ACTIVITY-BASED CHEMICAL PROTEOMICS PROFILING OF NATURAL PRODUCTS AND DRUG-LIKE SMALL MOLECULES

YANG PENGYU

(M. Sc., Chinese Academy of Sciences)

A THESIS SUBMITTED FOR THE DEGREE OF DOCTORATE OF PHILOSOPY

DEPARTMENT OF CHEMISTRY

NATIONAL UNIVERSITY OF SINGAPORE

2011

This thesis is dedicated to my parents, my wife and my daughter.

Acknowledgements

It is my great pleasure to take this opportunity to express my acknowledgements to all the people who have helped or encouraged me during my PhD project. First, I would like to express my deepest appreciation to my supervisor Prof. Yao Shao Q., for his support and guidance throughout the research. His continued support led me to the right way. His intensity, creativity, passion and dedication to science is admirable. Prof. Yao has allowed me great freedom in developing projects to work on in his lab and has always been supportive of the collaborations that led to much of my work. The diverse, interdisciplinary nature of Prof. Yao's research program is one of the things that drew me to his group in the first place, and I have learned much by being in such an environment

Special thanks to my collaborators for their excellent work including: Wang Min and Prof. He Cynthia Y. for their patience and support from my first days learning *T. brucei* to my final days working out these manuscripts; Mun Hong and Prof. James M. Lear for their contribution on library contruction.

Many researchers provided me their reagents or expertise. I would like to thank them here: Prof. Christopher J. Chang (University of California, Berkeley) for the bacterial His-AGT plasmid, mammalian plasmids FLAG-Cox8A-SNAP, FLAG-H2B-SNAP, FLAG-KDEL-SNAP, FLAG-NK1R-SNAP, mCherry-Cox8A, and mCherry-KDEL; Dr. Conor R. Caffrey and Prof. James H. McKerrow (University of California, San Francisco) for cruzain and rhodesain as well as anti-rhodesain and

anti-TbCatB; Prof. Siu Kwan Sze (Nanyang Technological University, Singapore) for his support for LC-MS/MS experiments.

I am truly grateful to all the labmates past and present. Especially I would like to thank Liu Kai, Raja, Wu Hao, Mingyu, Grace, Li Lin, Shen Yuan, Su Ling, and Mei Yin for their help and contribution. I would like to thank Hongyan, Haibin, Lay Pheng, Liqian, Candy, Jinyan, Su Ying, Xiamin, Chongjing, Zhenkun, and Xiaoyuan for their friendship. It is my great pleasure to study in such a friendly lab atmosphere.

I would like to thank the members of my thesis committee: Prof. Lu Yixin, Prof. Tan Choon Hong and Prof. James M. Lear for their input and helpful suggestions of my research projects.

I also acknowledge kind support from NUS for providing me research scholarship. Thanks also go to Department of Chemistry administrative staffs for their supports, especially Suriawati Bte Sa'Ad for her help on all kinds of issues.

Finally, I am grateful to my family and friends for their constant support and encouragement. I own my parents and parents-in-law for their endless love and support for my study. I am in debt to my brother and my sister, who takes care of my parents when I study abroad. I am deeply indebted to my wife Baiyun, for her endless love, support, and encouragement at all times. Without her support, this thesis work is impossible. My love to my daughter Guyu, cheers me up with her cute smile, her endless curiosity, her naughty requests, and sometimes even her cries. I would like to thank my friends – Zhibin, Lifa, Yunben, Yujun and their families – for their friendship and help.

Table of Contents

Summary

Assigning the cellular target(s) of bioactive small molecules, whether the compounds are discovered by cell-based phenotypic or target-based screens of chemical libraries, remains an ongoing challenge. The ability of accurately and thoroughly determining of SM–target interaction profiles as well as mapping of metabolic and signaling pathways on the proteomic scale would therefore be more illuminating, as it could provide invaluable biological insights for a drug candidate by both understanding the primary mechanism-of-action, and at the same time, side effects due to unexpected "off-target" interactions at a very early stage of drug development, which should help to reduce the attrition rate in development. In many cases such a capability could find new potential therapeutic value for an established drug as well as it could also offer strong clues for compound optimization in order to maximize the therapeutic potential and minimize potential cellular toxicity of a drug. The data may also serve to define previously unknown protein functions, based on the phenotypes induced by compounds. Recent advances in chemical proteomics (or activity-based proteomics), a multidisciplinary research area integrating biochemistry and cell biology with organic synthesis and mass spectrometry, have enabled a more direct and unbiased analysis of a drug's mechanism of action in the context of the proteomes as expressed in the target cell or the tissue of interest.

In this thesis, I describe the design and synthesis of Orlistat^{TM}-like natural product-based probes (Chapter 2, 3 & 4), K11777-like drug candidate-based probes (Chapter 5) and azanitrile-containing small molecules (Chapter 6), determination of

structure-activity relationships of these compounds, cellular target identification, validation and cellular localization in subsequent molecular biology and cell biological experimentsin both living mammalian cells and *Trypanosoma brucei* parasites.

List of Publications

(2007-2011)

1. P.-Y. Yang, M. Wang, K. Liu, M. H. Ngai, O. Sheriff, M. J. Lear, S. K. Sze, C. Y. He, S. Q. Yao, Parasite-Based Screening and Proteomic Profiling Reveals Orlistat, an FDA-Approved Drug, as a Potential Anti-*Trypanosoma brucei* Agent. *Chem. Eur. J.* **2012**, submitted.
2. P.-Y. Yang, M. Wang, H. Wu, L. Li, C. Y. He, S. Q. Yao, Design, Synthesis and Biological Evaluation of Potent Azadipeptide Nitrile Inhibitors and Activity-Based Probes as Promising Anti-*Trypanosoma brucei* Agents. *Chem. Eur. J.* **2012**, in press.
3. P.-Y. Yang, M. Wang, C. Y. He, S. Q. Yao, Activity-Based Proteome Profiling of Potential Cellular Targets of K11777 - a Clinical Cysteine Protease Inhibitor. *Chem. Commun.* **2012**, *48*, 835-837.
4. P.-Y. Yang, K. Liu, C. Zhang, G. Y. J. Chen, Y. Shen, M. H. Ngai, M. J. Lear, S. Q. Yao, Chemical Modification and Organelle-Specific Localization of Orlistat-Like Natural Product-Based Probes. *Chem. Asian. J.* **2011**, *6*, 2762-2775.
5. K. Liu, P.-Y. Yang, Z. Na, S. Q. Yao, Dynamic Profiling of Post-Translational Modifications on Newly Synthesized Proteins Using a Double Metabolic Incorporation Strategy. *Angew. Chem. Int. Ed.* **2011**, *50*, 6776-6781.

6. H. Wu, J. Ge, P.-Y. Yang, J. Wang, M. Uttamchandani, S. Q. Yao, A Peptide Aldehyde Microarray for High-Throughput Detection of Cellular Events. *J. Am. Chem. Soc.* **2011**, *133*, 1946-1954.
7. M. Hu, L. Li, H. Wu, Y. Su, P.-Y. Yang, M. Uttamchandani, Q.-H. Xu, S. Q. Yao, Multi-Color, One- and Two-Photon Imaging of Enzymatic Activities in Living Cells with Novel Fluorescently Quenched Activity-Based Probes (*q*ABPs). *J. Am. Chem. Soc.* **2011**, *133*, 12009-12020.
8. M. H. Ngai, P.-Y. Yang, K. Liu, Y. Shen, S. Q. Yao, M. J. Lear, Click-Based Synthesis and Proteomic Profiling of Lipstatin Analogues. *Chem. Commun.* **2010**, *46*, 8335-8337. (Cover Feature Article)
9. P.-Y. Yang, K. Liu, M. H. Ngai, M. J. Lear, M. R. Wenk, S. Q. Yao, Activity-Based Proteome Profiling of Potential Cellular Targets of Orlistat – an FDA-Approved Drug with Anti-Tumor Activities. *J. Am. Chem. Soc.* **2010**, *132*, 656-666. (Highlighted by Faculty of 1000 Biology)
10. L. P. Tan, H. Wu, P.-Y. Yang, K. A. Kalesh, X. Zhang, M. Hu, R. Srinivasan, S. Q. Yao, High-Throughput Discovery of Mycobacterium Tuberculosis Protein Tyrosine Phosphatase (MptpB) Inhibitors Using Click Chemistry. *Org. Lett.* **2009**, *11*, 5102-5105.
11. R. Srinivasan, L. P. Tan, H. Wu, P.-Y. Yang, K. A. Kalesh, S. Q. Yao, High-Throughput Synthesis of Azide Libraries Suitable for Direct "Click" Chemistry and in situ Screening. *Org. Biol. Chem.* **2009**, *7*, 1821-1828.

12. P.-Y. Yang, H. Wu, M. Y. Lee, A. Xu, R. Srinivasan, S. Q. Yao, Solid-Phase Synthesis of Azidomethylene Inhibitors Targeting Cysteine Proteases. *Org. Lett*. **2008**, *10*, 1881-1884.

13. S. L. Ng, P.-Y. Yang, K. Y.-T. Chen, R. Srinivasan, S. Q. Yao, "Click" synthesis of small-molecule inhibitors targeting caspases. *Org. Biomol. Chem.* **2008**, *6*, 844–847.

14. K. A. Kalesh, P.-Y. Yang, R. Srinivasan, S. Q. Yao, Click Chemistry as a High-Throughput Amenable Platform in Catalomics. *QSAR Comb. Sci.* **2007***, 26*, 1135–1144.

List of Abbreviations

Å	angstrom(s)
ABC	ammonium bicarbonate
ABPs	activity-base probes
ABPP	activity-based protein profiling
Ac	acetyl
ACC	7-aminocoumarin-4-acetic acid
ACN	acetonitrile
AGT	O^6-alkylguanine-DNA alkyltransferase
aq.	aqueous
Ar	aryl, argon
BINOL	1,1'-bi-2-naphthol
br	broad
BG	benzylguanine
Bn	benzyl
Boc	*tert*-butyl-oxycarbonyl
BSA	bovine serum albumin
BSF	bloodstream form
Cat	cathepsin
δ	chemical shift in ppm
Cbz	benzyloxycarbonyl
calcd	calculated
CuAAC	copper (I)-catalyzed azide-alkyne cycloaddition
DCC	*N*, *N'*-dicyclohexylcarbodiimide
DCM	dichloromethane
dd	doublet of doublet
°C	degrees Celsius
2D-GE	two-dimensional gel electrophoresis

DIAD	diisopropyl azodicarboxylate
DIC	*N*, *N'*-diisopropylcarbodiimide
DIEA	*N*, *N'*-diisopropylethylamine
DIGE	differential gel electrophoresis
DMAB	dimethylamine borane
DMAP	4-dimethylaminopyridine
DMEM	dulbecco's modified eagle medium
DMF	*N*, *N'*-dimethylformamide
DMP	Dess-Martin periodinane
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
EA	ethyl acetate
E. coli	*Escherichia coli*
ED_{50}	half-maximal effective dose
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	ethylenediaminetetracetic acid
ee	enantiomeric excess
EI	electron ionization
emPAI	exponentially modified protein abundance index
equiv.	equivalent(s)
ER	endoplasmic reticulum
ESI	electron spray ionization
Et	ethyl
FAS	fatty acid synthase
FBS	fetal bovine serum
FDA	US food and drug administration
FITC	fluorescein isothiocyanate
Fmoc	9-fluorenylmethoxycarbonyl

g	gram(s)
GI	gastrointestianl
GST	glutathione S-transferase
h	hour(s)
HAT	human African trypanosomiasis
His-AGT	hexahistidine-tagged AGT
HOBt	*N*-hydroxybenzotriazole
HPLC	high performance liquid chromatography
Hz	hertz
IC_{50}	half-maximal inhibitory concentration
ICAT	isotope coded affinity tag
IgG	immunoglobulin G
Imi	imidazole
IPI	international protein index
ISCF	isobutyl chloroformate
J	coupling constant
kDa	kiloDalton
KDEL	Lys-Asp-Glu-Leu (amino acid sequence)
LAH	lithium aluminum hydride
LC-MS/M	liquid chromatography tandem mass spectrometry
m	multiplet, meter(s)
M	molar
MALDI	matrix-assisted laser desorption/ionization
m-CPBA	*m*-chloroperbenzoic acid
Me	methyl
mg	milligram(s)
MHz	megahertz
μL	microliter(s)
μm	micrometer(s)

min	minute(s)
mL	milliliter(s)
mmol	millimole(s)
mM	millimolar
mol	mole(s)
m/z	mass-to-charge ratio
NHS	*N*-hydroxysuccinimide
NK1R	neurokinin 1 receptor
nM	nanomolar
NMM	*N*-methylmorpholine
NMR	nuclear magnetic resonance
p	para
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PBST	phosphate buffered saline with Tween-20
PCF	procyclic form
pH	hydrogen ion concentration
Ph	phenyl
PCR	polymerase chain reaction
PLS	protein localization sequence
ppm	parts per million
psi	pounds per square inch
PTM	post-translational modification
PVDF	polyvinyl difluoride
q	quartet
Q-TOF	quadropole-time-of-flight
RNA	ribonucleic acid
RNAi	RNA interference
s	singlet

sat	saturated
SAR	structure-activity relationship
SDS	sodium dodecyl sulfate
t	triplet
TAS-F	tris(dimethylamino)sulfonium difluorotrimethylsilicate
TBS	*tert*-butyldimethylsilyl
TBTA	tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine
tBu	*tert*-butyl
TCEP	tris(2-carboxyethyl) phosphine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
THL	tetrahydrolipstatin
TLC	thin layer chromatography
TMAL	tandem Mukaiyama-aldol lactonization
TMS	tetramethylsilane
TOF	time of flight
Ts	toluenesulfonic acid
UV	ultraviolet
VS	vinyl sulfone
w/v	weight to volume ratio
w/w	weight to weight ratio
XTT	2,3-bis-(2-methoxy-4-nitro-5-sulfophenyl)-2*H*-tetrazolium-5-carboxanilide
Z	benzyloxycarbonyl

List of Twenty Natural Amino Acids

Single Letter Code	*Three Letter Code*	*Full Name*
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartate
E	Glu	Glutamate
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

List of Figures

Chapter 4

Chapter 5

List of Schemes

List of Tables

Chapter 1

Introduction

1.1 Drug Target Identification – An Overview

Assigning the cellular target(s) of bioactive small molecules (SMs), whether the compounds are discovered by cell-based phenotypic or target-based screens of chemical libraries, remains an ongoing challenge, as testified to by the fact that the primary targets of various clinically approved drugs remain unknown. The ability of accurately and thoroughly determining of SM–target interaction profiles as well as mapping of metabolic and signaling pathways on the proteomic scale would therefore be more illuminating, as it could provide invaluable biological insights for a drug candidate both by understanding the primary mechanism-of-action, and at the same time, side effects due to unexpected "off-target" interactions at a very early stage of drug development, which should help to reduce the attrition rate in development. In many cases such a capability could find new potential therapeutic value for an established drug as well as it could also offer strong clues for compound optimization in order to maximize the therapeutic potential and minimize potential cellular toxicity of a drug. "Target ID" has been approached using a variety of genetic, computational ands biochemical methods, such as genome-wide genetic assays, expression profiling, bioinformatics and affinity chromatography.[1] While all of these methods can provide valuable information on the mode of action of drug candidates, each method has distinct advantages and disadvantages, for example, genome-wide suppression screens have not yet been performed in mammalian cells because analogous genome-wide heterozygous cell line collections do not exist; the same is true for various modifications of the yeast two-hybrid method.

To date, affinity chromatography combined with identification by mass spectrometry remains the most popular and straightforward method used to isolate the protein target(s).[2] Examples (Figure 1.1) include several key molecular targets, such as trapoxin (known as K-trap Affinity Matrix), FK-506, (*R*)-roscovitine ((*S*)-roscovitine used as negative control), withaferin A, (-)-diazonamide A (*seco*-analog used as negative control), stauprimide, QS11, thalidomide, pateamine A, ICG-001 and so on.[3] In this method, a chemical derivative of the bioactive small molecule is covalently immobilized on a solid matrix support (e.g., agarose, sepharose or magnetic beads). Then, the SM–linked beads are incubated with protein extracts, followed by extensive washing to remove nonspecifically bound proteins. Finally, the binding proteins are eluted with excess free drug or under denaturing conditions. Small molecule-immobilized chromatography is certainly a very useful tool to unravel unknown protein targets of natural products. However, this method typically involves a solid support that can only capture potential protein target(s) in vitro but not in living system, especially given the fact that certain SM–protein interactions may require conditions in the living systems (i.e., native cellular environments) that are not preserved in protein extracts.[4] Moreover, it is often challenging to prepare SM affinity reagents that retain the desired biological activity. To address these issues, a new chemical proteomics strategy, called activity-based protein profiling (ABPP), which makes use of small molecules that can covalently attach to catalytic residues in an enzyme active site, has recently been successfully applied to the elucidation of the cellular targets of bioactive small molecules and natural products. Several excellent

reviews[5] have been published recently in the field of ABPP describing the variety of probes and breadth of applications of this technology. Because this field is covered in detail elsewhere, only the basic concepts and advances in ABPP with a particular emphasis on target identification as well as in vivo imaging will be discussed in the following sections.

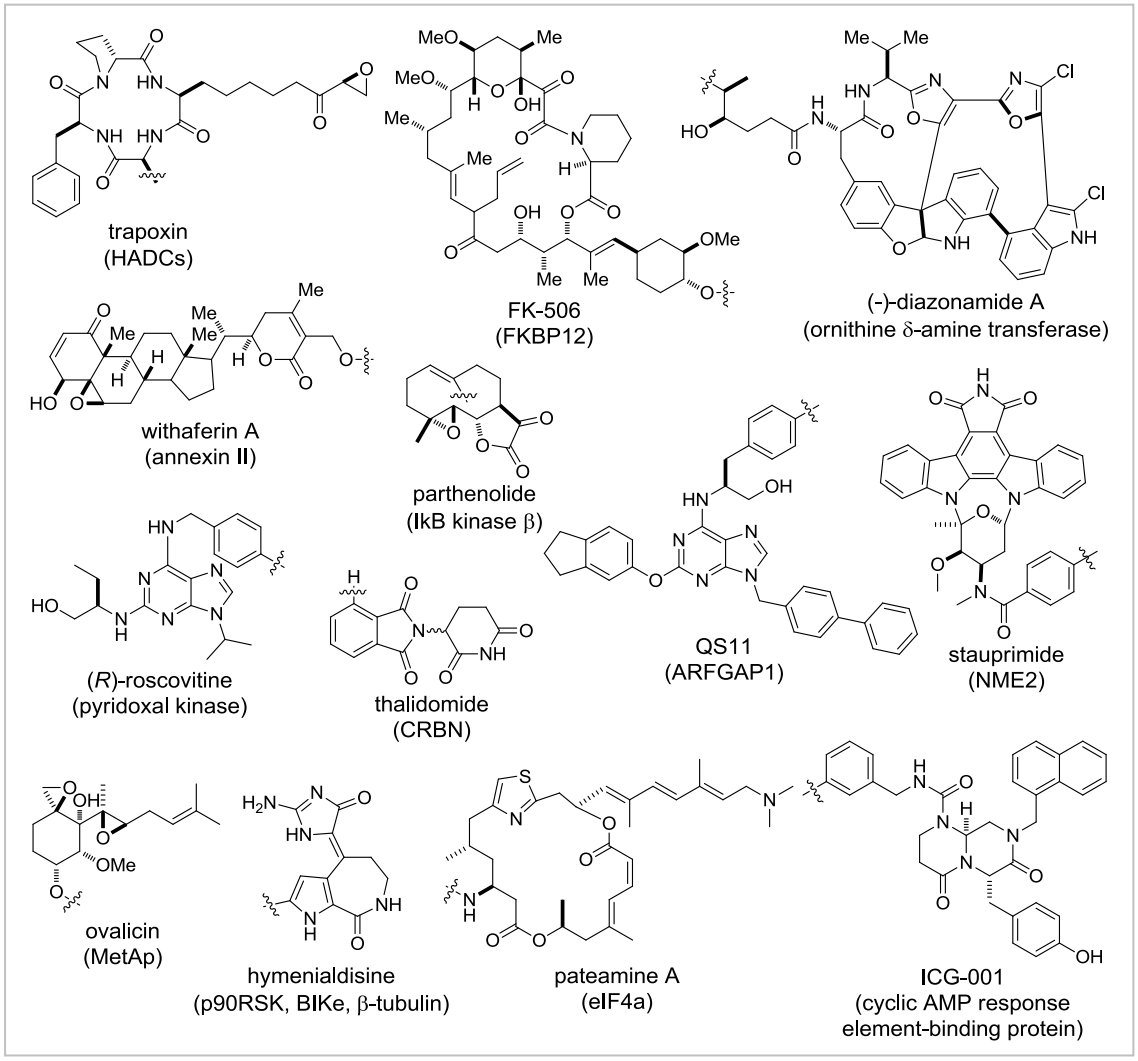


Figure 1.1 Representative structures of molecules whose target proteins (in brackets) were identified using affinity chromatography. The compounds of interest are modified and immobilized onto either a solid support or a biotin moiety (indicated by wavy bonds).

1.2 Activity-Based Protein Profiling (ABPP)

1.2.1 Introduction

The completion of the human genome sequencing project[6] has provided a wealth of new information that have shifted the focus of new biotechnologies from those that facilitate the characterization of single genes or proteins to methods that enable the rapid monitoring of all possible transcripts (so called genomics) or proteins (so called proteomics) contained in a living organism. Proteomics is the study of the proteome (defined as the entire complement of proteins expressed by a cell, tissue, or organism) and involves the technology used to identify and characterize the expression pattern and function of protein on a global scale.[7] In human, such information is crucial in order to design better diagnostic and therapeutic strategies.

Conventional proteomics approaches that heavily rely on two-dimensional gel electrophoresis (2DE) encounter difficulty detecting important fractions of the proteome, including membrane-associated and low-abundance proteins.[8] Therefore, over the years, several gel-free proteomics techniques have been developed to the use of isotopic tagging of proteins (ICAT) or peptides (iTRAQ) or metabolic incorporation of isotopically tagged amino acids (SILAC),[9] which enables the comparative analysis of protein expression levels by liquid chromatograph-tandem MS (LC-MS/MS). LC-based strategies have overcome some of the resolution problems of 2DE, providing improved access to proteins of low-abundance; however, these techniques, like 2DE, are restricted to detecting changes in protein abundance (part A of Figure 1.2). However, the abundance of a protein is not necessarily directly

coupled to its activity. Moreover, numerous posttranslational forms of protein regulation, including those governed by protein–protein interactions, remain undetected.[10] To address these limitations, a chemical proteomics strategy referred to as activity-based protein profiling (ABPP) has been introduced that makes use of synthetic small molecules that can be used to covalently modify a set of related enzymes in an activity dependent manner and subsequently allow their purification and/or identification. Therefore, such an approach could potentially distinguish active enzymes from their inactive zymogens or inhibitor.

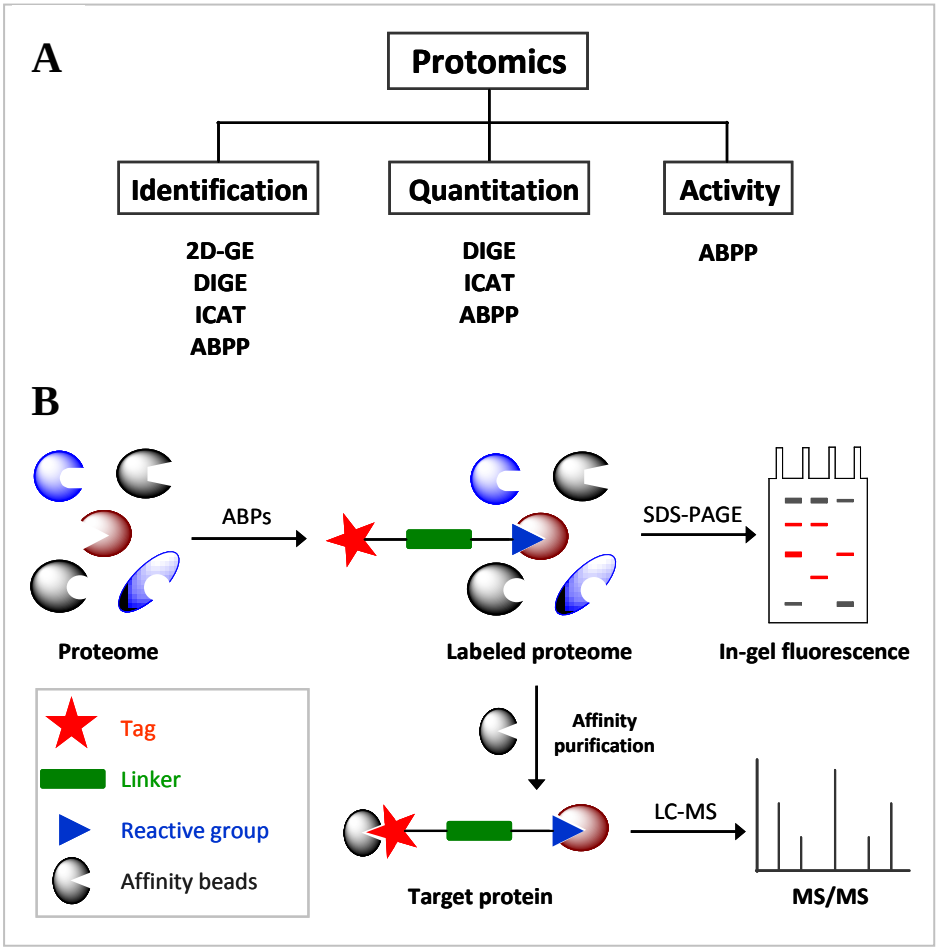


Figure 1.2 (A) Comparison of conventional proteomic approaches and ABPP. (B) Experimental workflow in a classical ABPP.

1.2.2 Activity-Based Probe Design

The small molecules used in ABPP, termed activity-based probes (ABPs),[5] generally consist of three basic elements: 1) a warhead (or an electrophilic trap) that can covalently attach to the active site of an enzyme, 2) a flexible linker that connects with 3) a tag that can be use for visualization (e.g. fluorophore, isotope), isolation and identification (e.g., biotin). In a typical ABPP proteomic experiment (part B of Figure 1.2), a proteome is labeled by ABP and probe-labeled proteins detected by either in-gel fluorescence scanning (for fluorophore-conjugated probes) or avidin enrichment, trypsin digestion and liquid chromatograpy–mass spectrometry (LC–MS) analysis (for biotinylated probes).

ABPs can be classified as either mechanism-based or affinity-based, depending on the mode by which the covalent linkage with the enzyme is established. Mechanism-based ABPs (also known as a directed approach), which are mainly based on irreversible enzyme inhibitors (or suicide inhibitors), use electrophilic warheads that specifically react with catalytic residues in the enzyme's active site.[11] Examples[12] include fluorophosphonate (FP) which reacts highly specifically with the enzymes of the serine hydrolase family, epoxide, diazomethyl ketone, acyloxymethyl ketone (AOMK), and α-haloketones showing a clear preference for cysteine proteases (Figure 1.3). Such an ABP could provide a fundamental understanding of the related enzyme's catalytic mechanism, substrate and/or inhibitor preference as well as functional roles. As noted, mechanism-based ABP is only applicable to targets that

possess a nucleophilic active site residue (mainly Ser, Cys or Lys) susceptible to covalent labeling by an electrophile. When this is lacking, three strategies have been developed for the design of affinity-based ABPs: (i) a photoreactive cross-linker (e.g., benzophenone, diazirine, aryl azide), which is conjugated to the probe that tightly binds within an active site, results in modification of potentially non-catalytic residues upon irradiation with UV light, (ii) a masked electrophile (also refered as to a latent electrophilic probe) that becomes activated upon cleavage by the enzyme, which means these reagents serve as bona fide enzyme substrates and only become into reactive probes after turnover. Examples include tyrosine phosphatase, glycosidase, and protease quinone methide probes, [13] and (iii) a novel ligand-directed tosyl (LDT) chemistry that can site-specifically introduce synthetic probes to the surface of proteins with concomitant release of the affinity ligands.[14]

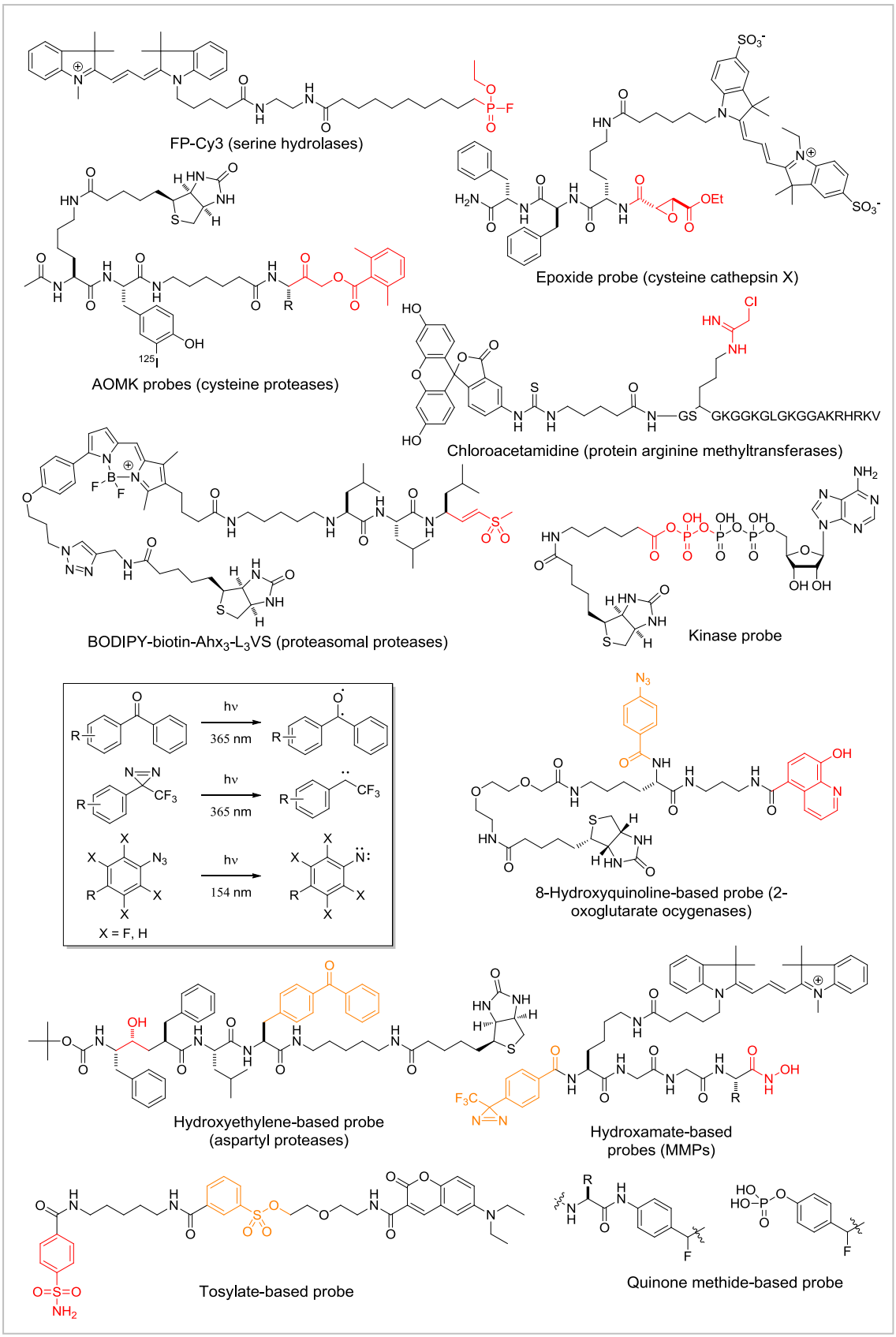


Figure 1.3 Structures of probes based on reactive groups (colored in red), photoreactive groups (colored in orange) and quinone methide or tosyl chemistry.

1.2.3 Bioorthogonal Chemistry in ABPP

A drawback encountered in classic ABPP, however, is the finding that the presence of the reporter entity generally diminishes or even precludes cellular uptake and distribution of probes, necessitating cell lysis prior to the labeling experiment. As mentioned before, certain SM–protein interactions may require conditions in the living systems that are not preserved in cell lysates. To overcome this shortcoming, several small latent ligation groups (mainly an alkyne or azide group) for a reporter tag have been introduced, with the idea that such constructs would likely to be cell-permeable. After labeling in living cells, attachment of the reporter tag may then be performed ex vivo through highly specific bioorthogonal ligation chemistry that show low reactivity towards other biomolecules, and can be easily conducted in aqueous media, such as the Staudinger ligation[15](Figure 1.4), and the copper catalyzed or strain-promoted azide alkyne cycloadditions (often referred to as “click” chemistry)[16]. However, the Staudinger ligation, initially developed to probe protein glycosylation events on cell surfaces, suffers from slow reaction kinetics and background phosphine oxidiation. CuAAC is a rapid reaction, but the Cu(I) catalyst is toxic to live cells, limiting its applications. The strain-promoted cycloaddition with fluorinated cyclooctyne reagents is currently the optimal bioorthogonal reaction for the azide, but the cyclooctynes are difficult to synthesize and still exhibit reactivity only one order of magnitude faster than the Staudinger ligation. These limitations have therefore promoted the recent expansion of the arsenal of other bioorthogonal chemical reactions used to label proteins in living systems, such as inverse-electron

demand Diels-Alder reactions between 1,2,4,5-tetrazines and strained dienophiles such as norbornene, cyclooctyne, and *trans*-cyclooctene.[17] The main advantages of this reaction are the following: there is no requirement of a catalyst; the bimolecular rate constants for reaction can be very high ($>10^3$ M^{-1} s^{-1}) with appropriate choice of tetrazine and strained dienophile. To date, the main limitation to the reaction compared to conventional azide/alkyne or Staudinger ligations is that the strained alkenes are rather large, perhaps preventing their metabolic incorporation. It should be noted that although alkenes are naturally found in lipids, fatty acids, cofactors and other natural products, most of other new bioorthogonal reactions employ activated alkenes that could be able to react selectively in the presence of endogenous olefins.

With these successful strategies, "tag-free" probes carrying small surrogate tags have been developed for various classes of enzymes with enhanced selectivity and improved pharmacological properties for selective labeling of protein targets in living cells (Figure 1.5). Recent examples include serine phospholipases, protein phosphatase methylesterase-1 (e.g., aza-β-lactams), transglutaminase-2 (e.g., 3-bromo-4,5-dihydroisoxazoles), α-haloacetamidines for dimethylarginine dimethylaminohydrolase (DDAH) or protein arginine deiminase (PAD), cytochrome P450s (i.e., aryl alkynes are oxidized to reactive ketene intermediates that inactivate P450s by covalent adduction to the protein), 14-3-3 ξ (e.g., isothiocyanate as an affinity bait), RSK kinase (e.g., fmk-pa, *f*luoro*m*ethyl *k*etone-*p*ropargyl*a*mine), and Abelson (Abl) tyrosine kinase (e.g., a derivative of Imatinib) as well as histone deacetylases (HDACs), and methionine aminopeptidase.[18] Of note, the decoupling of

detection tags from chemical probes alters and often improves the specificity of protein labeling compared to directly modified chemical probes, which is readily apparent with cysteine proteases[33], kinases[18h] and methionine aminopeptidases.[18k]

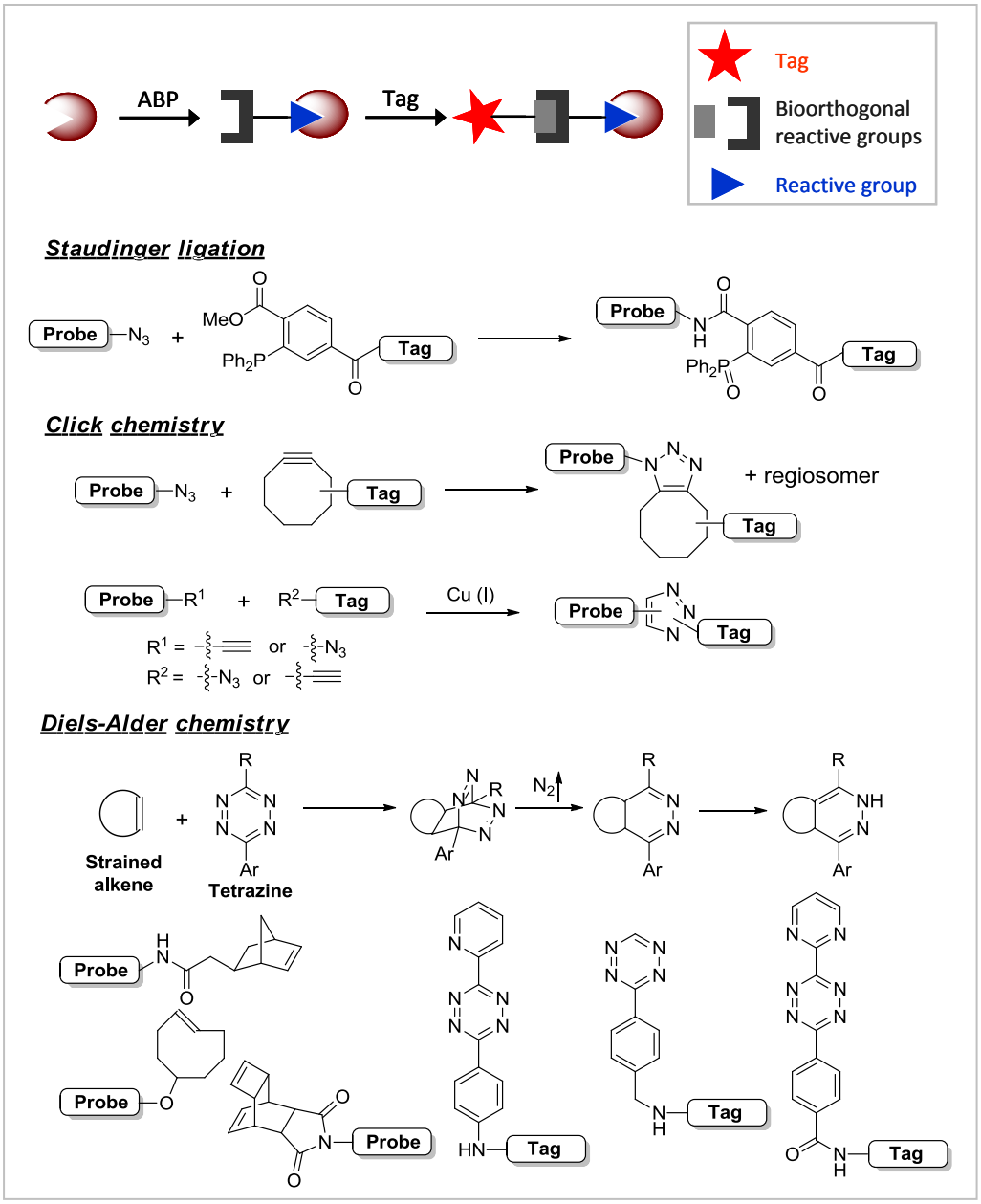


Figure 1.4 Two-step ABPP labeling assisted by the bioorthogonal reactions, such as the Staudinger ligation, click chemistry, and inverse-electron Diels-Alder reaction.

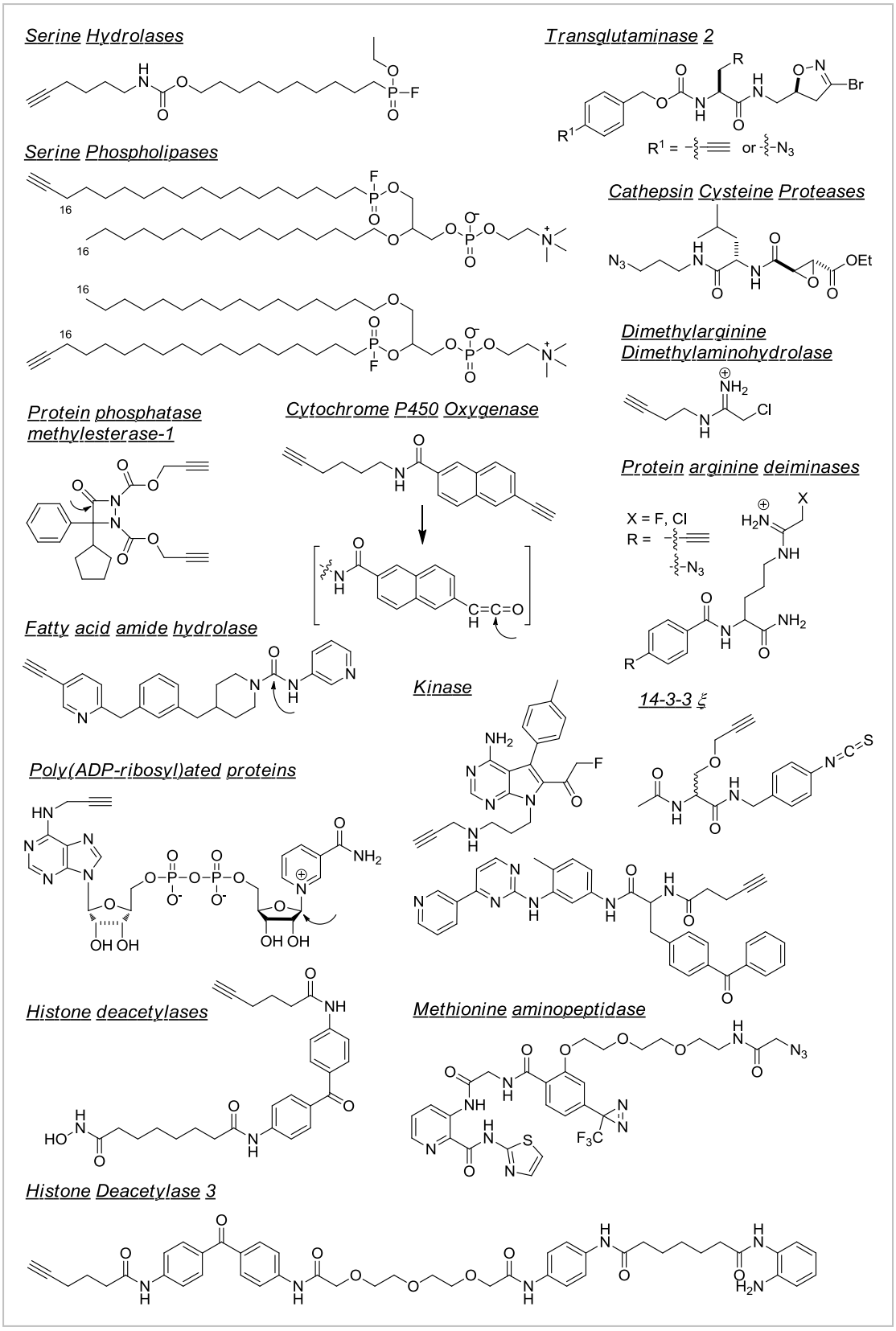


Figure 1.5 "Tag-free" probes for two-step ABPP labeling ordered by enzyme or enzyme classes they target.

Although the aforementioned directed-ABPs derived from known inhibitors are widely employed for targeting various classes of enzymes, non-directed ABPs are particularly useful for enzyme classes that do not have well-defined cognate affinity labels. A combinatorial or non-directed approach usually employs libraries of probes to identify "specific" protein-labeling events. Examples include sulfonate esters,[19] and α-chloroacetamides.[20] In addition to well-defined electrophiles (e.g. sulfonate ester, or α-chloroacetamide) serving as the warhead in non-directed approaches, reactive groups akin to those present in bioactive natural products like fumagillin, luminacin D and FR901464 have also been used to profile proteomes (Figure 1.6). Evans et al. used a 50-member natural product inspired spiroepoxide library for in situ profiling of mechanistic different enzyme classes and identified a compound,[4] MJE3 that inhibits breast cancer cell proliferation. Further analysis identified brain-type phosphoglycerate mutase-1 (PGAM-1), a key enzyme in glycolysis that converts 3-phosphoglycerate to 2-phosphoglycerate as cellular target for MJE3 and interestingly enough, PGAM-1 inhibition was observed only in intact cells, thereby implying PGAM1 as a regulator of cancer cell viability. These findings are of pharmacological significance since it provides straightforward ways to couple cellular phenotype with underlying biochemical mechanisms and further expands the utility of non-directed probes as invaluable tools to probe cellular functions in forward chemical genetic screens.

Similar approach was also applied to study the Hepatitis C Virus (HCV) replication by Pezacki and coworkers.[21] In this study, non-directed phenyl sulfonate

ester probe library with alkyne handle was used to selectively target a broad range of enzyme families that show differential activity during HCV replication. Most recently, Sieber group employing an ABP library of *trans*-β-lactones to different bacterial proteomes resulted in identification of over twenty enzymes with either a nucleophilic cysteine or serine residue in their active sites.[22] Among the identified enzymes, a key virulence regulator ClpP in *S. aureus* is of particular interest, as it is essential for the virulence of many pathogenic bacteria.

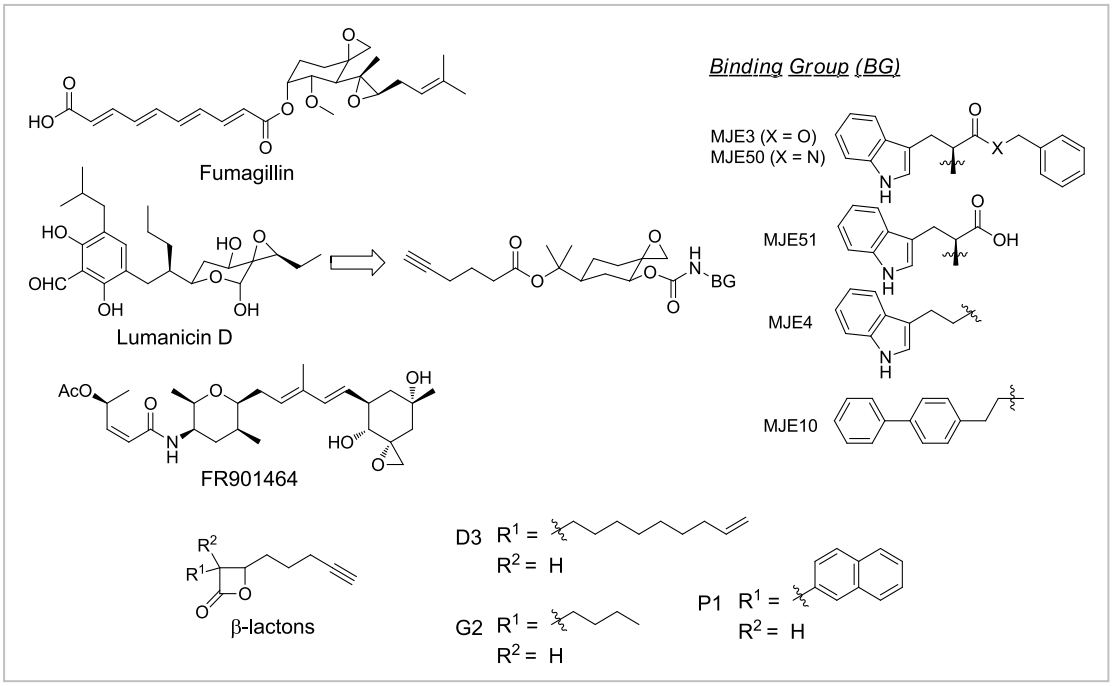


Figure 1.6 "Tag-free" non-directed ABPs inspired by natural products containing well-defined reactive groups.

Another major application of bioorthogonal chemistry in ABPP is linker systems that allow mild cleavage under conditions ideally orthogonal to functionalities present in the biological system. Biotinylated ABPs are often used for enrichment of captured enzymes, for instance, by pull-down with avidin-coated beads.

The main disadvantage of this approach, however, is the harsh denaturing conditions required to disrupt the biotin-avidin interation, as thus both endogenously biotinylated proteins and (denatured) avidin can contaminate the target proteins. To address this issue, several groups have developed linker systems that can be incorporated in a given ABP, or alternatively in a bioorthogonal reagent for two-step ABPP labeling, and that can be cleaved in a chemo-selective manner after affinity pull-down. Examples (Figure 1.7) include the disulfide, diazobenzene, bisaryl hydrazone, and the enzyme (e.g., TEV) cleavable or the photocleavable linker.[23] Clearly, synthetic chemistry has played a major role by providing additional orthogonal functional group and new cleavable linkers and improving the kinetics and selectivities of the reactions already at hand. Perhaps one way of doing that is to mine classic organic literatures.

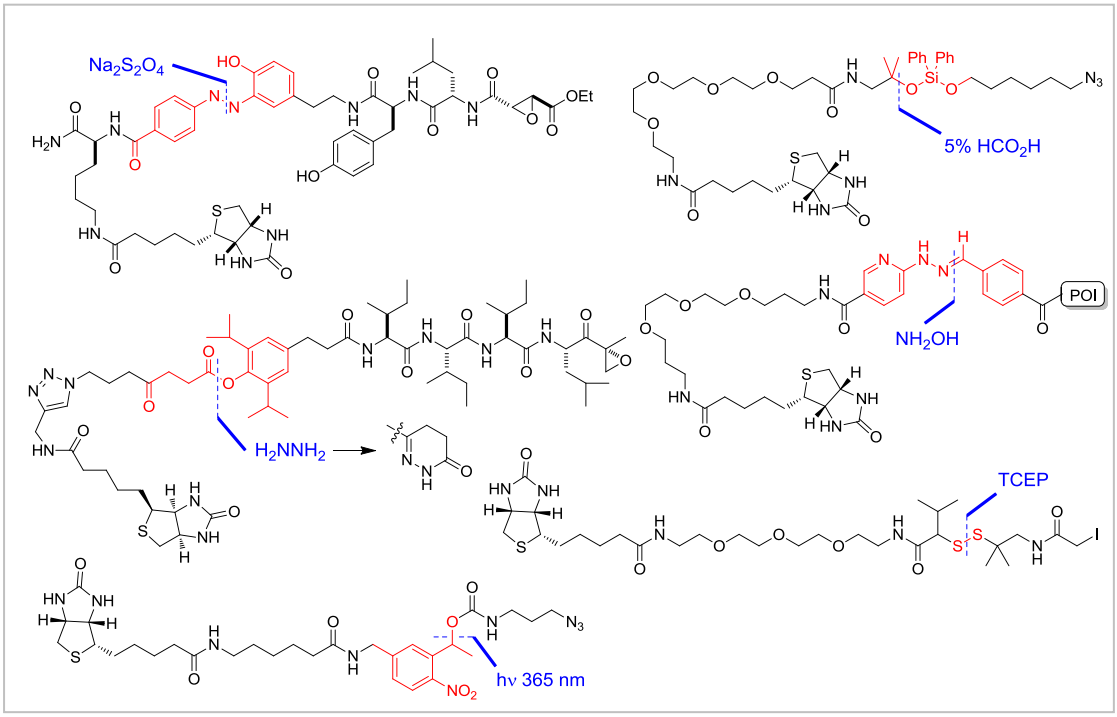


Figure 1.7 Cleavable linkers (colored in red) in ABPP that can be cleaved in a selective manner (colored in blue) after affinity pull-down.

1.2.4 Application of ABPP for Target ID of Natural Products

Natural products, and their derivatives and mimics, have rendered an enormous contribution to the treatment of diseases such as infentions and cancers over the past decades. Notable examples include antibiotics β-lactam penicillin G, the tetracyclines (e.g., tetracycline, chlortetracycline), the glycopeptides vancomycin, and the lipopetide daptomycin; Taxol (Paclitaxel), vinca alkaloids, and etoposide are plant-derived chemotherapeutic anti-cancer agents.[24] Natural products are biosynthesized or modified by proteins, and as such natural product have to bind them. However, the interaction of many bioactive natural products with their target proteins is often unknown. Recent advances in proteomics (e.g., chemical proteomics, expression proteomics), yeast three hybrid (Y3H) system, phage and mRNA display technologies have greatly accelerated target identification.[1b] In this section I overview the recent application of ABPP to elucidate the potential target(s) of a given natural product (Figure 1.8).

Wortmannin is a metabolite natural product produced by the fungus *Penicillium wortmannii* and has long been recognized as a highly specific inhibitor of phosphoinositide-3 kinase (PI-3K) enzyme superfamily. Wortmannin reacts in an irreversible manner with a critical lysine residue in the kinase domain of the PI- 3 kinases, and this lysine is conserved in the different PI-3 kinase and PIKK family members. Attachment of tags (e.g., BODIPY, rhodamine, or biotin) to the wortmannin scaffold has enabled both visualization and purification of target proteins in lysates and in living cells. Other than previously reported target PI-3 kinase and its related

kinases, Liu et al. identified Polo-like kinase 1 (PLK1) and Polo-like kinase 3 (PLK3), which were further confirmed by in vitro kinase assay with wortmannin.[25] Additionally, BODIPY-wortmannin, a cell permeable probe may provide a potentially tool for studying the biological activity of target kinases within living cells. Using similar approach appending a reporter group to natural product structure, many more ABPsbased on natural products have been developed and their target proteins have been identified. Examples include microcycins targeting serine/threonine protein phospharases in Jurkat cell lysates, cyclostreptin (-)-FR182877 targeting carboxylesterase-1 in the proteome derived from a variety of mouse tissues.[26a] Recently, Buey et al. found that cyclostreptin binds covalently to β-tubulin in cellular microtubules either with Thr220 (at the outer surface of a pore in the microtubule wall) or Asn228 (at the luminal paclitaxel site), but only Thr220 is modified when cyclostreptin interacts with unpolymerized tubulin.[26b] One further example is the showdomycin probe that can be used to identify the molecular targets in living Gram-positive and Gram-negative pathogenic bacteria, thereby explaining the antibacterial effects of showdomycin.[27]

Since HUN-7293 interacts reversibly with its targets, a photoreactive group (e.g., diazirine) has been incorporated into the HUN-7293 scaffold resulting in identification of the target protein, Sec61α, in endoplasmic reticulum (ER) microsome fraction.[28] Example of this type of affinity-based photo-cross-linking probe includes the vancomycin-based probe that revealed the specific labeling of two previously unknown targets, the bifunctional antolysin (ATL) in methicillin-resistant *S. aureus*

(MRSA) and a peptide ABC transporter (pABC) in vancomycin susceptible *E. faecalis* (VSE).[29] Staub et al. using probes derived from the naturally occurring antibiotics ampicillin, cephalosporin and aztreonam as well as some synthetic β-lactams to reveal a variety of penicillin binding proteins (PBPs) and additional enzymes, including a resistance-associated β-lactamase via tunning the core scaffold β-lactam ring.[30] Moreover, probes were also used to profile target enzymes in either antibiotic sensitive or resistant *S. aureus* strains (MRSA).

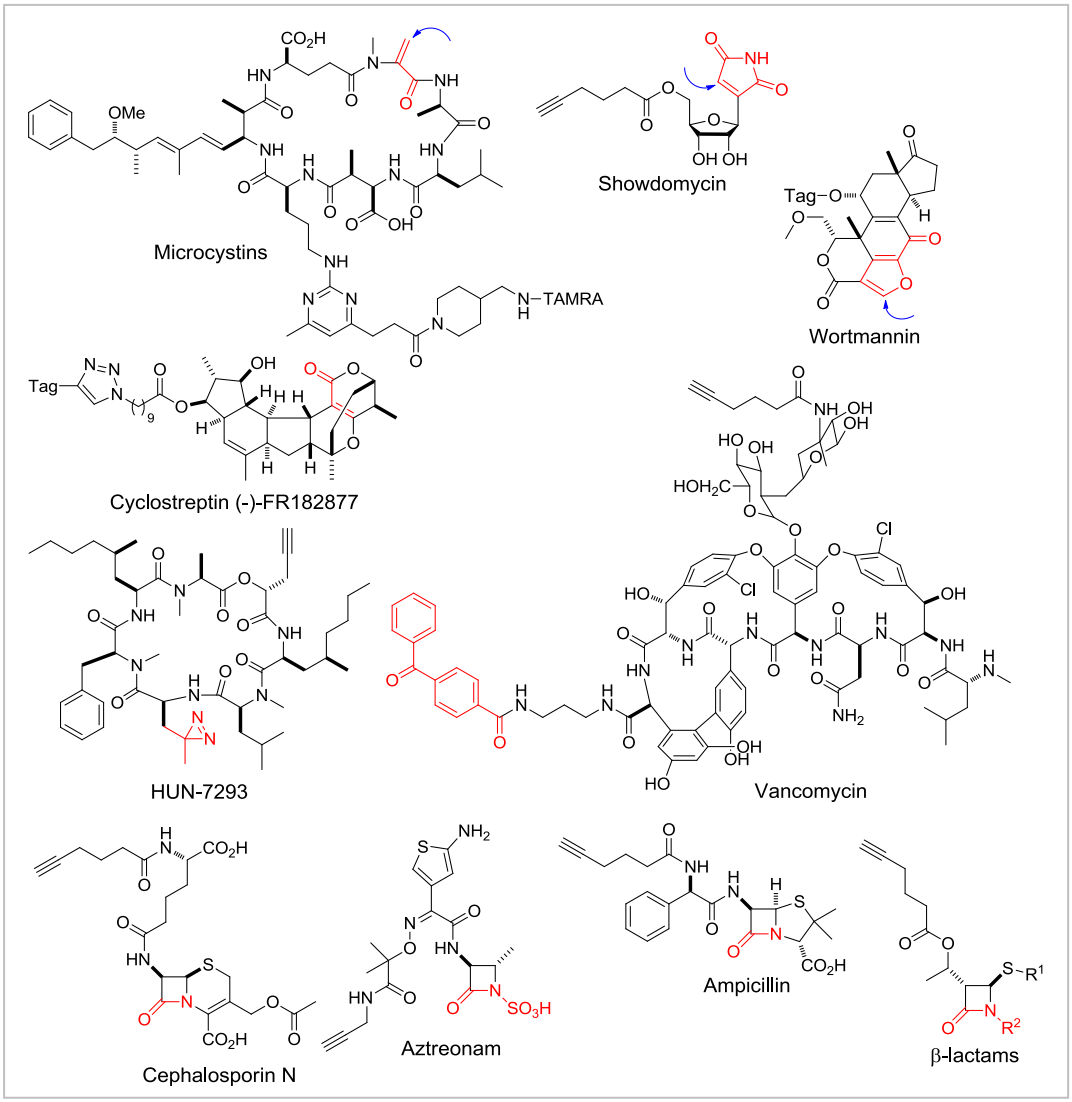


Figure 1.8 Examples of natural products with either a reactive group (arrow) or photoreactive group used as ABPs to identify their target proteins.

1.2.5 Application of ABPP for Imaging of Protein Activities

ABPs have also been adapted for demanding application such as monitoring of protein activity and revealing their spatial localization in real time in living systems. Several cell-permeable mechanism-based probes such as DCG-04, MGP140 and GB111 have been generated for visualization of cathepsin cysteine proteases in live cells and in vivo[31] (Figure 1.9). These probes contain either an epoxide (e.g., DCG-04, MGP140), or an acyloxymethylketone (AOMK) reactive electrophile (e.g., GB111) that covalently reacts with active site cysteine residue. Likewise, MV151 has been reported to label numerous active subunits of the proteasome and immunoproteasome in living cells. Additionally, administration of MV151 to mice allowed for in vivo labeling of proteasomes, which correlated with inhibition of proteasomal degradation in the affected tissues.[32] While the aforementioned cell-permeable probes have enabled the visualization of enzyme activity in live systems, these have limitations due to large fluorophore groups with their potential ability to alter the specificity of protein labeling or passive diffusion into cells with background fluorescence. To address this issue, the bioorthogonal reactions have been applied to facilitate imaging of protein activity. For instance, Hang et al. reported “tag-free” mechanism-based probe azido-E64 equipped with a small azide handle that can be conjugated with phosphine-biotin (p-biotin) via the Staudinger ligation, enabling the visualization active cathepsin B in *S. typhimurium*-infected primary macrophage and also allowing affinity enrichment of labeled proteins for MS-base proteomic analysis.[33] In addition to using bioorthogonal labeling reactions for visualizing

enzyme activities, Edgington et al. reported the AOMK-based caspase probe, tAB50-Cy5a containing a Tat peptide that makes use of multiple positively charged amino acids to carry attached cargo across membranes and has previously been used to increase cellular uptake.[34] The author used this probe to monitor the kinetics of apoptosis in multiple mouse models.

Recently, the development of quenched activity-based probes (qABPs) provides another alterative solution for reducing background fluorescence, whilst enabling live cell imaging of enzyme activity. Blum et al. reported cell-permeable AOMK-based probes that are modified with a fluorophore (e.g., BODIPY) and quencher (e.g., QSY7) such that covalent modification of an active cathepsin cysteine protease would result in a loss of the quenching group with the simultaneous recovery of the fluorescence ("turn-on").[35a] Furthermore, using quenched near-infra red ABPs (qNIRF-ABP) the authors were able to extend their approach to the in vivo non-invasive, whole-body imaging in mice bearing grafted tumours.[35b] Importantly the stable nature of the probes was useful for ex vivo biochemical identification of specific proteases as well as in monitoring small molecule inhibition of protease activity via fine-tuning of peptide specificity residues.

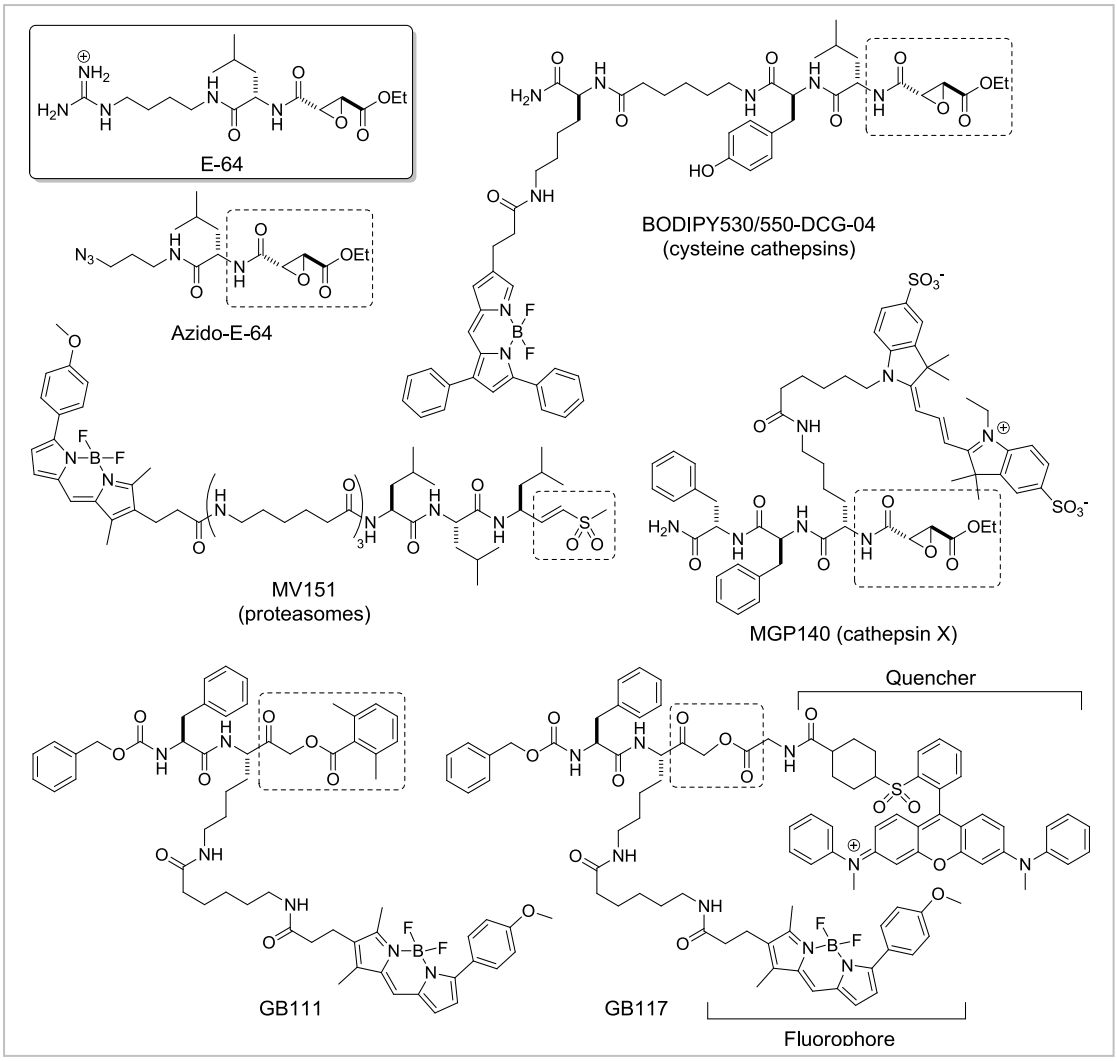


Figure 1.9 Examples of ABPs derive from suicide inhibitor motifs (see dashed boxes) that have been use to visualize their active target proteins in living systems.

While these qABPs have provided good sensitivity for detecting endogenous cysteine protease activity, one of the major limitations of the AOMK-based qABPs lies in that its target protein is deactivated upon labeling by qABP due to its suicide inhibitor nature. Thus, there is no signal amplification of the target enzymes, thereby limiting its sensitivity when screening low-abundance targets. Most recently, Hu et al. reported a new class of qABPs which are based on quinine methide chemistry for imaging of enzyme activity in live cells (Figure 1.10). Importantly, these qABPs

allow amplification of the target signal in situ, as they do not inactivate the enzyme target. Additionally, the authors also have also incorporated two-photon dyes into modular probe design, enabling firstly activity-based, fluorogenic two-photon imaging of enzyme activity.[36]

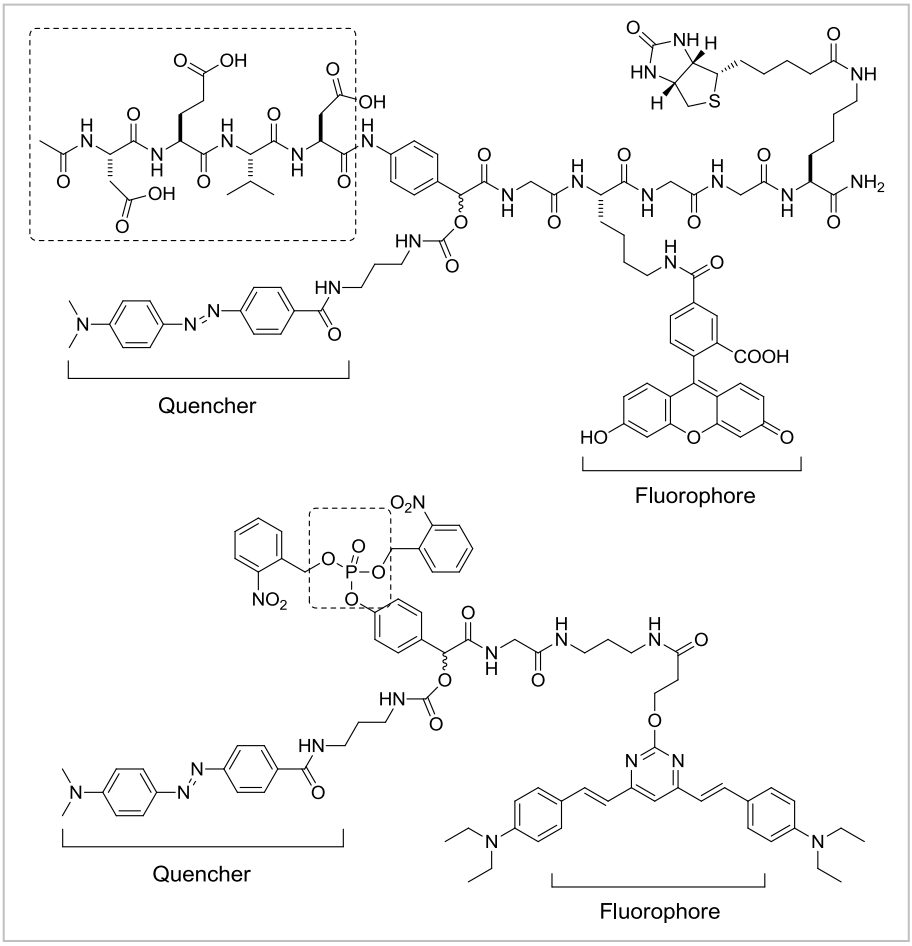


Figure 1.10 Structures of qABPs based on quinine methide chemistry that consist of the motifs recognized by active enzymes (see dashed boxes).

Collectively, all these examples bode well for the future, in which creative thinking from an organic chemistry background will continue to provide new tools and methods to gain insights into vital biological processes (so-called "chemical biology").

Chapter 2

Activity-Based Proteome Profiling of Potential Cellular Targets of Orlistat—An FDA-Approved Drug

Portions of the work presented here are taken from:

P.-Y. Yang, K. Liu, M. H. Ngai, M. J. Lear, M. R. Wenk, and S. Q. Yao, "Activity-Based Proteome Profiling of Potential Cellular Targets of Orlistat - an FDA-Approved Drug with Anti-Tumor Activities." *J. Am. Chem. Soc.* **2010**, *132*, 656-666. (Highlighted by Faculty of 1000 Biology)

Abstract

Orlistat™, or tetrahydrolipstatin (THL), is an FDA-approved antiobesity drug with potential antitumor activities. Cellular off-targets and potential side effects of Orlistat in cancer therapies, however, have not been extensively explored thus far. In this study, we report the total of synthesis of THL-like protein-reactive probes, in which extremely conservative modifications (i.e., an alkyne handle) were introduced in the parental THL structure to maintain the native biological properties of Orlistat, while providing the necessary functionality for target identification via the bio-orthogonal click chemistry. With these natural product-like, cell-permeable probes, we were able to demonstrate, for the first time, this chemical proteomic approach is suitable for the identification of previously unknown cellular targets of Orlistat. In addition to the expected fatty acid synthase (FAS), we identified a total of eight new targets, some of which were further validated by experiments including Western blotting, recombinant protein expression, and site-directed mutagenesis. Our findings have important implications in the consideration of Orlistat as a potential anticancer drug at its early stages of development for cancer therapy. Our strategy should be broadly useful for off-target identification against quite a number of existing drugs and/or candidates, which are also covalent modifiers of their biological targets.

2.1 Introduction

Drug discovery is a long and costly process, yet most drugs have side effects, ranging from simple nuisances to life-threatening complications.[37] Unanticipated effects of a drug, often revealed either during clinical trials or sometimes after the drug enters the market, could lead to termination of a drug development program/recall of the drug, or, in some rare cases where the effects are beneficial, new drug applications.[37b-c] Therefore one of the most critical steps in the drug discovery process is the effective identification of the so-called off-targets and anticipation of their potential side effects a priori.

Orlistat™, a saturated derivative of naturally occuring lipstatin, which was isolated from the bacterium *Streptomyces toxytricini*, is administered orally as an anti-obesity drug. It works primarily on pancreatic and gastric lipases within the gastrointestinal (GI) tract.[38] Pancreatic lipase is the key enzyme of dietary triglyceride absorption by hydrolyzing triglycerides into diglycerides and subsequently into monoglycerides and free fatty acids. Orlistat forms a covalent bond with the active site serine residue of pancreatic lipase, thereby blocking the hydrolysis of triglycerides from nutrition into free fatty acids (Figure 2.1). Recently, Orlistat was found to inhibit the thioesterase domain of fatty acid synthase (FAS), an enzyme essential for the growth of cancer cells, but not normal cells.[39,40] By effectively blocking the cellular FAS activity, Orlistat induces endoplasmic reticulum stress in tumor cells, inhibits endothelial cell proliferation and angiogenesis, and consequently delays tumor progression on a variety cancer cells, including prostate, breast, ovary,

and melanoma cancer cells.[39] As a result, this compound (as well as other orlistat-like analogs with improved potency and bioavailability[41]) is being pursued as a promising anti-cancer drug. Cellular off-targets and potential side effects of orlistat in cancer therapies, however, have not been extensively explored thus far.[42] Our long-term research goals focus on developing novel chemical proteomic strategies that enable large-scale studies of therapeutically relevant enzymes, as well as small molecules (i.e. potential drug candidates) that can modulate these enzymes' cellular activities.[43] In the current study, we set out to look for new cellular targets, including off-targets, of orlistat at its early stages of development for cancer therapy. Herein, we report, for the first time, by using a novel chemical proteomic approach, the identification and putative validation of several previously unknown cellular targets of orlistat.

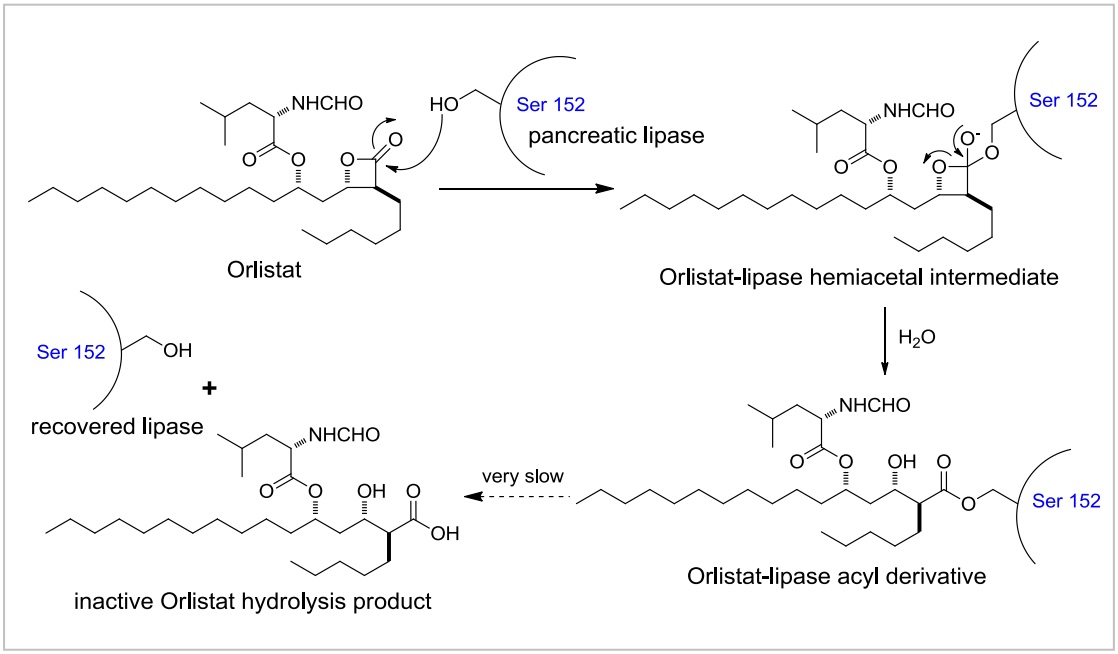


Figure 2.1. Schematic representation of inhibition of pancreatic lipase by Orlistat.

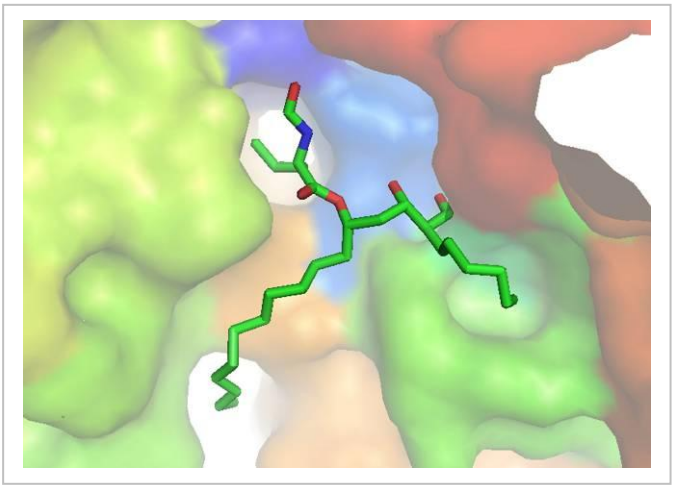


Figure 2.2. Surface representation of the complex (the thioesterase domain of FAS inhibited by Orlistat) highlighting the different binding channels and pockets (PDB entry 2PX6, image generated with PyMOL).

2.2 Results and Discussion

2.2.1 Design of Orlistat-like Probes

Our strategy is based on the well-established activity-based protein profiling (ABPP) approach,[5] by making use of THL-like protein-reactive probes **2-1**, **2-2** and **2-3** (i.e. THL-R, THL-L and THL-T, respectively, in Figure 2.3). We took advantage of several key properties known to THL in the design of our probes: (1) THL (being derived from a natural product) is cell-permeable, making our probes applicable for direct *in situ* cell-based screening; (2) THL reacts with its known cellular targets via a covalent reaction through its reactive β-lactone moiety and the nucleophilic active-site residue (e.g. Ser^{152} and Ser^{2308} in pancreatic lipase and FAS[39b], respectively) of the target protein, resulting in the formation of an isolatable protein/THL complex; (3) previous minor structural modifications at either the 16-carbon or 6-carbon aliphatic chains of THL did not significantly alter its native biological activities.[41b] Accordingly, probes **2-1**, and **2-2**, in which a terminal C-C triple bond was introduced into THL to replace the terminal C-C single bond of the aliphatic chains, were synthesized. We also synthesized probe **2-3** by substituting the CHO group in the *N*-formyl-L-leucine moiety of THL with a propiolic acyl group (Figure 2.3; shaded in pink). These extremely conservative modifications of introducing an alkyne handle in the parental THL structure were aimed at maintaining the native biological properties of Orlistat, whilst providing the necessary functionality for identification and characterization (i.e. imaging) of previously unknown cellular targets by downstream conjugation of the protein/probe complex to reporter tags via the bio-orthogonal click chemistry.[16]

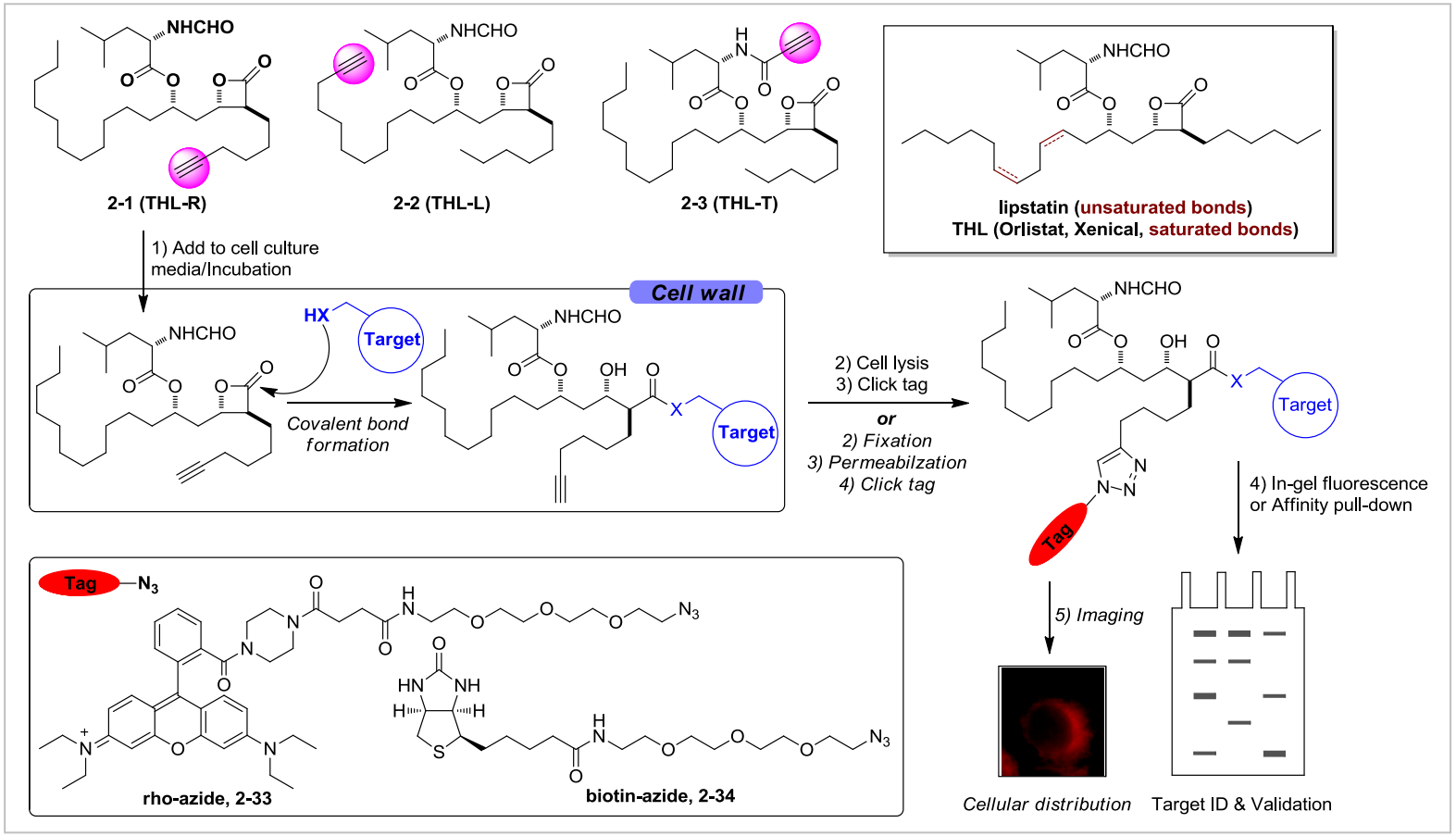
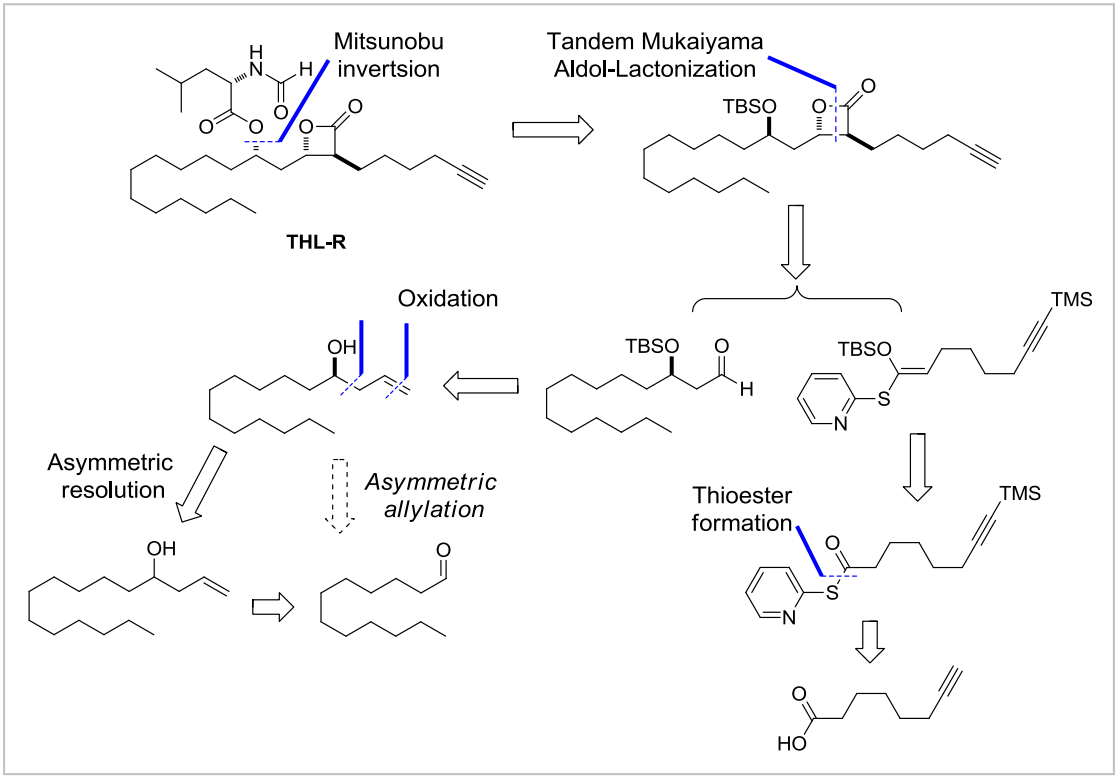


Figure 2.3. Overall strategy for cell-based proteome profiling of potential Orlistat targets using alkyne-containing, cell-permeable THL analogues, **2-1**, **2-2** & **2-3**. The alkyne handles in the probes are shaded (pink circle).

2.2.2 Retrosynthetic Analysis

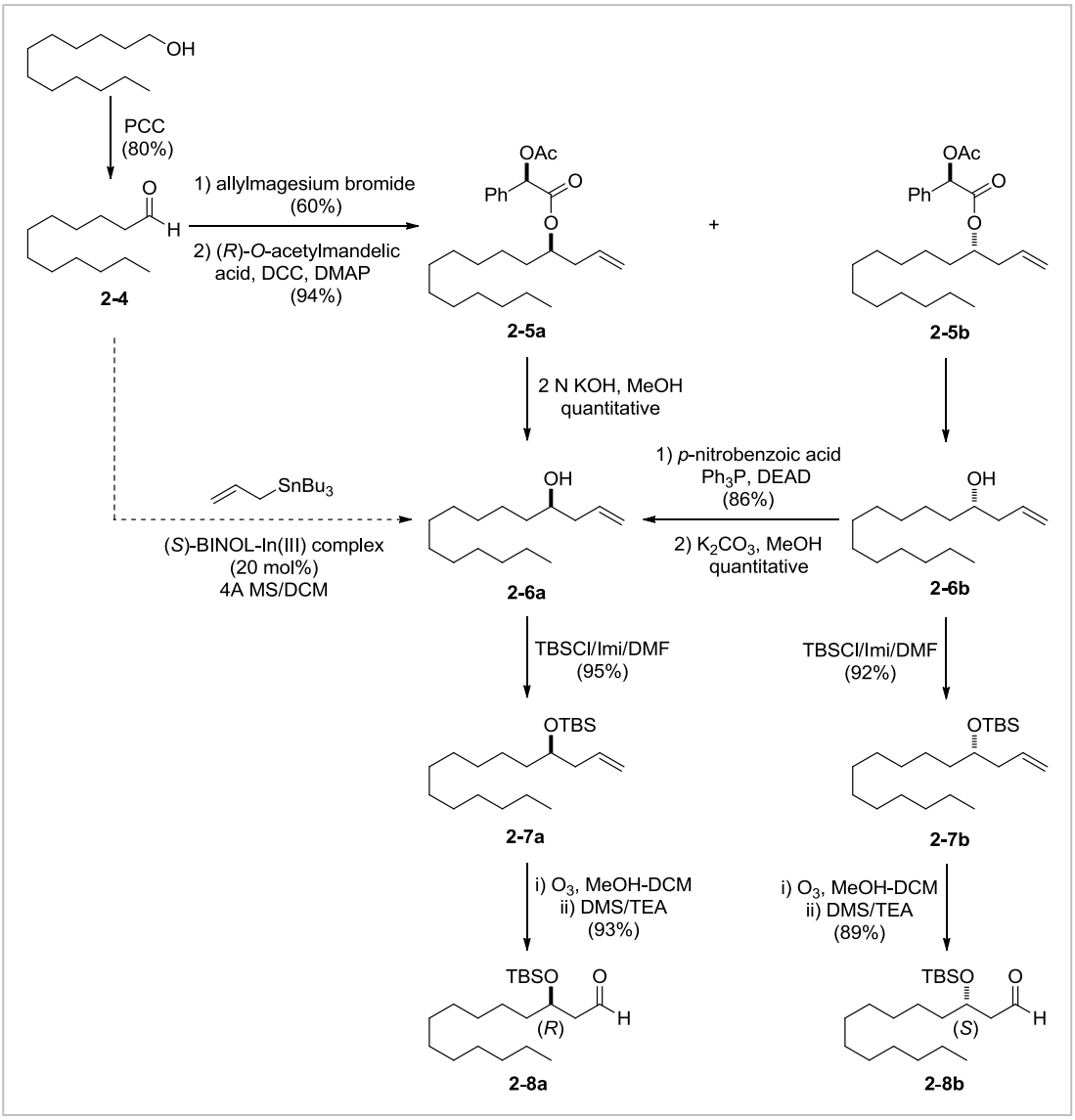
Numerous strategies for the total synthesis of THL have been reported.[41] Our retrosynthetic analysis is outlined in Scheme 2.1. THL-R can be readily derived from β-lactone intermediate via simple silyl deprotection and Mitsunobu esterification. *Trans*-β-lactone in turn could be obtained using a tandem Mukaiyama aldol-lactonization (TMAL) reaction of optically active TBS-protected β-hydroxy aldehyde and thiopyridyl ketene acetal as the key step. The TBS-protected β-hydroxy aldehyde can be produced from homoallylic alcohol in a straightforward 2-step sequence of TBS-protection and ozonolysis/oxidation. Chiral homoallylic alcohol would be prepared from asymmetric allylation of aldehyde or resolution of racemic homoallylic alcohol, which in turn is avaible through allylation of aldehyde.



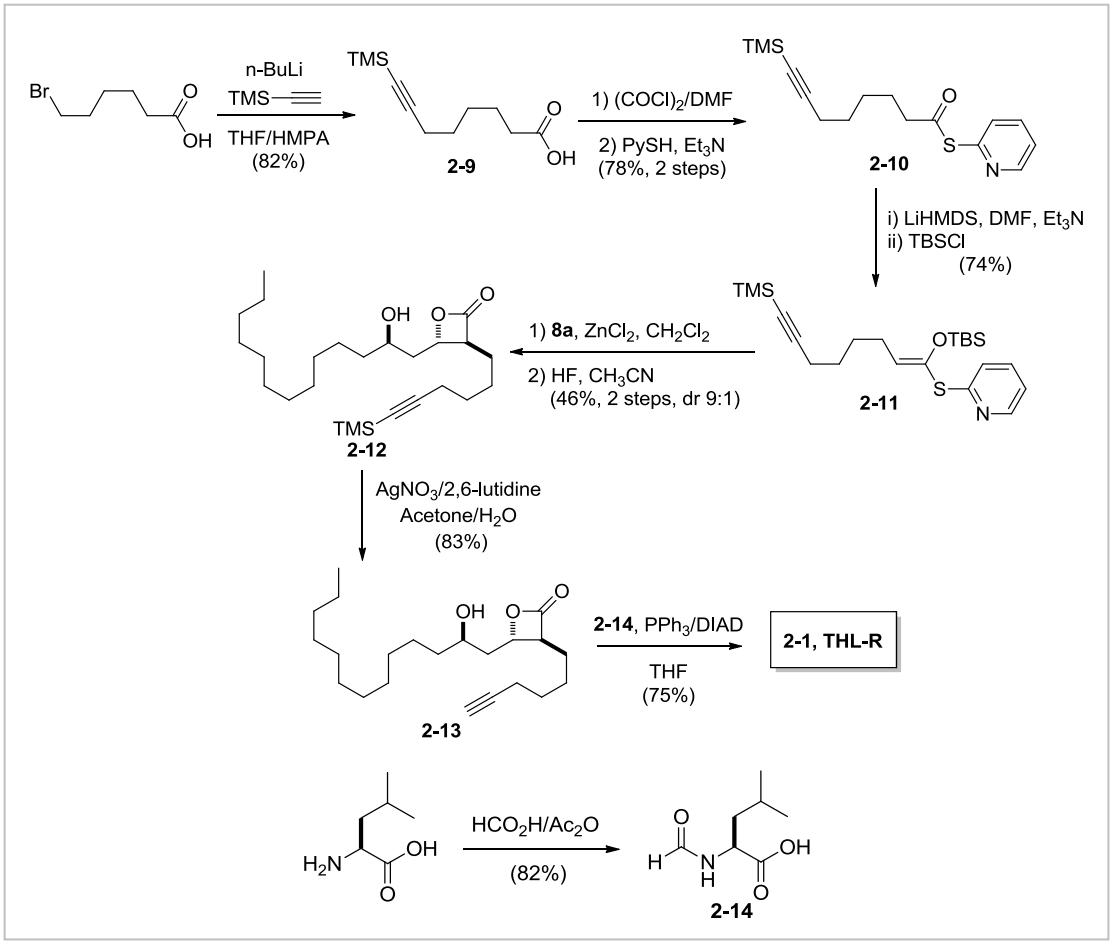
Scheme 2.1. Retrosynthetic analysis of THL-R (**2-1**).

2.2.3 Synthesi of Orlistat-like Probes

Synthesis of optically pure aldehyde **2-8a** relied on the asymmetric resolution of racemic allylic alcohols (Scheme 2.2). Briefly, lauraldehyde **2-4**, which was prepared from dodecan-1-ol by PCC oxidation and treated with allylmagnesium bromide followed by esterification with *O*-acetyl-D-mandelic acid gave the diastereomeric esters **2-5a** and **2-5b** which could be easily separated by chromatography. Deacylation of **2-5a** led to the optically pure homoallylic alcohol **2-6a**. The diastereomeric **2-5b** could be converted to the desired **2-6a** by desterification and inversion of configuration via a modified Mitsunobu reaction. The alcohol **2-6a** was then TBS-protected and the olefin oxidatively cleaved with ozonolysis to afford **2-8a**. The key $ZnCl_2$-mediated, tandem Mukaiyama aldol-lactonization (TMAL) between the β-hydroxy aldehyde **2-8a** and the thiopyridyl ketene acetal **2-11** (prepared in three steps involving alkylation of 6-bromohexanoic acid with ethynyltrimethylsilane, generation of thiopyridyl ester **2-10** and an enolization/silylation) giving the desired β-lactone as a mixture of diastereomers (~9:1 anti/syn) with complete selectivity for the *trans*-β-lactone. Following *O*-desilylation and silica gel separation, the enantiomerically pure γ-hydroxy β-lactone **2-12** was isolated. The *trans* configuration of the β-lactone core was unambiguously confirmed by coupling constant analysis ($J_{H2,\ H3}$ ~ 4.4 Hz).[41a] Subsequent *C*-desilylation with $AgNO_3$/2,6-lutidine gave terminal acetylene **2-13**, which was subjected to Mitsunobu conditions with *N*-formyl-L-leucine (**2-14**) to give the configurationally inverted product **2-1** (Scheme 2.3).



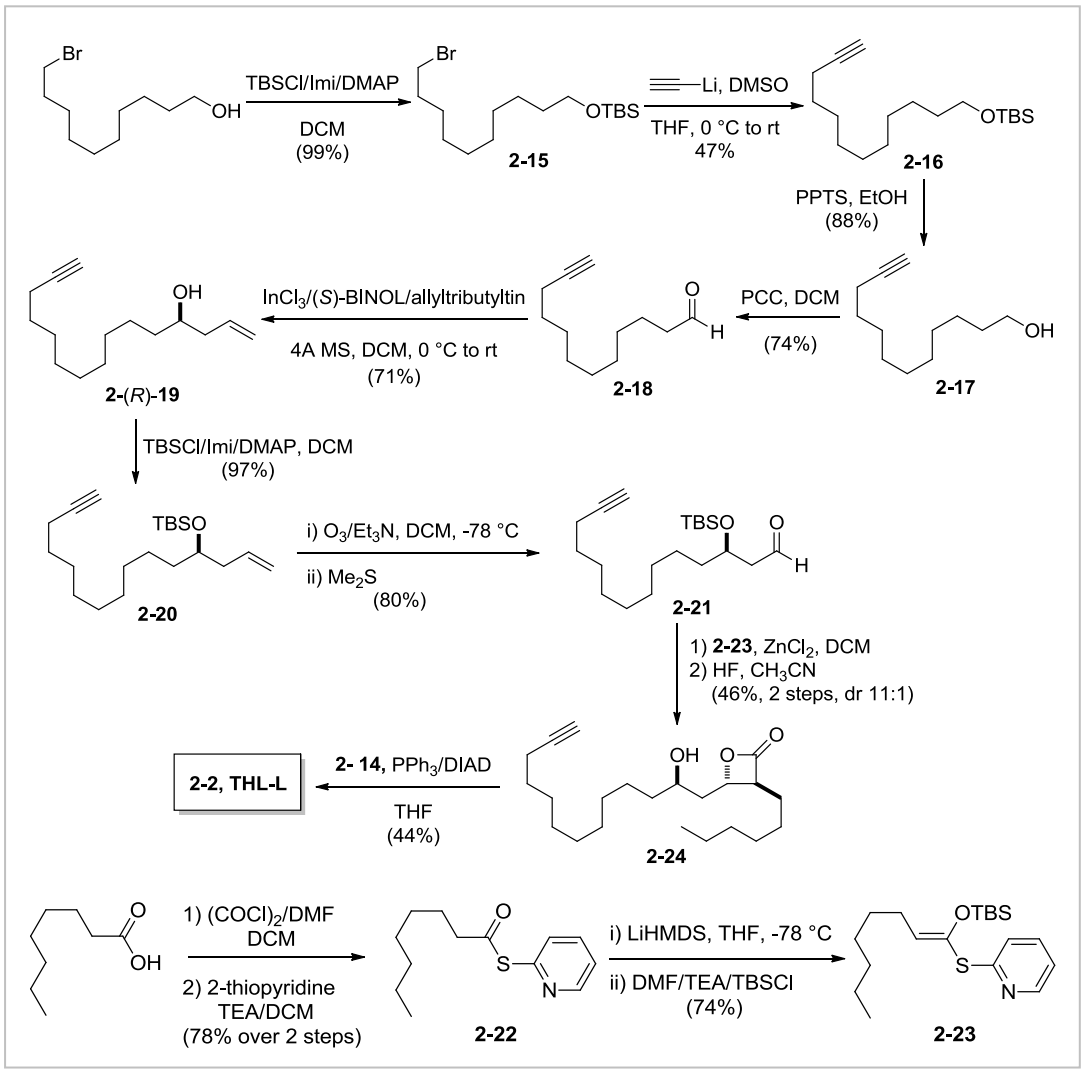
Scheme 2.2. Synthesis of **2-8a** and **2-8b.**



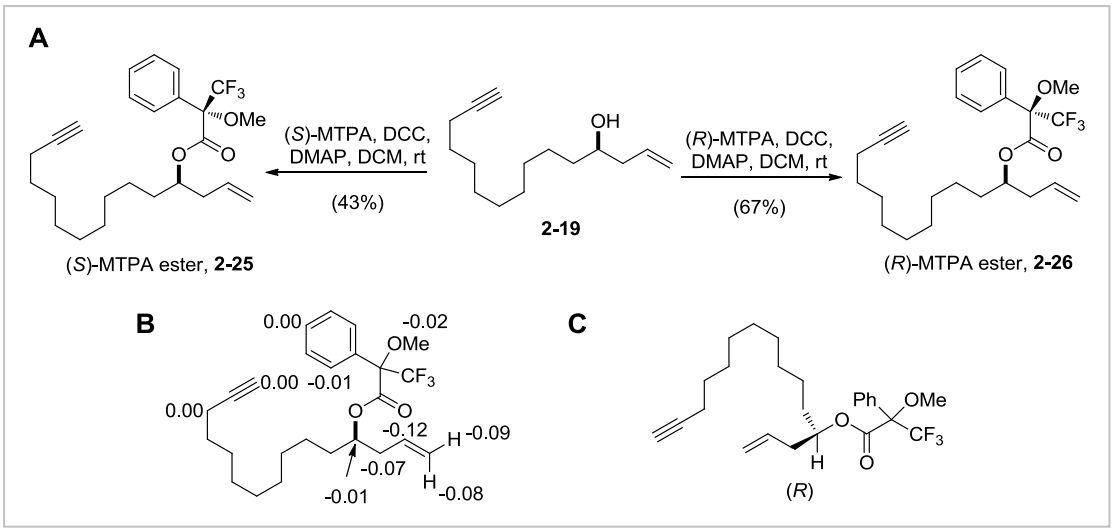
Scheme 2.3. Synthesis of THL-R **(2-1)**.

For the synthesis of **2-2** (THL-L), an alternative route to obtain the optically active aldehyde **2-21** was explored (Scheme 2.4). Starting from 10-bromo-1-decanol, TBS-protection, alkylation of lithium acetylene, TBS-deprotection, PCC oxidation, and asymmetric allylation with allyltributyl stannane in the presence of chiral (*S*)-BINOL-$InCl_3$ complex[44] to afford the chiral homoallylic alcohol **2-19**. The enantiomeric excess was determined to be 62.4% by ^{1}H NMR (500 MHz) analysis of its *S*-Mosher derivative at δ 5.63, and the absolute configuration of **2-19** was unequivocally determined as to be *R* using Mosher ester analysis[45] (Scheme 2.5).

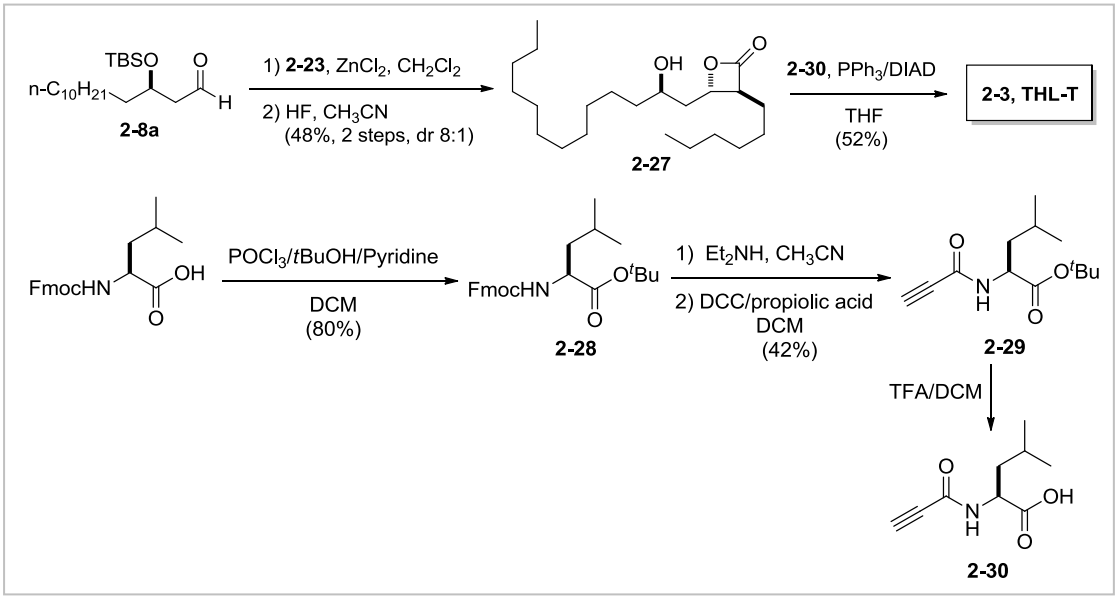
Subsequently, two steps sequence of TBS-protection and oxidative cleavage gave **2-21** which reacted with **2-23** under TMAL conditions to give **2-24** (~11:1 anti/syn) and eventually **2-2**. For the synthesis of **2-3** (THL-T), the known hydroxyl-β-lactone **2-27** (~8:1 anti/syn) was first prepared based on published procedures,[5a] then reacted with **2-30** (obtained in four steps from Fmoc-Leu-OH involving *O*-protection, Fmoc deprotection, DCC coupling, and *O*-deprotection) under Mitsunobu conditions, giving **2-3** in 52% yield (Scheme 2.6).



Scheme 2.4. Synthesis of THL-L (**2-2**).



Scheme 2.5. Determination of absolute configuration of **2-19**. (A) Synthesis of (*S*)- and (*R*)-Mosher esters **2-25** & **2-26**. (B) shows the $\Delta\delta^{SR}$ values from the ^{1}H NMR spectra of the (*S*)- and (*R*)-MTPA esters of homoallylic alcohol **2-19**. Based on model (**C**), the absolute configuration of homoallylic alcohol **2-19** was determined to be *R*.



Scheme 2.6. Synthesis of THL-T (**2-3**).

2.2.4 Effects on Cell Proliferation, Phosphorylation of eIf2α and Activation of Caspase-8

With three probes in hand, we next evaluated their biological activities against Orlistat (as a positive control). Three different types of cellular assays based

on previously established orlistat biology were used.[39a, 39c-e] First, the anti-proliferation activity of the four compounds (THL-R, THL-L, THL-T & Orlistat) against HepG2 cells (a human hepatocellular liver carcinoma cell line) were evaluated using the XTT assay (part A of Figure 2.4);[39a] all four compounds showed a dose-dependent inhibition of tumor cell proliferation over a 72 h time period with comparable potency. Second, we carried out comparative analysis of the compounds in their ability to induce phosphorylation of eIf2α in the prostate cancer PC-3 cells.[39c] Previous studies had shown that inhibition of FAS by Orlistat induces endoplasmic reticulum (ER) stress and results in the phosphorylation of the translation initiation factor eIF2α. As shown in part B of Figure 2.4, PC-3 cells treated with different amounts of each compound showed similarly elevated eIF2α phosphorylation. Because the phosphorylation of eIF2α leads to the inhibition of protein synthesis, we did metabolic labeling (illustrated in Figure 2.5) with azidohomoalanine (AHA, pioneered by Dieterich et al.[46]) to measure the levels of newly synthesized proteins in cells treated with orlistat, THL-R or cycloheximide (a well-known protein synthesis inhibitor). As shown in part C of Figure 2.4, greatly reduced levels of protein synthesis were observed in cells treated each of the three compounds. The inhibition of protein synthesis was dose-dependant, which is consistent with previous findings carried out using radiolabeled ^{35}S-methionine[39c]. Lastly, inhibition of FAS by Orlistat was previously shown to induce tumor cell apoptosis by activating caspase-8.[39e] We therefore tested the compounds against the invasive human breast cancer MCF-7 cells (part D of Figure 2.4); similar degrees of caspase-8 activation (as evidenced by the

appearance of the p41/43 bands corresponding to cleaved caspase-8) were observed against all four compounds (at either 5 or 25 μM). Collectively, these data show the introduction of a terminal alkyne handle at various designated locations in Orlistat scaffold did not noticeably affect its biological activities, thus **2-1**, **2-2** and **2-3** were indeed suitable chemical probes for cell-based proteome profiling and identification of previously unknown targets of orlistat.

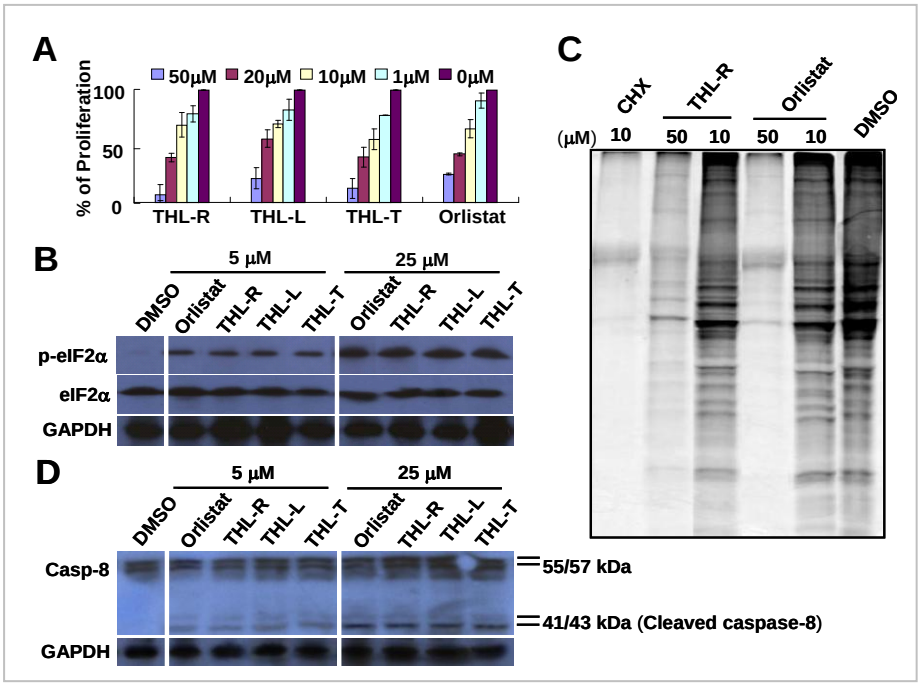


Figure 2.4. Biological evaluation of three Orlistat analogues. (A) Dose–dependent inhibition of HepG2 cell proliferation by Orlistat and three analogs (THL-R, THL-L and THL-T) using XTT assay. Data represent the average standard deviation for two trials. (B) Western blot analysis of eIF2α phosphorylation in PC-3 cells upon treatment with the four compounds. GAPDH was used as a loading control. (C) Inhibition of protein synthesis in HepG2 cells treated with the indicated concentrations of Orlistat, THL-R or CHX (cycloheximide, an inhibitor of protein biosynthesis) for 12 h and then pulsed with AHA (L-Azidohomoalanine) for 4 h. Cell lysates were prepared and subjected to click chemistry with a rhodamine-alkyne tag, SDS-PAGE analysis, and in-gel fluorescence scanning (fluorescent gel shown in grayscale). (D) Activation of caspase-8 in MCF-7 cells treated with the indicated concentrations of Orlistat/analogues.

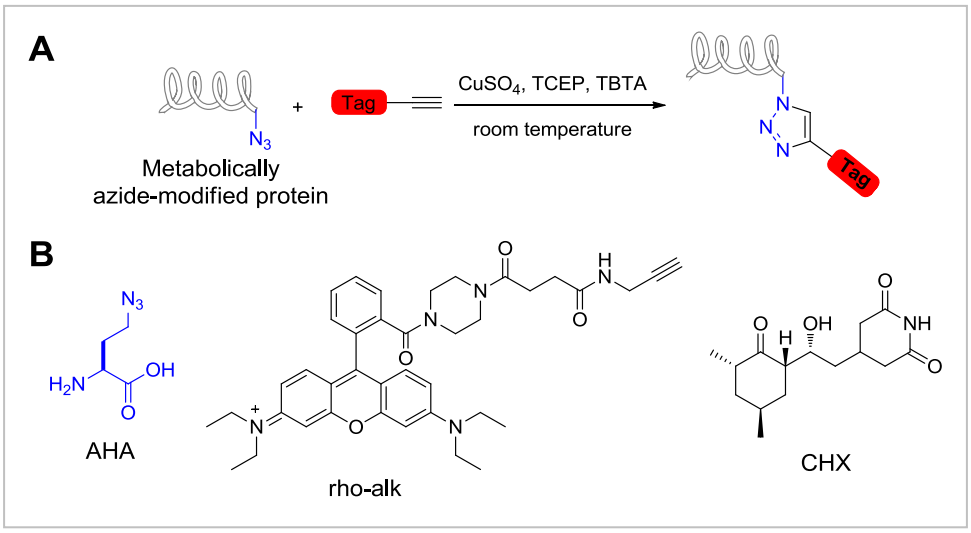


Figure 2.5. (A) Metabolic labeling with AHA and sequential click chemistry reactions with rho-alk allowing simultaneous visualization of the protein synthesis. (B) Structures of AHA, rho-alk and CHX used in this study.

2.2.5 In situ and in vitro Proteome Profiling

We next compared the in situ proteome reactivity profiles of the three probes to identify proteins that were covalently labeled by the probes in live HepG2 cells.[4] Probes (1-20 μM) were directly added to the cell culture medium, either alone or in the presence of 100 μM competing Orlistat. After 2 h, the cells were washed (to remove excessive probes), homogenized, incubated with rhodamine-azide (**2-33**) under click-chemistry conditions, separated by SDS/PAGE gel, and analyzed by in-gel fluorescence scanning (Figure 2.6). In addition to the expected FAS band, confirmed by treatment with an anti-FAS antibody (265 kDa, marked with red arrow in part C of Figure 2.7), we also observed a number of Orlistat-sensitive targets, that is, those labeled bands that were competed away by treatments with excessive Orlistat (marked with blue arrows and asterisks in parts A and B of Figure 2.6). Most of these

labeled bands were clearly visible even at a low probe concentration (e.g., 1 μM of THL-R; lane 1 in part A of Figure 2.6). Both THL-R and THL-L gave similar proteome labeling profiles, whereas THL-T consistently produced weaker labeled bands (possibly caused by inefficient click-chemistry conjugation due to inaccessibility of the alkyne handle located at the *N*-formyl-L-leucine end, which, based on X-ray structure of FAS/Orlistat complex, was buried deep into the protein[39b]).

To assess which nucleophilic residue of the labeled proteins might have been covalently modified by our probes, SDS-PAGE gels from the *in situ* labeling were subjected to in-gel treatment with hydroxylamine (NH_2OH), which preferentially cleaves thioesters, and esters to a lesser extent, under neutral pH conditions.[47] As shown in part C of Figure 2.6, the labeled FAS and most other Orlistat-sensitive bands showed a much reduced fluorescence signal upon treatment with NH_2OH, suggesting likely involvement of a cysteine/serine residue in the labeling between these proteins and THL (vide infra). It is interesting to note that the labeled FAS band was to some degree resistant to NH_2OH treatment, clearly indicating the formation of a more chemically stable ester linkage between Ser^{2308} in FAS[39b] and THL-R. In a related experiment, we also performed competitive ABPP with cerulenin, a known FAS inhibitor which irreversibly inactivates the β-ketoacyl-ACP synthase domain but not the thioesterase domain of FAS;[48] as expected (part D of Figure 2.6), Cerulenin did not abolish the labeling of the THL probes toward FAS as well as most other Orlistat-sensitive proteins. This indicates that the labeling of our probes is both target-

and domain-specific, and they may in the future be used to distinguish different domains of FAS. As shown in part E of Figure 2.6, when compared to *in situ* (live cell) labeling, similar proteome labeling profiles, albeit with significantly higher background labeling, were obtained when whole-cell lysates were used instead. This shows the importance of the *in situ* labeling (made possible by the cell-permeable property of the THL probes) as a prerequisite for accurate and specific target identification.

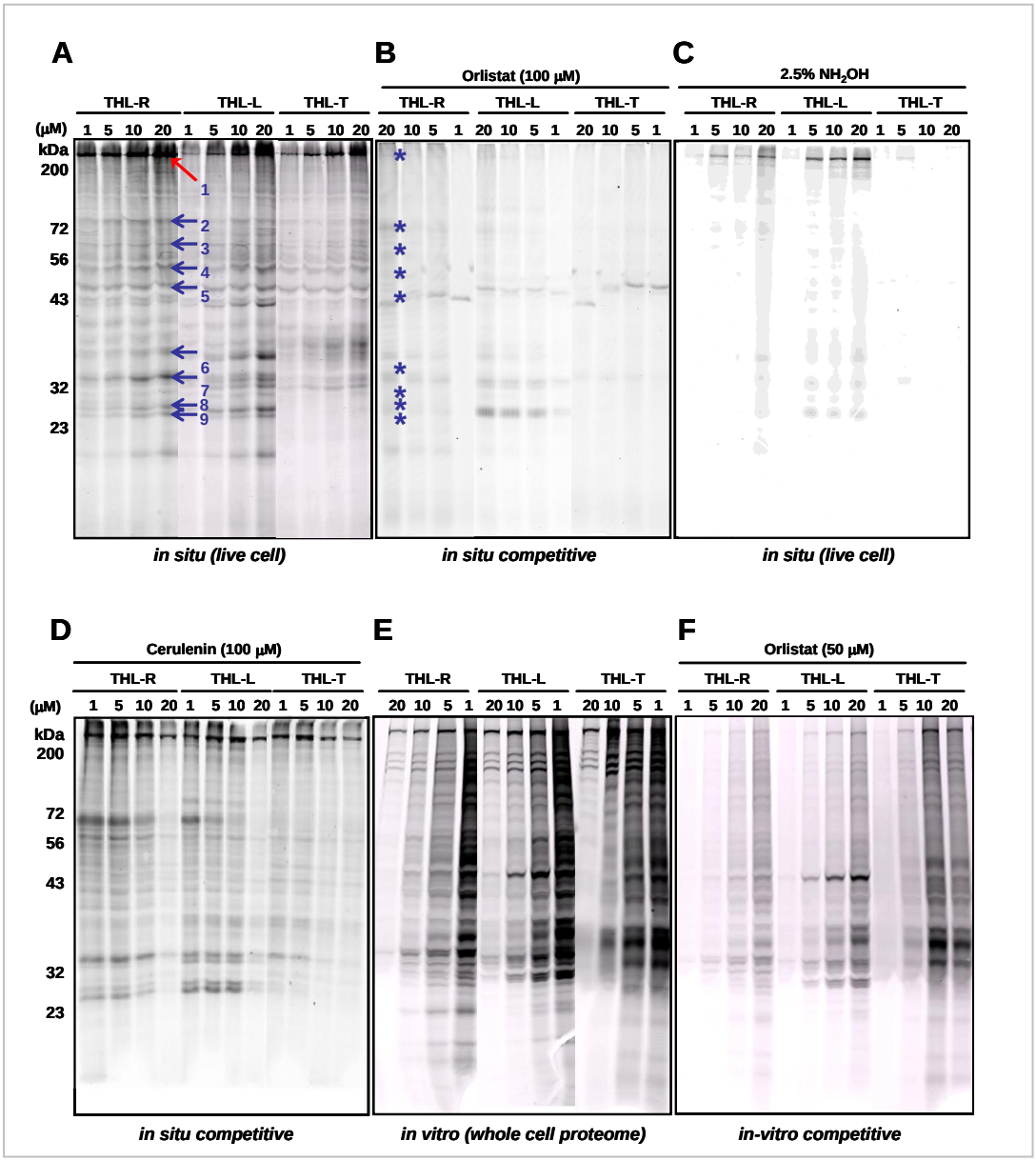


Figure 2.6. Comparison of *in situ* versus in vitro labeling profiles by THL analogs (THL-R, THL-L & THL-T) over a concentration range of 1-20 μM. (A) *In situ* proteome labeling of HepG2 cells, followed by click chemistry with a rhodamine-azide reporter tag (**2-33**), SDS/PAGE analysis, and in-gel fluorescence scanning (fluorescent gel shown in grayscale). Multiple THL-R-sensitive targets (marked with blue arrows) were isolated and identified in separate reactions using biotin-azide tag (**2-34**), followed by pull-down/MS procedure, and summarized in Table 2.1. (B) Specifically labeled targets of THL-R were defined as proteins whose labeling was completed by excess Orlistat (marked with asterisks). (C) Hydroxylamine sensitivity of proteins (with decreased fluorescence) in live HepG2 cells labeled with THL analogs (THL-R, THL-L & THL-T). (D) in situ competition experiment with Cerulenin. (E) in vitro labeling of HepG2 whole cell proteome. (F) in situ competition experiment with Orlistat.

2.2.6 Target Identification and Validation

Subsequently, the labeled protein extract was enriched (following click-chemistry conjugation with biotin-azide **2-34**) by avidin-agarose beads, separated by SDS-PAGE gel, confirmed by strepavidin blotting, subjected to in-gel trypsin digestion, and identified by MS/MS analysis. In addition to FAS, eight new proteins were identified (numbered 2 to 9, in part A of Figure 2.7 & Table 2.1), of which one is an unnamed protein. Two of the proteins, GAPDH and β-tubulin, are house-keeping genes constitutively expressed in most cell lines but are known to be expressed more highly in cancer cells.[49] Both GAPDH (a dehydrogenase) and β-tubulin (a hydrolase) possess nucleophilic active-site cysteine residues. It is therefore not surprising they were targets of Orlistat and their labeling was reversed by NH_2OH treatment. It should be noted that tubulins are well-known targets of anticancer drugs (e.g., taxol[50]). Three other proteins identified, RPL7a, RPL14, and RPS9, are ribosomal proteins. They are known to be implicated in protein synthesis, control of cellular transformation, tumor growth, aggressiveness and metastasis. Overexpression of these proteins had previously been reported in colon, brain liver, breast and prostate carcinoma.[51a-d] In retrospect, our earlier findings that protein synthesis was greatly inhibited in HepG2 cells treated with Orlistat, as shown in part C of Figure 2.4, may also imply that these ribosomal proteins are probable cellular targets of the drug. The remaining two proteins, Annexin A2 and Hsp90AB1, are involved in cell proliferation/division and protein degradation, respectively.

Table 2.1. Proteins identified by mass spectrometry with THL-R.[a]

#	protein name	accession No.	calc/obsd mass Da/kDa	# Pep.	protein score	E value	ion score	localization	protein funtion
2	Hsp90AB1	Gi39644662	75 088/72	18	136	2e-007	43	cytoplasm, nucleus	molecule chaperone with ATPase acitivity; stress response
3	unnamed	Gi194390424	56 156/56	8	83	0.038	50	cytoplasm, nucleus	protein biosynthesis; translational elongation
4	β-tubulin	Gi29788785	49 671/52	17	536	3.2e-049	116	nucleus	GTPase acitivity and cell division cell division
5	Annexin A2	Gi4757756	38 808/39	16	185	2.5e-012	38	cytoplasm, nucleus	calcium binding; cell division
6	GAPDH	Gi7669492	36 201/36	17	296	2e-023	62	cytoplasm, nucleus	glycolysis; energy production
7	RPL7a	Gi4506661	30 148/30	11	202	4.9e-014	72	cytoplasm, ribosome	biogenesis; protein synthesis
8	RPL14	Gi7513316	23 902/23	9	149	9.8e-009	65	cytoplasm, ribosome	biogenesis; protein synthesis
9	RPS9	Gi14141193	22 635/22	16	213	3.9e-015	47	cytoplasm, ribosome	biogenesis; protein synthesis

[a]. Representative MS/MS profile for RPL7a was shown in Appendix Two.

In an effort to further validate the MS results, we carried out pull-down and Western blotting experiments with the respective antibodies (part C of Figure 2.7);[52] results confirmed all proteins tested (including FAS) were indeed positively labeled by THL-R, and thus likely are true cellular targets of orlistat. Four of the proteins (β-tubulin, GAPDH, RPL7a, & RPL14) were taken for additional validation experiments. First, the c-Myc fusions of these proteins were transiently expressed in HEK-293T cells (part B of Figure 2.7), labeled (by THL-R *in situ*), immune-purified (with c-Myc agarose beads), subjected to click chemistry (with rhodamine-azide **2-33**), and analyzed by SDS-PAGE (part D of Figure 2.7); results unambiguously confirmed that all four proteins were fluorescently labeled by the probe.

To confirm whether labeling of the probe is active site-directed, GAPDH, whose active site residue Cys^{151} had previously been characterized,[16b] was taken as an example for further site-directed mutagenesis experiments. GAPDH is a multifunction protein and well known for its primary role as a glycolytic enzyme. Recently, increasing evidence has suggested that this enzyme is also involved in a variety of activities which are unrelated to energy production, including membrane fusion, microtubule bundling, DNA repair, and apoptosis.[49b] An active site (Cys^{151} to Ala) mutant of GAPDH was therefore generated, transiently expressed and purified from HEK-293T cells, labeled with THL-R, reacted with rhodamine-azide **2-33**, then finally analyzed by SDS-PAGE followed by in-gel fluorescence scanning (part E of Figure 2.7); as expected, probe THL-R only labeled the wild-type GAPDH and not the Cys151Ala mutant, thus confirming Cys^{151} as the residue in GAPDH being

covalently labeled by THL-R. This also confirms the active site-directed nature of the probe.

It should be noted that, in a series of recent reports as described in Chapter 1,[22] Sieber et al. had developed activity-based probes based on small molecule libraries containing a reactive β-lactone moiety. The authors concluded that β-lactones are promising privileged structures and could be used to identify a variety of mechanistically distinct enzymes. Our Orlistat-based probes, though structurally much more complex, appear to behave similarly and are capable of targeting a number of other cellular proteins in addition to FAS. Even though the exact physiological roles of these proteins in connection with Orlistat and its pharmacological effects have not been established from this study, we believe these putative cellular targets of Orlistat should be carefully evaluated when one considers using Orlistat in cancer therapy.

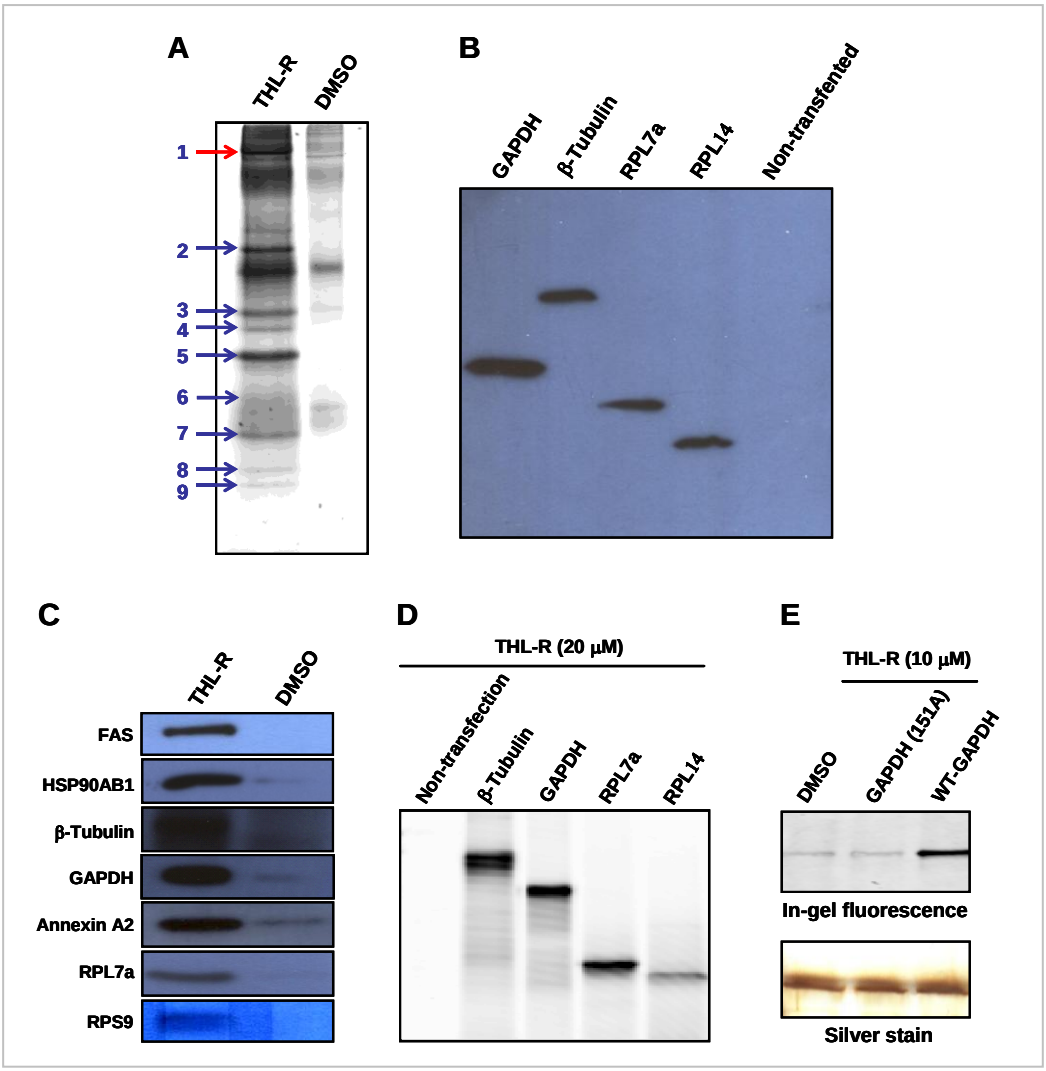


Figure 2.7. Affinity pull-down and target validation of the identified "hits". (A) Silver staining of gels of pulled-down fractions from THL-R-labeled or DMSO-treated HepG2 live cell lysates. (B) Western blotting of recombinantly expressed proteins with anti-c-Myc. (C) Western blot analysis of pulled-down fractions of HepG2 live cells treated with THL-R (or DMSO as negative controls; right lanes) with their respective antibodies. Biotin-azide **2-34** was used in the click chemistry with avidin-agarose beads for pull-down experiments. (D) *In situ* labeling of recombinantly expressed β-tubulin, GAPDH, RPL7a & RPL14 by THL-R (fluorescent gel shown in grayscale). (E) Comparative analysis of labeling of wild-type GAPDH and Cys^{151} Ala mutant (upper panel; fluorescent gel shown in grayscale); (lower panel, silver-stained gel) comparable amounts of protein loading were demonstrated in all three lanes.

2.2.7 Cellular Imaging

To demonstrate the utility of our cell-permeable probes for cellular imaging of Orlistat targets, we performed fluorescence microscopy to visualize probe-treated cells (Figure 2.8). Live HepG2 cells were first treated with THL-R (**2-1**), fixed with PFA, permeabilized with Triton X-100, conjugated to rhodamine-azide (**2-33**) by click chemistry and imaged (colored in Red). Immunofluorescence was also carried out on the same cells to visualize the localization of endogenous FAS (colored in Green in top panels). Minimal fluorescence was observed in samples treated with only DMSO, whereas in THL-R treated cells, fluorescence was mostly distributed in endoplasmic reticulum (ER). Note that endogenous FAS was mostly cytosolic (which includes ER). Thus our imaging results are consistent with previous findings that inhibition of FAS with Orlistat induces ER stress specifically in tumor cells.[3c] We take note, however, that, at its current state, THL-R might not be the most suitable chemical probe for bioimaging of FAS, as it also labels a number of other cellular proteins (as evidenced from our studies herein).

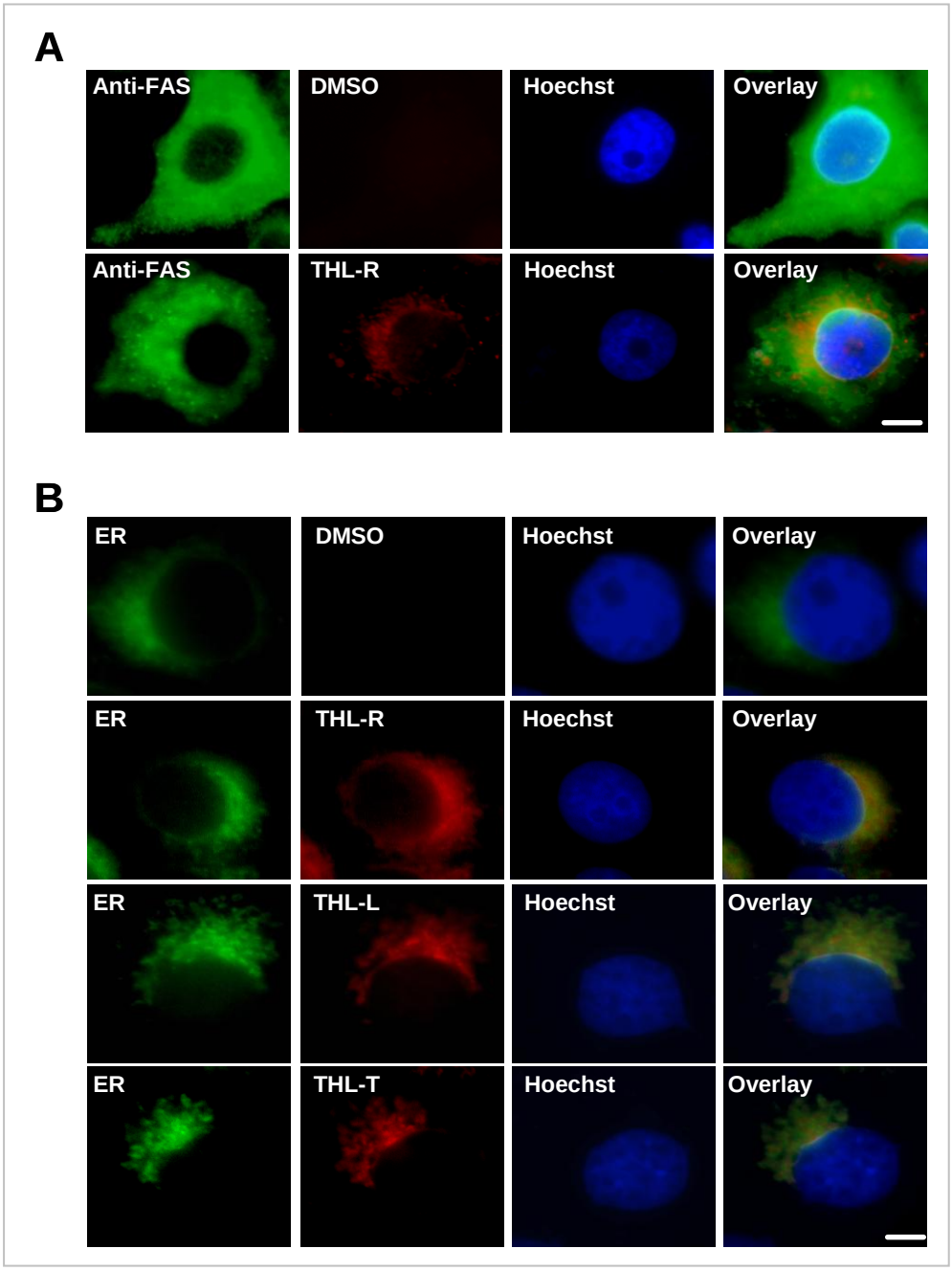


Figure 2.8. Cellular imaging of HepG2 cells treated with probes. (A) Immunofluorescence analysis of FAS in HepG2 cells not labeled in the absence or labeled in the presence of THL-R (20 μM). (B) Fluorescence microsocopy of HepG2 cells labeled with 20 μM probes (i.e., THL-R, THL-L, and THL-T) or DMSO. Live cells were treated with probe for 2 h, fixed, permeabilized, reacted with rhodamine-azide (10 μM; colored in red), incubated with anti-FAS primary antibodies, and stained with FITC-conjugated IgG antibody (green) or ER-Tracher™ Green (glibenclamide BODIPY® FL), and followed by Hoechst (blue) stain. Images were acquired by an Olympus IX71 inverted microscope, equipped with a 60X oil objective (NA 1.4, WD 0.13 mm) and CoolSNAP HQ CCD camera (Roper Scientific, Tucson, AZ, USA). Scale Bar = 8 μm. All images were acquired similarly.

2.3 Conclusion

We have developed a novel chemical proteomic approach that enabled, for the first time, identification and putative validation of several previously unknown cellular targets of Orlistat. The potential of these cell-permeable probes to be used as future imaging probes has also been explored. Whereas further studies are needed to better understand the exact relevance of Orlistat and its pharmacological effects in relation to these newly identified cellular targets, our findings point to a likely scenario that these proteins might be potential off-targets of Orlistat. It is also possible that the antitumor activities of Orlistat might have originated from the drug's ability to inhibit both FAS (as previously reported[39]) and some of these newly identified targets. In either case, our findings have important implications in consideration of Orlistat as a potential anticancer drug. Finally, our strategy should be broadly useful for off-target identification against quite a number of existing drugs and/or drug candidates, which are also covalent modifiers of their biological targets.[53]

Chapter 3

Chemical Modification and Organelle-Specific Localization of Orlistat-Like Natural-Product-Based Probes

Portions of the work presented here are taken from:

P.-Y. Yang, K. Liu, C. Zhang, G. Y. J. Chen, Y. Shen, M. H. Ngai, M. J. Lear, and S. Q. Yao, "Chemical Modification and Organelle-Specific Localization of Orlistat-Like Natural-Product-Based Probes." *Chem. Asian. J.* **2011**, *6*, 2762-2775 .

Abstract

Orlistat™, also known as tetrahydrolipstatin (THL), is an FDA-approved anti-obesity drug with potential anti-cancer activity. In Chapter 2, we developed a chemical proteomic approach, based on the Orlistat-like probe (THL-R) for large-scale identification of unknown cellular targets of Orlistat in human hepatocytes. In this chapter, we report the chemical synthesis and biological evaluation of an expanded set of Orlistat-like compounds, with the intention to further dissect and manipulate potential cellular targets of Orlistat. In doing so, we carried out proteome-wide activity-based profiling and large-scale pull-down/LCMS analysis of these compounds in live HepG2 cells, and successfully identified many putative cellular targets for Orlistat and its structural analogues. By qualitatively assessing the spectra counts of potential protein hits against each of the seventeen Orlistat analogues, we obtained both common and unique targets of these probes. Our results revealed that subtle structural modifications of Orlistat led to noticeable changes in both the cellular potency and target profiles of the drug. In order to further improve the cellular activity of Orlistat, we successfully applied the well-established AGT/SNAP-tag technology to our cell-permeable, benzylguanine (BG)-containing Orlistat variant (**3-4**). We showed that the drug could be delivered and effectively retained in different sub-cellular organelles of living cells. This strategy may provide a general and highly effective chemical tool for the potential sub-cellular targeting of small molecule drugs.

3.1 Introduction

Accurate and thorough determination of drug–target interaction is of great importance in drug discovery. It offers invaluable biological insights for a drug candidate, such as its mode of action both in vitro and in vivo. It also provides clues for compound optimization in order to maximize the therapeutic potential and minimize potential cellular toxicity of a drug.[1,54] Recent advances in chemical proteomics, a multidisciplinary research area integrating biochemistry and cell biology with organic synthesis and mass spectrometry, have enabled a more direct and unbiased analysis of a drug's mechanism of action in the context of the proteomes as expressed in the target cell or the tissue of interest.[1c,55] However, at present, few methods are available for large-scale profiling of drug-protein interaction in an *in vivo* setting.[56] We described (in Chapter 2) a chemical proteomic method that makes use of natural-product–like small-molecule probes for proteome-wide profiling of putative drug targets in living cells.[57] That strategy is based on the concept of activity-based protein profiling (ABPP), originally coined by Cravatt group and further developed by other groups.[5] In our studies, terminal-alkyne–containing probes based on Orlistat™ or tetrahydrolipstatin (THL, Figure 3.1), an FDA-approved anti-obesity drug, were developed.[58,59]

Orlistat works primarily on pancreatic and gastric lipases within the gastrointestinal (GI) tract. Recently, Orlistat and other lipstatin-based natural products have shown promising anti-tumor activities by potently inhibiting human fatty acid

synthase (FAS) in tumor cells.[39,60] FAS is an enzyme essential for the survival of cancer cells.

In our approach, extremely conservative modifications (i.e., an alkyne handle) were introduced in the parental Orlistat structure to provide the necessary functionality for target identification via Cu(I)-catalyzed 1,3-dipolar cycloaddition (click chemistry) or copper catalyzed azide–alkyne cycloadditions (CuAAC),[16] whilst maintaining the native biological properties of the drug. When combined with affinity purification and MS/MS analysis, we were able to identify and putatively validate several previously unknown cellular targets of Orlistat, including GAPDH, β-tubulin, Annexin A2, Hsp90AB1 and three ribosomal proteins (RPL7a, RPL14 & RPS9) as described in Chapter 2. Although some of them are highly expressed endogenously, and their presence might be explained primarily by their relative cellular abundance, such information is important to better understand the full biological effects of a drug.

Having said this, there are still a number of challenges that prevent Orlistat from being considered as a potential anti-tumor drug; it has poor solubility and bioavailability, and, most importantly, it lacks sufficient potency and specificity. To begin to address some of these issues, we and others had previously focused on the development of highly efficient chemistry which enable facile synthesis of Orlistat analogues by performing various chemical modifications at the aliphatic side chains around the β-lactone core of the drug.[41b] In our preliminary studies using different Orlistat-like compounds, we found a direct linkage between chemical modifications of Orlistat and its cellular activities.[57b] However, no further studies were carried out to

delineate the potential cellular targets of these analogs and their biological consequences. Another way to achieve good in vivo efficacy of bioactive compounds is to deliver them into the sub-cellular compartment of cells where the target enzyme resides. In this way, an increase in the effective concentration of the inhibitor (i.e. improved potency) with a better selectivity would be expected. In fact, such sub-cellular targeting has recently been demonstrated successfully for several inhibitors and smallmolecule probes in cultured cells.[61] Herein, we describe further chemical and cellular modifications of Orlistat with the ultimate aim to improve its cellular activities. By using an expanded set of new Orlistat-like probes, we performed in situ proteome-profiling experiments following treatment of live HepG2 cells with our probes (Figure 3.2). Large-scale pull-down/LCMS identifications of potential cellular targets were subsequently conducted. From these studies, both common and unique protein targets of Orlistat-like analogues were identified. In order to further improve Orlistat's cellular activities, we developed a "cellular modification" strategy to deliver our cell-permeable Orlistat variant to different sub-cellular organelles in living cells (Figure 3.5). This strategy was successfully implemented by taking advantage of the well-established AGT/SNAP-tag technology.[62]

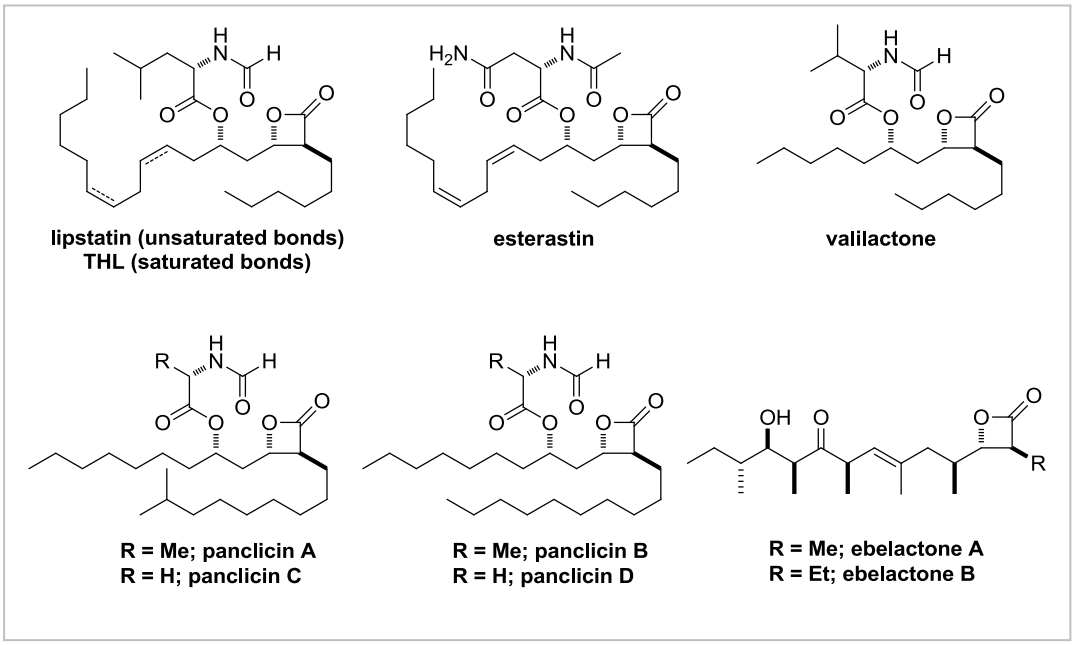


Figure 3.1. Representative structures of lipstatin family of natural products possessing trans-3,4-disubstitued-β-lactones.

3.2 Results and Discussion

3.2.1 Design of a Library of Orlistat-Like Natural Product-Based Probes

The design of Orlistat-like probes was based on the general structure of a number of naturally occurring congeners of lipstatin, such as lipstatin, esterastin, valilactone and panclicin A-D (Figure 3.1), and our previously reported three probes in Chapter 2, THL-R, THL-L and THL-T, in which a terminal C-C single bond in one of the aliphatic chains in the parental Orlistat structure was substituted with a C-C triple bond. This extremely conservative modification was previously shown to enable full retention of the native biological activities of Orlistat, and at the same time allow subsequent chemical profiling/target identification by the downstream conjugation of reporter tags via the bio-orthogonal click chemistry. We now wish to describe the

synthesis, anti-proliferation evaluation and proteomic profiling results of an expanded list of Orlistat analogs incorporating various degrees of molecular complexities (i.e., changes in the position of the terminal alkyne handle, the amino ester and chirality). This detailed investigation has refined our present understanding of the anti-proliferative role played by Orlistat and established certain structure-activity relationships (SARs) within this drug's structural motif from which next generations of designed analogs may emerge. To this end, sixteen new analogs of Orlistat were synthesized for the first time. Together with the three previously reported compounds (THL-R, THL-L & THL-R, hereafter renamed as **3-1a**, **3-2a** & **3-3a**, respectively) and two key intermediates (**2-13** and **2-24**), all twenty-one compounds were used in this study. In the analogue series **3-1b-j**, the left-hand aliphatic side chain remained intact, whilst the nature of amino ester and/or the chirality at C3, C4 and C6 positions were varied strategically. In the analogue series **3-2b-f**, the right-hand aliphatic side chain was kept constant, whilst the structure of the amino ester and the chirality at C3, C4 and C6 positions (**3-2f**, containing an *N*-formyl-L-leucine moiety) were altered. For the analogue series **3b**, further structural changes were made at the *N*-formyl group by the introduction of either a long alkyne-containing acyl group (i.e., **3-3b**). In the case of **3-3c**, two terminal alkyne handles were simultaneously introduced (top and right).

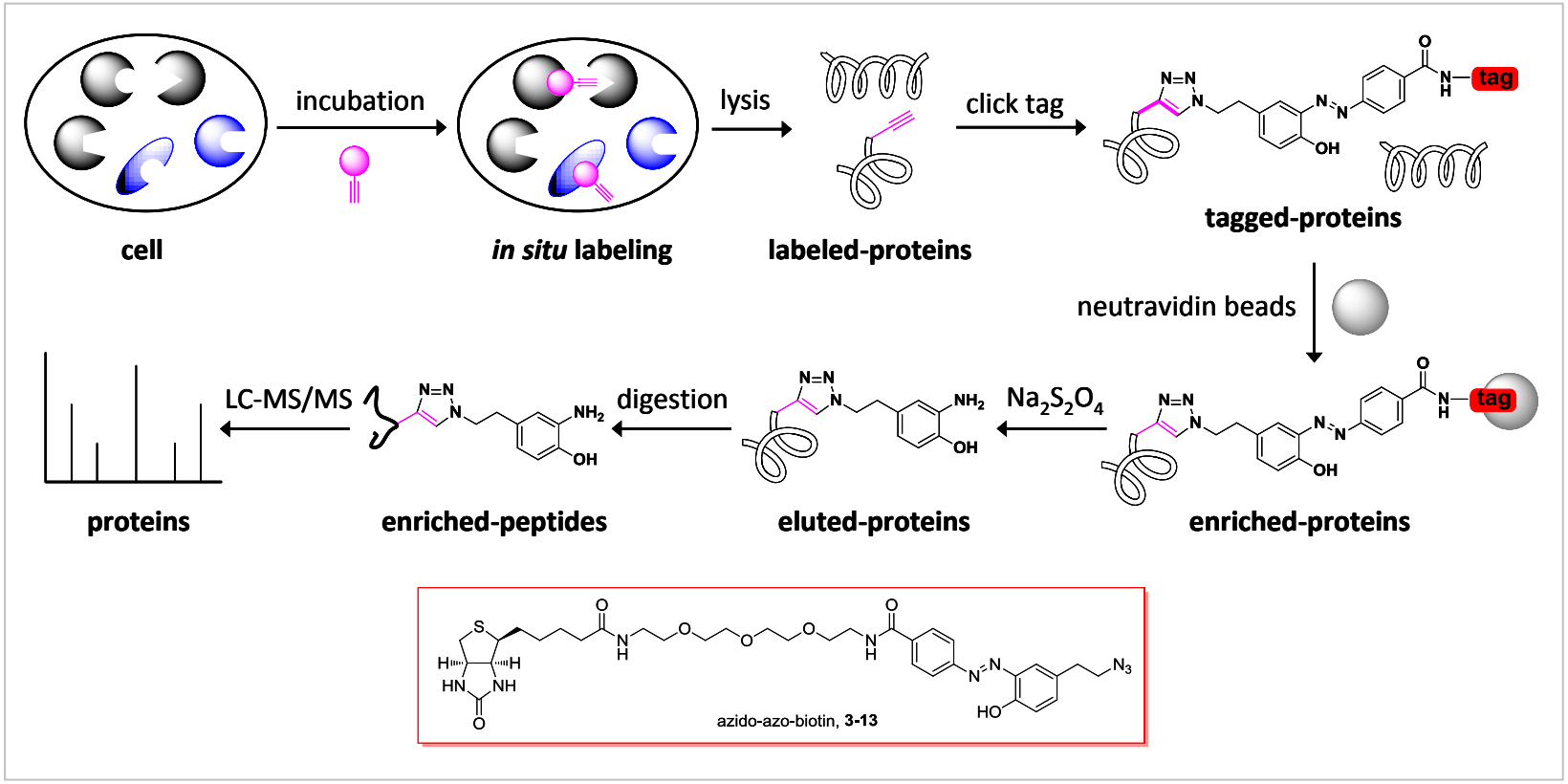
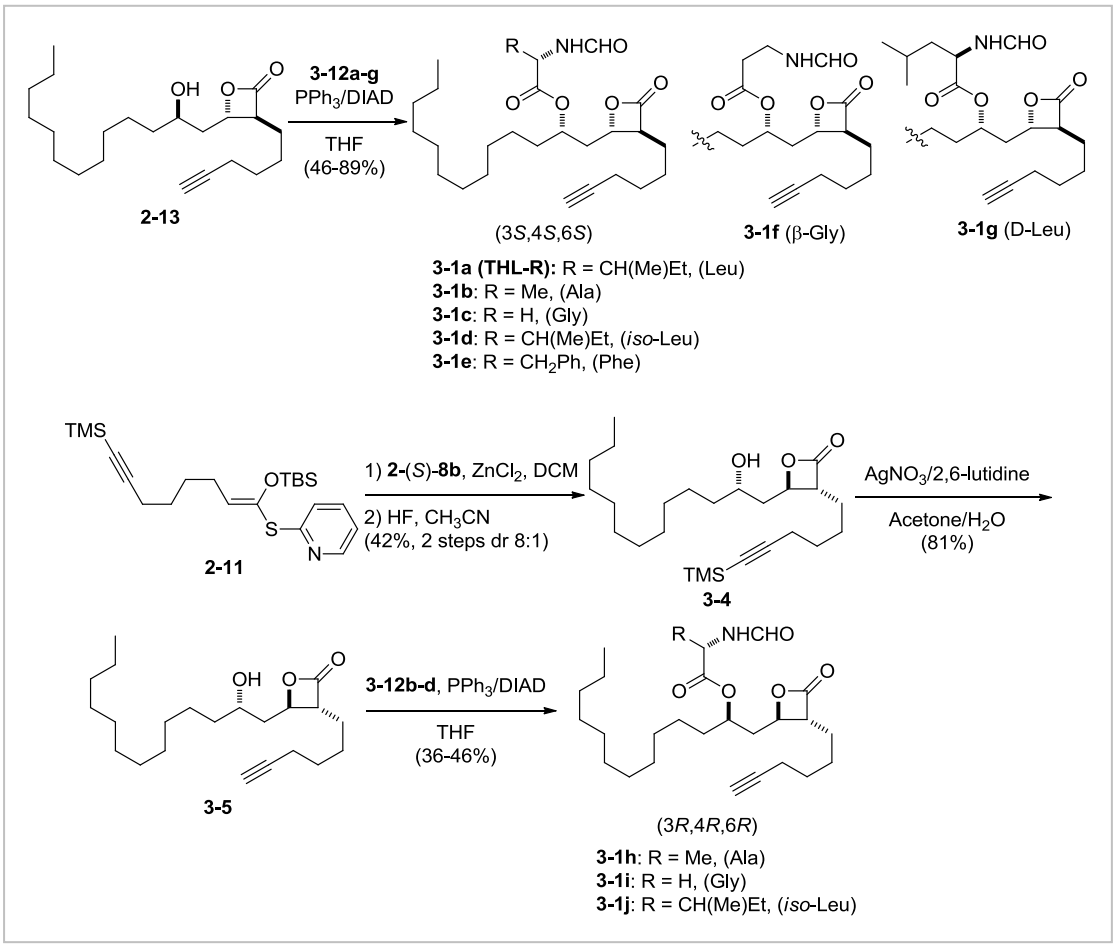


Figure 3.2. Overall workflow of the large-scale affinity pull-down/LCMS experiments.The chemically cleavable linker, **3-13**, used in study is shown in red.

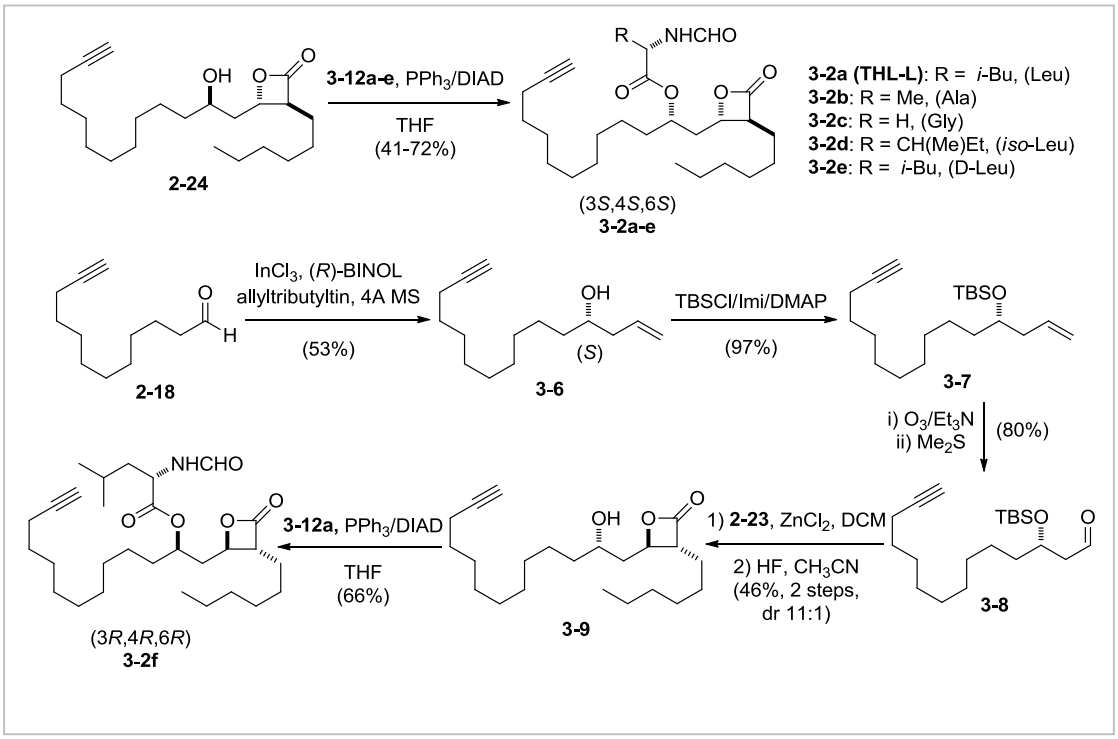
3.2.2 Synthesis of Sixteen Orlistat-Like Natural Product-Based Probes

As shown in Scheme 3.1, Scheme 3.2 and Scheme 3.3, the synthesis of these sixteen new compounds (**3-1b-j**, **3-2b-f** & **3-3b-c**) was based on previously described general strategies for the synthesis of **3-1a**, **3-2a** and **3-3a** (THL-R, THL-L & THL-T), respectively. Similar to **3-1a** and **3-2a**, compound **3-1b-g** and **3-2b-e** were conveniently synthesized from the key intermediate **2-13** and **2-14**, respectively, with various *N*-formylated amino acids (**3-14a-g**) under Mitsunobu conditions, thus giving rise to the corresponding configurationally inverted products in a single step. Synthesis of Orlistat analogs **3-1h-j** was started with previously prepared enantiopure aldehyde **2-8b** from the commercially available dodecan-1-ol. In this context, it is noteworthy that this synthetic sequence, which relied on the kinetic resolution of racemic allylic alcohols, was chosen because of its applicability to large-scale synthesis and convenient variation of the chirality at C-3, C-4 and C-6 positions. Briefly, the aldehyde **2-8b** was treated with thiopyridyl ketene acetal **2-11** under standard tandem Mukaiyama aldol-lactonization (TMAL) reaction conditions,[10] thus providing the silyloxy-β-lactone intermediates as a mixture of diastereomers (ca. 9:1 *anti/syn*) with complete selectivity for the trans-β-lactones. Subsequently, direct O-desilylation allowed isolation of the enantiomerically pure γ-hydroxyl β-lactones **3-4**. Subsequent C-desilylation with $AgNO_3$/2,6-lutidine gave the terminal acetylene **3-5**, which was subjected to Mitsunobu conditions with various *N*-formyl amino acids, thus giving rise to the corresponding configurationally inverted products **3-1h-j**. Synthesis of **3-2f** commenced with the optically active aldehyde **3-6**, which was obtained by

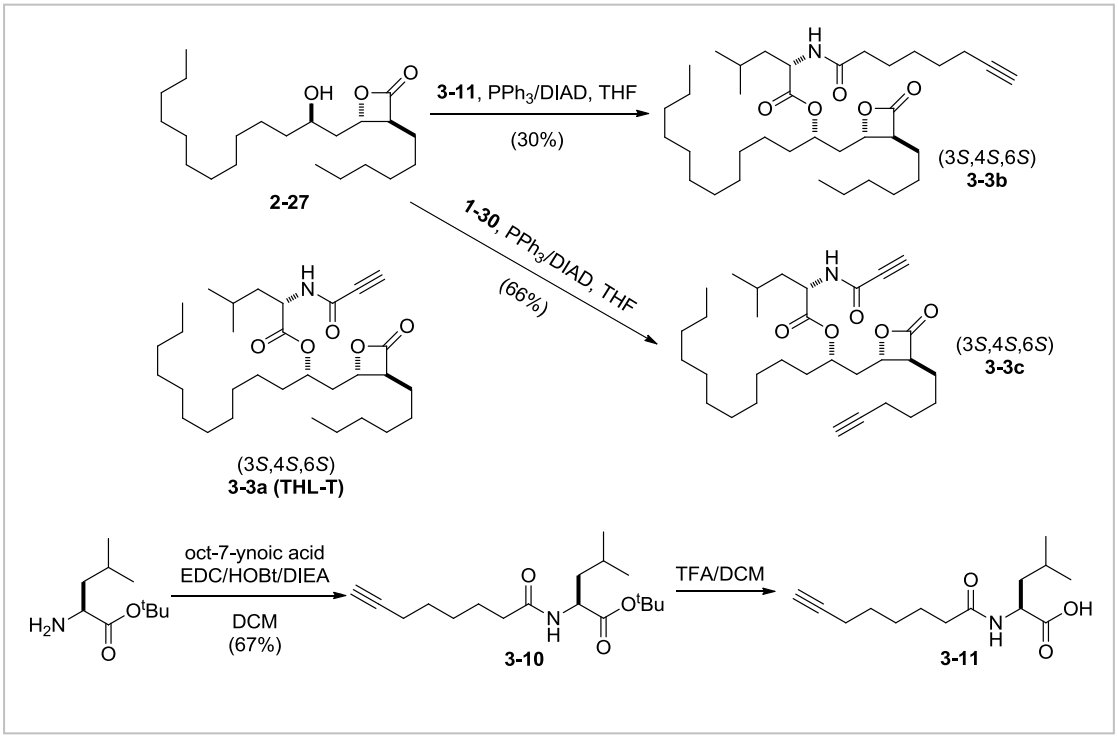
enantiomeric enrichment using an asymmetric allylation (allyltributyltin/(*R*)-BINOL/$InCl_3$) reaction. Following tert-butyldimethylsilyl (TBS) protection and oxidative cleavage, the desired aldehyde **3-8** was obtained, which, following TMAL and Mitsunobu reactions, led to the desired Orlistat analogs **3-2f**. For synthesis of **3-3b**, the known hydroxyl-β-lactone **2-27** was first prepared as previously described. Subsequent treatments with **3-11** under Mitsunobu conditions furnished the targeted analog **3-3b**. Compound **3-3c** was conveniently synthesized from the key intermediate **2-13** with **1-30** under Mitsunobu conditions.



Scheme 3.1. Synthesis of **3-1b-j**.



Scheme 3.2. Synthesis of **3-2b-f**.



Scheme 3.3. Synthesis of **3-3b-c**.

3.2.3 Biological Screening

With these sixteen new Orlistat analogs in hand (**3-1b-j**, **3-2b-f** & **3-3b-c**), together with the wild-type Orlistat, the three previously reported Orlistat analogues (**3-1a**, **3-2a** & **3-3a**) and two key intermediates (**2-13** and **2-24**), a total of twenty-two compounds were tested for their anti-proliferative activities against the human liver cancer line HepG2 using the XTT assay, as described in Chapter 2, followed by in situ proteome profiling and targeting identification using affinity pull-down/LCMS approaches. The overall workflow of the strategy is shown in Figure 3.2. As shown in Figure 3.3, nine of the compounds examined (**3-1a-c**, **3-1d**, **3-1f, 3-2a, 3-2d, 3-3a** & **3-3c**) displayed comparable activities to Orlistat in blocking the proliferation of HepG2 cancer cells. A closer examination of the data revealed a noteworthy structure-activity relationship (SAR) between the compounds tested and their anti-proliferative activities. In general, the chirality of the Orlistat analogs is important for activity. For example, in a direct head-to-head comparison, compounds with an inverted stereochemistry at C-3, C-4 and C-6 (*S* to *R*), i.e., **3-1b-d** versus **3-1h-j**, and **3-2a** versus **3-2f**, were shown to have significantly lower anti-proliferative activities. The stereochemistry of the C-α position in the *N*-formylated amino ester moiety of Orlistat also had a notable effect on its activity, as both **3-1a** and **3-2a** (which are structurally identical to Orlistat, except with a C-C to C≡C substitution at the left- and right-hand aliphatic chains, respectively) appeared to be more-active than their corresponding epimers, **3-1g**, and **3-2e**, respectively. The *N*-formylated amino ester moiety as a whole appeared to be essential for anti-proliferative activity, as

demonstrated by the lack of activity in the two hydroxyl lactone intermediates, **2-13** and **2-24** (in which the *N*-formylated amino ester was deleted completely). It should be noted that, in a previous study, the *N*-formyl group of Orlistat had also shown to be highly sensitive to the inhibition of FAS.[10b] The different anti-proliferative profiles displayed by Orlistat-like compounds likely have been accumulated through their inhibition against a multitude of cellular targets, including FAS, in HepG2 cells. Indeed, many drugs appear to work by synergistically targeting multiple proteins.[63]

3.2.4 In situ Proteome Profiling and Target Identification

We then carried out in situ proteome-reactivity profiles of all probes to identify potential protein targets which were covalently labeled by the probes in live HepG2 cells. Following previously described protocols with some modifications (summarized in Figure 3.2), each probe (20 μM) was added directly to the cell culture medium. After two hours of incubation, the cells were washed (to remove excessive probes), homogenized, incubated with rhodamine-azide (**2-33**) under click chemistry conditions, separated by SDS-PAGE gel, and analyzed by in-gel fluorescence scanning. As shown in part A of Figure 3.4, FAS in general was the main cellular target of these probes in HepG2 cells, as evident in the in-gel fluorescence scanning in which it showed up as the most strongly labeled band amongst all labeled bands. This is similar to the labeling profiles obtained from another class of Orlistat compounds previously reported.[57b] The two key intermediates, **2-13** and **2-24**, produced only weakly labeling profiles (including weak FAS labeling), again reinforcing the importance of

N-formylated amino ester moiety the Orlistat structure. However, covalent labeling of other cellular targets could still be readily detected in the fluorescence gel, with different probes generating marginally but noticeably different labeling profiles. This indicates our probes targeted both common and unique cellular targets in HepG2 cells, which might have accounted for the subtle differences observed in their anti-proliferative activities as well.

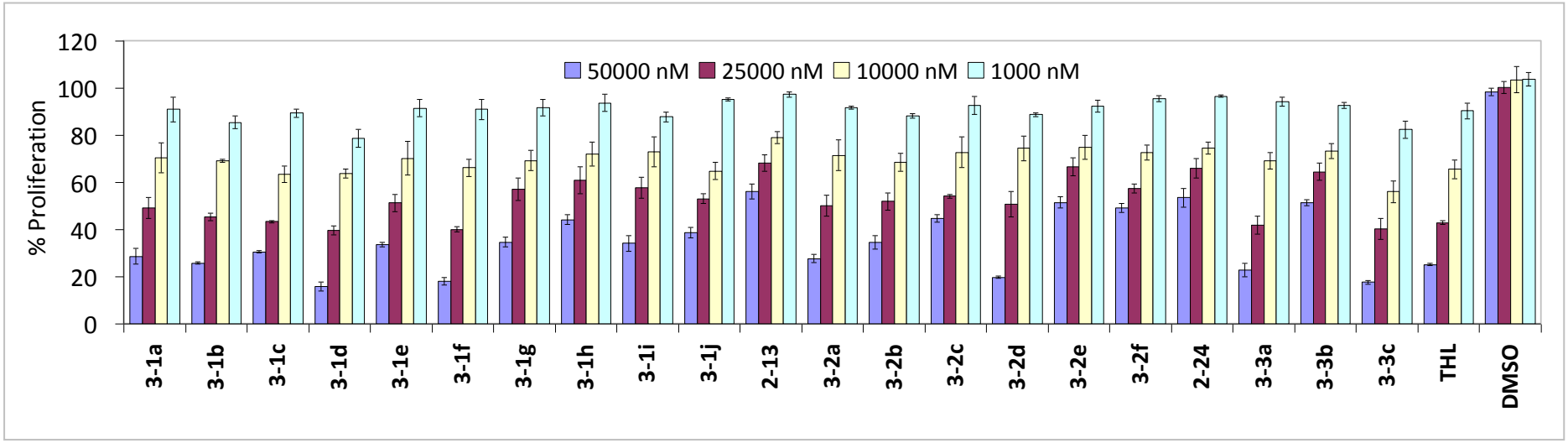


Figure 3.3. Dose-dependent inhibition of HepG2 cell-proliferation by the 21-member orlistat analogues library using an XTT assay. The data represent the average standard deviation for two trials. Wild-type Orlistat (positive control) and DMSO (negative control) were included in the assay.

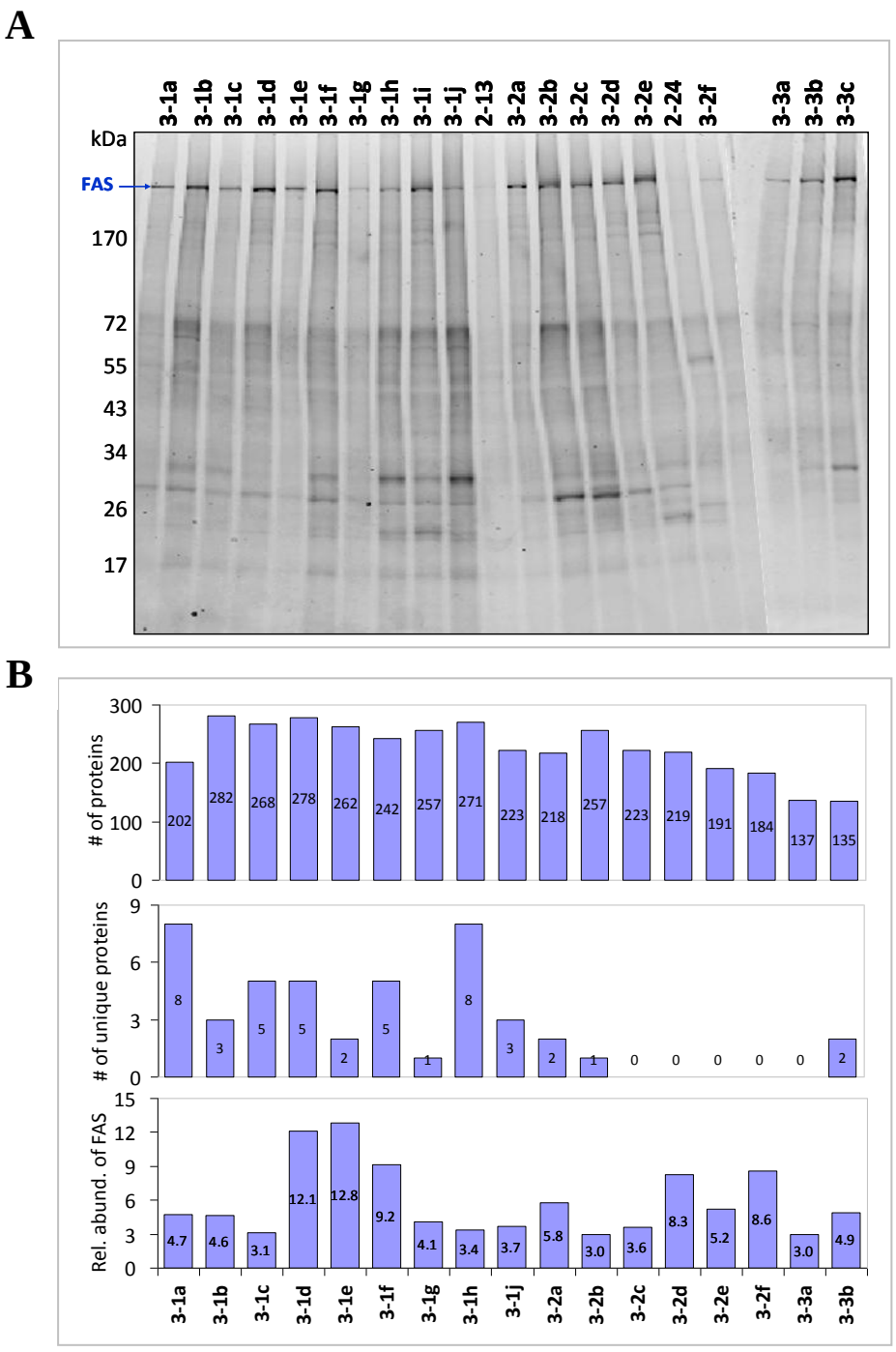


Figure 3.4. (A) *In situ* proteome-profiling of orlistat analogues against HepG2 cells. Probe-labeled proteins were detected by click-chemistry-mediated coupling to a rhodamine–azide reporter tag, followed by SDS-PAGE and in-gel fluorescence scanning. The arrow indicates the labeled FAS band. (B) Summary of proteins identified in all seventeen Orlistat-like compounds. Top: total number of proteins identified from each probe; Center: total number of unique proteins identified from each probe; Bottom: relative percentage abundance of FAS (calculated from the emPAI value for FAS in the experimental sample/the sum of the emPAI values for FAS in all seventeen samples) obtained from each probe.

To further identify these potential cellular targets, we subsequently performed large-scale affinity pull-down/LCMS analysis of protein lysates obtained from live HepG2 cells treated with each of the probes. In Chapter 2, we used a standard, non-cleavable biotin reporter tag to enrich covalently captured proteins and eluted them by boiling in the presence of an SDS-containing buffer. However, under such harsh denaturing conditions, both endogenously biotinylated and highly expressed cellular proteins (nonspecifically bound to the avidin beads) were still present in the eluent even after repeated washes, resulting in an unacceptable number of false positives and greatly interfering with the identification of low-abundance cellular targets. In this study, improvement was therefore made by making use of an azobenzene-functionalized biotinylated linker (compound **3-13** in Figure 3.2),[23e,64] which allows selective release of probe-labeled proteins under mild elution conditions (25 mM $Na_2S_2O_4$) that are compatible with mass spectrometric analysis. Consequently, fewer false positives and more reliable putative protein hits were obtained in our assay. Briefly as summarized in Figure 3.2, cellular lysates containing protein bands which have been covalently labeled by the Orlistat-like probe were treated with the cleavable azido-azo-biotin linker (**3-13**) under standard click chemistry conditions, precipitated with acetone, washed with methanol, solubilized with an SDS buffer, followed by affinity purification with Neutravidin beads and SDS-PAGE gel separation. The entire lane from each pull-down experiment was excised into 10 contiguous gel slices which were subsequently processed individually for in-gel trypsin digestion as described under Experimental Section. As negative controls, the entire large-scale proteomic

experiment (from cell treatment to LCMS analysis) was repeated with DMSO. Peptides obtained from each gel slice were eluted and subjected to nano-LC-MS/MS analysis. The LC-MS/MS data were searched against the IPI (International Protein Index) human protein database using an in-house MASCOT server for protein identification. All proteins were identified with a minimum score of 40 and at least two unique peptides. Based on these criteria, around 200 proteins on average were identified for each of the seventeen Orlistat probes used (see top graph; part B of Figure 3.4).

Details of the identified protein hits, including their accession numbers, emPAI values and protein masses, are listed in the Supplementary Table (in Appendix_CD). A total of sixty proteins were identified in all of our probes which may constitute the common cellular targets for these probes (for relevant references, see Appendix 1). These include well-known serine-type proteins, such as enolases (ENO1, ENO3), elongation factors (EEF2, EEF1A1 and EEF1G) and cysteine-type proteins, such as transferrin receptor protein (TFRC), ubiquitin-like modifier-activating enzyme (UBA1), and protein disulfide-isomerase (PDIA6), and several lipid metabolism related proteins, such as 3-hydroxyacyl-CoA dehydrogenase type-2 (HSD17B1), triosephosphate isomerase, and trans-2,3-enoyl-CoA reductase. Five of our previously identified targets against **3-1a** (FASN, GAPDH, β-tubulin, Annexin A2, and HSP90AB1), also emerged as common hits of our new probes. Not surprisingly, some other lipid and/or fatty acid metabolism related proteins, such as palmitoyl-protein thioesterase (PPT1), carnitine O-palmitoyltransferase (CPT1A), ATP-citrate synthase (ACLY), dihydrolipoyl dehydrogenase (DLD), acyl-CoA dehydrogenases (ACADM,

ACADVL), fatty aldehyde dehydrogenase (ALDH3A2), sterol O-acyltransferase (SOAT1), long-chain acyl-CoA synthetase (SLC27A2), acetyl-CoA acetyltransferases (ACAT1), lysophospholipid acyltransferase (MBOAT7), were also pulled-down by most (but not all) of our Orlistat probes.[65-71]

Recent advances in quantitative mass spectrometry have now allowed direct evaluation of drug-protein interaction based on the relative peak intensity (or the spectra count) of proteins/peptides obtained from a mass spectrometer.[72] Although semi-quantitative at best, this provides a very simple and quick method to evaluate the relative binding efficiency of a drug against its protein target.[73] Specifically, the so-called emPAI MS quantitation approach provides an estimate of protein concentration from the number of peptide sequences of a protein identified by tandem mass spectrometry.[73c,d] The abundance of an identified protein in a given pull-down sample relative to the abundance of the same protein in all pull-down samples was then calculated as the ratio of its emPAI value. These ratios where then plotted in a heat map for all protein/probe combinations (see Appendix Two). As shown in part B of Figure 3.4 (bottom graph), FAS which was successfully pulled down by all seventeen Orlistat probes showed varied emPAI values ranging from as low as 3.0 (with probe **3-3a**) to as high as 12.8 (with probe **3-1e**), thereby indicating small but relatively significant differences in the relative binding affinity between this protein and the different Orlistat probes. It should be noted that all MS-based results obtained herein (including all the putative protein hits identified and their relative emPAI values) should only be used as preliminary data. In our current study, improvement on our previous affinity pull-down

procedures was made. Nevertheless, due to the highly complex cellular environment, the intrinsic limitation of affinity pull-down assay and mass spectrometry,[72,73] false positives/non-specific proteins binders could be minimized but not eliminated entirely from our results. Consequently, proper follow-up studies and validation experiments will be needed before any biological conclusion is made for some of these protein hits.

Table 3.1. Potential unique proteins identified from MS[a]

gene symbol	protein name	probe	localization	function
HLA-A	HLA class I histocompatibility antigen, A-68 alpha chain	**3-1a**	membrane	host-virus interaction, immune response
PFN1	Profilin-1	**3-1a**	cytoplasm	cytoskeletal protein binding
RPL22	60S ribosomal protein L22	**3-1a**	cytoplasm, nucleus	protein metabolism
MIF	Macrophage migration inhibitory factor	**3-1a**	cytoplasm	cell proliferation, signal transduction
RPL12	60S ribosomal protein L12	**3-1a**	nucleus	protein metabolism
PC	Pyruvate carboxylase, mitochondrial	**3-1a**	mitochondrion	gluconeogenesis, lipid synthesis
PGM1	Phosphoglucomutase 1	**3-1a**	cytoplasm	glucose metabolism
BCAP31	B-cell receptor-associated protein 31	**3-1a**	ER, nucleus	transport, apoptosis
CLTCL1	Isoform 1 of Clathrin heavy chain 2	**3-1b**	cytoplasm	receptor-mediated endocytosis
PFKM	Isoform 2 of 6-phosphofructokinase, muscle type	**3-1b**	n/a	glycolysis
TPP2	Tripeptidyl-peptidase 2	**3-1b**	cytoplasm, nucleus	proteolysis
HLA-B	HLA class I histocompatibility antigen, B-59 alpha chain	**3-1c**	membrane, cytoplasm, nucleus	immune response
PTGES2	Prostaglandin E synthase 2	**3-1c**	membrane	fatty acid biosynthesis, lipid synthesis
ANP32A	Acidic leucine-rich nuclear phosphoprotein 32 family member A	**3-1c**	ER, nucleus, cytoplasm	tumor suppressor, transcriptional repression
ATP1B3	Sodium/potassium-transporting ATPase subunit beta-3	**3-1c**	membrane	transport
P4HB	cDNA FLJ59430, highly similar to Protein disulfide-isomerase	**3-1c**	ER	cell redox homeostasis
HLA-C	HLA class I histocompatibility antigen, Cw02 alpha chain	**3-1d**	membrane	immune response
EIF3E	Eukaryotic translation initiation factor 3 subunit E	**3-1d**	cytoplasm, nucleus	protein biosynthesis
NOP56	Nucleolar protein 56	**3-1d**	nucleus	ribosome biogenesis
MYOF	Isoform 5 of Myoferlin	**3-1d**	membrane	phospholipid binding

GNAI2	Guanine nucleotide-binding protein G(i), alpha-2 subunit	**3-1e**	membrane	signal transduction
SCARB1	Scavenger receptor class B member 1	**3-1e**	membrane	receptor for phospholipids, cholesterol ester, lipoproteins, phosphatidylserine
STAT3	Signal transducer and activator of transcription 3	**3-1f**	cytoplasm, nucleus	transcription regulation
SCAMP3	Secretory carrier-associated membrane protein 3	**3-1f**	membrane	transport
TMEM48	Isoform 3 of Nucleoporin NDC1	**3-1f**	nucleus	transport
MYH9	Isoform 2 of Myosin-9	**3-1f**	cytoplasm, nucleus	cytokinesis
SEC61A2	Sec61 alpha form 2 isoform b	**3-1f**	ER	transport
LMNA	Isoform C of Lamin-A/C	**3-1g**	nucleus	muscle organ development
BZW1	Basic leucine zipper and W2 domain-containing protein 1	**3-1h**	cytoplasm, nucleus	transcription regulation
DNAJC11	Isoform 3 of DnaJ homolog subfamily C member 11	**3-1h**	n/a	unknown
RAB4B	Ras-related protein Rab-4B	**3-1h**	membrane	vesicular trafficking
NIT2	Omega-amidase NIT2	**3-1h**	cytoplasm	hydrolase, metabolism
KIAA0368	KIAA0368 protein	**3-1h**	ER, centrosome	protein metabolism
GTF2I	Isoform 1 of General transcription factor II-I	**3-1h**	cytoplasm, nucleus	transcription regulation
ARF1	ADP-ribosylation factor 1	**3-1h**	Golgi	transport, signal transduction
TOP2B	DNA topoisomerase 2-beta	**3-1h**	nucleus	transcription regulation
LY6K	Lymphocyte antigen 6K	**3-1j**	membrane, nucleus	useful as a tumor biomarker
LYPLA2	Acyl-protein thioesterase 2	**3-1j**	cytoplasm	fatty acid biosynthesis, lipid synthesis
GNAO1	Guanine nucleotide-binding protein G(o) subunit alpha	**3-1j**	membrane	signal transduction
ATP2A1	SERCA1B of Sarcoplasmic/endoplasmic reticulum calcium ATPase 1	**3-2a**	ER	apoptosis in response to ER stress
SH3KBP1	SH3 domain-containing kinase-binding protein 1	**3-2a**	cytoplasm	signal transduction
PLEC1	Isoform 1 of Plectin	**3-2b**	cytoplasm	actin binding
ESYT2	Isoform 4 of Extended synaptotagmin-2	**3-3b**	membrane	unknown

[a] 44 potential unique proteins were listed. References are given in Appendix 2.

One of the questions we were interested to address was whether subtle chemical modifications of the Orlistat structure leads to changes of its cellular targets, which was already hinted from the earlier *in situ* proteome profiling experiments (part A of Figure 3.4). If so, one would expect unique proteins targeted by each of the seventeen Orlistat probes could be identified form our affinity pull-down/LCMS experiments. Upon closer examination of our MS results, we found a total of forty-four unique proteins from twelve of the seventeen Orlistat probe used (bottom graph in part B of Figure 3.4 and Table 3.1). Among these proteins are several HLA (human leukocyte antigen) class I antigens which are part of the genetic region known as the major histocompatibility complex (MHC) class I proteins,[74] and G proteins (guanine nucleotide-binding proteins) which belong to the larger group of enzymes called GTPases.[75] Specifically, HLA-A, HLA-B, HLA-C were targeted uniquely by **3-1a**, **3-1c** and **3-1d**, and GNAI2, GNAO1 were detected only with **3-1e** and **3-1j,** respectively. Interestingly, serine exopeptidase TPP2 (tripeptidyl-peptidase 2), which is essential for some MHC class I antigen presentation,[76] was selectively pulled-down by **3-1b**. STAT3 (signal transducer and activator of transcription 3) which has recently been identified as a tumor suppressor,[77] was uniquely pulled down by **3-1f**. NIT2, a omega-amidodicarboxylate amidohydrolase and a known tumor suppressor protein,[78] was pulled down only by **3-1h**. Again we caution that these MS-derived results are only preliminary findings, and needs to be further confirmed with suitable validation experiments. Nevertheless, our data concludes that subtle chemical modifications could

change selectivity profiles/cellular targets of Orlistat, and some of its analogs may be further developed into drugs targeting some of the unique proteins mentioned above.

3.2.5 Design and Synthesis of AGT/SNAP-Orlistat Bioconjugates as Organelle-Targetable Probes

Sub-cellular targeting of a drug, whereby the drug is delivered to a specific organelle inside living cells and subsequently retained there (thus greatly increasing the effective concentration of the drug), is expected to not only improve its in vivo efficacy but also minimize its off-target effects and cytotoxicity.[79] A number of drug-delivery vehicles, including the mannose cluster for mannose-receptor interaction targeting the endolysosomal pathway,[61a] cell-penetrating and cell-localization peptides,[61b,c] have been successfully used for sub-cellular delivery. We were particularly intrigued by the AGT/SNAP-tag method, invented by Johnsson and coworkers, which is widely used in the bioimaging field for highly efficient in vivo protein labeling by making use of the alkylation of Cys^{145} of human O^6-alkylguanine transferase (hAGT) with various O^6-benzylguanine suicide substrates to covalently label nuclear, cytosolic, cytoskeletal, and cell-surface recombinant proteins with fluorophores in *Escherichia coli*, yeast, and hAGT-deficient mammalian cells.[62, 80] Recently, the Johnsson group and other groups adopted this method for the sub-cellular delivery of small-molecule-based fluorescence sensors for the sensitive detection of zinc(II),[80b] calcium(II),[80c,d] and H_2O_2[80d] in cultured cells. We were interested to know whether the same concept could also be applied for sub-cellular

drug delivery. Because FAS expresses mostly in the endoplasmic reticulum (ER), modifying Orlistat to be delivered and retained in the ER should in principle greatly improve the drug's anti-tumor activity (e.g., both in potency and selectivity).

As shown in Figure 3.5, in a proof-of-principle study, we envisaged a drug-delivery system consisting of two critical components: a benzylguanine (BG)-derivatized Orlistat (**3-4**) and a model mammalian cell line over-expressing a SNAP protein fused to a desired protein localization sequence (PLS). Organelle-targetable probe **3-4** was designed on the basis of our previous findings that, having a terminal alkyne handle on the right-hand aliphatic chain and structural variations on the left-hand aliphatic chain of orlistat do not significantly interfere with the drug's inhibition on FAS. Therefore, we were hopeful that once **3-4** enters the cells and becomes conjugated by the SNAP protein (part B of Figure 3.5), the resulting AGT/SNAP-Orlistat conjugate would be locally "concentrated" and retained in the desired sub-cellular organelle (part A of Figure 3.5). Subsequently, sub-cellular protein targeting may take place. Compound **3-4** was conveniently prepared via click chemistry between benzylguanine azide (**3-16**), synthesized in two steps from 6-chloroguanin, and a previously reported compound **3-17**[57b] (Scheme 4.1).

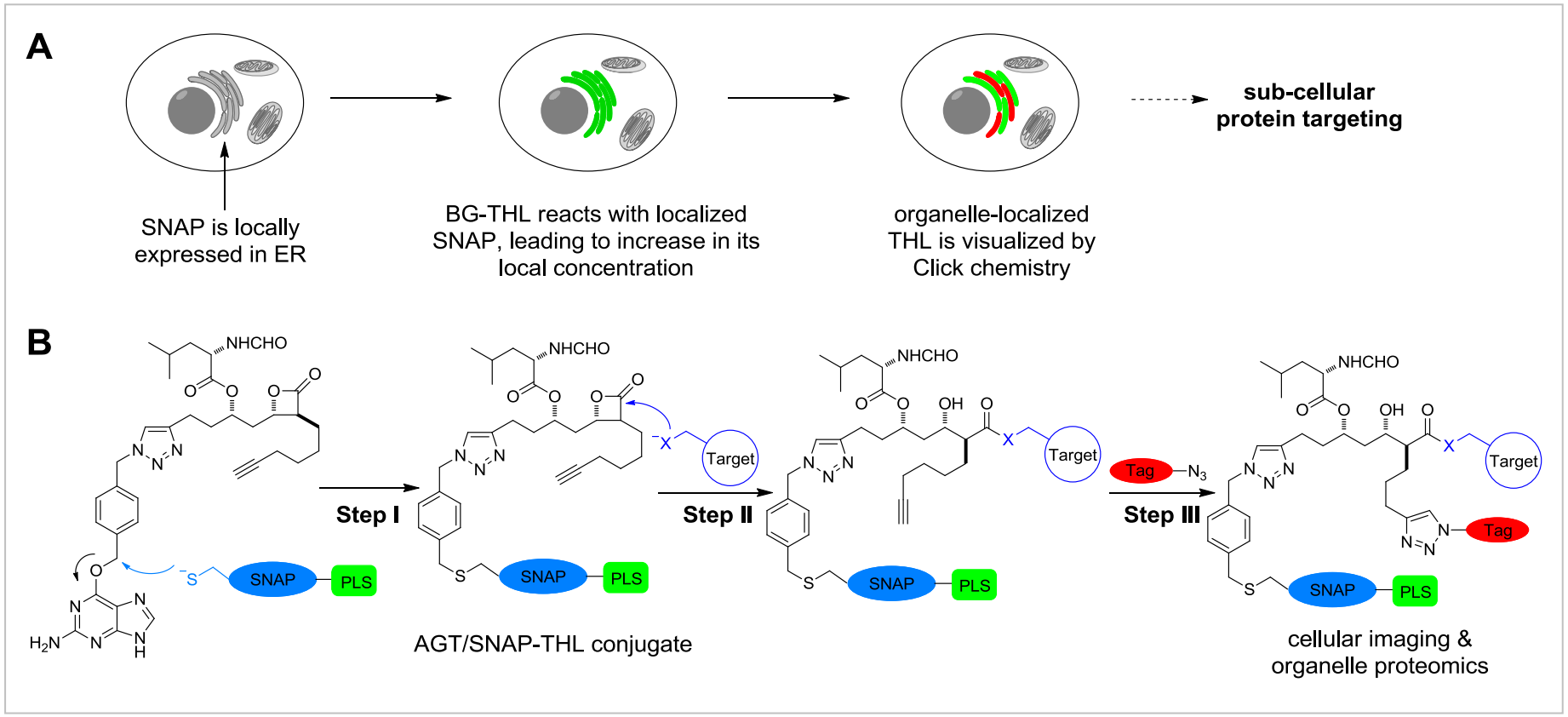
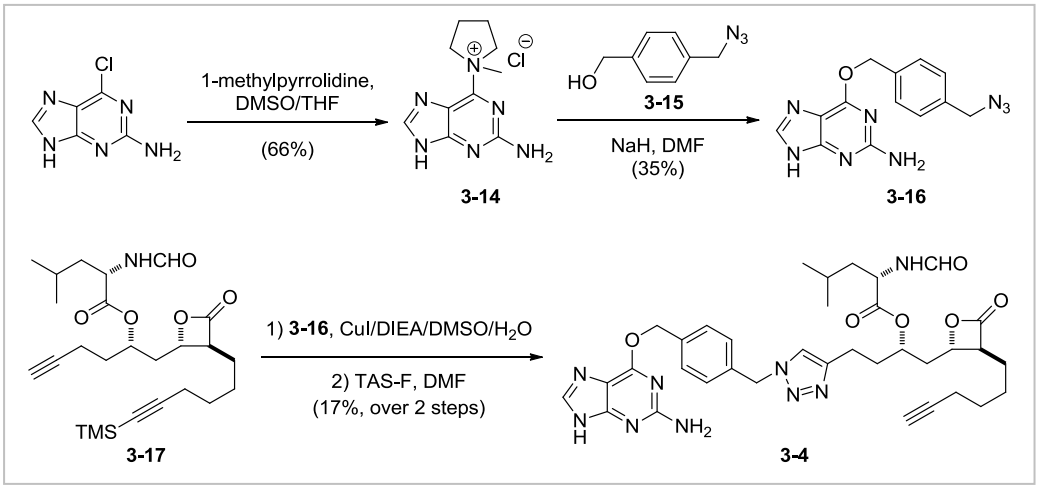


Figure 3.5. Design of AGT/SNAP-orlistat bioconjugates as a organelle-targetable probe. (A) Principle of AGT/SNAP-tag fusion strategy for sub-celllular protein targeting using a cell-permeable, benzylguanine (BG)-containing Orlistat analogue. (B) Bioconjugation reaction of AGT/SNAP+BG-Orlistat that occurs inside living cells, followed by sub-cellular protein targeting/bioimaging.



Scheme 3.4. Schematic showing of click chemistry-facilitated synthesis of the organelle-targeting, BG-containing Orlistat probe (**3-4**).

Next, we confirmed that BG-containing orlistat **3-4** could indeed serve as a substrate of AGT/SNAP and be successfully conjugated (i.e., part B of Figure 3.5, Step I). As shown in Figure 3.6, recombinantly purified hexahistidine-tagged AGT protein (His-AGT), pre-incubated with **3-4** in PBS for 4 h at room temperature, was shown to completely stop the fluoresence labeling between His-AGT and its natural substrate, **BGFL**, thereby indicating the successful formation of the desired AGT/SNAP-Orlistat (part B of Figure 3.5). Further confirmation was obtained from ESI-Q-TOF mass analusis (Table 3.2); a mass shift of 535 between the untreated His-AGT and (His-AGT + **3-4**) labeling reaction was observed, which is consistent with the calculated mass of the AGT/SNAP-orlistat conjugate. The reactive β-lactone ring in orlistat also remained intact after conjugation.

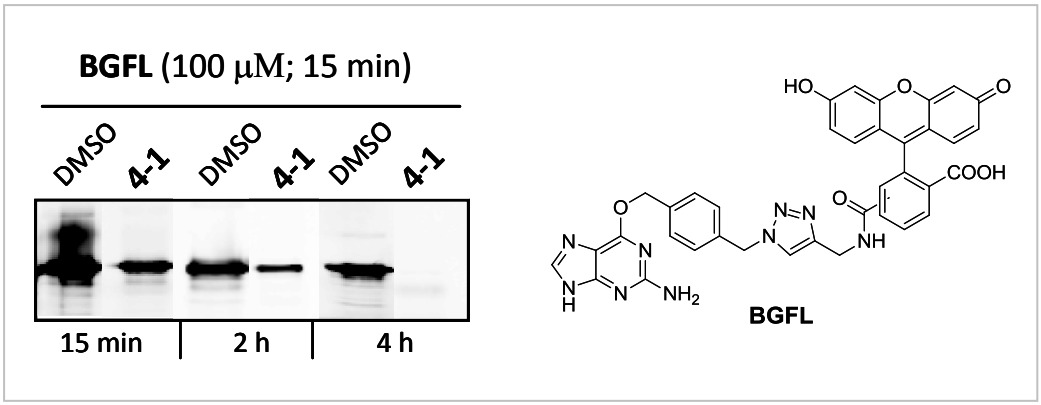


Figure 3.6. Competition assay for AGT labeling with **3-4**. Aliquots taken from the reaction mixture at the indicated time and incubated with BGFL (100 μM) for 15 minutes, followed by SDS-PAGE and in-gel fluorescence scanning.

Table 3.2. Observed Mass peaks for ESI-MS of His-AGT with or without 3-4[a].

sample	MW (Da)	mass difference
His-AGT	23606	
His-AGT + **3-4**	23701	535

[a]See Appendix 1 for original mass spectra.

Lastly, we assessed whether the direct application of **3-4** to live mammalian cells leads to successful localization and retention of the Orlistat probe in the desired sub-cellular organelle. However, we were unable to carry out experiments directly in HepG2 cancer cells, because an AGT-deficient cell line is currently not available (AGT is endogenously expressed in most mammalian cell lines. Its presence will lead to interference with our assay and high background conjugation/labelling.), and the intrinsically low transfection efficiency of HepG2 cells also precluded the accumulation of a sufficient amount of transiently expressed PLS-SNAP in the cells. Therefore, we used an AGT-deficient CHO-9 cell line instead. Sub-cellular expression of AGT-SNAP protein was performed by transient transfection of the CHO-9 cells

with mammalian DNA constructs containing the AGT-SNAP gene fused to different protein localization sequences (PLS; including mitochondria, ER & nucleus). Forty-eight hours later, after the proteins were successfully expressed and detected in the corresponding organelles (Figure 3.7; green channels), Compound **3-4** was directly applied to the media and the (SNAP + **3-4**) conjugation reaction was left to proceed for another four hours followed by extensive washing of the cells (to remove excessive probe). Subsequently, in situ click chemistry was carried out in these live cells, as previously described in Chapter 1, to visualize the localization/retention of **3-4** (part B of Figure 3.5, Step III). As shown in Figure 3.7 (red channels), the successful accumulation of **3-4** in the desired organlles (mitochondria, ER and nucleus), as guided by the localization of AGT-SNAP, was observed (see merged images of the green and red channels). Cells transfected with a control plasmid containing AGT/SNAP gene without a protein localization sequence displayed a uniform fluorescence labelling throughout the entire cell (data not shown). Taken together, these data show that Orlistat probe **3-4** could be successfully delivered/retained in sub-cellular organelles by using the AGT/SNAP-tag strategy.[12d] At present, owing to the lack of a suitable AGT-deficient HepG2 cell line, we were unable to carry out subsequent in situ proteome profiling/LCMS experiments for target identification.

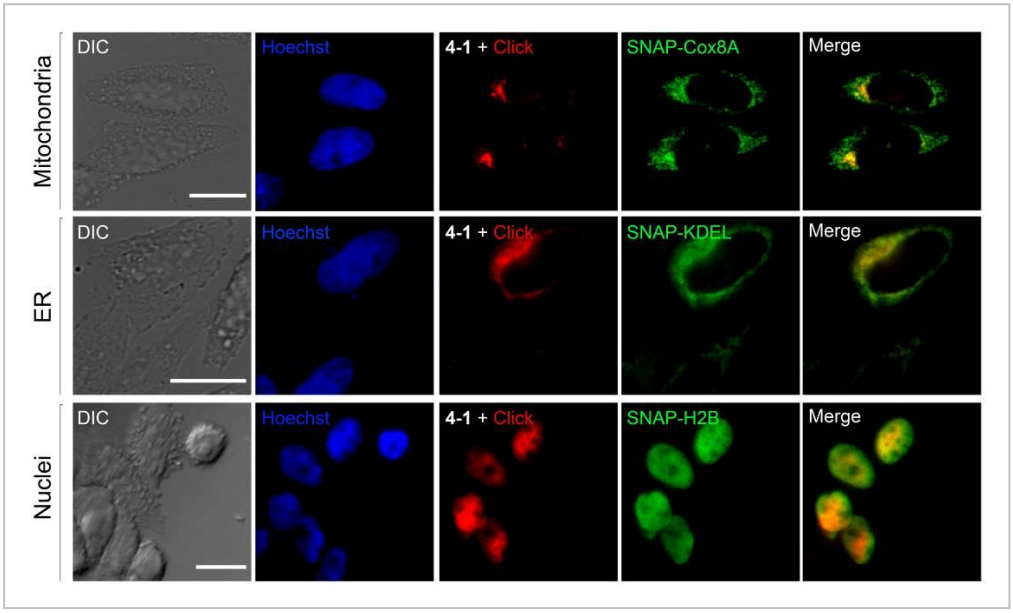


Figure 3.7. Images of CHO-9 cells expressing AGT-SNAP-tag in mitochondria (top), ER (center), or nuclei (bottom), then treated with compound **3-4**, fixed, permeabilized, clicked with rhodamine-azide (**2-33**), and imaged (red channels). Cell nuclei were stained with Hoechst (blue channels), anti-FLAG M2 primary antibody, followed by FITC-conjugated antimouse IgG antibody to detect AGT/SNAP protein (green channels). Red and Green channels were overlaid, giving merged images shown. Scale bar = 10 μm. PLS: Cox8A (mitochondria); KDEL (ER); H2B (nucleus).

3.3 Conclusion

We have synthesized and evaluated an expanded set of Orlistat-like compounds for their anti-proliferative activities in HepG2 cancer cells. Our results showed some of analogs (**3-1a-d**, **3-1f**, **3-2a**, **3-2d**, **3-3a** & **3-3c**) display comparable activities when compared to wild-type Orlistat, while others showed marginally weaker activities. Subsequent in situ proteome activity profiling, followed by large-scale affinity pull-down/LCMS identifications of putative protein hits enabled us to identify potential common and unique targets and assess their relative binding to our probes, thus providing some preliminary structure-activity-relationship (SAR) information.

This chemical proteomics approach may be used for discovery of new cellular targets of Orlistat and the design of better Orlistat analogs. It may provide a general method for both on- and off-target identification of other suitable drugs. Finally, in order to further improve the cellular activity of Orlistat, we have successfully applied the well-established AGT/SNAP-tag technology to a benzylguanine (BG)-containing Orlistat variant (**3**-**4**), and in a proof-of-concept bioimaging experiment, showed the drug could be delivered and effectively retained in different subcellular organelles of living cells. We believe the AGT/SNAP-tag strategy represents an interesting and viable tool for sub-cellular drug delivery and organelle targeting. It might even be used in future for in vivo organelle proteomics applications.[33]

Chapter 4

Parasite-Based Screening and Proteomic Profiling Reveal Orlistat™, an FDA-Approved Drug, as a Potential Anti -*Trypanosoma brucei* Agent

The content of this work has been distilled and submitted for publication. The authors are P.-Y. Yang, M. Wang, K. Liu, M. H. Ngai, O. Sherrif, M. J. Lear, S. K. Sze, C. Y. He, and S. Q. Yao.

Abstract

Trypanosoma brucei is a parasite that causes African sleeping sickness in humans and Nagana in livestock and is transmitted by the tsetse fly. There is an urgent need for development of new drugs against African trypanosomiasis due to the lack of vaccines and effective drugs. One potentially rapid and cost-effective strategy for discovering new trypanocidal drugs is to extend the indication of existing drugs already approved for other uses. Orlistat™, also known as tetrahydrolipstatin (THL), is an FDA-approved anti-obesity drug with potential anti-cancer activities. Its trypanocidal properties have not been sufficiently explored, and possible cellular targets of this drug has remained completely unknown in the two proliferating stages of *T. brucei*, the blood stream form (BSF) in the mammalian hosts and the procyclic form (PCF) in the insect vector. In Chapter 2 and 3, we developed a novel chemical proteomic approach, based on Orlistat-like small molecule probes, for large-scale proteome-wide identification of unknown cellular targets of Orlistat in human hepatocytes. Herein, we describe, for the first time, their biological evaluation in both BSF and PCF *T. brucei*. Furthermore, we carried out proteomic profiling of these compounds, and have successfully identifies many putative cellular targets of THL-R (a close mimic to Orlistat) in both parasitic forms, some of which are highly promising drug targets, potential diagnostic markers and possible vaccine candidates. In addition, we demonstrated that these Orlistat-like probes, when combined with the bio-orthogonal click chemistry and fluorescence microscopy, provide a unique and highly effective chemical tool to study drug uptake and distribution in parasites.

Given the economic challenges of de novo drug development for neglected diseases, we hope our findings will stimulate other research groups to acquire additional information necessary for conversion of orlistat-like compounds into new trypanocidal drugs.

4.1 Introduction

Human African trypanosomiasis (HAT; commonly known as sleeping sickness), caused by the protozoan parasite *Trypanosoma brucei*, is responsible for 50,000-70,000 deaths each year with increasing travel and migration within sub-Saharan Africa.[82] These trypanosomes undergo a complex lifecycle in both the mammalian bloodstream and insect vector, the tsetse fly (Figure 4.1). After a tsetse fly takes a blood meal on an infected mammal, the parasites multiply in the fly, progressing through procyclic, epimastigote, and metacyclic developmental stages in the insect gut and salivary glands. When the metacyclic trypanosomes pass from the fly to the mammalian host, they develop into long-slender bloodstream forms which multiply extracellularly in the blood. After some time, the parasites cross the blood-brain barrier to infect the central nervous system. At high parasite densities, short-stumpy bloodstream forms appear and are readily transmitted to the tsetse fly. To survive in different hosts, considerable adaptations in parasite morphology, surface composition, and metabolism including lipid and energy metabolism are required. Both procyclic and long-slender bloodstream forms of *T. b. brucei* are cultured in the laboratory. Similar to other neglected tropical diseases (NTDs), such as *Leishmania*

spp. and *Trypanosoma cruzi*, limited therapeutics for HAT and related diseases (e.g. Nagana in cattle) are available and of the drugs currently used (four drugs have been used so far to treat African trypanosomiasis, Figure 4.2), most are decades old.[82,83] Resistance and toxicity to current therapies make treatment increasingly problematic, and thus the development of new drugs is an important priority.

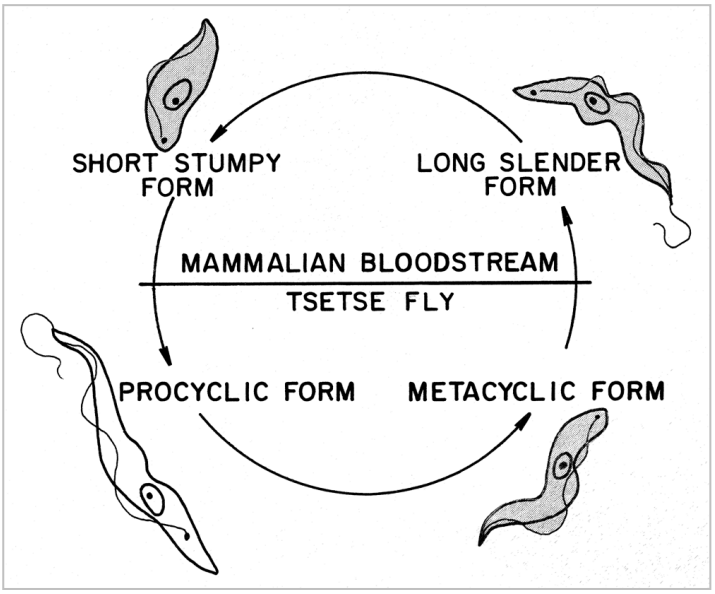


Figure 4.1. *Trypanosoma brucei* life-cycle. This simplified scheme does not show epimastigote and some other insect vector stages.

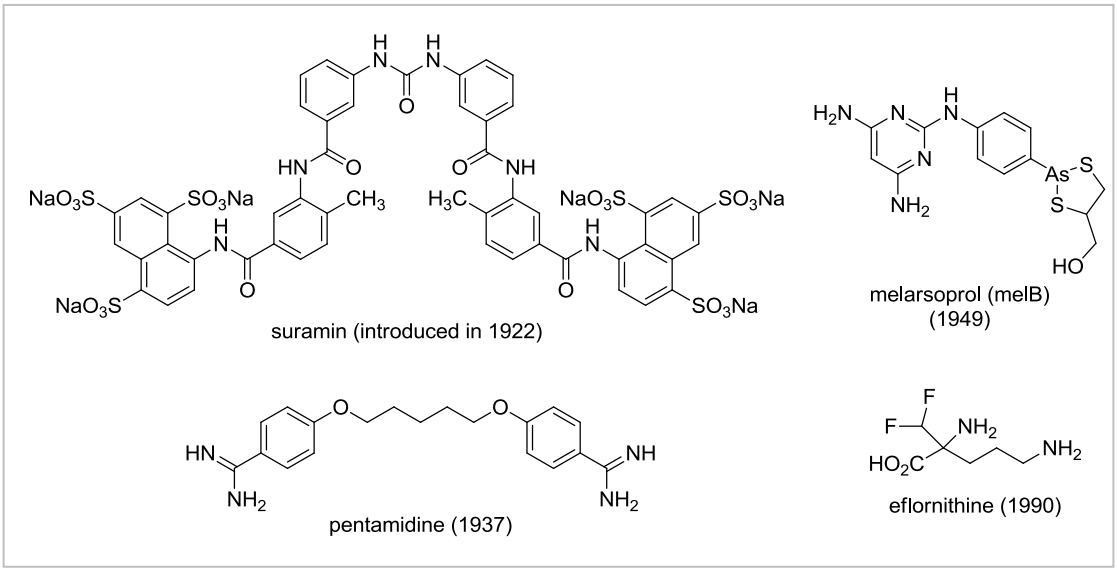


Figure 4.2. Current drugs for African trypanosomiasis.

One potentially rapid and cost-effective strategy for the discovery of new trypanocidal drugs is to explore existing drugs which possess well-characterized pharmacokinetic and safety profiles, and are already approved for other uses.[83] This 'piggy-backing' approach has historically been proven to be quite successful in bringing new therapies to the anti-parasite drug discovery pipeline. A recent example is Miltefosine, initially developed for breast cancer but now used for treating visceral leishmaniasis.[83d] Orlistat™ (marketed as Alli and Xenical; or THL), an FDA-approved anti-obesity drug which works primarily on pancreatic and gastric lipases within the gastrointestinal (GI) tract, was found to have potential anti-tumor (targeting fatty acid synthase, or FAS) and anti-mycobacterial activities.[84] In a more recent but very preliminary study using a high-throughput screening (HTS) assay, this drug together with many other tested compounds was found to also exhibit some trypanocidal activities.[85] However, no further study was carried out with live parasites and the molecular basis of this drug in the two replicating stages of *Trypanosoma brucei* remains completely unknown. Inspired by the concept of activity-based protein profiling (ABPP), which was initially coined by Cravatt and coworkers and further developed by others,[5] we previously described a chemical biology strategy which makes use of natural product-like small molecule probes for in situ proteome-wide profiling of putative drug targets. In this approach, extremely conservative modifications (i.e., an alkyne handle) were introduced in the parental Orlistat structure to provide the necessary functionality for target identification via Cu(II)-catalyzed 1,3-dipolar cycloaddition (click chemistry, CuAAC),[16] while

maintaining the native biological properties of the drug. Herein, we report the in situ proteome-wide profiling of our Orlistat-like probes in live BSF and PCF parasites for the first time. These compounds have been synthesized in Chapter 1 and 2 and first evaluated for their trypanocidal activities using a live parasite screening assay (Figure 4.3). A distinct killing kinetics was observed, with the Orlistat-like compounds killing BSF more efficiently than PCF parasites. Subsequent in situ chemical proteomic profiling was carried out, which led to the successful identification of both common and unique putative drug-binding targets of Orlistat in BSF and PCF. Bioimaging studies using selected Orlistat probes (i.e. THL-R) in combination with the bio-orthogonal click chemistry and azide-containing fluorophores enabled us, for the first time, to accurately visualize the cellular uptake and subsequent organelle-specific localization of these probes as well as associated morphological changes. From these studies, we tentatively concluded that, with its already well-established pharmacokinetic and safety profiles as an FDA-approved drug, Orlistat should be seriously considered as one of the most promising candidates for further anti-parasite drug development by pharmaceutical companies.

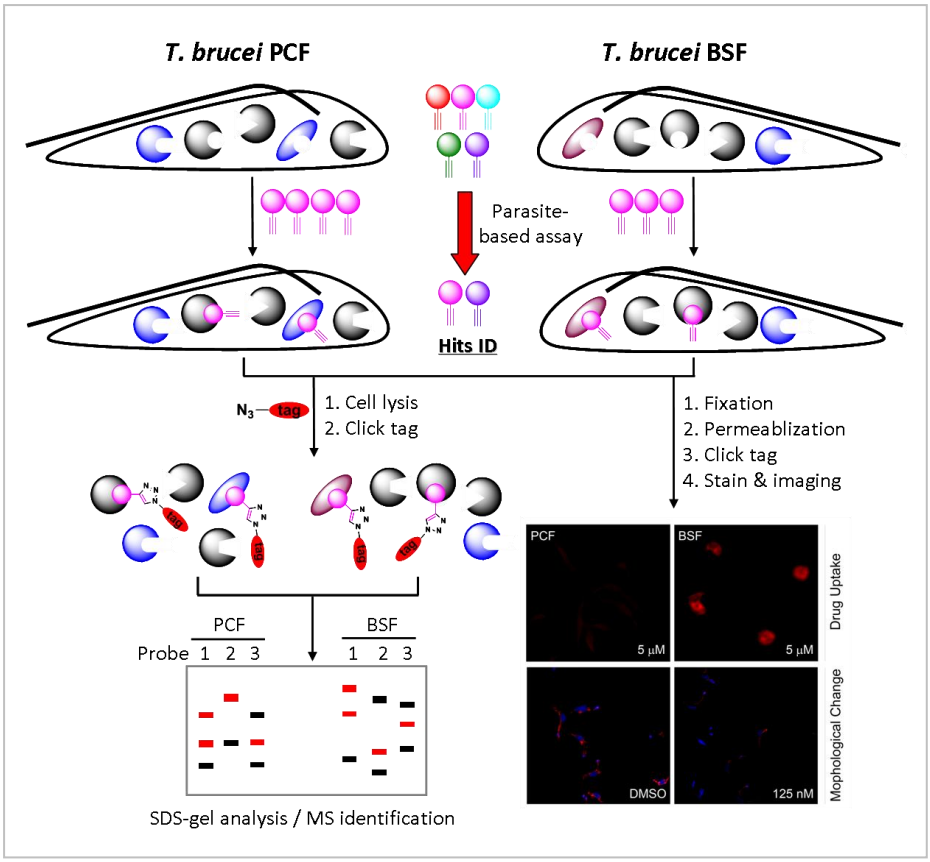


Figure 4.3. Comparative parasite-based screening and proteomic profiling of PCF and BSF with cell-permeable Orlistat-like probes. The alkyne handle of the probes enables the bioorthogonal CC to be subsequently carried out, leading to both target identification by using pull-down and LC-MS/MS techniques and imaging-based uptake studies of probes by fluorescence microscopy, respectively.

4.2 Results and Discussion

4.2.1 Trypanocidal Activities of Orlistat Compounds in PCF and BSF

We set out to consider three inter-related questions. First, does the introduction of a terminal alkyne handle in the Orlistat scaffold affect its trypanocidal activities? Second, could the relationship between the chemical structure of an Orlistat analogue and its trypanocidal activities (i.e., structure-activity relationship, or SAR) be established? Last, do these compounds have similar biological activities in both PCF and BSF trypanosomes? We first compared the cellular activity of THL-R (i.e.,

3-1a) and Orlistat against the BSF of *T. brucei*, which is the mammal-infecting form. All growth/inhibition studies were performed using the Guava ViaCount FACS (fluorescence-activated cell sorting) assay, which allows automated, high-throughput quantitative analysis of the killing effect of these compounds against live parasites (Figure 4.4); both compounds killed cultured parasites with comparable potency in a dose-dependent manner (for example, ~80% killing against BSF at 125 nM). These results confirm the imperceptible replacement of a terminal ethyl group by an ethynyl group, and that potential cellular targets identified from our subsequent in situ proteomic profiling using selected Orlistat analogs were likely targets of Orlistat itself (*vide infra*).

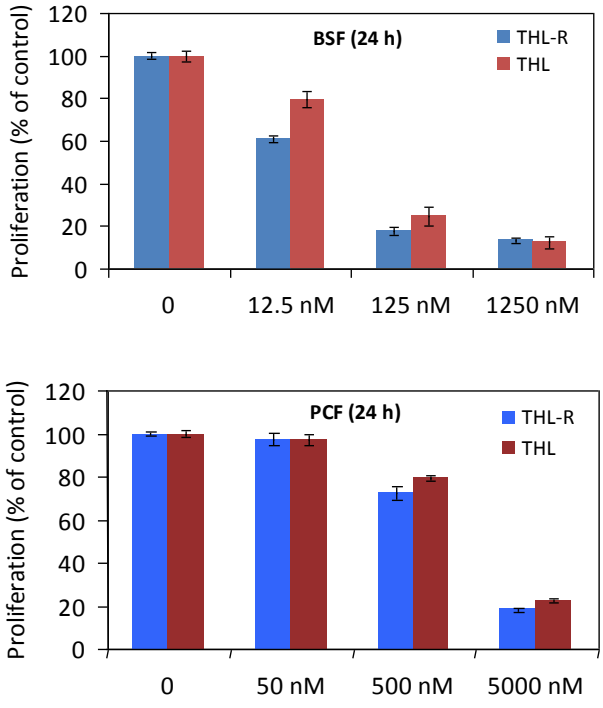


Figure 4.4. Concentration-dependent trypanocidal activities of THL and THL-R. ~1 × 10^5 *T. brucei* parasites per well were exposed to different concentrations of THL, THL-R and DMSO in 96-well plate format and then incubated at 37 ^{o}C (BSF) or 28 ^{o}C (PCF) for 24 h in 200 μL of growth medium. The effects of compounds on *T. brucei* were determined by Guava ViaCount assay. The values represent the averages of three independent experiments.

Next, we simultaneously screened a total of twenty-five potential trypanocidal compounds, twenty-two of which are Orlistat and its analogues/key intermediates (**3-1a-j**, **3-2a-f**, **3-3a-c**, **2-13** and **2-24**). The other three (Figure 4.5) are positive controls, including two well-known trypanocidal cysteine proteinase inhibitors, Z-Phe-Ala-CHN_2 (abbreviated as DMK)[86] and *N*-Mpip-Phe-Hph-VSPh (also known as K11777),[87] and a broad-spectrum lipase inhibitor MAFP.[88] Both Z-Phe-Ala-CHN_2 and K11777 are irreversible inhibitors of parasite cathepsin L and cathepsin B-like proteases and have shown promising trypanocidal activities. Notably, K11777 has passed through rodent, dog, and primate safety studies for Chagas disease caused by *Trypanosoma cruzi*. It is one of the most advanced trypanocidal agents known to-date and currently in late-stage preclinical development.[87c,d] DMSO was included in the assay as a negative control. All trypanocidal assays were initially carried out in triplicate at a single concentration (i.e. 125 nM against BSF; 5000 nM against PCF). Selected compounds were subsequently followed up with detailed dose-dependent studies.

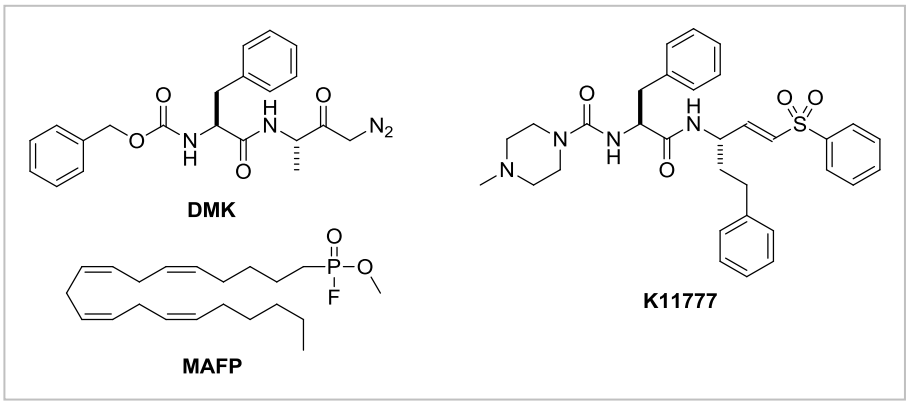


Figure 4.5. Structures of Z-Phe-Ala-CHN_2, K11777 & MAFP.

As shown in Figure 4.6, Orlistat and some of its analogs (**3-1a**, **3-1d**, **3-1e** & **3-1j**; highlighted with red bars) were found to be significantly more potent than all three above-mentioned positive controls, especially against BSF *T. brucei* (empty bars; data for PCF are represented with filled bars, respectively). It is also interesting to note that the trypanocidal activity of Orlistat and its analogues, did not appear to have originated from its inhibitory activity against lipases alone, as one might have expected (since orlistat is a known lipase inhibitor). For instance, MAFP is a broad-spectrum lipase inhibitor but displayed one of the lowest trypanocidal activities amongst all the compounds tested, suggesting that Orlistat and its analogues may function through other targets in *T. brucei*, in addition to parasitic lipases. A noteworthy structure-activity relationship (SAR) also emerged from our parasitescreening studies. In general, the *N*-formyl group and the position of the terminal ethynyl group are important for activity. Three of the compounds examined, i.e., **3-1a** (THL-R), **3-1d** and **3-1e**, were as potent as Orlistat, while **3-2a** (THL-L) and **3-3a** (THL-T) were comparatively less potent. The *N*-formylated amino ester moiety as a whole is essential for trypanocidal activity, as demonstrated by the lack of activity in the two hydroxyl lactone intermediates, **2-13** and **2-24** (in which the *N*-formylated amino ester was deleted completely). Similarly, a significant drop in the trypanocidal activity (by ~3 folds) was observed when the amino acid group in this moiety was changed from glycine to β-glycine (i.e. compare compounds **3-1c** and **3-1f**). The overall chirality of the drug did not appear to significantly affect its

trypanocidal activity. For example, in a direct head-to-head comparison, compounds with an inverted stereochemistry at C3, C4 and C6 (*S* to *R*) positions, i.e., **3-1b-d** versus **3-1h-j**, and **3-2a** versus **3-2f**, appeared to retain most of their antitrypanocidal activities. In fact, **3-1j**, together with **3-1a**, **3-1d** and **3-1e**, was one of the four most potent antitrypanocidal inhibitors identified in our assay against the BSF form of the parasites. The stereochemistry of the C-α position in the *N*-formylated amino ester moiety of Orlistat, however, had a much more notable effect on its activity, as both **3-1a** and **3-1b** (which are structurally identical to orlistat, except with a C-C to C≡C substitution at left- and right-hand aliphatic chains, respectively) appeared to be much more active than their corresponding epimers, **3-1g**, and **3-2e**, respectively. Interestingly, the relocation of the terminal alkyne handle from the right- to the left-hand aliphatic chain, i.e., **3-1b-d** versus **3-2b-d**, led to a general reduction in these compounds' trypanocidal activity. An increase in the chain length of the terminal alkyne located at the *N*-formyl end of Orlistat, i.e., **3-3a** versus **3-3b**, did not appear to dramatically affect the compound's activity.

Finally, the trypanocidal activity of all twenty-five compounds were examined and compared in both BSF and PCF forms of *T. brucei*, and our results clearly showed that all compounds tested were much more effective in killing BSF than PCF (Figure 4.6); a significantly higher dosage of the compounds (i.e., 5000 nM) was needed in order to achieve comparable killing effects in PCF. Further dose-dependent studies were carried out with selected compounds (THL and THL-R) against both BSF and PCF forms (Figure 4.7); results again showed that proliferation

of both life cycle stages of the trypanosomes was efficiently inhibited by both Orlistat (ED_{50} = 49.1 ±1.7 nM for BSF; and ED_{50} = 1.55 ±3.65 μM for PCF) and THL-R (ED_{50} = 46.2 ±3.7 nM for BSF; and ED_{50} = 1.45 ±3.56 μM for PCF). With such an effective trypanocidal profile, this immediately places Orlistat (as well as some of its analogues, i.e., THL-R) as one of the most potent drugs known to *T. brucei*. This, together with its already wellcharacterized pharmacokinetic and safety profiles (as an FDA-approved drug), bodes well for the argument that Orlistat should be immediately considered by pharmaceutical companies as one of the most promising trypanocidal drug candidates. It should be noted that most of the trypanocidal drugs currently under development focus on another highly similar parasitic protozoan *Trypanosoma cruzi*, which causes Chagas disease and infected at least 12 million people worldwide.[83]

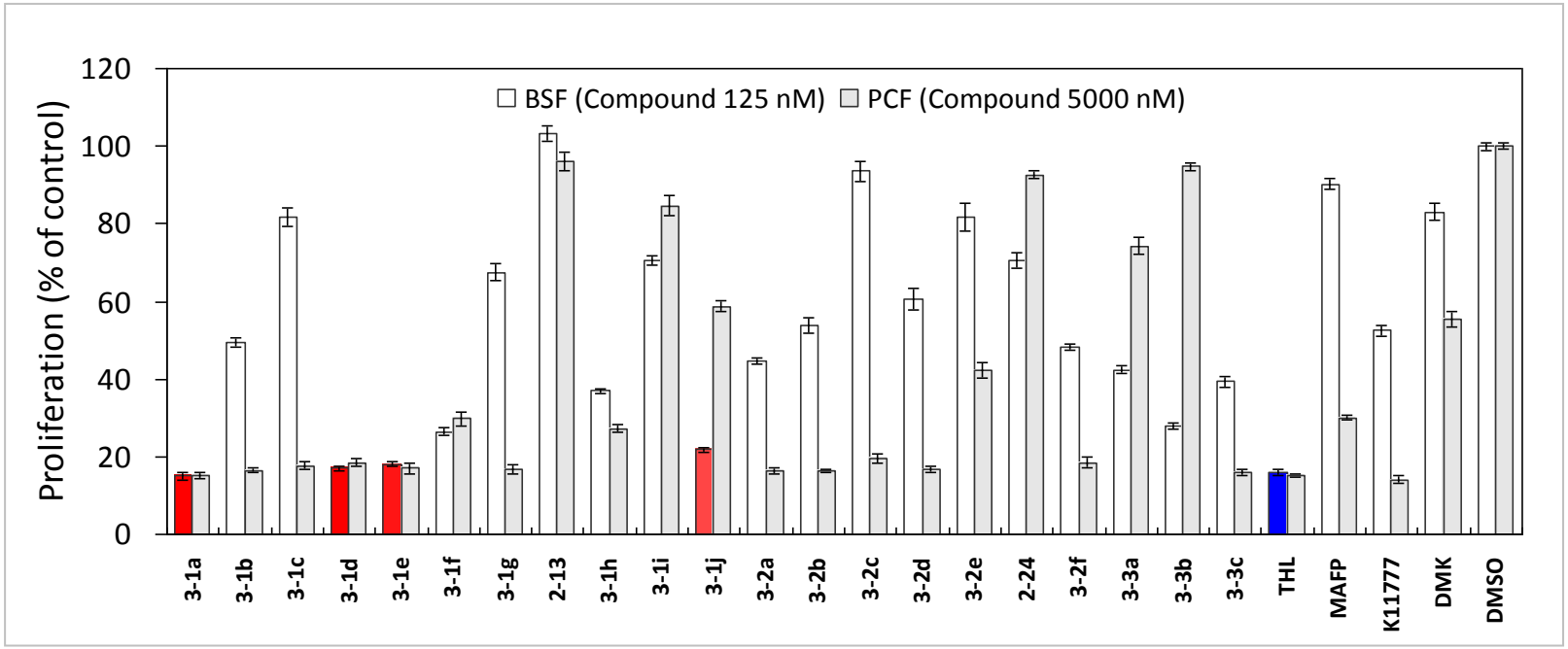


Figure 4.6. (A) A comparison of trypanocidal activities of the 21-member Orlistat library against BSF (white bars) and PCF (gray bars) after 24 h, respectively. All trypanocidal assays were initially carried out in triplicate at a single concentration (i.e. 125 nM for BSF and 5000 nM for PCF, respectively). Four probes with the highest trypanocidal activity against BSF are shown in red (**3-1a**, **3-1d**, **3-1e** & **3-1j**). BSF killing profile of Orlistat under identical conditions was shown in Blue. Results represent the average ± SD of three independent experiments.

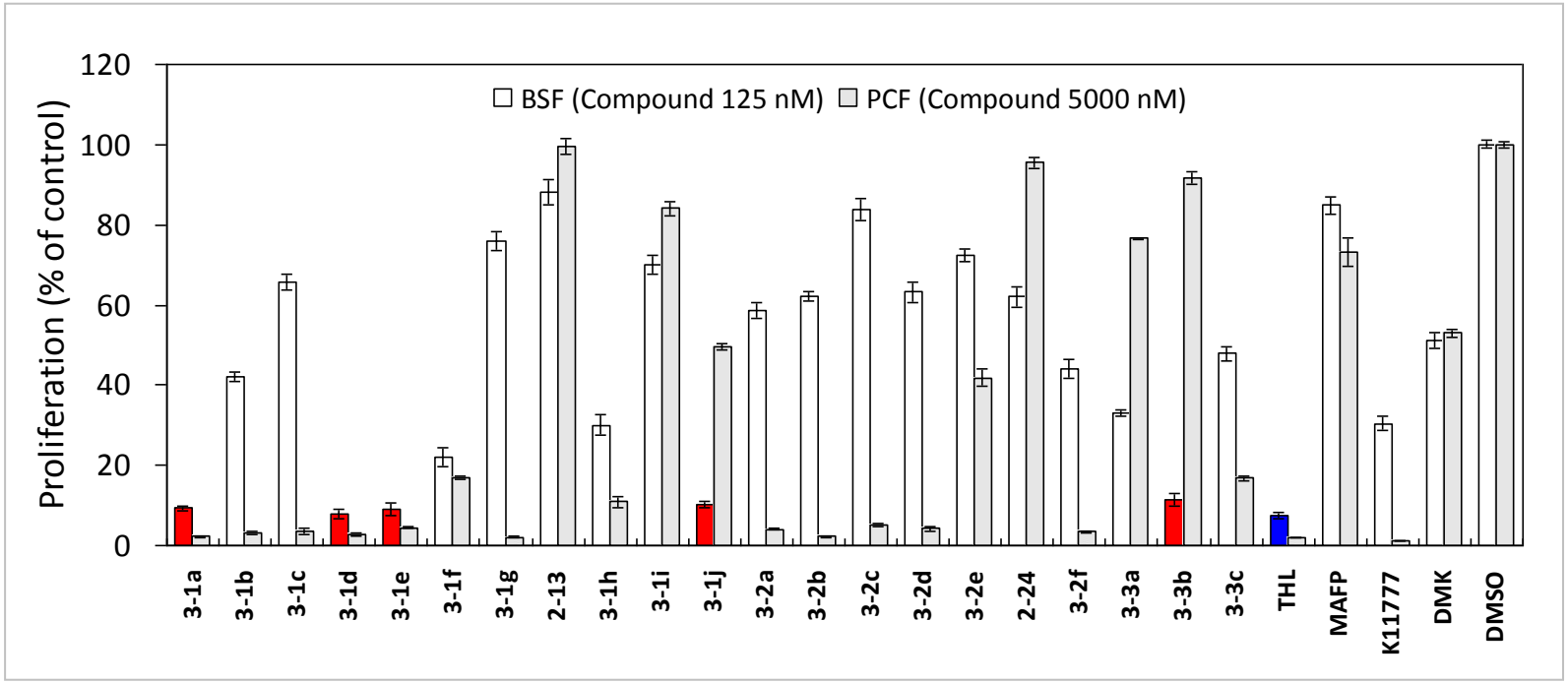


Figure 4.6. (B) A comparison of trypanocidal activitiesof the 21-member Orlistat library against BSF (white bars) and PCF (gray bars) after 48 h, respectively. Results represent the average ± SD of three independent experiments.

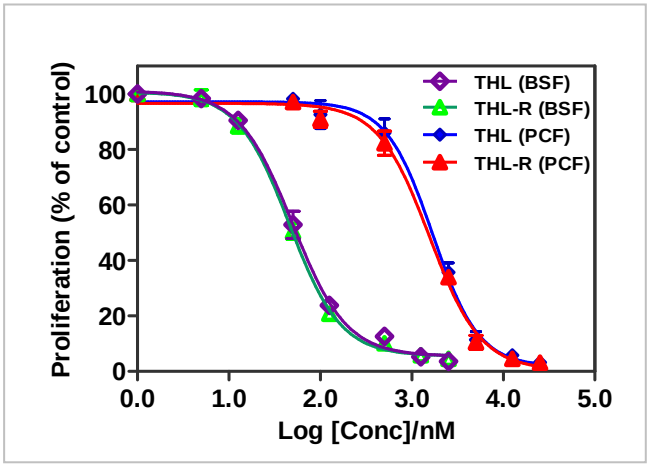


Figure 4.7. ED_{50} curves of Orlistat and THL-R (**3-1a**) against BSF and PCF. Orlistat (ED_{50} = 49.1 ± 1.7 nM for BSF; and ED_{50} = 1.55 ± 3.65 μM for PCF) and THL-R (ED_{50} = 46.2 ± 3.7 nM for BSF; and ED_{50} = 1.45 ± 3.56 μM for PCF).

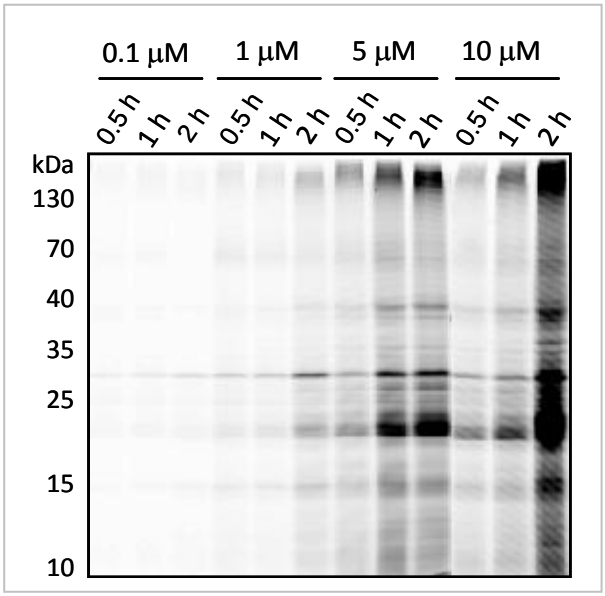
4.2.2 Comparative in Situ Proteomic Profiling of *T. brucei* parasites

The results above indicated that Orlistat and some of its analogs (i.e., **3-1a**, **3-1d**, **3-1e** & **3-1j**) conferred highly potent trypanocidal activities, especially against the blood stream form (BSF) of parasitic protozoan *Trypanosoma brucei*, which lives in the mammalian host. Comparatively, the same compounds killed the procyclic form (PCF) of the parasites less effectively. The observation that the trypanocidal activity of Orlistat was greater than that of the general lipase inhibitor MAFP suggests that, the molecular basis of Orlistat inhibition against *T. brucei* growth might have been originated from multiple sources, in addition to its suspected inhibition on endogenous parasitic lipases (Orlistat is a well-known lipase inhibitor).[89] What could be other possible molecular targets of this drug in *T. brucei*? To answer this question, we took advantage of the unique properties of our alkyne-containing Orlistat-like analogues. We had previously shown that the introduction of an alkyne handle in

Orlistat makes it possible for in situ proteome-wide profiling and identification of potential cellular targets of this drug via downstream conjugation of the protein/drug complex to reporter tags with the bio-orthogonal click chemistry. Although **3-1a**, **3-1d**, **3-1e** and **3-1j** all showed similar trypanocidal profiles against BSF *T. bruce*i, we chose **3-1a** (THL-R) for subsequent target identification as this compound is the closest structural mimic to Orlistat. To study whether in situ labeling in both intact cells was feasible, THL-R was added to the culture medium where parasites were grown, either alone or in the presence of a competing Orlistat (over a concentration range of 10-50 μM). At different time points the cells were harvested. The parasites were washed (to remove excessive probes), homogenized, incubated with rhodamine-azide (**2-33**) under click chemistry conditions, separated by SDS-PAGE gel, and analyzed by in-gel fluorescence scanning. As shown in Figure 4.8 and 4.9, the labeling was time- and dose-dependent. These results indicate that our probe can very easily enter the cellular compartments and when situated inside it selectively react with intracellular target(s). It is presently not clear how exactly the probe enter cellular compartments. Further, we compared the in situ proteome reactivity profiles of all twenty-one Orlistat probes against their cellular protein targets in live trypanosomes (Figure 4.10). In general, most probes (except **2-13** and **2-24**, which are key intermediates of Orlistat probes and lack the *N*-formyl amino ester moiety), despite obvious differences in their trypanocidal activities, showed similar in situ proteome reactivity profiles, indicating they have similar cellular targets in live parasites. On the other hand, the in situ proteome reactivity profiles between the BSF

and PCF for the same probe, though appearing similar for the most part, showed distinguishable differences (compare the two gels in Figure 4.10), pointing to the likelihood of the existence of both common and unique protein targets of Orlistat in the two different proliferating forms of the parasites. Probes **2-13** and **2-24** managed to only label a handful of proteins weakly (as shown in few labeled bands in their gel lanes), likely a reflection of their weak trypanocidal activity as well as the need for full retention of the Orlistat structure in future drug development. Most labeled bands in both the BSF and PCF profiling gels were Orlistat-sensitive, that is, their labeling was inhibited by the presence of a high-concentration competing Orlistat (Figure 4.9), while some of the labeled bands were not inhibited by MAFP (a broad-spectrum lipase inhibitor) and Cerulenin (a fatty acid synthase inhibitor), indicating that most labeled proteins are likely specific cellular targets of Orlistat, and some of which are not however lipase- and FAS-related.

A



B

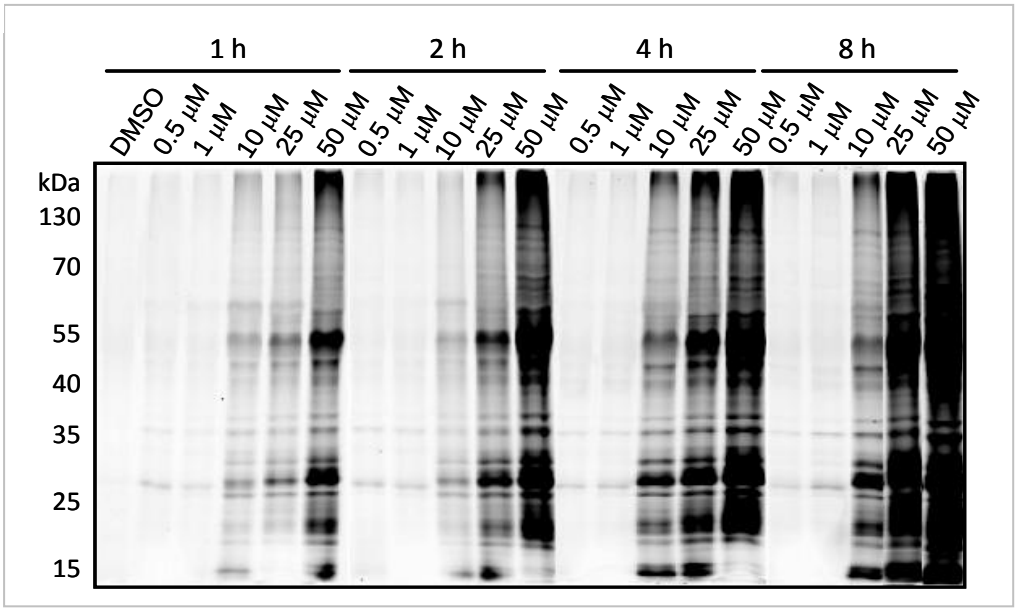


Figure 4.8. Dose-dependent and time course of *in situ* labeling/proteome profiling of the bloodstream forms (A), and procyclic forms (B) with THL-R (**3-1a**). Probe-labeled proteins were detected by click chemistry–mediated coupling to a rhodamine-azide reporter tag, followed by SDS-PAGE and in-gel fluorescence scanning.

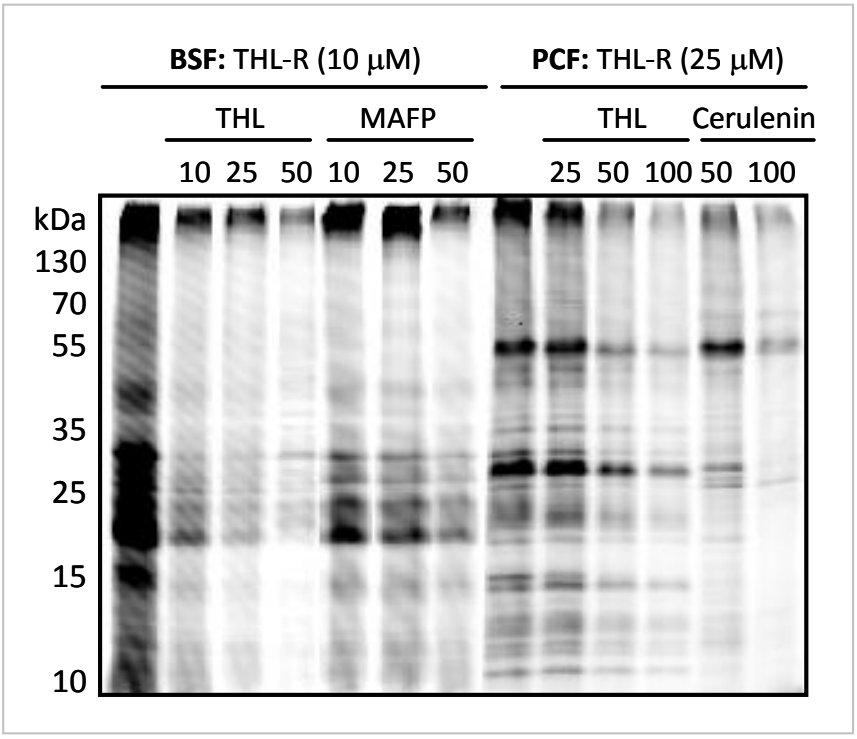


Figure 4.9. *In situ* competitive labeling *T. brucei* parasites with THL-R (**3-1a**) in the presence of different concentrations of Orlistat, MAFP or Cerulenin.

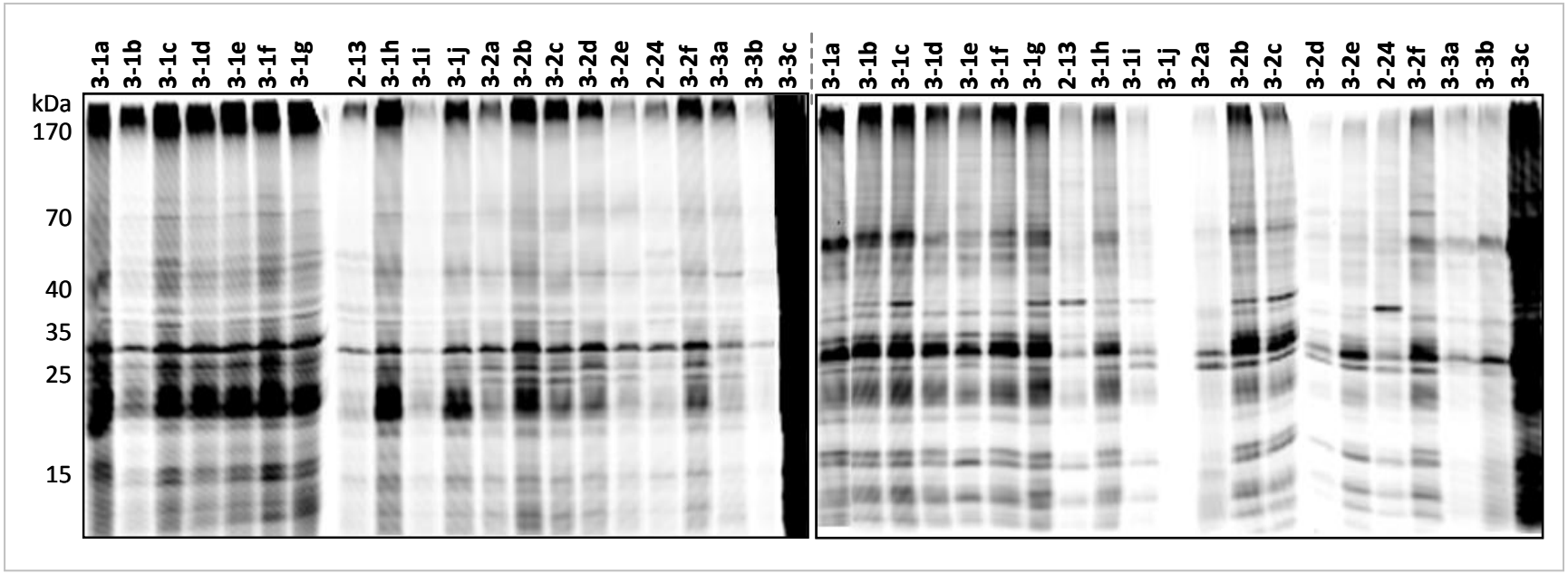


Figure 4.10. *In situ* proteomic profiling of Orlistat analogs against BSF (Left) and PCF (Right).

4.2.3 Putative Target Identification and Validation of Both BSF and PCF

Next, we performed large-scale proteomic analyses to identify Orlistat-targeted cellular proteins in both BSF and PCF parasites. Protein bands specifically labeled by **3-1a** (that is, those bands that were sensitive to Orlistat competition) in both BSF and PCF were enriched (following click-chemistry conjugation with biotin-azide **2-34**) by avidin-agarose beads, separated by one-dimensional SDS-PAGE and visualized by Coomassie staining or silver staining. The entire lane from each pull-down experiment was excised into 10 contiguous gel slices and each gel slice was individually processed for in-gel trypsin digestion as described under Experimental Section. As negative controls, the entire large-scale proteomic experiment (from cell treatment to LCMS analysis) was repeated with both BSF and PCF parasites treated with DMSO instead of **3-1a**. Peptides obtained from each gel slice were eluted and subjected to nano-LC-MS/MS analysis. The LC-MS/MS data were searched against a concatenated *T. brucei* database using an in-house MASCOT server (Matrix Science, London, UK) for protein identification. Two independent experiments identified 160 and 161 proteins in PCF and BSF samples, respectively. Of these, 29 proteins for PCF and 45 for BSF (or 18% and 27%, respectively) were hypothetical proteins, whilst 68 proteins were observed in both forms. Details of the protein identification including their score values, peptide matches and protein masses obtained on data analysis are listed in Appendix 2. The most prevalent proteins identified in both PCF and BSF of *T. brucei* were metabolic enzymes which are involved in the glycolytic pathway, amino acid metabolism, fatty

acid metabolism and lipid biosynthesis (part A of Figure 4.12). Cytoskeletal proteins that are generally known to be abundantly present in cells were also detected. In addition, some ribosomal proteins, mitochondrial proteins, and transporter proteins were identified. Various isoforms of histone proteins along with other nucleic acid-binding proteins were detected. Proteins involved in cell defense including variable surface glycoproteins and peroxidases were identified. We further grouped the identified proteins according to their known sub-cellular localizations, either theoretically predicted or experimentally confirmed (based on literature data mining) (part B of Figure 4.11); remarkably, 31-34% of all putative targets are mitochondrial proteins, 14-15% are glycosomal proteins and 12% are ER proteins, suggesting a major effect of Orlistat (and its analogues) on these organelles in *T. brucei.*

Among the proteins identified, some were inevitably non-specific protein binders of Orlistat caused by their "sticky" nature as well their high endogenous expression level, and they therefore likely offer few potential therapeutic values. Consequently, we focused our attention on other "hits" which posses known nucleophilic serine/cysteine residues in their active sites, as they are more likely to be specific targets of Orlistat. We also paid special attention to proteins which are potential diagnostic markers, drug targets and possible vaccine candidates. A deliberate effort was made to categorize potential Orlistat targets into both unique and common targets to BSF and PCF parasites. Of the interesting proteins positively identified in both BSF and PCF pulldowns are glycolytic enzymes including HK, ALD, PFK, TIM, GAPDH, glycerol-3-phosphate dehydrogenase [NAD+], PGK and

enolase.[89-96] Correct localization of glycosomal proteins had previously been shown to be essential for the survival of both BSF and PCF in *T. brucei*. Of the several mitochondrial proteins identified, alternative oxidase (AOX) is one of which detectable only in the BSF pull-down experiment. AOX functions as the sole terminal oxidase to re-oxidize NADH accumulated during glycosis,[97] and is essential for the survival of BSF parasites in their mammalian host. There is no known human homologue for this protein, making it a potentially attractive drug target for the treatment of trypanosomiasis. As shown in Table 4.1, many proteins identified only in the BSF pull-down are proteins invovled in lipid and fatty acid biosynthesis, including including GPI-PLC, inositol phosphorylceramide synthesis (SLS1), phosphatidyltransferase (PIS), essential neutral sphingomyelinase (TbnSMase), choline/ethanolamine phosphotransferase (CEPT), GPI transamidase component GAA1 (TbGAA1), GPI-anchor transamidase subunit 8 (GPI8), and glycosylphosphatidylinositol (GPI) anchor (TbGPI16). [98-102] While it is reasonable to assume these lipid-related proteins are true cellular targets of Orlistat in *T. brucei*, given this drug's well-characterized molecular targets in other biological systems. it is interesting to note they evaded our pull-down/target identification in PCF of *T. brucei*. We also identified proteins such as calpain,[86a,111] cathepsin L-like cysteine peptidase (CP),[111] cytoskeleton-associated protein CAP5.5 (CAP5.5),[112] metacaspase (MCA1)[113] and protein disulfide isomerase (BS2), all of which possess active cysteine residues and had previously been identified as targets of Orlistat and β-lactones. Sterol 14-alpha-demethylase (CYP51, also known as

3-oxo-5-alpha-steroid 4-dehydrogenase), a membrane-associated enzyme identified in both BSF and PCF pull-downs, has recently been validated as a potential target for antitrypanosomal therapy.[103] Among other potential drug targets, tryparedoxin peroxidase is a glutathione peroxiredoxin unique to trypanosomes.[104] This enzyme detoxifies peroxynitrite radicals produced by macrophages and hence plays an important role in successful evasion of host defense system. Gratifyingly, we identified two fatty acyl CoA synthetases (i.e., ASC1 and ASC3), which are the parasitic homologs of mammalian fatty acid synthetases (e.g. known targets of Orlistat) and whose cellular activities have recently been shown to be important for BSF parasites in RNA interference (RNAi)-based silencing experiments. Additionally, we identified several proteins, including fatty acyl CoA synthetase, protein disulfide isomerase, pretranslocation protein, membrane-bound acid phosphatase 1 (MBAP1), calreticulin, UDPglucose/ glycoprotein glucosyltransferase, glycerol-3-phosphate dehydrogenase, and spermidine synthase,[105-110] whose cellular activities have been shown to be important for BSF parasites in RNA interference (RNAi)-based silencing experiments. Nevertheless, owing to the highly complex cellular environment, the intrinsic limitation of affinity pull-down assay and mass spectrometry, false positives/non-specific proteins binders could be minimized but not eliminated entirely from our results. To confirm our pull-down/LC-MS/MS results, we selectively picked ten most likely “hits” of Orlistat from Table 4.1, and attempted to clone/express them in *E. coli.* as GST fusions. Five proteins that were successffully expressed in soluble forms were BS2, TbLysoPLA, MCA1, TbnSMase and GAPDH. Bacterial total lysates

over-expressed with these proteins, together with lysates over-expressed with GST only (a negative control), were directly labeled with **3**-**1a**, and subsequently analyzed by in-gel fluorescnece scanning following click chemistry with rhodamine-azide (part of C in Figure 4.11); all five proteins were successfully labeled by THL-R specifically, even in the presence of other endogenous bacterial protiens. GST alone was not labeled by THL-R (e.g. lane 1). Taken together, these results indicate that all five proteins were likely true cellular targets of Orlistat in *T. brucei*, in addition to the two fatty acyl CoA synthetases. Other "hits" identified from our pull-down experiments were also possible targets of orlistat, but could not be further validated in the present study, due to our difficulty in obtaining the recombiant proteins in bacteria. Nevertheless, our parasite-based screening, prteome profiling, target identification and validation results suggest that orlistat acts as a potent trypanocidal drug in *T. brucei* by targeting essential parasitic proteins invloved in lipid and fatty acid metabolic pathways, including lipases, fatty acyl CoA synthetases and other proteins possessing acitve serine/cysteine residues.

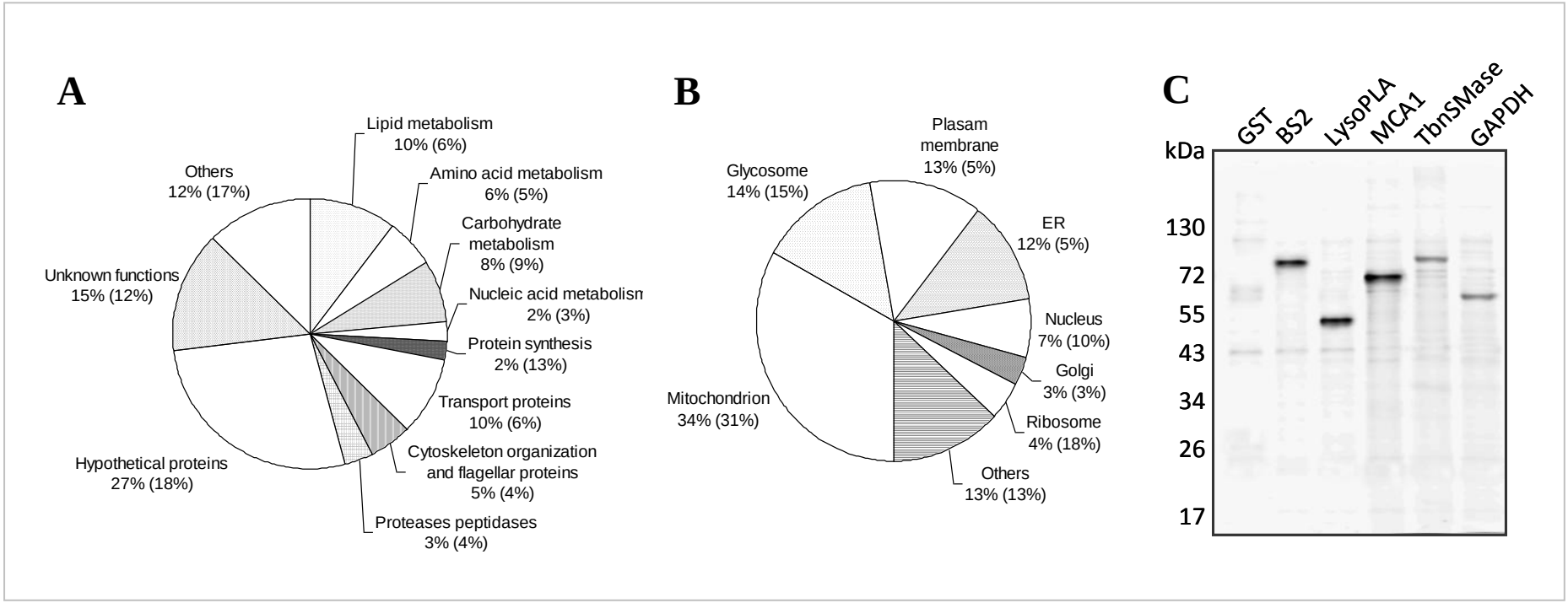


Figure 4.11. Functional classifications (A) and predicted/known sub-cellular localization (B) of identified proteins in BSF, as well as in PCF (parentheses). (C) Target validation experiments on five recombinantly expressed "hits". The labeling experiment was done with **3-1a** against the proteins expreesed in the whole *E. coli* proteome lysates. Equal amount of protein lysates and **3-1a** were used in each lane. The labeled lysates were subjected to click chemistry with rhodamine-azide, SDS-PAGE analysis, and in-gel fluorescence scanning (shown in grayscale). GST-expressing lysate (lane 1) was used as a negative control: no labeling of GST was observed. For the five "hits", all of them showed a strongly labeled, fluorescent band at the expected molecular weight.

Table 4.1. Representative proteins identified with THL-R in *Trypanosoma brucei*[a]

T. brucei gene	protein name	M_w / kDa	location	detection
Lipid metabolism				
Tb927.8.6390	lysophospholipase, putative,alpha/beta hydrolase, putative (LysoPLA)	30.09	n/a	both
Tb927.1.4830	phospholipase A1 (PLA1)	32.41	n/a	PCF
Tb09.211.3650	phospholipase A2-like protein, putative	49.91	n/a	BSF
Tb927.2.6000	glycosylphosphatidylinositol-specific phospholipase C (GPI-PLC)*	40.66	membrane	BSF
Tb11.01.6800	1-acyl-sn-glycerol-3-phosphate acyltransferase (GAT2)	30.36	membrane	both
Tb09.211.1030	inositol phosphorylceramide synthase (SLS1)*	37.97	golgi	BSF
Tb09.160.0530	CDP-diacylglycerol--inositol 3-phosphatidyltransferase**	24.11	ER	BSF
Tb927.5.3710	essential neutral sphingomyelinase (TbnSMase)*	65.02	ER	BSF
Tb10.6k15.1570	choline/ethanolamine phosphotransferase (CEPT)*	47.98	n/a	BSF
Tb10.100.0100	GPI transamidase component GAA1 (TbGAA1)*	51.31	membrane	BSF
Tb10.61.3060	GPI-anchor transamidase subunit 8 (GPI8)*	36.80	n/a	BSF
Tb10.70.2420	GPI inositol deacylase precursor (GPIdeAc)	61.43	membrane	BSF
Tb927.4.1920	GPI transamidase, putative (TbGPI16)	75.81	n/a	BSF
Tb11.01.4790	phospholipid:diacylglycerol acyltransferase-like protein	71.41	n/a	PCF
Tb927.6.1820	dolichyl pyrophosphate phosphatase*	22.24	ER	BSF
Tb927.3.1840	3-oxo-5-alpha-steroid 4-dehydrogenase**	33.36	membrane	both
Tb927.3.3580	lipophosphoglycan biosynthetic protein (LPG3)	87.77	cytoplasm	BSF
Fatty acid metabolism				
Tb927.7.4170	fatty acid elongase (ELO2)	30.49	ER	PCF
Tb927.7.4180	fatty acid elongase (ELO3)	33.87	ER	PCF
Tb09.160.2770	fatty acyl CoA syntetase 1 (ASC1)*	78.96	membrane	BSF
Tb09.160.2810	fatty acyl CoA synthetase 3 (ASC3)*	77.86	membrane	both

Tb927.8.7100	acetyl-CoA carboxylase	242.92	M	both
Proteolysis				
Tb927.8.8330	calpain, putative,cysteine peptidase**	98.46	n/a	both
Tb11.02.0730	metacaspase, cysteine peptidase, Clan CD, family C13 (MCA1)**	40.18	M	BSF
Tb927.6.1000	cysteine peptidase (CP), Clan CA, family C1, Cathepsin L-like**	48.34	n/a	BSF
Tb927.4.3950	cytoskeleton-associated protein CAP5.5, cysteine peptidase, Clan CA, family C2, (CAP5.5)**	94.65	n/a	PCF
Tb10.70.7090	serine carboxypeptidase (CBP1) precursor, serine peptidase, Clan SC, Family S10 (CBP1)	51.51	n/a	both
Tb927.10.8230	protein disulfide isomerase, bloodstream-specific protein 2 precursor (BS2)*	55.58	cytoplasm	BSF
Tb927.3.4910	signal peptide peptidase, aspartic peptidase, clan AD, family A22B	38.68	n/a	BSF
Tb10.406.0290	protein tyrosine phosphatase*	25.45	membrane	both
Mitochondrial proteins				
Tb10.6k15.3640	alternative oxidase (AOX)*	37.59	M	BSF
Tb11.02.5280	glycerol-3-phosphate dehydrogenase, mitochondrial*	66.99	M	both
Tb10.389.0890	pyruvate dehydrogenase E1 alpha subunit*	42.51	M	BSF
Tb11.02.0290	succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor*	53.1	M	PCF
Miscellaneous proteins				
Tb927.4.5010	calreticulin*	45.04	ER	BSF
Tb11.01.4701	membrane-bound acid phosphatase 1 precursor (MBAP1)*	59.38	lysosome	BSF
Tb927.4.2450	thioredoxin*	44.49	n/a	BSF
Tb11.02.5450	glucose-regulated protein 78, luminal binding protein 1 (BiP)*	71.44	ER	both
Tb11.02.4100	pretranslocation protein, alpha subunit, SEC61-like (pretranslocation process) protein*	53.65	ER	BSF
Tb927.6.4280	glyceraldehyde 3-phosphate dehydrogenase (GAPDH)**	43.87	G	BSF
Tb09.v1.0380	spermidine synthase (SpSyn)**	32.93	n/a	PCF

[a]. G, M and n/a represent glycosomal, mitochondrial and not available, respectively. Symbols in the protein name column: (*) sensitive to RNA interference; (**) putative drug target. The five “hits” chosen for further validation experiments are highlighted (in gray).

4.2.4 Cellular Uptake and Morphological Changes upon Drug Treatment

From our present studies, we have discovered a consistent trend, that is, Orlistat and its analogues killed BSF parasites much more effectively than PCF parasites. Our in situ proteomic profiling results have thus far suggested that Orlistat targets similar pathways/organelles in these two different parasites. Both common and unique molecular targets were successfully identified. While it is possible that Orlistat achieved much more efficient killing of BSF over PCF by selectively inhibiting some of the unique proteins present only in BSF (the investigation of this hypothesis is on-going), we wonder whether the difference in trypanocidal activity might also be due to differential cellular uptakes of this drug in the two different forms of parasites. With the alkyne-containing **3-1a**, which is an Orlistat mimic containing a extremely conservative C-C to C≡C structural mutation and possessing full antitrypanocidal activity, we were able to use it to accurately assess the cellular uptake and sub-cellular localization of Orlistat in both BSF and PCF parasites by fluorescence microscopy. In our bioimaging experiments, **3-1a**-treated live parasites were fixed, treated with the rhodamine-azide reporter tag under the bio-orthogonal click chemistry conditions then imaged. Immunofluorescence (IF) experiments were carried out, where applicable, to determine the sub-cellular localization of the drug, as well as subsequent morphological changes on the parasites (Figure 4.12 and 4.13). In a typical experiment, trypanosomes (1×10^5 cells/mL for either forms) were treated with 1-10 μM of **3-1a** for 2 hours. As shown in Figure 4.12 (**3-1a** was colored in Red and DAPI-stained nucleus and kinetoplast were colored in Blue), PCF parasites displayed

poor cellular uptakes of the drug - up to 10 μM of **3-1a** was needed in order to be sufficiently detected inside the parasites (e.g., right panels in Figure 4.12). On the other hand, the cellular uptake of the drug could be readily detected in BSF parasites even with only 1 μM of **3-1a** (left panels in Figure 4.12). Previous studies had suggested that BSF parasites exhibit much more rapid rates of endocytosis of lowdensity lipoproteins (LDLs), which carry phospholipids and cholesteryl esters, as compared to PCF parasites.[114] It is therefore intriguing to speculate that our Orlistat-like probes (i.e., **3-1a**), given their lipid-like chemical structure, may also enter the BSF parasites effectively via receptor–mediated endocytosis pathways. The same reason may also explain why our Orlistat-like compounds are much more potent trypanocidal agents in BSF than in PCF parasites. It should be highlighted that with the unique properties of these probes in bioimaging experiments (cell permeability, "click-able"fluorescence tagging), they allow convenient, rapid and highly sensitive monitoring of drug action in cells, and may be further developed into novel probes for positron emission tomography (PET) by clicking with reporter tags containing suitable radioisotopes (e.g., ^{18}F).

We further analyzed the morphological changes of different sub-cellular organelles in BSF and PCF *T. brucei* upon treatment with **3-1a** (125 nM and 2.5 μM for BSF and PCF, respectively, for 24 hours). Anti-TbBiP,[115] Anti-TbTrypanopain,[116] and Anti-TbSKL[117] antibodies were used to fluorescently stain the endoplasmic reticulum (ER), lysosomes, and glycosomes of the parasites, respectively. MitoTracker was used as the mitochondrion fluorescence marker. DAPI (2 μg/mL)

was used to stain the kinetoplast and nuclear DNA of the parasites. As shown in Figure 4.14, the normal network of the mitochondrion visualized with MitoTracker collapsed into a flattened sheet with a series of fenestrations upon drug treatments. The ER, glycosomes and lysosomes were barely visible compared with control cells (treated with DMSO only). These observations confirm that Orlistat has clear effects on the morphology of mitochondria, glycosomes, ER and lysosomes in both BSF and PCF cells, indicative of specific targeting of proteins localized to these organelles by the drug. These results correlate reasonably well with our above-described proteomic pull-down/LCMS data.

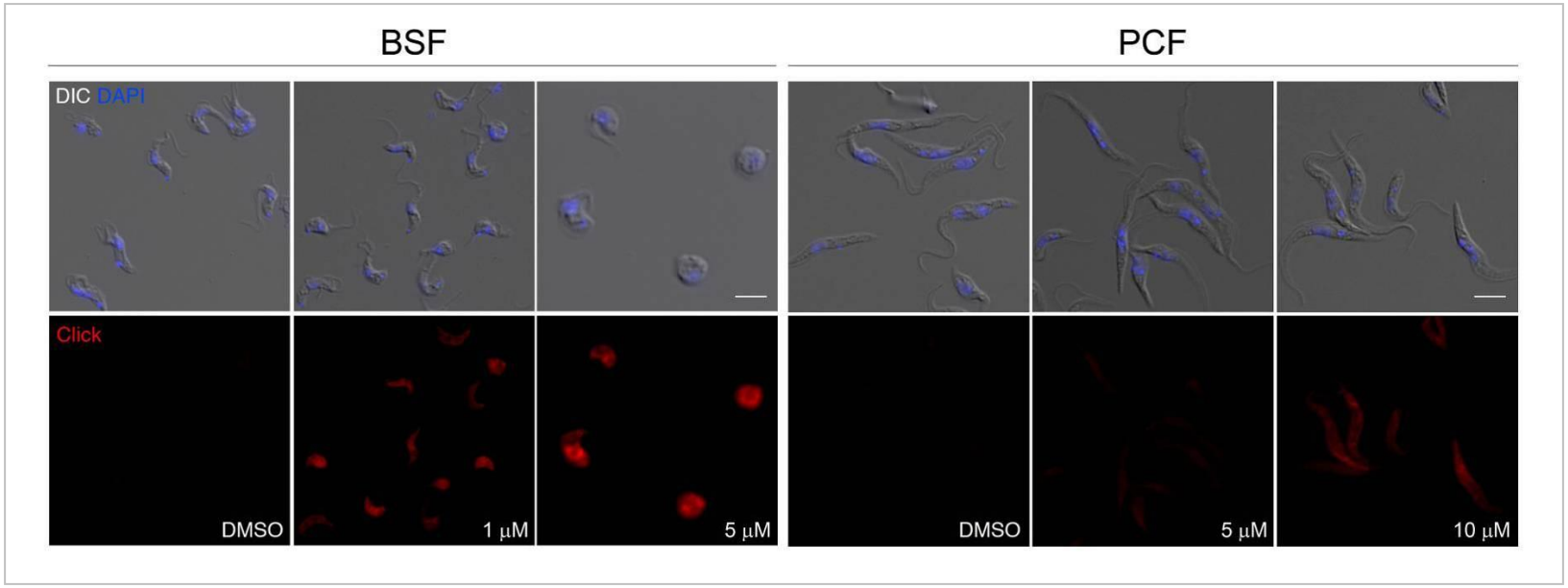
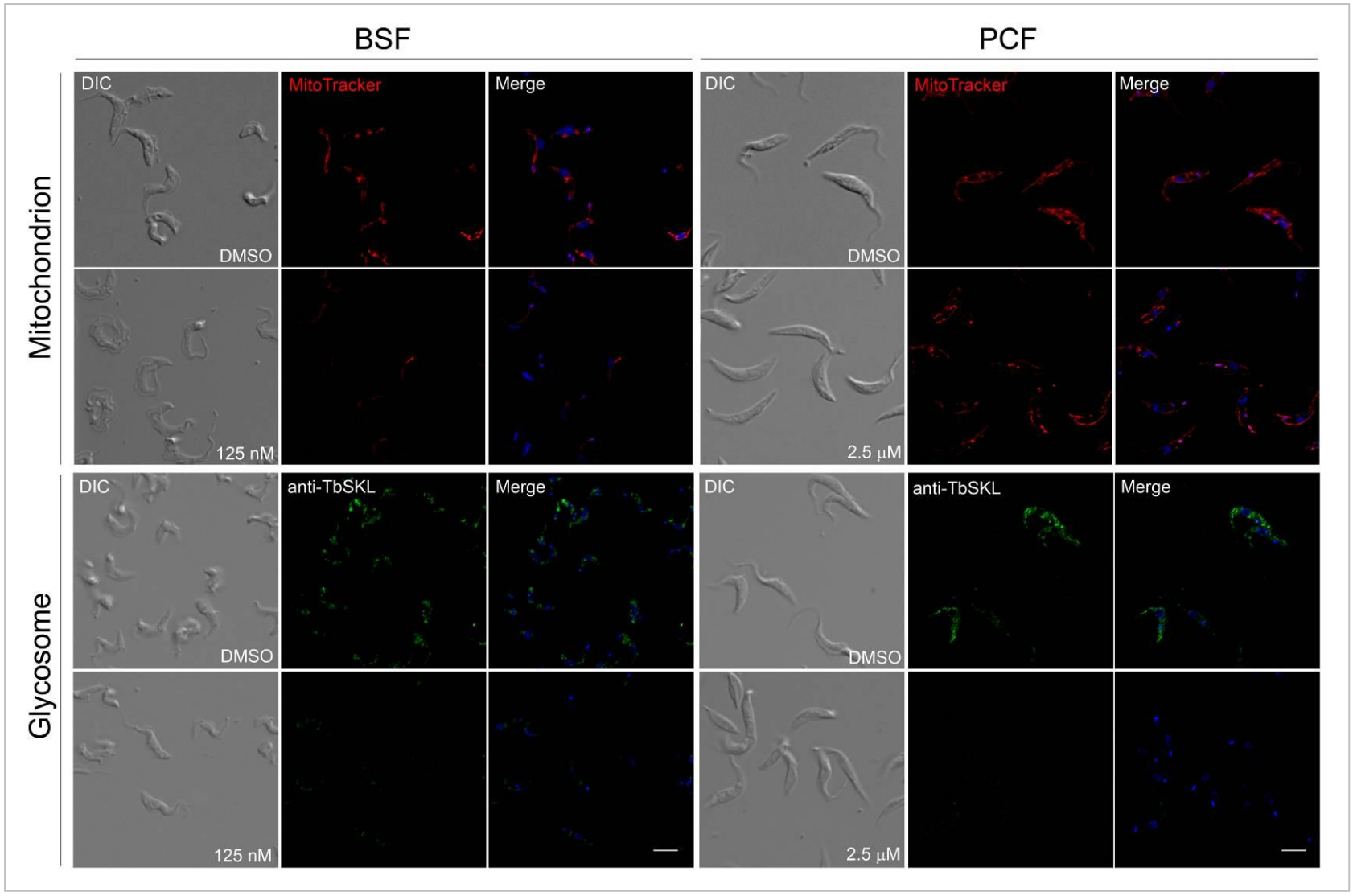


Figure 4.12. Cellular uptake of THL-R (**3-1a**) within the BSF and PCF *T. brucei*. Parasites (2×10^5 cells) were incubated with THL-R (**3-1a**) (at 0, 1 and 5 μM for BSF and 0, 5, and 10 μM for PCF, respectively) for 2 h, reacted with 10 μM of rhodamine–azide (**2-33**) under CuAAC conditions, and then imaged. (Top) DAPI stained, drug-treated parasites (with nucleus and kinetoplast colored in blue). (Bottom) Rhodamine channel showing cellular distribution of THL-R (**3-1a**) (colored in red). Scale bar represents 10 μm.


BSF
PCF
Mitochondrion
Glycosome
DIC
MitoTracker
Merge
anti-TbSKL
DMSO
125 nM
2.5 μM

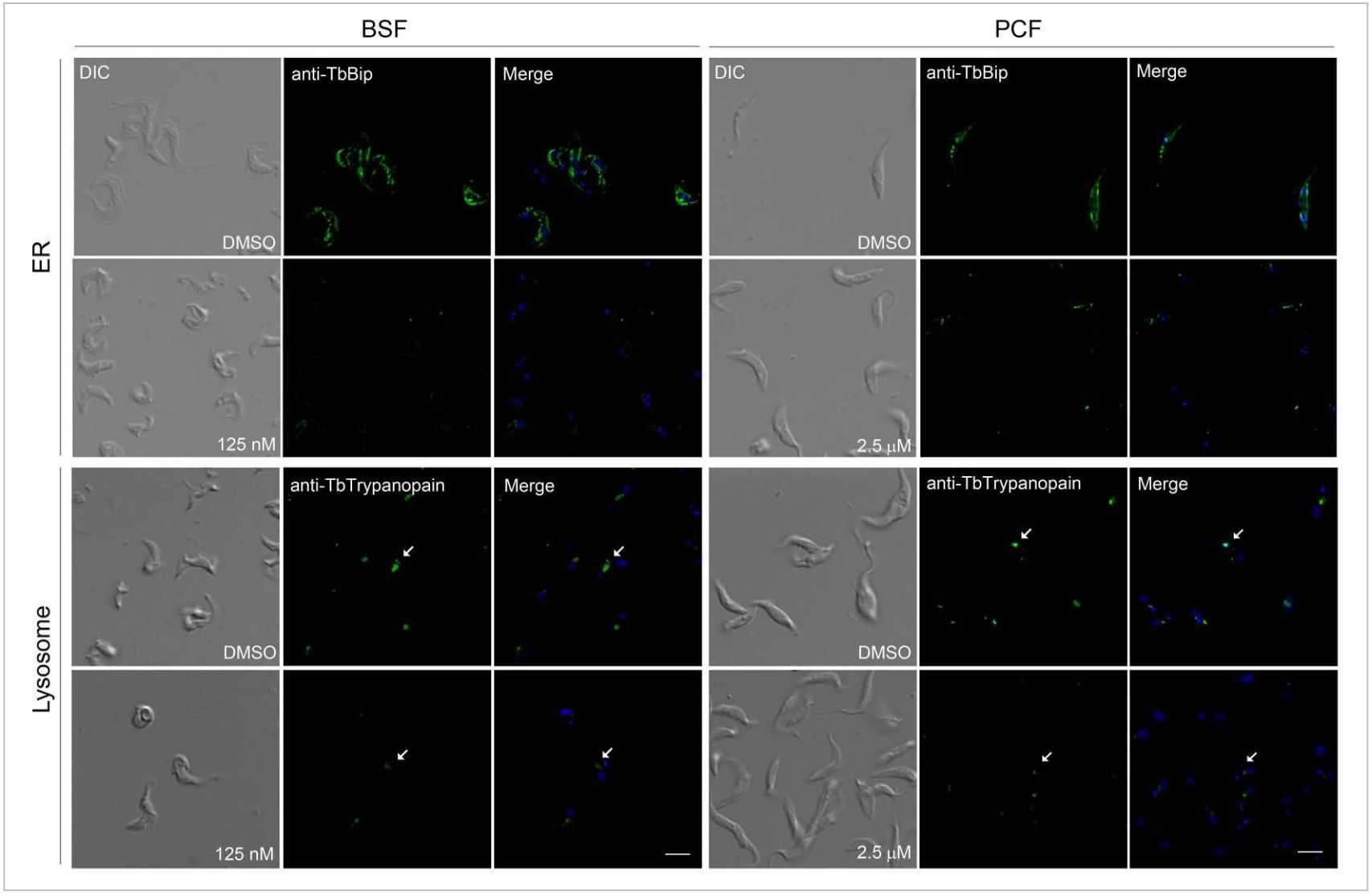

BSF
PCF
ER
Lysosome
DIC
anti-TbBip
Merge
DMSO
125 nM
2.5 μM
anti-TbTrypanopain

Figure 4.13. Morphological changes in BSF and PCF *T. brucei* treated with THL-R (**3-1a**) for 24 h (DMSO as negative controls). Immunofluorescence galleries of *T. brucei* parasites labeled with different sub-cellular organelle markers. MitoTracker Red CMXRos (Invitrogen) was used to stain the mitochondrion; anti-TbSKL was used to label glycosomes; anti-TbBiP was used to label ER; anti-TbTrypanopain was used to label lysosomes.All cells were counter-stained with DAPI to visualize kinetoplast and nucleus (colored in blue). Phase contrast images and merged fluorescence images are shown as labels indicate. Scale bar represents 10 μm.

4.3 Conclusion

We have evaluated our orlistat-like probes for their antitrypanocidal activities against both the bloodstream form (BSF) and procyclic form (PCF) of protozoan parasite *Trypanosoma brucei*. Our results showed Orlistat and some of its analogs (**3-1a**, **3-1d**, **3-1e** & **3-1j**) possess one of the most potent trypanocidal profiles known to *T. brucei*, especially against the BSF parasites. This, together with its already well-characterized pharmacokinetic and safety profiles as an FDA-approved drug, indicates that Orlistat should be immediately considered by pharmaceutical companies as one of the most promising trypanocidal drug candidates. Subsequent in situ proteomic profiling of the Orlistat-like probes enabled us to identify, for the first time, putative cellular targets of Orlistat in both BSF and PCF parasites, indicating organelle-specific targeting of the drug. Some of which were further validated by labeling of recombinantly expressed enzymes in *E. coli* lysate. Our results suggest that orlistat acts as a potent trypanocidal drug in *T. brucei* by targeting essential parasitic proteins invloved in lipid and fatty acid metabolic pathways, including lipases, fatty acyl CoA synthetases and other proteins possessing acitve serine/cysteine residues. With the unique design and properties of these Orlista-like

probes (e.g. cell permeability and click-ability), they allowed convenient, rapid and highly sensitive imaging-based experiments to be carried out with live parasites to study drug uptake and morphological changes. Our results confirmed much more efficient cellular uptakes of Orlistat-like compounds by BSF over PCF parasites. Clearly, in order to depict a better understanding of possible mechanisms and pharmacological effects of Orlistat in *T. brucei*, additional studies will be required for validation of other putative targets identified from our present study, and to further understand the biological consequency of some of these already-validated targets. Notwithstanding, from our current results, we concluded that Orlistat should be seriously considered as one of the most promising candidates for further anti-parasitic drug development.

Chapter 5

Activity-Based Chemical Proteomics Cellular Target Profiling of K11777, a Clinical Cysteine Protease Inhibitor

Portions of the work presented here are taken from:

P.-Y. Yang, M. Wang, C. Y. He, and S. Q. Yao, "Activity-Based Proteome Profiling of Potential Cellular Targets of K11777 - a Clinical Cysteine Protease Inhibitor." *Chem. Commun.* **2012**, *48*, 835-837.

Abstract

K11777 is a potent, irreversible inhibitor of parasite and mammalian cysteine protease. Its direct intracellular target proteins and potential side effects as a antiparasitic agent, however, have remained largely elusive to date. We report the design, synthesis and application of K11777-derived activity-based probes (ABPs) allowing *in situ* profiling and identification of potential cellular targets of K11777 in *Trypanosoma brucei*.

5.1 Introduction

More than a billion people suffer from neglected tropical diseases (NTDs), including malaria (caused by *Plasmodium falciparum*), sleeping sickness (aka Human African trypanosomiasis; caused by *Trypanosoma brucei*), Chagas' disease (aka American trypanosomiasis; caused by *Trypanosoma cruzi*), leishmaniasis, onchocerciasis, lymphatic filariasis and schistosomiasis.[83b,118] Most drugs available to date however are limited by parasite resistance and marked host toxicity. New strategies and targets that can effectively combat these parasitic diseases are therefore needed urgently. Equally important, from the standpoint of drug efficacy and safety, comprehensive cellular target profiling of a given drug candidate is becoming an integral step in the process of drug discovery. The knowledge of potential on- and off-targets of a drug at the earliest stages of its development will not only shed light on its potential success (or failure) during the clinical trials, but also provide invaluable insights into its mode of action and further optimizations.

One promising strategy for discovering small-molecule therapeutics for parasitic diseases has been to target the cathepsin L subfamily of the papain-like (clan CA, family C1) cysteine proteases such as cruzain, rhodesain (also called brucipain or trypanopain) and falcipains (FP-2 and -3).[119,120] Among various pharmacophores, vinyl sulfones have been widely studied as potential cysteine protease inhibitors and are entering the development pipeline as anti-parasitics (Figure 5.1).[121] One of these, K11777, which was developed from the predecessor K11002 by replacing the morpholine-urea ring in K11002 with an *N*-methylpiperazine (*N*-Mpip) for increased oral bioavailability and solubility in intestinal fluids, is currently in late-stage preclinical trials for Chagas disease.[122-124] Although it is clear that K11777 also demonstrated efficacy against other parasites as diverse as trypanosomes (i.e., *T. brucei*, *L. major*, *T. gondii*, *E. histolytica*) and schistosome bloodflukes (i.e., S. *mansoni*),[125-127 its mechanism of action is not well understood except for inhibition of the cathepsin L and/or B-like cysteine proteases in vitro. Little information is available concerning its other potential cellular targets and distribution.

I previously described a chemical profiling approach that makes use of drug-like probes for proteome-wide profiling of cellular targets in living cells. This method is based on activity-based protein profiling (ABPP), and large-scale LC-MS/MS analysis, allowing rapid determination of potential cellular targets of bioactive small molecules. Herein, I describe the first application of this method in live parasites. Several cell-permeable probes based on K11777 were designed, synthesized, and used to identify potential cellular targets of this drug candidate in *T.*

brucei.

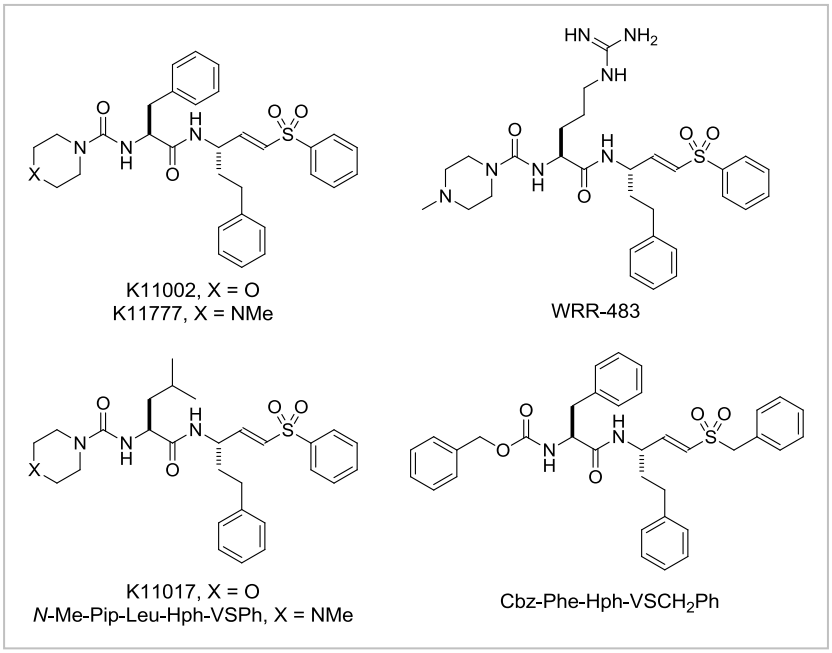


Figure 5.1. Structures of representative, anti-cruzain (K11002, K11777, WRR-483, & Cbz-Phe-Hph-$VSCH_2Ph$) and anti-malaria (K11017 & *N*-Me-Pip-Leu-Hph-VSPh) vinyl sulfones.

5.2 Results and Discussion

5.2.1 Design of K11777-like Probes

The design of K11777-like probes was based on the general structure of several trypanocidal vinyl sulfones (e.g., K11777, K11002 & Cbz-Phe-Hph-VSCH2Ph) and our previous experience with Orlistat-like probes. We took advantage of key properties of vinyl sulfones in our rational design of ABPs: (1) vinyl sulfones inhibit cysteine proteases via a Michael addition to form a covalent bond with the active-site cysteine residue (e.g., Cys 25 in the crystal structures of cruzain•K11777, rhodesain•K11777, & rhodesain•K11002; Cys 51 in FP-3•K11017);[121,128] (2) recent crystal structures of rhodesain with K11777 and

K11002 have shown that the *N*-Mpip /morpholine-urea ring didn't form any specific polar interactions with rhodesain and thus was assumed non-essential for inhibitor binding.[121,128] Accordingly, probes **5-1** and **5-2**, in which an C-C triple bond was introduced at the P3 position, were synthesized. We also synthesized probe **5-3** by introducing an C-C triple bond at the P1' position. Our experience was that such extremely conservative modifications would not compromise the native biological properties of mother compounds, whilst allowing subsequent target identification by downstream conjugation of reporter tags using bio-orthogonal click chemistry.

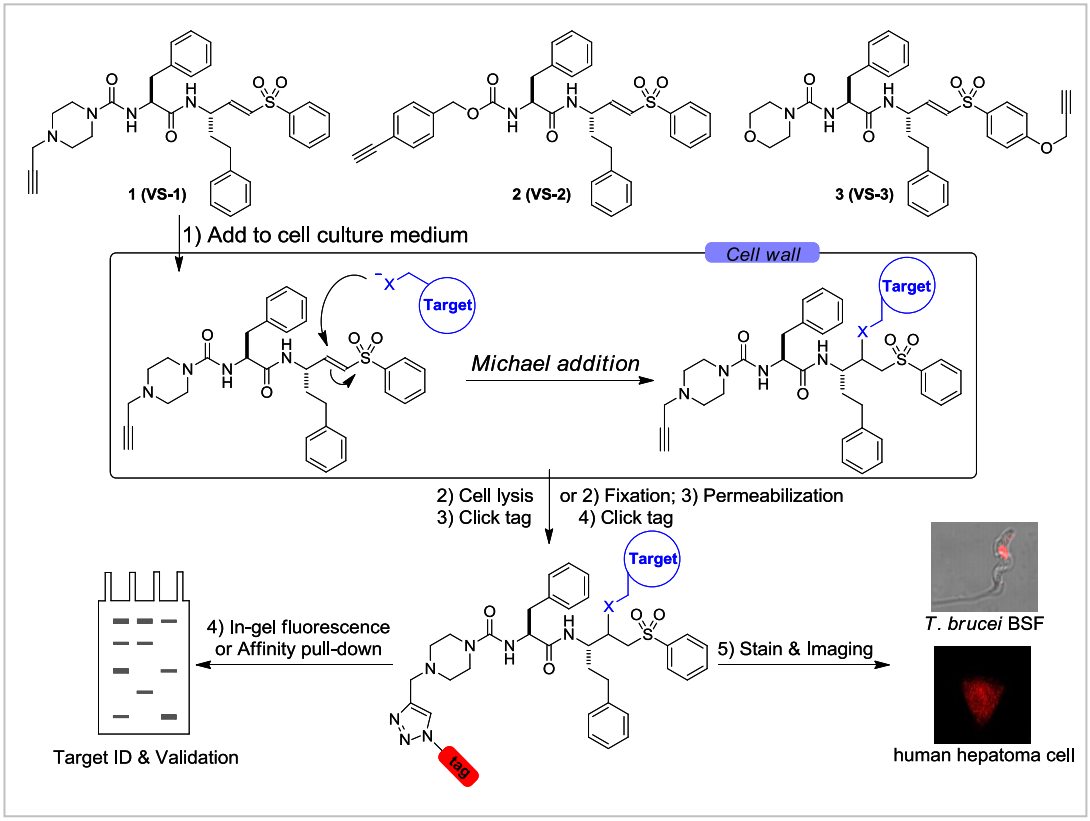


Figure 5.2. Structures of K11777-like probes (**5-1**, **5-2**, & **5-3**) and applications in *T. brucei* proteome profiling. Cell-permeable probes are directly incubated with living cells and covalently bind to the active sites of dedicated target proteins via a Michael addition. The cells are subsequently lysed, and the labeled proteins are "clicked" to a fluorescent rhodamine-azide dye for visualization/imaging or a biotin-azide for target identification by LCMS.

Superimposition of the rhodesain•K11777 and rhodesain•K11002 structures highlights structural differences that cause rhodesain to provide room for the branched *N*-Mpip. Molecular docking experiments were also carried out to ensure extra groups introduced in three analogues did not affect K11777 binding to its targets. As shown in Figure 5.3, the K11777 core binds expectedly to the enzyme pockets, with interactions closely matching those reported in the rhodesain•K11777 x-ray structure. The extra handles in the probes appeared to be projected toward the solvent-exposed surface, indicating minimal interference to the binding pockets.

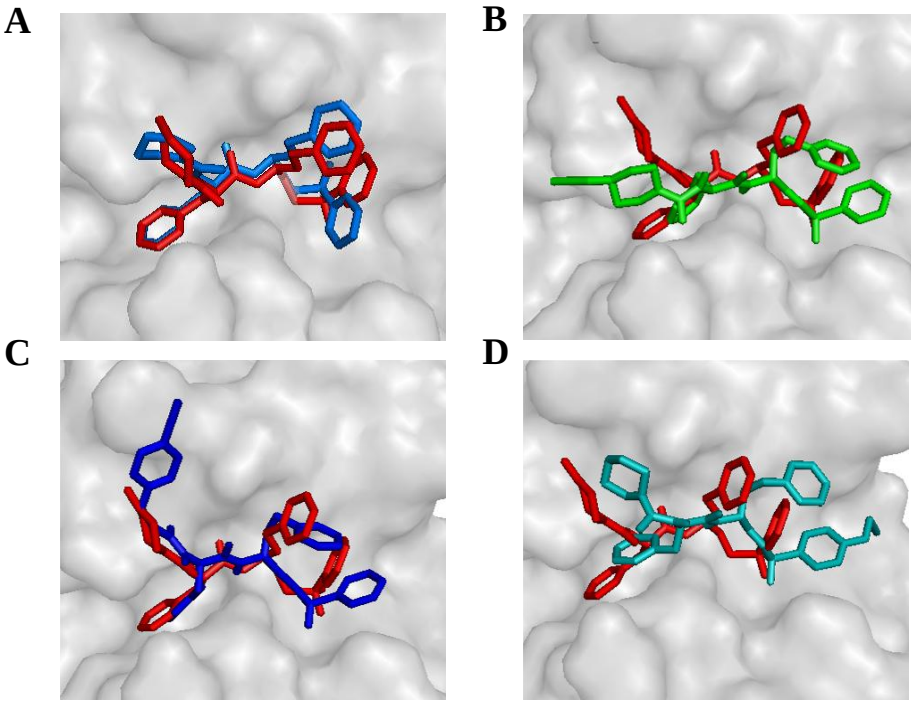
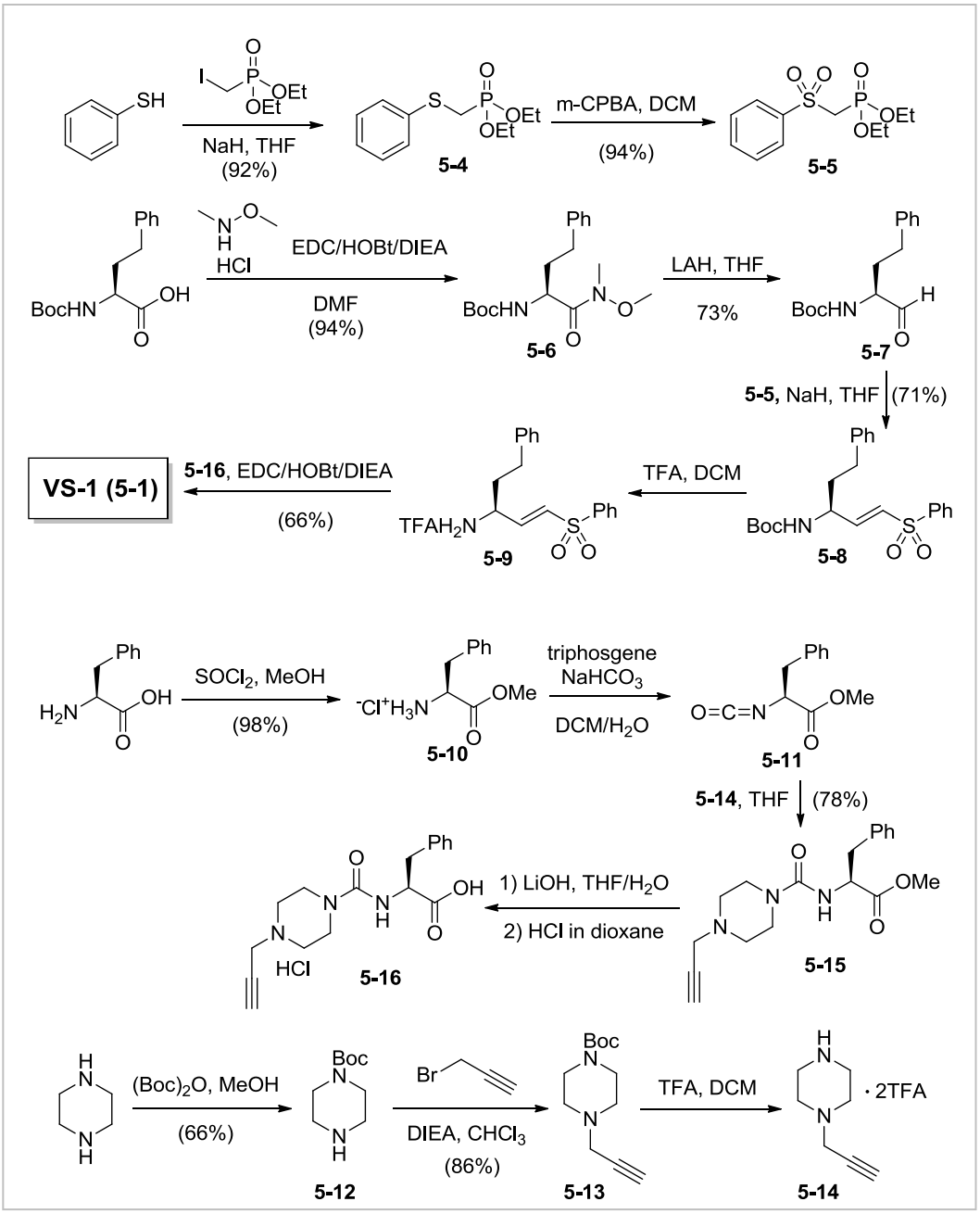


Figure 5.3. Superimposition of rhodesain•K11777 and rhodesain•K11002 (A). Surface representation of the docking of K11777 (in red) and VS-1 (in green, part B), VS-2 (in blue, part C) & VS-3 (in cyan, part D) in the binding pockets of rhodesain (the structure of rhodesain•K11777 has been described in the Protein Data Bank (PDB entry 2P7U), image generated with PyMOL.

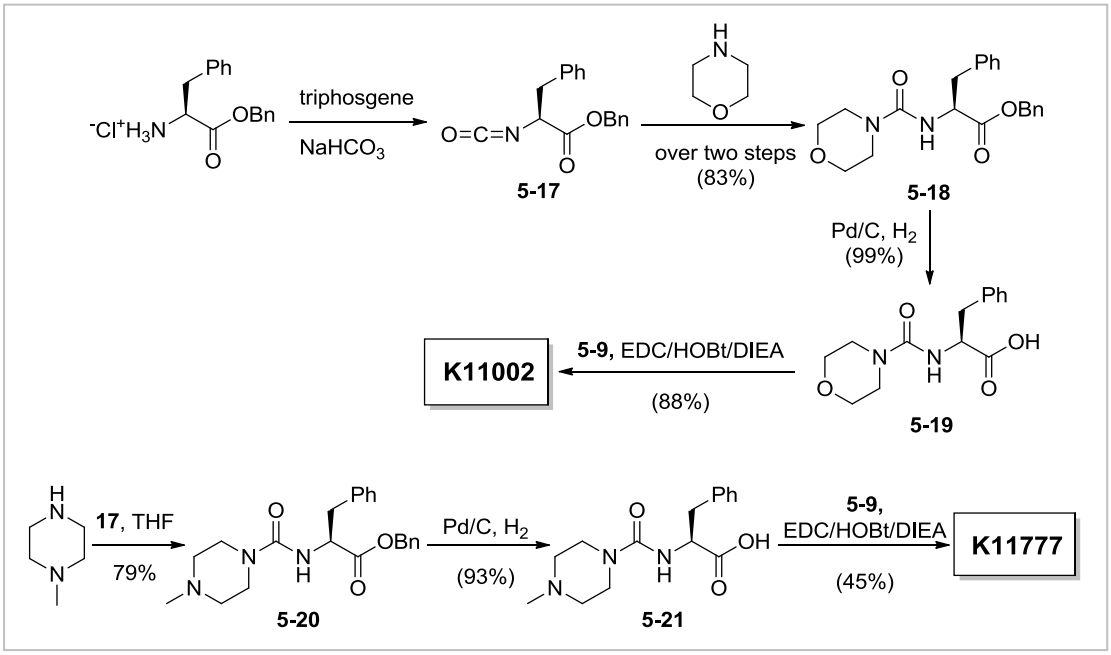
5.2.2 Synthesis of K11777-like Probes

The three analogues were synthesized by using appropriate modifications of previously described methods.[129] Reduction of the Weinreb amide **5-6** derived from commercially available Boc-homophenylalanine gave the corresponding aldehyde **5-7** in excellent yields. Horner–Wadsworth–Emmons olefination with phosphonate **5-5** (prepared in two steps involving alkylation of benzenethiol with diethyl iodomethylphosphonate and oxidation of sulfides **5-4** to sulfone **5-5** using m-CPBA) provided the vinyl sulfone **5-8** in 71% yield. Deprotection of the Boc group afforded amine **5-9** as a TFA salt. This material was coupled to **5-16** or **5-23** under standard peptide coupling conditions giving **5-1** (VS-1), and **5-2** (VS-2), repectively (K11777 and K11002 were also synthesized in parallel, see Scheme 5.2). Compound **5-16** was prepared from piperazine in five steps (only two purification steps), including mono–Boc protection of piperazine, *N*–alkylation with propargyl bromide, Boc deprotection, urea **5-15** formation via reaction of an intermediate isocyanate **5-11** (prepared in two steps involving methyl esterification of phenylalanine and generation of amino acid ester isocyanate using triphosgene) with **5-14**, methyl ester deprotection. Similar to **5-16**, compound **5-23** was conveniently synthesized from the intermediate isocyanate **5-11** with commercially available (4-ethynylphenyl)methanol, thus giving rise to the corresponding carbamate **5-22**, followed by the deprotection of methyl ester. For the synthesis of **5-3**, the known phosphonate **5-24** was first prepared as previously described.[130] Subsequent *O*–alkylation with propargyl bromide gave **5-25**. Following Horner–Wadsworth–Emmons reaction with aldehyde **6-7**, Boc deprotection, coupling

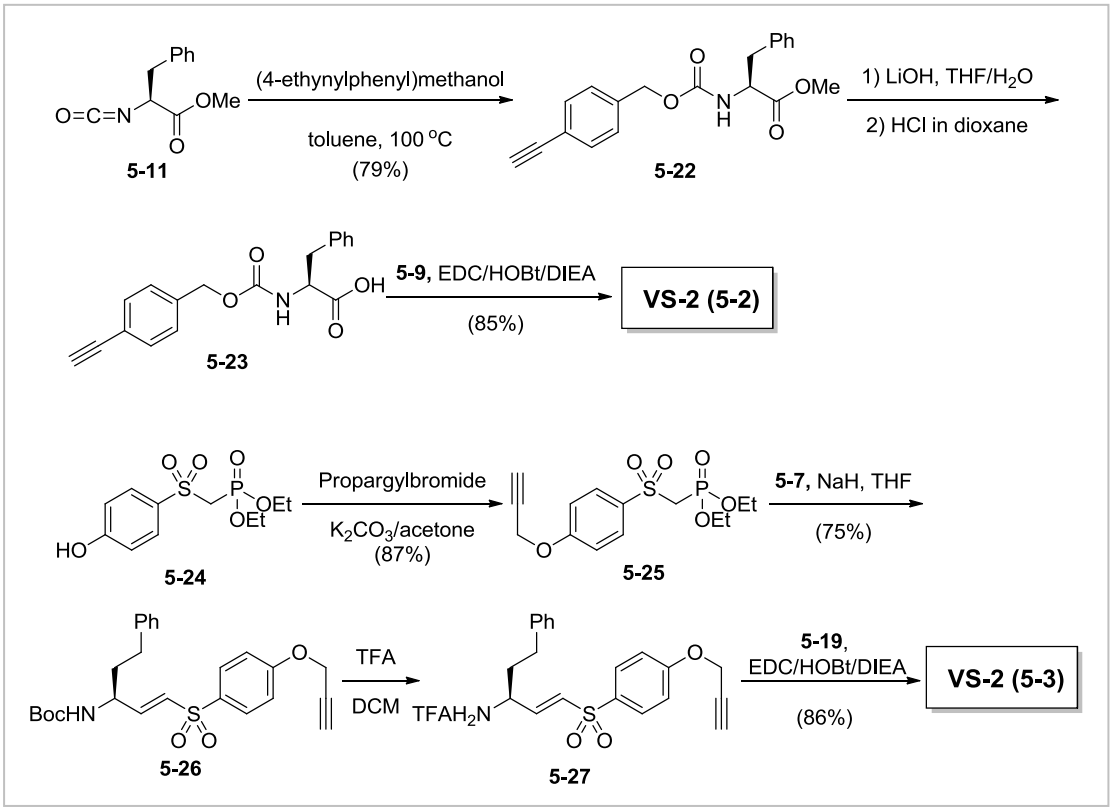
with **5-19**, which was obtained in three steps in 82% overall yield, furnished the targeted analogue **5-3** (VS-3).



Scheme 5.1. Synthesis of VS-1 (**5-1**).



Scheme 5.2. Synthesis of K11002 & K11777.



Scheme 5.3. Synthesis of VS-2 (**5-2**) & VS-3 (**5-3**).

5.2.3 Effects on Trypanocidal Activities

We first evaluated the anti-trypanocidal activities of vinyl sulfones VS-1, VS-2 & VS-3 (i.e., **5-1**, **5-2** & **5-3**), together with K11002 and K11777 in cell culture of both the bloodstream form (BSF) and the procyclic form (PCF) of *T. brucei* using the Guava ViaCount assay as described in Chapter 4, which represent two distinct stages of *T. brucei* during its complex life cycle alteration between mammalian hosts and insect vectors. In this context, it is worth noting that the BSF and PCF parasites differ extensively in morphology and metabolism. All vinyl sulfones blocked the parasite growth to a similar extent in a dose-dependent manner with ED_{50} values between 4 and 8 μM (Figure 5.4). Additionally, there was no noticeable difference in inhibition between BSF and PCF trypanosomes by all five compounds. These data show the introduction of a terminal alkyne handle at the indicated positions did not noticeably affect its trypanocidal activity. Therefore, all three probes should be suitable chemical probes for proteome profiling and cellular target identification of K11777. Only VS-1, however, was chosen for further comprehensive proteome profiling, LC-MS/MS and cellular bioimaging experiments based on its closest resemblance to K11777. The result was somewhat surprising since recent studies for cysteine proteases in the *T. brucei* show that cathepsin B like protease known as TbCatB, which is more abundantly expressed in the bloodstream versus the procyclic form of the parasite, is essential for parasite survival in both in vitro[131] and in vivo[132] models. In contrast, knockdown of brucipain by RNAi in vitro produced no detectable phenotypic changes.[131] As referenced above, the phenotypes associated with

interference of TbCatB are similar to those observed in the earlier study in which the cysteine protease inhibitor Z-Phe-Ala-CHN_2 (where Z is benzyloxycarbonyl) was used to kill the parasites. Thus, results from these in vitro RNAi studies of these two cathepsin-like proteases in *T. brucei* support the hypothesis that Z-Phe-Ala-CHN_2 kills the bloodstream form *T. brucei* by inhibiting TbCatB rather than brucipain. However, it remains to be clarified whether inhibition of the bloodstream form *T. brucei* development by vinyl sulfones (e.g., K11777) also might have been originated from TbCatB rather than brucipain, or from both of them.

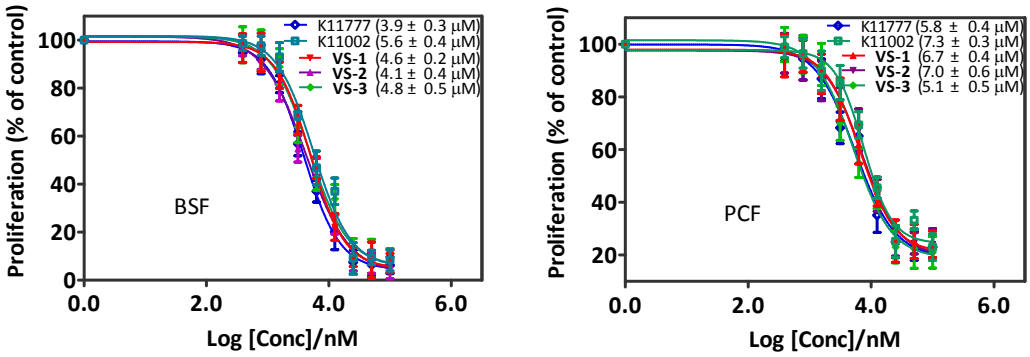


Figure 5.4. Dose-dependent trypanocidal effects of K11002, K11777 and three analogues (VS-1, VS-2 & VS-3) against bloodstream form (top) and procyclic form of *T. brucei* after 24 h. Results represent the average standard deviation for two independent trials.

5.2.4 In situ Proteome Profiling and Target Identification

Next, we compared the *in situ* proteome reactivity profiles of our probes against their cellular targets in live BSF and PCF (Figure 5.5). Generally, the three probes showed comparable in situ proteome reactivity profiles, thereby indicating they have similar cellular targets in the parasites. However, the reactivity profiles

between BSF and PCF for VS-1 (lanes 1 and 5 in part A of Figure 5.5), though similar, showed noticeable differences, suggesting the existence of both common and unique targets in the two forms. In addition to the expected rhodesain and TbCatB bands (~41 kDa and ~31 kDa indicated by arrows, respectively), which were subsequently verified by affinity pull-down/western blotting with the corresponding antibodies (part B of Figure 5.5), a number of other proteins were also covalently labeled by the three probes. These labeled bands were K11777-sensitive, that is, their labeling was blocked by the presence of a competing K11777 (lanes 2 and 6). Also evident in Figure 5.5A & 5.5B was that, VS-1 labeled only the mature active enzyme form of rhodesain (~41 kDa), but not its proform (~48 kDa).[140] In addition, TbCatB labeled by VS-1 was mostly detected in BSF, which is consistent with previous findings that TbCatB was up-regulated in BSF.[141] Subsequently, we performed large-scale proteomic analyses using VS-1 to identify potential cellular targets of K11777 in both parasite forms by affinity pull-down/LC-MS/MS experiments; all proteins were identified with a minimum protein score of 40 as well as at least two unique peptides, and results are summarized in Figure 5.5C, Table 5.1 and in Appendix 2. In total, 49 and 111 proteins were identified from BSF and PCF, respectively, 20 of which were from both parasite forms (see Appendix 2 for detailed protein ID). Among these proteins, some were inevitably non-specific protein binders caused by their "sticky" nature as well as their high endogenous expression level (e.g. cytoskeletal proteins and carbohydrate-metabolism-related proteins). We focused our attention on other candidates which possess known nucleophilic cysteine residues, because they are

more likely true cellular targets of K11777 (Table 5.1); in addition to the expected rhodesain and TbCatB, other proteins such as BS2, MCA4, AOX and two proteasome subunits were identified. It is worth noting that vinyl sulfones had previously been identified as potent and irreversible inhibitors of proteasomes (through modifications of their active-site threonine).[133,134] Thus, it is not surprising that two proteasome subunits were detected in PCF. It is also interesting to note that many more proteins were identified from PCF alone (91) than from BSF (29), and many of them were cytosolic.

We also confirmed that our probes were capable of labeling active cathepsin L in intact HepG2 cells. In doing so, we performed pull-down and Western blotting experiments with anti-cathepsin L antibody. Notably, cathepsin L is a lysosomal proteinase which is significantly involved in all stages of cancer progression, including growth, antigiogenesis, invasion, and and metastasis.[135,136] Recent interest in cathepsin L show that proteolysis by this enzyme is required for the entry and replication of the SARS and Ebola viruses in human cells.[137-139] Thus cathepsin L inhibitors have potential as novel anti-viral agents. As shown (Figure 5.6), cathepsin L labeled by VS-1 in HepG2 is the mature active enzymes (~30 kDa), as no labeling of its proform (~ 41 kDa) was detected.

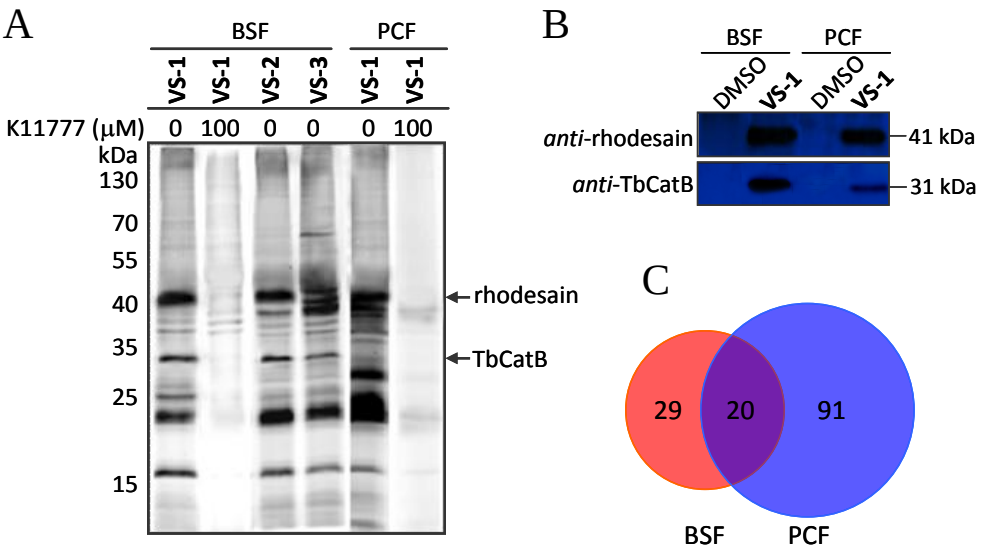


Figure 5.5. (A) *In situ* proteome profiling of probes (25 μM) against *T. brucei* BSF and PCF. K11777 was added to selected lanes in competition experiments. Putative bands of labeled rhodesain and TbCatB were indicated (with arrows). (B) Western blotting validation of rhodesain and TbCatB labeled with VS-1 (25 μM), following *in situ* labeling and affinity pull-down/WB experiments. Null pull-down experiments (with DMSO) were used as negative controls. (C) Venn diagram illustrating the number of proteins identified from *T. brucei* after *in situ* labeling and large-scale pull-down/LC-MS/MS experiments.

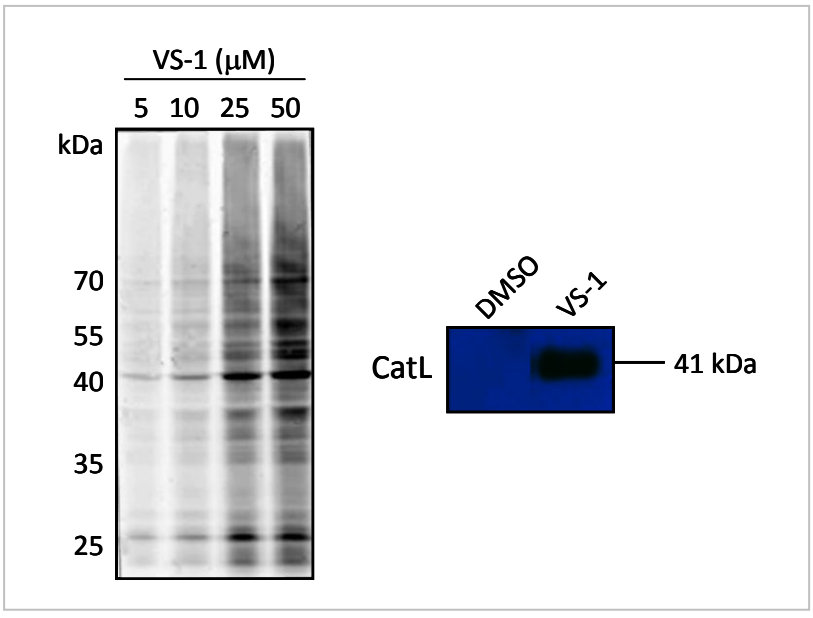


Figure 5.6. *In situ* proteome-profiling of VS-1 against HepG2 live cells and Western blotting analysis of pulled-down fractions treated with VS-1 (25 μM), or DMSO as negative controls with anti-cathepsin L antibody.

Table 5.2. Representative proteins identified with VS-1 in *Trypanosoma brucei*[a]

T. brucei gene	protein name	M_w / kDa	location	detection
Tb927.6.1000	cysteine peptidase precursor (CP), Clan CA, family C1, cathepsin L-like*	48.34	lysosome	both
Tb927.6.560	cysteine peptidase C (CPC), Clan CA, family C1, cathepsin B-like (TbCatB)**	37.22	lysosome	BSF
Tb927.10.8230	protein disulfide isomerase, bloodstream-specific protein 2 precursor (BS2)*	55.58	cytoplasm	BSF
Tb927.5.1810	lysosomal/endosomal membrane protein p67 (p67)*	72.73	lysosome	BSF
Tb10.70.5250	metacaspase MCA4, cysteine peptidase, Clan CD, family C13 (MCA4)**	38.97	nucleus	BSF
Tb927.10.7090	alternative oxidase (AOX)**	37.59	mitochondrion	BSF
Tb927.10.290	proteasome alpha 2 subunit*	25.36	cytoplasm	PCF
Tb927.10.6080	proteasome beta 5 subunit (PRCE)*	34.42	cytoplasm	PCF
Tb09.160.4250	tryparedoxin peroxidase (TRYP1)**	22.43	cytoplasm	PCF
Tb927.5.3350	iron superoxide dismutase**	26.82	mitochondrion	PCF
Tb927.10.7410	succinyl-CoA ligase [GDP-forming] beta-chain*	44.91	mitochondrion	PCF
Tb10.70.4740	enolase**	46.59	glycosome	PCF

[a] Symbols in the protein name column: (*) sensitive to RNA interference; (**) putative drug target.

5.2.5 Cellular Imaging

To visualize cellular localization of VS-1 and potential cellular targets of K11777, live parasites were directly treated with VS-1, fixed, and reacted *in sit*u with the rho-azide reporter, then imaged. Overall, no noticeable difference was observed in the probe uptake between BSF and PCF (Figure 5.7). As shown in Figure 5.8, VS-1 signals were detected mostly in the lysosome of BSF, where endogenous rhodesain resides. This lysosoal colocalization is important given that K11777 had little or no toxicity on mammalian presumably because vinyl sulfones become concentrated within the parasite,[143] and/or human cathepsins' redundancy and activation within other subcellular compartments.[144,145] In contrast, the probe was shown to be evenly distributed throughout PCF parasites with no specific subcellular localization. Similar results were obtained with HepG2 mammalian cells as shown in Figure 5.9. This result could be due to the fact that K11777 is a broad-spectrum cysteine protease inhibitor.[142] Our results correlate well with our above-described proteome profiling and affinity pull-down/LC-MS/MS data where different cellular targets were labeled in the two parasite forms, and many more cytosolic proteins were identified in PCF. The partial colocalization observed between VS-1 and rhodesain in BSF presumably also reflected the presence of other side targets of K11777. Collectively, these results established that VS-1 not only can be used for in situ proteome profiling and identification of potential cellular targets of K11777 in live parasites, it may also be used as an imaging probe and to study drug localization.

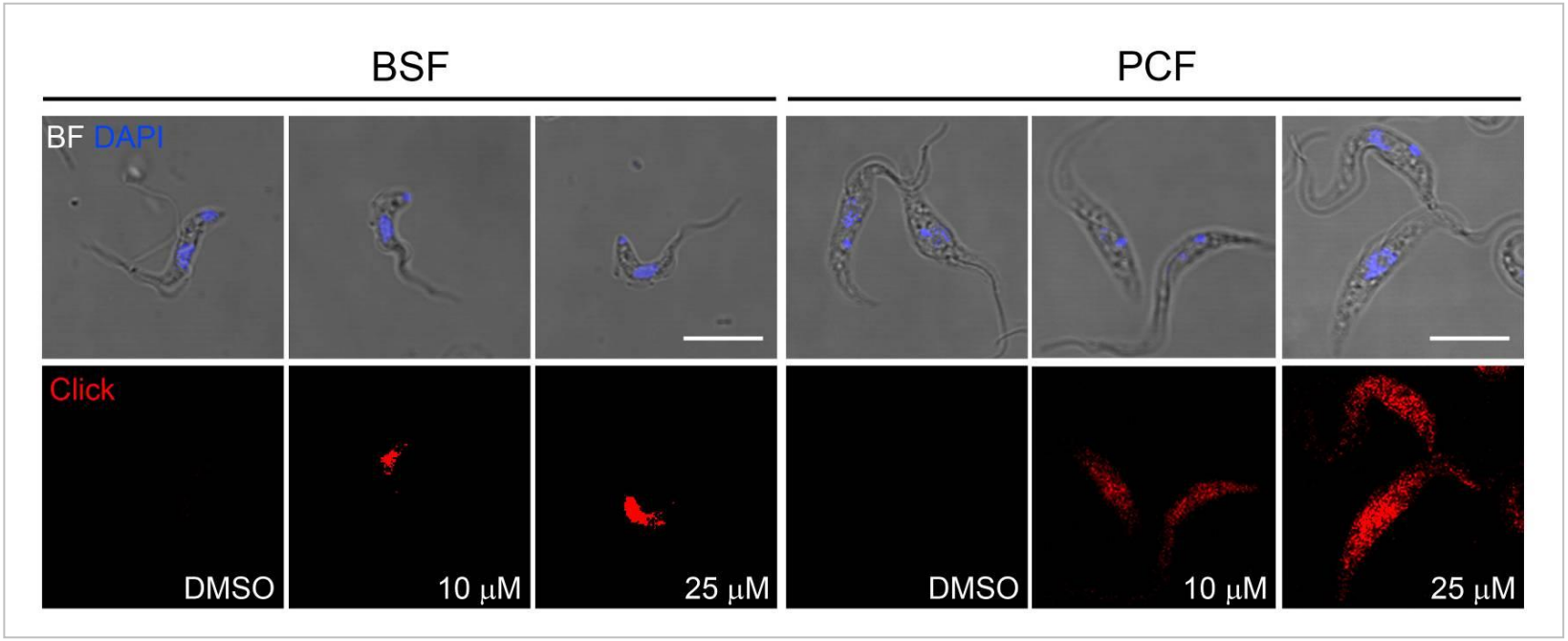


Figure 5.7. Cellular uptake of VS-1 within BSF and PCF *T. brucei*. Parasites (2×10^5 cells) were incubated with VS-1 (at 0, 10 & 25 μM, respectively) for 2 h, reacted with 10 μM of rhodamine–azide **2-33** under CuAAC conditions, and then imaged. DAPI stained (with nucleus and/or kinetoplast colored in blue); Rhodamine channel showing cellular distribution of VS-1 (colored in red). Scale bar represents 10 μm.

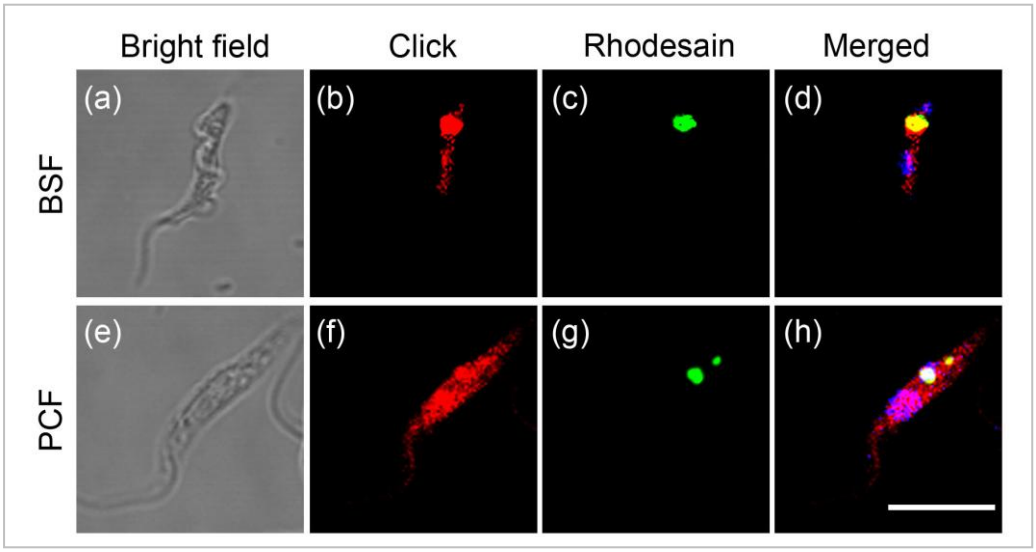


Figure 5.8. Confocal microscopy of VS-1 localization in BSF (top) and PCF (bottom). Live parasites were treated with VS-1 (25 μM) followed by imaging as described in ESI. Panels (a) and (e): bright-field images. Panels (b) and (f): 554 nm channel (pseudocolored in red) detecting cellular localization of VS-1. Panels (c) and (g): Immunofluorescence (IF) staining at 488 nm channel (pseudocolored in green) to detect cellular localization of rhodesain. Anti-rhodesain primary antibody and FITC-conjugated anti-rabbit IgG secondary antibody were used. Panels (d) and (h): merged images of panels (b) and (c), (f) and (g) together with nuclei (stained with DAPI; pseudocolored in blue). All images were acquired under the same settings. Scale bar = 10 μm.

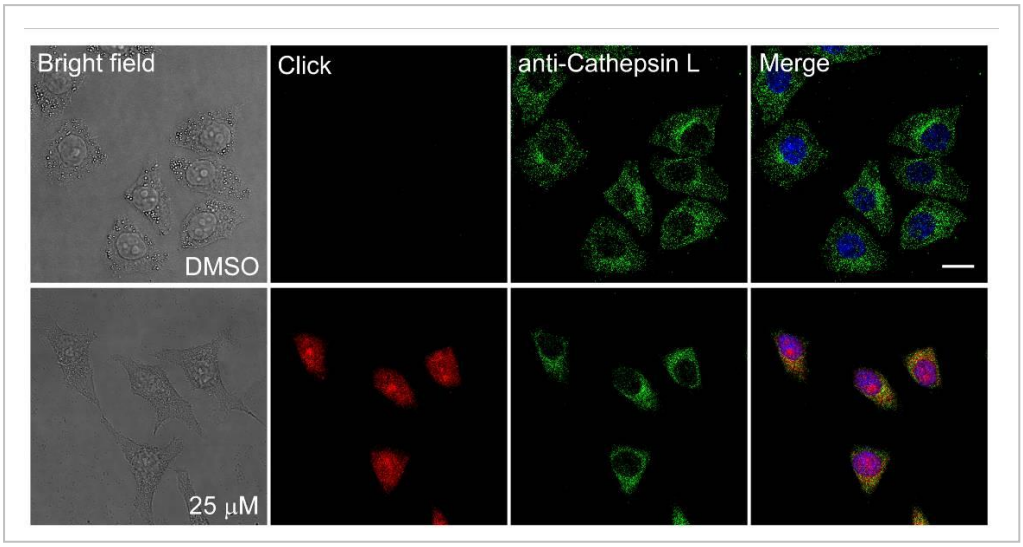


Figure 5.9. Immunofluorescence analysis of active cathepsin L in HepG2 cells labeled with VS-1. Scale bar represents 10 μm.

5.3 Conclusion

In conclusion, we have successfully synthesized and evaluated K11777-like probes for their anti-trypanocidal activities against both BSF and PCF *T. brucei*. Subsequent *in situ* proteome profiling of VS-1 enabled us to tentatively identify previously unknown cellular targets of K11777 in both parasite forms. Furthermore, we demonstrated the utility of our probe for live-parasite labeling and visualization of potential K11777-responsive targets (e.g. rhodesain). We believe that our observations may provide important information for investigating K11777 as a antiparasitic drug candidate other than *T. brucei*. Further studies are underway to validate some of the potential cellular targets identified from the present study.

Chapter 6

Design, Synthesis and Biological Evaluation of Potent Azadipeptide Nitrile Inhibitors and Activity-Based Probes as Promising Anti-*Trypanosoma brucei* Agents

Portions of the work presented here are taken from:

P.-Y. Yang, M. Wang, L. Li, H. Wu, C. Y. He, and S. Q. Yao, "Design, Synthesis and Biological Evaluation of Potent Azadipeptide Nitrile Inhibitors and Activity-Based Probes as Promising Anti-*Trypanosoma brucei* Agents." *Chem. Eur. J.* **2012**, in press.

Abstract

Trypanosoma cruzi and *Trypanosoma brucei* are parasites that cause Chagas disease and African sleeping sickness, respectively. There is an urgent need for development of new drugs against both diseases due to the lack of adequate cures and emerging drug resistance. One promising strategy for the discovery of small-molecule therapeutics against parasitic diseases has been to target the major cysteine proteases such as cruzain for *T. cruzi*, and rhodesain/TbCatB for *T. brucei*. Azadipeptide nitriles belong to a novel class of extremely potent cysteine protease inhibitors against papain-like proteases. We herein report the design, synthesis, and evaluation of a series of azanitrile containing compounds, most of which were shown to potently inhibit both recombinant cruzain and rhodesain at low nanomolar/picomolar ranges. A strong correlation between the potency of rhodesain inhibition (i.e. target-based screening) and the trypanocidal activity (i.e., whole organism-based screening) of the compounds was observed. To facilitate detailed studies of this important class of inhibitors, selected hit compounds from our screenings were chemically converted to activity-based probes (ABPs), which were subsequently used for in situ proteome profiling and cellular localization studies to further elucidate potential cellular targets (on and off) in both the disease-relevant bloodstream form (BSF) and the insect-residing procyclic form (PCF) of Trypanosoma brucei. Overall, the inhibitors presented herein show great promises as a new class of anti-trypanosome agents that possess better activities than existing drugs. The activity-based probes generated from this study could also serve as valuable tools for parasite-based proteome profiling

studies, as well as bioimaging agents for studies of cellular uptake and distribution of these drug candidates. Our studies therefore provide a good starting point for further development of these azanitrile-containing compounds as potential anti-parasitic agents.

6.1 Introduction

Chagas disease (or American trypanosomiasis) is caused by Trypanosoma cruzi and affects around 12 million people in Latin America, resulting in ~14,000 deaths per year.[146] More recently, due to immigrant carriers or infections from contaminated transfusions and organ transplants, Chagas disease has been reported in areas where it normally doesn't exist.[147] Human African trypanosomiasis (HAT, sleeping sickness) is caused by tsetse fly transmitted parasites of Trypanosoma brucei. Over 60 million people in sub-Saharan Africa are at risk of this disease, of which there are already 50,000-70,000 confirmed cases.[83e] Existing drugs used to treat these diseases are generally highly toxic and often ineffective, and drug resistance has developed in some cases.[148] With no immediate prospect of vaccines, there is a dire need to develop new trypanocidal agents with acceptable efficacy and safety profiles. One encouraging approach to novel anti-parasitic agents is the development of small molecule inhibitors targeting parasitic cysteine proteases such as cruzain (an essential protease for the survival of *T. cruzi*[149]), rhodesain (also known as brucipain, trypanopain) and more recently TbCatB (a cathepsin B like protease) in *T.*

brucei.[131] Cysteine proteases are fundamental to the metabolism of many parasites, and their inhibitors have been shown to kill protozoan parasites in both culture and animal models.

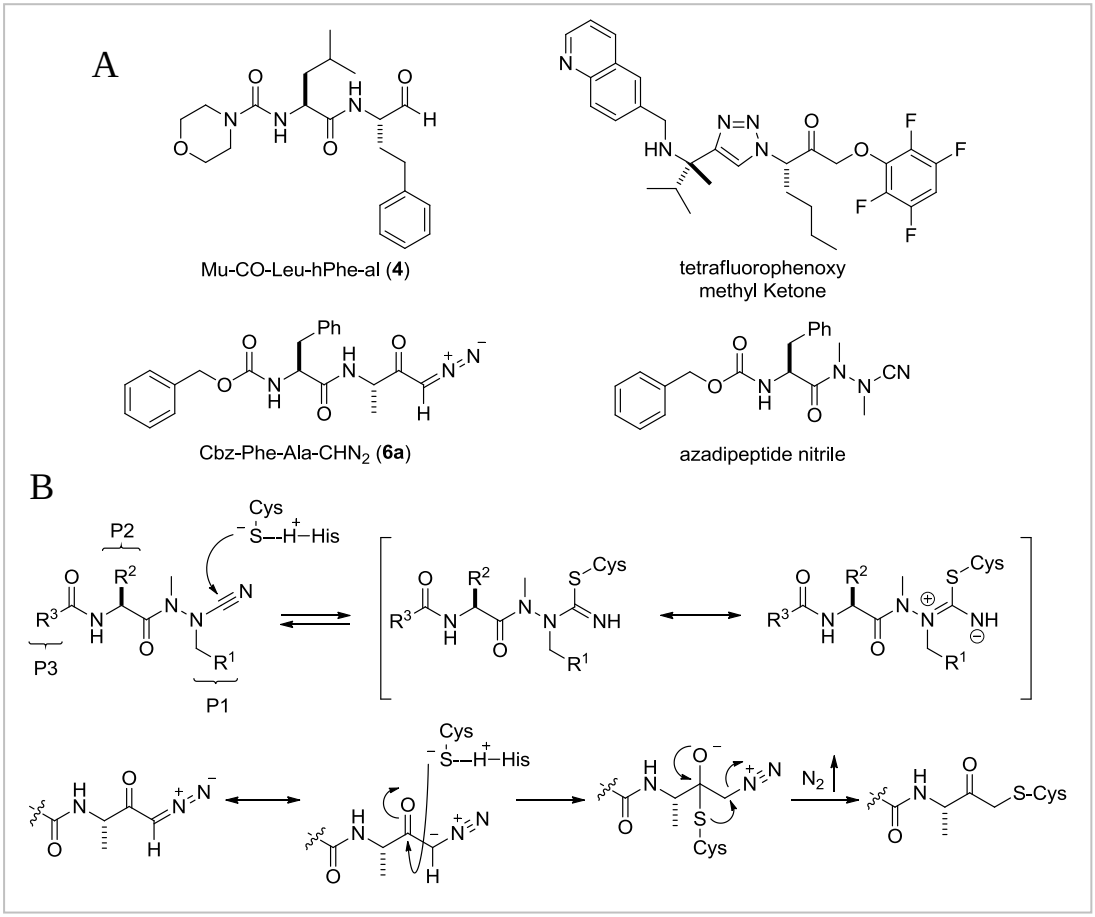


Figure 6.1. (A) Representative structures of anti-parasitic cysteine protease inhibitors. (B) Proposed inhibitory mechanism of cysteine proteases by azanitriles (top) and diazomethyl ketones (bottom).

One such inhibitor, N-Mpip-Phe-Hph-VSPh (also known as K11777; in Chapter 5) is currently in late-stage preclinical trials for Chagas disease.[124] K11777 was also shown to have potent in vitro and in vivo (mice were cured of infection) bioactivities against *T. brucei*,[126,150] presumably by targeting rhodesain, TbCatB, or both (since the inhibitor is non-selective). Another cysteine protease inhibitor,

Z-Phe-Ala-CHN_2 (**6-6a**; Figure 6.1A), has also been shown to be lethal to T. brucei in both in vitro and in vivo studies.[85,151] Vinyl sulfones, tetrafluorophenoxymethyl ketones, and diazomethyl ketones, all of which contain an electrophilic "warhead" that can covalently and irreversibly inactivate cruzain as a result of nucleophilic attack by the active-site cysteine (a representative example is shown in the bottom scheme of Figure 6.1B), are also known parasiticidal agents. Vinyl sulfones, however, are also well-known proteasome inhibitors.[133,134] Previous work in developing anti-malarial compounds showed that synthetic peptidyl aldehydes were potent reversible inhibitors of the cysteine protease hemoglobinases, falcipain-2 and falcipain-3. One such compound, Mu-CO-Leu-hPhe-al (Figure 6.1A), which exhibited strong inhibitory activity on falcipains and cultured Plasmodium falciparum parasites, was tested in a murine malaria model. Recently, Gütschow and co-workers reported proteolytically stable azadipeptide nitriles as a novel class of cysteine protease inhibitors that form covalent but reversible isothiosemicarbazide adducts with the active-site cysteine (Figure 6.1B; top).[152,153] Using cathepsin K as an example, the authors further demonstrated that structural optimizations of azadipeptide nitriles could result in selective cathepsin inhibitors. A subsequent study clearly demonstrated that azadipeptide nitriles also possessed potent anti-malarial activities.[154] In the current study, we have taken efforts to further optimize this class of inhibitors as potential agents against trypanosome parasites, with the ultimate aim to generate highly potent inhibitors that possess fewer off-targets and side effects (which are often associated with irreversible enzyme inhibitors).[155] Herein, we report the chemical synthesis and

biological evaluation of a series of aza-nitrile-containing inhibitors, as well as some activity-based probes derived from selected compounds. By first screening these compounds against purified recombinant cruzain and rhodesain (i.e., target-based screening), we found most of them were indeed highly potent inhibitors, with IC_{50} values in the low nanomolar/picomolar ranges. Since inhibitor sensitivities are often affected by assay conditions such as pH, ionic strength, and the presence of lipids or other proteins, we subsequently carried out whole organism-based screening and quantitatively verified the potencies of some of these compounds. A good correlation between the potency of rhodesain inhibition and the trypanocidal activity of these compounds was observed. Next, by combining the whole organism-based screening with activity-based protein profiling (ABPP) coupled with the bio-orthogonal click chemistry, we carried out comprehensive proteome profiling experiments against selected hit compounds in an effort to identify both on- and off-targets, which provided invaluable insights into the mode-of-action and further optimizations. Finally, fluorescence imaging experiments with the probes provided further information on the compound's cellular uptake and sub-cellular distribution (Figure 6.1C). Overall, the inhibitors presented have promise as trypanocidal agents with improved activity, and the probes can serve as valuable tools for assessing the uptake and distribution of the drug candidate and for examining the potential for drug resistance and toxicity.

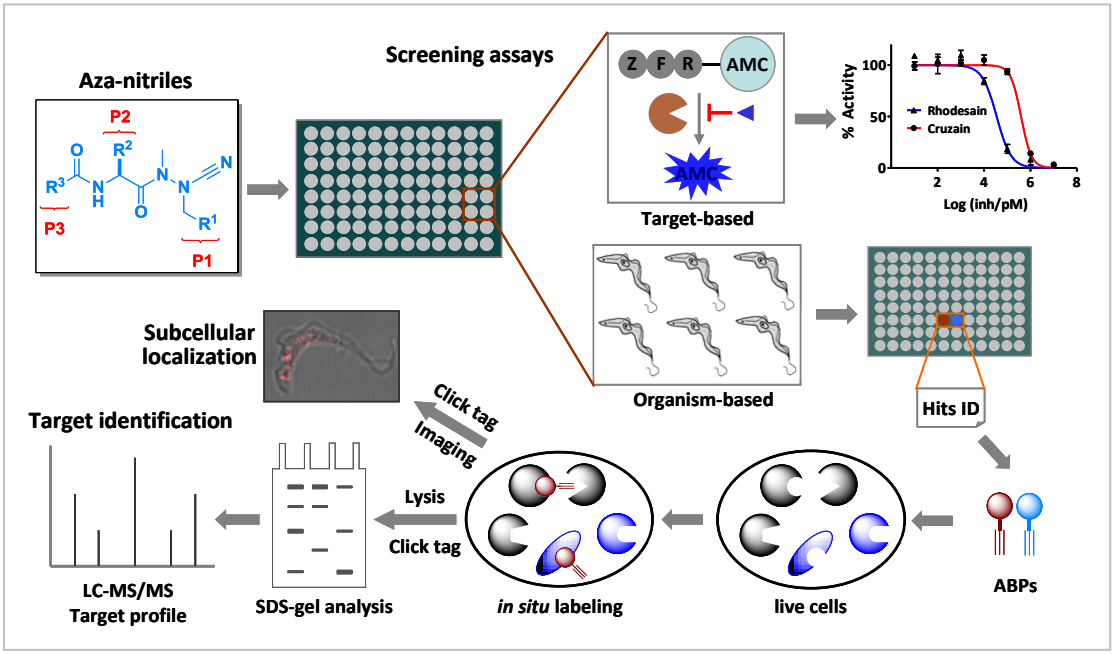


Figure 6.2. Overall workflow of chemical screens and characterizations of azanitriles that were screened with target-based (cruzain & rhodesain) and organism-based assays, and conversion of hit molecules into activity-based probes (ABPs). Putative identification of cellular targets (on and off) was done via in situ parasite-based proteome profiling. Reaction of the alkyne handle on the probe with rhodamine-azide via click chemistry allows visualization of target proteins by SDS-PAGE gel or fluorescence microscopy. Alternatively, with biotin-azide, the probe-labeled parasite proteome could be pulled-down/LC-MS/MS for target identification.

6.2 Results and Discussion

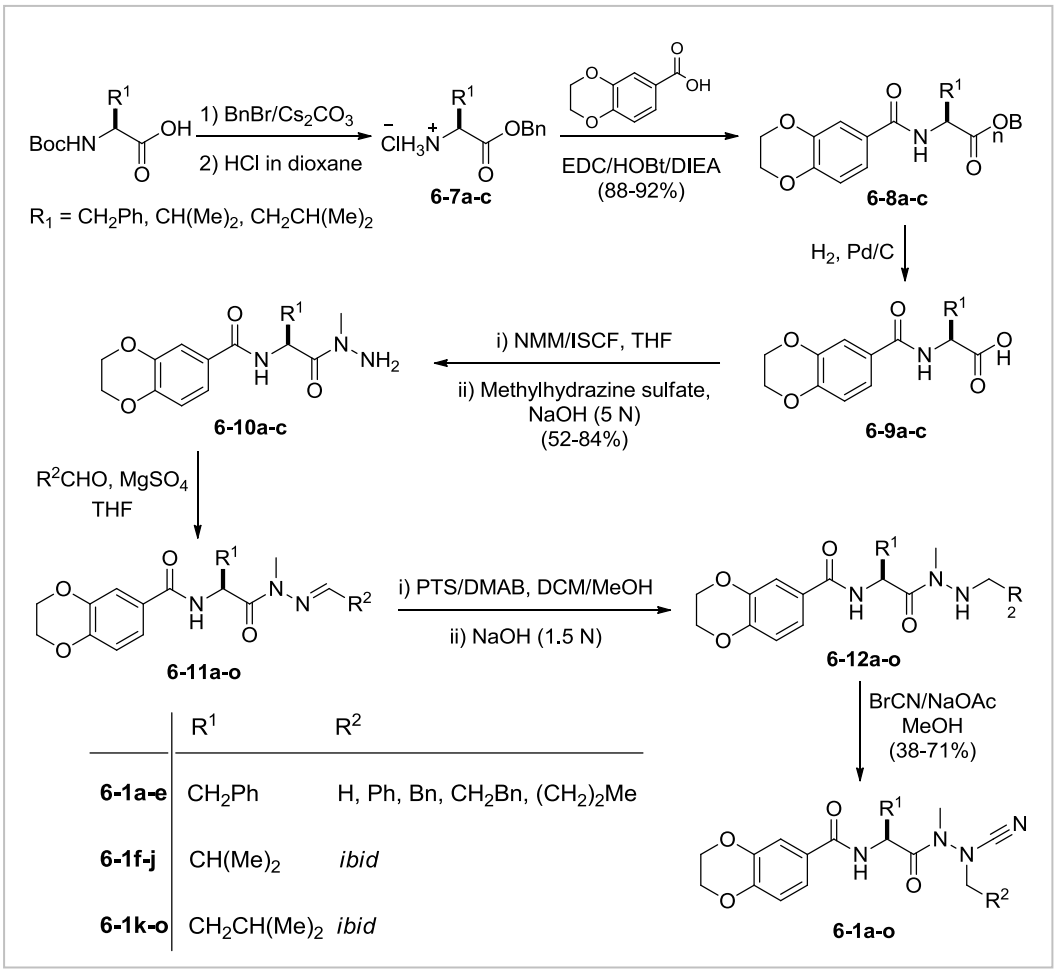
6.2.1 Design and Synthesis of Azadipeptide Nitriles

The isoelectronic replacement of the $C_\alpha H$ group by a nitrogen atom to yield azapeptides is a common structural modification in the chemistry of peptides and peptidomimetics.[156] This structural modification was first applied to the P_1 position of peptide nitriles in 2008.[152] The resulting azadipeptide nitriles were characterized as highly potent inhibitors of cysteine cathepsins (i.e., cathepsins L, S, and K). These

compounds possess a methylated P_2-P_1 peptide bond, which accounts for their stability towards proteolytic cleavage. Such azadipeptide nitriles also showed time-dependent inhibition, and the reversible formation of adducts with the active site cysteine was proposed (part B of Figure 6.1; top). The strategy we had undertaken in the current work was to build on these initial results to further demonstrate potent inhibition of cruzain and rhodesain by azanitrile containing compounds, as both enzymes are structurally homologous to cathepsin L. Previous studies showed that cruzain was effectively inhibited by vinyl sulfones bearing Leu or Phe residues at the P_2 position (e.g. K11777/K11002). However, the P_1 and P_3 groups were considered to be less important, due to their solvent exposure as seen in the protein structure.[121,128,157] In other words, the urea group at the P_3 position in K11777/K11002 did not actively participate in binding interactions with the enzyme; therefore this moiety could be replaced by a carbamate group, or an amide group.[140,158] The designed library was based on the scaffold presented in Scheme 6.1, in which R^3 was either 2,3-dihydro-1,4-benzodioxin-6-yl (resulting in an amide group at the P_3 position) or the carbamoyl groups, i.e., *t*BuO or BnO, as previously reported.[152] R^2 could be either phenyl, isopropyl or isobutyl, so that the amino acid at the P_2 position was either Phe, Val or Leu, respectively. The highest variability was introduced at the P_1 position with five different groups (i.e., Me, Bn, CH_2Bn, CH_2CH_2Bn, *n*-butyl), thereby enabling a systematic evaluation of the structure-activity relationship (SAR) between members of this library and their inhibitory property against cruzain/rhodesain. In this context, it is noteworthy that the

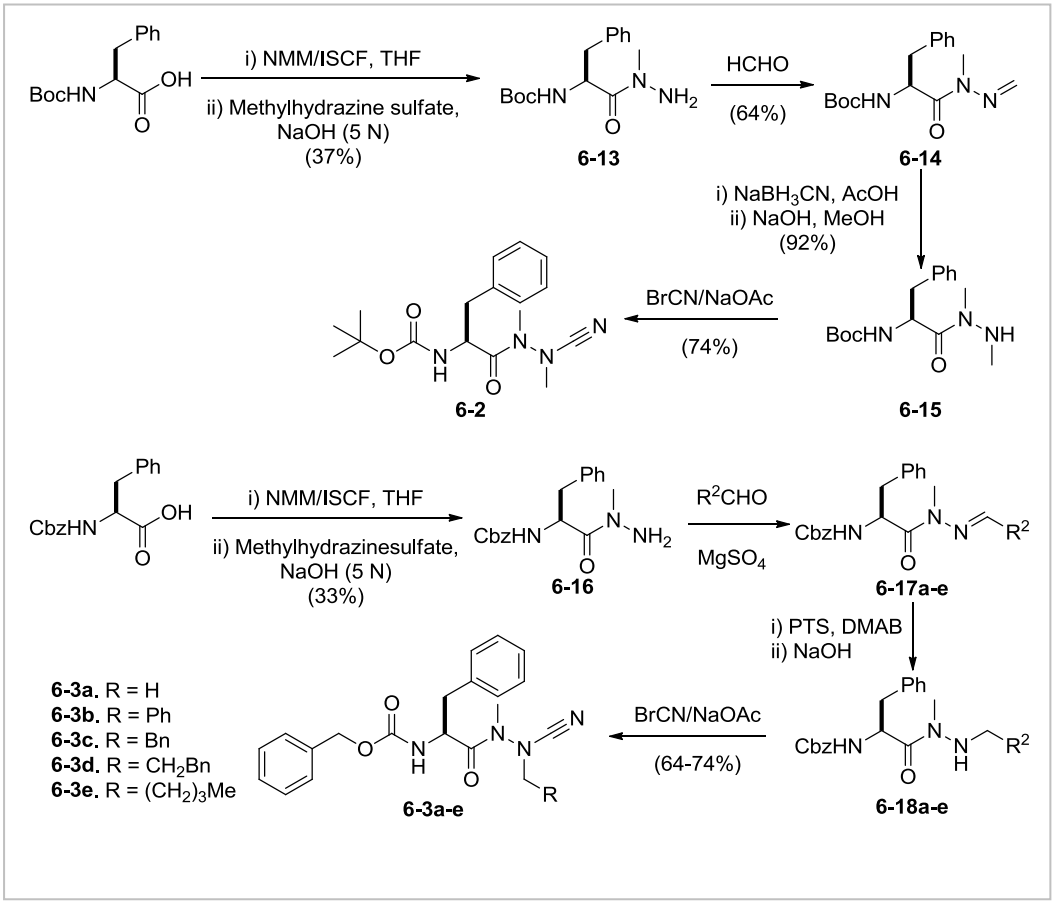
structural diversity of this library can be easily increased by further variations at P_1, P_2 and P_3 positions using the same chemistry developed herein.

As shown in Scheme 6.1, the synthesis of our library was accomplished by modifications of previously reported procedures for the preparation of **6-3b** and its derivatives (Scheme 6.2).[16] Starting from commercially available Boc-protected amino acids, we prepared the corresponding benzyl esters and then removed the amine protecting group using standard methods. Next, the coupling with 1,4-benzodioxane-6-carboxylic acid to give **6-8a–c**, whereupon the deprotection was carried out hydrogenolytically. The hydrazones **6-10a–c** were subjected to the reductive amination of various aldehydes using dimethylamine-borane complex as reducing regent to give the corresponding N^1-methyl-N^2-alkylhydrazides **6-12a–o**. Upon treatment with cyanogen bromide finished the desired aza-nitriles **6-1a–o** (Scheme 6.1). For the preparation of the analogue **6-2** (Scheme 6.2), the N^1-methyl-N^2-alkylhydrazide **6-15** was prepared by using sodium cyanoborohydride in the presence of acetic acid,[159] instead of dimethylamine–borane complex. Attempts to prepare the aldehyde **6-4** by oxidation of the corresponding primary alcohol with Dess-Martin periodinane (DMP) resulted in a complex mixture of products. Therefore, the Weinreb amide **6-22** was prepared and subsequently reduced with lithium aluminum hydride to furnish **6-4** (Scheme 6.3).

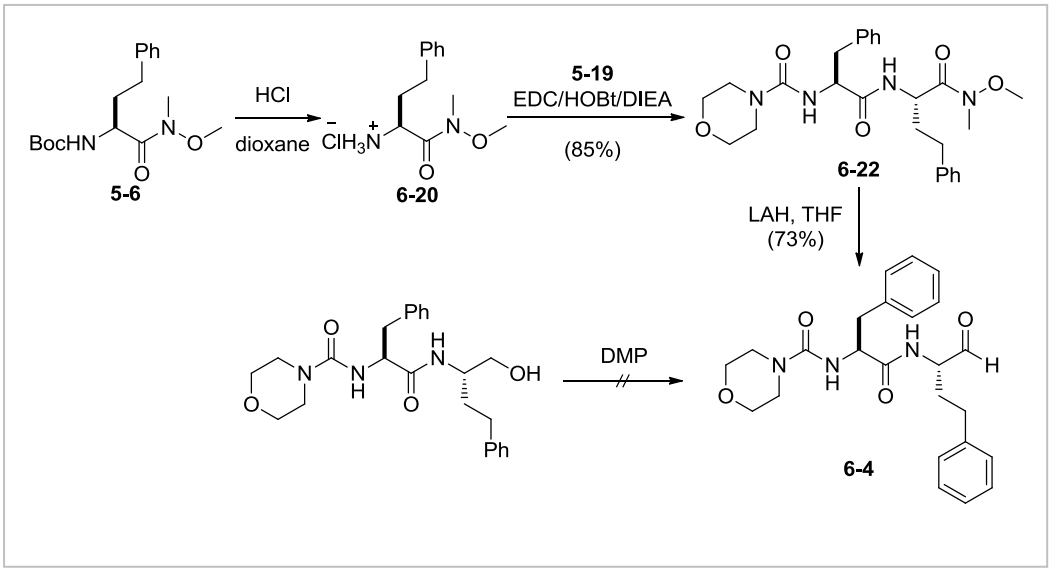


	R^1	R^2
6-1a-e	CH_2Ph	H, Ph, Bn, CH_2Bn, $(CH_2)_2Me$
6-1f-j	$CH(Me)_2$	*ibid*
6-1k-o	$CH_2CH(Me)_2$	*ibid*

Scheme 6.1. Synthesis of amide-based compounds **6-1a–o**.



Scheme 6.2. Synthesis of carbamate-based compounds **6-2** & **6-3a–3e**.



Scheme 6.3. Synthesis of the aldehyde **6-4**.

6.2.2 Biological Screening

With twenty-one aza-nitriles (**6-1a–o**, **6-2** & **6-3a–e**), the aldehyde **6-4** in hand, we next evaluated the inhibitory activity of the library against cruzain and rhodesain in fluorometric microplate assays using the substrate Z-Phe-Arg-AMC. For comparison, K11002 was also assayed under conditions. All compounds exhibited greater inhibitory capacity toward rhodesain than cruzain (Table 6.1). The anti-cruzain and anti-rhodesain activities were well correlated. This phenomenon is not surprising given the structural homology of these two enzymes. In general, compounds with Phe (**6-1a–e**) and Leu (**6-1k–o**) in P_2 were more active than compounds with Val (**6-1f–j**), whilst Leu compounds (**6-1k–o**) exhibited higher potency for cruzain. In P_1, methyl (**6-1a**, **6-1f** & **6-1k**) group yielded better results than others. Finally, the changes with P_3 confer a subtle effect on enzyme selectivity, arguing that P_3 may be a recognition element for other enzymes in the papain class than initially assumed. Indeed, a study in which a combination of optimized P_2 and P_3 substituents in azadipeptide nitriles led to a picomolar cathepsin K inhibitor with remarkable selectivity over other cathepsins was recently published.[153] Compound **6-2** (Figure 6.2), bearing a Phe side chain in P_2 and Boc group in P_3, respectively, was not only one of the strongest inhibitors of rhodesain (IC_{50} = 60 pM), but it also presented the highest selectivity toward this protease ($IC_{50\ Cuzi}/IC_{50\ Rhod}$ = 193). Compound **6-1a** was also a potent inhibitor of rhodesain (IC_{50} = 140 pM, $IC_{50\ Cuzi}/IC_{50\ Rhod}$ = 60). Another inhibitor with highest activity toward rhodesain was compound **6-1k** (IC_{50} = 60 pM), although this compound did not demonstrate selectivity ($IC_{50\ Cuzi}/IC_{50\ Rhod}$ = 5). As a control, IC_{50}

values of K11002 were calculated for rhodesain (IC_{50} = 0.36 nM) and cruzain (IC_{50} = 6.6 nM). Interestingly, compound **6-4**, with an incorporated aldehyde "warhead", displayed comparable potency as K11002/K11777 against both enzymes (IC_{50} values of 0.55 nM and 3.6 nM, respectively, against rhodesain and cruzain). Although we did not further evaluate the inhibitory activity of our compounds against TbCatB due to the unavailability of this enzyme, we were encouraged by the observation that some compounds show low picomolar IC_{50} values against rhodesain, thereby decided to investigate their activities against live parasites (i.e. whole organism-based screening). In this context, it is noteworthy that TbCatB has been shown to be essential for *T. brucei* survival based on RNA interference studies in vitro[131] and in mice[132]. Whether rhodesain is essential was unclear due to a lack of complete silencing (via RNAi) of target transcript, either in vitro[131] or in vivo[132] (40% proteolytic activity remained). Yet, even with the limited gene knockdown, rhodesain's importance to parasite survival was clear in that parasites exposed to RNAi were less efficient in penetrating an in vitro model of the blood-brain barrier and blocked by the cysteine protease inhibitor K11777.[150] As such, the combined data, to some extent, perhaps support rhodesain as a valuable drug target.[120]

We next evaluated the trypanocidal activities of these compounds in cell cultures of both BSF and PCF of *T. brucei* using a growth assay on a Guava PCA-96 system (Guava Technologies, USA) following the manufacturer's instructions. As shown in Figure 6.3 & 6.4, all compounds induced a potent, dose-dependent inhibition on parasite replication. As expected, most azanitriles were found to be

significantly more potent than the two positive controls, i.e., K11002 and compound **6-4**. Overall, there were similarities in the structure-activity relationship (SAR) in this series of compounds between the recombinant enzymes, particularly rhodesain, and *T. brucei*. Compounds **6-3a-e** with Phe at the P_2 position and Cbz at the P_3 position displayed the best parasite-killing activity against *T. brucei*. Similar to data obtained from rhodesain inhibition experiments, compounds having Phe (**6-1a-e**) and Leu (**6-1k-o**) at the P_2 position inhibited parasite growth better than compounds having Val (**6-1f-j**) at the same position. When compounds **6-1a-e** (with Phe at the P_2 position) and **6-1k-o** (with Leu at the P_2 position) were compared, no significant difference in anti-parasitic activity was observed. Notably, compound **6-2** which has a Boc group at the P_3 position and exhibited an IC_{50} value of 60 pM against rhodesain, was comparatively less potent in its anti-parasitic activity than compounds **6-1**a and **6-3a** (which bear an aromatic amide and a Cbz group at the P_3 position, respectively). This suggests that the P_3 residue is important for the trypanocidal activity of these aza-nitriles. For all compounds tested, there was no obvious difference in trypanocidal activity when the P_1 position was changed from a methyl group to other groups. This series of compounds, however, were able to block parasite growth much more effectively than K11002. These results did not entirely correlate with inhibition results obtained using recombinant rhodesain (Table 6.1), which might be due to differences in their cell permeability and/or the mechanism-of-action among these compounds. The exact reason is still under our active investigation. In order to quantitatively confirm the above screening results against some of the compounds tested across a

wider range of inhibitor concentrations, we subsequently obtained the ED_{50} values of selected compounds (**6-3b**, **6-3d**, **6-3e** and K11777/K11002) against both life cycle stages of *T. brucei* (Table 6.2). Interestingly, our results clearly showed that these azanitriles (e.g. **6-3b**, **6-3d** and **6-3e**) were ~2-fold more active in blocking PCF parasite growth (ED_{50} values of 0.4, 0.7 and 0.7 μM, respectively, after 24 h) than BSF parasite growth (ED_{50} values of 1.1, 1.1, and 1.4 μM, respectively). Importantly, all of them displayed stronger trypanocidal activity than K11777 and K11002 in both PCF and BSF of *T. brucei*. On the contrary, both K11777 and K11002 displayed a slightly stronger trypanocidal activity in BSF (ED_{50} = 0.9 and 5.6 μM) than in PCF (ED_{50} = 5.8 and 7.3 μM). In this context, it is worth noting that the *T. brucei* life cycles alternate between mammalian host stages and insect vector stages, which are coupled to extensive alterations in morphology and metabolism. Although the precise underlying mechanism for above differences is presently unclear, they appear again to be related to these compounds' differences in the cell permeability and/or cellular targets. Therefore, as described in the next sections, we had made an attempt to identify potential cellular targets (on and off) of some of the most potent compounds in *T. brucei* using activity-based protein profiling (ABPP) coupled with the bio-orthogonal click chemistry. To this end, we have successfully made cell-permeable activity-based probes (ABPs) and used them for parasite-based proteome profiling and bioimaging experiments (vide infra). It should also be noted that, due to the high similarity between *T. cruzi* and *T. brucei*, these compounds/probes are possibly effective againts *T. cruzi* as well.

Table 6.1. Inhibition cruzain and rhodesain by aza-nitriles 6-1a–o, 6-2, 6-3a–e, 6-4 & K11002.

Compd.	6-1a	6-1b	6-1c	6-1d	6-1e	6-1f	6-1g	6-1h	6-1i	6-1j	6-1k	6-1l
Rhodesain (IC_{50}/nM)	0.14 ±0.06	0.51 ±0.01	1.86 ±0.01	0.24 ±0.004	1.34 ±0.02	5.4 ±0.01	15.8 ±0.001	35.1 ±0.002	54.0 ±0.01	54.6 ±0.001	0.06 ±0.001	0.33 ±0.04
Cruzain (IC_{50}/nM)	8.5 ±0.002	19.5 ±0.01	101.6 ±0.005	8.1 ±0.008	52.2 ±0.02	18.8 ±0.003	133.6 ±0.02	396.9 ±0.003	410.6 ±0.001	410.6 ±0.01	0.34 ±0.004	4.2 ±0.02
Compd.	**6-1m**	**6-1n**	**6-1o**	**6-2**	**6-3a**	**6-3b**	**6-3c**	**6-3d**	**6-3e**	**6-4**	**K11002**	
Rhodesain (IC_{50}/nM)	0.71 ±0.001	0.14 ±0.02	0.33 ±0.02	0.06 ±0.008	0.4 ±0.01	0.19 ±0.02	0.81 ±0.009	2.0 ±0.002	0.06 ±0.008	0.55 ±0.005	0.36 ±0.02	
Cruzain (IC_{50}/nM)	9.8 ±0.004	2.5 ±0.02	8.4 ±0.002	11.6 ±0.02	3.4 ±0.005	8.4 ±0.01	22.7 ±0.003	115.7 ±0.002	3.1 ±0.006	3.6 ±0.002	6.6 ±0.001	

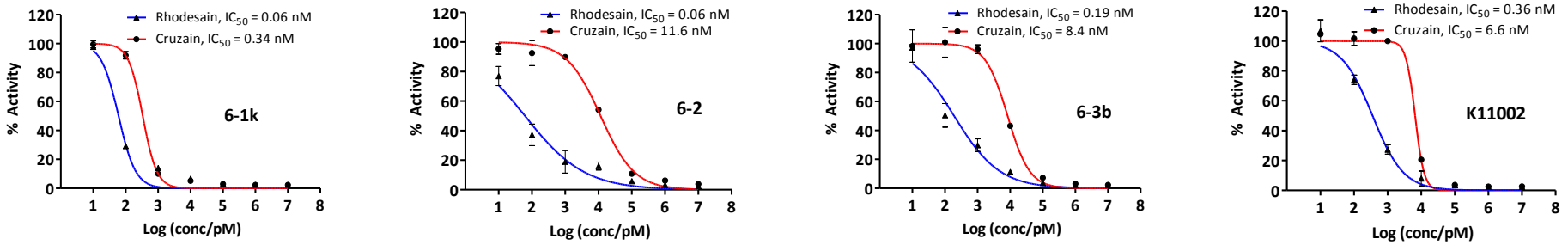


Figure 6.3. Representative IC_{50} curves for rhodesain and cruzain.

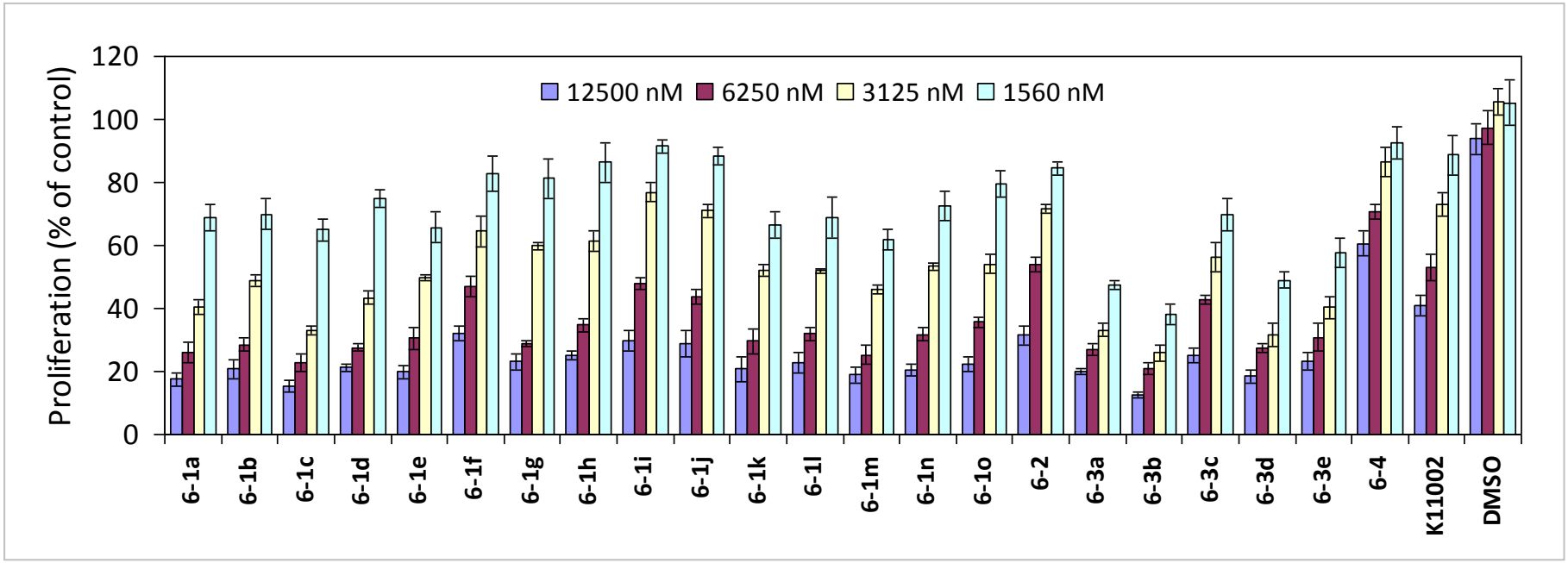


Figure 6.4. (A) Dose-dependent trypanocidal effects of aza-nitriles (**6-1a–o**, **6-2** & **6-3a–e**), the aldehyde **6-4** and vinyl sulfone K11002 against bloodstream forms of *T. brucei* after 24 h. Results represent the average standard deviation for two independent trials.

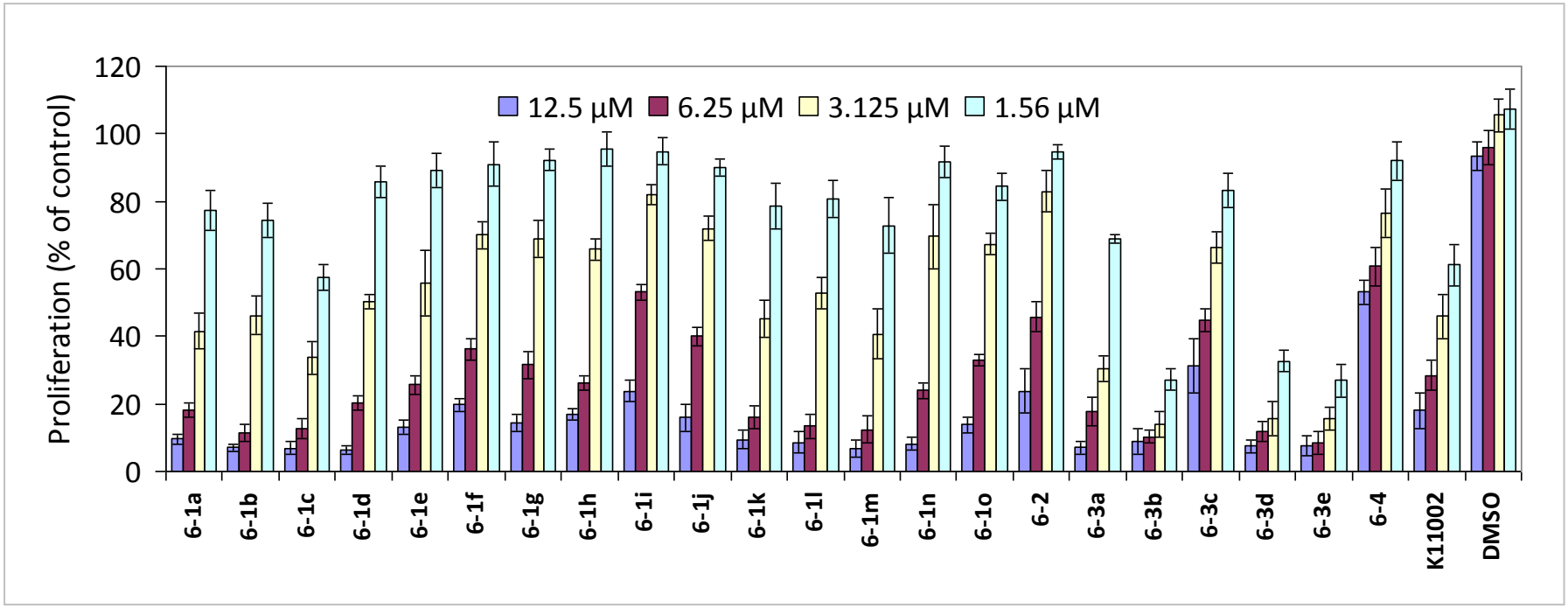


Figure 6.4. (B) Dose-dependent trypanocidal effects of aza-nitriles (**6-1a–1o**, **6-2** & **6-3a–e**), the aldehyde **6-4** and K11002 against bloodstream forms of *T. brucei* after 48 h. Results represent the average standard deviation for two independent trials.

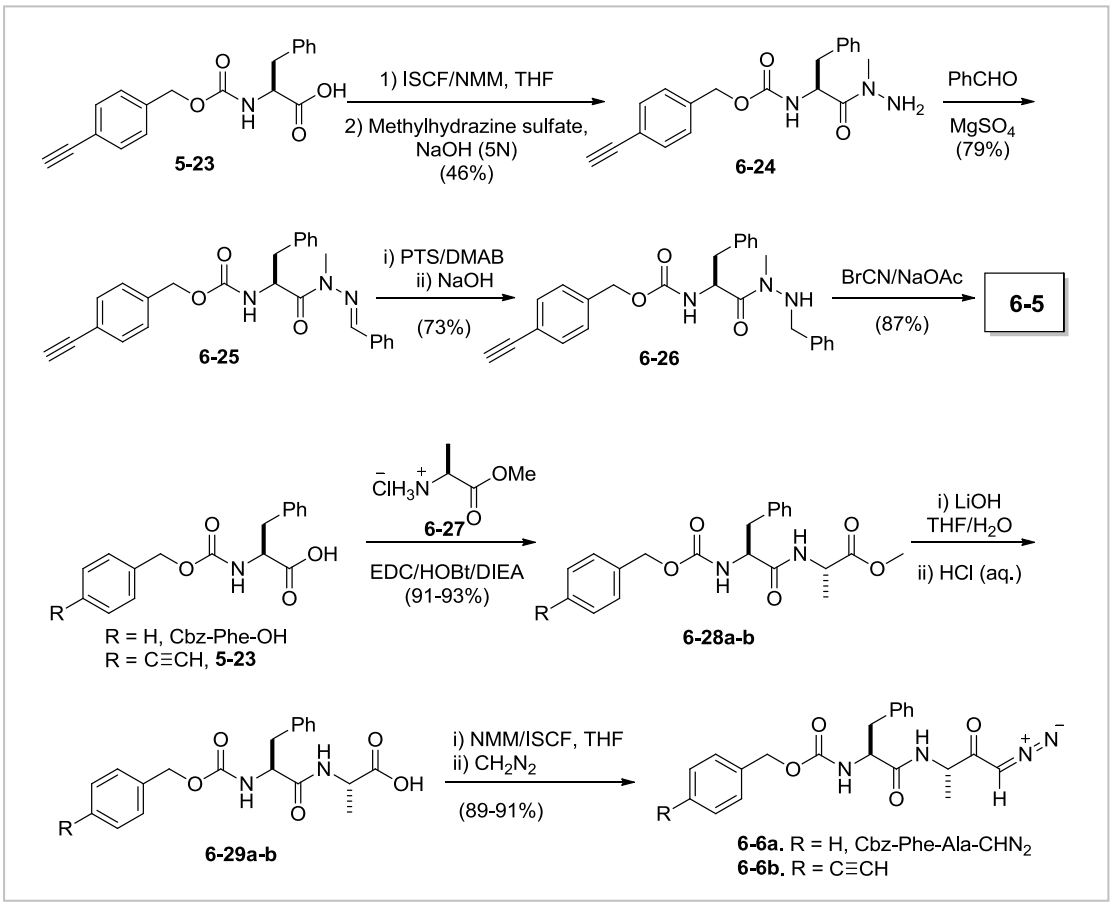
Table 6.2. ED_{50} values of compounds used in this study.

Compound	BSF (μM)	PCF (μM)
K11002	5.6	7.3
6-3b	1.1	0.4
6-3d	1.2	0.7
6-3e	1.4	0.7

6.2.3 Design and Synthesis of Activity-based Probes (ABPs)

Our screening results thus far showed some of the azanitriles were highly potent inhibitors of cruzain/rhodesain, and possessed excellent trypanocidal activity in *T. brucei* likely through inhibition of endogenous rhodesain/TbCatB activities, but they fell short in addressing the possibility that these compounds might also target other cellular proteins present in the parasite proteome (e.g. off-targets). To facilitate a more detailed understanding of this novel class of inhibitors, we converted two hit compounds identified from our screening results, **6-3b**, and Z-Phe-Ala-CHN_2 (**6-6a**) into activity-based probes (ABPs), which would enable parasite-based proteome profiling and large-scale pull-down/LC-MS/MS analysis to identify potential cellular targets (on and off) based on our previously established protocols. Briefly, two probes, **6-5** and **6-6b**, were designed and chemically synthesized (Scheme 6.5). In both probes, a small terminal alkyne was introduced at a suitable position (e.g. near the P_3 position) for subsequent click chemistry, followed by in-gel fluorescence and bioimaging/pull-down experiments. We had previously found, from an independent study with natural product-like probes, that such an extremely conservative modification was critical to ensure full retention of the native biological properties of

the parental compounds (e.g. **6-3b** and **6-6a**). Probes **6-5** and **6-6b** were conveniently prepared from the key intermediate **5-23**, which was obtained in four steps from phenylalanine. The subsequent transformations leading to **6-5** were similarly performed as in the case of **6-3b**. Similarly, compound **5-23** was coupled with L-alanine methyl ester hydrochloride (**6-27**) to give the corresponding dipeptide, whereupon subsequent hydrolysis was carried out under basic conditions, giving **6-29b**. Its activation with isobutyl chloroformate in the presence of *N*-methylmorpholine and the subsequent reaction of the mixed anhydride with diazomethane gave the diazomethylketone probe **6-6b**.



Scheme 6.5. Synthesis of Cbz-Phe-Ala-CHN_2 (**6-6a**) and two activity-based probes **6-5** and **6-6b.**

6.2.4 *In Situ* Proteome Profiling

We first confirmed that compounds **6-5** and **6-6b** retained the full biological activities as their parental compounds (part A & B of Figure 6.5); both the IC_{50} against recombinant rhodesain and dose-dependent inhibitory profiles against the growth of BSF/PCF of *T. brucei* were comparable to **6-3b** and **6-6a**, respectively. For example, probe **6-5** and compound **6-3b** displayed nearly identical IC_{50} values against rhodesain (0.21 and 0.19 nM, respectively; part A of Figure 6.5) and inhibitory profiles against *T. brucei* growth. Similar results were obtained for probe **6-6b** and compound **6-6a** (IC_{50} = 2.9 and 3.1 nM, respectively; for parasite inhibition profiles, compare part A of Figure 6.4 and part B of Figure 6.5). Taken together, these data showed that the introduction of a terminal alkyne handle in both **6-3b** and **6-6a** did not noticeably affect their trypanocidal activities, and confirmed **6-5** and **6-6b** were suitable ABPs for subsequent target identification/profiling studies. We next compared the *in situ* proteome reactivity profiles of two probes using live BSF and PCF by following a previously optimized procedure. Briefly, probes were directly added to the cell medium where parasites were grown. After two hours, the parasites were washed (to remove excessive probes), homogenized, incubated with rhodamine-azide under click chemistry conditions, separated by SDS-PAGE gel, and analyzed by in-gel fluorescence scanning (part C of Figure 6.5); the two probes in general showed comparable in situ proteome reactivity profiles in both forms of *T. brucei*, albeit with significant differences in relative reactivity (compare lanes 1 & 3 for BSF and lanes 5 & 7 for PCF). For example, 1 μM of **6-5** was able to give stronger

fluorescence-labeling profiles than 10 μM of **6-6b**. These results correlated well with the relative potency their trypanocidal activities. This observation, together with the fact that the fluorescence labeling of these two probes were efficiently inhibited by the addition of excessive **6-3b** and **6-6a** (lanes 2, 4, 6 & 8), indicates all the labeled proteins were likely true cellular targets of the probes, and both probes had similar cellular targets in the parasites. On a related note, as seen in lanes 1/5 (for probe **6-5**) and lanes 3/7 (for probe **6-6b**), the *in situ* proteome reactivity profiles between BSF and PCF of the parasites, though similar, showed noticeable differences, suggesting the existence of both common and unique targets in the two forms of parasites.

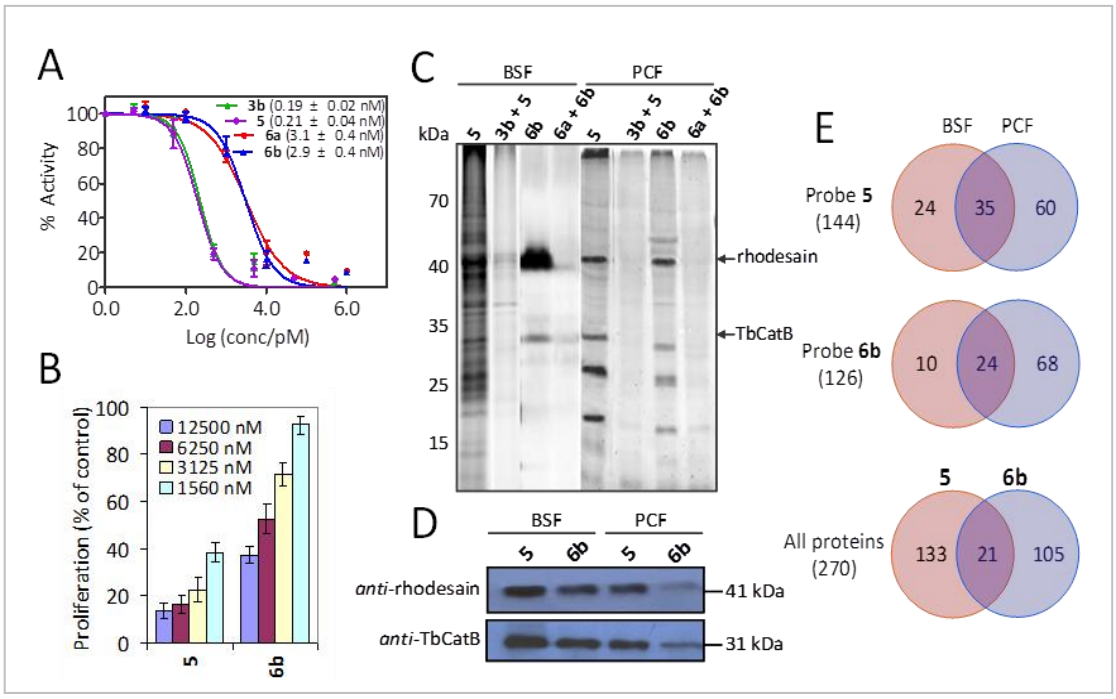


Figure 6.5. Biological evaluation of two ABPs (**6-5** and **6-6b**) in living *T. brucei*. (A) IC_{50} curves of probes and their parental compounds against rhodesain. (B) Dose-dependent, anti-trypanocidal effects of the probes against bloodstream forms of *T. brucei* after 24 h. Results represent the average standard deviation for duplicated independent trials. (C) *In situ* proteome-profiling of probes (**6-5**, 1μM; **6-6b**, 10 μM) against BSF and PCF of *T. brucei*. **6-3b** (10 μM) or **6-6a** (100 μM) was added to selected lanes in competition experiments. Putative bands of labeled rhodesain and

TbCatB were indicated (with arrows). (D) Western blotting validation of rhodesain and TbCatB labeled with **6-5** (1 μM) or **6-6b** (10 μM), following *in situ* labeling and affinity pull-down experiments. (E) Venn diagram illustrating the numbers of proteins identified from *T. brucei* with probe **6-5** (1 μM) and **6-6b** (10 μM), after *in situ* labeling and large-scale pull-down/LC-MS/MS experiments.

The covalent labeling of the two expected cellular targets, rhodesain and TbCatB (~41 kDa and ~31 kDa; part C of Figure 6.5), were unequivocally verified by affinity pull-down/western blotting with the corresponding antibodies (part D of Figure 6.5). The two probes labeled only the mature active enzyme form of rhodesain (~41 kDa), but not its proform (~48 kDa). In addition, the fluorescent TbCatB bands labeled by the two probes were detected more prominently in BSF than in PCF. This is consistent with previous findings that TbCatB was up-regulated in BSF. We also confirmed that our two newly synthesized activity-based probes were capable of labeling active cathepsin L in living HepG2 cells (Figure 6.6). As expected, only the mature active form of cathepsin L (~30 kDa) was labeled by both probes, whilst no labeling of its proforms (~41 kDa).

Finally, to identify other proteins that were also covalently labeled by the two probes, we performed large-scale proteomic analyses using **6-5** and **6-6b** in both parasite forms by affinity pull-down followed by LC-MS/MS experiments (Table 6.3). All proteins were identified with a minimum protein score of 30 as well as at least four unique peptides, and results are summarized in part E of Figure 6.5. In total, 59 & 95 proteins were identified with probe **6-5** from BSF and PCF, respectively, of

which 35 were detected in both parasitic forms. For probe **6-6b**, 34 and 92 proteins were identified in BSF and PCF, respectively, with 24 detected in both forms. When all proteins identified from both forms were pooled together, 133 and 105 of them were uniquely identified in BSF and PCF, respectively, and 21 were present in both forms. Among these proteins, some were non-specific protein binders caused by their "sticky" nature as well as their high endogenous expression level, such as carbohydrate-metabolism-related glycosomal proteins, although correct localization of glycosomal proteins has been shown to be essential for the survival of the bloodstream form as well as of the procyclic form of *T. brucei*. We therefore focused our attention on other candidate proteins having previously known nucleopholic cysteine residues in their active sites, as they are likely reactive towards both **6-5** and **6-6b**, and therefore may be true cellular targets of the corresponding parental compounds (**6-3b** and **6-6a**). As shown in Table 6.3, in addition to the expected rhodesain and TbCatB, other proteins such as CAP5.5, calpain-like cysteine peptidase, BS2, MCA4 and one proteosome subunit were identified only in BSF with either **6-5** or **6-6b**. Of note, sterol 14-alpha-demethylase (CYP51), which has been recently validated as a potential target for anti-trypanosomal therapy, was identified only with **6-6b** in both forms. Additionally, we have identified several proteins, including 3-ketoacyl-CoA thiolase, BILBO1, and succinyl-CoA synthetase alpha, all of which were previously reported to be essential for parasite survival from RNAi experiments. Finally, it should be noted that all LC-MS/MS-based results obtained above should only be used as preliminary data. Owing to the highly complex cellular environment

and the intrinsic limitation of affinity pull-down/mass spectrometric experiments, false positives/non-specific proteins binders could be minimized but not eliminated. Consequently, proper follow-up studies and validation experiments will be needed before any biological conclusions can be made for some of these protein hits.

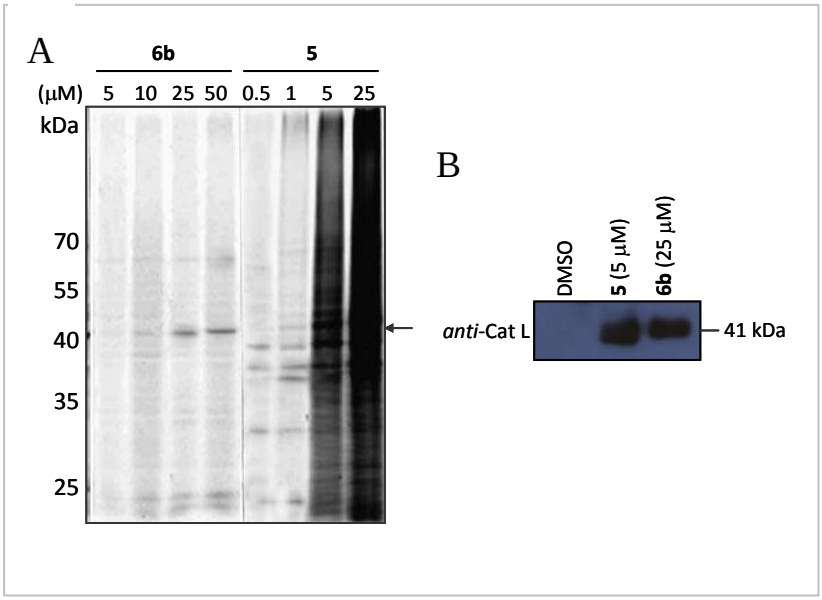


Figure 6.6. (A) Comparative studies of in situ labeling of HepG2 cells by **6-6b** & **6-5**. (B) Western blotting analysis of pulled-down fractions of HepG2 cells treated with **6-5** & DMSO with anti-cathepsin L.

Table 6.3. Representative proteins identified with **6-5** and **6-6b** in *Trypanosoma brucei*[a]

***T. brucei* gene**	**Protein name**	**Location**	**Detection**	
			6-5	**6-6b**
Tb927.6.1000	cysteine peptidase precursor (CP), Clan CA, family C1, Cathepsin L-like**	L	both	both
Tb927.6.560	cysteine peptidase C (CPC), Clan CA, family C1, Cathepsin B-like**	L	BSF	BSF
Tb927.4.3950	cytoskeleton-associated protein CAP5.5, putative,cysteine peptidase, Clan CA, family C2, putative (CAP5.5)**	n/a	both	/
Tb11.47.0035	calpain-like cysteine peptidase, cysteine peptidase, Clan CA, family C2, putative**	n/a	BSF	/
Tb927.3.3410	aspartyl aminopeptidase	n/a	BSF	/
Tb927.5.1810	lysosomal/endosomal membrane protein p67 (p67)	L	BSF	/
Tb11.01.1350	S-adenosylhomocysteine hydrolase	n/a	PCF	/
Tb927.6.950	cysteinyl-tRNA synthetase	C	PCF	/
Tb927.8.2540	3-ketoacyl-CoA thiolase*	M	PCF	/
Tb927.8.1990	tryparedoxin peroxidase (TRYP2)	M	PCF	/
Tb10.6k15.2290	protein disulfide isomerase, bloodstream- specific protein 2 precursor (BS2)*	C	/	BSF
Tb10.70.5250	metacaspase MCA4, cysteine peptidase, Clan CD, family C13 (MCA4)**	N	/	BSF
Tb927.6.1260	proteasome beta-1 subunit (TbPSB1)*	n/a	/	BSF
Tb11.02.0100	carboxypeptidase	n/a	/	PCF
Tb11.01.3960	flagellar protein essential for flagellar pocket biogenesis (BILBO1)*	F	/	PCF
Tb927.3.2230	succinyl-CoA synthetase alpha subunit*	M	/	PCF
Tb11.02.4080	sterol 14-alpha-demethylase (CYP51)**	n/a	/	PCF

[a] Symbols in the protein name column: (*) sensitive to RNA interference; (**) putative drug target.

6.2.5 Cellular Imaging Using Activity-based Probes

Having accomplished putative target identification using the corresponding ABPs, **6-5** and **6-6b**, we next investigated whether these probes could be used for drug uptake studies, as well as visualization of their sub-cellular distribution and potential cellular targets in live parasites (both BSF and PCF). To do this, live parasites were treated with probe **6-5** or **6-6b**, fixed, and reacted *in situ* with rhodamine-azide (Figure 6.7); consistent with in situ labeling profiles, as low as 0.1 μM of probe **6-5** was sufficient to stain the parasites strongly, rendering them visible under a confocal fluorescence microscope (panels b & f in part A of Figure 6.7), whereas 10 μM of **6-6b** was needed in order to produce fluorescent images with comparable probe labelling intensity (panels d & h in part A of Figure 6.7). There was no other noticeable difference in the cellular uptake of the two probes between BSF and PCF parasites. The sub-cellular distribution of the two probes, however, were notably different (part B of Figure 6.7); probe **6-6b** stained mostly lysosomal compartments of the BSF parasites, which was expected for a probe designed to target the endocytic cysteine proteases, rhodesain/TbCatB (panel j). This result was further confirmed by immunofluorescence (IF) experiments with the anti-rhodesain antibody (panel k). Interestingly, the same probe showed mostly cytosolic distribution in PCF parasites (panel n). Probe **6-5**, on the other hand, did not appear to entirely colocalize with rhodesain (panels d & h), and showed both lysosomal and cytosolic distribution in both BSF and PCF parasites (panels b & f). Surprisingly, staining of HepG2 cells (Human hepatocellular liver carcinoma) with the probes revealed that **6-5** (at l μM)

resided mostly in lysosomal compartments and colocalized with cathepsin L (Figure 6.8). On the other hand, even as high as 50 μM of **6**-**6b** only weakly stained HepG2 cells under similar conditions. Collectively, these results demonstrate that our newly synthesized ABPs could be used as potential cell-permeable bioimaging agents to study drug uptake and sub-cellular distribution.

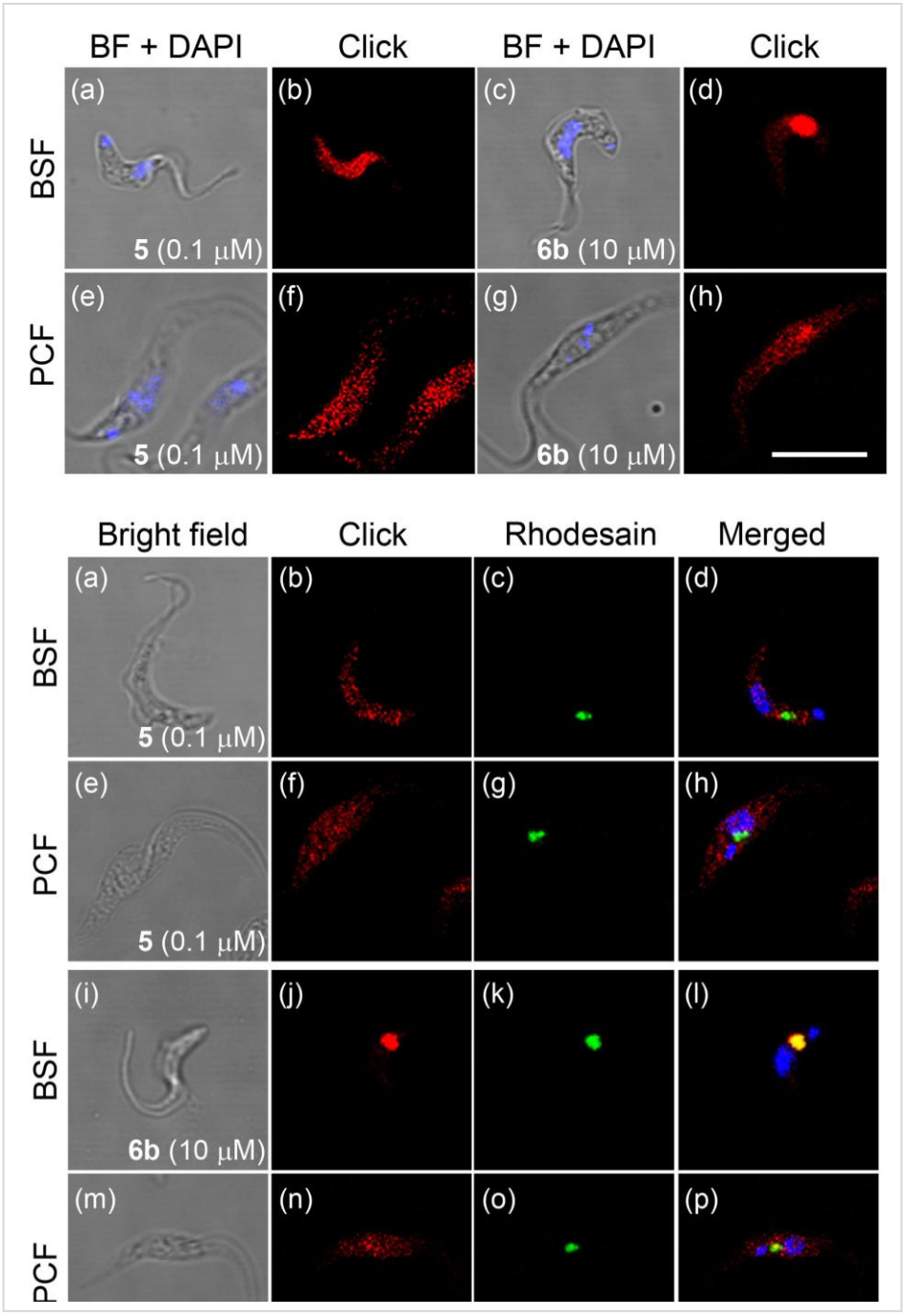


Figure 6.7. Fluorescence microscopy of cellular uptake (A) and sub-cellular localization (B) of the probes in in BSF and PCF of *T. brucei*. Live parasites were treated with 5 (0.1 μM) or 6b (10 μM) for 2 h, *in situ* reacted with rhodamine-azide (10 μM) under click chemistry conditions, and then imaged. (A) Panels a, c, e & g: merged images of bright-field with stained nuclei (DAPI; in blue); Panels b, d, f & h: 554 nm channel (pseudocolored in red) detecting cellular uptake of the probe. (B) Panels a, e, i & m: bright-field images; Panels b, f, j & n: 554 nm channel (pseudocolored in red) detecting cellular localization of the probe; Panels c, g, k & o:

Immunofluorescence (IF) staining at 488 nm channel (pseudocolored in green) to detect cellular localization of rhodesain. Anti-rhodesain primary antibody and FITC-conjugated anti-rabbit IgG secondary antibody were used; Panels d, h, l & p: merged images of panels b & c, f & g, j & k, n & o together with stained nuclei (DAPI; in blue). All images were acquired under the same settings. Scale bar = 10 μm.

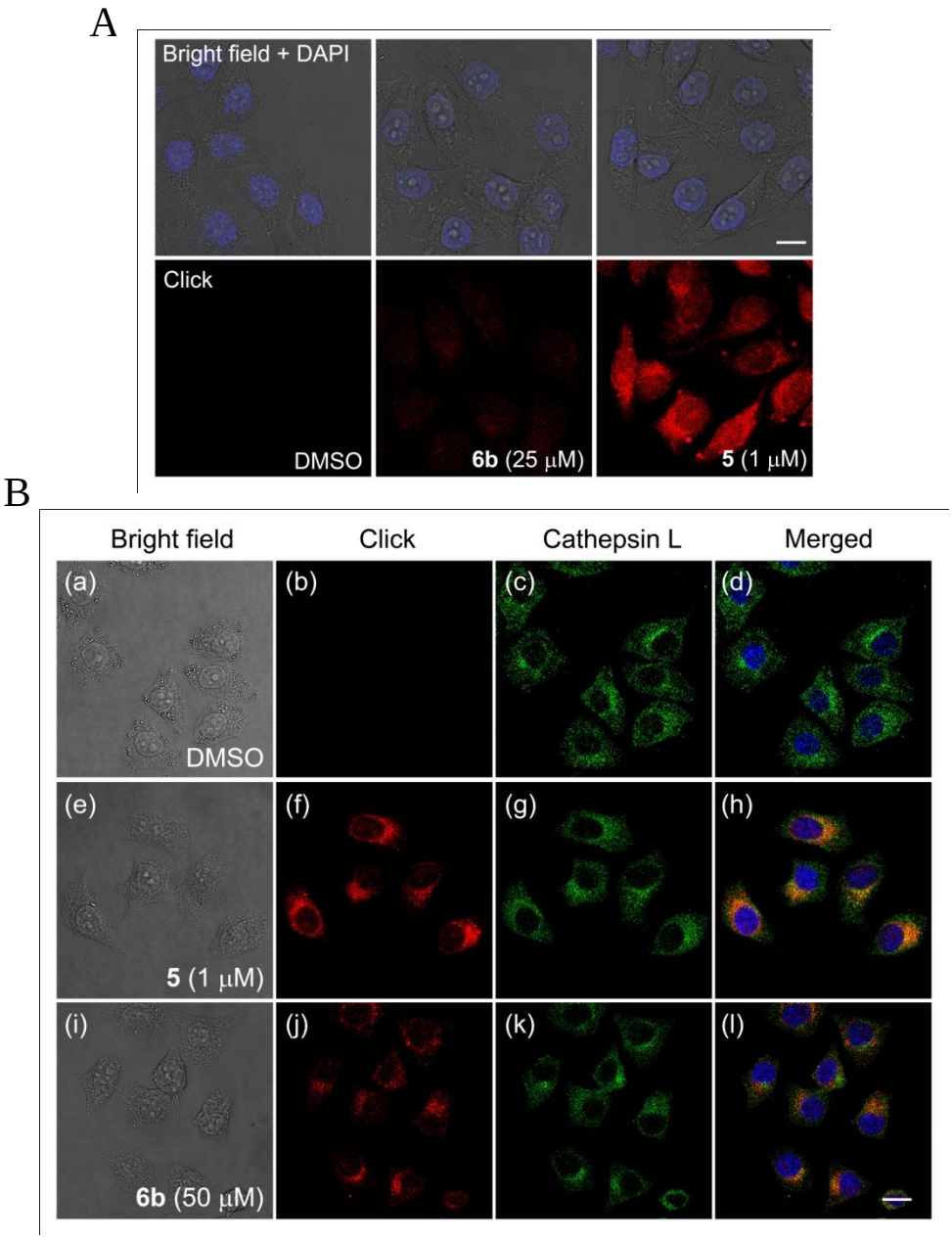


Figure 6.8. (A) Confocal microscope images of cellular uptake of **5** and **6b** within HepG2 live cells. (B) Confocal microscope images of cathepsin L in HepG2 cells treated with DMSO (top), **5** or **6b** and immunofluorescence staining. Panel (a), (e) and (i): Bright field images of the corresponding cells. Panel (b), (f) and (j): 554 nm channel (pseudocolored in red) detecting cellular localization of the corresponding

probe. Panel (c), (g) and (k): immunofluorescence staining at 488 nm channel (pseudocolored in green) using anti-cathepsin L primary antibody and FITC-conjugated anti-rabbit IgG secondary antibody detecting cellular localization of cathepsin L. Panel (d), (h) and (l): merged images of panels (b), (f) and (j), (c), (g) and (k) together with stained nuclei (with Hoechest; pseudocolored in blue). All images were acquired under the same settings. Scale bar = 10 μm.

6.3 Conclusion

In conclusion, we have synthesized a focused library of azanitrile-containing compounds and screened them for inhibitory activities against recombinant cruzain and rhodesain, as well as for their trypanocidal activities in both BSF and PCF of *T. brucei*. All compound exhibited excellent activities and showed promises as potential anti-parasitic agents. By comparing the structure-activity relationship (SAR) of these compounds, inhibition of rhodesain protease activity and trypanocidal activities were well correlated. The rational approach taken to design this novel class of compounds may provide a starting point for future development of compounds suitable for therapeutic applications, and be used to address selectivity profiles of different class of cysteine protease inhibitors (i.e., aza-nitriles vs diazomethylketones). Furthermore, based on the most potent hits from our screening results, two activity-based probes, **6-5** and **6-6b**, were subsequently developed. These cell-permeable probes enabled parasite-based proteome profiling/identification of potential cellular targets (on and off) of the parental compounds, and unequivocally confirmed that both rhodesain and TbCatB were targets of our compounds. The probes could be used in cell imaging experiments for further assessment of the drugs' uptake and their sub-cellular

distribution. Clearly, additional studies will be required for further target validation, as well as to determine drug efficacy and toxicity in animal models. Furthermore, we expect the azadipeptide nitrile inhibitors reported in this study will likely be potent against human cathepsins, due to the high sequence homology of these enzymes with rhodesain/cruzain. This, however, should not deter them from being considered as potential anti-parasitic agents. Previous studies had clearly indicated that preferential inhibition of parasite cysteine proteases over human cathespins is possible in infected human cells due to parasite localization, e.g. parasites reside in the cytoplasm of host cells, whereas the cathepsins are located in the less accessible lysosomes.

Chapter 7: Concluding Remarks

In this thesis, I have described the synthesis and biological characterization of orlistat-like natural product-based probes, K11777-like drug candidate-based probes and drug-like azadipeptide nitrile-based small molecules, determination of structure-activity relationships of these compounds, target identification and validation in subsequent molecular biology as well as cell biological experiments in both mammalian cells and *Trypanosoma brucei* parasites. It is hoped that the findings of this thesis would not only enable a more holistic understanding of orlistat from from antiobesity drug to new anticancer and trypanocidal agent as well as of K11777 as a drug candidate for parasitic infection, but also provide further directions on conversion of drugs (and drug candidates) or bioactive natural products to activity- or affinity-based probes in conjunction with mass spectrometry to comprehensively identify their direct binding targets (and off targets) on a proteome scale, thereby accelerating the clinical validation of drug candidates (drug activity and toxicity) and the discovery of the novel targets of hit compounds from, for instance, phenotypic high-throughput screening (HTS).

Chapter 8

Materials and Methods

8.1 General

8.1.1 Chemistry

All chemicals were purchased from commercial vendors and used without further purification, unless otherwise noted. Tetrahydrofuran (THF) and diether ether were distilled under an N_2 atmosphere over sodium benzophenone and used immediately. Dichloromethane (CH_2Cl_2) was distilled over CaH_2. Unless otherwise stated, all non-aqueous reactions were carried out under an atmosphere of dry argon or nitrogen in oven-dried glassware. Indictated reaction temperatures refer to those of the reaction bath, while room temperature (rt) is noted as 25 ^oC. Liquids and solutions were transferred via syringe. Reaction progress was monitored by TLC on pre-coated silica plates (Merck 60 F254, 250 μm thickness) and spots were visualized by ceric ammonium molybdate, basic $KMnO_4$, UV light or iodine. ^1H- and ^{13}C NMR spectra were recorded on a Bruker model Avance 300 MHz or DPX-300 MHz or DPX-500 MHz NMR spectrometer. Chemical shifts are reported in ppm relative to internal standard tetramethylsilane ($Si(CH_3)_4$ = 0.00 ppm) or the solvent resonance ($CDCl_3$ 7.26 ppm, 77.0 ppm, DMSO-d_6 2.50 ppm, 39.5 ppm for ^1H, ^{13}C, respectively). Data for ^1H NMR spectra are reported in terms of chemical shift (δ ppm), multiplicity, coupling constant (Hz), and number of protons. ^1H NMR coupling constants (J) are reported in Hertz (Hz) and multiplicity is indicated as follows: s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), br s (broad singlet), dd (doublet of doublet). Mass spectra were obtained on Shimadzu IT-TOF-MS or Shimadzu ESI-MS system. If need, products were purified via a Waters semipreparative HPLC equipped with a

Phenomenex Luna C18 reverse phase (5 μm, 30 mm × 75 mm) column. The mobile phase was a mixture of acetonitrile and H_2O each containing 0.1% trifluoroacetic acid.

8.1.2 Biological Assays

Cell-culture Conditions

HepG2 and HEK-293T cells were grown in DMEM (Invitrogen, Carlsbad, CA) containing 10% heat-inactivated fetal bovine serum (FBS; Gibco Invitrogen), 100 U/mL penicillin and 100 μg/mL streptomycin (Thermo Scientific, Rockford, IL) and maintained in a humidified 37 ^{o}C incubator with 5% CO_2. MCF-7 and PC-3 cells were maintained in RPMI 1640 medium supplemented with 10% FBS and 100 U/mL penicillin and 100 μg/mL streptomycin. CHO-9 cells were maintained in F12-Nutrient Mixture (Invitrogen) supplemented with 5% fetal bovine serum and 100 U/mL penicillin and 100 μg/mL streptomycin. The bloodstream forms of *T. brucei* were grown at 37 ℃ and 5% CO_2 in HMI-9 medium supplemented with 20% heat-inactivated fetal bovine serum (FBS). Procyclic forms YTAT 1.1 were grown at 28 ℃ and 5% CO_2 in Cunningham's medium supplemented with 15% heat-inactivated fetal bovine serum (FBS).

Cell Proliferation Assay (XTT)

Cell viability was determined using the XTT colorimetric cell proliferation kit (Roche) following manufacturer’s guidelines. Briefly, cells were grown to 20-30%

confluence (since they will reach ~ 90% confluence within 48 to 72 h in the absence of drugs) in 96-well plates under the conditions described above. The medium was aspirated, and then washed with PBS, and then treated, in duplicate, with 0.1 mL of the medium containing different concentrations of tested compounds. Compounds were applied from DMSO stocks whereby DMSO never exceeded 1% in the final solution. The same volume of DMSO was used as a negative control. Fresh medium, along with the compounds, were added every 24 h. After a total treatment time of 72 h, proliferation was assayed using the XTT colorimetric cell proliferation kit (Roche) following manufacturer's guidelines (read at 450 nm). Data represent the average ± s.d. for two trials.

Guava ViaCount Assay

Parasite number and percentage viability were determined in 96-well plate format using the Guava ViaCount assay on a Guava PCA-96 system (Guava Technologies, USA) following the manufacturer's instructions. Briefly, BSF and PCF trypanosomes were harvested in exponential growth phase and adjusted to a concentration of 1×10^5 parasites/mL in complete growth medium. For larger screens, the diluted trypanosomes were aliquoted (200 μL) in sterile 96-well flat white opaque culture plates (Greiner, Germany) using a Sciclone ALH 3000 Liquid Handling Workstation (Caliper Life Sciences, USA). For small-scale screens, diluted trypanosomes were dispensed manually using a multichannel pipette (BrandTech, USA). Compounds to be tested were serially diluted in DMSO by a Multidrop Combi

reagent dispenser (Thermo Scientific) for large screens requiring automation, or by a multichannel pipette for small screens conducted manually. Negative controls were performed using 1.25% DMSO. After 24 or 48 h of incubation, cell density and viability were evaluated using ViaCount assay on the Guava PCA-96 system. ED_{50} values were calculated by sigmoid curve fitting with GraphPad Prism 5.0 software (San Diego, USA). All data were collected in duplicate.

In Vitro Proteomic Profiling and In-Gel Fluorescence Scanning

To generate protein lysates, cells were washed twice with cold phosphate-buffered saline (PBS), and harvested, and collected by centrifugation. Cell pellets were resuspended in PBS and lysed by sonication. Protein concentration was determined by the Bradford assay. Cell lysates were diluted with PBS to achieve final concentration of ~ 1 mg/mL for labeling reactions. For in vitro proteomic profiling, probes were added to cell lysates (50 μg) in 50 μL PBS in the presence or absence of competitors. Unless indicated otherwise, samples were incubated for 2 h with varying concentrations of probe at room temperature. After incubation, 10 μL freshly premixed click chemistry reaction cocktail in PBS [rho-azide **2-33** (100 μM, 10 mM stock solution in DMSO), tris(2-carboxyethyl)phosphine hydrochloride (TCEP) (1 mM, 50 mM freshly prepared stock solution in deionized water), tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl] amine (TBTA) (100 μM, 10 mM stock solution in DMSO) and $CuSO_4$ (1 mM, 50 mM freshly prepared stock solution in deionized water)] was added and vortexed, then incubated for 2 h at room temperature

with gentle mixing. The reactions were terminated by the addition of pre-chilled acetone (0.5 mL), placed at -20 ℃ for 30 min and centrifuged at 13000 rpm for 10 min at 4 ℃ to precipitate proteins. The supernatant was discarded and the pellet washed two times with 200 μL of pre-chilled methanol. The protein pellets were allowed to air-dry for 10 min, resuspended in 25 μL 1×standard reducing SDS-loading buffer and heated for 10 min at 95 ℃; ~ 20 μg of protein was loaded per gel lane for separation by SDS-PAGE, then visualized by in-gel fluorescent scanning using a Typhoon 9410 Variable Mode Imager (Amersham) and images were analysed/quantitated with the ImageQuant^{TM} software (Amersham).

In Situ Proteomic Profiling and In-Gel Fluorescence Scanning

For mammalian cells, cells were grown to 80-90% confluence in 24-well plates under the conditions described above. The medium was removed, and then cells were washed twice with cold PBS, and treated with 0.5 mL of growth-medium-containing probe, with or without competitors. For *T. brucei* trypanosomes, parasites were plated into 6-well plates (PCF, 2 mL at ~1 × 10^7 cells/mL) or 25-mL cell culture flasks (BSF, 10 mL at ~2 × 10^6 cells/mL). Probes were applied from DMSO stocks whereby DMSO never exceeded 1% in the final solution. The same volume of DMSO was used as a negative control. After 2 h of incubation at 5% CO_2, the growth medium was aspirated, and cells were washed twice with PBS to remove the excessive probe, harvested, and pelleted by centrifugation. The cell pellet was resuspended in PBS (50 μL), homogenized by

sonication, and diluted to ~1 mg/mL with PBS. Probe targets were detected by click chemistry with a rho-azide **2-33**, SDS/PAGE analysis, and in-gel fluorescence scanning.

In Situ Pull-Down and Mass Spectrometry Identification

For in situ proteomic experiments, living mammalian cells or *T. brucei* parasites (T75 culturing flask) labeled in growth media with probe or DMSO (negative control) were harvested, washed, and homogenized in PBS. Fresh CuAAC reagents were added at the same concentrations as described above, except that biotin–azide **2-34** was substituted for rho–azide **2-33**. Acetone-precipitated and methanol-washed protein pellets were solubilized in PBS containing 0.1% (w/v) SDS by brief sonication. Insoluble materials were precipitated by centrifugation (13,000 g × 10 min) at 4 ^{o}C. The supernatants were then incubated with gentle shaking at 4 ^{o}C or overnight with Neutravidin agarose beads (50 μL/mg protein, Prod # 29204, Thermo Scientific, USA) which have been pre-washed twice with PBS. After centrifugation, the bead/complexes were washed extensively 4 times with 1% (w/v) SDS in PBS, three times with PBS, and twice with 250 mm ammonium bicarbonate (ABC). If need, beads were re-suspended in 500 μL of 8 M urea, and then 25 μL of 200 mM TCEP and 25 μL of 400 mM iodoacetamide were added for capping of the reactive cysteine residues. After 30 min, the beads were washed twice with 250 mm of ABC. Elution of bound proteins from beads was then performed twice using the boiling buffer (200 mM Tris pH 6.8, 400 mM DTT, 8% (w/v) SDS), then pooled. Protein samples were

concentrated using an YM-10 Centricon spin column (Millipore, USA). Following SDS-PAGE separation, protein bands were visualized by Coomassie blue/silver staining. Gel lanes corresponding to both DMSO-treataed and probe-treated samples were then each cut into 10 slices. Subsequent trypsin digestion (using In-Gel Trypsin Digestion Kit, Pierce Co., USA) and peptide extraction (with 50% acetonitrile and 1% formic acid) generated a total of 10 LCMS samples for each pull-down experiment. All samples were dried in vacuo and stored at -20 ^{o}C until LCMS analysis. Where necessary, each pull-down/LCMS experiment was performed in triplicate to ensure the high quality of MS hits obtained.

Each LCMS sample was re-suspended in 0.1% formic acid for mass spectrometry analysis as previously described. Briefly, peptides were separated and analyzed on a Shimadzu UFLC system (Shimadzu, Kyoto, Japan) coupled to an LTQ-FT Ultra (Thermo Electron, Germany). Mobile phase A (0.1% formic acid in H_2O) and mobile phase B (0.1% formic acid in acetonitrile) were used to establish the 60 min gradient comprised of 45 min of 5–35% B, 8 min of 35–50% B, and 2 min of 80% B, followed by reequilibrating at 5% B for 5 min. Peptides were then analyzed on LTQFT with an ADVANCE^{TM} CaptiveSpray^{TM} Source (Michrom BioResources, USA) at an electrospray potential of 1.5 kV. A gas-flow of 2 L/min, ion-transfer tube temperature of 180 ^{o}C and collision-gas pressure of 0.85 mTorr were used. The LTQ-FT was set to perform data acquisition in the positive ion mode as previously described except that the *m/z* range of 350–1600 was used in the full MS scan.[160a] The raw data were converted into the mgf format as described previously.[161b] The

database (76708 sequences, 33362815 residues) used for the Mascot search was a concatenated IPI protein database (or a concatenated *T. brucei* protein database). The database search was performed using an in-house Mascot server (version 2.2.07, Matrix Science, UK) with an MS tolerance of 10 ppm and MS/MS tolerance of 0.8 Da. Two missed cleavage sites of trypsin were allowed. Carbamidomethylation (C) was set as a fixed modification, and oxidation (M) and phosphorylation (S, T, and Y) were set as variable modifications.

Fluorescence Microscopy

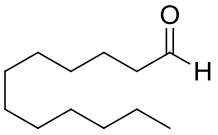
For mammalian cells, cells were seeded on sterile 12-mm glass coverslips contained in 24-well plates and grown for 24 h until ~ 60-70% confluence under culture conditions. After removing the medium and washing twice with PBS, cells were incubated in the growth medium containing probe at culture temperature and 5% CO_2 for 2 h. The cells were then washed twice with PBS, and fixed with 4% paraformaldehyde in PBS for 15 min at room temperature, washed with PBS (2 × 5 min with gentle agitation). For *T. brucei* parasites, trypanosomes (1×10^5 cells/mL for both forms) were incubated in growth medium containing probe at culture temperature and 5% CO_2 for 2 h. The parasites were then washed twice with PBS, and fixed with 4% paraformaldehyde in PBS for 15 min at room temperature and washed with PBS (2 × 5 min with gentle agitation), and then sedimented to poly-L-lysine-coated coverslips. In both cases, medium containing 1% DMSO was used as a negative control. Fixed cells were permeabilized with Triton-X 100 (0.1%

and 0.25% Triton-X 100 for mammalian cells and *T. brucei* parasites, respectively) in PBS for 15 min at room temperature, and washed with PBS (2 × 5 min with gentle agitation). The cells were blocked with 3% BSA in PBS for 30 min at room temperature, and washed with PBS (2 × 5 min with gentle agitation). The cells were then treated with a freshly pre-mixed click chemistry reaction solution [rho-azide **2-33** (10 μM final concentration from a 10 mM stock solution in DMSO), TCEP (1 mM final concentration from a 50 mM freshly prepared stock solution in deionized water), TBTA (100 μM final concentration from a 10 mM stock solution in DMSO), and $CuSO_4$ (1 mM final concentration from a 100 mM freshly prepared stock solution in deionized water)] in PBS for 1 h at room temperature. The cells were washed with PBS (1 × 5 min with gentle agitation), and cold methanol (1 × 5 min with gentle agitation), followed by 1% Tween-20 and 0.5 mM of EDTA in PBS (3 × 2 min with gentle agitation), and with PBS (2 × 5 min with gentle agitation). The cells were then incubated in PBS containing 0.25 μg/mL Hoechst or 2 μg/mL DAPI (DNA stain) for 15 min at room temperature, and washed with PBS (2 × 5 min with gentle agitation) and a final wash with deionized water (1 × 5 min with gentle agitation) before being mounted onto the Fluoromount G (Emsdiasum, USA). For immunofluorescence (IF) analysis, cells were then incubated with the corresponding primary antibodies and washed with PBS (3 × 5 min with gentle agitation), followed labeled with the corresponding secondary antibodies and a final wash with PBS (3 × 5 min with gentle agitation) before mounting.

8.2 Chapter 2

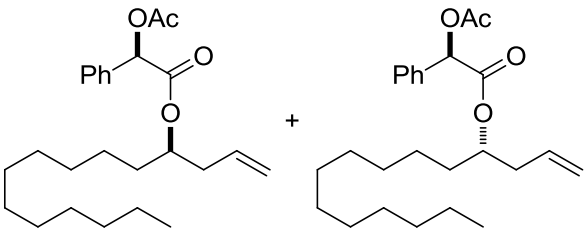
8.2.1 Chemical Syntheis

Dodecanal (2-4)



To a cooled (0 ℃) and stirred suspension of PCC (16.2 g, 75 mmol) in DCM (100 mL) was added dodecan-1-ol (11.2 mL, 50 mmol) in one lot. After being stirred for 3 h, when the reaction was complete, the reaction was quenched with anhydrous ether (100 mL). The mixture was stirred vigorously and the supernatant passed through a small pad of silica gel. The eluent was concentrated in vacuo to obtain **2-4** (8.3 g, 90%), which was purified by column chromatography on silica gel (5% EtOAc/hexanes). ^{1}H NMR (300 MHz, $CDCl_3$): δ 9.76 (t, *J* = 1.9 Hz, 1H), 2.41 (td, *J* = 1.8, 7.4 Hz, 2H), 1.58-1.65 (m, 2H), 1.27 (br s, 16H), 0.88 (t-like, 3H).

Synthesis of 2-5a and 2-5b



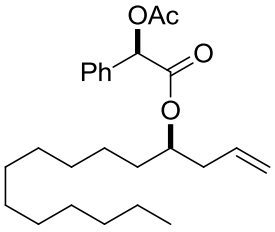
2-5a **2-5b**

To a stirred solution of **2-4** (5.18 g, 23.48 mmol) in dry THF (40 mL) cooled at -10 ^{o}C was added a 1 M solution of allylmagnesium bromide (35.2 mL, 35.2 mmol) in ether dropwise via syringe. The reaction mixture was stirred for 1 h at -10 ^{o}C, saturated aqueous NH_4Cl (50 mL) was added, and the mixture was partitioned between ether (300 mL) and water (30 mL). The organic layer was washed once with water (16 mL) and brine (30 mL) and dried (Na_2SO_4) and the solvent

removed in vacuo. Flash chromatography on silica gel (hexanes/EtOAc (24:1-16:1)) afforded 3.87 g (61%) of the homoallylic alcohol as a colorless oil which solidified below 0 ^{o}C. ^{l}H NMR (300 MHz, $CDCl_3$): δ 5.76-5.90 (m, 1H), 5.15 (br d, *J* = 2.9 Hz, 1H), 5.12 (s, 1H), 3.63-3.66 (m, 1H), 2.11-2.34 (m, 2H), 2.09 (br s, 1H), 1.63 (br s, 3H), 1.26 (s, 17H), 0.88 (t-like, 3H); ^{l3}C NMR (75 MHz, $CDCl_3$): δ 135.0, 118.0, 70.7, 42.0, 36.9, 31.9, 29.7, 29.6, 29.4, 25.7, 22.7, 14.1.

To a stirred solution of (*R*)-(-)-*O*-acetylmandelic acid (4.97 g, 25.6 mmol) and DMAP (1.24 g, 10.15 mmol) in dry DCM (46 mL) cooled at 0 ^{o}C was added dropwise a solution of the above homoallylic alcohol (4.6 g, 20.33 mmol) in DCM (46 mL) followed by a solution of DCC (5.25 g, 25.45 mmol) in DCM (40 mL) via cannula. The reaction mixture was stirred at 0 ^{o}C for 30 min, warmed to room temperature, and stirred for 1 h. The white precipitate was removed by filtration through a small pad of Silica gel and the pad washed with DCM (5 × 70 mL). The filtrate chromatography on silica gel (hexanes/benzene (1:9)) yielded 7.8 g (94%) of the mandelate esters **2-5a** and **2-5b** as colorless oils.

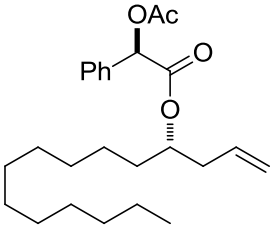
(4R)-4-[(2R)-Acetoxy-2-phenylacetoxy]pentadec-1-ene (2-5a)



^{l}H NMR (300 MHz, $CDCl_3$): δ 7.45-7.48 (m, 2H), 7.35-7.37 (m, 3H), 5.88 (s, 1H), 5.36-5.49 (m, 1H), 4.93 (app quint. *J* = 6.2 Hz, 1H), 4.75-4.81 (m, 2H), 2.15-2.19 (m, 5H), 1.54 (br s, 2H), 1.25 (br s, 18H), 0.88 (t-like, 3H); ^{l3}C NMR (75 MHz, $CDCl_3$): δ 170.2, 168.5, 134.0, 132.9, 129.1, 128.6, 127.6, 117.6, 75.0, 74.7, 38.3, 33.5, 31.9, 29.6, 29.5, 29.4

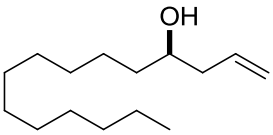
(2), 25.1, 22.7, 20.7, 14.1.

(R)-(S)-pentadec-1-en-4-yl 2-acetoxy-2-phenylacetate (2-5b)



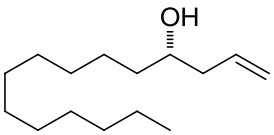
^{1}H NMR (300 MHz, $CDCl_3$): δ 7.45-7.49 (m, 2H), 7.35-7.37 (m, 3H), 5.88 (s, 1H), 5.66-5.80 (m, 1H), 5.08 (d, *J* = 6.4 Hz, 1H), 5.04 (s, 1H), 4.91 (quint., *J* = 6.2 Hz, 1H), 2.32 (t, *J* = 6.7 Hz, 2H), 2.17 (s, 3H), 1.07-1.45 (br m, 20H), 0.89 (t-like, 3H); ^{l3}C NMR (75 MHz, $CDCl_3$): δ 170.3, 168.6, 134.1, 132.9, 129.1, 128.6, 127.6, 117.9, 75.0, 74.7, 38.5, 33.2, 31.9, 29.6, 29.4, 29.3, 29.2, 24.7, 22.7, 20.7, 14.1.

(R)-4-Hydroxypentadec-1-en (2-6a)



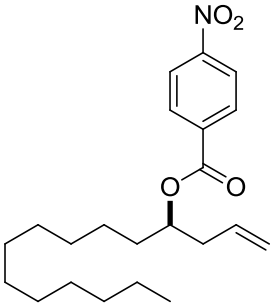
To a stirred solution of **2-5a** (3.15 g, 7.84 mmol) in MeOH (50 mL) was added dropwise 2 *N* KOH (19.6 mL, 39.21 mmol). The reaction mixture was heated at 75 ^{o}C for 6 h, cooled, and concentrated in vacuo. The residue was diluted with water (10 mL), acidified in cold 1 *N* HCl, and extracted with ether (3 × 100 mL). The combined organic extracts were washed with water (20 mL) and brine (20 mL) and dried (Na_2SO_4) and the solvent was removed in vacuo. Flash chromatography on silica gel (hexanes/EtOAc (16:1-12:l)) afforded 1.7 g (99%) of **2-6a** as a colorless oil, $[a]_D$ + 5.78^{o} (*c* 2.89, $CHCl_3$). The absolute configurations of resolved alcohols were assigned based on comparisons of optical rotations to literature values.

(*S*)-*Hydroxypentadec-l-ene (2-6b)*



Prepared according to the method of **2-6a** using **2-5b** (0.511 g, 1.27 mmol), and 2 *N* KOH (3.18 mL, 6.35 mmol) in MeOH (8 mL). Purification by flash chromatography on silica gel (hexanes/EtOAc (16:l- 12:l)) afforded 0.28 g (99%) of **2-6b** as a colorless oil, $[a]_D$ -6.63^o (*c* 1.69, $CHCl_3$).

Synthesis of 2-6a from 2-6b

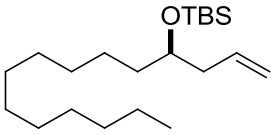


To a stirred solution of **2-6b** (1.13 g, 5.0 mmol) and Ph_3P (6.56 g, 25.0 mmol) in dry ether/toluene (3:l) was added *p*-nitrobenzoic acid (3.76 g, 22.5 mmol) followed by dropwise addition of DEAD (3.94 mL, 25 mmol) at room temperature. The reaction mixture was stirred for 30 min and the solvent was removed in vacuo to afford a residue, which after flash chromatograhy on silica gel (hexanes/EtOAc (49:l-39:1)) yielded 1.62 g (86%) of the *p*-nitrobenzoate ester as a colorless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.28 (quint. *J* = 2.0 Hz, 2H), 8.20 (dd, *J* = 2.1, 6.9 Hz, 2H), 5.74-5.89 (m, 1H), 5.06-5.23 (m, 3H), 2.45-2.49 (m, 2H), 1.68-1.77 (m, 2H), 1.25 (br s, 18H), 0.87 (t-like, 3H); ^{l3}C NMR (75 MHz, $CDCl_3$): δ 164.3, 150.5, 136.1, 133.3, 130.6, 123.5, 118.1, 75.4, 38.6, 33.6, 31.9, 29.6, 29.5 (2), 29.4, 29.3, 25.3, 22.7, 14.1.

Treatment of the above ester with K_2CO_3 in methanol afforded **2-6a** in quantitative yield. Briefly, to a stirred solution of the *p*-nitrobenzoate ester (1.61 g, 4.3 mmol) in MeOH (40 mL) was added K_2CO_3 (2.38 mL, 17.2 mmol) at 0 ^{o}C. The

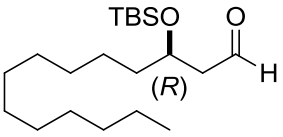
reaction mixture was stirred at room temperature for 12 h, filtered and concentrated in vacuo. The residue was diluted with water (40 mL), acidified in cold HOAc, and extracted with ethyl acetate (3 × 40 mL). The combined organic extracts were washed with water and brine and dried (Na_2SO_4), and the solvent was removed in vacuo. Flash chromatography on silica gel (hexanes/EtOAc (16:l- 12:l)) afforded 0.96 g (99%) of **2-6a** as a colorless oil.

(R)-tert-butyldimethyl(pentadec-1-en-4-yloxy)silane (2-7a)



To a solution of (*R*)-**2-6b** (1.62 g, 7.16 mmol) and imidazole (0.54 g, 7.85 mmol) in DMF (3.7 mL) was added TBSCl (1.19 g, 7.82mmol) as a solution of 4 mL of DMF at room temperature. The resulting mixture was stirred overnight, diluted with ether (50 mL), washed with brine, dried over $MgSO_4$, filtered, concentrated in vacuo, and finally purified by flash chromatography on silica gel (hexanes) to give 2.32 g (95%) of (*R*)-silyl ether **2-7a** as a colorless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 5.76-5.90 (m, 1H), 5.11-5.16 (m, 2H), 3.64 (br s, 1H), 2.09-2.34 (m, 2H), 1.46-1.65 (br s, 3H), 1.26 (s, 17H), 0.88 (br m, 12H), 0.06 (d, *J* = 5.0 Hz, 6H).

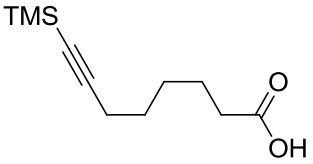
(R)-3-(tert-butyldimethylsilyloxy)tetradecanal (2-8a)



Ozone was bubbled into a stirred MeOH-DCM (1:1, 35 mL) solution of 1.1 g of (*R*)-silyl ether **2-7a** (3.2 mmol) at -78 ^{o}C until the blue color persisted for ~4 min. methyl sulfide (7.0 mL, 94.4 mmol) and

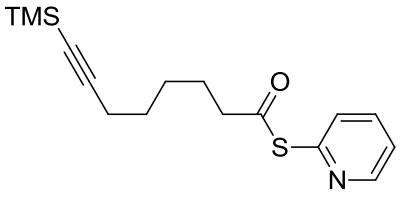
triethylamine (0.7 mL, 4.97 mmol) were added and the cooling bath was removed. The reaction mixture was warmed to room temperature, stirred overnight, concentrated in vacuo. Purification by flash chromatography on silica gel (EtOAc/hexanes 1:29) afforded 1.0 g (93%) of (*R*)-aldehyde **2-8a** as a colorless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 9.81 (t, *J* = 2.4 Hz, 1H), 4.17 (q, *J* = 5.8 Hz, 1H), 2.51 (dd, *J* = 2.4, 5.6 Hz, 2H), 1.51 (br m, 4H), 1.26 (br s, 18H), 0.88 (br m, 12H), 0.06 (d, *J* = 5.0 Hz, 6H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 202.5, 68.3, 50.8, 37.9, 31.9, 29.6, 29.5, 29.3, 25.8, 25.1, 22.7, 14.1, - 4.4, -4.7.

8-(trimethylsilyl)-oct-7-ynoic acid (2-9)



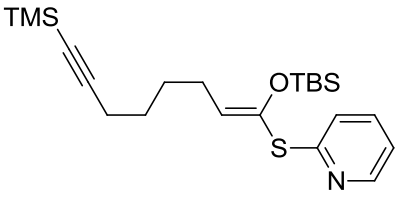
To a solution (trimethylsilyl)acethylene (3.4 mL, 24 mmol) in THF (20 mL) was added *n*-BuLi (16.5 mL, 26.4 mmol, 1.6M soln in hexanes) at –78 ^{o}C. After the mixture was stirred for 20 min, a solution of 6-bromohexanoic acid (2.6 g, 13.4 mmol) in dry THF (60 mL) and anhydrous HMPA (40 mL) was added via syringe over 10 min at –78 ^{o}C. The reaction mixture was stirred for 2 h, then quenched with saturated NH_4Cl solution, and diluted with ether, washed with brine, dried over Na_2SO_4, filtered, and concentrated in vacuo. The resulting residue was purified by flash column chromatography on silica gel (50% EtOAc/hexanes) to give **2-9** (2.32 g, 82%) as a colorless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 2.37 (t, *J* = 7.4 Hz, 2H), 2.23 (t, *J* = 6.72 Hz, 2H), 1.44-1.68 (m, 6H), 0.14 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 179.7, 107.2, 84.7, 33.9, 28.2(2), 24.2, 19.7, 0.2.

S-pyridin-2-yl 8-(trimethylsilyl)-oct-7-ynethioate (2-10)



To a stirred solution of acid **2-9** (2.27 g, 10.7 mmol) in DCM (30 mL) was added oxalyl chloride (1.4 mL, 16.0 mmol) at 0 ^{o}C. The reaction mixture was stirred for 2 h at ambient temperature. The solvent was evaporated under reduced pressure. To a stirred solution of the oil residue was added 2-thiopyridine (1.42 g, 12.8 mmol) in DCM (30 mL), followed by Et_3N (3.0 mL, 21.3 mmol) at 0 ^{o}C. The reaction mixture was stirred for 2 h at ambient temperature and then quenched with 1*N* HCl (30 mL) and neutralized by washing with saturated $NaHCO_3$ solution. The organic layers were combined and evaporated, then resulting yellow oil was purified by flash chromatography on silica gel (20% EtOAc/hexanes) to afford thiopyridyl ester **2-10** (2.55 g, 78%) as a yellow oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.61-8.63 (m, 1H), 7.71-7.77 (m, 1H), 7.61 (d, *J* = 7.9 Hz, 1H), 7.28-7.31 (m, 1H), 2.72 (t, *J* = 7.6 Hz, 2H), 2.32 (t, *J* = 6.9 Hz, 2H), 1.70-1.80 (m, 1H), 1.44-1.57 (m, 4H), 0.14 (s, 9H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 196.5, 151.6, 150.3, 137.0, 130.1, 123.4, 107.5, 84.3, 44.1, 28.7, 27.9, 25.3, 19.8, 0.2.

(E)-2-(1-(tert-butyldimethylsilyloxy)-8-(trimethylsilyl)-oct-1-en-7-ynylthio)-pyridine (2-11)

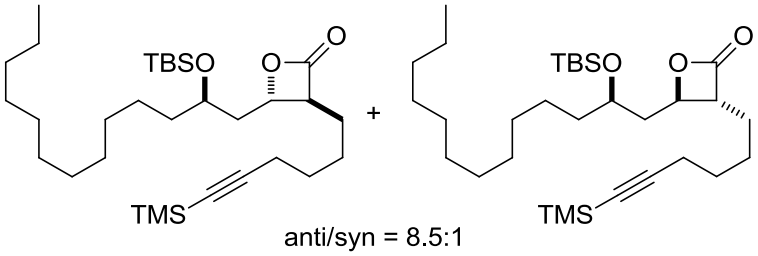


Thiopyridyl ester **2-10** (5.42 g, 17.9 mmol) was dried azeotropically with xylenes. After purging with

N_2, anhydrous DMF (1.57 g, 21.5 mmol) and anhydrous Et_3N (3.02 g, 21.5 mmol) in dry DCM (120 mL) was added and the mixture was cooled to −78 ^oC. After 10 min, LiHMDS (1.0 M, 41.3 mL, 41.3 mmol) was added dropwise. The reaction mixture was stirred for 30 min and then TBSCl (5.4 g, 35.8 mmol) was added dropwise to the reaction mixture at −78 ^oC. The reaction mixture was stirred for 2 h at −78 ^oC and then quenched with pH 7 phosphate buffer (20 mL × 2). The organic layer was separated, dried over Na_2SO_4 and evaporated. The resulting yellowish oil was purified by flash chromatography on silica gel (10% EtOAc/hexanes) to afford ketene acetal **2-11** (5.6 g, 74 %) as a yellow oil. ^1H NMR (500 MHz, $CDCl_3$): δ 8.40-8.50 (m, 1H), 7.54 (dt, *J* = 1.9, 7.6, Hz, 1H), 7.31-7.33 (d, *J* = 8.2 Hz, 1H), 6.98-7.01 (m, 1H), 5.39 (t, *J* = 7.0 Hz, 1H), 2.34 (dd, *J* = 6.9, 13.8 Hz, 2H), 2.20 (dd, *J* = 5.7, 12.6 Hz, 2H), 1.50-1.56 (m, 4H), 0.88 (s, 9H), 0.14 (s, 9H), 0.08(s, 6H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 174.1, 160.4, 149.4, 139.7, 136.5, 123.4, 121.5, 119.6, 107.3, 84.5, 35.9, 28.4, 28.3, 26.3, 25.7, 25.6, 25.0, 19.7, 18.1, 17.6, 0.2. -4.8; ESI-MS: (*m/z*) calcd for $C_{22}H_{37}NOSSi_2$ $[M+H]^+$ 420.213, found 420.244.

(3S,4S)-4-((R)-2-hydroxytridecyl)-3-(6-(trimethylsilyl)hex-5-ynyl)oxetan-2-one (2-12)

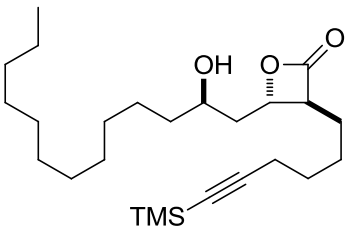
<u>TMAL reaction:</u>



$ZnCl_2$ (796 mg, 5.84 mmol) was fused under vacuum and then

allowed to cool to room temperature under a flow of nitrogen. 15 mL of dry DCM was then added via syringe followed by a solution of **2-8a** (547 mg, 2.92 mmol) in 5 mL of DCM and ketene acetal **2-11** (1.72 g, 4.08 mmol). The reaction mixture was stirred for 72 h at ambient temperature. 10 mL of pH 7 phosphate buffer were then added and the mixture was stirred vigorously for 45 min, filtered through a pad of celite, and washed with DCM. The organic filtrate was dried over Na_2SO_4, filtered, and concentrated. The residue was purified by flash chromatography on silica gel (10% EtOAc/hexanes) to provide the mixture of silyloxy-β-lactone diastereomers as white solids. ^{1}H NMR (500 MHz) analysis of the mixture of diastereomers indicated a diastereomeric ratio of 8.5:1. Without further purification, the mixture was used directly in the next step.

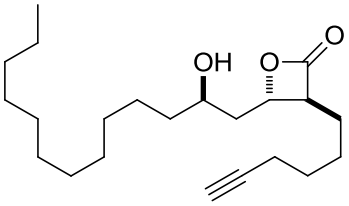
***O*-desilylation:**



To a stirred solution of the mixture of silyloxy-β-lactones in 48 mL of CH_3CN cooled to 0 ℃, 4.8 mL of HF (48%) was added dropwise. The mixture was stirred at 0 ℃ for 2 h, then allowed to warm to ambient temperature and stirred for an additional 8 h. The reaction mixture was diluted with 150 mL of Et_2O, quenched carefully with cold saturated $NaHCO_3$ solution, and washed with brine. The organic layer was dried over Na_2SO_4, filtered, and concentrated. The residue was purified by flash chromatography on silica gel (10% EtOAc/hexane) to provide the hydroxy-β-lactone **2-12** (360 mg, 46%) and a mixture of the two diastereomers (90

mg, 12%) as white solids (58% overall two steps). Spectroscopic data are reported for the major diastereomeric-β-lactone **2-12**. ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.50 (app quint., *J* = 4.4 Hz, 1H), 3.78-3.83 (m, 1H), 3.27 (dt, *J* = 3.8, 7.6 Hz, 1H), 2.24 (t, *J* = 7.0 Hz, 2H), 1.77-1.96 (m, 4H), 1.49-1.57 (m, 8H), 1.26 (br s, 18H), 0.88 (t, *J* = 7.0 Hz, 3H), 0.15 (s, 9H); ^{13}C-NMR (126 MHz, $CDCl_3$): δ 171.3, 106.7, 85.0, 75.6, 68.5, 56.5, 41.8, 38.2, 31.9, 29.6, 29.5, 29.3, 28.1, 27.3, 26.9, 25.9, 25.4, 22.7, 19.6, 14.1; ESI-MS: (*m/z*) calcd for $C_{25}H_{46}O_3Si$ $[M+H]^+$ 423.32, found 423.25.

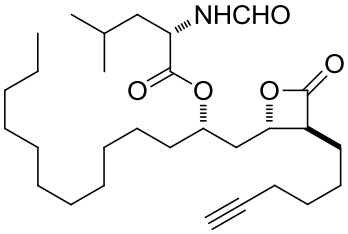
(3S,4S)-3-(hex-5-yn-1-yl)-4-((R)-2-hydroxytridecyl)oxetan-2-one (2-13)



***C*-desilylation:** To a solution of (trimethylsily1)-acetylene **2-12** (134 mg, 0.5 mmol) in 20 mL of a mixture of acetone/H_2O/2,6-lutidine (v/v/v 1: 1: 0.l) was added solid $AgNO_3$ (850 mg, 5 mmol, 10 equiv.). The white suspension was stirred vigorously for 4 h, and then 5 mL of a 1.0 M aqueous KH_2PO_4 solution was added. The resulting yellow slurry was stirred for an additional 30 min. Filtration of the reaction mixture through Celite removed most of the yellow precipitate. The filtrate was extracted with Et_2O, and the combined organic extracts were washed once with brine and dried over Na_2SO_4. Evaporation of the solvent under reduced pressure afforded a pale yellow oil. Purification by chromatography on silica gel (15:l hexanes/EtOAc) gave 81 mg of **2-13** (83%) as a sticky, colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.50 (app quint., *J* = 4.4 Hz, 1H), 3.78-3.83 (m, 1H), 3.28 (dt, *J* = 3.8, 8.2 Hz, 1H), 2.19-2.23 (m, 2H), 1.91-1.97 (m, 2H), 1.77-1.87 (m, 2H), 1.47-1.62 (m,

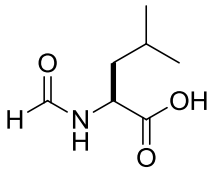
8H), 1.26 (br s, 18H) 0.88 (t, J = 6.3 Hz, 3H); ^{13}C-NMR (126 MHz, $CDCl_3$): δ 171.4, 84.0, 75.6, 68.7, 68.5, 56.4, 41.8, 38.2, 31.9, 29.6(2), 29.5(2), 29.3, 28.0, 27.2, 25.8, 25.4, 22.7, 18.1, 14.1; IT-TOF-MS: (*m/z*) calcd for $C_{22}H_{38}O_3$ $[M+H]^+$ 351.282, found 351.288.

(S)-(S)-1-((2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl)tridecan-2-yl 2-formamido-4-methylpentanoate (2-1, THL-R)



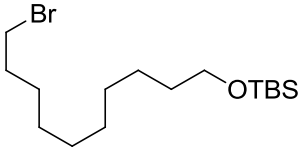
Mitsunobu reaction: β-lactone **2-13** (70 mg, 0.2 mmol), PPh_3 (79 mg, 0.3 mmol), and *N*-formyl-L-leucine (128 mg, 0.8 mmol) were placed in a RB-flask and azeotroped under vacuum with 0.5 mL of xylene for 30 min. Addition of 5.0 mL of dry THF was followed by cooling to 0 ℃, DIAD (60 μL, 0.3 mmol) was then added via syringe. The mixture was stirred at 0 ℃ for 10 min, allowed to warm to ambient temperature. The reaction was monitored by TLC. After the reaction was complete, the mixture was then concentrated in vacuo and polar impurities were separated by flash chromatography on silica gel (60:40, hexanes/EtOAc). Further purification of the residue by flash chromatography on silica gel (20% EtOAc/hexanes) yielded β-lactone **2-1**. (74 mg, 75%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.22 (s, 1H), 5.92 (d, J = 8.2 Hz, 1H), 5.01-5.06 (m, 1H), 4.68 (dt, J = 4.4, 8.8 Hz, 1H), 4.30 (q, J = 4.4 Hz, 1H), 3.23 (dt, J = 3.8, 7.6 Hz, 1H), 2.14-2.23 (m, 3H), 1.96 (t, J = 2.5 Hz, 1H), 1.95-2.04 (m, 1H), 1.55-1.82 (m, 11H), 1.25-1.27 (m, 18H), 0.96-0.98 (m, 6H), 0.88 (t, J = 7.0 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.9, 170.5, 160.6, 83.9, 74.7, 72.7, 68.8, 56.9, 49.7, 41.5, 38.7, 34.1, 31.9, 29.6, 29.5, 29.4, 29.3(2), 28.0, 27.1, 25.7, 25.1, 24.9, 22.9, 22.7, 21.9, 21.7, 18.1, 14.1; ESI-MS: (*m/z*) calcd for $C_{29}H_{49}NO_5$ $[M+H]^+$ 492.4, found 492.3.

(S)-(+)-N-Formyl leucine (2-14)



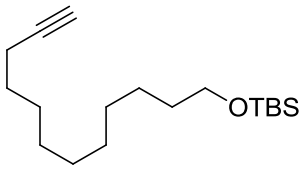
To a stirred solution of L-leucine (1.31 g, 10 mmol) in formic acid (50 mL) was added dropwise acetic anhydride (28 mL, 300 mmol) at 0 ^{o}C. After addition of the acetic anhydride, the solution was stirred at room temperature overnight. The solution was treated with water (40 mL). The solvent was removed under reduced pressure to yield a white residue. This residue was recrystallized from water to yield (*S*)-(+)-*N*-formyl leucine **2-14** (82%). ^{1}H NMR (300 MHz, DMSO-d_6) δ: 8.23 (s, 1H), 5.95 (d, *J* = 8.0 Hz, 1H), 4.10 (m, 1H), 1.70–1.92 (m, 3H), 0.97 (d, *J* = 6.2 Hz, 6H).

(10-Bromodecyloxy)(tert-butyl)dimethylsilane (2-15)



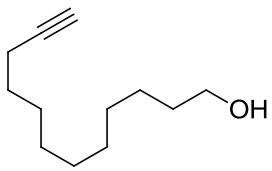
Imidazole (2.15 g, 31.6 mmol), DMAP (258 mg, 2.11 mmol) and TBSCl (3.81 g, 25.3 mmol) was added to a solution of 10-bromodecan-1-ol (5.0 g, 21.1 mmol) in DCM (60 mL) at 0 ℃. The reaction mixture was warmed to room temperature and stirred for 5 h. The organic phase was washed with water, brine, dried over anhydrous Na_2SO_4 and concentrated in vacuo. The crude product was then purified by flash column chromatography on silica gel (2% EtOAc/hexanes) to provide **2-15** (7.31 g, 99%) as a colourless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 3.60 (t, *J* = 6.6 Hz, 2H), 3.40 (t, *J* = 6.9 Hz, 2H), 1.85 (q, *J* = 6.9 Hz, 2H), 1.53-1.40 (m, 4H), 1.29 (br s, 10H), 0,89 (s, 9H), 0.05 (s, 6H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 63.3, 34.0, 32.9, 32.8, 29.5, 29.4, 28.7, 28.2, 26.0, 25.8, 18.4, -5.3; IT-TOF-MS: (*m/z*) calcd for $C_{16}H_{35}BrOSi$ $[M+1]^+$ 351.164, found 351.226.

tert-Butyl(dodec-11-ynyloxy)dimethylsilane (2-16)



To a stirred solution of **2-15** (6.0 g, 17.1 mmol) in THF (15 mL) was added dropwise a suspension of lithium acetylene (3.14 g, 34.1 mmol) in DMSO-THF (35 mL, 5:2) at 0 ℃. The reaction mixture was warmed to room temperature and stirred for 4.5 h, then quenched with saturated NH_4Cl solution (50 mL). The mixture was diluted with water and extracted with Et_2O. The combined organic layers were washed with brine, dried over anhydrous Na_2SO_4 and concentrated in vacuo. The crude alkyne was purified by flash column chromatography on silica gel (2% EtOAc/hexanes) to give **2-16** (2.38 g, 47%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.59 (t, *J* = 7.0 Hz, 2H), 2.18 (dt, *J* = 6.9, 2.5 Hz, 2H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.55-1.37 (m, 6H), 1.28 (br s, 10H), 0.89 (s, 9H), 0.04 (s, 6H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 84.8, 68.0, 63.3, 41.0, 32.9, 29.6, 29.44, 29.41, 29.1, 28.8, 28.5, 26.0, 25.8, 18.4, -5.25; ESI-MS: (*m/z*) calcd for $C_{18}H_{36}OSi$ $[M+H]^+$ 297.3, found 297.2.

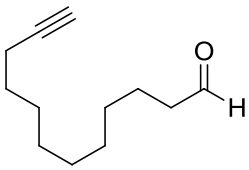
Dodec-11-yn-1-ol (2-17)



PPTS (510 mg, 2.0 mmol) was added to a solution of **2-16** (3.0 g, 10.1 mmol) in 95% EtOH (30 mL) and stirred at room temperature for 38 h. The reaction mixture was quenched with brine (30 mL) and extracted with Et_2O. The combined organic layers were washed with brine, dried over anhydrous Na_2SO_4 and concentrated in vacuo. The crude alcohol was purified by column chromatography on silica gel (10% EtOAc/hexanes)

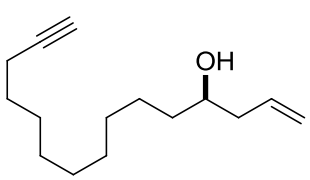
to yield **2-17** (1.619 g, 88%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 3.61 (t, *J* = 6.95, 2H), 2.16 (dt, *J* = 7.0, 2.5 Hz, 2H), 1.92 (s, 1H), 1.52 (m, 5H), 1.27 (br s, 11H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 84.7, 68.0, 63.0, 32.7, 29.5, 29.4, 29.3, 29.0, 28.7, 28.4, 25.7, 18.3.

Dodec-11-ynal (2-18)



To a solution of alcohol **2-17** (1.0 g, 5.5 mmol) in DCM (10 mL) was added a suspension of PCC (1.78 g, 8.3 mmol) in DCM (25 mL) at 0 ℃. The reaction mixture was warmed to room temperature and stirred for 3.5 h and then filtered through a pad of celite. The filtrate was concentrated under reduced pressure and the crude aldehyde was purified by column chromatography on silica gel (hexanes/EtOAc, 100:0 to 96:4) to provide aldehyde **2-18** (735 mg, 74%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.76 (t, *J* = 1.25 Hz, 1H), 2.42 (t, *J* = 7.6 Hz, 2H), 2.18 (dt, *J* = 7.0, 2.5 Hz, 2H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.62 (quin, *J* = 7.0 Hz, 2H), 1.52 (q, *J* = 7.0 Hz, 2H), 1.38 (m, 2H), 1.30 (br s, 8H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 202.9, 84.7, 68.1, 43.9, 29.3, 29.2, 29.1, 29.0, 28.7, 28.4, 22.0, 18.4.

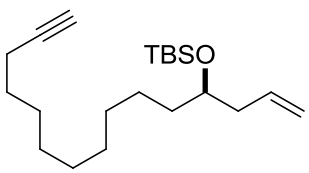
(R)-Pentadec-1-en-14-yn-4-ol (2-19)



(*S*)-BINOL (415 mg, 1.45 mmol) was added to a suspension of $InCl_3$ (292 mg, 1.32 mmol, azeotropically dried over THF (2 × 3 mL)) in DCM (15 mL) and stirred for 2 h at

room temperature. Allyltributyltin (4.37 g, 13.2 mmol) was added and stirred for 5 min, then the reaction mixture was cooled to -78 ℃ and stirred for 10 min. Aldehyde **2-18** (1.19 g, 6.6 mmol) in DCM (10 mL) was added dropwise and stirred at -78 ℃ for 4 h then warmed to room temperature overnight. Saturated $NaHCO_3$ solution (25 mL) was added to the reaction mixture and stirred for 30 min. The aqueous phase was extracted with DCM. The combined organic layers were washed with brine, dried over anhydrous Na_2SO_4 and concentrated under reduced pressure. The crude product was purified by column chromatography on silica gel (hexanes/EtOAc, 95:5) to give homoallylic alcohol **2-19** (1.06 g, 71%, 62.4% ee) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.79-5.87 (m, 1H), 5.15 (br d, *J* = 3.2 Hz, 1H), 5.12 (s, 1H), 3.62-3.67 (m, 1H), 2.28-2.33 (m, 1H), 2.11-2.19 (m, 3H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.29-1.57 (m, 16H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 134.9, 118.1, 84.8, 70.7, 68.0, 42.0, 36.8, 29.6, 29.5, 29.4, 29.1, 28.7, 28.5, 25.6, 18.4.

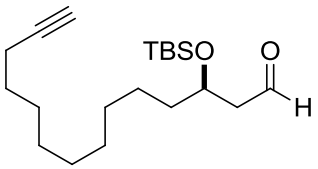
(R)-tert-Butyldimethyl(pentadec-1-en-14-yn-4-yloxy)silane (2-20)



Imidazole (848 mg, 12.5 mmol), DMAP (51 mg, 0.42 mmol) and TBSCl (1.13 g, 1.47 mmol) was added to a solution of alcohol **2-19** (980 mg, 4.15 mmol) in DCM (20 mL) at 0 ℃. The reaction mixture was warmed to room temperture and stirred for 7.5 h. The organic phase was washed with water, brine, dried over anhydrous Na_2SO_4 and concentrated in vacuo. The crude product was then purified by flash column chromatography on silica gel (hexanes/EtOAc, 96:4) to provide **2-20** (1.35 g, 97%) as

a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.77-5.85 (m, 1H), 5.03 (br d, *J* = 8.8 Hz, 1H), 5.01 (s, 1H), 3.67 (quint., *J* = 5.7 Hz, 1H), 2.16-2.22 (m, 4H), 1.94 (t, *J* = 2.5 Hz, 1H), 1.54 (q, *J* = 7.6 Hz, 1H), 1.26-1.41 (m, 15H), 0.89 (s. 9H), 0.04 (s, 6H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 135.5, 116.5, 84.8, 72.0, 68.0, 41.9, 36.8, 29.7, 29.6, 29.4, 29.1, 28.8, 28.5, 25.9, 25.7(2), 25.3, 18.4, 18.2, 18.1, -2.9.

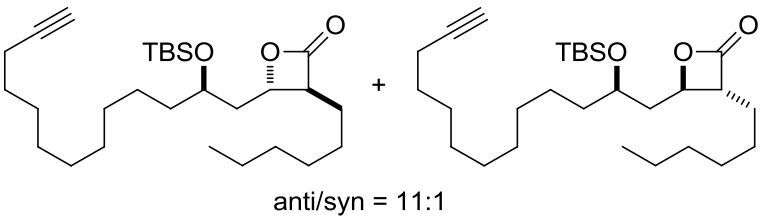
(R)-3-(tert-Butyldimethylsilyloxy)tetradec-13-ynal (2-21)



Ozone was bubbled through a solution of **2-20** (900 mg, 2.67 mmol) and Et_3N (0.77 mL, 5.34 mmol) in DCM (25 mL) cooled at -78 ℃. After reaction was complete (monitored by TLC), Me_2S (332 mg, 5.34 mmol) was added to the reaction mixture and warmed to room temperature. After stirring for 2 h, the solvent was removed under reduced pressure. The crude aldehyde was then purified by column chromatography on silica gel (hexanes/EtOAc, 100:0 to 9:1) to furnish aldehyde **2-21** (720 mg, 80%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.81 (t, *J* = 2.6 Hz, 1H), 4.17 (quint., *J* = 5.7 Hz, 1H), 2.51 (dd, *J* = 2.6, 5.0 Hz, 2H), 2.18 (dt, *J* = 2.5, 6.9 Hz, 2H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.29-1.55 (m, 16H), 0.88 (s, 9H), 0.07 (s, 3H), 0.06 (s, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 202.4, 84.8, 68.3, 68.0, 50.8, 37.8, 29.6, 29.5, 29.4, 29.0, 28.7, 28.5, 25.8, 25.1, 18.4, 18.0, -4.4, -4.7.

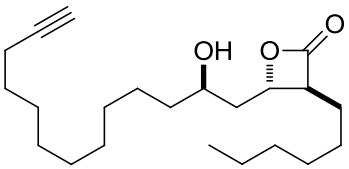
(3S,4S)-3-hexyl-4-((R)-2-hydroxytridec-12-ynyl)oxetan-2-one (2-24)

<u>TMAL reaction</u>:



Prepared according to the general TMAL reaction procedure using aldehyde **2-21** (544 mg, 1.61 mmol), $ZnCl_2$ (438 mg, 3.22 mmol), and ketene acetal **2-23** (740 mg, 2.42 mmol) in DCM (10 mL). The crude mixture was purified by column chromatography on silica gel (hexanes/EtOAc: 100:0 to 97:3) to provide a mixture of silyloxy-β-lactone diastereomers as a colourless oil. ^{1}H NMR (500 MHz) analysis of the mixture of diastereomers indicated that a diastereomeric ratio of 11:1. Without further purification, the mixture was used directly in the next step. ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.41-4.44 (m, 1H), 3.83-3.85 (m, 1H), 3.16-3.24 (m, 1H), 2.18 (dt, *J* = 2.5, 6.9 Hz, 3H), 1.94 (t, *J* = 2.5 Hz, 1H), 1.28-1.85 (m, 27H), 0.90 (s, 9H), 0.89 (t, *J* = 7.0 Hz, 3H), 0.07 (s, 6H).

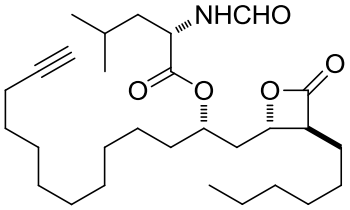
<u>*O*-desilylation</u>:



Prepared according to the general *O*-desilylation procedure using the above silyloxy-β-lactone (409 mg, 0.9 mmol), and 40% HF (780 μL) in CH_3CN (12 mL). The crude mixture was purified by flash chromatography on silica gel (hexanes/EtOAc, 100:0 to 9:1) to provide the hydroxy-β-lactone **2-24** (173 mg, 32%) and a mixture of the two diastereomers (42 mg, 8%) as a colourless oil (40% overall two steps). Spectroscopic data are reported for the major diastereomeric β-lactone **2-24**: ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.49 (dt, *J* = 4.4, 8.8 Hz, 1H), 3.78-3.83 (m, 1H),

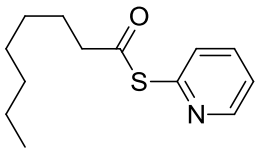
3.24-3.27 (m, 1H), 2.18 (dt, *J* = 2.5, 7.0 Hz, 2H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.72-1.92 (m, 4H), 1.29-1.53 (m, 25H), 0.88 (t, *J* = 7.0 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.6, 84.7, 75.6, 68.5, 68.0, 56.6, 41.8, 38.1, 31.5, 29.4, 29.0, 28.9, 28.7, 28.4, 27.7, 26.7, 25.4, 22.5, 18.4, 14.0; IT-TOF-MS:(m/z) calcd for $C_{22}H_{38}O_3$ $[M+H]^+$ 351.281, found 351.301.

(S)-((S)-1-((2S,3S)-3-hexyl-4-oxooxetan-2-yl)tridec-12-yn-2-yl)-2-formamido-4-methylpentanoate (2-2, THL-L)



Prepared according to the general Mitsunobu reaction procedure using β-lactone **2-24** (11 mg, 0.03 mmol), DIAD (23 μL, 0.12 mmol), *N*-formyl-L-leucine **2-14** (20 mg, 0.13 mmol), and triphenylphosphine (33 mg, 0.13 mmol) and in THF (1.0 mL). Purified by flash chromatography on silica gel (hexanes/EtOAc, 9:1 to 4:1) to give β-lactone **2-2** as a colourless oil (44%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.22 (s, 1H), 5.91 (br d, *J* = 8.2 Hz, 1H), 5.00-5.03 (m, 1H), 4.66-4.71 (m, 1H), 4.27-4.34 (m, 1H), 3.22 (dt, *J* = 4.4, 7.6 Hz, 1H), 2.18 (dt, *J* = 2.5, 6.9 Hz, 2H), 2.00 (dt, *J* = 4.4, 14.5 Hz, 1H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.28-1.85 (m, 30H), 0.96-0.99 (m, 6H), 0.89 (t, *J* = 7.6 Hz, 3H); IT-TOF-MS: (*m/z*) calcd for $C_{29}H_{49}NO_5$ $[M+H]^+$ 492.361, found 492.320.

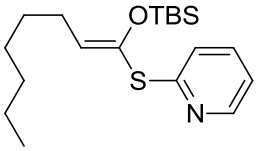
S-pyridin-2-yl octanethioate (2-22)[5]



To a stirred solution of ocanoic acid (1.54 g, 10.7 mmol) in dry DCM (30 mL) was added oxalyl chloride (1.4 mL, 16.0 mmol), and a catalytic amount of anhydrous DMF at 0 ^{o}C. The reaction mixture was stirred

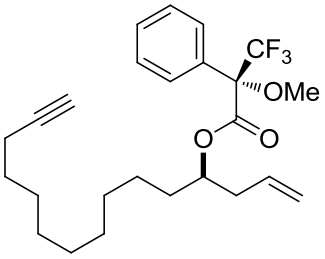
for 2 h at 25 ^{o}C. The solvent was evaporated under reduced pressure. To a stirred solution of the oil residue was added 2-thiopyridine (1.42 g, 12.8 mmol) in DCM (30 mL) followed by Et_3N (3.0 mL, 21.3 mmol) at 0 ^{o}C. The reaction mixture was stirred for 2 h at 25 ^{o}C and then quenched with 1*N* HCl (30 mL) and neutralized by washing with saturated $NaHCO_3$ solution (30 mL × 2). The organic layers were combined and evaporated, then resulting yellow oil was purified by flash chromatography on silica gel (20% EtOAc/hexanes) to afford thiol ester **2-22** (2.49 g, 98%) as a pure liquid.

(E)-2-(1-(tert-butyldimethylsilyloxy)oct-1-enylthio)pyridine (2-23)[6]



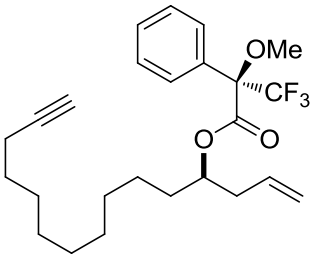
Anhydrous DMF was added to a solution of freshly prepared LiHMDS (5.05 mmol) in THF (10 mL) cooled at -78 ℃ followed by a solution of TBSCl (635 mg, 4.21 mmol) in THF (4 mL). After 10 min, a solution of thiol ester **2-22** (1.0 g, 4.21 mmol) in THF (4 mL) was added to the reaction mixture. The mixture was stirred at -78 ℃ for 20 min, then quenced with saturated NH_4Cl solution (10 mL). The crude was extracted with Et_2O (3 × 20 mL), the combined organic layers was washed with brine (30 mL), dried over anhydrous Na_2SO_4, and concentrated in vacuo. The residue was purified by flash chromatography on silica gel (5% EtOAc/hexanes) to provide thioketene acetal **2-23** (1.12 g, 74%) as a pale yellow oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 8.41 (d, *J* = 3.8 Hz, 1H), 7.54 (dt, *J* = 8.2, 1.7 Hz, 1H), 7.32 (d, *J* = 7.9 Hz, 1H), 6.99 (m, 1H), 5.40 (t, *J* = 7.0 Hz, 1H), 2.18 (t, *J* = 7.0 Hz, 2H), 1.31 (br s, 10H), 0.87 (s, 12H), 0.06 (s, 6H); IT-TOF-MS: (m/z) calcd for $C_{19}H_{33}NOSSi$ $[M+H]^+$ 352.205, found 352.182.

(S)-((R)-pentadec-1-en-14-yn-4-yl)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoate (2-25)



DCC (13 mg, 0.063 mmol) was added to a solution of (*S*)-MTPA (7.5 mg, 0.032 mmol) and DMAP (0.5 mg, 0.0042 mmol) in DCM (1.0 mL). After 5 min, homoallylic alcohol **2-19** (5.0 mg, 0.021 mmol) in DCM (0.5 mL) was added and the reaction mixture was left to stir at room temperature overnight. The reaction mixture was filtered through a pad of celite and the filtrate was concentrated under reduced pressure. The resulting crude product was purified by flash chromatography on silica gel (2% EtOAc/hexanes) to give (*S*)-Mosher ester **2-25** (1.2 mg, 43%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.54 (m, 2H), 7.39 (m, 3H), 5.60-5.68 (m, 1H), 5.14 (m, 1H), 5.03 (br d, *J* = 5.7 Hz, 1H), 5.00 (s, 1H), 3.54 (s, 3H), 2.35 (m, 2H), 2.18 (dt, *J* = 7.0, 2.5 Hz, 2H), 1.94 (t, *J* = 2.5 Hz, 1H), 1.52 (quin, *J* = 7.6 Hz, 2H), 1.26-1.37 (m, 14H); ESI-MS: (*m/z*) calcd for $C_{25}H_{33}F_2O_3$ $[M+Na]^+$ 461.2, found 461.1. The ee of homoallylic alcohol **2-19** was 62.4% as determined by ^{1}H NMR (500 MHz) analysis of its (*S*)-MTPA ester derivatives (4.2:1).

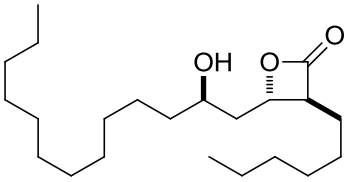
(R)-((R)-pentadec-1-en-14-yn-4-yl)-3,3,3-trifluoro-2-methoxy-2-phenylpropanoate (2-26)



Prepared according to the method of (*S*)-MTPA ester using homoallylic alcohol **2-19** (5.3 mg, 0.023 mmol), (*R*)-MTPA (11 mg, 0.046 mmol), DCC (14 mg, 0.069 mmol), and

DMAP (0.6 mg, 0.0046 mmol). Purification by flash chromatography on silica gel (2% EtOAc/hexanes) afforded (*R*)-MTPA ester **2-26** (6.7 mg, 67%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.54-7.56 (m, 2H), 7.38-7.40 (m, 3H), 5.72-5.80 (m, 1H), 5.15-5.18 (m, 1H), 5.11 (br d, *J* = 7.0 Hz, 1H), 5.09 (br s, 1H), 3.56 (s, 3H), 2.42 (t, *J* = 6.3 Hz, 2H), 2.18 (dt, *J* = 7.0, 2.5 Hz, 2H), 1.94 (t, *J* = 2.5 Hz, 1H), 1.52 (quin, *J* = 7.6 Hz, 2H), 1.20-1.37 (m, 14H); ESI-MS: (*m/z*) calcd for $C_{25}H_{33}F_2O_3$ $[M+Na]^+$ 461.2, found 461.1.

(3S, 4S)-3-hexyl-4-((R)-2-hydroxytridecyl)oxetan-2-one (2-27)

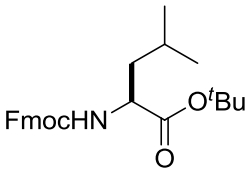


TMAL reaction: Prepared according to the general TMAL reaction procedure using aldehyde **2-8a** (432 mg, 1.27 mmol), $ZnCl_2$ (346 mg, 2.54 mmol) and ketene acetal **2-23** (652 mg, 1.85 mmol). Purification by flash chromatrography on silica gel (hexanes/EtOAc, 100:0 to 95:5) gave a mixture of silyloxy-β-lactone diastereomers as colourless oils, which was used directly in the next *O*-desilylation step.

***O*-desilylation:** Prepared according to the general *O*-desilylation reaction procedure using silyloxy-β-lactone (500 mg, 1.07 mmol) and 40% HF (2.5 mL). Purification by flash chromatrography on silica gel (hexanes/EtOAc, 100:0 to 96:4) gave the hydroxy-β-lactone **2-27** (174 mg, 38%) and a mixture of the two diastereomers (47 mg, 10%) as white solids (48% overall, 2 steps): ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.50 (m, 1H), 3.81 (m, 1H), 3.26 (dt, *J* = 7.6, 3.8 Hz, 1H), 1.95-1.26 (m, 35H), 0.88 (t, *J* = 6.9 Hz, 6H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.6, 75.6, 68.5,

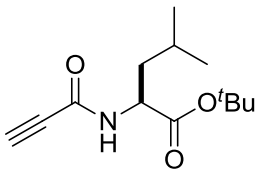
56.6, 41.8, 38.1, 31.9, 31.5, 29.63, 29.61, 29.55, 29.5, 29.3, 29.0, 27.7, 26.8, 25.4, 22.7, 14.1, 14.0; ESI-MS: (*m/z*) calcd for $C_{22}H_{42}O_3$ $[M+Na]^+$ 377.3, found: 377.2.

Fmoc-Leu-O^{t}Bu (2-28)



$POCl_3$ (695 mg, 4.53 mmol) was added to a solution of Fmoc-Leu-OH (1.0 g, 2.83 mmol), *tert*-butanol (2.1 g, 28.3 mmol) and pyridine (1.2 mL, 14.2 mmol) in anhydrous DCM (5 mL) at 0 ℃. The reaction mixture was warmed to room temperture and stirred for 4 h. The reaction mixture was washed with 1*N* HCl (10 mL), saturated $NaHCO_3$ solution (10 mL), brine (10 mL), dried over Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatography on silica gel (25% EtOAc/hexanes) to provide ester **2-28** (931 mg, 80%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.76 (d, *J* = 7.6 Hz, 2H), 7.60 (d, *J* = 7.0 Hz, 2H), 7.40 (t, *J* = 7.6 Hz, 2H), 7.31 (t, *J* = 7.6 Hz, 2H), 5.16 (br d, *J* = 8.2 Hz, 1H), 4,39 (d, *J* = 7.0 Hz, 2H), 4.28 (m, 1H), 4.23 (t, *J* = 7.0 Hz, 1H), 1.72-1.60 (m, 2H), 1.27 (br s, 10H), 0.96 (d, *J* = 6.3 Hz, 6H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.4, 155.9, 144.0, 143.9, 141.3, 127.7, 127.0, 125.1, 120.0, 81.9, 66.9, 53.1, 47.3, 42.1, 28.0, 24.8, 22.8, 22.0.

(S)-tert-butyl 4-methyl-2-propiolamidopentanoate (2-29)

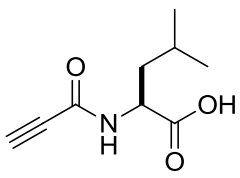


Et_2NH (1.1 mL, 11 mmol) was added to a solution of ester **2-28** (900 mg, 2.2 mmol) in CH_3CN (5 mL). The reaction mixture was stirred at room temperture overnight. The solvent was

evaporated under reduced pressure and the crude product was used without futher purification.

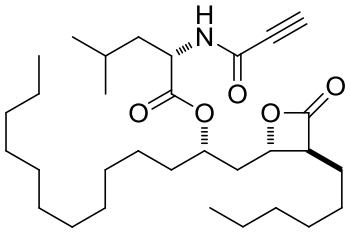
DCC (144 mg, 0.7 mmol) was added to a solution of propiolic acid (49 mg, 0.7 mmol) in DCM (1.0 mL) at 0 ℃ and stirred for 10 min. The unpurified H_2N-Leu-O^{t}Bu (144 mg, 0.7 mmol) in DCM (1.0 mL) was added dropwise to the reaction mixture. The reaction mixture was warmed to room temperture and stirred for 1 h. Then the reaction mixture was filtered through a pad of celite and the filtrate was concentrated under reduced pressure. The crude product was purified by flash chromatography on silica gel (25% EtOAc/hexanes) to provide amide **2-29** (70 mg, 42%) as a colourless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 6.47 (br s, 1H), 4.53 (m, 1H), 2.81 (s, 1H), 1.67-1.46 (m, 9H), 1.45 (s, 9H), 0.93 (d, *J* = 5.3 Hz, 6H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 171.4, 151.5, 82.4, 73.7, 51.6, 41.6, 27.9, 24.8, 22.7, 22.0.

(S)-4-methyl-2-propiolamidopentanoic acid (2-30)

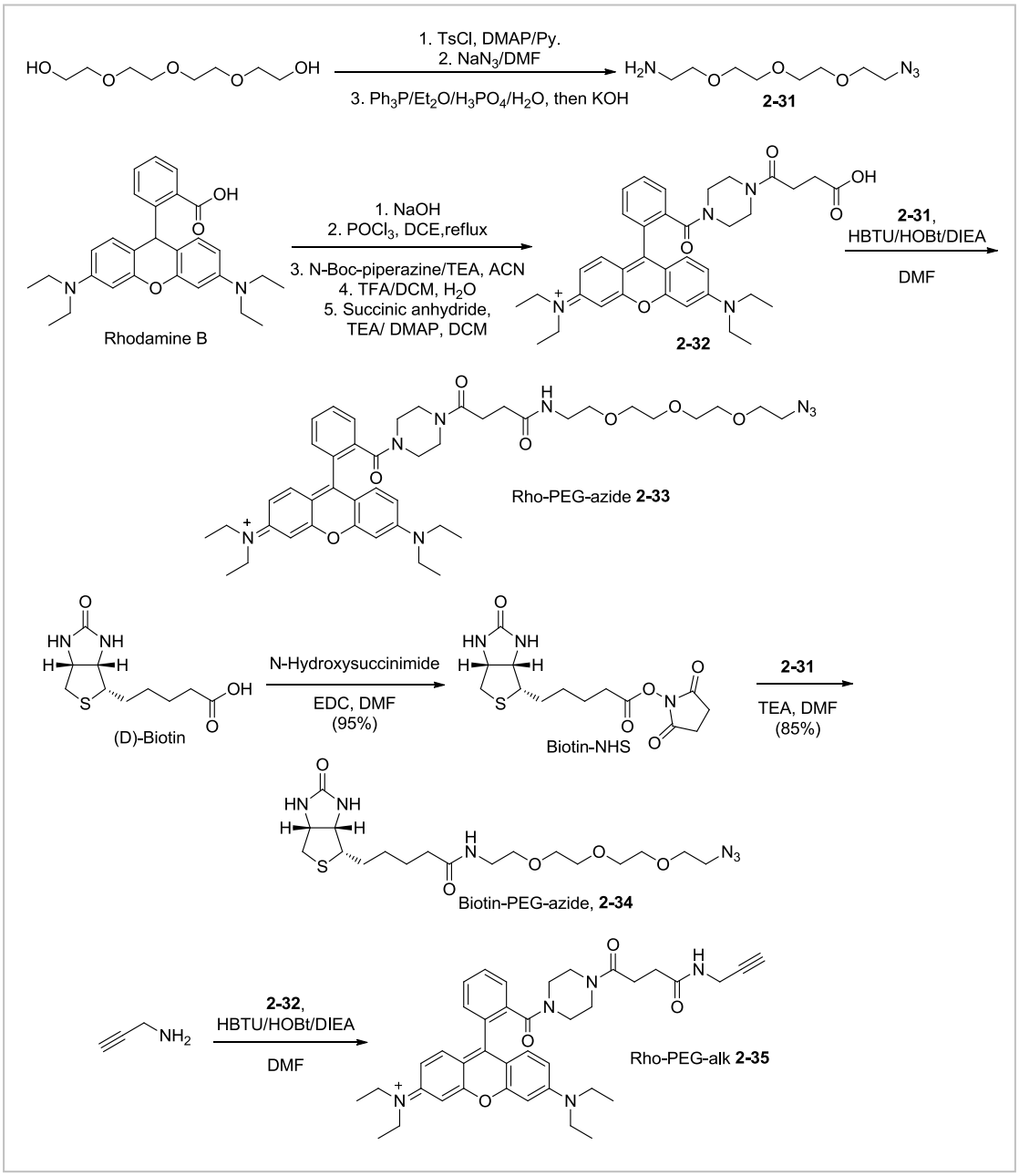


TFA (0.5 mL) was added to a solution of amide **2-29** (70 mg, 0.29 mmol) in DCM (0.5 mL) cooled at 0 ℃. The reaction mixture was warmed to room temperture and left to stir for 5 h. The solvent was evaporated under reduced pressure and the crude product was used without further purification. ^{1}H NMR (500 MHz, $CDCl_3$): δ 6.26 (br d, *J* = 7.6 Hz, 1H), 4.70 (m, 1H), 3.96 (br s, 1H), 2.87 (s, 1H), 1.74 (m, 2H), 1.63 (m, 1H), 0.97 (d, *J* = 5.1 Hz, 6H).

(R)-((S)-1-((2S,3S)-3-hexyl-4-oxooxetan-2-yl)tridecan-2-yl)-4-methyl-2-propiolamid opentanoate (2-3,THL-T)



Prepared according to the general Mitsunobu reaction procedure using β-lactone **2-27** (71 mg, 0.20 mmol), PPh_3 (68 mg, 0.26 mmol), acid **2-30** (46 mg, 0.25 mmol) and DIAD (52 μL, 0.26 mmol) in THF (3.0 mL). Purification by flash chromatrography on silica gel (hexanes/EtOAc, 100:0 to 85:15) afforded **2-3** (54 mg, 52%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 6.31 (br d, *J* = 7.6 Hz, 1H), 5.01-5.04 (m, 1H), 4.60 (dt, *J* = 4.4, 8.8 Hz, 1H), 4.26-4.30 (m, 1H), 3.21 (dt, *J* = 3.8, 7.6 Hz, 1H), 2.84 (s, 1H), 2.15 (dt, *J* = 7.6, 15.1 Hz, 1H), 1.99 (dt, *J* = 4.5, 15.7 Hz, 1H), 1.54-1.80 (m, 8H), 1.25 (br s, 25H), 0.96 (d, *J* = 6.3 Hz, 3H), 0.95 (d, *J* = 6.3 Hz, 3H), 0.874 (t, *J* = 7.0 Hz, 3H), 0.867 (t, *J* = 7.0 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.4, 170.7, 151.7, 83.3, 74.7, 74.1, 72.9, 57.0, 51.4, 41.2, 38.6, 34.0, 31.9, 31.4, 29.6, 29.5, 29.4, 29.3, 29.2, 28.9, 27.6, 26.7, 25.1, 24.9, 22.8, 22.6, 22.5, 21.7 14.1, 14.0; IT-TOF-MS: (*m/z*) calcd for $C_{31}H_{53}NO_5$ $[M+H]^+$ 520.392, found: 520.391.



N-(9-(2-(4-(1-azido-13-oxo-3,6,9-trioxa-12-azahexadecane)piperazine-1-carbonyl)phenyl)-6-(diethylamino)-3H-xanthen-3-ylidene)-N-ethylethanaminium (2-33, rho-N_3)

In a previously flame-dried under an argon atmosphere round-bottom flask equipped with a magnetic stir bar, was dissolved rhodamine B acid **2-32**[161] (31 mg, 0.05 mmol), HBTU (38 mg, 0.1 mmol), HOBt (14 mg, 0.1 mmol) in dry DMF (0.5

mL). DIEA (30 μL, 0.2 mmol) was added in one portion, and the reaction mixture was stirred at room temperature for 10 min. Amino azide linker **2-31**[162] (27 mg, 0.2 mmol) in DMF (0.5 mL) was then added in one portion and the reaction mixture was stirred at room temperature overnight. The solvent was evaporated under reduced pressure and the crude mixture was poured into 5% HCl (5 mL, saturated with brine), and extracted with DCM (3 × 5 mL). The combined organic phase was washed with saturated $NaHCO_3$ solution (2 × 5 mL), 5% HCl (2 × 5 mL, saturated with brine), and brine (10 mL), then dried over anhydrous Na_2SO_4, and concentrated to yield rho-N_3 **2-33** as a pure compound. IT-TOF-MS: (*m/z*) calcd for $C_{44}H_{59}N_8O_7$ 811.450, found 811.447.

***N*-(2-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)ethyl)-5-((3aR,4R,6aS)-2-oxohexahydro-1*H*-thieno[3,4-d]imidazol-4-yl)pentanamide (2-34, biotin-N_3)**

Triethylamine (0.1 mL, 0.71 mmol) was added to a solution of amino azide linker **2-31** (200 mg, 0.92 mmol) in dry DMF (5 mL). After the solution was stirred for 15 min, *N*-hydroxsuccinimidobiotin (biotin-NHS, 375 mg, 1.10 mmol) was added in one portion. The reaction mixture was stirred for 12 h at room temperature and then concentrated in vacuo to give a gel-like residue, which was purified by flash chromatography on silica gel (acetone/hexane-4:1) to afford 345 mg of biotin-N_3 **2-34** as a white solid (85%). IT-TOF-MS: (*m/z*) calcd for $C_{18}H_{32}N_6O_5S$ $[M+H]^+$ 445.216, found 445.222.

***N*-(6-(diethylamino)-9-(2-(4-(4-oxo-4-(prop-2-yn-1-ylamino)butanoyl)piperazine-1-**

carbonyl)phenyl)-3H-xanthen-3-ylidene)-N-ethylethanaminium (2-35, rho-alk)

Prepared according to the above similar procedure using rhodamine B acid **2-32** (31 mg, 0.05 mmol), propargylamine (15 μL, 0.2 mmol), HBTU (38 mg, 0.1 mmol), HOBt (14 mg, 0.1 mmol) and DIEA (30 μL, 0.2 mmol) in dry DMF (0.5 mL). IT-TOF-MS: (*m/z*) calcd for $C_{39}H_{46}N_5O_4$ 648.354, found 648.342.

8.2.2 Cell Biological Assays

Cell-culture conditions, cell-proliferation assay, in situ proteomic profiling, pull-down assay, and fluorescence microscopy were carried out using the general protocols as described in **8.1.2.**

8.2.2.1 Chemicals and Antibodies

Orlistat (98%), Tris(2-carboxyethyl) phosphine (TCEP), and the click chemistry ligand, tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) were purchased from Sigma-Aldrich. Antibody against FAS (Cat No. 610963) was from BD Transduction Labs (San Diego, CA). Antibodies against eIF2α (#9722), phospho-eIF2α (#9721), and cleaved caspase-8 (#9746) were from Cell Signaling Technologies (Beverly, MA). Antibodies against HSP90b (sc-1057), Annexin A2 (sc-1492) and RPL14 (KQ-16; sc-100826) were from Santa Cruz Biotechnology, Inc.. Antibodies against RPL7a (ab70753) and RPS9 (ab74711) were from Abcam. Hoechst33342 and ER-Tracker Green (glibenclamide BODIPY FL) were from Invitrogen.

8.2.2.2 Western Blotting

To monitor the effects of THL analogues on inducing phosphorylation of eIF2α, PC-3 cells were treated with indicated concentrations of Orlistat and THL analogues for 16 h. Samples from treated cells were then separated on 12% SDS-PAGE gel and further transferred to PVDF membranes. Membranes were blocked with 5% BSA in TBS. After blocking, membranes were incubated with anti-eIF2α (#9722 from cell signaling, 1/5000) or anti-phospho-eIF2α (#9721 from Cell Signaling, 1/2000). After incubation, membranes were washed with TBST for three times and then incubated with an appropriate secondary antibody [anti-mouse conjugated HRP (1/5000) or anti-rabbit conjugated HRP (1/5000)]. After secondary incubation, blots were washed again with TBST before the development with SuperSignal West Pico kit (Pierce).

To monitor the effects of THL analogues on inducing activation of caspase-8 pathway, MCF-7 cells were treated with indicated concentrations of Orlistat and THL analogues for 36 h. Samples from treated cells were then separated on a 12% SDS-PAGE gel and further transferred to PVDF membranes. Membranes were blocked with 5% BSA in TBS. After blocking, membranes were incubated with anti-caspase 8 (#9746 from Cell Signaling, 1/2000). After incubation, membranes were washed with TBST for three times, and then incubated with anti-mouse conjugated HRP (1/5000). After secondary incubation, blots were washed again with TBST before the development with SuperSignal West Pico kit (Pierce).

8.2.2.3 Measurement of Protein Synthesis

Live HepG2 cells were treated with the indicated concentrations of Orlistat, THL-R or CHX (cycloheximide, an inhibitor of protein biosynthesis) for 12 h, washed twice with PBS, and then pulsed with AHA (L-azidohomoalanine, 20 μM) for 4 h. Cells were collected, washed, and cell lysates were prepared and subjected to click chemistry with rhodamine-alkyne **2-35**, SDS-PAGE analysis, and in-gel fluorescence scanning.

8.2.2.4 Hydroxylamine Treatment of Gels

After the proteins were separated by SDS-PAGE gel, the gel was soaked in 40% MeOH, 10% acetic acid, shaking overnight at room temperature, washed with deionized water (2 × 5 min), and scanned for the prehydroxylamine treatment fluorescence. The gel was then soaked in PBS, shaking 1 h at room temperature, followed by boiling in neutralized hydroxylamine (Alfa Aesar) (2.5% final concentration) for 5 min, washing with deionized water (2 × 5 min), and soaking in 40% MeOH, 10% acetic acid, shaking overnight at room temperature. The gel was washed with deionized water (2 × 5 min) and scanned for the posthydroxylamine fluorescence.

8.2.2.5 Mass spectrometry Identification

After digestion, digested peptides were then extracted from the gel with 50%

acetonitrile and 1% formic acid. Tryptic peptide extracts were evaporated by speedvac and reconstituted with 10 μL 0.1% TFA, a volume of 2 μL of the peptide extracts were manually spotted onto a Prespotted AnchorChip MALDI target plate for MALDI-TOF Mass step with 10 μL of 10 mM ammonium phosphate in 0.1% TFA, and allowed to dry at ambient temperature. MALDI TOF mass spectra were recorded using Ultraflex III TOF/TOF mass spectrometer (Bruker Daltonics) with the compass 1.2 software package including flexControl 3.0 and flexAnalysis 3.0, calibrated with PAC peptide calibration standards. MS/MS analysis for the major peaks in PMF spectra were carried out by autoLIFT on the MALDI-TOF/TOF instrument. MS and MS/MS Peak lists with intensity value were submitted to Matrix Science Mascot server (http://www.matrixscience.com/) through *BioTools 3.0* (Bruker Daltonics) using database NCBInr 090202 version, variable modifications of carbamidometyl on cysteine (C) and oxidation on methione (M), allowing maxium of trypsin missed cleavage, peptide mass tolerance at 200 ppm; MS/MS mass tolerance of 0.7 Da.

8.2.2.6 Target Validation by Western Blotting

Pull-down sample from labeled lysates was separated on 12% SDS-PAGE gel together with pull-down sample from DMSO-treated, unlabeled lysates (negative controls). After SDS-PAGE gel separation, proteins were then transferred to a PVDF membrane and subsequently blocked with 2.5% (w/v) BSA/PBST. Membranes were incubated for 1 h at room temperature with the respective antibodies (i.e., anti-FAS, anti-RPL7a, anti-Annexin A2, anti-HSP90b, anti-RPS9, anti-GAPDH, and

anti-β-tubulin). After three times of washing with PBST, blots were further incubated with the appropriate secondary antibody for 1 h at room temperature. After incubation, the blot was washed again with PBST for 3 times and the SuperSignal West Pico kit (Pierce) was used to develop the blot.

8.2.2.7 Target Validation by Recombinant Protein Expression in HEK-293T Cells

Mammalian expression vectors overexpressing each of the four targets were purchased from Origene (USA). Vectors were transfected with Lipofectamine reagent (Qiagen) at 80% confluence. After 24 h of transfection, either nontransfected and transfected cell were incubated with THL-R. After incubation, small fractions of the samples were analyzed by Western blotting with 1/2000 anti-c-Myc antibody (Santa Cruz). The rest of the cells were lysed and immunopurified with c-Myc agarose beads (Santa Cruz). Eluted fractions were clicked with rho-azide **2-33** as previously described. After click chemistry, samples were concentrated, separated on a SDS-PAGE gel and scanned by in-gel fluorescence.

8.2.2.8 Identification of GAPDH Labeling Site

The mammalian expression vector overexpressing GAPDH was used as template for generating the GAPDH active-site mutant, using the Site-Direceted Mutagenesis System provided by Invitrogen (USA). Cys^{151} was mutated to Ala. Primers used for site-directed mutagenesis were designed following the vendor's protocols. They are shown below:

5’-ATCAGCAATGCCTCCGCCACCACCAACTGC-3’

5’-GCAGTTGGTGGTGGCGGAGGCATTGCTGAT-3’

Vectors overexpressing both wild-type and mutant GAPDH were transfected to HEK-293T cell lines. Recombinant protein was purified, labeled with THL-R and clicked with rho-azide **2-33** as previously described. Further, samples were separated on SDS-PAGE gel and fluorescently scanned with Typhoon Scanner. After scanning, gel was fixed and silver stained to visualize the total protein bands.

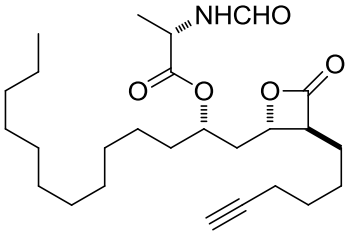
8.2.2.9 Cellular Imaging

For colocalizing in situ targets of THL analogues with FAS, cells were further incubated with anti-FASN primary antibodies (1:200) for 1 h at room temperature (or overnight at 4 ℃), and washed twice with PBS. The cells were incubated with FITC-conjugated antimouse IgG (1:500) for 1 h, washed again. Cells were stained with 1 μg/mL Hoechst for 10 min at room temperature, washed again before mounting. For colocalizing in situ targets of THL-R with ER, cells were further incubated with ER-Tracker Green (glibenclamide BODIPY FL) for 1 h at room temperature (or overnight at 4 ℃), and washed twice with PBS. Cells were stained with 1 μg/mL Hoechst for 10 min at room temperature, washed again before mounting.

8.3 Chapter 3

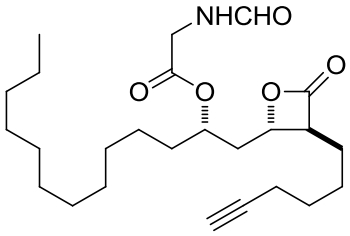
8.3.1 Chemical Syntheis

(S)-(S)-1-[(2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 2-formamidopropanoate (3-1b)



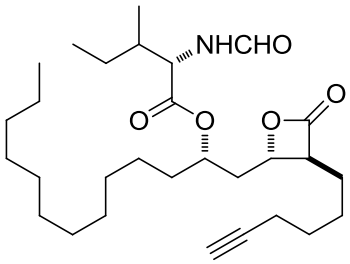
Prepared according to the representative Mitsunobu reaction procedure using β-lactone **2-13** (18 mg, 0.05 mmol), triphenylphosphine (20 mg, 0.075 mmol), PPh_3 (79 mg, 0.3 mmol), *N*-formyl-L-alanine (23 mg, 0.2 mmol) and DIAD (15 μL, 0.075 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (20% EtOAc/hexanes) gave **3-1b** (20 mg, 89%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.19 (s, 1H), 6.16 (br s, 1H), 5.07 (m, 1H), 4.64 (quin., *J* = 7.6 Hz, 1H), 4.32 (dt, *J* = 8.9, 4.4 Hz, 1H), 3.23 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.72 (dt, *J* = 6.3, 2.5 Hz, 2H), 7.16 (dt, *J* = 14.5, 4.4 Hz, 1H), 2.01 (dt, *J* = 14.5, 4.4 Hz, 1H), 1.95 (t, *J* = 2.5 Hz, 1H), 1.54-1.62 (m, 8H), 1.45 (d, *J* = 7.0 Hz, 3H), 1.25 (br s, 18H), 0.88 (t, *J* = 7.6 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.1, 170.4, 160.4, 83.8, 74.8, 72.9, 68.8, 60.4, 56.9, 47.1, 38.9, 34.2, 31.9, 29.6, 29.5, 29.4, 29.31. 29.27, 27.9, 27.1, 25.7, 25.1, 22.7, 21.0, 18.3, 18.1, 14.2, 14.1; ESI-MS: (*m/z*) calcd for $C_{26}H_{43}NO_5$ $[M+H]^+$ 450.3, found 450.3.

(S)-1-[(2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 2-formamidoacetate (3-1c)



Prepared according to the representative Mitsunobu reaction procedure using β-lactone **2-13** (26 mg, 0.075 mmol), triphenylphosphine (30 mg, 0.113 mmol), *N*-formyl-glycine (31 mg, 0.3 mmol), and DIAD (23 μL, 0.113 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (20% EtOAc/hexanes) gave **3-1c** (24 mg, 73%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.24 (s, 1H), 6.23 (br s, 1H), 4.34 (dt, *J* = 8.8, 4.4 Hz, 1H), 4.10 (dd, *J* = 18.3, 5.7 Hz, 1H), 4.02 (dd, *J* = 18.3, 5.7 Hz, 1H), 3.22 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.20 (dt, *J* = 6.3, 2.5 Hz, 2H), 2.18 (t, *J* = 8.2 Hz, 1H), 2.01 (dt, *J* = 14.1, 4.4 Hz, 1H), 1.95 (t, *J* = 2.5 Hz, 1H), 1.73-1.85 (m, 4H), 1.52-1.59 (m, 8H), 1.24 (br s, 18H), 0.86 (t, *J* = 6.9 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 170.5, 169.2, 161.0, 83.8, 74.9, 72.9, 68.8, 56.8, 40.1, 38.8, 34.0, 31.8, 29.54, 29.48, 29.37, 29.27, 29.24, 27.9, 27.0, 25.7, 25.1, 22.6, 18.0, 14.0; ESI-MS: (*m/z*) calcd for $C_{25}H_{41}NO_5$ $[M + H]^+$ 436.3, found 436.3.

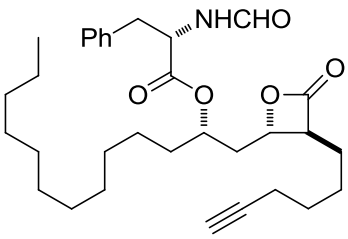
(2S)-(S)-1-[(2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 2-formamido-3-methylpentanoate (3-1d)



Prepared according to the representative Mitsunobu reaction procedure using β-lactone **2-13** (28 mg, 0.08 mmol), triphenylphosphine (31 mg, 0.12 mmol), *N*-formyl-L-isoleucine (51 mg, 0.32 mmol), and DIAD (24 μL, 0.12 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (20% EtOAc/hexanes) gave **3-1d** (21 mg, 54%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.23 (s, 1H), 6.10 (br d, *J* = 8.2 Hz, 1H), 5.02 (m, 1H), 4.64 (dd, *J* = 8.9,

5.1 Hz, 1H), 4.29 (dt, *J* = 8.9, 5.1 Hz, 1H), 3.24 (dt, *J* = 7.6, 4.5 Hz, 1H), 2.21 (dt, *J* = 6.3, 2.5 Hz, 2H), 2.18 (t, *J* = 7.6 Hz, 1H), 2.01 (dt, *J* = 10.1, 5.0 Hz, 1H), 1.95 (t, *J* = 2.5 Hz, 1H), 1.55-1.85 (m, 12H), 1.26 (br s, 17H), 0.96 (d, *J* = 6.9 Hz, 3H), 0.94 (t, *J* = 7.6 Hz, 3H), 0.87 (t, *J* = 7.6 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 170.9, 170.5, 160.7, 83.7, 74.4, 72.8, 68.7, 56.9, 55.4, 38.6, 37.6, 33.9, 31.9, 29.6, 29.5, 29.4, 29.29, 29.28, 27.9, 27.1, 25.7, 25.1, 24.9, 22.6, 18.1, 15.6, 14.1, 11.5; ESI-MS: (*m/z*) calcd for $C_{29}H_{46}NO_5$ $[M+H]^+$ 492.4, found 492.4.

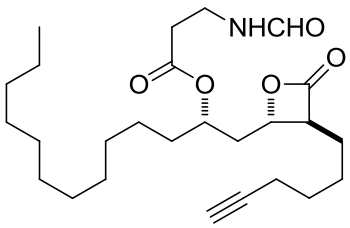
(S)-(S)-1-[(2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 2-formamido-3-phenylpropanoate (3-1e)



Prepared according to the representative Mitsunobu reaction procedure using β-lactone **2-13** (18 mg, 0.05 mmol), triphenylphosphine (20 mg, 0.075 mmol), *N*-formyl-L-phenylalanine (19 mg, 0.1 mmol), and DIAD (15 μL, 0.075 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (20% EtOAc/hexane) gave **3-1e** (12 mg, 46%) as white solids. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.17 (s, 1H), 7.26-7.32 (m, 3H), 7.17-7.19 (m, 2H), 6.10 (br d, *J* = 6.9 Hz, 1H), 4.99 (m, 1H), 4.90 (dd, *J* = 13.9, 7.3 Hz, 1H), 4.16-4.24 (m, 1H), 3.08-3.20 (m, 3H), 2.21 (dt, *J* = 6.3, 2.5 Hz, 2H), 2.18 (t, *J* = 7.6 Hz, 1H), 2.01 (dt, *J* = 10.1, 5.0 Hz, 1H), 1.95 (t, *J* = 2.5 Hz, 1H), 1.55-1.85 (m, 6H), 1.26 (br s, 20H), 0.88 (t, *J* = 6.95 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 170.8, 170.5, 160.5, 139.6, 129.3, 129.2, 128.7, 128.6, 127.3 (2), 83.9, 74.5, 72.9, 68.7, 56.8, 52.1, 38.5, 37.8, 33.9, 31.9, 29.6, 29.5, 29.4, 29.3 (2), 27.9, 27.1, 25.7, 25.1, 24.6, 22.6, 18.1, 14.1; ESI-MS: (*m/z*) calcd for

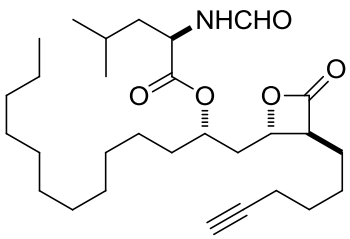
$C_{32}H_{47}NO_5$ $[M+H]^+$ 525.3, found 525.4.

(S)-1-[(2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 3-formamidopropanoate (3-1f)



Prepared according to the representative Mitsunobu reaction procedure using β-lactone **2-13** (35 mg, 0.1 mmol), triphenylphosphine (40 mg, 0.15 mmol), *N*-formyl-β-glycine (23 mg, 0.2 mmol), and DIAD (30 μL, 0.15 mmol) in 2.0 mL of THF. Purification by flash chromatography on silica gel (20% EtOAc/hexane) gave **3-1f** (30 mg, 67%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.15 (s, 1H), 6.24 (br s, 1H), 5.09 (dt, *J* = 6.3, 3.2 Hz, 1H), 4.36 (dt, *J* = 8.8, 4.4 Hz, 1H), 3.57 (quin., *J* = 5.6 Hz, 1H), 3.22 (dt, *J* = 6.9, 3.8 Hz, 1H), 2.55 (dd, *J* = 6.3, 5.0 Hz, 2H), 2.21 (dt, *J* = 6.9, 2.5 Hz, 2H), 2.10 (dt, *J* = 15.1, 8.2 Hz, 1H), 2.00 (dt, *J* = 15.1, 3.8 Hz, 1H), 1.96 (t, *J* = 2.5 Hz, 1H), 1.53-1.84 (m, 10H), 1.25 (br s, 18H), 0.87 (t, *J* = 7.0 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.0, 170.4, 161.2, 83.8, 76.0, 72.0, 68.8, 56.9, 39.0, 34.3, 33.4, 31.9, 29.6, 29.5, 29.4, 29.29, 29.26, 27.9, 27.0, 25.7, 25.2, 22.6, 18.0, 14.1; ESI-MS: (*m/z*) calcd for $C_{26}H_{43}NO_5$ $[M+H]^+$ 450.3, found 450.3.

(R)-(S)-1-[(2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 2-formamido-4-methylpentanoate (3-1g)

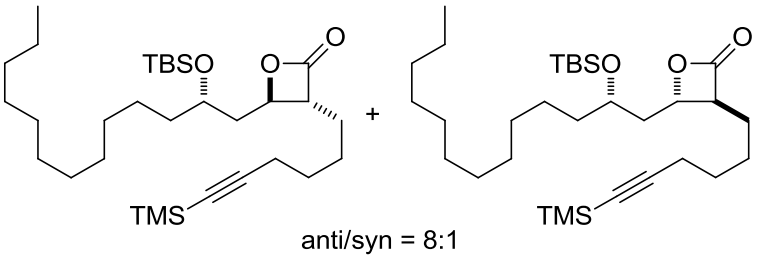


Prepared according to the representative Mitsunobu reaction procedure using β-lactone **2-13** (18 mg, 0.05 mmol), triphenylphosphine (20 mg, 0.075 mmol),

N-formyl-D-leucine (16 mg, 0.1 mmol), and DIAD (15 μL, 0.075 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (20% EtOAc/hexane) gave **3-1g** (18 mg, 73%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.20 (s, 1H), 5.93 (br d, *J* = 8.2 Hz, 1H), 5.02 (m, 1H), 4.63 (dt, *J* = 8.8, 3.0 Hz, 1H), 4.37 (m, 1H), 3.24 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.20 (m, 2H), 2.16 (t, *J* = 7.4 Hz, 1H), 2.01 (dt, *J* = 14.1, 4.4 Hz, 1H), 1.92 (t, *J* = 2.5 Hz, 1H), 1.50-1.84 (m, 10H), 1.25 (br s, 17H), 0.96 (d, *J* = 5.7 Hz, 6H), 0.87 (t, *J* = 7.6 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.2, 170.7, 160.7, 83.9, 74.4, 72.5, 68.7, 60.4, 56.8, 49.7, 41.3, 38.5, 36.6, 33.8, 31.9, 29.6, 29.5, 29.4, 29.3, 29.2, 28.0, 27.1, 25.7, 25.2, 24.9, 24.7, 22.8, 22.6, 21.8, 18.1, 14.2, 14.1; ESI-MS: (m/z) calcd for $C_{29}H_{46}NO_5$ $[M+H]^+$ 492.4, found 492.4.

(3R,4R)-4-((S)-2-hydroxytridecyl)-3-(6-(trimethylsilyl)hex-5-yn-1-yl)oxetan-2-one (3-4)

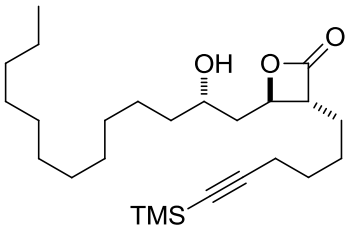
TMAL reaction:



Prepared according to the representative TMAL reaction procedure using $ZnCl_2$ (359 mg, 2.64 mmol), aldehyde **2-(*S*)-8b** (450 mg, 1.32 mmol) and ketene acetal **2-11** (831 mg, 1.98 mmol) in 5.0 mL of DCM. Purification by flash chromatography on silica gel (5% EtOAc/hexanes) to provide the mixture of silyloxy-β-lactone diastereomers as a colorless oil. ^{1}H NMR (500 MHz) analysis of the mixture of diastereomers indicated a diastereomeric ratio of 8:1. Without further purification, the mixture was used directly

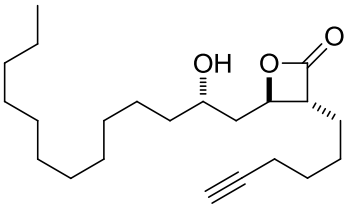
in the next step. ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.41 (m, 1H), 3.83 (m, 1H), 3.18 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.24 (m, 2H), 1.82-1.84 (m, 2H), 1.26 (br s, 26H), 0.90 (m, 12H), 0.15 (s, 9H), 0.07 (s, 6H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.4, 106.7, 84.9, 75.2, 68.6, 56.0, 41.9, 37.9, 31.9, 29.7, 29.63, 29.61, 29.57, 29.54, 29.3, 28.2, 27.3, 26.0, 25.9, 25.8, 24.6, 22.7, 19.6, 18.0, 14.1, 0.14, -4.3, -4.7.

***O*-<u>desilylation</u>:**



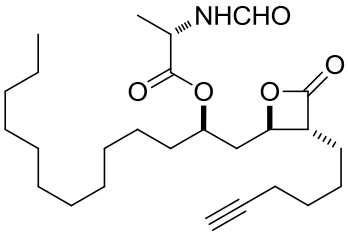
Prepared according to the representative *O*-desilylation reaction procedure using the above mixture of silyloxy-β-lactones (360 mg, 0.67 mmol) and 0.6 mL of HF (48%) in 2 mL of CH_3CN. Purification by flash chromatography on silica gel (5% EtOAc/hexanes) to provide the hydroxy-β-lactone **3-4** (167 mg, 30%) and a mixture of the two diastereomers (67 mg, 12%) as white solids (42% overall two steps). Spectroscopic data are reported for the major diastereomeric-β-lactone **3-4**. ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.50 (dt, *J* = 8.2, 4.4 Hz, 1H), 3.80 (m, 1H), 3.27 (dt, *J* = 8.2, 3.8 Hz, 1H), 2.13 (m, 2H), 1.93 (ddd, *J* = 13.9, 8.8, 3.2, 1H), 1.85 (dd, *J* = 9.5, 4.5 Hz, 1H), 1.82-1.76 (m, 2H), 1.55-1.38 (m, 7H), 1.25 (br s, 17H), 0.87 (t, *J* = 6.3 Hz, 3H), 0.14 (s, 9H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.4, 106.8, 85.0, 75.6, 68.5, 56.4, 41.8, 38.2, 31.9, 29.63, 29.61, 29.55, 29.48, 29.3, 28.1, 27.2, 25.9, 25.4, 25.4, 22.7, 19.6, 14.1.

(3R,4R)-3-(hex-5-yn-1-yl)-4-((S)-2-hydroxytridecyl)oxetan-2-one (3-5)



<u>*C*-desilylation</u>: Prepared according to the representative *C*-desilylation reaction procedure using **3-4** (80 mg, 0.19 mmol) and $AgNO_3$ (6.5 mg, 0.04 mmol in 1.0 mL of a mixture of acetone/H_2O/2,6-lutidine (3:1:0.l). Purification by chromatography on silica gel (10% EtOAc/hexanes) gave **3-5** (55 mg, 85%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.50 (dt, *J* = 8.8, 4.4 Hz, 1H), 3.78-3.84 (m, 1H), 3.29 (dt, *J* = 8.2, 3.8 Hz, 1H), 2.21 (dt, *J* = 6.3, 2.5 Hz, 2H), 1.95 (t, *J* = 2.5 Hz, 1H), 1.94 (ddd, *J* = 13.9, 8.2, 2.5 Hz, 1H), 1.78-1.87 (m, 3H), 1.48-1.62 (m, 8H), 1.26 (br s, 16H), 0.88 (t, *J* = 7.6 Hz, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): 171.3, 83.9, 75.6, 68.7, 68.5, 56.4, 41.8, 38.1, 31.9, 29.63, 29.6, 29.55, 29.49, 29.3, 28.0, 27.2, 25.8, 25.4, 22.7, 18.1, 14.1; ESI-MS: (*m/z*) calcd for $C_{22}H_{38}O_3$ $[M+Na]^+$ 373.3, found 373.1.

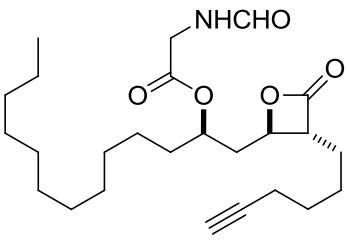
(S)-(R)-1-((2R,3R)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl)tridecan-2-yl 2-formamidopropanoate (3-1h)



Prepared according to the representative Mitsunobu reaction procedure using β-lactone **3-5** (14 mg, 0.04 mmol), triphenylphosphine (21 mg, 0.08 mmol), *N*-formyl-L-alanine (10 mg, 0.08 mmol), and DIAD (12 μL, 0.08 mmol) in 1.0 mL of THF. Purification by chromatography on silica gel (20% EtOAc/hexanes) yielded **3-1h** (7 mg, 36%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.18 (s, 1H), 6.17 (br s, 1H), 5.04-5.09 (m, 1H), 4.62 (quin, *J* = 7.6 Hz, 1H), 4.34-4.37 (m, 1H), 3.24 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.21 (dt, *J* = 6.3, 2.5 Hz, 2H), 2.19-2.14 (m, 1H), 2.03 (dt, *J* = 13.9, 5.0 Hz, 1H) 1.96 (t, *J* = 2.5 Hz, 1H), 1.84-1.54

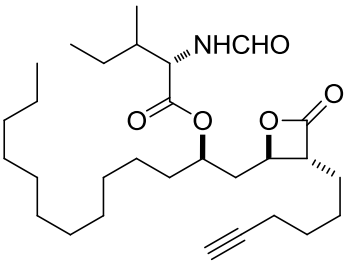
(m, 10 H), 1.45 (d, *J* = 6.9 Hz, 3H), 1.27 (br s, 17H), 0.88 (t, *J* = 7.6 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.2, 170.5, 160.4, 83.9, 74.6, 72.7, 68.7, 56.8, 47.1, 38.7, 33.9, 31.9, 29.6, 29.5, 29.4, 29.3, 29.2, 27.9, 27.1, 25.7, 25.2, 22.7, 18.2, 18.1, 14.1; ESI-MS: (*m/z*) calcd for $C_{26}H_{43}NO_5$ $[M+Na]^+$ 472.3, found 472.2.

(R)-1-[(2R,3R)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 2-formamidoacetate (3-1i)



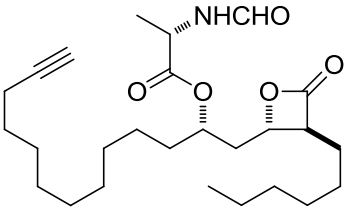
Prepared according to the representative Mitsunobu reaction procedure using β-lactone **3-5** (14 mg, 0.04 mmol), triphenylphosphine (21 mg, 0.08 mmol), *N*-formyl-L-glycine (10 mg, 0.1 mmol) and DIAD (12 μL, 0.08 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (15% EtOAc/hexanes) gave **3-1i** (8 mg, 46%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.26 (s, 1H), 6.11 (br s, 1H), 5.10-5.15 (m, 1H), 4.33-4.36 (m, 1H), 4.12 (dd, *J* = 19.0, 5.7 Hz, 1H), 4.04 (dd, *J* = 19.0, 5.0 Hz, 1H), 3.23 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.22 (dt, *J* = 7.0, 2.5 Hz, 2H), 2.1 (dt, *J* = 15.1, 7.6 Hz, 1H), 2.03 (dt, *J* = 14.5, 4.4, 1H), 1.96 (t, *J* = 2.5 Hz, 1H), 1.86-1.73 (m, 9H), 1.26 (br s, 14H), 0.88 (t, *J* = 7.6 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 170.5, 169.2, 160.9, 83.8, 74.9, 73.0, 68.8, 56.9, 40.1, 38.9, 34.1, 31.9, 29.6, 29.5, 29.4, 29.32, 29.26, 27.9, 25.7, 25.2 22.7, 18.1, 14.1; ESI-MS: (*m/z*) calcd for $C_{25}H_{41}NO_5$ $[M+H]^+$ 436.3, found 436.2.

(2S)-(R)-1-[(2R,3R)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 2-formamido-3-methylpentanoate (3-1j)



Prepared according to the representative Mitsunobu reaction procedure using β-lactone **3-5** (14 mg, 0.04 mmol), triphenylphosphine (21 mg, 0.08 mmol), *N*-formyl-L-isoleucine (14 mg, 0.08 mmol) and DIAD (12 μL, 0.08 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (15% EtOAc/hexanes) gave **3-1j** (8 mg, 41%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.23 (s, 1H), 6.16 (br s, 1H), 5.02-5.03 (m, 1H), 4.61 (dd, *J* = 8.2, 4.4 Hz, 1H), 4.34-4.37 (m, 1H), 3.23 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.20-2.15 (m, 3H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.60-1.25 (m, 30 H), 0.96 (d, *J* = 7.0 Hz, 2H), 0.93 (d, *J* = 7.6 Hz, 3H), 0.87 (t, *J* = 6.9 Hz, 3H); IT-TOF-MS: (*m/z*) calcd for $C_{29}H_{49}NO_5$ $[M+H]^+$ calcd 492.361, found 492.288.

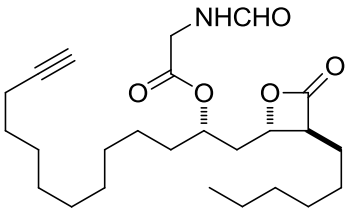
(S)-(S)-1-[(2S,3S)-3-hexyl-4-oxooxetan-2-yl]tridec-12-yn-2-yl 2-formamidopropanoate (3-2b)



Prepared according to the general Mitsunobu reaction procedure using β-lactone **2-24** (13 mg, 0.04 mmol), PPh_3 (47 mg, 0.18 mmol), *N*-formyl-L-alanine (21 mg, 0.18 mmol) and DEAD (27 μL, 0.17 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (hexanes/EtOAc, 100:0 to 4:1) to give **3-2b** (13 mg, 73%) as white solids. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.18 (s, 1H), 6.21 (br d, *J* = 6.3 Hz, 1H), 5.05-5.07 (m, 1H), 4.61-4.67 (m, 1H), 4.30 (m, 1H), 3.21 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.15 (dt, *J* = 7.0, 2.5 Hz, 3H), 2.00 (dt, *J* = 15.1, 4.5 Hz, 1H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.80-1.50 (m, 8H), 1.45 (d, *J* = 6.9 Hz, 3H), 1.39-1.27

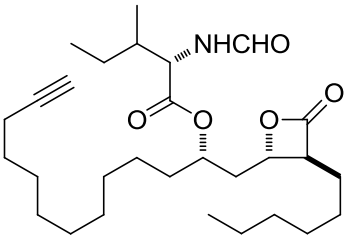
(m, 18H), 0.88 (t, *J* = 7.0 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.1, 170.7, 160.4, 84.7, 74.9, 73.0, 68.1, 57.1, 47.1, 38.9, 34.2, 31.4, 29.3, 29.2, 29.16, 29.0, 28.9, 28.7, 28.4, 27.6, 26.7, 25.1, 22.5, 18.3, 14.4, 14.0; TOF-MS: (*m/z*) calcd for $C_{26}H_{43}NO_5$ $[M+Na]^+$ 472.304, found 472.267.

(S)-1-[(2S,3S)-3-hexyl-4-oxooxetan-2-yl]tridec-12-yn-2-yl 2-formamidoacetate (3-2c)



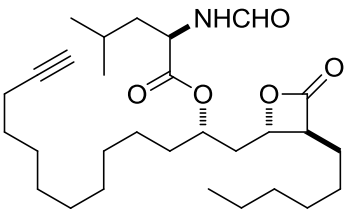
Prepared according to the general Mitsunobu reaction procedure using β-lactone **2-24** (12 mg, 0.03 mmol), PPh_3 (45 mg, 0.17 mmol), *N*-formyl-L-glycine (18 mg, 0.17 mmol) and DEAD (25 μL, 0.16 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (hexanes/EtOAc, 100:0 to 4:1) to give **3-2c** (15 mg, 70%) as a colourless oil: ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.04 (s, 1H), 6.21 (br s, 1H), 5.09-5.14 (m, 1H), 4.31-4.34 (m, 1H), 4.11 (dd, *J* = 18.3, 5.7 Hz, 1H), 4.02 (dd, *J* = 18.3, 5.1 Hz, 1H), 3.18-3.22 (m, 1H), 2.17 (dt, *J* = 6.9, 2.5 Hz, 2H), 2.10-2.15 (m, 1H), 2.01 (dt, *J* = 15.2, 4.2 Hz, 1H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.81-1.71 (m, 5H), 1.51 (quin., *J* = 7.0 Hz, 2H), 1.41-1.27 (m, 16H), 0.88 (t, *J* = 6.3 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 170.8, 169.2, 161.0, 84.7, 75.0, 73.0, 68.1, 57.0, 40.1, 38.9, 34.1, 31.4, 29.3, 29.2, 29.0, 28.9, 28.7, 28.4, 27.6, 26.7, 25.1, 22.5, 18.4, 14.0; TOF-MS: (*m/z*) calcd for $C_{25}H_{41}NO_5$ $[M+Na]^+$ 458.288, found 458.301.

(2S)-(S)-1-[(2S,3S)-3-hexyl-4-oxooxetan-2-yl]tridec-12-yn-2-yl 2-formamido-3-methylpentanoate (3-2d)



Prepared according to the general Mitsunobu reaction procedure using β-lactone **2-24** (12 mg, 0.03 mmol), PPh_3 (45 mg, 0.17 mmol), *N*-formyl-L-isoleucine (27 mg, 0.17 mmol) and DEAD (25 μL, 0.16 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (hexanes/EtOAc, 100:0 to 4:1) to give **3-2d** (11 mg, 63%) as white solids. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.25 (s, 1H), 6.05 (br d, *J* = 8.8 Hz, 1H), 5.00-5.05 (m, 1H), 4.66 (dd, *J* = 8.8, 5.1 Hz, 1H), 4.28 (dt, *J* = 7.6, 4.4 Hz, 1H), 3.23 (dt, *J* = 8.2, 4.4 Hz, 1H), 2.15 (dt, *J* = 14.5, 5.1 Hz, 3H), 2.01 (dt, *J* = 14.5, 5.1 Hz, 1H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.85-1.25 (m, 26H), 0.96 (d, *J* = 7.0 Hz, 3H), 0.95 (d, *J* = 7.0 Hz, 3H), 0.88 (t, *J* = 7.6H, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 170.9, 170.7, 160.7, 84.8, 74.6, 72.9, 68.1, 57.1, 55.4, 38.6, 37.6, 33.9, 31.5, 29.3, 29.2, 29.0, 28.9, 28.7, 28.4, 27.7, 26.7, 25.1, 24.9, 22.5, 18.4, 15.6, 14.0, 11.5; TOF-MS: (*m/z*) calcd for $C_{29}H_{49}NO_5$ $[M+H]^+$ 492.361, found 492.379.

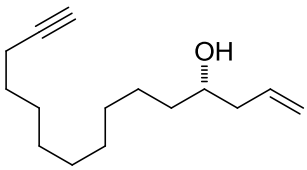
(R)-(S)-1-[(2S,3S)-3-hexyl-4-oxooxetan-2-yl]tridec-12-yn-2-yl 2-formamido-4-methylpentanoate (3-2e)



Prepared according to the general Mitsunobu reaction procedure using β-lactone **2-24** (15 mg, 0.04 mmol), PPh_3 (56 mg, 0.22 mmol), *N*-formyl-D-leucine (34 mg, 0.22 mmol) and DEAD (41 μL, 0.21 mmol) in 1.0 mL of THF. Purification by flash chromatography on silica gel (hexanes/EtOAc, 100:0 to 9:1) followed by second purification by flash chromatography on silica gel

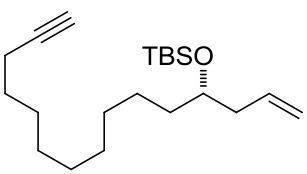
(hexanes/EtOAc, 9:1 to 4:1) to give **3-2e** (8 mg, 36%) as white solids. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.20 (s, 1H), 5.91 (br d, *J* = 5.3 Hz, 1H), 5.00-5.05 (m, 1H), 4.64-4.69 (m, 1H), 4.34-4.37 (m, 1H), 3.22 (dt, *J* =6.9, 4.4 Hz, 1H), 2.18 (dt, *J* = 6.9, 2.5 Hz, 1H), 2.02 (dt, *J* = 14.5, 4.5 Hz, 1H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.86-1.27 (m, 29H), 0.97 (d, *J* = 5.0 Hz, 6H), 0.88 (t, *J* = 7.0 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.2, 170.9, 160.6, 84.7, 74.6, 72.6, 68.1, 57.0, 49.6, 41.5, 38.5, 33.8, 31.5, 29.34, 29.32, 29.2, 29.0, 28.9, 28.7, 27.6, 26.7, 25.6, 25.2, 24.9, 22.8, 22.5, 21.9, 18.4, 14.0; TOF-MS: (*m/z*) calcd for $C_{29}H_{49}NO_5$ $[M+H]^+$ 492.361, found 492.344.

(S)-pentadec-1-en-14-yn-4-ol (3-6)



Prepared according to the general asymmetric allylation reaction procedure using aldehyde **2-18** (300 mg, 1.66 mmol), (*R*)-BINOL (106 mg, 0.37 mmol), $InCl_3$ [73 mg, 0.33 mmol, azeotropically dried over THF (2 ×3 mL)], allyltributyltin (1.10 g, 3.32 mmol) to provide (*S*)-homoallylic alcohol **3-6** (252 mg, 53%, 50% *ee*, determined using Mosher ester analysis) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.83 (m, 1H), 5.15 (br d, *J* = 3.2 Hz, 1H), 5.12 (d, *J* = 1.3 Hz, 1H), 3.64 (br s, 1H), 2.29 (m, 1H), 2.18 (dt, *J* = 7.0, 2.5 Hz, 2H), 2.19-2.11 (m, 3H), 1.54-1.29 (m, 17H).

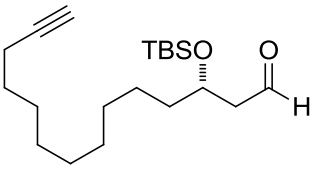
(S)-tert-butyldimethyl(pentadec-1-en-14-yn-4-yloxy)silane (3-7)



Prepared according to the general TBS-protection reaction procedure using homoallylic alcohol **3-6** (234 mg, 1.0 mmol),

imidazole (202 mg, 2.97 mmol), TBSCl (268 mg, 1.78 mmol) and DMAP (24 mg, 0.2 mmol) to provide (*S*)-silyl ether **3-7** (321 mg, 95%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.81 (m, 1H), 5.03 (dd, *J* = 9.5, 1.9 Hz, 1H), 5.01 (s, 1H), 3.68 (quin., *J* = 5.7 Hz, 1H), 2.19 (m, 5H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.54-1.28 (m, 17H), 0.89 (s, 9H), 0.05 (s, 6H).

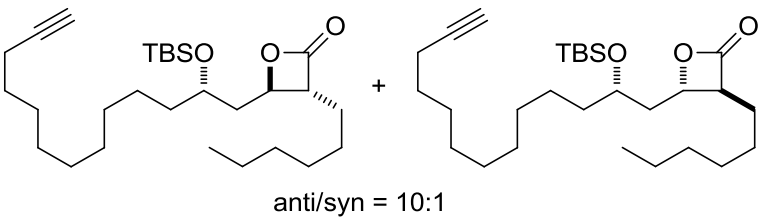
(S)-3-((tert-butyldimethylsilyl)oxy)tetradec-13-ynal (3-8)



Prepared according to the general ozonolysis reaction procedure using silyl ether **3-7** (320 mg, 0.95 mmol), and Et_3N (0.3 mL, 1.9 mmol) to provide aldehyde **3-8** (205 mg, 64%) as a colorless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 9.81 (s, 1H), 4.17 (quin, *J* = 5.7 Hz, 1H), 2.51 (dd, *J* = 6.3, 2.5 Hz, 2H), 2.18 (dt, *J* = 6.9, 2.5 Hz, 2H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.52-1.28 (m, 16H), 0.87 (s, 9H), 0.07 (s, 3H), 0.06 (s, 3H).

(3R,4R)-3-hexyl-4-((S)-2-hydroxytridec-12-yn-1-yl)oxetan-2-one (3-9)

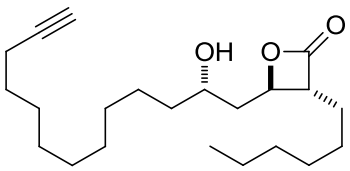
TMAL reaction:



Prepared according to the general TMAL reaction procedure using aldehyde **3-8** (200 mg, 0.59 mmol, $ZnCl_2$ (161 mg, 1.18 mmol) and ketene acetal **2-23** (272 mg, 0.89 mmol) in DCM (4.0 mL). Purification by flash chromatography on silica gel (hexanes/EtOAc, 100:0 to 95:5) gave the mixture of silyloxy-β-lactone

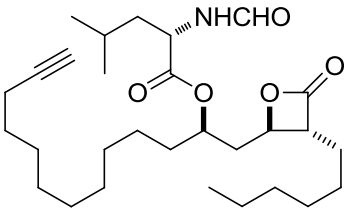
diastereomers as a colorless oil. Without further purification, the mixture was used directly in the next step. ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.42 (m, 1H), 3.83 (m, 1H), 3.18 (m, 1H), 2.18 (dt, *J* = 6.9, 2.5 Hz, 2H), 1.94 (t, *J* = 2.5 Hz, 1H), 1.87-1.72 (m, 3H), 1.62-1.28 (m, 25H), 0.90 (s, 9H), 0.89 (t, *J* = 6.9 Hz, 3H), 0.07 (s, 6H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.7, 84.8, 75.2, 68.6, 68.0, 56.1, 42.0, 37.9, 31.5, 29.7, 29.5, 29.4, 29.1, 28.9, 28.7, 28.5, 27.7, 26.7, 25.9, 25.8, 24.6, 22.5, 18.3, 18.0, 14.0, -4.3, -4.8.

<u>*O*-desilylation</u>:



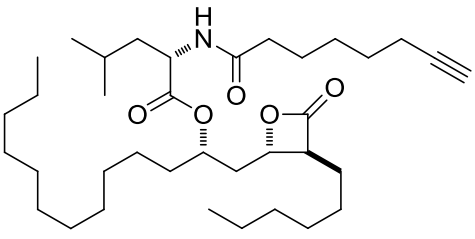
Prepared according to the general *O*-desilylation reaction procedure using silyloxy-β-lactone (190 mg, 0.42 mmol) and 40% HF (460 μL). Purification by flash chromatography on silica gel (hexanes/EtOAc, 100:0 to 96:4) gave the hydroxy-β-lactone **3-9** (68 mg, 37%) and a mixture of the two diastereomers (18 mg, 10%) as a colorless oil (47% overall, 2 steps). ^{1}H NMR (500 MHz, $CDCl_3$): δ 4.50 (m, 1H), 3.81 (m, 1H), 3.26 (dt, *J* = 8.1, 4.1 Hz, 1H), 2.18 (dt, *J* = 7.4, 2.6 Hz, 2H), 1.94 (t, *J* = 2.6 Hz, 1H), 1.92-1.74 (m, 4H), 1.55-1.29 (m, 24H), 0.88 (t, *J* = 6.6 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.6, 84.8, 68.5, 68.1, 56.6, 41.8, 38.1, 31.5, 29.4, 29.36, 29.0, 28.9, 28.7, 28.5, 27.7, 26.8, 25.4, 22.5, 18.4, 14.0; ESI-MS: (*m/z*) calcd for $C_{22}H_{38}O_3$ $[M+H]^+$ 351.3, found 351.2.

(S)-(R)-1-[(2R,3R)-3-hexyl-4-oxooxetan-2-yl]tridec-12-yn-2-yl 2-formamido-4-methylpentanoate (3-2f)



Prepared according to the general Mitsunobu reaction procedure using β-lactone **3-9** (15 mg, 0.04 mmol), PPh_3 (43 mg, 0.16 mmol), *N*-formyl-L-leucine (27 mg, 0.17 mmol) and DEAD (25 μL, 0.16 mmol) in THF (1.0 mL). Purification by flash chromatography on silica gel (hexanes/EtOAc, 100:0 to 4:1) to give **3-2f** (13 mg, 81%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$): δ 8.20 (s, 1H), 5.95 (br s, 1H), 5.02-5.04 (m, 1H), 4.66-4.69 (m, 1H), 4.26-4.31 (m, 1H), 3.21 (m, 1H), 2.17 (dt, *J* = 6.9, 2.5 Hz, 2H), 1.98-2.02 (m, 1H), 1.93 (t, *J* = 2.5 Hz, 1H) 1.65-1.82 (m, 6H), 1.48-1.58 (m, 4H), 1.28 (br s, 18H), 0.96 (d, *J* = 6.3 Hz, 6H), 0.88 (t, *J* = 6.3 Hz, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.2, 170.9, 160.6, 84.7, 74.5, 72.5, 68.0, 57.0, 47.6, 41.5, 38.7, 33.8, 31.4, 29.3, 29.2, 29.0, 28.9, 28.7, 28.4, 27.6, 26.7, 25.2, 24.9, 22.8, 22.5, 21.8, 18.3, 14.0; ESI-MS: (*m/z*) calcd for $C_{29}H_{49}NO_5$ $[M+Na]^+$ 514.4, found 514.3.

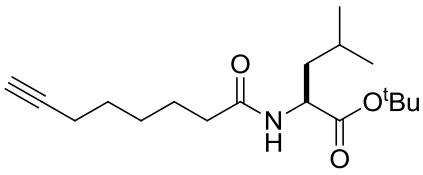
(S)-(S)-1-[(2S,3S)-3-hexyl-4-oxooxetan-2-yl]tridecan-2-yl 4-methyl-2-(oct-7-ynamido)pentanoate (3-3b)



Prepared according to the general Mitsunobu reaction procedure using β-lactone **2-17** (100 mg, 0.28 mmol), PPh_3 (110 mg, 0.42 mmol), acid **3-11** (106 mg, 0.42 mmol) and DIAD (77 μL, 0.39 mmol) in THF (5.0 mL). Purification by flash chromatrography on silica gel (hexanes/EtOAc, 100:0 to 80:20) to give **3-3b** (49 mg, 30%) as white solids. ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.77 (d, *J* = 8.2 Hz, 1H), 5.00 (m, 1H), 4.59 (dt, *J* = 8.8, 5.1 Hz,

1H), 4.27-4.30 (m, 1H), 3.21 (dt, *J* = 7.6, 3.8 Hz, 1H), 2.23 (t, *J* = 7.6 Hz, 2H), 2.19 (dt, *J* = 8.0, 2.5 Hz, 2H), 2.14 (t, *J* = 7.6 Hz, 2H), 1.98 (dt, *J* = 15.2, 4.4 Hz, 1H), 1.93 (t, *J* = 2.5 Hz, 1H), 1.71-1.82 (m, 2H), 1.42-1.67 (m, 14H), 1.26 (br s, 22H), 0.96 (d, *J* = 5.1 Hz, 3H), 0.95 (d, *J* = 5.1 Hz, 3H), 0.88 (t, *J* = 7.0 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.6, 172.5, 170.7, 84.4, 74.8, 72.4, 68.3, 57.1, 51.0, 41.7, 38.7, 36.3, 34.1, 31.9, 31.5, 29.61, 29.60, 29.5, 29.4, 29.3, 29.0, 28.3, 28.1, 27.7, 26.7, 25.1, 25.01, 25.0, 22.9, 22.7, 22.5, 21.9, 18.2, 14.1, 14.0; ESI-MS: (*m/z*) calcd for $C_{36}H_{63}NO_5$ $[M+H]^+$ 590.5; found 590.4.

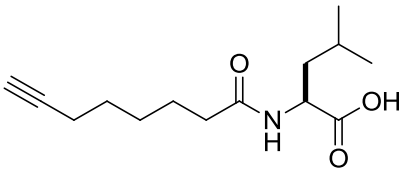
(S)-tert-butyl 4-methyl-2-(oct-7-ynamido)pentanoate (3-10)



To a solution of oct-7-ynoic acid (175 mg, 1.25 mmol) in DCM (7 cm) was added EDC•HCl (288 mg, 1.5 mmol), HOBt (203 mg, 1.5 mmol) and DIEA (0.30 mL, 2 mmol). After 10 min, NH_2-Leu-O*t*Bu[1] (281 mg, 1.5 mmol) in DCM (2 mL) was added dropwise. The reaction was stirred at rt for 21 h. The organic phase was extracted with an 1 *N* HCl solution (10 mL), saturated $NaHCO_3$ (10 mL), brine (10 mL) and dried over anhydrous Na_2SO_4. The residue was purified by flash chromatography on silica gel (hexanes/EtOAc, 4:1) to give amide **3-10** (258 mg, 67%) as a colorless oil. ^{1}H NMR (300 MHz, $CDCl_3$): δ 5.81 (d, *J* = 8.0 Hz, 1H), 4.53 (dt, *J* = 8.6, 5.4 Hz, 1H), 2.22 (t, *J* = 7.1 Hz, 2H), 2.19 (dt, *J* = 7.1, 2.6 Hz, 2H), 1.93 (t, *J* = 2.6 Hz, 1H), 1.49-1.71 (m, 9H), 1.46 (s, 9H), 0.95 (d, *J* = 6.4 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 172.5, 172.4, 84.4, 81.9, 68.3, 51.2, 42.1, 36.5, 28.3, 28.1, 28.0, 25.04,

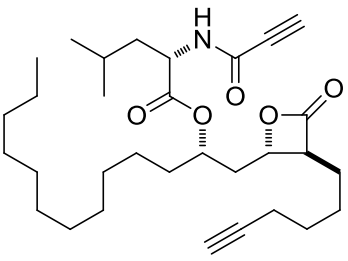
24.98, 22.8, 22.2, 18.3.

***(S)-4-methyl-2-(oct-7-ynamido)pentanoic acid (3-11*)**



TFA (2 mL) was added to a solution of **3-10** (233 mg, 0.75 mmol) in DCM (5 mL) cooled at 0 ℃. The reaction mixture was warmed to room temperture and left to stir for 5 h. The solvent was evaporated under reduced pressure and the crude product was used without further purification.

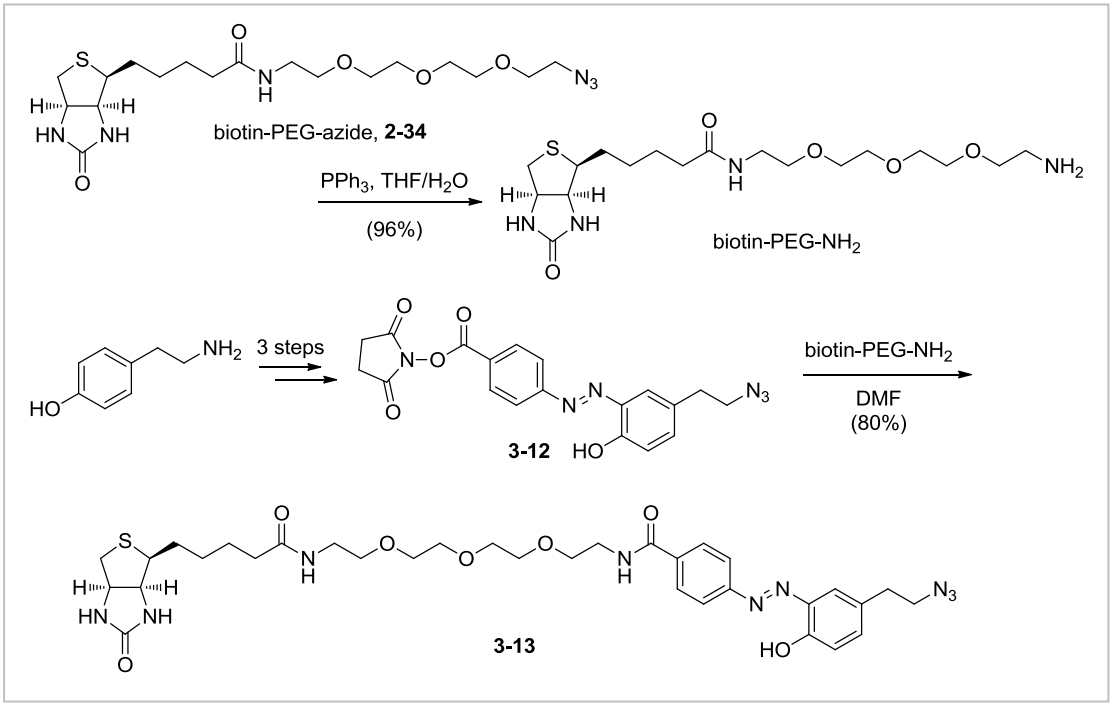
(S)-(S)-1-[(2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl]tridecan-2-yl 4-methyl-2-propiolamidopentanoate (3-3c)



Prepared according to the general Mitsunobu reaction procedure using β-lactone **2-27** (30 mg, 0.086 mmol), PPh_3 (34 mg, 0.13 mmol), acid **3-11** (24 mg, 0.13 mmol) and DIAD (26 μL, 0.13 mmol) in THF (1.0 mL). Purification by flash chromatrography on silica gel (10 % EtOAc/hexane) to give **3-3c** (19 mg, 66%) as white solids. ^{1}H NMR (500 MHz, $CDCl_3$): δ 6.25 (d, *J* = 8.2 Hz, 1H), 5.00-5.05 (m, 1H), 4.58-4.62 (m, 1H), 4.28-4.31 (m, 1H), 3.23 (dt, *J* = 3.8, 7.6 Hz, 1H), 2.85 (s, 1H), 2.22 (dt, *J* = 2.5, 6.3 Hz, 2H), 2.15 (t, *J* = 8.2 Hz, 1H), 2.01 (dt, *J* = 4.4, 15.2 Hz, 1H), 1.95 (t, *J* = 2.5 Hz, 1H), 1.73-1.85 (m, 2H), 1.63-1.68 (m, 3H), 1.54-1.68 (m, 8H), 1.25 (br s, 16H), 0.97 (d, *J* = 6.3 Hz, 3H), 0.96 (d, *J* = 6.3 Hz, 3H), 0.88 (t, *J* = 7.0 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 171.4, 170.4, 151.7, 83.9, 76.7, 74.6, 74.1, 72.8, 68.7, 56.9, 51.4, 41.2, 38.7, 34.0, 31.9, 29.6, 29.5, 29.4, 29.3

(2), 28.0, 27.1, 25.7, 25.1, 24.9, 22.8, 22.7, 21.8, 18.1, 14.1. ESI-MS: (*m/z*) calcd for $C_{31}H_{49}NO_5$ $[M+H]^+$ 516.4; found 516.1.

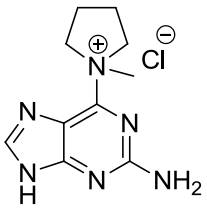
4-((E)-(5-(2-azidoethyl)-2-hydroxyphenyl)diazenyl)-N-(13-oxo-17-((3aS,4S,6aR)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)-3,6,9-trioxa-12-azaheptadecyl)benzamide (3-13)



To a solution of biotin-azide **2-34** (222 mg, 0.5 mmol) in THF (7 mL) and H_2O (2 mL) was added triphenyl phosphine (157 mg, 0.6 mmol) and stirred at rt for 12 h. The solvent was evaporated, and the residue was washed with ether (3 × 20 mL), DCM (2 × 10 mL) and dried in vacuo to yield biotin-PEG-NH_2 as a pale white solid (188 mg, 90%). Without further purification, the mixture was used directly in the next step. To a solution of compound **3-12**[64] (42 mg, 0.1 mmol) in anhydrous DMF (3 mL) was added biotin-PEG-NH_2 (42 mg, 0.1 mmol) in anhydrous DMF (2 mL). The reaction was complete after stirring at room temperature for 12 h and concentrated. The

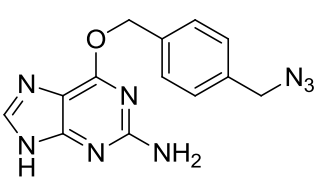
crude product was purified by flash chromatrography on silica gel (10% MeOH/DCM) to give **3-13** as yellow-colored solids (62 mg, 87%). IT-TOF-MS: (*m/z*) calcd for $C_{33}H_{45}N_9O_7S$ $[M+H]^+$ 712.316, found 712.329.

1-(2-amino-9H-purin-6-yl)-1-methylpyrrolidin-1-ium chloride (3-14)



To a solution of 6-chloro-guanine (1.0 g, 5.9 mmol) in DMSO/DMF (5 mL/40 mL) was added 1-methyl-pyrrolidin (1.4 mL, 13.2 mmol). After 18 h, 2 mL of acetone and hexane (1:1) were added to complete the precipitation. The solid was filtered, washed with ether and dried in vacu*o*, to give **3-14** (0.9 g, 66%). Without further purification, the mixture was used directly in the next step. ^{1}H NMR (300 MHz, DMSO-d_6): δ 13.46 (s, 1H), 8.34 (s, 1H), 7.10 (s, 2H), 4.60 (m, 2H), 4.01 (m, 2H), 3.65 (s, 3H), 2.24 (m, 2H), 2.06 (m, 2H); IT-TOF-MS: (*m/z*) calcd for $C_{10}H_{15}ClN_6$ $[M-Cl]^+$ 219.14, found 219.14.

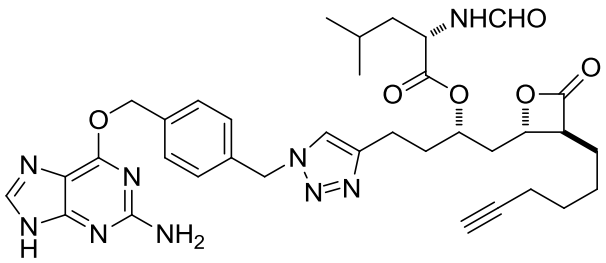
6-((4-(azidomethyl)benzyl)oxy)-9H-purin-2-amine (3-16)



To a solution of (4-(azidomethyl)phenyl)methanol **3-15** (333 mg, 2.04 mmol) in dry DMF (10 mL) was added NaH (60% dispersed in oil) (256 mg, 6.40 mmol) at 0 ℃. Then 570 mg (2.24 mmol) of **3-14** and 20 mg (0.16 mmol) of 4-dimethylaminopyridine were added subsequently. The resulting reaction mixture was stirred at rt for 4 h. The reaction was quenched by addition of water (1 mL). The solvent was removed under reduced pressure. The resulting residue was purified by chromatography with

dichloromethane:methanol (30:1) to give **3-16** (350 mg, 52.7%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 12.35 (bs, 1H), 7.81 (s, 1H), 7.53 (d, J = 8.1 Hz 2H), 7.39 (d, J = 7.8 Hz, 2H), 6.08 (s, 1H), 5.53 (s, 2H), 4.46 (s, 2H); 13C NMR (300 MHz, DMSO-d_6) δ 159.9, 137.2, 135.7, 128.9, 128.7, 66.8, 53.8; IT-TOF-MS: (*m/z*) calcd for $C_{13}H_{12}N_8O$ $[M+H]^+$ 297.1212, found 297.1121.

(S)-(S)-4-(1-(4-((2-amino-9H-purin-6-yl)methoxy)benzyl)-1H-1,2,3-triazol-4-yl)-1-((2S,3S)-3-(hex-5-yn-1-yl)-4-oxooxetan-2-yl)butan-2-yl 2-formamido-4-methylpentanoate (3-4)



To a mixture of di-alkyne β-lactone **3-17**[57b] (10.2 mg, 0.022 mmol) and BG-azide **3-16** (6.5 mg, 0.022 mmol) in a mixture of DMSO/H_2O (3/1, 400 μL) was added CuI (4.2 mg, 0.022 mmol) and DIPEA (8 μL, 0.044 mmol). The reaction mixture was stirred at room temperature for 72 h. The reaction mixture was then diluted with water (1 mL) and extracted with EtOAc (3 × 1 mL). The combined organic phase was extracted with 1M EDTA solution (3 × 1 mL), brine (2 mL) and dried over anhydrous Na_2SO_4. The product was used for next step without purification. LC-ESI-MS: (*m/z*) calcd for $C_{38}H_{52}N_9O_6Si$ $[M+H]^+$ 758.4, found 758.4.

To a solution of the above mixture in DMF (200 μL) was added a solution of tris(dimethylamino)sulfonium difluorotrimethylsilicate (TAS-F) (8.0 mg, 0.029 mmol) in DMF (200 μL). The reaction mixture was stirred at room temperature for 12 h. The

reaction mixture was diluted with water (500 μL) and the aqueous phase was extracted with EtOAc (3 × 1 mL). The combined organic phase was washed with brine (1 mL) and dried over anhydrous $NaSO_4$. The residue was purified by preparative HPLC (Phenomenex Luna C18, 5 μM, 50 × 3 mm; 0-10 min, 20-95% CH_3CN; 10-12 min, 100% CH_3CN; flow rate: 0.5 mL/min) to give **3-4** (2.6 mg, 17% yield over two steps) as a colourless oil. ^1H NMR (500 MHz, $CDCl_3$): δ 7.99 (s, 1H), 7.63 (s, 1H), 7.50 (d, *J* = 7.6 Hz, 2H), 7.41 (s, 1H), 7.24 (d, *J* = 7.6 Hz, 2H), 5.62 (d, *J* = 14.5 Hz, 1H), 5.52-5.55 (m, 1H), 5.43 (d, *J* = 12 Hz, 1H), 5.38 (d, *J* = 14.5 Hz, 1H), 4.94 (br s, 1H), 4.85 (br s, 1H), 4.49 (br s, 1H), 4.27 (dt, *J* = 8.2, 4.4 Hz, 1H), 3.20 (dt, *J* = 8.3, 3.8 Hz, 1H), 2.79 (t, *J* = 6.3 Hz, 1H), 2.17-2.20 (m, 2H), 2.09-2.15 (m, 2H), 2.00-2.07 (m, 2H), 1.95 (t, *J* = 2.5 Hz, 1H), 1.51-1.77 (m, 5H), 0.97 (d, *J* = 6.3 Hz, 3H), 0.94 (d, *J* = 6.3 Hz, 3H). IT-TOF-MS: (*m/z*) calcd for $C_{35}H_{43}N_9O_6$ $[M+H]^+$ 686.3415, found 686.3416.

8.3.2 Cell Biological Assays

Cell-culture conditions, cell-proliferation assay, in situ proteomic profiling, pull-down and mass spectrometry identification, and fluorescence microscopy were carried out using the general protocols as described in **7.1.2.**

8.3.2.1 Chemicals and Antibodies

Monoclonal ANTI-FLAG M2 antibody was purchased from Sigma. The bacterial His-AGT plasmid, mammalian plasmids FLAG-Cox8A-SNAP,

FLAG-H2B-SNAP, FLAG-KDEL-SNAP, FLAG-NK1R-SNAP, mCherry-Cox8A and mCherry-KDEL were generous gifts from Christopher J. Chang (University of California, Berkeley). AGT-deficient CHO-9 cell line was generous gift from Institute of Toxicology, University of Mainz (Germany). Recombinant expression and purification of His-AGT protein was as previously reported.[62d]

8.3.2.2 *In Vitro* His-AGT Labeling with (3-4) and Analysis by TOF-MS

His-AGT (2 μM) was incubated in PBS with 1 μM of **3-4** for different lengths of time (15 min, 2 h and 4 h) at room temperature, then incubated with BGFL (100 μM) for 15 min followed by SDS-PAGE and in-gel fluorescence scanning. The deconvoluted MS data were collected by ESI-TOF-MS.

8.3.2.3 Fluorescence Microscopy

AGT-deficient CHO-9 cells were seeded on coverslips and left to adhere for 24 h prior to transfection. Transient transfection was carried out by the standard protocol of Endofectin (Genecopoeia). Cells were incubated for 48 h prior to labeling. Then cells were incubated in growth medium containing **3-4** (30 μM) at culture temperature and 5% CO_2 for 4 h. The cells were then washed twice with PBS, and fixed with 4% paraformaldehyde in PBS for 15 min at room temperature and washed twice with PBS, and permeabilized with 0.1% Triton-X 100 in PBS for 15 min at room temperature, and washed twice with PBS. The cells were blocked with 2% bovine serum albumin (BSA) in PBS for 30 min at room temperature, and washed twice with PBS. The cells were then treated with a freshly pre-mixed click chemistry reaction solution in PBS for 1 h at room temperature. The cells were washed with PBS, cold methanol, 1% Tween-20 and 0.5 mM of EDTA in PBS, and PBS. The cells were then incubated in PBS containing 0.25 μg/mL of Hoechst for 15 min at room

temperature to stain the nuclear DNA, and washed twice with PBS and a final wash with deionized water before mounting. For colocalization of SNAP protein expressed, indirect immunofluorescent cytochemical staining was carried out according the manufacturer's instructions. Cells were blocked with 10% BSA in PBS for 30 min, and washed twice with PBS. Then the cells were incubated with monoclonal Anti-FLAG M2 antibody diluted in 1:2000 in 3% BSA in PBS for 2 h at 37 ^{o}C, and washed twice with PBS. Then secondary antibody goat-anti-mouse IgG-FITC (Santa Cruz) diluted in 1:100 in 3% BSA in PBS for 45 min at 37 ℃, and washed twice with PBS. Images were acquired using Observer Z1 (Zeiss, Germany) equipped with a 63X NA1.4 objective and a CoolSNAP HQ2 CCD camera (Photometrics, USA), or LSM 510 META (Zeiss, Germany) equipped with an EC Plan-Neofluar 100× NA1.3 objective.

8.4 Chapter 4

Cell-culture conditions, Guava ViaCount assay, in situ proteomic profiling, pull-down and mass spectrometry identification (*T. brucei* trypanosomes (~1-5 × 10^9 cells, ~5 mg) in growth media (1 × 10^7 cells/mL)), and fluorescence microscopy were carried out using the general protocols as described in **8.1.2.**

Chemicals and Antibodies. 4',6-diamidino-2-phenylindole (DAPI) was purchased from Sigma-Aldrich. MitoTracker Red CMXRos was from Invitrogen.

Fluorescence Microscopy

Drug uptake analysis. Trypanosomes (1×10^5 cells/mL for both forms) were incubated in growth medium containing different concentrations of THL-R at culture temperature and 5% CO_2 for 2 h. Medium containing 1% DMSO was used as negative control. The parasites were then washed twice with PBS, and fixed with 4% paraformaldehyde in PBS for 15 min at room temperature and washed twice with PBS, and then sedimented to poly-L-lysine-coated coverslips. Fixed cells were permeabilized with 0.25% Triton-X 100 in PBS for 15 min at room temperature, and washed twice with PBS. The cells were blocked with 3% BSA in PBS for 30 min at room temperature, and washed twice with PBS. The cells were then treated with a freshly premixed click chemistry reaction solution in PBS for 1 h at room temperature. The cells were washed with PBS, cold methanol, 1% Tween-20 and 0.5 mM EDTA in PBS, and PBS. The cells were then incubated in PBS containing 2 μg/mL DAPI for

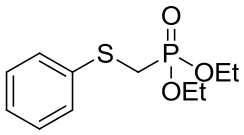
15 min at room temperature to stain kinetoplast and nuclear DNA, and washed twice with PBS and a final wash with deionized water before mounting with Fluoromount G (Emsdiasum, Fort Washington, PA). Images were acquired using Observer Z1 (Zeiss) equipped with a 63X NA1.4 objective and a CoolSNAP HQ2 CCD camera (Photometrics), or LSM 510 META (Zeiss) equipped with an EC Plan-Neofluar 100x NA1.3 objective.

Immunofluorescence analysis. Cells treated with THL-R for 24 h were attached to coverslips, fixed with 4% paraformaldehyde, blocked with 3% BSA in PBS and then incubated with corresponding antibodies. Anti-TbBiP, Anti-TbTrypanopain, and Anti-TbSKL were used to label the ER, the lysosome, and the glycosome, respectively. DAPI (2 μg/ml) was used to stain kinetoplast and nuclear DNA.

8.5 Chapter 5

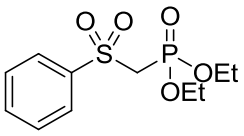
8.5.1 Chemical Syntheis

Diethyl phenylthiomethylphosphonate (5-4)



To a cooled (0 ℃) and stirred suspension of hexane-washed NaH (60% in mineral oil; 1.0 g, 24 mmol) in dry THF (100 mL) at 0 ℃ was added dropwise benzenethiol (2.0 mL, 20 mmol) via syringe. The mixture was stirred for an additional 30 min at 0 ℃ until effervescence ceased. Diethyl iodomethylphosphonate (4.0 mL, 22 mmol) was added dropwise and stirred for 12 h. A cold HCl solution (1 *N*) was added to break up the gelatinous emulsion until pH 6-7 was reached, and concentrated in vacuo. Diluted with H_2O (150 mL) and extracted with EtOAc (3 × 50 mL). The combined organic extracts were washed with saturated aqueous $NaHCO_3$ and brine, dried over Na_2SO_4, filtered and concentrated in vacuo. Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **5-4** as a colorless liquid (4.79 g, 92%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.30 (t, *J* = 7.1, 6H), 3.20 (d, *J* = 14.0 Hz, 2H), 4.09-4.20 (m, 4H), 7.20-7.33 (m, 3H), 7.42-7.46 (m, 2 H).

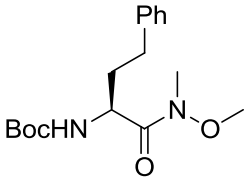
Diethyl phenylsulfonylmethylphosphonate (5-5)



To a solution of compound **5-4** (5.0 g, 19.2 mmol) in CH_2Cl_2 (100 mL) at 0 ℃ was added *m*-chloroperbenzoic acid (12.9 g of 77% *m*-CPBA, 57.2 mmol) over 1 h. The mixture was stirred overnight while warming to room temperature. The solution was then cooled to 0 ℃ and was treated

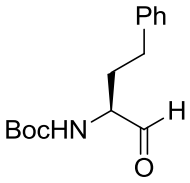
with NaOH (2 *N*) until pH 8-9 was reached. The organic phase was separated, dried over Na_2SO_4, filtered, and concentrated to dryness, giving the product **5-5** as a colorless oil (5.6 g, 94%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.30 (t, *J* = 7.1 Hz, 6H), 3.77 (d, *J* = 17.0 Hz, 2H), 4.11-4.21 (m, 4 H), 7.55-7.61 (m, 2 H), 7.65-7.71 (m, 1 H), 7.98-8.01 (m, 2 H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 1.30 (t, *J* = 7.1 Hz, 6H), 3.77 (d, *J* =17.0 Hz, 2H), 4.11-4.21(m, 4H), 7.55-7.61 (m, 2H), 7.65-7.71 (m, 1 H), 7.98-8.01 (m, 2 H).

(S)-tert-butyl [1-(methoxymethylcarbamoyl)-3-phenylpropyl]carbamate (5-6)



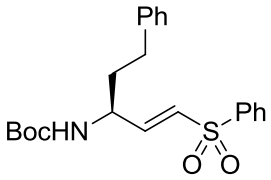
To a solution of (*S*)-Boc-Homophenylalanine (5.59 g, 20 mmol) in dry THF (100 mL) at 0 ℃ was added EDC•HCl (4.60 g, 24 mmol), HOBt (3.24 g, 12 mmol), *N*,*O*-dimethylhydroxylamine hydrochloride (2.34 g, 24 mmol) and DIPEA (5.2 mL, 30 mmol). The reaction was stirred at room temperature for 12 h, and concentrated in vacuo. Diluted with H_2O (150 mL) and extracted with EtOAc (3 × 50 mL). The combined organic extracts were washed with 1 wt% HCl, 20 wt% Na_2CO_3 and brine, dried over Na_2SO_4, filtered and concentrated in vacuo. The crude product was purified by column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give Boc-Hph-N(Me)OMe (**5-6**) as a white solid (6.20 g, 96%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.45 (s, 9H), 1.80-2.02 (m, 1H), 2.72 (m, 1H), 3.16 (s, 3H), 3.62 (s, 3H), 4.68 (br s, 1H), 5.23 (m, 1H), 7.15-7.31 (m, 5H).

(S)-tert-butyl (1-formyl-3-phenylpropyl)carbamate (Boc-Homophenylalaninal, Boc-HphH, 5-7)



To a solution of **5-6** (3.2 g, 10 mmol) in dry THF (50 mL) at 0 ℃ was added $LiAlH_4$ (0.45 g, 12 mmol) over 10 min, with vigorous stirring. The mixture was stirred for an additional 20 min at 0 ℃, whereupon cold water was carefully added until effervescence ceased. A cold HCl solution (1 *N*) was added to break up the gelatinous emulsion until pH 6~7 was reached. Diluted with H_2O (150 mL) and extracted with EtOAc (3 × 50 mL). The combined organic extracts were washed with saturated aqueous $NaHCO_3$ and brine, dried over Na_2SO_4, filtered and concentrated in vacuo. Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **5-7** as a white solid (1.92 g, 73%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.46 (s, 9H), 1.83-1.95 (m, 1H), 2.22 (m, 1H), 2.67 (t, *J* = 7.6 Hz, 2H), 4.24 (m, 1H), 5.09 (br s, 1H), 7.17-7.32 (m, 5H), 9.55 (s, 1H).

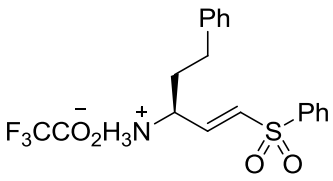
(S)-tert-butyl (3-benzenesulfonyl-1-phenethylallyl)carbamate (Boc-HphVSPh,5-8)



To a cooled (0 ℃) and stirred suspension of *n*-hexane-washed NaH (60% in mineral oil; 0.24 g, 6 mmol) in dry THF (50 mL) at 0 ℃ was added dropwise **5-5** (1.61 g, 5.5 mmol) in dry THF (10 mL) via syringe. The mixture was stirred for an additional 30 min at 0 ℃ and **5-7** (1.32 g, 5 mmol) in dry THF (10 mL) was added dropwise and stirred for 1 h. A cold 5 wt% $NaHSO_4$ solution was added to break up the gelatinous emulsion until pH 6~7 was reached, and concentrated in vacuo. Diluted with water (100 mL) and extracted

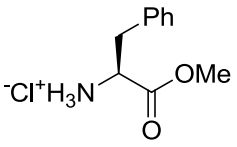
with EtOAc (3 × 25 mL). The combined organic extracts were washed with saturated aqueous $NaHCO_3$ and brine, dried over Na_2SO_4, filtered and concentrated under vacuum. Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **5-8** as a white foam (1.4 g, 70%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.40 (s, 9H), 1.83-1.94 (m, 2H), 2.62-2.70 (m, 2H), 4.36 (br s, 1H), 4.52 (br s, 1H), 6.43 (br d, *J* = 14.5 Hz, 1H), 6.87-6.90 (m, 1H), 7.13-7.30 (m, 5H), 7.51-7.61 (m, 3H), 7.61 (d, *J* = 6.0 Hz, 2H); IT-TOF-MS: (*m/z*) calcd for $C_{22}H_{27}NO_4S$ $[M+Na]^+$ 424.1661, found: 424.1575.

(S)-3-benzenesulfonyl-1-phenethylallylamine trifluoroacetate (TFA •HphVSPh, 5-9)



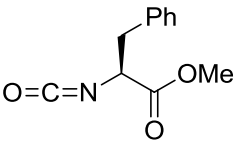
To a cooled (0 ℃) and stirred solution of **5-8** (1.2 g, 3 mmol) in DCM (15 mL) was added dropwise TFA (5 mL) via syringe. After being stirred for 2 h, Et_2O (100 mL) was added with stirring. The precipitate was filtered off, washed twice with Et_2O, and finally dried in vacuo to give 73.4 g (76%) of **5-9**. ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.34 (br s, 2H), 1.76-1.86 (m, 2H), 2.68 (t, *J* = 7.9 Hz, 2H), 3.53 (m, 1H), 6.49 (m, 1H), 6.98 (dd, *J* = 5.6, 14.9 Hz, 1H), 7.12-7.30 (m, 5H), 7.51-7.56 (m, 3H), 7.86 (d, *J* = 7.3 Hz, 2H). This material was pure enough to be used in the next step without further purification.

(S)-N-(4-chlorobenzylidene)phenylalanine methyl ester (HCl •Phe-OMe, 5-10)



To a cooled (0 ℃) and stirred suspension of phenylalanine (16.5 g, 100 mmol) in dry MeOH (150 mL) was added dropwise $SOCl_2$ (9 mL, 120 mmol) over a 1 h period. During the addition the mixture was cooled in ice-bath in order to keep the temperature < 5 ^oC. The clear solution was stirred for 12 h and subsequently for 2 h at 50 ^oC. After evaporation of the solvent under reduced pressure, Et_2O (100 mL) was added with stirring. The precipitate was filtered off, washed twice with ether, and finally dried in vacuo to give 21.6 g (98 %) of **5-10** as a white solid. This material was pure enough to be used in the next step without further purification.

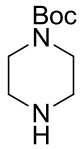
Methyl (S)-2-isocyanato-3-phenylpropanoate (OCN-PheOMe, 5-11)



To a solution of **5-10** (5.5 g, 25.5 mmol) in DCM (50 mL) at 0 °C was added saturated aqueous $NaHCO_3$ (50 mL) and triphosgene (2.52 g, 8.42 mmol) in a single portion with vigorous stirring. The reaction mixture was stirred at 0 °C for 15 min and then poured into a 250-mL separatory funnel. The organic layer was collected, and the aqueous layer is extracted with DCM (3 × 50 mL). The combined organic layers were washed with brine, dried (Na_2SO_4), vacuum filtered, and concentrated at reduced pressure using a rotary evaporator to give the product **5-11** as a colorless oil. ^1H NMR (300 MHz, $CDCl_3$): δ 3.03 (dd, *J* = 7.8, 13.8 Hz, 1H), 3.16 (dd, *J* = 4.8, 13.6 Hz, 1H), 3.81 (s, 3H), 4.27 (dd, *J* = 4.61, 7.8 Hz, 1H), 7.18-7.21 (m, 2H), 7.27-7.36 (m, 3H). This

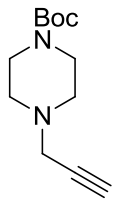
material was used in the next step without further purification, assuming a quantitative yield.

tert-butyl 1-piperazinecarboxylate (5-12)



To a solution of di-tert-butyl dicarbonate (5.80 g, 25.54 mmol) solution in 50 mL of dry MeOH was added drop-wise a solution of piperazine (4.0 g, 46.44 mmol) in 100 mL of dry MeOH previously cooled at 0 ℃. After 30 min, all ditert-butyl dicarbonate had been added and the mixture was warmed to room temperature. After 2 d, the solution was concentrated under reduced pressure and the crude solid was dissolved in 200 mL of Et_2O, and the white precipitate left was filtered off. The aqueous solution obtained by extracting the organic solution with 1 *N* citric acid (aq.) (3 × 100 mL) was washed with EtOAc (3 × 100 mL) and brought to pH ~ 11 by adding solid K_2CO_3. The turbid solution was extracted with EtOAc (3 × 100 mL) and dried over Na_2SO_4. The solution was concentrated under reduced pressure at 40 ℃ and stripped with DCM to yield clear oil which was crystallized into a white solid upon drying under reduced pressure. Yield: 71%; ^{1}H NMR (300 MHz, $CDCl_3$): δ 3.45-3.33 (m, 4H), 2.88-2.74 (m, 4H), 1.57 (s, 1H), 1.46 (s, 9H).

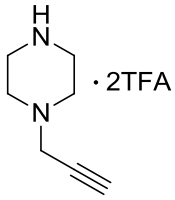
tert-Butyl 4-propargylpiperazine-1-carboxylate (5-13)



To a solution of **5-12** (1.86 g, 10 mmol) and diisopropylethylamine (1.9 mL, 11mmol) in $CHCl_3$ (50 mL) at 0 ℃ was added drop-wise a solution of propargyl bromide (80% in toluene, 1.2 mL, 10mmol) in $CHCl_3$ (50 mL).

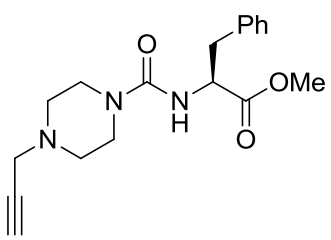
After the mixture was stirred for 24 h at room temperature, the solution obtained was washed with 5% $NaHCO_3$ (3 × 50mL), brine (2 × 50mL), and then dried over Na_2SO_4. The solution was filtered and evaporated to provide a brown oil. Purification by flash column chromatography on silica gel using 50% EtOAc in hexanes to give the product **5-13** as a yellow oil (1.4 g, 86%), which ultimately crystallized upon standing. ^{1}H NMR (500 MHz, $CDCl_3$): 1.49 (s, 9H), 2.26 (t, *J* = 2.5 Hz, 1H), 2.51 (t, *J* = 5.0 Hz, 4H), 3.32 (d, *J* = 2.55 Hz, 2H), 3.47 (t, *J* = 5.0 Hz, 4H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 29.10, 47.67, 52.32, 74.10, 79.10, 80.40, 155.39.

N-Propargylpiperazine•TFA salt (5-14)



To a solution of **5-13** (1.12g, 5 mmol) in CH_2Cl_2 (25 mL) at 0 ℃ was added trifluoroacetic acid (25mL). The solution was stirred at room temperature overnight, and then was evaporated to dryness in vacuum. Assumed quantitative yield of 1.16 g. The residue was suspended in 20 mL THF and used immediately without further purification.

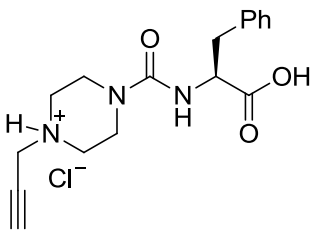
(S)-methyl 3-phenyl-2-(4-(prop-2-yn-1-yl)piperazine-1-carboxamido)propanoate (5-15)



To a solution of **5-14** (0.7 g, 2 mmol) in dry THF (10 mL) at 0 ℃ was added drop-wise a solution of DIEA (0.7 mL, 4 mmol) in 10 mL of dry THF previously cooled at 0 ℃. After 10 min, a solution of **5-11** (0.68 g, 2.4 mmol) in dry THF (10 mL) was added at 0 ℃. The mixture was stirred for 12 h and was

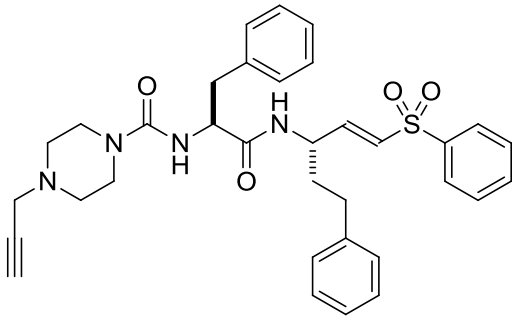
concentrated in vacuo to give a brown oil, and diluted with water (100 mL) and extracted with EtOAc (3 × 50 mL). The combined organic extracts were washed with saturated aqueous $NaHCO_3$ and brine, dried over Na_2SO_4, filtered and concentrated under vacuum. Purification by flash column chromatography on silica gel using 5 to 10% methanol in DCM gave the product **5-15** as a white solid (0.53 g, 80%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.49-2.56 (m, 4H), 3.08-3.16 (m, 2H), 3.31-3.43 (m, 6H), 3.72 (s, 3H), 4.77-4.80 (m, 1H), 4.81-4.91 (m, 1H), 7.10-7.11 (m, 2H), 7.23-7.30 (m, 3H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 38.97, 44.24, 47.53, 52.02, 52.87, 55.00, 74.23, 78.88, 127.68, 129.17, 129.96, 136.87, 157.07, 173.74; IT-TOF-MS: (*m/z*) calcd for $C_{18}H_{23}N_3O_3$ $[M+Na]^+$ 352.1739, found 352.1738.

(S)-3-phenyl-2-(4-(prop-2-yn-1-yl)piperazine-1-carboxamido)propanoic acid hydrochloride (5-16)



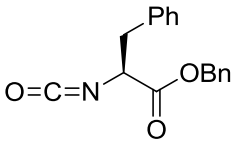
To a solution of **5-15** (0.6 g, 1.8 mmol) in THF (30 mL) at 0 ℃ was added drop-wise a solution of LiOH•H_2O (0.23 g, 5.5 mmol) in 10 mL of H_2O. The mixture was stirred for 4 h, and 4 *N* HCl in dioxane was then added slowly to adjust the pH of the reaction mixture to ~ 2 at 0 ℃. The resulting solution was evaporated in vacu*o*. The residue was washed with Et_2O (2 × 25 mL), dried in vacuo, and then lyophilized overnight to give the crude product **5-16**, along with a little LiCl, which was used directly in the following reaction without further purification, assuming a quantitative yield.

N-((S)-1-oxo-3-phenyl-1-(((S,E)-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino) propan-2-yl)-4-(prop-2-yn-1-yl)piperazine-1-carboxamide (5-1, VS-1)



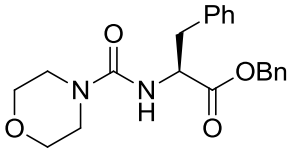
To a solution of **5-16** (215 mg, 0.6 mmol) in DMF (5 mL) was added EDC•HCl (115 mg, 0.6 mmol), HOBt (81 mg, 0.6 mmol) and DIEA (0.4 mL, 2.4 mmol). After 10 min, TFA•HphVSPh **5-9** (208 mg, 0.5 mmol) in DMF (5 mL) was added dropwise. The reaction was stirred at rt for 21 h. The resulting solution was evaporated *in vacuo* to give a brown oil, and diluted with water (50 mL) and extracted with DCM (3 × 50 mL). The combined organic extracts were washed with saturated aqueous $NaHCO_3$ and brine, dried over Na_2SO_4, filtered and concentrated under vacuum. Purification by flash column chromatography on silica gel using 5 to 10% methanol in DCM gave the product **5-1** as a white solid (165 mg, 55%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.72-1.75 (m, 1H), 1.82-1.84 (m, 1H), 2.44-2.56 (m, 7H), 3.01 (d, *J* = 7.6 Hz, 2H), 3.27-3.37 (m, 6H), 4.54-4.60 (m, 2H), 5.13 (d, *J* = 7.6 Hz, 1H), 6.09 (dd, *J* = 1.2, 15.1 Hz, 1H), 6.78 (dd, *J* = 4.9, 15.1 Hz, 1H), 6.87 (d, *J* = 8.3 Hz, 1H), 7.02 (d, *J* = 7.3 Hz, 2H), 7.11-7.24 (m, 8H), 7.54 (t, *J* = 7.6 Hz, 2H), 7.62 (t, *J* = 7.6 Hz, 1H), 7.84 (d, *J* = 7.6 Hz, 2H). ^{13}C NMR (126 MHz, $CDCl_3$): δ 31.77, 35.70, 36.49, 38.48, 43.71, 46.82, 49.10, 51.26, 56.01, 73.66, 78.14, 126.25, 127.14, 127.66, 128.37, 128.56, 128.72, 129.26, 129.30, 130.48, 133.47, 136.66, 140.23, 140.44, 145.66, 156.90, 171.86; IT-TOF-MS: (*m/z*) calcd for $C_{34}H_{38}N_4O_4S$ $[M+H]^+$ 599.2614, found 599.2545.

(S)-benzyl-2-isocyanato-3-phenylpropionate (OCN-PheOBzl, 5-17)



To a solution of (*S*)-Benzyl-2-amino-3-phenylpropionate hydrochloride (HCl•PheOBzl) (3.72 g, 12.75 mmol) in DCM (50 mL) at 0 ℃ was added saturated aqueous $NaHCO_3$ (50 mL) and triphosgene (1.25 g, 4.21 mmol) in a single portion with vigorous stirring. The reaction mixture was stirred at 0 ℃ for 15 min and then poured into a 250-mL separatory funnel. The organic layer was collected, and the aqueous layer is extracted with DCM (3 × 15 mL). The combined organic layers were washed with brine, dried (Na_2SO_4), vacuum filtered, and concentrated at reduced pressure using a rotary evaporator to give a colorless oil. The product, OCN-PheOBzl was used in the next step without further purification, assuming a quantitative yield.

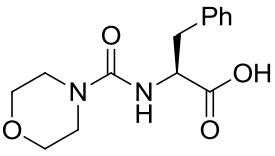
(S)-benzyl 2-(morpholine-4-carboxamido)-3-phenylpropanoate (5-18)



To a solution of **5-17** (3.59 g, 12.75 mmol) in dry THF (50 mL) at 0 ℃ was added morpholine (1.1 mL, 12.75 mmol). The mixture was stirred for 1 h and was concentrated in vacuo to a pale orange oil, and diluted with water (100 mL) and extracted with EtOAc (3 × 50 mL). The combined organic extracts were washed with HCl (1 *N*), saturated aqueous $NaHCO_3$ and brine, dried over Na_2SO_4, filtered and concentrated under vacuum. Purification by flash column chromatography on silica gel using 10 to 20% EtOAc in hexanes to give the product (*S*)-benzyl 2-(morpholine-4-carboxamido)-3-phenylpropanoate (Mu-PheOBzl, **5-18**) as a white

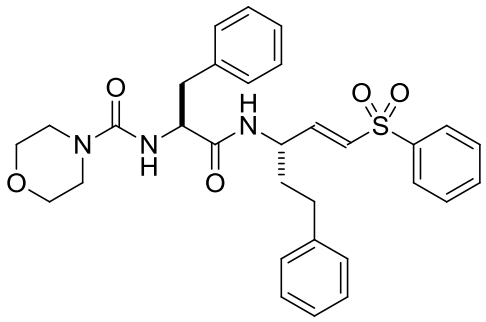
solid (3.9 g, 83% over two steps). ^{1}H NMR (300 MHz, $CDCl_3$): δ 3.11 (d, *J* = 5.3 Hz, 2H), 3.27-3.31 (m, 4H), 3.62-3.65 (m, 4H), 4.81-4.90 (m, 2H), 5.15 (dd, *J* = 12.3, 27.8 Hz, 2H), 6.99 (dd, *J* = 3.5, 7.0 Hz, 2H), 7.19-7.22 (m, 3H), 7.29-7.38 (m, 5H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 38.22, 43.92, 54.28, 66.41, 67.20, 127.01, 128.51, 128.58, 129.34, 135.19, 136.01, 156.66, 172.43.

(S)-2-(morpholine-4-carboxamido)-3-phenylpropanoic acid (5-19)



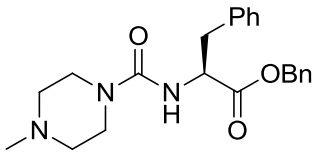
A solution of Mu-PheOBzl (**5-18**) (3.9 g, 10.6 mmol) in 1% HOAc/ethanol (100 mL) was charged with 10% palladium on active charcoal (Aldrich: 0.4 g). The solution in the Parr bottle was exposed to hydrogen on a Parr shaker (50 psi) for 12 h, filtered through Celite, and concentrated in vacuo. The residue was triturated with ether (100 mL) to remove residual ethanol and was reprecipitated from DCM/ether to give 2.94 g (99%) of (*S*)-2-(morpholine-4-carboxamido)-3-phenylpropanoic acid (Mu-PheOH, **5-19**). ^{1}H NMR (300 MHz, DMSO-d_6): δ 2.86-2.94 (m, 1H), 3.00-3.16 (m, 1H), 3.18-3.28 (m, 4H), 3.41-3.48 (m, 4H), 4.19-4.27 (m, 1H), 6.72 (d, *J* = 8.2 Hz, 1H), 7.17-7.30 (m, 5H). ^{13}C NMR (75 MHz, DMSO-d_6): δ 36.60, 43.94, 65.87, 126.24, 128.08, 129.15, 138.44, 157.37, 174.24. Without further purification, the mixture was used directly in the next step.

N-((S)-1-oxo-3-phenyl-1-(((S,E)-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino) propan-2-yl)morpholine-4-carboxamide (K11002)



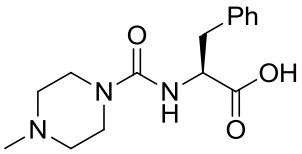
Prepared according to the similar procedure mentioned above by using **5-19** (290 mg, 1.04 mmol), **5-9** (420 mg, 1.0 mmol), EDC•HCl (190 mg, 1.0 mmol), HOBt (140 mg, 1.0 mmg), DIEA (0.34 mL, 2 mmol) in DMF (5 mL). Purification by flash column chromatography on silica gel using 25 to 50% EtOAc in hexanes to give K11002 as a white solid (450 mg, 80%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.74-1.90 (m, 2H), 2.55-2.60 (m, 2H), 3.25 (d, *J* = 4.1 Hz, 2H), 3.25-3.34 (m, 4H), 3.59-3.65 (m, 4H), 4.51 (m, 1H), 4.62 (m, 1H), 5.06 (m, 1H), 6.10 (dd, *J* = 1.65, 15.1 Hz, 1H), 6.79 (dd, *J* = 4.85, 15.1 Hz, 1H), 7.07 (d, *J* = 7.65 Hz, 2H), 7.15-7.28 (m, 8H), 7.57 (t, *J* = 7.8 Hz, 2H), 7.65 (t, *J* = 7.4 Hz, 1H), 7.87 (d, *J* = 7.5 Hz, 2H); IT-TOF-MS: (*m/z*) calcd for $C_{31}H_{35}N_3O_5S$ $[M+H]^+$ 562.2297, found 562.2629.

(S)-benzyl 2-(4-methylpiperazine-1-carboxamido)-3-phenylpropanoate (5-20)



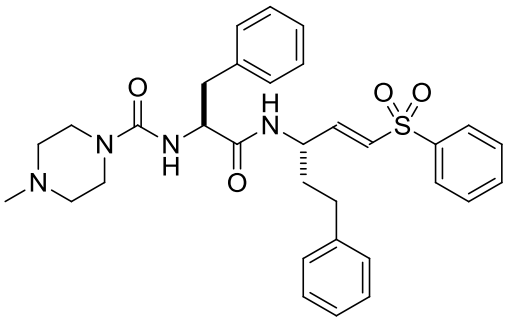
Prepared according to the similar procedure mentioned above by using (*S*)-benzyl-2-amino-3-phenylpropionate hydrochloride (HCl•PheOBzl, **5-17**) (5.84 g, 20 mmol), triphosgene (1.98 g, 6.67 mmol), and *N*-methylpiperazine (2.2 mL, 20 mmol). Purification by flash column chromatography on silica gel using 5 to 10% methanol in DCM to give 6.02 g (79%) the product MePip-PheOBzl (**5-20**) as a pale orange oil.

(S)-2-(4-methylpiperazine-1-carboxamido)-3-phenylpropanoic acid (5-21)



Prepared according to the similar procedure mentioned above by using MePip-PheOBzl (**5-20**) (7.5 g, 19.7 mmol), 10% Pd/C (0.75 g) in 1% HOAc/ethanol (50 mL) under 50 psi for 12 h. The compound was obtained as white solid (5.61 g, 98%). 1H NMR (300 MHz, DMSO-d_6): δ 2.15 (s, 3H), 2.18 (m, 4H), 2.84-2.94 (dd, *J* = 10.8, 15.1 Hz, 1H), 2.95-3.04 (dd, *J* = 5.0, 15.1 Hz, 1H), 3.14-3.30 (m, 4H), 4.17 (m, 1H), 6.65 (d, *J* = 8.0 Hz, 1H), 7.17-7.27 (m, 5H).

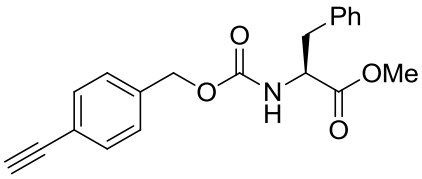
4-Methyl-N-((S)-1-oxo-3-phenyl-1-(((S,E)-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)propan-2-yl)piperazine-1-carboxamide (K11777)



Prepared according to the similar procedure mentioned above by using **5-21** (291 mg, 1.0 mmol), **5-9** (420 mg, 1.0 mmol), EDC•HCl (190 mg, 1.0 mmol), HOBt (140 mg, 1.0 mmg), DIEA (0.34 mL, 2 mmol) in DMF (5 mL). Purification by flash column chromatography on silica gel using 5 to 10% methanol in DCM to give K11777 as a white solid (260 mg, 45%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.76-1.88 (m, 2H), 2.27-2.35 (m, 7H), 2.55-2.60 (m, 2H), 3.05 (d, *J* = 7.5 Hz, 2H), 3.29-3.37 (m, 4H), 4.51 (m, 1H), 4.63 (m, 1H), 5.02 (m, 1H), 6.12 (dd, *J* = 1.55, 15.2 Hz, 1H), 6.79 (dd, *J* = 4.95, 15.1 Hz, 1H), 7.07 (d, *J* = 7.6 Hz, 2H), 7.15-7.30 (m, 8H), 7.57 (t, *J* = 7.9 Hz, 2H), 7.65 (t, *J* = 7.25 Hz, 1H), 7.87 (d, *J* = 1.3 Hz, 2H); IT-TOF-MS: (*m/z*) calcd for $C_{32}H_{38}N_4O_4S$ $[M+H]^+$ 575.2614, found 575.

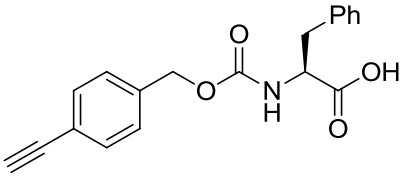
2600.

***(S)-methyl 2-((((4-ethynylbenzyl)oxy)carbonyl)amino)-3-phenylpropanoate (5-22*)**



To a solution of **5-11** (2.46 g, 12 mmol) in anhydrous toluene (25 mL) was added (4-ethynylphenyl)methanol (1.32 g, 10 mmol). The resulting solution was heated to 100 ^{o}C for 6 h and was concentrated in vacuo to a pale orange oil, and diluted with water (100 mL) and extracted with ether (3 × 50 mL). The combined organic extracts were washed with HCl (1 *N*), saturated aqueous $NaHCO_3$ and brine, dried over Na_2SO_4, filtered and concentrated under vacuum. Purification by flash column chromatography on silica gel using 10% EtOAc in hexanes to give **5-22** as a white solid (2.65 g, 79%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 3.03-3.16 (m, 2H), 3.72 (s, 3H), 4.62-4.69 (m, 1H), 5.03-5.12 (m, 2H), 5.34 (br d, *J* = 12.65 Hz, 2H), 7.08-7.11 (m, 2H), 7.21-7.31 (m, 5H), 7.46 (d, *J* = 13.7 Hz, 2H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 38.90, 53.01, 55.49, 67.02, 78.19, 83.98, 122.60, 127.85, 128.46, 129.30, 129.92, 132.93, 136.32, 137.71, 156.12, 172.59.

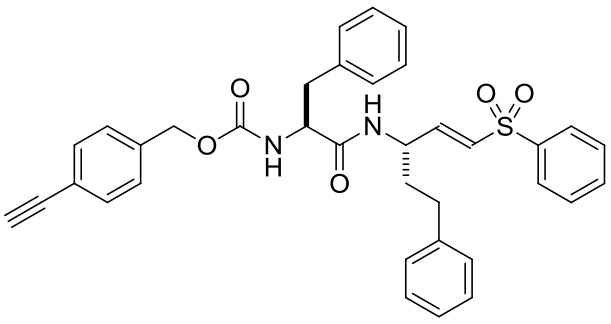
(S)-2-((((4-ethynylbenzyl)oxy)carbonyl)amino)-3-phenylpropanoic acid (5-23)



To a solution of **5-22** (3.4 g, 10 mmol) in THF (60 mL) at 0 ℃ was added dropwise an aqueous solution of LiOH (0.72 g, 30 mmol) in 20 mL of H_2O. The reaction was monitored by TLC (elution with hexane/EtOAc: 1/1) until

complete disappearance (approximately 2 h) of the starting ester. The reaction mixture was then acidified with 2 *N* HCl (to pH ~ 2) and was extracted with EtOAc (3 × 50). After drying over Na_2SO_4, filtration and evaporation of the organic phase, the compound was used directly in the following reaction without further purification (quantitative yield). ^{1}H NMR (300 MHz, DMSO-d_6): δ 2.84 (m, 1H), 2.88-3.11 (m, 1H), 4.17-4.23 (m, 2H), 4.99 (s, 2H), 7.20-7.31 (m, 2H), 7.45 (d, *J* = 13.45 Hz, 2H), 7.68 (d, *J* = 13.95 Hz, 1H); ^{13}C NMR (75 MHz, DMSO-d_6): δ 36.49, 55.52, 64.73, 80.86, 83.27, 121.01, 126.39, 127.54, 128.18, 129.09, 131.65, 137.86, 138.00, 155.88, 173.26.

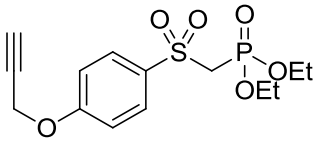
***4-Ethynylbenzyl((S)-1-oxo-3-phenyl-1-(((S,E)-5-phenyl-1-(phenylsulfonyl)pent-1-en-3-yl)amino)propan-2-yl)carbamate (5-2, VS-2*)**



Prepared according to the similar procedure mentioned above by using **5-23** (162 mg, 0.5 mmol), **5-9** (208 mg, 0.5 mmol), EDC•HCl (115 mg, 0.6 mmol), HOBt (81 mg, 0.6 mmg), DIEA (0.2 mL, 1.2 mmol) in DMF (5 mL). Purification by flash column chromatography on silica gel using 20% EtOAc in hexanes give **5-2** as a white solid (258 mg, 85%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.74-1.79 (m, 1H), 1.86-1.89 (m, 1H), 2.52-2.57 (m, 2H), 2.95-3.03 (m, 2H), 3.09 (s, 1H), 4.27 (m, 1H), 4.64 (dd, *J* = 3.5, 5.0 Hz, 1H), 5.05 (s, 2H), 5.23 (br d, *J* = 6.25 Hz, 1H), 5.79 (br d, *J* = 7.55 Hz, 1H), 6.04 (dd, *J* = 0.9, 15.1 Hz, 1H), 6.75 (dd, *J* = 4.7, 15.1 Hz, 1H), 7.03 (d, *J* = 7.2 Hz,

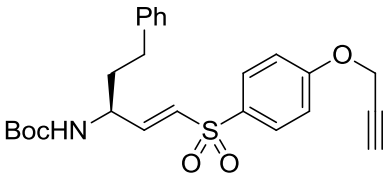
2H), 7.11 (d, *J* = 7.1 Hz, 2H), 7.16-7.28 (m, 8H), 7.44 (d, *J* = 8.1 Hz, 2H), 7.55 (t, *J* = 7.8 Hz, 2H), 7.63 (t, *J* = 7.35 Hz, 1H), 7.85 (d, *J* = 7.65 Hz, 2H); IT-TOF-MS: (*m/z*) calcd for $C_{36}H_{34}N_2O_5S$ $[M+H]^+$ 607.2188, found 607. 2078.

Diethyl (((4-(prop-2-yn-1-yloxy)phenyl)sulfonyl)methyl)phosphonate (5-25)



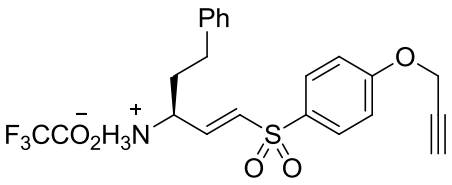
A mixture of 4-hydroxy-thiophenyl-methyl-diethyl-phosphonate sulfone **5-24**[130] (3.08 g, 10 mmol) and anhydrous K_2CO_3 (1.66 g, 12 mmol) in dry acetone (50 mL) was stirred at rt for 2 h. 80% of propargyl bromide in toluene (1.25 mL, 11 mmol) was added dropwise. The mixture was then stirred for 12 h, and TLC analysis indicated all the starting materials had been consumed. After removal of acetone under reduced pressure, the reaction mixture was poured into water (50 mL) and extracted with EtOAc (3 × 50 mL). The combined organic layer was washed successively with 1 *N* HCl, water, and brine. After drying over Na_2SO_4, filtration and evaporation of the organic phase, the compound was purified by flash column chromatography on silica gel using 50% EtOAc in hexanes give **5-25** as a white solid (3.01 g, 87%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.30 (t, *J* = 7.0 Hz, 6H), 2.57 (t, *J* = 2.4 Hz, 1H), 3.74 (d, *J* = 16.8 Hz, 2H), 4.15 (m, 4H), 4.78 (d, *J* = 2.4 Hz, 2H), 7.10-7.23 (m, 2H), 7.93-7.95 (m, 2H).

(S,E)-tert-butyl (5-phenyl-1-((4-(prop-2-yn-1-yloxy)phenyl)sulfonyl)pent-1-en-3-yl) carbamate (5-26)



Prepared according to the similar procedure mentioned above by using **5-24** (3.0 g, 8.67 mmol), **5-7** (2.1 g, 7.88 mmol), and NaH (60% in oil, 0.38 g, 9.5 mmol) in anhydrous THF (100 mL). Purification by flash column chromatography on silica gel using 20% EtOAc in hexanes give **5-26** as a white solid (2.67 g, 75%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.40 (s, 9H), 1.78-1.86 (m, 1H), 1.89-1.96 (m, 1H), 2.55 (t, *J* = 2.5 Hz, 1H), 2.62-2.70 (m, 2H), 4.35 (br s, 1H), 4.52 (br s, 1H), 4.76 (d, *J* = 1.9 Hz, 2H), 6.40 (d, *J* = 15.1 Hz, 1H), 6.84 (dd, *J* = 3.8, 14.5 Hz, 1H), 7.08 (t, *J* = 3.15 Hz, 2H), 7.14 (d, *J* = 6.95 Hz, 2H), 7.18-7.30 (m, 3H), 7.81 (dd, *J* = 2.55, 11.35 Hz, 2H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 28.23, 31.89, 35.96, 50.62, 56.03, 76.53, 77.30, 80.23, 115.44, 126.32, 128.32, 128.61, 129.84, 131.09, 132.67, 140.75, 145.23, 154.88, 161.42, 184.26.

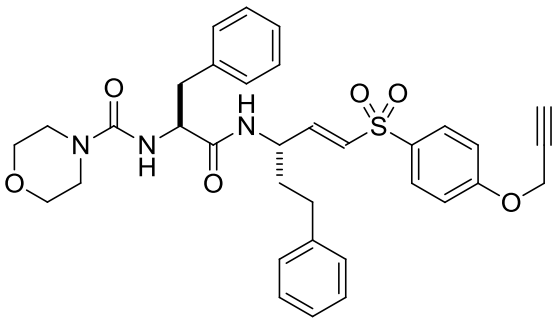
(S,E)-5-phenyl-1-((4-(prop-2-yn-1-yloxy)phenyl)sulfonyl)pent-1-en-3-amine trifluoroacetate (5-27)



Prepared according to the same procedure mentioned above by using **5-26** (2.28 g, 5.0 mmol) in 100 mL of TFA/DCM (1/1), followed by precipitation with Et_2O, filtered off, washed twice with Et_2O, and finally dried in vacuo to give 2.3 g (98%) of **5-27**. This material was pure enough to be used in the next step without further purification.

N-((S)-1-oxo-3-phenyl-1-(((S,E)-5-phenyl-1-((4-(prop-2-yn-1-yloxy)phenyl)sulfonyl

***)pent-1-en-3-yl)amino)propan-2-yl)morpholine-4-carboxamide (5-3, VS-3*)**



Prepared according to the same procedure mentioned above by using **5-19** (139 mg, 0.5 mmol), **5-27** (234 mg, 0.5 mmol), EDC•HCl (115 mg, 0.6 mmol), HOBt (81 mg, 0.6 mmg), DIEA (0.2 mL, 1.2 mmol) in DMF (5 mL). Purification by flash column chromatography on silica gel using 20% EtOAc in hexanes give **5-3** as a white solid (265 mg, 86%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.64-1.85 (m, 2H), 2.41 (t, *J* = 7.55 Hz, 2H), 2.59 (t, *J* = 1.8 Hz, 1H), 3.03-3.11 (m, 2H), 3.22-3.31 (m, 4H), 3.60-3.66 (m, 4H), 4.46 (dd, *J* = 7.6, 15.2 Hz, 1H), 4.60-4.62 (m, 1H), 4.78 (d, *J* = 2.55 Hz, 2H), 5.03 (d, *J* = 7.55 Hz, 1H), 6.39 (br s, 1H), 6.50 (br d, *J* = 15.15 Hz, 1H), 6.80 (dd, *J* = 5.0, 15.1 Hz, 1H), 7.05 (d, *J* = 6.95 Hz, 2H), 7.08 (dd, *J* = 1.9, 6.95 Hz, 2H), 7.20-7.30 (m, 8H), 7.80 (dd, *J* = 1.85, 6.9 Hz, 2H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 31.68, 35.48, 37.97, 43.92, 49.15, 56.04, 56.20, 66.27, 76.62, 77.32, 115.41, 126.25, 127.15, 128.32, 128.53, 128.79, 129.22, 129.84, 131.16, 132.63, 136.68, 140.37, 144.48, 157.23, 161.41, 171.64; IT-TOF-MS: (*m/z*) calcd for $C_{34}H_{37}N_3O_6S$ $[M+H]^+$ 616.2403, found 616. 2293.

8.5.2 Cell Biological Assays

Cell-culture conditions, Guava ViaCount assay, in situ proteomic profiling, pull-down and mass spectrometry identification, and fluorescence microscopy were

carried out using the general protocols as described in **8.1.2.**

8.5.2.1 Chemicals and Antibodies

Anti-cathepsin L (ab6314) was from Abcam. Anti-rhodesain ∆ C, and anti-TbcatB were generous gifts from James H. McKerrow (University of California, San Francisco).

8.5.2.2 Western Blotting

Pull-down samples from in situ labeling with VS-1 were separated on 12% SDS-PAGE gel together with pull-down sample from DMSO-treated (negative control). After SDS-PAGE gel separation, proteins were then transferred to a PVDF membrane and subsequently blocked with 3% (w/v) BSA/PBST overnight at 4^{o}C. Membranes were incubated for 1 h at room temperature with the respective antibodies (anti-cathepsin L for HepG2; anti-rhodesain ∆ C, or TbcatB for *T. brucei*), and washed with PBST (3 × 15 min with gentle agitation), then followed by incubation with a anti-mouse-IgG conjugated secondary antibody in the blocking buffer mentioned above. After wash with PBST (3 × 15 min with gentle agitation), the SuperSignal West Pico kit (Pierce) was used to develop the blot.

8.5.2.3 Fluorescence Microscopy

Confocal images were taken on a Leica TCS SP5X Confocal Microscope System equipped with Leica HCX PL APO 100×/1.40 oil objective, 405 nm Diode

laser, White laser (470 nm to 670 nm, with 1 nm increments, with 8 channels AOTF for simultaneous control of 8 laser lines, each excitation wavelength provides 1.5 mV, PMT detector range from 420 nm to 700 nm for steady state fluorescence. DAPI, FITC and rhodamine were excited with a krypton/argon laser at 405, 488 nm and 554 nm, respectively, and the emission was collected through a 420-470, 500-550 and 565-650 nm filters, respectively. Images were processed with Leica Application Suite Advanced Fluorescence (LAS AF).

8.6 Chapter 6

8.6.1 Chemical Syntheis

*General procedure for the preparation of **6-7a–c***

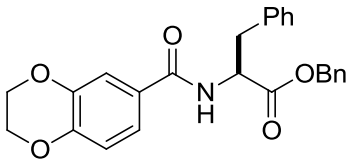
To a cooled (0 ℃) and stirred suspension of Boc-protected amino acid (25.0 mmol) in dry DMF (50 mL) was added Cs_2CO_3 (9.0 g, 27.5 mmol). After being stirred for 30 min, BnBr (3.0 mL, 25.0 mmol) was added via syringe. The mixture was stirred for 12 h, filtered and concentrated in vacuo. The residue was diluted with water (50 mL), and extracted with ether (3 × 50 mL). The combined organic extracts were washed twice with water and brine, dried (Na_2SO_4), and the solvent was removed in vacuo. This material was used in the next step without further purification. To a stirred solution of Boc-protected amino acid benzyl ester (20 mmol) in dry THF (25 mL) cooled at 0 ^oC was added a 4.0 M solution of HCl in 1,4-dioxane (15 mL) dropwise via syringe. After being stirred for 12 h, ether (150 mL) was added with stirring. The precipitate was filtered off, washed twice with ether, and finally dried in a vacuum desiccator (P_2O_5). This material was pure enough to be used in the next step without further purification.

*General procedure for the preparation of **6-8a–c***

To a cooled (0 ℃) solution of 1,4-benzodioxane-6-carboxylic acid acid (1.80 g, 10.0 mmol) in DMF (50 mL) was added EDC•HCl (2.31 g 12.0 mmol), HOBt (1.62 g, 12.0 mmol), **6-7a** (2.92 g, 10.0 mmol) and DIPEA (2.6 mL, 15.0 mmol). The reaction was stirred at room temperature for 18 h, and concentrated in vacuo. Diluted

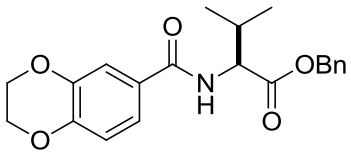
with water (150 mL) and extracted with EtOAc (3 × 50 mL). The combined organic extracts were washed with 1 wt% HCl, 20 wt% Na_2CO_3 and brine, dried over Na_2SO_4, filtered and concentrated under vacuum. The crude product was purified by column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the desired product **6-8a**.

(S)-benzyl 2-(2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamido)-3-phenylpropanoate (6-8a)



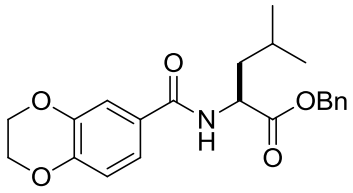
(white solid, 90%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 3.16-3.29 (m, 2H), 4.24-4.30 (m, 4H), 5.06-5.23 (m, 3H), 6.46 (d, *J* = 5.0 Hz, 1H), 6.87 (d, *J* = 15.0 Hz, 1H), 6.99-7.02 (m, 2H), 7.19-7.39 (m, 10H).

(S)-benzyl 2-(2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamido)-3-methylbutanoate (6-8b)



(white solid, 92%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 0.93 (d, *J* = 3.0 Hz, 3H), 0.97 (d, *J* = 6.0 Hz, 3H), 2.24-2.30 (m, 1H), 4.26-4.30 (m, 4H), 4.79-4.81 (m, 1H), 5.15-5.24 (m, 2H), 6.51 (d, *J* = 6.0 Hz, 1H), 6.89 (d, *J* = 6.0 Hz, 1H), 7.29-7.38 (m, 7H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 17.8, 19.0, 31.7, 57.4, 64.2, 64.6, 67.1, 116.6, 117.3, 120.5, 127.5, 128.4, 128.5, 128.6, 135.4, 143.4, 146.7, 166.6, 172.1.

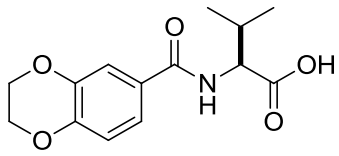
(S)-benzyl 2-(2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamido)-4-methylpentanoate (6-8c)



(white solid, 87%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 0.93-0.99 (m, 6H), 1.64-1.75 (m, 3H), 4.26-4.31 (m, 4H), 4.79-4.91 (m, 1H), 5.19 (m, 2H), 6.43 (d, *J* = 6.0 Hz, 1H), 6.88 (dd, *J* = 3.0, 6.0 Hz, 1H), 7.27-7.38 (m, 7H).

General procedure for the preparation of ***6-9a–c***

A solution of **6-8b** (2.96 g, 8.0 mmol) in a mixture of 100 mL ethanol/1 mL HOAc was prepared and transferred to a Parr bottle. Under a nitrogen atmosphere, a portion of 10% Pd/C (300 mg) was added to the solution in the Parr bottle. The Parr bottle was then installed on a Parr shaker apparatus and then pressurized with hydrogen to about 50 psi and was shaken until the hydrogen uptake ceased. The reaction mixture was filtered through Celite to remove the catalyst and the filtrate was evaporated under reduced pressure to give the product **6-9b** as a white solid (2.1 g, 94%). This material was pure enough to be used in the next step without further purification.

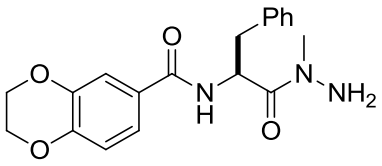


^{1}H NMR (300 MHz, DMDO-d_6): δ 0.93 (d, *J* = 6.0 Hz, 3H), 0.95 (d, *J* = 6.0 Hz, 3H), 2.11-2.20 (m, 1H), 4.21-4.29 (m, 5H), 6.92 (d, *J* = 9.0 Hz, 1H), 7.45 (t, *J* = 3.0 Hz, 2H), 8.21 (d, *J* = 9.0 Hz, 1H); ^{13}C NMR (75 MHz, DMDO-d_6): δ 18.8, 19.3, 29.5, 58.4, 64.0, 64.4, 104.5, 116.6, 116.7, 121.1, 127.1, 142.8, 146.1, 165.9, 173.2.

General procedure for the preparation of ***6-10a–c***

To a stirred solution of **6-9a** (1.6 g, 4.9 mmol) in dry THF (25 mL) cooled at -25 ^{o}C was added dropwise *N*-methylmorpholine (NMM) (0.65 mL, 5.9 mmol), and isobutyl chloroformate (ISCF) (0.77 mL, 5.9 mmol) via syringe. The reaction was stirred for 1 h and filtered to remove the NMM•HCl salt. Methylhydrazine sulfate (1.5 g, 10 mmol) was dissolved in H_2O (1 mL) and 5 *N* NaOH (4 mL) was added under ice cooling. Methylhydrazine solution was added dropwise to the filtrate at −25 ^{o}C. The reaction was allowed to warm to room temperature and stirred for 24 h. After evaporation of the solvent, the aqueous residue was extracted with EtOAc (3 × 20 mL). The combined organic extracts were washed with H_2O, saturated $NaHCO_3$, brine, dried over Na_2SO_4 and concentrated in vacuo. Purification by flash column chromatography on silica gel using 50 to 80% EtOAc in hexanes to give the desired product **6-10a**.

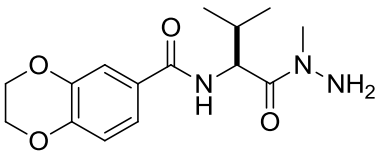
(S)-N-(1-(1-methylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-10a)



(white solid, 1.46 g, 84%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.67 (br d, *J* = 12.0 Hz, 2H), 3.09 (s, 3H), 3.35 (s, 2H), 4.25-4.17 (m, 4H), 5.91-5.98 (m, 1H), 6.86 (d, *J* = 6.0 Hz, 2H), 7.21-7.3 (m, 6H).

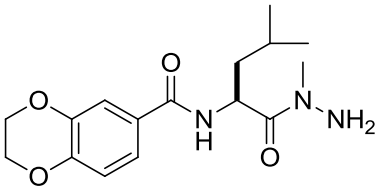
(S)-N-(3-methyl-1-(1-methylhydrazinyl)-1-oxobutan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-10b)

(white solid, 1.16 g, 84%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 0.94-1.03 (m, 6H),



2.23-2.20 (m, 1H), 3.21 (s, 3H), 4.27 (d, *J* = 6.0 Hz, 4H), 5.64 (dd, *J* = 6.0, 9.0 Hz, 1H), 6.73 (br d, *J* = 9.0 Hz, 1H), 6.89 (dd, *J* = 3.0, 6.0 Hz, 1H), 7.30-7.38 (m, 2H).

(S)-N-(4-methyl-1-(1-methylhydrazinyl)-1-oxopentan-2-yl)-2,3-dihydrobenzo[b][1,4] dioxine-6-carboxamide (6-10c)



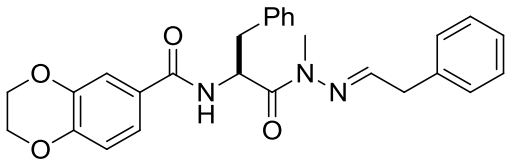
(white solid, 1.01 g, 52%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 0.94 (d, *J* = 6.0 Hz, 3H), 0.98 (d, *J* = 3.0 Hz, 3H), 1.6-1.76 (m, 3H), 3.76 (s, 3H), 4.28 (br s, 4H), 4.79-4.86 (m, 1H), 6.56 (br s, 1H), 6.86-6.89 (m, 1H), 7.29-7.35 (m, 2H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 22.0, 22.8, 24.9, 41.7, 51.0, 52.3, 64.1, 64.5, 116.6, 117.2, 120.5, 127.1, 143.3, 146.6, 163.3, 173.8.

General procedure for the preparation of ***6-11a–o***

To a stirred solution of **6-10a** (178 mg, 0.5 mmol) in dry THF (5 mL) was added benzeneacetaldehyde (0.11 mL, 1.0 mmol), followed by anhydrous $MgSO_4$ (240 mg, 2.0 mmol). After stirring at room temperature for 4 h, one more equivalent (0.11 mL, 1.0 mmol) of benzeneacetaldehyde was added and stirring was continued for 12 h, filtered and concentrated in vacuo. Purification by flash column chromatography (silica gel) using 20 to 50% EtOAc in hexanes to give the desired product **6-11c**.

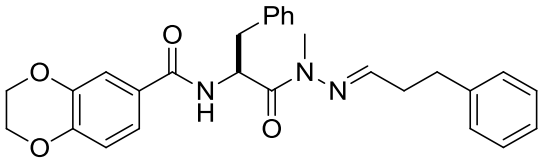
(S,E)-N-(1-(1-methyl-2-(2-phenylethylidene)hydrazinyl)-1-oxo-3-phenylpropan-2-yl

)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-11c)



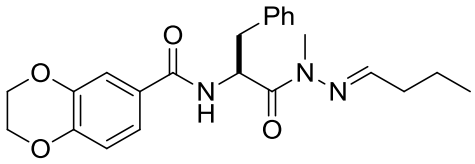
(white solid, 204 mg, 89%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 3.08-3.27 (m, 5H), 3.67 (d, *J* = 6.0 Hz, 2H), 4.25-4.29 (m, 4H), 5.98-6.05 (m, 1H), 6.86 (d, *J* = 9.0 Hz, 2H), 7.07-7.25 (m, 12H).

(S,E)-N-(1-(1-methyl-2-(3-phenylpropylidene)hydrazinyl)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-11d)



(white solid, 215 mg, 91%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 2.63-2.70 (m, 2H), 2.91 (t, *J* = 9.0 Hz, 2H), 3.01-3.20 (m, 5H), 4.25-4.29 (m, 4H), 5.95-6.01 (m, 1H), 6.85 (d, *J* = 9.0 Hz, 2H), 7.04-7.07 (m, 3H), 7.18-7.32 (m, 9H).

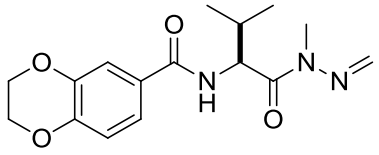
(S,E)-N-(1-(1-methyl-2-pentylidenehydrazinyl)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-11e)



(white solid, 178 mg, 87%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.00 (t, *J* = 9.0 Hz, 3H), 1.58-1.67 (m, 2H), 2.29-2.36 (m, 2H), 3.08-3.27 (m, 5H), 4.24-4.30 (m, 4H), 5.96-6.02 (m, 1H), 6.86 (d, *J* = 8.4 Hz, 2H), 7.05-7.11 (m, 3H), 7.18-7.32 (m, 4H).

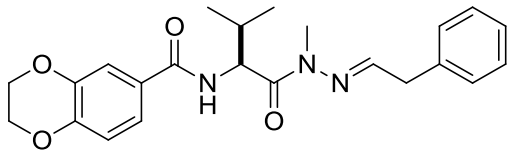
(S)-N-(3-methyl-1-(1-methyl-2-methylenehydrazinyl)-1-oxobutan-2-yl)-2,3-dihydrob

enzo[b][1,4]dioxine-6-carboxamide (6-11f)



(white solid, 126 mg, 79%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 0.91 (d, *J* = 6.0 Hz, 3H), 1.00 (d, *J* = 6.0 Hz, 3H), 2.15-2.26 (m, 1H), 3.23 (s, 3H), 4.26-4.30 (m, 4H), 5.78 (dd, *J* = 6.0, 9.0 Hz, 1H), 6.55 (d, *J* = 9.0 Hz, 3H), 6.69-6.87 (m, 2H), 6.88 (d, *J* = 9.0 Hz, 1H), 7.31-7.39 (m, 2H).

(S,E)-N-(3-methyl-1-(1-methyl-2-(2-phenylethylidene)hydrazinyl)-1-oxobutan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-11h)



(white solid, 172 mg, 84%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 0.89 (d, *J* = 6.9 Hz, 3H), 0.99 (d, *J* = 6.9 Hz, 3H), 2.14-2.25 (m, 1H), 3.22 (s, 3H), 3.71 (d, *J* = 6.0 Hz, 2H), 4.26-4.31 (m, 4H), 5.76 (dd, *J* = 3.0, 9.0 Hz, 1H), 6.81 (d, *J* = 9.0 Hz, 1H), 6.89 (d, *J* = 8.4 Hz, 1H), 7.17 (t, *J* = 5.4 Hz, 1H), 7.24-7.35 (m, 6H).

General procedure for the preparation of ***6-12a–o***

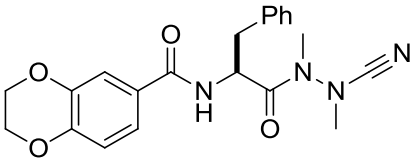
To a cooled (0 ℃) solution of **6-11a** (184 mg, 0.4 mmol) in DCM (4 mL) was added a solution of *p*-toluenesulfonic acid (460 mg, 2.4 mmol) in DCM/MeOH (3:1; 8 mL), and dimethylamine borane (DMAB; 40 mg, 0.64 mmol). The mixture was allowed to react at room temperature for 2 h before additional 1.6 equivalents DMAB (40 mg, 0.64 mmol) were added. After 2 h, 1.5 *N* NaOH (10 mL) was added

and stirring was continued for 30 min. The volume was reduced in vacuo, and the aqueous residue was extracted with DCM (3 × 20 mL). The combined organic layers were washed with H_2O and brine, and the solvent was dried (Na_2SO_4), filtered and concentrated under vacuum. This material was pure enough to be used in the next step without further purification.

*General procedure for the preparation of **6-1a–o***

To a stirred solution of **6-12a** (185 mg, 0.5 mmol) in dry MeOH (10 mL) was added anhydrous sodium acetate (164 mg, 2.0 mmol) and cyanogen bromide (106 mg, 1.0 mmol). The mixture was stirred at room temperature for 12 h, two additional equivalents of cyanogen bromide (106 mg, 1.0 mmol) were added and stirring was continued for 24 h. The solvent was removed under reduce pressure, and the oily residue was suspended in H_2O (10 mL). A pH of 1-2 was adjusted (10% $KHSO_4$), it was extracted with EtOAc (5 × 20 mL), and the combined organic layers were washed with H_2O, saturated $NaHCO_3$ and brine. The solvent was dried (Na_2SO_4) and evaporated. Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes give the product **6-1a** as a white solid (146 mg, 74%).

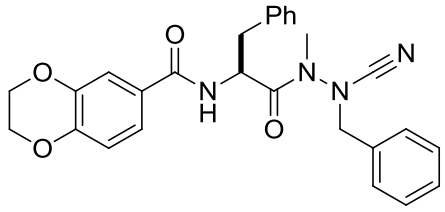
(S)-N-(1-(2-cyano-1,2-dimethylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1a)



^{1}H NMR (500 MHz, $CDCl_3$): δ 3.00-3.23 (m, 2H), 3.24 (s, 3H), 3.33 (s, 3H), 4.26 (dd, *J* = 5.0, 15.0

Hz, 4H), 5.41-5.46 (m, 1H), 6.71 (d, *J* = 5.0 Hz, 1H), 6.84 (d, *J* = 10.0 Hz, 1 H), 7.18-7.34 (m, 7H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 29.7, 30.5, 37.9, 41.1, 50.7, 64.1, 64.5, 113.5, 116.6, 117.2, 120.6, 126.7, 127.3, 128.8, 129.2, 129.6, 135.6, 143.3, 146.8, 166.7, 173.4; LC-IT-TOF/MS: (*m/z*) calcd for $C_{21}H_{22}N_4O_4$ $[M+H]^+$ 395.1641, found 395.1654.

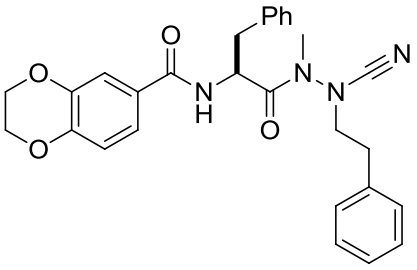
(S)-N-(1-(2-benzyl-2-cyano-1-methylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1b)



Prepared according to the general procedure using **6-12b** (102 mg, 0.23 mmol), sodium acetate (75 mg, 0.9 mmol), and cyanogen bromide (96 mg, 0.9 mmol) in 10 mL of dry methanol to afford 84 mg (79%) of **6-1b** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.97-3.14 (m, 2H), 3.16 (s, 3H), 4.23-4.27 (m, 4H), 4.64 (s, 2H), 5.52-5.56 (m, 1H), 6.84-6.87 (m, 2H), 7.17-7.53 (m, 13H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 32.3, 37.8, 51.3, 58.9, 64.1, 64.5, 112.3, 115.9, 116.6, 117.2, 120.6, 126.8, 127.2, 127.4, 128.2, 128.7, 128.8, 129.1, 129.3, 129.4, 129.5, 129.7, 130.3, 131.5, 135.8, 143.3, 146.7, 166.8, 173.8; LC-IT-TOF/MS: (*m/z*) calcd for $C_{27}H_{26}N_4O_4$ $[M+H]^+$ 471.1954, found 471.1968.

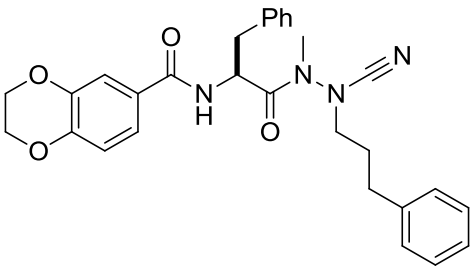
(S)-N-(1-(2-cyano-1-methyl-2-phenethylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1c)

Prepared according to the general procedure using **6-12c** (138 mg, 0.3 mmol), sodium acetate (100 mg, 1.2 mmol), and cyanogen bromide (128 mg, 1.2 mmol) in 10 mL



of dry methanol to afford 118 mg (81%) of **6-1c** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.85-2.89 (m, 2H), 3.06-3.16 (m, 2H), 3.18 (s, 3H), 3.71-3.76 (m, 1H), 3.83-3.89 (m, 1H), 4.26 (dd, *J* = 5.0, 10.0 Hz, 4H), 5.11-5.16 (m, 1H), 6.38 (d, *J* = 10.0 Hz, 1H), 6.85 (d, *J* = 10.0 Hz, 1H), 7.03 (d, *J* = 10.0 Hz, 1H), 7.15-7.36 (m, 11H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 31.6, 32.9, 37.7, 50.6, 54.8, 64.2, 64.5, 112.3, 116.6, 117.3, 120.5, 126.9, 127.2, 127.3, 128.8, 128.9, 129.0, 129.2, 129.5, 135.4, 136.2, 143.4, 146.8, 166.6, 173.4; LC-IT-TOF/MS: (*m/z*) calcd for $C_{28}H_{28}N_4O_4$ $[M+H]^+$ 485.2111, found 485.2102.

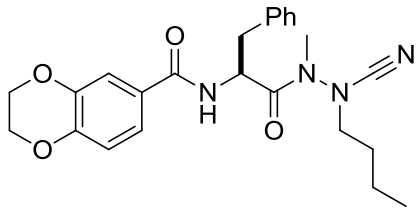
(S)-N-(1-(2-cyano-1-methyl-2-(3-phenylpropyl)hydrazinyl)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1d)



Prepared according to the general procedure using **6-12d** (160 mg, 0.34 mmol), sodium acetate (112 mg, 1.4 mmol), and cyanogen bromide (143 mg, 1.4 mmol) in 10 mL of dry methanol to afford 123 mg (73%) of **6-1d** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.10-2.16 (m, 2H), 2.74-2.86 (m, 2H), 3.00-3.04 (m, 2H), 3.22 (s, 3H), 3.45-3.55 (m, 2H), 4.26 (dd, *J* = 5.0, 10.0 Hz, 4H), 5.44-5.48 (m, 1H), 6.63 (d, *J* = 10.0 Hz, 1H), 6.84 (d, *J* = 10.0 Hz, 1H), 7.13-7.33 (m, 11H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 28.2, 31.6, 32.6, 38.0, 50.8, 53.6, 64.2, 64.6, 112.5, 116.6, 117.3, 120.6, 126.4, 126.9, 127.4, 128.4, 128.7, 128.8, 129.2, 129.6, 135.7, 140.0, 143.4, 146.8, 166.6, 168.4, 173.5; LC-IT-TOF/MS: (*m/z*) calcd for $C_{29}H_{30}N_4O_4$ $[M+H]^+$ 499.2267,

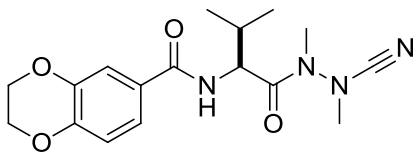
found 499.2329.

(S)-N-(1-(2-butyl-2-cyano-1-methylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1e)



Prepared according to the general procedure using **6-12e** (155 mg, 0.38 mmol), sodium acetate (123 mg, 1.5 mmol), and cyanogen bromide (160 mg, 1.5 mmol) in 10 mL of dry methanol to afford 113 mg (69%) of **6-1e** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 0.99 (t, *J* = 10.0 Hz, 3H), 1.47-1.56 (m, 2H), 1.74-1.83 (m, 2H), 2.96-3.21 (m, 2H), 3.26 (s, 3H), 3.42- 3.54 (m, 2H), 4.25 (d, *J* = 10.0 Hz, 4H), 5.39-5.43 (m, 1H), 6.79-6.87 (m, 2H), 7.18-7.33 (m, 6H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 13.7, 19.6, 19.8, 28.6, 31.5, 37.7, 50.9, 53.8, 64.1, 64.5, 112.5, 116.6, 117.2, 120.6, 126.9, 127.2, 128.7, 129.2, 129.5, 135.8, 143.3, 146.7, 166.6, 173.5; LC-IT-TOF/MS: (*m/z*) calcd for $C_{24}H_{28}N_4O_4$ $[M+H]^+$ 437.2111, found 437.2102.

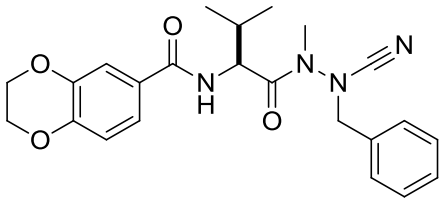
(S)-N-(1-(2-cyano-1,2-dimethylhydrazinyl)-3-methyl-1-oxobutan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1f)



Prepared according to the general procedure using **6-12f** (106 mg, 0.33 mmol), sodium acetate (107 mg, 1.3 mmol), and cyanogen bromide (138 mg, 1.3 mmol) in 10 mL of dry methanol to afford 80 mg (70%) of **6-1f** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.04 (d, *J* = 10.0 Hz, 3H), 1.07 (d, *J* = 10.0 Hz, 3H), 2.09-2.17 (m, 1H), 3.24 (s, 3H), 3.32 (s, 3H), 4.26-4.30 (m, 4H), 5.10-5.13 (m, 1H),

6.57 (d, *J* = 10.0 Hz, 1H), 6.88 (d, *J* = 5.0 Hz, 1H), 7.29-7.35 (m, 2H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 18.1, 19.6, 30.2, 31.4, 41.2, 54.0, 64.2, 64.5, 113.9, 116.6, 117.3, 120.6, 127.0, 143.4, 146.8, 166.9, 173.0; LC-IT-TOF/MS: (*m/z*) calcd for $C_{17}H_{22}N_4O_4$ $[M+H]^+$ 347.1641, found 347.347.1648.

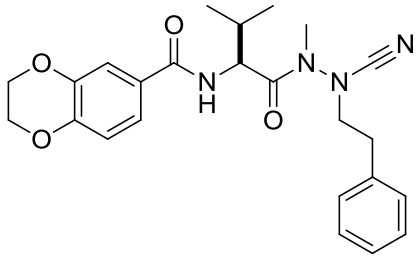
(S)-N-(1-(2-benzyl-2-cyano-1-methylhydrazinyl)-3-methyl-1-oxobutan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1g)



Prepared according to the general procedure using **6-12g** (140 mg, 0.35 mmol), sodium acetate (115 mg, 1.4 mmol), and cyanogen bromide (149 mg, 1.4 mmol) in 10 mL of dry methanol to afford 115 mg (78%) of **6-1g** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.03 (d, *J* = 10.0 Hz, 3H), 1.05 (d, *J* = 10.0 Hz, 3H), 2.10-2.17 (m, 1H), 3.13 (s, 3H), 4.29 (dd, *J* = 5.0, 10.0 Hz, 4H), 4.62 (dd, *J* = 15.0, 20.0 Hz, 2H), 5.26-5.30 (m, 1H), 6.54 (d, *J* = 10.0 Hz, 1H), 6.91 (d, *J* = 5.0 Hz, 1H), 7.30-7.44 (m, 6H), 7.55 (d, *J* = 10.0 Hz, 1H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 18.0, 19.7, 31.5, 32.2, 54.2, 59.0, 64.2, 64.6, 112.7, 116.7, 117.3, 120.6, 127.2, 129.3, 129.7, 130.3, 131.6, 143.4, 146.8, 167.0, 174.1; LC-IT-TOF/MS: (*m/z*) calcd for $C_{23}H_{26}N_4O_4$ $[M+H]^+$ 423.1594, found 423.1606.

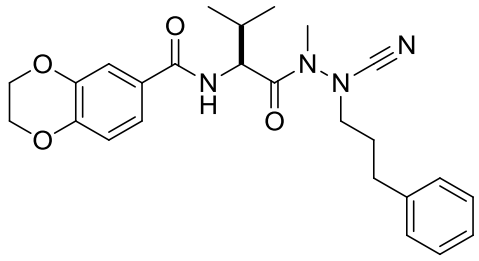
(S)-N-(1-(2-cyano-1-methyl-2-phenethylhydrazinyl)-3-methyl-1-oxobutan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1h)

Prepared according to the general procedure using **6-12h** (136 mg, 0.33 mmol), sodium acetate (108 mg, 1.3 mmol), and cyanogen bromide (141 mg, 1.3 mmol) in 10



mL of dry methanol to afford 97 mg (67%) of **6-1h** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 0.89 (t, *J* = 10.0 Hz, 6H), 1.94-2.04 (m, 1H), 3.06-3.13 (m, 2H), 3.19 (s, 3H), 3.60-3.73 (m, 1H), 3.84-3.89 (m, 1H), 4.27-4.30 (m, 4H), 4.72 (t, *J* = 10.0, 1H), 6.36 (d, *J* = 10.0 Hz, 1H), 6.89 (d, *J* = 10 Hz, 1H), 7.23-7.38 (m, 7H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 17.9, 19.5, 31.0, 31.4, 32.9, 54.1, 54.8, 64.2, 64.6, 112.6, 116.6, 117.3, 120.5, 127.1, 127.2, 128.8, 129.2, 136.3, 143.4, 146.7, 166.9, 174.0; LC-IT-TOF/MS: (*m/z*) calcd for $C_{24}H_{28}N_4O_4$ $[M+H]^+$ 437.2111, found 437.2121.

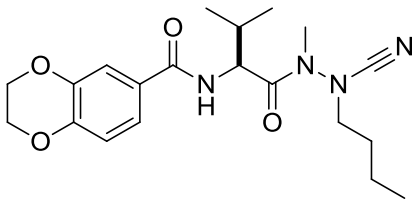
(S)-N-(1-(2-cyano-1-methyl-2-(3-phenylpropyl)hydrazinyl)-3-methyl-1-oxobutan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1i)



Prepared according to the general procedure using **12i** (127 mg, 0.3 mmol), sodium acetate (98 mg, 1.2 mmol), and cyanogen bromide (127 mg, 1.2 mmol) in 10 mL of dry methanol to afford 101 mg (75%) of **6-1i** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.02 (d, *J* = 5.0 Hz, 3H), 1.08 (d, *J* = 5.0 Hz, 3H), 2.09-2.18 (m, 3H), 2.78-2.85 (m, 2H), 3.21 (s, 3H), 3.43-3.51 (m, 2H), 4.26 (dd, *J* = 5.0, 10.0 Hz, 4H), 5.18 (dd, *J* = 5.0, 10.0 Hz, 1H), 6.54 (d, *J* = 10.0 Hz, 1H), 6.89 (d, *J* = 10 Hz, 1H), 7.19-7.34 (m, 7H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 17.8, 19.7, 28.1, 31.38, 31.42, 32.6, 53.6, 53.9, 64.2, 64.5, 113.7, 116.6, 117.3, 120.5, 126.3, 127.2, 128.3, 128.6, 140.1, 143.4, 146.7, 166.8, 174.0; LC-IT-TOF/MS: (*m/z*) calcd for $C_{25}H_{30}N_4O_4$ $[M+H]^+$ 451.2267, found

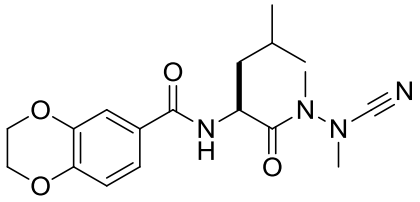
451.2284.

(S)-N-(1-(2-butyl-2-cyano-1-methylhydrazinyl)-3-methyl-1-oxobutan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1j)



Prepared according to the general procedure using **6-12j** (124 mg, 0.34 mmol), sodium acetate (112 mg, 1.4 mmol), and cyanogen bromide (144 mg, 1.4 mmol) in 10 mL of dry methanol to afford 94 mg (71%) of **6-1j** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.02 (d, *J* = 5.0 Hz, 3H), 1.08 (d, *J* = 5.0 Hz, 3H), 1.49-1.58 (m, 2H), 1.73-1.81 (m, 2H), 2.13-2.19 (m, 1H), 3.24 (s, 3H), 3.37-3.50 (m, 2H), 4.26-4.30 (m, 4H), 5.14 (dd, *J* = 5.0, 10.0 Hz, 1H), 6.54 (d, *J* = 5.0 Hz, 1H), 6.89 (d, *J* = 10.0 Hz, 1 H), 7.28-7.34 (m, 2H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 13.7, 17.6, 19.7, 19.8, 53.8, 54.0, 64.2, 64.5, 112.7, 116.6, 117.3, 120.5, 127.2, 143.4, 146.7, 166.8, 173.9; LC-IT-TOF/MS: (*m/z*) calcd for $C_{20}H_{28}N_4O_4$ $[M+H]^+$ 389.2111, found 389.2134.

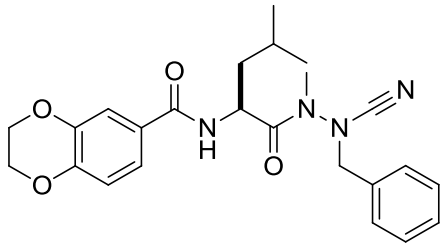
(S)-N-(1-(2-cyano-1,2-dimethylhydrazinyl)-4-methyl-1-oxopentan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1k)



Prepared according to the general procedure using **6-12k** (118 mg, 0.35 mmol), sodium acetate (115 mg, 1.4 mmol), and cyanogen bromide (149 mg, 1.4 mmol) in 10 mL of dry methanol to afford 97 mg (77%) of **6-1k** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.00 (d, *J* = 5.0 Hz, 3H), 1.05 (d, *J* = 5.0 Hz, 3H),

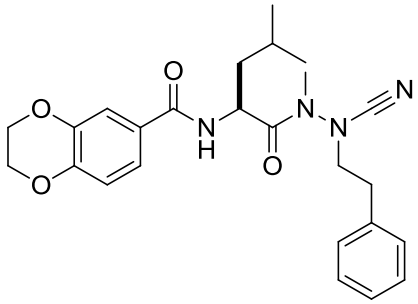
1.56-1.81 (m, 3H), 3.23 (s, 3H), 3.34 (s, 3H), 4.26-4.30 (m, 4H), 5.23-5.28 (m, 1H), 6.58 (d, *J* = 10.0 Hz, 1H), 6.88 (d, *J* = 10.0 Hz, 1H), 7.27-7.32 (m, 2H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 21.5, 32.3, 25.1, 29.7, 30.5, 41.0, 41.3, 48.2, 64.2, 64.5, 113.6, 116.6, 117.3, 120.6, 126.8, 143.4, 146.8, 167.0, 174.7; LC-IT-TOF/MS: (*m/z*) calcd for $C_{18}H_{24}N_4O_4$ $[M+H]^+$ 361.1798, found 361.1808.

(S)-N-(1-(2-benzyl-2-cyano-1-methylhydrazinyl)-4-methyl-1-oxopentan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1l)



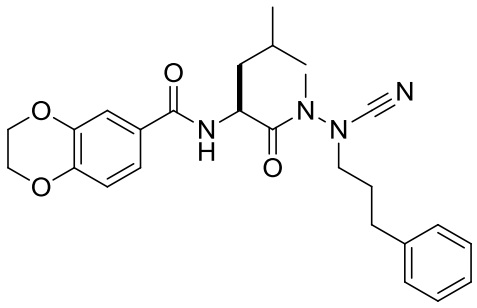
Prepared according to the general procedure using **6-12l** (132 mg, 0.32 mmol), sodium acetate (106 mg, 1.3 mmol), and cyanogen bromide (138 mg, 1.3 mmol) in 10 mL of dry methanol to afford 131 mg (81%) of **6-1l** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 0.96 (t, *J* = 5.0 Hz, 6H), 1.50-1.81 (m, 3H), 3.14 (s, 3H), 4.28 (dd, *J* = 5.0, 15.0 Hz, 4H), 4.63 (s, 2H), 5.41-5.45 (m, 1H), 6.67 (d, *J* = 10.0 Hz, 1H), 6.89 (d, *J* = 10.0 Hz, 1H), 7.31-7.57 (m, 7H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 21.4, 23.3, 25.0, 32.3, 41.3, 48.4, 58.9, 64.2, 64.6, 112.4, 116.7, 117.3, 120.7, 127.0, 127.4, 128.8, 129.3, 129.8, 130.4, 131.5, 143.4, 146.8, 167.1, 175.1; LC-IT-TOF/MS: (*m/z*) calcd for $C_{24}H_{28}N_4O_4$ $[M+H]^+$ 437.2111, found 437.2124.

(S)-N-(1-(2-cyano-1-methyl-2-phenethylhydrazinyl)-4-methyl-1-oxopentan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1m)



Prepared according to the general procedure using **6-12m** (132 mg, 0.31 mg), sodium acetate (98 mg, 1.2 mmol), and cyanogen bromide (127 mg, 1.2 mmol) in 10 mL of dry methanol to afford 110 mg (79%) of **6-1m** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 0.91 (d, *J* = 5.0 Hz, 3H), 0.93 (d, *J* = 5.0 Hz, 3H), 1.48-1.69 (m, 3H), 3.11-3.14 (m, 2H), 3.16 (s, 3H), 3.69-3.74 (m, 1H), 3.81-3.87 (m, 1H), 4.25-4.29 (m, 4H), 5.11-5.15 (m, 1H), 6.58 (d, *J* = 5.0 Hz, 1H), 6.87 (d, *J* = 10.0 Hz, 1H), 7.23-7.35 (m, 7H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 21.5, 23.2, 25.1, 31.6, 33.0, 41.1, 48.3, 54.9, 64.2, 64.5, 112.4, 116.6, 117.2, 120.6, 127.0, 127.2, 128.8, 129.0, 136.1, 143.3, 146.7, 167.0, 175.0; LC-IT-TOF/MS: (*m/z*) calcd for $C_{25}H_{30}N_4O_4$ $[M+H]^+$ 451.2267, found 451.2260.

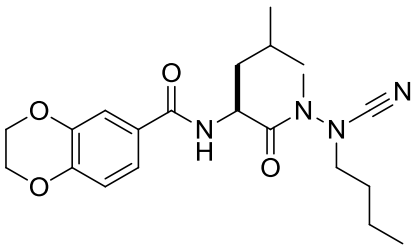
***(S)-N-(1-(2-cyano-1-methyl-2-(3-phenylpropyl)hydrazinyl)-4-methyl-1-oxopentan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1n*)**



The general was followed using **6-12n** (123 mg, 0.28 mmol), sodium acetate (92 mg, 1.1 mmol), and cyanogen bromide (117, 1.1 mmol) in 10 mL of dry methanol to afford 95 mg (73%) of **6-1n** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 0.98 (d, *J* = 5.0 Hz, 3H), 1.03 (d, *J* = 5.0 Hz, 3H), 1.57-1.67 (m, 2H), 1.80-1.81 (m, 1H), 2.11-2.17 (m, 2H), 2.77-2.86 (m, 2H), 3.21 (s, 3H), 3.47-3.56 (m, 2H), 4.25-4.29 (m, 4H), 5.28-5.33 (m, 1H), 6.68 (d, *J* = 10.0 Hz, 1H), 6.88 (d, *J* = 5.0 Hz, 1 H), 7.21-7.32 (m, 7H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 21.5, 23.3, 25.1, 28.2, 31.5, 32.6, 41.2, 48.3, 53.4, 61.1, 64.5, 112.5,

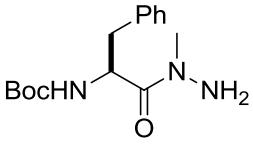
116.6, 117.2, 120.6, 126.4, 127.0, 128.3, 128.6, 140.1, 143.3, 146.7, 166.9, 175.0; LC-IT-TOF/MS: (*m/z*) calcd for $C_{26}H_{32}N_4O_4$ $[M+H]^+$ 465.2424, found 465.2638.

(S)-N-(1-(2-butyl-2-cyano-1-methylhydrazinyl)-4-methyl-1-oxopentan-2-yl)-2,3-dihydrobenzo[b][1,4]dioxine-6-carboxamide (6-1o)



Prepared according to the general procedure using **6-12o** (132 mg, 0.35 mmol), sodium acetate (115 mg, 1.4 mmol), and cyanogen bromide (148 mg, 1.4 mmol) in 10 mL of dry methanol to afford 94 mg (67%) of **6-1o** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 0.98-1.05 (m, 9H), 1.50-1.82 (m, 7H), 3.23 (s, 3H), 3.41-3.53 (m, 2H), 4.26-4.30 (m, 4H), 5.24-5.29 (m, 1H), 6.58 (d, *J* = 10.0 Hz, 1H), 6.88 (d, *J* = 5.0 Hz, 1H), 7.28-7.32 (m, 2H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 13.7, 19.8, 21.4, 23.3, 25.1, 28.7, 31.5, 41.3, 48.3, 53.7, 64.2, 64.5, 112.6, 116.6, 117.2, 120.6, 127.0, 143.4, 146.7, 166.9, 174.9; LC-IT-TOF/MS: (*m/z*) calcd for $C_{21}H_{30}N_4O_4$ $[M+H]^+$ 403.2267, found 403.2314.

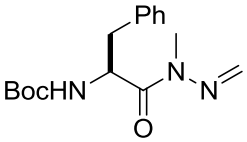
(S)-tert-butyl 1-(1-methyl-2-methylenehydrazinyl)-1-oxo-3-phenylpropan-2-yl) carbamate (6-13)



Prepared according to the general procedure using Boc-Phe-OH (6.63 g, 25.0 mmol), *N*-methylmorpholine (NMM) (3.3 mL, 30.0 mmol), isobutyl chloroformate (ISCF) (3.9 mL, 30.0 mmol) and methylhydrazine sulfate (4.32 g, 30.0 mmol). Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product

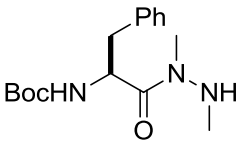
6-13 as a colourless oil (2.71 g, 37%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.38 (s, 9H), 2.93-2.98 (m, 2H), 3.05 (s, 3H), 3.43 (s, 2H), 5.39 (d, *J* = 9.0 Hz, 1H), 5.46-5.53 (m, 1H), 7.19-7.30 (m, 5H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 28.2, 38.3, 39.8, 50.7, 79.1, 126.4, 128.0, 129.2, 155.1, 173.7; LC-IT-TOF/MS: (*m/*z) calcd for $C_{15}H_{23}N_3O_3$ $[M+H]^+$ 294.1739, found 294. 1748.

(S)-tert-butyl (1-(1-methyl-2-methylenehydrazinyl)-1-oxo-3-phenylpropan-2-yl) carbamate (6-14)



Prepared according to the general procedure using **6-13** (0.9 g, 3.0 mmol) and formaldehyde 37% solution (0.5 mL, 6.0 mmol). Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **6-14** as a colourless oil (0.6 g, 64%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.41 (s, 9H), 2.87-2.94 (m, 1H), 3.08-3.15 (m, 1H), 3.21 (s, 3H), 5.29 (d, *J* = 8.2 Hz, 1H), 5.05-5.17 (m, 1H), 6.48 (d, *J* = 10.5 Hz, 1H), 6.66 (d, *J* = 10.5 Hz, 1H), 7.16-7.31 (m, 5H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 26.8, 28.3, 39.1, 52.2, 79.3, 126.5, 128.1, 129.4, 136.8, 155.1, 173.4; LC-IT-TOF/MS: (*m/z*) calcd for $C_{16}H_{23}N_3O_3$ $[M+H]^+$ 306.1739, found 306. 1784.

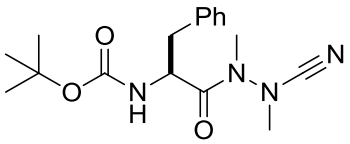
(S)-tert-butyl (1-(1,2-dimethylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-15)



To a cooled (0 ℃) solution of **6-14** (1.53 g, 5.0 mmol) in dry THF (15 mL) was added acetic acid (9 mL), and $NaBH_3CN$ (0.79 g, 12.5 mmol). The reaction was allowed to warm to room

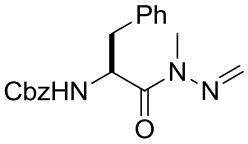
temperature and stirred overnight. After evaporation of solvent, resulting residue was cooled to 0 ^{o}C and saturated $NaHCO_3$ was added. Aqueous solution was extracted with EtOAc (3 × 50 mL). The combined organic layers were washed with brine, dried over Na_2SO_4 and concentrated in vacuo. Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **6-15** as a colourless oil (1.41 g, 92%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.42 (s, 9H), 2.47 (d, *J* = 3.8 Hz, 3H), 2.81 (s, 1H), 2.89-3.01 (m, 2H), 3.06 (s, 3H), 5.27 (d, *J* = 8.7Hz, 1H), 5.40-5.47 (m, 1H), 7.21-7.32 (m, 5H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 28.3, 31.8, 35.4, 39.9, 51.0, 79.3, 126.5, 128.2, 129.5, 137.2, 174.0; LC-IT-TOF/MS: (*m/z*) calcd for $C_{16}H_{25}N_3O_3$ $[M+H]^+$ 308.1896, found 308.1992.

(S)-tert-butyl (1-(2-cyano-1,2-dimethylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-2)



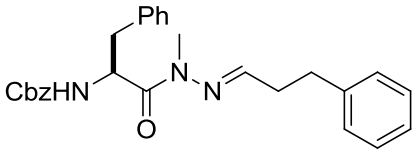
Prepared according to the general procedure using **6-15** (307 mg, 1.0 mmol), anhydrous sodium acetate (230 mg, 3 mmol), and cyanogen bromide (350 mg, 3.3 mmol) in 25 mL of dry MeOH to afford **6-2** as a pale yellow solid (246 mg, 74%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.37 (s, 9H), 2.84-2.31 (m, 2H), 3.20 (s, 3H), 3.24 (s, 3H), 4.99-5.08 (m, 2H), 7.20-7.33 (m, 5H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 28.2, 30.4, 38.5, 41.2, 51.4, 80.0, 113.4, 127.1, 128.6, 129.2, 129.5, 135.6, 155.3, 173.6; LC-IT-TOF/MS: (*m/z*) calcd for $C_{17}H_{24}N_4O_3$ $[M+H]^+$ 333.1848, found 333.1896.

(S)-benzyl (1-(1-methylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-17a)



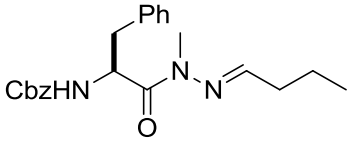
Prepared according to the general procedure using **6-16** (327 mg, 1.0 mmol) and formaldehyde (0.2 mL, 2.0 mmol). Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **6-17a** as a white solid (207 mg, 61%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.91 (d, *J* = 6.0 Hz, 1H), 2.94 (d, *J* = 6.0 Hz, 1H), 3.18 (s, 3H), 5.05 (dd, *J* = 6.0, 12.0 Hz, 2H), 5.51 (d, *J* = 6.0 Hz, 1H), 5.65 (dd, *J* = 6.0, 9.0 Hz, 1H), 6.47 (d, *J* = 6.0 Hz, 1H), 6.65 (d, *J* = 6.0 Hz, 1H), 7.11 (d, *J* = 3.0 Hz, 2H), 7.18-7.35 (m, 8H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 26.9, 39.0, 52.8, 66.7, 126.7, 128.0, 128.2, 128.4, 129.4, 129.7, 136.6, 155.7, 173.0.

(S,E)-benzyl (1-(1-methyl-2-(3-phenylpropylidene)hydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-17d)



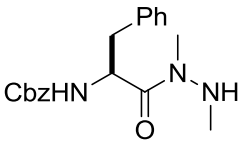
(white solid, 393 mg, 89%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 2.60-2.68 (m, 2H), 2.83-3.03 (m, 4H), 3.14 (s, 3H), 5.05 (dd, *J* = 3.0 Hz, 2H), 5.61 (t, *J* = 9.0 Hz, 1H), 7.01-7.31 (m, 15H).

(S,E)-benzyl (1-(1-methyl-2-pentylidenehydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-17e)



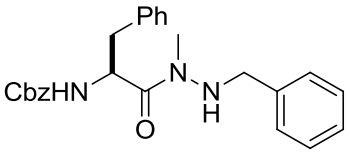
(white solid, 360 mg, 91%), ^{1}H NMR (300 MHz, $CDCl_3$): δ 0.99 (t, *J* = 6.0 Hz, 3H), 1.57-1.64 (m, 4H), 2.27-2.34 (m, 2H), 2.88-3,15 (m, 2H), 3.17 (s, 3H), 5.05 (d, *J* = 3.0 Hz, 2H), 5.55-5.64 (m, 2H), 7.02-7.36 (m, 11H).

(S)-benzyl (1-(1,2-dimethylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-18a)



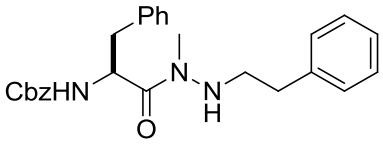
Prepared according to the general procedure using **6-17a** (68 mg, 0.2 mmol), *p*-toluenesulfonic acid (230 mg, 1.2 mmol), and dimethylamine borane (DMAB; 40 mg, 0.64 mmol). The product was directly used in the next reduction step, without further purification. This compound was pure enough to give satisfactory NMR data. ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.46 (s, 3H), 2.91-3.00 (m, 2H), 3.02 (s, 3H), 5.01-5.10 (m, 2H), 5.47 (t, *J* = 3.0 Hz, 1H), 7.16-7.37 (m, 10H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 31.8. 35.4, 35.7, 39.7, 51.5, 66.6, 126.7, 127.9, 128.0, 128.2, 128.5, 128.6, 129.5, 136.6, 136.9, 155.9, 173.7; LC-IT-TOF/MS: (m/z) calcd for $C_{19}H_{23}N_3O_3$ $[M+H]^+$ 342.1739, found 342.1780.

(S)-benzyl (1-(2-benzyl-1-methylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-18b)



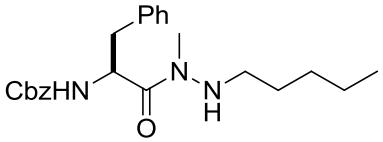
^{1}H NMR (300 MHz, $CDCl_3$): δ 2.91-3.00 (m, 2H), 3.16 (s, 3H), 3.74 (dd, *J* = 6.0, 9.0 Hz, 1H), 3.88 (dd, *J* = 6.0, 9.0 Hz, 1H), 5.07 (dd, *J* = 3.0, 6.0 Hz, 2H), 5.53 (t, *J* = 6.0 Hz, 2H), 7.14-7.33 (m, 15H); LC-IT-TOF/MS: (*m/z*) calcd for $C_{25}H_{27}N_3O_3$ $[M+H]^+$ 418.2052, found 418.1938.

(S)-benzyl
(1-(1-methyl-2-phenethylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-18c)



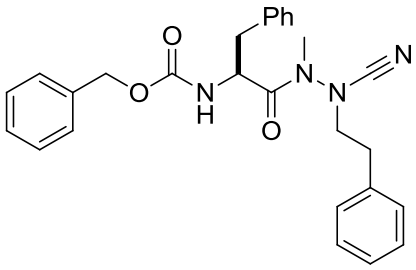
^{1}H NMR (300 MHz, $CDCl_3$): δ 2.71 (br d, *J* = 6.0 Hz, 2H), 2.83 (br d, *J* = 6.0 Hz, 2H), 2.94 (s, 3H), 2.98-3.03 (m, 2H), 4.99-5.10 (m, 2H), 5.42-5.59 (m, 2H), 7.10-7.30 (m, 15H); LC-IT-TOF/MS: (*m/z*) calcd for $C_{26}H_{29}N_3O_3$ $[M+H]^+$ 432.2209, found 432.2276.

(S)-benzyl (1-(1-methyl-2-pentylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-18e)



^{1}H NMR (300 MHz, $CDCl_3$): δ 0.90 (t, *J* = 6.0 Hz, 3H), 1.33-1.35 (m, 4H), 2.59-2.98 (m, 6H), 3.01 (s, 3H), 5.04-5.10 (m, 2H), 5.46-5.51 (m, 2H), 7.15-7.33 (m, 10H); LC-IT-TOF/MS: (*m/z*) calcd for $C_{23}H_{31}N_3O_3$ $[M + H]^+$ 398.2365, found 398.2274.

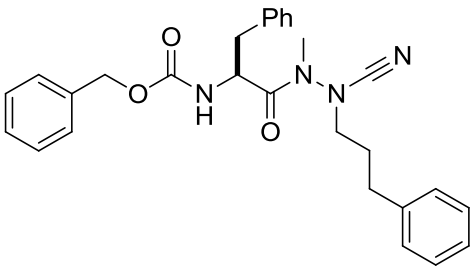
(*S*)-benzyl (1-(2-cyano-1-methyl-2-phenethylhydrazinyl)-1-oxo-3-phenylpropan-2-yl)carbamate (6-3c)



Prepared according to the general procedure using **6-18c** (44 mg, 0.1 mmol), anhydrous sodium acetate (23 mg, 0.3 mmol), and cyanogen bromide (35 mg, 1.1 mmol) in dry MeOH (5 mL) at 40 ℃ for 24 h. Purification by flash column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **6-3c** as a white solid (246 mg, 74%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.68-2.81 (m, 2H), 2.94-2.3.11 (m, 2H), 3.15 (s, 3H), 3.68 (t, *J* = 5.0 Hz, 2H), 4.79 (br d, *J* = 5 Hz, 1H), 4.98-5.14 (m, 2H), 5.23-5.27 (m, 1H), 6.99-7.30 (m, 15H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 31.4, 32.8, 38.0, 51.8, 54.9, 67.0, 112.1, 127.2, 127.9, 128.1, 128.5, 128.7, 128.8, 128.9, 129.1, 129.4, 135.1, 136.1,

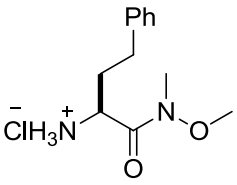
155.9, 173.4; LC-IT-TOF/MS: (*m/z*) calcd for $C_{27}H_{28}N_4O_3$ $[M+H]^+$ 457.2161, found 457.2217.

(S)-benzyl (1-(2-cyano-1-methyl-2-(3-phenylpropyl)hydrazinyl)-1-oxo-3-phenylpropan-2-yl) carbamate (6-3d)



Prepared according to the general procedure using **6-18d** (112 mg, 0.25 mmol), sodium acetate (82 mg, 1.0 mmol), and cyanogen bromide (106 mg, 1.0 mmol) in 10 mL of dry methanol to afford 82 mg (72%) of **6-3d** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.09-2.11 (m, 2H), 2.65-2.90 (m, 4H), 3.08-3.16 (m, 2H), 3.19 (s, 3H), 3.32-3.37 (m, 1H), 3.44-3.49 (m, 1H), 4.98-5.15 (m, 2H), 5.32 (br s, 1H), 7.14-7.34 (m, 15H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 28.1, 31.5, 32.5, 38.4, 52.2, 53.6, 67.0, 112.3, 126.4, 127.3, 127.9, 128.2, 128.3, 128.5, 128.7, 128.8, 129.2, 129.5, 135.4, 139.9, 155.8, 173.4; LC-IT-TOF/MS: (*m/z*) calcd for $C_{28}H_{30}N_4O_3$ $[M+H]^+$ 471.2318, found 471.2365.

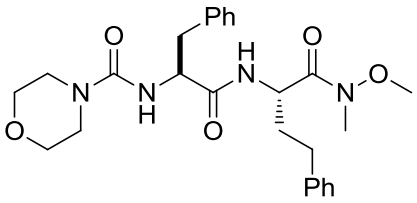
HCl•Hph-N(Me)OMe (6-20)



Prepared according to the general deprotection procedure using Boc-Hph-N(Me)OMe (**5-6**) (1.6 g, 5 mmol) and 4.0 M HCl in 1,4-dioxane (15 mL) in dry THF (10 mL) for 6 h. Dry ether (100 mL) was added with stirring, and the precipitate was filtered off, washed twice with ether, and finally dried in a vacuum desiccator (P_2O_5). This material was pure enough

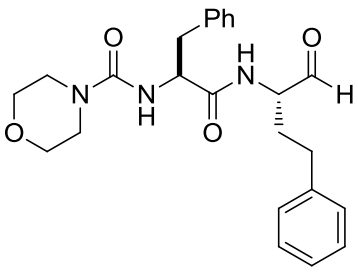
to be used in the next step without further purification.

Mu-Phe-Hph-N(Me)OMe (6-22)



Prepared according to the general coupling reation procedure using **5-19** (0.56 g, 2 mmol), EDC•HCl (0.42 g, 2.2 mmol), HOBt (0.32 g, 2.2 mmol), **6-20** (0.52 g, 2 mmol) and DIPEA (0.51 mL, 3 mmol) in DMF (10 mL) for 12 h, and concentrated in vacuo. The crude product was purified by column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **6-22** as a white solid (0.82 g, 85%). ^{1}H NMR (300 MHz, $CDCl_3$): δ 1.85-2.06 (m, 2H), 2.56-2.64 (m, 2H), 3.08-3.10 (m, 2H), 3.16 (s, 3H), 3.27-3.33 (m, 4H), 3.62-3.66 (m, 7H), 4.66 (dd, *J* = 6.7, 13.7 Hz, 2H), 5.93 (brs, 1H), 5.09 (d, *J* = 7.2 Hz, 1H), 6.83 (d, *J* = 7.9 Hz, 1H), 7.14-7.26 (m, 10H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 31.52, 32.11, 33.99, 38.64, 44.00, 49.03, 55.39, 61.49, 66.41, 126.03, 126.94, 128.36, 128.49, 128.56, 129.43, 136.81, 140.99, 156.93, 171.68.

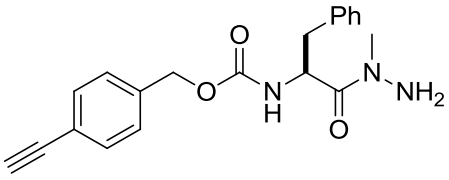
Mu-Phe-Hph-al (6-4)



To a solution of **6-22** (480 mg, 1 mmol) in dry THF (25 mL) at 0 ℃ was added $LiAlH_4$ (57 mg, 1.5 mmol) with vigorous stirring. The mixture was stirred for an additional 20 min at 0 ℃, whereupon cold water was carefully added until effervescence ceased. A cold HCl solution (1 *N*) was added to

break up the gelatinous emulsion until pH 6~7 was reached. Diluted with H_2O (50 mL) and extracted with EtOAc (3 × 25 mL). The combined organic extracts were washed with saturated aqueous $NaHCO_3$ and brine, dried over Na_2SO_4, filtered and concentrated in vacuo. Purification by flash column chromatography on silica gel using 50 to 75% EtOAc in hexanes to give 350 mg (73%) of **6-4** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.79-1.85 (m, 1H), 2.10-2.13 (m, 1H), 2.50-2.60 (m, 2H), 3.11-3.13 (m, 2H), 3.25-3.36 (m, 4H), 3.57-3.62 (m, 4H), 4.25-4.31 (m, 1H), 4.76 (dd, *J* = 10.0, 15.0 Hz, 1H), 5.60 (t, *J* = 10.0 Hz, 1H), 7.11-7.42 (m, 10H), 9.31 (s, 1H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 30.1, 31.2, 38.7, 43.9, 55.7, 58.3, 66.2, 126.3, 127.0, 128.3, 128.48, 128.52, 129.25, 129.31, 136.7, 140.3, 157.1, 172.9, 199.1; LC-IT-TOF/MS: (*m/z*) calcd for $C_{24}H_{29}N_3O_4$ $[M+H]^+$ 424.2158, found 424.2164.

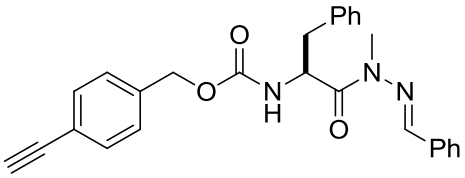
(S)-4-ethynylbenzyl 1-(1-methylhydrazinyl)-1-oxo-3-phenylpropan-2-ylcarbamate (6-24)



Prepared according to the general procedure using **5-23** (1.62 g, 5 mmol), *N*-methylmorpholine (NMM) (0.65 mL, 6 mmol), isobutyl chloroformate (ISCF) (0.78 mL, 6 mmol), and methylhydrazine solution (methylhydrazine sulfate (1.5 g, 10 mmol), H_2O (1 mL) and 5 *N* NaOH (4 mL)) in dry THF (50 mL) for 24 h. Purification by flash column chromatography on silica gel using 30 to 50% EtOAc in hexanes to give the product **6-24** as a white solid (1.15 g, 65%). ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.27-2.34 (m, 2H), 2.89 (dd, *J* = 7.1, 13.9 Hz,

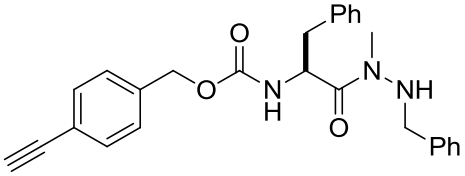
1H), 3.11 (dd, *J* = 5.1, 13.9 Hz, 1H), 3.16 (s, 3H), 5.01 (d, *J* = 12.6 Hz, 1H), 5.05 (d, *J* = 12.3 Hz, 1H), 5.55 (d, *J* = 9.2 Hz, 1H), 5.58-5.64 (m, 1H), 7.02 (t, *J* = 5.2 Hz, 1H), 7.07-7.12 (m, 2H), 7.14-7.25 (m, 3H), 7.25-7.37 (m, 5H).

(S,E)-4-ethynylbenzyl 1-(2-benzylidene-1-methylhydrazinyl)-1-oxo-3-phenylpropan-2-ylcarbamate (6-25)



Prepared according to the general procedure using **6-24** (176 mg, 0.5 mmol), benzaldehyde (0.33 mL, 3.0 mmol), and anhydrous $MgSO_4$ (0.6 g, 5 mmol). Purification by flash column chromatography (silica gel) using 10 to 20% EtOAc in hexanes to give the product **6-25** as a white solid (190 mg, 87%). ^{1}H NMR (300 MHz, DMSO-d_6): δ 3.00-3.07 (m, 1H), 3.16-3.22 (m, 1H), 3.35 (s, 3H), 5.06 (dd, *J* = 12.7, 19.7 Hz, 2H), 5.64 (m, 1H), 5.77-5.84 (m, 1H), 7.12-7.27 (m, 7H), 7.41-7.47 (m, 6H), 7.47-7.68 (m, 3H); ^{13}C NMR (75 MHz, $CDCl_3$): δ 28.85, 39.92, 53.73, 66.78, 78.04, 84.10, 123.40, 127.46, 128.04, 128.37, 128.97, 129.52, 130.05, 132.90, 134.85, 137.20, 138.11, 141.00, 173.40.

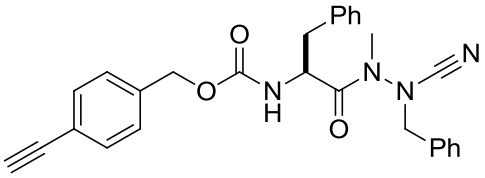
(S)-4-ethynylbenzyl 1-(2-benzyl-1-methylhydrazinyl)-1-oxo-3-phenylpropan-2-ylcarbamate (6-26)



Prepared according to the general procedure using **6-25** (110 mg, 0.25 mmol), *p*-toluenesulfonic acid (285 mg, 1.5 mmol), and dimethylamine borane (DMAB; 48 mg, 0.8

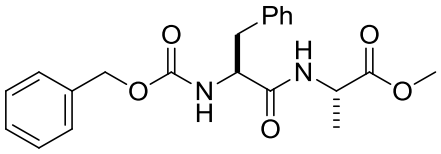
mmol). The product was used in the next step without further purification.

(S)-4-ethynylbenzyl 1-(2-benzyl-2-cyano-1-methylhydrazinyl)-1-oxo-3-phenyl-propan-2-ylcarbamate (6-5)



Prepared according to the general procedure using **6-26** (88 mg, 0.2 mmol), anhydrous sodium acetate (66 mg, 0.8 mmol), and cyanogen bromide (85 mg, 0.8 mmol) in 5 mL of dry MeOH to afford 81 (87%) mg of **6-5** as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$): δ 2.83-2.87 (m, 2H), 3.09 (s, 1H), 3.12 (s, 3H), 4.52 (dd, *J* = 10.0, 15.0 Hz, 2H), 5.02-5.08 (m, 4H), 5.38 (d, *J* = 10.0 Hz, 1H), 7.13-7.48 (m, 14H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 32.3, 38.4, 53.5, 59.0, 66.4, 77.6, 83.2, 112.2, 127.3, 127.7, 128.7, 129.2, 129.3, 129.5, 129.7, 129.8, 130.2, 131.2, 132.2, 135.4, 136.8, 155.8, 173.4; LC-IT-TOF/MS: (*m/z*) calcd for $C_{28}H_{26}N_4O_3$ $[M+H]^+$ 467.2005, found 467.2104.

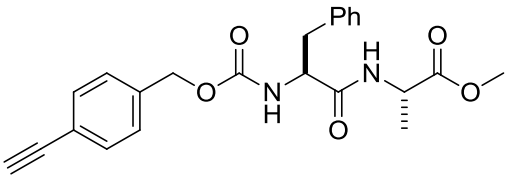
(S)-methyl 2-((S)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido) propanoate (6-28a)



Prepared according to the general procedure using Cbz-Phe-OH (0.75 g, 2.5 mmol), EDC•HCl (0.55 g, 3.0 mmol), HOBt (0.4 g, 3.0 mmol), **6-27** (0.35 g, 2.5 mmol) and DIPEA (0.85 mL, 5 mmol) in DMF (10 mL) for 12 h. The crude product was purified by column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **6-28a** as a white solid (0.87 g, 91%). ^{1}H NMR

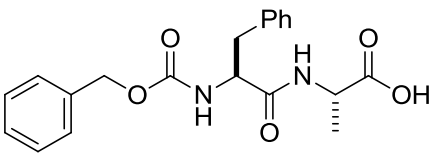
(500 MHz, $CDCl_3$): δ 1.33 (d, *J* = 5.0 Hz, 3H), 3.03-3.13 (m, 2H), 3.71 (s, 3H), 4.43-4.53 (m, 2H), 5.08 (s, 2H), 5.34 (br s, 1H), 6.36 (br s, 1H), 7.18-7.36 (m, 10H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 18.3, 38.5, 48.2, 52.4, 56.0, 67.1, 127.1, 128.0, 128.2, 128.5, 128.7, 129.3, 136.2, 170.3, 172.8.

(S)-methyl 2-((S)-2-((((4-ethynylbenzyl)oxy)carbonyl)amino)-3-phenyl-propanamido)propanoate (6-28b)



Prepared according to the general procedure using **5-23** (0.8 g, 2.5 mmol), EDC•HCl (0.55 g, 3.0 mmol), HOBt (0.4 g, 3.0 mmol), **6-27** (0.35 g, 2.5 mmol) and DIPEA (0.85 mL, 5 mmol) in DMF (10 mL) for 12 h. The crude product was purified by column chromatography on silica gel using 20 to 50% EtOAc in hexanes to give the product **6-28b** as a white solid (0.95 g, 93%). ^1H NMR (500 MHz, $CDCl_3$): δ 1.33 (d, *J* = 5.0 Hz, 3H), 3.06-3.08 (m, 3H), 3.71 (s, 3H), 4.44-4.53 (m, 2H), 5.06 (m, 2H), 5.45 (br s, 1H), 6.43 (br s, 1H), 7.17-7.29 (m, 7H), 7.46 (d, *J* = 5.0 Hz, 2H); ^{13}C NMR (126 MHz, $CDCl_3$): δ 18.3, 38.6, 48.1, 52.4, 56.0, 66.4, 77.5, 83.3, 121.9, 127.1, 127.7, 128.6, 129.3, 132.2, 136.2, 170.4, 172.8.

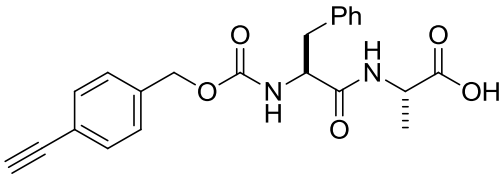
(S)-2-((S)-2-(((benzyloxy)carbonyl)amino)-3-phenylpropanamido)propanoic acid (6-29a)



Prepared according to the general procedure using **6-28a** (0.7 g, 1.8 mmol), LiOH•H_2O (0.23 g, 5.4 mmol) in THF/H_2O (3/1, 12 mL) for 2 h. The

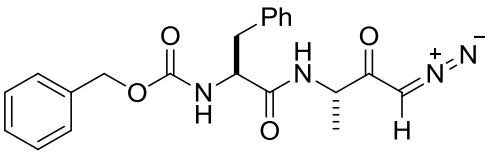
product **6-29a** was pure enough to be used in the next step without further purification. ^{1}H NMR (500 MHz, DMSO-d_6): δ 1.31 (d, J = 10.0 Hz, 3H), 2.72 (dd, J = 10.0, 15.0 Hz, 1H), 3.02 (dd, J = 10.0, 15.0 Hz, 1H), 4.21-4.31 (m, 2H), 4.93 (s, 2H), 7.16-7.34 (m, 9H), 8.33 (d, J = 5.0 Hz, 1H), 12.53 (s, 1H); ^{13}C NMR (126 MHz, DMSO-d_6): δ 17.1, 37.4, 47.5, 55.8, 65.1, 126.2, 127.4, 127.6, 128.0, 129.2, 137.0, 138.1, 155.8, 171.4, 174.0.

(S)-2-((S)-2-((((4-ethynylbenzyl)oxy)carbonyl)amino)-3-phenylpropanamido)propanoic acid (6-29b)



Prepared according to the general procedure using **6-28b** (0.6 g, 1.5 mmol), LiOH•H_2O (0.19 g, 4.5 mmol) in THF/H_2O (3/1, 12 mL) for 2 h. The product **6-29b** was pure enough to be used in the next step without further purification. ^{1}H NMR (500 MHz, DMSO-d_6): δ 1.31 (d, J = 10.0 Hz, 3H), 2.72 (t, J = 10.0 Hz, 1H), 3.02 (t, J = 10.0 Hz, 1H), 4.17 (s, 1H), 4.22-4.30 (m, 2H), 4.94 (s, 2H), 7.19-7.31 (m, 6H), 7.43 (d, J = 5.0 Hz, 2H), 7.51 (d, J = 5.0 Hz, 1H), 8.33 (d, J = 5.0 Hz, 1H), 12.56 (s, 1H); ^{13}C NMR (126 MHz, DMSO-d_6): δ 17.1, 37.4, 47.5, 55.9, 64.6, 80.8, 83.2, 126.2, 127.4, 128.0, 129.2, 131.6, 138.0, 155.7, 171.3. 174.0.

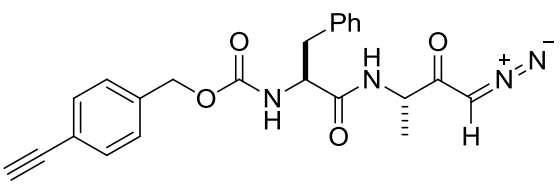
Cbz-Phe-Ala-CHN$_2$ (6-6a)



To a stirred solution of **6-29a** (0.37 g, 1.0 mmol) in dry THF (10 mL) cooled at -25 ^{o}C

was added dropwise *N*-methylmorpholine (NMM, 0.15 mL, 1.2 mmol), and isobutyl chloroformate (ISCF, 0.17 mL, 1.2 mmol) via syringe. The reaction was stirred for 1 h at -15 ^{o}C and added of a solution of ~0.25 mmol/mL diazomethane in ether at -78 ^{o}C (15 mL). The reaction was stirred for 4 h until to 0 ^{o}C. After evaporation of the solvent, the aqueous residue was extracted with EtOAc (3 × 20 mL). The combined organic extracts were washed with H_2O, saturated $NaHCO_3$, brine, dried over Na_2SO_4 and concentrated in vacuo. Purification by flash column chromatography on silica gel using 25 to 50% EtOAc in hexanes to give the product **6-6a** as a white solid (0.35 g, 91%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 1.33 (d, *J* = 5.0 Hz, 3H), 2.78 (dd, *J* = 10.0, 15.0 Hz, 1H), 3.03 (dd, *J* = 10.0, 15.0 Hz, 1H), 4.26-4.32 (m, 2H), 4.95 (s, 2H), 5.79 (s, 1H), 7.19-7.34 (m, 9H), 7.56 (d, *J* = 10.0 Hz, 1H), 8.45 (d, *J* = 5.0 Hz, 1H); ^{13}C NMR (126 MHz, DMSO-d_6): δ 16.9, 37.4, 52.2, 55.9, 65.2, 126.2, 127.4, 127.6, 128.0, 128.2, 129.3, 137.0, 137.9, 155.8, 171.3, 195.1; LC-IT-TOF/MS: (*m/z*) calcd for $C_{21}H_{22}N_4O_4$ $[M+H]^+$ 395.1641, found 395.1588.

4-ethynylbenzyl ((S)-1-(((S)-4-diazo-3-oxobutan-2-yl)amino)-1-oxo-3-phenyl-propan-2-yl)carbamate (6-6b, alk-DMK)



Prepared according to the general procedure using **6-29b** (0.39 g, 1.0 mmol), NMM (0.15 mL, 1.2 mmol), ISCF (0.17 mL, 1.2 mmol), and the diazomethane-ether solution. Purification by flash column chromatography on silica gel using 25 to 50% EtOAc in hexanes to give the product **6-6b** as a white solid (0.37 g, 89%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 1.22

(d, *J* = 10.0 Hz, 3H), 2.78 (dd, *J* = 5.0, 105.0 Hz, 1H), 3.03 (dd, *J* = 5.0, 10.0 Hz, 1H), 4.18 (s, 1H), 4.26-4.32 (m, 2H), 4.96 (s, 2H), 5.80 (s, 1H), 7.16-7.32 (m, 6H), 7.44 (d, *J* = 5.0 Hz, 2H), 7.62 (d, *J* = 10.0 Hz, 1H), 8.45 (d, *J* = 5.0 Hz, 1H); ^{13}C NMR (126 MHz, DMSO-d_6): δ 16.9, 37.4, 52.2, 55.9, 64.7, 80.8, 83.2, 120.9, 126.2, 127.5, 128.0, 129.3, 131.6, 137.9, 138.0, 155.7, 171.3, 195.1; LC-IT-TOF/MS: (*m/z*) calcd for $C_{23}H_{22}N_4O_4$ $[M+H]^+$ 419.1641, found 419.1760.

8.6.2 Cell Biological Assays

Cell-culture conditions, Guava ViaCount assay, in situ proteomic profiling, pull-down and mass spectrometry identification, and fluorescence microscopy were carried out using the general protocols as described in **7.1.2.**

8.6.2.1 Enzyme Assays

Cruzain and rhodesain were assayed in freshly prepared 100 mM acetate buffer pH 5.5, containing 5 mM dithiothreitol (DTT) and 0.001% Triton X-100. 25 μL of reagents (cruzain, or rhodesain at 4 nM final concentration) was added to a 384-well black plate (Greiner, Germany) that contained 1 μL of test compound (concentration ranging from 10 μM to 10 pM). The enzyme-compound mixture was incubated for 5 min at room temperature. Then 25 μL of substrate Z-FR-AMC (Bachem, 10 μM final concentration, in the same buffer solution as above) was added to the enzyme-compound mixture to initiate the reaction. The rate of increase in fluorescence, resulting from the proteolytic cleavage of the substrate leading to the

release of fluorogenic AMC was monitoring as with an automated microtiter plate spectrofluorimeter (BioTek Synergy 4 Fluorescence Plate Reader) with fluorescence readout setting of excitation at 355 nM and emission at 460 nM. Controls were performed using enzyme alone, enzyme with DMSO vehicle, enzyme in the presence of the previously known, highly effective irreversible inhibitor (Mu-Phe-HphVSPh, K11002), and reversible inhibitor (Mu-Phe-Hph-al, **6-4**). IC_{50} values were determined by sigmoid dose-response curve fitting with GraphPad Prism software using inhibitor concentrations in the linear portion of a plot of inhibition versus log[I] (seven concentrations tested with at least two in the linear range). Data are the means of at least two independent assays.

8.6.2.2 Western Blotting

Pull-down samples from in situ labeling with **6-5**, or alk-DMK were separated on 12% SDS-PAGE gel together with pull-down sample from DMSO-treated (negative control). After SDS-PAGE gel separation, proteins were then transferred to a PVDF membrane and subsequently blocked with 3% (w/v) BSA/PBST overnight at 4^{o}C. Membranes were incubated for 1 h at room temperature with the respective antibodies (anti-cathepsin L for HepG2, anti-rhodesainΔC for *T. brucei*), and washed with PBST (3 × 15 min with gentle agitation), then followed by incubation with a anti-mouse-IgG conjugated secondary antibody in the blocking buffer mentioned above. After wash with PBST (3 × 15 min with gentle agitation), the SuperSignal West Pico kit (Pierce) was used to develop the blot.

Chapter 9

References

[1] a) G. C. Terstappen, C. Schlupen, R. Raggiaschi, G. Gaviraghi, *Nat. Rev. Drug Discov.* **2007**, *6*, 891–903; b) J. N. Chan, C. Nislow, A. Emili, *Trends Pharmacol. Sci.* **2010**, *31*, 82–88; c) U. Rix, G. Superti-Furga, *Nat. Chem. Biol.* **2009**, *5*, 616–624; d) C. H. Ho, J. Piotrowski, S. J. Dixon, A. Baryshnikova, M. Costanzo, C. Boone, *Curr. Opin. Chem. Biol.* **2011**, *15*, 66–78; d) C. C. Wang, K. W. Cheng, Q.-Y. He, F. Chen, *Proteomics Clin. Appl.* **2008**, *2*, 338–354; e) B. Lomenick, R. W. Olsen, J. Huang, *ACS Chem. Biol.* **2011**, 6, 34–46.

[2] a) B. J. Leslie, P. J. Hergenrother, *Chem. Soc. Rev.* **2008**, *37*, 1347–1360; b) S. Sato, A. Murata, T. Shirakawa, M.Uesugi, *Chem. Biol.* **2010**, *17*, 616–23; c) K. W. Cheng, C. C. Wang, M. Wang, Q.-Y. He, F. Chen, *Mass Spectrom. Rev.* **2010**, *29*, 126–155.

[3] a) J. Taunton, C. A. Hassig and S. L. Schreiber, *Science* **1996**, *272*, 408–411; b) M. W. Harding, A. Galat, D. E. Uehling, S. L. Schreiber, *Nature* **1989**, *341*, 758–760; c) S. Bach, M. Knockaert, J. Reinhardt, O. Lozach, S. Schmitt, B. Baratte, M. Koken, S. P. Coburn, L. Tang, T. Jiang, D.-C. Liang, H. Galons, J.-F. Dierick, L. A. Pinna, F. Meggio, F. Totzke, C. Schächtele, A. S. Lerman, A. Carnero, Y. Wan, N. Gray, L. Meijer, *J. Biol. Chem.* **2005**, *280*, 31208–31219; d) R. R. Falsey, M. T. Marron, G. M. K. B. Gunaherath, N. Shirahatti, D. Mahadevan, A. A. L. Gunatilaka, L. Whitesell, *Nat. Chem. Biol.* **2006**, *2*, 33–38; e) G. Wang, L. Shang, A. W. G. Burgett, P. G. Harran, X. Wang, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 2068–2073; f) S. Zhu, H. Wurdak, J. Wang, C. A. Lyssiotis, E. C. Peters, C. Y. Cho, X. Wu, P. G. Schultz, *Cell Stem Cell* **2009**, *4*, 416–26; g) Q. Zhang, M. B.

Major, S. Takanashi, N. D. Camp, N. Nishiya, E. C. Peters, M. H. Ginsberg, P. G. Schultz, R. T. Moon, S. Ding, Proc. *Natl. Acad. Sci. USA* **2007**, *104*, 7444–7448.; h) T. Ito, H. Ando, T. Suzuki, T. Ogura, K. Hotta, Y. Imamura, Y. Yamaguchi, H. Handa, *Science* **2010**, *327*, 1345–1350; h) W. K. Low, Y. Dang, T. Schneider-Poetsch, Z. Shi, N. S. Choi, W. C. Merrick, D. Romo, J. O. Liu, *Mol. Cell*. **2005**, *20*, 709–22; i) B. H. B. Kwok, B. Koh, M. I. Ndubuisi, M. Elofsson, C. M. Crews, *Chem. Biol*. **2001**, *8*, 759–766; j) K. H. Emami, C. Nguyen, H. Ma, D. H. Kim, K. W. Jeong, M. Eguchi, R. T. Moon, J.-L. Teo, S. W. Oh, H. Y. Kim, S. H. Moon, J. R. Ha, M. Kahn, *Proc. Natl. Acad. Sci.USA* **2004**, *101*, 12682–12687; k) E. C. Griffith, Z. Su, B. E. Turk, S. Chen, Y. H. Chang, Z. Wu, K. Biemann, J. O. Liu, *Chem. Biol*. **1997**, *4*, 461–471; l) Y. Wan, W. Hur, C. Y. Cho, Y. Liu, F. J. Adrian, O. Lozach, S. Bach, T. Mayer, D. Fabbro, L. Meijer, N. S. Gray, *Chem. Biol.* **2004**, *11*, 247–259.

[4] M. J. Evans, A. Saghatelian, E. J. Sorenson, B. F. Cravatt, *Nat. Biotechnol.* **2005**, *23*, 1303–1307.

[5] a) M. J. Evans, B. F. Cravatt, *Chem. Rev.* **2006**, *106*, 3279–3301; b) M. Fonović, M. Bogyo, *Curr. Pharm. Des.* **2007**, *13*, 253–261; c) M. Uttamchandani, J. Li. H. Sun, S. Q. Yao, *ChemBioChem* **2008**, *9*, 667–675; d) Fonović, M.; Bogyo, M. *Expert Rev. Proteomics* **2008**, *5*, 721–730; e) B. F. Cravatt, A. T. Wright, J. W. Kozarich, *Annu. Rev. Biochem.* **2008**, *77*, 383–414.

[6] a) E. S. Lander, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, et al., *Nature* **2001**, *409*, 860–921; b) J. C. Venter, M. D. Adams, E. W. Myers, P. W.

Li, R. J. Mural, G. G. Sutton, et al., *Science* **2001**, *291*, 1304–1351.

[7] A. Saghatelian, B. F. Cravatt, *Nat. Chem. Biol.* **2005**, *1*, 130–142.

[8] S. P. Gygi, G. L.Corthals, Y. Zhang, Y. Rochon, R. Aebersold, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 9390–9395.

[9] a) S. P. Gygi, B. Rist, S. A. Gerber, F. Turecek, M. H. Gelb, R. Aebersold, *Nat. Biotechnol.* **1999**, *17*, 994–999; b) P. L. Ross, Y. N. Huang, J. N. Marchese, B. Williamson, K. Parker, S. Hattan, N. Khainovski, S. Pillai, S. Dey, S. Daniels, S. Purkayastha, P. Juhasz, S. Martin, M. Bartlet-Jones, F. He, A. Jacobson, D. J. Pappin, *Mol. Cell Proteomics* **2004**, *3*, 1154–1169; c) S. E. Ong, B. Blagoev, I. Kratchmarova, D. B. Kristensen, H. Steen, A. Pandey, M. Mann, *Mol. Cell. Proteomics* 2002, *1*, 376–386.

[10] B. Kobe, B. E. Kemp, *Nature* **1999**, *402*, 373–376.

[11] J. C. Powers, J. L. Asgian, O. D. Ekici, K. E. James, *Chem. Rev.* **2002**, *102*, 4639–4750.

[12] a) Y. Liu, M. P. Patricelli, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 14694–14699; b) M. G. Paulick, M. Bogyo, *ACS Chem. Biol.* **2011**, *6*, 563–572; c) O. Obianyo, C. P. Causey, J. E. Jones, P. R. Thompson, *ACS Chem. Biol.* **2011**, DOI: 10.1021/cb2001473; d) D. Kato, K. M. Boatright, A. B. Berger, T. Nazif, G. Blum, C. Ryan, K. A. H. Chehade, G. S. Salvesen, M. Bogyo, *Nat. Chem. Biol.* **2005**, *1*, 33–38; e) K. A. Kalesh, L. P. Tan, K. Lu, L. Gao, J. Wang, S. Q. Yao, *Chem. Commun.* **2010**, *46*, 589–591; f) M. P. Patricelli, A. K. Szardenings, M. Liyanage, T. K. Nomanbhoy, M. Wu, H. Weissig, A. Aban, D. Chun, S. Tanner, J.

W. Kozarich, *Biochemistry* **2007**, *46*, 350–358; g) M. Verdoes , B. I. Florea , U. Hillaert , L. I. Willems , W. A. van der Linden , M. Sae-Heng , D. V. Filippov , A. F. Kisselev , G. A. van der Marel , H. S. Overkleeft, *ChemBioChem* **2008**, *9*, 1735 –1738; h) D. Rotili, M. Altun, A. Kawamura, A.Wolf, R. Fischer, I. K. Leung, M. M. Mackeen, Y. M. Tian, P. J. Ratcliffe, A. Mai, B. M. Kessler, C. J. Schofield, *Chem. Biol*. **2011**, *18*, 642–654; i) E. W. Chan, S. Chattopadhaya, R. C. Panicker, X. Huang, S. Q. Yao, *J. Am. Chem. Soc.* **2004**, *126*, 14435–14446; j) Y. M. Li, M. Xu, M. T. Lai, Q. Huang, J. L. Castro, J. DiMuzio-Mower, T. Harrison, C. Lellis, A. Nadin, J. G. Neduvelil, R. B. Register, M. K. Sardana, M. S. Shearman, A. L. Smith, X. P. Shi, K. C. Yin, J. A. Shafer, S. J. Gardell, *Nature* **2000**, *405*, 689–694.

[13] a) L. C. Lo, T. L. Pang, C. H. Kuo, Y. L. Chiang, H. Y. Wang, J. J. Lin, *J. Proteome Res.* **2002**, *1*, 35–40; b) R. Srinivasan, X. Huang, S. L. Ng, S. Q. Yao, *ChemBioChem.* **2006**, *7*, 32–36; c) C. P. Lu, C. T. Ren, S. H. Wu, C. Y. Chu, L. C. Lo, *ChemBioChem* **2007**, *8*, 2187–2190; d) T. Komatsu, K. Kikuchi, H. Takakusa, K. Hanaoka, T. Ueno, M. Kamiya, Y. Urano, T. Nagano, *J. Am. Chem. Soc.* **2006**, *128*, 15946–15947.

[14] a) S. Tsukiji, M. Miyagawa, Y. Takaoka, T. Tamura, I. Hamachi, *Nat. Chem. Biol.* **2009**, *5*, 341–343; b) S. Tsukiji, H. Wang, M. Miyagawa, T. Tamura, Y. Takaoka, I. Hamachi, *J. Am. Chem. Soc.* **2009**, *131*, 9046–9054.

[15] E. Saxon, C. R. Bertozzi, *Sicence* **2000**, *287*, 2007–2010.

[16] a) S. K. Mamidyala, M. G. Finn, *Chem. Soc. Rev.* **2010**, *39*, 1252–1261; b) E. M.

Sletten, C. R. Bertozzi, *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998; e) M. Meldal, C. W. Tornøe *Chem. Rev.* **2008**, *108*, 2952–3015; d) K. A. Kalesh, H. Shi, J. Ge, S. Q. Yao, *Org. Biomol. Chem.* **2010**, *8*, 1749–1762; e) K. A. Kalesh, P. -Y. Yang, R. Srinivasan, S. Q. Yao, *QSAR Comb. Sci.* **2007**, *26*, 1135–1144.

[17] a) N. K. Devaray, R. Weissleder, S. A. Hilderbrand, *Bioconjugate Chem.* **2008**, *19*, 2297–2299; b) N. K. Devaray, R. Upadhyay, J. B. Haun, S. A. Hilderbrand, R. Weissleder, *Angew. Chem. Int. Ed.* **2009**, *48*, 7013–7016; c) N. K. Devaray, S. Hilderbrand, R. Upadhyay, R. Mazitschek, R .Weissleder, *Angew. Chem. Int. Ed.* **2010**, *49*, 2869–2872.

[18] a) S. E. Tully, B. F. Cravatt, *J. Am. Chem. Soc.* **2010**, *132*, 3264–3265; b) D. A. Bachovchin, J. T. Mohr, A. E. Speers, C. Wang, J. M. Berlin, T. P. Spicer, V. Fernandez-Vega, P. Chase, P. S. Hodder, S. C. Schürer, D. K. Nomura, H. Rosen, G. C. Fu, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6811–6816; c) L. Dafik, C. Khosla, *Chem. Biol.* **2011**, 18, 58–66; d) Y. Wang, S. Hu, W. Fast, *J. Am. Chem. Soc.* **2009**, *131*, 15096–15097; e) Y. Luo, B. Knuckley, M. Bhatia, P. J. Pellechia, P. R. Thompson, *J. Am. Chem. Soc.* **2006**, *128*, 14468–14469; f) A. T. Wright, J. D. Song, B. F. Cravatt, *J. Am. Chem. Soc.* **2009**, *131*, 10692–10700; g) K. D. Park, D. Kim, O. Reamtong, C. Eyers, S. J. Gaskell, R. Liu, H. Kohn, *J. Am. Chem. Soc.* **2011**, *131*, 11320–11330; h) M. S. Cohen, H. Hadjivassiliou, J. Taunton, *Nat. Chem. Biol.* **2007**, *3*, 156–160; i) C. M. Salisbury, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1171–1176; j) C. M. Salisbury, B. F. Cravatt, *J. Am. Chem. Soc.* **2008**, *130*, 2184–2194; k) W. W. Qiu, J. Xu, J. Y. Li, J.

Li, F. J. Nan, *Chembiochem* **2007**, *8*, 1351–1358; l) H. Jiang, J. H. Kim, K. M. Frizzell, W. L. Kraus, H. Lin, *J. Am. Chem. Soc.* **2010**, *132*, 9363–9372.

[19] K. T. Barglow, B. F. Cravatt, *Chem. Biol.* **2004**, *11*, 1523–1531.

[20] M. J. Evans, A. Saghatelian, E. J. Sorensen, B. F. Cravatt, *Nat. Biotechnol.* **2005**, *23*, 1303–1307.

[21] R. Singaravelu, D. R. Blais, C. S. McKay, J. P. Pezacki, *Proteome Sci.* **2010**, *8*, 5.

[22] a) T. Böttcher, S. A. Sieber, *Angew. Chem. Int. Ed.* **2008**, *47*, 4600–4603; b) T. Böttcher, S. A. Sieber, *J. Am. Chem. Soc.* 2008, *130*, 14400–14401; c) T. Böttcher, S. A. Sieber, *ChemBioChem* **2009**, *10*, 663–666; d) T. Böttcher, S. A. Sieber, *ChemMedChem* **2009**, *4*, 1260–1263.

[23] a) C. A. Gartner, J. E. Elias, C. E. Bakalarski, S. P. Gygi, *J. Proteome Res.* **2007**, *6*, 1482–1491; b) S. H. L. Verhelst, M. Fonovic, M. Bogyo, *Angew. Chem. Int. Ed.* **2007**, *46*, 1284–1286; c) A. Dirksen, S. Yegneswaran, P. E. Dawson, *Angew. Chem. Int. Ed.* **2010**, *49*, 2023–2027; d) A. E. Speers, B. F. Cravatt, *J. Am. Chem. Soc.* **2005**, *127*, 10018–10019; e) Y. Y. Yang, M. Grammel, A. S. Raghavan, G. Charron, H. C. Hang, *Chem. Biol.* **2010**, *17*, 1212–1222; f) J. Szychowski, A. Mahdavi, J. J. Hodas, J. D. Bagert, J. T. Ngo, P. Landgraf, D. C. Dieterich, E. M. Schuman, D. A. Tirrell, *J. Am. Chem. Soc.* **2010**, *132*, 18351–18360.

[24] a) K. C. Nicolaou, J. S. Chen, D. J. Edmonds, A. A. Estrada, *Angew. Chem. Int. Ed.* **2009**, *48*, 660–719; b) G. M. Cragg, P. G. Grothaus, D. J. Newman, *Chem. Rev.* **2009**, *109*, 3012–3043.

[25] a) Y. Liu, K. R. Shreder, W. Gai, S. Corral, D. K. Ferris, J. S. Rosenblum, *Chem. Biol.* **2005**, *12*, 99–107; b) Y. Liu, N. Jiang, J. Wu, W. Dai, J. S. Rosenblum, *J. Biol. Chem.* **2007**, *282*, 2505–2511.

[26] a) G. C. Adam, C. D. Vanderwal, E. J. Sorensen, B. F. Cravatt, *Angew. Chem. Int. Ed.* **2003**, *42*, 5480–5484; b) R. M. Buey, E. Calvo, I. Barasoain, O. Pineda, M. C. Edler, R. Matesanz, G. Cerezo, C. D. Vanderwal, B. W. Day, E. J. Sorensen, J. A. López, J. M. Andreu, E. Hamel, J. F. Díaz, *Nat. Chem. Biol.* **2007**, *3*, 117–125.

[27] T. Böttcher, S. A. Sieber, *J. Am. Chem. Soc.* **2010**, *132*, 6964–6972.

[28] A. L. MacKinnon, J. L. Garrison, R. S. Hegde, J. Taunton, *J. Am. Chem. Soc.* **2007**, *129*, 14560–14561.

[29] J. Eirich, R. Orth, S. A. Sieber, *J. Am. Chem. Soc.* **2011**, *133*, 12144–12153.

[30] a) I. Staub, S. A. Sieber, *J. Am. Chem. Soc.* **2008**, *130*, 13400–13409; b) I. Staub, S. A. Sieber, *J. Am. Chem. Soc.* **2009**, *131*, 6271–6276.

[31] a) D. Greenbaum, A. Baruch, L. Hayrapetian, Z. Darula, A. Burlingame, K. F. Medzihradszky, M. Bogyo, *Mol. Cell. Proteomics* **2002**, *1*, 60–68; b) G. Blum, G. von Degenfeld, M. J. Merchant, H. M. Blau, M. Bogyo, *Nat. Chem. Biol.* **2007**, *3*, 668– 677; c) A. Watzke, G. Kosec, M. Kindermann, V. Jeske, H. P. Nestler, V. Turk, B. Turk, K. U. Wendt, *Angew. Chem. Int. Ed.* **2008**, *47*, 406–409.

[32] M. Verdoes, B. I. Florea, V. Menendez-Benito, C. J. Maynard, M. D. Witte, W. A. van der Linden, et al., *Chem. Biol.* **2006**, *13*, 1217–1226.

[33] H. C. Hang, J. Loureiro, E. Spooner, A. W. van der Velden, Y. M. Kim, A. M. Pollington, R. Maehr, M. N. Starnbach, H. L. Ploegh, *ACS Chem. Biol.* **2006**, *1*,

713–723.

[34] L. E. Edgington, A. B. Berger, G. Blum, V. E. Albrow, M. G. Paulick, N. Lineberry, et al., *Nat. Med.* **2009**, *15*, 967–973.

[35] a) G. Blum, S. R. Mullins, K. Keren, M. Fonovic, C. Jedeszko, M. J. Rice, B. F. Sloane, M. Bogyo, *Nat. Chem. Biol.* **2005**, *1*, 203–209; b) G. Blum, G. von Degenfeld, M. J. Merchant, H. M. Blau, M. Bogyo, *Nat. Chem. Biol.* **2007**, *3*, 668–677.

[36] M. Hu, L. Li, H. Wu, Y. Su, P.-Y. Yang, M. Uttamchandani, Q.-H. Xu, S.Q. Yao, *J. Am. Chem. Soc.* **2008**, *133*, 12009–12020.

[37] a) K. M. Giacomini, R. M. Krauss, D. M. Roden, M. Eichelbaum, M. R. Hayden, Y. Nakamura, *Nature* **2007**, *446*, 975–977; b) M. Campillos, M. Kuhn, A. -C. Gavin, L. J. Jensen, P. Bork, *Science* **2008**, *321*, 263–266; c) S. Crunkhorn, *Nat. Rev. Drug Discov.* **2008**, *7*, 729–729.

[38] R. Guerciolini, *Int. J. Obes. Relat. Metab. Disord.* **1997**, *21 Suppl 3*, S12-23.

[39] a) S. J. Kridel, F. Axelrod, N. Rozenkrantz, J. W. Smith, *Cancer Res.* **2004**, *64*, 2070–2075; b) C. W. Pemble, L. C. Johnson, S. J. Kridel, W. T. Lowther, *Nat. Struct. Mol. Biol.* **2007**, *14*, 704–709; c) J. L. Little, F. B. Wheeler, D. R. Fels, C. Koumenis, S. J. Kridel, *Cancer Res.* **2007**, *67*, 1262–1269; d) L. M. Knowles, F. Axelrod, C. D. Browne, J. W. Smith, *J. Biol. Chem.* **2004**, *279*, 30540–30545; e) L. M. Knowel, C. Yang, A. Osterman, J. W. Smith, *J. Biol. Chem.* **2008**, *283*, 31378–31384; f) J. A. Menendez, L. Vellon, R. Lupu, *Ann. Oncol.* **2005**, *16*, 1253–1267; g) J. L. Little, F. B. Wheeler, C. Koumenis, S. J. Kridel, *Mol. Cancer*

Ther. **2008**, *7*, 3816–3824; h) S. J. Kridel, W. T. Lowther, C. W. Pemble, *Expert Opin. Investig. Drugs* **2007**, *16*, 1817–1829; i) J. A. Menendez, L. Vellon, R. Lupu, *Exp. Biol. Med.* **2005**, *230*, 151–154; j) M. A. Carvalho, K. G. Zecchin, F. Seguin, D. C. Bastos, M. Agostini, A. L. C. A. Rangel, S. S. Veiga, H. F. Raposo, H. C. F. Oliveira, M. Loda, R. D. Coletta, E. Graner, *Int. J. Cancer* **2008**, *123*, 2557–2565; k) J. A. Menendez, L. Vellon, R. Lupu, *Int. J. Gynecol Cancer* **2006**, *16*, 219–221; l) C. D. Browne, E. J. Hindmarsh, J. W. Smith, *FASEB J.* **2006**, *20*, 2027–2035; m) F. Cheng, Q. Wang, M. Chen, F. A. Quiocho, J. Ma, *Proteins* **2008**, *70*, 1228–1234.

[40] J. A. Menendez, R. Lupu, *Nat. Rev. Cancer* **2007**, *7*, 763–777.

[41] a) G. Ma, M. Zancanella, Y. Oyola, R. D. Richardson, J. W. Smith, D. Romo, *Org. Lett.* **2006**, 8, 4497–4500, and references cited therein; b) R. D. Richardson, G. Ma, Y. Oyola, M. Zancanella, L. M. Knowel, P. Cieplak, D. Romo, J. W. Smith, *J. Med. Chem.* **2008**, *51*, 5285–5296.

[42] T. D. Filippatos, C. S. Derdemezis, I. F. Gazi, E. S. Nakou, D. P. Mikhailidis, M. S. Elisaf, *Drug Saf.* **2008**, *31*, 53–65.

[43] M. Uttamchandani, C. H. S. Lu, S. Q. Yao, *Acc. Chem. Res.* **2009**, *42*, 1183–1192.

[44] Y. -C. Teo, K. -T. Tan, T. -P. Loh, *Chem. Commun.* **2005**, 1318–1320.

[45] T. R. Hoye, C. S. Jeffrey, F. Shao, *Nat. Protoc.* **2007**, *2*, 2451–2458.

[46] D. C. Dieterich, A. J. Link, J. Graumann, D. A. Tirrell, E. M. Schuman, *Proc. Natl.Acad. Sci. U.S.A.* **2006**, *103*, 9482–9487.

[47] G. Charron, M. M. Zhang, J. S. Yount, J. Wilson, A. S. Raghavan, E. Shamir, H. C. Hang, *J. Am. Chem. Soc.* **2009**, *131*, 4967–4975.

[48] M. Martin Moche, G. Schneider, P. Edwards, K. Dehesh, Y. Lindqvist, *J. Biol. Chem.* **1999**, *274*, 6031–6034.

[49] a) R. Mori, Q. Wang, K. D. Danenberg, J. K. Pinski, P. V. Danenberg, *Prostate* **2008**, *68*, 1555–1560; b) M. S. Phadke, N. F. Krynetskaia, A. K. Mishra, E. Krynetskiy, *J. Pharmacol. Exp. Therapeut.* **2009**, *331*, 77–86.

[50] D. G. Kingston, D. J. Newman, *Curr. Opin. Drug Discov. Dev.* **2007**, *10*, 130–144.

[51] a) Y. Zhu, H. Lin, Z. Li, M. Wang, J. Luo, *Breast Cancer Res. Treat.* **2001**, 69, 29–38; b) Y. Liu, X. Zhu, S. Liao, Q. Tang, K. Liu, X. Guan, J. Zhang, Z. Feng, *Oncol. Rep.* **2007**, *18*, 943–951; c) Y. Wang, D. Cheong, S. Chan, S. C. Hooi, *Int. J. Oncol.* **2000**, *16*, 757–762; d) M. H. Vaarala, K. S. Porvari, A. P. Kyllonen, M. V. J. Mustonen, O. Lukkarinen, P. T. Vihko, *Int. J. Cancer* **1998**, *78*, 27–32.

[52] Commerically available antibody for RPL14 failed to detect even endogenous RPL14, and was thus not pursued further in our experiments.

[53] a) M. H. Potashman, M. E. Duggan, *J. Med. Chem.* **2009**, *52*, 1231–1246; b) C. Drahl, B. F. Cravatt, E. J. Sorensen, *Angew. Chem., Int. Ed.* **2005**, *44*, 5788–5809; c) D. Mukherji, J. Spicer, *Exp. Opin. Invest. Drugs* **2009**, *18*, 293–301.

[54] L. Sleno, A. Emili, *Curr. Opin. Chem. Biol.* **2008**, *12*, 46–54.

[55] a) B. F. Cravatt, G. M. Simon, J. R. 3rd. Yates, *Nature* **2007**, *450*, 991–1000; b) M. Bantscheff, A. Scholten, A. J. Heck, *Drug Discovery Today* **2009**, *14*,

1021–1029; d) K. Wierzba, M. Muroi, H. Osada, *Curr. Opin. Chem. Biol.* **2011**, *15*, 57–65.

[56] a) T. Bottcher, M. Pitscheider, S. A. Sieber, *Angew. Chem. Int. Ed.* **2010**, *49*, 2680–2698; b) E. E. Carlson, *ACS Chem. Biol.* **2010**, *5*, 639–653.

[57] a) P.-Y. Yang, K. Liu, M. H. Ngai, M. J. Lear, M. R. Wenk, S. Q. Yao, J. *Am. Chem. Soc.* **2010**, *132*, 656–666; b) M. H. Ngai, P.-Y. Yang, K. Liu, Y. Shen, S. Q. Yao, M. J. Lear, *Chem. Commun.* **2010**, *46*, 8335–8837.

[58] a) Z. Wang, C. Gu, T. Colby, T. Shindo, R. Balamurugan, H. Waldmann, M. Kaiser, R. A. van der Hoorn, *Nat. Chem. Biol.* **2008**, *4*, 557–563; b) F. J. Dekker, O. Rocks, N. Vartak, S. Menninger, C. Hedberg, R. Balamurugan, S. Wetzel, S. Renner, M. Gerauer, B. Schölermann, M. Rusch, J. W. Kramer, D. Rauh, G. W. Coates, L. Brunsveld, P. I. Bastiaens, H. Waldmann, *Nat. Chem. Biol.* **2010**, *6*, 449–456.

[59] US Food and Drug Administration. Orlistat (marketed as Alli and Xenical) Information (online). Available from URL: http://www.fda.gov/Drugs/DrugSafety/PostmarketDrugSafetyInformationforPatientsandProviders/ucm180076.htm (Accessed 2010 Nov. 15).

[60] C. P. Pallasch, J. Schwamb, S. Königs, A. Schulz, S. Debey, D. Kofler, J. L. Schultze, M. Hallek, A. Ultsch, C. M. Wendtner, *Leukemia* **2008**, *22*, 585–585.

[61] a) W. Hillaert, M. Verdoes, B. I. Florea, N. Saragliadis, K. L. L. Habets, J. Kuiper, S. Van Calenbergh, F. Ossendorp, G. A. van der Marel, C. Driessen, H. S. Overkleeft, *Angew. Chem., Int. Ed.* **2009**, *48*, 1629–1632; b) M. P. Pereira, S. O.

Kelley, *J. Am. Chem. Soc.* **2011**, *133*, 3260-3263; c) Y. Loh, H. Shi, M. Hu, S. Q. Yao, *Chem. Commun.* **2010**, *46*, 8407-8409; d) D. Srikun, A. E. Albers, C. I. Nam, A. T. Iavaron, C. J. Chang, *J. Am. Chem. Soc.* **2010**, *132*, 4455–4465.

[62] a) A. Keppler, S. Gendreizig, T. Gronemeyer, H. Pick, H. Vogel, K. Johnsson, *Nat. Biotechnol.* **2003**, *21*, 86–89; b) A. Juillerat, T. Gronemeyer, A. Keppler, S. Gendreizig, H. Pick, H. Vogel, K. Johnsson, *Chem. Biol.* **2003**, *10*, 313–317; c) A. Keppler, M. Kindermann, S. Gendreizig, H. Pick, H. Vogel, K. Johnsson, *Methods* **2004**, *32*, 437–444; d) A. Gautier, A. Juillerat, C. Heinis, I. R. Correa, M. Kindermann, F. Beaufils, K. Johnsson, *Chem. Biol.* **2008**, *15*, 128–136.

[63] a) S. Wong, J. McLaughlin, D. Cheng, C. Zhang, K. M. Shokat, O. N. Witte, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 17456–17461; b) C. Kung, D. M. Kenski, S. H. Dickerson, R. W. Howson, L. F. Kuyper, H. D. Madhani, K. M. Shokat, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 3587–3592.

[64] Y. Y. Yang, J. M. Ascano, H. C. Hang, *J. Am. Chem. Soc.* **2010**, *132*, 3640–3641.

[65] Z. Ma, C. Chu, D. J. Cheng, *J. Lipid Res.* **2009**, *50*, 2131–2135.

[66] C. A. Brautigam, J. L. Chuang, D. R. Tomchick, M. Machius, D. T. Chuang, *J. Mol. Biol.* **2005**, *350*, 543–552.

[67] J. L. Periago, M. L. Pita, M. A. Sanchez del Castillo, G. Caamaño, M. D. Suárez, *Lipids* **1989**, *24*, 383–388.

[68] G. B. Kudolo, M. J. Harper, *J. Lipid Mediat. Cell Signal.* **1995**, *11*, 145–158.

[69] R. Sundaramoorthy, E. Micossi, M. S. Alphey, V. Germain, J. H. Bryce, S. M. Smith, G. A. Leonard, W. N. Hunter, *J. Mol. Biol.* **2006**, *359*, 347–357.

[70] a) K. Moreau, E. Dizin, H. Ray, C. Luquain, E. Lefai, F. Foufelle, M. Billaud, G. M. Lenoir, N. D. Venezia, *J. Biol. Chem.* **2006**, *281*, 3172–3181; b) S. Y. Bu, M. T. Mashek, D. G. Mashek, *J. Biol. Chem.* **2009**, *284*, 30474–30483.

[71] a) J. Y. Lu, L. A. Verkruyse, S. L. Hofmann, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10046–10050; b) A. Dahlqvist, U. Stål, M. Lenman, A.Banas, M. Lee, L. Sandager, H. Ronne, S. Stymne, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6487–6492.

[72] a) S. E. Ong, M. Mann, *Nat. Chem. Biol.* **2005**, *1*, 252–262; b) .M. Bantscheff, M. Schirle, G. Sweetman, J. Rick, B. Kuster, *Anal. Bioanal. Chem.* **2007**, *389*, 1017–1031.

[73] a) W. Zhu, J. W. Smith, C. M. Huang, *J. Biomed. Biotechnol.* **2010**, *2010*, 840518; b) H. Liu, R. G. Sadygov, J. R. Yates 3rd, *Anal. Chem.* **2004**, *76*, 4193–4201; c) Y. Ishihama, Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber, M. Mann, *Mol. Cell. Proteomics* **2005**, *4*, 1265–1272; d) K. Shinoda, M. Tomita, Y. Ishihama, *Bioinformatics*, **2010**, *26*, 576–577.

[74] a) P. J. R. Goulder, D. I. Watkins, *Nat. Rev. Immunol.* **2008**, *8*, 619–630; b) Y. Kawashima, K. Pfafferott, J. Frater, P. Matthews, R. Payne, M. Addo, H. Gatanaga, M. Fujiwara, A. Hachiya, H. Koizumi, N. Kuse, S. Oka, A. Duda, A. Prendergast, H. Crawford, A. Leslie, Z. Brumme, C. Brumme, T. Allen, C. Brander, R. Kaslow, J. Tang, E. Hunter, S. Allen, J. Mulenga, S. Branch, T. Roach, M. John, S. Mallal, A. Ogwu, R. Shapiro, J. G. Prado, S. Fidler, J. Weber, O. G. Pybus, P. Klenerman, T. Ndung'u, R. Phillips, D. Heckerman, P. R. Harrigan, B. D. Walker, M. Takiguchi, P. Goulder, *Nature* **2009**, *458*, 641–645; c) P. Kiepiela,

A. J. Leslie, I. Honeyborne, D. Ramduth, C. Thobakgale, S. Chetty, P. Rathnavalu, C. Moore, K. J. Pfafferott, L. Hilton, P. Zimbwa, S. Moore, T. Allen, C. Brander, M. M. Addo, M. Altfeld, I. James, S. Mallal, M. Bunce, L. D. Barber, J. Szinger, C. Day, P. Klenerman, J. Mullins, B. Korber, H. M. Coovadia, B. D. Walker, P. J. Goulder, *Nature* **2004**, *432*, 769–775.

[75] a) S. R. Neves, P. T. Ram, R. Iyengar, *Science* **2002**, *296*, 1636–1639; b) J. C. Migeon, S. L. Thomas, N.M. Nathanson, *J. Biol. Chem.* **1994**, *269*, 29146–29152; c) J. F. Foley, S. P. Singh, M. Cantu, L. Chen, H. H. Zhang, J. M. Farber, *J. Biol. Chem.* **2010**, *285*, 35537–35550.

[76] J. Diekmann, E. Adamopoulou, O. Beck, G. Rauser, S. Lurati, S. Tenzer, H. Einsele, H. G. Rammensee, H. Schild, M. S. Topp, *J. Immunol.* **2009**, *183*, 1587–1597.

[77] N. de la Iglesia, G. Konopka, S. V. Puram, J. A. Chan, R. M. Bachoo, M. J. You, D. E. Levy, R. A. Depinho, A. Bonni A, *Genes. Dev.* **2008**, *22*, 449–462.

[78] S. Jaisson, M. Veiga-da-Cunha, E. Van Schaftingen, *Biochimie* **2009**, *91*, 1066–1071.

[79] a) L. Rajendran, H. J. Knölker, K. Simons, *Nat. Rev. Drug Discov.* **2010**, *9*, 29–42; b) L. Rajendran, A. Schneider, G. Schlechtingen, S. Weidlich, J. Ries, T. Braxmeier, P. Schwille, J. B. Schulz, C. Schroeder, M. Simons, G. Jennings, H. J. Knolker, K. Simons, *Science* **2008**, *14*, 242–243.

[80] a) A. Keppler, H. Pick, C. Arrivoli, H. Vogel, K. Johnsson, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9955–9959; b) E. Tomat, E. M. Nolan, J. Jaworski, S. J. Lippard,

J. Am. Chem. Soc. **2008**, *130*, 15776–15777; c) M. Bannwarth, I. R. Correa, M. Sztretye, S. Pouvreau, C. Fellay, A. Aebischer, L. Royer, E. Rios, K. Johnsson, *ACS Chem. Biol.* **2009**, *4*, 179–190; d) M. Kamiya, K. Johnsson, *Anal. Chem.* **2010**, *82*, 6472–6479.

[81] a) S. W. Taylor, E. Fahy, S. S. Ghosh, *Trends Biotechnol.* **2003**, *21*, 82–88; b) A. Y. Andreyev, Z. Shen, Z. Guan, A. Ryan, E. Fahy, S. Subramaniam, C. R. Raetz, S. Briggs, E. A. Dennis, *Mol. Cell Proteomics.* **2010**, *9*, 388–402.

[82] a) World Health Organization. *Wkly. Epidemiol. Rec.* **2006**, *81*, 71–80; b) P. P. Simarro, J. Jannin, P. Cattand, *PLoS Med.* **2008**, *5*, e55; c) S. Besteiro, M. P. Barrett, L. Riviere, F. Bringaud, *Trends Parasitol.* **2005**, *21*, 185–191.

[83] a) R. Pink, A. Hudson, M.-A. Mouries, M. Bendig, *Nat. Rev. Drug Disc.* **2005**, *4*, 727–740; b) A. R. Renslo, J. H. McKerrow, *Nat. Chem. Biol.* **2006**, *2*, 701–710; c) C. Chong, D. J. Sullivan Jr., *Nature* **2007**, *448*, 745–646; d) S. Sundar, T. K. Jha C. P. Thakur, J. Engel, H. Sindermann, C. Fischer, K. Junge, A. Bryceson, J. Berman, *N. Engl. J. Med.* **2002**, *347*, 1739–1746.

[84] a) S. K. Parker, R. M. Barkley, J. G. Rino, M. L. Vasil, *PLoS One.* **2009**, *4*, e4281; b) P. K. Crellin, J. P. Vivian, J. Scoble, F. M. Chow, N. P. West, R. Brammananth, N. I. Proellocks, A. Shahine, J. Le Nours, M. C. Wilce, W. J. Britton, R. L. Coppel, J. Rossjohn, T. Beddoe, *J. Biol. Chem.* **2010**, *285*, 30050–30060; c) K. L. Low, P. S. Rao, G. Shui, A. K. Bendt, K. Pethe, T. Dick, M. R. Wenk, *J. Bacteriol.* **2009**, *191*, 5037–5043.

[85] Z. B. Mackey, A. M. Baca, J. P. Mallari, A B. psel, A. Shelat, E. J. Hansell, P. K.

Chiang, B. Wolff, K. R. Guy, J. Williams, J. H. McKerrow, *Chem. Biol. Drug Des.* **2006**, *67*, 355–363.

[86] a) S. Scory, C. R. Caffrey, Y. D. Stierhof, A. Ruppel, D. Steverding, *Exp. Parasitol.* **1999**, *91*, 327–333; b) Z. B. Mackey, T. C. O'Brien, D. C. Greenbaum, R. B. Blank, J. H. McKerrow, *J. Biol. Chem.* **2004**, *279*, 48426–48433.

[87] a) J. C. Engel, P. S. Doyle, I. Hsieh, J. H. McKerrow, *J. Exp. Med.* **1998**, *188*, 725–734; b) P. S. Doyle, Y. M. Zhou, J. C. Engel, J. H. McKerrow, *Antimicrob. Agents Chemother.* **2007**, *51*, 3932–3939; c) J. H. McKerrow, P. S. Doyle, J. C. Engel, L. M. Podust, S. A. Robertson, R. Ferreira, T. Saxton, M. Arkin, I. D. Kerr, L. S. Brinen, C. S. Craik, *Mem. Inst. Oswaldo Cruzi* **2009**, *104*, 263–269; d) K. Brak, P. S. Doyle, J. H. McKerrow, J. A. Ellman, *J. Am. Chem. Soc.* **2006**, *130*, 6404–6410.

[88] a) Y. C. Lio, L. J. Reynolds, J. Balsinde, E. A. Dennis, Biochim. Biophys. Acta **1996**, 1302, 556–560; b) E. L. Ridgley, L. Ruben, Mol. Biochem. Parasitol. **2001**, 114, 29–40.

[89] a) M. A. Albert, J. R. Haanstra, V. Hannacert, J. Van Roy, *J. Biol. Chem.* **2005**, *280*, 28306–28315; b) J. W. Chambers, M. L. Fowler, M. T. Morris, J. C. Morris, *Mol. Biochem. Parasitol.* **2008**, *158*, 202–207.

[90] C. L. Verlinde, V. Hannaert, C. Blonski, M.Willson, J. J. Périé, L. A. Fothergill-Gilmore, F. R. Opperdoes, M. H. Gelb, W. G. Hol, P. A. Michels, *Drug Resist. Updat.* **2001**, *4*, 50–65.

[91] P. S. Kessler, M. Parsons *J. Biol. Chem.* **2005**, *280*, 9030–9036.

[92] S. Helfert, A. M. Estévez, B. Bakker, P. Michels, C. Clayton, *Biochem. J.* **2001**, *357*, 117–125.

[93] A. J. Cáceres, P. A. Michels, V. Hannaert, *Mol. Biochem. Parasitol.* **2010**, *169*, 50–54.

[94] a) I. Z. Zubrzycki, *Biophys. J.* **2002**, *82*, 2906–2915; b) S. Marché, P. A. Michels, F. R. Opperdoes, *Mol. Biochem. Parasitol.* **2000**, *106*, 83–91.

[95] C. Subramaniam, P. Veazey, S. Redmond, J. Hayes-Sinclair, E. Chambers, M. Carrington, K. Gull, K. Matthews, D. Horn, M. C. Field, *Eukaryot. Cel.* **2006**, *5*, 1539–1549.

[96] V. Hannaert, M. A. Albert, D. J. Rigden, M. T. da Silva Giotto, O. Thiemann, R. C. Garratt, J. Van Roy, F. R. Opperdoes, P. A. Michels, *Eur. J. Biochem.* **2003**, *270*, 3205–3213.

[97] K. Nakamura, S. Fujioka, S. Fukumoto, N. Inoue, K. Sakamoto, H. Hirata, Y. Kido, Y. Yabu, T. Suzuki, Y. Watanabe, H. Saimoto, H. Akiyama, K. Kita, *Parasitol. Int.* **2010**, *59*, 560–564.

[98] O. Hanrahan, H. Webb, R. O'Byrne, E. Brabazon, A. Treumann, J. D. Sunter, M. Carrington, H. P. Voorheis, *PLoS Pathog.* **2009**, *5*, e1000468.

[99] a) S. S. Sutterwala, F. F. Hsu, E. S. Sevova, K. J. Schwartz, K. Zhang, P. Key, J. Turk, S. M. Beverley, J. D. Bangs, Mol. Microbiol. **2008**, 70, 281–296; b) E. S. Sevova, J. D.Bangs, Mol. Biol. Cell **2009**, 20, 4739–4750.

[100] a) S. A. Young, T. K. Smith, *Mol. Microbiol.* **2010**, *76*, 1461–1482; b) F. Gibellini, W. N. Hunter, T. K. Smith, *Mol. Microbiol.* **2009**, *73*, 826–843.

[101] S. S. Sutterwala, C. H. Creswell, S. Sanyal, A. K. Menon, J. D. Bangs, Eukaryot. Cell **2007**, 6, 454–464.

[102] a) T. K.Smith, P. Bütikofer, *Mol. Biochem. Parasitol.* **2010**, *172*, 66–79; b) G. Richmond, T. K. Smith, *Mol. Microbiol.* **2007**, *63*, 1078–1095.

[103] a) G. I. Lepesheva, R. D. Ott, T. Y. Hargrove, Y. Y. Kleshchenko, I. Schuster, W. D. Nes, G. C. Hill, F. Villalta, M. R.Waterman, *Chem. Biol.* **2007**, *14*, 1283–1293; b) G. I. Lepesheva, H.-W. Park, T. Y. Hargrove, B. Vanhollebeke, Z. Wawrzak, J. M. Harp, M. Sundaramoorthy, W. David Nes, E. Pays, M. Chaudhuri, F. Villalta, M. R.Waterman, *J. Biol. Chem.* **2010**, *285*, 1773–1780.

[104] a) J. Rubotham, K. Woods, J. A. Garcia-Salcedo, E. Pays, D. P. Nolan, *J. Biol. Chem.* **2005**, *280*, 10410–10418; b) M. A.Comini, R. L. Krauth-Siegel, L. Flohé, *Biochem. J.* **2007**, *402*, 43–49.

[105] a) D. W. Jiang, K. A. Werbovetz, A. Varadhachary, R. N. Cole, P. T. Englund, *Mol. Biochem. Parasitol.* **2004**, *135*, 149–152; b) Y. S. Morita, K. S. Paul, P. T. Englund, *Science* **2000**, *288*, 140–143; c) L. Rivière, P. Moreau, S. Allmann, M. Hahn, M. Biran, N. Plazolles, J. M. Franconi, M. Boshart, F. Bringaud, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 12694–12699; d) D.W. Jiang, P. T. Englund, *Biochem. J.* **2001**, *358*, 757–761; e) D. W. Jiang, R. Ingersoll, P. J. Myler, P. T. Englund, *Exp. Parasitol.* **2000**, *96*, 16–22.

[106] J. Rubotham, K. Woods, J. A. Garcia-Salcedo, E. Pays, D. P. Nolan, *J. Biol. Chem.* **2005**, *280*, 10410–10418.

[107] M. Engstler, F. Weise, K. Bopp, C.G. Grünfelder, M. Günzel, N. Heddergott,

P. Overath, *J. Cell Sci.* **2005**, *118*, 2105–2118.

[108] M. C. Field, T. Sergeenko, Y. N. Wang, S. Böhm, M. Carrington, *PLoS One.* **2010**, *5*, e8468.

[109] C. Ebikeme, J. Hubert, M. Biran, G. Gouspillou, P. Morand, N. Plazolles, F. Guegan, P. Diolez, J. M. Franconi, J. C. Portais, F. Bringaud, *J. Biol. Chem.* **2010**, *285*, 32312–32324.

[110] S. C. Roberts, Y. Jiang, A. Jardim, N. S. Carter, O. Heby, B. Ullman, *Mol. Biochem. Parasitol.* **2001**, *115*, 217–226.

[111] W. Liu, K. Apagyi, L. McLeavy, K. Ersfeld, *Mol. Biochem. Parasitol.* **2010**, *169*, 20–26.

[112] S. Olego-Fernandez, S. Vaughan, M. K. Shaw, K. Gull, M. L. Ginger, *Protist.* **2009**, *160*, 576–590.

[113] M. J. Helms, A. Ambit, P. Appleton, L. Tetley, G. H. Coombs, J. C. Mottram, *J. Cell Sci.* **2006**, *119*, 1105–1117.

[114] I. Coppens, P. J. Courtoy, *Annu. Rev. Microbiol.* **2000**, *54*, 129–156.

[115] J. D. Bangs, L. Uyetake, M. J. Brickman, A. E. Balber, J. C. Boothroyd, *J. Cell Sci.* **1993**, *105*, 1101–1113.

[116] R. F. Peck, A. M. Shiflett, K. J. Schwartz, A. McCann, S. L. Hajduk, J. D. Bangs, *Mol. Microbiol.* **2008**, *68*, 933–946.

[117] G. A. Keller, S. Krisans, S. J. Gould, J. M. Sommer, C. C. Wang, W. Schliebs, W. Kunau, S. Brody, S. Subramani, *J. Cell Biol.* **1991**, *114*, 893–904.

[118] S. Nwaka, A. Hudson, *Nat. Rev. Drug Discov.* **2006**, *5*, 941–955.

[119] P. S. Doyle, M. Sajid, T. O'Brien, K. DuBois, J. C. Engel, Z. B. Mackey, S. Reed, *Curr. Pharm. Des.* **2008**, *14*, 889–900.

[120] C. R. Caffrey, D. Steverding, *Mol. Biochem. Parasitol.* **2009**, *167,* 12–19.

[121] I. D. Kerr, C. J. Farady, R. Marion, M. Richert, M. Sajid, K. C. Pandey, C. R. Caffrey, J. Legac, E. Hansell, J. H. McKerrow, C. S. Craik, P. J. Rosenthal, L. S. Brinen, *J. Biol. Chem.* **2009**, *284*, 25697–25703.

[122] J. C. Engel, P. S. Doyle, I. Hsieh, J. H. McKerrow, *J. Exp. Med.* **1998**, *188*, 725–734.

[123] S. C. Barr, K. L. Warner, B. G. Kornreic, J. Piscitelli, A. Wolfe, L. Benet, J. H. McKerrow, *Antimicrob. Agents Chemother.* **2005**, *49*, 5160–5161.

[124] P. S. Doyle, Y. M. Zhou, J. C. Engel, J. H. McKerrow, *Antimicrob. Agents Chemother.* **2007**, *51*, 3932–3939.

[125] M. H. Abdulla, K. C. Lim, M. Sajid, J. H. McKerrow, C. R. Caffrey, *PLoS Med.*, **2007**, *4*: e14.

[126] C. R. Caffrey, S. Scory, D. Steverding, *Curr. Drug Targets* **2000**, *1*, 155–162.

[127] S. G. Mel éndez-L ópez, S. Herdman, K. Hirata, M. H. Choi, Y. Choe, C. Craik, C. R. Caffrey, E. Hansell, B. Ch ávez-Mungu á, Y. T. Chen, W. R. Roush, J. McKerrow, L. Eckmann, J. Guo, S. L. Jr Stanley, S. L. Reed, *Eukaryot Cell*, **2007**, *6*, 1130–1136.

[128] I. D. Kerr, P. Wu, R. Marion-Tsukamaki, Z. B. Mackey, L. S. Brinen, *PloS Negl. Trop. Dis.* **2010**, *4*: e701.

[129] R. Somoza, H. Zhan, K. K. Bowman, L. Yu, K. D. Mortara, J. T. Palmer, J. M.

Clark, M. E. McGrath, *Biochemistry*, **2000**, *39*, 12543–12551.

[130] G. Wang, M. Uttamchandani, G. Y. J. Chen, S. Q. Yao, *Org. Lett.* **2003**, *5*, 737–740.

[131] Z. B. Mackey, T. C. O'Brien, D. C. Greenbaum, R. B. Blank, J. H. McKerrow, *J. Biol. Chem.* **2004**, *279*, 48426–48433.

[132] M. H. Abdulla, T. O'Brien, Z. B. Mackey, M. Sajid, D. J. Grab, J. H. McKerrow, *PLoS Negl. Trop. Dis.*, **2008**, *2*, e298.

[133] M. Bogyo, J. S. McMaster, M. Gaczynska, D. Tortorella, A. L. Goldberg, H. Ploegh, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6629–6634.

[134] T. McCormack, W. Baumeister, L. Grenier, C. Moomaw, L. Plamondon, B. Pramanik, C. Slaughter, F. Soucy, R. Stein, F. Zuhl, L. Dick, *J. Biol. Chem.* **1997**, *272*, 26103–26109.

[135] T. Nomura, N. Katunuma, *J. Med. Invest.* **2005**, *52*, 1–9.

[136] M. Lankelma, D. M. Voorend, T. Barwari, J. Koetsveld, A. H. Van der Spek, A. P. De Porto, G. Van Rooijen, C. J. Van Noorden, *Life Sci.* **2010**, *86*, 225–233.

[137] Chandran, N. J. Sullivan, U. Felbor, S. P. Whelan, J. M. Cunningham, *Science* **2005**, *308*, 1643–1645.

[138] C. T. Pager, R.E. Dutch, *J. Virol.* **2005**, *79*, 12714–12720.

[139] G. Simmons, D. N. Gosalia, A. J. Rennekamp, J. D. Reeves, S. L. Diamond, P. Bates, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 11876–11881.

[140] C. R. Caffrey, E. Hansell, K. D. Lucas, L. S. Brinen, A. Alvarez Hernandez, J. Cheng, S. L. Gwaltney II, W. R. Roush, Y.-D. Stierhof, M. Bogyo, D. Steverding,

J. H. McKerrow, *Mol. Biochem. Parasitol.* **2001**, *118*, 61–73.

[141] T. C. O'Brien, Z. B. Mackey, R. D. Fetter, Y. Choe, A. J. O'Donoghue, M. Zhou, C. S. Craik, C. R. Caffrey, J. H. McKerrow, *J. Biol. Chem.* **2008**, *283*, 28934–28943.

[142] G. Paulick, M. Bogyo, *ACS Chem. Biol.* **2011**, *6*, 563–572.

[143] J. H. McKerrow, J. C. Engel, C. R. Caffrey, *Bioorg Med. Chem.* **1999**, *7*, 639–644.

[144] C. J. Engel, P. S. Doyle, J. Palmer, I. Hsieh, D. F. Bainton, J. H. McKerrow, *J. Cell. Sci.* **1998**, *111*, 597–606.

[145] K. Brix, A. Dunkhorst, K. Mayer, S. Jordans, *Biochimie* **2008**, *90*, 194–207.

[146] P. J. Hotez, D. H. Molyneux, A. Fenwick, J. Kumaresan, S. E. Sachs, J. D. Sacha, L. Savioli, *N. Engl. J. Med.* **2007**, *357*, 1018–1027.

[147] C. Young, P. Losikoff, A. Chawla, L. Glasser, E. Forman, *Transfusion* **2007**, *47*, 540–544.

[148] A. Cavalli, M. L. Bolognesi, *J. Med. Chem.* **2009**, *52*, 7339–7459.

[149] S. C. Barr, K. L. Warner, B. G. Kornreic, J. Piscitelli, A. Wolfe, L. Benet, J. H. McKerrow, *Antimicrob. Agents Chemother.* **2005**, *49*, 5160–5161.

[150] O. V. Nikolskaia, A. P. de A Lima, Y. V. Kim, J. D. Lonsdale-Eccles, T. Fukuma, J. Scharfstein, D. J. Grab, *J. Clin. Invest.* **2006**, *116*, 2739–2747.

[151] S. Scory, Y. D. Stierhof, C. R. Caffrey, D. Steverding, *Kintoplastid Biol. Dis.* **2007**, *6*, 2.

[152] R. Löser, M. Frizler, K. Schilling, M. Gütschow, *Angew. Chem. Int. Ed.* **2008**,

47, 4331–4334.

[153] M. Frizler, F. Lohr, N. Furtmann, J. Kläs, M. Gütschow, *J. Med. Chem.* **2011**, *54*, 396–400.

[154] R. Löser, J. Gut, P. J. Rosenthal, M. Frizler, M. Gütschow, K. T. Andrews, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 252–255.

[155] J. Singh, R. U. Petter, T. A. Baillie, A. Whitty, *Nat. Rev. Drug Discov.* **2011**, *10*, 307–317.

[156] A. Zega, *Curr. Med. Chem.* **2005**, *12*, 589–597.

[157] L. S. Brinen, E. Hansell, J. Cheng, W. R. Roush, J. H. McKerrow, R. J. Fletterick, *Structure* **2000**, *8*, 831–840.

[158] P. Jaishankar, E. Hansell, D.-M. Zhao, P. S. Doyle, J. H. McKerrow, A. R. Renslo, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 624–628.

[159] R. Löser, K. Schilling, E. Dimmig, M. Gütschow, *J. Med. Chem.* 2005, *48*, 7688–7707.

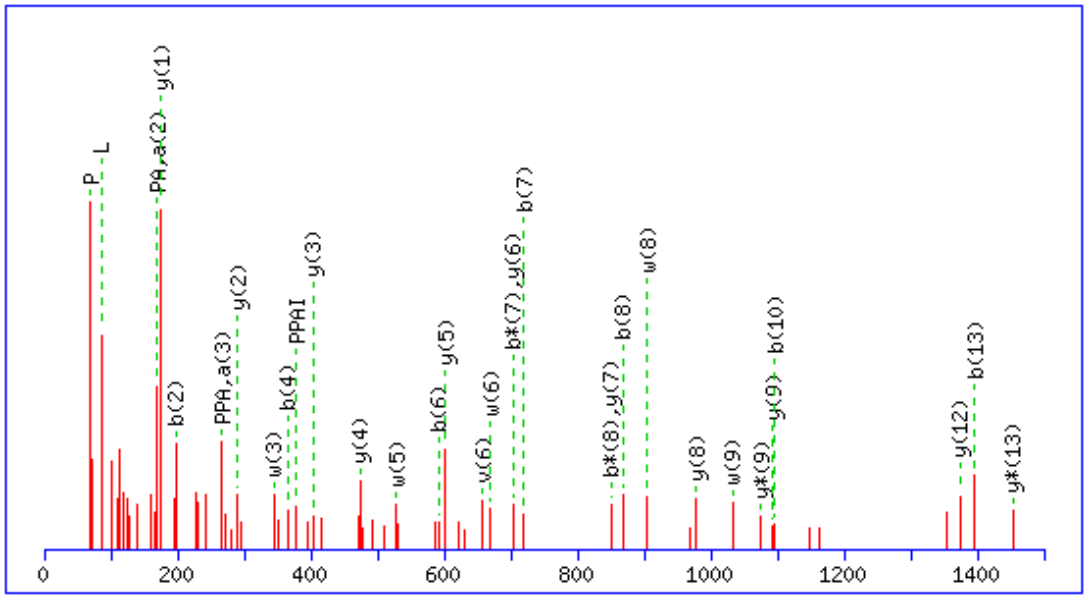
[160] a) C. S. Gan, T. Guo, H. Zhang, S. K. Lim, S. K. Sze, *J. Proteome Res.* **2008**, *7*, 4869–4877; b) P. Hao, T. Guo, X. Li, S. S. Adav, J. Yang, M. Wei, S. K. Sze, *J. Proteome Res.* **2010**, *9*, 3520–3526.

[161] T. Nguyen, M. B. Francis, *Org. Lett.* **2003**, *5*, 3245–3248.

[162] A. W. Schwabacher, J. W. Lane, M. W. Schiesher, K. M. Leigh, C. W. Johnson, *J. Org. Chem.* **1999**, *63*, 1727–1729.

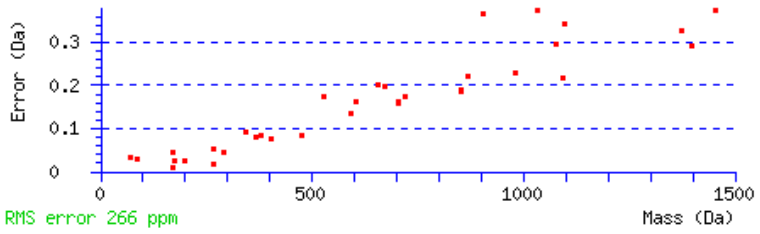
Appendix One

MS/MS Fragmentation of VPPAINQFTQALDR
Found in gi|4506661, ribosomal protein L7a [Homo sapiens]
Match to Query 58: 1568.930124 from (1569.937400,1+)
MaldiWellID: 55833, SpectrumID: 156280,



MONOISOTOPIC mass of neutral peptide Mr(calc): 1568.83
Fixed modifications: Carbamidomethyl (C)
Ions Score: 72, Expect: 0.00038
Matches (Bold Red): 36/217 fragment ions using 34 most intense peaks

#	Immon.	a	a*	a^0	b	b*	b^0	Seq.	v	w	w'	y	y*	y^0	#
1	72.08	72.08			100.08			V							14
2	**70.07**	**169.13**			**197.13**			P	1428.72	1427.73		1470.77	**1453.74**	1452.76	13
3	**70.07**	**266.19**			294.18			P	1331.67	1330.67		**1373.72**	1356.69	1355.71	12
4	44.05	337.22			**365.22**			A	1260.63			1276.66	1259.64	1258.65	11
5	**86.10**	450.31			478.30			I	1147.55	1160.57	1174.59	1205.63	1188.60	1187.62	10
6	87.06	564.35	547.32		**592.35**	575.32		N	1033.51	**1032.51**		**1092.54**	**1075.52**	1074.53	9
7	101.07	692.41	675.38		**720.40**	**703.38**		Q	905.45	**904.45**		**978.50**	961.47	960.49	8
8	120.08	839.48	822.45		**867.47**	**850.45**		F	758.38			**850.44**	833.42	832.43	7
9	74.06	940.53	923.50	922.51	968.52	951.49	950.51	T	**657.33**	**670.35**	672.33	**703.37**	686.35	685.36	6
10	101.07	1068.58	1051.56	1050.57	**1096.58**	1079.55	1078.57	Q	529.27	**528.28**		**602.33**	585.30	584.32	5
11	44.05	1139.62	1122.59	1121.61	1167.62	1150.59	1149.61	A	458.24			**474.27**	457.24	456.26	4
12	**86.10**	1252.70	1235.68	1234.69	1280.70	1263.67	1262.69	L	345.15	**344.16**		**403.23**	386.20	385.22	3
13	88.04	1367.73	1350.71	1349.72	**1395.73**	1378.70	1377.72	D	230.12	229.13		**290.15**	273.12	272.14	2
14	129.11							R	74.02	73.03		**175.12**	158.09		1



Chapter 2_Figure S1. Representative MS/MS profile for RPL7a.

Chapter 3_ Relevant references of **Table 3.1**

1. For HLA-A, also see: M. Williams, J. F. Roeth, M. R. Kasper, R. I. Fleis, C. G. Przybycin, K. L. Collins, *J. Virol.* **2002**, *76*, 12173–121784.
2. For PFN1, see: R. Gieselmann, D. J. Kwiatkowski, P. A. Janmey, W. Witke, *Eur. J. Biochem.* **1995**, *229*, 621–628.
3. For RPL22, see: G. Nucifora, C. R. Begy, R. Erickson, H. A. Drabkin, J. D. Rowley, *Proc. Nat. Acad. Sci. USA* **1993**, *90*, 7784–7788.
4. For MIF, see: L. Leng, C. N. Metz, Y. Fang, J. Xu, S. Donnelly, J. Baugh, T. Delohery, Y. Chen, R. A. Mitchell, R. Bucala, *J. Exp. Med.* **2003**, *197*, 1467–1476.
5. For RPL12, see: M. Cuccurese, G. Russo, A. Russo, C. Pietropaolo, *Nucleic Acids Res.* **2005**, *33*, 5965–5977.
6. For PC, see: S. Jitrapakdee, M. St Maurice, I. Rayment, W. W. Cleland, J. C. Wallace, P. V. Attwood, *Biochem. J.* **2008**, *413*, 369–387.
7. For PGM1, see: D. B. Whitehouse, W. Putt, J. U. Lovegrove, K. Morrison, M. Hollyoake, M. F. Fox, D. A. Hopkinson, Y. H. Edwards, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 411–415.
8. For BCAP31, see: B. Wang, M. Nguyen, D. G. Breckenridge, M. Stojanovic, P. A. Clemons, S. Kuppig, G. C. Shore, *J. Biol. Chem.* **2003**, *278*, 14461-14468.
9. For CLTCL1, see: S. Vassilopoulos, C. Esk, S. Hoshino, B. H. Funke, C. -Y. Chen, A. M. Plocik, W. E. Wright, R. Kucherlapati, F. M. Brodsky, *Science* **2009**, *324*, 1192–1196.

10. For PFKM, see: M. Garcí a, A. Pujol, A. Ruzo, E. Riu, J. Ruberte, A. Arbós, A. Serafí n, B. Albella, J. E. Felí u, F. Bosch, *PLoS Genet.* **2009**, *5*, e1000615.

11. For TPP2, see: E. Geier, G. Pfeifer, M. Wilm, M. Lucchiari-Hartz, W. Baumeister, K. Eichmann, G. Niedermann, *Science* **1999**, *283*, 978–981.

12. For HLA-B, also see: J. M. Johnson, C. Nicot, J. Fullen, V. Ciminale, L. Casareto, J. C. Mulloy, S. Jacobson, G. Franchini, *J. Virol.* **2001**, *75*, 6086–6094.

13. For PTGES2, see: K.Watanabe, H. Ohkubo, H. Niwa, N. Tanikawa, N. Koda, S. Ito, Y. Ohmiya, *Biochem. Biophys. Res. Commun.* **2003**, *306*, 577–581.

14. For ANP32A, see: a) J. Bai, J. R. Brody, S. S. Kadkol, G. R. Pasternack, *Oncogene* **2001**, *20*, 2153–2160; b) M. Cvetanovic, R. J. Rooney, J. J. Garcia, N. Toporovskaya, H. Y. Zoghbi, P. Opal, *EMBO Rep.* **2007**, *8*, 671–677.

15. For ATP1B3, see: N. Malik, V. A. Canfield, M. C. Beckers, P.Gros, R. Levenson, *J. Biol. Chem.* **1996**, *271*, 22754–22758.

16. For P4HB, see: R. A. Lumb, N. J. Bulleid, *EMBO J.* **2002**, *21*, 6763–6770.

17. For HLA-C, also see: J. Zemmour, J. E. Gumperz, W. H. Hildebrand, F. E. Ward, S. G. Marsh, R. C. Williams, P. Parham, *Tissue Antigens.* **1992**, *39*, 249–257.

18. For EIF3E, see: B. Siridechadilok, C. S. Fraser, R. J. Hall, J. A. Doudna, E. Nogales, *Science* **2005**, *310*, 1513–1515.

19. For NOP56, see: T. Gautier, T. Bergè s, D. Tollervey, E. Hurt, *Mol. Cell Biol.* **1997**, *17*, 7088–7098.

20. For MYOF, see: P. N. Bernatchez, L. Acevedo, C. Fernandez-Hernando, T. Murata, C. Chalouni, J. Kim, H. Erdjument-Bromage, V. Shah, J. -P. Gratton, E. M. McNally, P. Tempst, W. C. Sessa, *J. Biol. Chem.* **2007**, *282*, 30745–30753.

21. For GNAI2, also see: C. R. Beals, C. B. Wilson, R. M. Perlmutter, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7886–7890.

22. For SCARB1, see: E. Scarselli, H. Ansuini, R. Cerino, R. M. Roccasecca, S. Acali, G. Filocamo, C. Traboni, A. Nicosia, R. Cortese, A. Vitelli, *EMBO J.* **2002**, *21*, 5017–5025.

23. For STAT3, see: A. Dreuw, S. Radtke, S. Pflanz, B. E. Lippok, P. C. Heinrich, H. M. Hermanns, *J. Biol. Chem.* **2004**, *279*, 36112–36120.

24. For TMEM48, see: J. Mansfeld, S. Guettinger, L. A. Hawryluk-Gara, N. Pante, M. Mall, V. Galy, U. Haselmann, P. Muehlhaeusser, R. W. Wozniak, I. W. Mattaj, U. Kutay, W. Antonin, *Mol. Cell* **2006**, *22*, 93–103.

25. For LMNA, see: G. Bonne, M. R. Di Barletta, S. Varnous, H. -M. Becane, E.-H. Hammouda, L. Merlini, F. Muntoni, C. R. Greenberg, F. Gary, J.-A. Urtizberea, D. Duboc, M. Fardeau, D. Toniolo, K. Schwartz, *Nat. Genet.* **1999**, *21*, 285–288.

26. For BZWL, see: P. Mitra, P. S. Vaughan, J. L. Stein, G. S. Stein, A. J. van Wijnen, *Biochemistry* **2001**, *40*, 10693–10699.

27. For RAB4B, see: H. He, F. Dai, L. Yu, X. She, Y. Zhao, J. Jiang, X. Chen, S. Zhao, *Gene Expr.* **2002**, *10*, 231–242.

28. For NIT2, see: B. F. Krasnikov, C. -H. Chien, R. Nostramo, J. T. Pinto, E. Nieves, M. Callaway, J. Sun, K. Huebner, A. J. L. Cooper, *Biochimie* **2009**, *91*, 1072–1080.

29. For KIAA0368, see: C. Gorbea, G. M.Goellner, K. Teter, R. K. Holmes, M. Rechsteiner, *J. Biol. Chem.* **2004**, *279*, 54849–54861.

30. For GTF2I, see: C. D. Novina, S. Kumar, U. Bajpai, V. Cheriyath, K. Zhang, S. Pillai, H. H. Wortis, A. L. Roy, *Mol. Cell. Biol.* **1999**, *19*, 5014–5024.

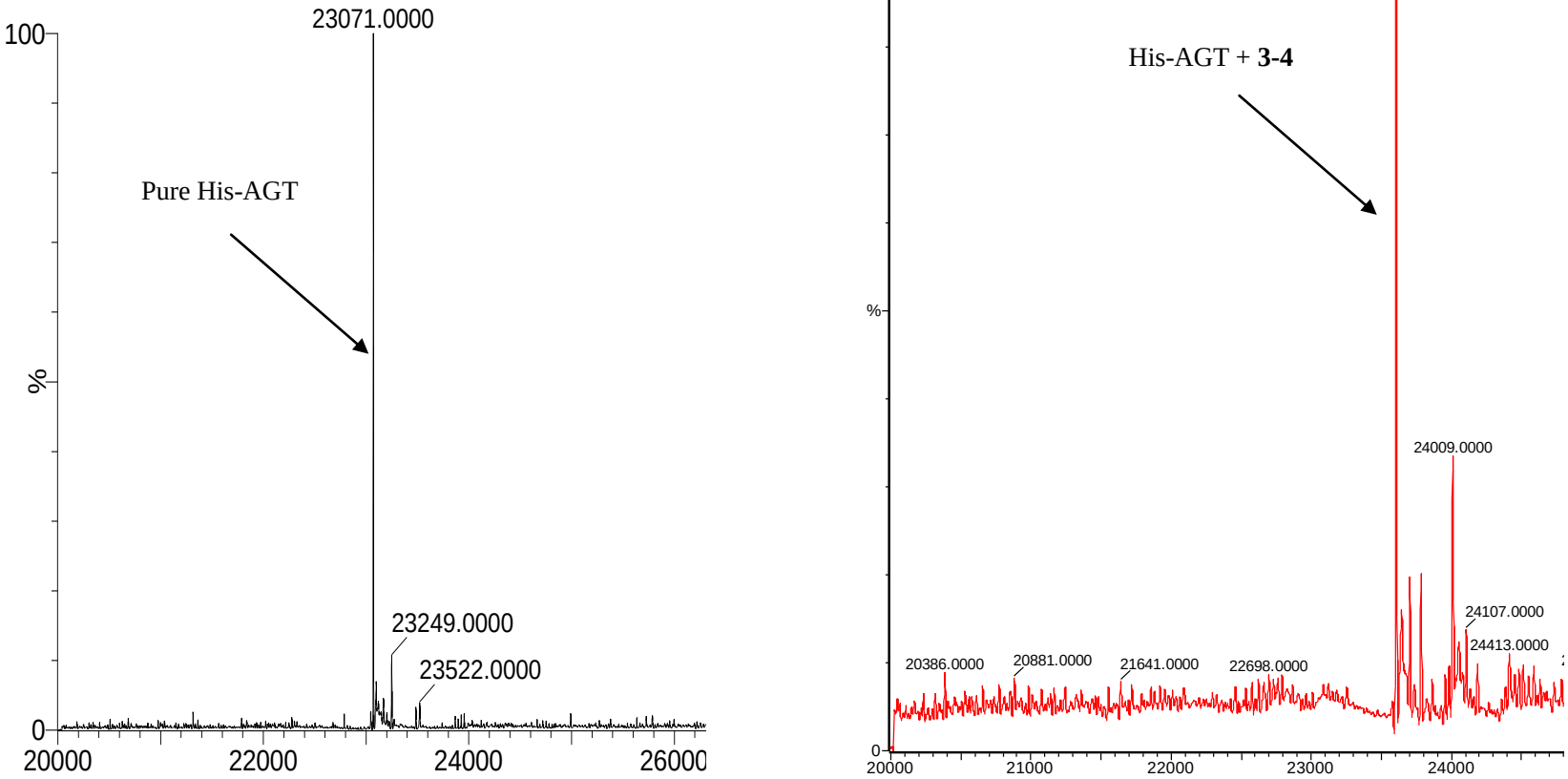
31. For LY6K, see: N. Ishikawa, A. Takano, W. Yasui, K. Inai, H. Nishimura, H. Ito, Y. Miyagi, H. Nakayama, M. Fujita, M. Hosokawa, E. Tsuchiya, N. Kohno, Y. Nakamura, Y. Daigo, *Cancer Res.* **2007**, *67*, 11601–11611

32. For GNAO1, also see: S. Lavu, J. Clark, R. Swarup, K. Matshushima, K. Paturu, J. Moss, H.-F. Kung, *Biochem. Biophys. Res. Commun.* **1988**, *150*, 811–815.

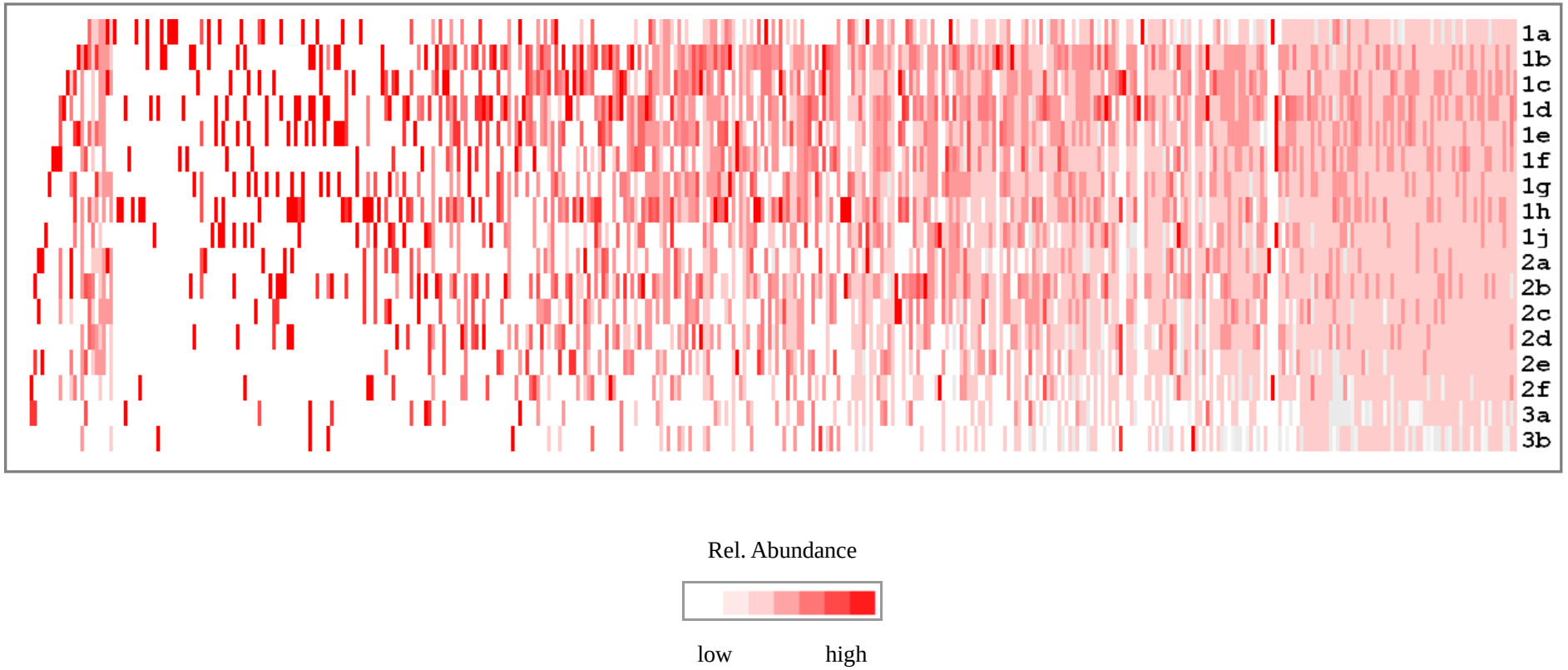
33. For ATP2A1, see: M. Chami, B. Oulès, G. Szabadkai, R. Tacine, R. Rizzuto, P. Paterlini-Bréchot, *Mol. Cell.* **2008**, *32*, 641–651.

34. For SH3KBP1, see: a) Y. Aissouni, G. Zapart, J. L. Iovanna, I. Dikic, P. Soubeyran, *Biochem. Biophys. Res. Commun.* **2005**, *338*, 808–814; b) T. Narita, T. Nishimura, K. Yoshizaki, T. Taniyama, *Exp. Cell Res.* **2005**, *304*, 256–264.

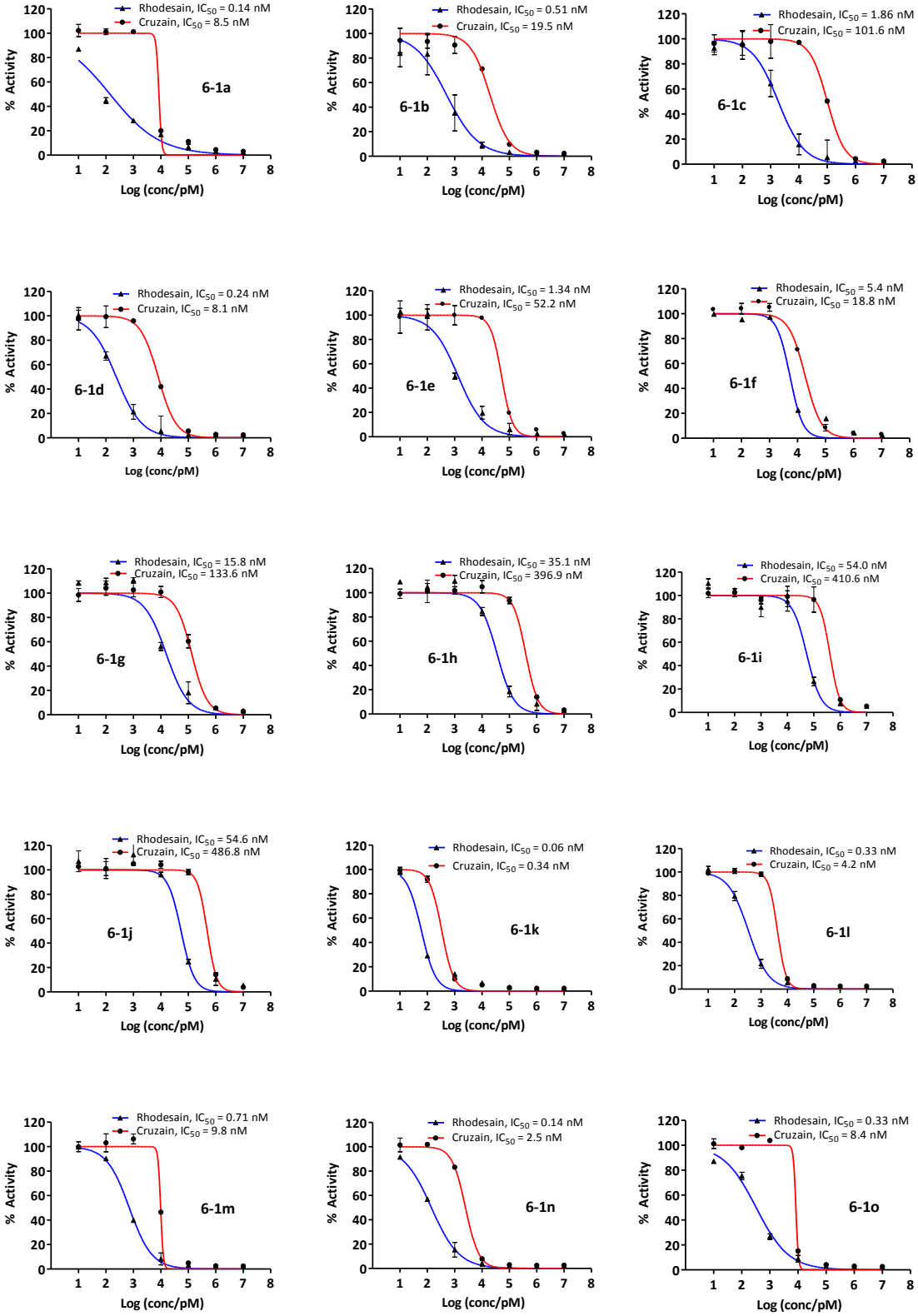
35. For PLEC1, see: C. Abrahamsberg, P. Fuchs, S. Osmanagic-Myers, I. Fischer, F. Propst, A. Elbe-Bürger, G. Wiche, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 18449–18454.

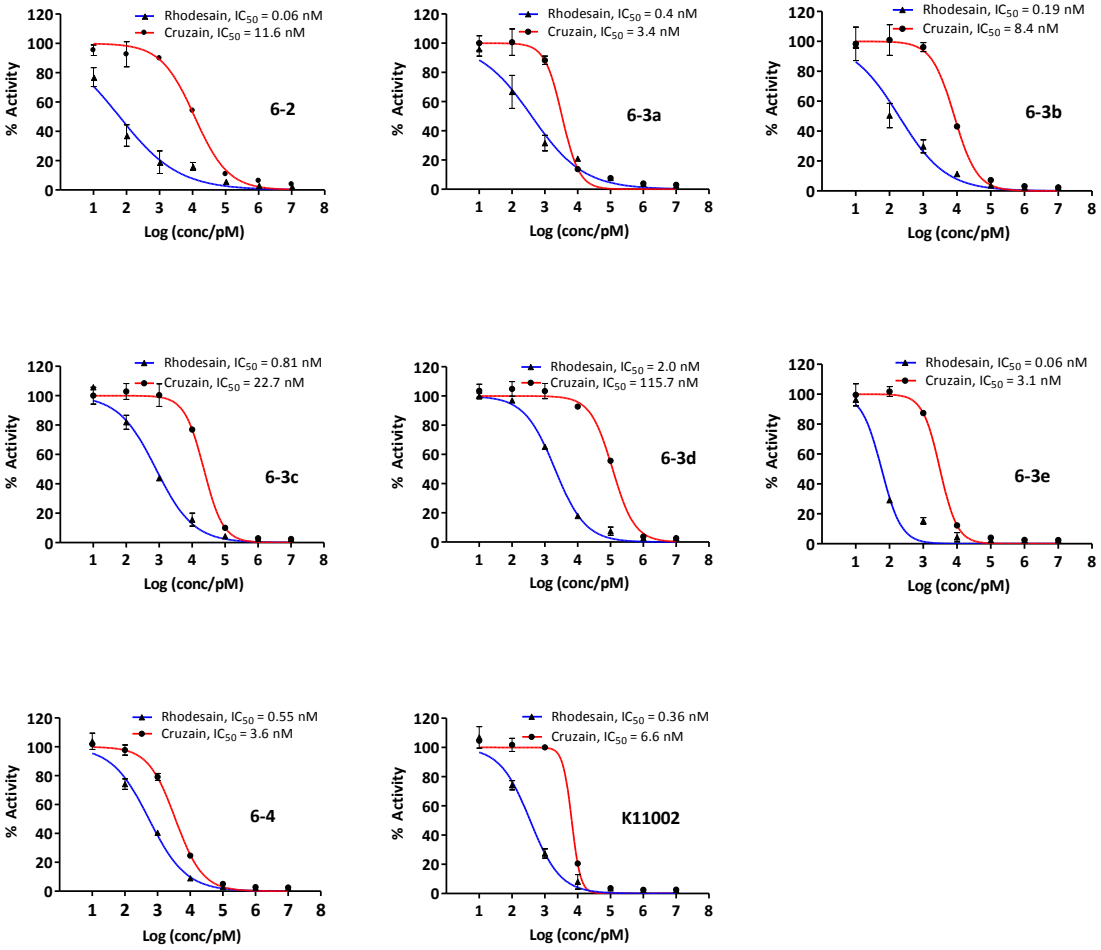


Chapter 3_Figure S1. Observed Mass peaks for ESI-MS of His-AGT with or without **3-4**

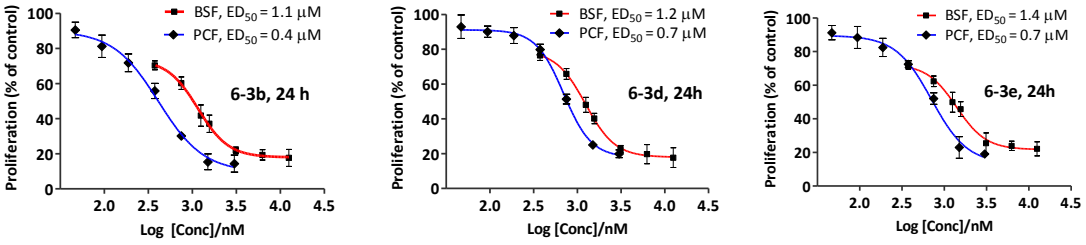


Chapter 3_Figure S2. A full heatmap showing relative abundance of each protein identified by the probes from the pull-down/LCMS experiments. The resulting relative abundance is color-indicated according to their intensity (white for low abundance; red for high abundance).

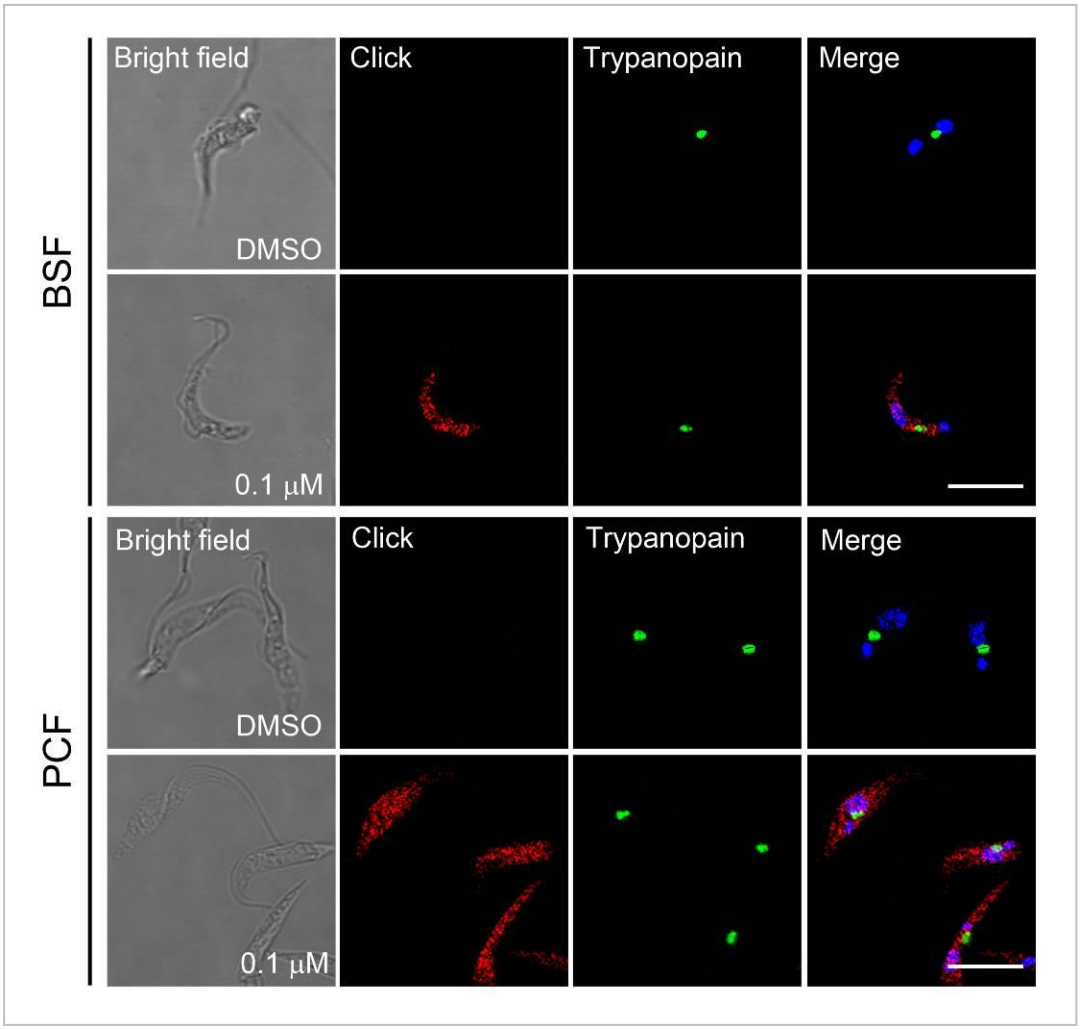

Rhodesain, IC50 = 0.14 nM
Cruzain, IC50 = 8.5 nM
6-1a
Rhodesain, IC50 = 0.51 nM
Cruzain, IC50 = 19.5 nM
6-1b
Rhodesain, IC50 = 1.86 nM
Cruzain, IC50 = 101.6 nM
6-1c
Rhodesain, IC50 = 0.24 nM
Cruzain, IC50 = 8.1 nM
6-1d
Rhodesain, IC50 = 1.34 nM
Cruzain, IC50 = 52.2 nM
6-1e
Rhodesain, IC50 = 5.4 nM
Cruzain, IC50 = 18.8 nM
6-1f
Rhodesain, IC50 = 15.8 nM
Cruzain, IC50 = 133.6 nM
6-1g
Rhodesain, IC50 = 35.1 nM
Cruzain, IC50 = 396.9 nM
6-1h
Rhodesain, IC50 = 54.0 nM
Cruzain, IC50 = 410.6 nM
6-1i
Rhodesain, IC50 = 54.6 nM
Cruzain, IC50 = 486.8 nM
6-1j
Rhodesain, IC50 = 0.06 nM
Cruzain, IC50 = 0.34 nM
6-1k
Rhodesain, IC50 = 0.33 nM
Cruzain, IC50 = 4.2 nM
6-1l
Rhodesain, IC50 = 0.71 nM
Cruzain, IC50 = 9.8 nM
6-1m
Rhodesain, IC50 = 0.14 nM
Cruzain, IC50 = 2.5 nM
6-1n
Rhodesain, IC50 = 0.33 nM
Cruzain, IC50 = 8.4 nM
6-1o
% Activity
Log (conc/pM)



Chapter 6_Figure S1. IC_{50} values of **6–1a-o, 6–2, 6–3a-e, 6–4** and K11002 against Rhodesain (blue) and Cruzain (red).



Chapter 6_Figure S2. ED_{50} values of **6–3b, 6–3d,** and **6–3e** against PCF (blue) and BSF (red).

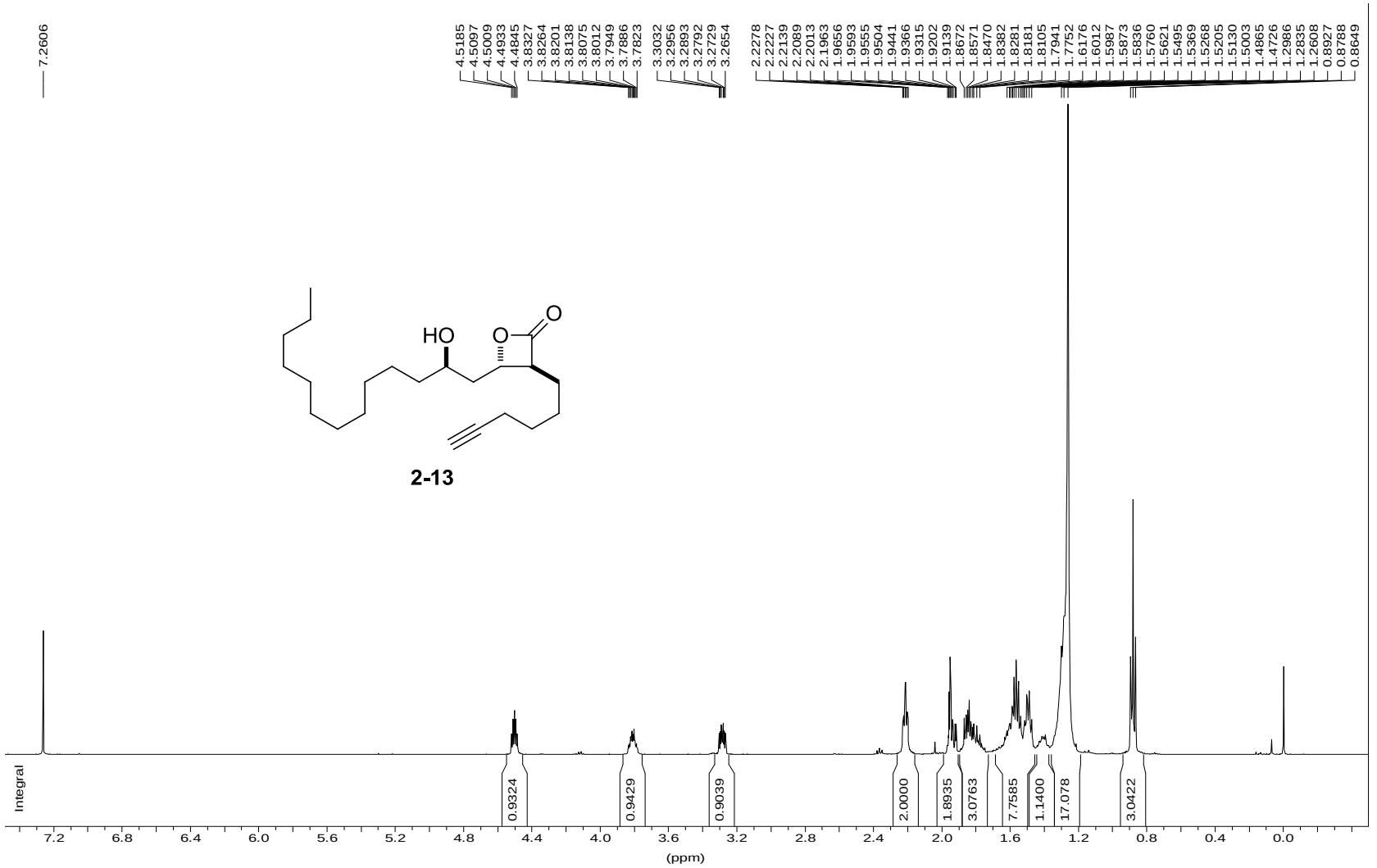


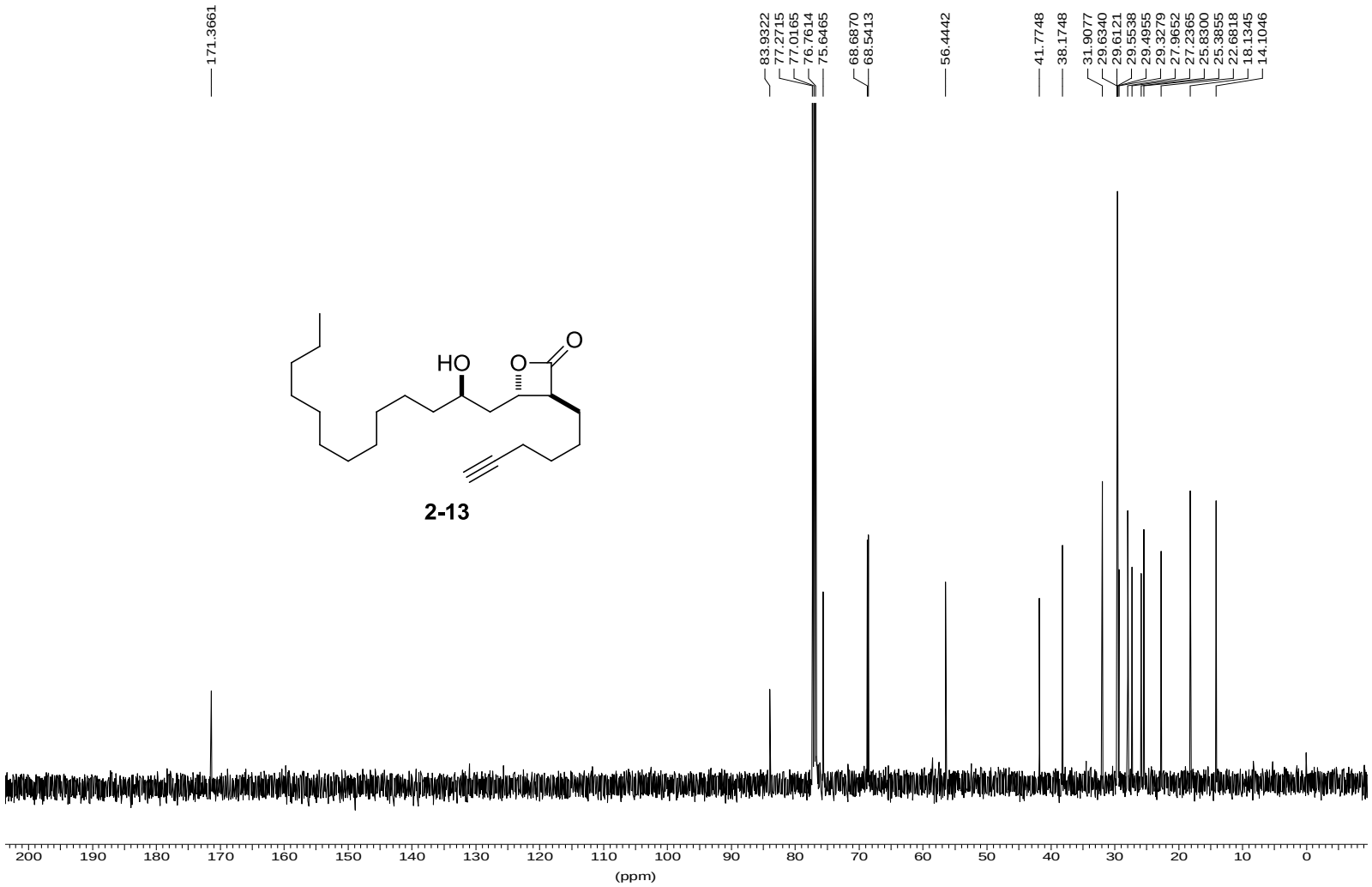
Chapter 6_Figure S3. Immunofluorescence analysis of brucipain in BSF and PCF *T. brucei* treated with DMSO and **6-5** for 2 h. Live cells were treated with **6-5** or 1% DMSO, fixed, permeabilized, reacted with rhodamine-azide (red), incubated with anti-Trypanopain, visualized by staining with FITC-conjugated goat anti-rabbit IgG antibody. All cells were counter-stained with DAPI to visualize kinetoplast and/or nucleus (blue). Phase contrast images and merged fluorescence images are shown as labels indicate. Scale bar represents 10 μm.

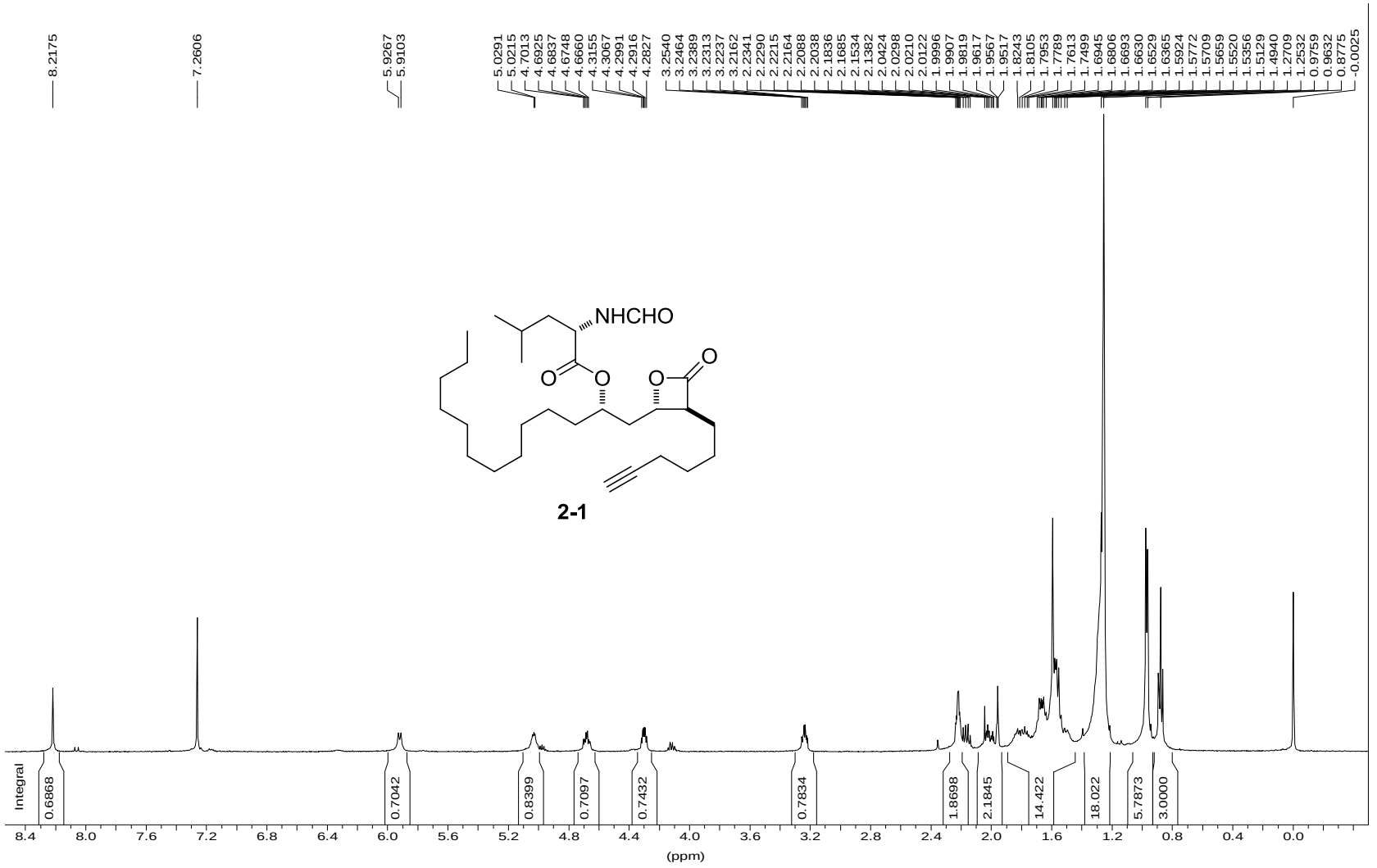
Appendix 2

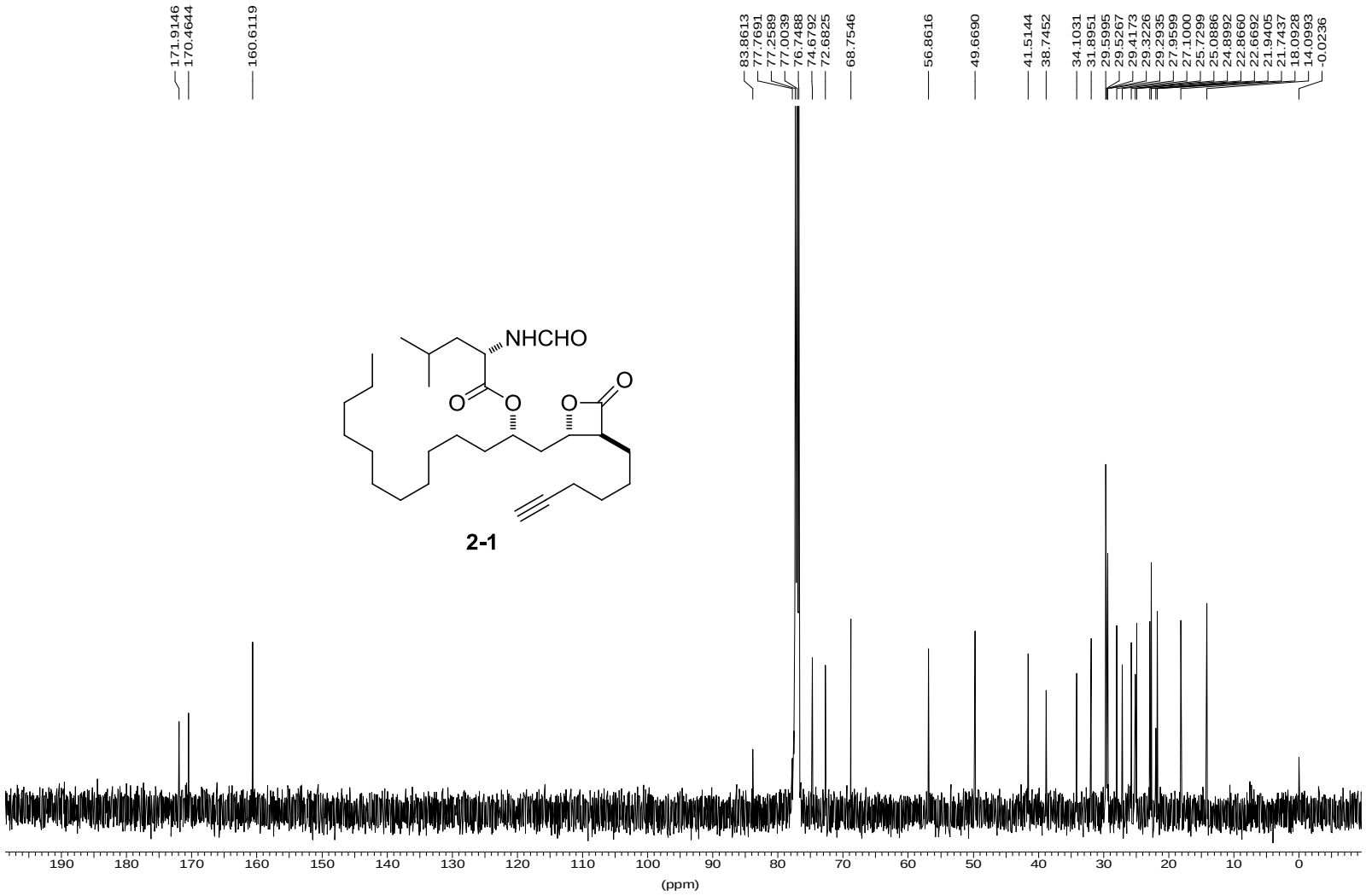
^{1}H– and ^{13}C NMR spectra of all the studied compounds (THL analogues; K11777 analogues; aza-nitrile– and diazomethyl-ketone containing inhibitors and probes) are listed in this Appendix.

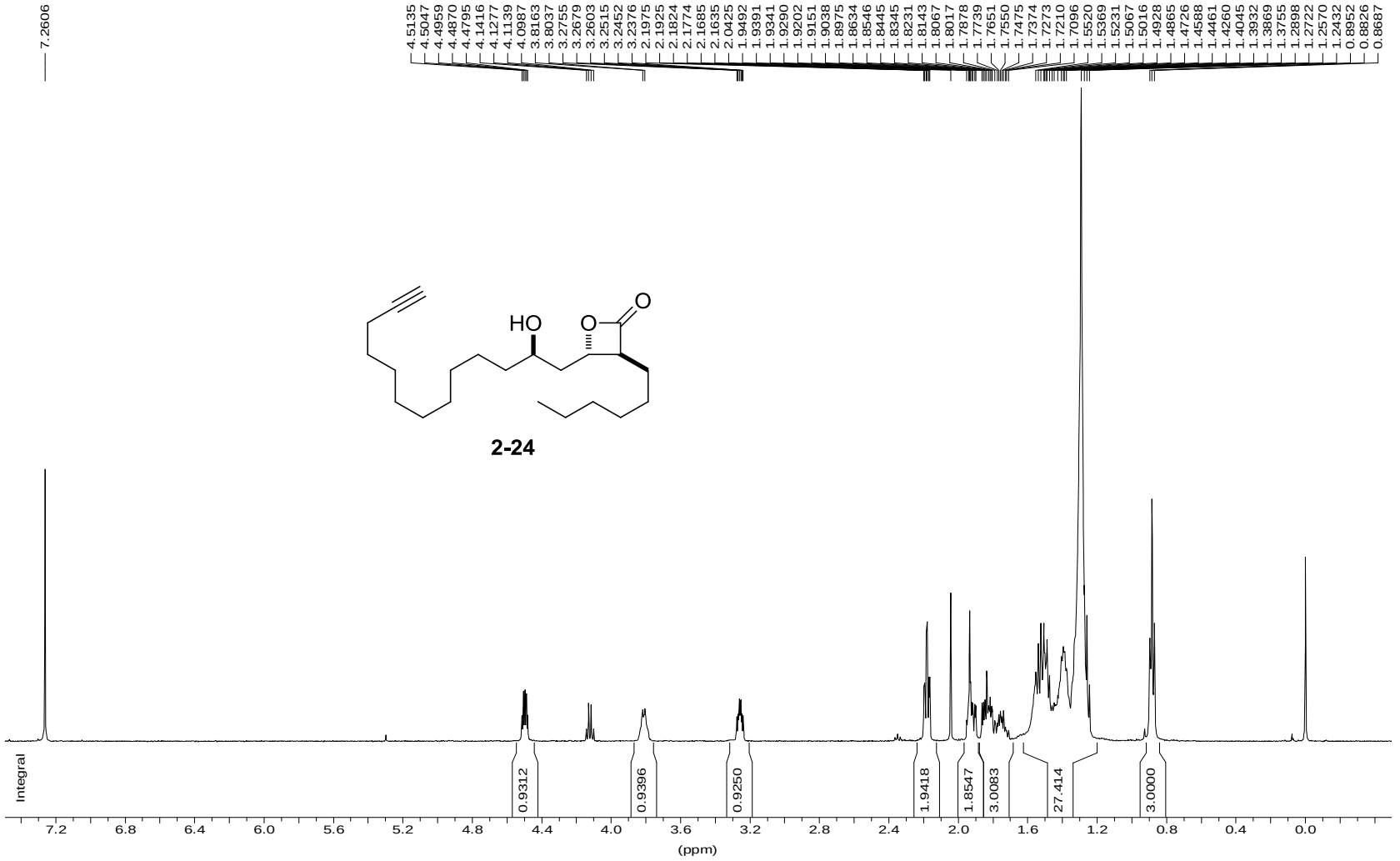
^{1}H– and ^{13}C NMR spectra of the intermediate compounds and representative LCMS of the studied compounds are listed on Appendix_CD.

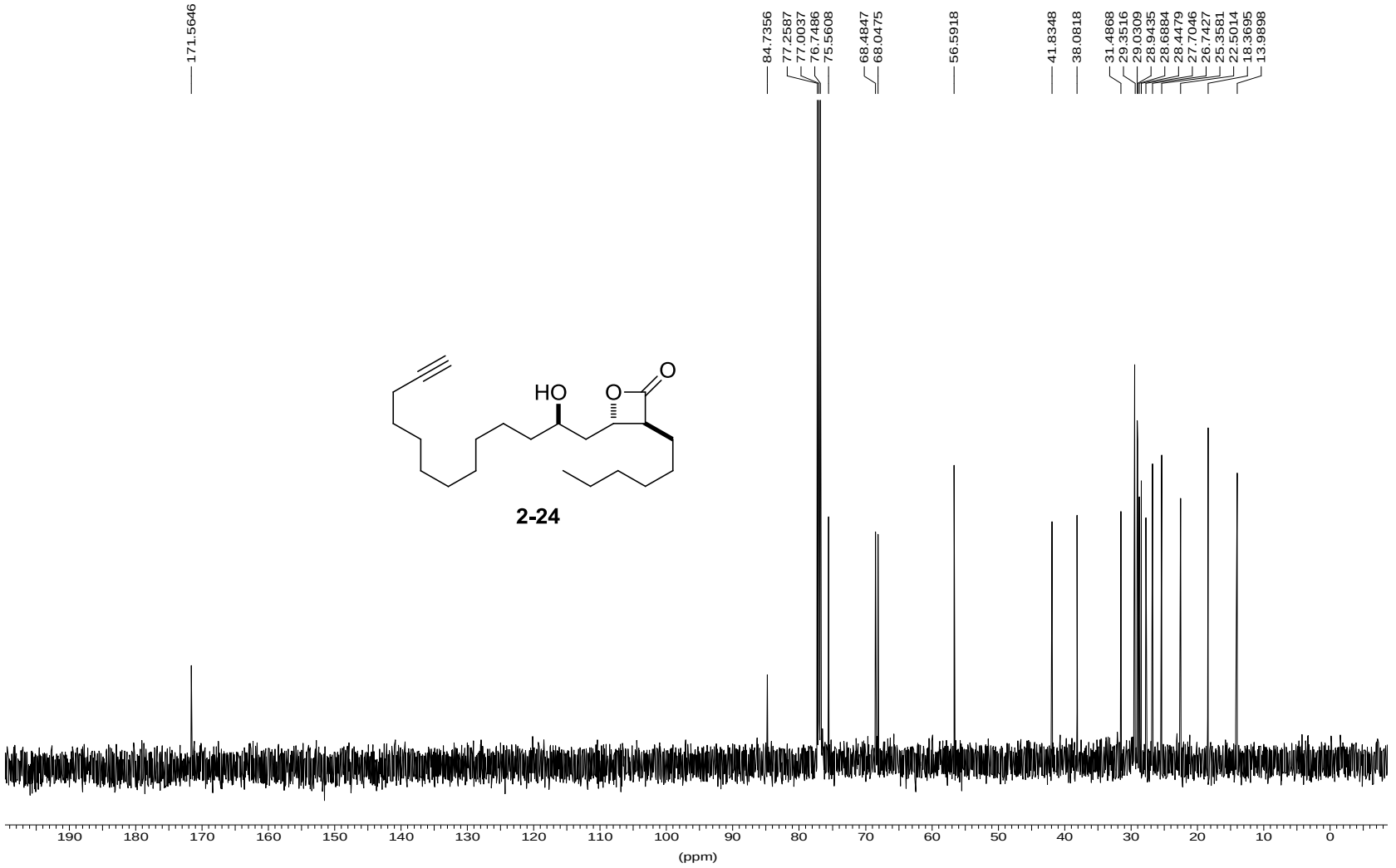

7.2606
4.5185
4.5097
4.5009
4.4933
4.4845
3.8327
3.8264
3.8201
3.8138
3.8075
3.8012
3.7949
3.7886
3.7823
3.3032
3.2956
3.2893
3.2792
3.2729
3.2654
2.2278
2.2227
2.2139
2.2089
2.2013
2.1963
1.9656
1.9593
1.9555
1.9504
1.9441
1.9366
1.9315
1.9202
1.9139
1.8672
1.8571
1.8470
1.8382
1.8281
1.8181
1.8105
1.7941
1.7752
1.6176
1.6012
1.5987
1.5873
1.5836
1.5760
1.5621
1.5495
1.5369
1.5268
1.5205
1.5130
1.5003
1.4865
1.4726
1.2986
1.2835
1.2608
0.8927
0.8788
0.8649
HO
O
O
2-13
Integral
0.9324
0.9429
0.9039
2.0000
1.8935
3.0763
7.7585
1.1400
17.078
3.0422
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

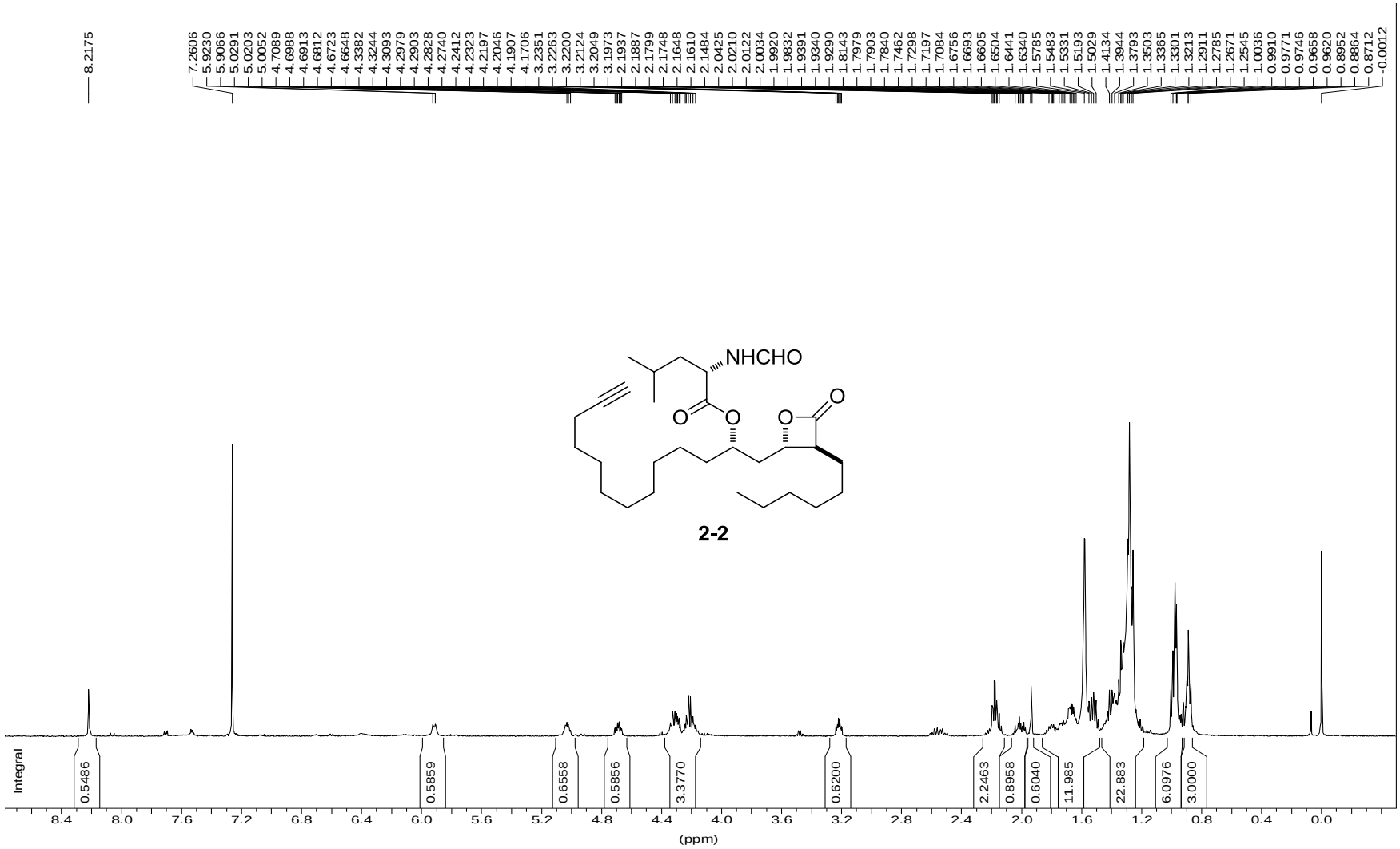

171.3661
83.9322
77.2715
77.0165
76.7614
75.6465
68.6870
68.5413
56.4442
41.7748
38.1748
31.9077
29.6340
29.6121
29.5538
29.4955
29.3279
27.9652
27.2365
25.8300
25.3855
22.6818
18.1345
14.1046
HO
O
O
2-13
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)

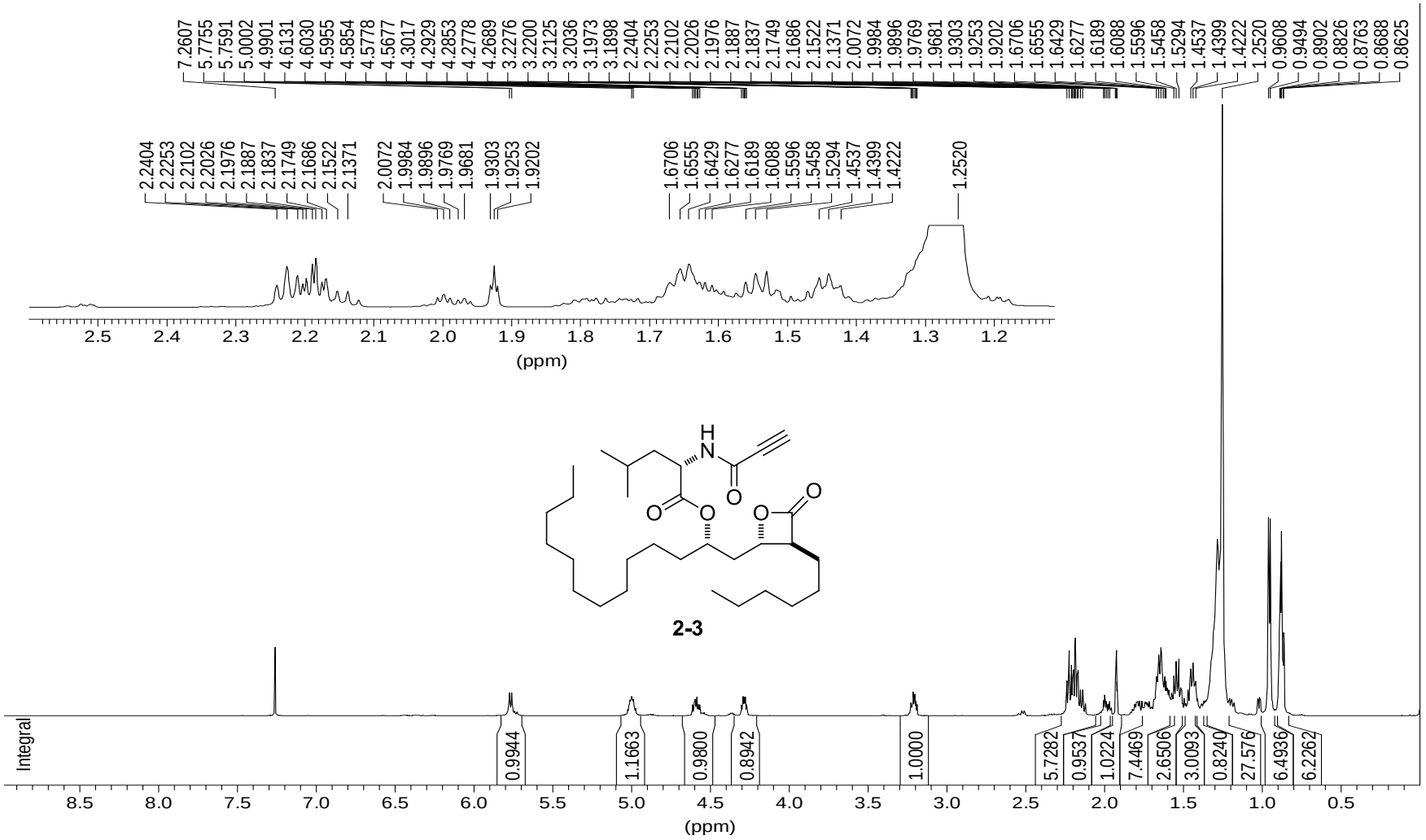

8.2175
7.2606
5.9267
5.9103
5.0291
5.0215
4.7013
4.6925
4.6837
4.6748
4.6660
4.3155
4.3067
4.2991
4.2916
4.2827
3.2540
3.2464
3.2389
3.2313
3.2237
3.2162
2.2341
2.2290
2.2215
2.2164
2.2088
2.2038
2.1836
2.1685
2.1534
2.1382
2.0424
2.0298
2.0210
2.0122
1.9996
1.9907
1.9819
1.9617
1.9567
1.9517
1.8243
1.8105
1.7953
1.7789
1.7613
1.7499
1.6945
1.6806
1.6693
1.6630
1.6529
1.6365
1.5924
1.5772
1.5709
1.5659
1.5520
1.5356
1.5129
1.4940
1.2709
1.2532
0.9759
0.9632
0.8775
-0.0025
NHCHO
O
O
O
O
2-1
Integral
0.6868
0.7042
0.8399
0.7097
0.7432
0.7834
1.8698
2.1845
14.422
18.022
5.7873
3.0000
8.4
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

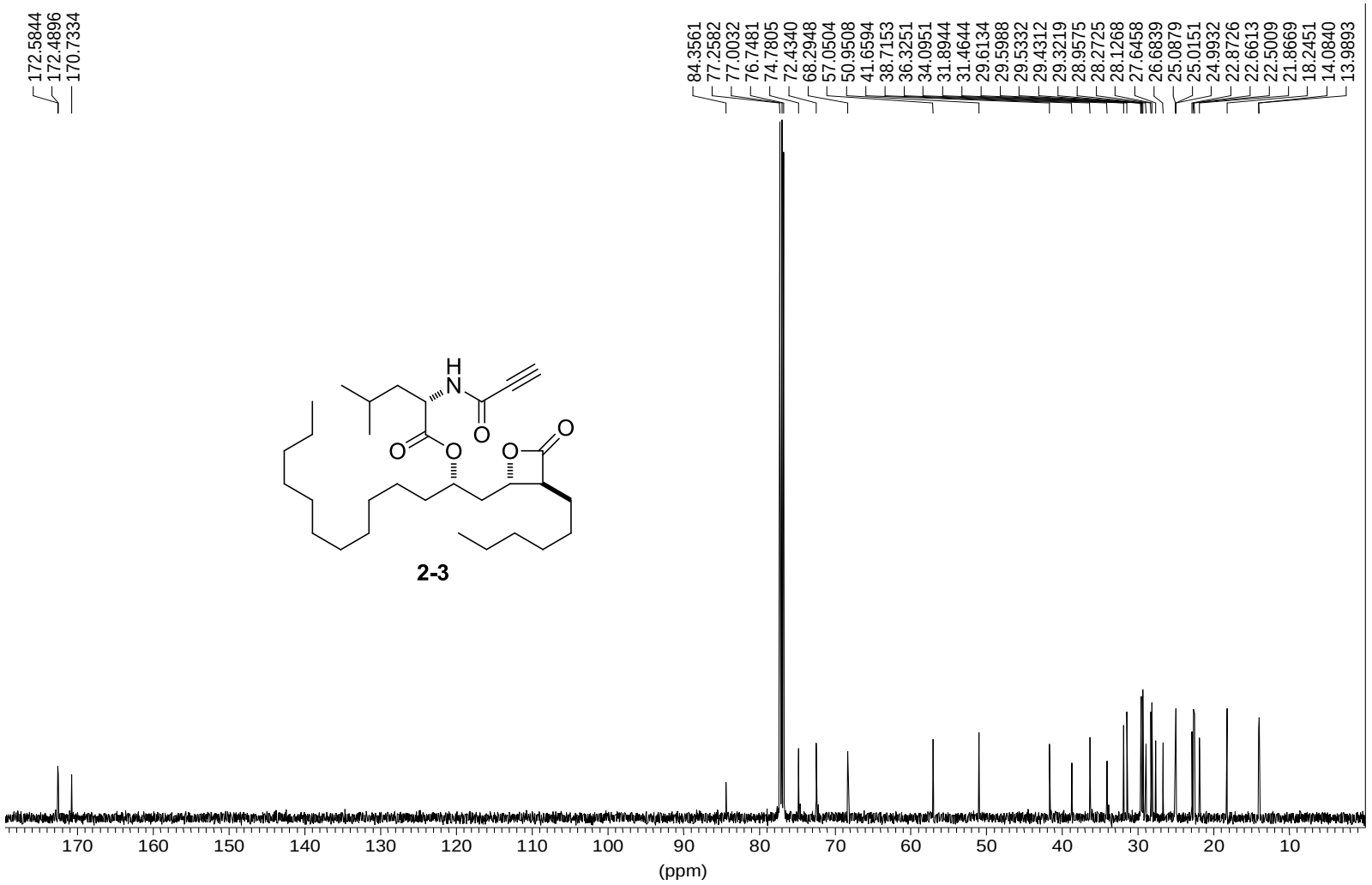

171.9146
170.4644
160.6119
83.8613
77.7691
77.2589
77.0039
76.7488
74.6792
72.6825
68.7546
56.8616
49.6690
41.5144
38.7452
34.1031
31.8951
29.5995
29.5267
29.4173
29.3226
29.2935
27.9599
27.1000
25.7299
25.0886
24.8992
22.8660
22.6692
21.9405
21.7437
18.0928
14.0993
-0.0236
NHCHO
O
O
O
O
2-1
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)

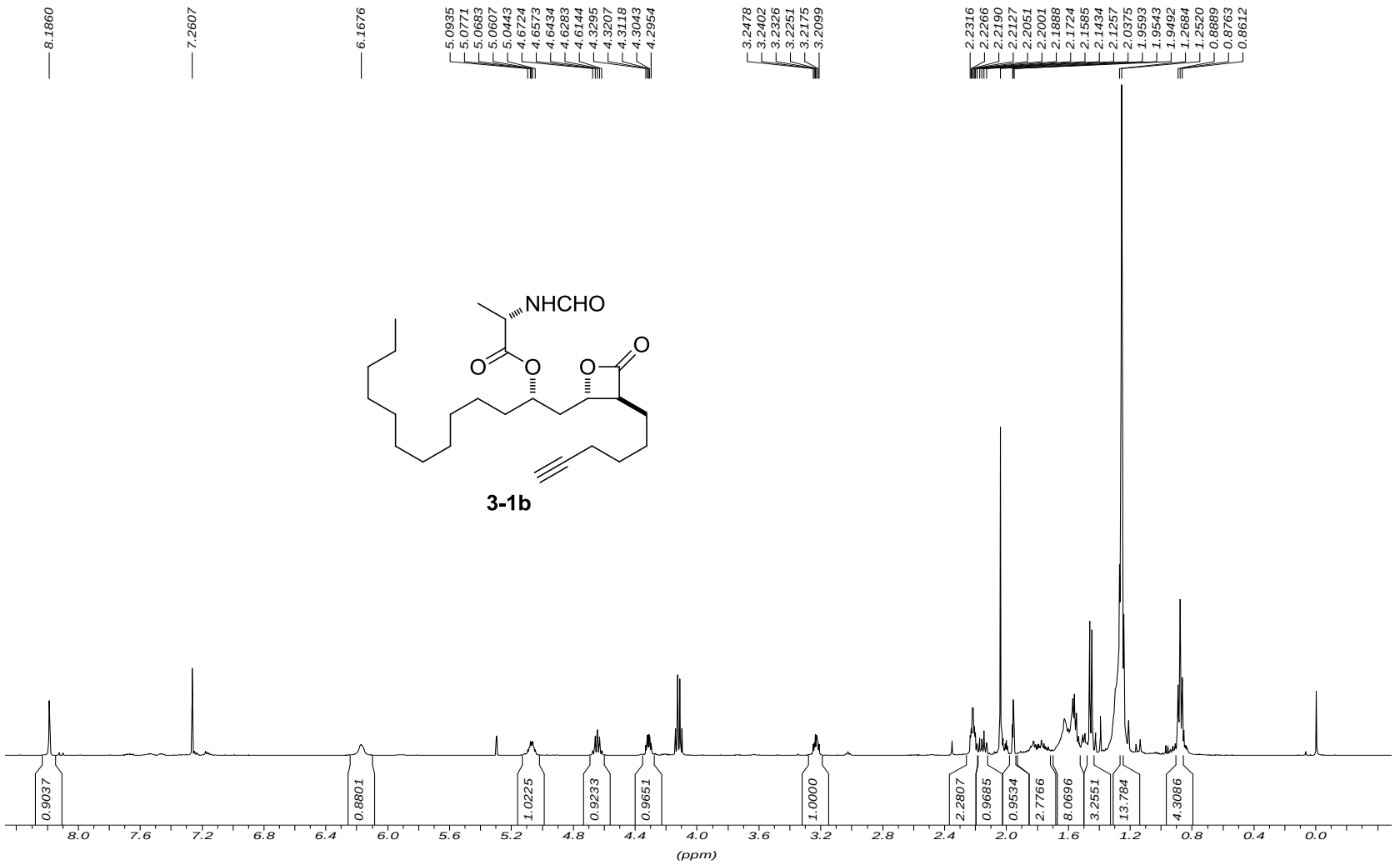

7.2606
4.5135
4.5047
4.4959
4.4870
4.4795
4.1416
4.1277
4.1139
4.0987
3.8163
3.8037
3.2755
3.2679
3.2603
3.2515
3.2452
3.2376
2.1975
2.1925
2.1824
2.1774
2.1685
2.1635
2.0425
1.9492
1.9391
1.9341
1.9290
1.9202
1.9151
1.9038
1.8975
1.8634
1.8546
1.8445
1.8345
1.8231
1.8143
1.8067
1.8017
1.7878
1.7739
1.7651
1.7550
1.7475
1.7374
1.7273
1.7210
1.7096
1.5520
1.5369
1.5231
1.5067
1.5016
1.4928
1.4865
1.4726
1.4588
1.4461
1.4260
1.4045
1.3932
1.3869
1.3755
1.2898
1.2722
1.2570
1.2432
0.8952
0.8826
0.8687
HO
O
O
2-24
Integral
0.9312
0.9396
0.9250
1.9418
1.8547
3.0083
27.414
3.0000
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

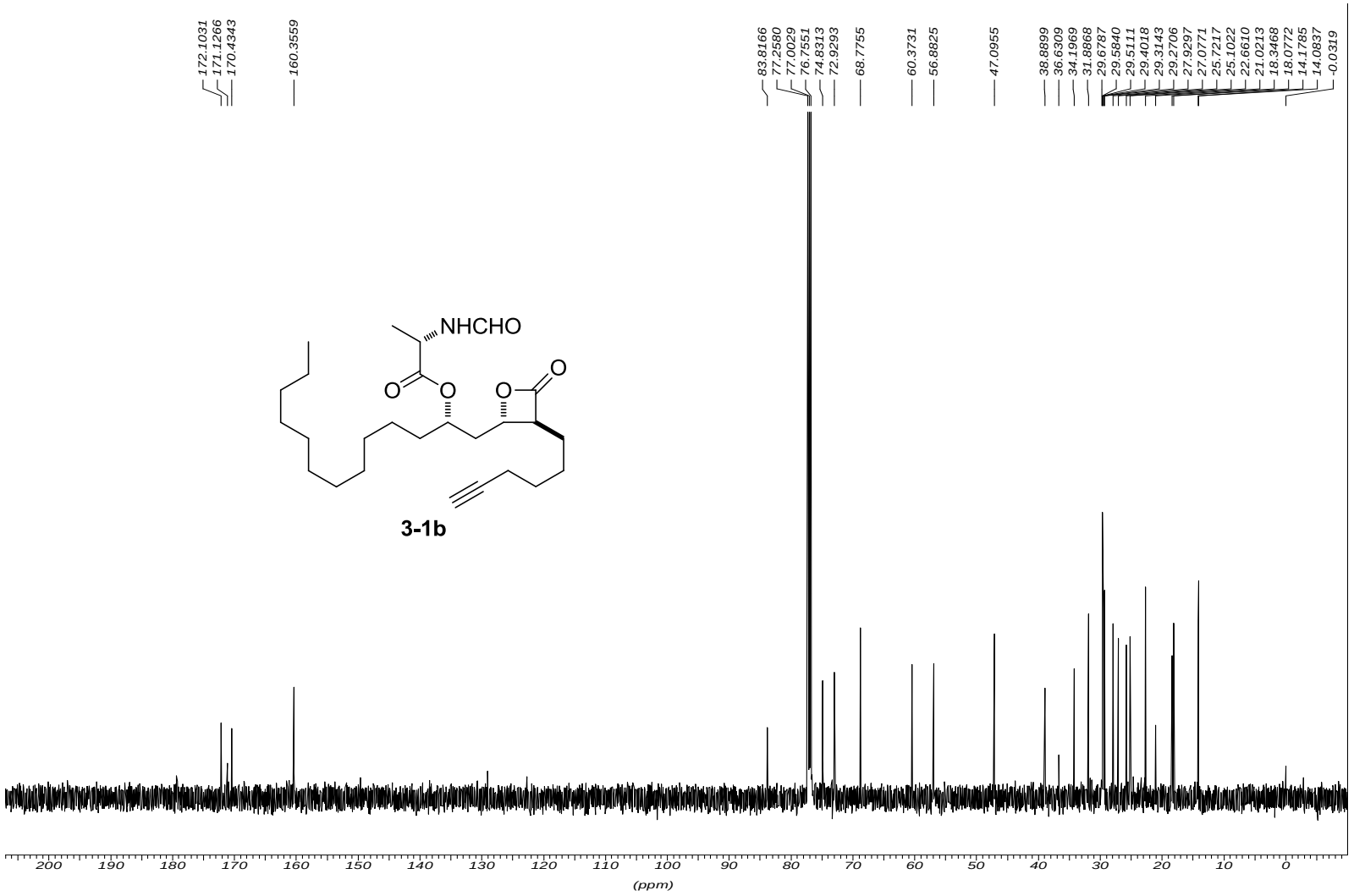

171.5646
84.7356
77.2587
77.0037
76.7486
75.5608
68.4847
68.0475
56.5918
41.8348
38.0818
31.4868
29.3516
29.0309
28.9435
28.6884
28.4479
27.7046
26.7427
25.3581
22.5014
18.3695
13.9898
HO
O
O
2-24
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)


8.2175
7.2606
5.9230
5.9066
5.0291
5.0203
5.0052
4.7089
4.6988
4.6913
4.6812
4.6723
4.6648
4.3382
4.3244
4.3093
4.2979
4.2903
4.2828
4.2740
4.2412
4.2323
4.2197
4.2046
4.1907
4.1706
3.2351
3.2263
3.2200
3.2124
3.2049
3.1973
2.1937
2.1887
2.1799
2.1748
2.1648
2.1610
2.1484
2.0425
2.0210
2.0122
2.0034
1.9920
1.9832
1.9391
1.9340
1.9290
1.8143
1.7979
1.7903
1.7840
1.7462
1.7298
1.7197
1.7084
1.6756
1.6693
1.6605
1.6504
1.6441
1.6340
1.5785
1.5483
1.5331
1.5193
1.5029
1.4134
1.3944
1.3793
1.3503
1.3365
1.3301
1.3213
1.2911
1.2785
1.2671
1.2545
1.0036
0.9910
0.9771
0.9746
0.9658
0.9620
0.8952
0.8864
0.8712
-0.0012
NHCHO
O
O
O
O
2-2
Integral
0.5486
0.5859
0.6558
0.5856
3.3770
0.6200
2.2463
0.8958
0.6040
11.985
22.883
6.0976
3.0000
8.4
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

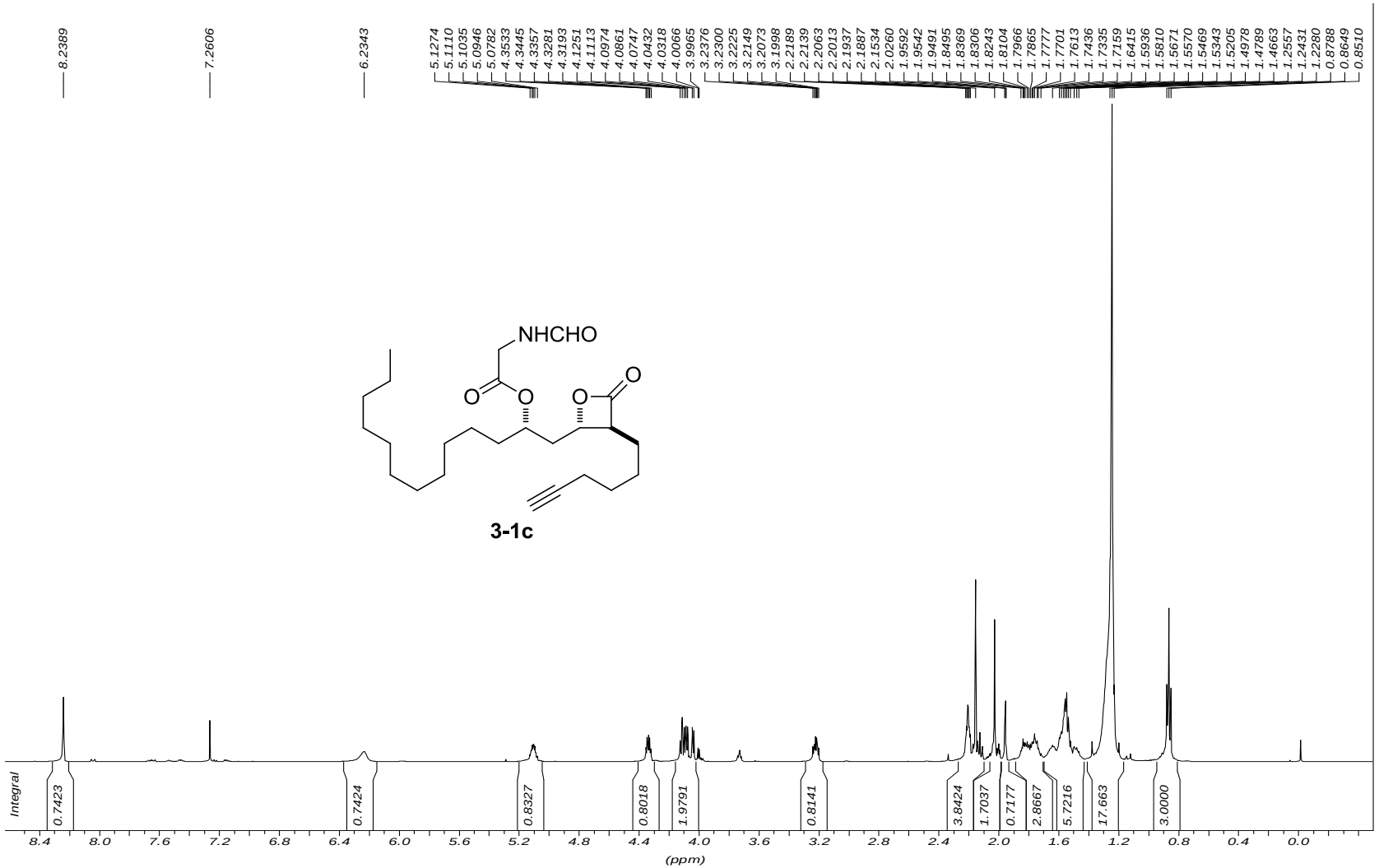

7.2607 5.7755 5.7591 5.0002 4.9901 4.6131 4.6030 4.5955 4.5854 4.5778 4.5677 4.3017 4.2929 4.2853 4.2778 4.2689 3.2276 3.2200 3.2125 3.2036 3.1973 3.1898 2.2404 2.2253 2.2102 2.2026 2.1976 2.1887 2.1837 2.1749 2.1686 2.1522 2.1371 2.0072 1.9984 1.9896 1.9769 1.9681 1.9303 1.9253 1.9202 1.6706 1.6555 1.6429 1.6277 1.6189 1.6088 1.5596 1.5458 1.5294 1.4537 1.4399 1.4222 1.2520 0.9608 0.9494 0.8902 0.8826 0.8763 0.8688 0.8625
2.5 2.4 2.3 2.2 2.1 2.0 1.9 1.8 1.7 1.6 1.5 1.4 1.3 1.2
(ppm)
H N O O O O O
2-3
Integral
0.9944 1.1663 0.9800 0.8942 1.0000 5.7282 0.9537 1.0224 7.4469 2.6506 3.0093 0.8240 27.576 6.4936 6.2262
8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
(ppm)

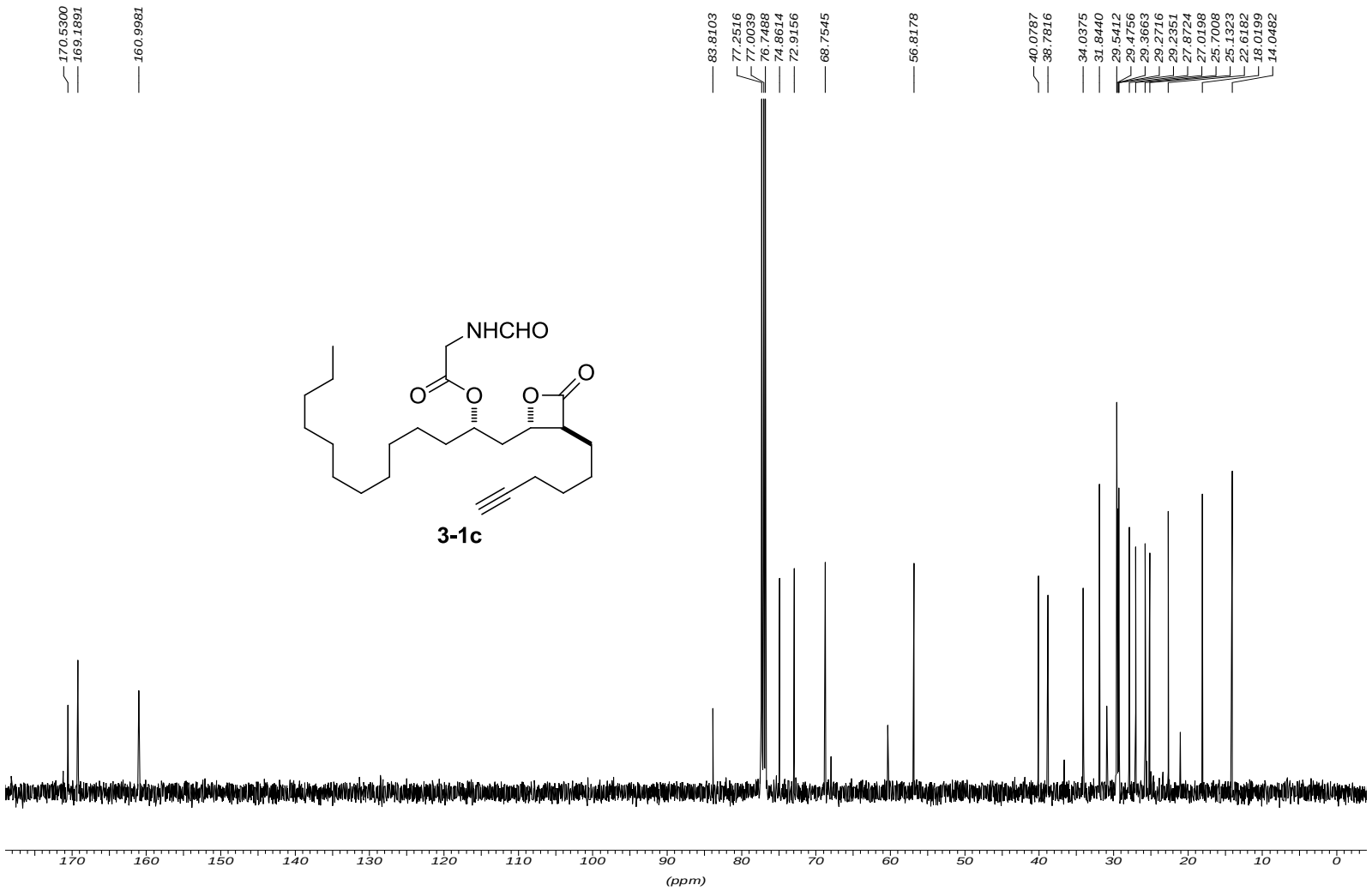

172.5844
172.4896
170.7334
84.3561
77.2582
77.0032
76.7481
74.7805
72.4340
68.2948
57.0504
50.9508
41.6594
38.7153
36.3251
34.0951
31.8944
31.4644
29.6134
29.5988
29.5332
29.4312
29.3219
28.9575
28.2725
28.1268
27.6458
26.6839
25.0879
25.0151
24.9932
22.8726
22.6613
22.5009
21.8669
18.2451
14.0840
13.9893
H
N
O
O
O
O
O
2-3
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
(ppm)

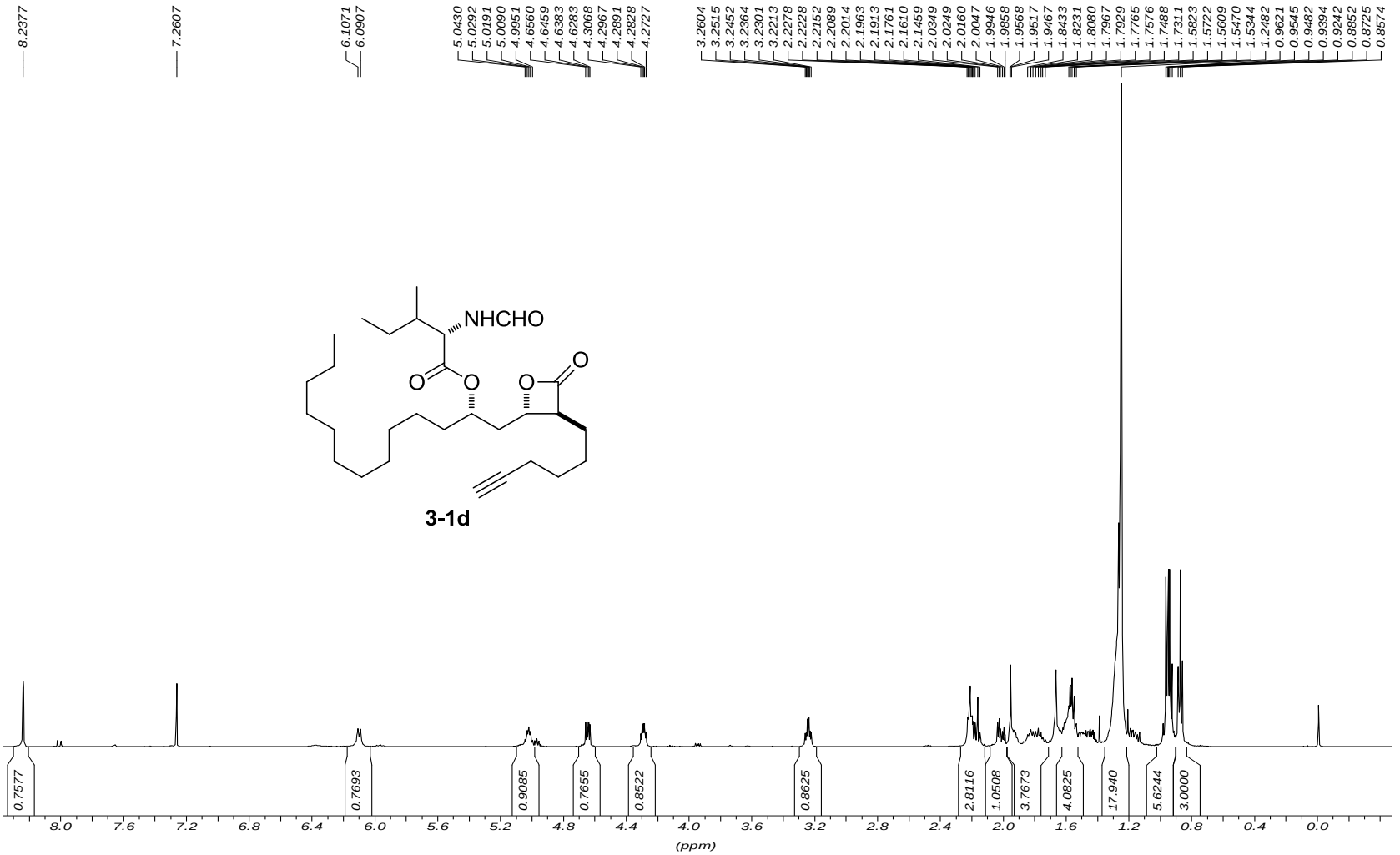

8.1860
7.2607
6.1676
5.0935
5.0771
5.0683
5.0607
5.0443
4.6724
4.6573
4.6434
4.6283
4.6144
4.3295
4.3207
4.3118
4.3043
4.2954
3.2478
3.2402
3.2326
3.2251
3.2175
3.2099
2.2316
2.2266
2.2190
2.2127
2.2051
2.2001
2.1888
2.1724
2.1585
2.1434
2.1257
2.0375
1.9593
1.9543
1.9492
1.2684
1.2520
0.8889
0.8763
0.8612
NHCHO
O
O
O
O
3-1b
0.9037
0.8801
1.0225
0.9233
0.9651
1.0000
2.2807
0.9685
0.9534
2.7766
8.0696
3.2551
13.784
4.3086
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

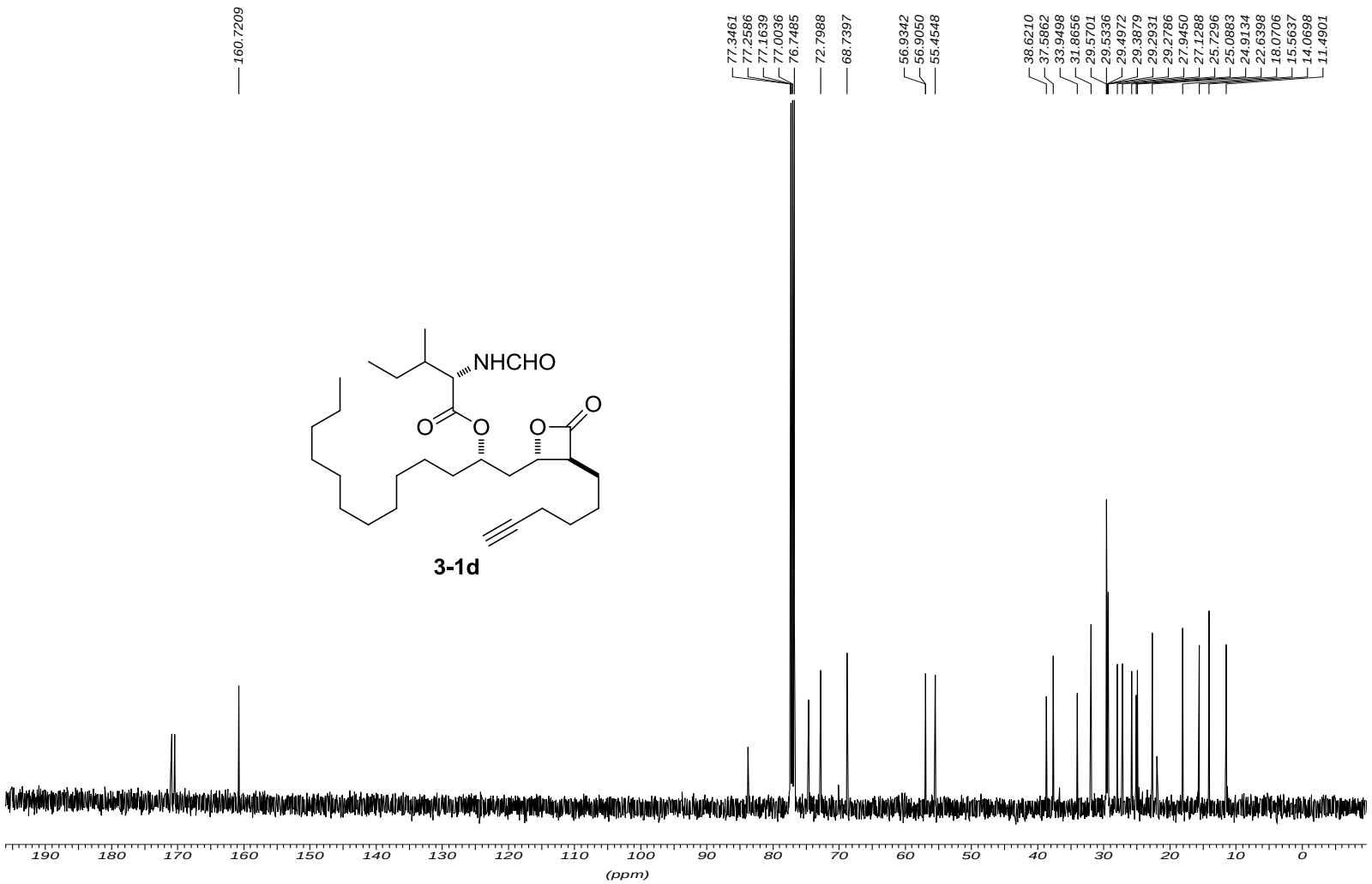

172.1031
171.1266
170.4343
160.3559
83.8166
77.2580
77.0029
76.7551
74.8313
72.9293
68.7755
60.3731
56.8825
47.0955
38.8899
36.6309
34.1969
31.8868
29.6787
29.5840
29.5111
29.4018
29.3143
29.2706
27.9297
27.0771
25.7217
25.1022
22.6610
21.0213
18.3468
18.0772
14.1785
14.0837
-0.0319
NHCHO
O
O
O
O
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)

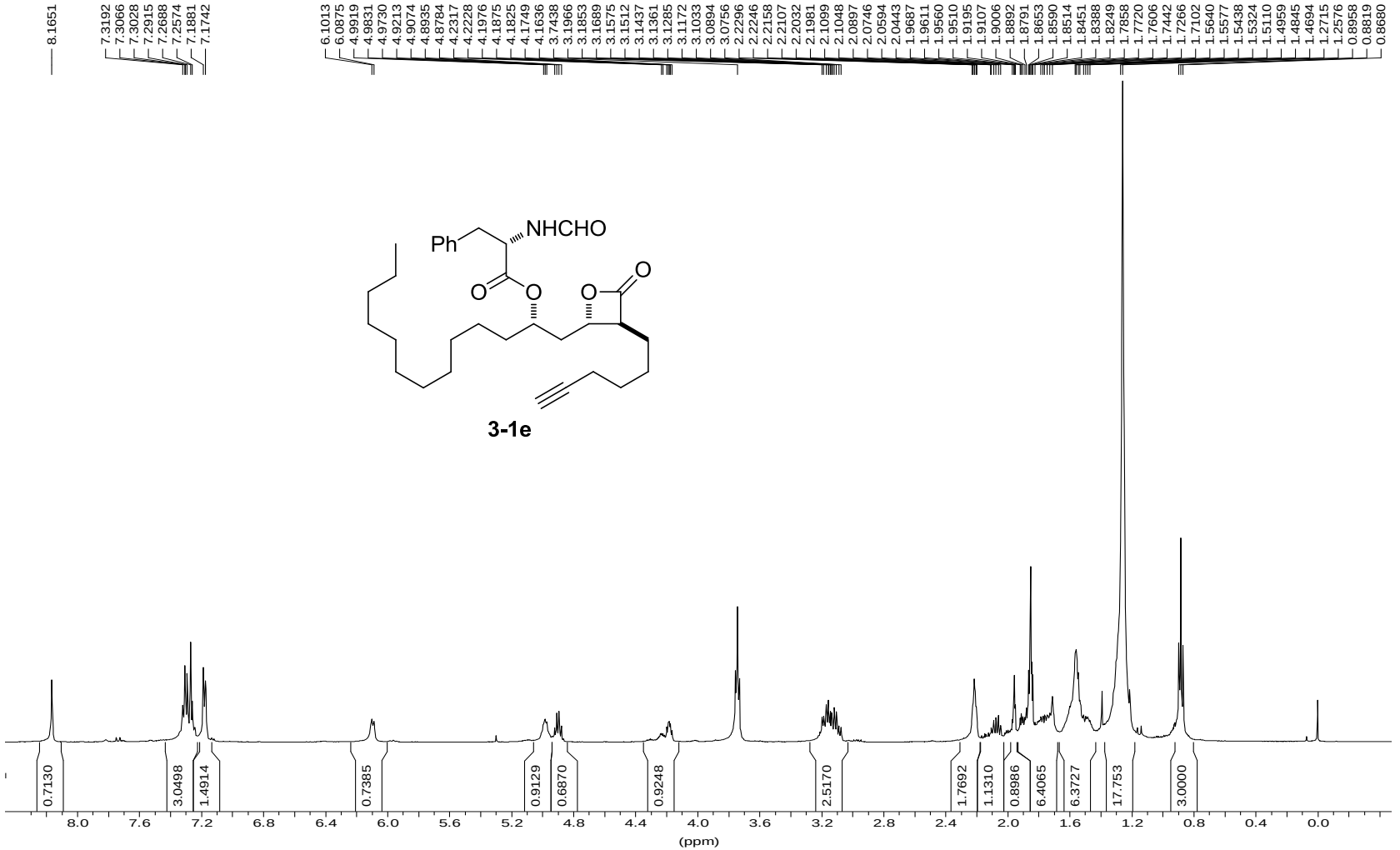
3-1b

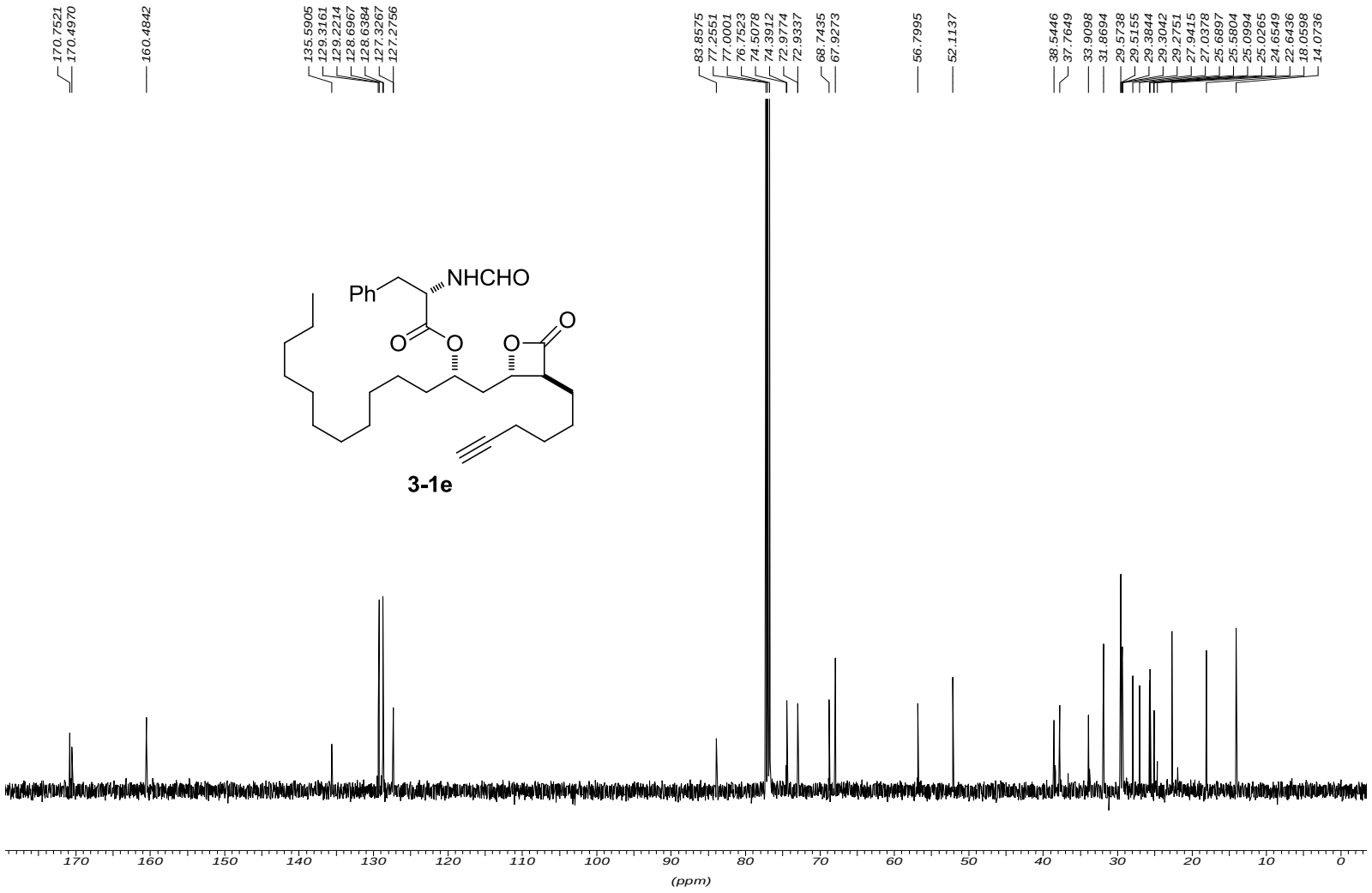

8.2389
7.2606
6.2343
5.1274
5.1110
5.1035
5.0946
5.0782
4.3533
4.3445
4.3357
4.3281
4.3193
4.1251
4.1113
4.0974
4.0861
4.0747
4.0432
4.0318
4.0066
3.9965
3.2376
3.2300
3.2225
3.2149
3.2073
3.1998
2.2189
2.2139
2.2063
2.2013
2.1937
2.1887
2.1534
2.0260
1.9592
1.9542
1.9491
1.8495
1.8369
1.8306
1.8243
1.8104
1.7966
1.7865
1.7777
1.7701
1.7613
1.7436
1.7335
1.7159
1.6415
1.5936
1.5810
1.5671
1.5570
1.5469
1.5343
1.5205
1.4978
1.4789
1.4663
1.2557
1.2431
1.2280
0.8788
0.8649
0.8510
NHCHO
O
O
O
O
3-1c
Integral
0.7423
0.7424
0.8327
0.8018
1.9791
0.8141
3.8424
1.7037
0.7177
2.8667
5.7216
17.663
3.0000
8.4
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

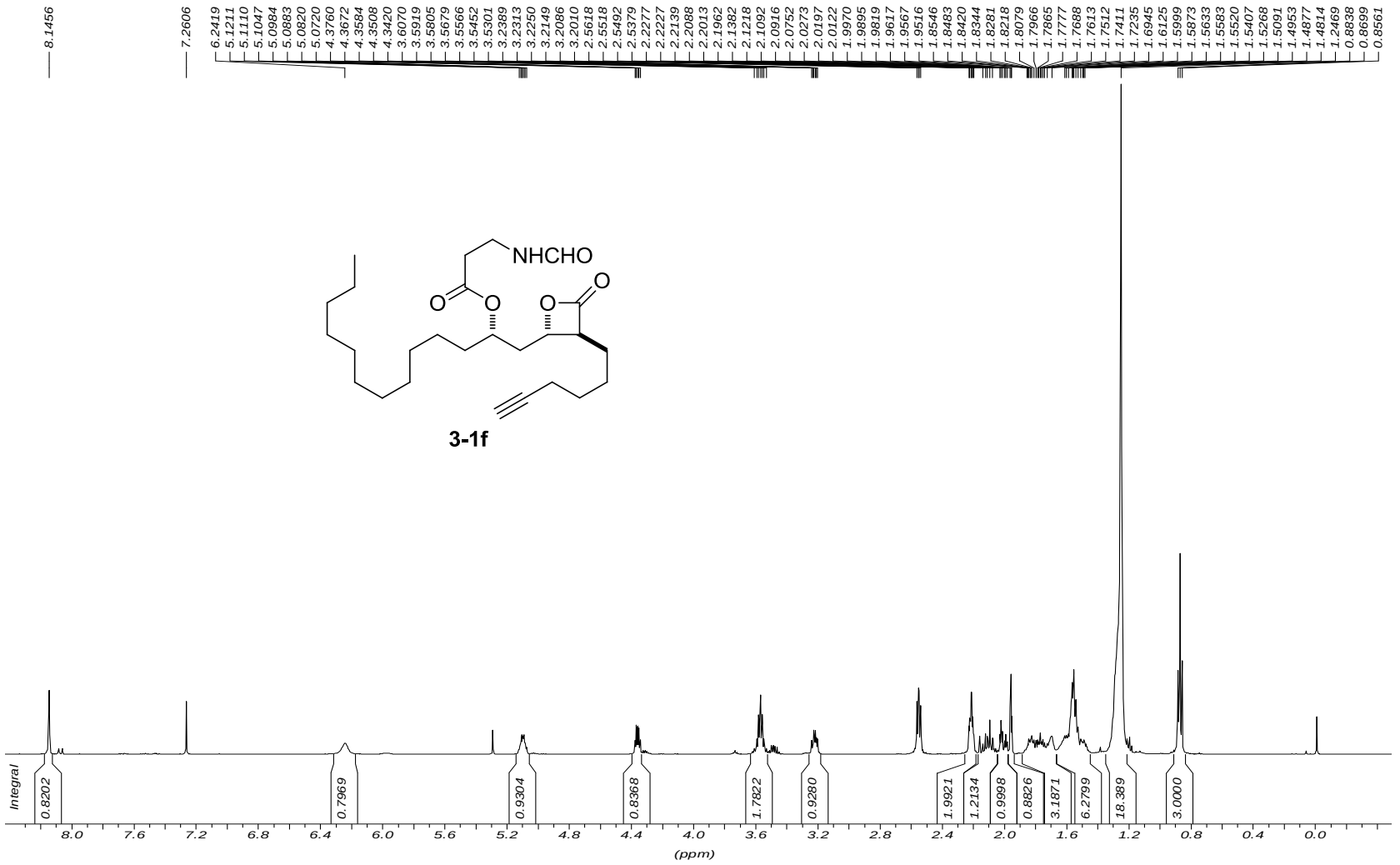

170.5300
169.1891
160.9981
83.8103
77.2516
77.0039
76.7488
74.8614
72.9156
68.7545
56.8178
40.0787
38.7816
34.0375
31.8440
29.5412
29.4756
29.3663
29.2716
29.2351
27.8724
27.0198
25.7008
25.1323
22.6182
18.0199
14.0482
NHCHO
O
O
O
O
3-1c
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)

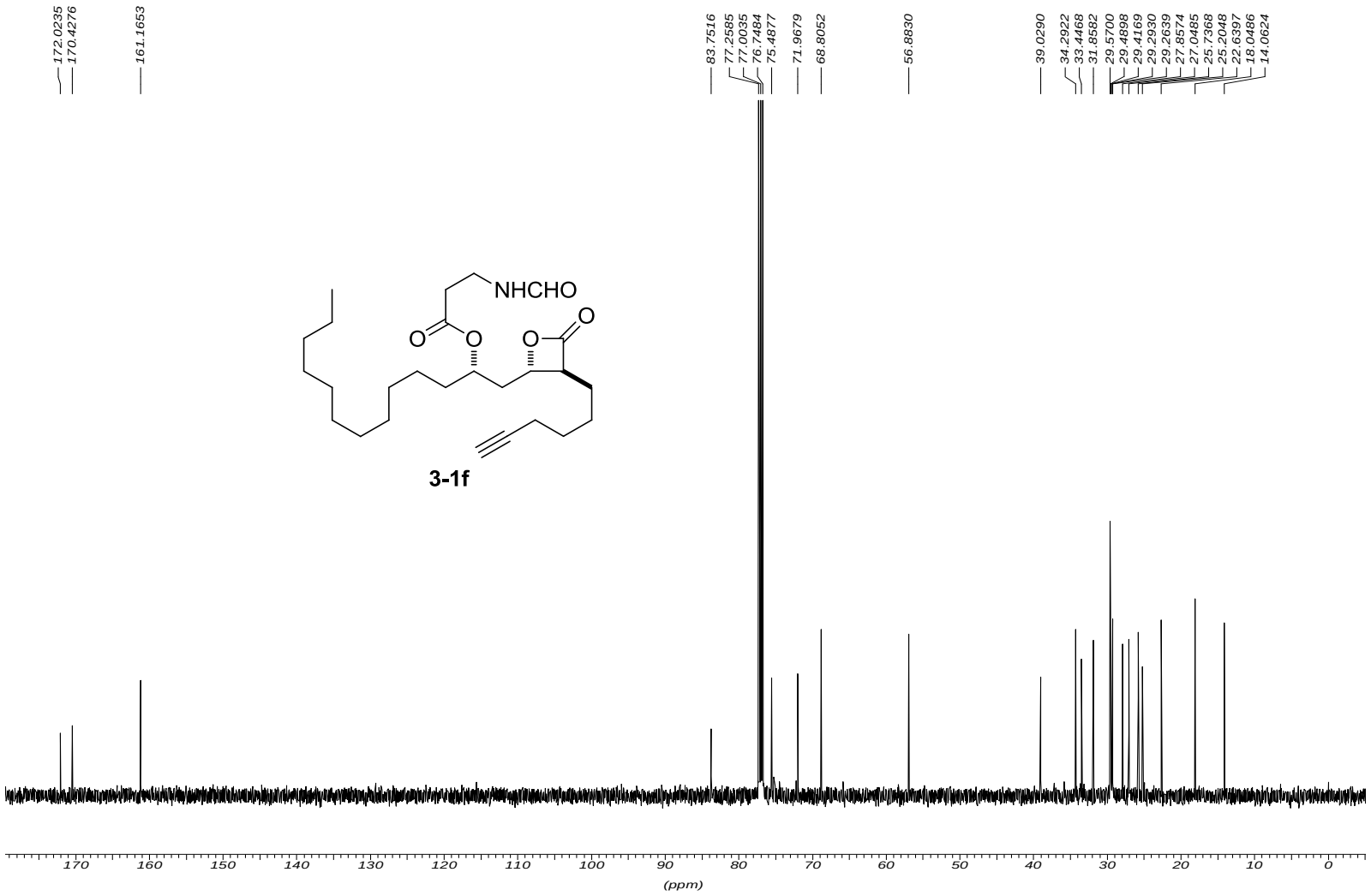

8.2377
7.2607
6.1071
6.0907
5.0430
5.0292
5.0191
5.0090
4.9951
4.6560
4.6459
4.6383
4.6283
4.3068
4.2967
4.2891
4.2828
4.2727
3.2604
3.2515
3.2452
3.2364
3.2301
3.2213
2.2278
2.2228
2.2152
2.2089
2.2014
2.1963
2.1913
2.1761
2.1610
2.1459
2.0349
2.0249
2.0160
2.0047
1.9946
1.9858
1.9568
1.9517
1.9467
1.8433
1.8231
1.8080
1.7967
1.7929
1.7765
1.7576
1.7488
1.7311
1.5823
1.5722
1.5609
1.5470
1.5344
1.2482
0.9621
0.9545
0.9482
0.9394
0.9242
0.8852
0.8725
0.8574
NHCHO
O
O
O
O
3-1d
0.7577
0.7693
0.9085
0.7655
0.8522
0.8625
2.8116
1.0508
3.7673
4.0825
17.940
5.6244
3.0000
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

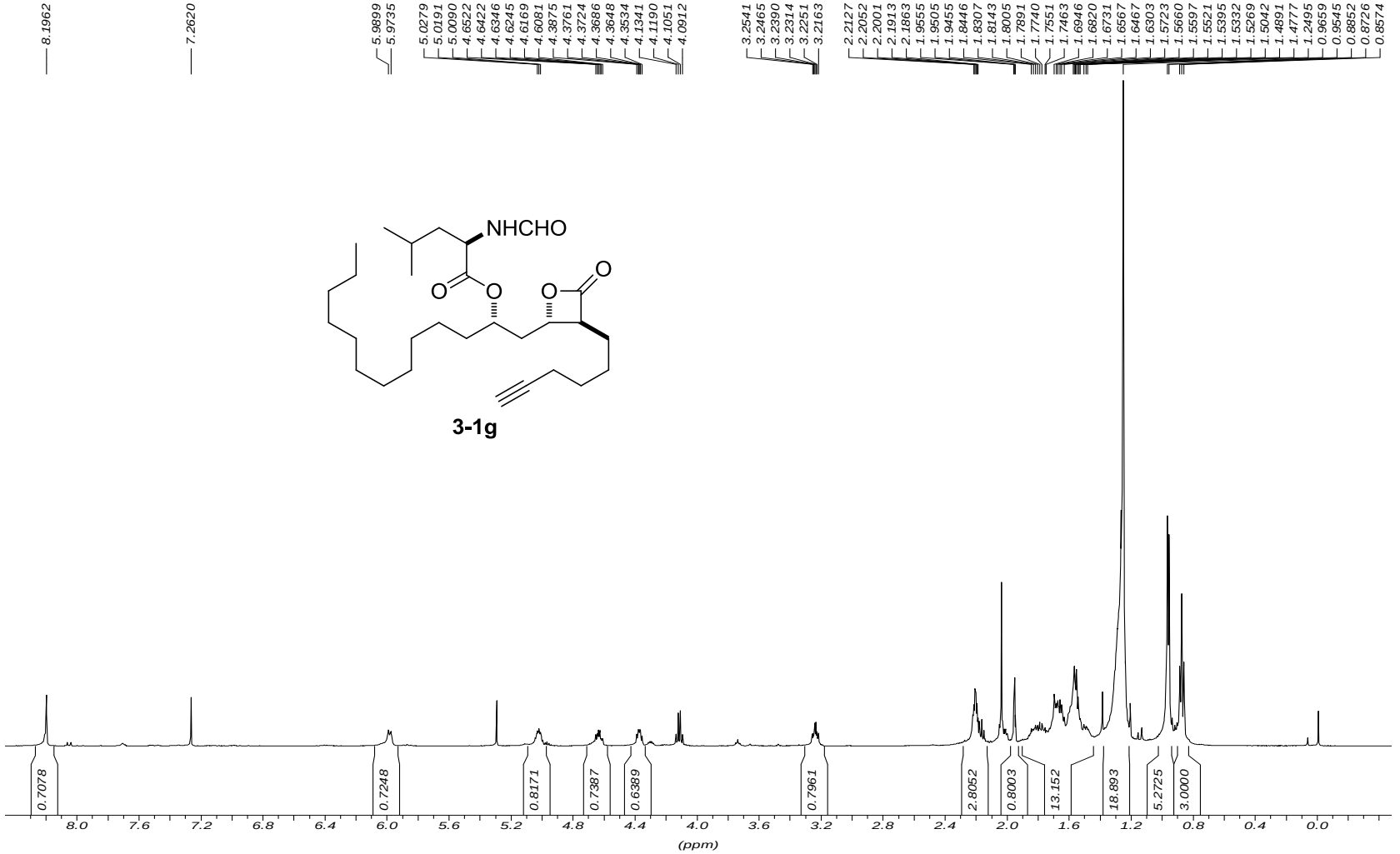

160.7209
77.3461
77.2586
77.1639
77.0036
76.7485
72.7988
68.7397
56.9342
56.9050
55.4548
38.6210
37.5862
33.9498
31.8656
29.5701
29.5336
29.4972
29.3879
29.2931
29.2786
27.9450
27.1288
25.7296
25.0883
24.9134
22.6398
18.0706
15.5637
14.0698
11.4901
NHCHO
O
O
O
O
3-1d
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)


8.1651
7.3192
7.3066
7.3028
7.2915
7.2688
7.2574
7.1881
7.1742
6.1013
6.0875
4.9919
4.9831
4.9730
4.9213
4.9074
4.8935
4.8784
4.2317
4.2228
4.1976
4.1875
4.1825
4.1749
4.1636
3.7438
3.1966
3.1853
3.1689
3.1575
3.1512
3.1437
3.1361
3.1285
3.1172
3.1033
3.0894
3.0756
2.2296
2.2246
2.2158
2.2107
2.2032
2.1981
2.1099
2.1048
2.0897
2.0746
2.0594
2.0443
1.9687
1.9611
1.9560
1.9510
1.9195
1.9107
1.9006
1.8892
1.8791
1.8653
1.8590
1.8514
1.8451
1.8388
1.8249
1.7858
1.7720
1.7606
1.7442
1.7266
1.7102
1.5640
1.5577
1.5438
1.5324
1.5110
1.4959
1.4845
1.4694
1.2715
1.2576
0.8958
0.8819
0.8680
Ph
NHCHO
O
O
O
O
3-1e
0.7130
3.0498
1.4914
0.7385
0.9129
0.6870
0.9248
2.5170
1.7692
1.1310
0.8986
6.4065
6.3727
17.753
3.0000
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

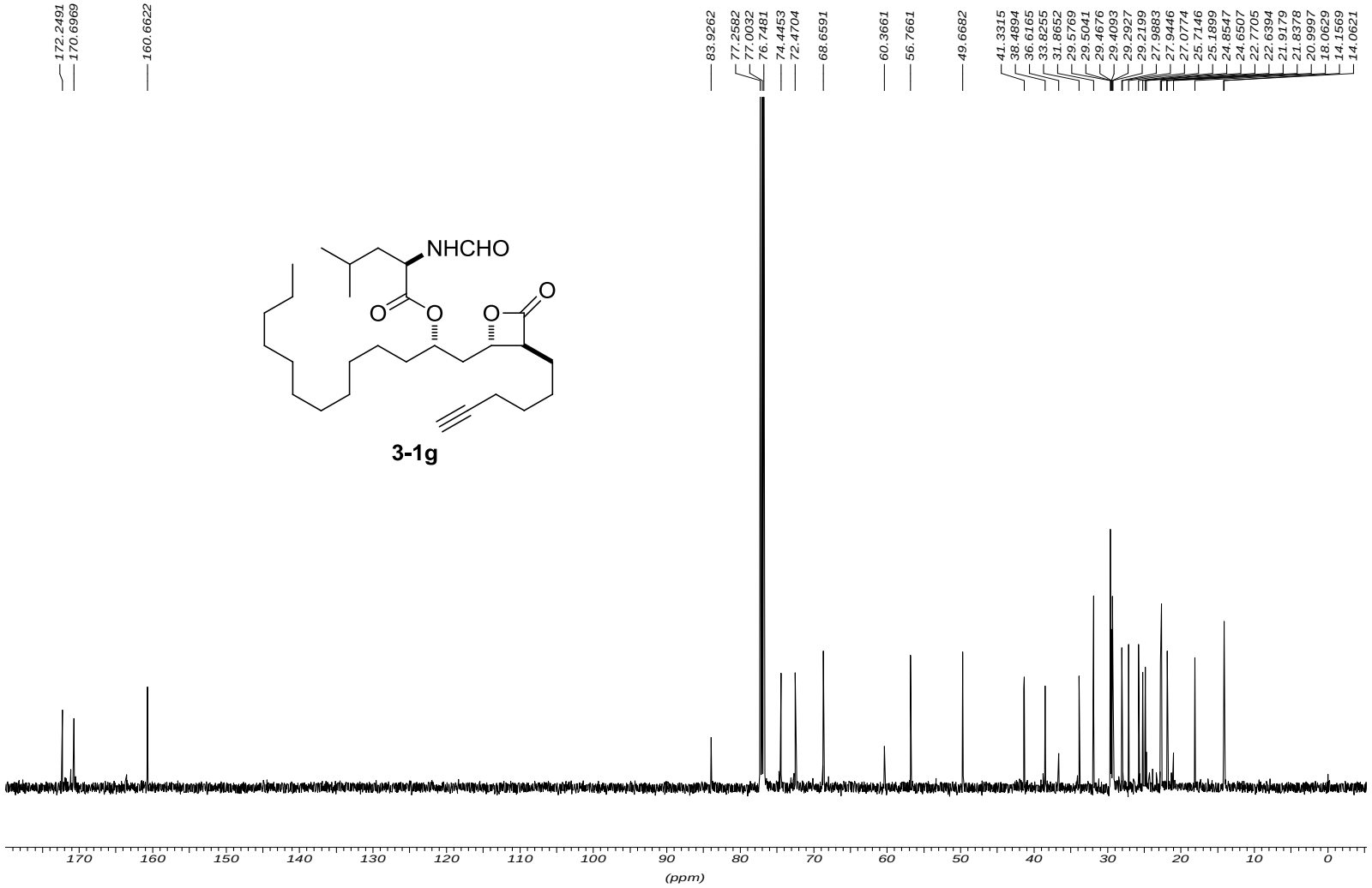

170.7521
170.4970
160.4842
135.5905
129.3161
129.2214
128.6967
128.6384
127.3267
127.2756
83.8575
77.2551
77.0001
76.7523
74.5078
74.3912
72.9774
72.9337
68.7435
67.9273
56.7995
52.1137
38.5446
37.7649
33.9098
31.8694
29.5738
29.5155
29.3844
29.3042
29.2751
27.9415
27.0378
25.6897
25.5804
25.0994
25.0265
24.6549
22.6436
18.0598
14.0736
NHCHO
Ph
O
O
O
O
3-1e
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)


8.1456
7.2606
6.2419
5.1211
5.1110
5.1047
5.0984
5.0883
5.0820
5.0720
4.3760
4.3672
4.3584
4.3508
4.3420
3.6070
3.5919
3.5805
3.5679
3.5566
3.5452
3.5301
3.2389
3.2313
3.2250
3.2149
3.2086
3.2010
2.5618
2.5518
2.5492
2.5379
2.2277
2.2227
2.2139
2.2088
2.2013
2.1962
2.1382
2.1218
2.1092
2.0916
2.0752
2.0273
2.0197
2.0122
1.9970
1.9895
1.9819
1.9617
1.9567
1.9516
1.8546
1.8483
1.8420
1.8344
1.8281
1.8218
1.8079
1.7966
1.7865
1.7777
1.7688
1.7613
1.7512
1.7411
1.7235
1.6945
1.6125
1.5999
1.5873
1.5633
1.5583
1.5520
1.5407
1.5268
1.5091
1.4953
1.4877
1.4814
1.2469
0.8838
0.8699
0.8561
NHCHO
O
O
O
O
3-1f
Integral
0.8202
0.7969
0.9304
0.8368
1.7822
0.9280
1.9921
1.2134
0.9998
0.8826
3.1871
6.2799
18.389
3.0000
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

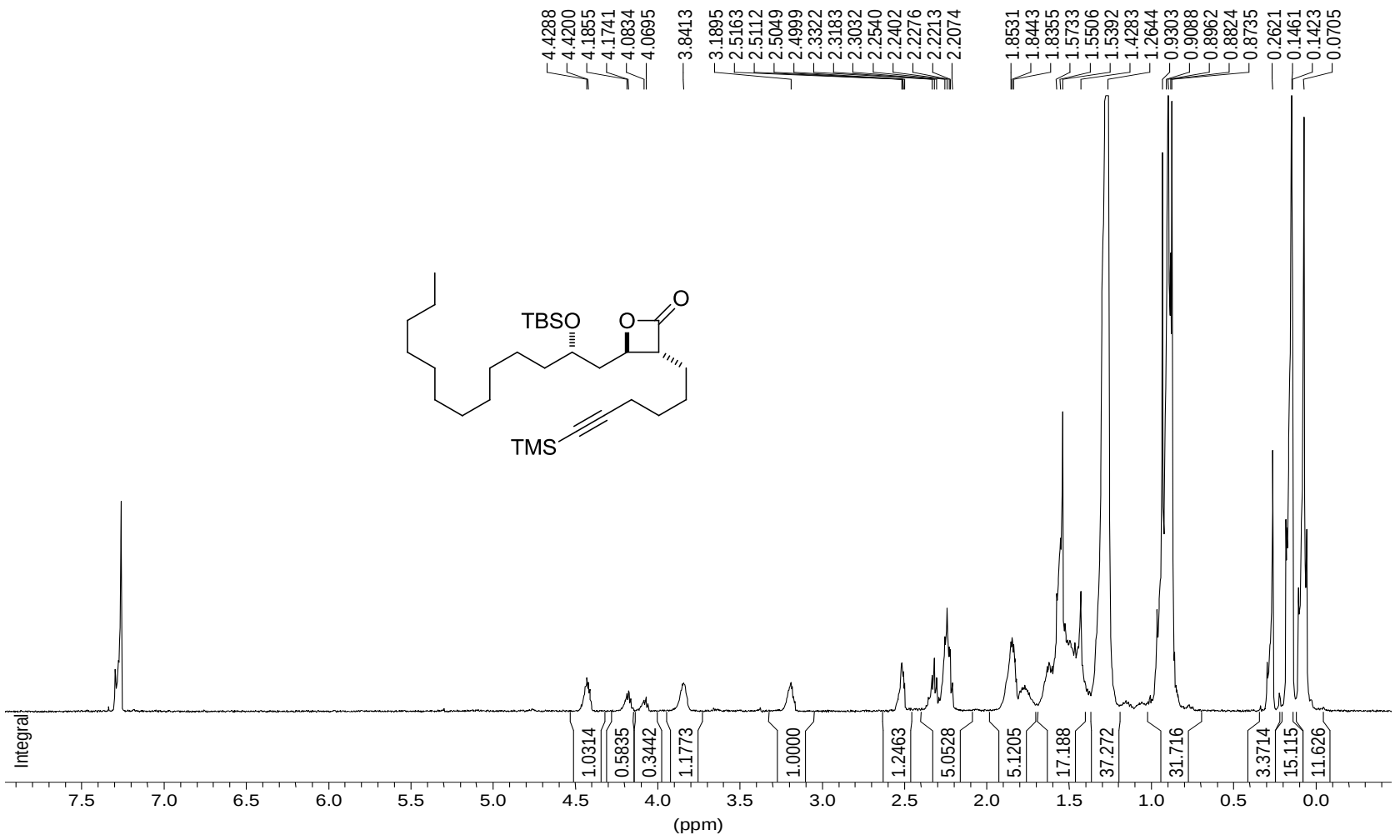

172.0235
170.4276
161.1653
83.7516
77.2585
77.0035
76.7484
75.4877
71.9679
68.8052
56.8830
39.0290
34.2922
33.4468
31.8582
29.5700
29.4898
29.4169
29.2930
29.2639
27.8574
27.0485
25.7368
25.2048
22.6397
18.0486
14.0624
NHCHO
O
O
O
O
3-1f
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)

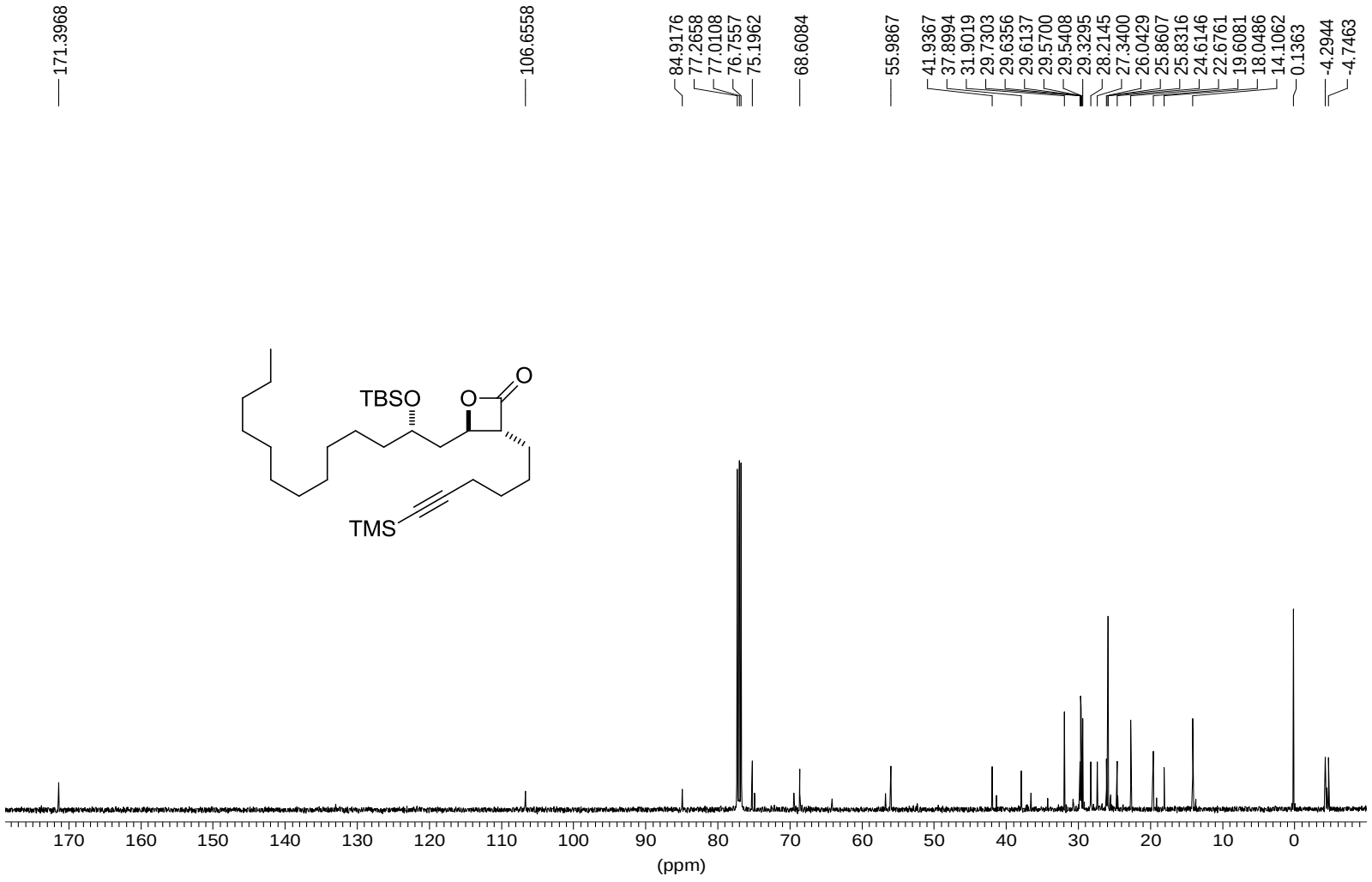

8.1962
7.2620
5.9899
5.9735
5.0279
5.0191
5.0090
4.6522
4.6422
4.6346
4.6245
4.6169
4.6081
4.3875
4.3761
4.3724
4.3686
4.3648
4.3534
4.1341
4.1190
4.1051
4.0912
3.2541
3.2465
3.2390
3.2314
3.2251
3.2163
2.2127
2.2052
2.2001
2.1913
2.1863
1.9555
1.9505
1.9455
1.8446
1.8307
1.8143
1.8005
1.7891
1.7740
1.7551
1.7463
1.6946
1.6820
1.6731
1.6567
1.6467
1.6303
1.5723
1.5660
1.5597
1.5521
1.5395
1.5332
1.5269
1.5042
1.4891
1.4777
1.2495
0.9659
0.9545
0.8852
0.8726
0.8574
NHCHO
O
O
O
O
0.7078
0.7248
0.8171
0.7387
0.6389
0.7961
2.8052
0.8003
13.152
18.893
5.2725
3.0000
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)

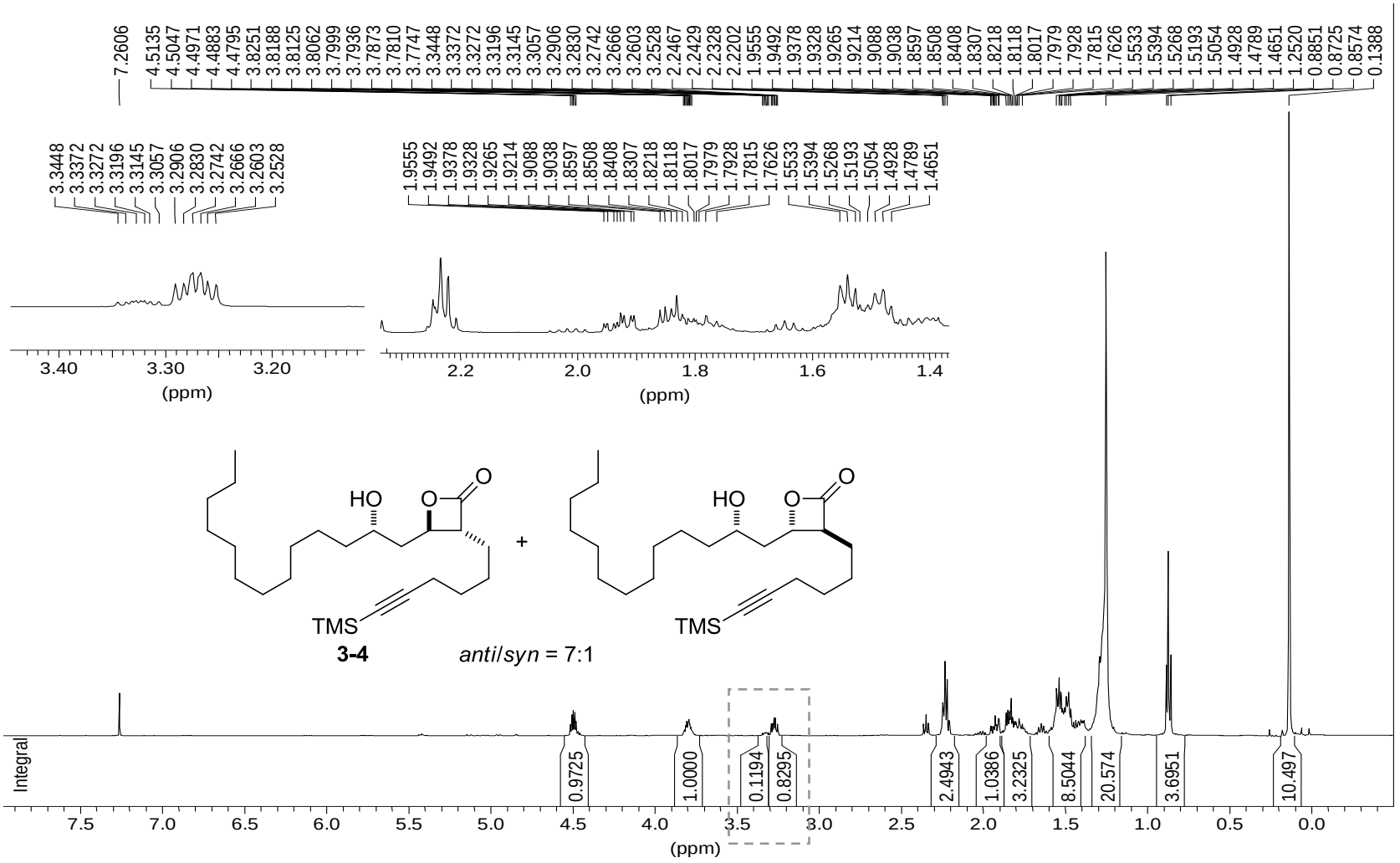
3-1g

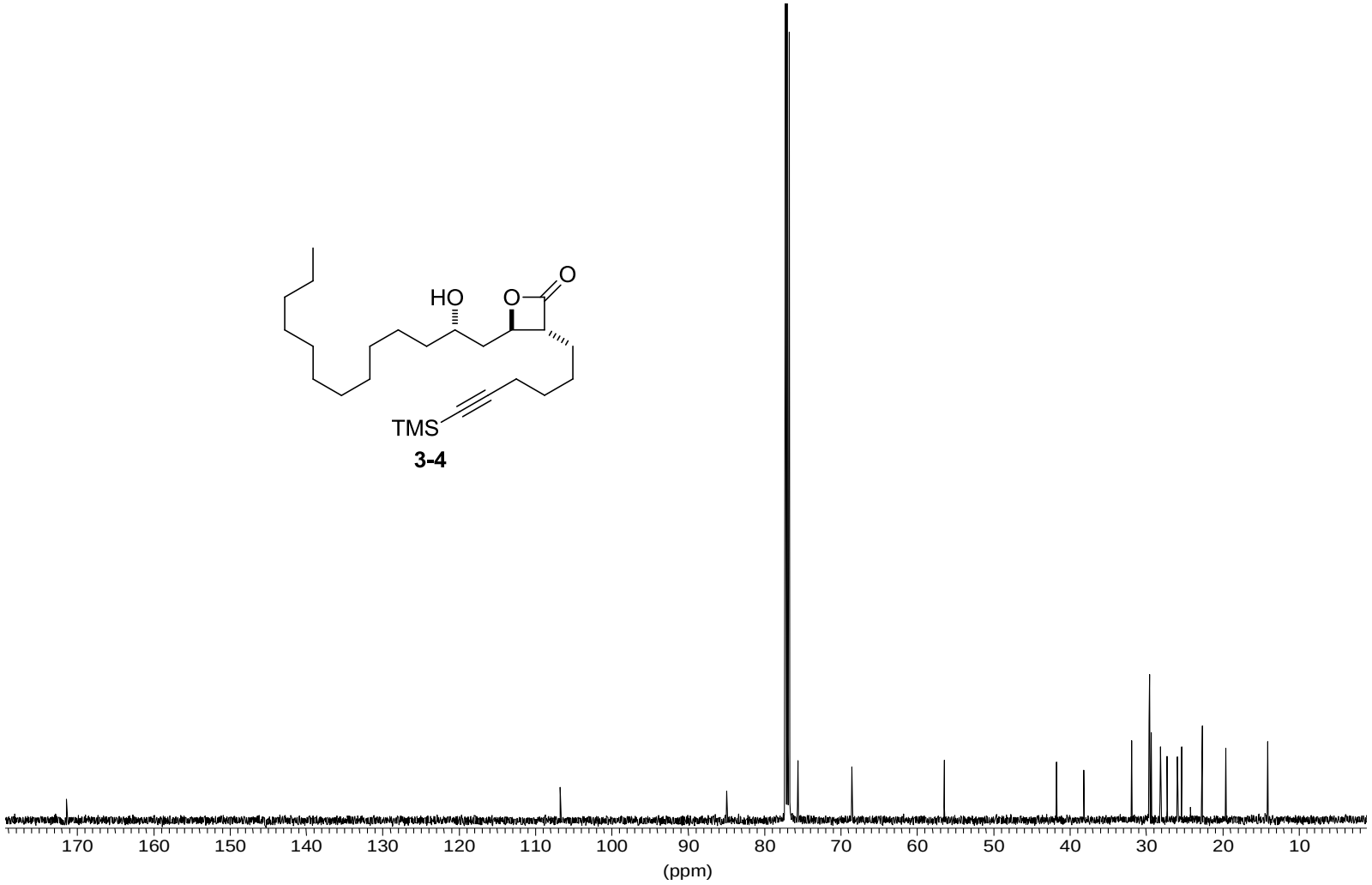

172.2491
170.6969
160.6622
83.9262
77.2582
77.0032
76.7481
74.4453
72.4704
68.6591
60.3661
56.7661
49.6682
41.3315
38.4894
36.6165
33.8255
31.8652
29.5769
29.5041
29.4676
29.4093
29.2927
29.2199
27.9883
27.9446
27.0774
25.7146
25.1899
24.8547
24.6507
22.7705
22.6394
21.9179
21.8378
20.9997
18.0629
14.1569
14.0621
NHCHO
O
O
O
O
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)

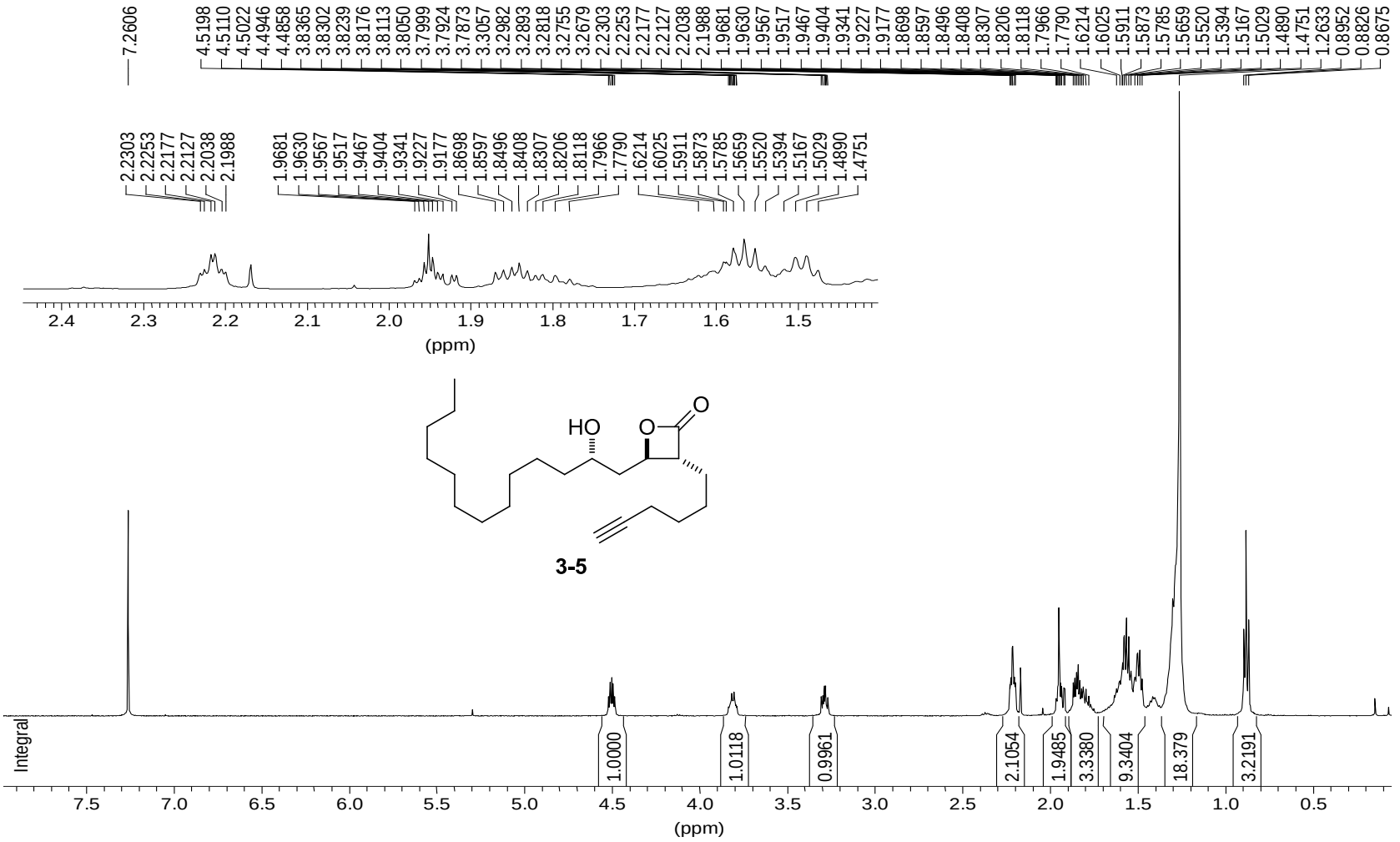
3-1g

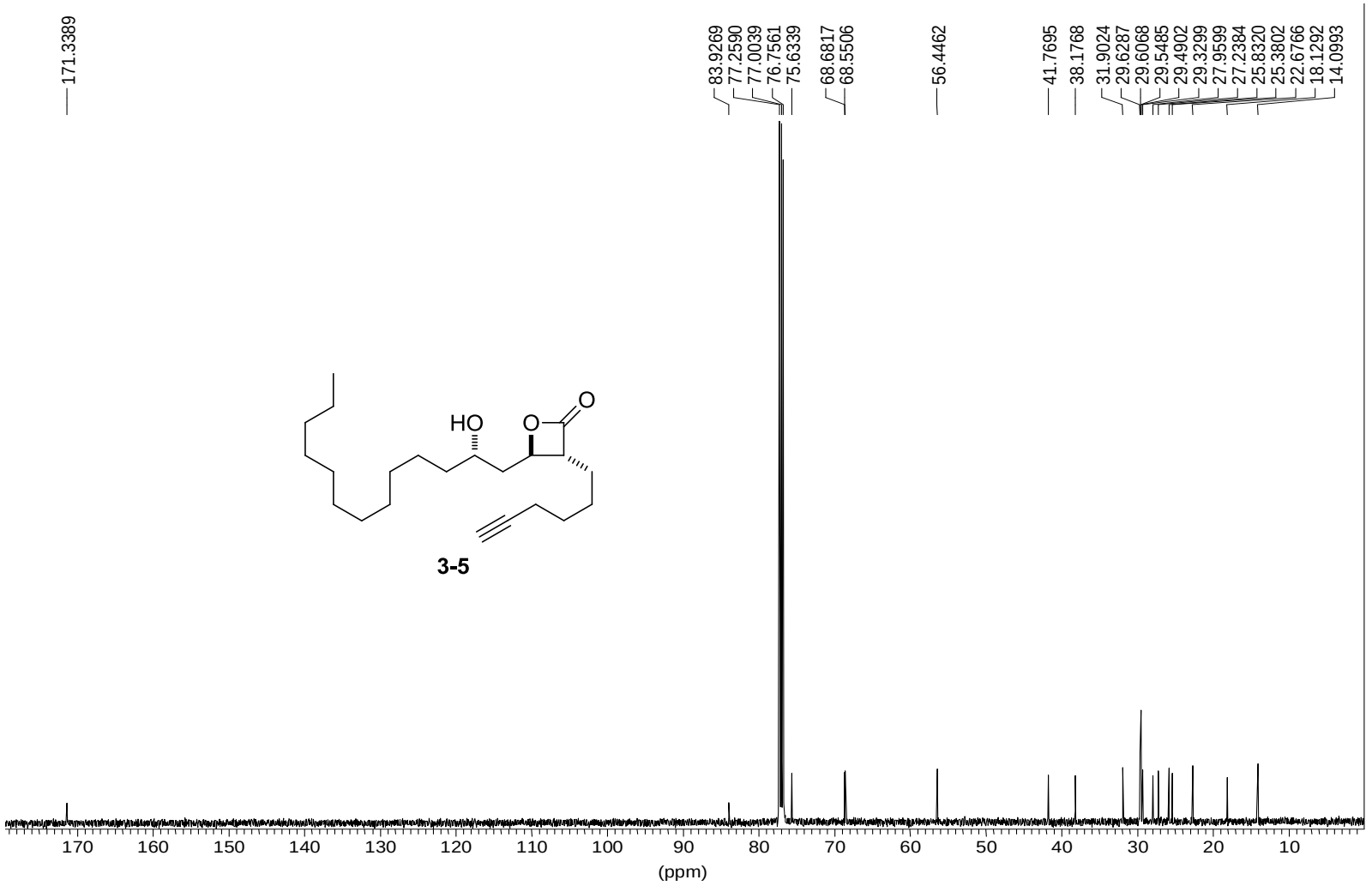

4.4288
4.4200
4.1855
4.1741
4.0834
4.0695
3.8413
3.1895
2.5163
2.5112
2.5049
2.4999
2.3322
2.3183
2.3032
2.2540
2.2402
2.2276
2.2213
2.2074
1.8531
1.8443
1.8355
1.5733
1.5506
1.5392
1.4283
1.2644
0.9303
0.9088
0.8962
0.8824
0.8735
0.2621
0.1461
0.1423
0.0705
TBSO
O
O
TMS
Integral
1.0314
0.5835
0.3442
1.1773
1.0000
1.2463
5.0528
5.1205
17.188
37.272
31.716
3.3714
15.115
11.626
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
(ppm)

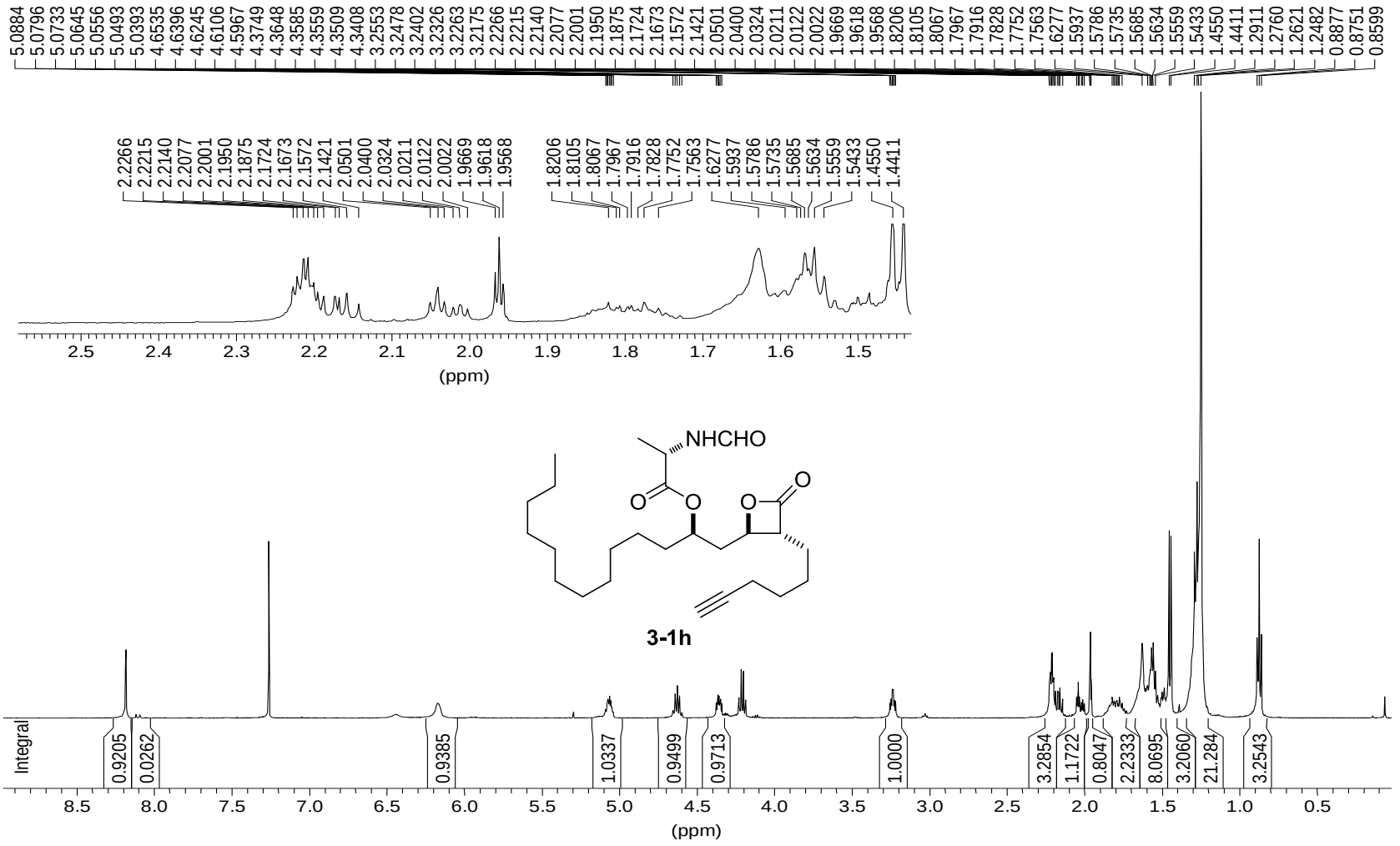

171.3968
106.6558
84.9176
77.2658
77.0108
76.7557
75.1962
68.6084
55.9867
41.9367
37.8994
31.9019
29.7303
29.6356
29.6137
29.5700
29.5408
29.3295
28.2145
27.3400
26.0429
25.8607
25.8316
24.6146
22.6761
19.6081
18.0486
14.1062
0.1363
-4.2944
-4.7463
TBSO
O
O
TMS
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)

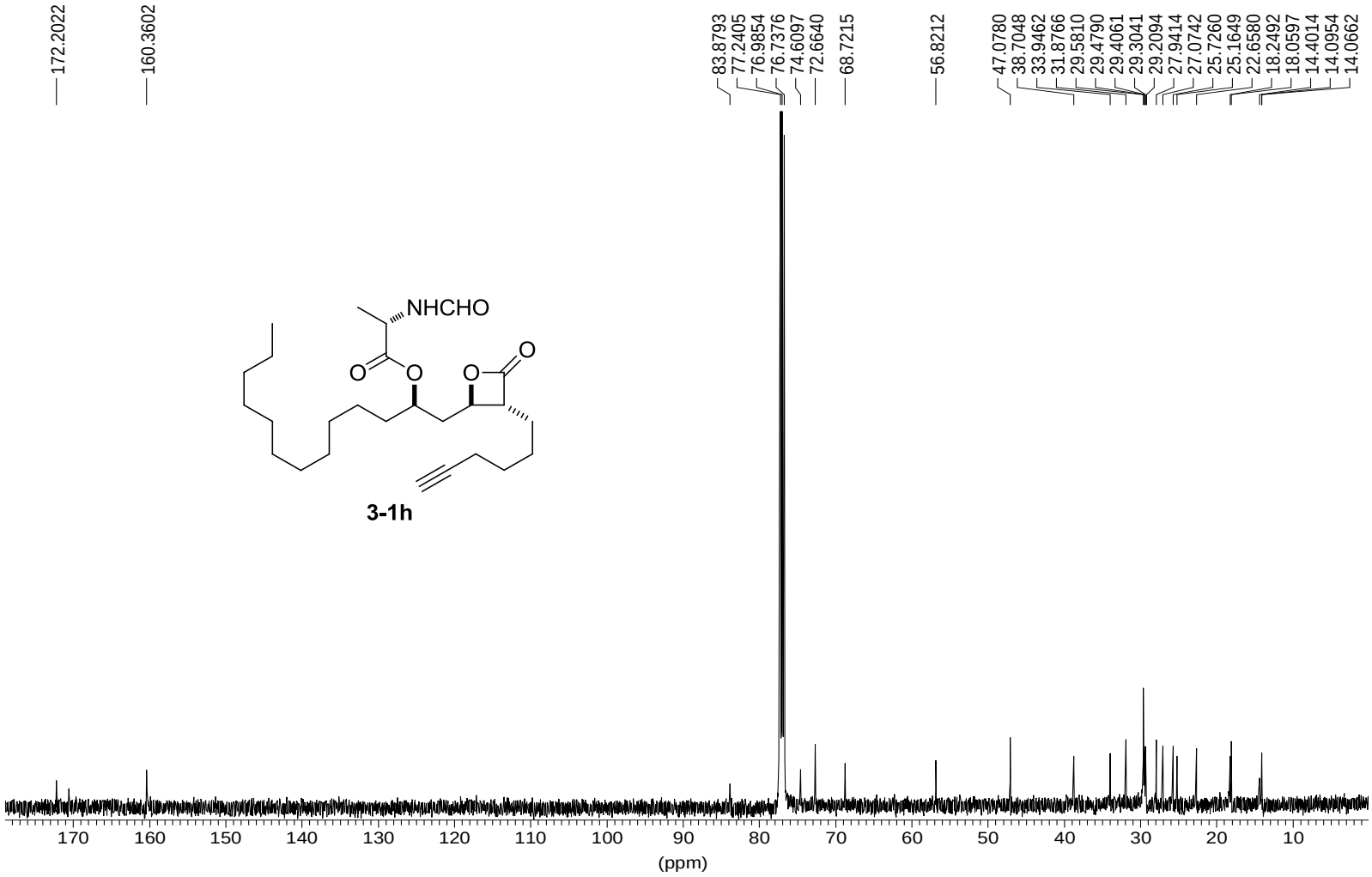

7.2606
4.5135
4.5047
4.4971
4.4883
4.4795
3.8251
3.8188
3.8125
3.8062
3.7999
3.7936
3.7873
3.7810
3.7747
3.3448
3.3372
3.3272
3.3196
3.3145
3.3057
3.2906
3.2830
3.2742
3.2666
3.2603
3.2528
2.2467
2.2429
2.2328
2.2202
1.9555
1.9492
1.9378
1.9328
1.9265
1.9214
1.9088
1.9038
1.8597
1.8508
1.8408
1.8307
1.8218
1.8118
1.8017
1.7979
1.7928
1.7815
1.7626
1.5533
1.5394
1.5268
1.5193
1.5054
1.4928
1.4789
1.4651
1.2520
0.8851
0.8725
0.8574
0.1388
3.40
3.30
3.20
(ppm)
2.2
2.0
1.8
1.6
1.4
(ppm)
HO
O
O
TMS
+
HO
O
O
TMS
3-4
anti/syn = 7:1
Integral
0.9725
1.0000
0.1194
0.8295
2.4943
1.0386
3.2325
8.5044
20.574
3.6951
10.497
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
(ppm)

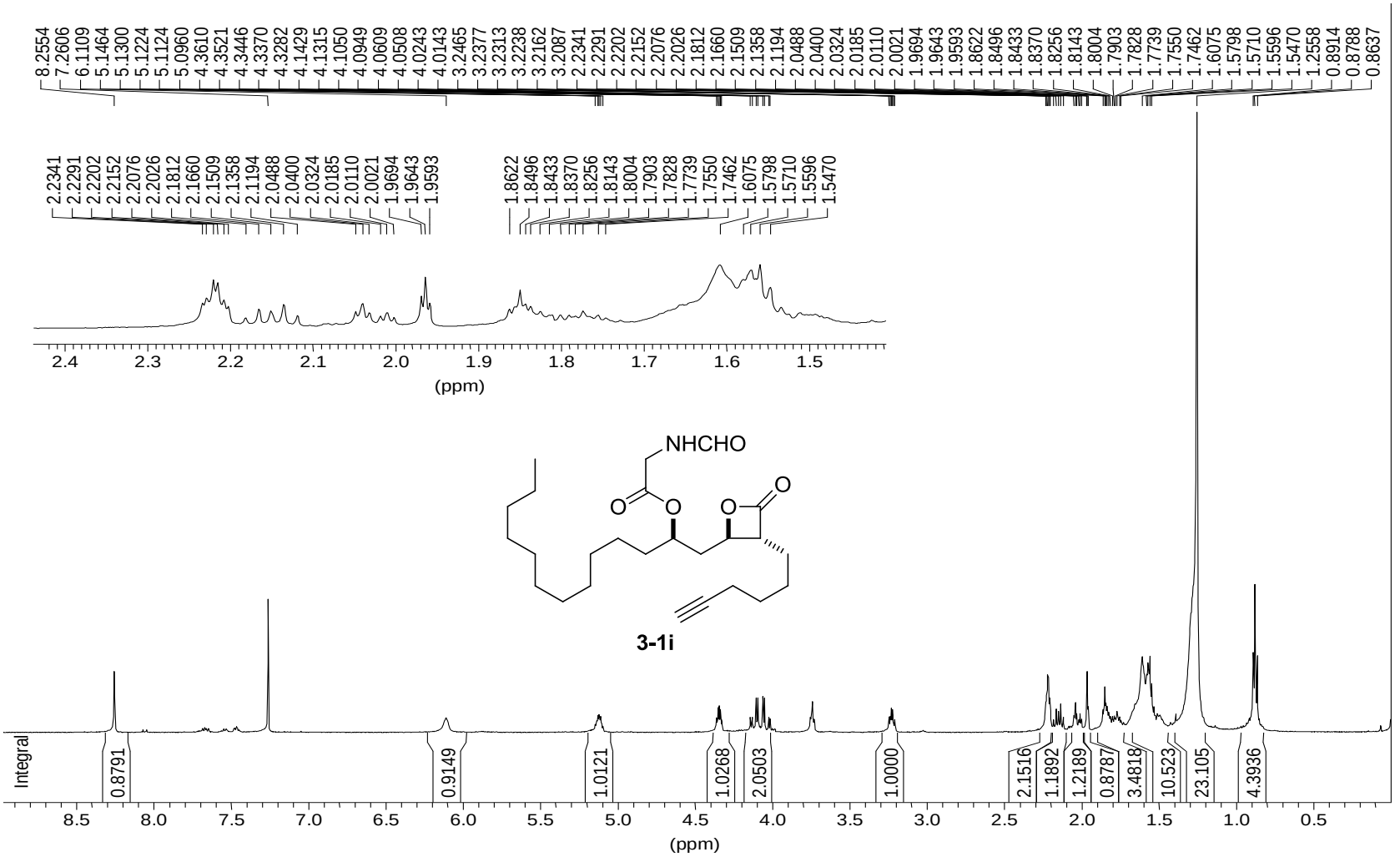

HO
O
O
TMS
3-4
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
(ppm)

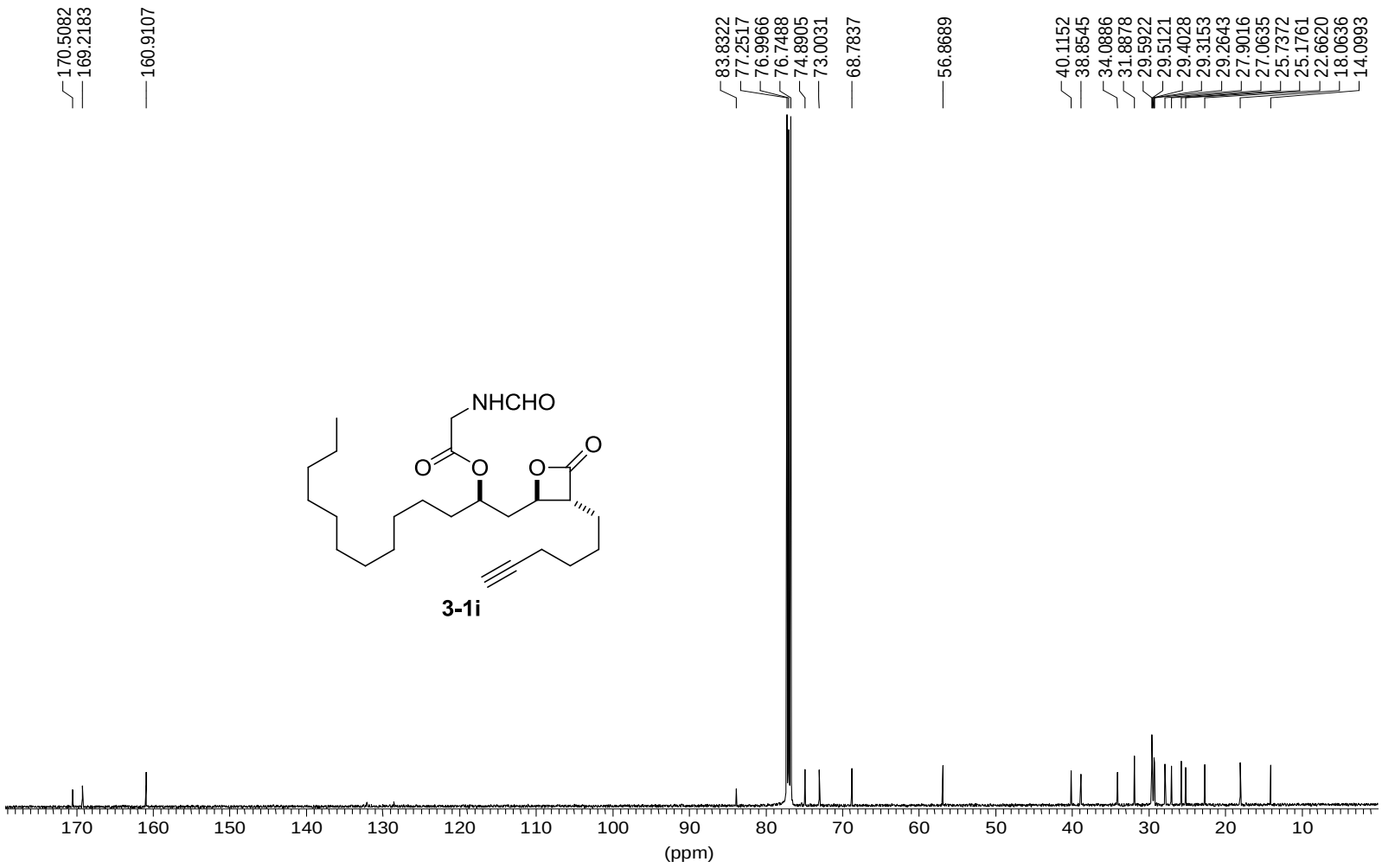

7.2606
4.5198
4.5110
4.5022
4.4946
4.4858
3.8365
3.8302
3.8239
3.8176
3.8113
3.8050
3.7999
3.7924
3.7873
3.3057
3.2982
3.2893
3.2818
3.2755
3.2679
2.2303
2.2253
2.2177
2.2127
2.2038
2.1988
1.9681
1.9630
1.9567
1.9517
1.9467
1.9404
1.9341
1.9227
1.9177
1.8698
1.8597
1.8496
1.8408
1.8307
1.8206
1.8118
1.7966
1.7790
1.6214
1.6025
1.5911
1.5873
1.5785
1.5659
1.5520
1.5394
1.5167
1.5029
1.4890
1.4751
1.2633
0.8952
0.8826
0.8675
2.4
2.3
2.2
2.1
2.0
1.9
1.8
1.7
1.6
1.5
(ppm)
HO
O
O
3-5
Integral
1.0000
1.0118
0.9961
2.1054
1.9485
3.3380
9.3404
18.379
3.2191
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
(ppm)

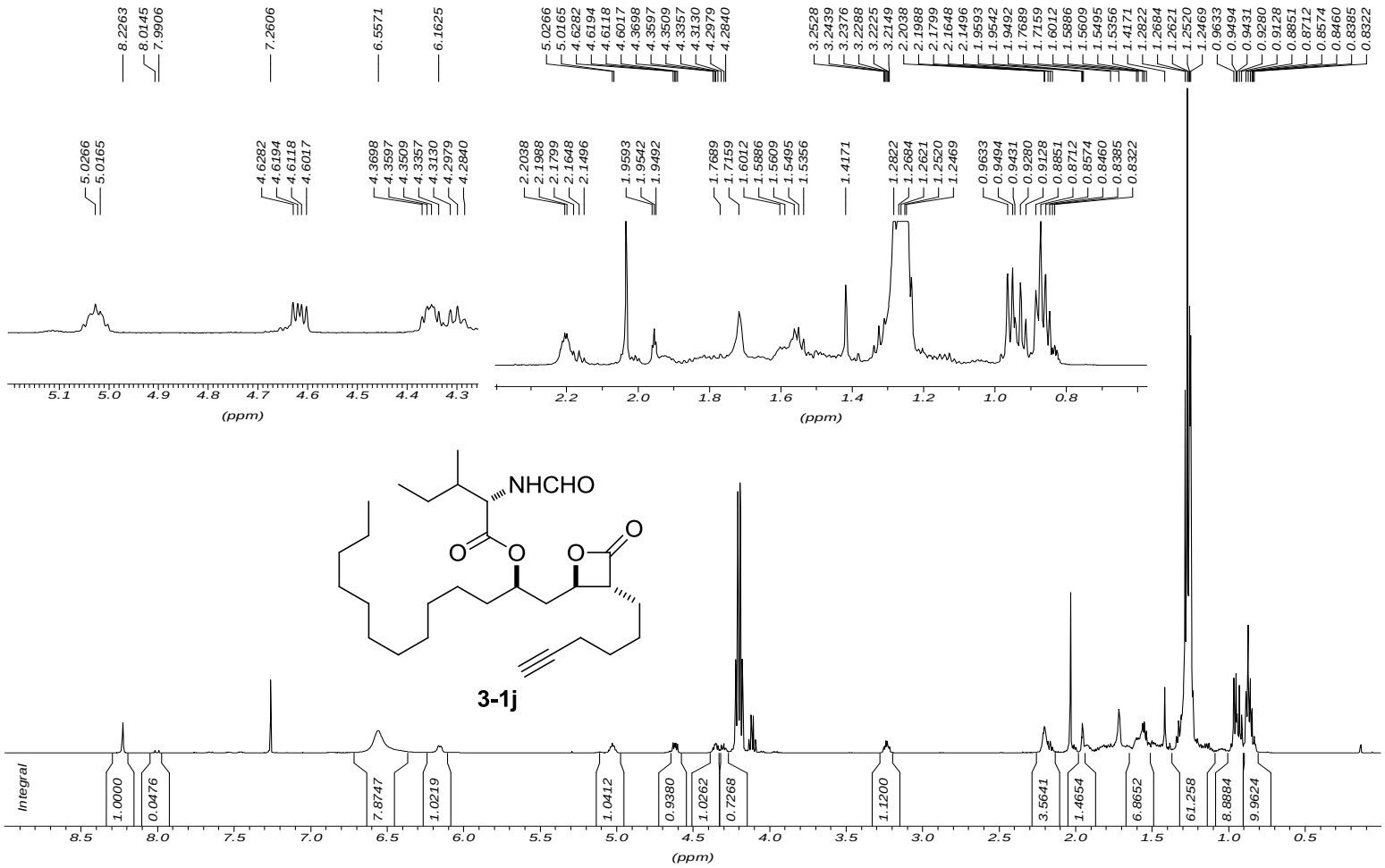

171.3389
83.9269
77.2590
77.0039
76.7561
75.6339
68.6817
68.5506
56.4462
41.7695
38.1768
31.9024
29.6287
29.6068
29.5485
29.4902
29.3299
27.9599
27.2384
25.8320
25.3802
22.6766
18.1292
14.0993
HO
O
O
3-5
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
(ppm)

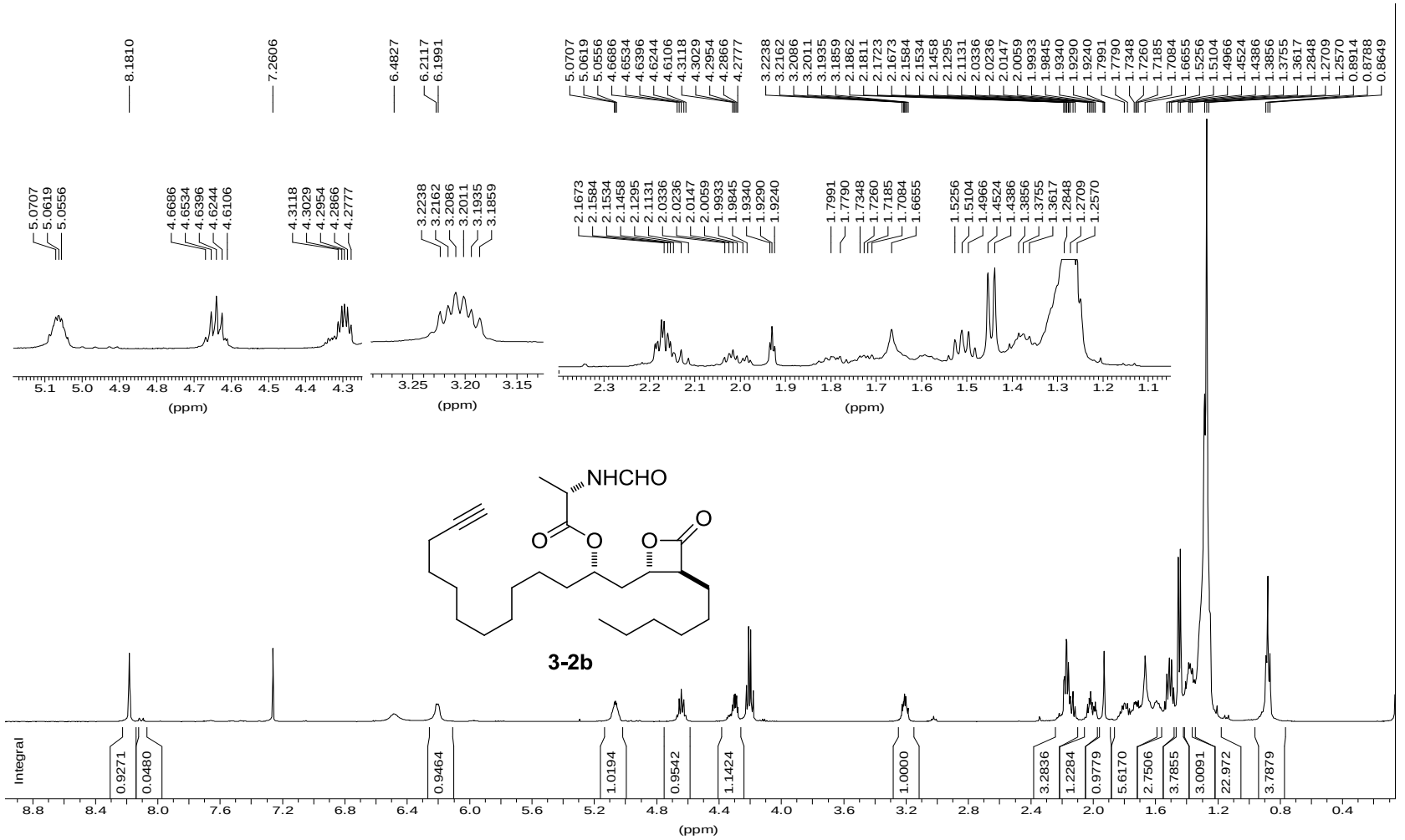

5.0884 5.0796 5.0733 5.0645 5.0556 5.0493 5.0393 4.6535 4.6396 4.6245 4.6106 4.5967 4.3749 4.3648 4.3585 4.3559 4.3509 4.3408 3.2553 3.2478 3.2402 3.2326 3.2263 3.2175 2.2266 2.2215 2.2140 2.2077 2.2001 2.1950 2.1875 2.1724 2.1673 2.1572 2.1421 2.0501 2.0400 2.0324 2.0211 2.0122 2.0022 1.9669 1.9618 1.9568 1.8206 1.8105 1.8067 1.7967 1.7916 1.7828 1.7752 1.7563 1.6277 1.5937 1.5786 1.5735 1.5685 1.5634 1.5559 1.5433 1.4550 1.4411 1.2911 1.2760 1.2621 1.2482 0.8877 0.8751 0.8599
2.5 2.4 2.3 2.2 2.1 2.0 1.9 1.8 1.7 1.6 1.5
(ppm)
NHCHO
O
O
O
O
3-1h
Integral
0.9205 0.0262 0.9385 1.0337 0.9499 0.9713 1.0000 3.2854 1.1722 0.8047 2.2333 8.0695 3.2060 21.284 3.2543
8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
(ppm)

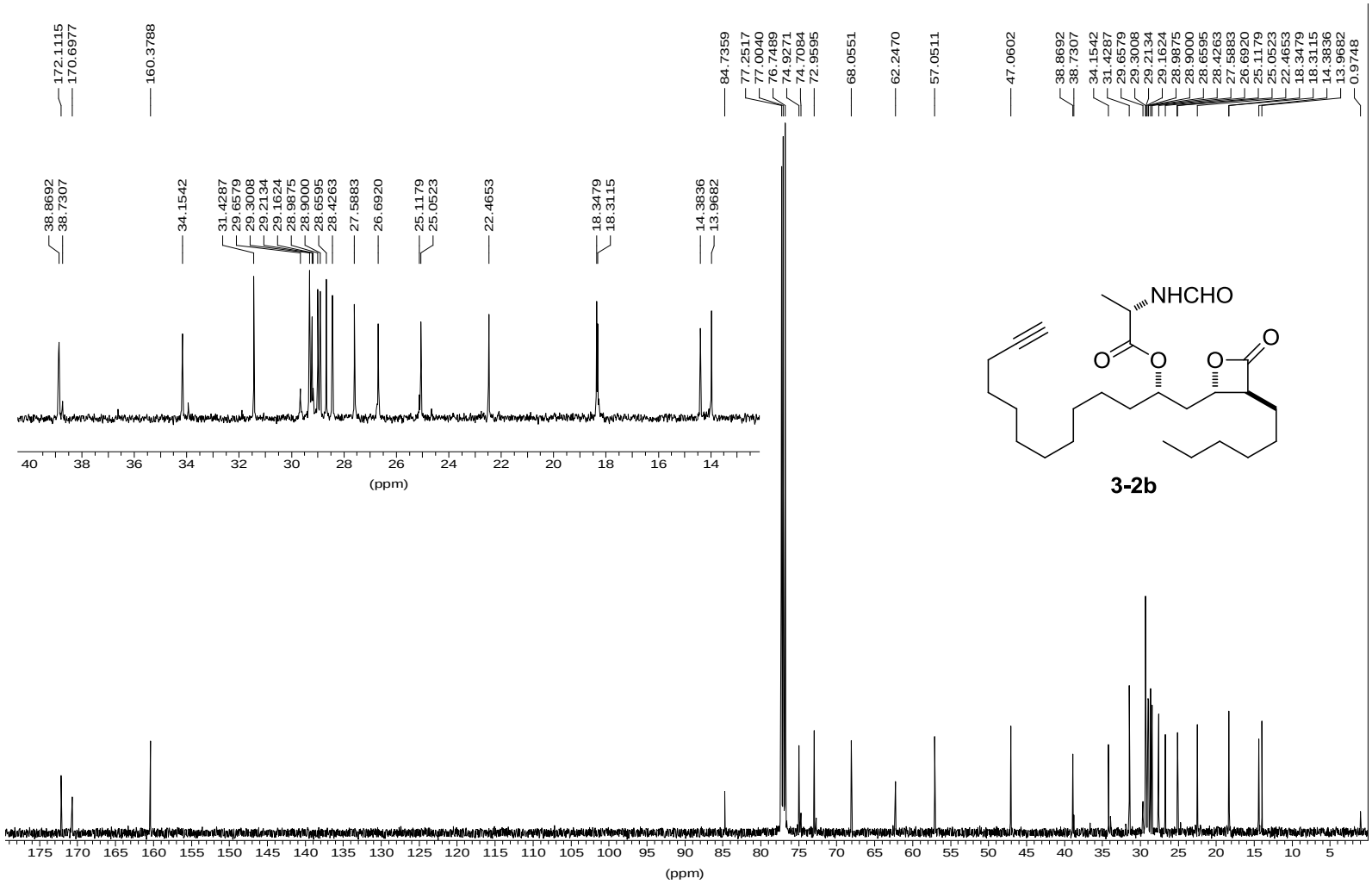

172.2022
160.3602
83.8793
77.2405
76.9854
76.7376
74.6097
72.6640
68.7215
56.8212
47.0780
38.7048
33.9462
31.8766
29.5810
29.4790
29.4061
29.3041
29.2094
27.9414
27.0742
25.7260
25.1649
22.6580
18.2492
18.0597
14.4014
14.0954
14.0662
NHCHO
O
O
O
O
3-1h
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
(ppm)

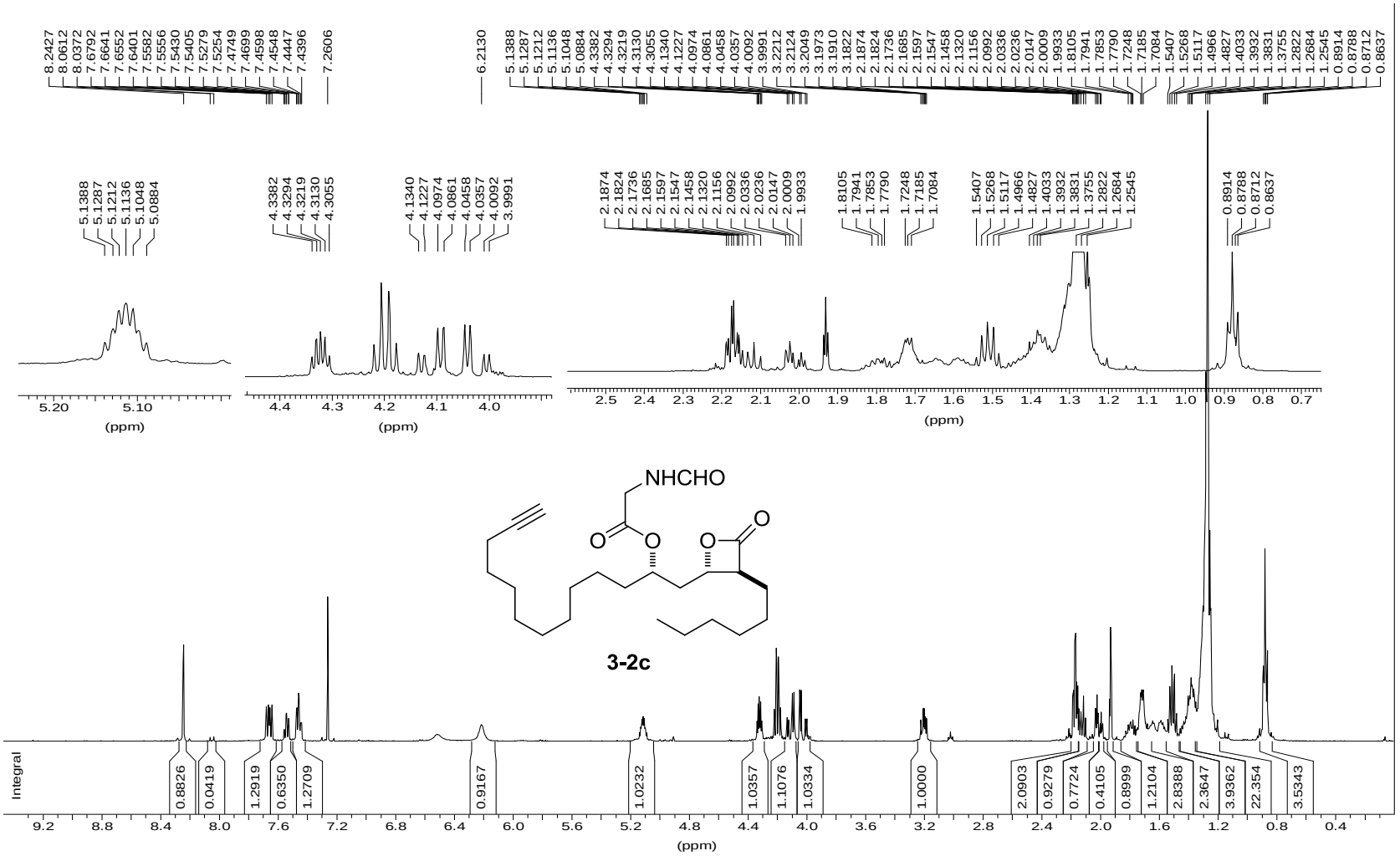

8.2554
7.2606
6.1109
5.1464
5.1300
5.1224
5.1124
5.0960
4.3610
4.3521
4.3446
4.3370
4.3282
4.1429
4.1315
4.1050
4.0949
4.0609
4.0508
4.0243
4.0143
3.2465
3.2377
3.2313
3.2238
3.2162
3.2087
2.2341
2.2291
2.2202
2.2152
2.2076
2.2026
2.1812
2.1660
2.1509
2.1358
2.1194
2.0488
2.0400
2.0324
2.0185
2.0110
2.0021
1.9694
1.9643
1.9593
1.8622
1.8496
1.8433
1.8370
1.8256
1.8143
1.8004
1.7903
1.7828
1.7739
1.7550
1.7462
1.6075
1.5798
1.5710
1.5596
1.5470
1.2558
0.8914
0.8788
0.8637
2.4
2.3
2.2
2.1
2.0
1.9
1.8
1.7
1.6
1.5
(ppm)
NHCHO
O
O
O
O
3-1i
Integral
0.8791
0.9149
1.0121
1.0268
2.0503
1.0000
2.1516
1.1892
1.2189
0.8787
3.4818
10.523
23.105
4.3936
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
(ppm)

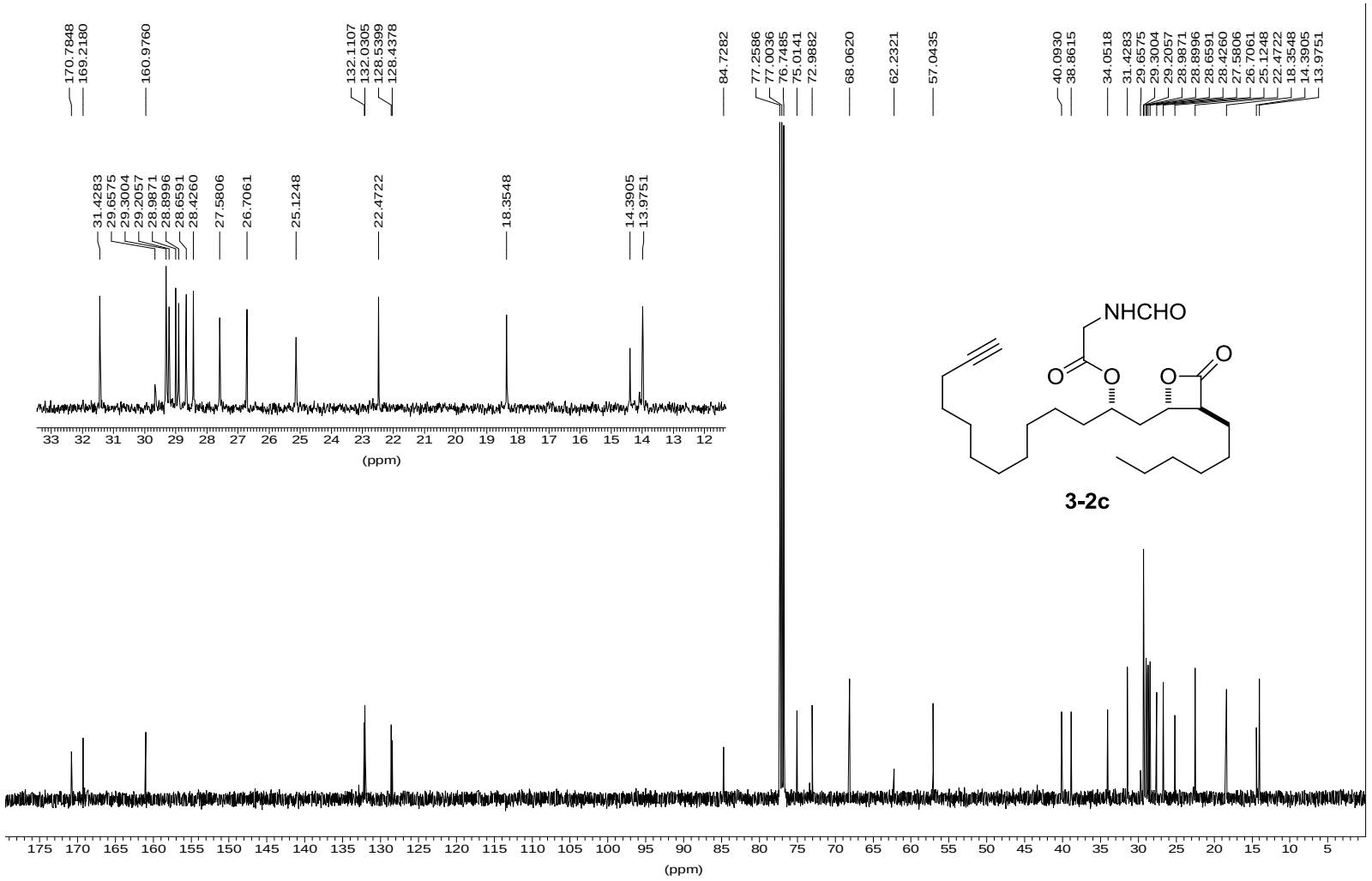

170.5082
169.2183
160.9107
83.8322
77.2517
76.9966
76.7488
74.8905
73.0031
68.7837
56.8689
40.1152
38.8545
34.0886
31.8878
29.5922
29.5121
29.4028
29.3153
29.2643
27.9016
27.0635
25.7372
25.1761
22.6620
18.0636
14.0093
NHCHO
O
O
O
O
3-1i
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
(ppm)

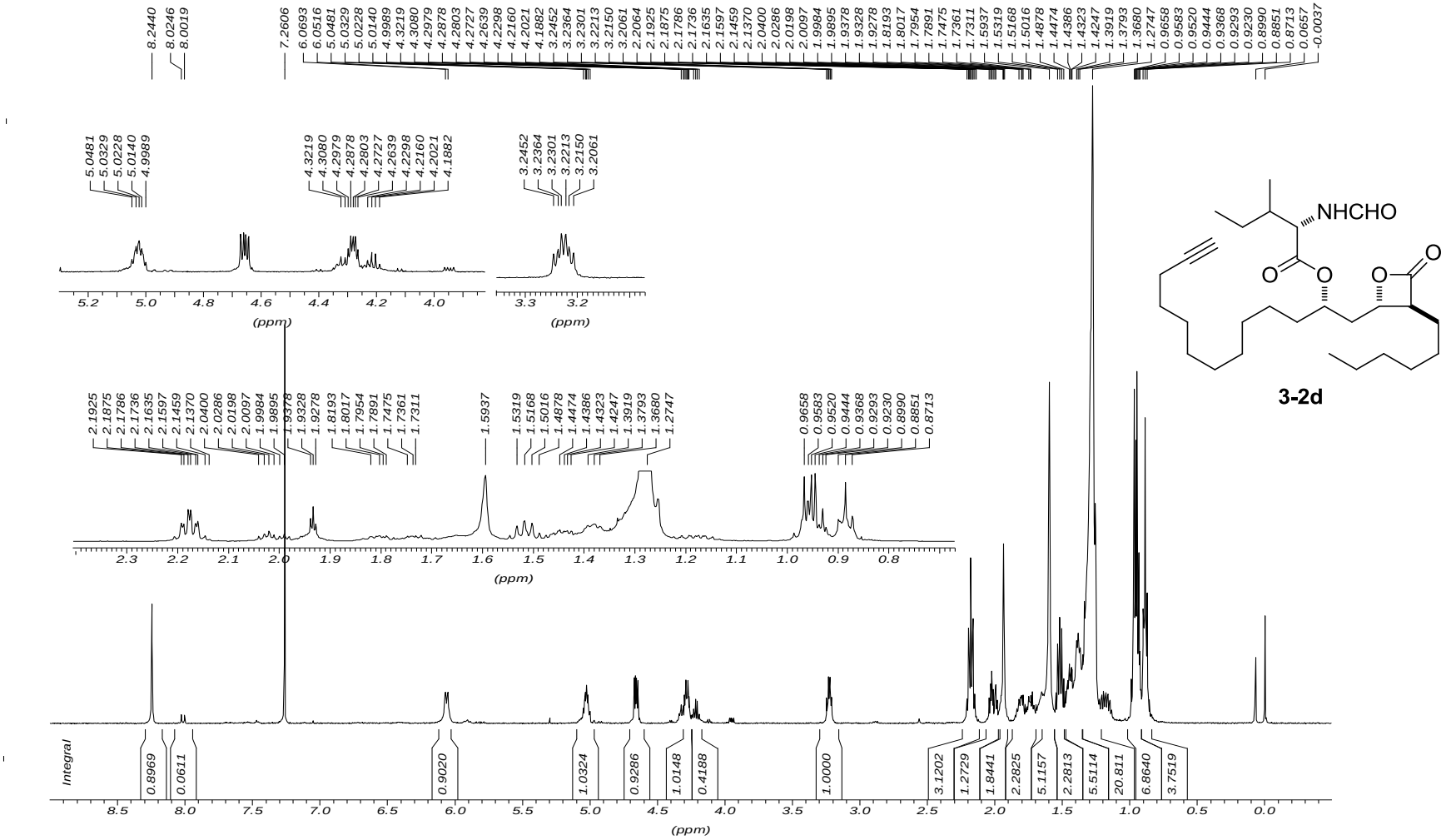

8.2263
8.0145
7.9906
7.2606
6.5571
6.1625
5.0266
5.0165
4.6282
4.6194
4.6118
4.6017
4.3698
4.3597
4.3509
4.3357
4.3130
4.2979
4.2840
3.2528
3.2439
3.2376
3.2288
3.2225
3.2149
2.2038
2.1988
2.1799
2.1648
2.1496
1.9593
1.9542
1.9492
1.7689
1.7159
1.6012
1.5886
1.5609
1.5495
1.5356
1.4171
1.2822
1.2684
1.2621
1.2520
1.2469
0.9633
0.9494
0.9431
0.9280
0.9128
0.8851
0.8712
0.8574
0.8460
0.8385
0.8322
(ppm)
NHCHO
O
O
O
O
3-1j
Integral
1.0000
0.0476
7.8747
1.0219
1.0412
0.9380
1.0262
0.7268
1.1200
3.5641
1.4654
6.8652
61.258
8.8884
9.9624
(ppm)

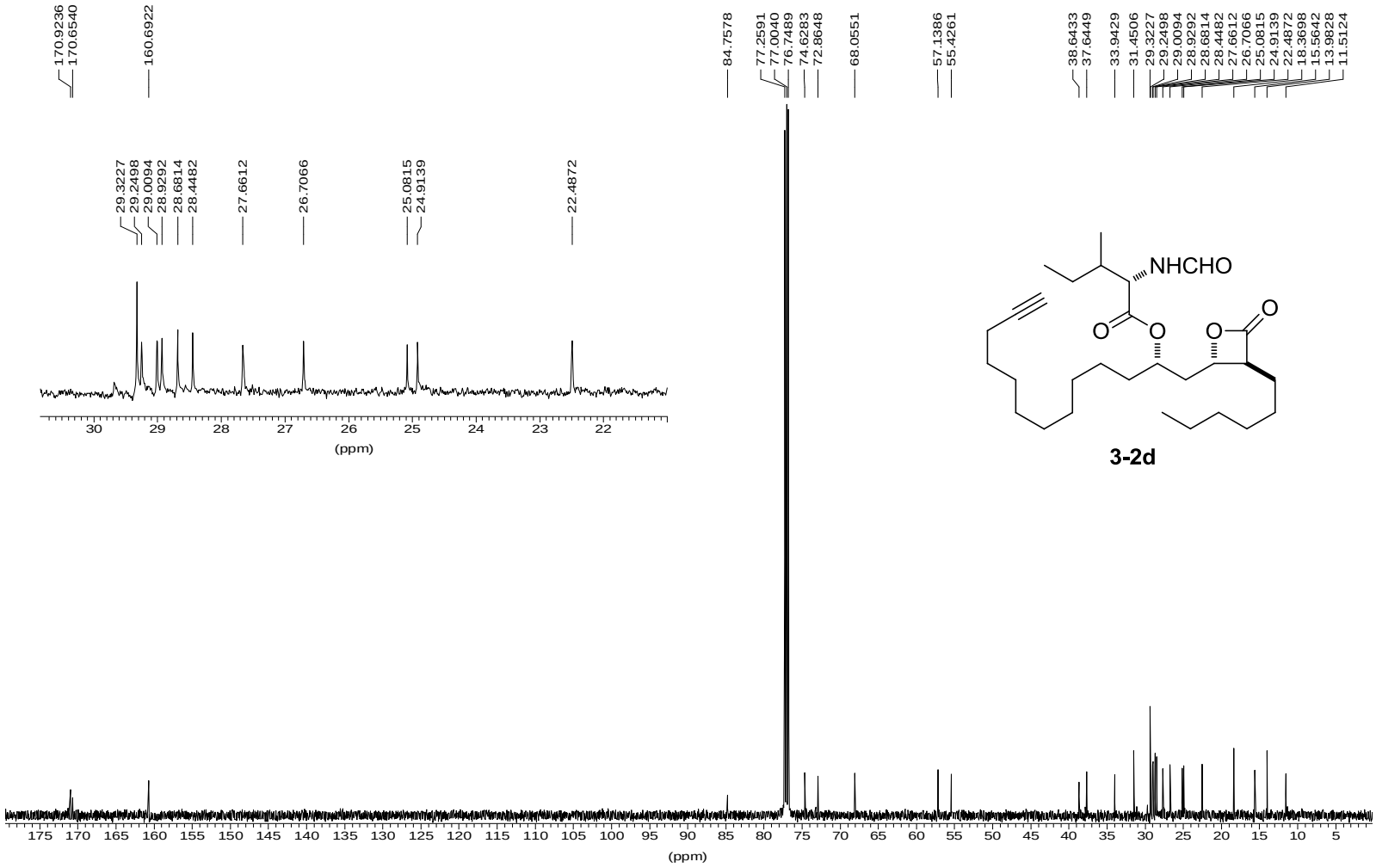

NHCHO
O
O
O
O
3-2b
Integral
0.9271
0.0480
0.9464
1.0194
0.9542
1.1424
1.0000
3.2836
1.2284
0.9779
5.6170
2.7506
3.7855
3.0091
22.972
3.7879
(ppm)


172.1115
170.6977
160.3788
84.7359
77.2517
77.0040
76.7489
74.9271
74.7084
72.9595
68.0551
62.2470
57.0511
47.0602
38.8692
38.7307
34.1542
31.4287
29.6579
29.3008
29.2134
29.1624
28.9875
28.9000
28.6595
28.4263
27.5883
26.6920
25.1179
25.0523
22.4653
18.3479
18.3115
14.3836
13.9682
0.9748
(ppm)
NHCHO
O
O
O
O
3-2b

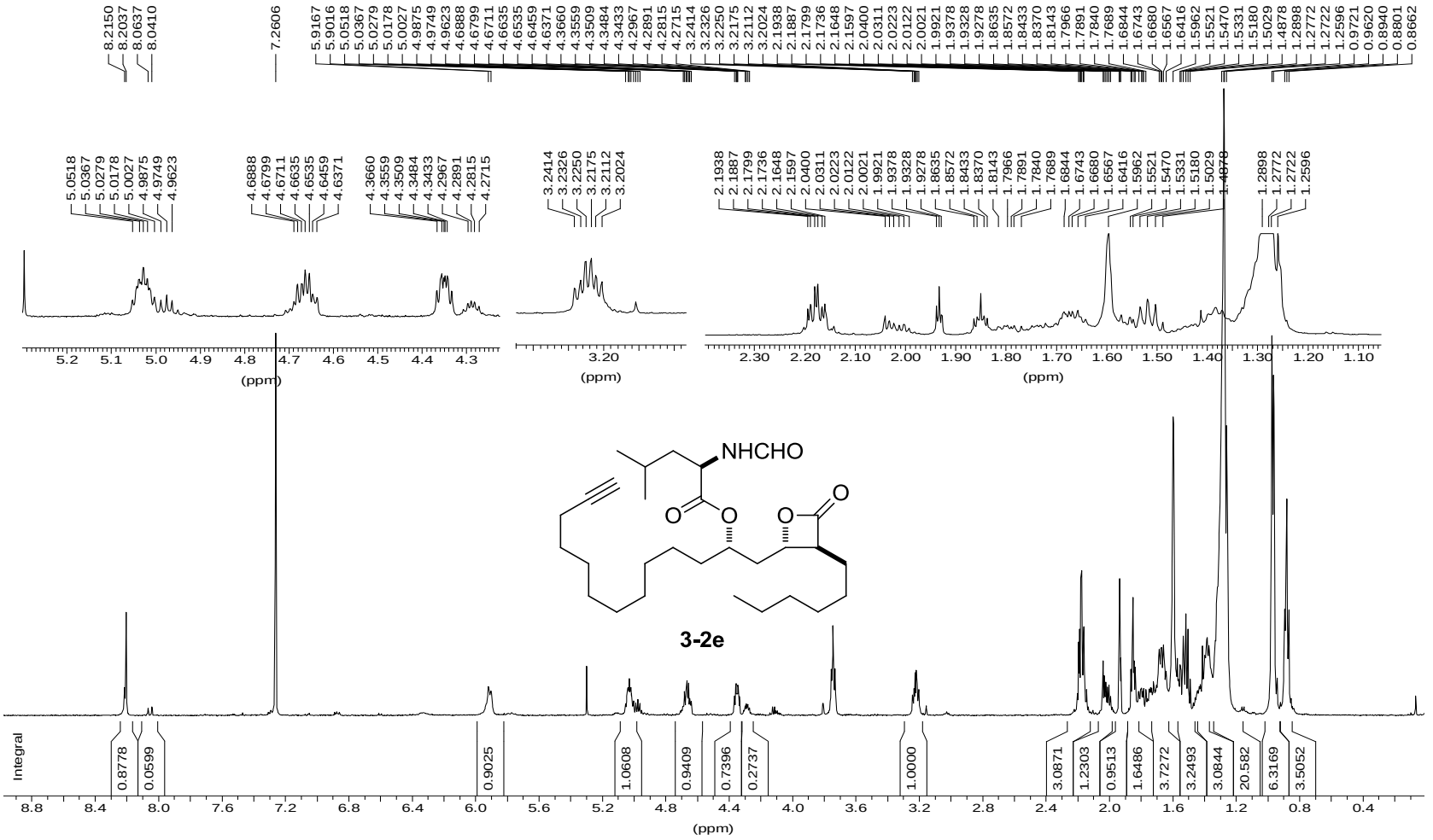

NHCHO
O
O
O
O
3-2c
(ppm)
Integral

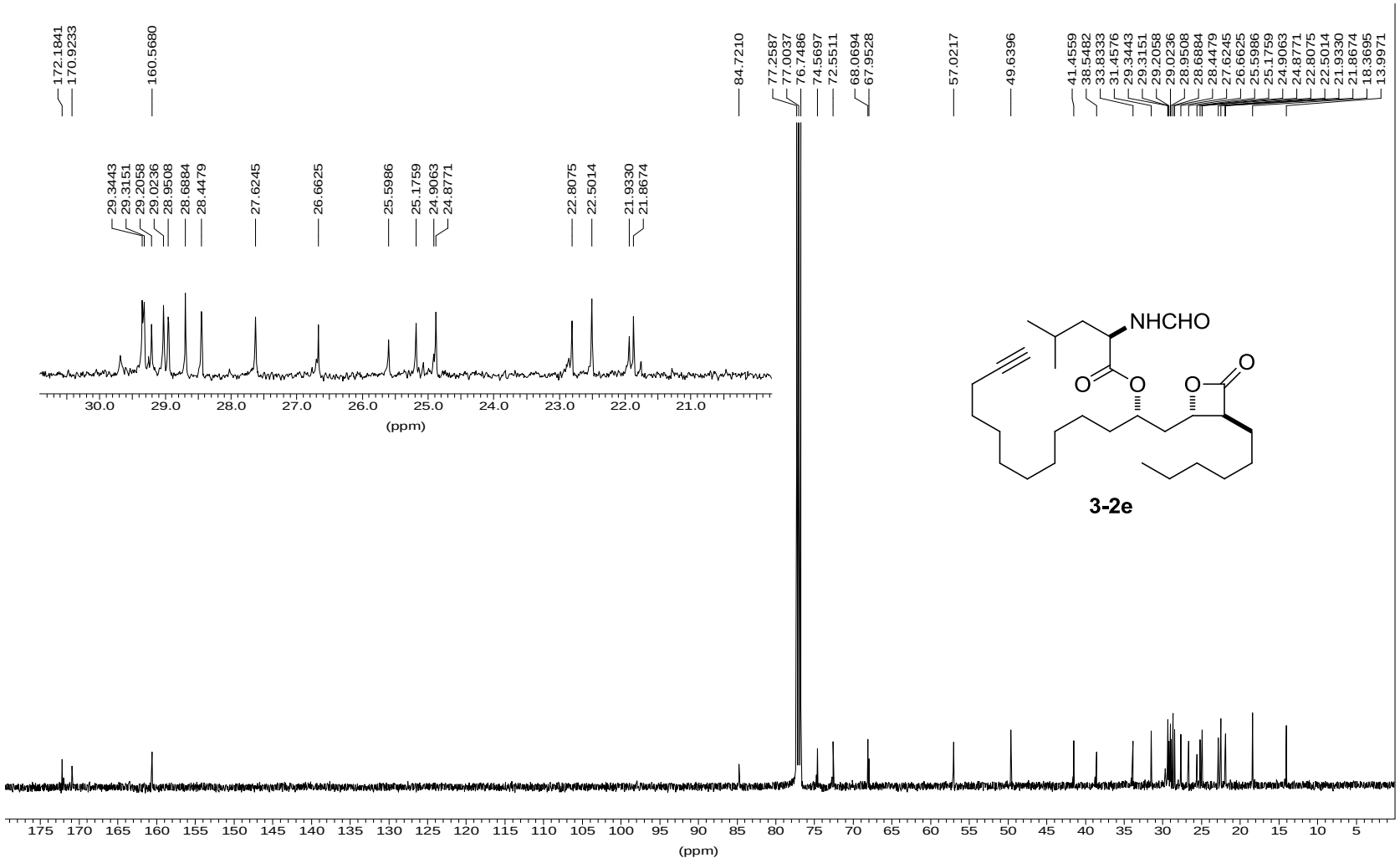

170.7848
169.2180
160.9760
132.1107
132.0305
128.5399
128.4378
84.7282
77.2586
77.0036
76.7485
75.0141
72.9882
68.0620
62.2321
57.0435
40.0930
38.8615
34.0518
31.4283
29.6575
29.3004
29.2057
28.9871
28.8996
28.6591
28.4260
27.5806
26.7061
25.1248
22.4722
18.3548
14.3905
13.9751
(ppm)
NHCHO
O
O
O
O
3-2c

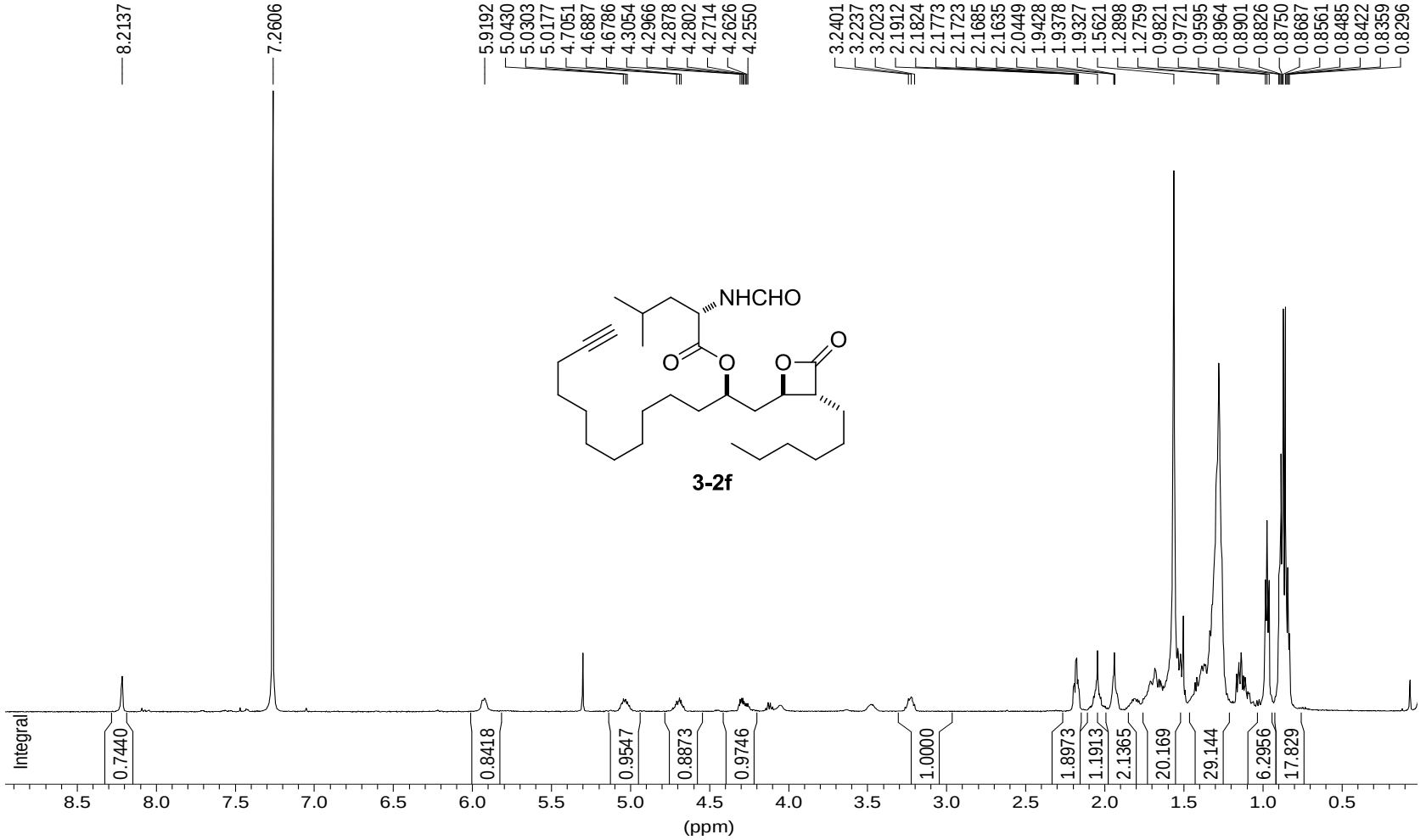

NHCHO
O
O
O
O
3-2d
Integral
(ppm)

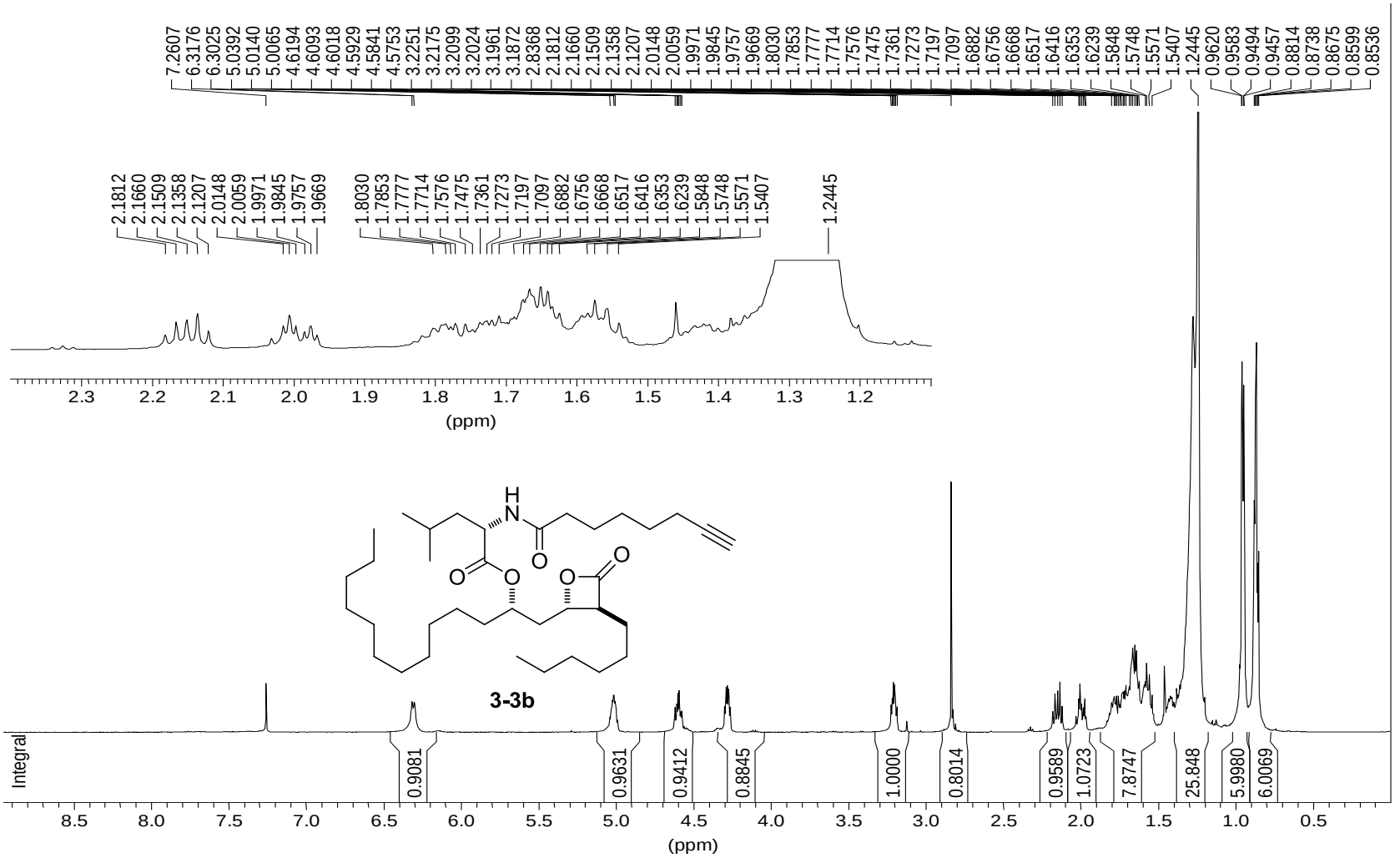

170.9236
170.6540
160.6922
84.7578
77.2591
77.0040
76.7489
74.6283
72.8648
68.0551
57.1386
55.4261
38.6433
37.6449
33.9429
31.4506
29.3227
29.2498
29.0094
28.9292
28.6814
28.4482
27.6612
26.7066
25.0815
24.9139
22.4872
18.3698
15.5642
13.9828
11.5124
NHCHO
O
O
O
O
(ppm)
(ppm)

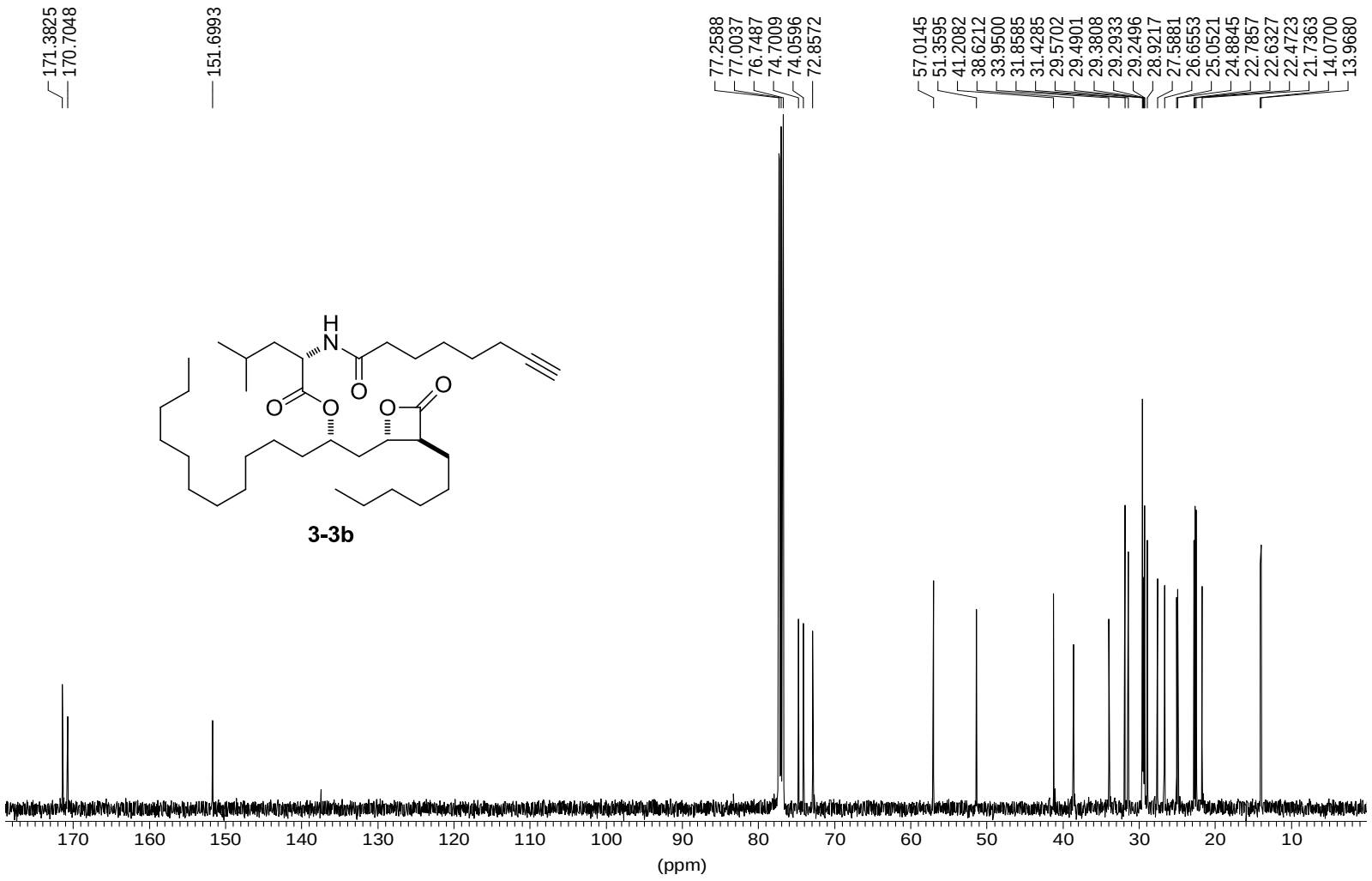
3-2d

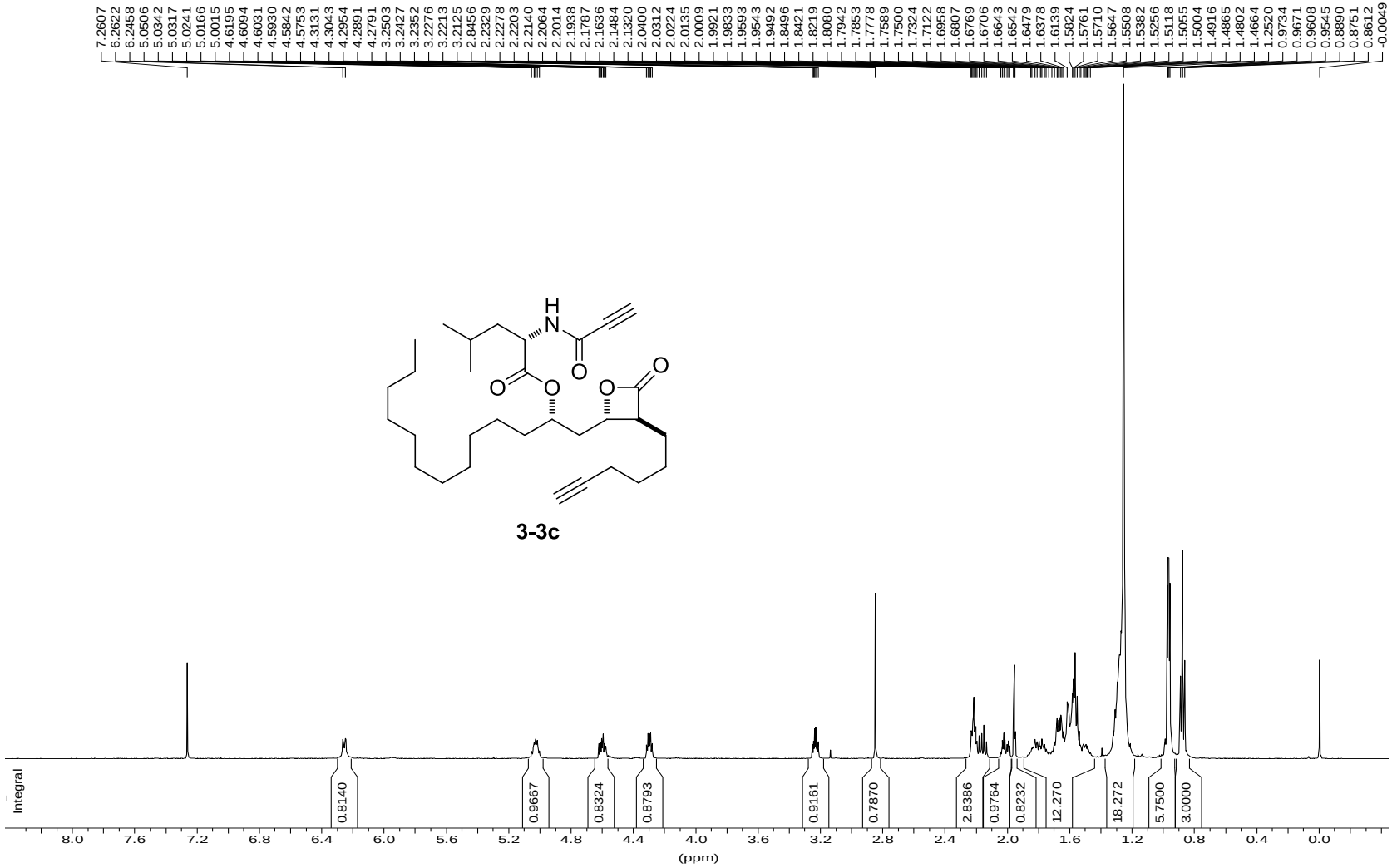

NHCHO
O
O
O
O
3-2e
Integral
0.8778
0.0599
0.9025
1.0608
0.9409
0.7396
0.2737
1.0000
3.0871
1.2303
0.9513
1.6486
3.7272
3.2493
3.0844
20.582
6.3169
3.5052
(ppm)

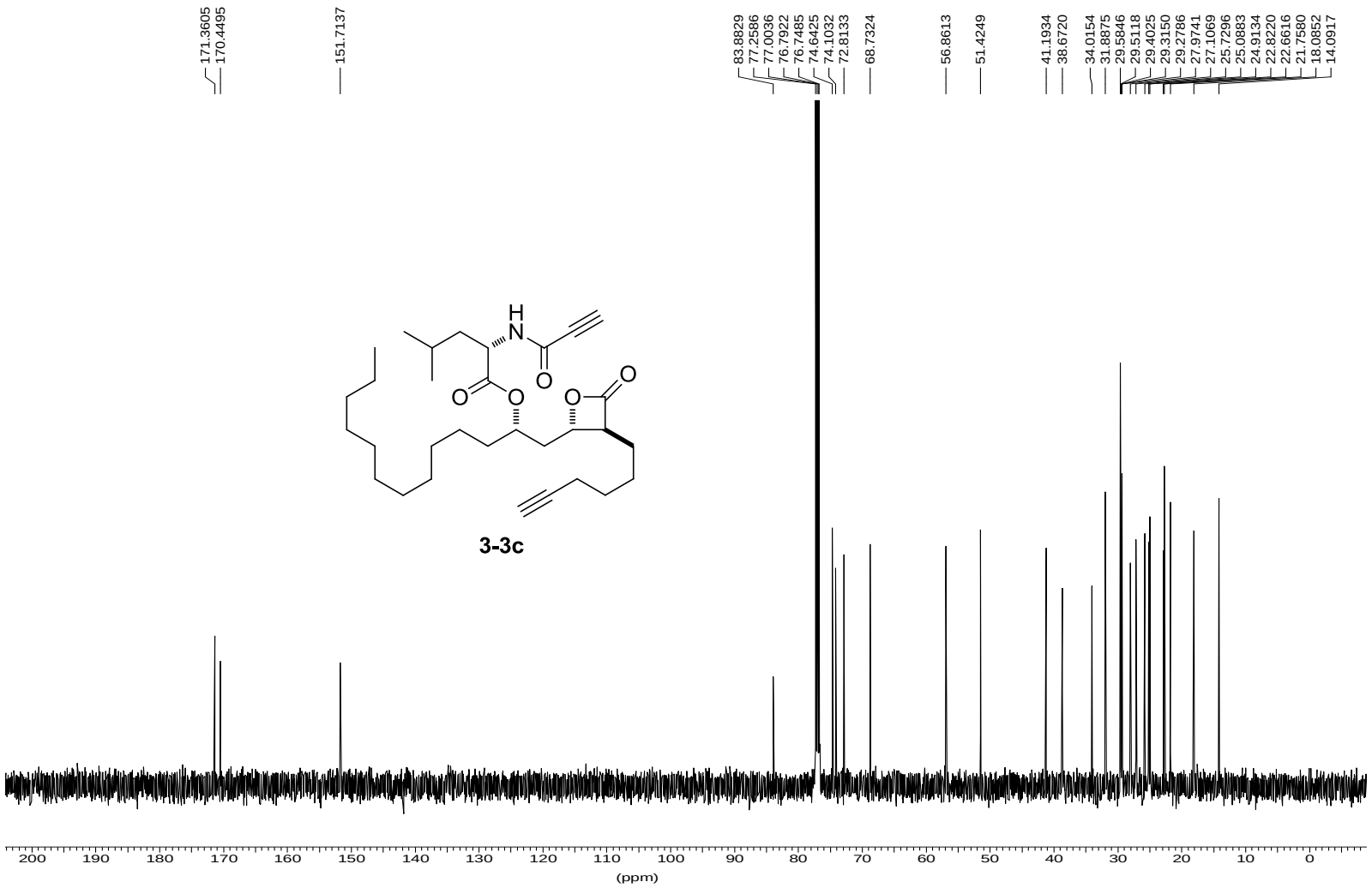

172.1841
170.9233
160.5680
84.7210
77.2587
77.0037
76.7486
74.5697
72.5511
68.0694
67.9528
57.0217
49.6396
41.4559
38.5482
33.8333
31.4576
29.3443
29.3151
29.2058
29.0236
28.9508
28.6884
28.4479
27.6245
26.6625
25.5986
25.1759
24.9063
24.8771
22.8075
22.5014
21.9330
21.8674
18.3695
13.9971
(ppm)
NHCHO
O
O
O
O
3-2e

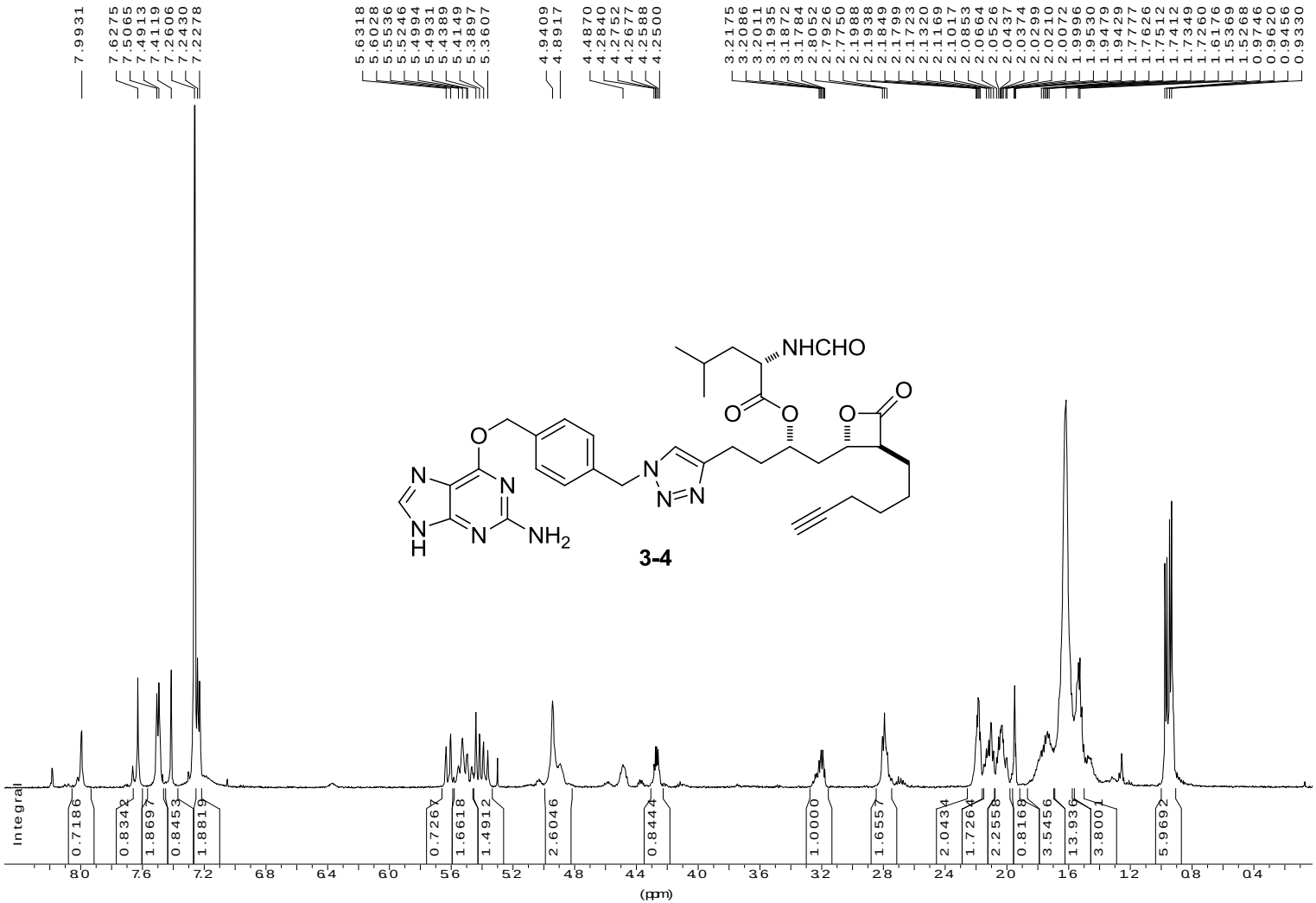

8.2137
7.2606
5.9192
5.0430
5.0303
5.0177
4.7051
4.6887
4.6786
4.3054
4.2966
4.2878
4.2802
4.2714
4.2626
4.2550
3.2401
3.2237
3.2023
2.1912
2.1824
2.1773
2.1723
2.1685
2.1635
2.0449
1.9428
1.9378
1.9327
1.5621
1.2898
1.2759
0.9821
0.9721
0.9595
0.8964
0.8901
0.8826
0.8750
0.8687
0.8561
0.8485
0.8422
0.8359
0.8296
NHCHO
O
O
O
O
3-2f
Integral
0.7440
0.8418
0.9547
0.8873
0.9746
1.0000
1.8973
1.1913
2.1365
20.169
29.144
6.2956
17.829
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
(ppm)


7.2607
6.3176
6.3025
5.0392
5.0140
5.0065
4.6194
4.6093
4.6018
4.5929
4.5841
4.5753
3.2251
3.2175
3.2099
3.2024
3.1961
3.1872
2.8368
2.1812
2.1660
2.1509
2.1358
2.1207
2.0148
2.0059
2.0059
1.9971
1.9845
1.9757
1.9669
1.8030
1.7853
1.7777
1.7714
1.7576
1.7475
1.7361
1.7273
1.7197
1.7097
1.6882
1.6756
1.6668
1.6517
1.6416
1.6353
1.6239
1.5848
1.5748
1.5571
1.5407
1.2445
(ppm)
3-3b
Integral
0.9081
0.9631
0.9412
0.8845
1.0000
0.8014
0.9589
1.0723
7.8747
25.848
5.9980
6.0069
(ppm)

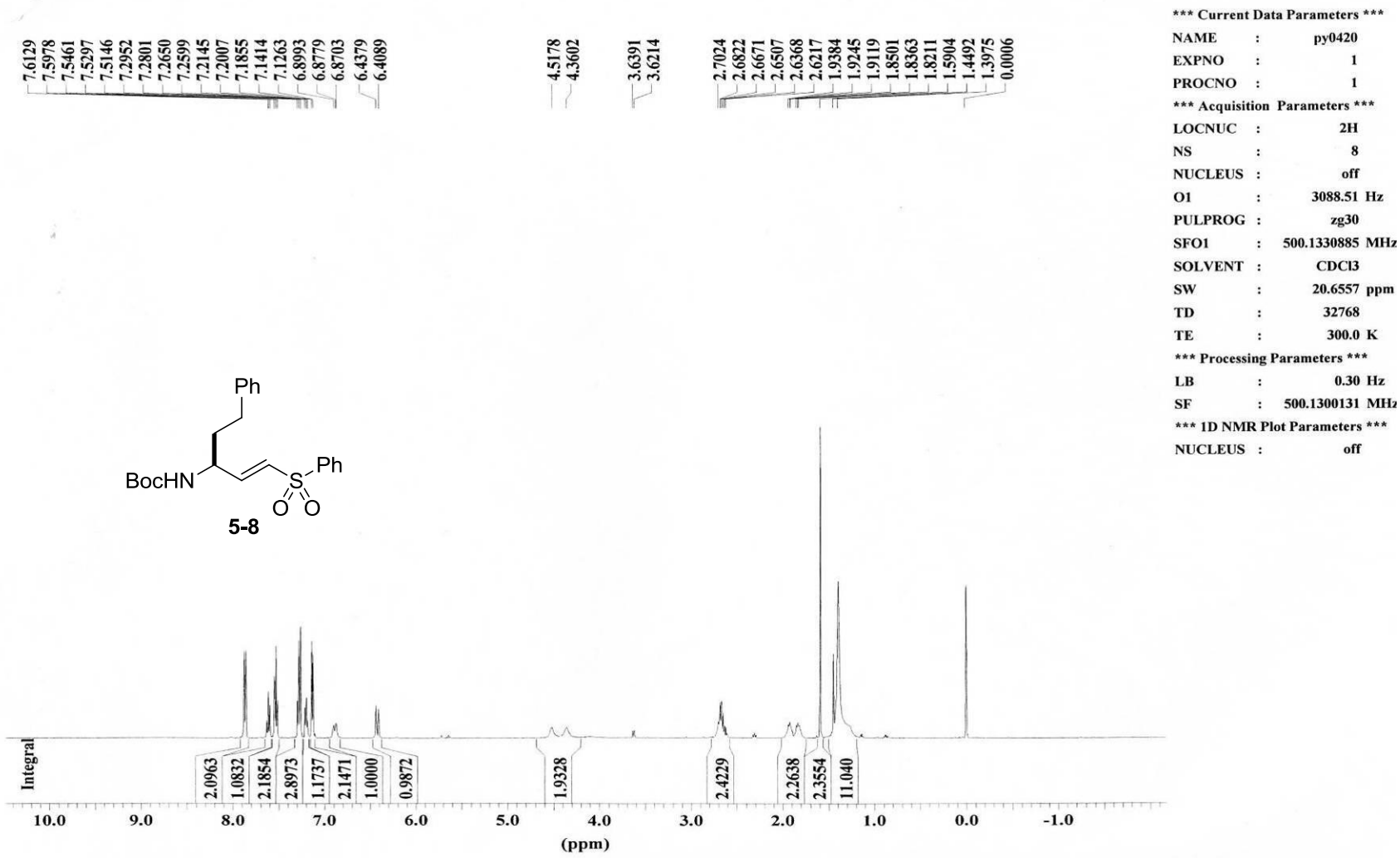

171.3825
170.7048
151.6993
77.2588
77.0037
76.7487
74.7009
74.0596
72.8572
57.0145
51.3595
41.2082
38.6212
33.9500
31.8585
31.4285
29.5702
29.4901
29.3808
29.2933
29.2496
28.9217
27.5881
26.6553
25.0521
24.8845
22.7857
22.6327
22.4723
21.7363
14.0700
13.9680
H
N
O
O
O
O
O
3-3b
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
(ppm)


7.2607
6.2622
6.2458
5.0506
5.0342
5.0317
5.0241
5.0166
5.0015
4.6195
4.6094
4.6031
4.5930
4.5842
4.5753
4.3131
4.3043
4.2954
4.2891
4.2791
3.2503
3.2427
3.2352
3.2276
3.2213
3.2125
2.8456
2.2329
2.2278
2.2203
2.2140
2.2064
2.2014
2.1938
2.1787
2.1636
2.1484
2.1320
2.0400
2.0312
2.0224
2.0135
2.0009
1.9921
1.9833
1.9593
1.9543
1.9492
1.8496
1.8421
1.8219
1.8080
1.7942
1.7853
1.7778
1.7589
1.7500
1.7324
1.7122
1.6958
1.6807
1.6769
1.6706
1.6643
1.6542
1.6479
1.6378
1.6139
1.5824
1.5761
1.5710
1.5647
1.5508
1.5382
1.5256
1.5118
1.5055
1.5004
1.4916
1.4885
1.4802
1.4664
1.2520
0.9734
0.9671
0.9608
0.9545
0.8890
0.8751
0.8612
-0.0049
H
N
O
O
O
O
O
3-3c
Integral
0.8140
0.9667
0.8324
0.8793
0.9161
0.7870
2.8386
0.9764
0.8232
12.270
18.272
5.7500
3.0000
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
0.0
(ppm)


171.3605
170.4495
151.7137
83.8829
77.2586
77.0036
76.7922
76.7485
74.6425
74.1032
72.8133
68.7324
56.8613
51.4249
41.1934
38.6720
34.0154
31.8875
29.5846
29.5118
29.4025
29.3150
29.2786
27.9741
27.1069
25.7296
25.0883
24.9134
22.8220
22.6616
21.7580
18.0852
14.0917
H
N
O
O
O
O
O
3-3c
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
(ppm)


7.9931
7.6275
7.5065
7.4913
7.4119
7.2606
7.2430
7.2278
5.6318
5.6028
5.5536
5.5246
5.4994
5.4931
5.4389
5.4149
5.3897
5.3607
4.9409
4.8917
4.4870
4.2840
4.2752
4.2677
4.2588
4.2500
3.2175
3.2086
3.2011
3.1935
3.1872
3.1784
2.8052
2.7926
2.7750
2.1988
2.1938
2.1849
2.1799
2.1723
2.1320
2.1169
2.1017
2.0853
2.0664
2.0526
2.0437
2.0374
2.0299
2.0210
2.0072
1.9996
1.9530
1.9479
1.9429
1.7777
1.7626
1.7512
1.7412
1.7349
1.7260
1.6176
1.5369
1.5268
0.9746
0.9620
0.9456
0.9330
NHCHO
O
O
O
O
O
N
N
N
N
N
H
NH2
3-4
Integral
0.7186
0.8342
1.8697
0.8453
1.8819
0.7267
1.6618
1.4912
2.6046
0.8444
1.0000
1.6557
2.0434
1.7264
2.2558
0.8169
3.5456
13.936
3.8001
5.9692
8.0
7.6
7.2
6.8
6.4
6.0
5.6
5.2
4.8
4.4
4.0
3.6
3.2
2.8
2.4
2.0
1.6
1.2
0.8
0.4
(ppm)

BocHphVSPh in CDCl3 1H AMX500



*** Current Data Parameters ***

NAME	:	py0420
EXPNO	:	1
PROCNO	:	1

*** Acquisition Parameters ***

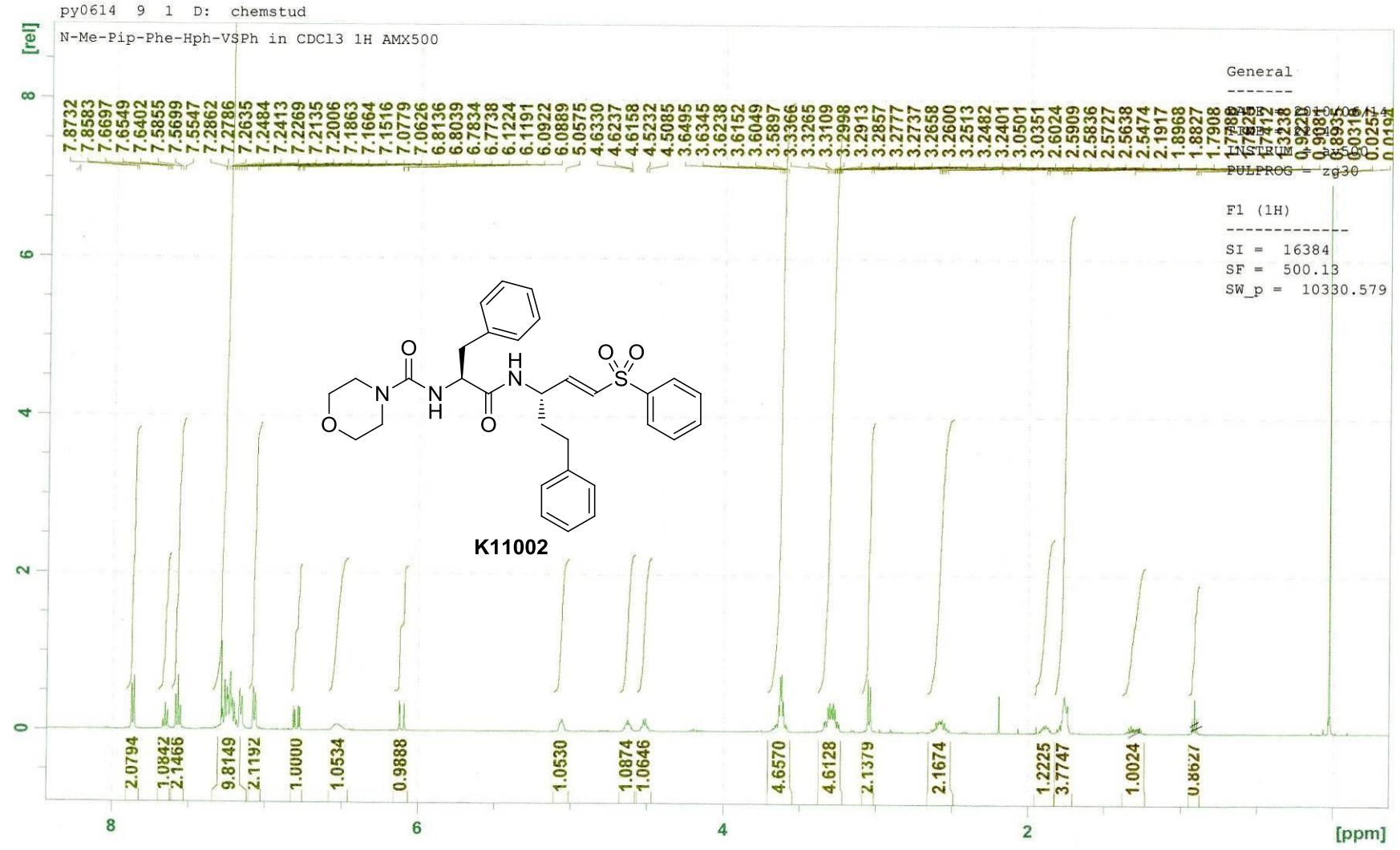
LOCNUC	:	2H
NS	:	8
NUCLEUS	:	off
O1	:	3088.51 Hz
PULPROG	:	zg30
SFO1	:	500.1330885 MHz
SOLVENT	:	CDCl3
SW	:	20.6557 ppm
TD	:	32768
TE	:	300.0 K

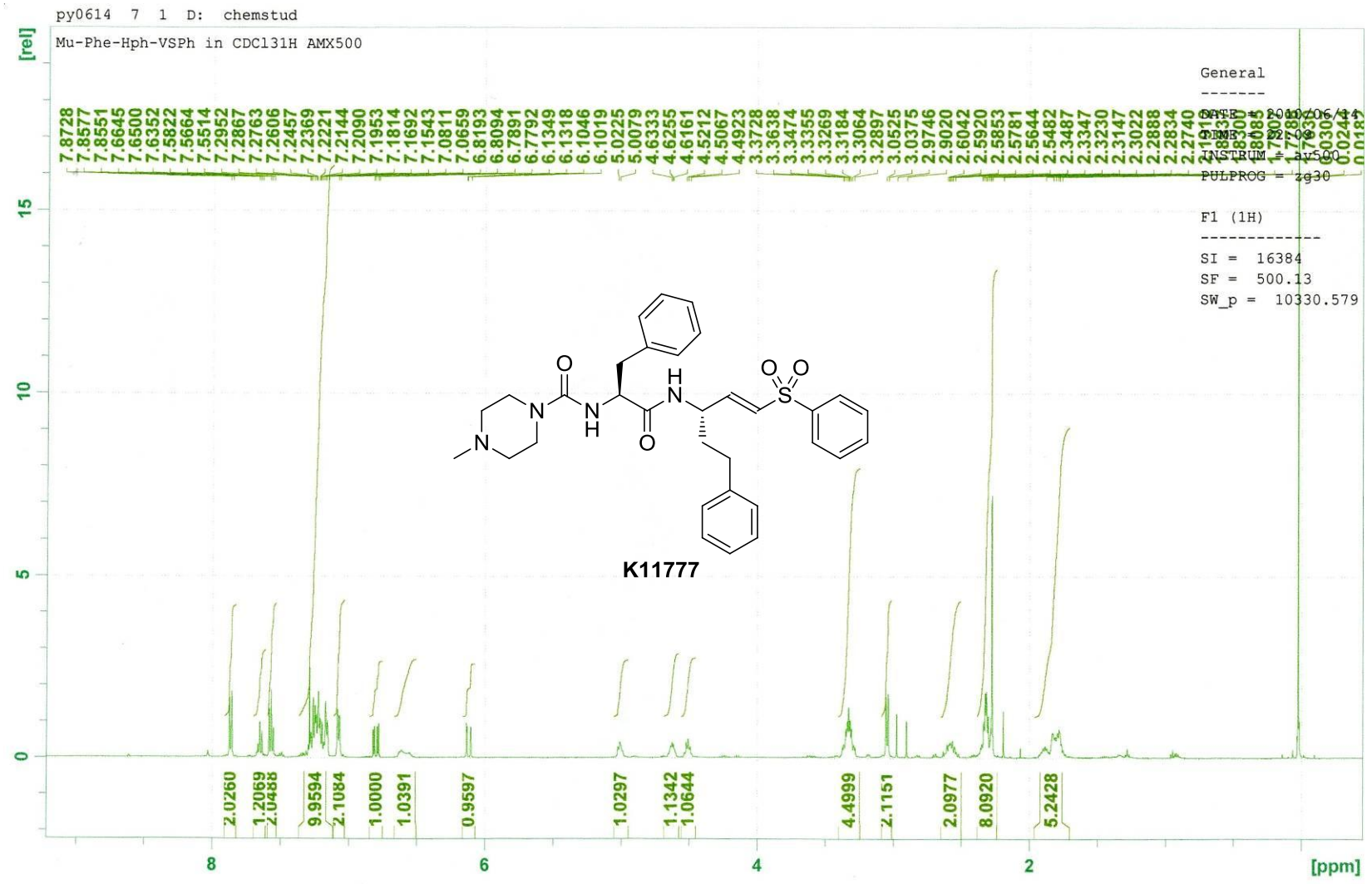
*** Processing Parameters ***

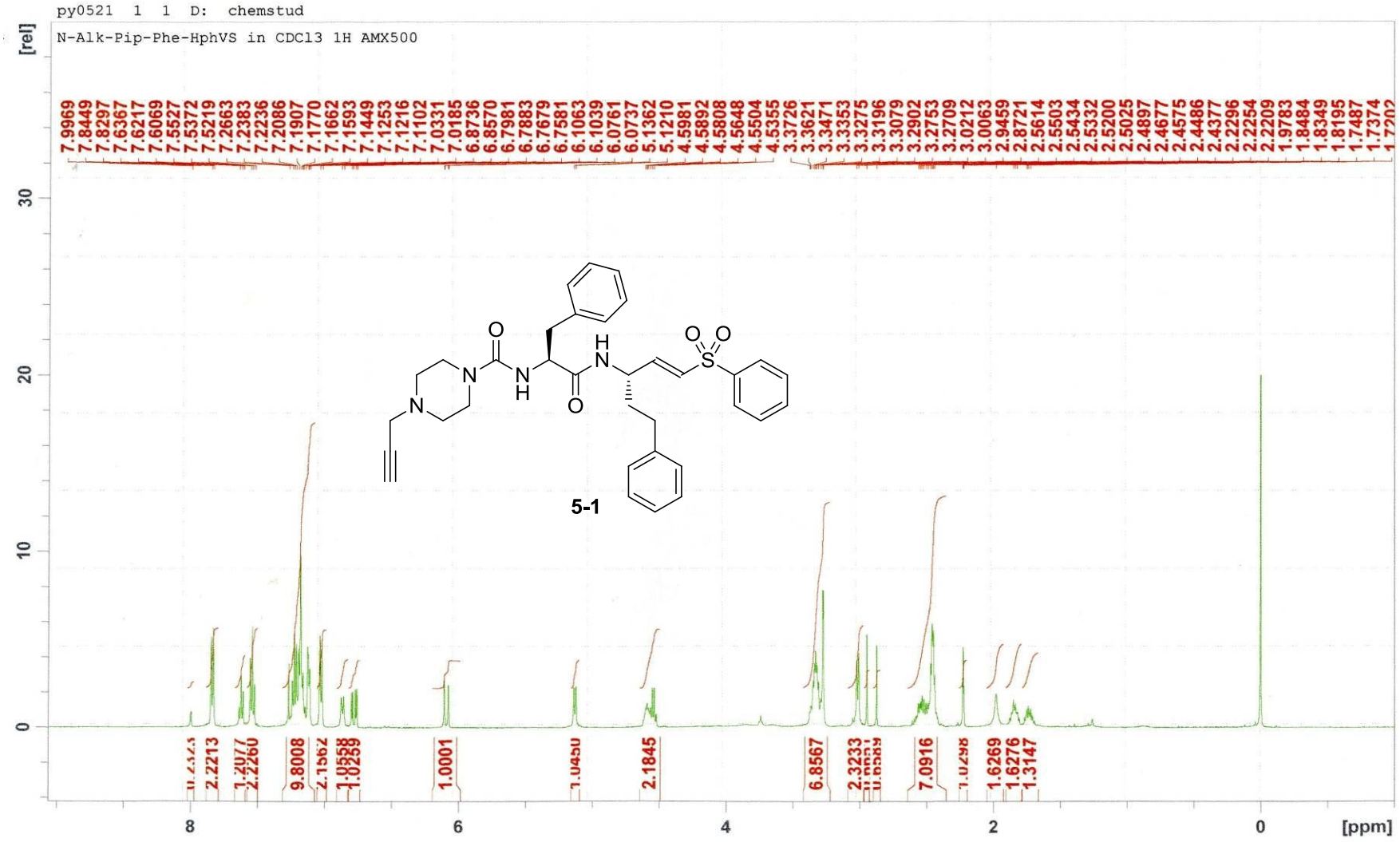
LB	:	0.30 Hz
SF	:	500.1300131 MHz

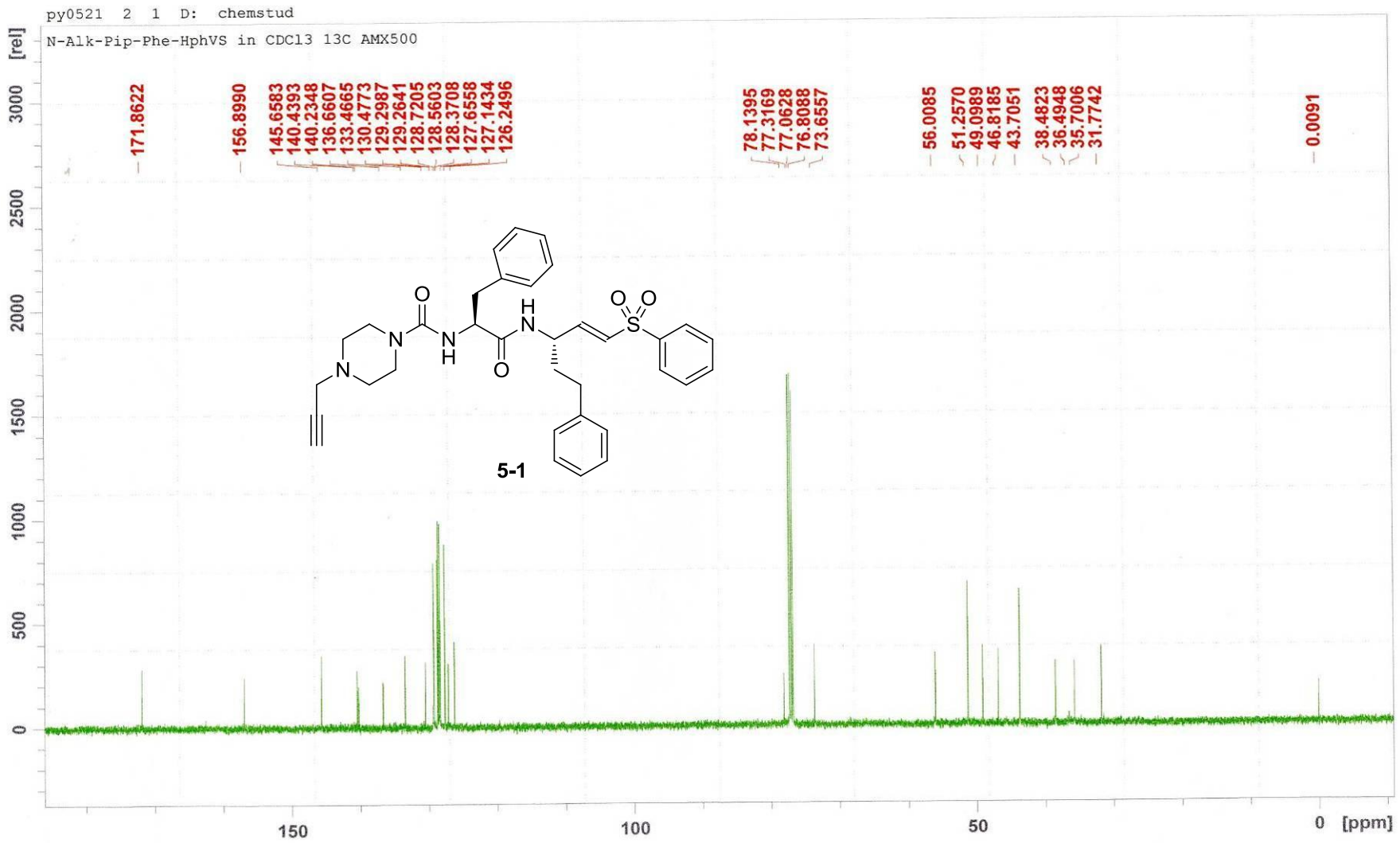
*** 1D NMR Plot Parameters ***

NUCLEUS	:	off

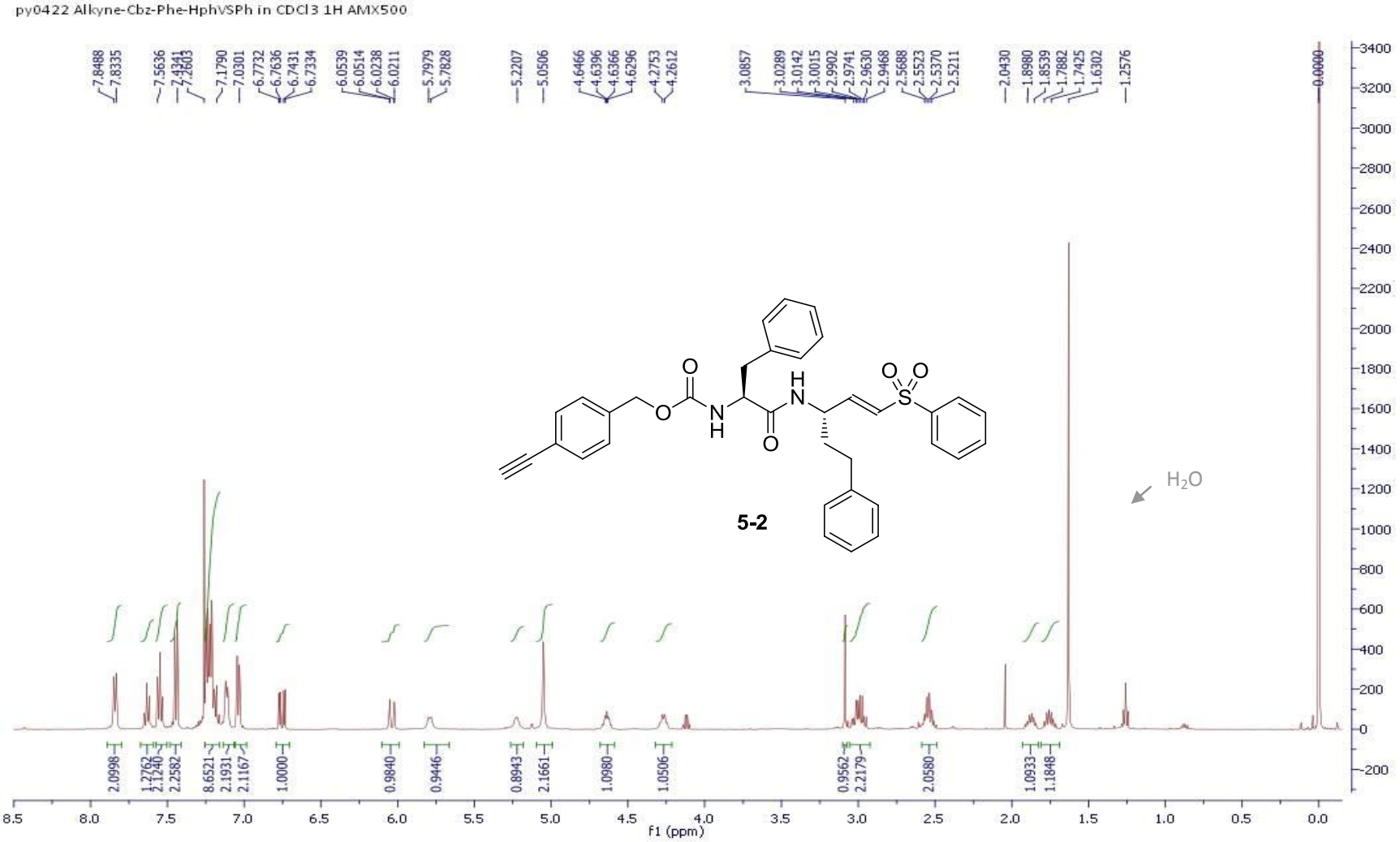

py0614 9 1 D: chemstud
N-Me-Pip-Phe-Hph-VSPh in CDCl3 1H AMX500
[rel]
General
F1 (1H)
SI = 16384
SF = 500.13
SW_p = 10330.579
PULPROG = zg30
K11002
[ppm]


py0614 7 1 D: chemstud
Mu-Phe-Hph-VSPh in CDCl31H AMX500
[rel]
General
DATE = 2010/06/14
INSTRUM = av500
PULPROG = zg30
F1 (1H)
SI = 16384
SF = 500.13
SW_p = 10330.579
K11777
[ppm]

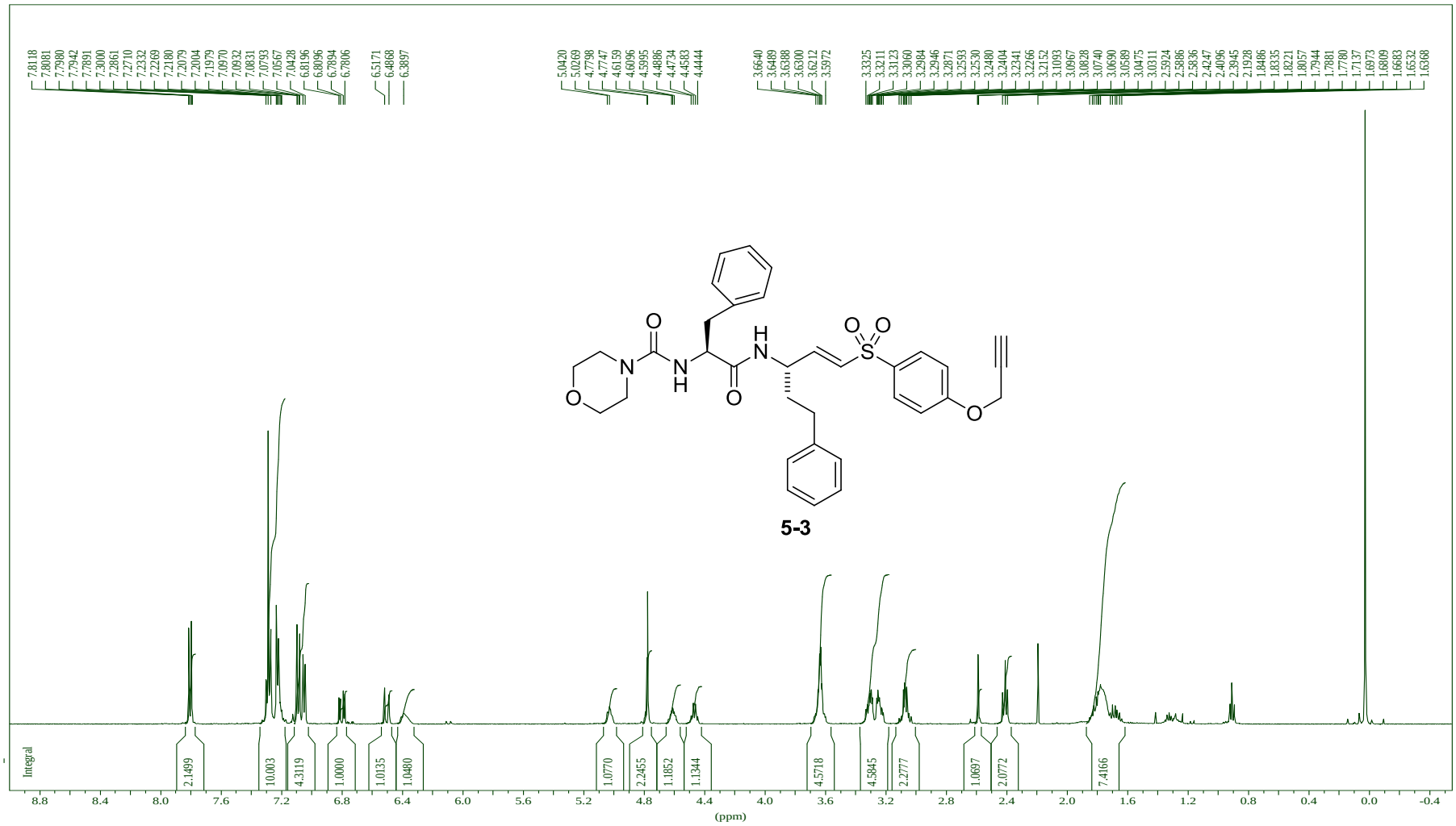

py0521 1 1 D: chemstud
N-Alk-Pip-Phe-HphVS in CDCl3 1H AMX500
[rel]
30
20
10
0
8
6
4
2
0
[ppm]
5-1
9.8008
1.0001
2.1845
6.8567
7.0916
1.6269
1.6276
1.3147

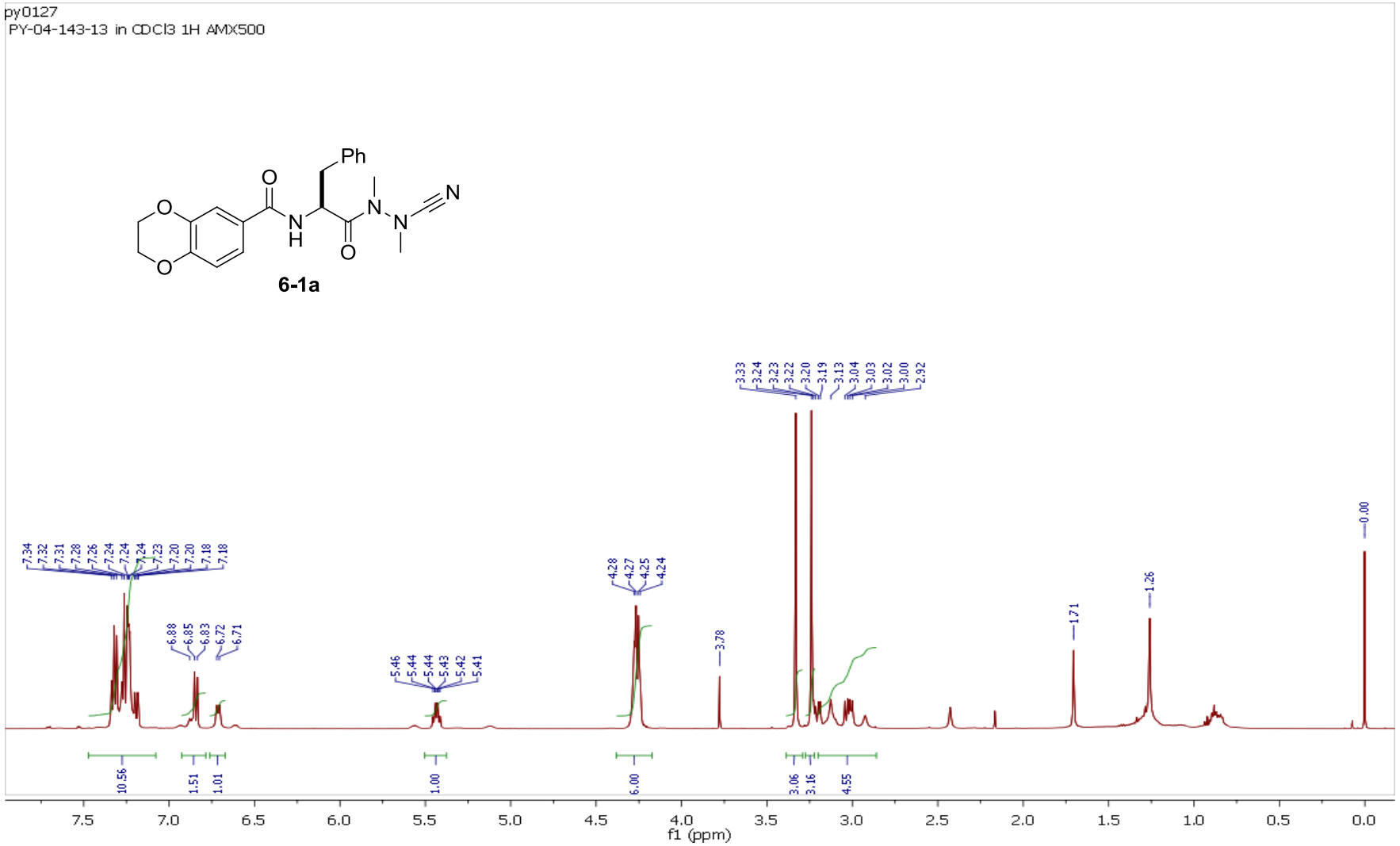

py0521 2 1 D: chemstud
N-Alk-Pip-Phe-HphVS in CDCl3 13C AMX500
[rel]
3000
2500
2000
1500
1000
500
0
171.8622
156.8990
145.6583
140.4393
140.2348
136.6607
133.4665
130.4773
129.2987
129.2641
128.7205
128.5603
128.3708
127.6558
127.1434
126.2496
78.1395
77.3169
77.0628
76.8088
73.6557
56.0085
51.2570
49.0989
46.8185
43.7051
38.4823
36.4948
35.7006
31.7742
0.0091
150
100
50
0
[ppm]

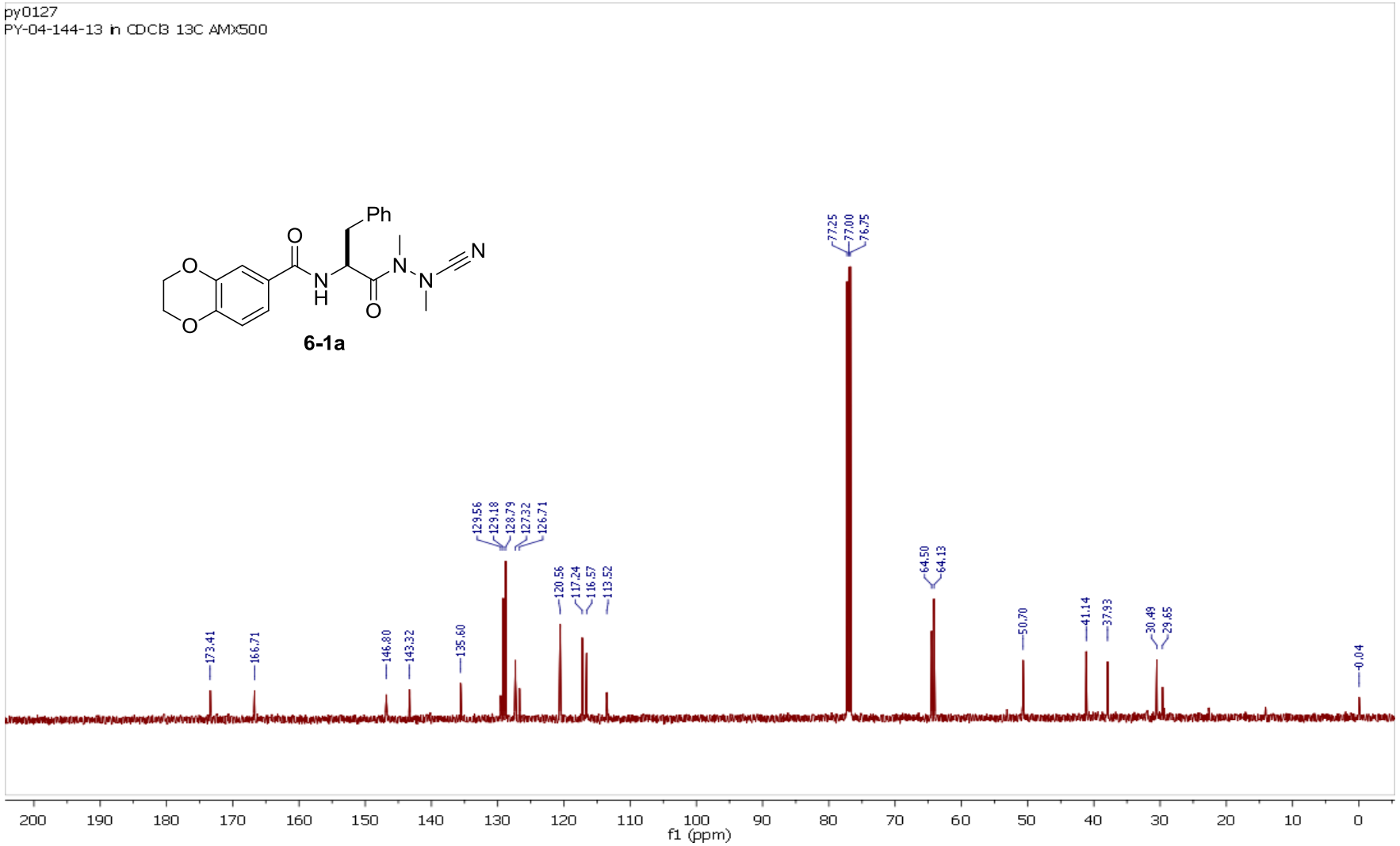
5-1

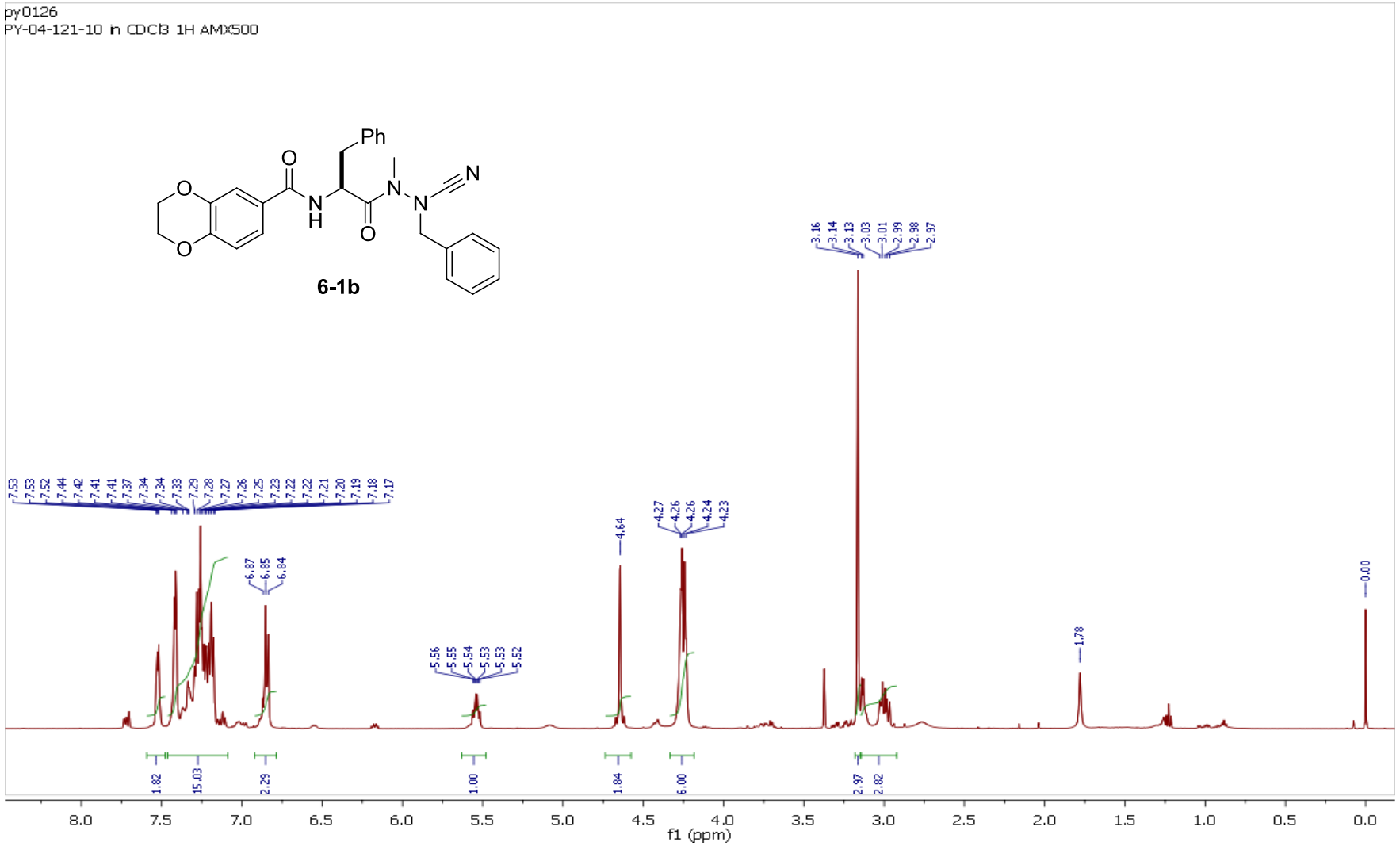

py0422 Alkyne-Cbz-Phe-HphVSPh in CDCl3 1H AMX500
5-2
H2O
f1 (ppm)

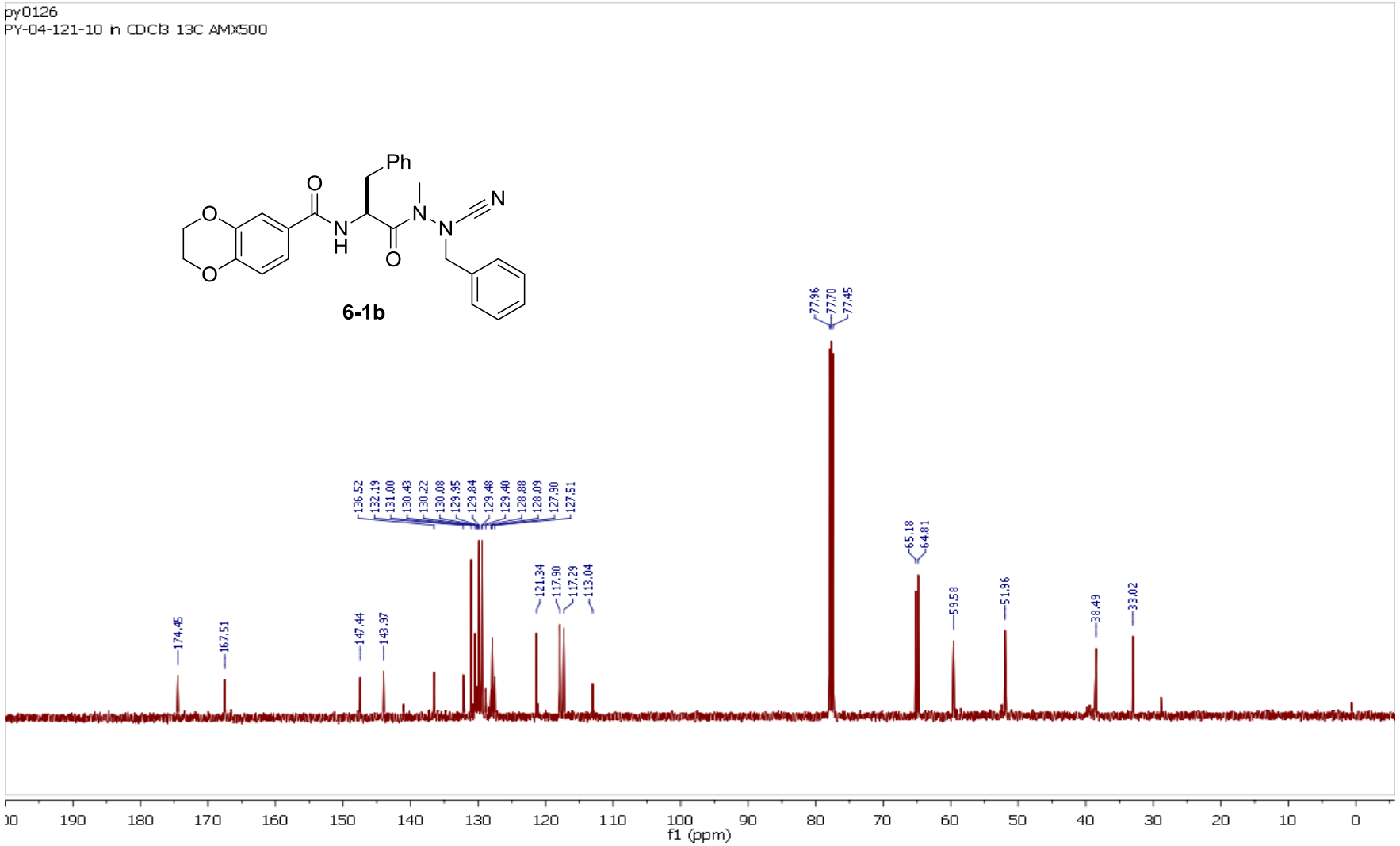
Mu-Phe-Hph-VS-OAlk in CDCl31H AMX500

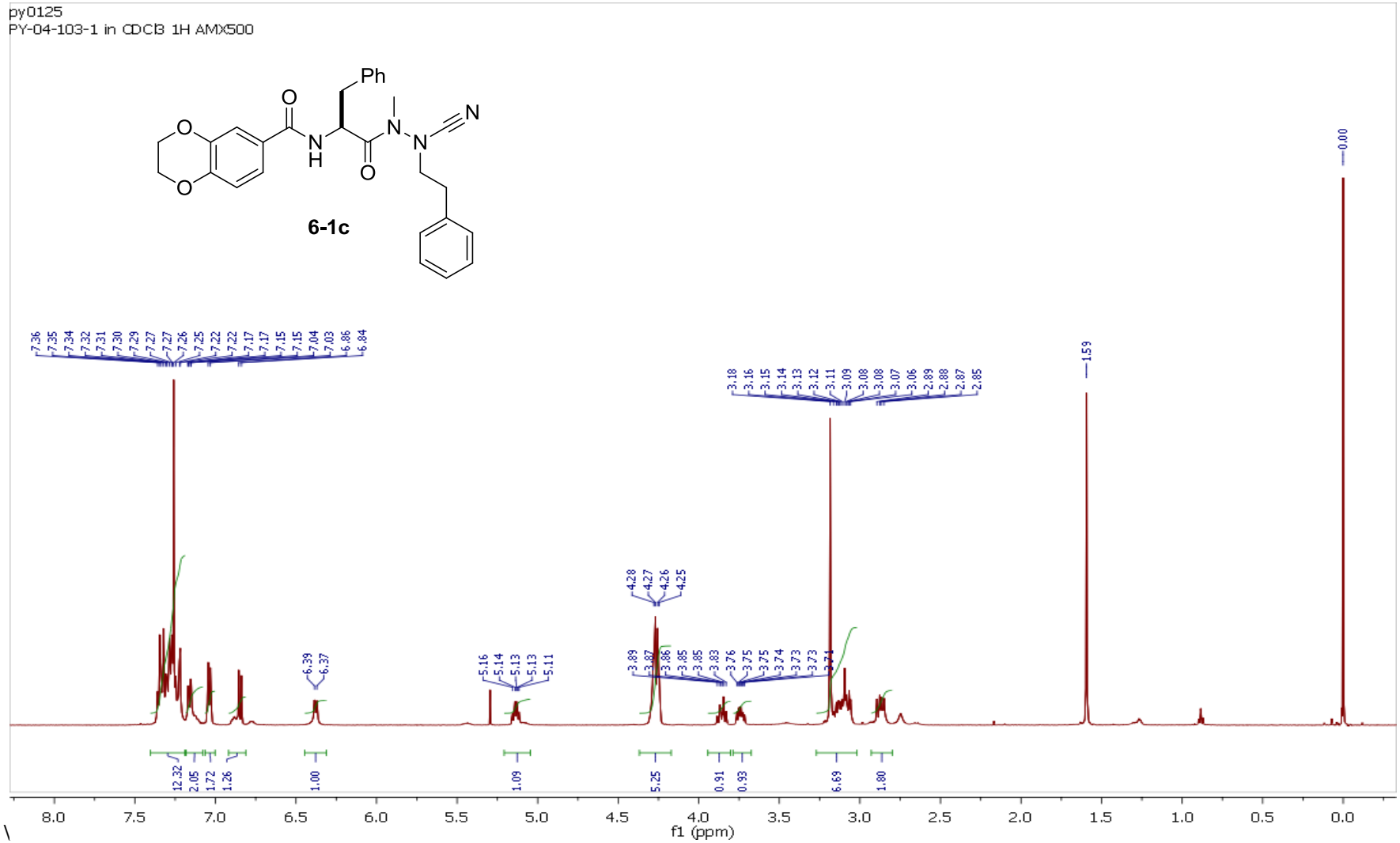


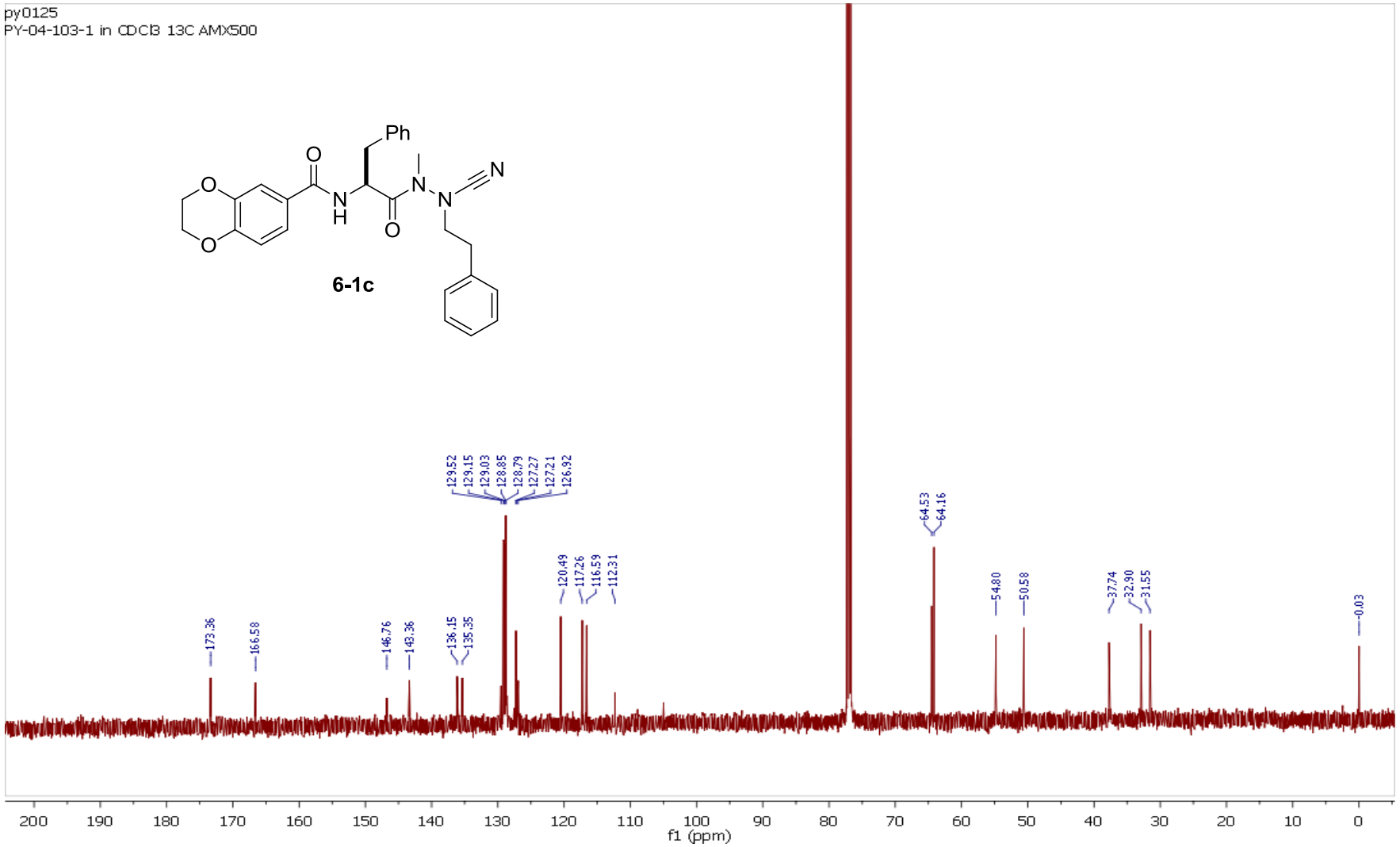

py0127
PY-04-143-13 in CDCl3 1H AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1a
7.34
7.32
7.31
7.28
7.26
7.24
7.24
7.24
7.23
7.20
7.20
7.18
7.18
6.88
6.85
6.83
6.72
6.71
5.46
5.44
5.44
5.43
5.42
5.41
4.28
4.27
4.25
4.24
3.78
3.33
3.24
3.23
3.22
3.20
3.19
3.13
3.04
3.03
3.02
3.00
2.92
1.71
1.26
0.00
10.56
1.51
1.01
1.00
6.00
3.06
3.16
4.55
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

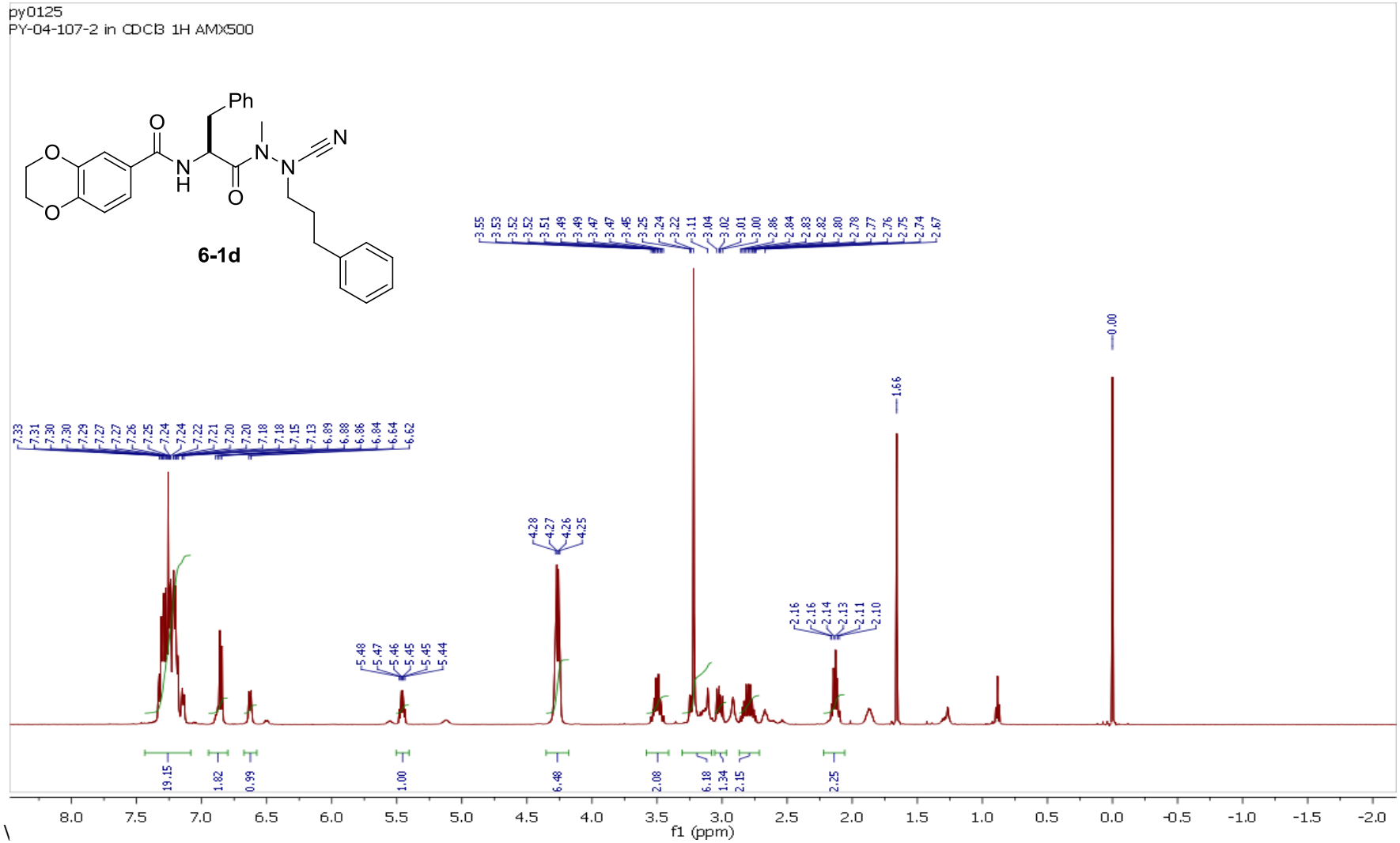

py0127
PY-04-144-13 in CDCl3 13C AMX500
Ph
O
N
H
N
N
N
O
O
O
6-1a
77.25
77.00
76.75
129.56
129.18
128.79
127.32
126.71
120.56
117.24
116.57
113.52
64.50
64.13
50.70
41.14
37.93
30.49
29.65
173.41
166.71
146.80
143.32
135.60
0.04
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

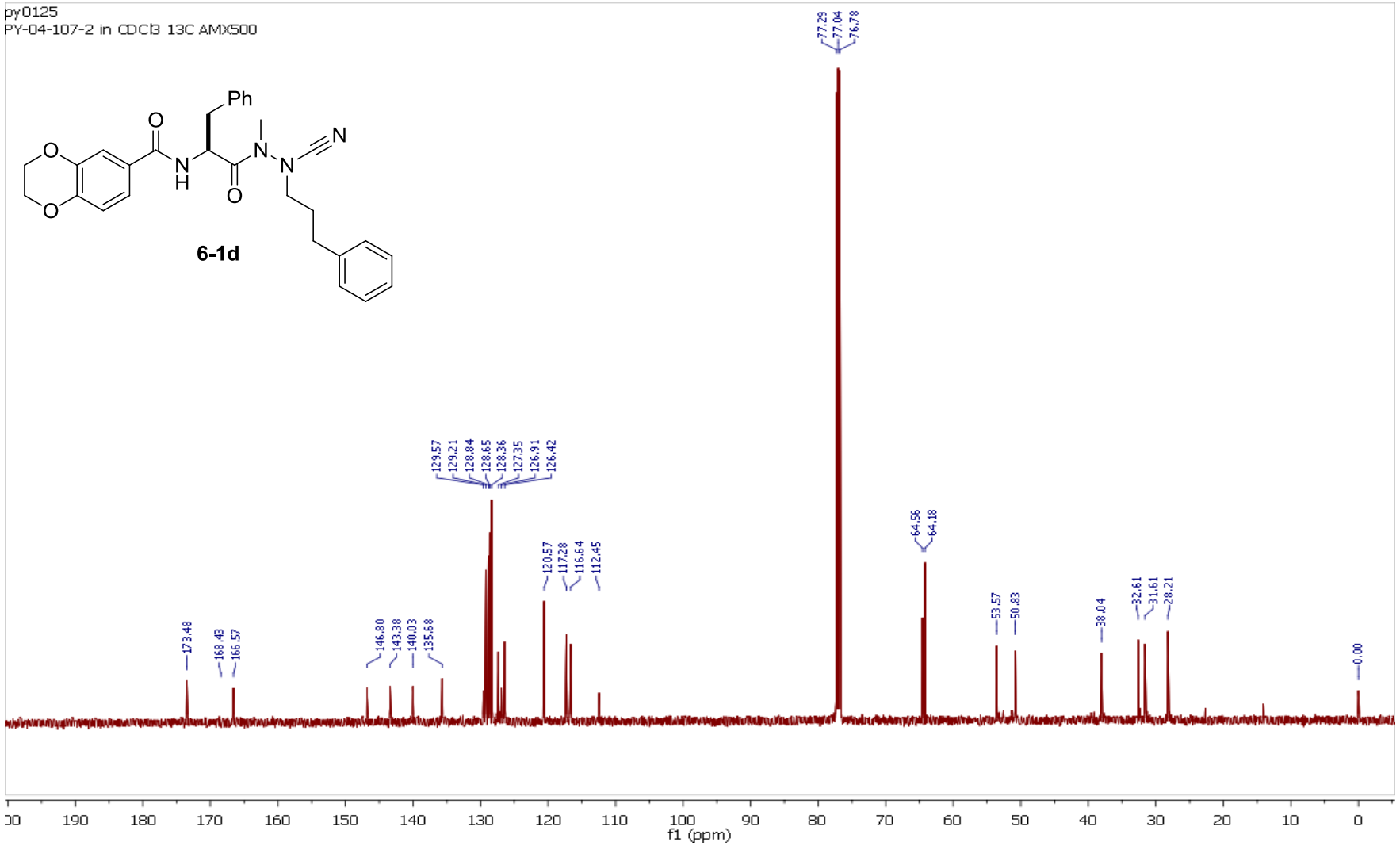

py0126
PY-04-121-10 in CDCl3 1H AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1b
7.53
7.53
7.52
7.44
7.42
7.41
7.41
7.37
7.34
7.34
7.33
7.29
7.28
7.27
7.26
7.25
7.23
7.22
7.22
7.21
7.20
7.19
7.18
7.17
6.87
6.85
6.84
5.56
5.55
5.54
5.53
5.53
5.52
4.64
4.27
4.26
4.26
4.24
4.23
3.16
3.14
3.13
3.03
3.01
2.99
2.98
2.97
1.78
0.00
1.82
15.03
2.29
1.00
1.84
6.00
2.97
2.82
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

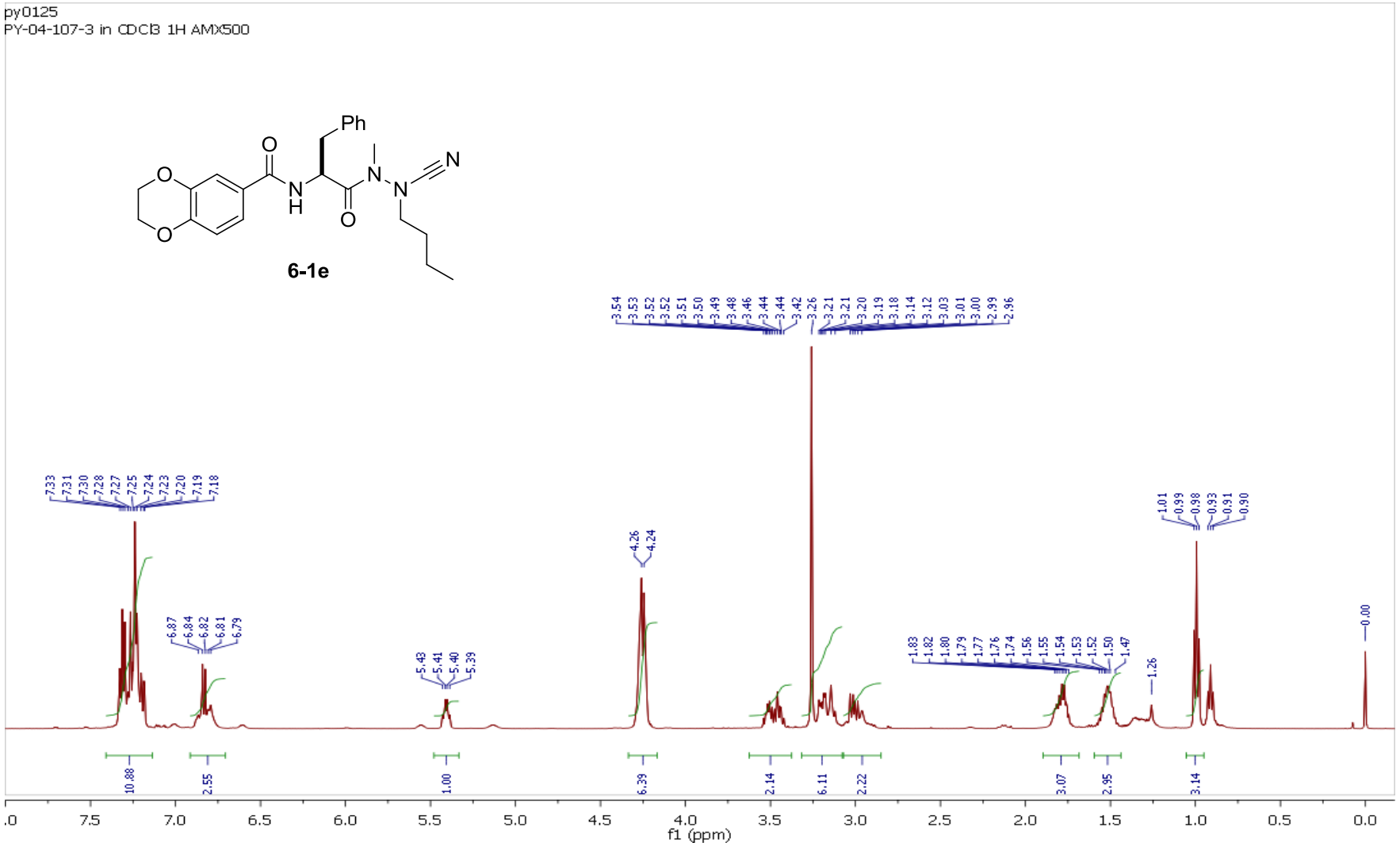

py0126
PY-04-121-10 in CDCl3 13C AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1b
77.96
77.70
77.45
136.52
132.19
131.00
130.43
130.22
130.08
129.95
129.84
129.48
129.40
128.88
128.09
127.90
127.51
121.34
117.90
117.29
113.04
65.18
64.81
59.58
51.96
38.49
33.02
174.45
167.51
147.44
143.97
f1 (ppm)

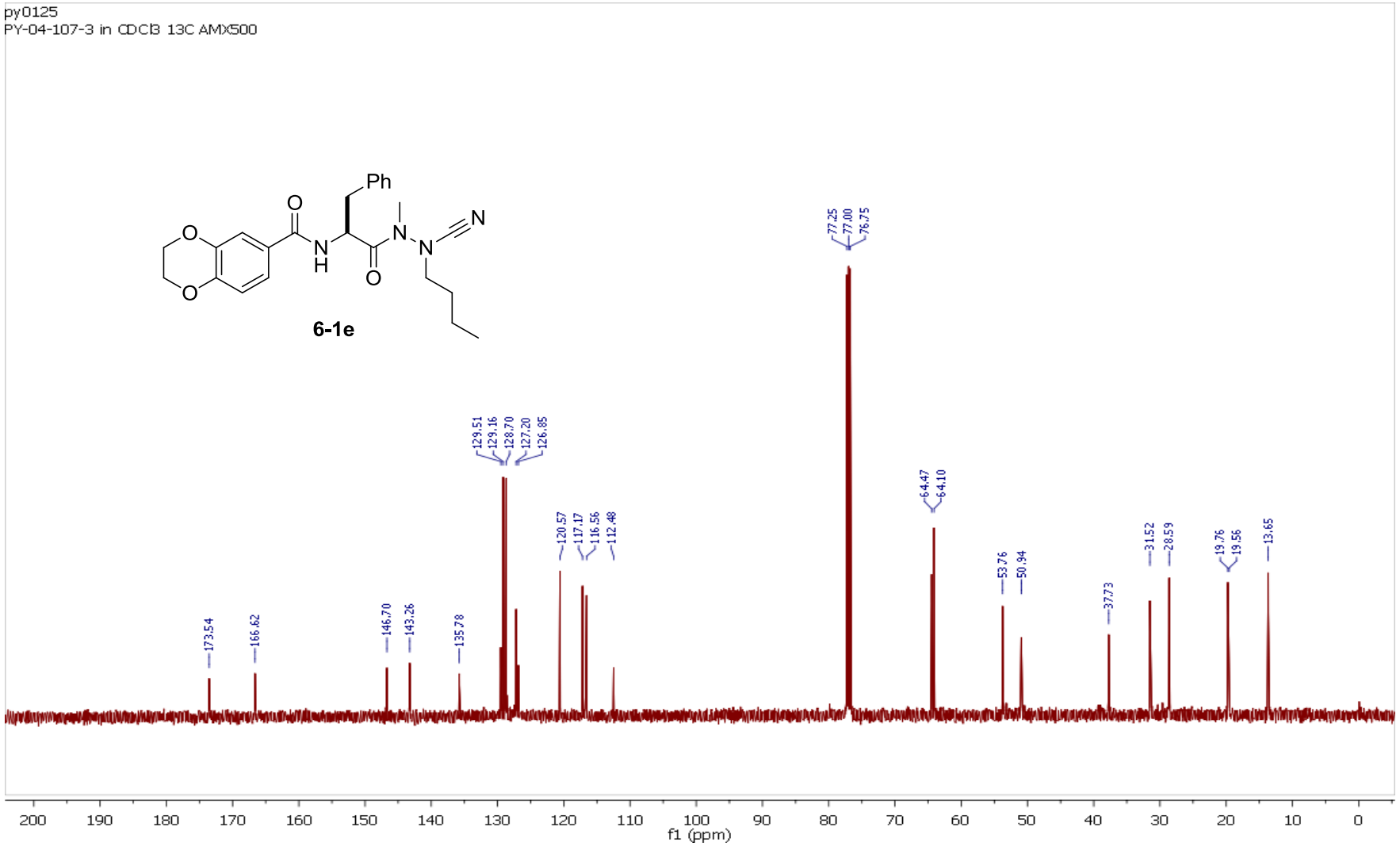

py0125
PY-04-103-1 in CDCl3 1H AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1c
7.36
7.35
7.34
7.32
7.31
7.30
7.29
7.27
7.27
7.26
7.25
7.22
7.22
7.17
7.17
7.15
7.15
7.04
7.03
6.86
6.84
6.39
6.37
5.16
5.14
5.13
5.13
5.11
4.28
4.27
4.26
4.25
3.89
3.87
3.86
3.85
3.85
3.83
3.76
3.75
3.75
3.74
3.73
3.73
3.71
3.18
3.16
3.15
3.14
3.13
3.12
3.11
3.09
3.08
3.08
3.07
3.06
2.89
2.88
2.87
2.85
1.59
0.00
12.32
2.05
1.72
1.26
1.00
1.09
5.25
0.91
0.93
6.69
1.80
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


py0125
PY-04-103-1 in CDCl3 13C AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1c
129.52
129.15
129.03
128.85
128.79
127.27
127.21
126.92
173.36
166.58
146.76
143.36
136.15
135.35
120.49
117.26
116.59
112.31
64.53
64.16
54.80
50.58
37.74
32.90
31.55
0.03
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

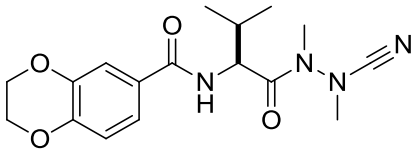

py0125
PY-04-107-2 in CDCl3 1H AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1d
3.55
3.53
3.52
3.52
3.51
3.49
3.49
3.47
3.47
3.45
3.25
3.24
3.22
3.11
3.04
3.02
3.01
3.00
2.86
2.84
2.83
2.82
2.80
2.78
2.77
2.76
2.75
2.74
2.67
0.00
1.66
7.33
7.31
7.30
7.30
7.29
7.27
7.27
7.26
7.25
7.24
7.24
7.22
7.21
7.20
7.20
7.18
7.18
7.15
7.13
6.89
6.88
6.86
6.84
6.64
6.62
4.28
4.27
4.26
4.25
5.48
5.47
5.46
5.45
5.45
5.44
2.16
2.16
2.14
2.13
2.11
2.10
19.15
1.82
0.99
1.00
6.48
2.08
6.18
1.34
2.15
2.25
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
f1 (ppm)


py0125
PY-04-107-2 in CDCl3 13C AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1d
77.29
77.04
76.78
129.57
129.21
128.84
128.65
128.36
127.35
126.91
126.42
120.57
117.28
116.64
112.45
64.56
64.18
53.57
50.83
38.04
32.61
31.61
28.21
173.48
168.43
166.57
146.80
143.38
140.03
135.68
0.00
f1 (ppm)

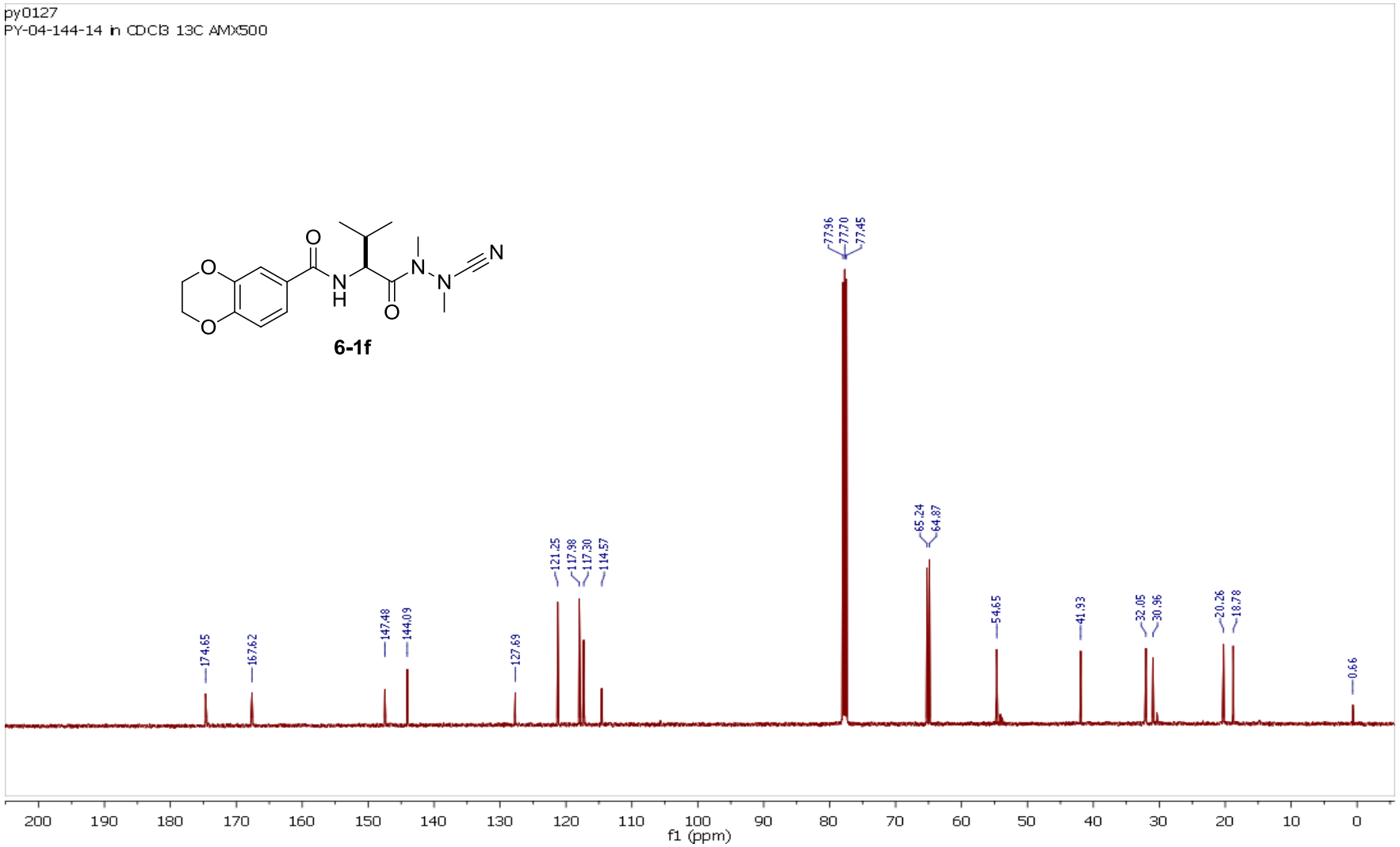

py0125
PY-04-107-3 in CDCl3 1H AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1e
7.33
7.31
7.30
7.28
7.27
7.25
7.24
7.23
7.20
7.19
7.18
6.87
6.84
6.82
6.81
6.79
5.43
5.41
5.40
5.39
4.26
4.24
3.54
3.53
3.52
3.52
3.51
3.50
3.49
3.48
3.46
3.44
3.44
3.42
3.26
3.21
3.21
3.20
3.19
3.18
3.14
3.12
3.03
3.01
3.00
2.99
2.96
1.83
1.82
1.80
1.79
1.77
1.76
1.74
1.56
1.55
1.54
1.53
1.52
1.50
1.47
1.26
1.01
0.99
0.98
0.93
0.91
0.90
0.00
10.88
2.55
1.00
6.39
2.14
6.11
2.22
3.07
2.95
3.14
f1 (ppm)

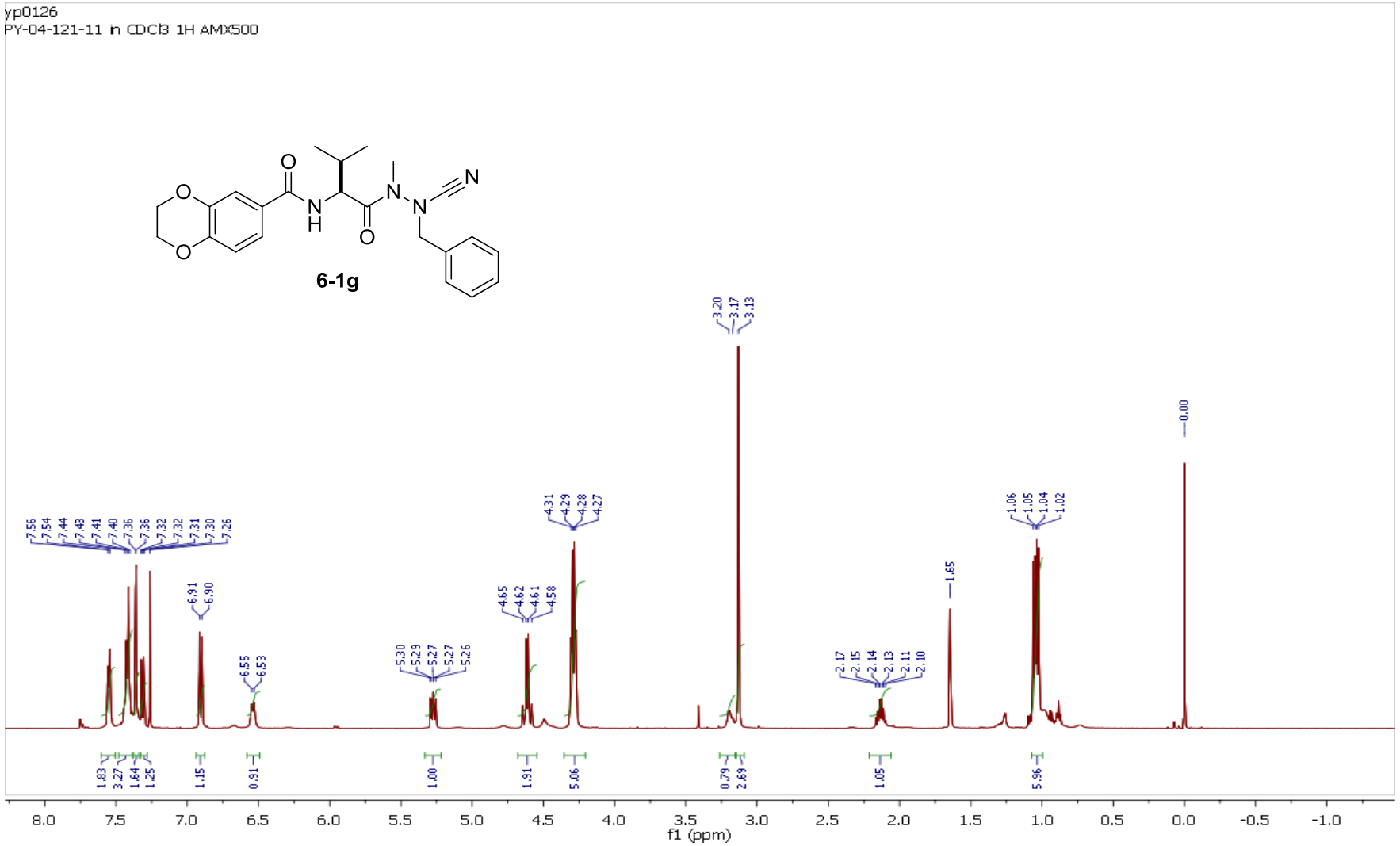

py0125
PY-04-107-3 in CDCl3 13C AMX500
Ph
O
N
H
O
N
N
N
O
O
6-1e
77.25
77.00
76.75
129.51
129.16
128.70
127.20
126.85
120.57
117.17
116.56
112.48
64.47
64.10
53.76
50.94
37.73
31.52
28.59
19.76
19.56
13.65
173.54
166.62
146.70
143.26
135.78
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

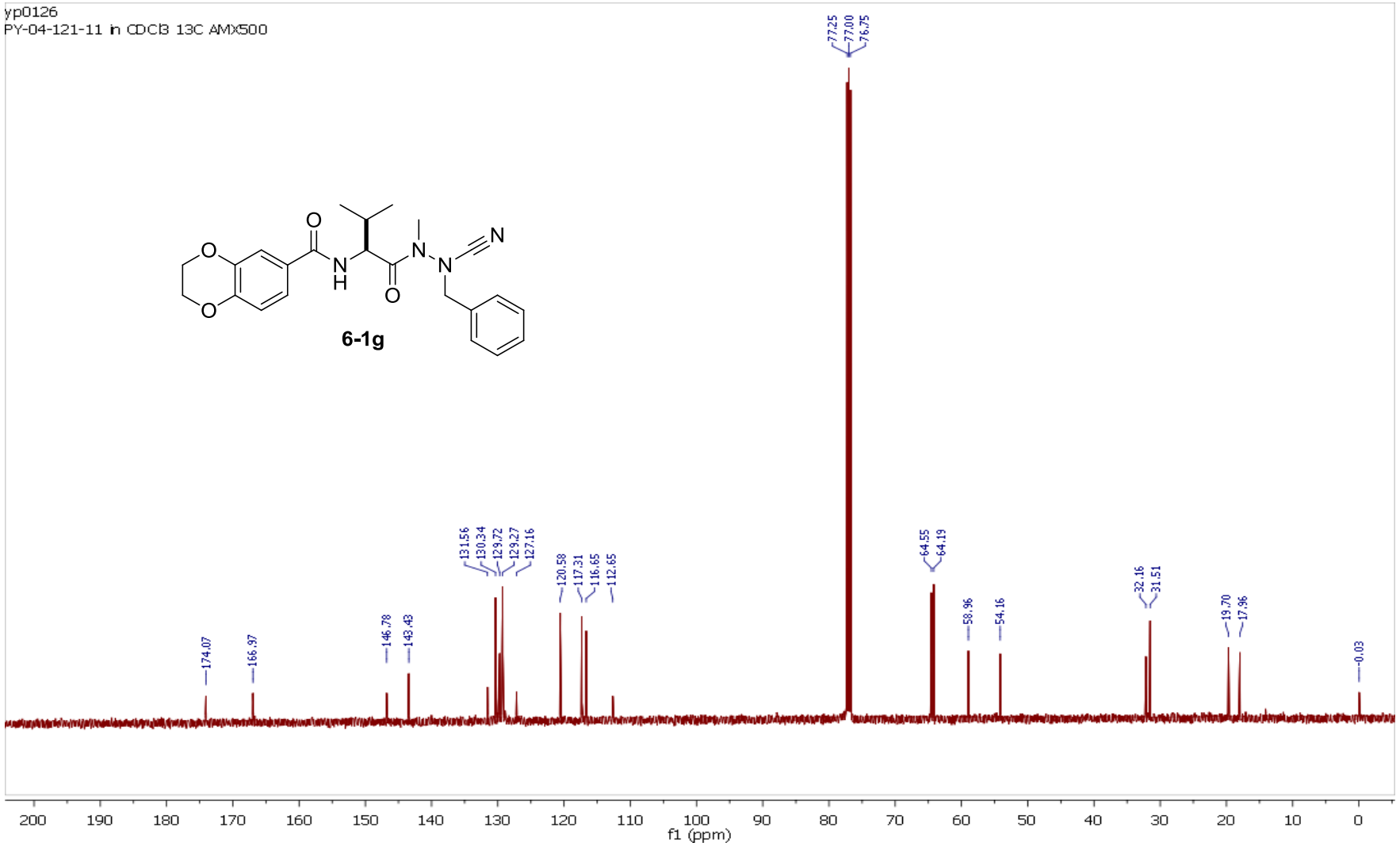
py0127
PY-04-144-14 in CDCl3 1H AMX500

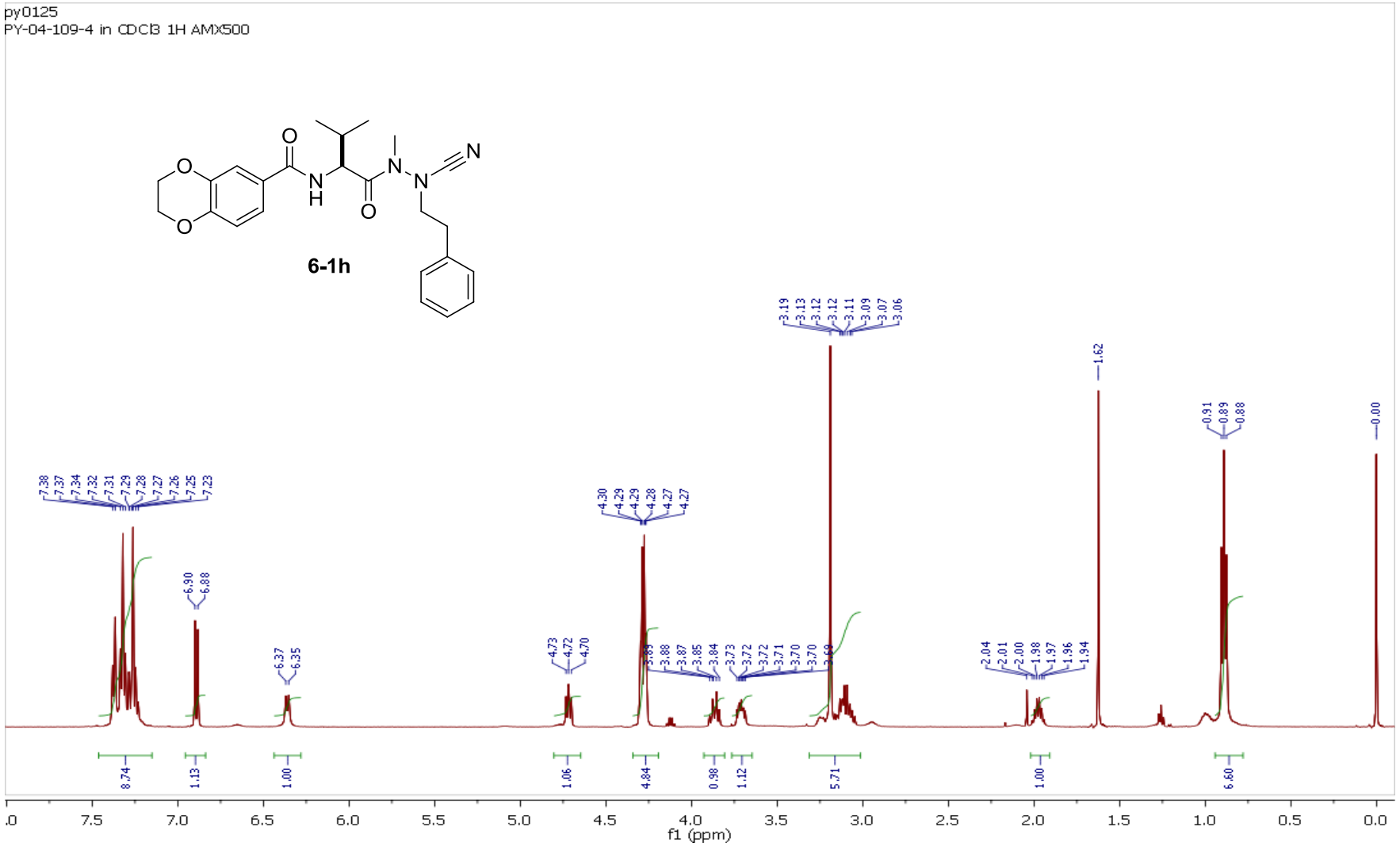


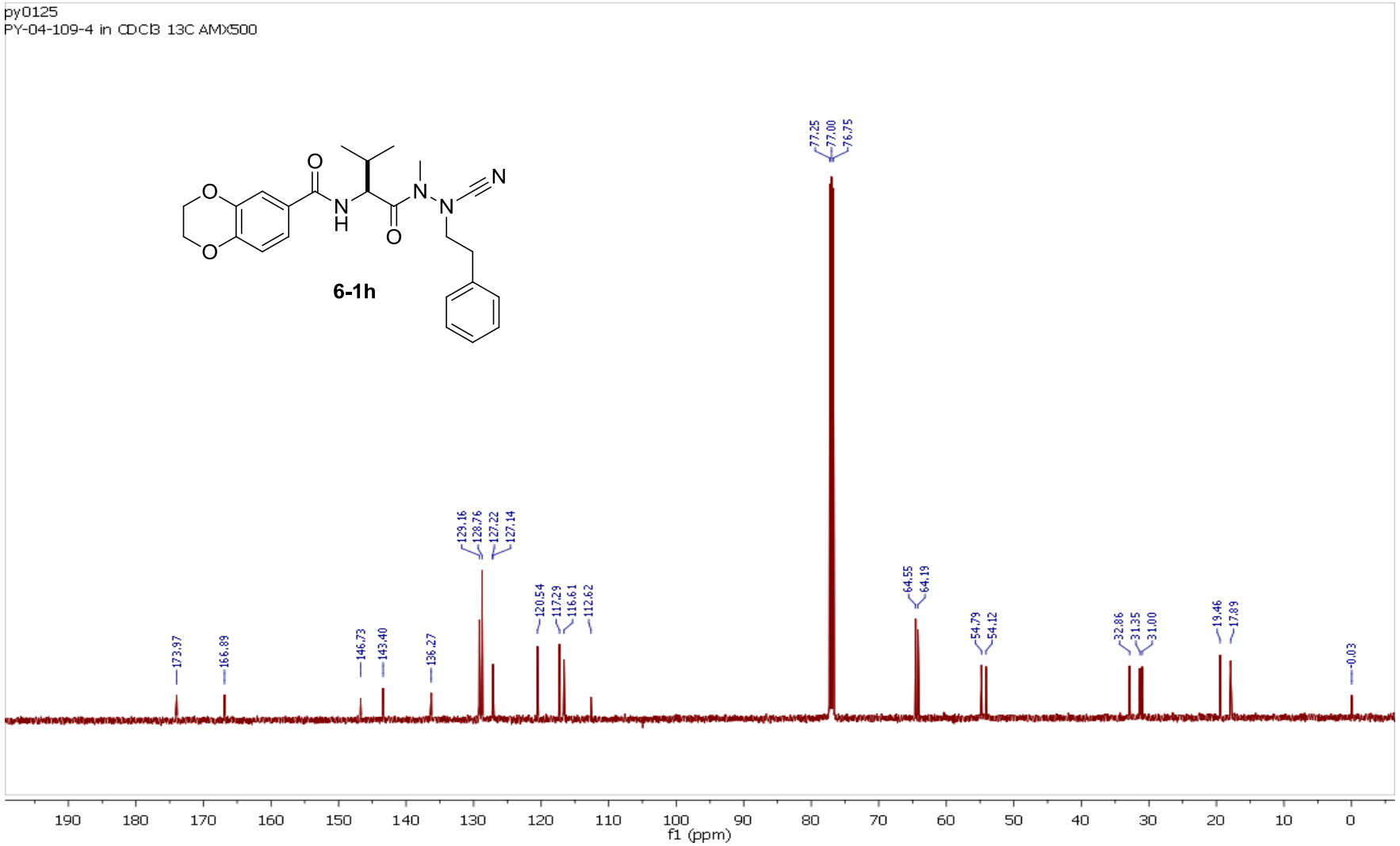
6-1f

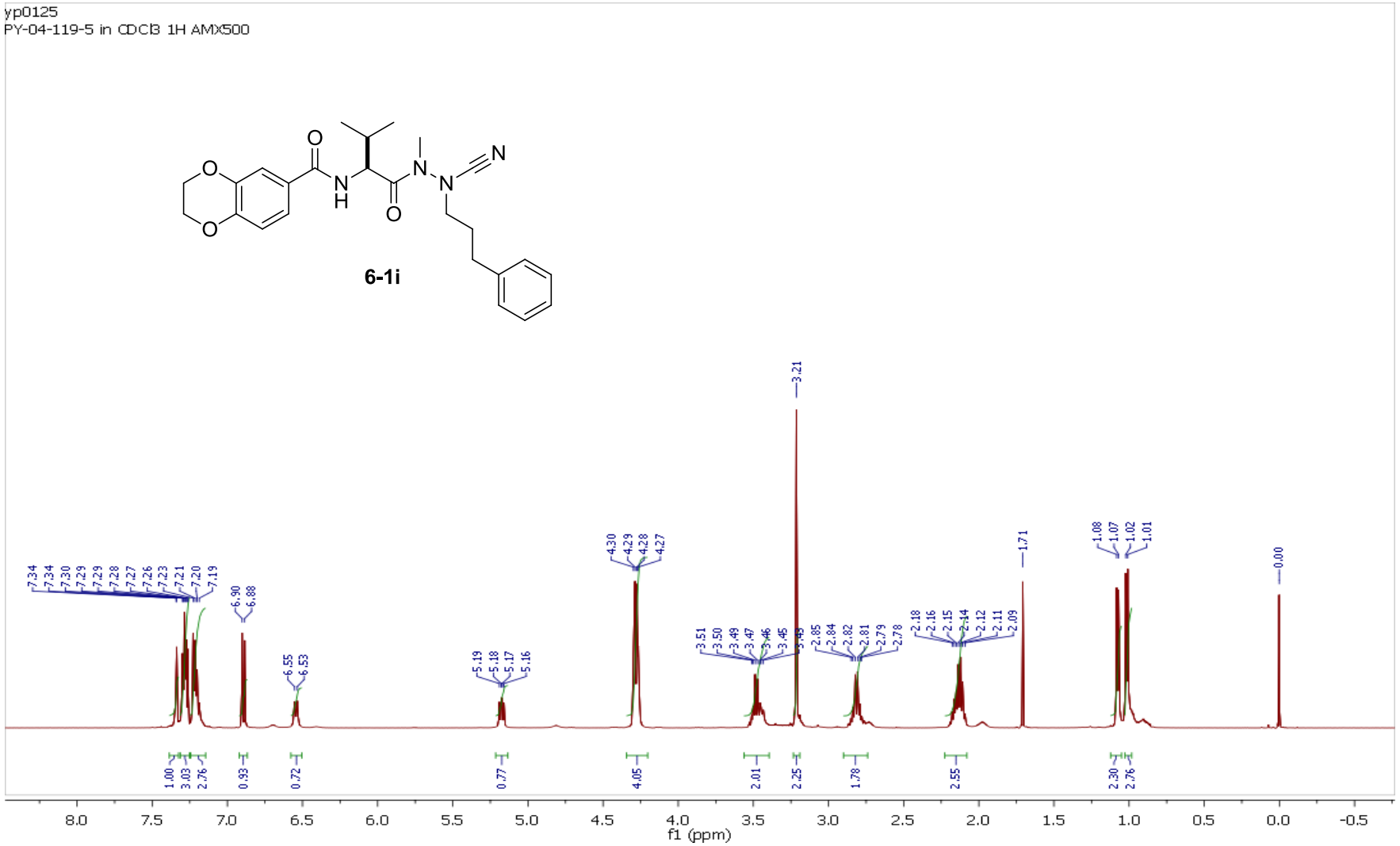

py0127
PY-04-144-14 in CDCl3 13C AMX500
6-1f
77.96
77.70
77.45
65.24
64.87
121.25
117.98
117.30
114.57
174.65
167.62
147.48
144.09
127.69
54.65
41.93
32.05
30.96
20.26
18.78
0.66
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

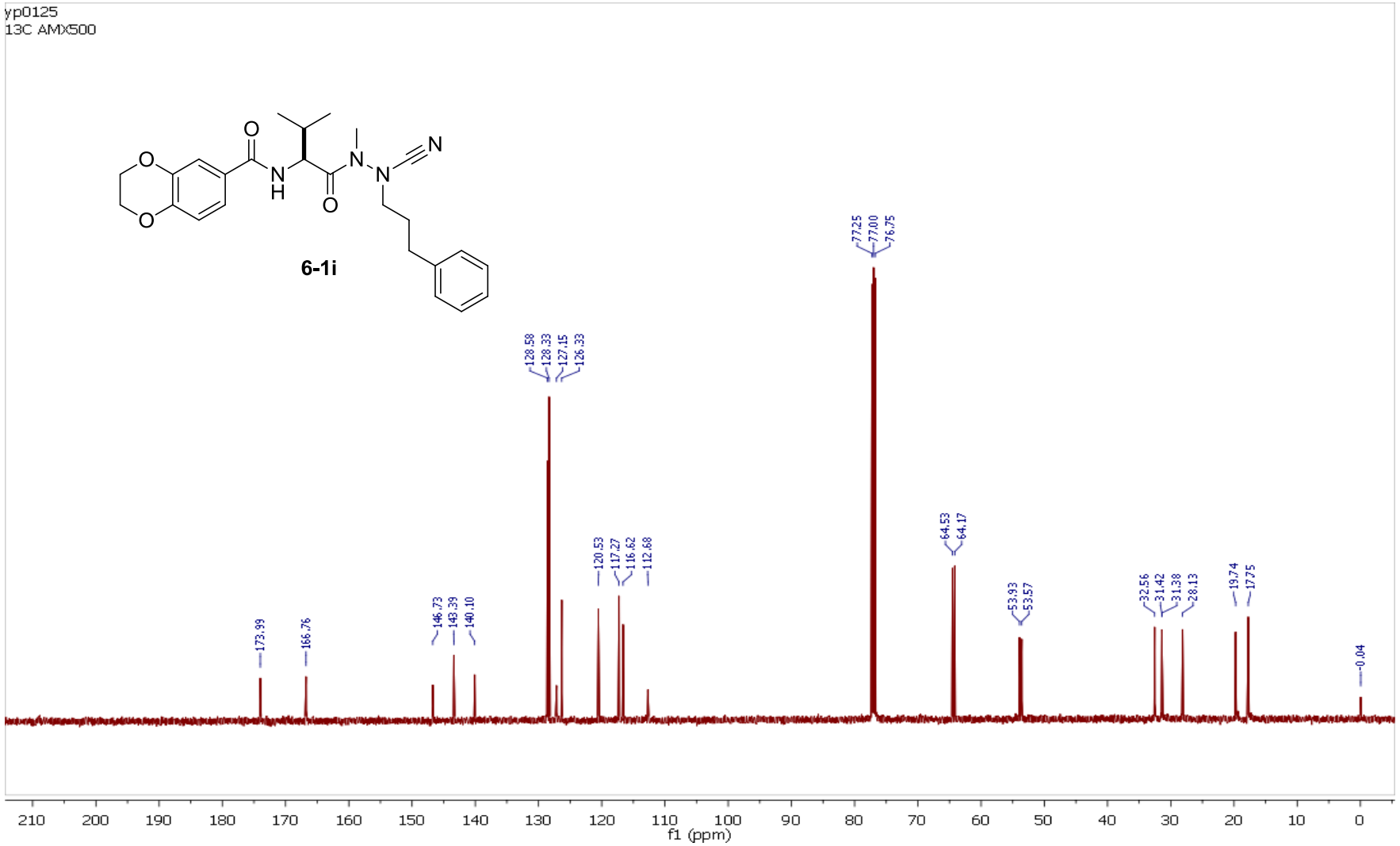

yp0126
PY-04-121-11 in CDCl3 1H AMX500
6-1g
f1 (ppm)

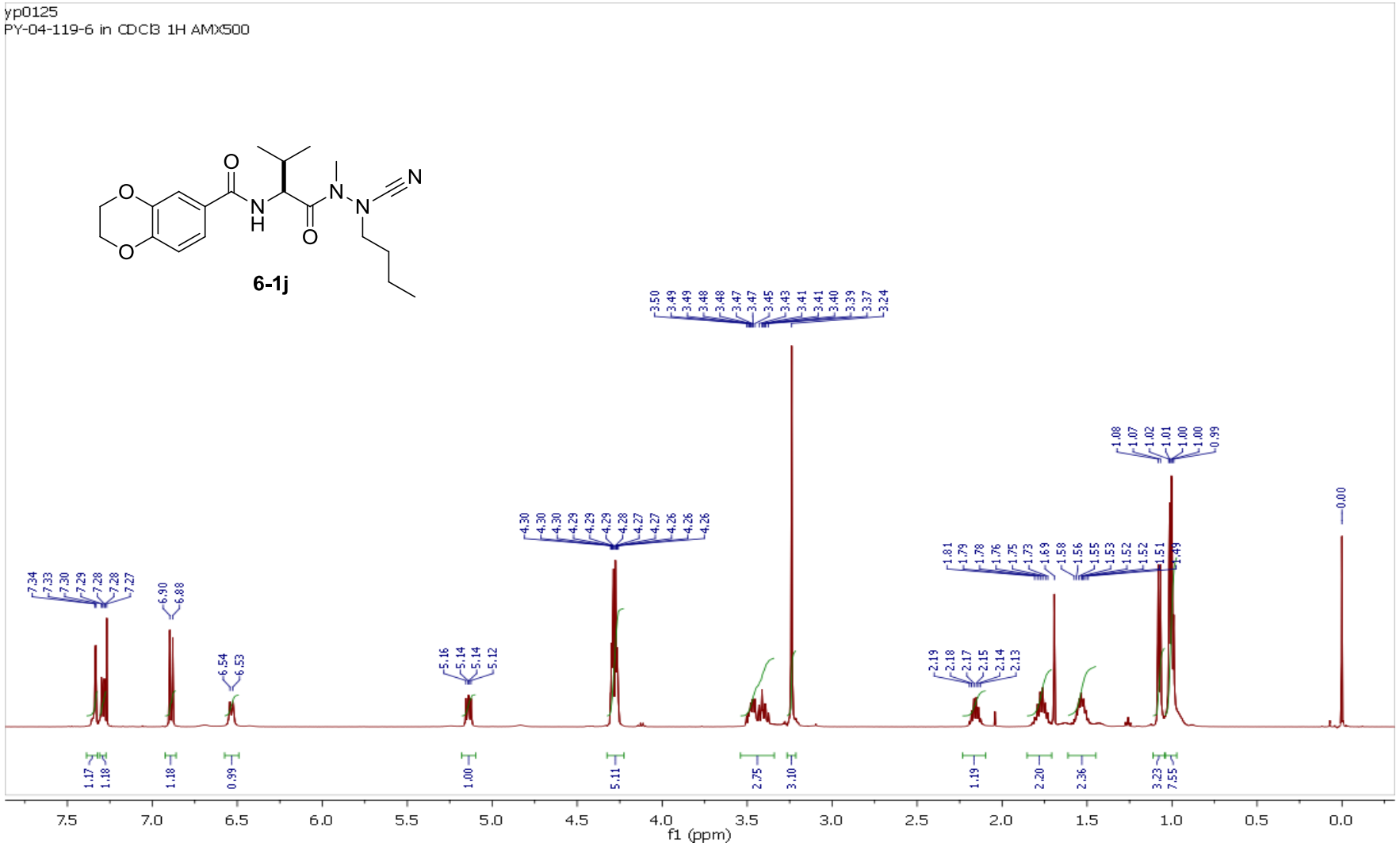

yp0126
PY-04-121-11 in CDCl3 13C AMX500
6-1g
77.25
77.00
76.75
174.07
166.97
146.78
143.43
131.56
130.34
129.72
129.27
127.16
120.58
117.31
116.65
112.65
64.55
64.19
58.96
54.16
32.16
31.51
19.70
17.96
-0.03
f1 (ppm)

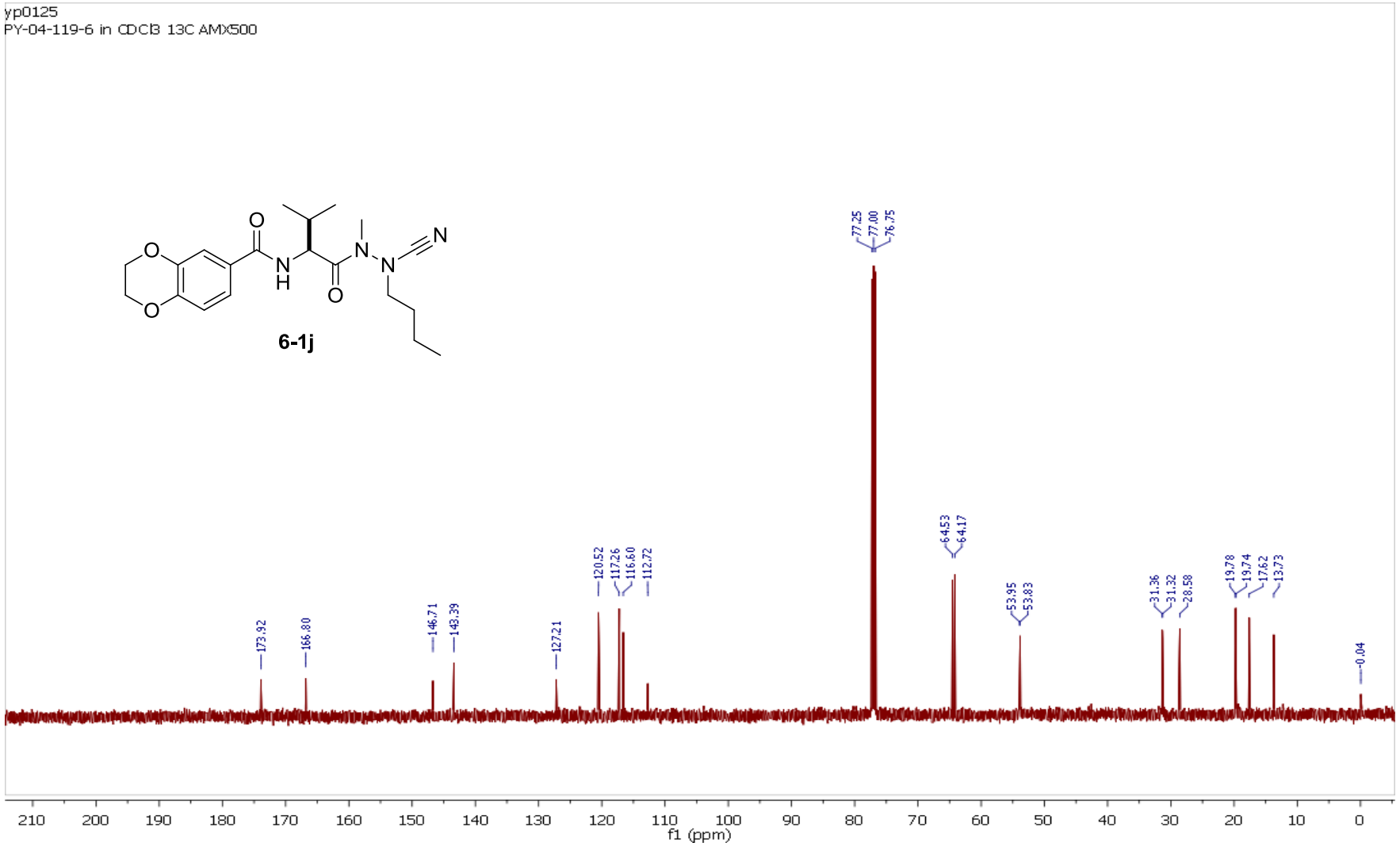

py0125
PY-04-109-4 in CDCl3 1H AMX500
6-1h
f1 (ppm)

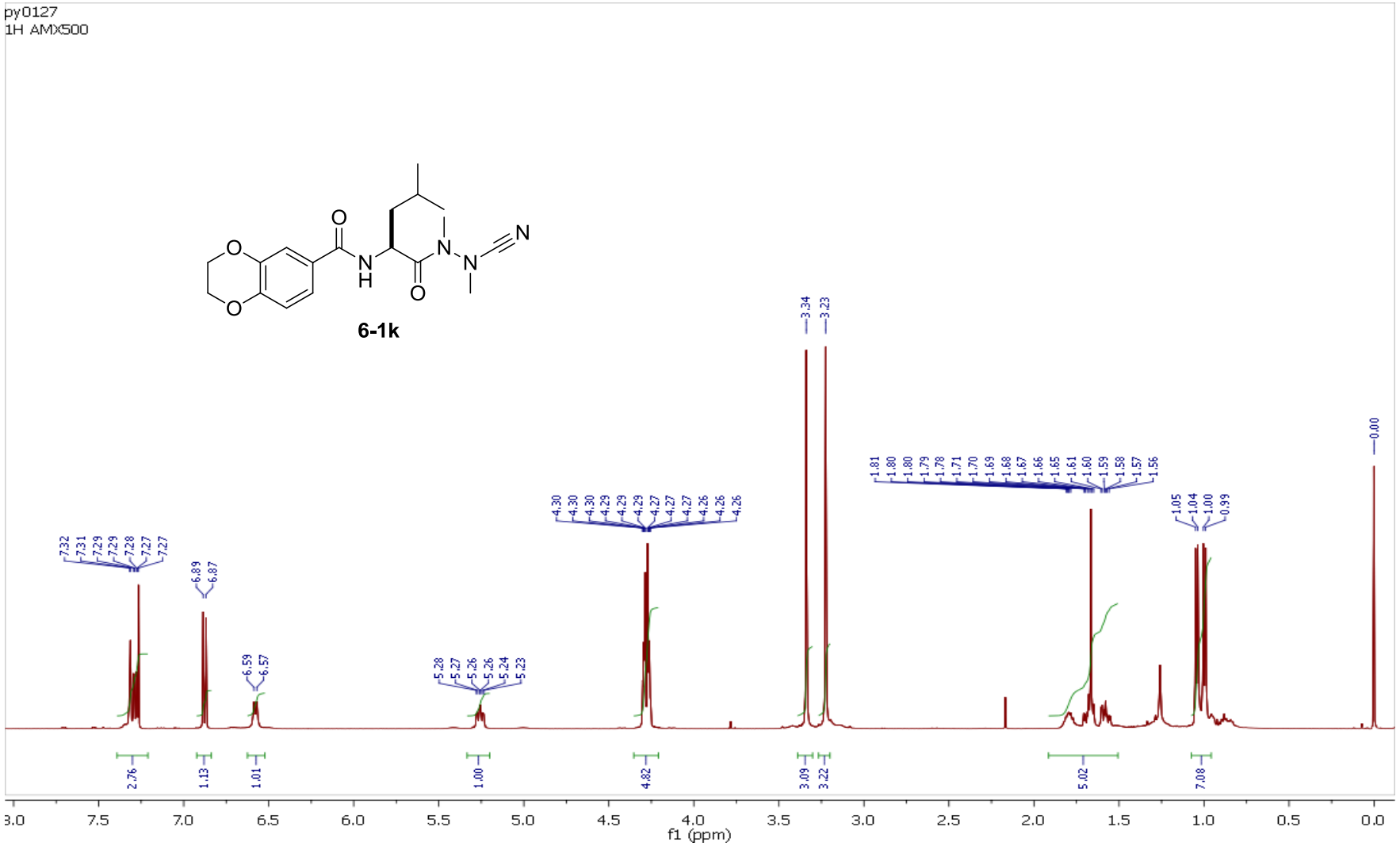

py0125
PY-04-109-4 in CDCl3 13C AMX500
6-1h
77.25
77.00
76.75
173.97
166.89
146.73
143.40
136.27
129.16
128.76
127.22
127.14
120.54
117.29
116.61
112.62
64.55
64.19
54.79
54.12
32.86
31.35
31.00
19.46
17.89
-0.03
f1 (ppm)

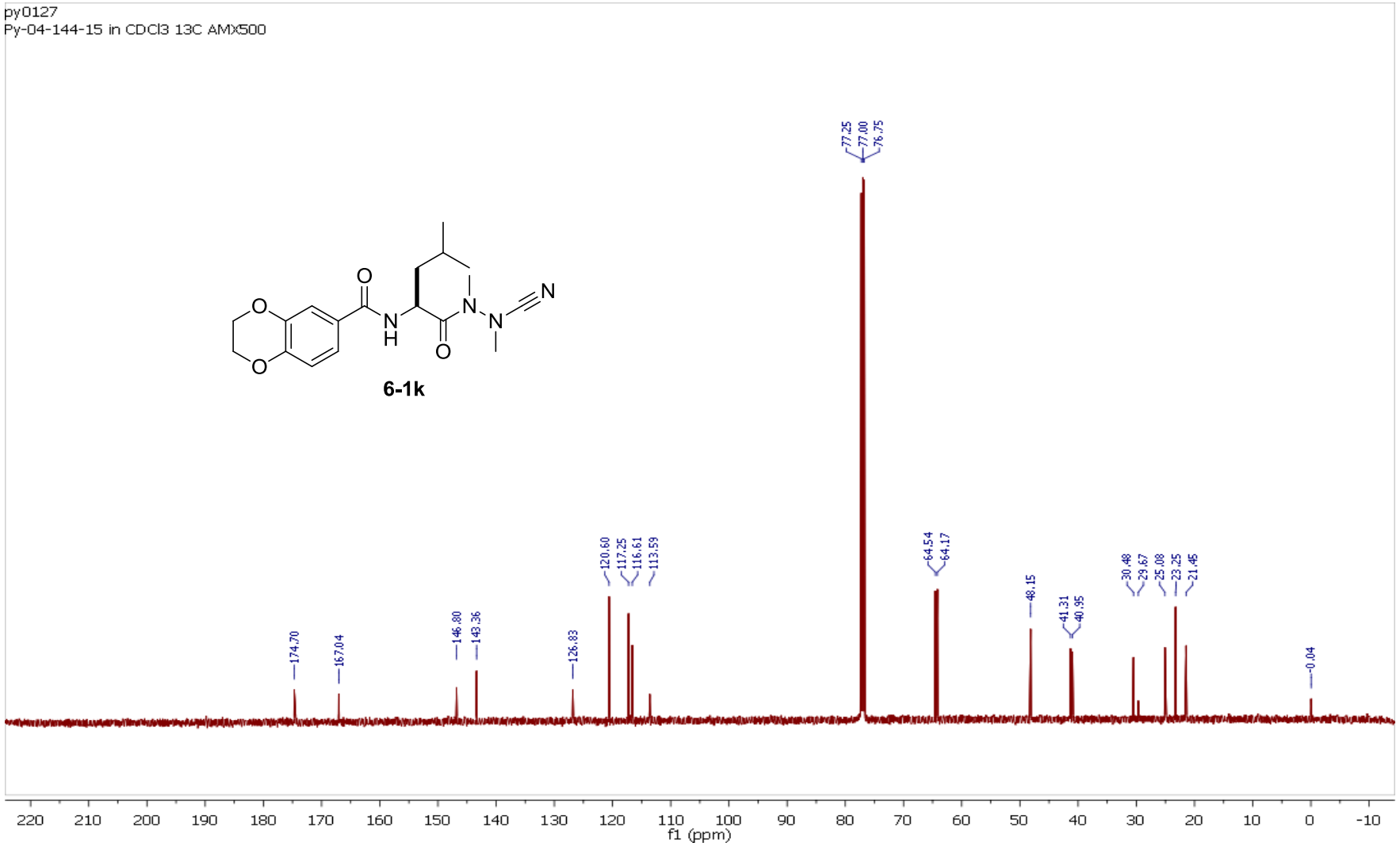

yp0125
PY-04-119-5 in CDCl3 1H AMX500
6-1i
7.34
7.34
7.30
7.29
7.29
7.28
7.27
7.26
7.23
7.21
7.20
7.19
6.90
6.88
6.55
6.53
5.19
5.18
5.17
5.16
4.30
4.29
4.28
4.27
3.51
3.50
3.49
3.47
3.46
3.45
3.21
2.85
2.84
2.82
2.81
2.79
2.78
2.18
2.16
2.15
2.14
2.12
2.11
2.09
1.71
1.08
1.07
1.02
1.01
0.00
1.00
3.03
2.76
0.93
0.72
0.77
4.05
2.01
2.25
1.78
2.55
2.30
2.76
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
f1 (ppm)

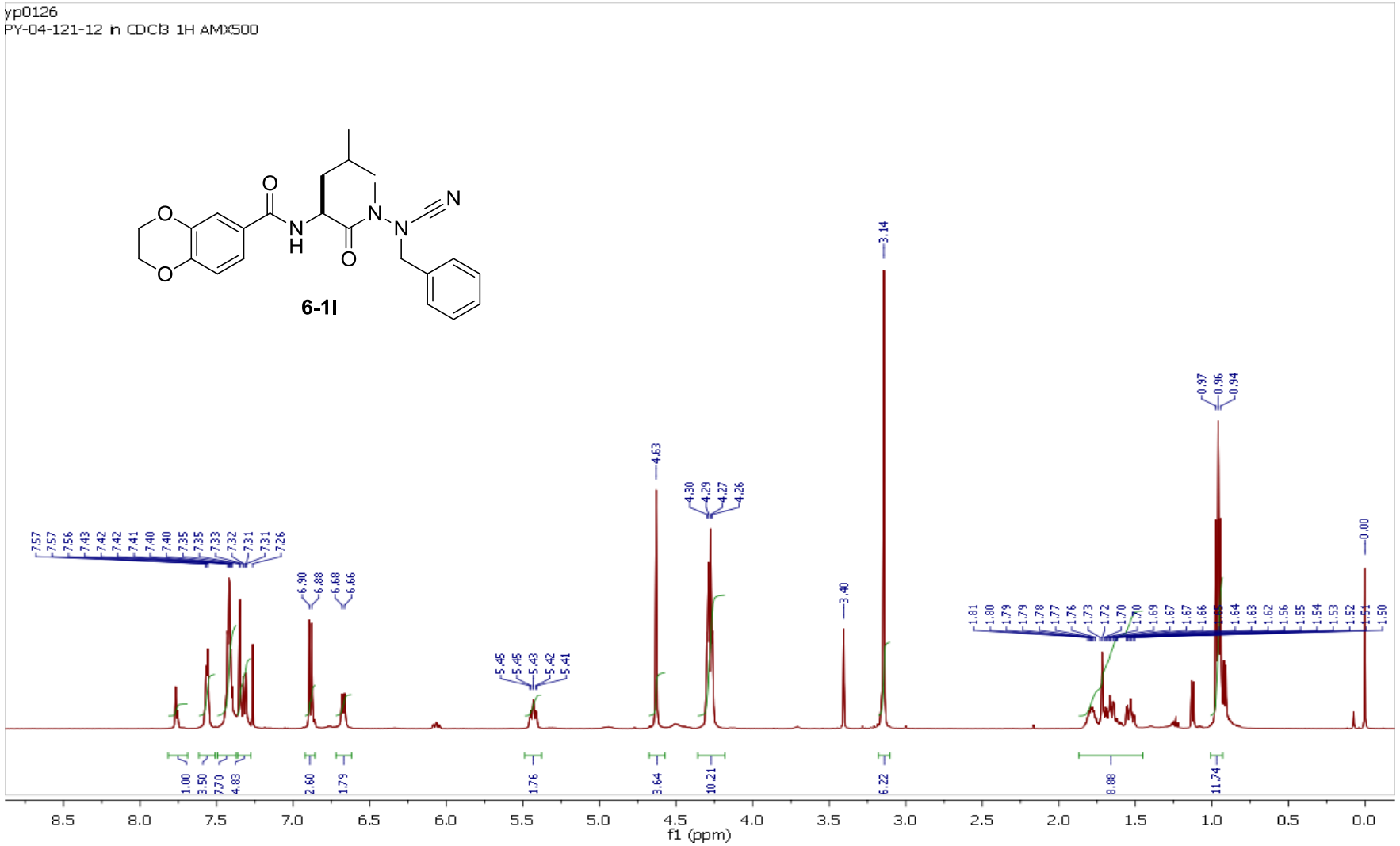

yp0125
13C AMX500
6-1i
173.99
166.76
146.73
143.39
140.10
128.58
128.33
127.15
126.33
120.53
117.27
116.62
112.68
77.25
77.00
76.75
64.53
64.17
53.93
53.57
32.56
31.42
31.38
28.13
19.74
17.75
0.04
f1 (ppm)

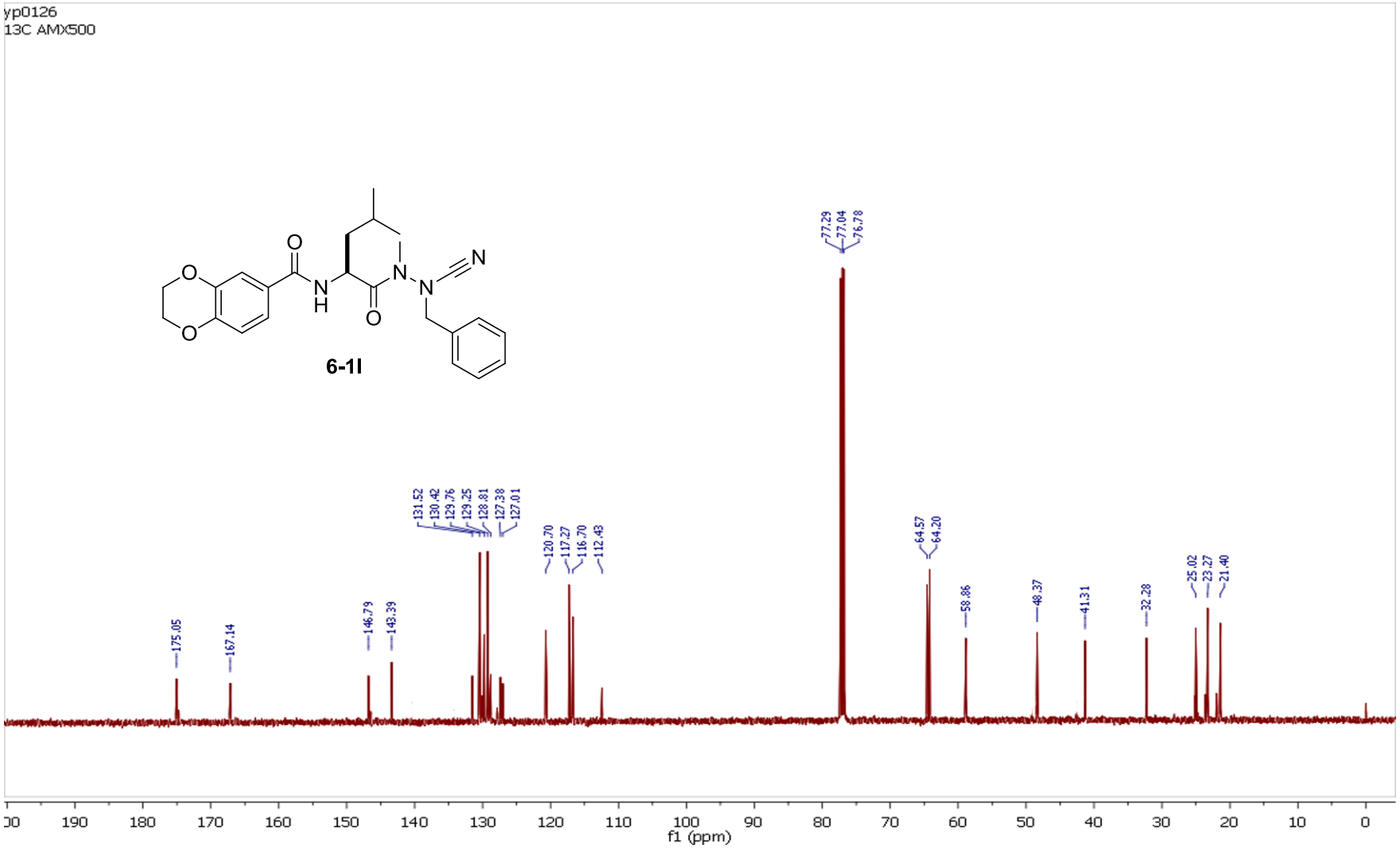

yp0125
PY-04-119-6 in CDCl3 1H AMX500
6-1j
f1 (ppm)

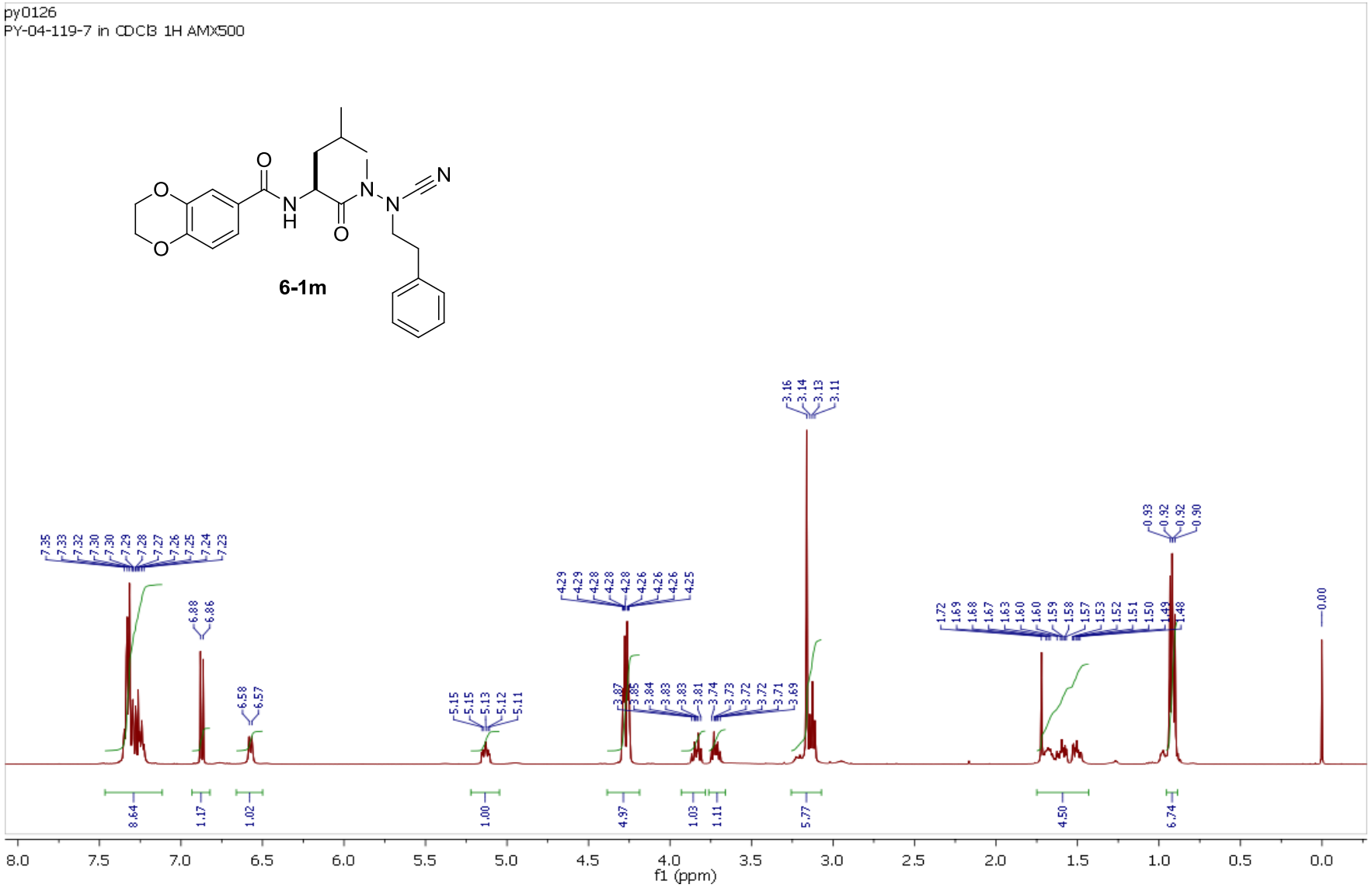

yp0125
PY-04-119-6 in CDCl3 13C AMX500
6-1j
77.25
77.00
76.75
64.53
64.17
53.95
53.83
31.36
31.32
28.58
19.78
19.74
17.62
13.73
-0.04
173.92
166.80
146.71
143.39
127.21
120.52
117.26
116.60
112.72
f1 (ppm)

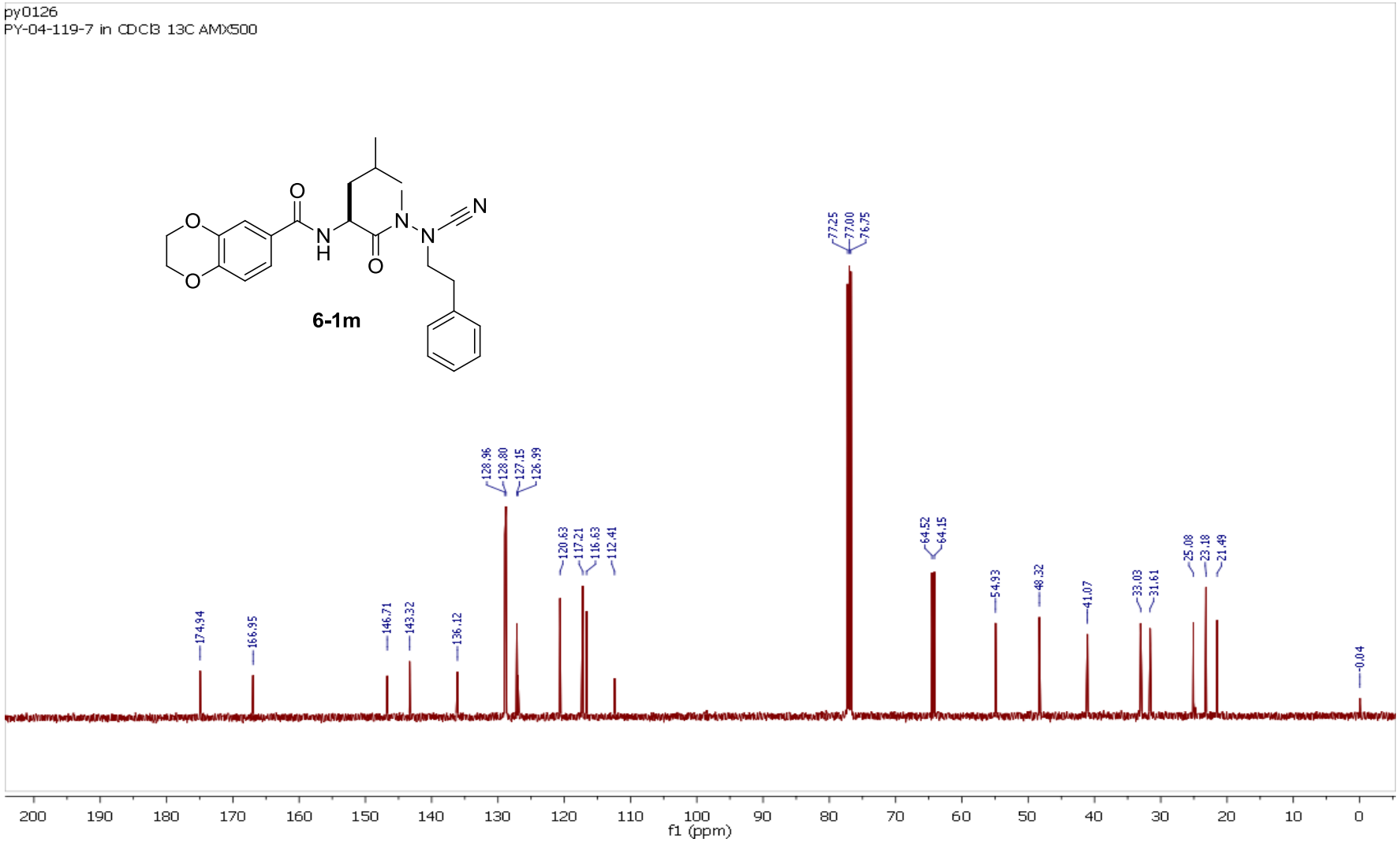

py0127
1H AMX500
6-1k
f1 (ppm)

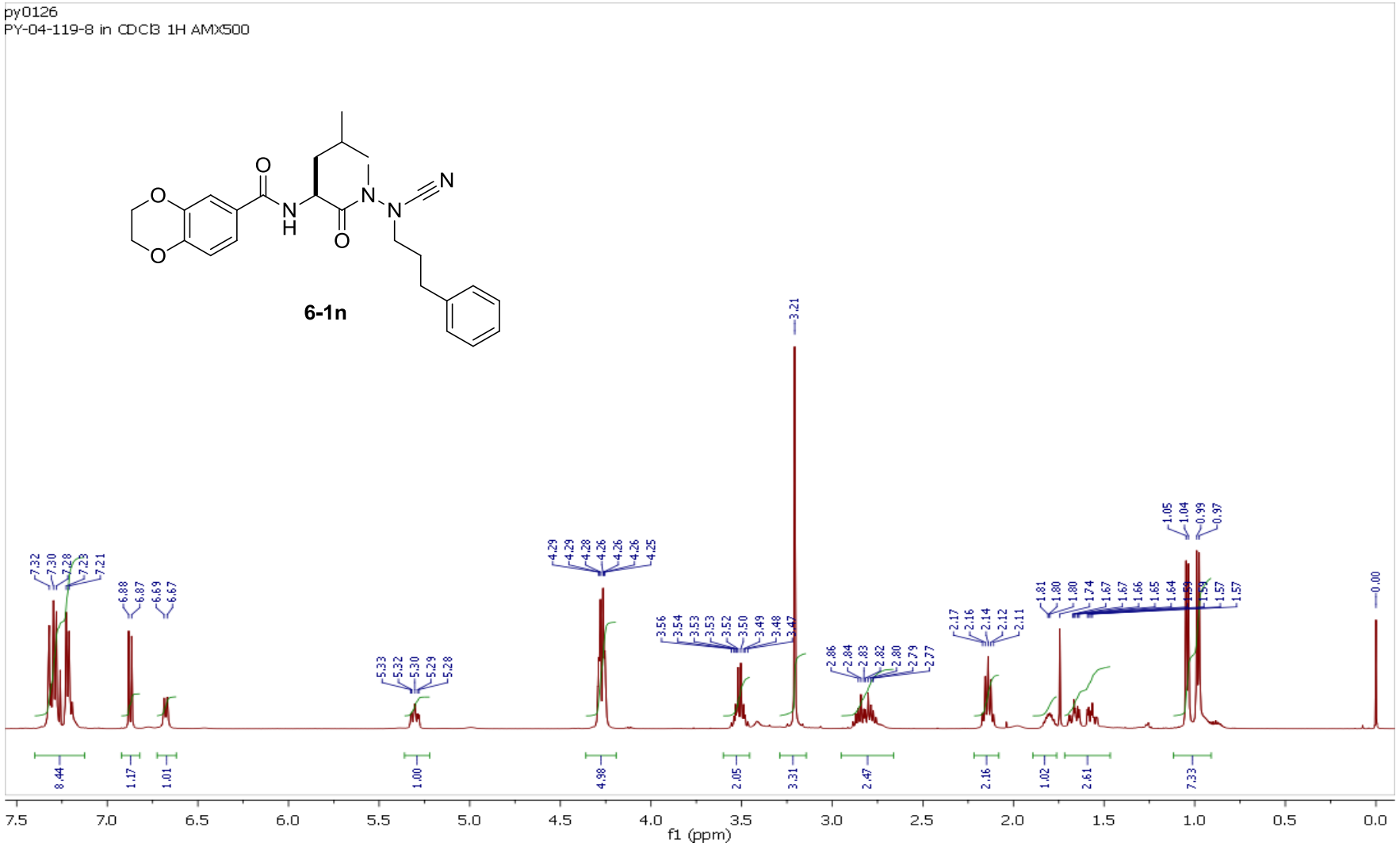

py0127
Py-04-144-15 in CDCl3 13C AMX500
6-1k
77.25
77.00
76.75
174.70
167.04
146.80
143.36
126.83
120.60
117.25
116.61
113.59
64.54
64.17
48.15
41.31
40.95
30.48
29.67
25.08
23.25
21.45
-0.04
f1 (ppm)

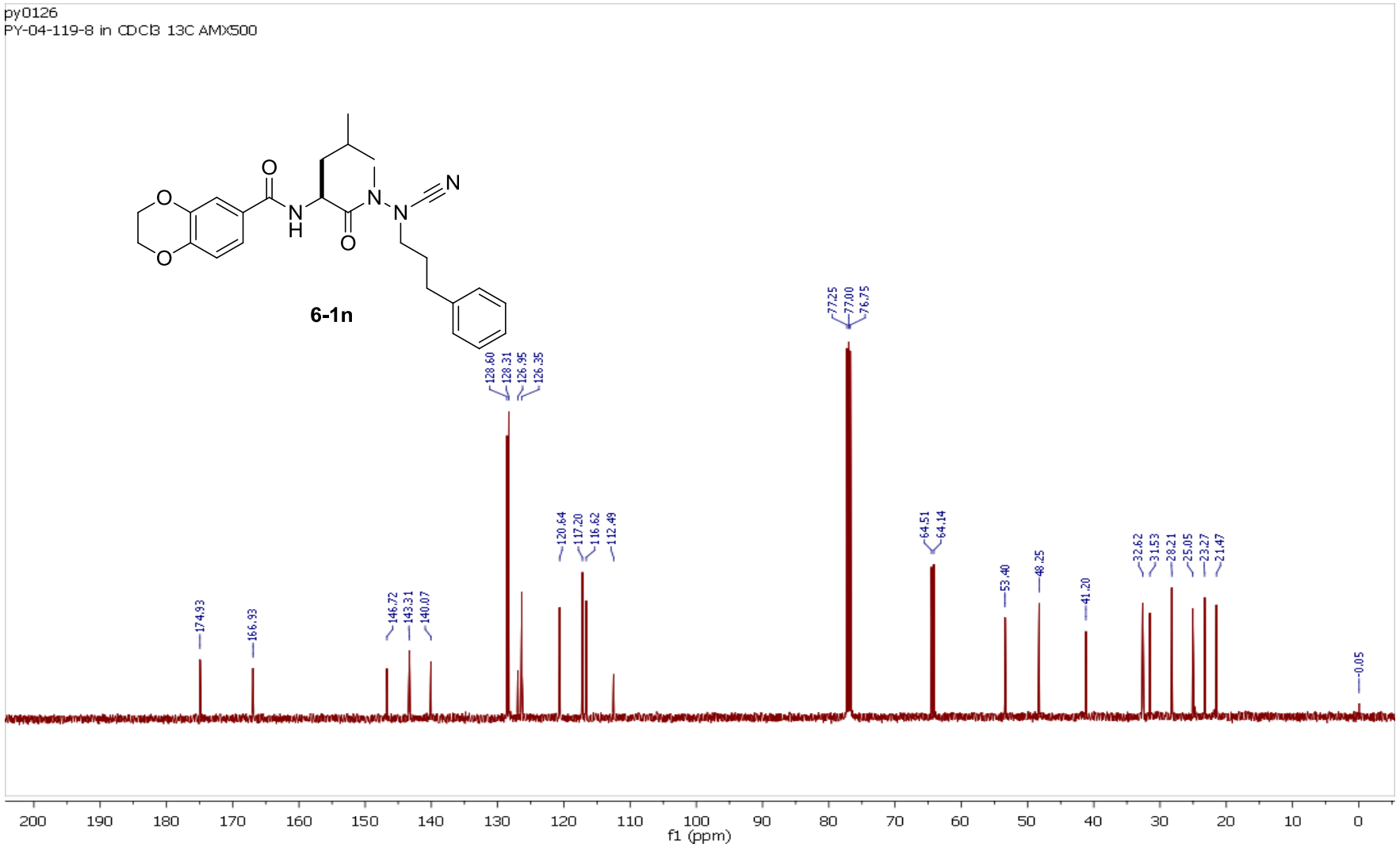

yp0126
PY-04-121-12 in CDCl3 1H AMX500
6-1l
f1 (ppm)

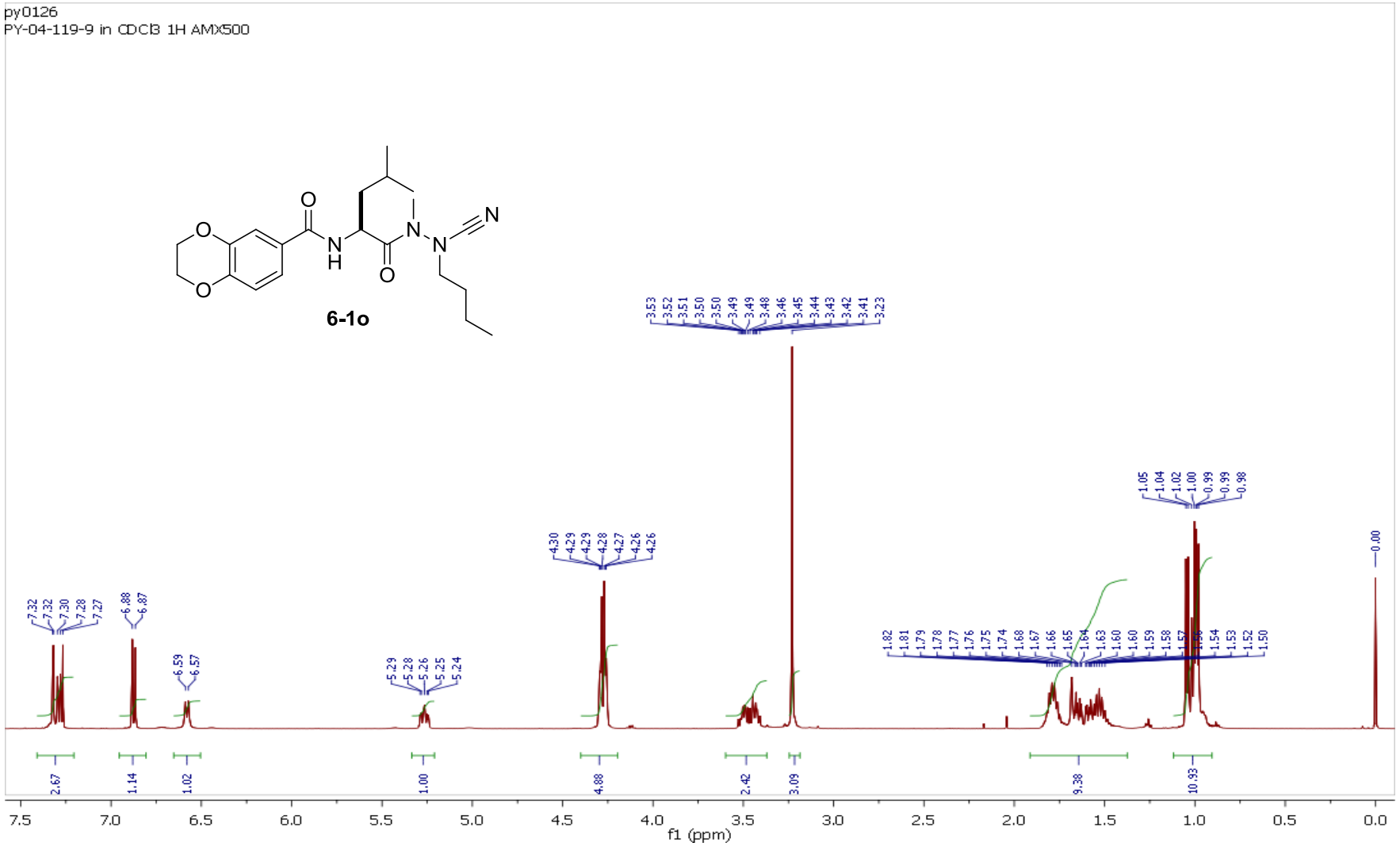

yp0126
13C AMX500
6-1l
77.29
77.04
76.78
131.52
130.42
129.76
129.25
128.81
127.38
127.01
120.70
117.27
116.70
112.43
64.57
64.20
58.86
48.37
41.31
32.28
25.02
23.27
21.40
175.05
167.14
146.79
143.39
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

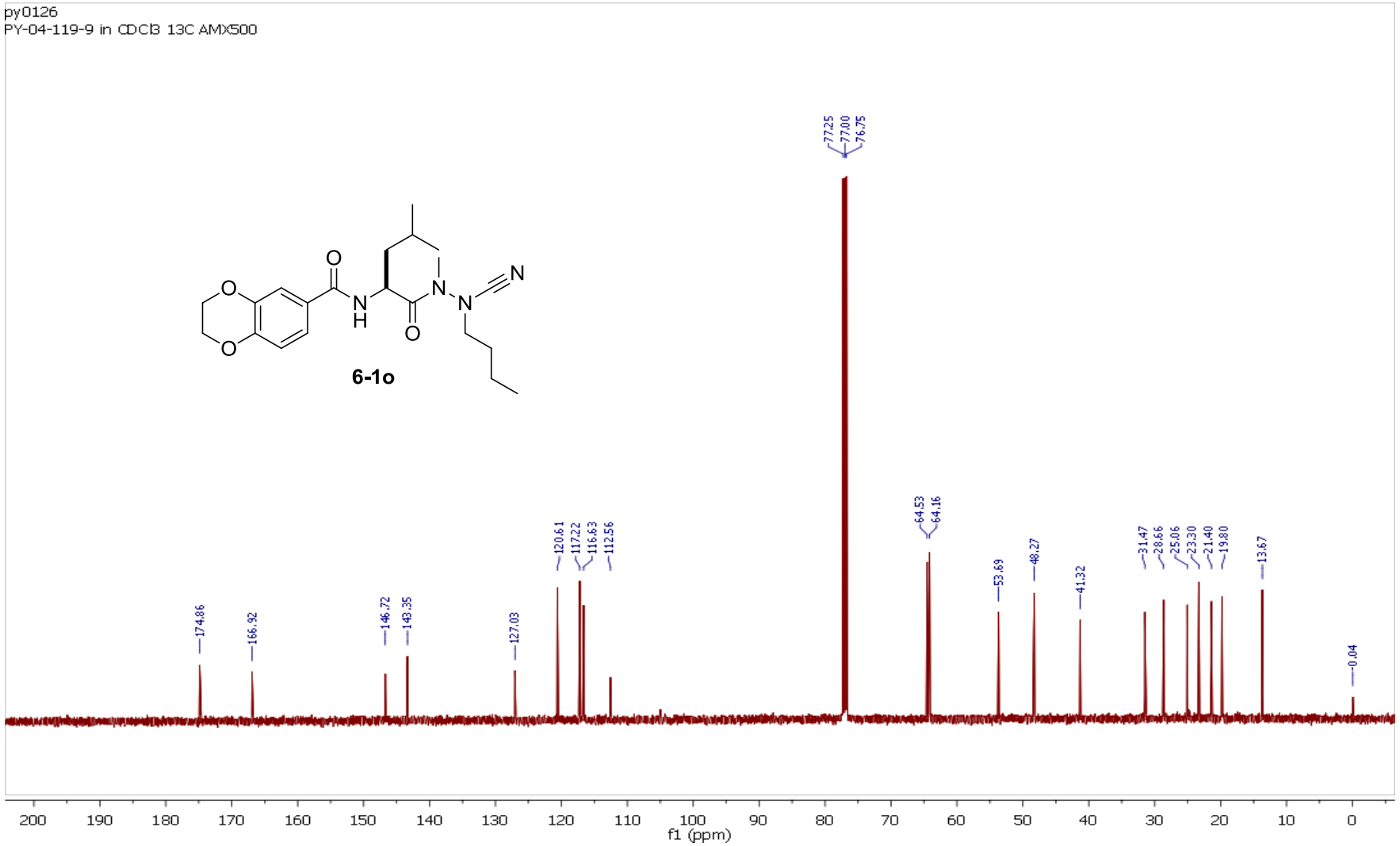

py0126
PY-04-119-7 in CDCl3 1H AMX500
6-1m
f1 (ppm)

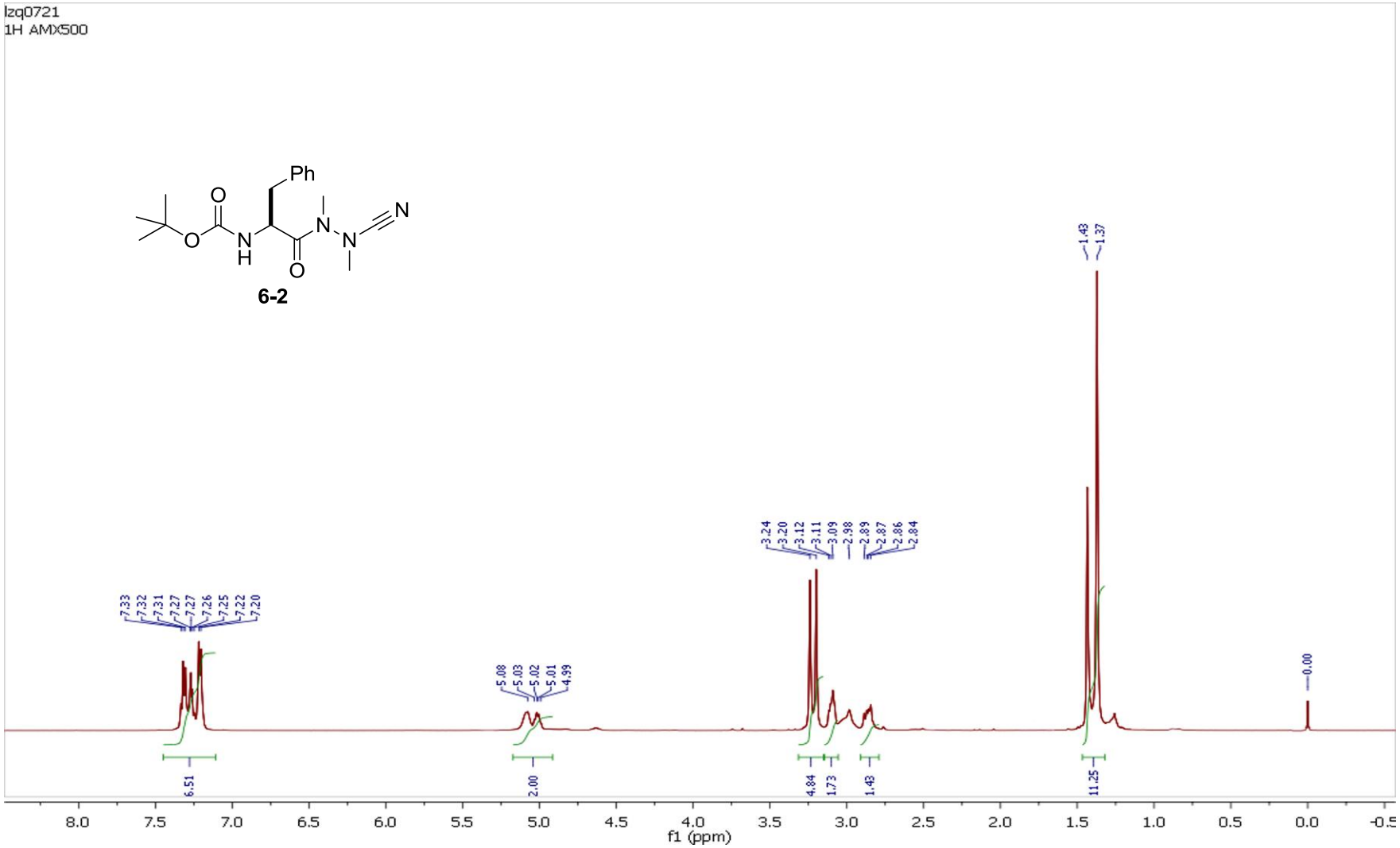

py0126
PY-04-119-7 in CDCl3 13C AMX500
6-1m
77.25
77.00
76.75
128.96
128.80
127.15
126.99
120.63
117.21
116.63
112.41
64.52
64.15
54.93
48.32
41.07
33.03
31.61
25.08
23.18
21.49
174.94
166.95
146.71
143.32
136.12
0.04
f1 (ppm)

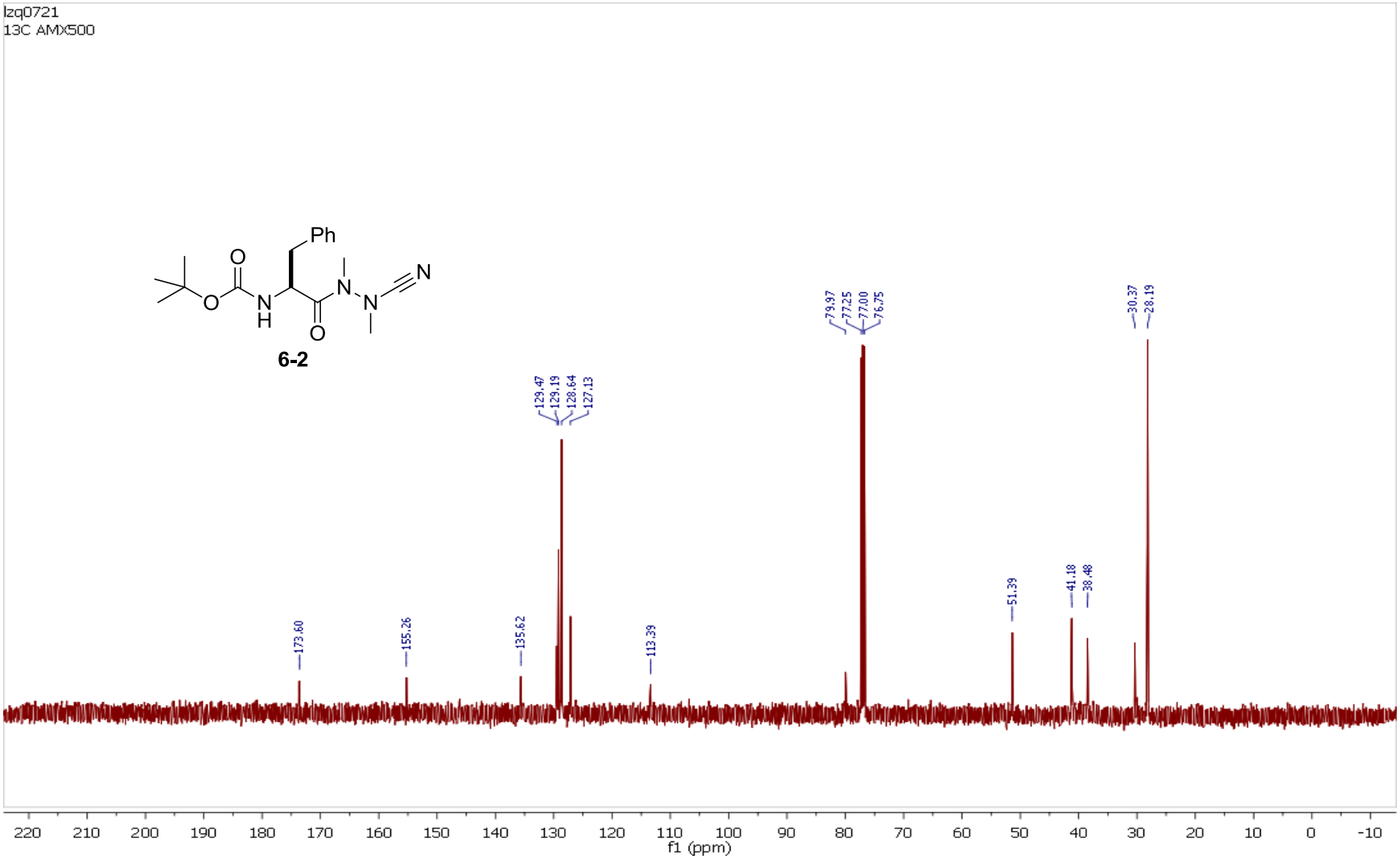

py0126
PY-04-119-8 in CDCl3 1H AMX500
6-1n
f1 (ppm)

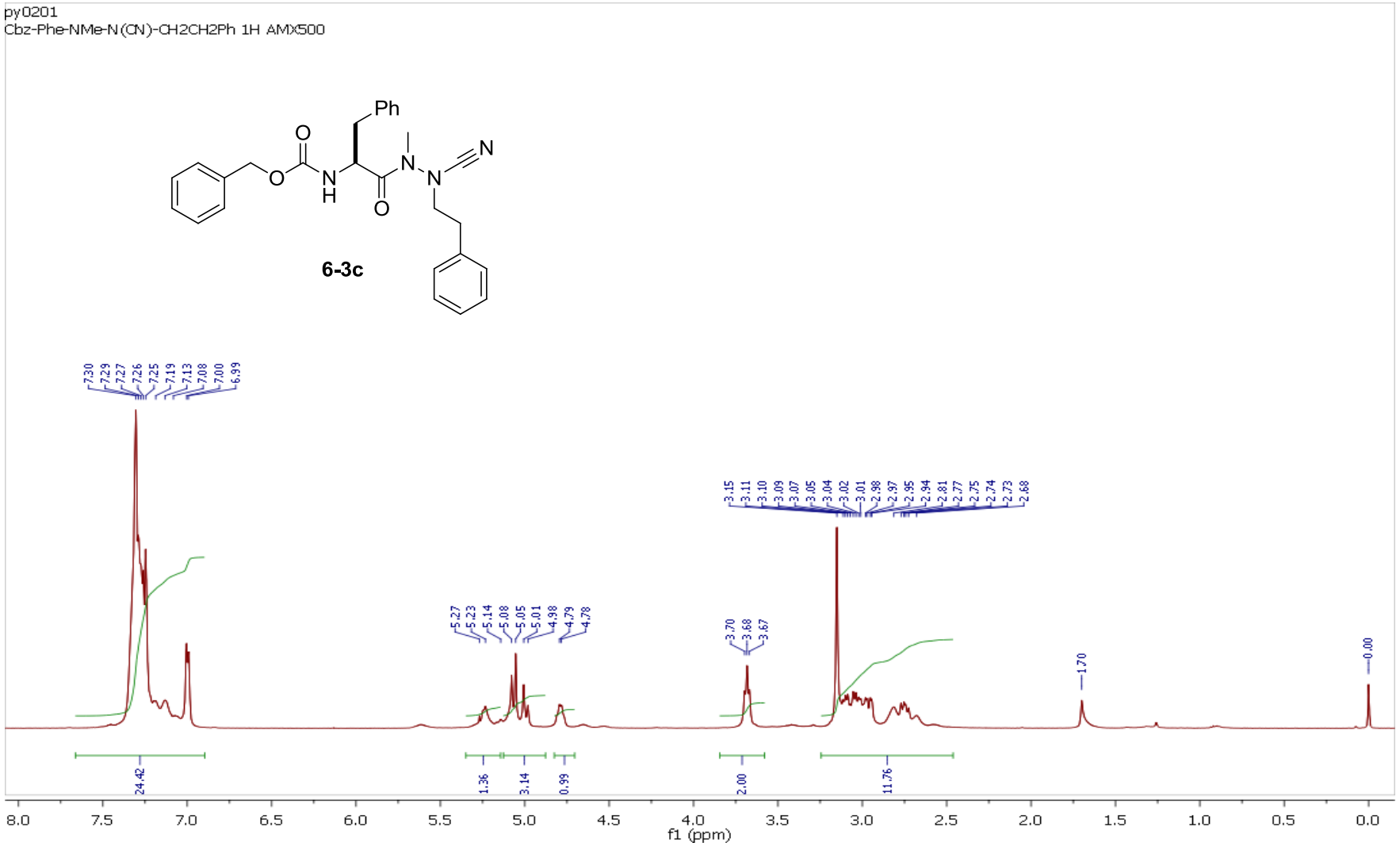

py0126
PY-04-119-8 in CDCl3 13C AMX500
6-1n
174.93
166.93
146.72
143.31
140.07
128.60
128.31
126.95
126.35
120.64
117.20
116.62
112.49
77.25
77.00
76.75
64.51
64.14
53.40
48.25
41.20
32.62
31.53
28.21
25.05
23.27
21.47
-0.05
f1 (ppm)

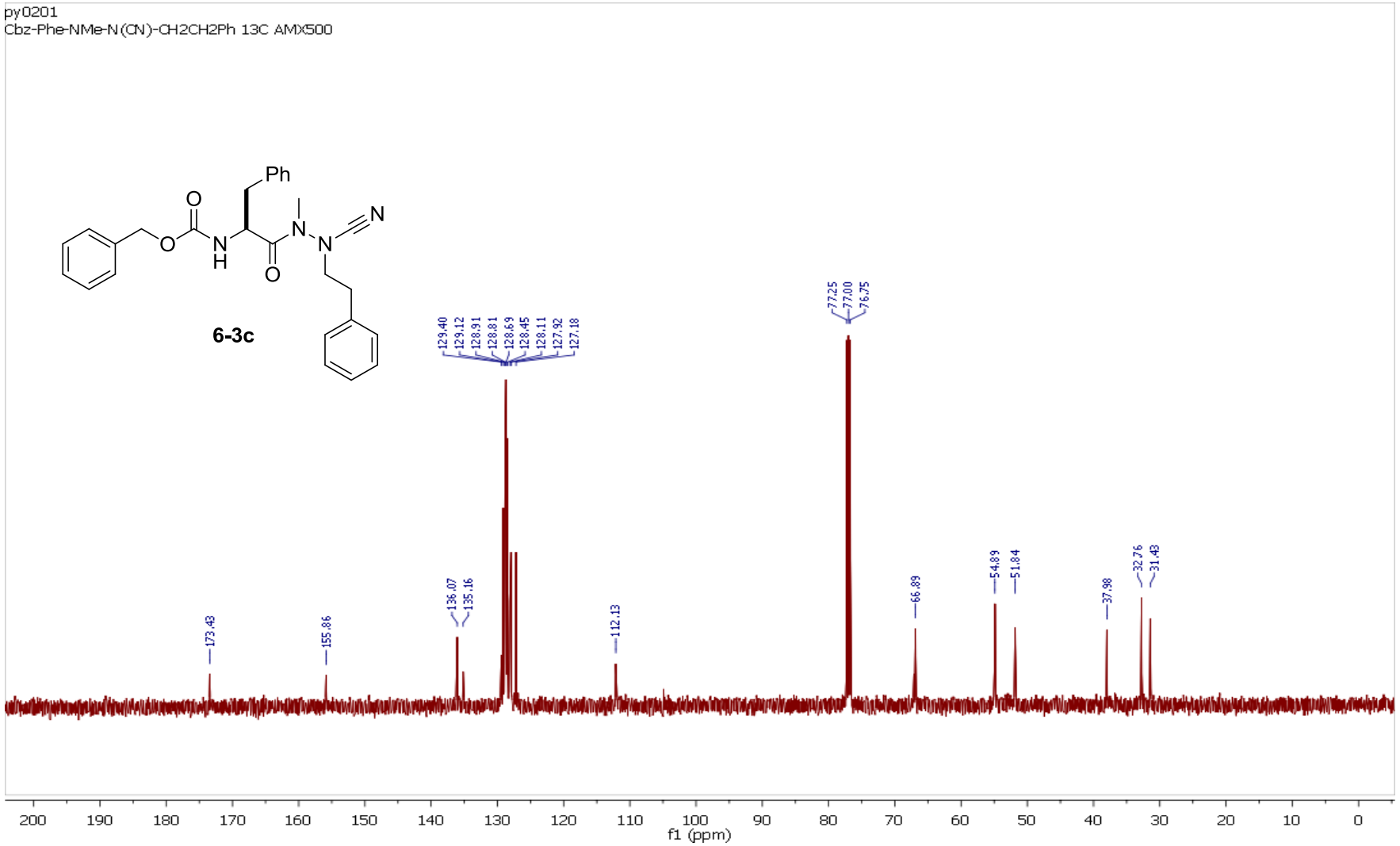

py0126
PY-04-119-9 in CDCl3 1H AMX500
6-1o
f1 (ppm)

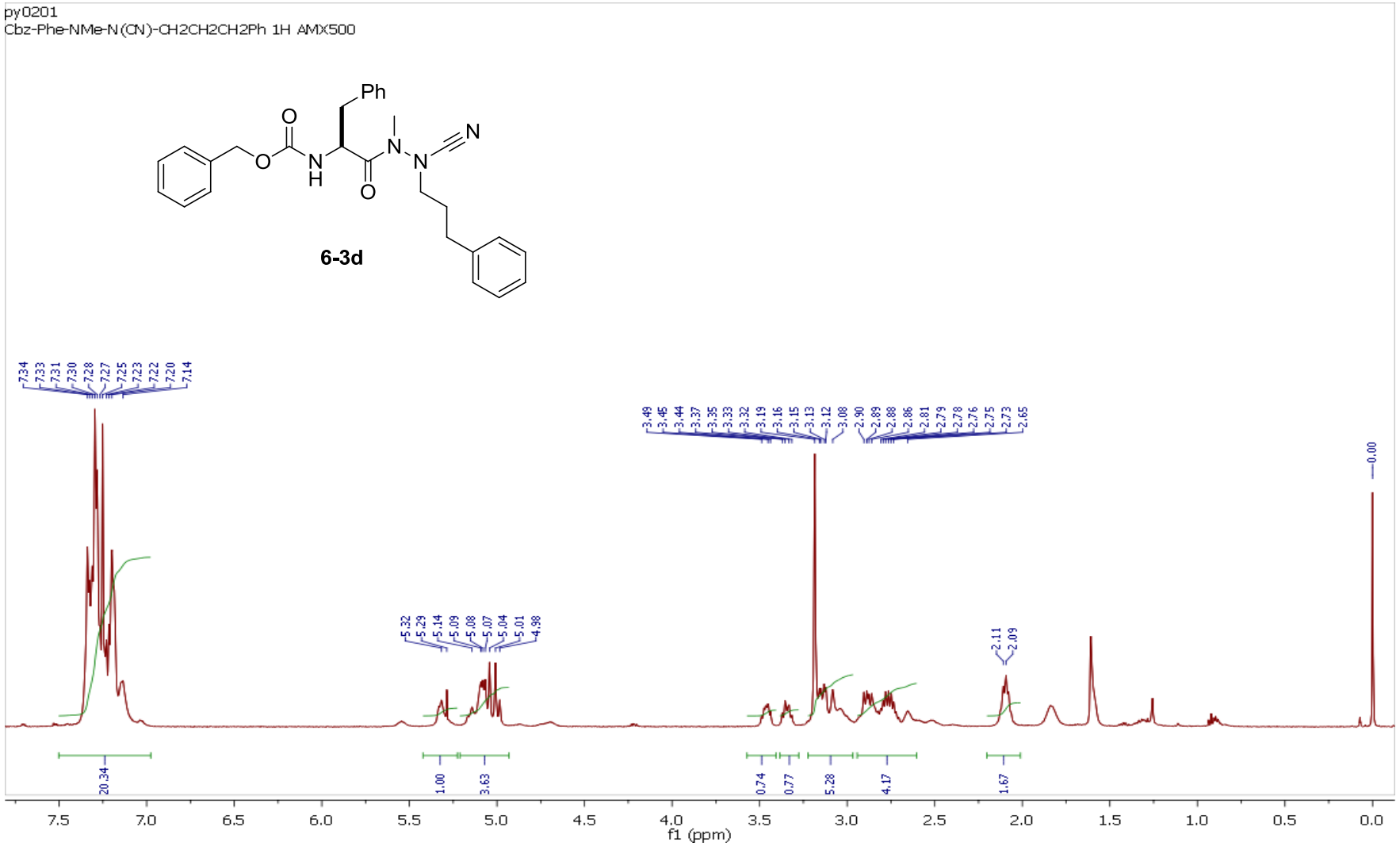

py0126
PY-04-119-9 in CDCl3 13C AMX500
6-1o
77.25
77.00
76.75
174.86
166.92
146.72
143.35
127.03
120.61
117.22
116.63
112.56
64.53
64.16
53.69
48.27
41.32
31.47
28.66
25.06
23.30
21.40
19.80
13.67
0.04
200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

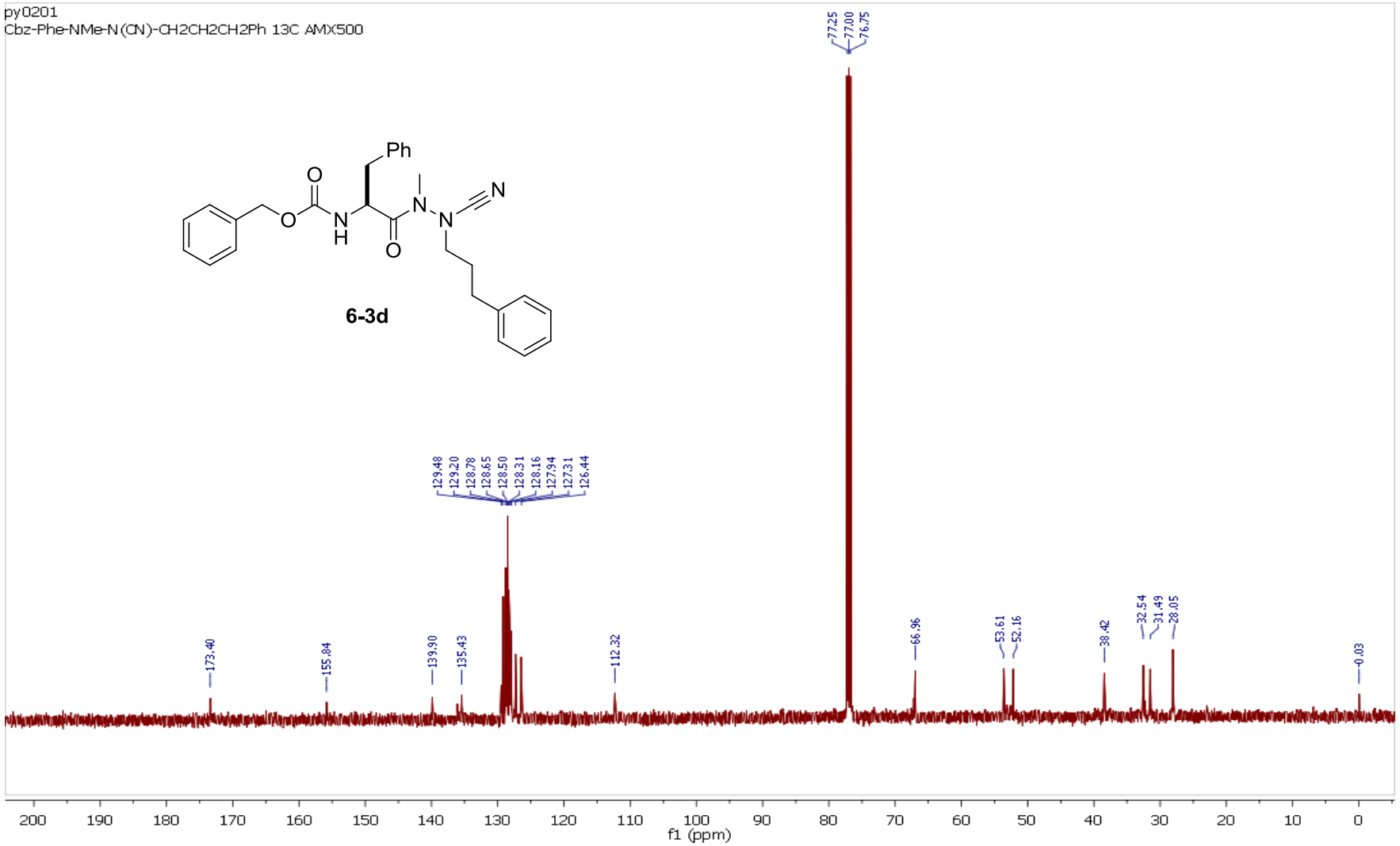

lzq0721
1H AMX500
Ph
O
O
N
H
N
N
N
O
6-2
7.33
7.32
7.31
7.27
7.27
7.26
7.25
7.22
7.20
5.08
5.03
5.02
5.01
4.99
3.24
3.20
3.12
3.11
3.09
2.98
2.89
2.87
2.86
2.84
1.43
1.37
0.00
6.51
2.00
4.84
1.73
1.43
11.25
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
f1 (ppm)

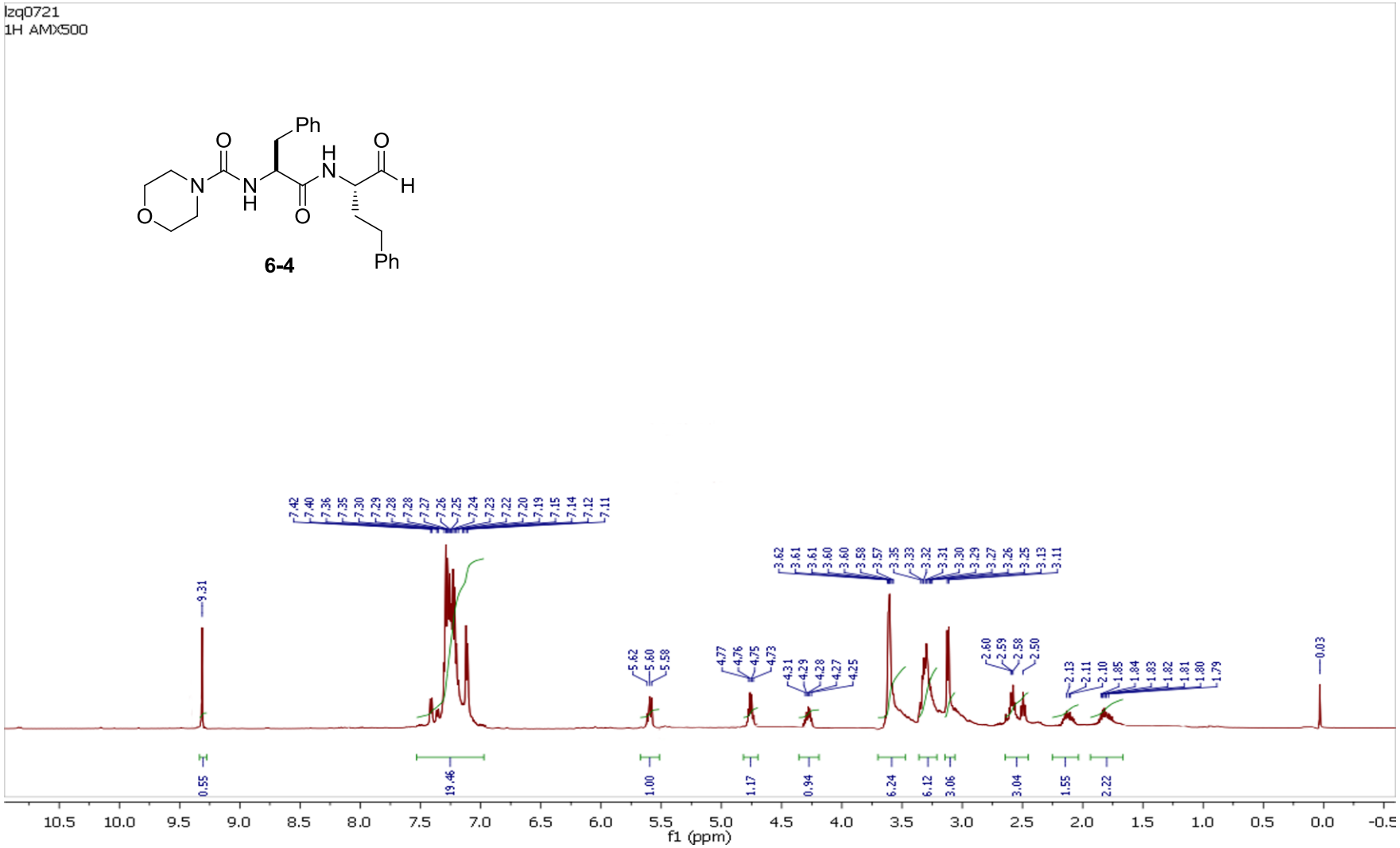

lzq0721
13C AMX500
Ph
O
N
H
O
N
N
N
6-2
173.60
155.26
135.62
129.47
129.19
128.64
127.13
113.39
79.97
77.25
77.00
76.75
51.39
41.18
38.48
30.37
28.19
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

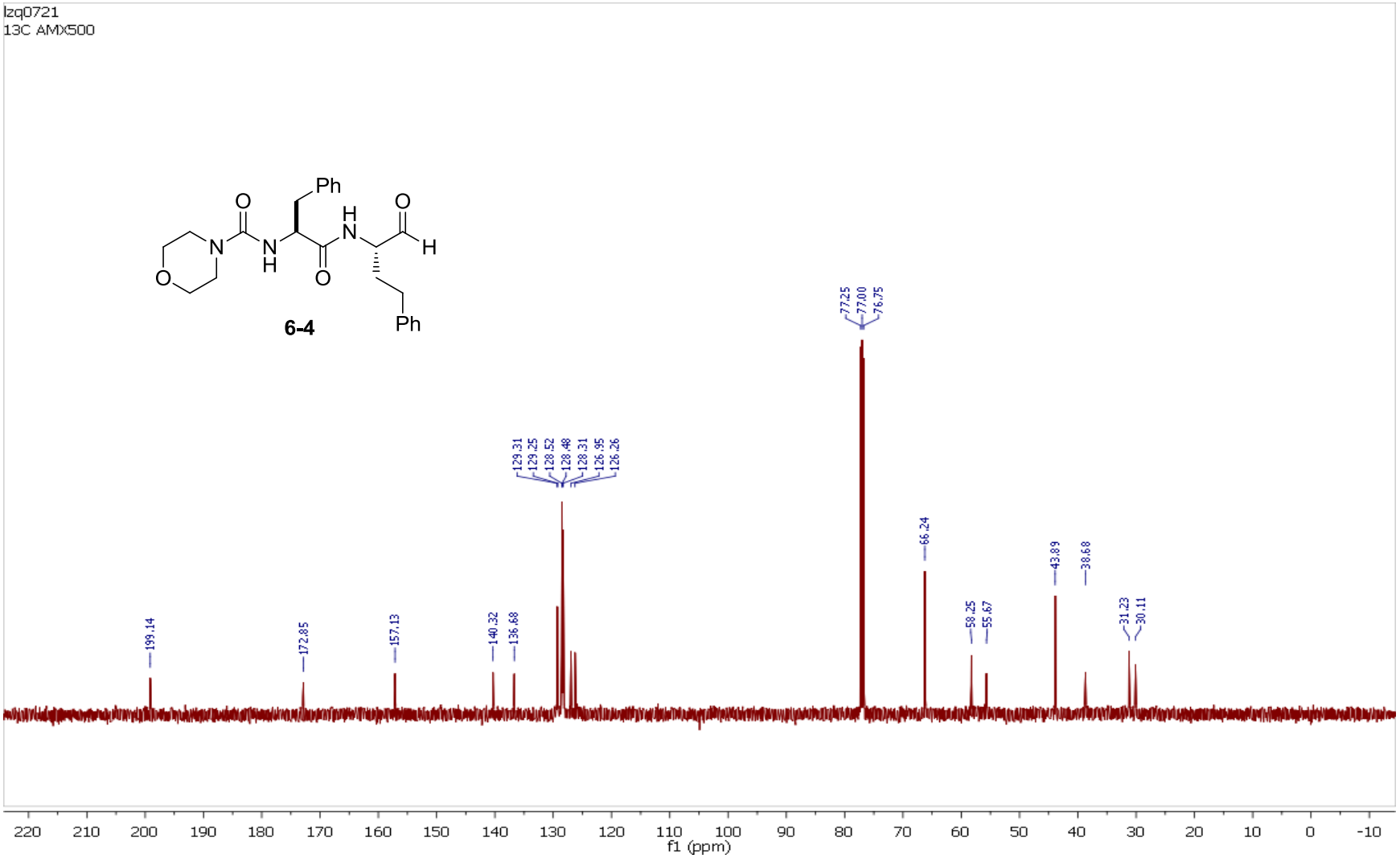

py0201
Cbz-Phe-NMe-N(CN)-CH2CH2Ph 1H AMX500
Ph
O
O
N
H
O
N
N
N
6-3c
7.30
7.29
7.27
7.26
7.25
7.19
7.13
7.08
7.00
6.99
5.27
5.23
5.14
5.08
5.05
5.01
4.98
4.79
4.78
3.70
3.68
3.67
3.15
3.11
3.10
3.09
3.07
3.05
3.04
3.02
3.01
2.98
2.97
2.95
2.94
2.81
2.77
2.75
2.74
2.73
2.68
1.70
0.00
24.42
1.36
3.14
0.99
2.00
11.76
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

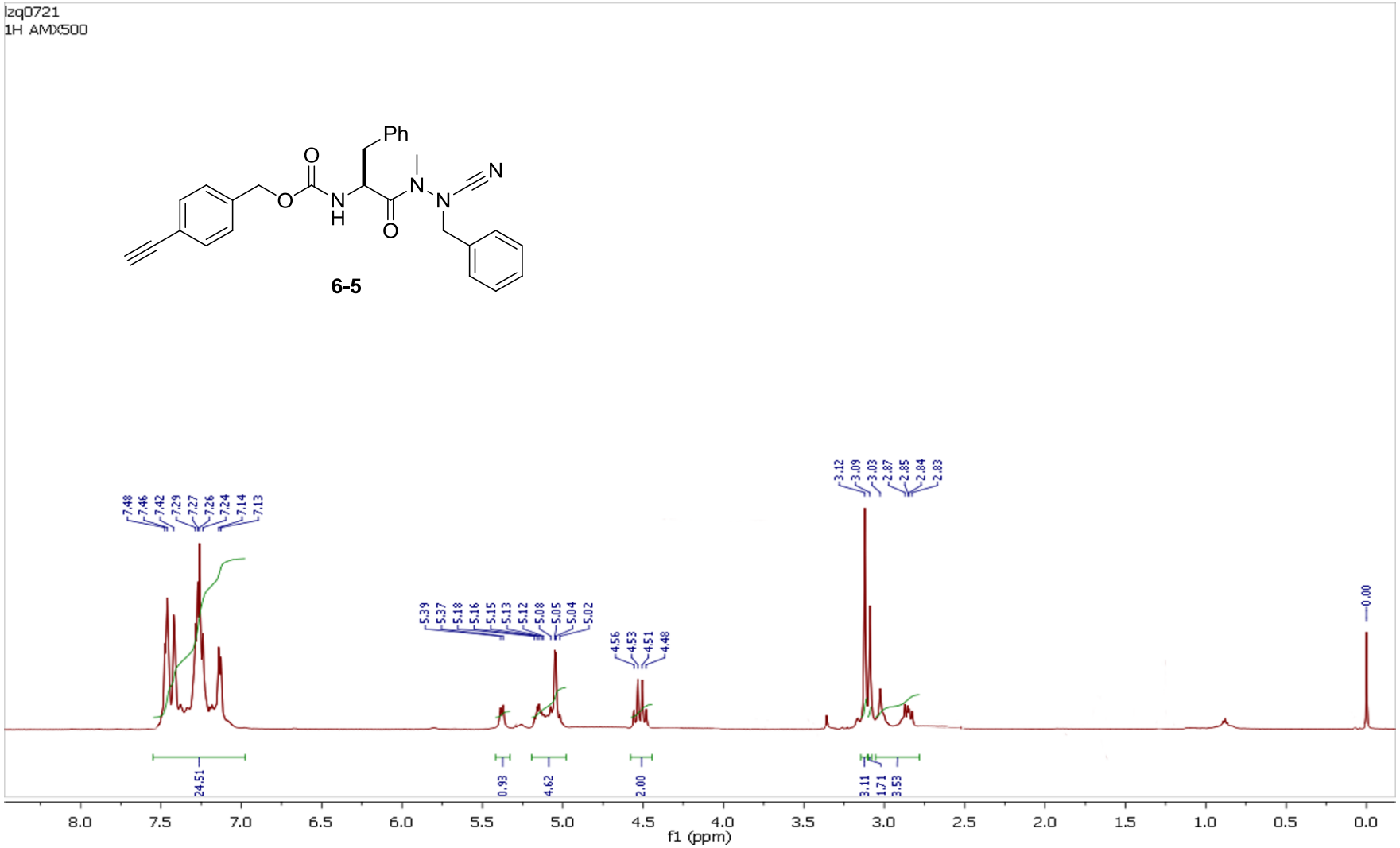

py0201
Cbz-Phe-NMe-N(CN)-CH2CH2Ph 13C AMX500
Ph
O
O
N
H
O
N
N
N
6-3c
129.40
129.12
128.91
128.81
128.69
128.45
128.11
127.92
127.18
77.25
77.00
76.75
173.43
155.86
136.07
135.16
112.13
66.89
54.89
51.84
37.98
32.76
31.43
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

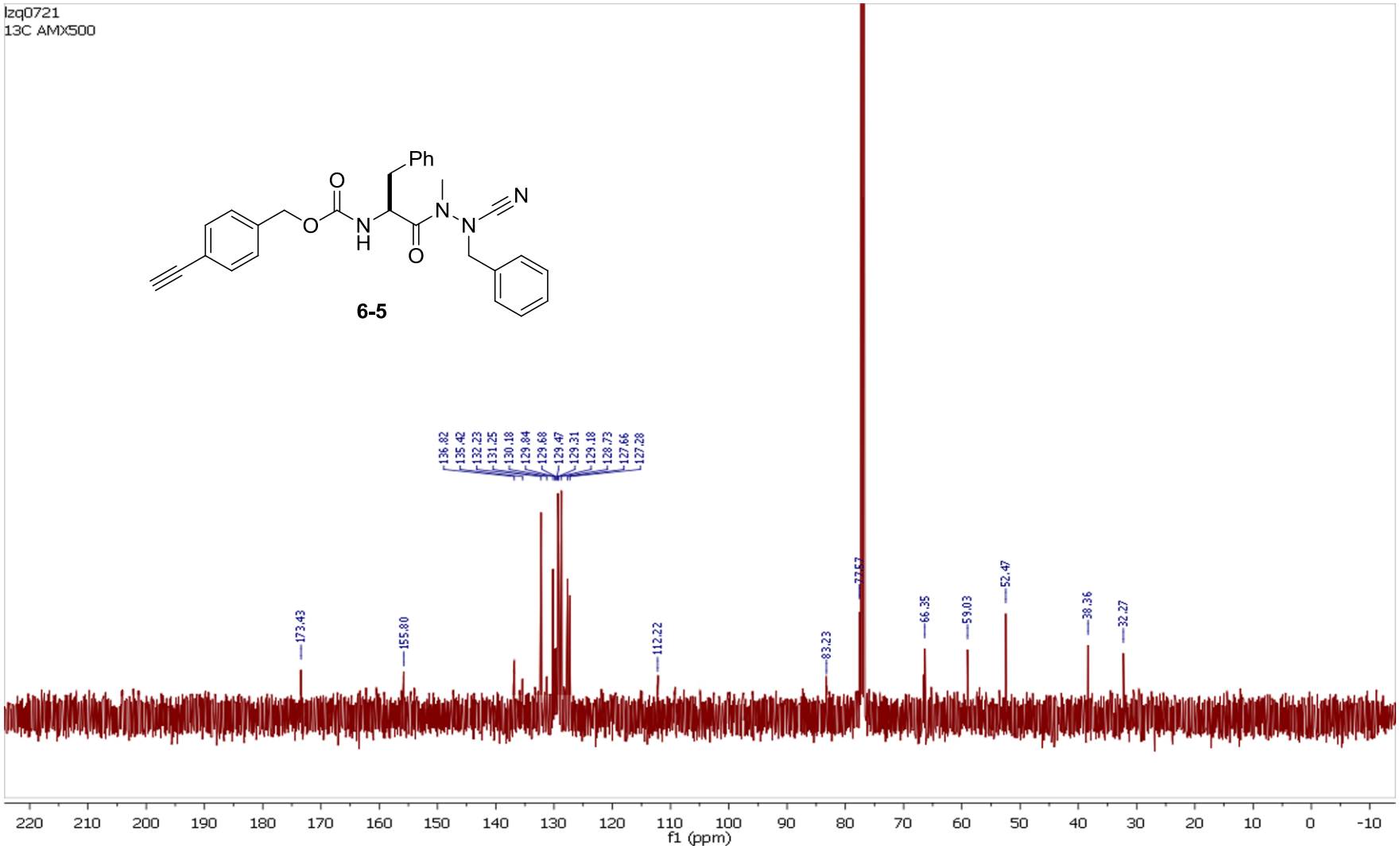

py0201
Cbz-Phe-NMe-N(CN)-CH2CH2CH2Ph 1H AMX500
Ph
O
O
N
H
O
N
N
N
6-3d
7.34
7.33
7.31
7.30
7.28
7.27
7.25
7.23
7.22
7.20
7.14
5.32
5.29
5.14
5.09
5.08
5.07
5.04
5.01
4.98
3.49
3.45
3.44
3.37
3.35
3.33
3.32
3.19
3.16
3.15
3.13
3.12
3.08
2.90
2.89
2.88
2.86
2.81
2.79
2.78
2.76
2.75
2.73
2.65
2.11
2.09
0.00
20.34
1.00
3.63
0.74
0.77
5.28
4.17
1.67
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

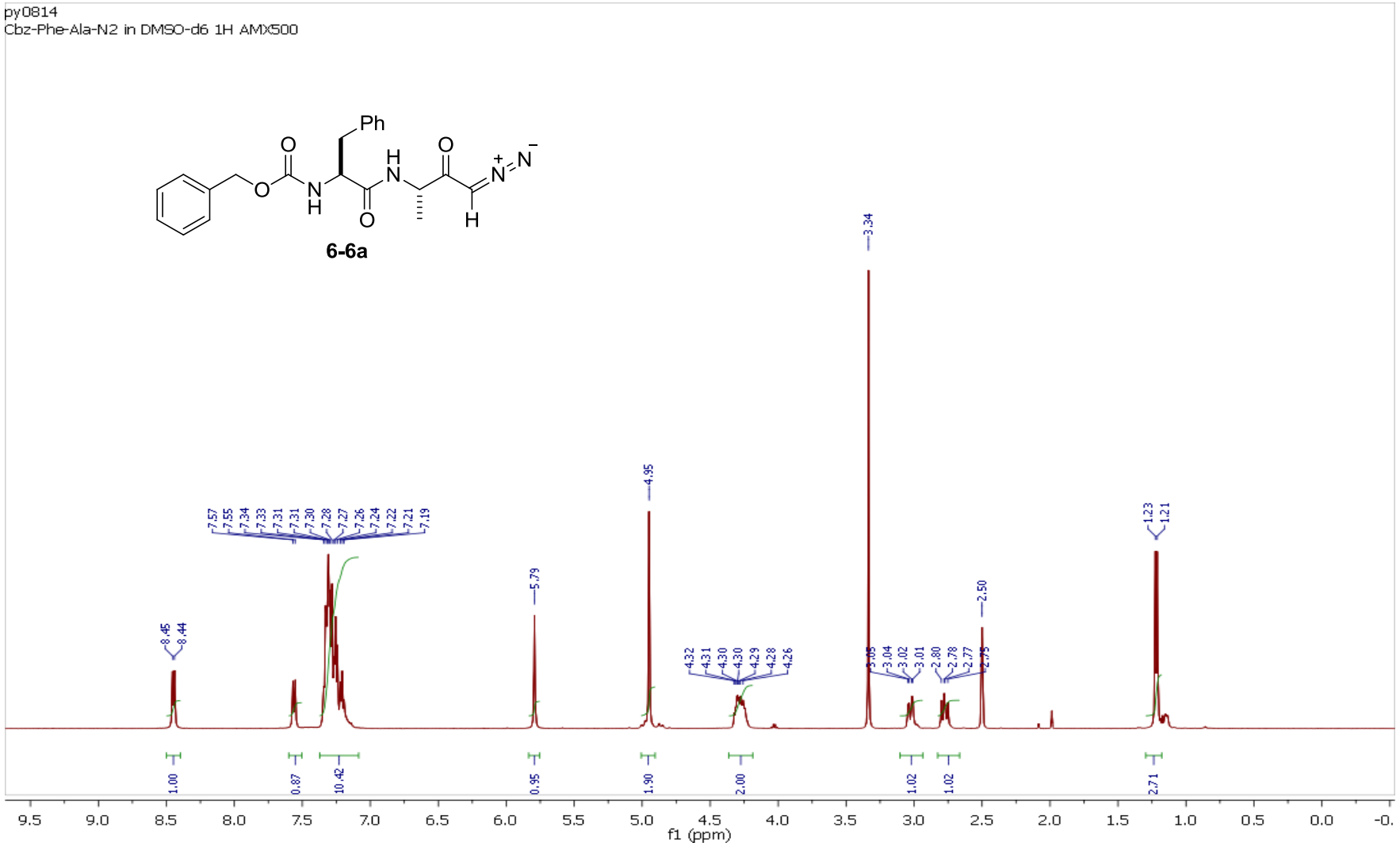

py0201
Cbz-Phe-NMe-N(CN)-CH2CH2CH2Ph 13C AMX500
Ph
O
O
N
H
N
N
N
O
6-3d
77.25
77.00
76.75
129.48
129.20
128.78
128.65
128.50
128.31
128.16
127.94
127.31
126.44
173.40
155.84
139.90
135.43
112.32
66.96
53.61
52.16
38.42
32.54
31.49
28.05
-0.03
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

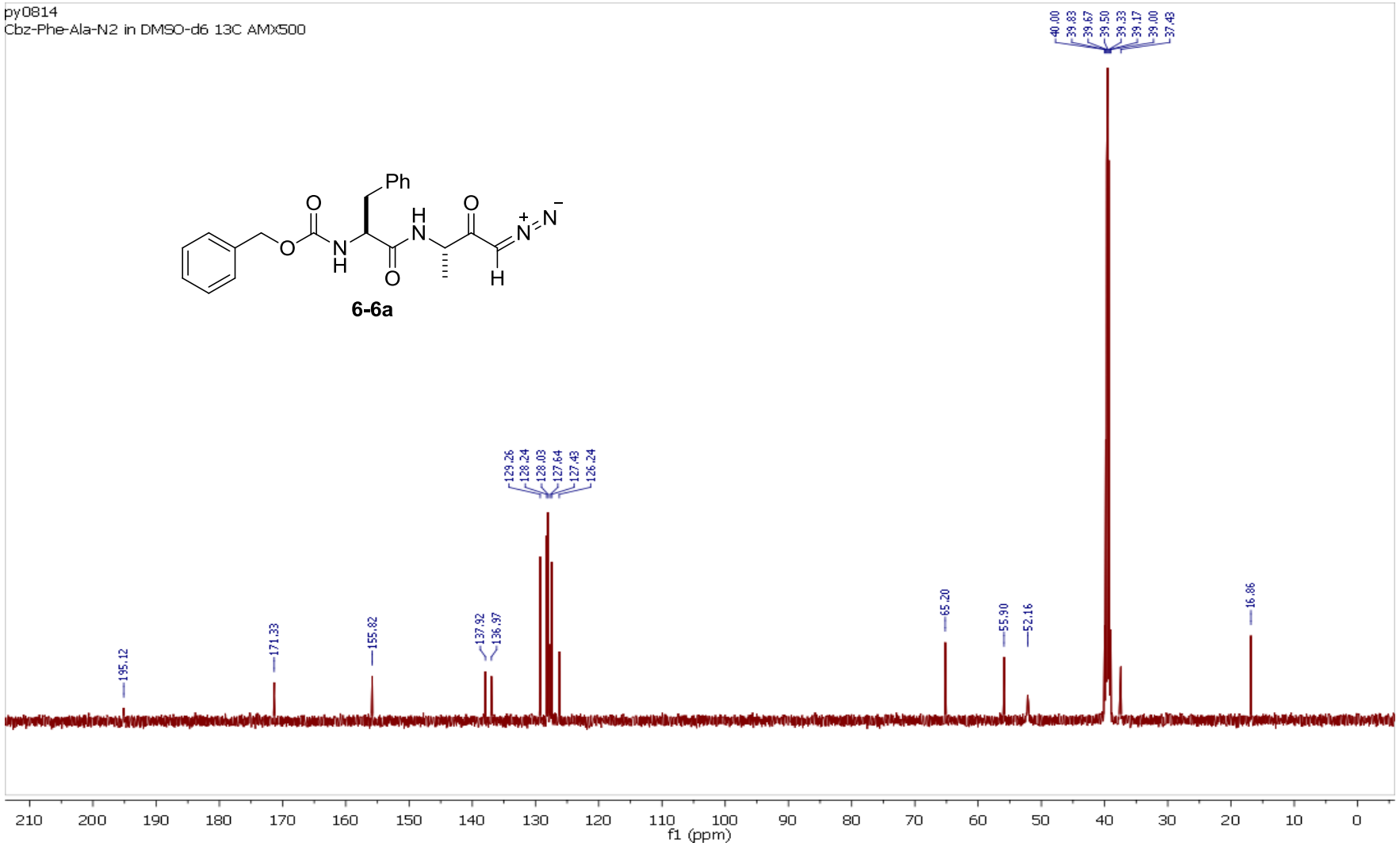

lzq0721
1H AMX500
Ph
O
N
O
H
N
N
H
O
O
H
Ph
6-4
f1 (ppm)

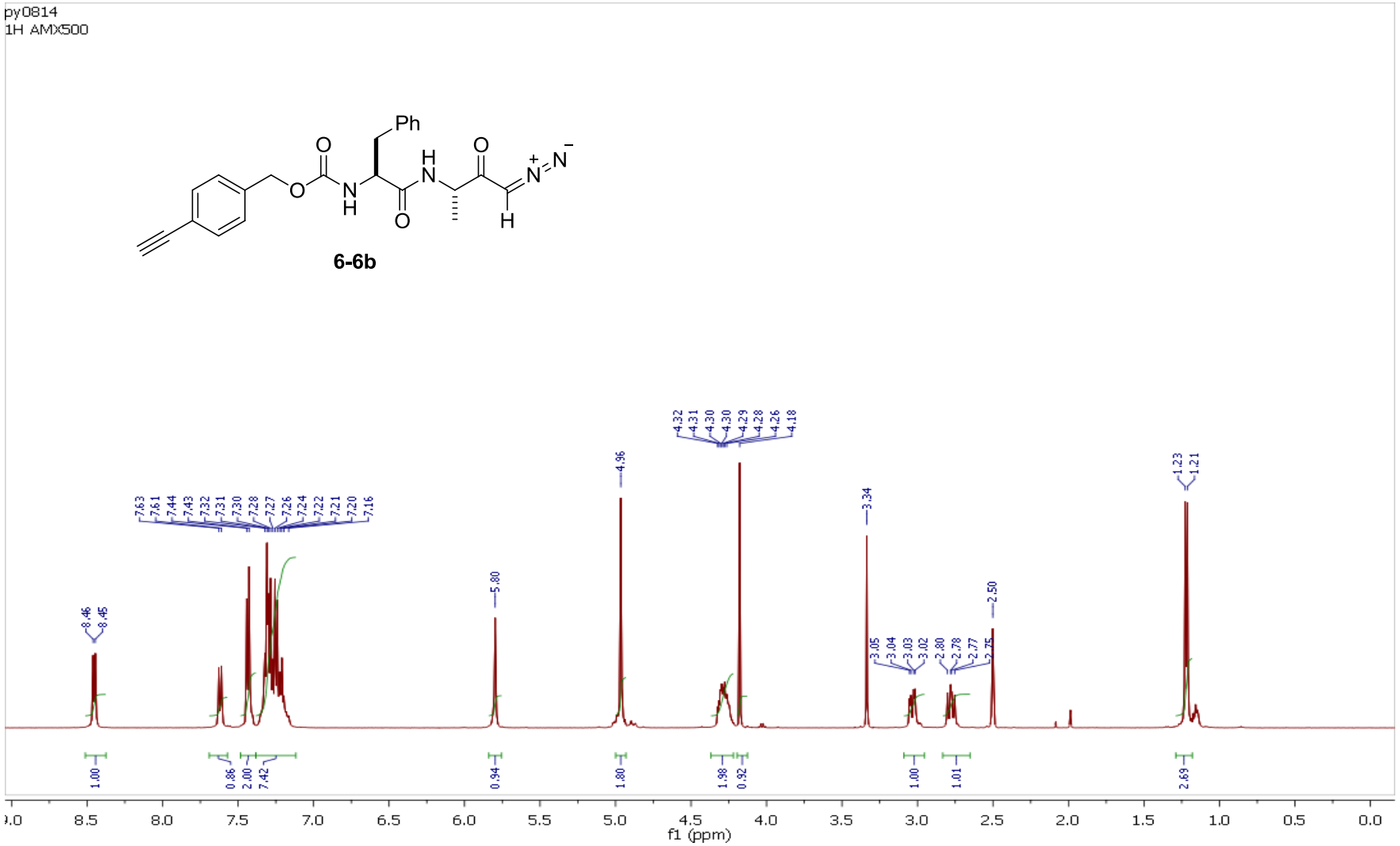

lzq0721
13C AMX500
Ph
O
H
N
O
N
H
N
O
H
O
Ph
6-4
77.25
77.00
76.75
129.31
129.25
128.52
128.48
128.31
126.95
126.26
66.24
43.89
38.68
58.25
55.67
31.23
30.11
199.14
172.85
157.13
140.32
136.68
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

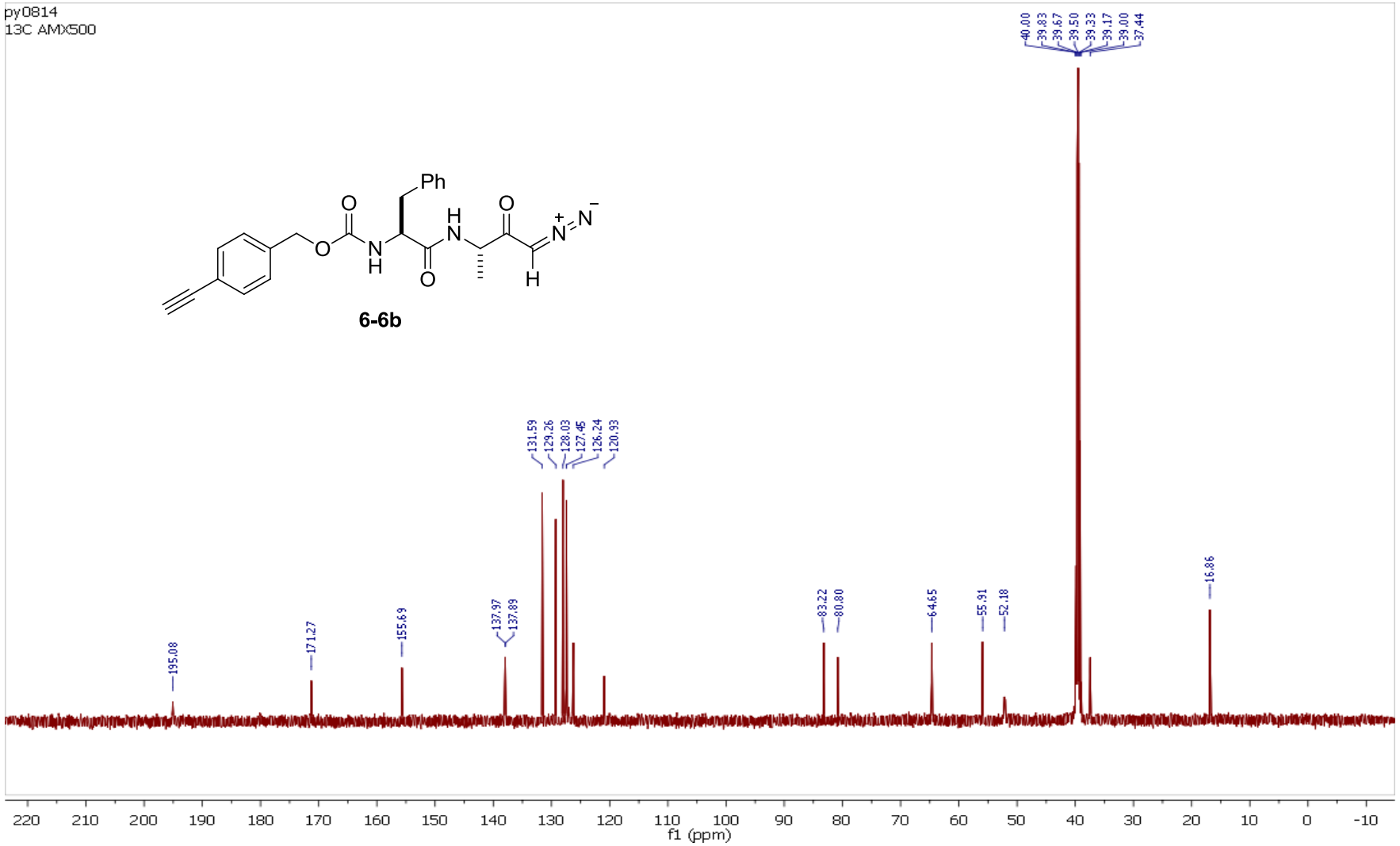

lzq0721
1H AMX500
Ph
O
O
N
H
O
N
N
N
6-5
7.48
7.46
7.42
7.29
7.27
7.26
7.24
7.14
7.13
5.39
5.37
5.18
5.16
5.15
5.13
5.12
5.08
5.05
5.04
5.02
4.56
4.53
4.51
4.48
3.12
3.09
3.03
2.87
2.85
2.84
2.83
0.00
24.51
0.93
4.62
2.00
3.11
1.71
3.53
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)


lzq0721
13C AMX500
Ph
O
O
N
H
O
N
N
N
6-5
136.82
135.42
132.23
131.25
130.18
129.84
129.68
129.47
129.31
129.18
128.73
127.66
127.28
173.43
155.80
112.22
83.23
77.67
66.35
59.03
52.47
38.36
32.27
220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)


py0814
Cbz-Phe-Ala-N2 in DMSO-d6 1H AMX500
Ph
O
O
H
N
N
O
O
H
+
N
N
−
6-6a
3.34
4.95
5.79
2.50
1.23
1.21
8.45
8.44
7.57
7.55
7.34
7.33
7.31
7.31
7.30
7.28
7.27
7.26
7.24
7.22
7.21
7.19
4.32
4.31
4.30
4.30
4.29
4.28
4.26
3.05
3.04
3.02
3.01
2.80
2.78
2.77
2.75
1.00
0.87
10.42
0.95
1.90
2.00
1.02
1.02
2.71
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
-0.
f1 (ppm)


py0814
Cbz-Phe-Ala-N2 in DMSO-d6 13C AMX500
Ph
O
O
H
N
O
N
H
O
+
N
N
−
H
6-6a
40.00
39.83
39.67
39.50
39.33
39.17
39.00
37.43
129.26
128.24
128.03
127.64
127.43
126.24
195.12
171.33
155.82
137.92
136.97
65.20
55.90
52.16
16.86
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)


py0814
1H AMX500
Ph
O
N
H
O
H
N
O
H
N+
N−
6-6b
8.46
8.45
7.63
7.61
7.44
7.43
7.32
7.31
7.30
7.28
7.27
7.26
7.24
7.22
7.21
7.20
7.16
5.80
4.96
4.32
4.31
4.30
4.30
4.29
4.28
4.26
4.18
3.34
3.05
3.04
3.03
3.02
2.80
2.78
2.77
2.75
2.50
1.23
1.21
1.00
0.86
2.00
7.42
0.94
1.80
1.98
0.92
1.00
1.01
2.69
f1 (ppm)


py0814
13C AMX500
Ph
O
O
H
N
O
N
N
O
H
6-6b
40.00
39.83
39.67
39.50
39.33
39.17
39.00
37.44
131.59
129.26
128.03
127.45
126.24
120.93
195.08
171.27
155.69
137.97
137.89
83.22
80.80
64.65
55.91
52.18
16.86
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)