidransky, E.; Nalls, M. A.; Aasly, J. O.; Aharon-Peretz, J.; Annesi, G.; Barbosa, E. R.; Bar-Shira, A.; Berg, D.; Bras, J.; Brice, A.; et al. Multicenter Analysis of Glucocerebrosidase Mutations in Parkinson's Disease. *New England Journal of Medicine* **2009**, *361* (17), 1651-1661.

(141) Do, J.; McKinney, C.; Sharma, P.; Sidransky, E. Glucocerebrosidase and its relevance to Parkinson disease. *Molecular Neurodegeneration* **2019**, *14* (1).

(142) Lepack, A. E.; Werner, C. T.; Stewart, A. F.; Fulton, S. L.; Zhong, P.; Farrelly, L. A.; Smith, A. C. W.; Ramakrishnan, A.; Lyu, Y.; Bastle, R. M.; et al. Dopaminylation of histone H3 in ventral tegmental area regulates cocaine seeking. *Science* **2020**, *368* (6487), 197-201.

(143) Fulton, S. L.; Mitra, S.; Lepack, A. E.; Martin, J. A.; Stewart, A. F.; Converse, J.; Hochstetler, M.; Dietz, D. M.; Maze, I. Histone H3 dopaminylation in ventral tegmental area underlies heroin-induced transcriptional and behavioral plasticity in male rats. *Neuropsychopharmacology* **2022**, 10.1038/s41386-022-01279-4.

(144) Girault, J.-A. Epigenetic tinkering with neurotransmitters. *Science* **2020**, *368* (6487), 134-135.

(145) Rowe, D. B.; Le, W.; Smith, R. G.; Appel, S. H. Antibodies from patients with Parkinson's disease react with protein modified by dopamine oxidation. *J Neurosci Res* **1998**, *53* (5), 551-558.

(146) Zhang, W.; Phillips, K.; Wielgus, A. R.; Liu, J.; Albertini, A.; Zucca, F. A.; Faust, R.; Qian, S. Y.; Miller, D. S.; Chignell, C. F.; et al. Neuromelanin Activates Microglia and Induces Degeneration of Dopaminergic Neurons: Implications for Progression of Parkinson's Disease. *Neurotoxicity Research* **2011**, *19* (1), 63-72.

(147) Tan, E. K.; Chao, Y. X.; West, A.; Chan, L. L.; Poewe, W.; Jankovic, J. Parkinson disease and the immune system - associations, mechanisms and therapeutics. *Nat Rev Neurol* **2020**, *16* (6), 303-318.

(148) Bonam, S. R.; Muller, S. Parkinson's disease is an autoimmune disease: A reappraisal. *Autoimmun Rev* **2020**, *19* (12), 102684.

(149) Lindestam Arlehamn, C. S.; Dhanwani, R.; Pham, J.; Kuan, R.; Frazier, A.; Rezende Dutra, J.; Phillips, E.; Mallal, S.; Roederer, M.; Marder, K. S.; et al. α-Synuclein-specific T cell reactivity is associated with preclinical and early Parkinson's disease. *Nature Communications* **2020**, *11* (1).

(150) Bindoli, A.; Rigobello, M. P.; Galzigna, L. Toxicity of aminochromes. *Toxicology letters* **1989**, *48* (1), 3-20.

(151) Arriagada, C.; Dagnino-Subiabre, A.; Caviedes, P.; Armero, J. M.; Caviedes, R.; Segura-Aguilar, J. Studies of aminochrome toxicity in a mouse derived neuronal cell line: is this toxicity mediated via glutamate transmission? *Amino Acids* **2000**, *18* (4), 363-373.

(152) Van Laar, V. S.; Dukes, A. A.; Cascio, M.; Hastings, T. G. Proteomic analysis of rat brain mitochondria following exposure to dopamine quinone: Implications for Parkinson disease. *Neurobiology of Disease* **2008**, *29* (3), 477-489.

(153) Van Laar, V. S.; Mishizen, A. J.; Cascio, M.; Hastings, T. G. Proteomic identification of dopamine-conjugated proteins from isolated rat brain mitochondria and SH-SY5Y cells. *Neurobiology of Disease* **2009**, *34* (3), 487-500.

(154) Brues, A. M. Biological hazards and toxicity of radioactive isotopes. *The Journal of Clinical Investigation* **1949**, *28* (6), 1286-1296.

(155) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal Chemistry: Fishing for Selectivity in a Sea of Functionality. *Angewandte Chemie International Edition* **2009**, *48* (38), 6974-6998.

(156) Post, M. R.; Sulzer, D. The chemical tools for imaging dopamine release. *Cell Chemical Biology* **2021**, *28* (6), 748-764.

(157) Chiueh, C. C.; Zukowska-Grojec, Z.; Kirk, K. L.; Kopin, I. J. 6-Fluorocatecholamines as false adrenergic neurotransmitters. *Journal of Pharmacology and Experimental Therapeutics* **1983**, *225* (3), 529-533.

(158) Yanagisawa, D.; Oda, K.; Inden, M.; Morikawa, S.; Inubushi, T.; Taniguchi, T.; Hijioka, M.; Kitamura, Y.; Tooyama, I. Fluorodopa is a Promising Fluorine-19 MRI Probe for Evaluating Striatal Dopaminergic Function in a Rat Model of Parkinson's Disease. *Journal of Neuroscience Research* **2017**, *95* (7), 1485-1494.

(159) Rodriguez, P. C.; Pereira, D. B.; Borgkvist, A.; Wong, M. Y.; Barnard, C.; Sonders, M. S.; Zhang, H.; Sames, D.; Sulzer, D. Fluorescent dopamine tracer resolves individual dopaminergic synapses and their activity in the brain. *Proceedings of the National Academy of Sciences* **2013**, *110* (3), 870-875.

(160) Pereira, D. B.; Schmitz, Y.; Mészáros, J.; Merchant, P.; Hu, G.; Li, S.; Henke, A.; Lizardi-Ortiz, J. E.; Karpowicz, R. J.; Morgenstern, T. J.; et al. Fluorescent false neurotransmitter reveals functionally silent dopamine vesicle clusters in the striatum. *Nature Neuroscience* **2016**, *19* (4), 578-586.

(161) Hu, G.; Henke, A.; Karpowicz, R. J.; Sonders, M. S.; Farrimond, F.; Edwards, R.; Sulzer, D.; Sames, D. New Fluorescent Substrate Enables Quantitative and High-Throughput Examination of Vesicular Monoamine Transporter 2 (VMAT2). *ACS Chemical Biology* **2013**, *8* (9), 1947-1954.

(162) Paz, M.; Flückiger, R.; Boak, A.; Kagan, H.; Gallop, P. M. Specific detection of quinoproteins by redox-cycling staining. *Journal of Biological Chemistry* **1991**, *266* (2), 689-692.

(163) Yu, G.; Liu, H.; Zhou, W.; Zhu, X.; Yu, C.; Wang, N.; Zhang, Y.; Ma, J.; Zhao, Y.; Xu, Y. In vivo protein targets for increased quinoprotein adduct formation in aged substantia nigra. *Experimental neurology* **2015**, *271*, 13-24.

(164) Elgawish, M. S.; Kishikawa, N.; Ohyama, K.; Kuroda, N. Characterization of quinone derived protein adducts and their selective identification using redox cycling based chemiluminescence assay. *Journal of Chromatography A* **2015**, *1403*, 96-103.

(165) Mazzulli, J. R.; Burbulla, L. F.; Krainc, D.; Ischiropoulos, H. Detection of Free and Protein-Boundortho-Quinones by Near-Infrared Fluorescence. *Analytical Chemistry* **2016**, *88* (4), 2399-2405.

(166) Adamek, V.; Liu, X.-C.; Zhang, Y. A.; Adamkova, K.; Scouten, W. H. New aliphatic boronate ligands for affinity chromatography. *Journal of Chromatography A* **1992**, *625* (2), 91-99.

(167) António, J. P. M.; Russo, R.; Carvalho, C. P.; Cal, P. M. S. D.; Gois, P. M. P. Boronic acids as building blocks for the construction of therapeutically useful bioconjugates. *Chemical Society Reviews* **2019**, *48* (13), 3513-3536.

(168) Ma, W.; Liu, H.-T.; Long, Y.-T. Monitoring Dopamine Quinone-Induced Dopaminergic Neurotoxicity Using Dopamine Functionalized Quantum Dots. *ACS Applied Materials & Interfaces* **2015**, *7* (26), 14352-14358.

(169) Liu, X.; Liu, J. Biosensors and sensors for dopamine detection. *View* **2021**, *2* (1), 20200102.

(170) Pradhan, T.; Jung, H. S.; Jang, J. H.; Kim, T. W.; Kang, C.; Kim, J. S. Chemical sensing of neurotransmitters. *Chem. Soc. Rev.* **2014**, *43* (13), 4684-4713.

(171) Patriarchi, T.; Cho, J. R.; Merten, K.; Howe, M. W.; Marley, A.; Xiong, W.-H.; Folk, R. W.; Broussard, G. J.; Liang, R.; Jang, M. J.; et al. Ultrafast neuronal imaging of dopamine dynamics with designed genetically encoded sensors. *Science* **2018**, *360* (6396), eaat4422.

(172) Patriarchi, T.; Mohebi, A.; Sun, J.; Marley, A.; Liang, R.; Dong, C.; Puhger, K.; Mizuno, G. O.; Davis, C. M.; Wiltgen, B.; et al. An expanded palette of dopamine sensors for multiplex imaging in vivo. *Nature Methods* **2020**, *17* (11), 1147-1155.

(173) Secor, K. E.; Glass, T. E. Selective Amine Recognition: Development of a Chemosensor for Dopamine and Norepinephrine. *Organic Letters* **2004**, *6* (21), 3727-3730.

(174) Ganguly, M.; Mondal, C.; Jana, J.; Pal, A.; Pal, T. Selective Dopamine Chemosensing Using Silver-Enhanced Fluorescence. *Langmuir* **2014**, *30* (14), 4120-4128.

(175) Gabrielli, L.; Carril, M.; Padro, D.; Mancin, F. Multimodal 19 F NMR Dopamine Detection and Imaging with a Nanoparticle-Based Displacement Assay. *Chemistry - A European Journal* **2018**, *24* (49), 13036-13042.

(176) Zhao, L.; Xin, X.; Ding, P.; Song, A.; Xie, Z.; Shen, J.; Xu, G. Fluorescent oligomer as a chemosensor for the label-free detection of Fe3+ and dopamine with selectivity and sensitivity. *Analytica Chimica Acta* **2016**, *926*, 99-106.

(177) Zhang, X.; Zhu, Y.; Li, X.; Guo, X.; Zhang, B.; Jia, X.; Dai, B. A simple, fast and low-cost turn-on fluorescence method for dopamine detection using in situ reaction. *Analytica chimica acta* **2016**, *944*, 51-56.

(178) Wightman, R. M.; May, L. J.; Michael, A. C. Detection of Dopamine Dynamics in the Brain. *Analytical Chemistry* **1988**, *60* (13), 769A-793A.

(179) Venton, B. J.; Cao, Q. Fundamentals of fast-scan cyclic voltammetry for dopamine detection. *The Analyst* **2020**, *145* (4), 1158-1168.

(180) Robinson, D. L.; Venton, B. J.; Heien, M. L. A. V.; Wightman, R. M. Detecting Subsecond Dopamine Release with Fast-Scan Cyclic Voltammetry in Vivo. *Clinical Chemistry* **2003**, *49* (10), 1763-1773.

(181) Spanos, M.; Gras-Najjar, J.; Letchworth, J. M.; Sanford, A. L.; Toups, J. V.; Sombers, L. A. Quantitation of Hydrogen Peroxide Fluctuations and Their Modulation of Dopamine Dynamics in the Rat Dorsal Striatum Using Fast-Scan Cyclic Voltammetry. *ACS Chemical Neuroscience* **2013**, *4* (5), 782-789.

(182) Bergstrom, B. P.; Garris, P. A. 'Passive stabilization' of striatal extracellular dopamine across the lesion spectrum encompassing the presymptomatic phase of Parkinson's disease: a voltammetric study in the 6-OHDA-lesioned rat. *Journal of Neurochemistry* **2003**, *87* (5), 1224-1236.

(183) Zhang, L.; Le, W.; Xie, W.; Dani, J. A. Age-related changes in dopamine signaling in Nurr1 deficient mice as a model of Parkinson's disease. *Neurobiology of Aging* **2012**, *33* (5), 1001.e1007-1001.e1001.

(184) Dunn, A. R.; Stout, K. A.; Ozawa, M.; Lohr, K. M.; Hoffman, C. A.; Bernstein, A. I.; Li, Y.; Wang, M.; Sgobio, C.; Sastry, N.; et al. Synaptic vesicle glycoprotein 2C (SV2C) modulates dopamine release and is disrupted in Parkinson disease. *Proceedings of the National Academy of Sciences* **2017**, *114* (11), E2253-E2262.

(185) Cragg, S.; Clarke, D.; Greenfield, S. Real-time dynamics of dopamine released from neuronal transplants in experimental Parkinson's disease. *Experimental neurology* **2000**, *164* (1), 145-153.

(186) Garris, P. A.; Walker, Q. D.; Wightman, R. M. Dopamine release and uptake rates both decrease in the partially denervated striatum in proportion to the loss of dopamine terminals. *Brain research* **1997**, *753* (2), 225-234.

(187) Fields, S. Proteomics in genomeland. *Science* **2001**, *291* (5507), 1221-1224.

(188) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. Protein sequencing by tandem mass spectrometry. *Proceedings of the National Academy of Sciences* **1986**, *83* (17), 6233-6237.

(189) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. *Analytical Chemistry* **2013**, *85* (11), 5288-5296.

(190) Cox, J.; Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **2008**, *26* (12), 1367-1372.

(191) Chen, S.-H.; Li, C.-W. Detection and Characterization of Catechol Quinone-Derived Protein Adducts Using Biomolecular Mass Spectrometry. *Frontiers in Chemistry* **2019**, *7*, Mini Review.

(192) Ciccimaro, E.; Blair, I. A. Stable-isotope dilution LC–MS for quantitative biomarker analysis. *Bioanalysis* **2010**, *2* (2), 311-341.

(193) Zhang, F.; Dryhurst, G. Effects of L-Cysteine on the Oxidation Chemistry of Dopamine: New Reaction Pathways of Potential Relevance to Idiopathic Parkinson's Disease. *Journal of Medicinal Chemistry* **1994**, *37* (8), 1084-1098.

(194) Fornstedt, B.; Pileblad, E.; Carlsson, A. In Vivo Autoxidation of Dopamine in Guinea Pig Striatum Increases with Age. *Journal of Neurochemistry* **1990**, *55* (2), 655-659.

(195) Spencer, J. P. E.; Jenner, P.; Daniel, S. E.; Lees, A. J.; Marsden, D. C.; Halliwell, B. Conjugates of Catecholamines with Cysteine and GSH in Parkinson's Disease: Possible Mechanisms of Formation Involving Reactive Oxygen Species. *Journal of Neurochemistry* **2002**, *71* (5), 2112-2122.

(196) Badillo-Ramírez, I.; Saniger, J. M.; Rivas-Arancibia, S. 5-S-cysteinyl-dopamine, a neurotoxic endogenous metabolite of dopamine: Implications for Parkinson's disease. *Neurochemistry international* **2019**, *129*, 104514.

(197) De Lau, L. M.; Breteler, M. M. Epidemiology of Parkinson's disease. *The Lancet Neurology* **2006**, *5* (6), 525-535.

(198) Lee, S.-J. Origins and Effects of Extracellular α-synuclein: Implications in Parkinson's Disease. *Journal of Molecular Neuroscience* **2008**, *34* (1), 17-22.

(199) Braak, H.; Ghebremedhin, E.; Rüb, U.; Bratzke, H.; Del Tredici, K. Stages in the development of Parkinson's disease-related pathology. *Cell and Tissue Research* **2004**, *318* (1), 121-134.

(200) Braak, H.; Del Tredici, K.; Rüb, U.; De Vos, R. A.; Steur, E. N. J.; Braak, E. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of aging* **2003**, *24* (2), 197-211.

(201) Kim, S.; Kwon, S.-H.; Kam, T.-I.; Panicker, N.; Karuppagounder, S. S.; Lee, S.; Lee, J. H.; Kim, W. R.; Kook, M.; Foss, C. A. Transneuronal propagation of pathologic α-synuclein from the gut to the brain models Parkinson's disease. *Neuron* **2019**, *103* (4), 627-641. e627.

(202) Klingelhoefer, L.; Reichmann, H. Pathogenesis of Parkinson disease—the gut–brain axis and environmental factors. *Nature Reviews Neurology* **2015**, *11* (11), 625-636.

(203) Surmeier, D. J.; Guzman, J. N.; Sanchez-Padilla, J.; Goldberg, J. A. The Origins of Oxidant Stress in Parkinson's Disease and Therapeutic Strategies. *Antioxidants & Redox Signaling* **2011**, *14* (7), 1289-1301.

(204) Dehay, B.; Martinez-Vicente, M.; Ramirez, A.; Perier, C.; Klein, C.; Vila, M.; Bezard, E. Lysosomal dysfunction in Parkinson disease. *Autophagy* **2012**, *8* (9), 1389-1391.

(205) Abou-Sleiman, P. M.; Muqit, M. M.; Wood, N. W. Expanding insights of mitochondrial dysfunction in Parkinson's disease. *Nature Reviews Neuroscience* **2006**, *7* (3), 207-219.

(206) Exner, N.; Lutz, A. K.; Haass, C.; Winklhofer, K. F. Mitochondrial dysfunction in Parkinson's disease: molecular mechanisms and pathophysiological consequences. *The EMBO journal* **2012**, *31* (14), 3038-3062.

(207) Satake, W.; Nakabayashi, Y.; Mizuta, I.; Hirota, Y.; Ito, C.; Kubo, M.; Kawaguchi, T.; Tsunoda, T.; Watanabe, M.; Takeda, A. Genome-wide association study identifies common variants at four loci as genetic risk factors for Parkinson's disease. *Nature genetics* **2009**, *41* (12), 1303-1307.

(208) Klein, C.; Westenberger, A. Genetics of Parkinson's disease. *Cold Spring Harbor perspectives in medicine* **2012**, *2* (1), a008888.

(209) Deng, H.; Wang, P.; Jankovic, J. The genetics of Parkinson disease. *Ageing research reviews* **2018**, *42*, 72-85.

(210) Surmeier, D. J.; Obeso, J. A.; Halliday, G. M. Selective neuronal vulnerability in Parkinson disease. *Nature Reviews Neuroscience* **2017**, *18* (2), 101-113.

(211) Sulzer, D.; Surmeier, D. J. Neuronal vulnerability, pathogenesis, and Parkinson's disease. *Movement Disorders* **2013**, *28* (6), 715-724.

(212) Liebler, D. C. Protein Damage by Reactive Electrophiles: Targets and Consequences. *Chemical Research in Toxicology* **2008**, *21* (1), 117-128.

(213) 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* (7325), 790-795.

(214) Ford, B.; Bateman, L. A.; Gutierrez-Palominos, L.; Park, R.; Nomura, D. K. Mapping proteome-wide targets of glyphosate in mice. *Cell chemical biology* **2017**, *24* (2), 133-140.

(215) Matthews, M. L.; He, L.; Horning, B. D.; Olson, E. J.; Correia, B. E.; Yates, J. R.; Dawson, P. E.; Cravatt, B. F. Chemoproteomic profiling and discovery of protein electrophiles in human cells. *Nature Chemistry* **2017**, *9* (3), 234-243.

(216) Wang, J.; Zhang, C.-J.; Zhang, J.; He, Y.; Lee, Y. M.; Chen, S.; Lim, T. K.; Ng, S.; Shen, H.-M.; Lin, Q. Mapping sites of aspirin-induced acetylations in live cells by quantitative acid-cleavable activity-based protein profiling (QA-ABPP). *Scientific Reports* **2015**, *5* (1), 7896.

(217) Tian, C.; Sun, R.; Liu, K.; Fu, L.; Liu, X.; Zhou, W.; Yang, Y.; Yang, J. Multiplexed Thiol Reactivity Profiling for Target Discovery of Electrophilic Natural Products. *Cell Chemical Biology* **2017**, *24* (11), 1416-1427.e1415.

(218) Wang, C.; Weerapana, E.; Blewett, M. M.; Cravatt, B. F. A chemoproteomic platform to quantitatively map targets of lipid-derived electrophiles. *Nature Methods* **2014**, *11* (1), 79-85.

(219) Parvez, S.; Long, M. J. C.; Poganik, J. R.; Aye, Y. Redox Signaling by Reactive Electrophiles and Oxidants. *Chemical Reviews* **2018**, *118* (18), 8798-8888.

(220) Liu, X.; Long, M. J. C.; Aye, Y. Proteomics and Beyond: Cell Decision-Making Shaped by Reactive Electrophiles. *Trends in Biochemical Sciences* **2019**, *44* (1), 75-89.

(221) Hurben, A. K.; Ge, P.; Bouchard, J. L.; Doran, T. M.; Tretyakova, N. Y. Photocaged dicarbonyl probe provides spatiotemporal control over protein glycation. *Chemical Communications* **2022**, *58* (6), 855-858.

(222) Lin, H.-Y.; Haegele, J. A.; Disare, M. T.; Lin, Q.; Aye, Y. A Generalizable Platform for Interrogating Target- and Signal-Specific Consequences of Electrophilic

Modifications in Redox-Dependent Cell Signaling. *Journal of the American Chemical Society* **2015**, *137* (19), 6232-6244.

(223) Long, M. J.; Miranda Herrera, P. A.; Aye, Y. Hitting the bullseye: endogenous electrophiles show remarkable nuance in signaling regulation. *Chemical Research in Toxicology* **2022**.

(224) Zhao, Y.; Long, M. J.; Wang, Y.; Zhang, S.; Aye, Y. Ube2V2 is a rosetta stone bridging redox and ubiquitin codes, coordinating DNA damage responses. *ACS central science* **2018**, *4* (2), 246-259.

(225) Zhao, Y.; Herrera, P. A. M.; Chang, D.; Hamelin, R.; Long, M. J.; Aye, Y. Function-guided proximity mapping unveils electrophilic-metabolite sensing by proteins not present in their canonical locales. *Proceedings of the National Academy of Sciences* **2022**, *119* (5).

(226) Asad, N.; McLain, D. E.; Condon, A. F.; Gore, S.; Hampton, S. E.; Vijay, S.; Williams, J. T.; Dore, T. M. Photoactivatable Dopamine and Sulpiride to Explore the Function of Dopaminergic Neurons and Circuits. *ACS Chemical Neuroscience* **2020**, *11* (6), 939-951.

(227) Escorihuela, J.; Looijen, W. J. E.; Wang, X.; Aquino, A. J. A.; Lischka, H.; Zuilhof, H. Cycloaddition of Strained Cyclic Alkenes and Ortho-Quinones: A Distortion/Interaction Analysis. *The Journal of Organic Chemistry* **2020**, *85* (21), 13557-13566.

(228) Dorsey, E.; Elbaz, A.; Nichols, E.; Abd-Allah, F.; Abdelalim, A.; Adsuar, J.; Ansha, M.; Brayne, C.; Choi, J.; Collado-Mateo, D. GBD 2016 Parkinson's Disease

Collaborators. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* **2018**, *17* (11), 939-953.

(229) Zecca, L.; Fariello, R.; Riederer, P.; Sulzer, D.; Gatti, A.; Tampellini, D. The absolute concentration of nigral neuromelanin, assayed by a new sensitive method, increases throughout the life and is dramatically decreased in Parkinson's disease. *FEBS Letters* **2002**, *510* (3), 216-220.

(230) Cheng, H.-C.; Ulane, C. M.; Burke, R. E. Clinical progression in Parkinson disease and the neurobiology of axons. *Annals of Neurology* **2010**, *67* (6), 715-725.

(231) Chang, D.; Nalls, M. A.; Hallgrímsdóttir, I. B.; Hunkapiller, J.; Van Der Brug, M.; Cai, F.; Kerchner, G. A.; Ayalon, G.; Bingol, B.; Sheng, M.; et al. A meta-analysis of genome-wide association studies identifies 17 new Parkinson's disease risk loci. *Nature Genetics* **2017**, *49* (10), 1511-1516.

(232) Herrera, A.; Muñoz, P.; Steinbusch, H. W. M.; Segura-Aguilar, J. Are Dopamine Oxidation Metabolites Involved in the Loss of Dopaminergic Neurons in the Nigrostriatal System in Parkinson's Disease? *ACS Chemical Neuroscience* **2017**, *8* (4), 702-711.

(233) Zecca, L.; Zucca, F.; Costi, P.; Tampellini, D.; Gatti, A.; Gerlach, M.; Riederer, P.; Fariello, R.; Ito, S.; Gallorini, M. The neuromelanin of human substantia nigra: structure, synthesis and molecular behaviour. *Advances in Research on Neurodegeneration* **2003**, 145-155.

(234) Zhou, Z. D.; Lim, T. M. Dopamine (DA) induced irreversible proteasome inhibition via DA derived quinones. *Free Radical Research* **2009**, *43* (4), 417-430.

(235) Tyanova, S.; Temu, T.; Sinitcyn, P.; Carlson, A.; Hein, M. Y.; Geiger, T.; Mann, M.; Cox, J. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nature Methods* **2016**, *13* (9), 731-740.

(236) Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics* **2014**, *13* (9), 2513-2526.

(237) Liu, Z.; Hu, B.-H.; Messersmith, P. B. Acetonide protection of dopamine for the synthesis of highly pure N-docosahexaenoyldopamine. *Tetrahedron Letters* **2010**, *51* (18), 2403-2405.

(238) Fukuyama, T.; Jow, C.-K.; Cheung, M. 2-and 4-Nitrobenzenesulfonamides: Exceptionally versatile means for preparation of secondary amines and protection of amines. *Tetrahedron Letters* **1995**, *36* (36), 6373-6374.

(239) Asanuma, M.; Miyazaki, I.; Ogawa, N. Dopamine-or L-DOPA-induced neurotoxicity: the role of dopamine quinone formation and tyrosinase in a model of Parkinson's disease. *Neurotoxicity research* **2003**, *5* (3), 165-176.

(240) Jimenez, M.; Garcia-Carmona, F.; Garcia-Canovas, F.; Iborra, J.; Lozano, J.; Martinez, F. Chemical intermediates in dopamine oxidation by tyrosinase, and kinetic studies of the process. *Archives of biochemistry and biophysics* **1984**, *235* (2), 438-448.

(241) Xie, H. R.; Hu, L. S.; Li, G. Y. SH-SY5Y human neuroblastoma cell line: in vitro cell model of dopaminergic neurons in Parkinson's disease. *Chin Med J (Engl)* **2010**, *123* (8), 1086-1092.

(242) Rampersad, S. N. Multiple Applications of Alamar Blue as an Indicator of Metabolic Function and Cellular Health in Cell Viability Bioassays. *Sensors* **2012**, *12* (9), 12347-12360.

(243) Monti, D. A.; Zabrecky, G.; Kremens, D.; Liang, T.-W.; Wintering, N. A.; Cai, J.; Wei, X.; Bazzan, A. J.; Zhong, L.; Bowen, B.; et al. N-Acetyl Cysteine May Support Dopamine Neurons in Parkinson's Disease: Preliminary Clinical and Cell Line Data. *PLOS ONE* **2016**, *11* (6), e0157602.

(244) Kolb, H. C.; Finn, M.; Sharpless, K. B. Click chemistry: diverse chemical function from a few good reactions. *Angewandte Chemie International Edition* **2001**, *40* (11), 2004-2021.

(245) Pierce, E. N.; Piyankarage, S. C.; Dunlap, T.; Litosh, V.; Siklos, M. I.; Wang, Y.-T.; Thatcher, G. R. J. Prodrugs Bioactivated to Quinones Target NF-κB and Multiple Protein Networks: Identification of the Quinonome. *Chemical Research in Toxicology* **2016**, *29* (7), 1151-1159.

(246) Thomas, P. D.; Campbell, M. J.; Kejariwal, A.; Mi, H.; Karlak, B.; Daverman, R.; Diemer, K.; Muruganujan, A.; Narechania, A. PANTHER: a library of protein families and subfamilies indexed by function. *Genome research* **2003**, *13* (9), 2129-2141.

(247) Szklarczyk, D.; Franceschini, A.; Wyder, S.; Forslund, K.; Heller, D.; Huerta-Cepas, J.; Simonovic, M.; Roth, A.; Santos, A.; Tsafou, K. P.; et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Research* **2015**, *43* (D1), D447-D452.

(248) Andreu, C. I.; Woehlbier, U.; Torres, M.; Hetz, C. Protein disulfide isomerases in neurodegeneration: From disease mechanisms to biomedical applications. *FEBS Letters* **2012**, *586* (18), 2826-2834.

(249) Hastings, T. G. The role of dopamine oxidation in mitochondrial dysfunction: implications for Parkinson's disease. *Journal of Bioenergetics and Biomembranes* **2009**, *41* (6), 469-472.

(250) Wang, X.; Thomas, B.; Sachdeva, R.; Arterburn, L.; Frye, L.; Hatcher, P. G.; Cornwell, D. G.; Ma, J. Mechanism of arylating quinone toxicity involving Michael adduct formation and induction of endoplasmic reticulum stress. *Proceedings of the National Academy of Sciences* **2006**, *103* (10), 3604-3609.

(251) Conn, K. J.; Gao, W.; McKee, A.; Lan, M. S.; Ullman, M. D.; Eisenhauer, P. B.; Fine, R. E.; Wells, J. M. Identification of the protein disulfide isomerase family member PDIp in experimental Parkinson's disease and Lewy body pathology. *Brain research* **2004**, *1022* (1-2), 164-172.

(252) Kasprowicz, J.; Kuenen, S.; Miskiewicz, K.; Habets, R. L. P.; Smitz, L.; Verstreken, P. Inactivation of clathrin heavy chain inhibits synaptic recycling but allows bulk membrane uptake. *Journal of Cell Biology* **2008**, *182* (5), 1007-1016.

(253) Ebanks, K.; Lewis, P. A.; Bandopadhyay, R. Vesicular dysfunction and the pathogenesis of Parkinson's disease: clues from genetic studies. *Frontiers in Neuroscience* **2020**, 1381.

(254) Bateman, L. A.; Zaro, B. W.; Miller, S. M.; Pratt, M. R. An Alkyne–Aspirin Chemical Reporter for the Detection of Aspirin-Dependent Protein Modification in Living Cells. *Journal of the American Chemical Society* **2013**, *135* (39), 14568-14573.

(255) Tsukidate, T.; Li, Q.; Hang, H. C. Targeted and proteome-wide analysis of metabolite–protein interactions. *Current opinion in chemical biology* **2020**, *54*, 19-27.

(256) Counihan, J. L.; Ford, B.; Nomura, D. K. Mapping proteome-wide interactions of reactive chemicals using chemoproteomic platforms. *Current opinion in chemical biology* **2016**, *30*, 68-76.

(257) Yang, Y.; Fonović, M.; Verhelst, S. H. L. Cleavable Linkers in Chemical Proteomics Applications. 10.1007/978-1-4939-6439-0_14Springer New York, 2017; pp 185-203.

(258) Klán, P.; Šolomek, T.; Bochet, C. G.; Blanc, A.; Givens, R.; Rubina, M.; Popik, V.; Kostikov, A.; Wirz, J. Photoremovable Protecting Groups in Chemistry and Biology: Reaction Mechanisms and Efficacy. *Chemical Reviews* **2013**, *113* (1), 119-191.

(259) Kand, D.; Liu, P.; Navarro, M. X.; Fischer, L. J.; Rousso-Noori, L.; Friedmann-Morvinski, D.; Winter, A. H.; Miller, E. W.; Weinstain, R. Water-Soluble BODIPY Photocages with Tunable Cellular Localization. *Journal of the American Chemical Society* **2020**, *142* (11), 4970-4974.

(260) Rubinstein, N.; Liu, P.; Miller, E. W.; Weinstain, R. meso-Methylhydroxy BODIPY: a scaffold for photo-labile protecting groups. *Chemical Communications* **2015**, *51* (29), 6369-6372.

(261) Rubner, S.; Scharow, A.; Schubert, S.; Berg, T. Selective degradation of polo-like kinase 1 by a hydrophobically tagged inhibitor of the Polo-box domain. *Angewandte Chemie* **2018**, *130* (52), 17289-17293.

(262) Shi, Y.; Koh, J. T. Light-Activated Transcription and Repression by Using Photocaged SERMs. *ChemBioChem* **2004**, *5* (6), 788-796.

(263) Mahmoodi, M. M.; Abate-Pella, D.; Pundsack, T. J.; Palsuledesai, C. C.; Goff, P. C.; Blank, D. A.; Distefano, M. D. Nitrodibenzofuran: A one-and two-photon sensitive protecting group that is superior to brominated hydroxycoumarin for thiol caging in peptides. *Journal of the American Chemical Society* **2016**, *138* (18), 5848-5859.

(264) Il'ichev, Y. V.; Wirz, J. Rearrangements of 2-nitrobenzyl compounds. 1. Potential energy surface of 2-nitrotoluene and its isomers explored with ab initio and density functional theory methods. *The Journal of Physical Chemistry A* **2000**, *104* (33), 7856-7870.

(265) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; et al. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chemical Biology* **2008**, *3* (6), 373-382.

(266) Holmes, C. P. Model studies for new o-nitrobenzyl photolabile linkers: Substituent effects on the rates of photochemical cleavage. *The Journal of organic chemistry* **1997**, *62* (8), 2370-2380.

(267) Montgomery Jr, E. B. Heavy metals and the etiology of Parkinson's disease and other movement disorders. *Toxicology* **1995**, *97* (1-3), 3-9.

(268) Dexter, D.; Carayon, A.; Javoy-Agid, F.; Agid, Y.; Wells, F.; Daniel, S.; Lees, A.; Jenner, P.; Marsden, C. Alterations in the levels of iron, ferritin and other trace metals in Parkinson's disease and other neurodegenerative diseases affecting the basal ganglia. *Brain* **1991**, *114* (4), 1953-1975.

(269) Smith, L. M.; Kelleher, N. L. Proteoform: a single term describing protein complexity. *Nature Methods* **2013**, *10* (3), 186-187.

(270) Aebersold, R.; Agar, J. N.; Amster, I. J.; Baker, M. S.; Bertozzi, C. R.; Boja, E. S.; Costello, C. E.; Cravatt, B. F.; Fenselau, C.; Garcia, B. A.; et al. How many human proteoforms are there? *Nature Chemical Biology* **2018**, *14* (3), 206-214.

(271) Smith, L. M.; Kelleher, N. L. Proteoforms as the next proteomics currency. *Science* **2018**, *359* (6380), 1106-1107.

(272) Smith, L. M.; Agar, J. N.; Chamot-Rooke, J.; Danis, P. O.; Ge, Y.; Loo, J. A.; Paša-Tolić, L.; Tsybin, Y. O.; Kelleher, N. L. The Human Proteoform Project: Defining the human proteome. *Science Advances* **2021**, *7* (46).

(273) Karve, T. M.; Cheema, A. K. Small changes huge impact: the role of protein posttranslational modifications in cellular homeostasis and disease. *Journal of amino acids* **2011**, *2011*.

(274) Spoel, S. H. Orchestrating the proteome with post-translational modifications. *Journal of Experimental Botany* **2018**, *69* (19), 4499-4503.

(275) Seo, J.-W.; Lee, K.-J. Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *BMB Reports* **2004**, *37* (1), 35-44.

(276) Deribe, Y. L.; Pawson, T.; Dikic, I. Post-translational modifications in signal integration. *Nature Structural & Molecular Biology* **2010**, *17* (6), 666-672.

(277) Jennings, E. Q.; Fritz, K. S.; Galligan, J. J. Biochemical genesis of enzymatic and non-enzymatic post-translational modifications. *Molecular Aspects of Medicine* **2021**, 101053.

(278) Ramazi, S.; Zahiri, J. Posttranslational modifications in proteins: resources, tools and prediction methods. *Database (Oxford)* **2021**, *2021*.

(279) Beltrao, P.; Bork, P.; Krogan, N. J.; Noort, V. Evolution and functional cross-talk of protein post-translational modifications. *Molecular Systems Biology* **2013**, *9* (1), 714.

(280) Harmel, R.; Fiedler, D. Features and regulation of non-enzymatic post-translational modifications. *Nature Chemical Biology* **2018**, *14* (3), 244-252.

(281) McKerrow, J. H. Non-enzymatic, post-translational, amino acid modifications in ageing. A brief review. *Mechanisms of Ageing and Development* **1979**, *10* (6), 371-377.

(282) Zheng, Q.; Maksimovic, I.; Upad, A.; David, Y. Non-enzymatic covalent modifications: a new link between metabolism and epigenetics. *Protein & cell* **2020**, *11* (6), 401-416.

(283) Kosmachevskaya, O. V.; Topunov, A. F. Nonenzymatic Reactions in Metabolism: Their Role in Evolution and Adaptation. *Applied Biochemistry and Microbiology* **2021**, *57* (5), 543-555.

(284) Bollong, M. J.; Lee, G.; Coukos, J. S.; Yun, H.; Zambaldo, C.; Chang, J. W.; Chin, E. N.; Ahmad, I.; Chatterjee, A. K.; Lairson, L. L.; et al. A metabolite-derived protein

modification integrates glycolysis with KEAP1–NRF2 signalling. *Nature* **2018**, *562* (7728), 600-604.

(285) Zheng, Q.; Omans, N. D.; Leicher, R.; Osunsade, A.; Agustinus, A. S.; Finkin-Groner, E.; D'Ambrosio, H.; Liu, B.; Chandarlapaty, S.; Liu, S.; et al. Reversible histone glycation is associated with disease-related changes in chromatin architecture. *Nature Communications* **2019**, *10* (1).

(286) Chaudhuri, J.; Bains, Y.; Guha, S.; Kahn, A.; Hall, D.; Bose, N.; Gugliucci, A.; Kapahi, P. The Role of Advanced Glycation End Products in Aging and Metabolic Diseases: Bridging Association and Causality. *Cell Metabolism* **2018**, *28* (3), 337-352.

(287) Phillips, S. A.; Thornalley, P. J. The formation of methylglyoxal from triose phosphates. Investigation using a specific assay for methylglyoxal. *European Journal of Biochemistry* **1993**, *212* (1), 101-105.

(288) Sjoblom, N. M.; Kelsey, M. M. G.; Scheck, R. A. A Systematic Study of Selective Protein Glycation. *Angewandte Chemie* **2018**, *130* (49), 16309-16314.

(289) Lo, T. W.; Selwood, T.; Thornalley, P. J. The reaction of methylglyoxal with aminoguanidine under physiological conditions and prevention of methylglyoxal binding to plasma proteins. *Biochem Pharmacol* **1994**, *48* (10), 1865-1870.

(290) Frischmann, M.; Bidmon, C.; Angerer, J.; Pischetsrieder, M. Identification of DNA Adducts of Methylglyoxal. *Chemical Research in Toxicology* **2005**, *18* (10), 1586-1592.

(291) Lai, S. W. T.; Lopez Gonzalez, E. J.; Zoukari, T.; Ki, P.; Shuck, S. C. Methylglyoxal and Its Adducts: Induction, Repair, and Association with Disease. *Chem Res Toxicol* **2022**, *35* (10), 1720-1746.

(292) Kold-Christensen, R.; Johannsen, M. Methylglyoxal Metabolism and Aging-Related Disease: Moving from Correlation toward Causation. *Trends Endocrinol Metab* **2020**, *31* (2), 81-92.

(293) Wang, T.; Streeter, M. D.; Spiegel, D. A. Generation and characterization of antibodies against arginine-derived advanced glycation endproducts. *Bioorganic & medicinal chemistry letters* **2015**, *25* (21), 4881-4886.

(294) Wang, T.; Douglass, E. F.; Fitzgerald, K. J.; Spiegel, D. A. A "Turn-On" Fluorescent Sensor for Methylglyoxal. *Journal of the American Chemical Society* **2013**, *135* (33), 12429-12433.

(295) Yang, M.; Fan, J.; Zhang, J.; Du, J.; Peng, X. Visualization of methylglyoxal in living cells and diabetic mice model with a 1,8-naphthalimide-based two-photon fluorescent probe. *Chemical Science* **2018**, *9* (33), 6758-6764.

(296) Sibbersen, C.; Palmfeldt, J.; Hansen, J.; Gregersen, N.; Jørgensen, K. A.; Johannsen, M. Development of a chemical probe for identifying protein targets of α-oxoaldehydes. *Chemical Communications* **2013**, *49* (38), 4012-4014.

(297) Zheng, Q.; Maksimovic, I.; Upad, A.; Guber, D.; David, Y. Synthesis of an Alkynyl Methylglyoxal Probe to Investigate Nonenzymatic Histone Glycation. *The Journal of Organic Chemistry* **2020**, *85* (3), 1691-1697.

(298) Dobler, D.; Ahmed, N.; Song, L.; Eboigbodin, K. E.; Thornalley, P. J. Increased dicarbonyl metabolism in endothelial cells in hyperglycemia induces anoikis and impairs angiogenesis by RGD and GFOGER motif modification. *Diabetes* **2006**, *55* (7), 1961-1969.

(299) Fang, X.; Fu, Y.; Long, M. J. C.; Haegele, J. A.; Ge, E. J.; Parvez, S.; Aye, Y. Temporally Controlled Targeting of 4-Hydroxynonenal to Specific Proteins in Living Cells. *Journal of the American Chemical Society* **2013**, *135* (39), 14496-14499.

(300) Long, M. J. C.; Rogg, C.; Aye, Y. An Oculus to Profile and Probe Target Engagement In Vivo: How T-REX Was Born and Its Evolution into G-REX. *Accounts of Chemical Research* **2021**, *54* (3), 618-631.

(301) Long, M. J. C.; Zhao, Y.; Aye, Y. Neighborhood watch: tools for defining locale-dependent subproteomes and their contextual signaling activities. *RSC Chemical Biology* **2020**, *1* (2), 42-55.

(302) Delomba, W. C.; Stone, E. A.; Alley, K. A.; Iannarone, V.; Tarsis, E.; Ovaska, S.; Ovaska, T. V. Utilization of the Thorpe–Ingold Effect in the Synthesis of Cyclooctanoid Ring Systems via Anionic 6-exo-dig Cyclization/Claisen Rearrangement Sequence. *The Journal of Organic Chemistry* **2020**, *85* (15), 9464-9474.

(303) Reddy, C. R.; Patil, A. D.; Mohammed, S. Z. Oxa-[3+3] annulation of MBH-carbonates of propiolaldehydes with α-nitro/bromo ketones to access 2H-pyrans. *Chemical Communications* **2020**, *56* (52), 7191-7194.

(304) Brinson, R. G.; Jones, P. B. Caged trans-4-Hydroxy-2-nonenal. *Organic Letters* **2004**, *6* (21), 3767-3770.

(305) Jiang, X.; Zhang, C.; Chen, J.; Choi, S.; Zhou, Y.; Zhao, M.; Song, X.; Chen, X.; Maletić-Savatić, M.; Palzkill, T.; et al. Quantitative Real-Time Imaging of Glutathione with Subcellular Resolution. *Antioxidants & Redox Signaling* **2019**, *30* (16), 1900-1910.

(306) Akram, M. Mini-review on Glycolysis and Cancer. *Journal of Cancer Education* **2013**, *28* (3), 454-457.

(307) Chan, C.-M.; Huang, D.-Y.; Huang, Y.-P.; Hsu, S.-H.; Kang, L.-Y.; Shen, C.-M.; Lin, W.-W. Methylglyoxal induces cell death through endoplasmic reticulum stress-associated ROS production and mitochondrial dysfunction. *Journal of Cellular and Molecular Medicine* **2016**, *20* (9), 1749-1760.

(308) Parvez, S.; Long, M. J. C.; Lin, H.-Y.; Zhao, Y.; Haegele, J. A.; Pham, V. N.; Lee, D. K.; Aye, Y. T-REX on-demand redox targeting in live cells. *Nature Protocols* **2016**, *11* (12), 2328-2356.

(309) Galligan, J. J.; Wepy, J. A.; Streeter, M. D.; Kingsley, P. J.; Mitchener, M. M.; Wauchope, O. R.; Beavers, W. N.; Rose, K. L.; Wang, T.; Spiegel, D. A. Methylglyoxal-derived posttranslational arginine modifications are abundant histone marks. *Proceedings of the National Academy of Sciences* **2018**, *115* (37), 9228-9233.

(310) Chang, T.-K.; Shravage, B. V.; Hayes, S. D.; Powers, C. M.; Simin, R. T.; Wade Harper, J.; Baehrecke, E. H. Uba1 functions in Atg7-and Atg3-independent autophagy. *Nature cell biology* **2013**, *15* (9), 1067-1078.

(311) Lv, Z.; Williams, K. M.; Yuan, L.; Atkison, J. H.; Olsen, S. K. Crystal structure of a human ubiquitin E1-ubiquitin complex reveals conserved functional elements essential for activity. *J Biol Chem* **2018**, *293* (47), 18337-18352.

(312) Milo, R.; Jorgensen, P.; Moran, U.; Weber, G.; Springer, M. BioNumbers—the database of key numbers in molecular and cell biology. *Nucleic Acids Research* **2010**, *38* (suppl_1), D750-D753.

(313) McKenna, S. M.; Fay, E. M.; McGouran, J. F. Flipping the Switch: Innovations in Inducible Probes for Protein Profiling. *ACS Chem Biol* **2021**, *16* (12), 2719-2730.

(314) Lacoursiere, R. E.; Hadi, D.; Shaw, G. S. Acetylation, Phosphorylation, Ubiquitination (Oh My!): Following Post-Translational Modifications on the Ubiquitin Road. *Biomolecules* **2022**, *12* (3).

(315) Groen, E. J.; Gillingwater, T. H. UBA1: at the crossroads of ubiquitin homeostasis and neurodegeneration. *Trends in molecular medicine* **2015**, *21* (10), 622-632.

(316) Merk, O.; Speit, G. Significance of formaldehyde-induced DNA–protein crosslinks for mutagenesis. *Environmental and Molecular Mutagenesis* **1998**, *32* (3), 260-268.

(317) Murata-Kamiya, N.; Kamiya, H. Methylglyoxal, an endogenous aldehyde, crosslinks DNA polymerase and the substrate DNA. *Nucleic Acids Res* **2001**, *29* (16), 3433-3438.

(318) Michaelson-Richie, E. D.; Loeber, R. L.; Codreanu, S. G.; Ming, X.; Liebler, D. C.; Campbell, C.; Tretyakova, N. Y. DNA− protein cross-linking by 1, 2, 3, 4-diepoxybutane. *Journal of proteome research* **2010**, *9* (9), 4356-4367.

(319) Loeber, R.; Rajesh, M.; Fang, Q.; Pegg, A. E.; Tretyakova, N. Cross-linking of the human DNA repair protein O 6-alkylguanine DNA alkyltransferase to DNA in the presence of 1, 2, 3, 4-diepoxybutane. *Chemical research in toxicology* **2006**, *19* (5), 645-654.

(320) Gherezghiher, T. B.; Ming, X.; Villalta, P. W.; Campbell, C.; Tretyakova, N. Y. 1, 2, 3, 4-Diepoxybutane-induced DNA–protein cross-linking in human fibrosarcoma (HT1080) cells. *Journal of proteome research* **2013**, *12* (5), 2151-2164.

(321) Barker, S.; Weinfeld, M.; Zheng, J.; Li, L.; Murray, D. Identification of mammalian proteins cross-linked to DNA by ionizing radiation. *Journal of Biological Chemistry* **2005**, *280* (40), 33826-33838.

(322) Barker, S.; Weinfeld, M.; Murray, D. DNA–protein crosslinks: their induction, repair, and biological consequences. *Mutation Research/Reviews in Mutation Research* **2005**, *589* (2), 111-135.

(323) Zhitkovich, A.; Voitkun, V.; Kluz, T.; Costa, M. Utilization of DNA-protein cross-links as a biomarker of chromium exposure. *Environmental Health Perspectives* **1998**, *106* (suppl 4), 969-974.

(324) Kloster, M.; Kostrhunova, H.; Zaludova, R.; Malina, J.; Kasparkova, J.; Brabec, V.; Farrell, N. Trifunctional dinuclear platinum complexes as DNA− protein cross-linking agents. *Biochemistry* **2004**, *43* (24), 7776-7786.

(325) Ewig, R. A.; Kohn, K. W. DNA damage and repair in mouse leukemia L1210 cells treated with nitrogen mustard, 1, 3-bis (2-chloroethyl)-1-nitrosourea, and other nitrosoureas. *Cancer research* **1977**, *37* (7_Part_1), 2114-2122.

(326) Michaelson-Richie, E. D.; Ming, X.; Codreanu, S. G.; Loeber, R. L.; Liebler, D. C.; Campbell, C.; Tretyakova, N. Y. Mechlorethamine-induced DNA–protein cross-linking in human fibrosarcoma (HT1080) cells. *Journal of proteome research* **2011**, *10* (6), 2785-2796.

(327) Baker, J. M.; Parish, J.; Curtis, J. P. DNA-DNA and DNA-protein crosslinking and repair in Neurospora crassa following exposure to nitrogen mustard. *Mutation Research/DNA Repair Reports* **1984**, *132* (5-6), 171-179.

(328) Ewig, R. A.; Kohn, K. W. DNA-protein cross-linking and DNA interstrand cross-linking by haloethylnitrosoureas in L1210 cells. *Cancer research* **1978**, *38* (10), 3197-3203.

(329) Wickramaratne, S.; Ji, S.; Mukherjee, S.; Su, Y.; Pence, M. G.; Lior-Hoffmann, L.; Fu, I.; Broyde, S.; Guengerich, F. P.; Distefano, M. Bypass of DNA-protein cross-links conjugated to the 7-deazaguanine position of DNA by translesion synthesis polymerases. *Journal of Biological Chemistry* **2016**, *291* (45), 23589-23603.

(330) Oleinick, N. L.; Chiu, S.; Ramakrishnan, N.; Xue, L. Y. The formation, identification, and significance of DNA-protein cross-links in mammalian cells. *The British journal of cancer. Supplement* **1987**, *8*, 135.

(331) Wickramaratne, S.; Boldry, E. J.; Buehler, C.; Wang, Y.-C.; Distefano, M. D.; Tretyakova, N. Y. Error-prone translesion synthesis past DNA-peptide cross-links conjugated to the major groove of DNA via C5 of thymidine. *Journal of Biological Chemistry* **2015**, *290* (2), 775-787.

(332) Tretyakova, N. Y.; Michaelson-Richie, E. D.; Gherezghiher, T. B.; Kurtz, J.; Ming, X.; Wickramaratne, S.; Campion, M.; Kanugula, S.; Pegg, A. E.; Campbell, C. DNA-reactive protein monoepoxides induce cell death and mutagenesis in mammalian cells. *Biochemistry* **2013**, *52* (18), 3171-3181.

(333) Tretyakova, N. Y.; Groehler IV, A.; Ji, S. DNA–protein cross-links: formation, structural identities, and biological outcomes. *Accounts of chemical research* **2015**, *48* (6), 1631-1644.

(334) Groehler IV, A.; Degner, A.; Tretyakova, N. Y. Mass Spectrometry-Based Tools to Characterize DNA–Protein Cross-Linking by Bis-Electrophiles. *Basic & clinical pharmacology & toxicology* **2017**, *121*, 63-77.

(335) Ruijs, M.; Van Andel, R.; Oshima, J.; Madan, K.; Nieuwint, A.; Aalfs, C. Atypical progeroid syndrome: an unknown helicase gene defect? *American Journal of Medical Genetics Part A* **2003**, *116* (3), 295-299.

(336) Lessel, D.; Vaz, B.; Halder, S.; Lockhart, P. J.; Marinovic-Terzic, I.; Lopez-Mosqueda, J.; Philipp, M.; Sim, J. C.; Smith, K. R.; Oehler, J. Mutations in SPRTN cause early onset hepatocellular carcinoma, genomic instability and progeroid features. *Nature genetics* **2014**, *46* (11), 1239-1244.

(337) Lopez-Mosqueda, J.; Maddi, K.; Prgomet, S.; Kalayil, S.; Marinovic-Terzic, I.; Terzic, J.; Dikic, I. SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. *Elife* **2016**, *5*, e21491.

(338) Rabbani, N.; Thornalley, P. J. Measurement of methylglyoxal by stable isotopic dilution analysis LC-MS/MS with corroborative prediction in physiological samples. *Nature protocols* **2014**, *9* (8), 1969-1979.

(339) Thornalley, P. J. Pharmacology of methylglyoxal: formation, modification of proteins and nucleic acids, and enzymatic detoxification-a role in pathogenesis and antiproliferative chemotherapy. *General Pharmacology: The Vascular System* **1996**, *27* (4), 565-573.

(340) Kilhovd, B. K.; Juutilainen, A.; Lehto, S.; Rönnemaa, T.; Torjesen, P. A.; Hanssen, K. F.; Laakso, M. Increased serum levels of methylglyoxal-derived hydroimidazolone-

AGE are associated with increased cardiovascular disease mortality in nondiabetic women. *Atherosclerosis* **2009**, *205* (2), 590-594.

(341) Agalou, S.; Ahmed, N.; Babaei-Jadidi, R.; Dawnay, A.; Thornalley, P. J. Profound mishandling of protein glycation degradation products in uremia and dialysis. *Journal of the American Society of Nephrology* **2005**, *16* (5), 1471-1485.

(342) Giacco, F.; Du, X.; D'Agati, V. D.; Milne, R.; Sui, G.; Geoffrion, M.; Brownlee, M. Knockdown of glyoxalase 1 mimics diabetic nephropathy in nondiabetic mice. *Diabetes* **2014**, *63* (1), 291-299.

(343) El-Osta, A.; Brasacchio, D.; Yao, D.; Pocai, A.; Jones, P. L.; Roeder, R. G.; Cooper, M. E.; Brownlee, M. Transient high glucose causes persistent epigenetic changes and altered gene expression during subsequent normoglycemia. *The Journal of experimental medicine* **2008**, *205* (10), 2409-2417.

(344) Thornalley, P. J. Protein and nucleotide damage by glyoxal and methylglyoxal in physiological systems-role in ageing and disease. *Drug metabolism and drug interactions* **2008**, *23* (1-2), 125-150.

(345) Bidmon, C.; Frischmann, M.; Pischetsrieder, M. Analysis of DNA-bound advanced glycation end-products by LC and mass spectrometry. *Journal of Chromatography B* **2007**, *855* (1), 51-58.

(346) Petrova, K. V.; Millsap, A. D.; Stec, D. F.; Rizzo, C. J. Characterization of the Deoxyguanosine–Lysine Cross-Link of Methylglyoxal. *Chemical Research in Toxicology* **2014**, *27* (6), 1019-1029.

(347) Zheng, Q.; Omans, N. D.; Leicher, R.; Osunsade, A.; Agustinus, A. S.; Finkin-Groner, E.; D'Ambrosio, H.; Liu, B.; Chandarlapaty, S.; Liu, S. Reversible histone glycation is associated with disease-related changes in chromatin architecture. *Nature Communications* **2019**, *10* (1), 1289.

(348) Vaz, B.; Popovic, M.; Newman, J. A.; Fielden, J.; Aitkenhead, H.; Halder, S.; Singh, A. N.; Vendrell, I.; Fischer, R.; Torrecilla, I. Metalloprotease SPRTN/DVC1 orchestrates replication-coupled DNA-protein crosslink repair. *Molecular cell* **2016**, *64* (4), 704-719.

(349) Rasheed, S.; Nelson-Rees, W. A.; Toth, E. M.; Arnstein, P.; Gardner, M. B. Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer* **1974**, *33* (4), 1027-1033.

(350) Maskey, R. S.; Kim, M. S.; Baker, D. J.; Childs, B.; Malureanu, L. A.; Jeganathan, K. B.; Machida, Y.; Van Deursen, J. M.; Machida, Y. J. Spartan deficiency causes genomic instability and progeroid phenotypes. *Nature communications* **2014**, *5* (1), 5744.

(351) Zhitkovich, A.; Costa, M. A simple, sensitive assay to detect DNA–protein cromlinks in intact cells and in vivo. *Carcinogenesis* **1992**, *13* (8), 1485-1489.

(352) Sherman, B. T.; Hao, M.; Qiu, J.; Jiao, X.; Baseler, M. W.; Lane, H. C.; Imamichi, T.; Chang, W. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic acids research* **2022**, *50* (W1), W216-W221.

(353) Sun, Y.; Miller Jenkins, L. M.; Su, Y. P.; Nitiss, K. C.; Nitiss, J. L.; Pommier, Y. A conserved SUMO pathway repairs topoisomerase DNA-protein cross-links by engaging ubiquitin-mediated proteasomal degradation. *Science Advances* **2020**, *6* (46), eaba6290.

(354) Ming, X.; Groehler IV, A.; Michaelson-Richie, E. D.; Villalta, P. W.; Campbell, C.; Tretyakova, N. Y. Mass spectrometry based proteomics study of cisplatin-Induced DNA–protein cross-linking in human fibrosarcoma (HT1080) cells. *Chemical research in toxicology* **2017**, *30* (4), 980-995.

(355) Loeber, R. L.; Michaelson-Richie, E. D.; Codreanu, S. G.; Liebler, D. C.; Campbell, C. R.; Tretyakova, N. Y. Proteomic analysis of DNA− protein cross-linking by antitumor nitrogen mustards. *Chemical research in toxicology* **2009**, *22* (6), 1151-1162.

(356) Lee, J.-y.; Song, J.; Kwon, K.; Jang, S.; Kim, C.; Baek, K.; Kim, J.; Park, C. Human DJ-1 and its homologs are novel glyoxalases. *Human molecular genetics* **2012**, *21* (14), 3215-3225.

(357) Morgenstern, J.; Fleming, T.; Schumacher, D.; Eckstein, V.; Freichel, M.; Herzig, S.; Nawroth, P. Loss of glyoxalase 1 induces compensatory mechanism to achieve dicarbonyl detoxification in mammalian Schwann cells. *Journal of Biological Chemistry* **2017**, *292* (8), 3224-3238.

(358) Schumacher, D.; Morgenstern, J.; Oguchi, Y.; Volk, N.; Kopf, S.; Groener, J. B.; Nawroth, P. P.; Fleming, T.; Freichel, M. Compensatory mechanisms for methylglyoxal detoxification in experimental & clinical diabetes. *Molecular metabolism* **2018**, *18*, 143-152.

(359) Vander Jagt, D. L.; Hunsaker, L. A. Methylglyoxal metabolism and diabetic complications: roles of aldose reductase, glyoxalase-I, betaine aldehyde dehydrogenase and 2-oxoaldehyde dehydrogenase. *Chemico-biological interactions* **2003**, *143*, 341-351.

(360) Vander Jagt, D. L.; Robinson, B.; Taylor, K.; Hunsaker, L. Reduction of trioses by NADPH-dependent aldo-keto reductases. Aldose reductase, methylglyoxal, and diabetic complications. *Journal of Biological Chemistry* **1992**, *267* (7), 4364-4369.

(361) Maskey, R. S.; Flatten, K. S.; Sieben, C. J.; Peterson, K. L.; Baker, D. J.; Nam, H.-J.; Kim, M. S.; Smyrk, T. C.; Kojima, Y.; Machida, Y. Spartan deficiency causes accumulation of Topoisomerase 1 cleavage complexes and tumorigenesis. *Nucleic acids research* **2017**, *45* (8), 4564-4576.

(362) Sambrook, J.; Russell, D. W. Purification of nucleic acids by extraction with phenol: chloroform. *Cold Spring Harbor Protocols* **2006**, *2006* (1), pdb. prot4455.

(363) Groehler IV, A.; Villalta, P. W.; Campbell, C.; Tretyakova, N. Covalent DNA–protein cross-linking by phosphoramide mustard and nornitrogen mustard in human cells. *Chemical research in toxicology* **2016**, *29* (2), 190-202.

(364) Huang, D. W.; Sherman, B. T.; Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **2009**, *4* (1), 44-57.

(365) Huang, D. W.; Sherman, B. T.; Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research* **2009**, *37* (1), 1-13.

(366) Mélèse, T.; Xue, Z. The nucleolus: an organelle formed by the act of building a ribosome. *Current opinion in cell biology* **1995**, *7* (3), 319-324.

(367) Stützer, A.; Welp, L. M.; Raabe, M.; Sachsenberg, T.; Kappert, C.; Wulf, A.; Lau, A. M.; David, S.-S.; Chernev, A.; Kramer, K.; et al. Analysis of protein-DNA interactions in chromatin by UV induced cross-linking and mass spectrometry. *Nature Communications* **2020**, *11* (1), 5250.

(368) Demarse, N. A.; Ponnusamy, S.; Spicer, E. K.; Apohan, E.; Baatz, J. E.; Ogretmen, B.; Davies, C. Direct binding of glyceraldehyde 3-phosphate dehydrogenase to telomeric DNA protects telomeres against chemotherapy-induced rapid degradation. *Journal of molecular biology* **2009**, *394* (4), 789-803.

(369) White, M. R.; Khan, M. M.; Deredge, D.; Ross, C. R.; Quintyn, R.; Zucconi, B. E.; Wysocki, V. H.; Wintrode, P. L.; Wilson, G. M.; Garcin, E. D. A dimer interface mutation in glyceraldehyde-3-phosphate dehydrogenase regulates its binding to AU-rich RNA. *Journal of Biological Chemistry* **2015**, *290* (3), 1770-1785.

(370) Demarse, N. A.; Ponnusamy, S.; Spicer, E. K.; Apohan, E.; Baatz, J. E.; Ogretmen, B.; Davies, C. Direct binding of glyceraldehyde 3-phosphate dehydrogenase to telomeric DNA protects telomeres against chemotherapy-induced rapid degradation. *J Mol Biol* **2009**, *394* (4), 789-803.

(371) Long, M. J.; Huang, K.-T.; Aye, Y. The not so identical twins:(dis) similarities between reactive electrophile and oxidant sensing and signaling. *Chemical Society Reviews* **2021**, *50* (22), 12269-12291.

(372) Schalkwijk, C.; Stehouwer, C. Methylglyoxal, a highly reactive dicarbonyl compound, in diabetes, its vascular complications, and other age-related diseases. *Physiological reviews* **2020**, *100* (1), 407-461.

(373) Vaca, C. E.; Fang, J.-L.; Conradi, M.; Hou, S.-M. Development of a 32P-postlabelling method for the analysis of 2'-deoxyguanosine-3'-monophosphate and DNA adducts of methylglyoxal. *Carcinogenesis* **1994**, *15* (9), 1887-1894.

(374) Oya, T.; Hattori, N.; Mizuno, Y.; Miyata, S.; Maeda, S.; Osawa, T.; Uchida, K. Methylglyoxal modification of protein: chemical and immunochemical characterization of methylglyoxal-arginine adducts. *Journal of Biological Chemistry* **1999**, *274* (26), 18492-18502.

(375) Chakraborty, S.; Karmakar, K.; Chakravortty, D. Cells producing their own nemesis: understanding methylglyoxal metabolism. *IUBMB life* **2014**, *66* (10), 667-678.

(376) Sibbersen, C.; Schou Oxvig, A.-M.; Bisgaard Olesen, S.; Nielsen, C. B.; Galligan, J. J.; Jørgensen, K. A.; Palmfeldt, J.; Johannsen, M. Profiling of methylglyoxal blood metabolism and advanced glycation end-product proteome using a chemical probe. *ACS Chemical Biology* **2018**, *13* (12), 3294-3305.

(377) Coukos, J. S.; Moellering, R. E. Methylglyoxal forms diverse mercaptomethylimidazole crosslinks with thiol and guanidine pairs in endogenous metabolites and proteins. *ACS Chemical Biology* **2021**, *16* (11), 2453-2461.

(378) Chen, X.; Liu, Y.; Kong, L.; Wen, Z.; Wang, W.; Wang, C. Quantitative Chemoproteomic Profiling of Protein Cross-Links Induced by Methylglyoxal. *ACS Chemical Biology* **2022**, *17* (8), 2010-2017.

(379) Coukos, J. S.; Lee, C. W.; Pillai, K. S.; Liu, K. J.; Moellering, R. E. Widespread, Reversible Cysteine Modification by Methylglyoxal Regulates Metabolic Enzyme Function. *ACS Chemical Biology* **2022**.

(380) Rabbani, N.; Thornalley, P. J. Dicarbonyl stress in cell and tissue dysfunction contributing to ageing and disease. *Biochemical and biophysical research communications* **2015**, *458* (2), 221-226.

(381) Rabbani, N.; Thornalley, P. J. Advanced glycation end products in the pathogenesis of chronic kidney disease. *Kidney international* **2018**, *93* (4), 803-813.

(382) Rabbani, N.; Xue, M.; Thornalley, P. J. Activity, regulation, copy number and function in the glyoxalase system. *Biochemical Society transactions* **2014**, *42* (2), 419-424.

(383) Distler, M. G.; Palmer, A. A. Role of Glyoxalase 1 (Glo1) and methylglyoxal (MG) in behavior: recent advances and mechanistic insights. *Frontiers in genetics* **2012**, *3*, 250.

(384) Wang, X.; He, S.; Sun, J. M.; Delcuve, G. P.; Davie, J. R. Selective association of peroxiredoxin 1 with genomic DNA and COX-2 upstream promoter elements in estrogen receptor negative breast cancer cells. *Mol Biol Cell* **2010**, *21* (17), 2987-2995.

(385) Vilanova, B.; Fernández, D.; Casasnovas, R.; Pomar, A.; Alvarez-Idaboy, J.; Hernández-Haro, N.; Grand, A.; Adrover, M.; Donoso, J.; Frau, J. Formation mechanism of glyoxal-DNA adduct, a DNA cross-link precursor. *International journal of biological macromolecules* **2017**, *98*, 664-675.

(386) Shaham, J.; Bomstein, Y.; Meltzer, A.; Kaufman, Z.; Palma, E.; Ribak, J. DNA-protein crosslinks, a biomarker of exposure to formaldehyde—in vitro and in vivo studies. *Carcinogenesis* **1996**, *17* (1), 121-126.

(387) Nicholls, C.; Li, H.; Liu, J. P. GAPDH: a common enzyme with uncommon functions. *Clinical and Experimental Pharmacology and Physiology* **2012**, *39* (8), 674-679.

(388) Tristan, C.; Shahani, N.; Sedlak, T. W.; Sawa, A. The diverse functions of GAPDH: views from different subcellular compartments. *Cellular signalling* **2011**, *23* (2), 317-323.

(389) Colell, A.; Green, D. R.; Ricci, J. E. Novel roles for GAPDH in cell death and carcinogenesis. *Cell Death & Differentiation* **2009**, *16* (12), 1573-1581.

(390) Maessen, D. E.; Stehouwer, C. D.; Schalkwijk, C. G. The role of methylglyoxal and the glyoxalase system in diabetes and other age-related diseases. *Clinical science* **2015**, *128* (12), 839-861.

(391) Aubert, G.; Lansdorp, P. M. Telomeres and aging. *Physiological reviews* **2008**, *88* (2), 557-579.

(392) Lee, H. J.; Howell, S. K.; Sanford, R. J.; Beisswenger, P. J. Methylglyoxal can modify GAPDH activity and structure. *Ann N Y Acad Sci* **2005**, *1043*, 135-145.

(393) Rodrigues, D. C.; Harvey, E. M.; Suraj, R.; Erickson, S. L.; Mohammad, L.; Ren, M.; Liu, H.; He, G.; Kaplan, D. R.; Ellis, J.; et al. Methylglyoxal couples metabolic and translational control of Notch signalling in mammalian neural stem cells. *Nat Commun* **2020**, *11* (1), 2018.

(394) Stingele, J.; Bellelli, R.; Boulton, S. J. Mechanisms of DNA–protein crosslink repair. *Nature Reviews Molecular Cell Biology* **2017**, *18* (9), 563-573.

(395) Stingele, J.; Bellelli, R.; Alte, F.; Hewitt, G.; Sarek, G.; Maslen, S. L.; Tsutakawa, S. E.; Borg, A.; Kjær, S.; Tainer, J. A. Mechanism and regulation of DNA-protein crosslink repair by the DNA-dependent metalloprotease SPRTN. *Molecular cell* **2016**, *64* (4), 688-703.

(396) Nakano, T.; Morishita, S.; Katafuchi, A.; Matsubara, M.; Horikawa, Y.; Terato, H.; Salem, A. M.; Izumi, S.; Pack, S. P.; Makino, K. Nucleotide excision repair and homologous recombination systems commit differentially to the repair of DNA-protein crosslinks. *Molecular cell* **2007**, *28* (1), 147-158.

(397) Nakano, T.; Katafuchi, A.; Matsubara, M.; Terato, H.; Tsuboi, T.; Masuda, T.; Tatsumoto, T.; Pack, S. P.; Makino, K.; Croteau, D. L. Homologous recombination but not nucleotide excision repair plays a pivotal role in tolerance of DNA-protein cross-links in mammalian cells. *Journal of Biological Chemistry* **2009**, *284* (40), 27065-27076.

(398) Duxin, J. P.; Dewar, J. M.; Yardimci, H.; Walter, J. C. Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell* **2014**, *159* (2), 346-357.

(399) Tirado-Magallanes, R.; Rebbani, K.; Lim, R.; Pradhan, S.; Benoukraf, T. Whole genome DNA methylation: beyond genes silencing. *Oncotarget* **2017**, *8* (3), 5629.

(400) Moore, L. D.; Le, T.; Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* **2013**, *38* (1), 23-38.

(401) Mulholland, C. B.; Traube, F. R.; Ugur, E.; Parsa, E.; Eckl, E.-M.; Schönung, M.; Modic, M.; Bartoschek, M. D.; Stolz, P.; Ryan, J. Distinct and stage-specific

contributions of TET1 and TET2 to stepwise cytosine oxidation in the transition from naive to primed pluripotency. *Scientific reports* **2020**, *10* (1), 1-12.

(402) Seiler, C. L.; Fernandez, J.; Koerperich, Z.; Andersen, M. P.; Kotandeniya, D.; Nguyen, M. E.; Sham, Y. Y.; Tretyakova, N. Y. Maintenance DNA methyltransferase activity in the presence of oxidized forms of 5-methylcytosine: structural basis for ten eleven translocation-mediated DNA demethylation. *Biochemistry* **2018**, *57* (42), 6061-6069.

(403) Jacobs, A. L.; Schär, P. DNA glycosylases: in DNA repair and beyond. *Chromosoma* **2012**, *121* (1), 1-20.

(404) Wu, H.; Zhang, Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes & development* **2011**, *25* (23), 2436-2452.

(405) Hu, L.; Lu, J.; Cheng, J.; Rao, Q.; Li, Z.; Hou, H.; Lou, Z.; Zhang, L.; Li, W.; Gong, W. Structural insight into substrate preference for TET-mediated oxidation. *Nature* **2015**, *527* (7576), 118-122.

(406) Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **2009**, *324* (5929), 930-935.

(407) Yang, J.; Bashkenova, N.; Zang, R.; Huang, X.; Wang, J. The roles of TET family proteins in development and stem cells. *Development* **2020**, *147* (2), dev183129.

(408) Ko, M.; An, J.; Pastor, W. A.; Koralov, S. B.; Rajewsky, K.; Rao, A. TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunological reviews* **2015**, *263* (1), 6-21.

(409) Wu, J.; Li, H.; Shi, M.; Zhu, Y.; Ma, Y.; Zhong, Y.; Xiong, C.; Chen, H.; Peng, C. TET1-mediated DNA hydroxymethylation activates inhibitors of the Wnt/β-catenin signaling pathway to suppress EMT in pancreatic tumor cells. *Journal of Experimental & Clinical Cancer Research* **2019**, *38* (1), 1-17.

(410) Lemonnier, F.; Poullot, E.; Dupuy, A.; Couronné, L.; Martin, N.; Scourzic, L.; Fataccioli, V.; Bruneau, J.; Cairns, R. A.; Mak, T. W. Loss of 5-hydroxymethylcytosine is a frequent event in peripheral T-cell lymphomas. *haematologica* **2018**, *103* (3), e115.

(411) Ko, M.; Huang, Y.; Jankowska, A. M.; Pape, U. J.; Tahiliani, M.; Bandukwala, H. S.; An, J.; Lamperti, E. D.; Koh, K. P.; Ganetzky, R. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **2010**, *468* (7325), 839-843.

(412) Li, Z.; Cai, X.; Cai, C.-L.; Wang, J.; Zhang, W.; Petersen, B. E.; Yang, F.-C.; Xu, M. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood, The Journal of the American Society of Hematology* **2011**, *118* (17), 4509-4518.

(413) Moran-Crusio, K.; Reavie, L.; Shih, A.; Abdel-Wahab, O.; Ndiaye-Lobry, D.; Lobry, C.; Figueroa, M. E.; Vasanthakumar, A.; Patel, J.; Zhao, X. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer cell* **2011**, *20* (1), 11-24.

(414) Xu, Y.-p.; Lv, L.; Liu, Y.; Smith, M. D.; Li, W.-C.; Tan, X.-m.; Cheng, M.; Li, Z.; Bovino, M.; Aubé, J. Tumor suppressor TET2 promotes cancer immunity and immunotherapy efficacy. *The Journal of clinical investigation* **2019**, *129* (10), 4316-4331.

(415) Yu, S.; Yin, Y.; Hong, S.; Cao, S.; Huang, Y.; Chen, S.; Liu, Y.; Guan, H.; Zhang, Q.; Li, Y. TET1 is a tumor suppressor that inhibits papillary thyroid carcinoma cell migration and invasion. *International journal of endocrinology* **2020**, *2020*.

(416) Coltro, G.; Mangaonkar, A. A.; Lasho, T. L.; Finke, C. M.; Pophali, P.; Carr, R.; Gangat, N.; Binder, M.; Pardanani, A.; Fernandez-Zapico, M. Clinical, molecular, and prognostic correlates of number, type, and functional localization of TET2 mutations in chronic myelomonocytic leukemia (CMML)—a study of 1084 patients. *Leukemia* **2020**, *34* (5), 1407-1421.

(417) Rivas, M. P.; Aguiar, T. F. M.; Fernandes, G. R.; Caires-Júnior, L. C.; Goulart, E.; Telles-Silva, K. A.; Cypriano, M.; De Toledo, S. R. C.; Rosenberg, C.; Carraro, D. M. TET upregulation leads to 5-hydroxymethylation enrichment in hepatoblastoma. *Frontiers in genetics* **2019**, *10*, 553.

(418) Wu, M.-Z.; Chen, S.-F.; Nieh, S.; Benner, C.; Ger, L.-P.; Jan, C.-I.; Ma, L.; Chen, C.-H.; Hishida, T.; Chang, H.-T. Hypoxia Drives Breast Tumor Malignancy through a TET–TNFα–p38–MAPK Signaling AxisHypoxia Drives Breast Tumor Malignancy. *Cancer research* **2015**, *75* (18), 3912-3924.

(419) Huang, H.; Jiang, X.; Li, Z.; Li, Y.; Song, C.-X.; He, C.; Sun, M.; Chen, P.; Gurbuxani, S.; Wang, J. TET1 plays an essential oncogenic role in MLL-rearranged

leukemia. *Proceedings of the National Academy of Sciences* **2013**, *110* (29), 11994-11999.

(420) Bamezai, S.; Demir, D.; Pulikkottil, A. J.; Ciccarone, F.; Fischbein, E.; Sinha, A.; Borga, C.; Te Kronnie, G.; Meyer, L.-H.; Mohr, F. TET1 promotes growth of T-cell acute lymphoblastic leukemia and can be antagonized via PARP inhibition. *Leukemia* **2021**, *35* (2), 389-403.

(421) Chou, W.-C.; Chou, S.-C.; Liu, C.-Y.; Chen, C.-Y.; Hou, H.-A.; Kuo, Y.-Y.; Lee, M.-C.; Ko, B.-S.; Tang, J.-L.; Yao, M. TET2 mutation is an unfavorable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics. *Blood, The Journal of the American Society of Hematology* **2011**, *118* (14), 3803-3810.

(422) Guan, Y.; Tiwari, A. D.; Phillips, J. G.; Hasipek, M.; Grabowski, D. R.; Pagliuca, S.; Gopal, P.; Kerr, C. M.; Adema, V.; Radivoyevitch, T. A Therapeutic Strategy for Preferential Targeting of TET2-Mutant and TET Dioxygenase–Deficient Cells in Myeloid NeoplasmsSynthetic Lethality by TET Dioxygenase Inhibition. *Blood cancer discovery* **2021**, *2* (2), 146-161.

(423) Ganesan, A. Epigenetic drug discovery: a success story for cofactor interference. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2018**, *373* (1748), 20170069.

(424) Rose, N. R.; McDonough, M. A.; King, O. N.; Kawamura, A.; Schofield, C. J. Inhibition of 2-oxoglutarate dependent oxygenases. *Chemical Society reviews* **2011**, *40* (8), 4364-4397.

(425) Clifton, I. J.; McDonough, M. A.; Ehrismann, D.; Kershaw, N. J.; Granatino, N.; Schofield, C. J. Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *Journal of inorganic biochemistry* **2006**, *100* (4), 644-669.

(426) Yeoh, K. K.; Chan, M. C.; Thalhammer, A.; Demetriades, M.; Chowdhury, R.; Tian, Y.-M.; Stolze, I.; McNeill, L. A.; Lee, M. K.; Woon, E. C. Dual-action inhibitors of HIF prolyl hydroxylases that induce binding of a second iron ion. *Organic & biomolecular chemistry* **2013**, *11* (5), 732-745.

(427) Mole, D. R.; Schlemminger, I.; McNeill, L. A.; Hewitson, K. S.; Pugh, C. W.; Ratcliffe, P. J.; Schofield, C. J. 2-oxoglutarate analogue inhibitors of HIF prolyl hydroxylase. *Bioorganic & medicinal chemistry letters* **2003**, *13* (16), 2677-2680.

(428) Singh, A. K.; Zhao, B.; Liu, X.; Wang, X.; Li, H.; Qin, H.; Wu, X.; Ma, Y.; Horne, D.; Yu, X. Selective targeting of TET catalytic domain promotes somatic cell reprogramming. *Proceedings of the National Academy of Sciences* **2020**, *117* (7), 3621-3626.

(429) Baell, J. B.; Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *Journal of medicinal chemistry* **2010**, *53* (7), 2719-2740.

(430) Dahlin, J. L.; Walters, M. A. How to triage PAINS-full research. *Assay and drug development technologies* **2016**, *14* (3), 168-174.

(431) Chua, G. N.; Wassarman, K. L.; Sun, H.; Alp, J. A.; Jarczyk, E. I.; Kuzio, N. J.; Bennett, M. J.; Malachowsky, B. G.; Kruse, M.; Kennedy, A. J. Cytosine-based TET enzyme inhibitors. *ACS medicinal chemistry letters* **2019**, *10* (2), 180-185.

(432) Seiler, C. L.; Song, J. M.; Fernandez, J.; Abrahante, J. E.; Kono, T. J.; Chen, Y.; Ren, Y.; Kassie, F.; Tretyakova, N. Y. Epigenetic changes in alveolar type II lung cells of A/J mice following intranasal treatment with lipopolysaccharide. *Chemical research in toxicology* **2019**, *32* (5), 831-839.

(433) Yan, F.; Shen, N.; Pang, J.; Xie, D.; Deng, B.; Molina, J.; Yang, P.; Liu, S. Restoration of miR-101 suppresses lung tumorigenesis through inhibition of DNMT3a-dependent DNA methylation. *Cell death & disease* **2014**, *5* (9), e1413-e1413.

(434) Bian, K.; Chen, F.; Humulock, Z. T.; Tang, Q.; Li, D. Copper inhibits the AlkB family DNA repair enzymes under Wilson's disease condition. *Chemical research in toxicology* **2017**, *30* (10), 1794-1796.

(435) Bleijlevens, B.; Shivarattan, T.; Sedgwick, B.; Rigby, S. E.; Matthews, S. J. Replacement of non-heme Fe (II) with Cu (II) in the α-ketoglutarate dependent DNA repair enzyme AlkB: Spectroscopic characterization of the active site. *Journal of inorganic biochemistry* **2007**, *101* (7), 1043-1048.

(436) Robertson, K. D. DNA methylation and human disease. *Nature Reviews Genetics* **2005**, *6* (8), 597-610.

(437) Kulis, M.; Esteller, M. DNA methylation and cancer. *Advances in genetics* **2010**, *70*, 27-56.

(438) Kohli, R. M.; Zhang, Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **2013**, *502* (7472), 472-479.

(439) Rasmussen, K. D.; Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes & development* **2016**, *30* (7), 733-750.

(440) Lio, C.-W. J.; Yue, X.; López-Moyado, I. F.; Tahiliani, M.; Aravind, L.; Rao, A. TET methylcytosine oxidases: new insights from a decade of research. *Journal of biosciences* **2020**, *45*, 1-14.

(441) Yin, X.; Xu, Y. Structure and function of TET enzymes. *DNA methyltransferases-role and function* **2016**, 275-302.

(442) Jin, S. G.; Zhang, Z. M.; Dunwell, T. L.; Harter, M. R.; Wu, X.; Johnson, J.; Li, Z.; Liu, J.; Szabó, P. E.; Lu, Q.; et al. Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. *Cell Rep* **2016**, *14* (3), 493-505.

(443) Ko, M.; An, J.; Bandukwala, H. S.; Chavez, L.; Aijö, T.; Pastor, W. A.; Segal, M. F.; Li, H.; Koh, K. P.; Lähdesmäki, H.; et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature* **2013**, *497* (7447), 122-126.

(444) Fraietta, J. A.; Nobles, C. L.; Sammons, M. A.; Lundh, S.; Carty, S. A.; Reich, T. J.; Cogdill, A. P.; Morrissette, J. J. D.; DeNizio, J. E.; Reddy, S.; et al. Disruption of TET2 promotes the therapeutic efficacy of CD19-targeted T cells. *Nature* **2018**, *558* (7709), 307-312.

(445) Koivunen, P.; Lee, S.; Duncan, C. G.; Lopez, G.; Lu, G.; Ramkissoon, S.; Losman, J. A.; Joensuu, P.; Bergmann, U.; Gross, S. Transformation by the (R)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. *Nature* **2012**, *483* (7390), 484-488.

(446) Chen, L.-L.; Morcelle, C.; Cheng, Z.-L.; Chen, X.; Xu, Y.; Gao, Y.; Song, J.; Li, Z.; Smith, M. D.; Shi, M. Itaconate inhibits TET DNA dioxygenases to dampen inflammatory responses. *Nature Cell Biology* **2022**, *24* (3), 353-363.

(447) Hu, L.; Li, Z.; Cheng, J.; Rao, Q.; Gong, W.; Liu, M.; Shi, Y. G.; Zhu, J.; Wang, P.; Xu, Y. Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation. *Cell* **2013**, *155* (7), 1545-1555.

(448) Sudhamalla, B.; Wang, S.; Snyder, V.; Kavoosi, S.; Arora, S.; Islam, K. Complementary steric engineering at the protein–ligand interface for analogue-sensitive TET oxygenases. *Journal of the American Chemical Society* **2018**, *140* (32), 10263-10269.

(449) Hausinger, R. P. Fe (II)/α-ketoglutarate-dependent hydroxylases and related enzymes. *Critical reviews in biochemistry and molecular biology* **2004**, *39* (1), 21-68.

(450) Weirath, N. A.; Hurben, A. K.; Chao, C.; Pujari, S. S.; Cheng, T.; Liu, S.; Tretyakova, N. Y. Small Molecule Inhibitors of TET Dioxygenases: Bobcat339 Activity Is Mediated by Contaminating Copper (II). *ACS Medicinal Chemistry Letters* **2022**, *13* (5), 792-798.

(451) Baell, J.; Walters, M. A. Chemistry: Chemical con artists foil drug discovery. *Nature* **2014**, *513* (7519), 481-483.

(452) Guan, Y.; Hasipek, M.; Jiang, D.; Tiwari, A. D.; Grabowski, D. R.; Pagliuca, S.; Kongkiatkamon, S.; Patel, B.; Singh, S.; Parker, Y. Eltrombopag inhibits TET dioxygenase to contribute to hematopoietic stem cell expansion in aplastic anemia. *The Journal of clinical investigation* **2022**, *132* (4).

(453) Tiwari, A. D.; Guan, Y.; Grabowski, D. R.; Maciejewski, J. P.; Jha, B. K.; Phillips, J. G. SAR insights into TET2 catalytic domain inhibition: Synthesis of 2-Hydroxy-4-Methylene-pentanedicarboxylates. *Bioorg Med Chem* **2021**, *39*, 116141.

(454) Anastas, J. N.; Zee, B. M.; Kalin, J. H.; Kim, M.; Guo, R.; Alexandrescu, S.; Blanco, M. A.; Giera, S.; Gillespie, S. M.; Das, J. Re-programing chromatin with a bifunctional LSD1/HDAC inhibitor induces therapeutic differentiation in DIPG. *Cancer cell* **2019**, *36* (5), 528-544. e510.

(455) Smalley, J. P.; Cowley, S. M.; Hodgkinson, J. T. Bifunctional HDAC therapeutics: one drug to rule them all? *Molecules* **2020**, *25* (19), 4394.

(456) Jenke, R.; Reßing, N.; Hansen, F. K.; Aigner, A.; Büch, T. Anticancer therapy with HDAC inhibitors: mechanism-based combination strategies and future perspectives. *Cancers* **2021**, *13* (4), 634.

(457) An, J.; Rao, A.; Ko, M. TET family dioxygenases and DNA demethylation in stem cells and cancers. *Exp Mol Med* **2017**, *49* (4), e323.

(458) Luo, X.; Liu, Y.; Kubicek, S.; Myllyharju, J.; Tumber, A.; Ng, S.; Che, K. H.; Podoll, J.; Heightman, T. D.; Oppermann, U. A selective inhibitor and probe of the cellular functions of Jumonji C domain-containing histone demethylases. *Journal of the American Chemical Society* **2011**, *133* (24), 9451-9456.

(459) Arora, J. S.; Kaur, N.; Iv, O. P. Chemoselective N-acylation via condensations of N-(benzoyloxy) amines and α-ketophosphonic acids under aqueous conditions. *The Journal of Organic Chemistry* **2008**, *73* (16), 6182-6186.

(460) Jäger, S.; Rasched, G.; Kornreich-Leshem, H.; Engeser, M.; Thum, O.; Famulok, M. A versatile toolbox for variable DNA functionalization at high density. *Journal of the American Chemical Society* **2005**, *127* (43), 15071-15082.

(461) Rai, D.; Johar, M.; Manning, T.; Agrawal, B.; Kunimoto, D. Y.; Kumar, R. Design and studies of novel 5-substituted alkynylpyrimidine nucleosides as potent inhibitors of mycobacteria. *Journal of medicinal chemistry* **2005**, *48* (22), 7012-7017.

(462) Robert, H.; Andrew, K. D. Synthesis of N1-unsubstituted 5-alkynylcytosine and derivatives thereof. *Heterocycles: an international journal for reviews and communications in heterocyclic chemistry* **2006**, *68* (7), 1325-1328.

(463) Jin, B.; Li, Y.; Robertson, K. D. DNA methylation: superior or subordinate in the epigenetic hierarchy? *Genes Cancer* **2011**, *2* (6), 607-617.

(464) Melamed, P.; Yosefzon, Y.; David, C.; Tsukerman, A.; Pnueli, L. Tet enzymes, variants, and differential effects on function. *Frontiers in cell and developmental biology* **2018**, *6*, 22.

(465) Abla, H.; Sollazzo, M.; Gasparre, G.; Iommarini, L.; Porcelli, A. M. The multifaceted contribution of α-ketoglutarate to tumor progression: An opportunity to exploit? In *Seminars in cell & developmental biology*, 2020; Elsevier: Vol. 98, pp 26-33.

(466) Baksh, S. C.; Finley, L. W. Metabolic coordination of cell fate by α-ketoglutarate-dependent dioxygenases. *Trends in cell biology* **2021**, *31* (1), 24-36.

(467) Yu, B.; Edstrom, W. C.; Benach, J.; Hamuro, Y.; Weber, P. C.; Gibney, B. R.; Hunt, J. F. Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **2006**, *439* (7078), 879-884.

(468) Velázquez-Campoy, A.; Ohtaka, H.; Nezami, A.; Muzammil, S.; Freire, E. Isothermal titration calorimetry. *Current protocols in cell biology* **2004**, *23* (1), 17.18. 11-17.18. 24.

(469) Schasfoort, R. B. *Handbook of surface plasmon resonance*; Royal Society of Chemistry, 2017.

(470) Wienken, C. J.; Baaske, P.; Rothbauer, U.; Braun, D.; Duhr, S. Protein-binding assays in biological liquids using microscale thermophoresis. *Nature communications* **2010**, *1* (1), 100.

(471) Guan, Y.; Tiwari, A. D.; Hasipek, M.; Grabowski, D.; Lindner, D.; Phillips, J. G.; Mian, O.; Xu, M.; Maciejewski, J. P.; Jha, B. K. Therapeutic targeting of TET-dioxygenase deficiency in myeloid malignancies. *Blood* **2021**, *138*, 3985.

(472) Xu, Q.; Wang, C.; Zhou, J.-X.; Xu, Z.-M.; Gao, J.; Sui, P.; Walsh, C. P.; Ji, H.; Xu, G.-L. Loss of TET reprograms Wnt signaling through impaired demethylation to promote lung cancer development. *Proceedings of the National Academy of Sciences* **2022**, *119* (6), e2107599119.

(473) Alrehaili, A. A.; Gharib, A. F.; Alghamdi, S. A.; Alhazmi, A.; Al-Shehri, S. S.; Hagag, H. M.; Alsaeedi, F. A.; Alhuthali, H. M.; Raafat, N.; Etewa, R. L. Evaluation of TET Family Gene Expression and 5-Hydroxymethylcytosine as Potential Epigenetic Markers in Non-small Cell Lung Cancer. *in vivo* **2023**, *37* (1), 445-453.

(474) Wang, Z.; Du, M.; Yuan, Q.; Guo, Y.; Hutchinson, J. N.; Su, L.; Zheng, Y.; Wang, J.; Mucci, L. A.; Lin, X. Epigenomic analysis of 5-hydroxymethylcytosine (5hmC)

reveals novel DNA methylation markers for lung cancers. *Neoplasia* **2020**, *22* (3), 154-161.

(475) Yokoyama, S.; Higashi, M.; Tsutsumida, H.; Wakimoto, J.; Hamada, T.; Wiest, E.; Matsuo, K.; Kitazono, I.; Goto, Y.; Guo, X. TET1-mediated DNA hypomethylation regulates the expression of MUC4 in lung cancer. *Genes & Cancer* **2017**, *8* (3-4), 517.

(476) Forloni, M.; Gupta, R.; Nagarajan, A.; Sun, L.-S.; Dong, Y.; Pirazzoli, V.; Toki, M.; Wurtz, A.; Melnick, M. A.; Kobayashi, S. Oncogenic EGFR represses the TET1 DNA demethylase to induce silencing of tumor suppressors in cancer cells. *Cell reports* **2016**, *16* (2), 457-471.

(477) Filipczak, P. T.; Leng, S.; Tellez, C. S.; Do, K. C.; Grimes, M. J.; Thomas, C. L.; Walton-Filipczak, S. R.; Picchi, M. A.; Belinsky, S. A. p53-Suppressed Oncogene TET1 Prevents Cellular Aging in Lung CancerTET1 as a p53-Suppressed Lung Cancer Oncogene. *Cancer research* **2019**, *79* (8), 1758-1768.

(478) Meissner, A.; Gnirke, A.; Bell, G. W.; Ramsahoye, B.; Lander, E. S.; Jaenisch, R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic acids research* **2005**, *33* (18), 5868-5877.

(479) Mann, M. Functional and quantitative proteomics using SILAC. *Nature Reviews Molecular Cell Biology* **2006**, *7* (12), 952-958.

(480) Chahrour, O.; Cobice, D.; Malone, J. Stable isotope labelling methods in mass spectrometry-based quantitative proteomics. *Journal of pharmaceutical and biomedical analysis* **2015**, *113*, 2-20.

(481) Langdon, S. R.; Ertl, P.; Brown, N. Bioisosteric replacement and scaffold hopping in lead generation and optimization. *Molecular informatics* **2010**, *29* (5), 366-385.

(482) Olesen, P. H. The use of bioisosteric groups in lead optimization. *Current opinion in drug discovery & development* **2001**, *4* (4), 471-478.

(483) Mateus, A.; Määttä, T. A.; Savitski, M. M. Thermal proteome profiling: unbiased assessment of protein state through heat-induced stability changes. *Proteome science* **2016**, *15* (1), 1-7.

(484) Ha, J.; Park, H.; Park, J.; Park, S. B. Recent advances in identifying protein targets in drug discovery. *Cell Chemical Biology* **2021**, *28* (3), 394-423.

(485) Békés, M.; Langley, D. R.; Crews, C. M. PROTAC targeted protein degraders: the past is prologue. *Nat Rev Drug Discov* **2022**, *21* (3), 181-200.

(486) Li, X.; Song, Y. Proteolysis-targeting chimera (PROTAC) for targeted protein degradation and cancer therapy. *J Hematol Oncol* **2020**, *13* (1), 50.

(487) Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S. E.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat Protoc* **2015**, *10* (5), 733-755.

A. Appendix

Table A.1. Abbreviations

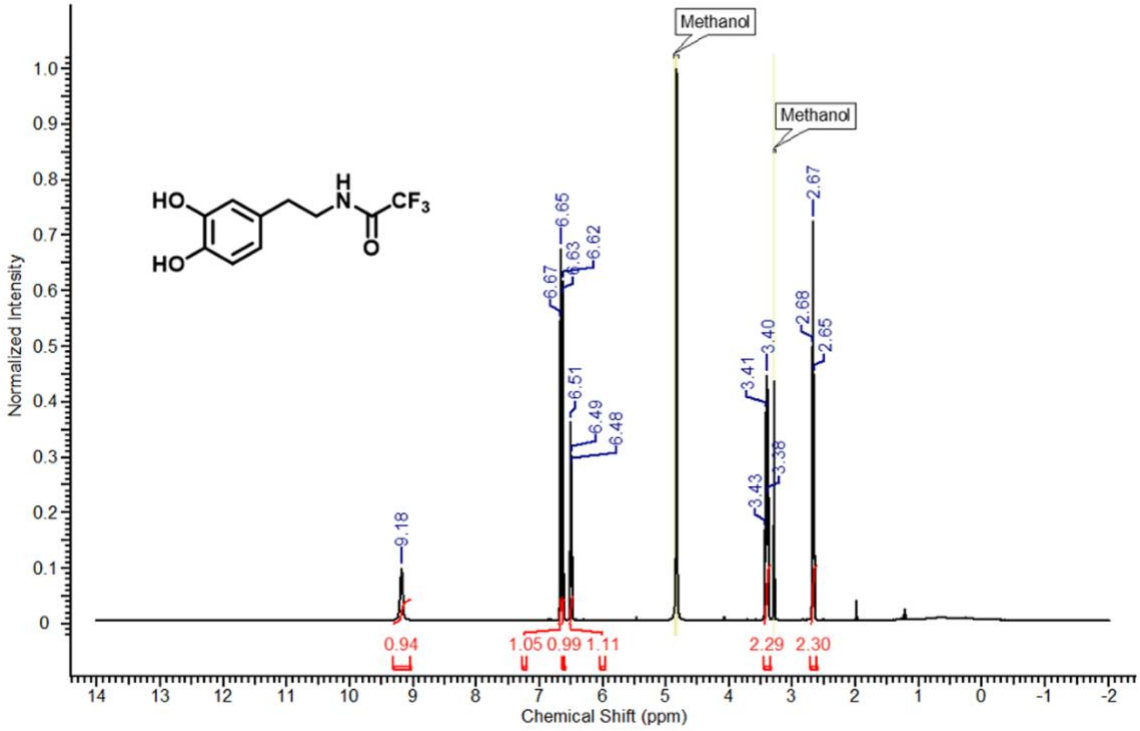
A	Adenosine or adenine
Ac_2O	Acetic anhydride
ACN	Acetonitrile
AcOH	Acetic acid
AGEs	Advanced glycation end- products
ANOVA	Analysis of variance
Arg	Arginine
BCA	Bicinchoninic acid
Boc	Tert-butoxy carbonyl
BSA	Bovine serum albumin
C	Cytidine or cytosine
CuAAC	Copper catalyzed alkyne azide cycloaddition
Cys	Cysteine
DA	Dopamine
DAC	Dopaminochrome
DAPI	4′,6-diamidino-2-phenylindole
DAQ	Dopaquinone
DCM	Dichloromethane
DHI	5,6-dihydroxyindole
DIPEA	N,N-diisopropylethylamine
DMEM	Dulbecco's modified eagle media
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DPA	Dopamine protein adduct
DPC	DNA protein crosslink
EDCI HCl	N-ethyl-N′-carbodiimide hydrochloride
EDTA	Ethylenediaminetetraacetic acid
ESI-MS	Electrospray ionization mass spectrometry
EtOAc	Ethyl acetate
EtOH	Ethanol
FA	Formic acid

FAM	Fluorescein amidite
FBS	Fetal bovine serum
Fmoc	Fluorenyl-9-methoxycarbonyl
G	Guanosine or guanine
HEK293	Human embryonic kidney cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
hmC	5-Hydroxymethylcytosine
HPLC	High performance liquid chromatography
L-DOPA	3,4dihydroxyphenylalanine
LC-MS	Liquid chromatography-mass spectrometry
LFQ	Label-free quantification
Lys	Lysine
MALDI-TOF/MS	Matrix-assisted laser desorption/ionization-time of flight mass spectrometry
mC	5-methylcytosine
MEM	Minimum essential media
MeOH	Methanol
MGO	Methylglyoxal
MW	Molecular weight
NBT	Nitroblue tetrazolium
nIRF	Near-infrared fluorescence
NMR	Nuclear magnetic resonance
PBS	Phosphate buffered saline
PD	Parkinson’s Disease
Pd/C	Palladium (0) over carbon
PMT	Post translational modification
pTSA	p-Toluenesulfonic acid
RES	Reactive electrophile species
RFU	Relative fluorescent unit
RNA	Ribonucleic acid
RP-HPLC	Reverse phase high performance liquid chromatography
RT	Room temperature
SAR	Structure activity relationship
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SEM	Standard error of the mean

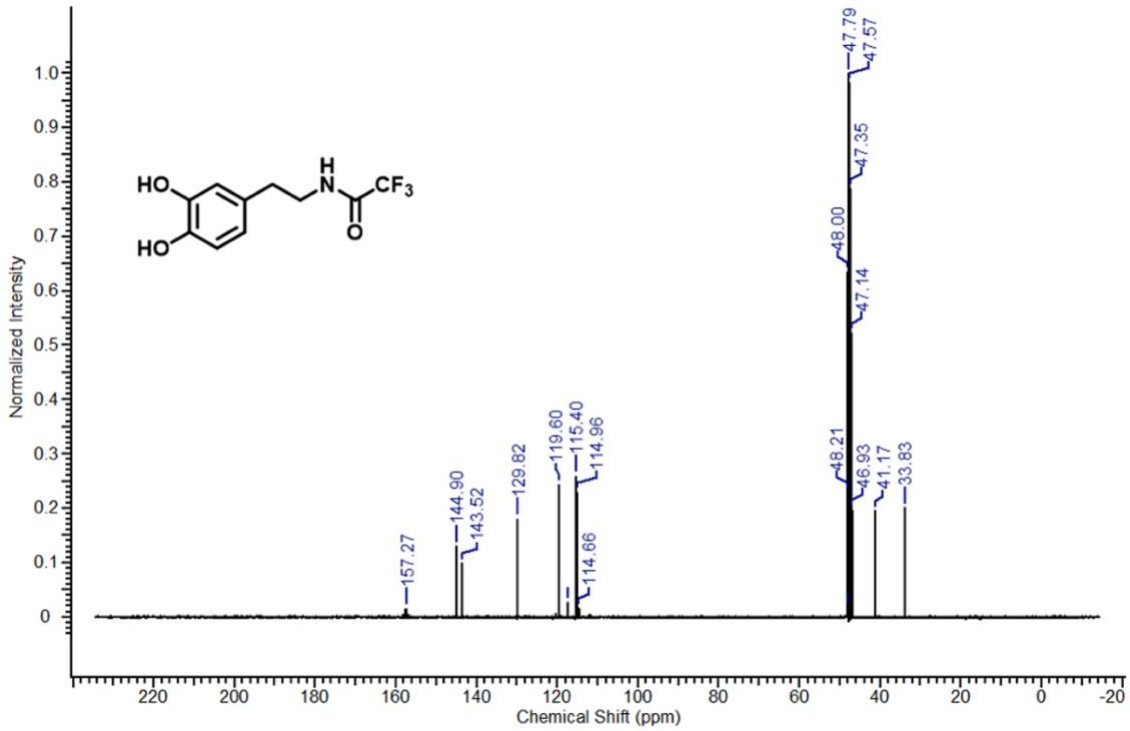
SNc	Substantia nigra pars compacta
T	Thymidine or thymine
TBAF	Tetrabutylammonium fluoride
TBAI	Tetrabutylammonium iodide
TBS	Tris-buffered saline
TE	Tris EDTA
TEA	triethylamine
TET	Ten-eleven translocase
TFA	Trifluoroacetic acid
TH	Tyrosine hydroxylase
THF	Tetrahydrofuran
THPTA	Tris-hydroxypropyltriazolylmethylamine)
TIPS	Triisopropylsilane
TLC	Thin layer chromatography
TMR	Tetramethylrhodamine
Tris	Tris(hydroxymethyl)aminomethane
Tyr	Tyrosine
U	Uridine or uracil
UV	Ultraviolet
WT	Wildtype

A.1 NMR Characterization of Synthesized Molecules

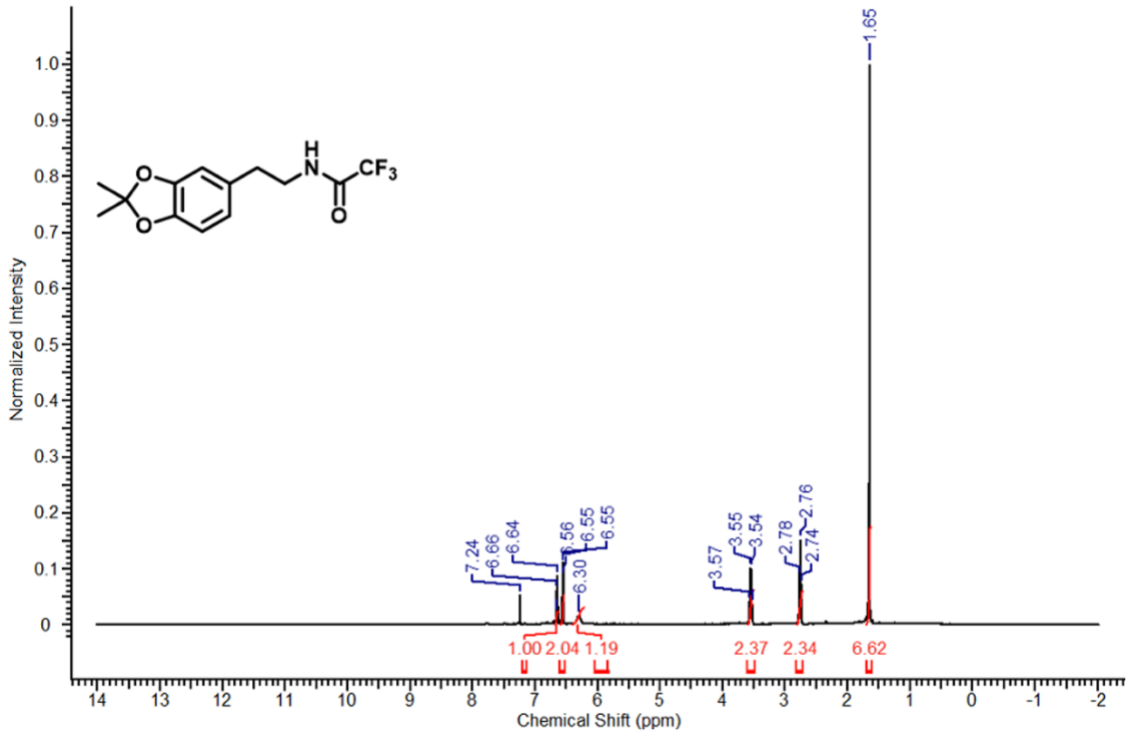
N-(3,4-dihydroxyphenethyl)-2,2,2-trifluoroacetamide



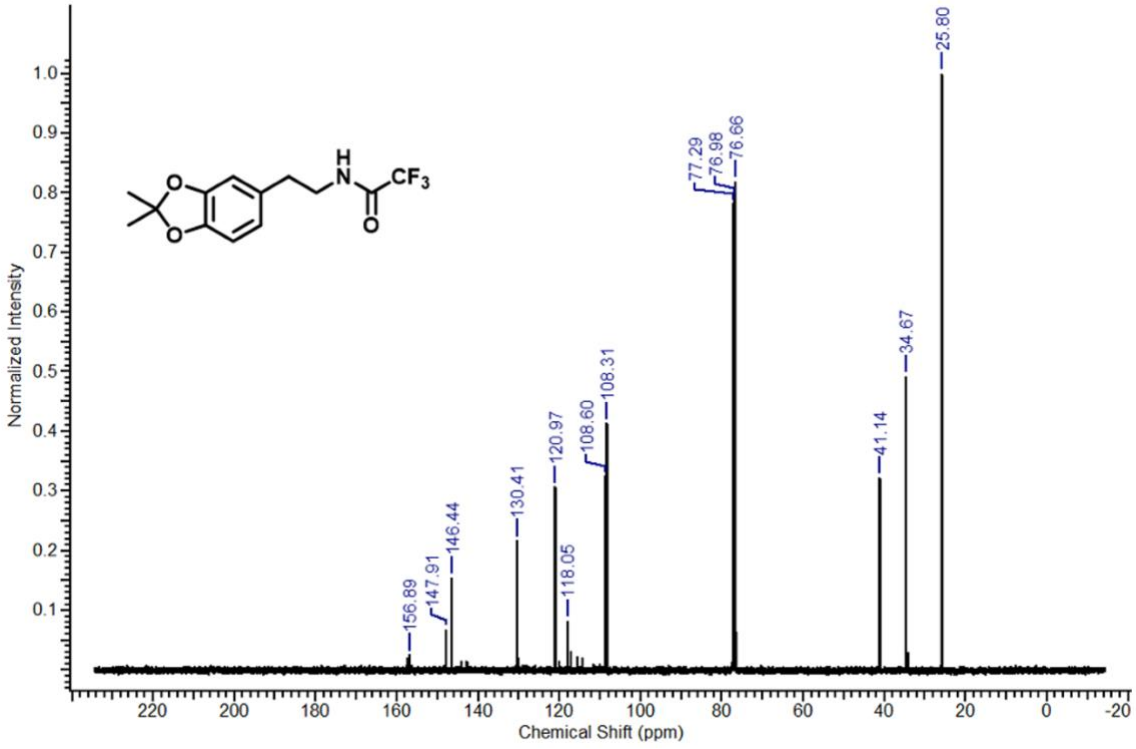
N-(3,4-dihydroxyphenethyl)-2,2,2-trifluoroacetamide



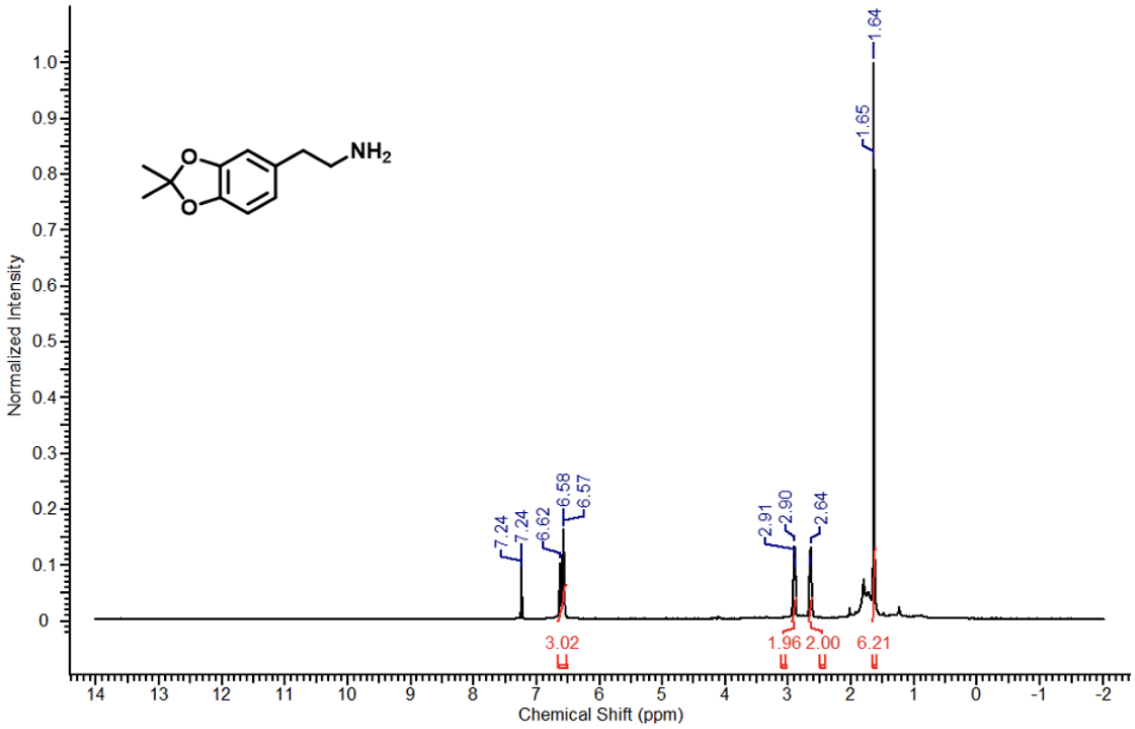
N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-2,2,2-trifluoroacetamide



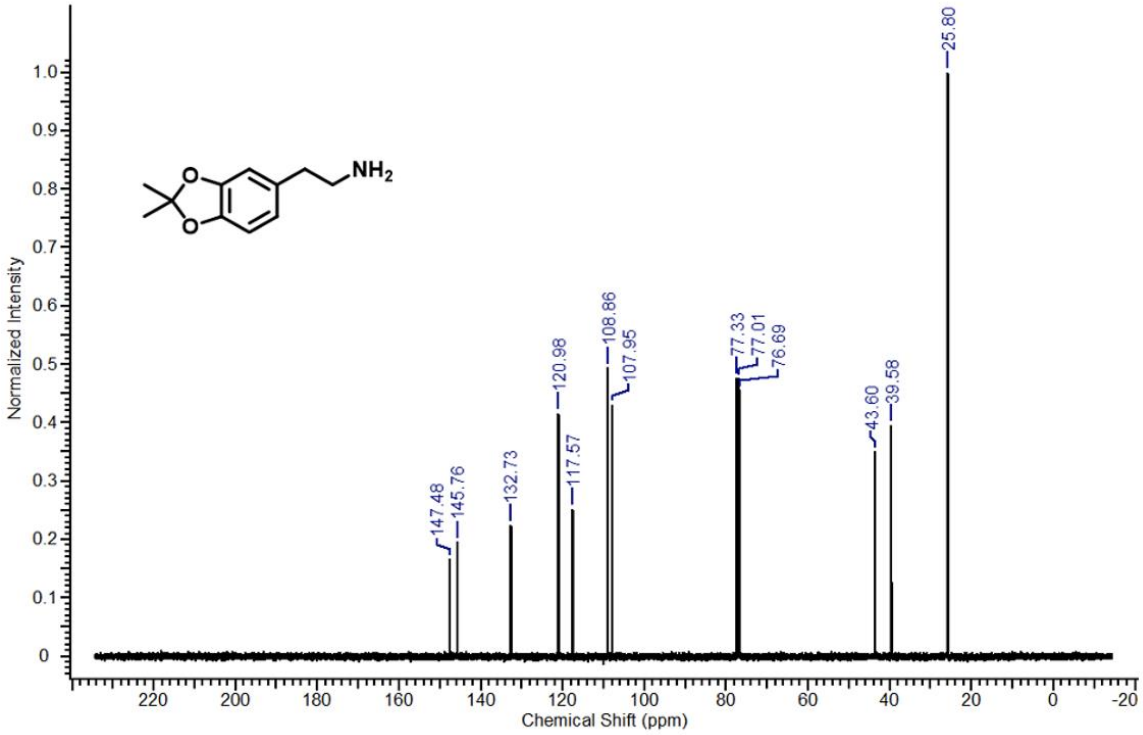
N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-2,2,2-trifluoroacetamide



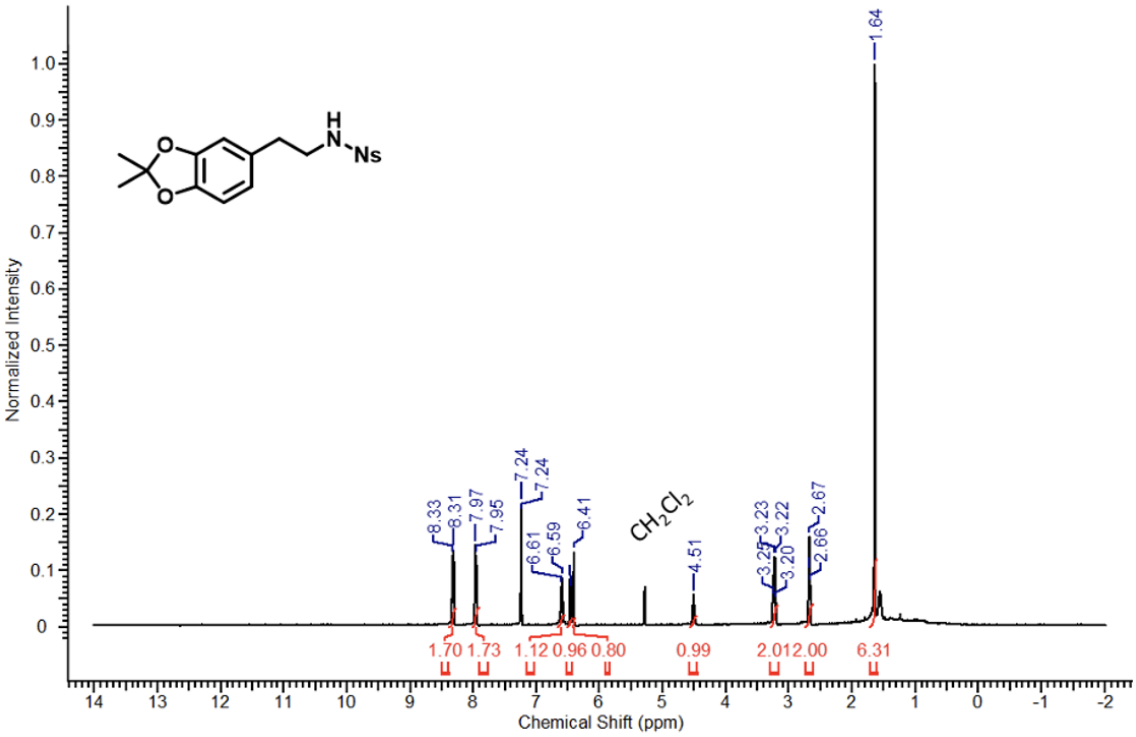
2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethan-1-amine



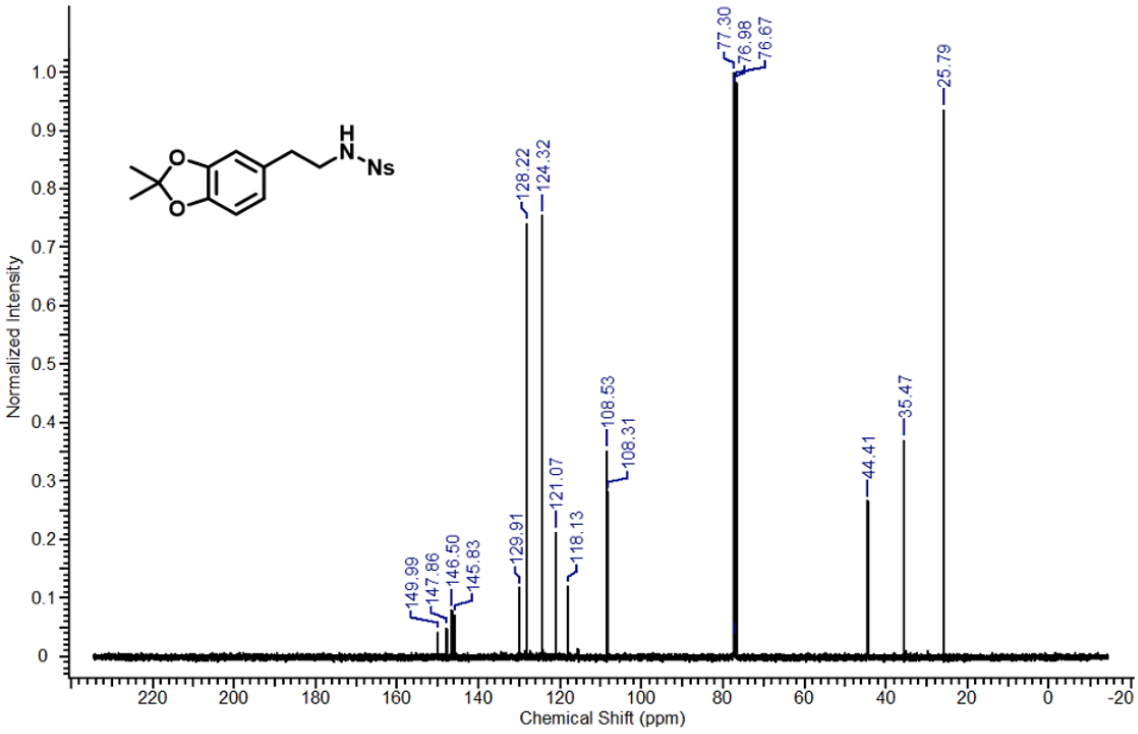
2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethan-1-amine



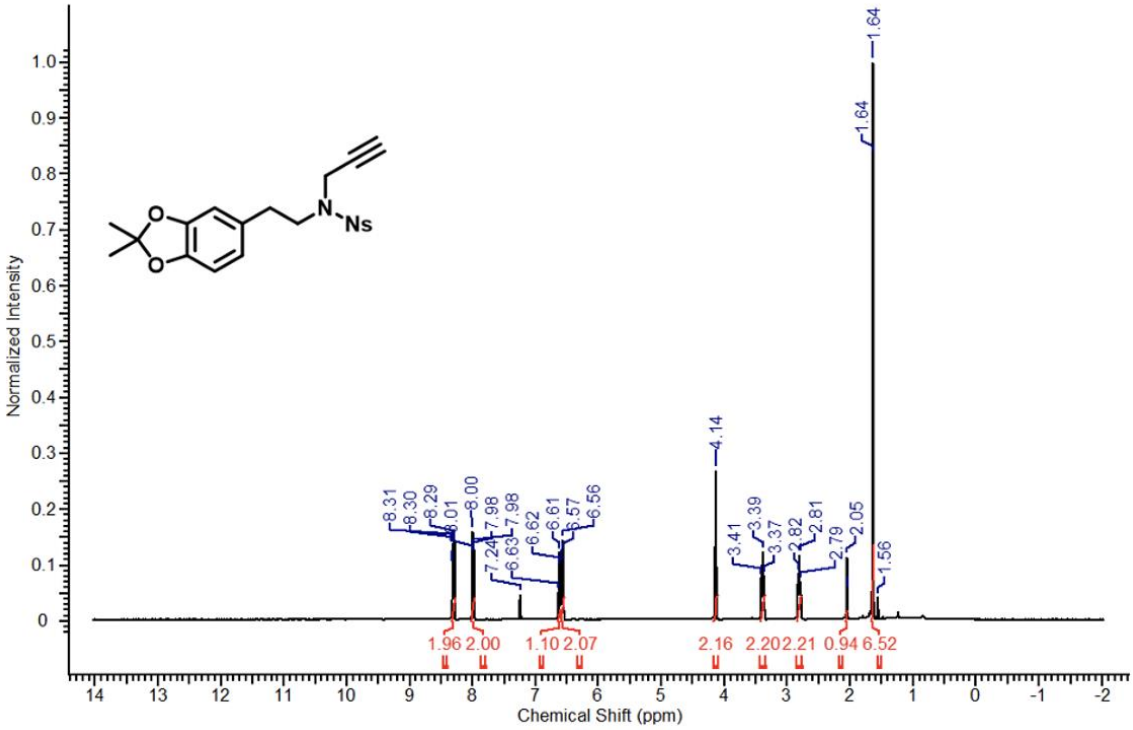
N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-4-nitrobenzenesulfonamide



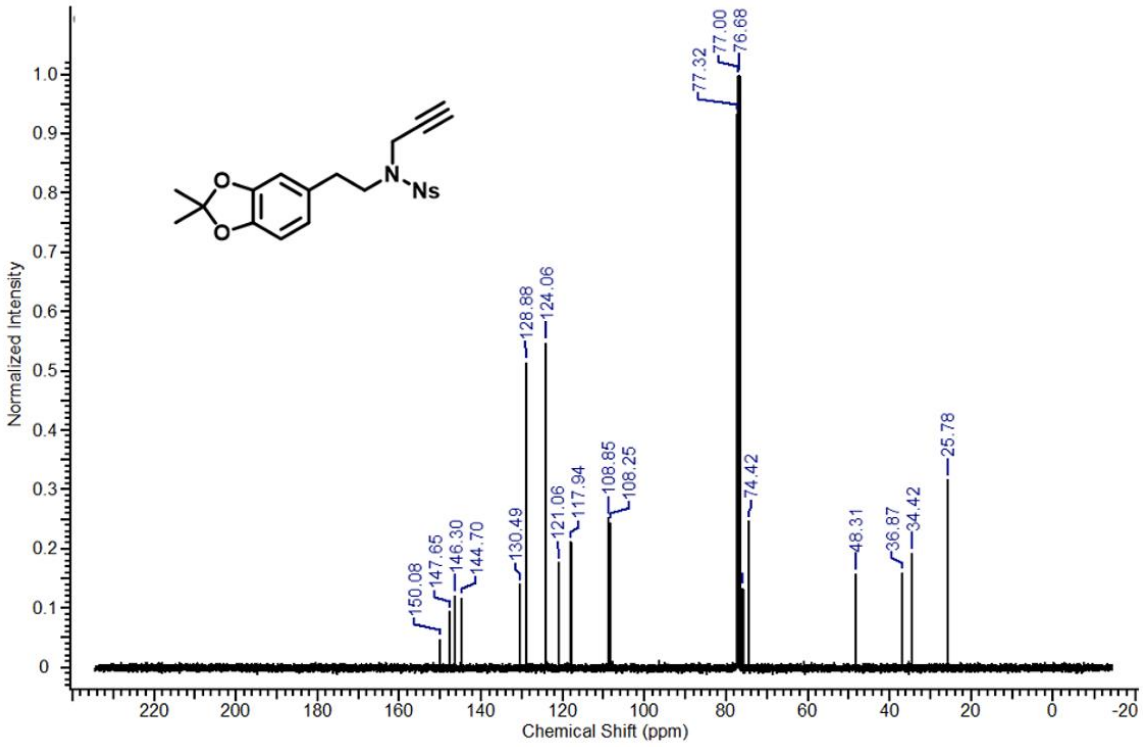
N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-4-nitrobenzenesulfonamide



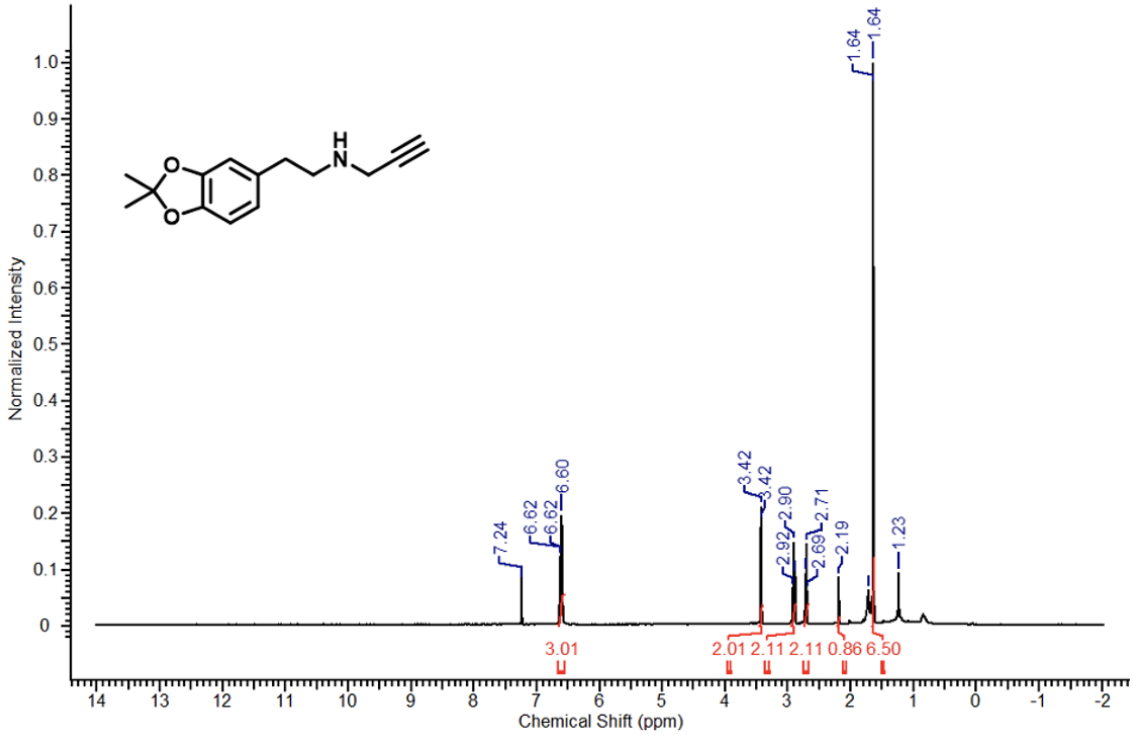
N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-4-nitro-*N*-(prop-2-yn-1-yl)benzenesulfonamide



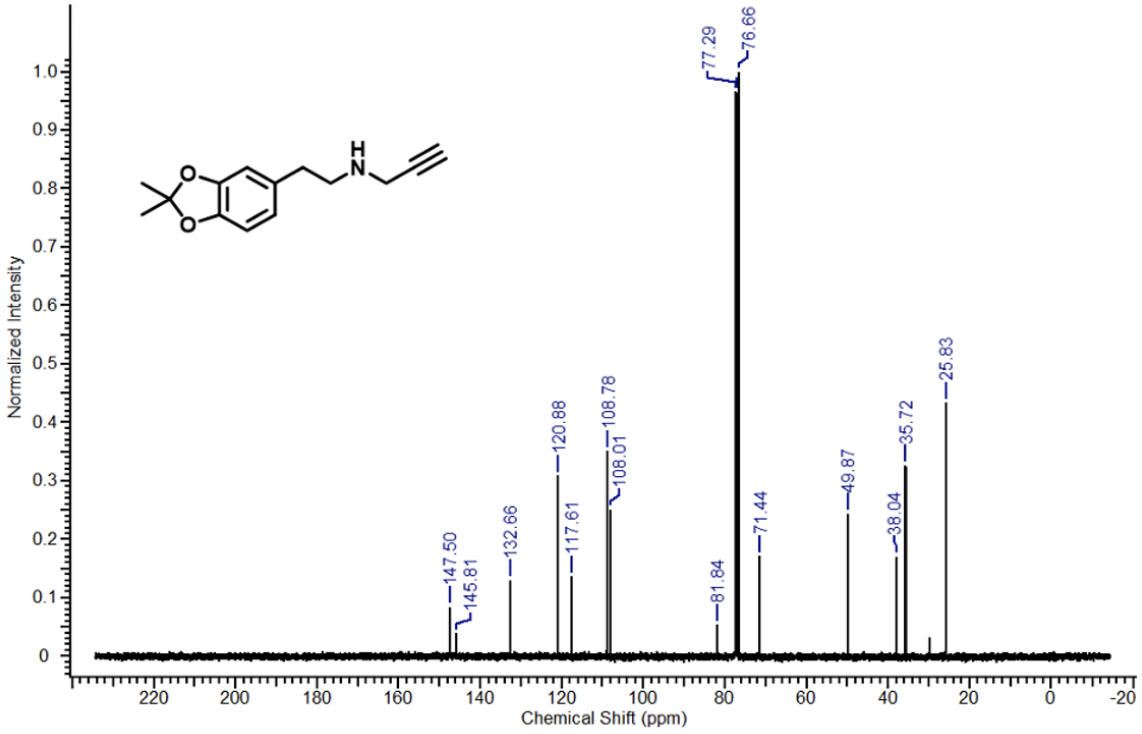
N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)-4-nitro-*N*-(prop-2-yn-1-yl)benzenesulfonamide



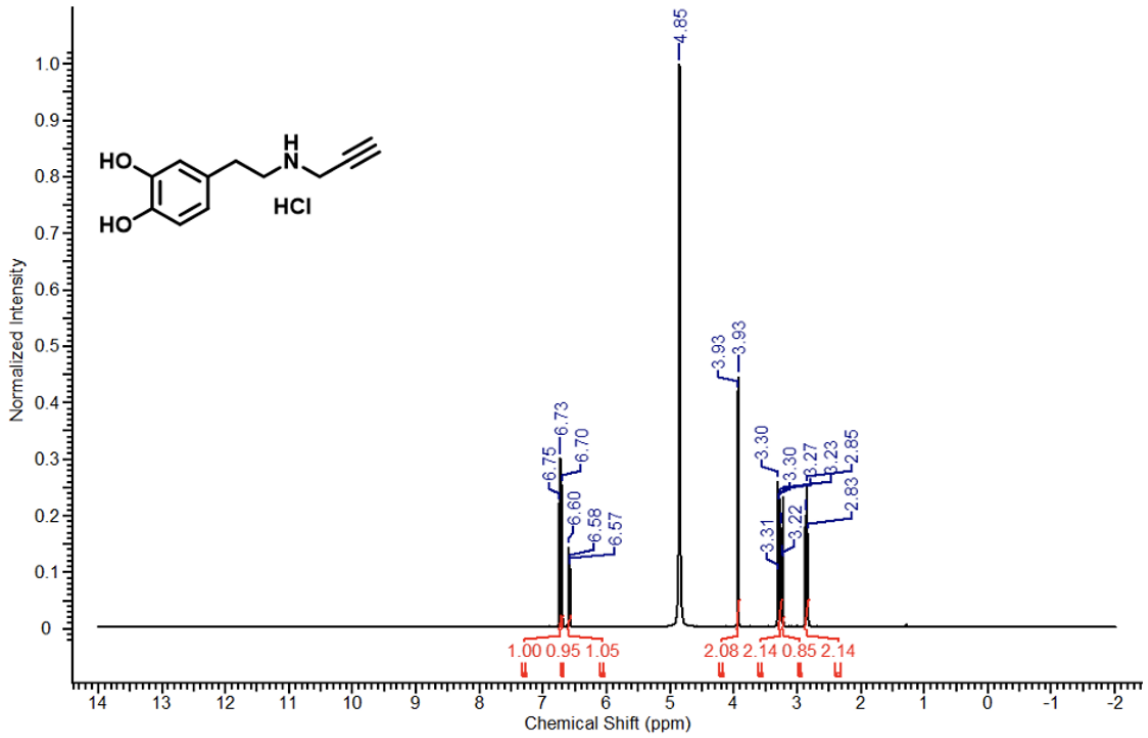
N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)prop-2-yn-1-amine



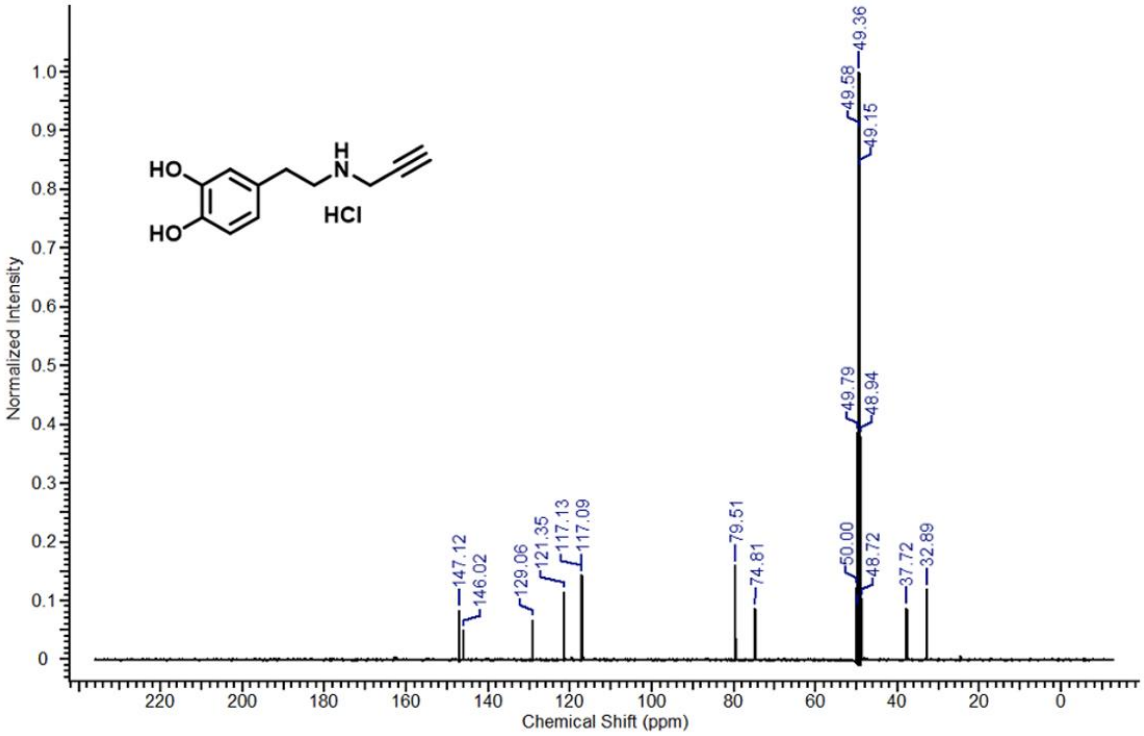
N-(2-(2,2-dimethylbenzo[1,3]dioxol-5-yl)ethyl)prop-2-yn-1-amine



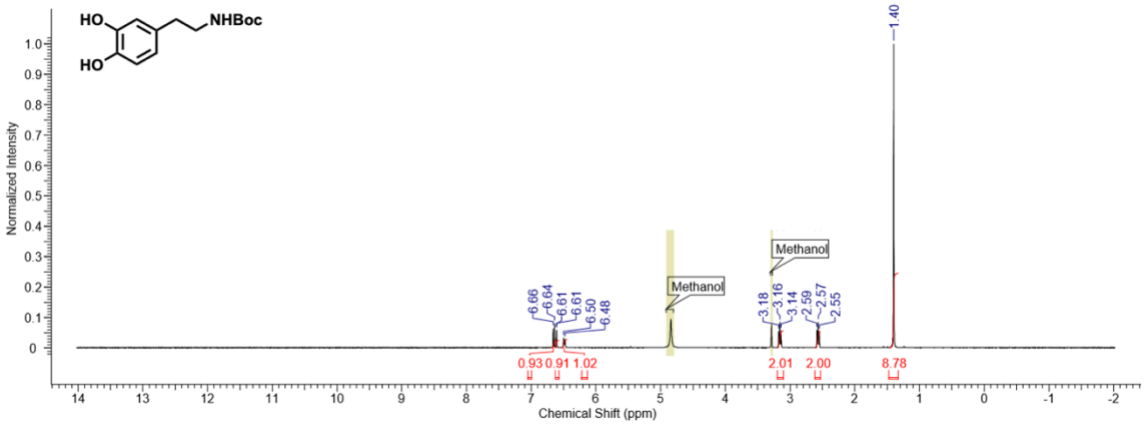
4-(2-(prop-2-yn-1-ylamino)ethyl)benzene-1,2-diol



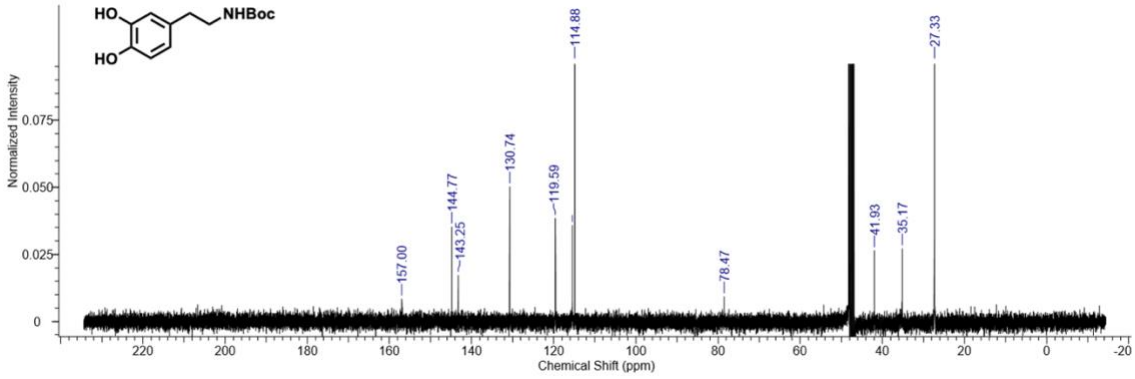
4-(2-(prop-2-yn-1-ylamino)ethyl)benzene-1,2-diol



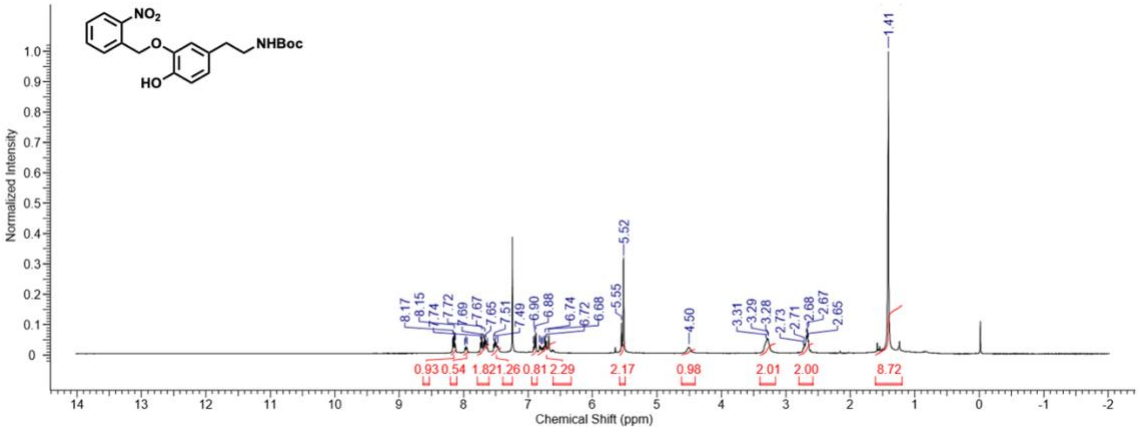
tert-butyl (3,4-dihydroxyphenethyl)carbamate



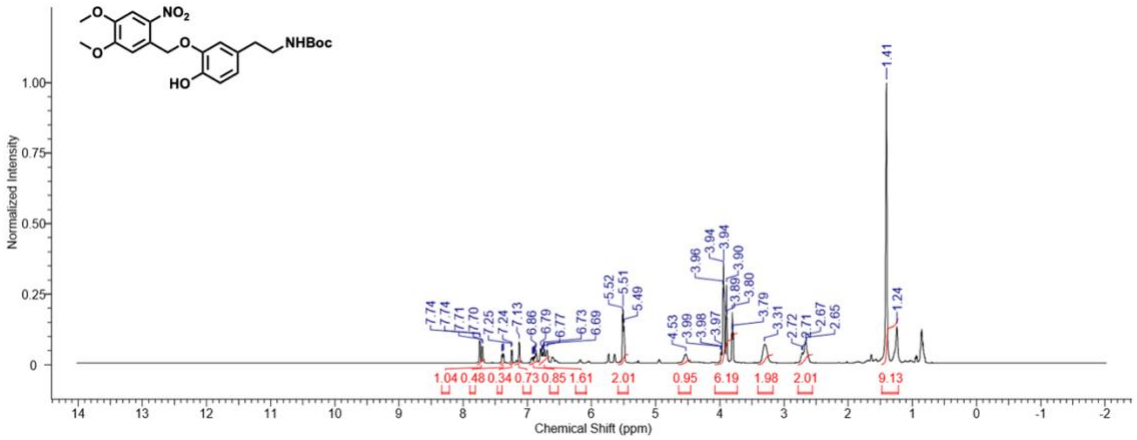
tert-butyl (3,4-dihydroxyphenethyl)carbamate



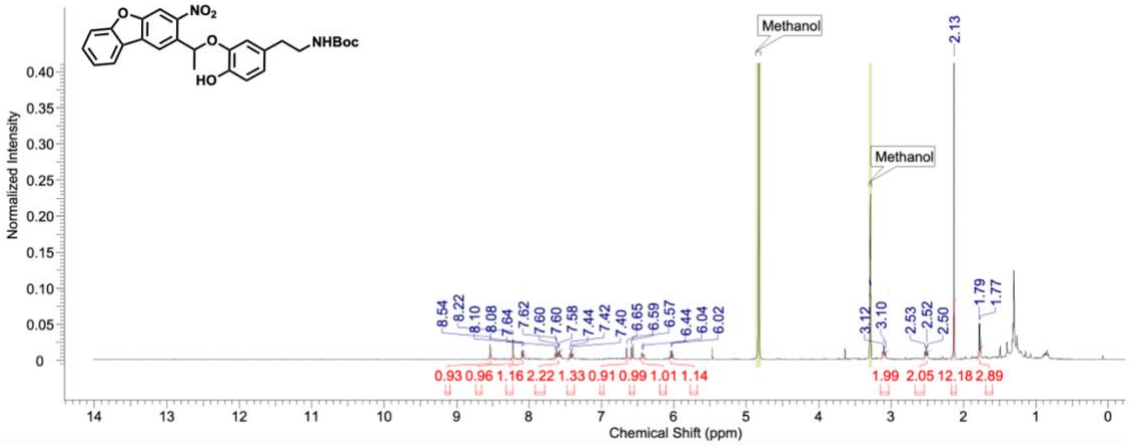
tert-butyl (4-hydroxy-3-((2-nitrobenzyl)oxy)phenethyl)carbamate



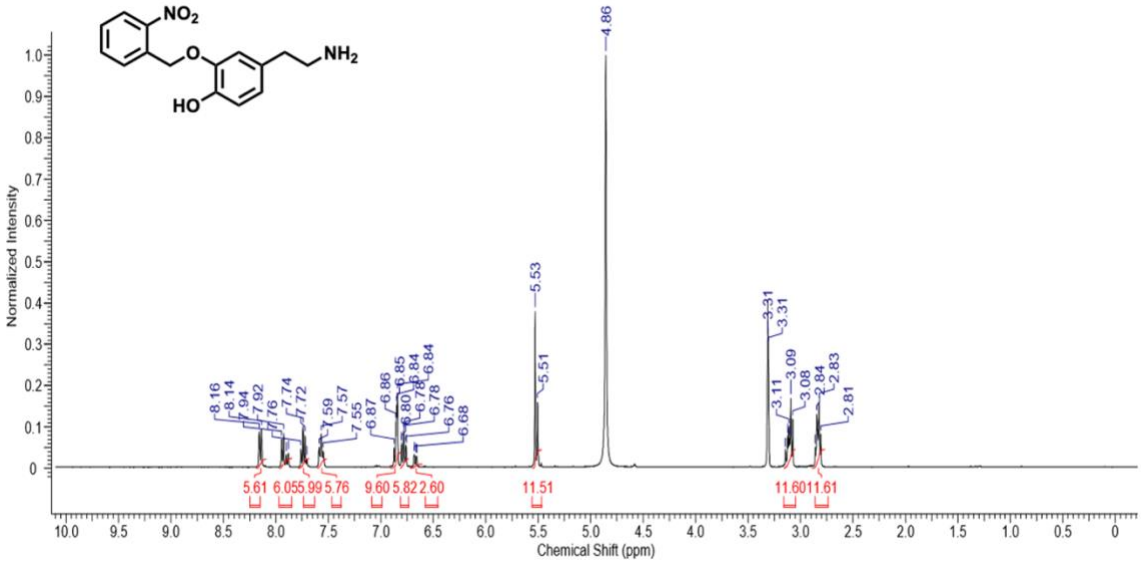
tert-butyl (3-((4,5-dimethoxy-2-nitrobenzyl)oxy)-4-hydroxyphenethyl)carbamate



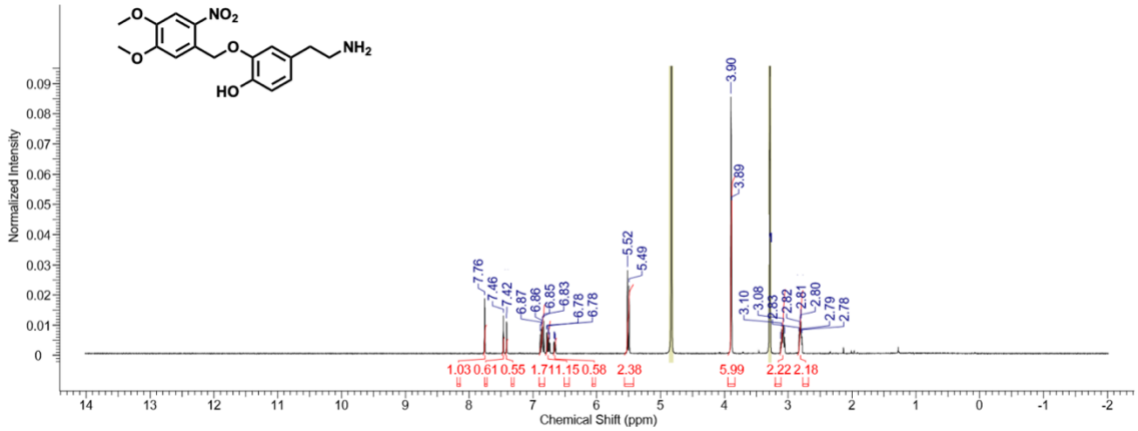
tert-butyl (4-hydroxy-3-(1-(3-nitrodibenzo[*b*,*d*]furan-2-yl)ethoxy)phenethyl)carbamate



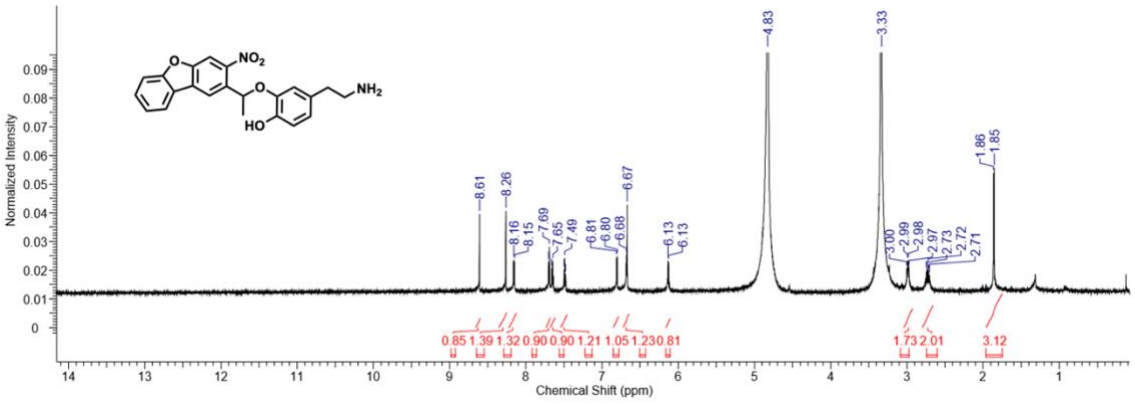
4-(2-aminoethyl)-2-((2-nitrobenzyl)oxy)phenol



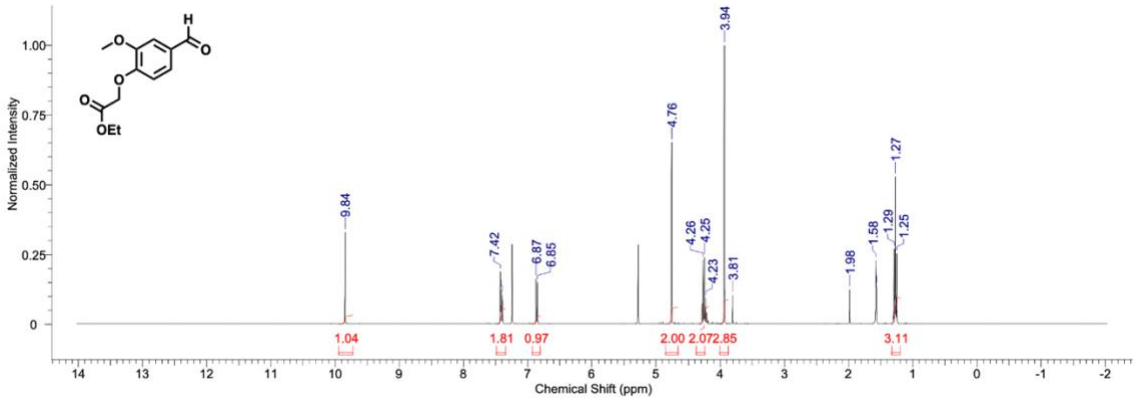
4-(2-aminoethyl)-2-((4,5-dimethoxy-2-nitrobenzyl)oxy)phenol



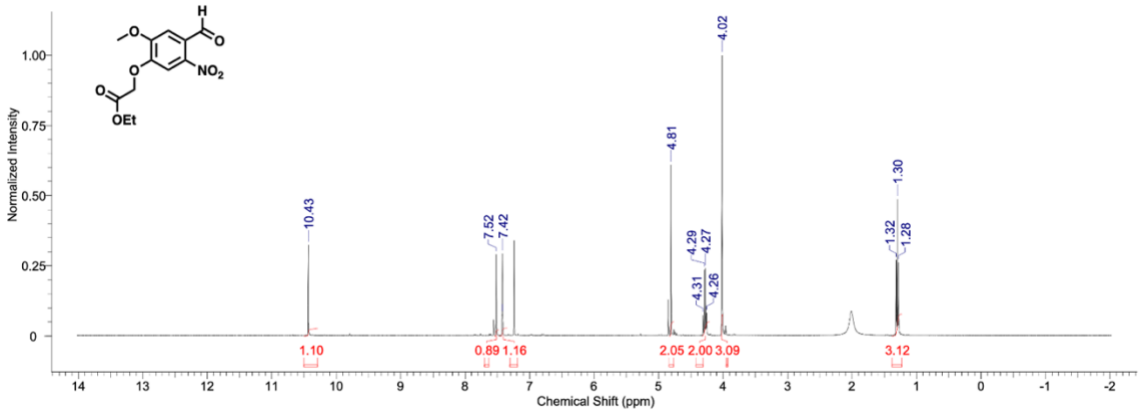
4-(2-aminoethyl)-2-(1-(3-nitrodibenzo[*b*,*d*]furan-2-yl)ethoxy)phenol



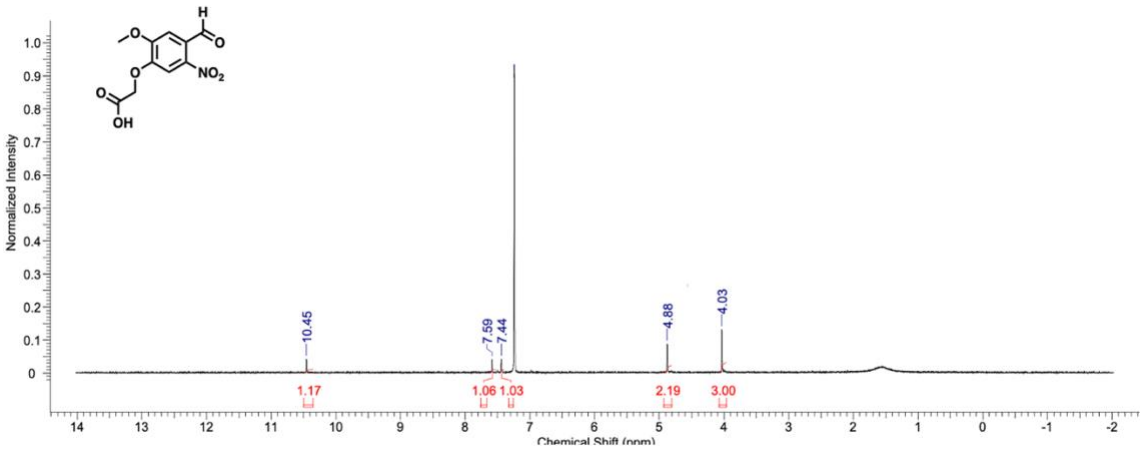
ethyl 2-(4-formyl-2-methoxyphenoxy)acetate



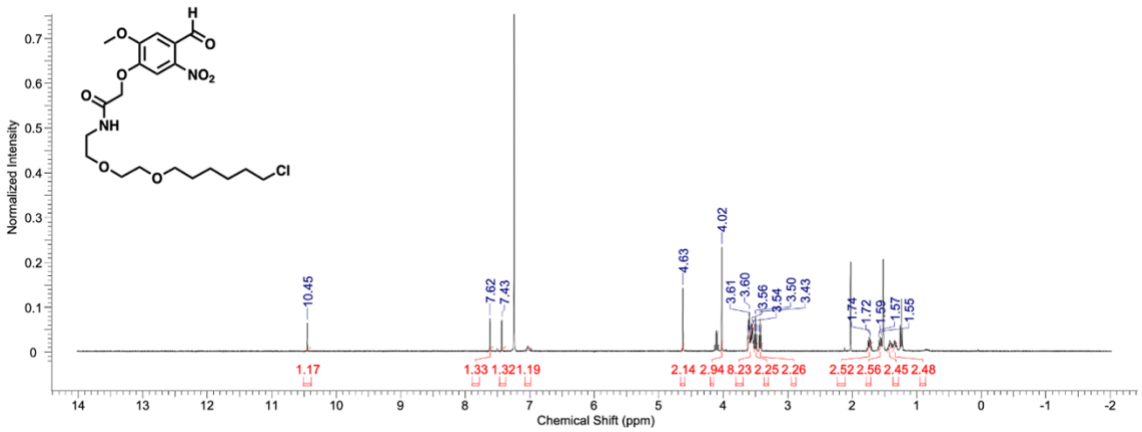
ethyl 2-(4-formyl-2-methoxy-5-nitrophenoxy)acetate



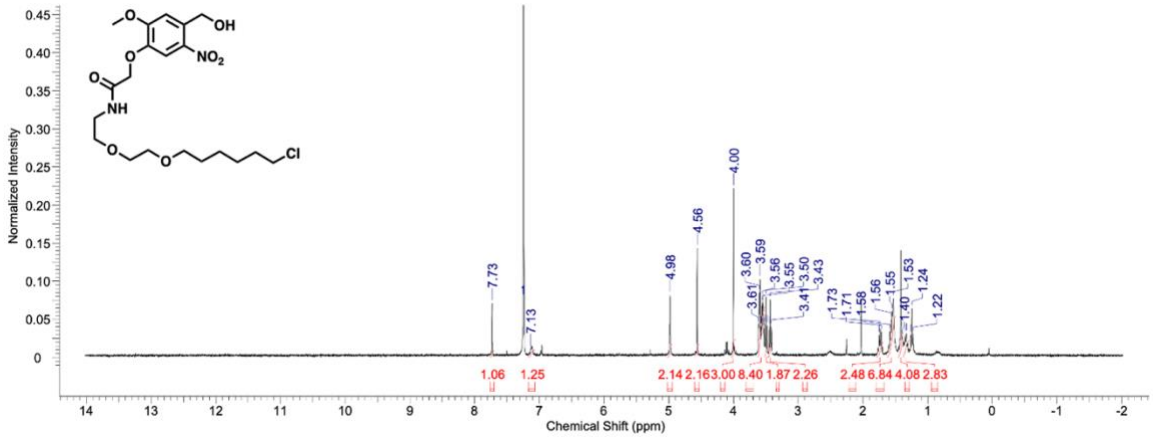
2-(4-formyl-2-methoxy-5-nitrophenoxy)acetic acid



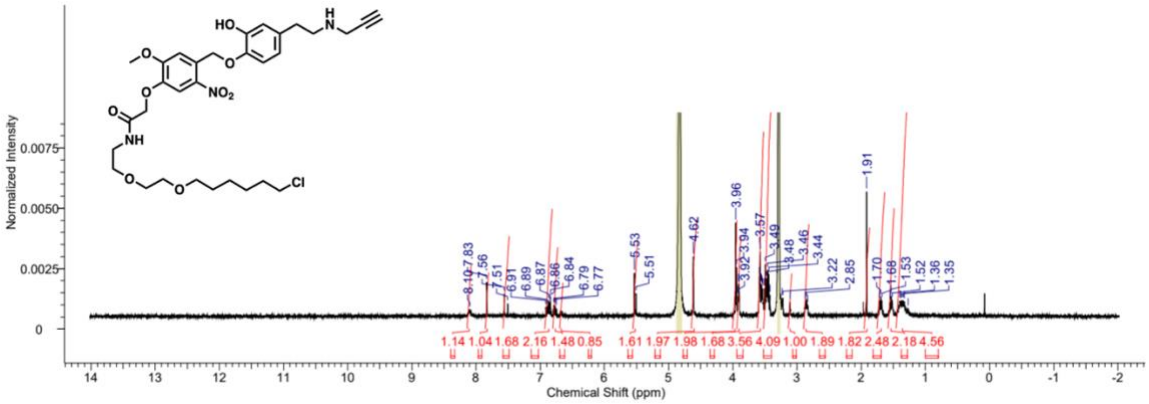
N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-(4-formyl-2-methoxy-5-nitrophenoxy)acetamide



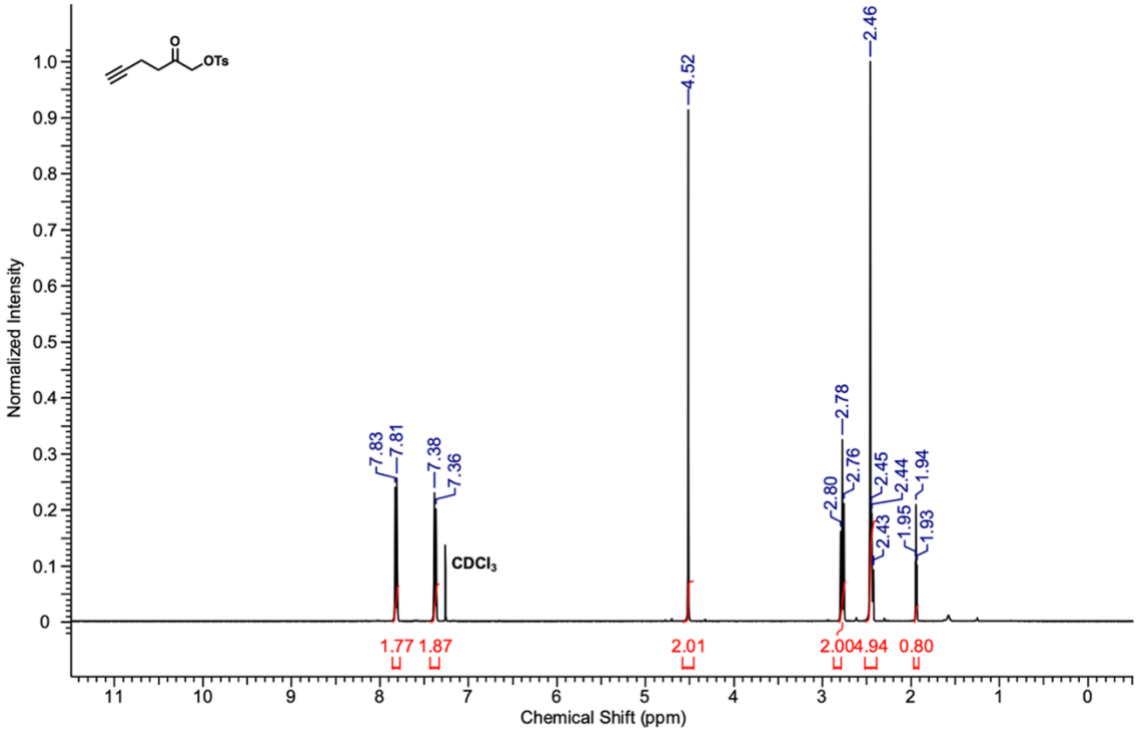
N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-(4-(hydroxymethyl)-2-methoxy-5-nitrophenoxy)acetamide



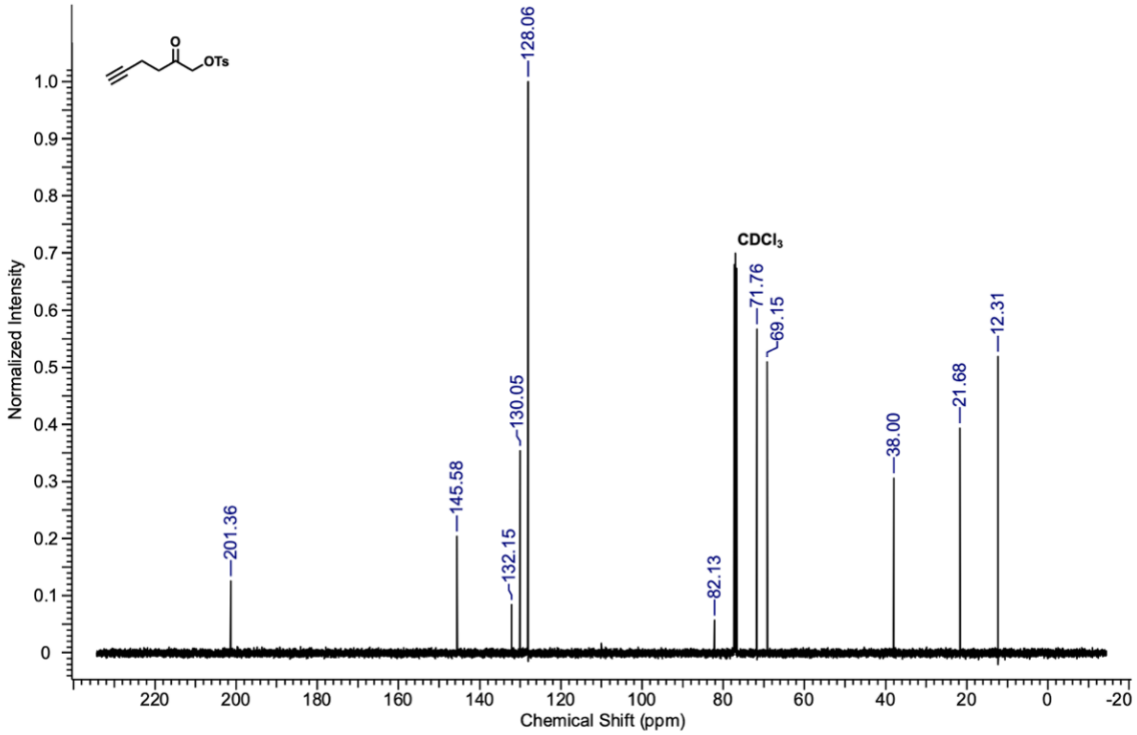
N-(2-(2-((6-chlorohexyl)oxy)ethoxy)ethyl)-2-(4-((2-hydroxy-4-(2-(prop-2-yn-1-ylamino)ethyl)phenoxy)methyl)-2-methoxy-5-nitrophenoxy)acetamide



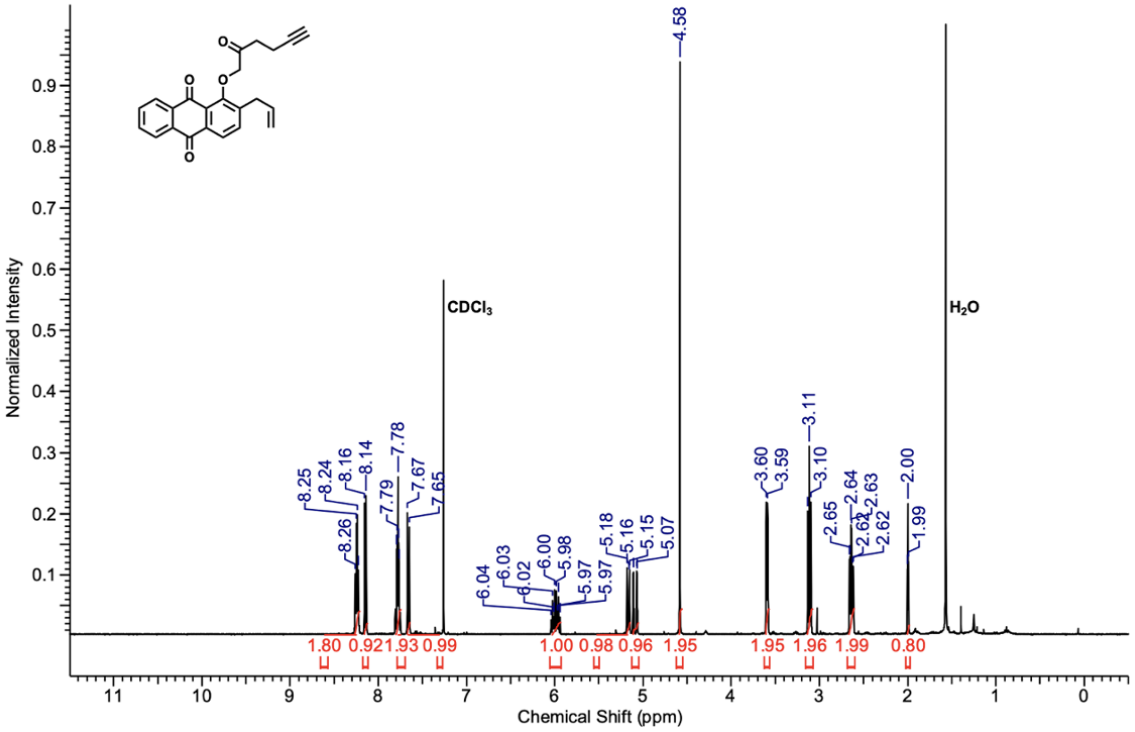
2-Oxohex-5-yn-1-yl 4-methylbenzenesulfonate



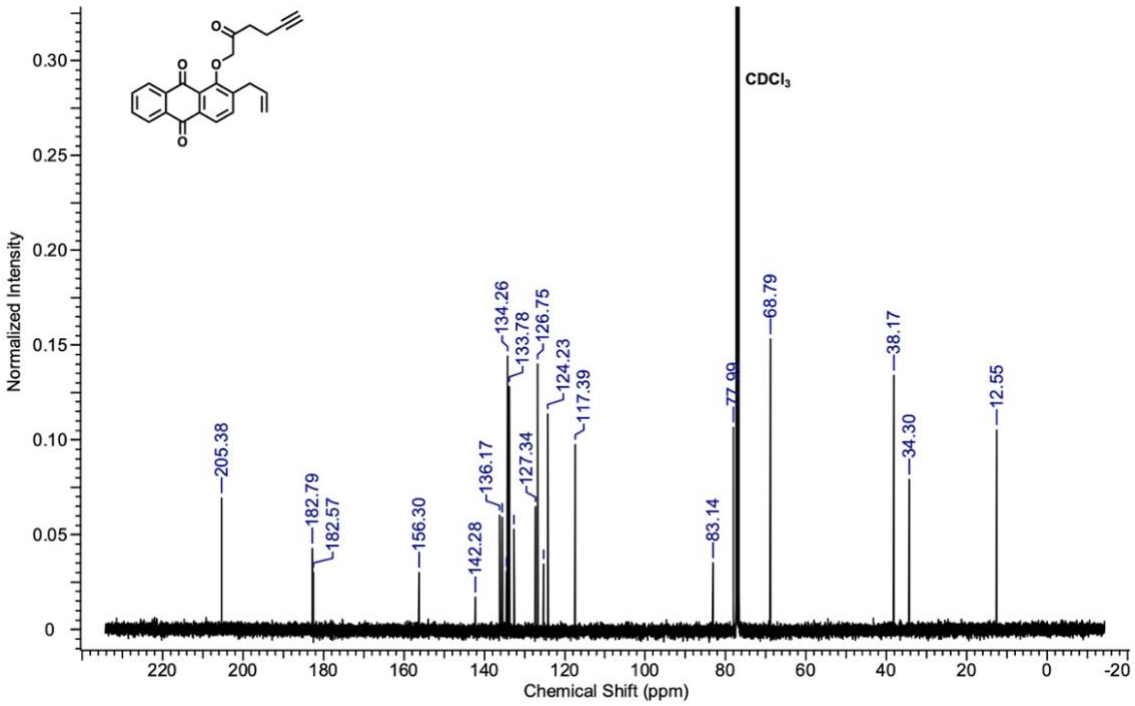
2-Oxohex-5-yn-1-yl 4-methylbenzenesulfonate



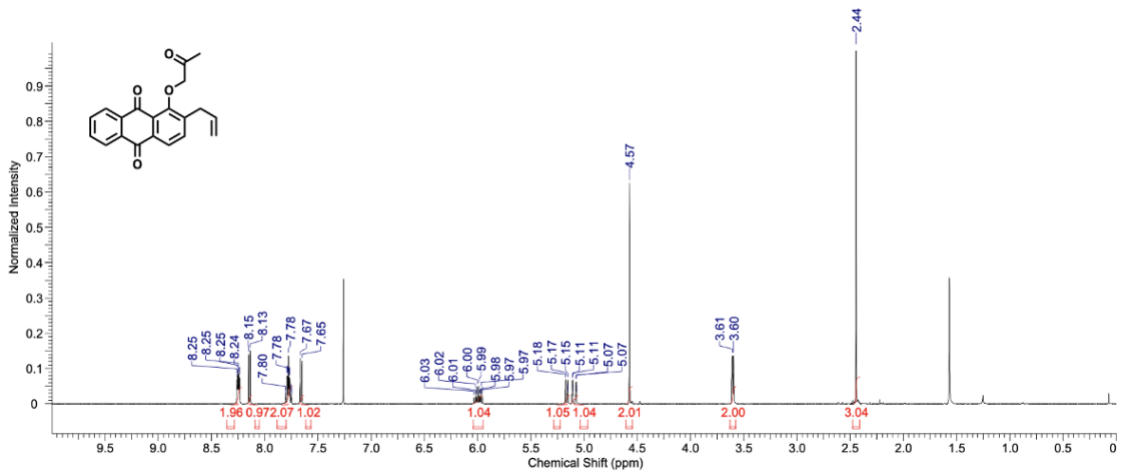
2-Allyl-1-[(2-oxohex-5-yn-1-yl)oxy]anthracene-9,10-dione



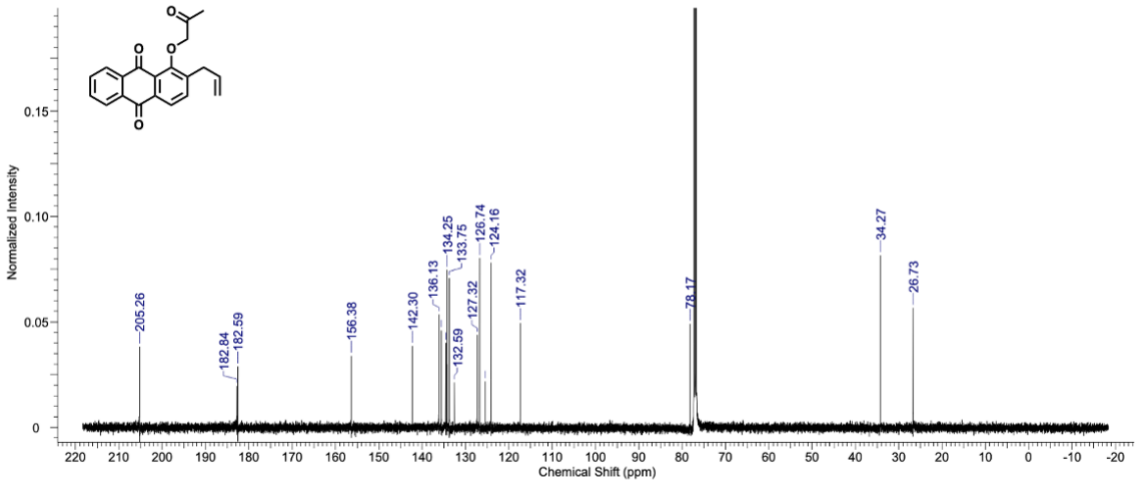
2-Allyl-1-[(2-oxohex-5-yn-1-yl)oxy]anthracene-9,10-dione



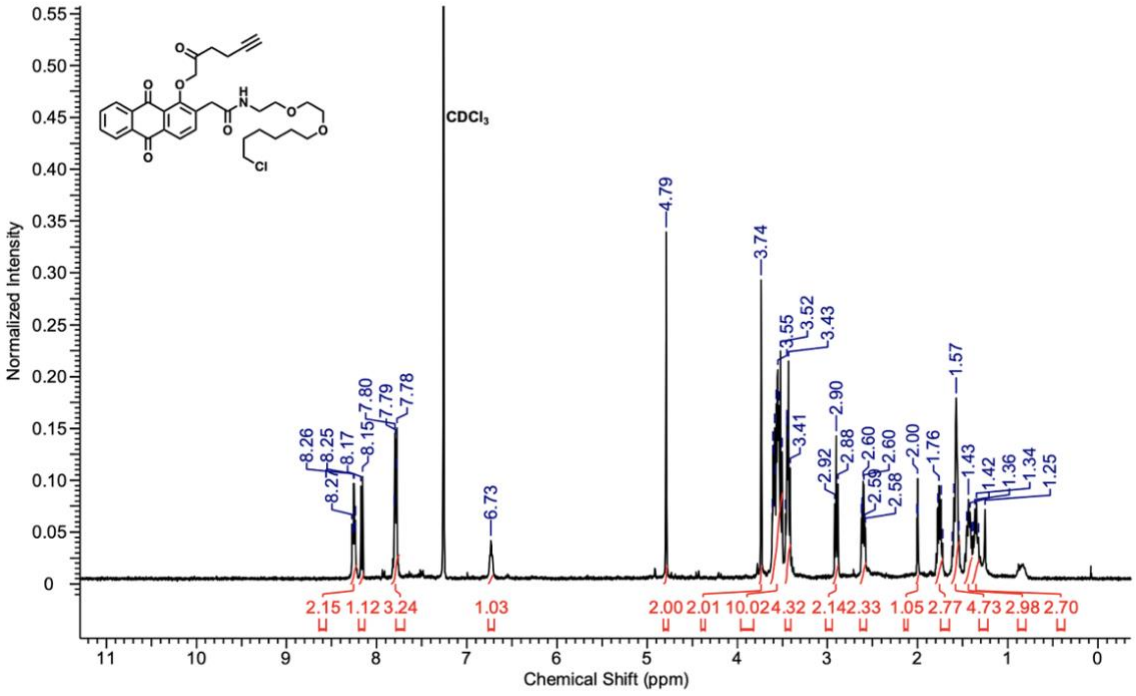
2-allyl-1-(2-oxopropoxy)anthracene-9,10-dione



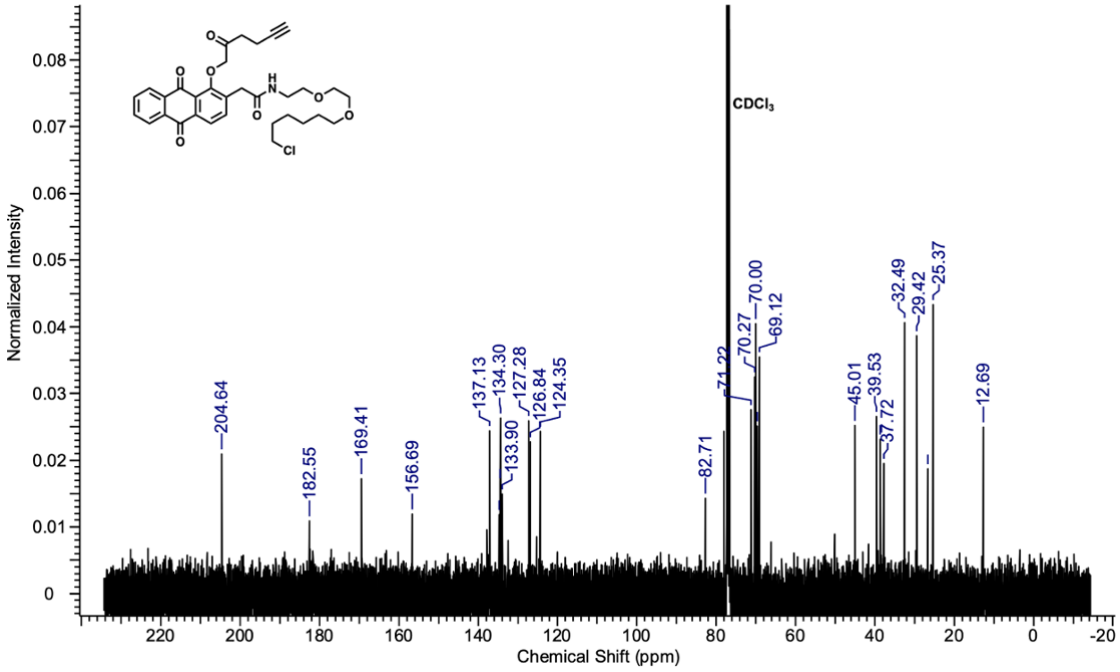
2-allyl-1-(2-oxopropoxy)anthracene-9,10-dione



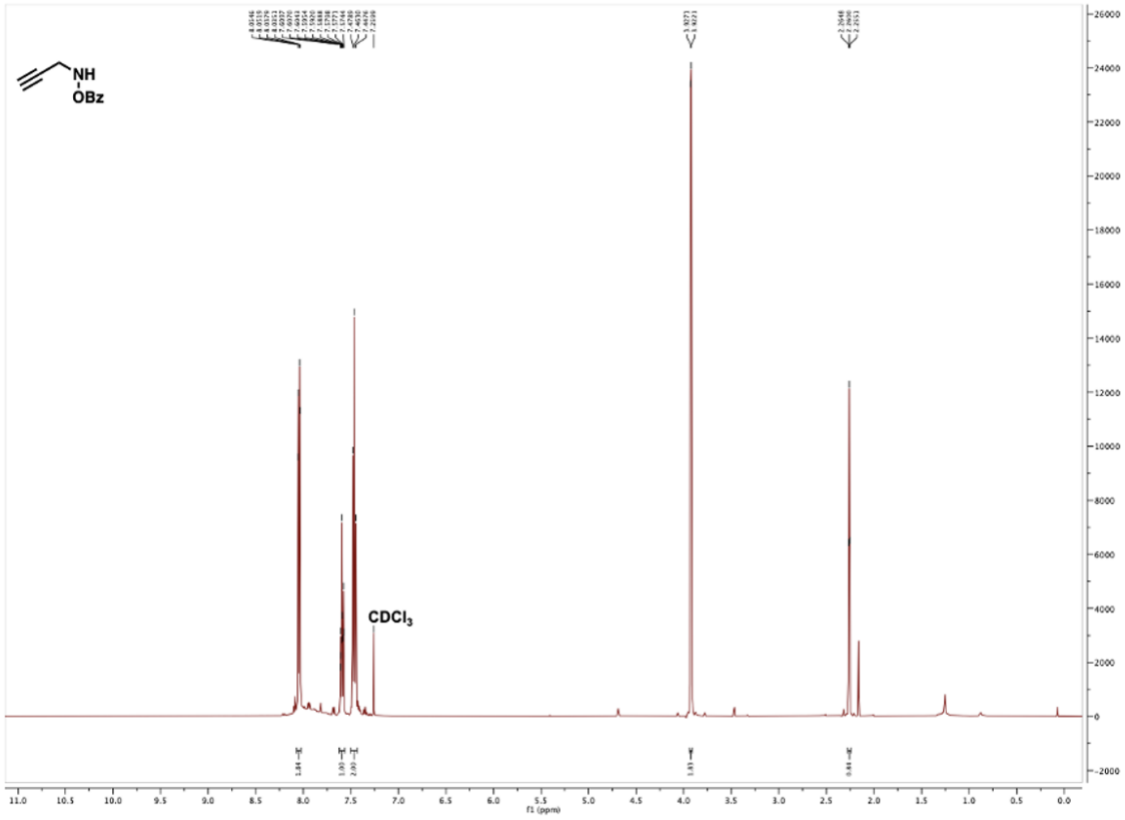
N-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethyl)-2-(9,10-dioxo-1-((2-oxohex-5-yn-1-yl)oxy)-9,10-dihydroanthracen-2-yl)acetamide



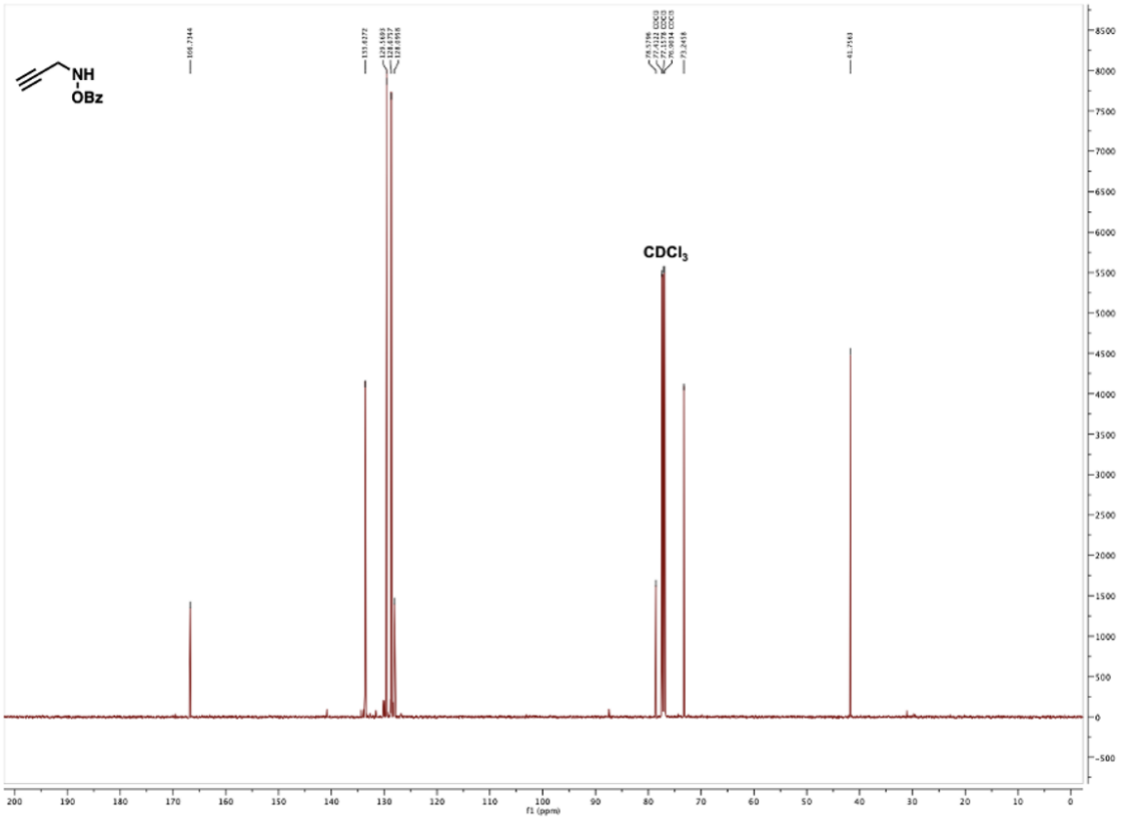
N-(2-(2-((6-Chlorohexyl)oxy)ethoxy)ethyl)-2-(9,10-dioxo-1-((2-oxohex-5-yn-1-yl)oxy)-9,10-dihydroanthracen-2-yl)acetamide



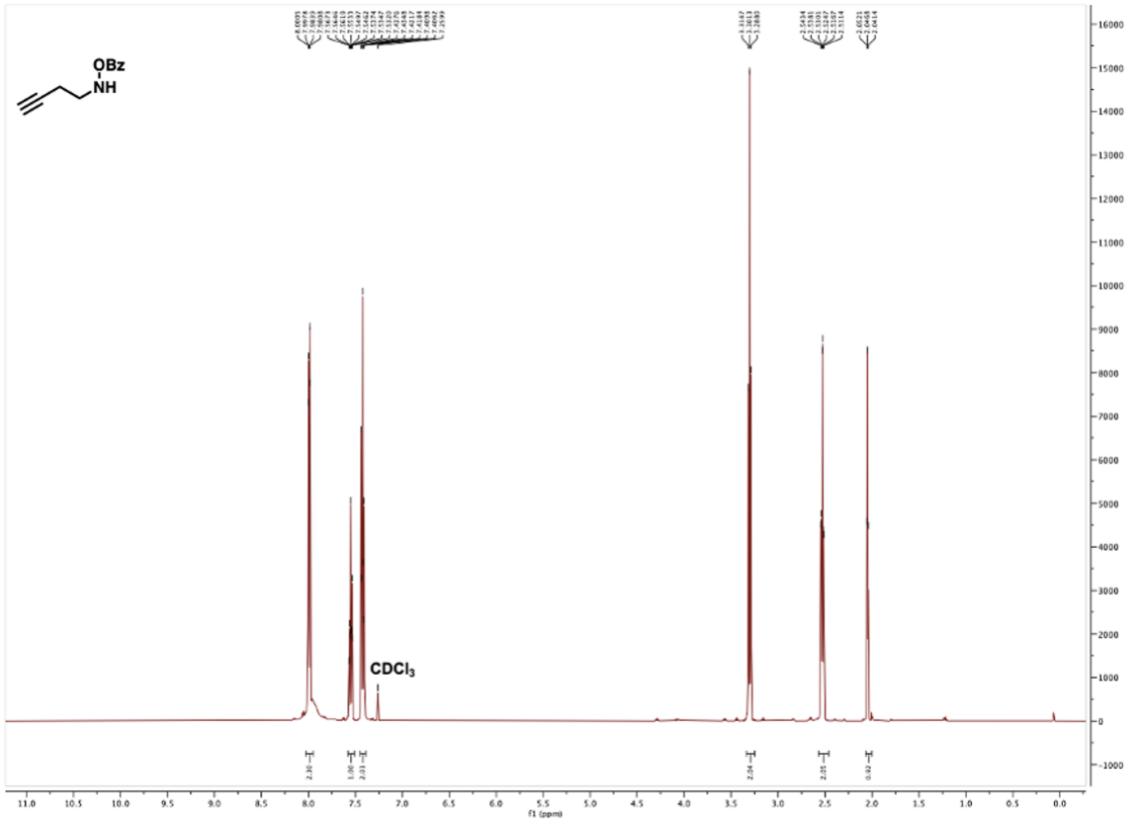
O-benzoyl-N-(prop-2-yn-1-yl)hydroxylamine



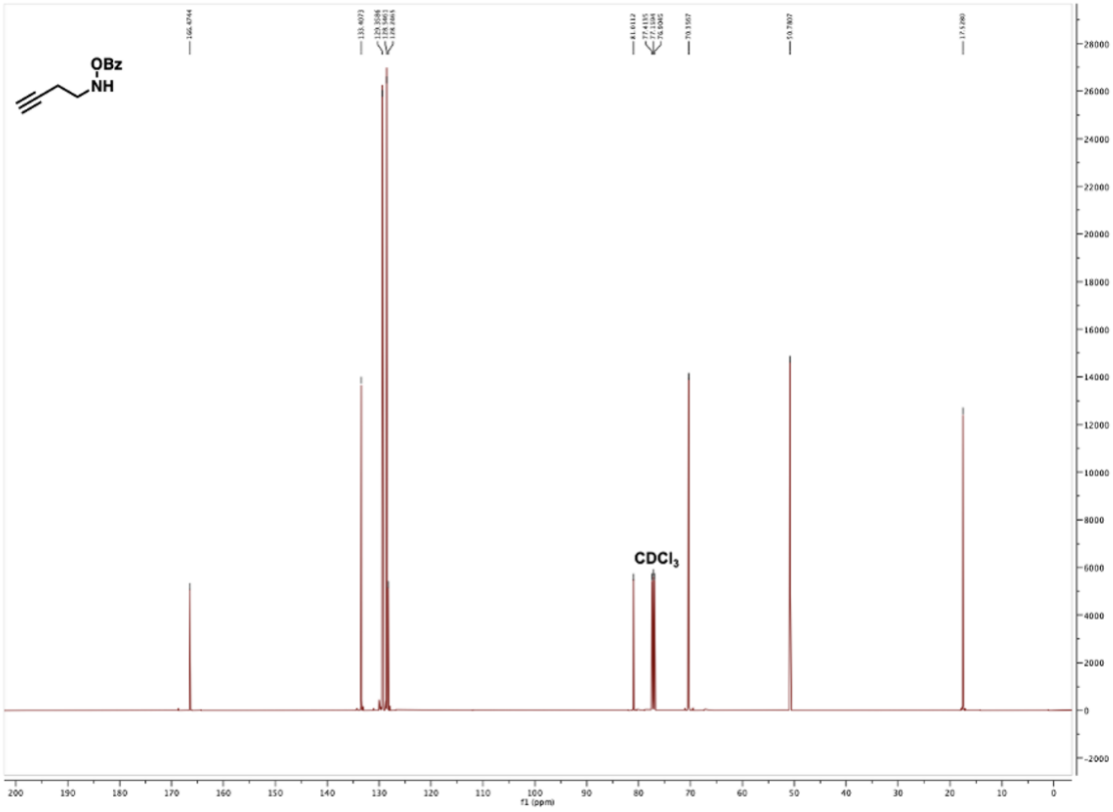
O-benzoyl-N-(prop-2-yn-1-yl)hydroxylamine



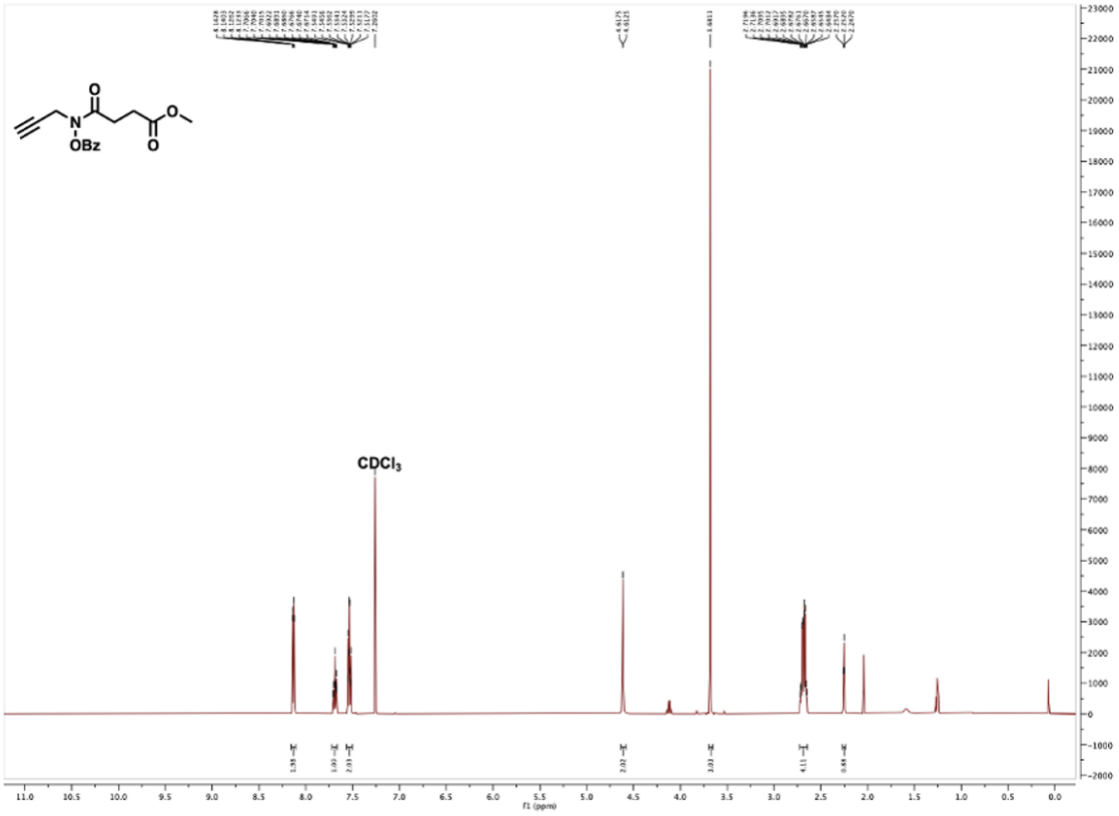
O-benzoyl-N-(but-3-yn-1-yl)hydroxylamine



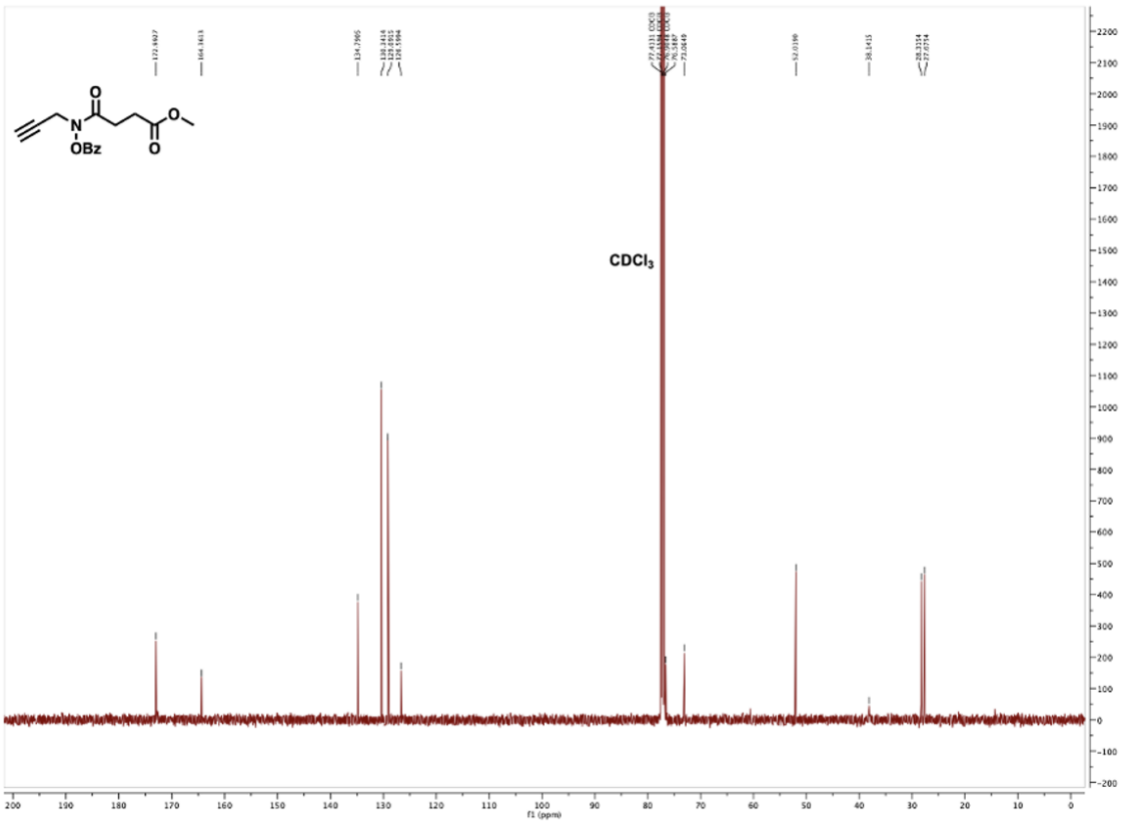
O-benzoyl-N-(but-3-yn-1-yl)hydroxylamine



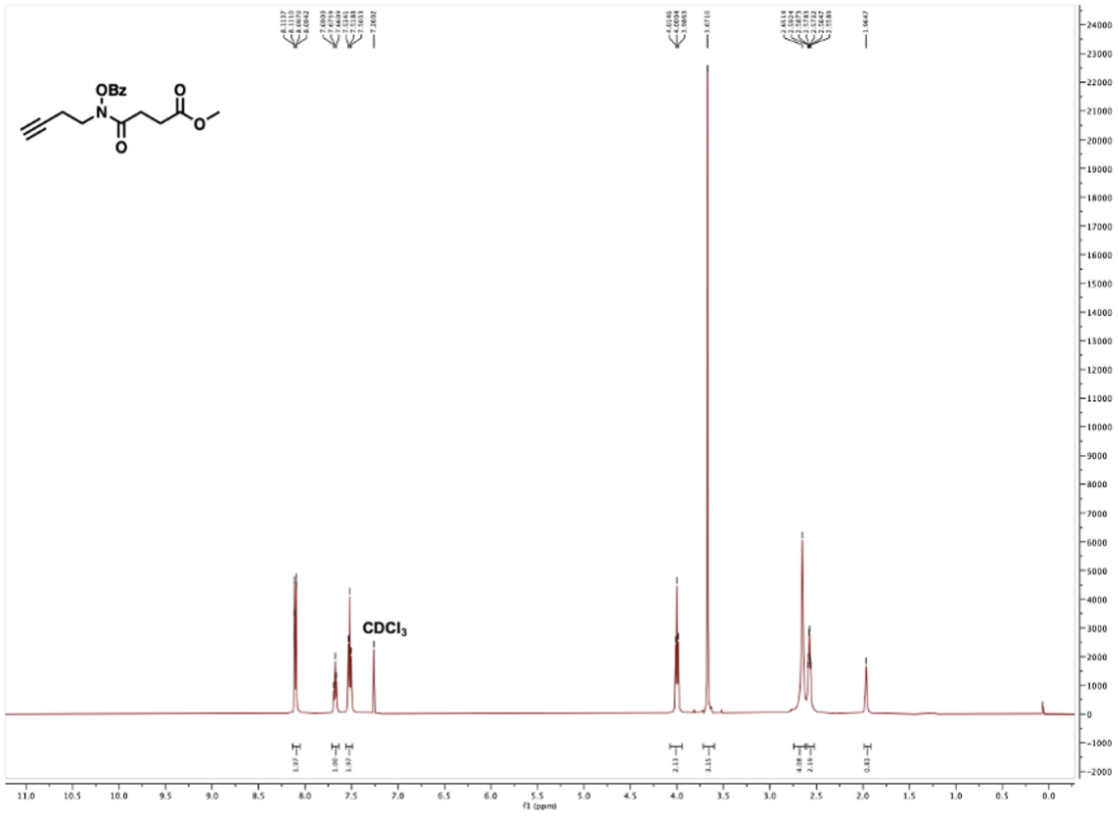
methyl 4-((benzoyloxy)(prop-2-yn-1-yl)amino)-4-oxobutanoate



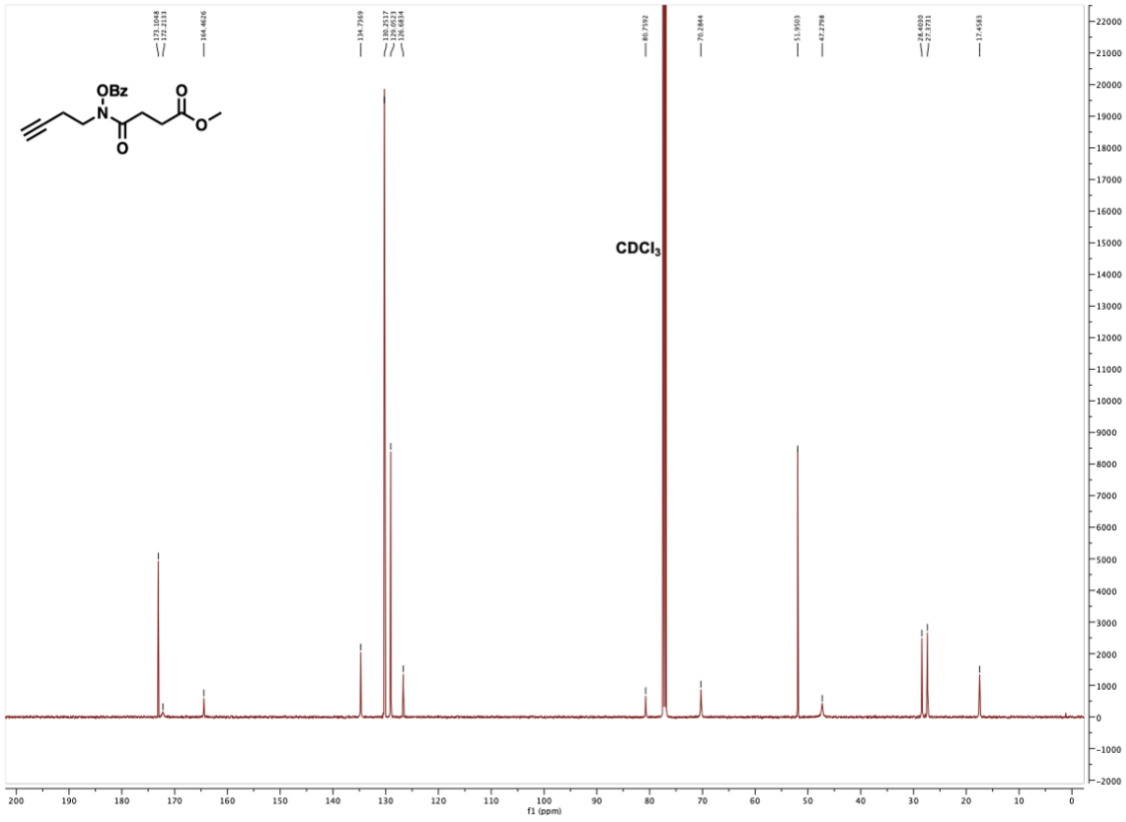
methyl 4-((benzoyloxy)(prop-2-yn-1-yl)amino)-4-oxobutanoate



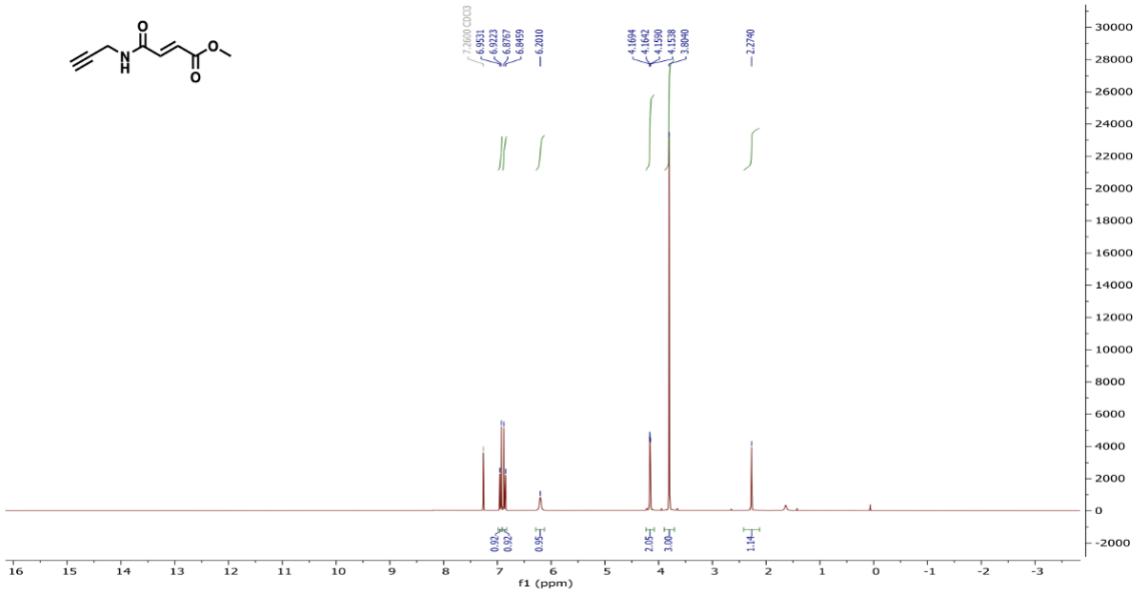
methyl 4-((benzoyloxy)(but-3-yn-1-yl)amino)-4-oxobutanoate



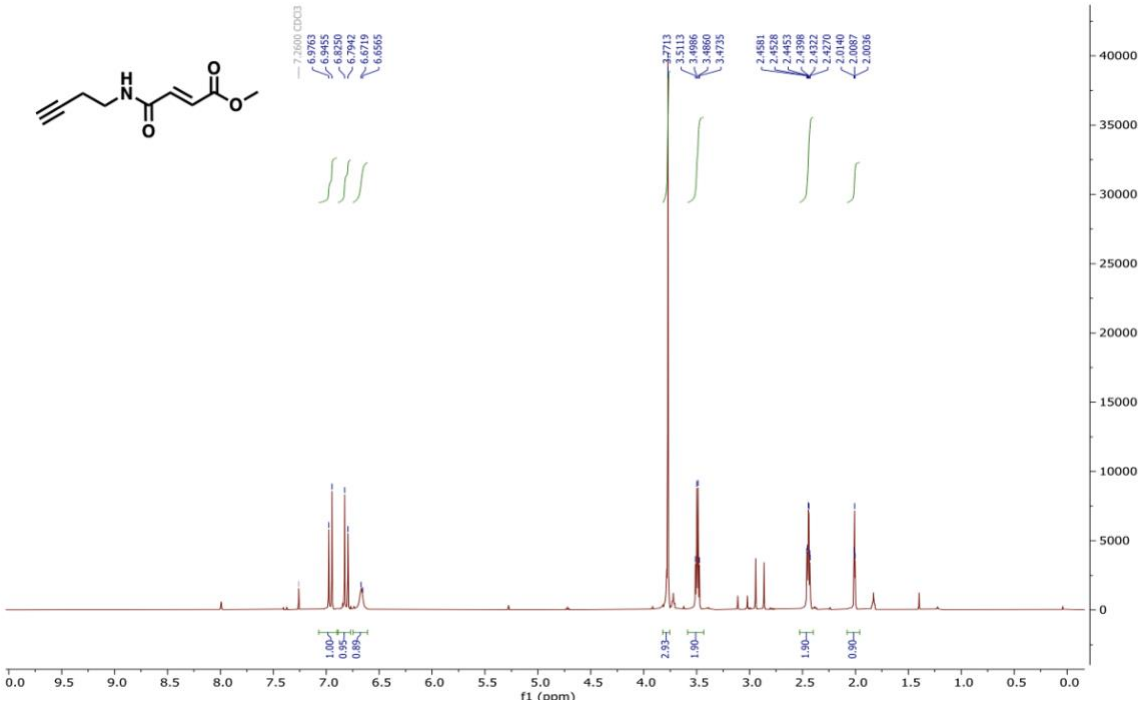
methyl 4-((benzoyloxy)(but-3-yn-1-yl)amino)-4-oxobutanoate



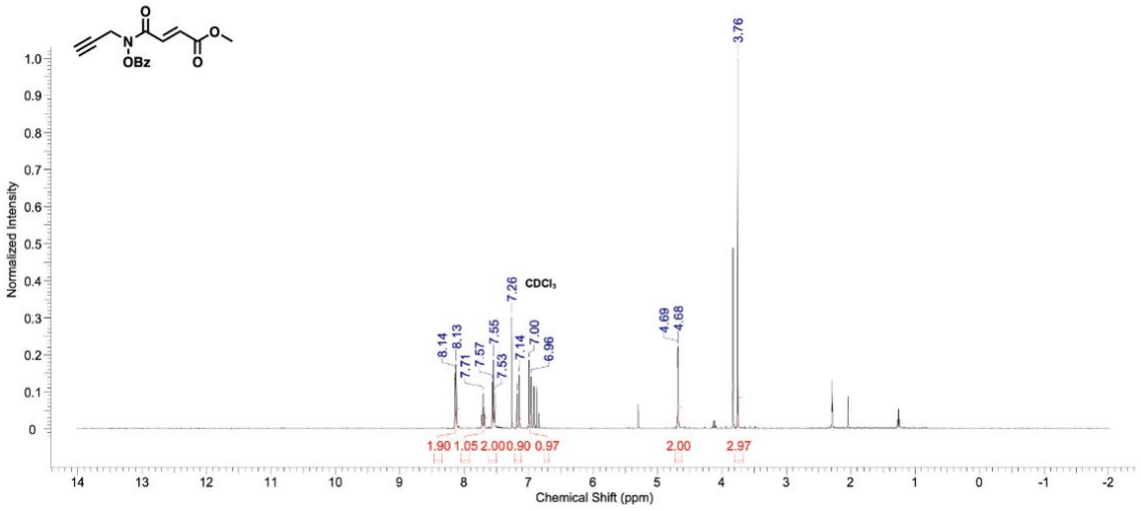
methyl (*E*)-4-(but-3-yn-1-ylamino)-4-oxobut-2-enoate



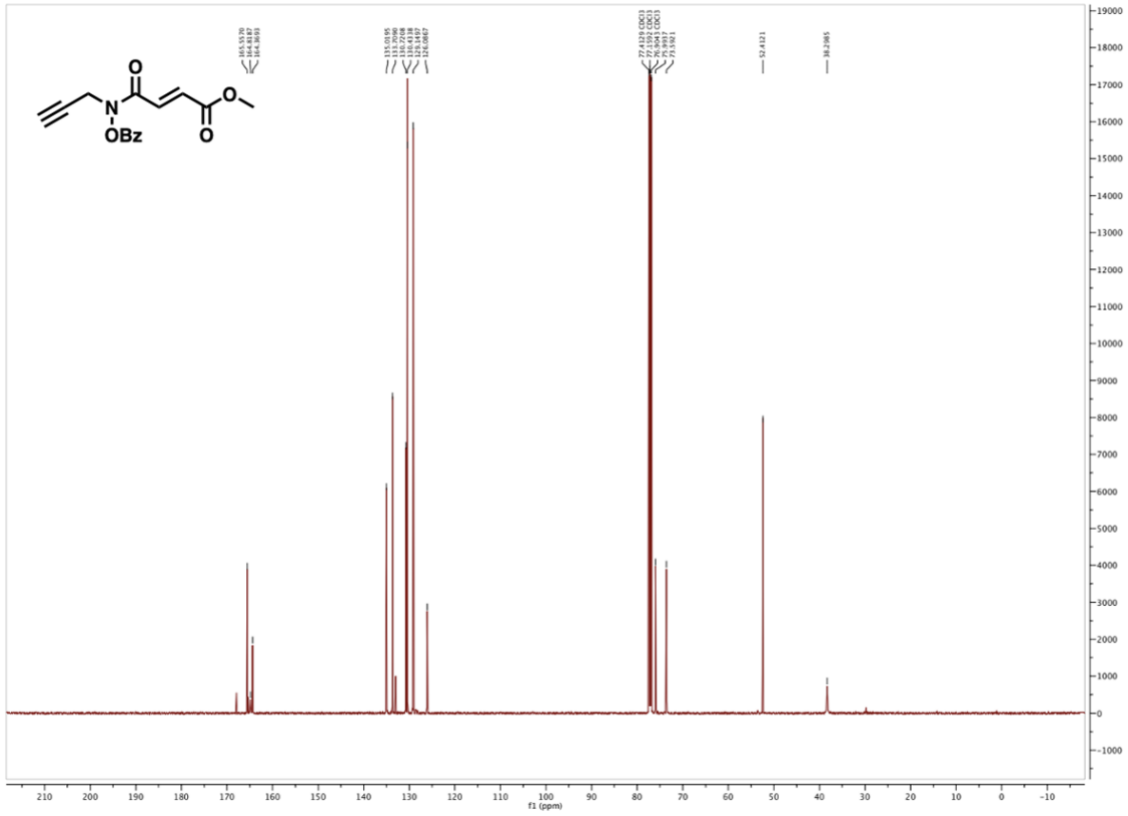
methyl (E)-4-(but-3-yn-1-ylamino)-4-oxobut-2-enoate



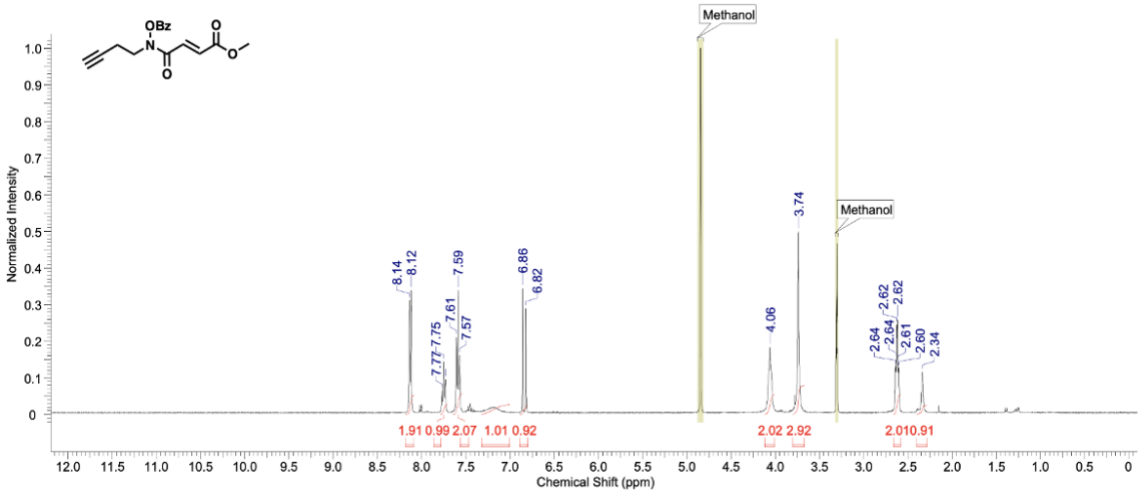
methyl (*E*)-4-oxo-4-(prop-2-yn-1-ylamino)but-2-enoate



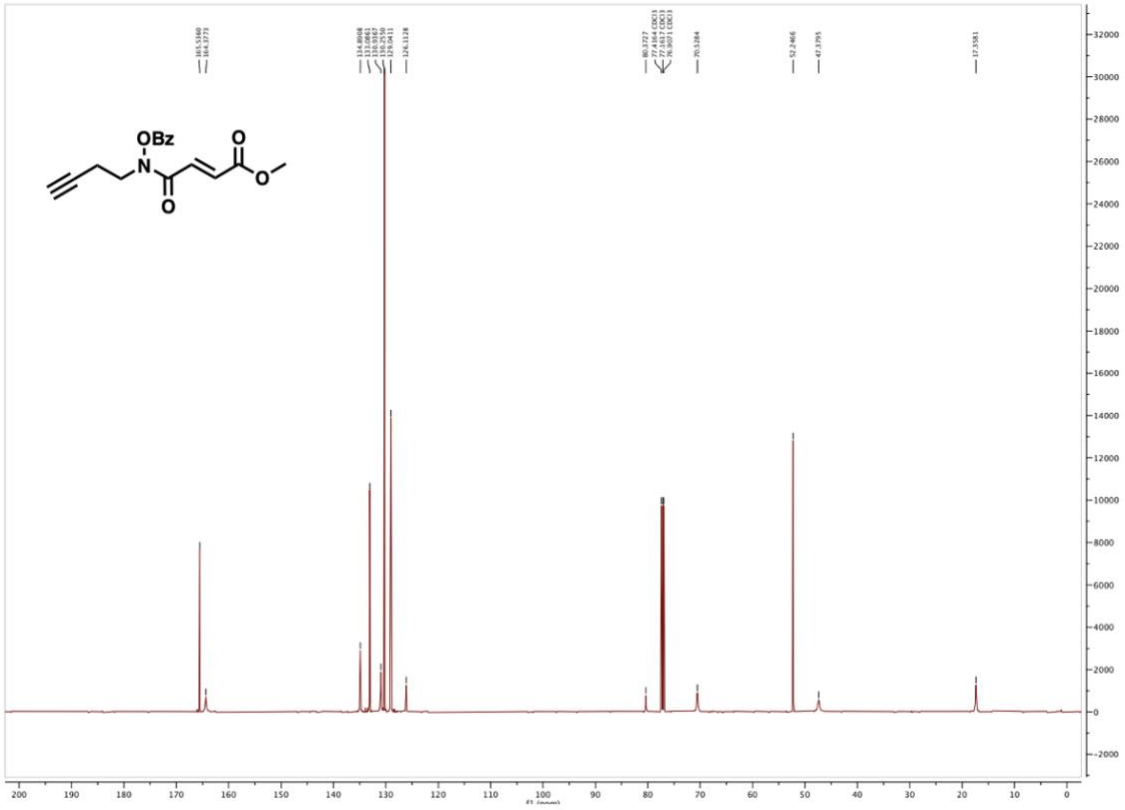
methyl (*E*)-4-oxo-4-(prop-2-yn-1-ylamino)but-2-enoate



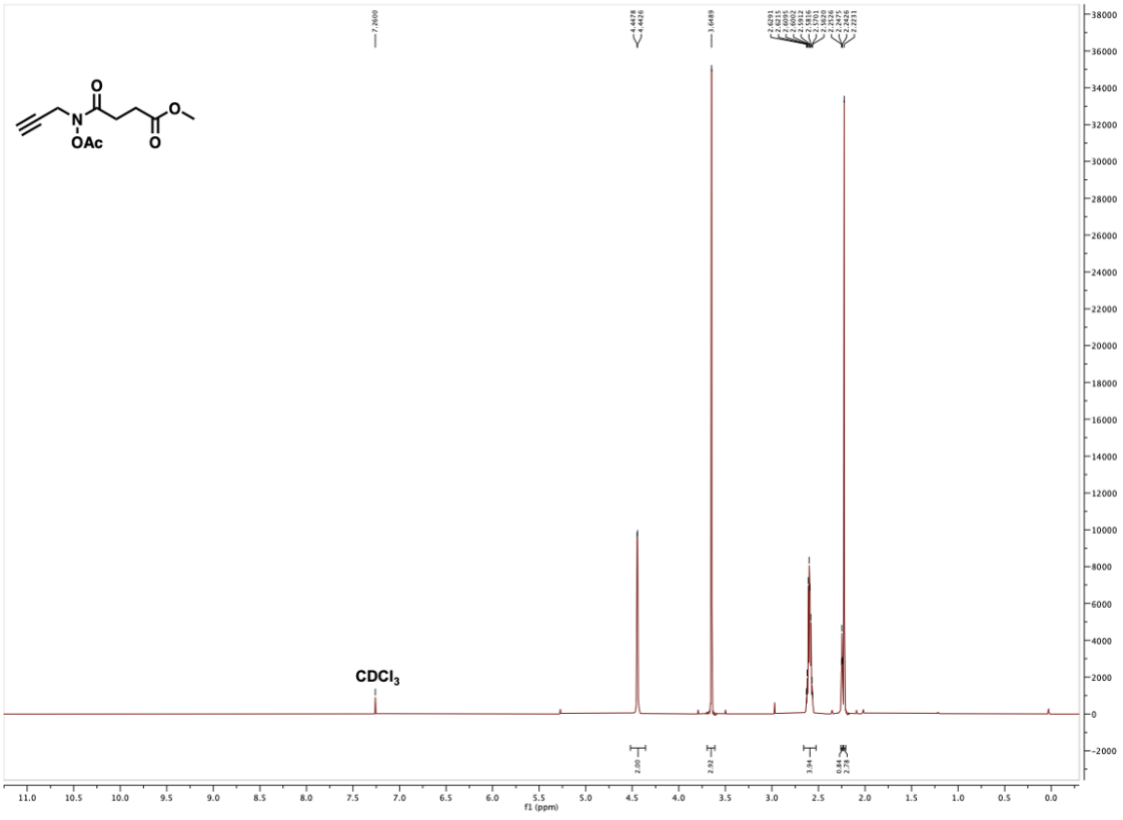
methyl (*E*)-4-(but-3-yn-1-ylamino)-4-oxobut-2-enoate



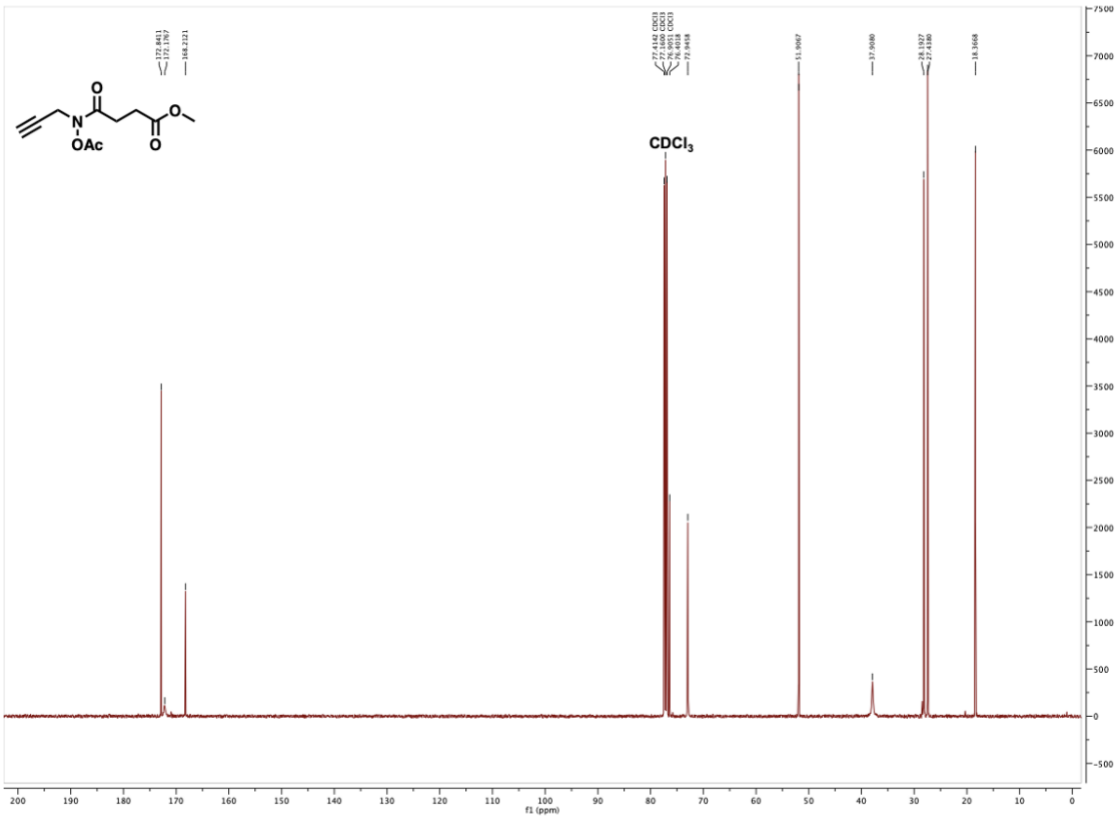
methyl (*E*)-4-(but-3-yn-1-ylamino)-4-oxobut-2-enoate



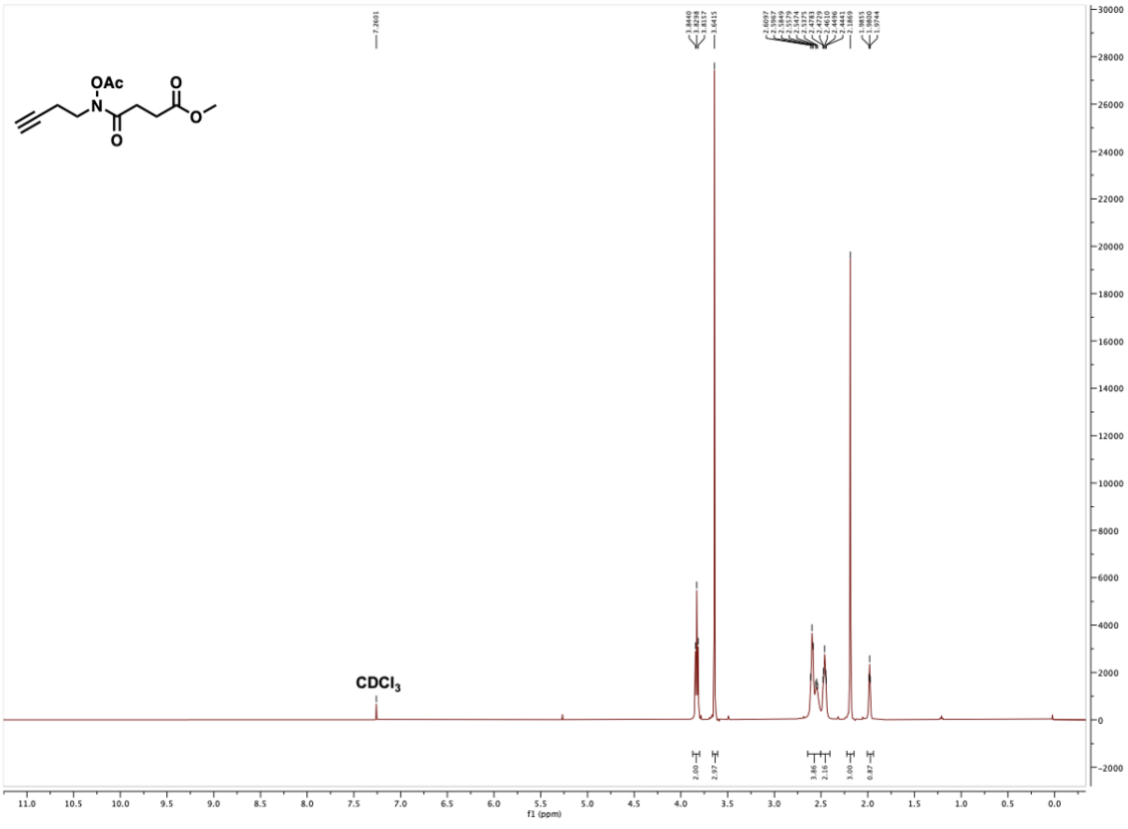
methyl 4-(acetoxy(prop-2-yn-1-yl)amino)-4-oxobutanoate



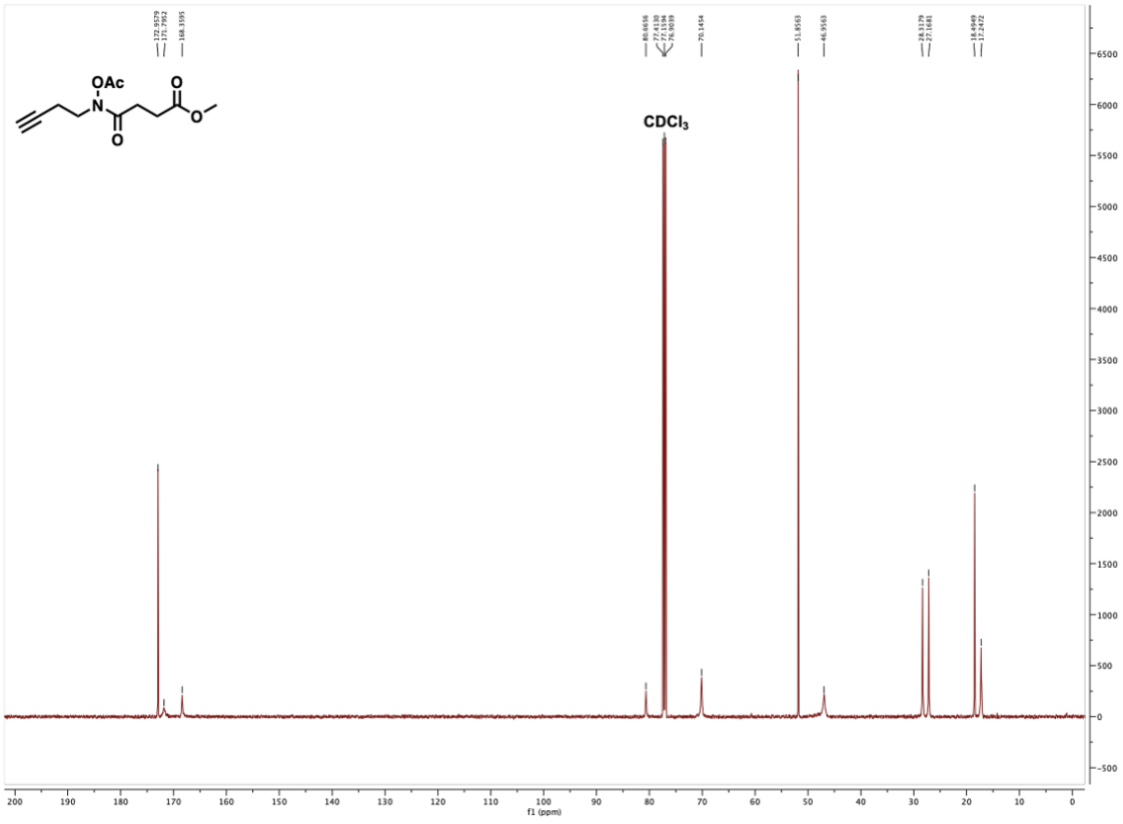
methyl 4-(acetoxy(prop-2-yn-1-yl)amino)-4-oxobutanoate



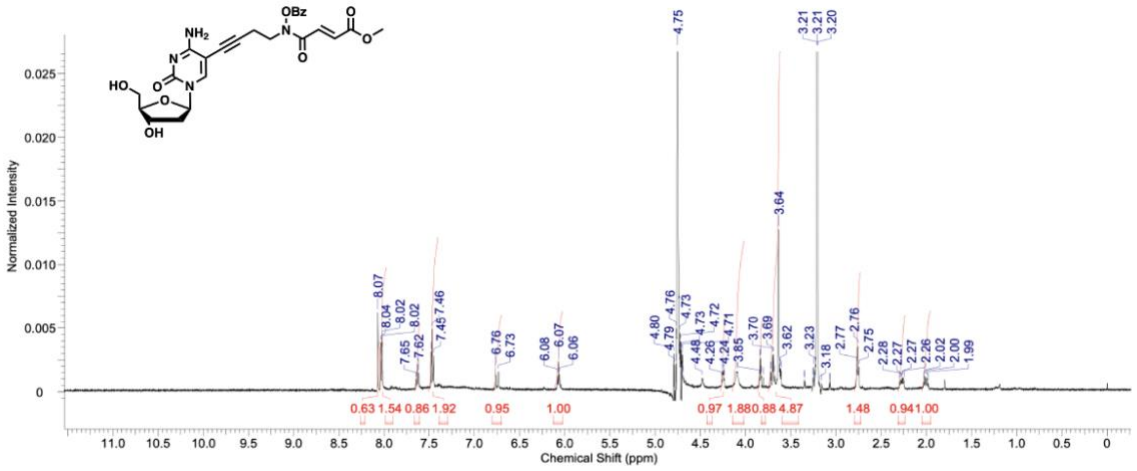
methyl 4-(acetoxy(but-3-yn-1-yl)amino)-4-oxobutanoate



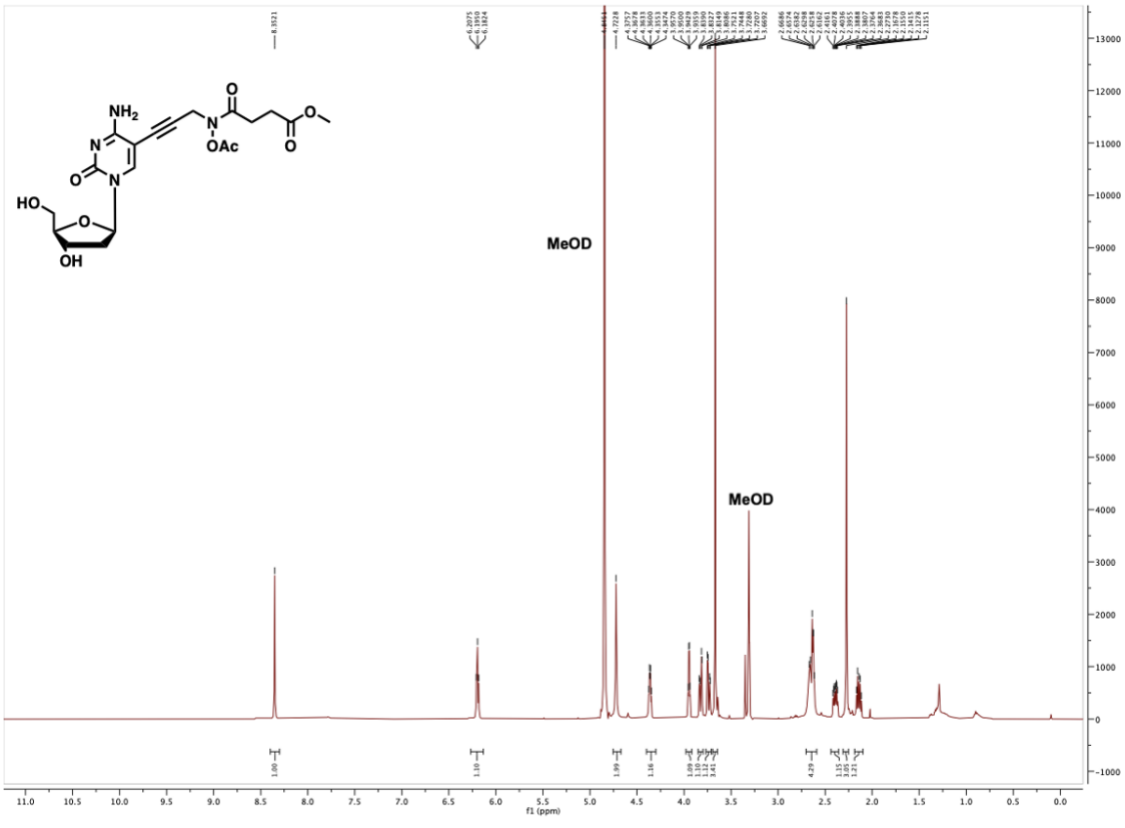
methyl 4-(acetoxy(but-3-yn-1-yl)amino)-4-oxobutanoate



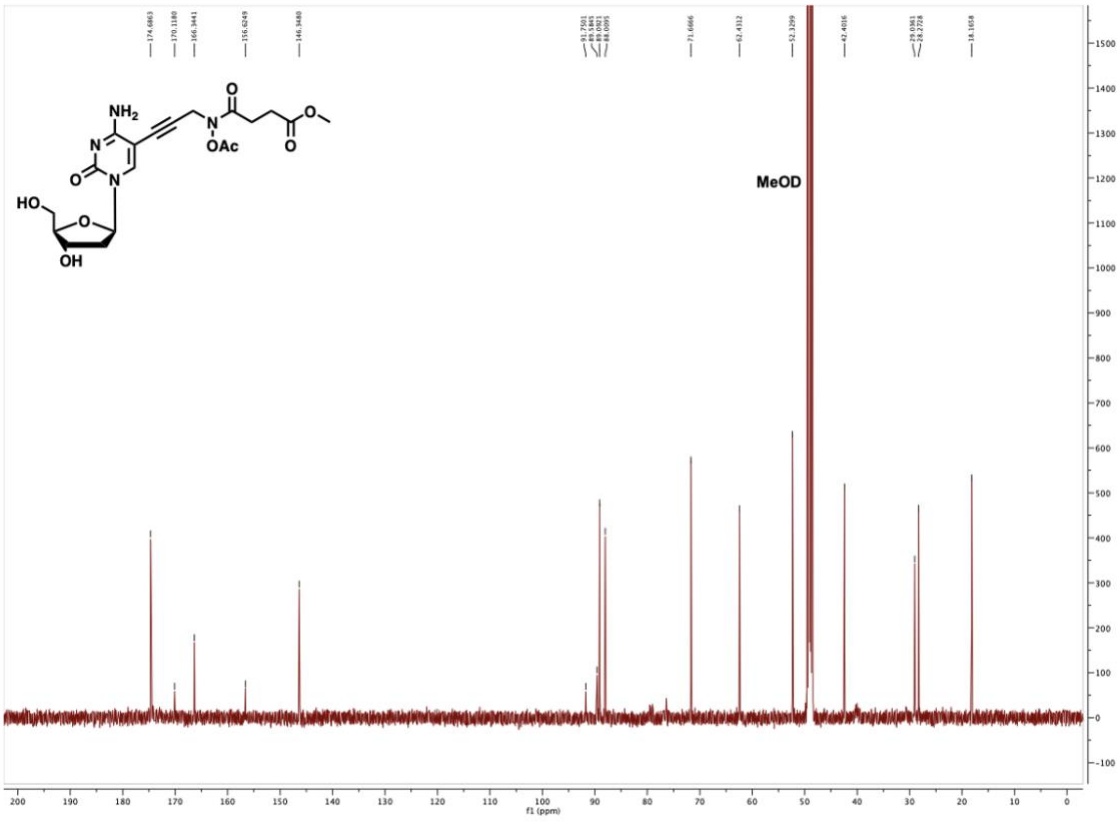
methyl (*E*)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(benzoyloxy)amino)-4-oxobut-2-enoate



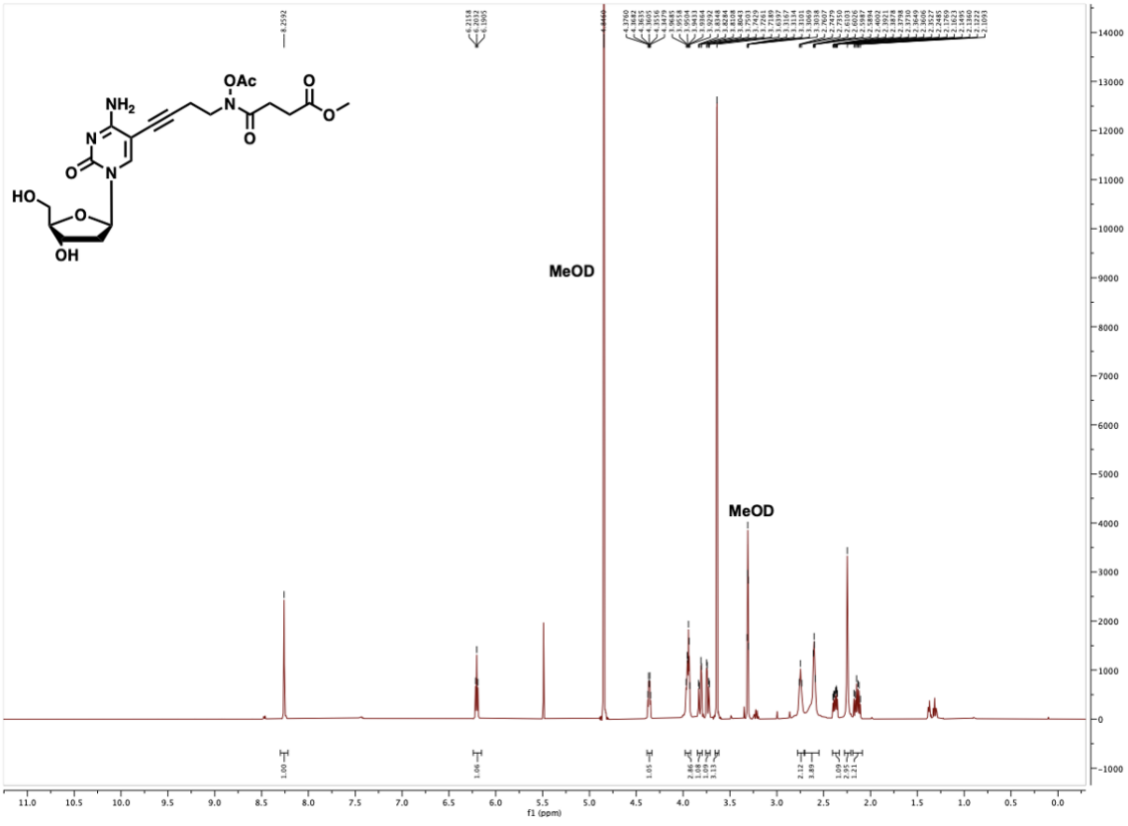
methyl 4-(acetoxy(3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)amino)-4-oxobutanoate



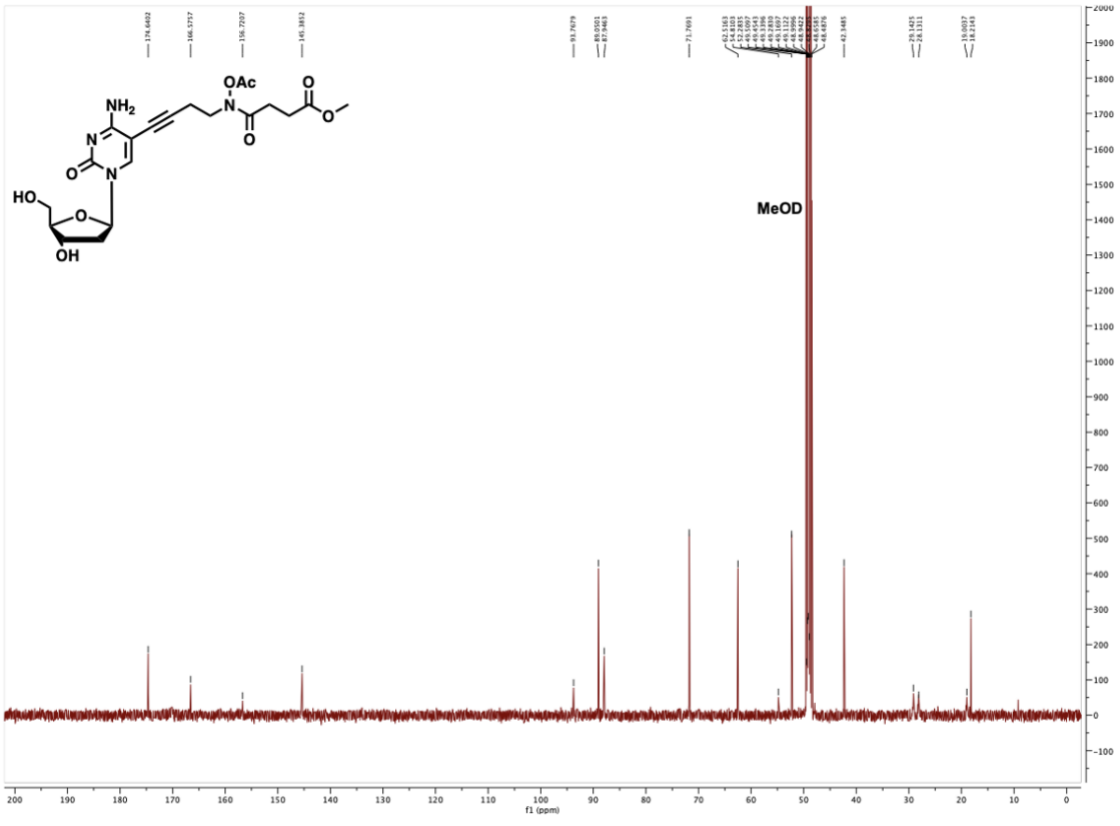
methyl 4-(acetoxy(3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)amino)-4-oxobutanoate



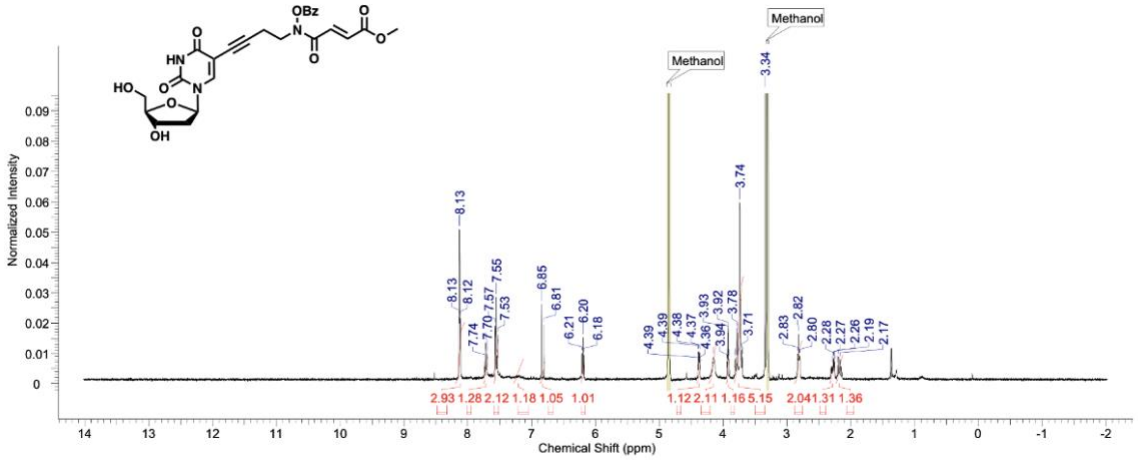
methyl 4-(acetoxy(4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobutanoate



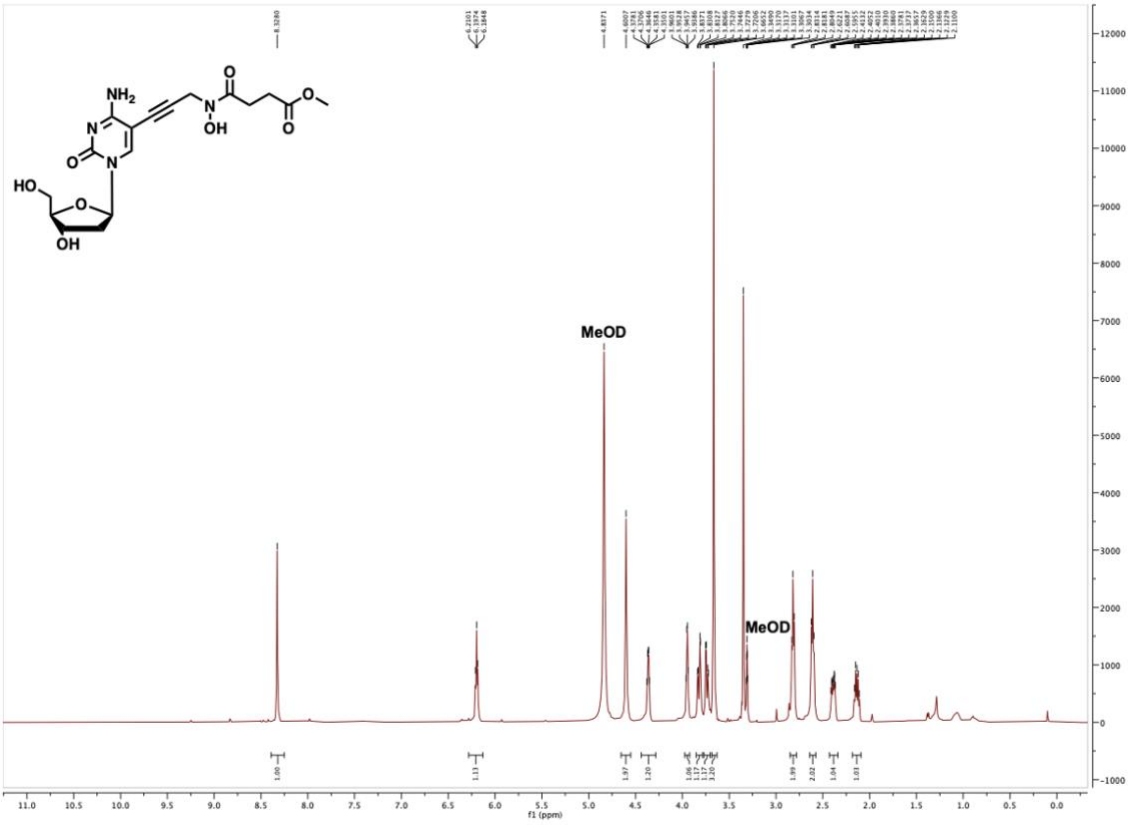
methyl 4-(acetoxy(4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobutanoate



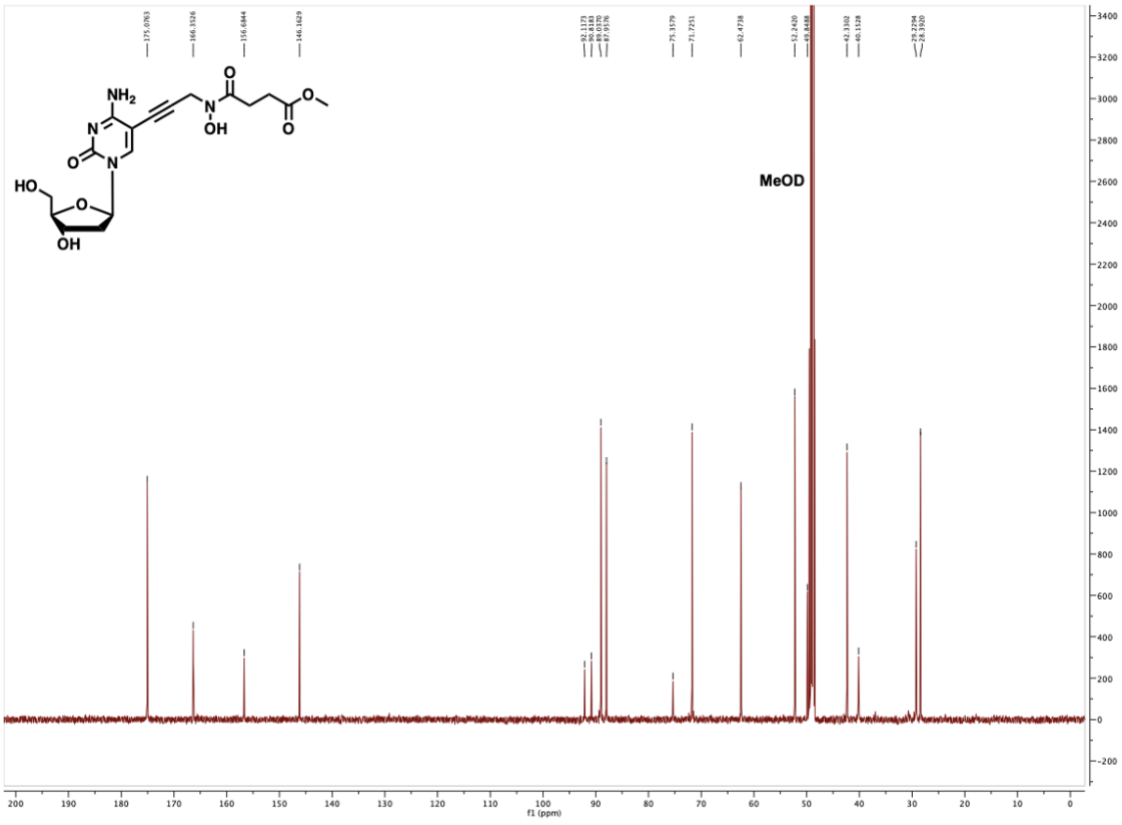
methyl (*E*)-4-((benzoyloxy)(4-(1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoate



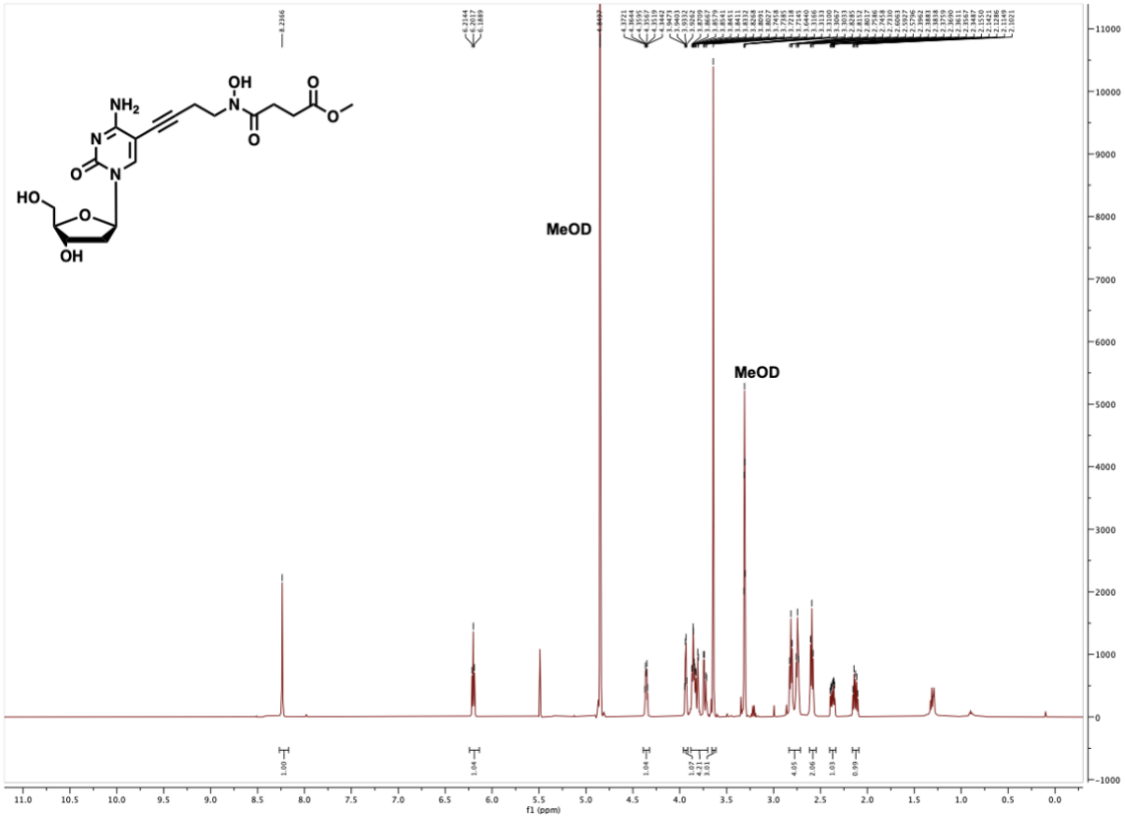
methyl 4-((3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)(hydroxy)amino)-4-oxobutanoate



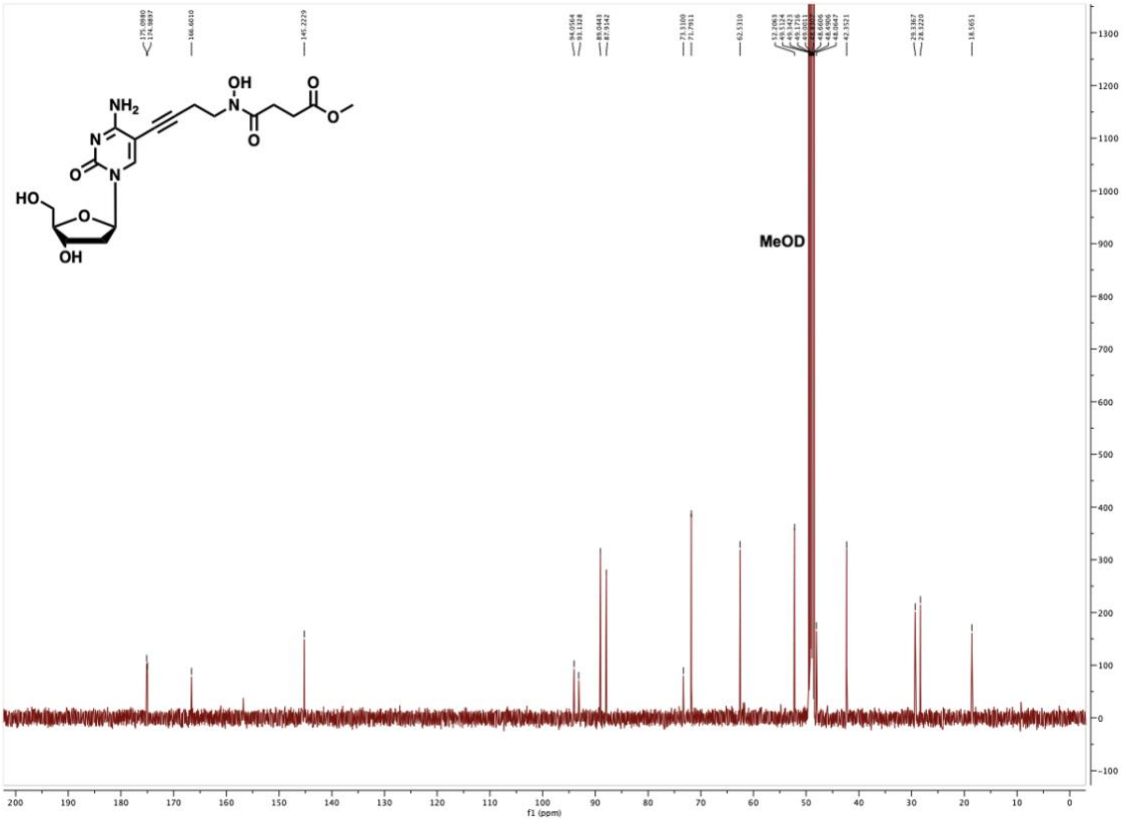
methyl 4-((3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)(hydroxy)amino)-4-oxobutanoate



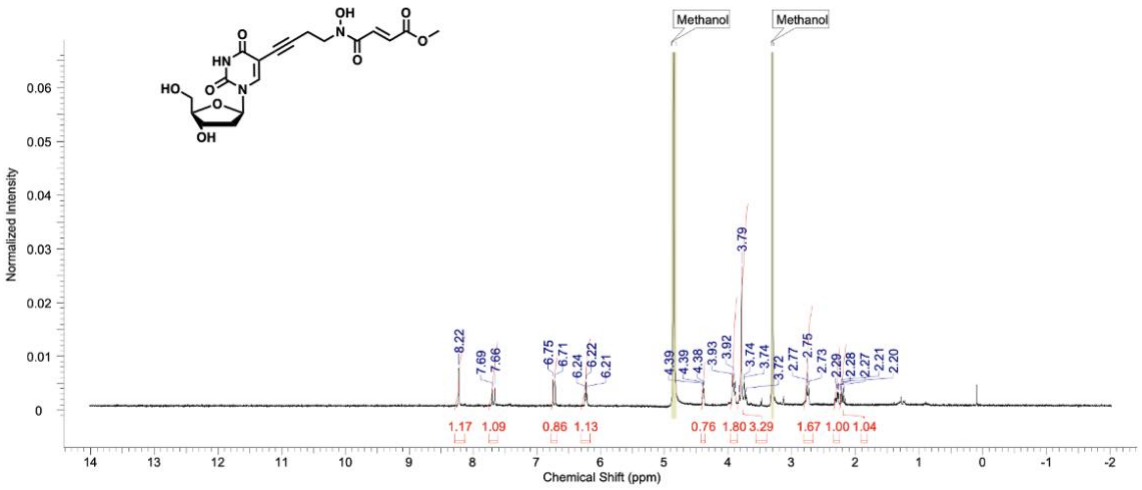
methyl 4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobutanoate



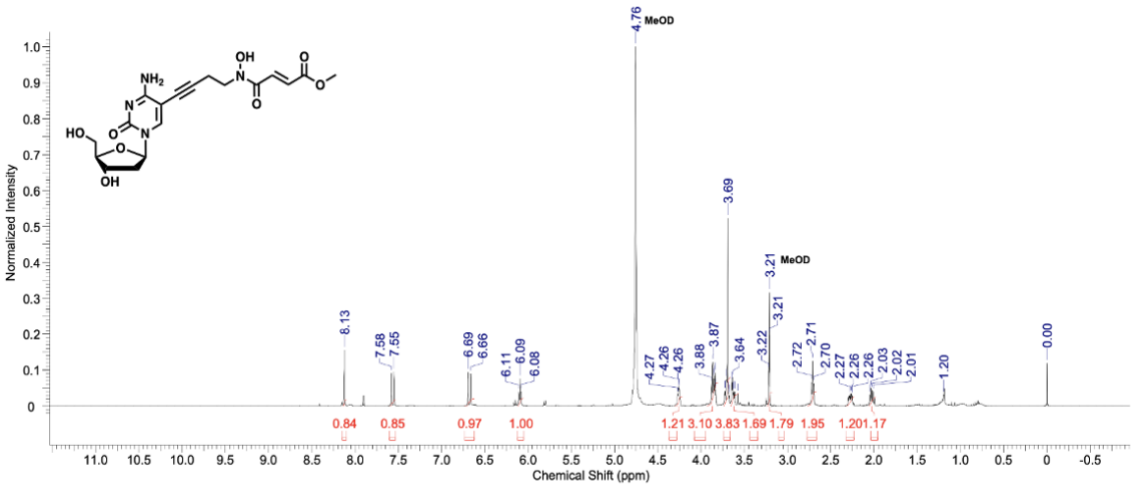
methyl 4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobutanoate



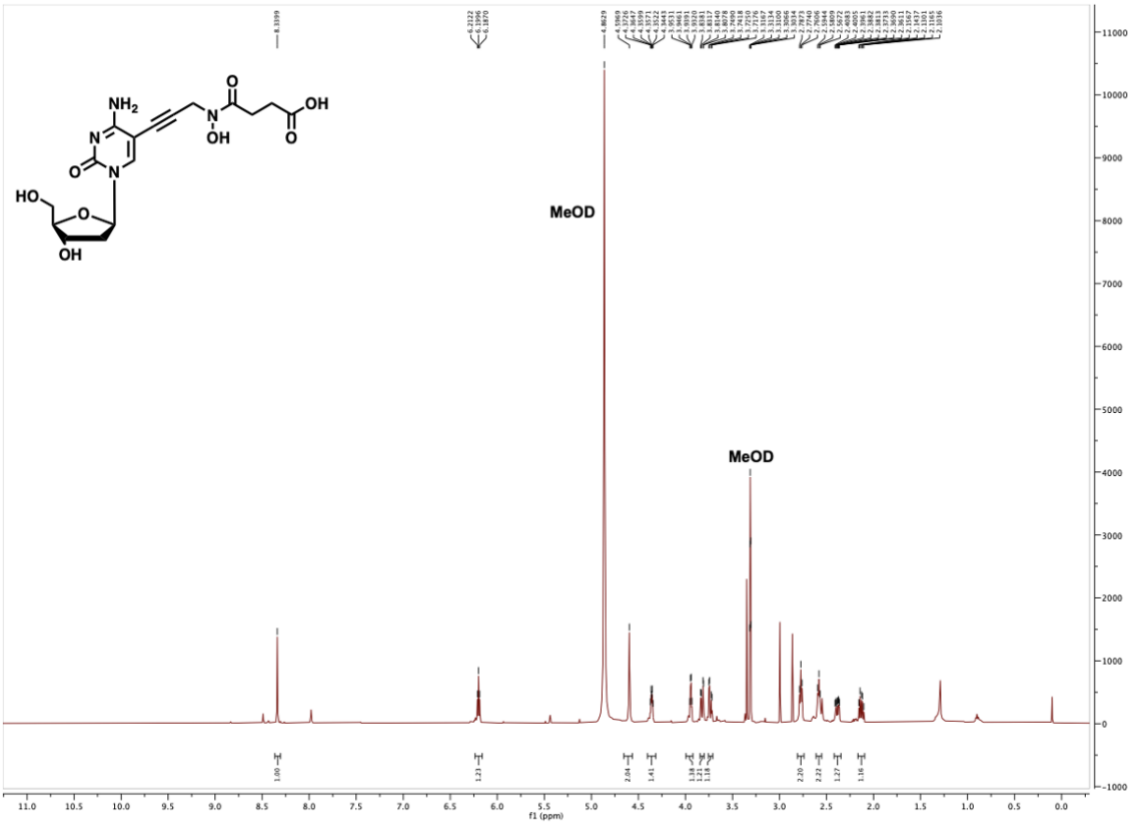
methyl (E)-4-(hydroxy(4-(1-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoate



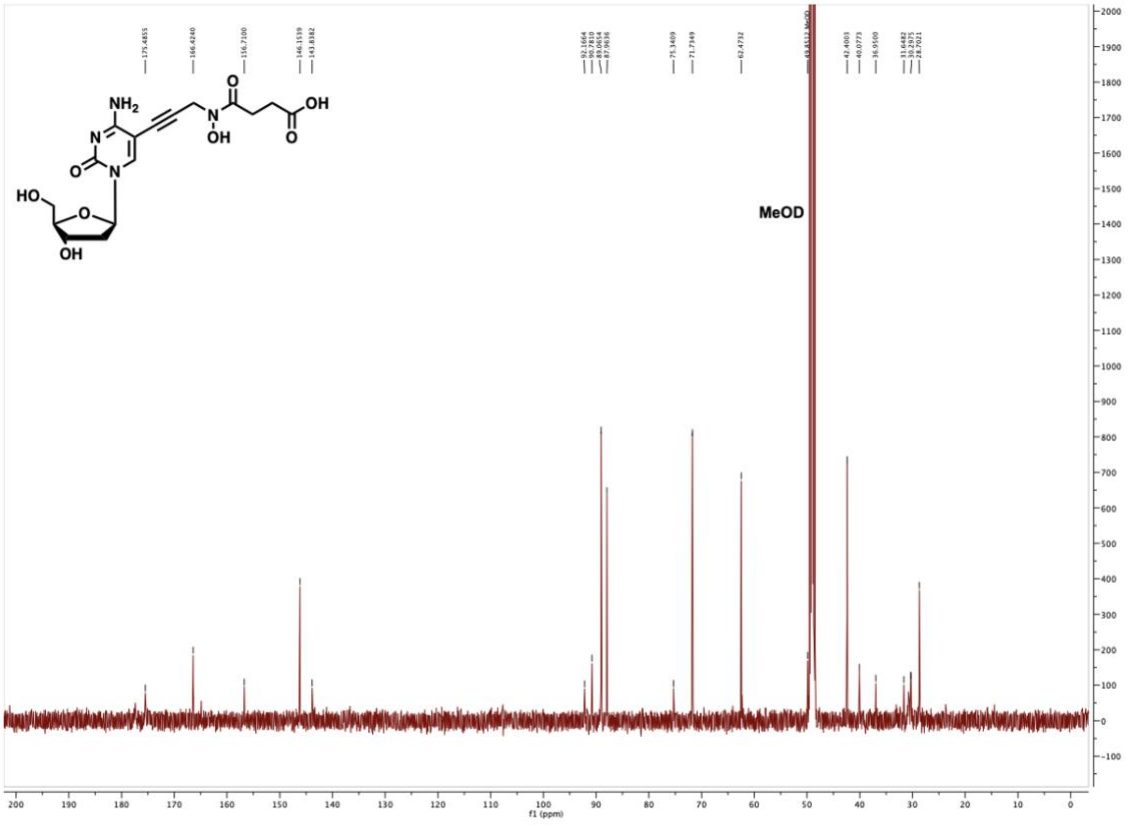
methyl (*E*)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobut-2-enoate



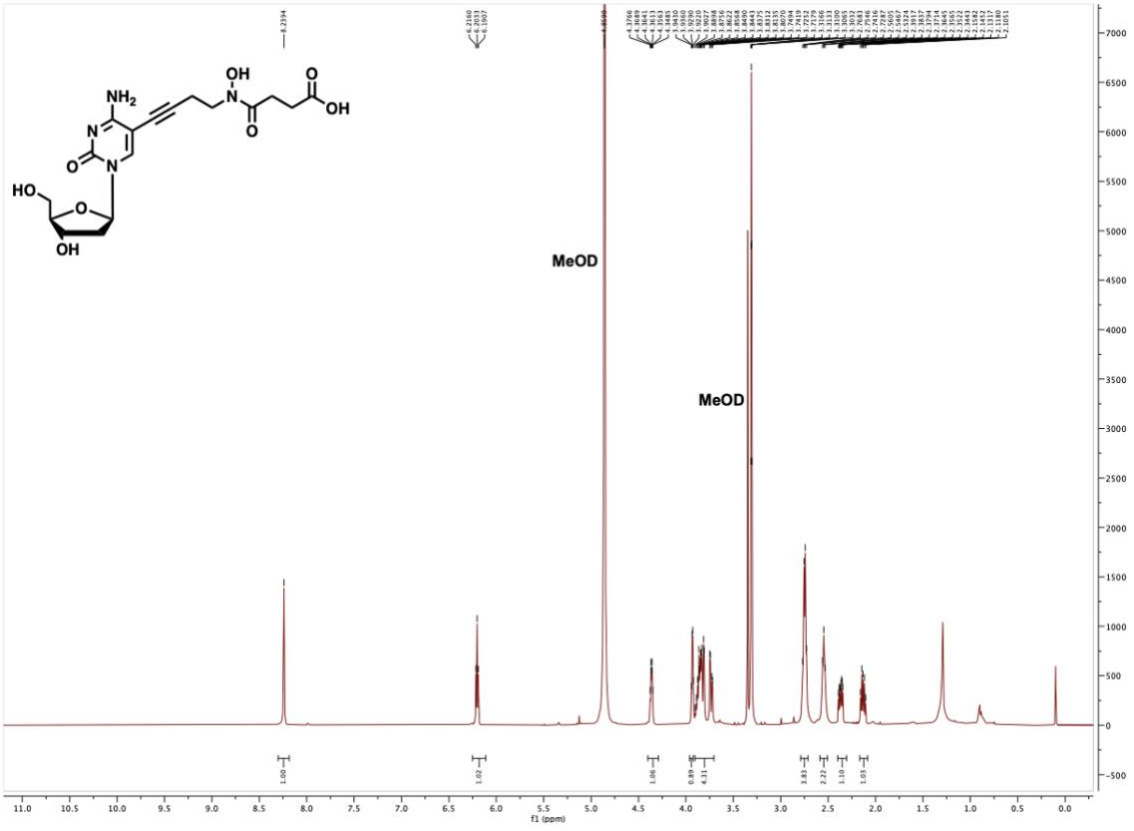
4-((3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)(hydroxy)amino)-4-oxobutanoic acid



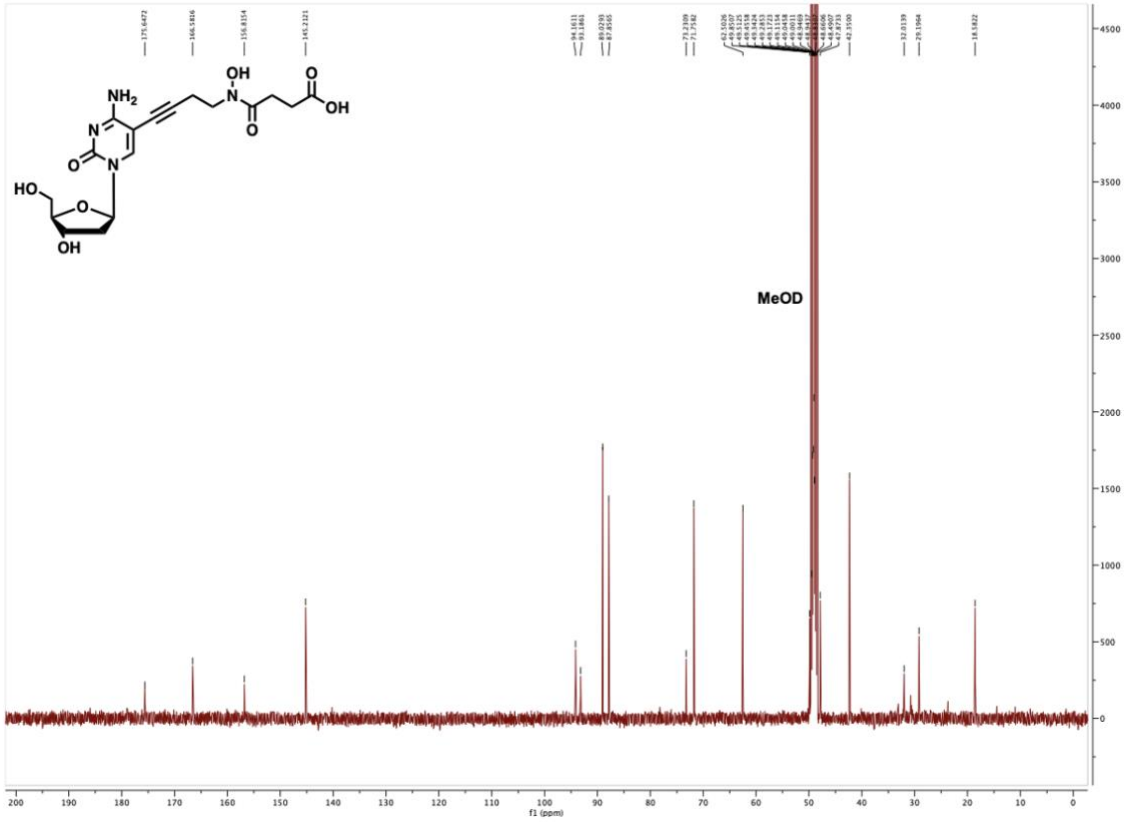
4-((3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)prop-2-yn-1-yl)(hydroxy)amino)-4-oxobutanoic acid



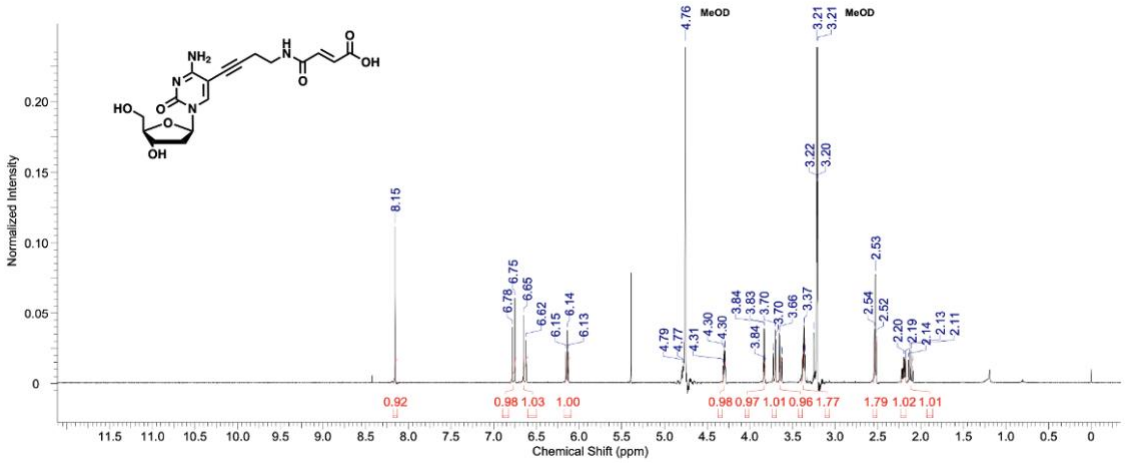
4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobutanoic acid



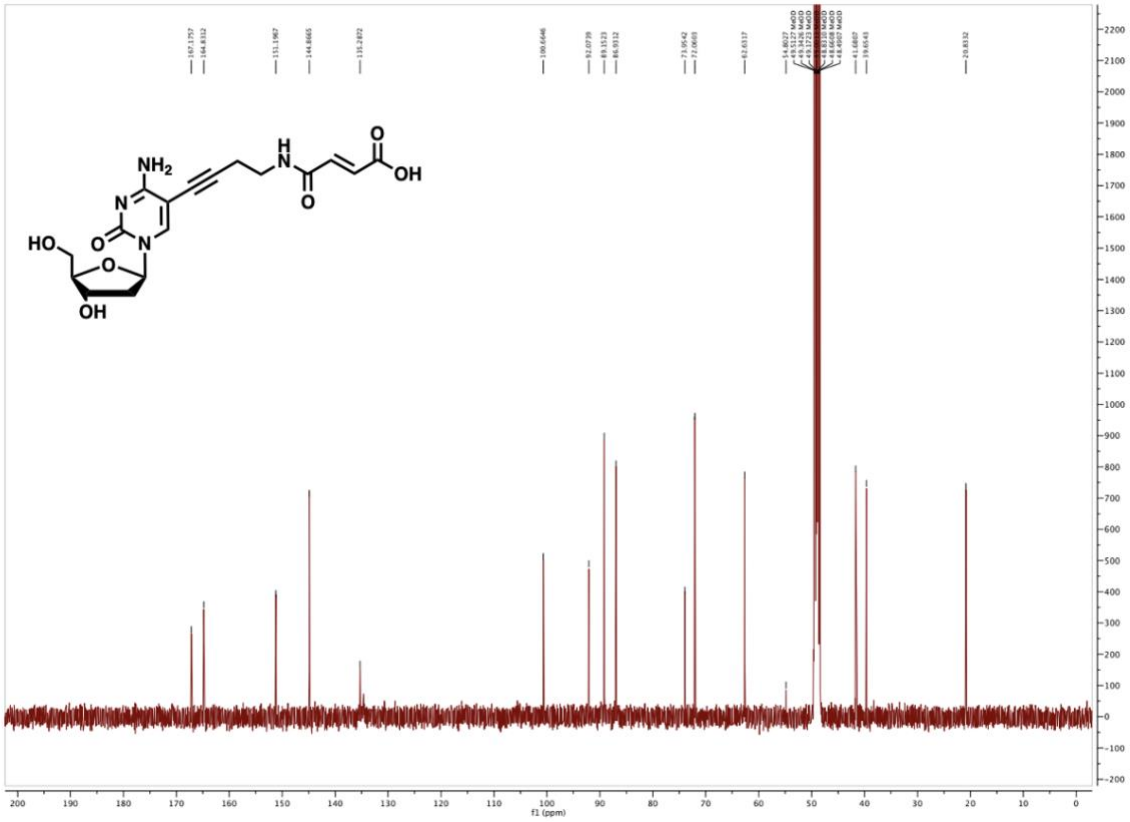
4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobutanoic acid



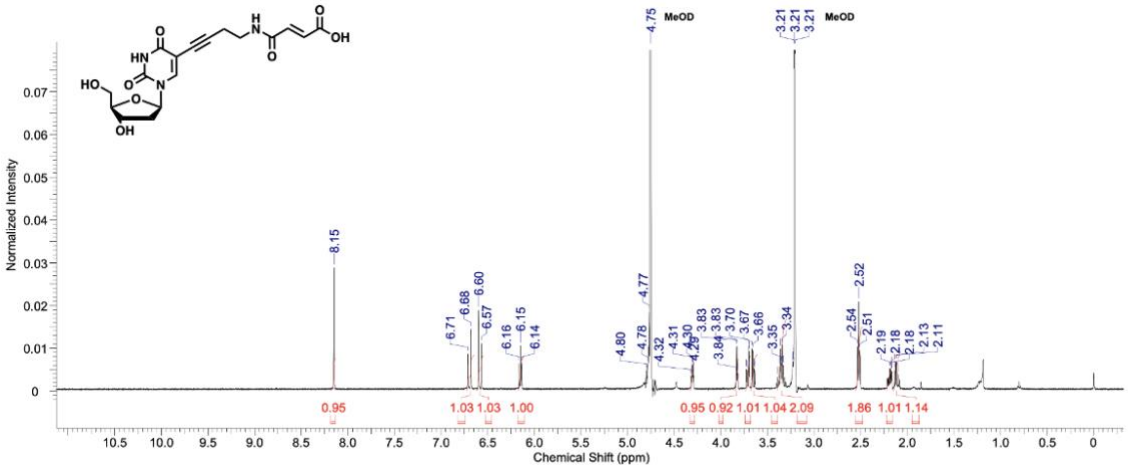
E)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoic acid



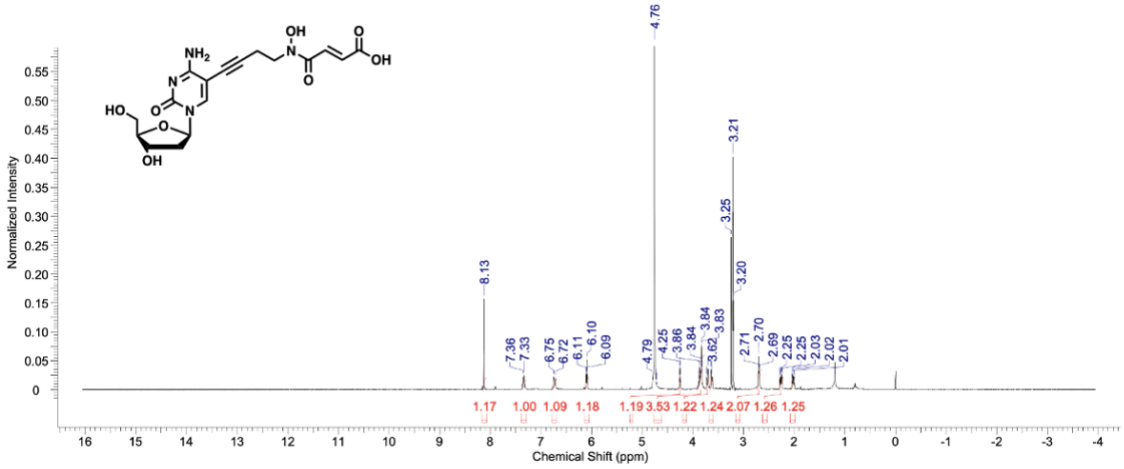
E)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoic acid



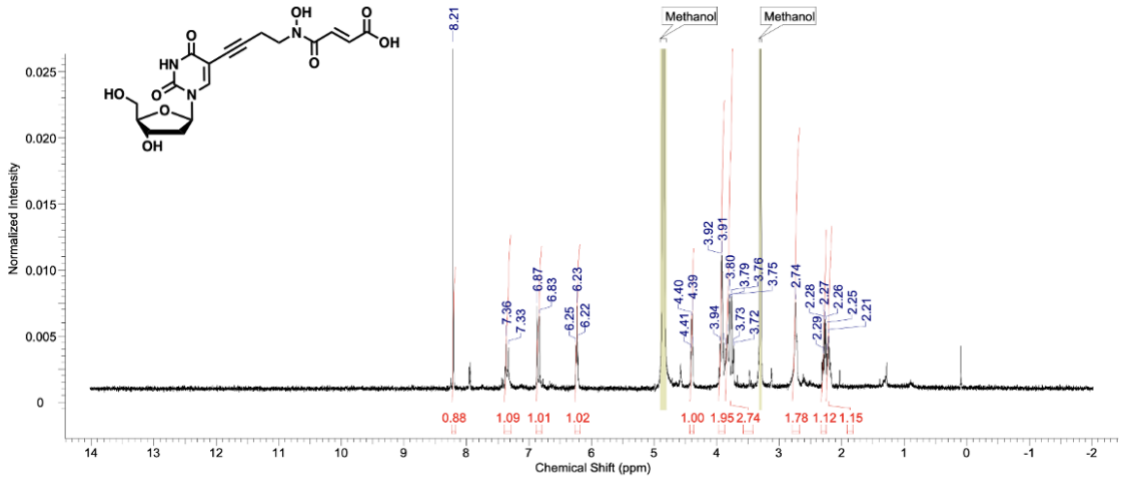
(*E*)-4-((4-(1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoic acid



(*E*)-4-((4-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)but-3-yn-1-yl)(hydroxy)amino)-4-oxobut-2-enoic acid



(*E*)-4-(hydroxy(4-(1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl)but-3-yn-1-yl)amino)-4-oxobut-2-enoic acid



4-((3-(4-amino-1-((2*R*,4*S*,5*R*)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-2-oxo-1,2-dihydropyrimidin-5-yl)propyl)(hydroxy)amino)-4-oxobutanoic acid