

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

Driving new applications of bioorthogonal 1,2,4-triazines and cyclopropenones

Permalink

<https://escholarship.org/uc/item/8f85m4kq>

Author

Nguyen, Sean Son

Publication Date

2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Driving new applications of bioorthogonal 1,2,4-triazines and cyclopropenones

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Chemistry

by

Sean S. Nguyen

Dissertation Committee:
Professor Jennifer A. Prescher, Chair
Professor Robert C. Spitale
Professor Chang C. Liu

2020

Chapter 1 © Springer Nature Limited
Chapter 2 © Royal Society of Chemistry
Chapter 3 © Royal Society of Chemistry
Chapter 4 © American Chemical Society
All other materials © 2020 Sean S. Nguyen

DEDICATION

To my parents, my siblings, Liz and Beller
for their unconditional love and support

TABLE OF CONTENTS

	Page
LIST OF FIGURES	iv
LIST OF SCHEMES	vii
LIST OF TABLES	viii
ACKNOWLEDGEMENTS	ix
CURRICULUM VITAE	x
ABSTRACT TO THE DISSERTATION	xii
CHAPTER 1: Developing bioorthogonal probes to span a spectrum of reactivity	1
CHAPTER 2: Driving new applications of bioorthogonal 1,2,4-triazines	52
CHAPTER 3: Bioorthogonal cyclopropenones as chemically triggered crosslinkers	106
CHAPTER 4: Butenolide synthesis from functionalized cyclopropenones	153
APPENDIX A: NMR spectra for Chapters 2 and 3	221
APPENDIX B: NMR spectra for Chapter 4	269

LIST OF FIGURES

	Page
Figure 1-1. Translating reactions from round-bottom flasks to living systems.	3
Figure 1-2. Tuning transformations for biological application.	8
Figure 1-3. Cyclopropenone (CpX) derivatives exhibit unique reactivities.	11
Figure 1-4. An expansive set of reagents for dipolar cycloaddition.	15
Figure 1-5. A collection of strained alkenes for IEDDA reactions.	18
Figure 1-6. Steric modifications tune cyclopropene reactivity.	21
Figure 1-7. Triple component protein labeling.	25
Figure 2-1. Overall scheme for the bioorthogonal chemical reporter strategy.	53
Figure 2-2. Triazines are versatile bioorthogonal motifs.	55
Figure 2-3. Autodocking studies with triazine amino acids in the <i>E. coli</i> PheRS active site	57
Figure 2-4. Triazine motifs function in biologically relevant environments.	59
Figure 2-5. Metabolic labeling studies in BL21(DE3) <i>E. coli</i> cells.	61
Figure 2.6. Growth studies in <i>E. coli</i> auxotrophic for phenylalanine.	62
Figure 2-7. Full chemical structures of probes used in metabolic labeling studies.	63
Figure 2-8. Metabolic labeling studies using amino acid 2.3 in <i>E. coli</i> auxotrophs.	64
Figure 2-9. Flow cytometry analysis of amino acid 2.3 in mammalian cells.	65
Figure 2-10. Distortion/interaction analysis of factors controlling mutually orthogonal cycloadditions.	67
Figure 2-11. Orthogonal [4+2] cycloadditions enable dual protein labelling.	68
Figure 2-12. ESI-MS analysis of labeled Nluc-Triazine.	69
Figure 2-13. ESI-MS of orthogonal [4+2] cycloadditions on proteins.	70

Figure 2-14. Orthogonal [4+2] cycloadditions were performed in cell lysate.	71
Figure 2-15. Tricomponent, one-pot bioorthogonal reaction.	72
Figure 2-16. LC-MS traces of starting bioorthogonal reagents and corresponding ligated adducts.	73
Figure 3-1. Cyclopropenones as chemically triggered crosslinkers.	108
Figure 3-2. Chemically triggered CpO trapping with serine.	110
Figure 3-3. Chemically triggered CpO trapping with cysteine.	110
Figure 3-4. Chemically triggered CpO trapping with lysine.	111
Figure 3-5. Chemically triggered CpO trapping with tyrosine.	111
Figure 3-6. Chemically triggered crosslinking of split luciferase probes.	112
Figure 3-7. SmBiT-CpO 3.8 and LgBiT associate to form a functional, light-emitting enzyme.	113
Figure 3-8. LC-MS characterization of SmBiT-CpO peptides.	114
Figure 3-9. Phosphine-triggered crosslinking tracks with affinity.	115
Figure 3-10. SmBiT-CpO 3.10 does not form covalent adducts with HEWL.	116
Figure 3-11. Crosslinking is time-dependent with phosphine 3.2 .	116
Figure 3-12. Improved crosslinking with an optimized trigger.	117
Figure 3-13. Slower crosslinking reactions were observed with less nucleophilic phosphine 3.2 .	119
Figure 3-14. LgBiT residues K124 and K136 are not necessary for crosslinking.	120
Figure 3-15. Site of Nluc crosslink analyzed via tryptic digest.	120
Figure 3-16. Diminished crosslinking is observed with LgBiT mutants Y16F and S28A.	121
Figure 3-17. Dissociation constants for SmBiT-CpO 3.10 and either native (WT) or mutant LgBiTs	122
Figure 3-18. Comparison between CpO and diazirine crosslinking	123

Figure 3-19. Using CpO crosslinkers can be applied to capture RNA-mediated interactions.	125
Figure 4-1. Butenolides in natural products and recent methodologies.	155
Figure 4-2. CpOs react with bioorthogonal phosphines to reveal ketene ylides.	156
Figure 4-3. Mechanistic studies involving butenolide formation.	163

LIST OF SCHEMES

	Page
Scheme 2-1. Mutually orthogonal bioorthogonal cycloadditions.	56
Scheme 2-2. Synthesis of triazine noncanonical amino acids.	60
Scheme S2-1. Synthesis of triazine-carbonate probe 2.10 .	88
Scheme S2-2. Synthesis of triazine-maleimide probe S-2.10 .	95
Scheme 4-1. General synthesis of functionalized alkynes with masked hydroxymethyl tethers.	160
Scheme 4-2. Diverse butenolides were synthesized from substituted CpOs.	161

LIST OF TABLES

	Page
Table 2-1. Grid coordinates used for docking amino acids 2.3 and 2.4 in the PheRS	77
Table 2-2. Autoinduction expression media components.	86
Table 3-1. Phosphine-Mediated Crosslinking of CpO with Amino Acids.	109
Table 3-2. Synthesis of CpO-adenosine analogues via S _N Ar Transformations.	125
Table 3-3. Alternative strategy to access CpO-adenosine analogues.	126
Table 3-4. Forward and reverse primers used to generate mutant LgBiT plasmids.	132
Table 3-5. Primers used to generate gene inserts and vector backbones.	132
Table 4-1. Optimization of butenolide cyclization.	158

ACKNOWLEDGEMENTS

I need to first thank my advisor, Professor Jennifer Prescher. I have learned so much on how to think more efficiently as a scientist, how to plan experiments to get to the most important question, how to be a better storyteller, and how to curb frustration and failure without losing my sanity. For all of this, and so much more, I am indebted to you. Thank you for having faith in me as a scientist, and for being such a supportive and kind mentor.

To past and present members of the Prescher lab, thank you for your wonderful discussions and support. There are a few people I would like to especially thank. Dr. David Row, you taught me everything there was to know about organic synthesis. You were the beacon of scientific rigor, and you taught me how to extract every ounce of information out of an experiment even if it didn't work. You are the scientist I strive to be every day. Science aside, you are also an amazing friend. I miss jamming out to Lamb of God with you, or nerding out with you about The Last of Us. I hope we can hang out again soon. To Dr. Krysten Jones, all of the molecular biology skills I know today were because of your mentorship. Thank you for your guidance. To Kevin Ng, thanks for being a wonderful friend. I will miss our hotpot dinners, our rants about labwork, and most importantly, the breakfast club. Last, I need to thank Zi Yao. I know we joke all the time about hating each other, but I genuinely could not have gotten to the end without you. You are an incredible scientist, and you're never afraid to ask questions or propose ideas, no matter how crazy they sound. And despite that apathetic façade you put up, I know you're a good person. Thank you for everything you've done for me, and I really hope we can stay in touch.

I would like extend my gratitude to two organizations for funding support. Thank you to the Rowland Endowed Chair and Fellowship Fund for awarding me with the inaugural F. Sherwood Rowland Graduate Research Fellowship. I also want to thank Allergan for supporting my research through the Allergan Graduate Research Fellowship.

I would be remiss if I did not thank the directors of the incredible facilities here at UCI. The breadth and depth of your knowledge is world-class, and my projects would not have been possible without your guidance. To Ben Katz and Felix Grun, thank you for your help on all things related to mass spectrometry. I also need to thank Phil Dennison for running a perfect NMR facility.

To my girlfriend, Liz, thank you for incredible support during this journey. You have sacrificed so much to make graduate school easier for me. I appreciate everything you've done, and I'm so lucky to have you by my side. I couldn't have done this without you and the little one. I love you, and I'm excited to see where we go next!

Last, to my parents, thank you for your unconditional love and support. You both fled a war-torn country to a foreign place, who greeted you only with stigmatism for being "boat people". You have given up everything, including your own happiness at times, to ensure that your children would have a brighter future. Nothing I do will ever repay the debt I owe you. For as long as I am alive, I will do everything I can to make sure your sacrifices were not in vain. I love you both so much, and I am forever grateful for everything you have done, and continue to do, for me.

CURRICULUM VITAE

EDUCATION

University of California, Irvine. Irvine, CA 2015 – 2020
Ph.D., Chemistry

University of California, Berkeley. Berkeley, CA 2011 – 2015
B.S. with High Honors in Chemical Biology
Minor in Bioengineering

RESEARCH EXPERIENCE

University of California, Irvine. Irvine, CA 2015 – 2020
Graduate Research
Advisor: Professor Jennifer A. Prescher

University of California, Berkeley. Berkeley, CA 2013 – 2015
Undergraduate Research
Honors Thesis: The Use of Lanthanide Metals in Magnetic Resonance Imaging and Small Molecule Reporter Systems
Advisor: Professor Kenneth N. Raymond

HONORS AND AWARDS

2020 Harold W. Moore Award (UCI)
2019 F. Sherwood Rowland Graduate Research Fellowship (UCI)
2019 Allergan Graduate Research Fellowship
2019 ACS BIOL Division Outstanding Poster Award
2018 ACS BIOL Division Best Poster Award
2018 RSC *Chem. Sci.* Research Presentation Award
2018 ACS BIOL Division Travel Award
2016 NSF GRFP Honorable Mention
2015 Senior Undergraduate Research Award (UCB)

PUBLICATIONS

Scinto, S. L.; Bilodeau, D. A.*; Hincapie, R.*; Lee, W.*; Nguyen, S. S.*; Xu, M.*; amEnde, C. W.; Finn, M. G.; Lang, K.; Lin, Q.; Pezacki, J. P.; Prescher, J. A.; Robillard, M. S.; Fox, J. M. "Bioorthogonal chemistry." *Nat. Rev. Methods Primers*. *In revision*.

Row, R. D.*; Nguyen, S. S.*; Ferreira, A. J.*; Prescher, J. A. "Chemically triggered crosslinking with bioorthogonal cyclopropanones." *Chem. Commun.* **2020**, 56, 10883–10886.

Nguyen, S. S.*; Prescher, J. A. "Developing bioorthogonal probes to span a spectrum of reactivities." *Nat. Rev. Chem.* **2020**, *4*, 476–489.

Nguyen, S. S.*; Ferreira, A. F.*; Long, Z. G.; Row, R. D.; Heiss, T. K.; Dorn, R. S.; Prescher, J. A. "Butenolide synthesis from functionalized cyclopropenones." *Org. Lett.* **2019**, *21*, 8695–8699.

Nguyen, S. S.*; Kamber, D. N.*; Liang, Y.; Liu, F.; Briggs, J.; Shih, H.-W.; Houk, K. N.; Prescher, J. A. "Isomeric 1,2,4-triazines exhibit distinct profiles of reactivity." *Chem. Sci.* **2019**, *10*, 9109–9114.

Pailloux, S. L.; Nguyen, S.; Zhou, S.; Hom, M. E.; Keyser, M. N.; Smiles, D.; Raymond, K. N. "Synthesis and Chemical Reactivity of a 6-Me-3,2-Hydroxypyridinone Dithiazolidine with Primary Amines: A Route to New Hexadentate Chelators for Hard Metal(III) Ions." *J. Heterocyclic Chem.* **2016**, *53*, 1065–1073.

SELECTED PRESENTATIONS

Poster Presentations

Cyclopropenones as chemically triggered crosslinkers 2019
Gordon Research Conference (GRC): Bioorganic Chemistry, Andover NH

Compatible bioorthogonal reactions for multicomponent protein labeling 2018
Genetic Code Expansion Conference 2018, Corvallis, OR

Compatible bioorthogonal reactions for multicomponent protein labeling 2018
Genetic Code Expansion Conference 2018
255th ACS National Meeting and Exposition, New Orleans, LA

Oral Presentations

Cyclopropenones as chemically triggered crosslinkers 2019
Vertex Day, UCI Dept. of Pharm. Sci., Irvine, CA

Compatible bioorthogonal reactions for multicomponent protein labeling 2018
University of California Chemical Symposium 2018, Lake Arrowhead, CA.

ABSTRACT OF THE DISSERTATION

Driving new applications of bioorthogonal 1,2,4-triazines and cyclopropenones

By

Sean S. Nguyen

Doctor of Philosophy in Chemistry

University of California, Irvine, 2020

Professor Jennifer A. Prescher, Chair

Bioorthogonal chemistry enables researchers to study biomolecules in their native environments without perturbing endogenous cellular processes. Over the past two decades, significant strides have been made in developing new, and refining existing, bioorthogonal reactions. The continuously expanding toolbox has opened avenues for tackling increasingly complex biological questions. Despite these advances, limitations remain. The majority of bioorthogonal chemistries follow similar mechanisms, and thus cannot be used simultaneously in multicomponent labeling applications. Bioorthogonal chemistry has also seen little use outside of biomolecule labeling and pull-down applications. To push bioorthogonal chemistry into new directions, I took advantage of two I scaffolds previously developed in the lab: the 1,2,4-triazine and the cyclopropenone.

In Chapter 1, I summarize recent advances made in the field of bioorthogonal chemistry. I focus on how physical organic principles can guide the iterative refinement of existing bioorthogonal transformations. Along the way, I also highlight advances

made in developing collections of bioorthogonal chemistries that can be used in multicomponent labeling applications. In Chapter 2, I describe the use of bioorthogonal 1,2,4-triazines as surrogates for aromatic amino acid residues. These scaffolds are isosteric to phenylalanine, and in principle could be metabolized into the proteome by endogenous cellular machinery. The unique reactivity profile of 1,2,4-triazines also makes them strong candidates for developing mutually orthogonal bioorthogonal reactions. My efforts to develop a triplet set of bioorthogonal reactions is also highlighted. In Chapter 3, I discuss the use of bioorthogonal cyclopropanones as chemically triggered crosslinkers to capture biomolecule interactions. This reaction is robust enough to be conducted in the presence of bacterial cellular lysate. Chapter 4 showcases the use of the cyclopropanone as a synthon for accessing butenolide scaffolds. Butenolides are found in a variety of natural products, but existing routes to access them are not generalizable. The methodology is organocatalytic, and has a broad tolerance for functional groups and substitution patterns.

This thesis describes the use of bioorthogonal chemistry in underexplored applications. Outside of biomolecule labeling, these transformations can have utility in areas such as chemical crosslinking and synthetic methodology. Both the cyclopropanone and the 1,2,4-triazine are prime candidates for pushing the frontiers of bioorthogonal chemistry. These scaffolds exhibit remarkable stability in biological environments, and possess unique manifolds of reactivity that can be exploited in multiple settings outside of traditional biomolecule tagging.

Chapter 1: Developing Bioorthogonal Probes to Span a Spectrum of Reactivities

Adapted from published work, with permission from Springer Nature: Nguyen, S. S.; Prescher, J. A. *Nat. Rev. Chem.* **2020**, *4*, 476–489.

1.1 Introduction

A comprehensive view of living systems requires tools and methods to probe biomolecules in their native habitats. Fluorescent proteins and other genetic tags have long been used in this capacity [1]. Although powerful, such tools are not amenable to direct monitoring of non-proteinaceous targets, including small-molecule metabolites. The need for more generalizable platforms has spurred the development of bioorthogonal chemistries — highly selective reactions that can be used to covalently tag targets in live cells and, in some cases, living organisms. For decades, bioorthogonal reactions have been used to visualize and profile diverse biomolecules. These studies have provided fundamental insights into various aspects of cell and organismal biology. Such studies also revealed limitations in the bioorthogonal toolkit, inspiring the development of more selective and finely tuned probes.

Central to all applications of bioorthogonal chemistry are reactions that are chemoselective and compatible with living systems. The development of such transformations can be difficult. The solvent and temperature are fixed, and the reactions must proceed in the midst of a multitude of interfering functionalities. The reactions often cannot be accelerated simply by heating up the subject or adding more reagent. Thus, the canonical rules for fine-tuning chemistries in round-bottom flasks often fail to translate to physiological environments, where few parameters can be varied [2]. Nonetheless, generations of chemists have been inspired to control bond

formation in live cells and organisms. Their efforts have provided transformations that can be executed without detriment in living systems.

This Chapter assesses the development of bioorthogonal reactions, with an emphasis on how mechanistic insights have driven the field. Like other areas of chemistry, no 'one-size-fits-all' transformation exists. Rather, each reaction has its pros and cons, with limitations that continue to propel new advances. In the first section, we provide a brief history of bioorthogonal chemistry and introduce the common tactics that facilitated early probe development. The bulk of the Chapter then showcases recent achievements in bioorthogonal reaction design. We also highlight efforts to engineer bioorthogonal chemistries that can be used concurrently for multicomponent labelling. These, and other, innovations will continue to expand the collection of biocompatible and mutually orthogonal reagents.

1.2 Setting the stage: Classic bioorthogonal transformations

Chemists rely on robust and versatile reactions to craft complex molecules. Synthetic routes are drafted with reagent accessibility, yields and selectivities in mind. Ideally, the transformations are fast, selective and applicable to a broad range of substrates. In reality, most reactions are not universally efficient and require tweaks to temperature, pH or stoichiometry in different contexts. Catalysts and solvents are also varied to optimize yields. Limitations in reaction scope often become the inspiration for new transformations. This iterative cycle of reaction discovery and refinement has provided a compendium of methods for controlling bond formation in various contexts.

In some cases, hundreds, if not thousands, of specialized reagents have been developed to address shortcomings in reaction scope.

Iterative refinement has also been used to tune reagents for use in living systems, with certain considerations in mind (Figure 1-1) [2-3]. Bioorthogonal functional groups must be both kinetically and metabolically stable, yet react rapidly with complementary probes under physiological conditions. The reactions must also be tolerant of water and other biological functional groups [4-6]. The constraints imposed by cells and tissues exclude many organic transformations. Several biological applications also demand reagents that impart a minimal steric 'footprint' [7-8]. Thus, developing chemistries that feature small reagents is another important goal.

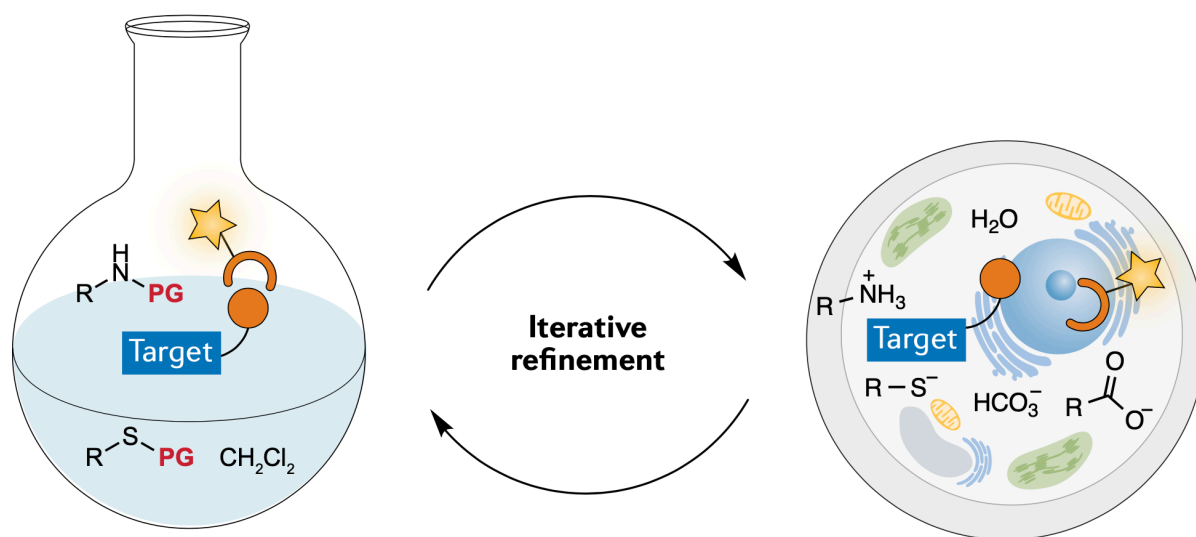


Figure 1-1. Translating reactions from round-bottom flasks to living systems. Reactants (orange circle and arc) used in cells and tissues must be compatible with cellular functionality. Challenging applications continue to drive the refinement of bioorthogonal reactions. Optimization in either flasks or cells invokes similar principles, although variables that are easily controlled in flasks are often invariant in living systems.

So, where does one begin? Hunting for unusual functional groups in heterologous organisms is a good starting point. Microorganisms, plants and other species often have access to chemistries and functionalities not present in mammalian cells. These motifs can survive in biological environments and are immediately orthogonal to mammalian systems, making them attractive candidates for bioorthogonal chemistry. A classic example is the alkyne, a motif that is present in numerous microbial metabolites [9], but absent in higher eukaryotes. The alkyne has been applied as a bioorthogonal motif in numerous settings [2]. Popular bioorthogonal functional groups also include some unlikely candidates from synthetic chemistry [10-11]. Organic azides and strained alkynes were traditionally viewed as too unstable for use in living systems. However, careful tuning has provided reagents now recognized among the gold standards in the field. These examples provided important lessons for subsequent reagent development [12]. In this section, we provide a brief perspective on how some seemingly “misfit” functional groups became stalwarts of the two most common classes of bioorthogonal chemistry: polar reactions and cycloadditions (Figure 1-2). We highlight early obstacles and key advances in the development of these probes. Initial successes provided a roadmap for continued reagent refinement.

1.2.1 Polar reactions

Aldehydes and ketones were among the first reagents employed as bioorthogonal labels. These carbonyl groups were attractive for biomolecule tagging, owing to their small size and compatibility in living systems [13]. Moreover, these electrophiles also readily condense with α -nucleophiles, such as hydrazides and aminoxy groups (Figure

1-2Ab). Aldehyde and ketone condensations have been used to label a wide variety of biomolecules [14-17]. However, the reactions are pH sensitive and quite slow in physiological settings. Aniline catalysts [18-21] can boost the rates, although the transformations remain difficult to execute in cellular environments.

Although aldehydes and ketones have been less widely used in cells, only a few other bioorthogonal functional groups rival their minimal size. Among the most influential has been the organic azide. This motif is abiotic and comprises just three atoms. Azides are remarkably inert in biological settings but exhibit unique manifolds of reactivity. For example, they can react with soft nucleophiles, including triarylphosphines (via Staudinger reduction). This reaction proceeds via an aza-ylide intermediate, which can be intercepted by a neighbouring electrophile. Bertozzi and colleagues capitalized on this feature, installing an ester on the phosphine to trap the aza-ylide (Figure 1.2Ad). The transformation ultimately linked the two reactants via an amide bond. This variant – termed the Staudinger ligation – was amenable to tagging azides in a variety of complex environments, including live cells [22]. Early applications featured glycans and post-translational modifications, although many other biomolecules have since been targeted [23-27]. The Staudinger ligation was also the first reaction of its kind to be used in live rodents, a particularly demanding environment [28].

The versatility of the Staudinger ligation propelled an entire field of reaction development. Many early studies were directed at increasing ligation speed [4-5]. Although the Staudinger ligation was robust enough to tag glycans and other biomolecules in vivo, the slow rate proved challenging for imaging studies in rodents and higher organisms. Large amounts of reagent were required to drive covalent bond

formation on reasonable timescales. Such quantities were not easily achieved, owing to the limited water solubility of many phosphine probes. These shortcomings generally sidelined the *in vivo* application of Staudinger ligation but helped to inspire other types of transformations with organic azides.

1.2.2 Cycloadditions

Cycloadditions are popular bioorthogonal transformations, owing to their exquisite chemoselectivity (Figure 1-2Ba). One of the earliest exploited was [3+2] cycloaddition with azides. In addition to being mild electrophiles, organic azides are 1,3-dipoles that are subject to react with alkynes. As alkynes are rare in higher eukaryotes, they are suitably orthogonal. However, azide–alkyne cycloadditions typically require high temperatures or pressures to proceed, conditions that are not biocompatible. A key breakthrough was the recognition that the cycloaddition could be accelerated using Cu(I) (Figure 1.2Bb). The copper-catalyzed azide-alkyne cycloaddition (CuAAC) is ubiquitous in chemical biology and other disciplines [29-30], and is still inspiring new transformations [31]. The cytotoxicity of Cu(I) [32-34] has precluded the *in vivo* application of CuAAC in some cases, although more biofriendly catalysts have been developed. Many feature water-soluble ligands that stabilize Cu(I) and prevent the formation of reactive oxygen species [35-38]. Polytriazole ligands, in particular, enabled CuAAC reactions to be conducted in live cells and zebrafish [36].

The constraints posed by copper catalysts also inspired chemists to devise new strategies to ligate 1,3-dipoles. One of the most fruitful approaches relied on strain energy – bending the alkyne from its normal linear geometry [39]. The smallest of the

stable cycloalkynes, cyclooctyne, was found to react with azides in the absence of a copper catalyst [40]. This work gave rise to an entire family of strain-promoted azide–alkyne cycloadditions (SPAACs) (Figure 1-2Bc). Such reactions have been widely used in complex biological settings, including plants [41], *C. elegans* [42], and mice [43-46]. Although SPAAC reactions minimize toxicity in living systems, some of the alkynes react with biological thiols [47]. Moreover, the fastest reactions observed with the most reactive strained alkynes plateau at rates of $\sim 1 \text{ M}^{-1} \text{ s}^{-1}$, which are not suitable for some *in vivo* applications.

Other cycloadditions have been developed to address the need for fast-acting, biocompatible reagents. One notable class comprises inverse electron-demand Diels–Alder (IEDDA) cycloadditions (Figure 1-2Bd). The IEDDA reaction between tetrazine and *trans*-cyclooctene (TCO), in particular, has gained prominence (Figure 1-2Be) [11,48-50]. This reaction pair boasts rates up to 10^6 times faster than the most reactive SPAAC pair. Such rapid reactions have widespread use in cells and living organisms, as only small amounts of material are required to drive covalent bond formation. The kinetic profile of the tetrazine-TCO cycloaddition has enabled the exploration of new imaging platforms *in vivo*, including PET imaging [51-52] and other modalities [53-55]. Tetrazine-TCO reactions have also been widely employed to tag biomolecules, release prodrugs [56-59], and activate enzymes [60-62].

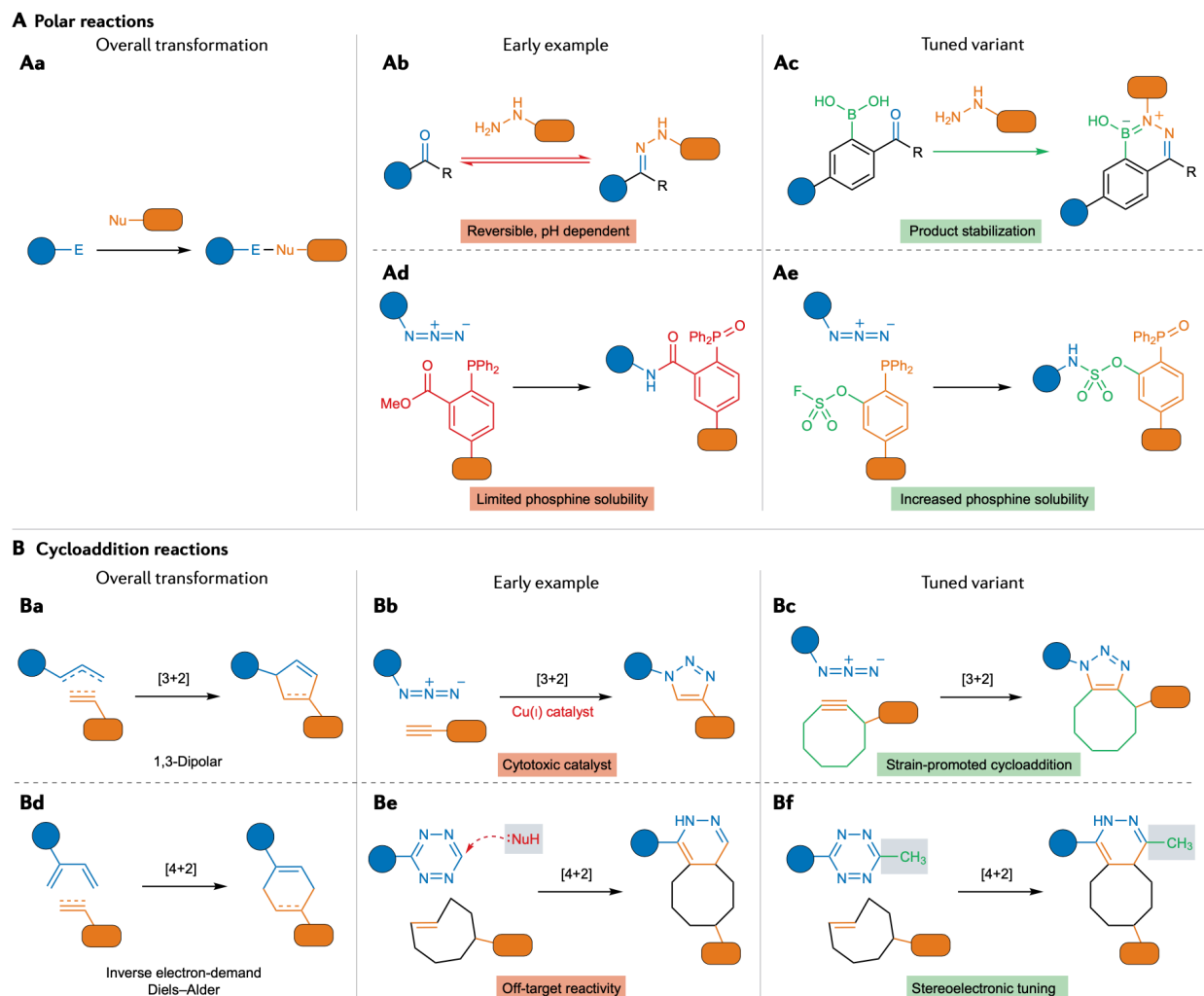


Figure 1-2. Tuning transformations for biological application. Bioorthogonal chemistries largely fall into two categories: polar reactions and cycloadditions. Examples of how limitations (highlighted in red) have spurred the development of new reactions are shown.

1.3 Fine-tuning bioorthogonal reagents and reactions

Despite remarkable achievements in bioorthogonal reaction design over the past two decades, limitations remain. Many reagents are insufficiently stable for use in the most demanding biological settings. Others react too slowly to transition to in vivo applications. And although many of the reactions work well on single targets, most cannot be used in combination to label multiple targets, owing to cross-reactivity

concerns. Simultaneous tracking of multiple biomolecules thus remains a long-standing challenge. Collectively, these limitations underscore the need to not only develop new reactions but also to fine-tune existing chemistries. The next sections highlight recent approaches to address voids in the bioorthogonal toolbox and to broaden the diversity of reactivities (Figure 1-2).

1.3.1 Tuning polar reagents and reactions

As noted above, some of the earliest bioorthogonal transformations involved condensation reactions with carbonyls and α -nucleophiles. The reversible nature of these reactions presented challenges in biomolecule labelling. Efforts to forge more stable adducts have been undertaken, including the use of proximal stabilizing groups [63-69]. A key example features boronic acids (Figure 1-2Ac). These functional groups can coordinate ligation adducts, including hydrazones [64] and oximes [65-66], preventing hydrolysis. Boronic acids have also been used to stabilize ligation adducts formed from other α -nucleophiles [67-69], as well as tag N-terminal cysteine residues on peptides and small proteins [70-71].

Polar reactions with azides have also been tuned for specific applications. The Staudinger ligation was an early target for optimization, with efforts focused on increasing rates and reagent solubility. Introducing electron-withdrawing groups on the azide or electron-donating groups on the phosphine increased the ligation speed. For example, electron-deficient aryl azides with fluorine [72] and chlorine [73] substituents exhibited faster rates compared to their unsubstituted counterparts. Interestingly, the key aza-ylide formed in these cases was a stable adduct, obviating the need for an

electrophilic trap on the phosphine. Halogenated aryl azides have since been used to label glycans and proteins in live cells [72-73]. Generating suitable electron-rich phosphines for Staudinger ligation has been less straightforward. More nucleophilic scaffolds can reduce disulfide bonds in proteins and are prone to rapid oxidation. Despite this liability, numerous phosphine probes have been tuned for cellular, and other, applications [74-76].

Phosphine reagents have also been tuned for improved biocompatibility. For example, Ren and co-workers synthesized phosphines comprising a fluorosulfate group (Figure 1-2Ae) [77]. This handle functions similarly to the ester trap in the original Staudinger ligation, by intercepting the aza-ylide. After ejection of fluoride and hydrolysis, the ligation provides an aryl sulfamate ester. The modified phosphine displayed markedly improved water solubility compared to initial Staudinger ligation probes. The sulfamate ester products also mimic phosphate backbones present in many biomolecules and metabolites.

Entirely different phosphorus nucleophiles have also been examined for azide ligation. Phosphites [78-80] and phosphonites [81-83], for example, react similarly with aryl azides compared to the Staudinger ligation probes. The resulting phosphoramidate and phosphonamidite adducts (respectively) can serve as surrogates for phosphate ester backbones found in nucleotides and other biomacromolecules.

Tuned phosphines have been exploited in other chemical contexts. For example, phosphines can react with Michael acceptors to form stable phosphonium adducts. This transformation has been employed to label biomolecules. In one study, tris(2-carboxyethyl)phosphine (TCEP), a water-soluble reductant widely used in biology, was

reacted with electron-deficient alkenes. This reaction was used to study protein glycosylation in live cells [84] and crotonylation patterns on histone proteins [85].

A more recent addition to the phosphine ligation kit comprises cyclopropenone derivatives (CpXs) [86-89]. In the case of cyclopropenones (X = O), the ligation proceeds via an initial Michael-type addition, followed by ring-opening to form a reactive ketene-ylide (Figure 1-3). The ketene can be trapped by pendant nucleophiles on the phosphine to furnish α,β -unsaturated carbonyls. Mono-substituted CpOs were prone to react with biologically relevant thiols [86], but tuning has provided more bioinert scaffolds comprising dialkyl motifs [87]. These latter reagents can be used in live cells.

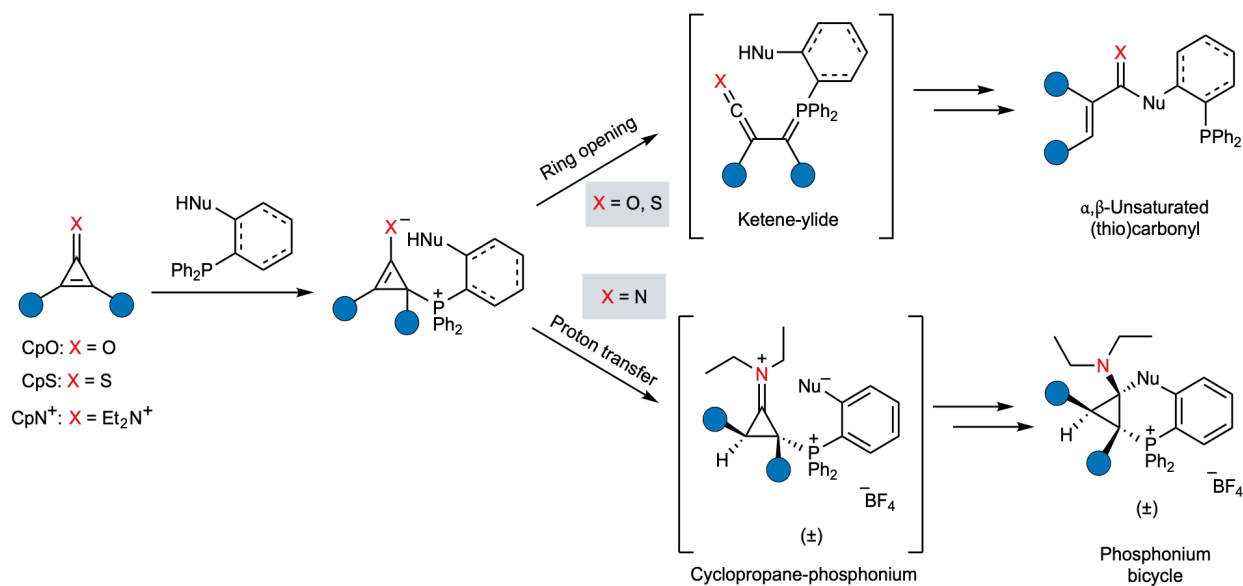


Figure 1-3. Cyclopropenone (CpX) derivatives exhibit unique reactivities. Analogs featuring oxygen or sulfur atoms (CpO and CpS, respectively) react with phosphines via ketene-ylide intermediates. By contrast, nitrogenous scaffolds (CpN⁺) react to form phosphonium bicycles.

The reactivity of CpOs can be further modulated through heteroatom replacement, a common strategy for probe tuning in bioorthogonal chemistry. The sulfur variant of CpO, cyclopropenethione (CpS), also reacts rapidly with phosphines via thioketene-ylide intermediates [88]. The thiocarbonyl products are useful biophysical

probes to study protein stability and function [90-91]. Nitrogen CpO heteroanalogs, cyclopropeniminium (CpN^+) ions, were also explored as bioorthogonal reagents. Interestingly, CpN^+ ions react with phosphines via a different mechanism than that of CpO or CpS derivatives. The initial Michael addition provides an enamine, which undergoes proton transfer instead of ring opening to give phosphonium bicycles [89] (Figure 1-3). Thus, small changes to a core scaffold can have profound effects on reactivity.

1.3.2 Tuning dipoles for bioorthogonal cycloaddition

The azide-alkyne cycloaddition remains one of the most popular 1,3-dipolar cycloadditions for examining biological systems. The success of this ligation has served as inspiration for exploring new 1,3-dipoles (Figure 1-4a). One example includes nitrones [92], reagents that can react more rapidly than azides with certain alkynes [93-95]. Nitron-alkyne cycloadditions, similar to their azide-alkyne counterparts, can be accelerated by copper catalysis. Such reactions have been used to image sugar metabolism and receptor-ligand interactions in mammalian cells, as well as peptidoglycan structures in bacteria [96-98].

Another newcomer to the bioorthogonal dipole set is sydnone [99]. This 1,3-dipole reacts with alkynes to afford pyrazole adducts. Although initial applications required copper catalysis [100], subsequent work identified scaffolds that were also capable of copper-free reactions with strained alkynes [101]. Sydnones have undergone further modification to modulate their reactivity for biological application. Chlorine substituents were found to increase reaction rates (30-fold) with a popular strained

alkyne, bicyclo[6.1.0]nonyne (BCN) [102]. Fluorine substitution further boosted reaction rates of sydnone with a variety of strained alkynes [103].

Close relatives of organic azides, diazo compounds have been similarly tuned as dipoles for bioorthogonal application [104]. The small size of the diazo group makes it attractive for biomolecule labeling, but such motifs were long thought to be too reactive for cellular use. However, diazo motifs are stable when in conjugation with aryl systems or electron-withdrawing groups (e.g., esters, amides) [105-106]. They ligate strained alkynes with second-order rate constants similar to azides. Diazo-cyclooctyne reactions have been used to tag cellular glycans among other targets [107-108]. Importantly, the resulting pyrazole adducts are stable to a number of biological nucleophiles [109]. Further diazo tuning has provided scaffolds that react with acyclic electron-deficient alkenes, a transformation that can be performed in the presence of azides [110].

More reactive 1,3-dipoles have also been harnessed for bioorthogonal application. Most are masked until an external trigger (often light) is applied, enabling the reactant to be released “on demand” [106, 111-113]. One such class of dipoles comprises nitrile imines. These motifs react robustly with strained alkenes such as norbornene [114] and cyclopropene [115-116], but they are prone to rapid hydrolysis in the absence of a ligation partner. Nitrile imines can be caged as tetrazoles or sydnone, functional groups that are more stable in biological environs. UV irradiation can liberate the reactive species. The starting tetrazole chromophores can be tuned for photolysis – and thus nitrile imine release – using different wavelengths of light [117-118]. This added layer of control has inspired the exploration of other “photo-click” reactions to

expand the compendium of bioorthogonal chemistries [119-121]. Nitrile imines have also been extensively tuned via electronic [122] and steric modification [123].

1.3.3 Tuning dipolarophiles for bioorthogonal cycloaddition

Tuning bioorthogonal cycloadditions is perhaps best exemplified in the context of the strained cycloalkynes as dipolarophiles [10,12]. Many efforts have been directed at modifying ring strain, with an eye towards increasing reaction rates or *in cellulo* stability. Examples include modulation of ring size or conformation [124-131], electronic perturbation [132-135], installation of endocyclic heteroatoms [136-138], and combinations thereof (Figure 1.4b) [139-140]. There have also been significant efforts to improve the water solubility of these relatively greasy probes. Towards this end, cycloalkynes featuring sulfamate backbones, as well as larger heterocyclic derivatives (up to 12-membered rings) have also been explored [141-146]. The heteroatom variants were generally more water soluble and stable. However, the benefits of ring expansion came at the cost of rate, a trade-off that is prevalent in bioorthogonal reagent tuning [2]. With the reactivity-stability axis in mind, one of the more impactful cycloalkynes has been BCN [131]. This scaffold has increased strain energy compared to the original cyclooctyne (due to the fused cyclopropane ring) [147], but is remarkably stable in cellular environments. Coupled with its synthetic accessibility, BCN has become a staple strained alkyne for cycloaddition chemistries in a number of fields. In a recent example, BCN-fluorophore conjugates were used for super-resolution imaging in live cells (Figure 1.4c,d) [148]. FtsZ, a protein involved in bacterial cell division, was enzymatically outfitted with an azide. Subsequent treatment with cell-permeable BCN

derivatives (comprising photoswitchable rhodamines) enabled protein localization patterns to be observed.

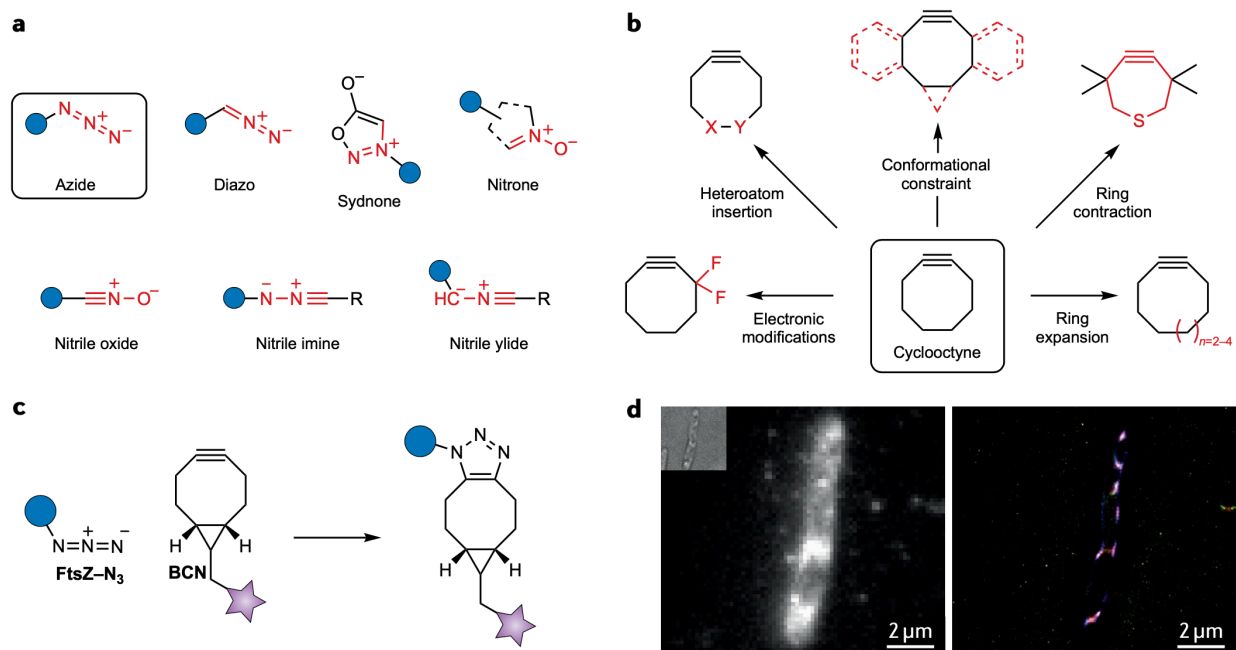


Figure 1-4. An expansive set of reagents for dipolar cycloaddition. (A) Various 1,3 dipoles are available for bioorthogonal labeling. Organic azides were among the first to be exploited for cycloadditions in cells. More recently, diazo, sydnone, and nitron motifs have been developed for biomolecule labeling. Highly reactive dipoles (including nitrile oxides, imines, and ylides) are also viable ligation partners and are typically produced “on demand” via photolysis or chemical uncaging. (B) Several classes of strained alkynes are available for bioorthogonal cycloaddition. These reagents derive from cyclooctyne, the first-reported variant. Extensive modifications to cyclooctyne provide a suite of reagents with varying reactivities. (C) Example application of bioorthogonal cycloadditions. FtsZ, a bacterial protein involved in cellular division, was outfitted with an azide handle. Subsequent ligation with BCN-photoswitchable fluorophore conjugates enabled super-resolution microscopy in live *E. coli*. (D) Diffraction-limited fluorescence (left) and super-resolution (right) microscopy images of FtsZ localization in live *E. coli*. The inset shows a bright-field image of the bacterial cell. Images were adapted with permission from ref. [148], The American Chemical Society.

Efforts to modulate cycloalkyne reactivity have been bolstered by computation. Calculations can readily predict combinations of steric and electronic features to tune scaffolds for desired outcomes. One well-established approach for modulating cycloaddition partners relies on the Distortion/Interaction model [149-150]. This analysis

computes the activation barrier for a given reaction by calculating the difference between a distortion energy (i.e., the energy required for reactants to adopt their ideal transition state geometries) and an interaction energy (i.e., favorable orbital overlap between the two reaction partners). The calculated activation energy is then correlated to a predicted rate constant, which can ultimately guide reagent tuning. One of the earliest demonstrations of the Distortion/Interaction model in bioorthogonal reagent design involved a series of biarylazacyclooctynone (BARAC) analogues. This study revealed structural features that impeded BARAC reactivity with azides and set the stage for improved cyclooctyne development [151]. Similar computational studies have been used to fine-tune sydnone-cycloalkyne reactions [152] and other bioorthogonal cycloadditions [149, 153-154].

1.3.4 Tuning dienophiles for bioorthogonal cycloaddition

Parallel developments in the realm of IEDDA have expanded the number of bioorthogonal chemistries in recent years. Reactions featuring TCO, in particular, have found widespread use in cells and *in vivo* [11]. Such transformations have also been the targets of extensive reagent tuning [155-159]. Many efforts have focused on pushing the kinetics of the tetrazine-TCO ligation (Figure 1-5). Increasing TCO strain was hypothesized to boost reaction rate, similar to the strained alkynes. Toward this end, Fox and coworkers designed scaffolds wherein TCO was forced to adopt a half-chair conformation (e.g., d-TCO [155] and s-TCO [155]). Such motifs were predicted to be 5.6-5.9 kcal mol⁻¹ higher in energy relative to the more stable crown conformation of

TCO [160]. The more reactive d-TCO and s-TCO variants display bimolecular rate constants as high as $3 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$.

Similar to the strained alkynes, ring contraction strategies have been employed in the context of strained alkenes. Recently, *trans*-cycloheptene (TCH) analogs have been reported as voracious dienophiles in IEDDA cycloadditions with tetrazines [161-162]. TCH readily isomerizes under ambient conditions but can be isolated as a stable complex with Ag(I). In addition, incorporation of an endocyclic silicon atom provided a more stable seven-membered cycloalkene, sila-*trans*-cycloheptene (SiTCH). The longer Si-C bond lengths in SiTCH relieved some ring strain, enabling facile isolation. Computational studies further revealed that the activation barrier for SiTCH reactivity with diphenyltetrazine was significantly lower compared to s-TCO, enabling rapid ligation. The second-order rate constant for SiTCH and a model tetrazine was $1.14 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$, the fastest bioorthogonal ligation on record. Despite their impressive rates, though, TCH and SiTCH degrade rapidly in the presence of cellular thiols [162].

1,3-Disubstituted cyclopropenes, alternative dienophiles for IEDDA reactions, have also been tuned for a variety of applications. We and others reported that these scaffolds react with tetrazines in biological environments [163-164]. Compared to TCO, the reaction between tetrazines and cyclopropenes is much slower. However, the small size and cellular stability of cyclopropenes have rendered them attractive for interrogating biomolecules in complex environments [165]. One notable example from the Chin lab showcased a cyclopropene-lysine analog to monitor nascent protein biosynthesis in *Drosophila* [166].

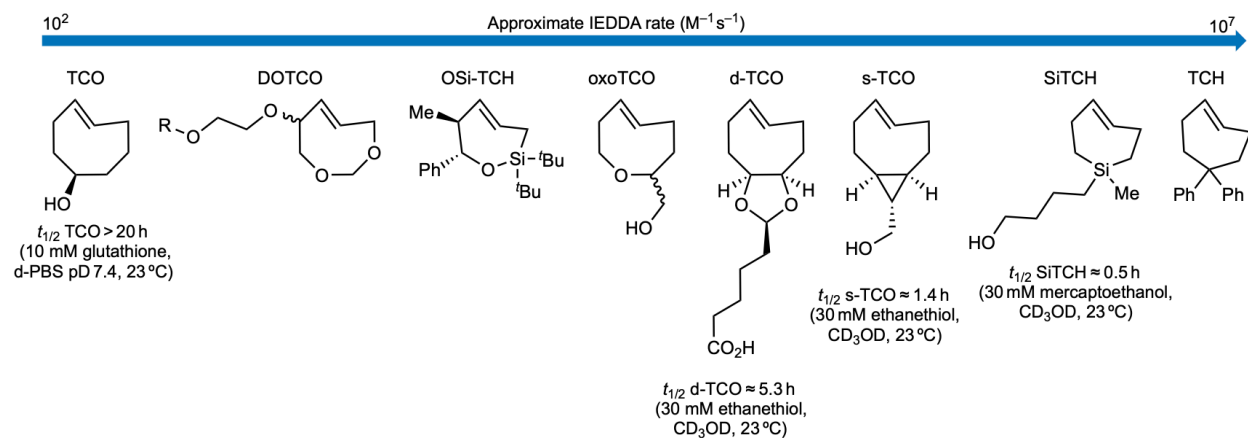


Figure 1-5. A collection of strained alkenes for IEDDA reactions. These reactions exhibit a range of rates with tetrazines. *trans*-Cyclooctene (TCO) was the original member of this group. TCO has since been modified to access faster and more biocompatible reagents. Some of the scaffolds feature endocyclic heteroatoms (DOTCO [156], OSi-TCH [161], oxoTCO [158]), conformational constraints (d-TCO [155], s-TCO [155]), and contracted rings (SiTCH [162], TCH [162]). In general, the fastest reacting scaffolds exhibit the shortest half-lives in the presence of biological thiols.

The reactivity profile of the cyclopropene can be dramatically influenced by steric tuning. The Distortion/Interaction model predicted that increasing steric bulk at C3 on cyclopropene would drastically reduce its reactivity with tetrazines. Diminished reactivity was attributed to geometric constraints in the transition state. The poor orbital overlap between the cycloaddition partners was predicted to slow the ligation. Indeed, no reaction was observed even when 3,3-disubstituted cyclopropenes were subjected to a variety of functionalized tetrazines. Such cyclopropenes still reacted with nitrile imines, though, and this differential reactivity was exploited to label two proteins in tandem [116].

Further cyclopropene tuning was achieved upon introduction of spirocycles at C3. Based on crystallographic data, cyclopropene **1.2** exhibited a reduced bond angle between the two C³-groups compared to parent cyclopropene **1.1** (Figure 1-6). The decreased bond angle drove the substituents away from the incoming nitrile imine,

increasing ligation rates by 15-fold [167]. Interestingly, the spirocyclic cyclopropenes also exhibited reactivity with some tetrazines, and have been used to label cell-surface receptors in live cells [168]. Heteroatoms within the spirocycle further boosted cyclopropene reactivity by lowering LUMO energies [169]. Spirocyclic cyclopropenes have recently been outfitted with light-sensitive cages. Upon uncaging, the scaffolds become available for IEDDA ligation. These masked reagents can be used for biomolecule labeling with spatiotemporal control, but their syntheses remain challenging [170-171].

Tetrazines have also been explored as reactants for other biocompatible cycloadditions. Recent work has featured transformations with isonitriles. These dienophiles react with tetrazines via [4+1] cycloaddition to provide imine products. With primary and secondary isonitriles, the imines undergo facile tautomerization and hydrolysis to liberate amines. Such isonitriles have proven useful as cages for amino fluorophores and small-molecule drugs [172-173]. Hydrolysis can be mitigated with appropriately tuned isonitriles [174-175]. For example, tertiary scaffolds react with tetrazines to form stable ligation adducts, with no subsequent reactions [175]. Because of their small size and versatility, isonitriles have been used to label biomolecules, including proteins [176] and glycans [177].

Simple vinyl groups have been shown to undergo IEDDA reactions with tetrazine probes. The small size of the vinyl motif is attractive for biomolecule tagging strategies. However, even with electron-rich vinyl reagents, most of the alkenes examined to date suffer from low aqueous solubility and slow rates [178-179]. Vinylboronic acids (VBAs) react significantly faster with tetrazines. The ligation liberates boronic acid, a major

driving force for the reaction [180]. Even more rapid cycloadditions can be achieved with electron-rich VBAs and tetrazines outfitted with hydroxy groups. These latter groups coordinate the boronic acid motifs [181-182]. The VBA-tetrazine ligation has been used to profile the efficacy of proteasome inhibitors in live cells [183]. Many other cyclic [184-185] and acyclic [186] alkenes have been similarly explored as tetrazine ligation partners [187].

1.3.5 Tuning dienes for bioorthogonal cycloaddition

The second half of the IEDDA reaction, the diene (most often, tetrazine), has also been thoroughly modified to achieve altered stability and reactivity profiles (Figure 1.2Bf) [62, 182, 188-191]. In one example, the stability of the tetrazine probe was modulated using electronic perturbations. An amino acid comprising a tetrazine motif was initially found to hydrolyze in the cellular milieu. The hydrolysis was facilitated by a labile secondary amine linkage. The electron-donating amine group also slowed IEDDA rates. Simple removal of the amine handle addressed both of these limitations [190]. In another example, researchers synthesized a panel of tetrazines for amine uncaging with TCO. The tetrazines were screened with model enzymes bearing TCO-caged amino acids. The best tetrazines facilitated near quantitative uncaging in under 4 minutes with as little as 50 μ M reagent [62].

The rates that some tetrazine ligations achieve are unparalleled. Such rates, though, often come at the cost of probe stability. Many of the most reactive tetrazines are known to react with thiols, which can lead to off-target effects and high background labeling in cells [192]. To address this liability, some groups have drawn inspiration from

caged bioorthogonal reagents. These efforts are focused on liberating the reactive tetrazine and take advantage of the redox properties of the motif. The reduced form, dihydrotetrazine, is unreactive towards dienophiles and stable in biological contexts. Dihydrotetrazines can then be oxidized to tetrazines *in situ* via enzymatic [193], or photocatalytic [193] approaches. Such strategies provide an avenue to employ even the most reactive tetrazines in biological environments. Caged tetrazines have also been employed to decorate electrode surfaces with biomolecules [194].

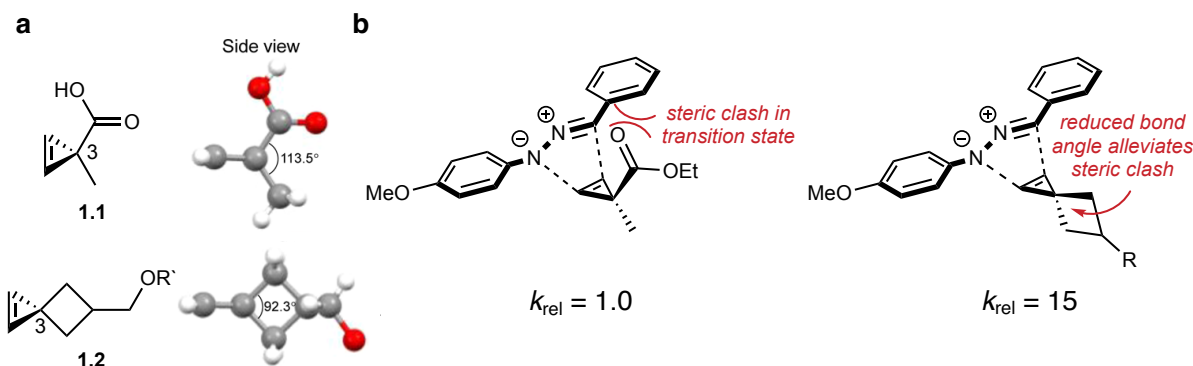


Figure 1-6. Steric modifications tune cyclopropene reactivity. (A) The angle between C³-substituents is reduced in spirocyclic cyclopropene **1.2** compared to parent compound **1.1**. (B) The decreased bond angle minimizes steric clashes in the transition state, resulting in faster reactions with nitrile imines. Part (A) was adapted with permission from ref. [167], The American Chemical Society.

The instability of some tetrazine motifs inspired pursuits of less electron-deficient dienes, including 1,2,4-triazines. Like its tetrazine counterpart, the triazine reacts via IEDDA with strained π -systems such as TCO and BCN [195-196]. While the rates of these reactions are markedly slower, the triazine is stable in the presence of biological thiols for over 24 hours at elevated pH. These scaffolds have since been installed in proteins [195,197], as well as enzymatically appended onto DNA *in vitro* [198-199]. 1,2,4-Triazines have also been electronically tuned to achieve faster kinetic

profiles [200-201]. For example, electron-withdrawing pyridinium groups have been used to improve cycloaddition rates. These scaffolds have been recently used to label mitochondria in live cells [200]. Subtle steric modifications to the triazine core also provided altered modes of reactivity with strained alkynes. In a recent study, the Distortion/Interaction model predicted that triazine substitutions at C3 and C6 would diminish reactivity with sterically encumbered strained alkynes [202]. Such reactions were predicted to remain facile with C⁵-substituted isomers, as steric clashes were minimized at the bond-forming centers (i.e., C3 and C6). The predictions were confirmed experimentally via simultaneous, dual labeling of two protein targets.

The number of dienes that can participate in IEDDA cycloadditions is also growing. A nitrogenous heterocycle, 4*H*-pyrazole, was recently shown to ligate the strained alkyne BCN [203]. At the outset, the pyrazole required further tuning to elicit robust reactivity. Addition of fluorine substituents was hypothesized to impart negative hyperconjugative effects on the pyrazole ring. This tuning would increase the antiaromatic character of the scaffold, resulting in destabilization of the pyrazole ring and more rapid ligation. Computational analyses verified that a *gem*-difluoro group decreased the LUMO values of the 4*H*-pyrazole, resulting in fast reactivity with BCN. Such antiaromaticity considerations along a reaction coordinate could be more generally exploited in bioorthogonal reaction design.

Another class of dienes comprises *ortho*-quinones. These motifs react with dienophiles via strain-promoted oxidation-controlled cyclooctyne–1,2-quinone (SPOCQ) cycloadditions [204]. Early iterations involved converting 1,2-catechols to the corresponding quinones using an exogenous oxidant [205]. SPOCQ reactions were

orders of magnitude faster than azide-alkyne cycloadditions (with second-order rate constants of $\sim 500 \text{ M}^{-1} \text{ s}^{-1}$), on par with some tetrazine-TCO ligations. Later studies demonstrated that genetically encodable tyrosine tags could be selectively oxidized to quinones *in situ* using tyrosinase. These motifs could then be used to append small molecules for antibody-drug conjugate formation. The strategy enabled control over the location and number of warheads that were appended to the antibody [206]. *ortho*-Quinones were also found to react efficiently with strained alkenes, such as cyclopropenes [207]. However, the need to generate *ortho*-quinones *in situ* can impose constraints. *ortho*-Quinoline quinone methides were explored as more versatile alternatives, as these scaffolds can be generated *in situ* without any external triggers in biological environments. The motifs were found to react robustly with vinyl thioethers via hetero Diels–Alder cycloaddition. The unusual thioacetal adduct formed was found to be stable in aqueous solution at various pH values [208-209].

1.4 Combining mutually orthogonal bioorthogonal reactions

As evident from above, the past decade has seen a surge in the number of transformations available for biological application. Despite the expanded toolkit, it remains challenging to apply more than one reaction at a time [210]. New reaction development has largely focused on labeling single targets. Identifying collections of compatible bioorthogonal chemistries would enable multicomponent labeling studies, and allow a broader set of biological processes to be examined. Finding such combinations of reactions has historically been challenging, as many popular reagents

cross-react with one another. The search for orthogonal reactions has accelerated in recent years, aided by computational tools and reaction tuning.

Bioorthogonal reactions that feature unique mechanisms are well suited for multicomponent labeling studies. Reagents with distinct modes of reactivity can often mitigate incompatibility issues. For example, azide-alkyne cycloadditions can be used in tandem with hydrazine/ketone condensations [211], various IEDDA reagents [163, 178, 212-214], some 1,3-dipoles [98, 110], and other motifs [180, 215-217]. Efforts to employ three mutually compatible bioorthogonal groups have also been pursued [202, 213, 218-220]. One recent example featured azide-, cyclopropene-, and alkyne-containing sugars to study the heterogeneity of glycan metabolism in plant cells [219]. The ligations could not be performed simultaneously, though, owing to cross-reactivities among the reaction partners. Cumbersome washes were also required between ligations, eroding temporal resolution. Only three studies to date have been able to achieve simultaneous triple labeling [221-223]. One notable example features two tetrazines, one that is sterically encumbered and reacts selectively with a small isonitrile. The second tetrazine ligates TCO in a typical IEDDA cycloaddition. The tetrazine reactions were combined with an azide/strained alkyne pair in a triple labeling experiment. Three model proteins, containing either a bulky tetrazine, a less encumbered tetrazine, or an azide were mixed and reacted with isonitrile-, TCO-, and strained-alkyne fluorophores. The matched adducts were detected by in-gel fluorescence, with no evidence of cross-reactivity (Figure 1-7) [223]. While this study showcased triple component labeling in a model context, the reagents should be applicable in other biological settings.

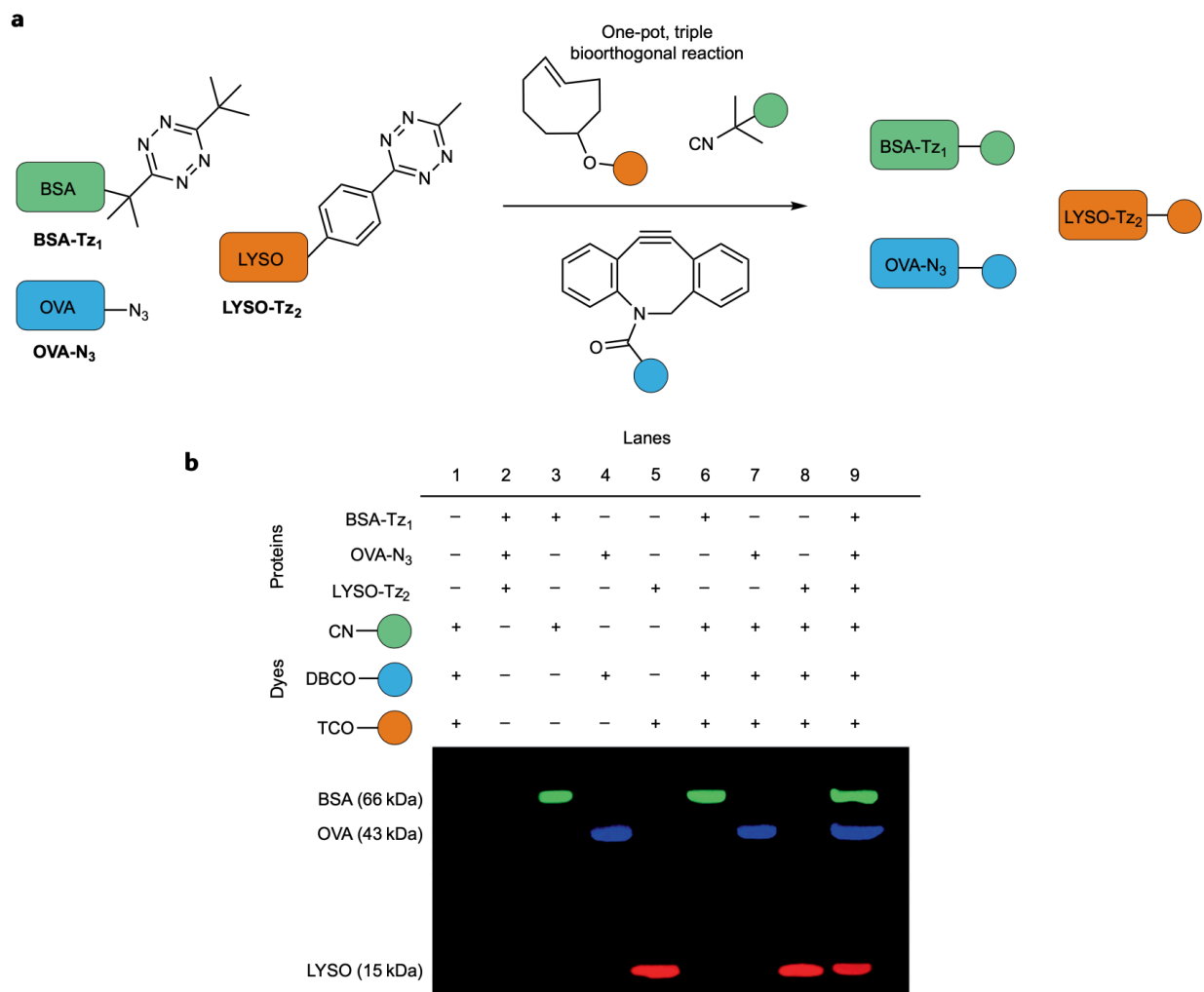


Figure 1-7. Triple component protein labeling. (A) Three mechanistically distinct reactions — [4+1] tetrazine (Tz₁)–isonitrile cycloaddition, [4+2] tetrazine (Tz₂)–*trans*-cyclooctene (TCO) cycloaddition and [3+2] azide–dibenzocyclooctyne (DBCO) cycloaddition — were used to simultaneously label three model proteins (bovine serum albumin (BSA), ovalbumin (OVA) and lysozyme (LYSO)). (B) The covalent protein adducts were visualized through gel electrophoresis followed by fluorescence scanning. No cross-reactivities were detected. Part (B) was adapted with permission from ref. [223], Wiley-VCH.

1.5 Exploring new genres of reactivity

Identifying additional genres of bioorthogonal chemistry will continue to bolster multicomponent labeling studies. Recent efforts to explore new areas of chemical space – coupled with additional tuning of existing ligations – are proving fruitful. Boron reagents are gaining traction. In recent work, a diboron probe was found to react

selectively with *N*-oxides. Importantly, the ligation proceeded efficiently inside cells, one of the harshest biological environments [224]. The unique mechanism of this reaction is further compatible with several existing bioorthogonal chemistries.

Perhaps the most noteworthy developments in polar reagent design involve sulfur(VI) fluorides. These motifs are remarkably stable in cellular environments, and react robustly with oxygen and nitrogen nucleophiles through sulfur(VI) fluoride exchange (SuFEx) chemistry [225-226]. SuFEx reagents have been used for drug design [227-228], activity-based profiling [229] and other protein targeting studies [230]. They have also been employed for examining protein-protein interactions via proximity-driven crosslinking [231-232]. SuFEx electrophiles also provide a mild and convenient method to introduce bioorthogonal azides onto a variety of amine targets, further exemplifying the utility of these motifs [233].

Explorations into new cycloaddition platforms are also expanding the bioorthogonal toolkit. One example involves the quadricyclane ligation [221]. This reaction is a formal $[2\sigma+2\sigma+2\pi]$ cycloaddition with Ni-bis(dithiolene) complexes, a rather unique mode of reactivity. Quadricyclanes are highly strained molecules (~ 80 kcal mol⁻¹ of strain energy [234]), but can survive in aqueous conditions and in the presence of thiols for extended periods. Quadricyclanes have even been used for intracellular labeling applications via genetic code expansion [235]. The novelty of the quadricyclane reaction is likely to inspire continued exploration of other cycloaddition manifolds for bioorthogonal application.

Bioorthogonal platforms of reactivity have also expanded to include transition metals. These metals are virtually absent in the cellular milieu and can forge stable C-C

bonds, making them attractive for selective labeling applications. Focused reviews on this topic have been covered recently, but a few examples are described below [236-237]. Water soluble Au(I) complexes have been developed to drive covalent bond formation in aqueous media [238]. Gold nanoparticles have also been employed for imaging in zebrafish [239]. More reactive Au(III) species can be harnessed for selective amidation reactions, although the metal complexes must be associated with scaffolding proteins [240]. Inspired by copper catalysis, several groups have optimized ligands to form water soluble, minimally toxic palladium complexes. Some have been used for Suzuki-Miyaura [241-242] couplings in cells, along with Sonogashira reactions [243-244], and alkynyl-carbamate deprotections [245-246]. Palladium nanoparticles [247] and assemblies [248] have also been developed for conducting various transformations *in cellulo*. Similar advances in ruthenium chemistry [249-251] and iridium photocatalysis [252] are enabling biomolecule targeting. New developments in protein bioconjugation are also being leveraged for more general bioorthogonal application. One example features aryl diazonium ions [253-254]. These electrophiles react robustly with electron-rich aromatic side chains such as tyrosine. Site-specific modification with aryl diazoniums was historically challenging owing to the abundance of tyrosine residues in proteins. Further exploration revealed a selective reaction between 5-hydroxytryptophan (5HTP) and aryl diazonium reagents. The additional electron density in 5HTP enabled the use of less reactive diazonium reagents, preventing background labeling of other aromatic side chains. This reaction has since been used to label recombinant proteins and antibody derivatives [255], and additional targets are

anticipated. Further tuning of other protein-labeling strategies is likely to provide more bioorthogonal reagents [256-257].

1.6 Introduction to the Thesis

At its core, bioorthogonal chemistry is focused on controlling reactive functional groups in harsh environments. How does one design a pair of reagents to form a single adduct, in the confines of a cell or whole organism? The strict requirements imposed on these reactions have often forced researchers to explore unconventional handles. Historic work provided an initial set of reagents, many of which are still employed for examining biomolecules *in vivo*.

The toolbox has greatly expanded in recent years via iterative tuning of established probes. Some common themes from these studies have emerged. For example, the rates and stabilities of first-generation tools can be modulated for particular applications, drawing on common physical organic chemistry principles. The impacts of certain modifications can often be predicted computationally. Computational analyses can also be invaluable to the hunt for collections of mutually compatible transformations.

After decades of work on bioorthogonal chemistries, there is still no “one-size-fits-all” reaction. Rather, a spectrum of reactivities exists and the challenge lies in knowing how best to apply the probes. Recent successes in tuning highly reactive chemical handles suggest that other “fringe” functional groups can be harnessed for understanding biology. There is a further need for not just single-component reactions, but also collections of bioorthogonal chemistries that can be used in tandem [210]. The

continued exploration of unique modes of reactivity [256-260] will be useful in this regard.

This thesis showcases new applications of a bioorthogonal reaction developed in the Prescher lab between cyclopropenones (CpOs) and phosphines [86-87]. Upon incubation with a phosphine probe, CpOs can undergo ring-opening to unveil a highly electrophilic ketene-ylide intermediate. This reactive species can be trapped both inter- and intramolecularly with appropriate nucleophiles, and is the crux of the projects described herein. In Chapter 2, the unique reactivity of CpOs and phosphines was exploited to develop a collection of three mutually orthogonal bioorthogonal reactions. Because the majority of popular bioorthogonal reactions are cycloadditions, the CpO-phosphine ligation presented a unique opportunity to develop mutually exclusive bioorthogonal reactions. These ligations could be performed in a one-pot fashion without any evidence of cross-reactivity between probes. In Chapter 3, the CpO-phosphine reaction was optimized to capture protein-protein interactions with temporal resolution. Whereas previous iterations of the ligation featured intramolecular trapping, this rendition of the reaction was optimized for intermolecular events. The resulting crosslinking strategy was applied to a split protein system, and could be performed even in the presence of cell lysate. Finally, in Chapter 4, the CpO-phosphine reaction was co-opted to develop a methodology for accessing substituted butenolides. The optimized conditions featured phosphine organocatalysis, and was applied to over 20 substrates bearing various functional handles. Collectively, this work will expand the capabilities of bioorthogonal reactions in previously underexplored applications.

1.7 References

- [1]. Rodriguez, E. A.; Campbell, R. E.; Lin, J. Y.; Lin, M. Z.; Miyawaki, A.; Palmer, A. E.; Shu, X.; Zhang, J.; Tsien, R. Y. *Trends Biochem. Sci.* **2017**, *42*, 111–129.
- [2]. Row, R. D.; Prescher, J. A. *Acc. Chem. Res.* **2018**, *51*, 1073–1081.
- [3]. Tian, Y.; Lin, Q. *ACS Chem. Biol.* **2019**, *14*, 2489–2496.
- [4]. Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974–6998.
- [5]. Lang, K.; Chin, J. W. *Chem. Rev.* **2014**, *114*, 4764–4806.
- [6]. Carrell, T.; Vrabel, M.; Yang, M.; Yang, Y.; Chen, P. R.; Dommerholt, J.; Rutjes, F. P. T. J.; van Delft, F. L.; Herner, A.; Lin, Q.; Wu, H.; Devaraj, N. K.; Kath-Schorr, S., *Cycloadditions in Bioorthogonal Chemistry*. Springer International Publishing: Switzerland, 2016; Vol. 374.
- [7]. Andresen, M.; Schmitz-Salue, R.; Jakobs, S. *Mol. Biol. Cell* **2004**, *15*, 5616–5622.
- [8]. Grammel, M.; Hang, H. C. *Nat. Chem. Biol.* **2013**, *9*, 475–484.
- [9]. Chai, Q.-Y.; Yang, Z.; Lin, H.-W.; Han, B.-N. *Mar. Drugs* **2016**, *14*, 216.
- [10]. Jewett, J. C.; Bertozzi, C. R. *Chem. Soc. Rev.* **2010**, *39*, 1272–1279.
- [11]. Selvaraj, R.; Fox, J. M. *Curr. Opin. Chem. Biol.* **2013**, *17*, 753–760.
- [12]. Dommerholt, J.; Rutjes, F. P. J. T.; van Delft, F. L. *Top. Curr. Chem.* **2016**, *374*, 16.
- [13]. Kölmel, D. K.; Kool, E. T. *Chem. Rev.* **2017**, *117*, 10358–10376.
- [14]. Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125–1128.
- [15]. Datta, D.; Wang, P.; Carrico, I. S.; Mayo, S. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2002**, *124*, 5652–5653.

- [16]. Rabuka, D.; Rush, J. S.; deHart, G. W.; Wu, P.; Bertozzi, C. R. *Nat. Protoc.* **2012**, *7*, 1052–1067.
- [17]. Hardisty, R. E.; Kawasaki, F.; Sahakyan, A. B.; Balasubramanian, S. *J. Am. Chem. Soc.* **2015**, *137*, 9270–9272.
- [18]. Dirksen, A.; Dirksen, S.; Hackeng, T. M.; Dawson, P. E. *J. Am. Chem. Soc.* **2006**, *128*, 15602–15603.
- [19]. Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem. Int. Ed.* **2006**, *45*, 7581–7584.
- [20]. Dirksen, A.; Dawson, P. E. *Bioconjugate Chem.* **2008**, *19*, 2543–2548.
- [21]. Larsen, D.; Kietrys, A. M.; Clark, S. A.; Park, H. S.; Ekebergh, A.; Kool, E. T. *Chem. Sci.* **2018**, *9*, 5252–5259.
- [22]. Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007–2010.
- [23]. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 19–24.
- [24]. Saxon, E.; Luchansky, S. J.; Hang, H. C.; Yu, C.; Lee, S. C.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2002**, *124*, 14893–14902.
- [25]. Cai, J.; Li, X.; Yue, X.; Taylor, J. S. *J. Am. Chem. Soc.* **2004**, *126*, 16324–16325.
- [26]. Hannoush, R. N.; Sun, J. *Nat. Chem. Biol.* **2010**, *6*, 498–506.
- [27]. Chuh, K. N.; Pratt, M. R. *Curr. Opin. Chem. Biol.* **2015**, *24*, 27–37.
- [28]. Prescher, J. A.; Dube, D. H.; Bertozzi, C. R. *Nature* **2004**, *430*, 873–877.
- [29]. Meldal, M.; Tornøe, C. W. *Chem. Rev.* **2008**, *108*, 2952–3015.
- [30]. Thirumurugan, P.; Matosiuk, D.; Jozwiak, K. *Chem. Rev.* **2013**, *113*, 4905–4979.

- [31]. Lampkowski, J. S.; Villa, J. K.; Young, T. S.; Young, D. D. *Angew. Chem. Int. Ed.* **2015**, *54*, 9343–9346.
- [32]. Gaetke, L. M.; Chow, C. K. *Toxicology* **2003**, *189*, 147–163.
- [33]. Soares, E. V.; Hebbelinck, K.; Soares, H. M. V. M. *Can. J. Microbiol.* **2003**, *49*, 336–343.
- [34]. Kennedy, D. C.; McKay, C. S.; Legault, M. C. B.; Danielson, D. C.; Blake, J. A.; Pegoraro, A. F.; Stolow, A.; Mester, Z.; Pezacki, J. P. *J. Am. Chem. Soc.* **2011**, *133*, 17993–18001.
- [35]. del Amo, D. S.; Wang, W.; Jiang, H.; Besanceney, C.; Yan, A. C.; Levy, M.; Liu, Y.; Marlow, F. L.; Wu, P. *J. Am. Chem. Soc.* **2010**, *132*, 16893–16899.
- [36]. Besanceney-Webler, C.; Jiang, H.; Zheng, T.; Feng, L.; del Amo, D. S.; Wang, W.; Klivansky, L. M.; Marlow, F. L.; Liu, Y.; Wu, P. *Angew. Chem. Int. Ed.* **2011**, *50*, 8051–8056.
- [37]. Uttamapinant, C.; Tangpeerachaikul, A.; Grecian, S.; Clarke, S.; Singh, U.; Slade, P.; Gee, K. R.; Ting, A. Y. *Angew. Chem. Int. Ed.* **2012**, *51*, 5852–5856.
- [38]. Yang, M.; Jalloh, A. S.; Wei, W.; Zhao, J.; Wu, P.; Chen, P. R. *Nat. Commun.* **2014**, *5*, 4981.
- [39]. Wittig, G.; Krebs, A. *Chem. Ber.* **1961**, *94*, 3260–3275.
- [40]. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046–15047.
- [41]. Rydahl, M. G.; Hansen, A. R.; Kračun, S. K.; Mravec, J. *Front. Plant Sci.* **2018**, *9*, 581.

- [42]. Yuet, K. P.; Doma, M. K.; Ngo, J. T.; Sweredoski, M. J.; Graham, R. L. J.; Moradian, A.; Hess, S.; Schuman, E. M.; Sternberg, P. W.; Tirrell, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 2705–2710.
- [43]. tom Dieck, S.; Kochen, L.; Hanus, C.; Heumüller, M.; Bartnik, I.; Nassim-Assir, B.; Merk, K.; Mosler, T.; Garg, S.; Bunse, S.; Tirrell, D. A.; Schuman, E. M. *Nat. Methods* **2015**, *12*, 411–414.
- [44]. Ernst, R. J.; Krogager, T. P.; Maywood, E. S.; Zanchi, R.; Beránek, V.; Elliott, T. S.; Barry, N. P.; Hastings, M. H.; Chin, J. W. *Nat. Chem. Biol.* **2016**, *12*, 776–778.
- [45]. Alvarez-Castelao, B.; Schanzenbächer, C. T.; Hanus, C.; Glock, C.; tom Dieck, S.; Dörrbaum, A. R.; Bartnik, I.; Nassim-Assir, B.; Ciirdaeva, E.; Mueller, A.; Dieterich, D. C.; Tirrell, D. A.; Schuman, E. M. *Nat. Biotechnol.* **2017**, *35*, 1196–1201.
- [46]. Krogager, T. P.; Ernst, R. J.; Elliott, T. S.; Calo, L.; Beránek, B.; Ciabatti, E.; Spillantini, M. G.; Tripodi, M.; Hastings, M. H.; Chin, J. W. *Nat. Biotechnol.* **2018**, *36*, 156–159.
- [47]. van Geel, R.; Pruijn, G. J. M.; van Delft, F. L.; Boelens, W. C. *Bioconjugate Chem.* **2012**, *23*, 392–398.
- [48]. Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518–13519.
- [49]. Taylor, M. T.; Blackman, M. L.; Dmitrenko, O.; Fox, J. M. *J. Am. Chem. Soc.* **2011**, *133*, 9646–9649.
- [50]. Devaraj, N. K.; Upadhyay, R.; Haun, J. B.; Hilderbrand, S. A.; Weissleder, R. *Angew. Chem. Int. Ed.* **2009**, *48*, 7013–7016.

- [51]. Wang, M.; Svatunek, D.; Rohlfing, D.; Liu, Y.; Wang, H.; Giglio, B.; Yuan, H.; Wu, Z.; Li, Z.; Fox, J. *Theranostics* **2016**, *6*, 887–895.
- [52]. Billaud, E. M. F.; Belderbos, S.; Cleeren, F.; Maes, W.; Van De Wouwer, M.; Koole, M.; Geukens, N.; Bormans, G. *Bioconjugate Chem.* **2017**, *28*, 2915–2920.
- [53]. Rossin, R.; van den Bosch, S. M.; ten Hoeve, W.; Versteegen, R. M.; Lub, J.; Robillard, M. S. *Bioconjugate Chem.* **2013**, *24*, 1210–1217.
- [54]. Läppchen, T.; Rossin, R.; van Mourik, T. R.; Gruntz, G.; Hoeben, F. J. M.; Versteegen, R. M.; Janssen, H. M.; Lub, J.; Robillard, M. S. *Nucl. Med. Biol.* **2017**, *55*, 19–26.
- [55]. Bilton, H. A.; Ahmad, Z.; Janzen, N.; Czorny, S.; Valliant, J. F. *J. Vis. Exp.* **2017**, *120*, e55188.
- [56]. Versteegen, R. M.; Rossin, R.; ten Hoeve, W.; Janssen, H. M.; Robillard, M. S. *Angew. Chem. Int. Ed.* **2013**, *52*, 14112–14116.
- [57]. Rossin, R.; Versteegen, R. M.; Wu, J.; Khasanov, A.; Wessels, H. J.; Steenbergen, E. J.; ten Hoeve, W.; Janssen, H. M.; van Onzen, A. H. A. M.; Hudson, P. J.; Robillard, M. S. *Nat. Commun.* **2018**, *9*, 1484.
- [58]. Versteegen, R. M.; ten Hoeve, W.; Rossin, R.; de Geus, M. A. R.; Janssen, H. M.; Robillard, M. S. *Angew. Chem. Int. Ed.* **2018**, *57*, 10494–10499.
- [59]. Mejia Oneto, J. M.; Khan, I.; Seebald, L.; Royzen, M. *ACS Cent. Sci.* **2016**, *2*, 476–482.
- [60]. Zhang, G.; Li, J.; Fan, X.; Liu, Y.; Zheng, S.; Ge, Y.; Chen, P. R. *ACS Cent. Sci.* **2016**, *2*, 325–331.

- [61]. van der Gracht, A. M. F.; de Geus, M. A. R.; Camps, M. G. M.; Ruckwardt, T. J.; Sarris, A. J. C.; Bremmers, J.; Maurits, E.; Pawlak, J. B.; Posthoorn, M. M.; Bongers, K. M.; Filippov, D. V.; Overkleeft, H. S.; Robillard, M. S.; Ossendorp, F.; van Kasteren, S. I. *ACS Chem. Biol.* **2018**, *13*, 1569–1576.
- [62]. Fan, X.; Ge, Y.; Lin, F.; Yang, Y.; Zhang, G.; Ngai, W. S. C.; Lin, Z.; Zheng, S.; Wang, J.; Zhao, J.; Li, J.; R., C. P. *Angew. Chem. Int. Ed.* **2016**, *55*, 14046–14050.
- [63]. Schmidt, P.; Zhou, L.; Tishinov, K.; Zimmerman, K.; Gillingham, D. *Angew. Chem. Int. Ed.* **2014**, *53*, 10928–10931.
- [64]. Dilek, O.; Lei, Z.; Mukherjee, K.; Bane, S. *Chem. Commun.* **2015**, *51*, 16992–16995.
- [65]. Schmidt, P.; Stress, C.; Gillingham, D. *Chem. Sci.* **2015**, *6*, 3329–3333.
- [66]. Meadows, M. K.; Roesner, E. K.; Lynch, V. M.; James, T. D.; Anslyn, E. V. *Org. Lett.* **2017**, *19*, 3179–3182.
- [67]. Bandyopadhyay, A.; Cambray, S.; Gao, J. *J. Am. Chem. Soc.* **2017**, *139*, 871–878.
- [68]. Akgun, B.; Li, C.; Hao, Y.; Lambkin, G.; Derda, R.; Hall, D. G. *J. Am. Chem. Soc.* **2017**, *139*, 14285–14291.
- [69]. Gu, H.; Chio, T. I.; Lei, Z.; Staples, R. J.; Hirschi, J. S.; Bane, S. *Org. Biomol. Chem.* **2017**, *15*, 7543–7548.
- [70]. Bandyopadhyay, A.; Cambray, S.; Gao, J. *Chem. Sci.* **2016**, *7*, 4589–4593.
- [71]. Faustino, H.; Silva, M. J. S. A.; Veiros, L. F.; Bernardes, G. J. L.; Gois, P. M. P. *Chem. Sci.* **2016**, *7*, 5052–5058.
- [72]. Sundhoro, M.; Jeon, S.; Park, J.; Ramström, O.; Yan, M. *Angew. Chem. Int. Ed.* **2017**, *56*, 12117–12121.

- [73]. Meguro, T.; Terashima, N.; Ito, H.; Koike, Y.; Kii, I.; Yoshida, S.; Hosoya, T. *Chem. Commun.* **2018**, *54*, 7904–7907.
- [74]. Saneyoshi, H.; Ochikubo, T.; Mashimo, T.; Hatano, K.; Ito, Y.; Abe, H. *Org. Lett.* **2014**, *16*, 30–33.
- [75]. Luo, J.; Liu, Q.; Morihiro, K.; Deiters, A. *Nat. Chem.* **2016**, *8*, 1027–1034.
- [76]. Lukasak, B.; Morihiro, K.; Deiters, A. *Sci. Rep.* **2019**, *9*, 1470.
- [77]. Ren, G.; Zheng, Q.; Wang, H. *Org. Lett.* **2017**, *19*, 1582–1585.
- [78]. Böhrsch, V.; Serwa, R.; Majkut, P.; Krause, E.; Hackenberger, C. P. R. *Chem. Commun.* **2010**, *46*, 3176–3178.
- [79]. Serwa, R.; Majkut, P.; Horstmann, B.; Swiecicki, J.-M.; Gerrits, M.; Krause, E.; Hackenberger, C. P. R. *Chem. Sci.* **2010**, *1*, 596–602.
- [80]. Nischan, N.; Kasper, M.-A.; Mathew, T.; Hackenberger, C. P. R. *Org. Biomol. Chem.* **2016**, *14*, 7500–7508.
- [81]. Vallée, M. R. J.; Majkut, P.; Wilkening, I.; Weise, C.; Müller, G.; Hackenberger, C. P. R. *Org. Lett.* **2011**, *13*, 5440–5443.
- [82]. Kasper, M.-A.; Glanz, M.; Oder, A.; Schmieder, P.; von Kries, J. P.; Hackenberger, C. P. R. *Chem. Sci.* **2019**, *10*, 6322–6329.
- [83]. Kasper, M.-A.; Glanz, M.; Stengl, A.; Penkert, M.; Klenk, S.; Sauer, T.; Schumacher, D.; Helma, J.; Krause, E.; Cardoso, M. C.; Leonhardt, H.; Hackenberger, C. P. R. *Angew. Chem. Int. Ed.* **2019**, *58*, 11625–11630.
- [84]. Lee, Y.-J.; Kurra, Y.; Liu, W. R. *ChemBioChem* **2016**, *17*, 456–461.
- [85]. Bos, J.; Muir, T. W. *J. Am. Chem. Soc.* **2018**, *140*, 4757–4760.
- [86]. Shih, H.-W.; Prescher, J. A. *J. Am. Chem. Soc.* **2015**, *137*, 10036–10039.

- [87]. Row, R. D.; Shih, H.-W.; Alexander, A. T.; Mehl, R. A.; Prescher, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 7370–7375.
- [88]. Row, R. D.; Prescher, J. A. *Org. Lett.* **2018**, *20*, 5614–5617.
- [89]. Heiss, T. K.; Prescher, J. A. *J. Org. Chem.* **2019**, *84*, 7443–7448.
- [90]. Newberry, R. W.; VanVeller, B.; Guzei, I. A.; Raines, R. T. *J. Am. Chem. Soc.* **2013**, *135*, 7843–7846.
- [91]. Walters, C. R.; Szantai-Kas, D. M.; Zhang, Y.; Reinert, Z. E.; Horne, S. W.; Chenoweth, D. M.; Petersson, E. J. *Chem. Sci.* **2017**, *8*, 2868–2877.
- [92]. McKay, C. S.; Moran, J.; Pezacki, J. P. *Chem. Commun.* **2010**, *46*, 931–933.
- [93]. McKay, C. S.; Chigrinova, M.; Blake, J. A.; Pezacki, J. P. *Org. Biomol. Chem.* **2012**, *10*, 3066–3070.
- [94]. MacKenzie, D. A.; Pezacki, J. P. *Can. J. Chem.* **2014**, *92*, 337–340.
- [95]. Gunawardene, P. N.; Luo, W.; Polgar, A. M.; Corrigan, J. F.; Workentin, M. S. *Org. Lett.* **2019**, *21*, 5547–5551.
- [96]. McKay, C. S.; Blake, J. A.; Cheng, J.; Danielson, D. C.; Pezacki, J. P. *Chem. Commun.* **2011**, *47*, 10040–10042.
- [97]. Sherratt, A. R.; Chigrinova, M.; McKay, C. S.; Beaulieu, L.-P. B.; Rouleau, Y.; Pezacki, J. P. *RSC Adv.* **2014**, *4*, 46966–46969.
- [98]. Sherratt, A. R.; Chigrinova, M.; MacKenzie, D. A.; Rastogi, N. K.; Ouattara, M. T. M.; Pezacki, A. T.; Pezacki, J. P. *Bioconjugate Chem.* **2016**, *27*, 1222–1226.
- [99]. Decuypère, E.; Plougastel, L.; Audisio, D.; Taran, F. *Chem. Commun.* **2017**, *53*, 11515–11527.

- [100]. Kolodych, S.; Rasolofonjatovo, E.; Chaumontet, M.; Nevers, M.-C.; Créminon, C.; Taran, F. *Angew. Chem. Int. Ed.* **2013**, *52*, 12056–12060.
- [101]. Wallace, S.; Chin, J. W. *Chem. Sci.* **2014**, *5*, 1742–1744.
- [102]. Plougastel, L.; Koniev, O.; Specklin, S.; Decuypere, E.; Créminon, C.; Buisson, D.-A.; Wagner, A.; Kolodych, S.; Taran, F. *Chem. Commun.* **2014**, *50*, 9376–9378.
- [103]. Liu, H.; Audisio, D.; Plougastel, L.; Decuypere, E.; Buisson, D.-A.; Koniev, O.; Kolodych, S.; Wagner, A.; Elhabiri, M.; Krzyczmonik, A.; Forsback, S.; Solin, O.; Gouverneur, V.; Taran, F. *Angew. Chem. Int. Ed.* **2016**, *55*, 12073–12077.
- [104]. Mix, K. A.; Aronoff, M. R.; Raines, R. T. *ACS Chem. Biol.* **2016**, *11*, 3233–3244.
- [105]. Moran, J.; McKay, C. S.; Pezacki, J. P. *Can. J. Chem.* **2011**, *89*, 148–151.
- [106]. Sanders, B. C.; Friscourt, F. e. e.; Ledin, P. A.; Mbua, N. E.; Arumugam, S.; Guo, J.; Boltje, T. J.; Popik, V. V.; Boons, G.-J. *J. Am. Chem. Soc.* **2011**, *133*, 949–957.
- [107]. Josa-Culleré, L.; Wainman, Y. A.; Brindle, K. M.; Leeper, F. J. *RSC Adv.* **2014**, *4*, 52241–52244.
- [108]. Anderson, K. A.; Aronoff, M. R.; McGrath, N. A.; Raines, R. T. *J. Am. Chem. Soc.* **2015**, *137*, 2412–2415.
- [109]. McGrath, N. A.; Raines, R. T. *Chem. Sci.* **2012**, *3*, 3237–3240.
- [110]. Aronoff, M. R.; Gold, B.; Raines, R. T. *Org. Lett.* **2016**, *18*, 1538–1541.
- [111]. Lim, R. K. V.; Lin, Q. *Chem. Commun.* **2010**, *46*, 7993–7995.
- [112]. King, M.; Wagner, A. *Bioconjugate Chem.* **2014**, *25*, 825–839.
- [113]. Schäfer, R. J. B.; Monaco, M. R.; Li, M.; Tirla, A.; Rivera-Fuentes, P.; Wennemers, H. *J. Am. Chem. Soc.* **2019**, *141*, 18644–18648.
- [114]. Yu, Z.; Lim, R. K. V.; Lin, Q. *Chem. Eur. J.* **2010**, *16*, 13325–13329.

- [115]. Yu, Z.; Pan, Y.; Wang, Z.; Wang, J.; Lin, Q. *Angew. Chem. Int. Ed.* **2012**, *51*, 10600–10604.
- [116]. Kamber, D. N.; Nazarova, L. A.; Liang, Y.; Lopez, S. A.; Patterson, D. M.; Shih, H.-W.; Houk, K. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2013**, *135*, 13680–13683.
- [117]. An, P.; Yu, Z.; Lin, Q. *Chem. Commun.* **2013**, *49*, 9920–9922.
- [118]. Yu, Z.; Ohulchansky, T. Y.; An, P.; Prasad, P. N.; Lin, Q. *J. Am. Chem. Soc.* **2013**, *135*, 16766–16769.
- [119]. Tasdelen, M. A.; Yagci, Y. *Angew. Chem. Int. Ed.* **2013**, *52*, 5930–5938.
- [120]. Li, J.; Kong, H.; Huang, L.; Cheng, B.; Qin, K.; Zheng, M.; Yan, Z.; Zhang, Y. *J. Am. Chem. Soc.* **2018**, *140*, 14542–14546.
- [121]. Zhang, X.; Wu, X.; Jiang, S.; Gao, J.; Yao, Z.; Deng, J.; Zhang, L.; Yu, Z. *Chem. Commun.* **2019**, *55*, 7187–7190.
- [122]. Lim, R. K. V.; Lin, Q. *Acc. Chem. Res.* **2011**, *44*, 828–839.
- [123]. An, P.; Lewandowski, T. M.; Erbay, T. G.; Liu, P.; Lin, Q. *J. Am. Chem. Soc.* **2018**, *140*, 4860–4868.
- [124]. de Almeida, G.; Sletten, E. M.; Nakamura, H.; Palaniappan, K. K.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2012**, *51*, 2443–2447.
- [125]. King, M.; Baati, R.; Wagner, A. *Chem. Commun.* **2012**, *48*, 9308–9309.
- [126]. de Almeida, G.; Townsend, L. C.; Bertozzi, C. R. *Org. Lett.* **2013**, *15*, 3038–3041.
- [127]. Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G.-J. *Angew. Chem. Int. Ed.* **2008**, *47*, 2253–2255.
- [128]. Stöckmann, H.; Neves, A. A.; Stairs, S.; Ireland-Zecchini, H.; Brindle, K. M.; Leeper, F. J. *Chem. Sci.* **2011**, *2*, 932–936.

- [129]. Friscourt, F.; Ledin, P. A.; Mbuja, N. E.; Flanagan-Steet, H. R.; Wolfert, M. A.; Steet, R.; Boons, G.-J. *J. Am. Chem. Soc.* **2012**, *134*, 5381–5389.
- [130]. Varga, B. R.; Kállay, M.; Hegyi, K.; Béni, S.; Kele, P. *Chem. Eur. J.* **2012**, *18*, 822–828.
- [131]. Dommerholt, J.; Schmidt, S.; Temming, R.; Hendriks, L. J. A.; Rutjes, F. P. J. T.; van Hest, J. C. M.; Lefebber, D. J.; Friedl, P.; van Delft, F. L. *Angew. Chem. Int. Ed.* **2010**, *49*, 9422–9425.
- [132]. Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem. Biol.* **2006**, *1*, 644–648.
- [133]. Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 16793–16797.
- [134]. Sletten, E. M.; Nakamura, H.; Jewett, J. C.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 11799–11805.
- [135]. Li, W.; Zou, J.; Zhu, S.; Mao, X.; Tian, H.; Wang, X. *Chem. Eur. J.* **2019**, *25*, 10328–10332.
- [136]. Sletten, E. M.; Bertozzi, C. R. *Org. Lett.* **2008**, *10*, 3097–3099.
- [137]. Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 3688–3690.
- [138]. Debets, M. F.; van Berkel, S. S.; Schoffelen, S.; Rutjes, F. P. J. T.; van Hest, J. C. M.; van Delft, F. L. *Chem. Commun.* **2010**, *46*, 97–99.
- [139]. Gold, B.; Shevchenko, N. E.; Bonus, N.; Dudley, G. B.; Alabugin, I. V. *J. Org. Chem.* **2012**, *77*, 75–89.

- [140]. Gold, B.; Batsomboon, P.; Dudley, G. B.; Alabugin, I. V. *J. Org. Chem.* **2014**, *79*, 6221–6232.
- [141]. Kaneda, K.; Naruse, R.; Yamamoto, S. *Org. Lett.* **2017**, *19*, 1096–1099.
- [142]. Burke, E. G.; Gold, B.; Hoang, T. T.; Raines, R. T.; Schomaker, J. M. *J. Am. Chem. Soc.* **2017**, *139*, 8029–8037.
- [143]. Ni, R.; Mitsuda, N.; Kashiwagi, T.; Igawa, K.; Tomooka, K. *Angew. Chem. Int. Ed.* **2015**, *54*, 1190–1194.
- [144]. Del Grosso, A.; Galanopoulos, L.-D.; Chiu, C. K. C.; Clarkson, G. J.; O'Connor, P. B.; Wills, M. *Org. Biomol. Chem.* **2017**, *15*, 4517–4521.
- [145]. Mistry, A.; Knighton, R. C.; Forshaw, S.; Dualeh, Z.; Parker, J. S.; Wills, M. *Org. Biomol. Chem.* **2018**, *16*, 8965–8975.
- [146]. Harris, T.; dos Passos Gomes, G.; Ayad, S.; Clark, R. J.; Lobodin, V. V.; Tuscan, M.; Hanson, K.; Alabugin, I. V. *Chem* **2017**, *3*, 629–640.
- [147]. Meier, H.; Schuh-Popitz, C.; Peiersen, H. *Angew. Chem. Int. Ed.* **1981**, *20*, 270–271.
- [148]. Ho, S. H.; Tirrell, D. A. *ACS Cent. Sci.* **2019**, *5*, 1911–1919.
- [149]. Liu, F.; Liang, Y.; Houk, K. N. *Acc. Chem. Res.* **2017**, *50*, 2297–2308.
- [150]. Bickelhaupt, F. M.; Houk, K. N. *Angew. Chem. Int. Ed.* **2017**, *56*, 10070–10086.
- [151]. Gordon, C. G.; Mackey, J. L.; Jewett, J. C.; Sletten, E. M.; Houk, K. N.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2012**, *134*, 9199–9208.
- [152]. Tao, H.; Liu, F.; Zeng, R.; Shao, Z.; Zou, L.; Cao, Y.; Murphy, J. M.; Houk, K. N.; Liang, Y. *Chem. Commun.* **2018**, *54*, 5082–5085.

- [153]. Liu, F.; Paton, R. S.; Kim, S.; Liang, Y.; Houk, K. N. *J. Am. Chem. Soc.* **2013**, *135*, 15642-15649.
- [154]. Liu, F.; Liang, Y.; Houk, K. N. *J. Am. Chem. Soc.* **2014**, *136*, 11483–11493.
- [155]. Darko, A.; Wallace, S.; Dmitrenko, O.; Machovina, M. M.; Mehl, R. A.; Chin, J. W.; Fox, J. M. *Chem. Sci.* **2014**, *5*, 3770–3776.
- [156]. Kozma, E.; Nikić, I.; Varga, B. R.; Aramburu, I. V.; Kang, J. H.; Fackler, O. T.; Lemke, E. A.; Kele, P. *ChemBioChem* **2016**, *17*, 1518–1524.
- [157]. Knorr, G.; Kozma, E.; Herner, A.; Lemke, E. A.; Kele, P. *Chem. Eur. J.* **2016**, *22*, 8972–8979.
- [158]. Lambert, W. D.; Scinto, S. L.; Dmitrenko, O.; Boyd, S. J.; Magboo, R.; Mehl, R. A.; Chin, J. W.; Fox, J. M.; Wallace, S. *Org. Biomol. Chem.* **2017**, *15*, 6640–6644.
- [159]. Siegl, S. J.; Vázquez, A.; Dzijak, R.; Dračinský, M.; Galeta, J.; Rampmaier, R.; Klepetářová, B.; Vrabel, M. *Chem. Eur. J.* **2018**, *24*, 2426–2432.
- [160]. Bach, R. D. *J. Am. Chem. Soc.* **2009**, *131*, 5233–5243.
- [161]. Santucci, J., III; Sanzone, J. R.; Woerpel, K. A. *Eur. J. Org. Chem.* **2016**, *2016*, 2933–2943.
- [162]. Fang, Y.; Zhang, H.; Huang, Z.; Scinto, S. L.; Yang, J. C.; am Ende, C.; Dmitrenko, O.; Johnson, D. S.; Fox, J. M. *Chem. Sci.* **2018**, *9*, 1953–1963.
- [163]. Patterson, D. M.; Nazarova, L. A.; Xie, B.; Kamber, D. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2012**, *134*, 18638–18643.
- [164]. Yang, J.; Šečkutè, J.; Cole, C. M.; Devaraj, N. K. *Angew. Chem. Int. Ed.* **2012**, *51*, 7476–7479.

- [165]. Ravasco, J. M. J. M.; Monteiro, C. M.; Trindade, A. F. *Org. Chem. Front.* **2017**, *4*, 1167–1198.
- [166]. Elliot, T. S.; Townsley, F. M.; Bianco, A.; Ernst, R. J.; Sachdeva, A.; Elsässer, S. J.; Davis, L.; Lang, K.; Pisa, R.; Greiss, S.; Lilley, K. S.; Chin, J. W. *Nat. Biotechnol.* **2014**, *32*, 465–472.
- [167]. Yu, Z.; Lin, Q. *J. Am. Chem. Soc.* **2014**, *136*, 4153–4156.
- [168]. Ramil, C. P.; Dong, M.; An, P.; Lewandowski, T. M.; Yu, Z.; Miller, L. J.; Lin, Q. *J. Am. Chem. Soc.* **2017**, *139*, 13376–13386.
- [169]. An, P.; Wu, H.-Y.; Lewandowski, T. M.; Lin, Q. *Chem. Commun.* **2018**, *54*, 14005–14008.
- [170]. Kumar, P.; Jiang, T.; Li, S.; Zainul, O.; Laughlin, S. T. *Org. Biomol. Chem.* **2018**, *16*, 4081–4085.
- [171]. Kumar, P.; Zainul, O.; Camarda, F. M.; Jiang, T.; Mannone, J. A.; Huang, W.; Laughlin, S. T. *Org. Lett.* **2019**, *21*, 3721–3725.
- [172]. Tu, J.; Xu, M.; Parvez, S.; Peterson, R. T.; Franzini, R. M. *J. Am. Chem. Soc.* **2018**, *140*, 8410–8414.
- [173]. Tu, J.; Svatunek, D.; Parvez, S.; Eckvahl, H. J.; Xu, M.; Peterson, R. T.; Houk, K. N.; Franzini, R. M. *Chem. Sci.* **2020**, *11*, 169–179.
- [174]. Xu, M.; Deb, T.; Tu, J.; Franzini, R. M. *J. Org. Chem.* **2019**, *84*, 15520–15529.
- [175]. Stöckmann, H.; Neves, A. A.; Stairs, S.; Brindle, K. M.; Leeper, F. J. *Org. Biomol. Chem.* **2011**, *9*, 7303–7305.
- [176]. Chen, Y.; Wu, K.-L.; Tang, J.; Loredó, A.; Clements, J.; Pei, J.; Peng, Z.; Gupta, R.; Fang, X.; Xiao, H. *ACS Chem. Biol.* **2019**, *14*, 2793–2799.

- [177]. Stairs, S.; Neves, A. A.; Stöckmann, H.; Wainman, Y. A.; Ireland-Zecchini, H.; Brindle, K. M.; Leeper, F. J. *ChemBioChem* **2013**, *14*, 1063–1067.
- [178]. Niederwieser, A.; Späte, A.-K.; Nguyen, L. D.; Jüngst, C.; Reutter, W.; Wittmann, V. *Angew. Chem. Int. Ed.* **2013**, *52*, 4265–4268.
- [179]. Lee, Y.-J.; Kurra, Y.; Yang, Y.; Torres-Kolbus, J.; Deiters, A.; Liu, W. R. *Chem. Commun.* **2014**, *50*, 13085–13088.
- [180]. Eising, S.; Lelivelt, F.; Bongers, K. M. *Angew. Chem. Int. Ed.* **2016**, *55*, 12243–12247.
- [181]. Eising, S.; Xin, B.-T.; Kleinpenning, F.; Heming, J. J. A.; Florea, B. I.; Overkleeft, H. S.; Bongers, K. M. *ChemBioChem* **2018**, *19*, 1648–1652.
- [182]. Eising, S.; Engwerda, A. H. J.; Riedijk, X.; Bickelhaupt, F. M.; Bongers, K. M. *Bioconjugate Chem.* **2018**, *29*, 3054–3059.
- [183]. Eising, S.; van der Linden, N. G. A.; Kleinpenning, F.; Bongers, K. M. *Bioconjugate Chem.* **2018**, *29*, 982–986.
- [184]. Engelsma, S. B.; Willems, L. I.; van Paaschen, C. E.; van Kasteren, S. I.; van der Marel, G. A.; Overkleeft, H. S.; Filippov, D. V. *Org. Lett.* **2014**, *16*, 2744–2747.
- [185]. de Montes, E. G.; Jiménez-Moreno, E.; Oliveira, B. L.; Navo, C. D.; Cal, P. M. S. D.; Jiménez-Osés, G.; Robina, I.; Moreno-Vargas, A. J.; Bernardes, G. J. L. *Chem. Sci.* **2019**, *10*, 4515–4522.
- [186]. Oliveira, B. L.; Guo, Z.; Boutureira, O.; Guerreiro, A.; Jiménez-Osés, G.; Bernardes, G. J. L. *Angew. Chem. Int. Ed.* **2016**, *55*, 14683–14687.
- [187]. Oliveira, B. L.; Guo, Z.; Bernardes, G. J. L. *Chem. Soc. Rev.* **2017**, *46*, 4811–5174.

- [188]. Chen, W.; Wang, D.; Dai, C.; Hamelberg, D.; Wang, B. *Chem. Commun.* **2012**, 48, 1736–1738.
- [189]. Wang, D.; Chen, W.; Zheng, Y.; Dai, C.; Wang, K.; Ke, B.; Wang, B. *Org. Biomol. Chem.* **2014**, 12, 3950–3955.
- [190]. Blizzard, R. J.; Backus, D. R.; Brown, W.; Bazewicz, C. G.; Li, Y.; Mehl, R. A. *J. Am. Chem. Soc.* **2015**, 137, 10044–10047.
- [191]. Lambert, W. D.; Fang, Y.; Mahapatra, S.; Huang, Z.; am Ende, C. W.; Fox, J. M. *J. Am. Chem. Soc.* **2019**, 141, 17068–17074.
- [192]. Šečkutė, J.; Devaraj, N. K. *Curr. Opin. Chem. Biol.* **2013**, 17, 761–767.
- [193]. Zhang, H.; Trout, W. S.; Liu, S.; Andrade, G. A.; Hudson, D. A.; Scinto, S. L.; Dicker, K. T.; Li, Y.; Lazouski, N.; Rosenthal, J.; Thorpe, C.; Jia, X.; Fox, J. M. *J. Am. Chem. Soc.* **2016**, 138, 5978–5983.
- [194]. Ehret, F.; Wu, H.; Alexander, S. C.; Devaraj, N. K. *J. Am. Chem. Soc.* **2015**, 137, 8876–8879.
- [195]. Kamber, D. N.; Liang, Y.; Blizzard, R. J.; Liu, F.; Mehl, R. A.; Houk, K. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2015**, 137, 8388–8391.
- [196]. Horner, K. A.; Valette, N. M.; Webb, M. E. *Chem. Eur. J.* **2015**, 21, 14376–14381.
- [197]. Baalman, M.; Ziegler, M. J.; Werther, P.; Wilhelm, J.; Wombacher, R. *Bioconjugate Chem.* **2019**, 30, 1405–1414.
- [198]. Peewasan, K.; Wagenknecht, H.-A. *ChemBioChem* **2017**, 18, 1473–1476.
- [199]. Reisacher, U.; Groitl, B.; Strasser, R.; Cserép, G. B.; Kele, P.; Wagenknecht, H.-A. *Bioconjugate Chem.* **2019**, 30, 1773–1780.

- [200]. Siegl, S. J.; Dzijak, R.; Vázquez, A.; Pohl, R.; Vrabel, M. *Chem. Sci.* **2017**, *8*, 3593–3598.
- [201]. Kozhevnikov, V. N.; Deary, M. E.; Mantso, T.; Panayiotidis, M. I.; Sims, M. T. *Chem. Commun.* **2019**, *55*, 14283–14286.
- [202]. Kamber, D. N.; Nguyen, S. S.; Liu, F.; Briggs, J. S.; Shih, H.-W.; Row, R. D.; Long, Z. G.; Houk, K. N.; Liang, Y.; Prescher, J. A. *Chem. Sci.* **2019**, *10*, 9109–9114.
- [203]. Levandowski, B. J.; Abularrage, N. S.; Houk, K. N.; Raines, R. T. *Org. Lett.* **2019**, *21*, 8492–8495.
- [204]. Bruins, J. J.; Albada, B.; van Delft, F. *Chem. Eur. J.* **2018**, *24*, 4749–4756.
- [205]. Borrmann, A.; Fatunsin, O.; Dommerholt, J.; Jonker, A. M.; Löwik, D. W. P. M.; van Hest, J. C. M.; van Delft, F. L. *Bioconjugate Chem.* **2015**, *26*, 257–261.
- [206]. Bruins, J. J.; Westphal, A. H.; Albada, B.; Wagner, K.; Bartels, L.; Spits, H.; van Berkel, W. J. H.; van Delft, F. L. *Bioconjugate Chem.* **2017**, *28*, 1189–1193.
- [207]. Gahtory, D.; Sen, R.; Kuzmyn, A. R.; Escorihuela, J.; Zuilhof, H. *Angew. Chem. Int. Ed.* **2018**, *57*, 10118–10122.
- [208]. Li, Q.; Dong, T.; Liu, X.; Lei, X. *J. Am. Chem. Soc.* **2013**, *135*, 4996–4999.
- [209]. Minard, A.; Liano, D.; Wang, X.; Antonio, M. D. *Bioorg. Med. Chem.* **2019**, *27*, 2298–2305.
- [210]. Shih, H.-W.; Kamber, D. N.; Prescher, J. A. *Curr. Opin. Chem. Biol.* **2014**, *21*, 103–111.
- [211]. Chang, P. V.; Prescher, J. A.; Hangauer, M. J.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2007**, *129*, 8400–8401.

- [212]. Karver, M. R.; Weissleder, R.; Hilderbrand, S. A. *Angew. Chem. Int. Ed.* **2012**, *51*, 920–922.
- [213]. Schoch, J.; Staudt, M.; Samanta, A.; Wiessler, M.; Jäschke, A. *Bioconjugate Chem.* **2012**, *23*, 1382–1386.
- [214]. Hudak, J. E.; Alvarez, D.; Skelly, A.; von Adrian, U. H.; Kasper, D. L. *Nat. Microbiol.* **2017**, *2*, 17099.
- [215]. Zhang, X.; Dong, T.; Li, Q.; Liu, X.; Li, L.; Chen, S.; Lei, X. *ACS Chem. Biol.* **2015**, *10*, 1676–1683.
- [216]. Narayanam, M. K.; Liang, Y.; Houk, K. N.; Murphy, J. M. *Chem. Sci.* **2016**, *7*, 1257–1261.
- [217]. Shao, Z.; Liu, W.; Tao, H.; Liu, F.; Zeng, R.; Champagne, P. A.; Cao, Y.; Houk, K. N.; Liang, Y. *Chem. Commun.* **2018**, *54*, 14089–14092.
- [218]. Winz, M.-L.; Linder, E. C.; Becker, J.; Jäschke, A. *Chem. Commun.* **2018**, *54*, 11781–11784.
- [219]. Simon, C.; Lion, C.; Spriet, C.; Baldacci-Cresp, F.; Hawkins, S.; Biot, C. *Angew. Chem. Int. Ed.* **2018**, *57*, 16665–16671.
- [220]. Schart, V. F.; Hassenrück, J.; Späte, A.-K.; Dold, J. E. G. A.; Fahrner, R.; Wittmann, V. *ChemBioChem* **2019**, *20*, 166–171.
- [221]. Sletten, E. M.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2011**, *133*, 17570–17573.
- [222]. Italia, J. S.; Addy, P. S.; Erickson, S. B.; Peeler, J. C.; Weerapana, E.; Chatterjee, A. *J. Am. Chem. Soc.* **2019**, *141*, 6204–6212.

- [223]. Tu, J.; Svatunek, D.; Parvez, S.; Liu, A. C.; Levandowski, B. J.; Eckvahl, H. J.; Peterson, R. T.; Houk, K. N.; Franzini, R. M. *Angew. Chem. Int. Ed.* **2019**, *58*, 9043–9048.
- [224]. Kim, J.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2015**, *54*, 15777–15781.
- [225]. Dong, J.; Krasnova, L.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2014**, *53*, 9430–9448.
- [226]. Barrow, A. S.; Smedley, C. J.; Zheng, Q.; Li, S.; Dong, J.; Moses, J. E. *Chem. Soc. Rev.* **2019**, *48*, 4731–4758.
- [227]. Zheng, Q.; Woehl, J. L.; Kitamura, S.; Santos-Martins, D.; Smedley, C. J.; Li, G.; Forli, S.; Moses, J. E.; Wolan, D. W.; Sharpless, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 18808–18814.
- [228]. Mortenson, D. E.; Brighty, G. J.; Plate, L.; Bare, G.; Chen, W.; Li, S.; Wang, H.; Cravatt, B. F.; Forli, S.; Powers, E. T.; Sharpless, K. B.; Wilson, I. A.; Kelly, J. W. *J. Am. Chem. Soc.* **2018**, *140*, 200–210.
- [229]. Shannon, A.; Gu, C.; McLaughlin, C. J.; Kaiser, M.; van der Hoorn, R. A. L.; Weerapana, E. *ChemBioChem* **2012**, *13*, 2327–2330.
- [230]. Hett, E. C.; Xu, H.; Geoghegan, K. F.; Gopalsamy, A.; Kyne, R. E., Jr.; Menard, C. A.; Narayanan, A.; Parikh, M. D.; Liu, S.; Roberts, L.; Robinson, R. P.; Tones, M. A.; Jones, L. H. *ACS Chem. Biol.* **2015**, *10*, 1094–1098.
- [231]. Wang, N.; Yang, B.; Fu, C.; Zhu, H.; Zheng, F.; Kobayashi, T.; Liu, J.; Li, S.; Ma, C.; Wang, P. G.; Wang, Q.; Wang*, L. *J. Am. Chem. Soc.* **2018**, *140*, 4995–4999.
- [232]. Yang, B.; Wu, H.; Schnier, P. D.; Liu, Y.; Liu, J.; Wang, N.; DeGrado, W. F.; Wang, L. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 11162–11167.

- [233]. Meng, G.; Guo, T.; Ma, T.; Zhang, J.; Shen, Y.; Sharpless, K. B.; Dong, J. *Nature* **2019**, *574*, 86–89.
- [234]. Kabakoff, D. S.; Bünzil, J.-C. G.; Oth, J. F. M.; Hammond, W. B.; Berson, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1510–1512.
- [235]. Tomlin, F. M.; Gordon, C. G.; Han, Y.; Wu, T. S.; Sletten, E. M.; Bertozzi, C. R. *Bioorg. Med. Chem.* **2018**, *26*, 5280–5290.
- [236]. Martínez-Calvo, M.; Mascareñas, J. L. *Coord. Chem. Rev.* **2018**, *359*, 57–79.
- [237]. Jang, S.-Y.; Murale, D. P.; Kim, A. D.; Lee, J.-S. *ChemBioChem* **2019**, *20*, 1498–1507.
- [238]. Vidal, C.; Tomás-Gamasa, M.; Destito, P.; López, F.; Mascareñas, J. L. *Nat. Commun.* **2018**, *9*, 1913.
- [239]. Pérez-López, A. M.; Rubio-Ruiz, B.; Sebastián, V.; Hamilton, L.; Adam, C.; Bray, T. L.; Irusta, S.; Brennan, P. M.; Lloyd-Jones, G. C.; Sieger, D.; Santamaría, J.; Unciti-Broceta, A. *Angew. Chem. Int. Ed.* **2017**, *56*, 12548–12552.
- [240]. Tsubokura, K.; Vong, K. K. H.; Pradipta, A. R.; Ogura, A.; Urano, S.; Tahara, T.; Nozaki, S.; Onoe, H.; Nakao, Y.; Sibgatullina, R.; Kurbangalieva, A.; Watanabe, Y.; Tanaka, K. *Angew. Chem. Int. Ed.* **2017**, *56*, 3579–3584.
- [241]. Chalker, J. M.; Wood, C. S. C.; Davis, B. G. *J. Am. Chem. Soc.* **2009**, *131*, 16346–16347.
- [242]. Clavadetscher, J.; Indrigo, E.; Chankeshwara, S. V.; Lilienkampf, A.; Bradley, M. *Angew. Chem. Int. Ed.* **2017**, *56*, 6864–6868.
- [243]. Li, N.; Lim, R. K. V.; Edwardraja, S.; Lin, Q. *J. Am. Chem. Soc.* **2011**, *133*, 15316–15319.

- [244]. Li, N.; Ramil, C. P.; Lim, R. K. V.; Lin, Q. *ACS Chem. Biol.* **2015**, *10*, 379–384.
- [245]. Li, J.; Yu, J.; Zhao, J.; Wang, J.; Zheng, S.; Lin, S.; Chen, L.; Yang, M.; Jia, S.; Zhang, X.; Chen, P. R. *Nat. Chem.* **2014**, *6*, 352–361.
- [246]. Wang, J.; Cheng, B.; Li, J.; Zhang, Z.; Hong, Y.; Chen, X.; Chen, P. R. *Angew. Chem. Int. Ed.* **2015**, *54*, 5364–5368.
- [247]. Yusop, R. M.; Unciti-Broceta, A.; Johansson, E. M. V.; Sánchez-Martín, R. M.; Bradley, M. *Nat. Chem.* **2011**, *3*, 239–243.
- [248]. Wang, F.; Zhang, Y.; Du, Z.; Ren, J.; Qu, X. *Nat. Commun.* **2018**, *9*, 1209.
- [249]. Tomás-Gamasa, M.; Martínez-Calvo, M.; Couceiro, J. R.; Mascareñas, J. L. *Nat. Commun.* **2016**, *7*, 12538.
- [250]. Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642–9643.
- [251]. Lin, Y. A.; Boutureira, O.; Lercher, L.; Bhushan, B.; Paton, R. S.; Davis, B. G. *J. Am. Chem. Soc.* **2013**, *135*, 12156–12159.
- [252]. Bloom, S.; Liu, C.; Kölmel, D. K.; Qiao, J. X.; Zhang, Y.; Poss, M. A.; Ewing, W. R.; MacMillan, D. W. C. *Nat. Chem.* **2018**, *10*, 205–211.
- [253]. Hooker, J. M.; Kovacs, E. W.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 3718–3719.
- [254]. Sengupta, S.; Chandrasekaran, S. *Org. Biomol. Chem.* **2019**, *17*, 8308–8329.
- [255]. Addy, P. S.; Erickson, S. B.; Italia, J. S.; Chatterjee, A. *J. Am. Chem. Soc.* **2017**, *139*, 11670–11673.
- [256]. Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Toste, F. D.; Chang, C. J. *Science* **2017**, *355*, 597–602.

- [257]. Taylor, M. T.; Nelson, J. E.; Suero, M.; Gaunt, M. J. *Nature* **2018**, *562*, 563–568.
- [258]. Alvarez-Dorta, D.; Thobie-Gautier, C.; Croyal, M.; Bouzelha, M.; Mével, M.; Deniaud, D.; Boujtita, M.; Gouin, S. G. *J. Am. Chem. Soc.* **2018**, *140*, 17120–17126.
- [259]. Geng, J.; Li, W.; Zhang, Y.; Thottappillil, N.; Clavadetscher, J.; Lilienkampf, A.; Bradley, M. *Nat. Chem.* **2019**, *11*, 578–586.
- [260]. Vidal, C.; Tomás-Gamasa, M.; Gutiérrez-González, A.; Mascareñas, J. L. *J. Am. Chem. Soc.* **2019**, *141*, 5125–5129.

Chapter 2: Driving new applications of bioorthogonal 1,2,4-triazines

Adapted from published work, with permission from the Royal Society of Chemistry: Kamber, D. N.; Nguyen, S. S.; Liu, F.; Briggs, J. S.; Shih, H.-W.; Row, R. D.; Long, Z. G.; Houk, K. N.; Liang, Y.; Prescher, J. A. *Chem. Sci.* **2019**, *10*, 9109–9114.

2.1 Abstract

As discussed in Chapter 1, the size and stability of a bioorthogonal reporter is a key consideration when choosing the right chemistry for a given application. Some of the most popular reagents are sterically demanding and have limited stability in cellular environments. Bioorthogonal chemistry has also largely been limited to studying single targets, owing to cross-reactivity between probes. In this Chapter, I discuss how the bioorthogonal 1,2,4-triazines can fill both of these voids.

2.2 Introduction

Interactions with host cells govern the pathology of many infectious agents, including the causative agents of botulism and cholera [1-2]. The initial stages of these infections, along with pathogen persistence, are driven by the secretion of virulence factors into the host cell [3]. Proteins and small peptides comprise a large fraction of the secreted effectors, and these biomolecules can disrupt normal cell function and promote infection [4]. While effector proteins have been identified from several organisms, a complete inventory of such factors remains unknown in most cases [5-6]. The interactions of secreted factors with host biomolecules, and the timing of these events, are also poorly understood. This is, in part, due to the lack of general molecular tools for tracking large collections of biomolecules in their native habitat.

In recent years, the bioorthogonal chemical reporter strategy has emerged as a powerful tool to visualize and identify secreted biomolecules in complex environments (Figure 2-1) [7]. In this strategy, target biomolecules can be labeled with building blocks containing unique chemical motifs using the cell's metabolic machinery [8-11]. The reporters must be small enough to maneuver through biosynthetic pathways, and stable enough to avoid degradation in the cell. Once incorporated, the reporters can be reacted with a second probe to label the target for visualization or isolation. Like the reporters, the secondary probes must also be inert to the cellular environment, and react only with their respective coupling partner. The bioorthogonal chemical reporter strategy has been widely used to visualize and profile entire proteomes, including those secreted from pathogens [12-13]. For example, Tirrell and coworkers used azidohomoalanine, homopropargylglycine, and azidonorleucine to examine proteins secreted by *Yersinia enterocolitica* and other pathogens [14-16]. These noncanonical amino acid reporters serve as surrogates for methionine, and are well tolerated by native enzymes because they closely resemble the native substrate. The labeled microbial proteins can then be detected using copper-catalyzed “click” chemistry.

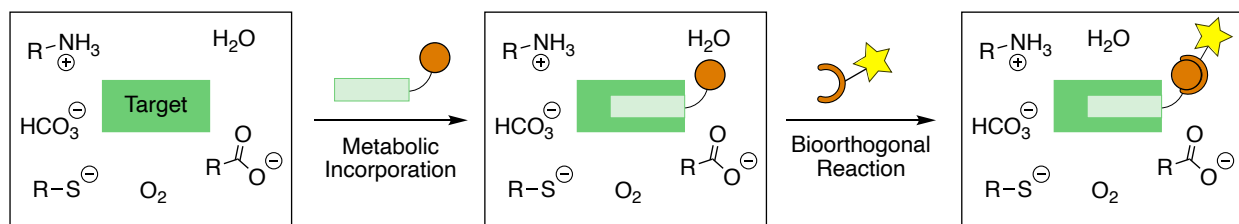


Figure 2-1. Overall scheme for the bioorthogonal chemical reporter strategy. A building block reporter (e.g. a noncanonical amino acid) is incorporated into a biomolecule target (e.g. a protein) via metabolism. The reporter can be detected using a bioorthogonal reaction.

While azido and alkynyl amino acids have been widely employed in proteome labeling, they are not without limitations. Azides are prone to degrade to amines in the highly reducing conditions of the cell [7]. Terminal alkynes require cytotoxic copper catalysts for bioorthogonal detection, limiting their practicality in living systems [17]. Although strained alkynes mitigate this requirement, they are susceptible to thiol-mediated degradation, and are too sterically demanding for use in native biological pathways [18-20]. Having access to additional small, stable noncanonical amino acids would be useful for protein labeling applications. Probes that exhibit unique modes of reactivity would also be beneficial for studies involving multiple secreted epitopes.

In 2015, the Prescher lab showed that 1,2,4-triazines are viable scaffolds for the bioorthogonal chemical reporter strategy. These motifs react robustly with strained π -systems, such as *trans*-cyclooctene (TCO, Figure 2-2A). Triazines are remarkably stable in the presence of thiols, a requirement for cellular use [21-22]. This observation is in stark contrast to the 1,2,4,5-tetrazine, a well-known and structurally related bioorthogonal handle that degrades rapidly in the presence of biological thiols [23]. The triazine's structural similarity to native aromatic amino acids makes it attractive for tagging newly synthesized proteins via native biosynthetic pathways. Previous work in the lab showed that triazine **1** could be utilized by engineered biosynthetic enzymes, but was too sterically demanding to be incorporated in endogenous pathways [21]. Smaller triazine amino acids **2–4** (Figure 2-2C) could potentially be more useful in this regard. Indeed, several small, aromatic amino acid derivatives such as 4-fluorophenylalanine and pyridyl amino acids are known to be metabolized and incorporated into bacterial proteins [24-25].

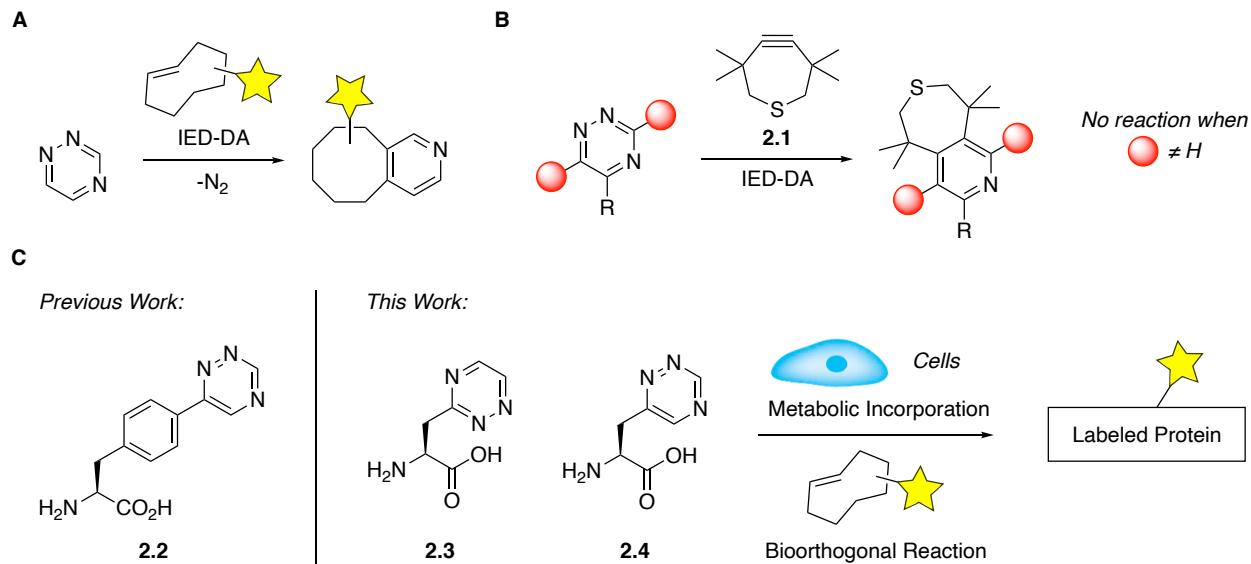
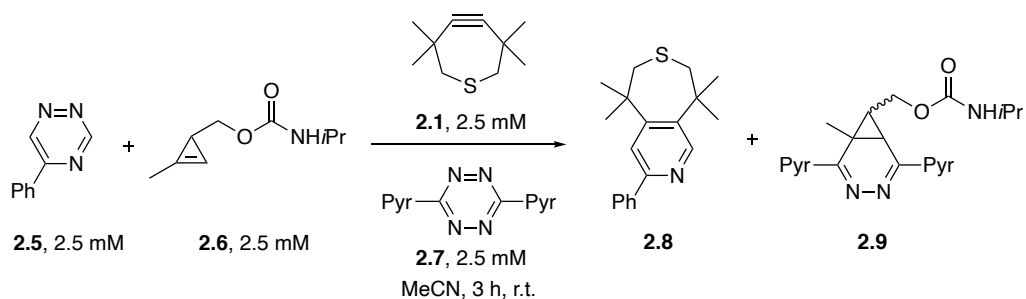


Figure 2-2. Triazines are versatile bioorthogonal motifs. A) Triazine scaffolds have been shown to react robustly with *trans*-cyclooctene (TCO) via an inverse electron-demand Diels–Alder cycloaddition (IED-DA). B) Isomeric 1,2,4-triazines exhibit unique reactivities. Tuning the steric bulk around the heterocycle affords selective reactivity with different strained π -systems, such as tetramethylthiacycloheptyne **2.1** (TMTH). C) A panel of triazine amino acids. Eliminating the additional phenyl ring in noncanonical amino acid **2.1** affords minimalist motifs. Varying the substitution around the triazine scaffold would take full advantage of the differences in reactivity with TCO and TMTH for potential in tandem experiments.

In addition to its small size, the triazine also displays unique profiles of reactivity with depending on substitution pattern. Previous work by Dr. David Kamber showed that 5-substituted triazines react robustly with sterically encumbered alkynes such as tetramethylthiacycloheptyne (TMTH) (Figure 2-2B). Such scaffolds display no signs of reactivity with less reactive π -systems, such as the cyclopropene (Cp). This mode of reactivity set the stage for orthogonal reaction development between other popular bioorthogonal chemistries, such as the ligation between 1,2,4,5-tetrazines and Cps. Dr. Kamber demonstrated that the bioorthogonal [4+2] cycloadditions could be conducted in a one-pot fashion without evidence of cross reactivity. However, this result was obtained using organic solvent (Scheme 2-1). Whether the reactions could be used simultaneously in biological environments remained to be determined.

Scheme 2-1. Mutually orthogonal bioorthogonal cycloadditions.



Below I describe my efforts to translate 1,2,4-triazines and their associated reactions to cellular contexts. Specifically, I examined whether the triazine motif could be used as a surrogate for phenylalanine (Phe). I synthesized triazine noncanonical amino acids and evaluated their metabolic incorporation in bacteria. I also investigated whether the dual [4+2] cycloadditions could be performed in aqueous solution on the surface of proteins. Additionally, I determined whether a third reaction could be incorporated to create a triplet of mutually orthogonal, bioorthogonal reactions. To date, only two triple-component ligation strategies have been reported [26-27].

2.3 Results and Discussion

2.3.1 Triazine amino acids as surrogates for phenylalanine

The study of proteome dynamics has benefitted greatly from noncanonical amino acids. These tools allow researchers to differentiate newly synthesized proteins in response to changes in the cellular environment. Such probes typically feature bioorthogonal handles, such that newly synthesized proteins can be selectively tagged for downstream proteomic analyses. To be successful, the noncanonical amino acids must be incorporated throughout the entire proteome, usually at a specific residue [28].

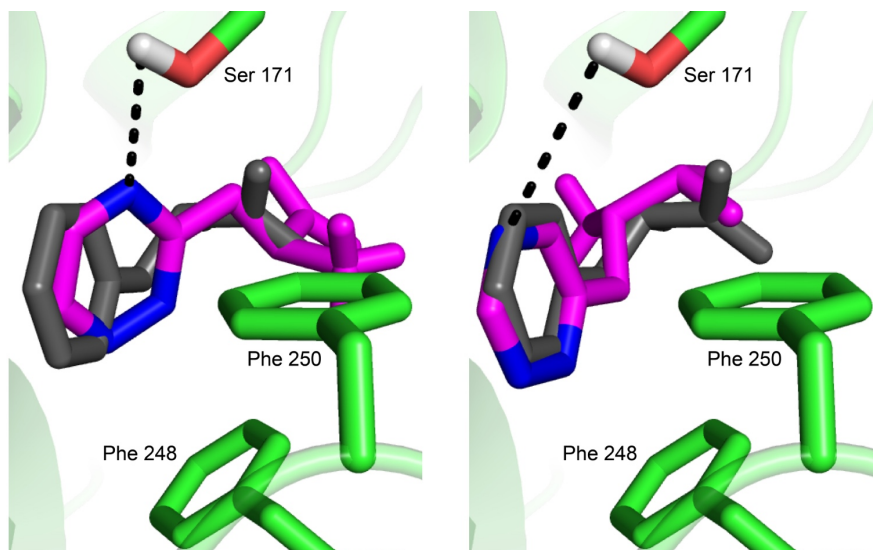


Figure 2-3. Autodocking studies with triazine amino acids **2.3** (left) and **2.4** (right) in the *E. coli* PheRS active site (PDB ID: 3PCO) [29]. In addition to a potential hydrogen bond interaction with Ser 171, π -stacking interactions may be possible between the triazine amino acids and two key phenylalanine residues (Phe 248, 250) that reside in the active site.

The most stringent enzymes involved in this process have historically been the aminoacyl-tRNA synthetases (AARS). These enzymes are responsible for transporting amino acids to the ribosome, and typically have high affinity for their cognate substrate. This specificity emphasizes the need for minimally perturbed reporters.

We hypothesized that triazine amino acids would be ideal surrogates for aromatic amino acids. The small size and increased aqueous stability of triazine scaffolds gave us confidence that the candidate amino acids could be metabolized by the cell's native machinery. We reasoned that the AARS most relevant to triazine metabolism would be the phenylalanine aminoacyl-tRNA synthetase (PheRS), due to the triazine's structural similarity to the natural substrate. Indeed, modeling studies of triazine amino acids **2.3** and **2.4** using Autodock suggested that these motifs could retain favorable interactions in the active site of the PheRS (Figure 2-3) [29-30].

To capitalize on the unique features of triazines as metabolic reporters, a panel of “minimalist” triazine amino acids were designed (Figure 2-2C). The core structures differed from triazine probes developed previously in the group, which comprised only aryl substituents. We first sought to determine whether alkyl-substituted triazine scaffolds were stable in biological conditions. To this end, triazine reporters were appended nonspecifically to the model protein lysozyme using carbonate **2.10** (Scheme S2-1) and verified by mass spectrometry (Figures 2-4A and 2-4B). The labeled proteins were then reacted with a *trans*-cyclooctene probe bearing a biotin moiety (TCO-biotin) and analyzed via SDS-PAGE and Western blot. As depicted in Figure 2-4C-D, sufficient signal over background was observed after 2 h at 37 °C, and the reaction was dose-dependent with respect to the TCO-biotin. Gratifyingly, the triazine-TCO reaction showed no nonspecific labeling in the presence of cell lysate, further demonstrating its utility in complex biological environments (Figure 2-4E).

After verifying the stability of alkyl-triazines, I proceeded to synthesize amino acids **2.3** and **2.4** using Negishi couplings between iodoalanine and functionalized triazines [31-33]. Commercially available L-serine methyl ester was protected with an Fmoc group. This base-labile protecting group was chosen due to the instability of 1,2,4-triazines under strongly acidic conditions. Compound **2.11** was iodinated using triphenylphosphine and iodine to access **2.12** (Scheme 2-2A). The triazine portions of the amino acids were derived from 3-amino-1,2,4-triazine (Scheme 2-2B). A Sandmeyer reaction followed by a Negishi coupling to **2.12** gave the fully protected 3-substituted triazine amino acid **2.14**. Subsequent deprotection and HPLC purification gave amino acid **2.3**. For the 6-substituted amino acid, electrophilic substitution with bromine

followed by a Sandmeyer reaction gave triazine **2.15**. This product was not isolated, as it degraded upon purification. Triazine **2.15** was taken on directly to the Negishi coupling. The resulting amino acid **2.16** was deprotected to give amino acid **2.4**.

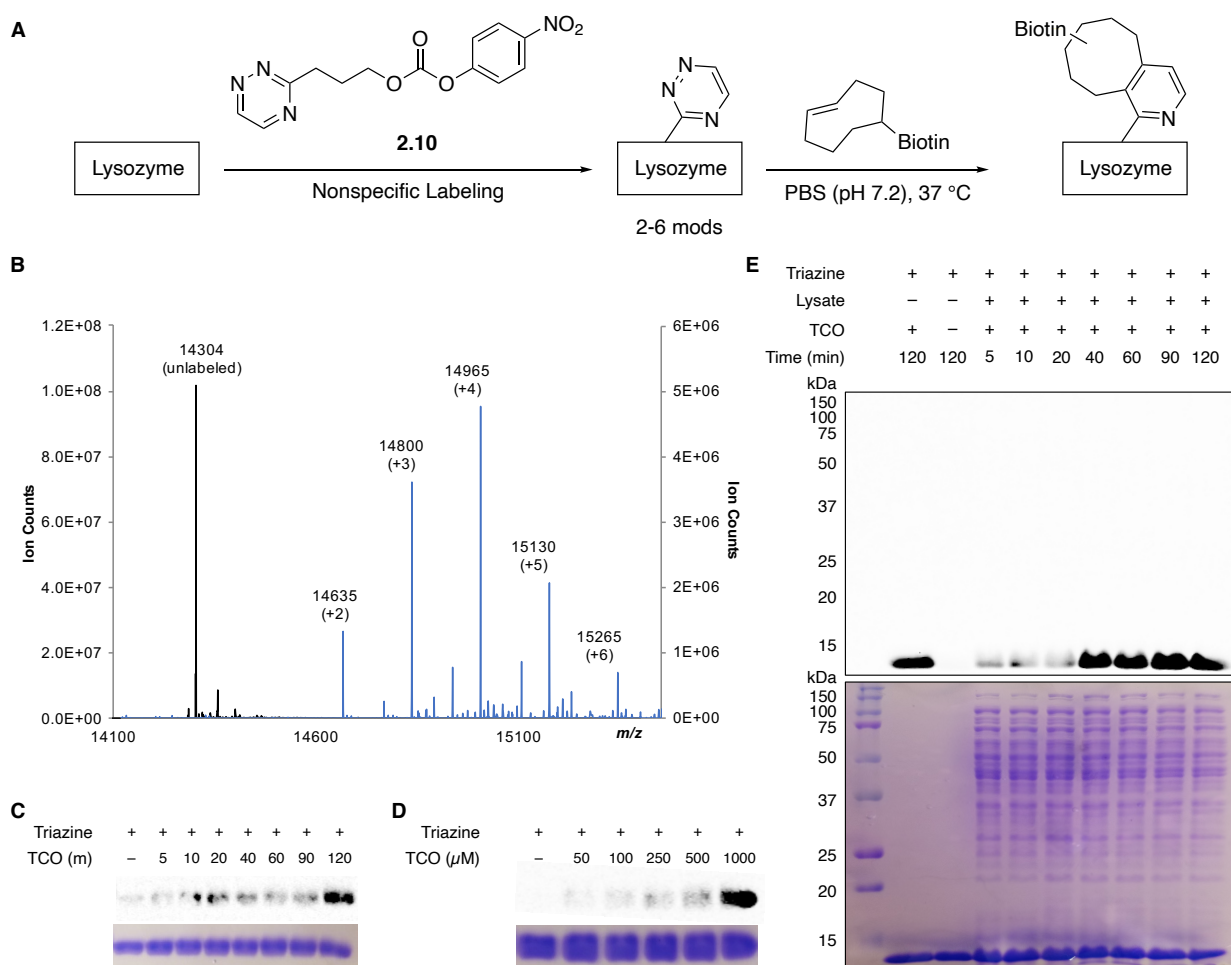
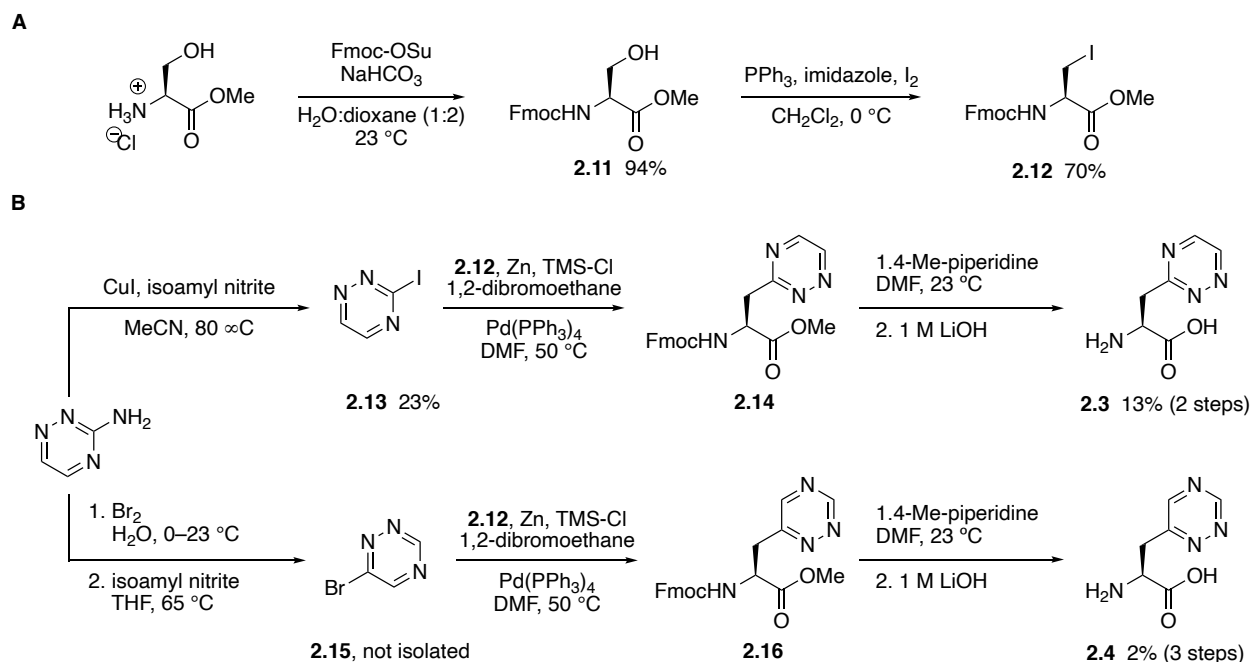


Figure 2-4. Triazine motifs function in biologically relevant environments. A) Carbonate **2.10** was used to label surface-exposed lysine residues on the protein lysozyme. B) Triazine attachment was verified using mass spectrometry. C) Western blot analysis of triazine-labeled lysozyme treated with TCO-biotin (1 mM) for 0–2 h at 37 °C. D) Western blot analysis of triazine-labeled lysozyme treated with up to 1 mM TCO-biotin for 2 h at 37 °C. E) Western blot analysis of cell lysate (15 μ g) doped with triazine-labeled lysozyme (5 μ g) and reacted with TCO-biotin (1 mM) for 5–120 min at 37 °C. Protein loading was verified with Coomassie staining.

Scheme 2-2. Synthesis of triazine noncanonical amino acids. A) Synthesis of the amino acid fragment. B) Synthesis of the triazine fragment and subsequent coupling.



The aqueous stabilities of amino acids **2.3** and **2.4** were evaluated to determine whether these scaffolds would be suitable for use in biological studies. The compounds were first incubated in buffered solutions (pH 7.4) at elevated temperatures, and monitored via ^1H NMR spectroscopy. No degradation was observed over seven days (See Appendix A). The amino acids were also subjected to high concentrations of L-glutathione, conditions that mimic the harsh intracellular environments encountered *in vivo*. After 24 h, the triazine amino acids showed no signs of degradation, suggesting that they are suitable for use in cellular studies (See Appendix A).

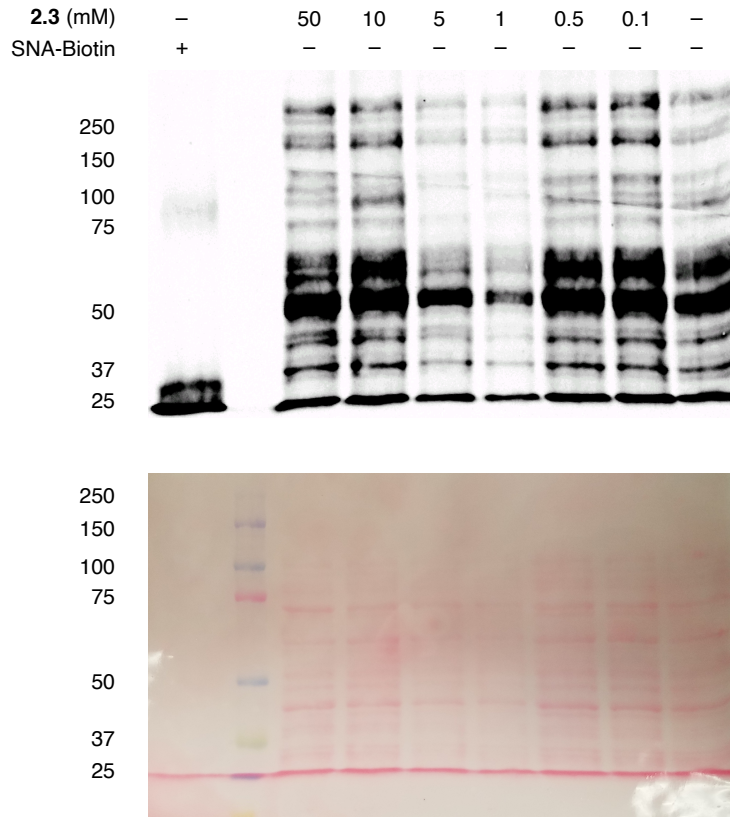


Figure 2-5. Metabolic labeling studies in BL21(DE3) *E. coli* cells. Cells were incubated in LB media with or without amino acid **2.3**. After 24 h the cells were lysed, and the supernatant was collected. Isolated proteins were labeled with TCO-biotin (1 mM final concentration) for 24 h at room temperature. The labeled proteins were separated by gel electrophoresis and analyzed via Western blot (top). Protein loading was confirmed using Ponceau S stain (bottom). Sambucus Nigra Lectin-Biotin (SNA-Biotin) was used as a positive control for streptavidin binding.

Metabolic labeling studies were performed in bacterial cells to determine if amino acid **2.3** could be incorporated into proteins using native biosynthetic machinery. In an initial experiment, BL21(DE3) *E. coli* cells were incubated in rich media containing amino acid **2.3** or no amino acid supplement. After 24 h, the cells were harvested, lysed, and reacted with TCO-biotin for 24 h. The resulting protein samples were analyzed via Western blot. No signal was observed over background, suggesting that amino acid **2.3** is unable to compete with the large pools of endogenous substrate for incorporation (Figure 2-5). To minimize competition between Phe and amino acid **2.3**,

an analogous labeling experiment was repeated using AF-IQ cells, a strain of *E. coli* that is auxotrophic for Phe [25]. Cells were incubated in M9 minimal media supplemented with Phe as a positive control, various concentrations of amino acid **2.3**, or no amino acid. Periodic OD₆₀₀ measurements were taken to monitor cell growth. Over the course of 24 h, though, no significant growth was observed in the cultures containing amino acid **2.3** (Figure 2-6).

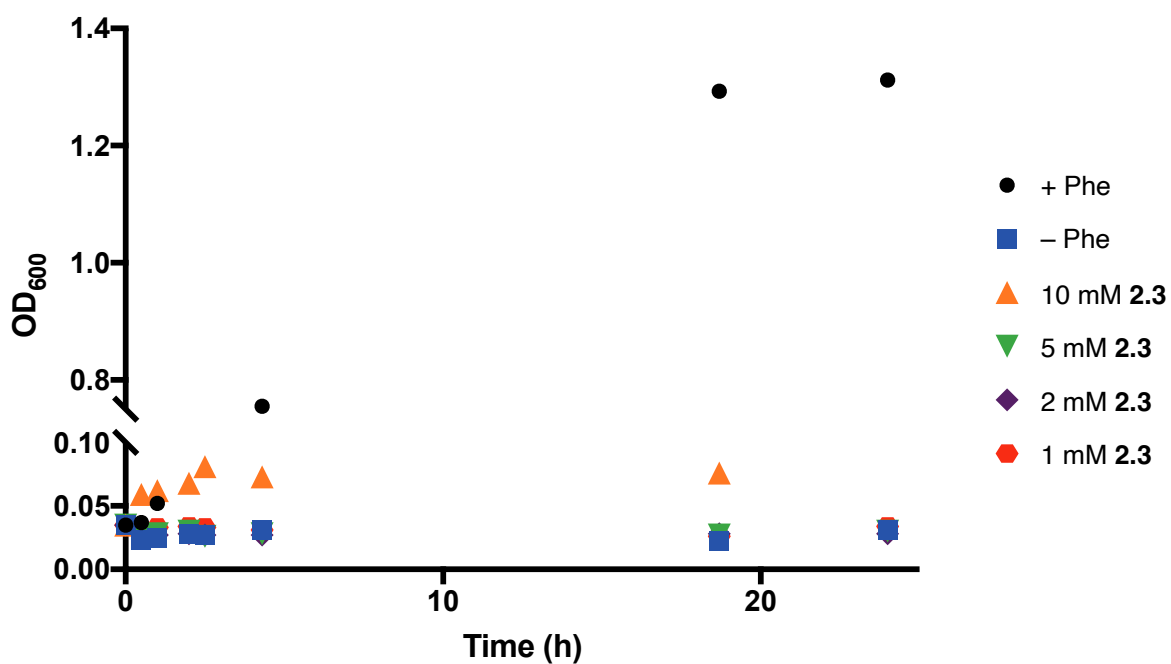


Figure 2-6. AF-IQ *E. coli* cells were treated with minimal media supplemented with phenylalanine, triazine amino acid **2.3**, or no amino acid. OD₆₀₀ measurements were taken periodically over 28 h.

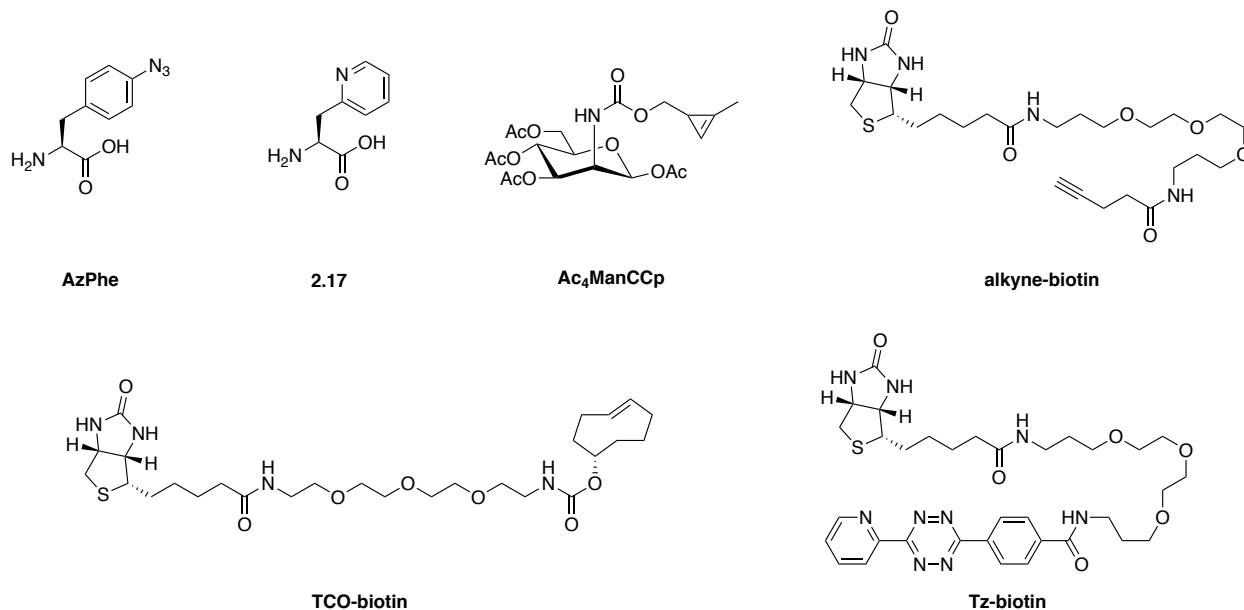


Figure 2-7. Full chemical structures of 4-azidophenylalanine (AzPhe), pyridyl amino acid **2.17**, cyclopropene sugar Ac₄ManCCP, alkyne-biotin, TCO-biotin and Tz-biotin.

While these experiments suggest that triazine amino acid **2.3** is not effectively processed by native *E. coli* machinery, it could be a viable substrate for well-known mutant strains. Tirrell and coworkers have demonstrated that a single point mutation to the PheRS enzyme (A294G, denoted as PheRS*) can improve its ability to incorporate noncanonical amino acids by relaxing the specificity of the active site [24,34]. Relevant to my work, PheRS* was able to process a related pyridyl amino acid (**2.17**, Figure 2-7). To test whether the same synthetase could tolerate amino acid **2.3**, I employed an auxotrophic *E. coli* strain expressing PheRS*, as well as murine dihydrofolate reductase (mDHFR), both under the control of the *lac* promoter [24]. The cells were incubated in M9 minimal media in the presence of pyridyl amino acid **2.17** (positive control), amino acid **2.3**, Phe or no amino acid supplement. Western blot analysis showed no signal over background, suggesting that the mutant PheRS is also not tolerant of amino acid **2.3**

(Figure 2-8A). Similar results were obtained with a mutant PheRS known to incorporate other Phe analogues (Figure 2-8B) [35].

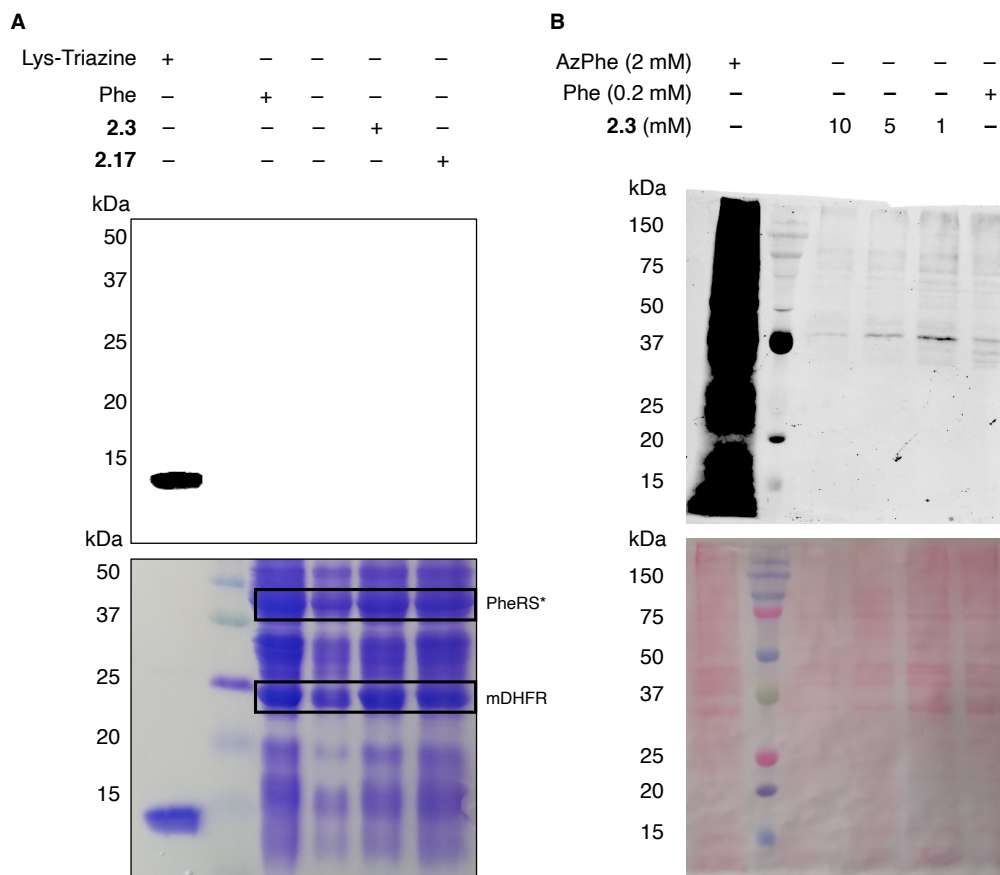


Figure 2-8. A) Metabolic labeling studies in AF-IQ[pQE-FS] *E. coli* cells. Cells were incubated in media containing Phe, triazine amino acid **2.3**, or pyridyl amino acid **2.17** (250 mg L^{-1}). Expression of the mutant PheRS and mDHFR was induced with isopropyl β -D-1-thiogalactopyranoside (IPTG, 1 mM final). The cells were lysed after a 4 h induction period. Isolated proteins were labeled with TCO-biotin (1 mM final). The labeled proteins were analyzed by SDS-PAGE and Western blot. Coomassie staining was used to verify protein loading. B) Metabolic labeling studies in KY33[pKPY514] *E. coli* cells suggest triazine amino acid **2.3** is not incorporated via a mutant PheRS [35]. Cells were incubated in media containing various concentrations of triazine amino acid **2.3**. Expression of the mutant PheRS was induced with IPTG (1 mM final). The cells were lysed after a 3 h incubation period. Isolated proteins were labeled with TCO-biotin (1 mM final) or alkyne-biotin (100 μM final). The labeled proteins were separated by gel electrophoresis and transferred to nitrocellulose prior to Ponceau S staining.

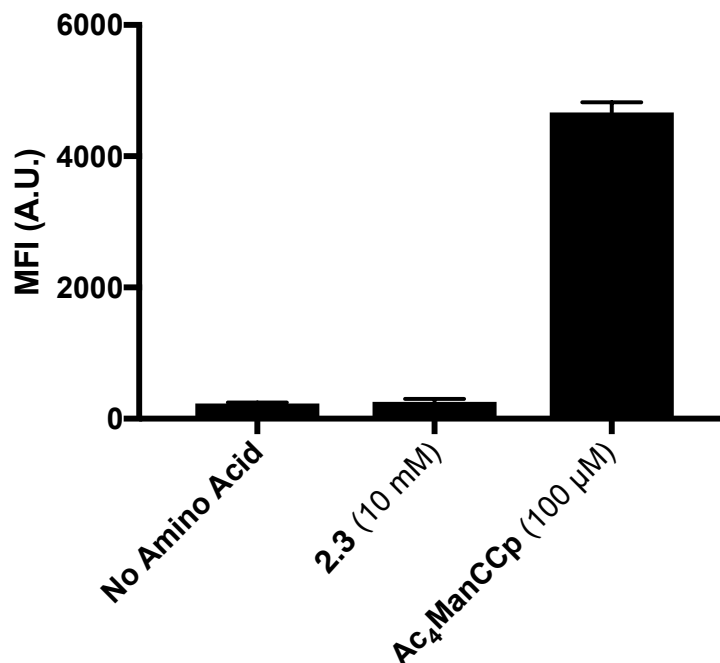


Figure 2-9. Flow cytometry analysis suggests that amino acid **2.3** is not tolerated by mammalian cells. HEK293 cells were incubated with amino acid **2.3** (10 mM), cyclopropene sugar **Ac₄ManCCp** [10] as a positive control (100 μM) or no supplement. Cells were fixed, permeabilized, and treated with TCO-biotin (1 mM) or a tetrazine-biotin conjugate (**Tz-biotin**, 10 μM). Cells were then stained with an avidin-APC conjugate and analyzed. Mean fluorescence intensities (MFI) of each cell population. Error bars represent the standard deviation of the mean for three labeling reactions.

Although *E. coli* are well-characterized model systems, their cell wall could pose a challenge for the triazine amino acid **2.3** to cross into the intracellular space. We reasoned that mammalian cells would offer a less stringent barrier for triazine amino acids to cross, owing to their lack of cell walls [36]. To this end, HEK293 cells were treated with amino acid **2.3** for 72 h. The cells were fixed, permeabilized, treated with the TCO-biotin conjugate, and analyzed by flow cytometry (Figure 2-9) [10]. Similar to the labeling studies in bacterial cells, no significant signal was observed over background. Collectively, these metabolic labeling studies suggest that no incorporation occurred in both mammalian and bacterial systems using native machinery. Ongoing work is being dedicated to exploring cell-free protein synthesis as a strategy for

installing triazine amino acids [37]. We are also screening engineered PheRS variants that may tolerate triazine amino acids **2.3** and **2.4**. The ability of the triazine scaffold to serve as a proxy for aromatic amino acids will open new avenues for studying complex protein environments and biophysical phenomena.

2.3.2 Developing a triplet of mutually orthogonal bioorthogonal reactions

As mentioned above, the unique reactivity profile of isomeric triazines suggested immediate opportunities for mutually orthogonal reaction development. Indeed, 5-substituted triazines reacted robustly with TMTH (**2.1**), even in the presence of tetrazine and cyclopropene probes. The two cycloadditions could be executed simultaneously without evidence of cross-reactivity (Scheme 2.1). The exquisite mutual orthogonality of the reagents arises from the interplay of intrinsic reactivities and steric factors [38-39]. Tetrazines are more reactive than triazines with bioorthogonal dienophiles due to the lower LUMO+1 energies of the π -systems [40-42]. This intrinsic order of reactivity is manifested in the “Distortion/Interaction-Activation Strain Model” [43] analysis shown in Figure 2-10. The sterically unhindered cyclopropene reacts most slowly with triazine **2.5**, and fastest with the tetrazine **2.7**. These differences in reactivity arise from increased interaction energies that parallel the increasing electrophilicity across the series. The same behavior is observed with other sterically unhindered dienophiles, such as the highly reactive TCO. With a sterically hindered dienophile (e.g., TMTH), though, the opposite trend is observed. The more hindered tetrazine reacts most slowly with TMTH (**2.1**), while the less hindered 5-phenyl triazine (**2.5**) ligates most rapidly. TMTH reactivity is mainly controlled by distortion energies that are strongly influenced

by steric considerations. Balancing electronic interaction energies (that enhance reactivity) with steric effects (that increase distortion energies and thus decrease reactivity) is a general approach for developing mutually orthogonal bioorthogonal reactions [39, 44-45].

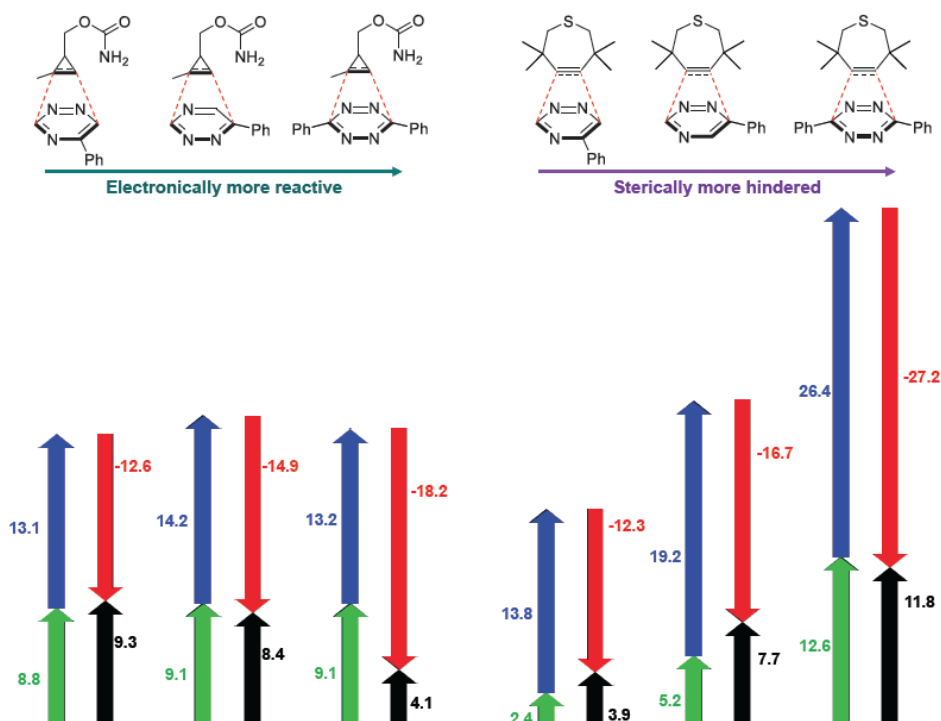


Figure 2-10. Distortion/interaction analysis of factors controlling mutually orthogonal cycloadditions. Black arrows are activation potential energies, green and blue arrows are distortion energies of dienophile and diene, respectively, and red arrows are interactions energies. All values are given in kcal/mol.

Encouraged by the computational and experimental analyses, we examined whether the orthogonal cycloadditions could be used in more complex settings. As mentioned previously, the dual cycloadditions were only performed in organic solvent. To transition these chemistries to more biologically relevant environments, we attempted the labeling reactions in concert with two model proteins (Figure 2-11, GFP and NanoLuciferase, Nluc). We attached a single 5-substituted triazine to Nluc using

cysteine-maleimide chemistry (Figure 2-12, Scheme S2-2). In this case, Nluc was engineered to harbor a single cysteine at residue 180 [46]. The resulting conjugate (**Nluc-Triazine**) was readily ligated with TMTH (**2.1**, Figure 2-13). We further prepared a cyclopropene-GFP (**GFP-Cp**) conjugate via genetic code expansion (Figure 2-13) [47]. The model proteins were combined 1:1 in PBS (pH 7.3, 2 μ M final concentration). The mixture was then treated with TMTH (**2.1**) and tetrazine **2.7**. After 3 h, full consumption of the starting proteins was observed via mass spectrometry (Figure 2-13). No cross-reactivities were observed, suggesting that the cycloadditions can be used to label two biological targets simultaneously. The orthogonal [4+2] cycloadditions could also be performed in the presence of lysate, conditions that closely mimic cellular environments (Figure 2-14). The cycloadditions were also compatible with polar bioorthogonal reactants, such as the cyclopropenone (CpO)-phosphine ligation. The unique reactivity of the CpO ligation placed it outside of the cycloaddition regime, immediately enabling simultaneous, triple-component ligations (Figures 2-15 and 2-16). Collections of three mutually orthogonal reactions are rare [26-27, 48].

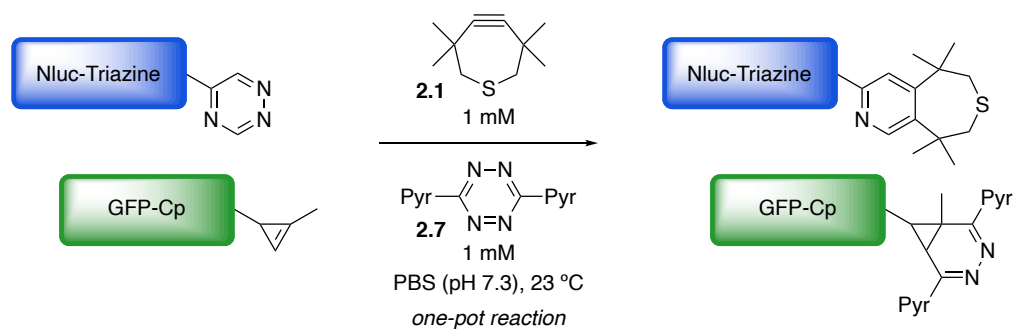


Figure 2-11. Orthogonal [4+2] cycloadditions enable dual protein labelling. Nluc-Triazine and GFP-Cp were mixed 1:1 in PBS (pH 7.3, 2 mM), and subsequently treated with TMTH (9) and tetrazine 3 (1 mM).

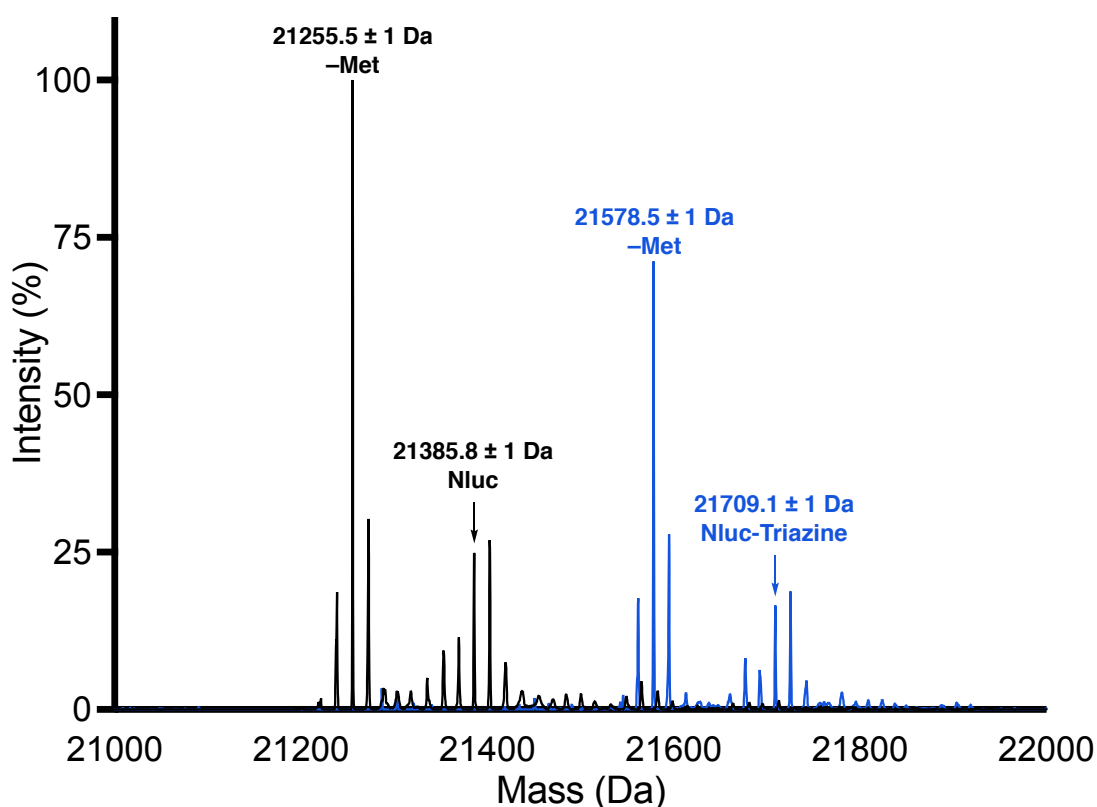


Figure 2-12. ESI-MS analysis of unlabeled **Nluc** (G180C, black) and **Nluc-Triazine** (blue). The expected masses were observed, in addition to masses corresponding to proteins with oxidized methionine residues and/or N-terminal methionine deletions. The observed and expected mass values for each species are tabulated below.

Species	Expected Mass (Da)	Observed Mass (Da)
Nluc – Met + 2 Met[O]	21239.3	21238.7±1
Nluc – Met + 3 Met[O]	21255.3	21255.5±1
Nluc – Met + 4 Met[O]	21271.3	21272.5±1
Nluc + 1 Met[O]	21354.5	21353.2±1
Nluc + 2 Met[O]	21370.5	21369.5±1
Nluc + 3 Met[O]	21386.5	21385.8±1
Nluc + 4 Met[O]	21402.5	21402.6±1
Nluc + 5 Met[O]	21418.5	21419.6±1
Nluc-Triazine – Met + 2 Met[O]	21562.4	21561.8±1
Nluc-Triazine – Met + 3 Met[O]	21578.4	21578.5±1
Nluc-Triazine – Met + 4 Met[O]	21594.4	21595.0±1
Nluc-Triazine + 1 Met[O]	21677.6	21676.9±1
Nluc-Triazine + 2 Met[O]	21693.6	21692.3±1
Nluc-Triazine + 3 Met[O]	21709.6	21709.1±1
Nluc-Triazine + 4 Met[O]	21725.6	21725.5±1
Nluc-Triazine + 5 Met[O]	21741.6	21742.5±1

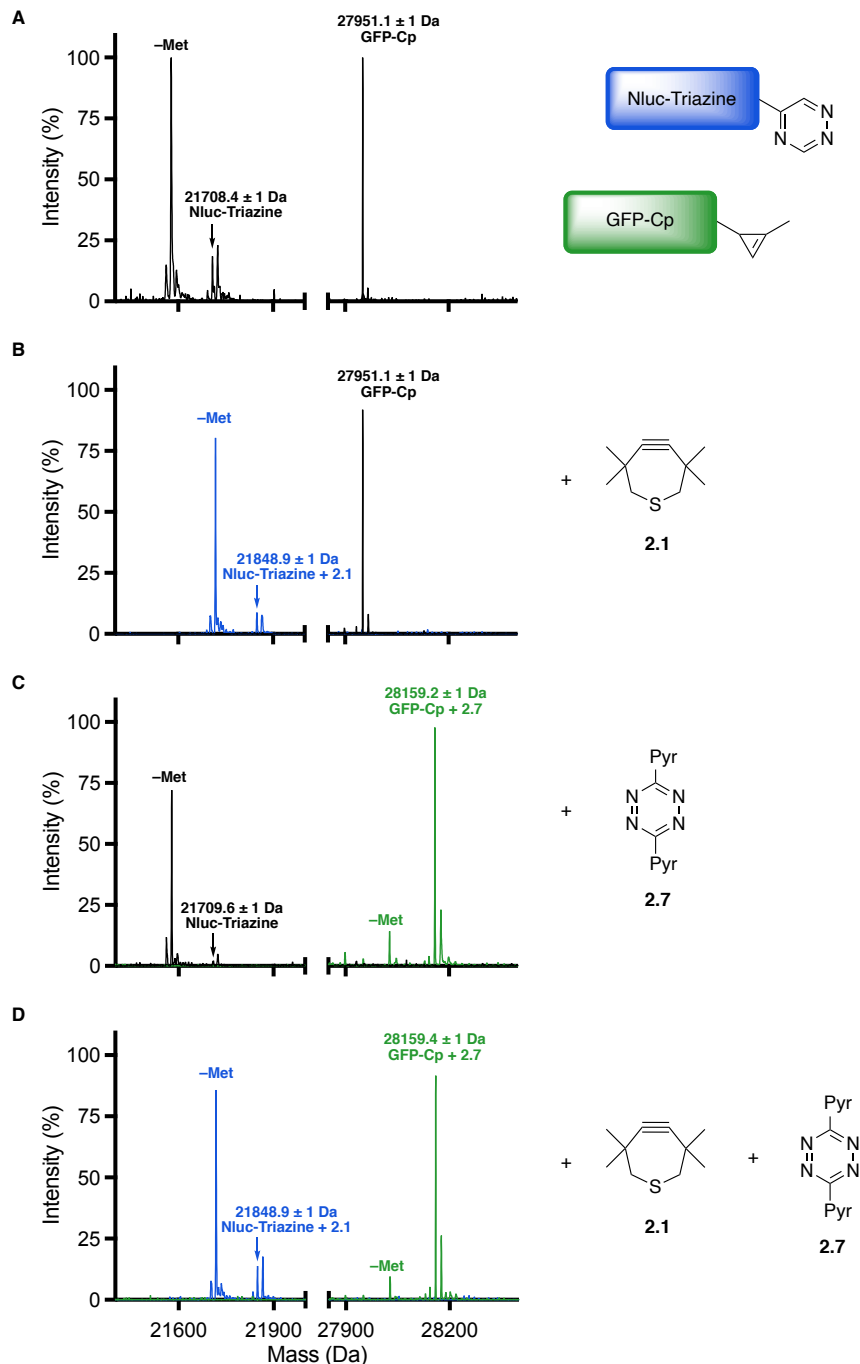


Figure 2-13. ESI-MS of orthogonal [4+2] cycloadditions on proteins. (A) Masses of starting proteins before exposure to small molecule probes (black). (B) **Nluc-Triazine** reacts upon treatment with TMTH **2.1** (blue) with a peak at 21848.9±1 Da, which corresponds to the cycloadduct plus three oxidized methionine residues (expected mass 21849.7 Da). A secondary peak at 21718.0±1 Da corresponds to the cycloadduct, plus three oxidized methionine residues, and cleavage of the *N*-terminal methionine residue (expected mass 21718.5 Da). **GFP-Cp** remains unreacted (expected mass 27951.4 Da). (C) **GFP-Cp** reacts upon treatment with tetrazine **2.7** (green) with a major peak at 28159.2±1 Da (expected mass 28159.5 Da). A secondary peak at 28027.7±1 Da corresponds to loss of the *N*-terminal methionine residue (expected mass 28028.3 Da). **Nluc-Triazine** is unreacted. (D) Addition of both small molecules leads to quantitative conversion to two distinct cycloadducts. No cross-reactivity adducts are observed.

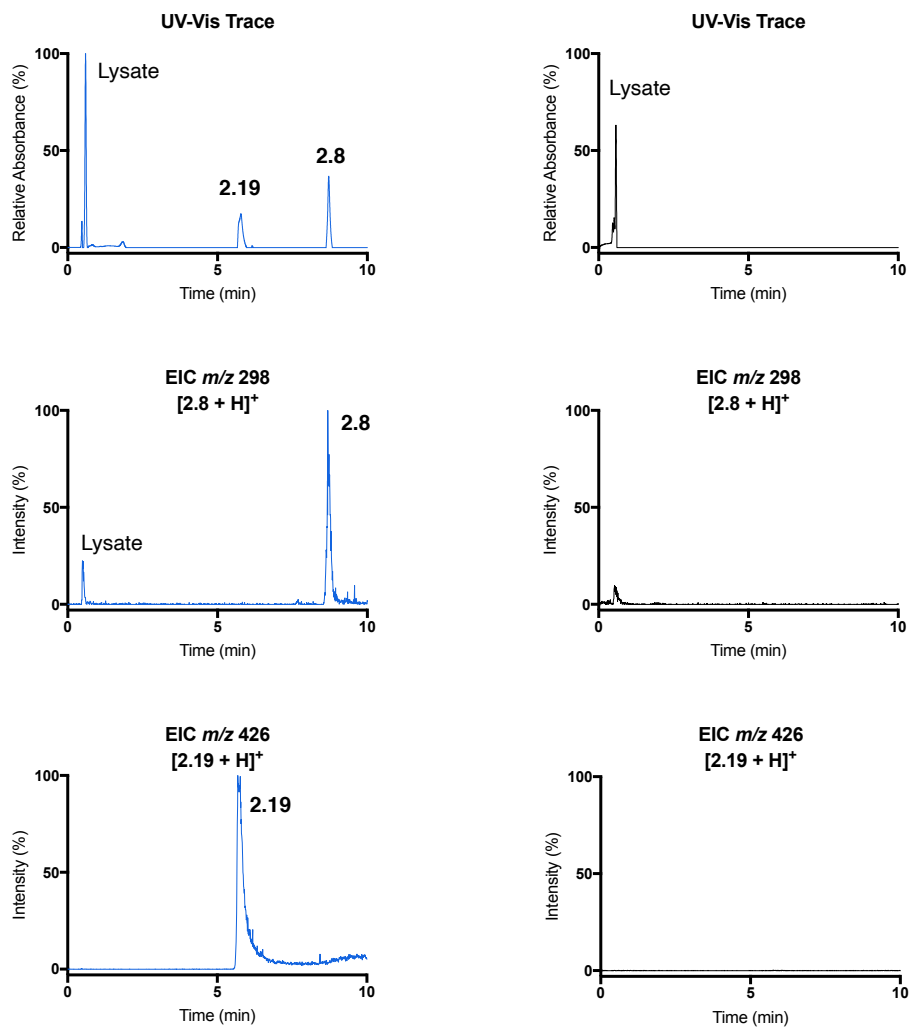
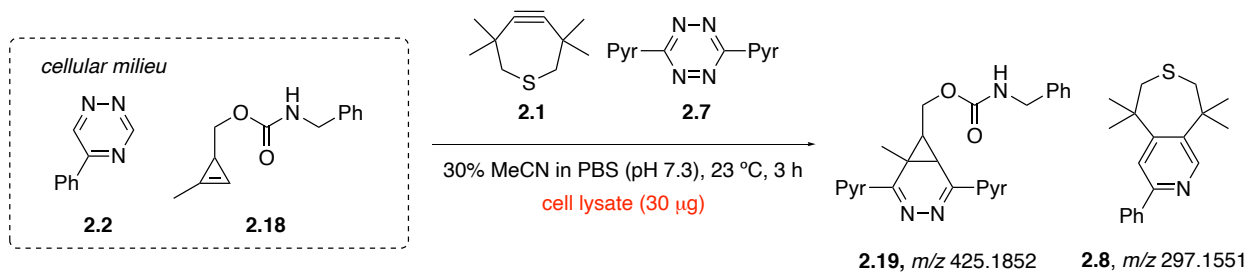


Figure 2-14. Orthogonal [4+2] cycloadditions were performed in cell lysate. (A) Triazine **2.2**, Cp **2.18**, tetrazine **2.7** and TMTH **2.1** were converted to their respective cycloadducts in the presence of cell lysate (blue). (B) No adducts were observed in the absence of the requisite reaction partners (black). The reactions were analyzed by absorbance (top) and extracted ion chromatographs (middle, *m/z* 298; bottom, *m/z* 426).

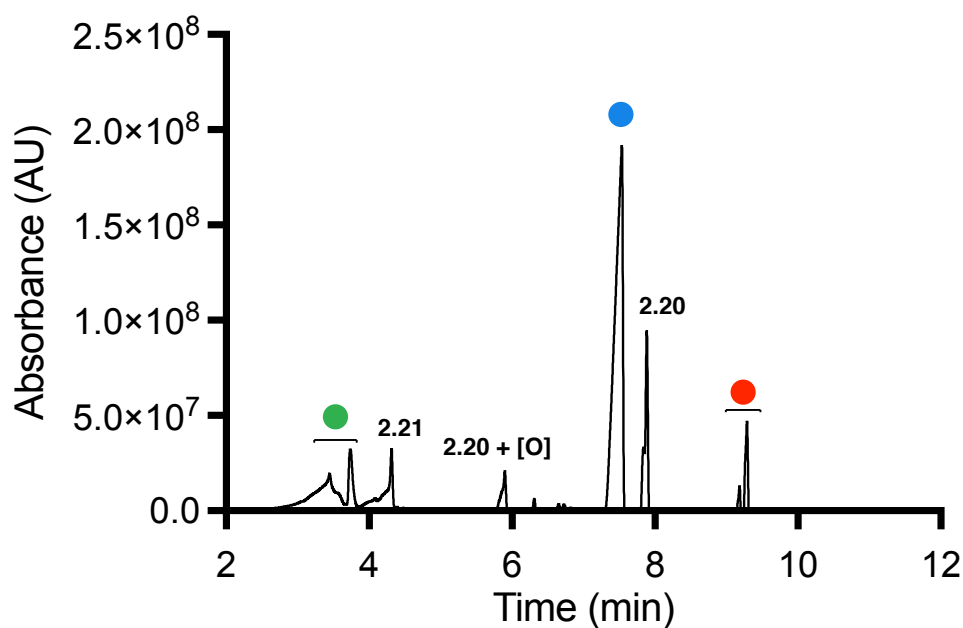
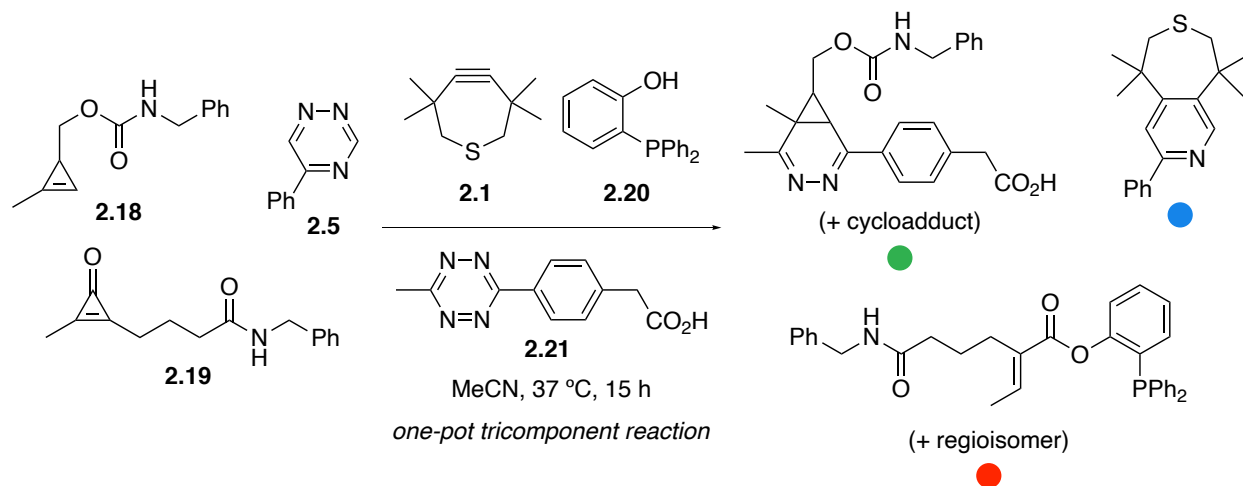


Figure 2-15. Tricomponent, one-pot bioorthogonal reaction. Triazine **2.5**, Cp **2.18**, cyclopropanone **2.19**, TMTH **2.1**, phosphine **2.20** and tetrazine **2.21** were mixed at equimolar concentrations in MeCN (1 mM final concentration for all reagents). The reaction was monitored by LC-MS. Three distinct ligation products were observed, corresponding to the matched reaction partners. No cross-reactivity was observed. Individual traces and reference reactions are shown in Figure 2-14 below.

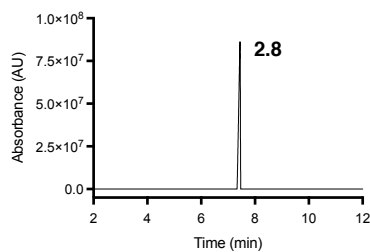
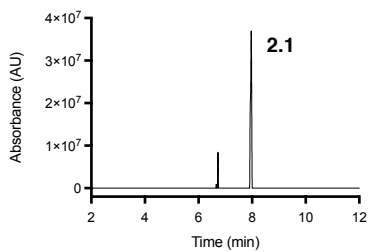
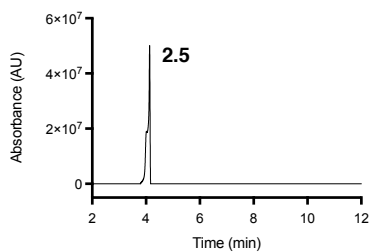
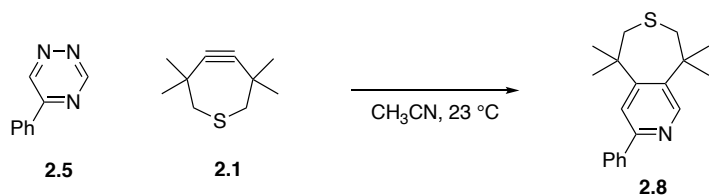
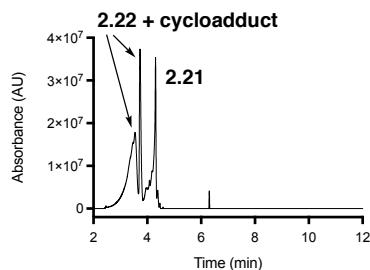
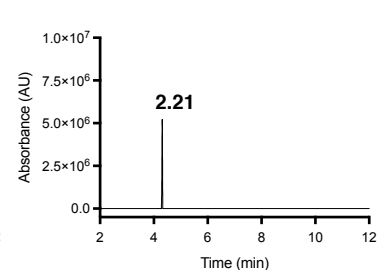
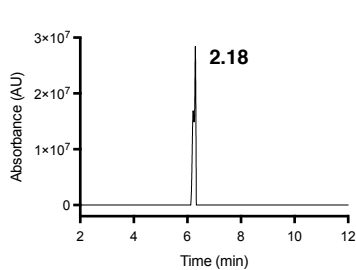
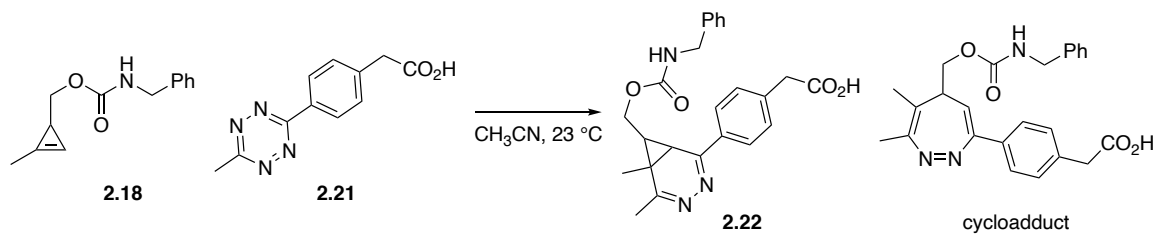
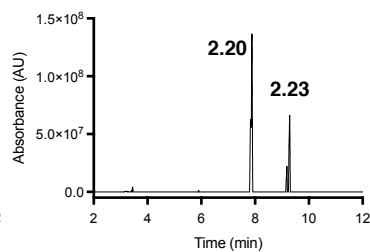
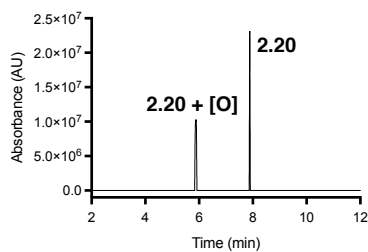
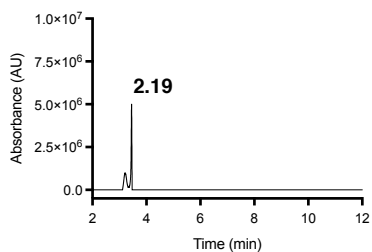
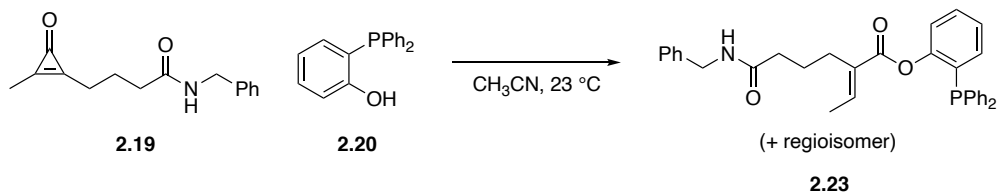
A**B****C**

Figure 2-16. LC-MS traces of starting bioorthogonal reagents and corresponding ligated adducts. (A) Triazine **2.5** and TMTH **2.1** were mixed in a 1:1 ratio (1 mM final concentration) in MeCN, then the reaction was monitored by LC-MS for the formation of ligation adduct **2.8**. (B) Cp **2.18** and tetrazine **2.21** were mixed in a 1:1 ratio (1 mM final concentration) in MeCN, then the reaction was monitored by LC-MS for the formation of ligation adduct **2.22**. Product **2.22** can further rearrange over the course of the experiment [49]. (C) Cyclopropanone **2.19** and phosphine **2.20** were mixed in a 1:1 ratio (1 mM final concentration) in MeCN, then the reaction was monitored by LC-MS for the formation of ligation adduct **2.23** and its regioisomer. Phosphine **2.20** oxidized over the course of the reaction (**2.20** + [O]). In all cases, the reaction trace is shown on the far right. The first two traces correspond to starting materials only.

2.4 Conclusions

The size and stability of the 1,2,4-triazine made it an ideal candidate for metabolic labeling applications. Its potential as a phenylalanine surrogate opens interesting avenues for exploring dynamic biophysical processes, such as protein folding. Key to the success of this goal was the development of small, alkyl-triazine amino acids with robust aqueous stability. In this work, I verified that alkyl-substituted 1,2,4-triazine scaffolds were stable in aqueous environments, and reacted robustly with their TCO reaction partner in the presence of cell lysate. A panel of “minimalist” triazine amino acid probes were synthesized and shown to be stable in aqueous environments. They were evaluated for their ability to serve as surrogates for Phe through proteome labeling studies in bacterial and mammalian systems.

We also identified isomeric triazine scaffolds that exhibit unique cycloaddition profiles. Computational and experimental analyses of triazine reactivity were performed. Isomeric triazines were found to react robustly with TCO, but only the least sterically encumbered isomers (5-substituted) reacted with other dienophiles. Notably, 5-substituted 1,2,4-triazines reacted efficiently with TMTH, one of the most sterically encumbered strained alkynes reported to date. The cycloaddition was successfully used in combination with another popular IED-DA reaction, the tetrazine ligation with 1-

methylcyclopropene. These mutually compatible reactions can be used in tandem to tag protein targets in biologically relevant environments. Future work will address the need for functional TMTM conjugates. TMTM has been historically difficult to outfit with fluorophores and other reporter groups, although new strategies for derivatization are being pursued. A panel of easily accessible reagents will further enhance multi-component labeling applications.

2.5 Materials and Methods

2.5.1 General Information

All reagents were obtained from commercial sources and used without further purification. Reactions were run under ambient conditions, unless otherwise indicated. Tetrahydrofuran (THF), diethyl ether (Et₂O), dichloromethane (CH₂Cl₂), dimethylformamide (DMF), and acetonitrile (MeCN) were degassed with argon and run through two 4 x 36 inch columns of anhydrous neutral A-2 (8 x 14 mesh; LaRoche Chemicals; activated under a flow of argon at 350 °C for 12 h). Thin-layer chromatography was performed using Silica Gel 60 F₂₅₄-coated glass plates (0.25 mm thickness), and visualization was performed with KMnO₄ stain and/or UV irradiation. Chromatography was accomplished with 60 Å (240–400 mesh) silica gel, commercially available from Sorbent Technologies. Organic solutions were concentrated under reduced pressure using a Büchi rotary evaporator. HPLC purifications were performed on a Varian ProStar equipped with 325 a Dual Wavelength UV-Vis detector. Analytical runs were performed using an Agilent C18 Scalar column (4.6 x 150 mm, 5 µm) with a 1 mL/min flow rate, and visualized with 210 nm wavelength. NMR spectra were collected

on a Bruker DRX400 (400 MHz ^1H , 100 MHz ^{13}C , 376.5 MHz ^{19}F), a Bruker DRX500 equipped with a cryo probe (500 MHz ^1H , 125.7 MHz ^{13}C), or a Bruker AVANCE600 equipped with a cryo probe (600 MHz ^1H , 150 MHz ^{13}C). All spectra were collected at 298 K. High-resolution mass spectrometry was performed by the University of California, Irvine Mass Spectrometry Center. Protein mass spectrometry (ESI-MS) experiments were performed on a Waters Xevo LV-G2 Q-TOF instrument.

2.5.2 Molecular Docking Studies

Conformational predictions of triazines **2.3** and **2.4** were performed using *Autodock Vina* [30]. The *E. coli* PheRS crystal structure was prepared for docking by removing all water molecules and cocrystallized ligands [29]. The AARS input file was generated using *Autodock Tools* [50]. All residues in the AARS were kept rigid, and grid coordinates were generated to define the sampling space for conformational analysis (Table 2-1). The unnatural amino acid ligands were constructed in Spartan Student Edition, and molecular geometries were optimized using the B3LYP(6-31G*) basis set. Rotational and torsional degrees of freedom for the minimized ligands were defined in *Autodock Tools*. The exhaustiveness parameter was set to 100. The threshold for potential polar contacts was defined as any interaction that was separated by less than 4 Å, the upper limit for a typical hydrogen bond.

Table 2-1. Grid coordinates used for docking amino acids **2.3** and **2.4** in the PheRS and predicted binding energies.

	Amino Acid 2.3	Amino Acid 2.4
Center_X	-25.431	-25.431
Center_Y	-10.192	-10.192
Center_Z	54.417	54.417
Size_X	42	42
Size_Y	18	18
Size_Z	30	30
Exhaustiveness	100	100
Binding Energy (kcal/mol)	-6.8	-6.5

2.5.3 M9 Minimal Media Recipes

M9 Minimal Medium A: In double-distilled water, 6.8 g/L sodium phosphate dibasic anhydrous, 3 g/L potassium phosphate monobasic, 0.5 g/L sodium chloride, 1 g/L ammonium chloride, 1 mM magnesium sulfate heptahydrate, 0.1 mM calcium chloride, 5 mg/L thiamine hydrochloride, 0.2% (w/v) glucose, 30 mg/L chloramphenicol, and 20 mg/L of the following reagents: L-alanine, L-cysteine, L-aspartic acid, L-glutamic acid, glycine, L-histidine, L-isoleucine, L-lysine, L-leucine, L-methionine, L-asparagine, L-proline, L-glutamine, L-arginine, L-serine, L-threonine, L-valine, L-tryptophan.

M9 Minimal Medium B: In M9 minimal medium B, add 20 mg/L of the following reagents: L-phenylalanine, L-tyrosine.

M9 Minimal Medium C: Follow recipe for M9 minimal medium A, but add 40 mg/mL ampicillin instead of 30 mg/mL chloramphenicol.

M9 Minimal Medium D: In double-distilled water, 6.8 g/L sodium phosphate dibasic anhydrous, 3 g/L potassium phosphate monobasic, 0.5 g/L sodium chloride dihydrate, 1 g/L ammonium chloride, 2 mM magnesium sulfate heptahydrate, 0.1 mM calcium chloride, 35 mg/L thiamine hydrochloride, 20 mM glucose, and 40 mg/L of the following reagents: kanamycin sulfate, L-alanine, L-cysteine, L-aspartic acid, L-glutamic acid, glycine, L-histidine, L-isoleucine, L-lysine, L-leucine, L-methionine, L-asparagine, L-proline, L-glutamine, L-arginine, L-serine, L-threonine, L-valine, L-tryptophan, L-tyrosine.

M9 Minimal Medium E: In M9 minimal medium D, 40 mg/L L-phenylalanine.

2.5.4 Cell Lysis Methods

Method A: Cells were harvested via centrifugation (4700 rpm, 10 min, 4 °C). The cells were resuspended in PBS (pH 7.4, 100 µL) and treated with Halt™ Protease Inhibitor Cocktail (ThermoFisher Scientific, 1X final concentration) and phenylmethylsulfonyl fluoride (PMSF, Gold Biotechnology, 0.5 mM final concentration). The cells were then lysed via sonication using a Sonic Vibra Cell VC 130 Ultrasonic Cell Disruptor, and the debris was pelleted by centrifugation at 14,500 rpm for 10 min at 4 °C. Protein concentrations in the remaining supernatant were quantified using a Pierce™ BCA Protein Assay Kit (Thermo Fisher).

Method B: Cells were harvested by centrifugation (4700 rpm, 15 min, 4 °C), and the pellets were stored at then lysed using a urea buffer (8 M urea, 100 mM NaH₂PO₄, 10 mM Tris-base, pH = 8.0) and treated with Halt™ Protease Inhibitor Cocktail

(ThermoFisher Scientific, 1X final concentration) and phenylmethylsulfonyl fluoride (PMSF, Gold Biotechnology, 0.5 mM final concentration). The samples were subjected to five freeze/thaw cycles using a dry ice/acetone slurry, and the debris was pelleted by centrifugation (14400 rpm, 20 min, 4 °C). Protein concentrations in the resulting supernatant were normalized to OD₆₀₀ measurements after the 4 h induction period.

2.5.5 Labeling Studies Using BL21(DE3) E. coli cells

An overnight culture of BL21(DE3) *E. coli* cells was grown in LB media supplemented with 40 mg/L kanamycin sulfate. Aliquots of the overnight culture (500 µL) were treated with amino acid **2.3** (0.1, 0.5, 1, 5, 10, and 50 mM) or no amino acid. The resulting cultures were incubated at 37 °C with shaking (250 rpm) for 24 h. Lysis method A was used to prepare samples for Western blot analysis.

2.5.6 Labeling Studies Using AF-IQ E. coli cells

An overnight culture of AF-IQ *E. coli* cells [25] was grown in freshly prepared M9 minimal media B. The cells were pelleted by centrifugation (4700 rpm, 10 min, 4 °C) and washed twice with ice-cold 0.9% (w/v) sodium chloride. The cells were resuspended in minimal media A (5 mL), and aliquots of the resulting culture (100 µL) were used to inoculate either minimal media B, minimal media A, or minimal media A supplemented with amino acid **2.3** (1, 2, 5, and 10 mM). The cells were incubated at 37 °C with shaking (250 rpm) for 28 h. OD₆₀₀ measurements were taken periodically over the course of the experiment.

2.5.7 Labeling Studies Using AF-IQ[pQE-FS] *E. coli* cells

An overnight culture of AF-IQ[pQE-FS] *E. coli* cells [25] was grown in LB media supplemented with 40 mg/L ampicillin and 30 mg/mL chloramphenicol. The cells were pelleted by centrifugation (4700 rpm, 10 min, 4 °C) and washed twice with ice-cold 0.9% (w/v) sodium chloride. The cells were resuspended in minimal media B, and aliquots of the resulting culture (150 µL) were used to inoculate fresh minimal media B (2 mL). The cultures were incubated at 37 °C with shaking (250 rpm) until reaching a target OD₆₀₀ of 0.8–1.0. The cells were pelleted by centrifugation (4700 rpm, 10 min, 4 °C) and washed twice with ice-cold 0.9% (w/v) sodium chloride. The cultures were resuspended in minimal media B (ampicillin instead of chloramphenicol), minimal media C, or minimal media C supplemented with either amino acid **2.3** (250 mg/L) or **2.17** (250 mg/L). The cells were incubated at 37 °C with shaking (250 rpm) for 20 min. Protein expression was induced by the addition of 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After 4 h of agitation at 37 °C, cells were harvested by centrifugation (4700 rpm, 15 min, 4 °C), and the pellets were stored at –20 °C. Lysis method B was used to prepare samples for Western blot analysis.

2.5.8 Labeling Studies Using KY33[pKPY514] *E. coli* cells

An overnight culture of KY33[pKPY514] *E. coli* cells [35] was grown in LB media supplemented with 40 mg/L kanamycin sulfate. Aliquots of the overnight culture (125 µL) were used to inoculate freshly prepared M9 minimal media E (5 mL). The cells were incubated at 37 °C with shaking (250 rpm) until reaching an OD₆₀₀ of 0.5. Cells were pelleted by centrifugation (4700 rpm, 15 min, 4 °C) and washed twice with ice-cold 0.9%

(w/v) sodium chloride. The resulting pellets were resuspended in M9 minimal medium D supplemented with either 4-azidophenylalanine (2 mM), amino acid **2.3** (1, 2, 5, and 10 mM) or phenylalanine (0.2 mM). The cells were incubated at 37 °C with shaking (250 rpm) for an additional 30 min. Then, expression of the mutant PheRS was induced by the addition of 1 M IPTG to a final concentration of 1 mM. After 3 h of agitation at 37 °C, cells were harvested by centrifugation (4700 rpm, 15 min, 4 °C), and the pellets were stored at –20 °C. Lysis method A was used to prepare samples for Western blot analysis.

2.5.9 Western Blot Analyses

Protein samples (50 µg protein in 100 µL total volume) were treated with TCO-biotin (1 mM) for 24 h at 23 °C for BL21(DE3), or 2 h at 37 °C for auxotrophic strains. Azide-containing proteins were treated for 1 h at 23 °C with a “click” chemistry cocktail comprising the following reagents: alkyne-biotin (100 µM); sodium ascorbate (1 mM); tris[(1-benzyl-1-H-1,2,3-triazol-4-yl)methyl]amine (TBTA, 100 µM); CuSO₄•5H₂O (1 mM). To precipitate proteins, ice-cold methanol (1 mL) was added and the samples were stored at –80 °C overnight. Protein precipitates were centrifuged at 14,500 rpm for 10 min at 4 °C, aspirated, and air-dried for 1 h. The protein isolates were then resuspended in 10 µL buffer (4% SDS, 150 mM NaCl, 50 mM triethanolamine, pH 7.4) and treated with SDS-PAGE loading buffer (10 µL of a 2X stock containing 20% glycerol, 0.2% bromophenol blue, 0.4% β-mercaptoethanol). The samples were heated at 95 °C for 5 min, separated by gel electrophoresis using 10% polyacrylamide gels, and then electroblotted to nitrocellulose membranes (0.2 µm; Bio-Rad). Transfer

efficiency was analyzed with Ponceau S staining. The membranes were rinsed with water and incubated with blocking buffer (7% bovine serum albumin in PBS containing 1% Tween® 20, PBS-T) for 1 h at 23 °C, followed by IRDye® 800CW streptavidin (LI-COR Biosciences; 1:10,000 dilution in blocking buffer) for 2 h at 23 °C. The membranes were subsequently washed with PBS-T (every 10 min x 6) and PBS (pH 7.4, every 5 min x 2), then imaged using an Odyssey infrared imaging system (LI-COR Biosciences, Odyssey Version 3.0).

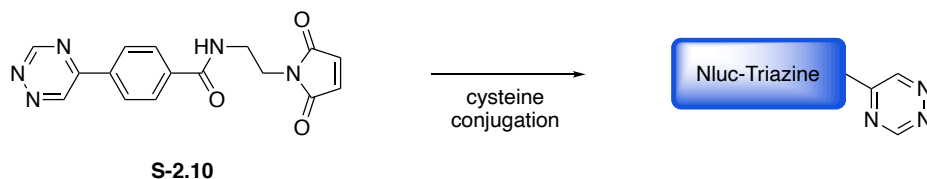
For labeling studies using AF-IQ[pQE-FS] *E. coli* cells, the samples were treated with TCO-biotin (1 mM final) and diluted with PBS to a total volume of 50 µL, then incubated at 37 °C for 5–120 min. The reactions were quenched with 3,6-di-2-pyridyl-1,2,4,5-tetrazine (1 µL of 50 mM DMSO stock) and analyzed via SDS-PAGE using 15% acrylamide gels and Western blot as outlined above.

2.5.10 Metabolic Labeling Studies in Mammalian Cells

Metabolic labeling studies were performed following a procedure outlined by Patterson, *et al.*, with some modifications [10]. HEK293 cells were plated at ~400,000 cells/well in 2 mL DMEM media (Corning) supplemented with 10% FBS, penicillin (100 U/mL), and streptomycin (100 µg/mL). Cells were incubated with triazine amino acid **2.3** (10 mM) or cyclopropene sugar Ac₄ManCCp (100 µM) for 72 h in a 5% CO₂, water-saturated incubator at 37 °C. The cells were rinsed with PBS containing 1% bovine serum albumin (FACS buffer, 3 x 500 µL), then fixed with a 4% solution of paraformaldehyde in FACS buffer (100 µL, 15 min, 0 °C). The cells were subsequently pelleted (3400 rpm, 3 min, 0 °C), washed with FACS buffer (3 x 500 µL), and

permeabilized with a 0.2% solution of Tween® 20 in PBS (200 μ L, 30 min, 23 $^{\circ}$ C). The cells were pelleted (3400 rpm, 3 min, 0 $^{\circ}$ C), washed with FACS buffer (3 x 500 μ L), and reacted with either TCO-biotin (1 mM, 1 h, 37 $^{\circ}$ C) or Tz-biotin (10 μ M, 30 min, 37 $^{\circ}$ C). The cells were pelleted (3400 rpm, 3 min, 0 $^{\circ}$ C), washed with FACS buffer (3 x 500 μ L), and stained with streptavidin-APC (eBioscience, 1:100 dilution in FACS buffer) for 1 h on ice in the dark. The cells were pelleted (3400 rpm, 3 min, 0 $^{\circ}$ C) and washed with additional FACS buffer (3 x 500 μ L), then analyzed by flow cytometry on an LSRII flow cytometer (BD Biosciences). For each sample, data were acquired for 5,000 events. Cells were analyzed in triplicate. Cellular fluorescence data were analyzed using FloJo software (Tree Star, Inc.).

2.5.11 Preparation of triazine-labeled of NanoLuciferase (Nluc-Triazine)

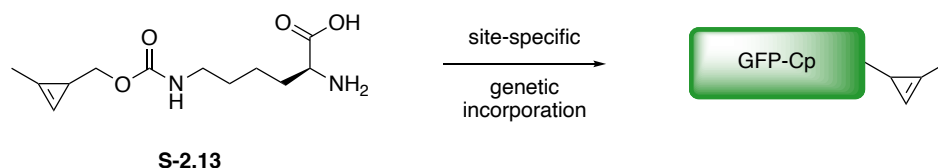


E. coli BL21(DE3) cells were transformed with pCold-His6-NLuc(G180C) [46] and used to inoculate 50 mL of LB broth containing ampicillin (40 μ g/mL). The culture was incubated at 37 $^{\circ}$ C with shaking (225 rpm) for 16 hr. A 15-mL aliquot of the starter culture was used to inoculate 1 L of LB broth containing ampicillin (40 μ g/mL), and the resulting culture was incubated at 37 $^{\circ}$ C with shaking (225 rpm). At $OD_{600} = 0.6$, the flask was cooled in an ice water bath for 20 min. Protein expression was induced by adding 1 M isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM, and the culture was incubated at 16 $^{\circ}$ C with shaking (225 rpm) for 20 h.

Cells were collected via centrifugation (3500 x *g* for 10 min at 4 °C). The pellet was resuspended in 30 mL of buffer (50 mM sodium phosphate, 20 mM imidazole, pH 7.8) and treated with Halt™ Protease Inhibitor Cocktail (Thermo Scientific). The cells were lysed via sonication, and the lysate was centrifuged (4500 x *g* for 45 min at 4 °C). The lysate was passed through a 0.22 µm PES membrane syringe filter (Olympus), and the desired protein was purified by Ni²⁺-affinity chromatography (eluting with 250 mM imidazole in 50 mM sodium phosphate buffer). Fractions containing the desired protein were combined and dialyzed against 50 mM sodium phosphate (pH 7.8) overnight. The final protein concentration was determined using a Jasco V-730 UV-Vis spectrophotometer and a predicted extinction coefficient of 24750 M⁻¹ cm⁻¹ (ProtParam).

Dithiothreitol (DTT) was added to the Nluc (G180C) sample from above (1 mM final concentration). The sample was diluted to a final protein concentration of 50 µM with phosphate buffered saline (PBS, pH 7.3), then treated with triazine **S-2.10** (10 mM stock in DMSO, 1 mM final concentration). The sample was incubated at room temperature for 2 h, and aliquots were analyzed by mass spectrometry. Once complete conversion was observed, the sample was concentrated via spin filtration (3 kDa MW cutoff, Millipore) and washed with PBS (pH 7.3, 3 x 400 µL). After the final wash, the sample was concentrated to a final volume of 50 µL and stored at -20 °C. Protein concentrations were determined as previously described.

2.5.12 Preparation of GFP-Cp



E. coli TOP10 cells transformed with pULTRA-WTPyIRS/pBAD-GFP-150 were used to inoculate 3 mL of LB broth containing spectinomycin (50 $\mu\text{g}/\text{mL}$) and ampicillin (100 $\mu\text{g}/\text{mL}$). After 16 h, the starter culture (250 μL) was used to inoculate autoinduction expression media (AIM, 25 mL, Table 2-2) containing 1 mM **S-2.13**, spectinomycin (50 $\mu\text{g}/\text{mL}$), and ampicillin (100 $\mu\text{g}/\text{mL}$) [51]. The culture was incubated at 37 $^{\circ}\text{C}$ with shaking (225 rpm) for 48 h.

Cells were collected via centrifugation (4500 rpm for 20 min at 4 $^{\circ}\text{C}$). The pellet was resuspended in 5 mL of PBS (pH 7.3), then treated with phenylmethylsulfonyl fluoride (PMSF, 500 μM final concentration) and a protease inhibitor cocktail (Sigma Aldrich). The cells were lysed via sonication and the lysate was centrifuged (14500 rpm for 30 min at 4 $^{\circ}\text{C}$). The lysate was passed through a 0.45 μm PES membrane syringe filter (Olympus) and was added to ProfinityTM IMAC resin (BioRad, 400 μL bed volume). The slurry was gently rocked for 1 h at 4 $^{\circ}\text{C}$, then washed with wash buffer (20 mM imidazole in PBS, pH 7.3, 8 mL). **GFP-Cp** was eluted using 1 mL of elution buffer (250 mM imidazole in PBS, pH 7.3). The protein was concentrated via spin filtration (3 kDa MW cutoff, Millipore), washed with PBS (pH 7.3, 500 μL), then concentrated again. This was repeated three times to remove excess imidazole. On the final wash, the protein was concentrated to a final volume of 40 μL . The final protein concentration was

determined by measuring the absorbance value at 488 nm using a Jasco V-730 UV-Vis spectrophotometer and an extinction coefficient of 88300 M⁻¹ cm⁻¹ [52]. Successful incorporation of amino acid **S-2.13** was verified through mass spectrometry.

Table 2-2. Autoinduction expression media components.^a

Aspartate (5%) ^b	2.5 mL
Glycerol (40%) ^b	0.625 mL
Glucose (40%) ^b	62.5 μL
L-Arabinose (20%) ^b	0.125 mL
Lactose (10%) ^b	0.1 mL
25X 18 Amino Acid Mix	2 mL
50X M-Salts	1 mL
1 M MgSO ₄	0.1 mL
5000X Trace Metals	10 μL

^aFinal volume of 50 mL, pH adjusted to 7.3

^bPercentage weight by volume

2.5.13 One-pot, dual labeling reaction on model proteins

Nluc-Triazine and **GFP-Cp** were mixed 1:1 in PBS (pH 7.3, 2 μM final concentration for each protein conjugate). The resulting mixture was treated with TMTH **2.1** (25 mM stock in DMSO, 1 mM final concentration) and tetrazine **2.7** (10 mM stock in DMSO, 1 mM final concentration). The solution was briefly vortexed and then incubated at room temperature for 3 h. Aliquots of the reaction solution (10 μL) were analyzed by mass spectrometry over time. Signal intensities of the individual cycloadducts were normalized to their respective unreacted starting materials.

2.5.14 One-pot, dual labeling reaction in cell lysate

An overnight culture of *E. coli* TOP10 cells (25 mL, LB Broth) was pelleted (3500 x g for 10 min at 4 °C) and re-suspended in PBS (pH 7.3, 200 μL). The suspension was

treated with PMSF (500 μ M final concentration) and a protease cocktail inhibitor (Sigma Aldrich), then lysed via sonication. The lysate was clarified (14500 rpm for 30 min at 4 $^{\circ}$ C) and the supernatant was collected. Total protein concentrations were determined using a Pierce[™] BCA protein assay kit (Thermo Fisher).

Clarified bacterial cell lysate (30 μ g) was treated with triazine **2.5** (10 mM stock in MeCN), cyclopropene **2.18** (10 mM stock in MeCN), tetrazine **2.21** (10 mM stock in MeCN), and TMTH **2.1** (50 mM stock in MeCN). The mixtures were diluted with PBS (pH 7.3) to a final volume of 50 μ L (1 mM final concentration for all bioorthogonal reagents), then allowed to stand at ambient temperature for 3 h. The reactions were diluted with 50% MeCN in H₂O (50 μ L) and concentrated via spin filtration (14500 rpm for 30 min at 4 $^{\circ}$ C) using a 3 kDa MW cutoff spin filter (Millipore) to a final volume of \sim 50 μ L. The filtrates were analyzed by LC-MS.

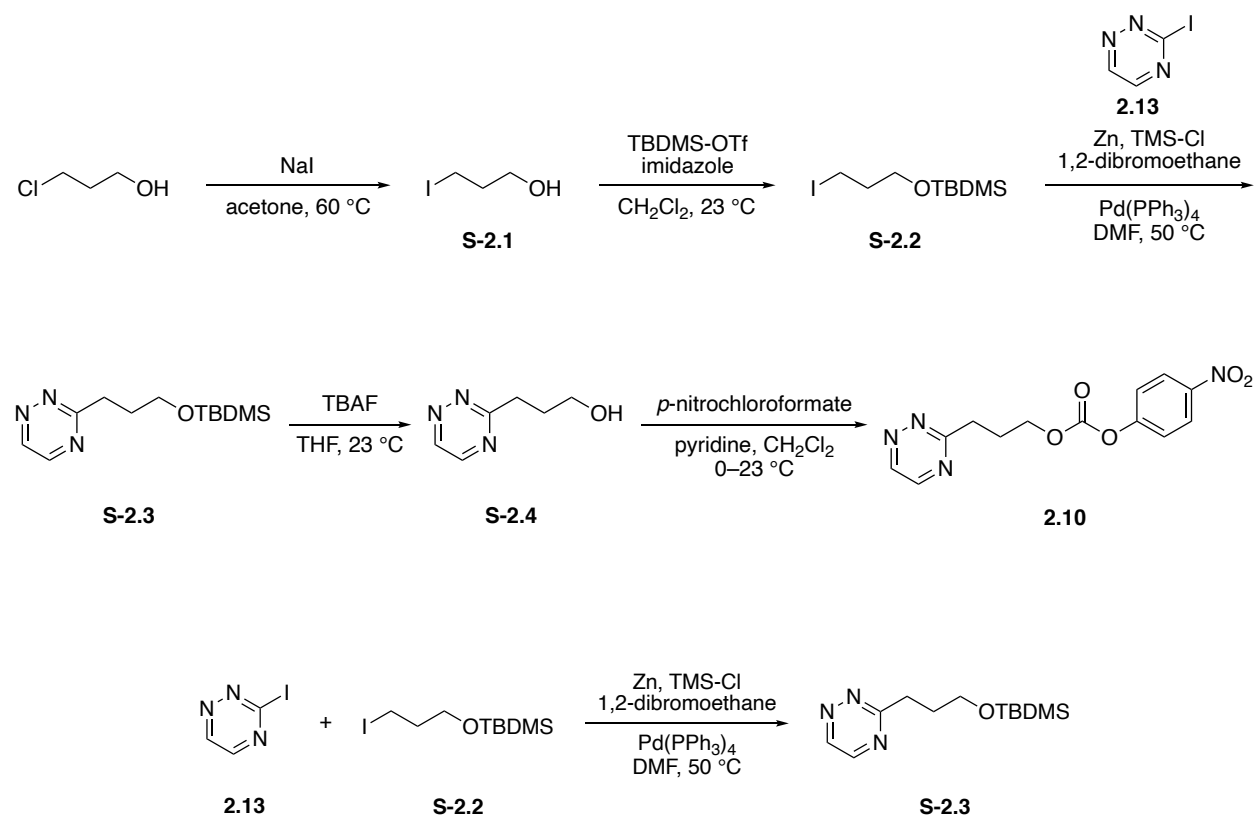
2.5.15 One-pot, triple labeling reaction

Triazine **2.5** (10 mM stock in MeCN), cyclopropene **2.18** (10 mM stock in MeCN), cyclopropeneone **2.19**, TMTH **2.1** (50 mM stock in MeCN), tetrazine **2.21** (10 mM stock in MeCN), and phosphine **2.20** (10 mM stock in MeCN) were mixed at equimolar ratios, then diluted with MeCN to a final volume of 200 μ L (1 mM final concentration for all reagents). The reaction was allowed to stand at ambient temperature for 18 h, then analyzed by LC-MS. Control reactions were conducted analogously using different reaction partners.

2.5.16 Synthetic Procedures

Compounds **2.1** [53], **2.5** [54], **2.6** [55], **2.11** [56], **2.12** [57], **2.13** [21], **2.15** [21], **2.21** [58], **S-2.1** [59], **S-2.2** [60], **S-2.5** [61], **S-2.9** [62], **S-2.11** [10], **S-2.12** [63], and **S-2.13** [47] were synthesized as previously reported.

Scheme S2-1. Synthesis of triazine-carbonate probe **2.10**.

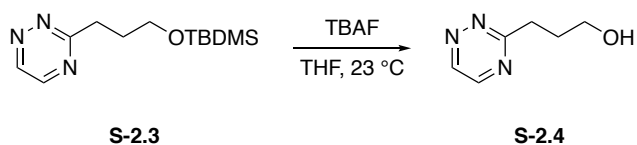


3-(3-((*tert*-Butyldimethylsilyl)oxy)propyl)-1,2,4-triazine (**S-2.3**)

To a 25-mL round-bottom flask was added zinc powder (0.329 g, 5.03 mmol), and the zinc was gently flame-dried under vacuum. After cooling to ambient temperature, anhydrous DMF (0.7 mL) was added, followed by 1,2-dibromoethane (0.03 mL, 0.3 mmol). The Zn slurry was stirred at 80 °C for one min, then cooled to ambient temperature. This cycle was repeated once more. Chlorotrimethylsilane (0.01 mL, 0.08

mmol.) was added, and the slurry was stirred at 80 °C for 30 min. After cooling to ambient temperature, a solution of **S-2.2** (0.232 g, 0.772 mmol) in anhydrous DMF (0.7 mL) was added dropwise. The resulting slurry was stirred at 45 °C until TLC indicated full consumption of **S-2.2** (2:1 hexanes:EtOAc). The reaction mixture was cooled to ambient temperature. Pd(PPh₃)₄ (92.3 mg, 0.0799 mmol, 10 mol %) was added, followed by a solution of **2.13** in anhydrous DMF (1.6 mL). The reaction mixture was stirred at 50 °C for 16 h. The resulting dark brown solution was diluted with EtOAc and washed with brine (4 x 100 mL). The organic layer was dried with MgSO₄, filtered and concentrated *in vacuo*. The crude brown oil was dry-loaded on Celite and purified by flash column chromatography (eluting with 35% EtOAc in hexanes) to give **S-2.3** (51 mg, 0.20 mmol, 26%) as a dark orange oil.

¹H NMR (500 MHz, CDCl₃): δ 9.11 (d, *J* = 2.3 Hz, 1H), 8.55 (d, *J* = 2.3 Hz, 1H), 3.74 (t, *J* = 6.2 Hz, 2H), 3.22 (d, *J* = 7.7 Hz, 2H), 2.11 (quint, *J* = 6.9 Hz, 2H), 0.87 (s, 9H), 0.03 (s, 6H). ¹³C NMR (125 MHz, CDCl₃): δ 170.9, 148.8, 147.7, 62.5, 34.2, 31.2, 26.1, 18.5, -5.2. HRMS (ESI+) *m/z* calculated for C₁₂H₂₃N₃OSiNa [M+Na]⁺ 276.1508, found 276.1513.

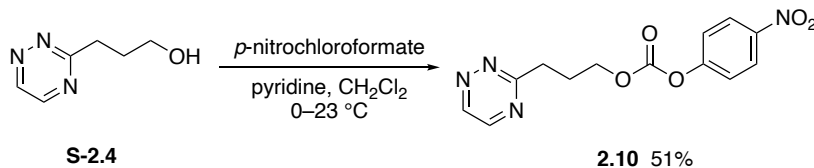


3-(1,2,4-Triazin-3-yl)propan-1-ol (**S-2.4**)

To a scintillation vial containing **S-2.3** (50.7 mg, 0.200 mmol) was added anhydrous THF (0.7 mL). Tetrabutylammonium fluoride (TBAF, 1 M solution in THF, 0.30 mL, 0.30 mmol) was added, and the reaction was stirred at ambient temperature

for 22 h. The reaction mixture was concentrated *in vacuo*, and the resulting brown oil was purified by flash column chromatography (eluting with 2% MeOH in CH₂Cl₂) to give **S-2.4** (16 mg, 0.11 mmol, 58%) as a yellow oil.

¹H NMR (500 MHz, CDCl₃): δ 9.13 (d, *J* = 2.2 Hz, 1H), 8.57 (d, *J* = 2.3 Hz, 1H), 3.77 (t, *J* = 6.1 Hz, 2H), 3.28 (t, *J* = 7.3 Hz, 2H), 2.16 (quint, *J* = 6.7 Hz, 2H), 2.00 (br s, 1H). ¹³C NMR (125 MHz, CDCl₃): δ 170.6, 148.8, 147.8, 62.1, 34.2, 30.8. HRMS (ESI+) *m/z* calculated for C₆H₉N₃ONa [M+Na]⁺ 162.0643, found 162.0651.

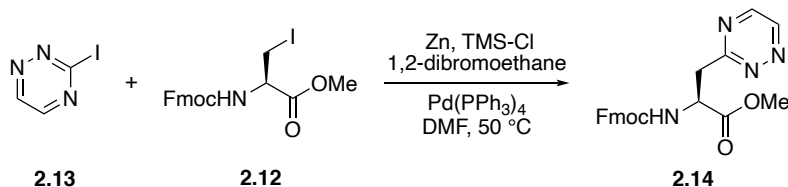


3-(1,2,4-Triazin-3-yl)propyl (4-nitrophenyl) carbonate (**2.10**)

To a flame-dried 10-mL round-bottom flask was added **S-2.4** (10.5 mg, 0.0754 mmol) in anhydrous CH₂Cl₂ (0.5 mL). Anhydrous pyridine (0.04 mL, 0.5 mmol) was added, and the solution was cooled to 0 °C. Then, 4-nitrophenylchloroformate (55.6 mg, 0.276 mmol) was added, and the solution was stirred at 0 °C for 20 min. The solution was slowly warmed to ambient temperature and allowed to stir for an additional 2 h. The reaction mixture was diluted with CH₂Cl₂ (5 mL) and H₂O (5 mL) and the layers were separated. The aqueous layer was washed with CH₂Cl₂ (3 x 5 mL), and the combined organic layers were dried with MgSO₄, filtered and concentrated *in vacuo*. The crude yellow solid was purified by flash column chromatography (eluting with 2% MeOH in CH₂Cl₂) to give **2.10** (7.6 mg, 25 μmol, 33%) as a light orange oil. Trace amounts of

grease and 4-nitrophenol were observed, likely from hydrolysis of the carbonate product during purification.

^1H NMR (500 MHz, CDCl_3): δ 9.16 (s, 1H), 8.60 (s, 1H), 8.28 (d, $J = 9.1$ Hz, 2H), 7.38 (d, $J = 9.1$ Hz, 2H), 4.44 (t, $J = 6.3$ Hz, 2H), 3.33 (t, $J = 7.4$ Hz, 2H), 2.40 (quint, $J = 6.9$ Hz, 2H). ^{13}C NMR (125 MHz, CDCl_3): δ . 155.6, 152.6, 149.0, 148.1, 145.2, 125.5, 122.0, 68.6, 33.6, 26.7. HRMS (ESI+) m/z calculated for $\text{C}_{13}\text{H}_{12}\text{N}_4\text{O}_5$ $[\text{M}+\text{Na}]^+$ 327.0705, found 327.0703.



***N*-[(9-Fluorenylmethoxy)carbonyl]-L-3-(1,2,4-triazin-3-yl)propanoate (2.14)**

To a round-bottom flask was added zinc powder (0.831 g, 12.7 mmol), and the zinc was gently flame-dried under vacuum. After cooling to ambient temperature, anhydrous DMF (2 mL) was added, followed by 1,2-dibromoethane (0.07 mL, 0.8 mmol). The Zn slurry was stirred at 80 °C for one min, then cooled to ambient temperature. This cycle was repeated once more. Chlorotrimethylsilane (0.02 mL, 0.1 mmol.) was added, and the slurry was stirred at 80 °C for 30 min. After cooling to ambient temperature, a solution of **2.12** (0.957 g, 2.12 mmol) in anhydrous DMF (7.5 mL) was added dropwise. The resulting slurry was stirred at 45 °C until TLC indicated full consumption of **2.12** (2:1 hexanes:EtOAc). The reaction mixture was then cooled to ambient temperature. To a separate, flame-dried round-bottom flask was added Pd(PPh₃)₄ (0.253 g, 0.219 mmol) and a solution of compound **2.13** (0.530 g, 2.54 mmol)

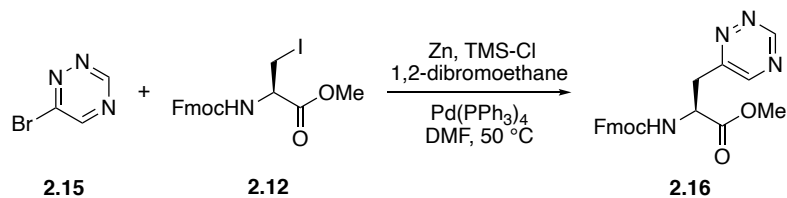
in anhydrous DMF (5 mL). The reaction mixture was stirred at ambient temperature under N₂ for 10 min. The resulting yellow suspension was added dropwise to the Zn slurry, and the reaction mixture was stirred at 50 °C for 16 h. The reaction mixture was then diluted with CH₂Cl₂ (75 mL) and washed with brine (4 x 100 mL). The organic layer was dried with MgSO₄ and filtered, then dry-loaded onto Celite and passed through a silica plug (eluting with 20-50% EtOAc in hexanes) to give a yellow oil. Product **2.14** contained inseparable triphenylphosphine oxide impurities, and was used without further purification.



(S)-2-Amino-3-(1,2,4-triazin-3-yl)propanoic acid (2.3)

To a scintillation vial containing **2.14** (0.227 g, 0.561 mmol) was added 20% 4-methylpiperidine in DMF (5 mL). The reaction mixture was stirred at ambient temperature under N₂ until full consumption of starting material was observed by TLC (1:1 hexanes:EtOAc, ninhydrin stain). A solution of LiOH (1 M, 0.73 mL, 0.73 mmol) was added dropwise, and the reaction was stirred for 5 h at ambient temperature. The reaction mixture was diluted with H₂O (5 mL) and washed with CH₂Cl₂ (3 x 15 mL). The aqueous layer was concentrated and purified by HPLC, eluting with a gradient of 0-95% MeCN in H₂O + 0.1% HCl over 30 min. The desired fractions were collected and lyophilized to give the chloride salt of **2.3** (60 mg, 0.29 mmol, 13% over two steps) as a light yellow solid.

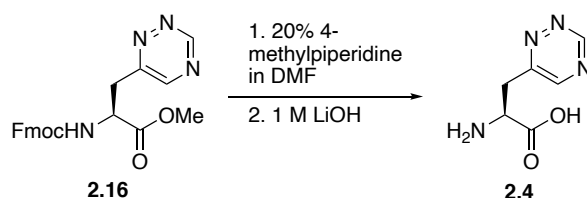
^1H NMR (400 MHz, D_2O): δ 9.27 (d, J = 2.5 Hz, 1H), 8.86 (d, J = 2.5 Hz, 1H), 4.09 (dd, J = 7.8, 5.3 Hz, 1H), 3.62 (dd, J = 15.5, 5.3 Hz, 1H), 3.53 (dd, J = 15.5, 7.8 Hz, 1H). ^{13}C NMR (125 MHz, D_2O): δ 167.0, 151.2, 148.4, 110.0, 54.2, 39.4. HRMS (ESI+) m/z calculated for $\text{C}_6\text{H}_8\text{N}_4\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 191.0545, found 191.0554.



***N*-[(9-Fluorenylmethoxy)carbonyl]-L-3-(1,2,4-triazin-6-yl)propanoate (2.14)**

To a round-bottom flask was added zinc powder (0.196 g, 3.00 mmol), and the zinc was gently flame-dried under vacuum. After cooling to ambient temperature, anhydrous DMF (0.45 mL) was added, followed by 1,2-dibromoethane (0.02 mL, 0.2 mmol). The Zn slurry was stirred at 80 °C for one min, then cooled to ambient temperature. This cycle was repeated once more. Chlorotrimethylsilane (0.01 mL, 0.08 mmol) was added, and the slurry was stirred at 80 °C for 30 min. After cooling to ambient temperature, a solution of **2.12** (0.217 g, 0.48 mmol) in anhydrous DMF (2 mL) was added dropwise. The resulting slurry was stirred at 45 °C until TLC indicated full consumption of **2.12** (2:1 hexanes:EtOAc). The reaction mixture was cooled to ambient temperature. $\text{Pd}(\text{PPh}_3)_4$ (57.1 mg, 0.0494 mmol, 10 mol %) was added, followed by the solution of **2.15** diluted in anhydrous DMF (0.5 mL). The reaction mixture was stirred at 50 °C for 16 h. The reaction mixture was diluted with CH_2Cl_2 (75 mL) and washed with brine (4 x 100 mL). The organic layer was dried with MgSO_4 , filtered and concentrated *in vacuo*. The brown crude oil was dry-loaded onto Celite and passed through a silica

plug (eluting with 25-40% EtOAc in hexanes) to give a yellow oil. Product **2.16** contained inseparable triphenylphosphine oxide impurities, and was used without further purification.

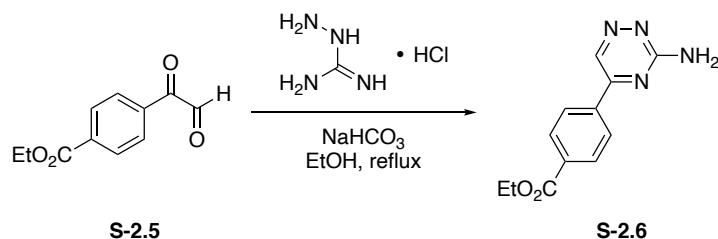
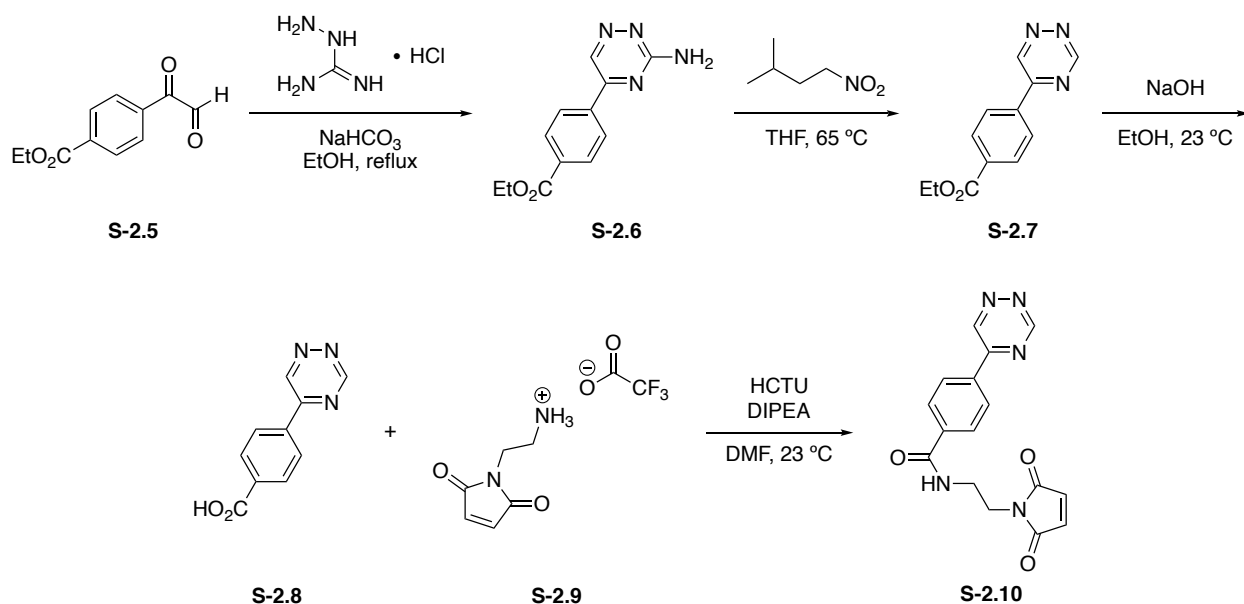


(S)-2-Amino-3-(1,2,4-triazin-6-yl)propanoic acid (**2.4**)

To a scintillation vial containing **2.16** was added 20% 4-methylpiperidine in DMF (1 mL). The reaction mixture was stirred at ambient temperature under N₂ until full consumption of starting material was observed by TLC (1:1 hexanes:EtOAc, ninhydrin stain). A solution of LiOH (1 M, 0.19 mL, 0.19 mmol) was added dropwise, and the reaction was stirred at ambient temperature for 9 h. The reaction mixture was diluted with H₂O (3 mL) and washed with CH₂Cl₂ (3 x 5 mL). The aqueous layer was concentrated *in vacuo* and purified by HPLC, eluting with a gradient of 0-95% MeCN in H₂O + 0.1% HCl over 30 min. The desired fractions were collected and lyophilized to give the chloride salt of **2.4** (10 mg, 0.05 mmol, 2% over three steps) as a light yellow solid.

¹H NMR (500 MHz, D₂O): δ 9.57 (d, *J* = 1.4 Hz, 1H), 8.80 (d, *J* = 1.3 Hz, 1H), 3.98 (app t, *J* = 6.7 Hz, 1H), 3.63–3.52 (m, 2H). ¹³C NMR (125 MHz, D₂O): δ 159.7, 155.67, 155.65, 151.9, 54.6, 36.1. HRMS (ESI⁻) *m/z* calculated for C₆H₇N₄O₂ [M-H]⁻ 167.0569, found 167.0577.

Scheme S2-2. Synthesis of triazine-maleimide probe **S-2.10**.

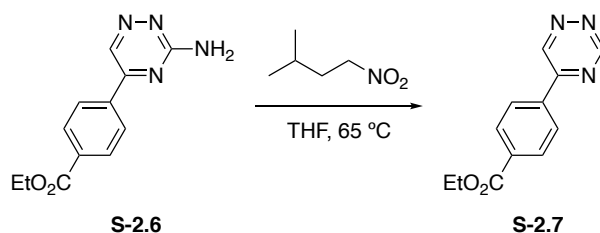


Ethyl 4-(3-amino-1,2,4-triazin-5-yl)benzoate (S-2.6)

To a round-bottom flask was added **S-2.5** (0.912 g, 4.42 mmol) and EtOH (60 mL). Aminoguanidine hydrochloride (0.491 g, 4.44 mmol) and NaHCO₃ (1.12 g, 13.3 mmol) were added, and the mixture was stirred at reflux overnight. The resulting green solution was diluted with H₂O (40 mL) and EtOAc (40 mL), and the organic layer was washed with H₂O (3 x 75 mL). The aqueous layer was extracted with EtOAc (3 x 50 mL), and the combined organic layers were dried with MgSO₄. The solution was filtered and concentrated *in vacuo*, then dry-loaded on silica and purified by flash

chromatography (eluting with 10–20% EtOAc in CH₂Cl₂) to give **S-2.6** (73.8 mg, 7%) as a bright yellow solid.

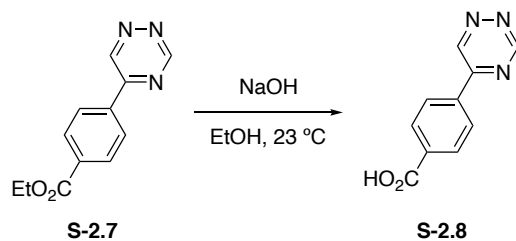
¹H NMR (400 MHz, (CD₃)₂SO): δ 9.28 (s, 1H), 8.31 (d, *J* = 7.6 Hz, 2H), 8.12 (d, *J* = 7.5 Hz, 2H), 7.38 (br s, 2H), 4.36 (q, *J* = 7.0 Hz, 2H), 1.35 (t, *J* = 7.0 Hz, 3H). ¹³C NMR (125 MHz, (CD₃)₂SO): δ 165.2, 163.1, 153.8, 138.4, 137.2, 132.4, 129.7, 127.6, 61.2, 14.2. HRMS (ESI⁺) *m/z* calcd. for C₁₂H₁₃N₄O₂ [M+H]⁺ 245.1039, found 245.1045.



Ethyl 4-(1,2,4-triazin-5-yl)benzoate (**S-2.7**)

To a flame-dried Schlenk tube under N₂ was added a solution of **S-2.6** (70.9 mg, 0.290 mmol) in anhydrous THF (10 mL). Isoamyl nitrite (0.39 mL, 2.9 mmol) was added, and the Schlenk tube was sealed and stirred at 65 °C for 16 h. The yellow solution was diluted with EtOAc (20 mL) and was washed with H₂O (2 x 30 mL) and brine (1 x 30 mL). The aqueous layer was extracted with EtOAc (1 x 30 mL), and the combined organic layers were dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude oil was dry-loaded on silica and purified by flash chromatography (eluting with 10% EtOAc in hexanes) to give **S-2.7** (27 mg, 41%) as a yellow solid.

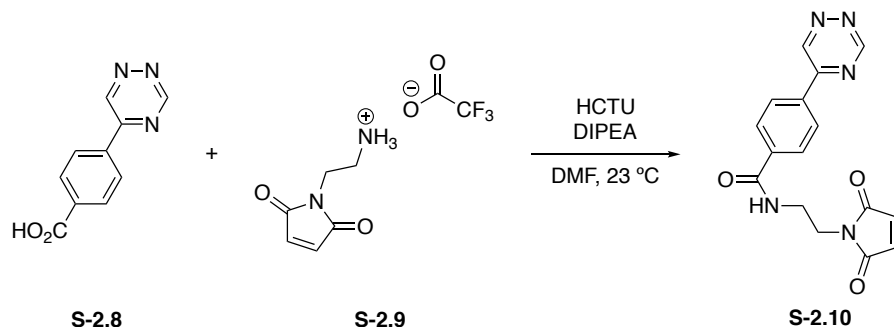
¹H NMR (400 MHz, CDCl₃): δ 9.77 (d, *J* = 2.0 Hz, 1H), 9.74 (d, *J* = 2.0 Hz, 1H), 8.27–8.22 (m, 4H), 4.43 (q, *J* = 7.1 Hz, 2H), 1.43 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (125 MHz, CDCl₃): δ 165.8, 157.7, 154.7, 147.0, 134.2, 130.7, 127.7, 61.7, 29.8, 14.4. HRMS (ESI⁺) *m/z* calcd. for C₁₂H₁₂N₃O₂ [M+H]⁺ 230.0930, found 230.0929.



4-(1,2,4-Triazin-5-yl)benzoic acid (**S-2.8**)

To a round-bottom flask was added **S-2.7** (23.6 mg, 0.102 mmol) and EtOH (5 mL). A solution of sodium hydroxide (0.52 mL of a 1 M solution) was added, and the solution was stirred at room temperature under N₂ overnight. The reaction was acidified with 1 M HCl to pH 2, then diluted with H₂O (20 mL) and EtOAc (20 mL). The aqueous layer was extracted with EtOAc (2 x 20 mL), and the combined organic layers were dried with MgSO₄. The mixture was filtered and the solvent was evaporated *in vacuo*. The crude solid was dry-loaded on silica and purified by flash chromatography (eluting with 1–5% MeOH in CH₂Cl₂) to give **S-2.8** (7.6 mg, 37%) as a beige solid.

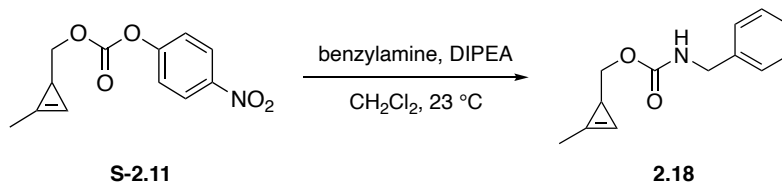
¹H NMR (400 MHz, (CD₃)₂SO): δ 13.35 (br s, 1H), 10.18 (d, *J* = 1.9 Hz, 1H), 9.86 (d, *J* = 2.0 Hz, 1H), 8.44 (d, *J* = 8.4 Hz, 2H), 8.14 (d, *J* = 8.0 Hz, 2H). ¹³C NMR (125 MHz, (CD₃)₂SO): δ 166.7, 157.3, 153.7, 147.6, 136.8, 130.0, 127.9. HRMS (ESI⁺) *m/z* calcd. for C₁₀H₆N₃O₂ [M–H][–] 200.0460, found 200.0455.



***N*-(2-(2,5-Dioxo-2,5-dihydro-1H-pyrrol-1-yl)ethyl)-4-(1,2,4-triazin-5-yl)benzamide (S-2.10)**

To an oven-dried round-bottom flask under N₂ was added a solution of **S-2.8** (5.8 mg, 0.0288 mmol) in dry DMF (0.5 mL). Diisopropylethylamine (0.03 mL, 0.2 mmol) and HCTU (15.4 mg, 0.0372 mmol) were added. After 2 min, **S-2.9** (12.3 mg, 0.0484 mmol) in dry DMF (0.25 mL) was added dropwise, and the solution was stirred under N₂ overnight. The reaction was diluted with CH₂Cl₂ (20 mL), and the organic layer was washed with 1 M LiCl (4 x 20 mL). The combined organic layers were dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude yellow solid was purified by flash chromatography (eluting with 1% MeOH in CH₂Cl₂) to give **S-2.10** (2.9 mg, 31%) as a yellow solid.

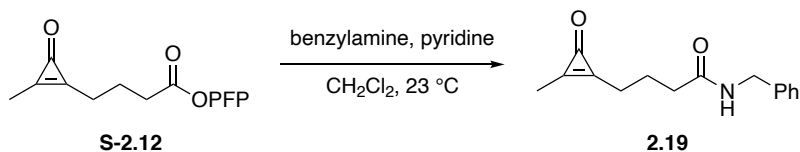
¹H NMR (600 MHz, (CD₃)₂SO): δ 10.17 (d, *J* = 1.8 Hz, 1H), 9.83 (d, *J* = 1.8 Hz, 1H), 8.78 (t, *J* = 5.9 Hz, 1H), 8.41 (d, *J* = 8.3 Hz, 2H), 7.94 (d, *J* = 8.3 Hz, 2H), 7.02 (s, 2H), 3.61 (t, *J* = 5.6 Hz, 2H), 3.44 (apparent q, *J* = 5.7 Hz, 2H). ¹³C NMR (150 MHz, (CD₃)₂SO): δ 171.1, 165.7, 157.3, 153.8, 147.6, 137.8, 135.4, 134.6, 128.0, 127.7, 37.7, 37.1. HRMS (ESI⁺) *m/z* calcd. for C₁₆H₁₃N₅O₃Na [M+Na]⁺ 346.0916, found 346.0911.



(2-Methylcycloprop-2-en-1-yl)methyl benzylcarbamate (2.18)

To a flame-dried round-bottom flask was added **S-2.11** (83 mg, 0.33 mmol) and anhydrous CH₂Cl₂ (2 mL). *N,N*-Diisopropylethylamine (DIPEA, 180 μL, 1.04 mmol) and benzylamine (110 μL, 1.04 mmol) were added, and the reaction was stirred at ambient temperature for 16 h. The mixture was diluted with CH₂Cl₂ (15 mL), then washed with saturated NaHCO₃ (3 x 20 mL) and brine (2 x 20 mL). The organic layer was dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude oil was purified on silica (eluting with 10–25% EtOAc in hexanes), then further purified by HPLC (10–90% MeCN in H₂O over 15 min). The desired fractions were combined and lyophilized to give **2.18** as a white solid (39 mg, 54%).

¹H NMR (500 MHz, CDCl₃) δ 7.35–7.27 (m, 5H), 6.56 (s, 1H), 4.96 (br s, 1H), 4.38 (d, *J* = 5.9 Hz, 2H), 4.01–3.98 (m, 1H), 3.97–3.93 (dd, *J* = 11.0, 5.2 Hz, 1H), 2.13 (s, 3H), 1.67–1.64 (m, 1H). ¹³C NMR (150 MHz, CDCl₃) δ 157.0, 138.9, 128.8, 127.7, 127.6, 102.3, 72.6, 45.2, 17.4, 11.8. HRMS (ESI⁺) *m/z* calculated for C₁₃H₁₅NO₂Na [M+Na]⁺ 240.1001, found 240.0995.



N-Benzyl-4-(2-methyl-3-oxocycloprop-1-en-1-yl)butanamide (2.19)

To a flame-dried round-bottom flask was added **S-2.12** (63 mg, 0.20 mmol) and anhydrous CH₂Cl₂ (3 mL). Anhydrous pyridine (24 μL, 0.30 mmol) and benzylamine (11 μL, 0.10 mmol) were added, and the reaction was stirred at ambient temperature overnight. The resulting solution was diluted with H₂O (20 mL), acidified to pH ~ 1 with 1 M HCl, and extracted with CH₂Cl₂ (3 x 10 mL). The organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude oil was purified on silica (eluting with 0–100% acetone in EtOAc), then further purified by HPLC (10–90% MeCN in H₂O over 15 min). The desired fractions were combined and lyophilized to give **2.19** as a white solid (11 mg, 23%).

¹H NMR (600 MHz, CDCl₃) δ 7.32–7.29 (m, 2H), 7.26–7.23 (m, 3H), 6.74 (br s, 1H), 4.41 (d, *J* = 5.8 Hz, 2H), 2.61 (t, *J* = 6.9 Hz, 2H), 2.38 (t, *J* = 7.0 Hz, 2H), 2.24 (s, 3H), 2.02 (quint, *J* = 7.0 Hz, 2H). ¹³C NMR (150 MHz, CDCl₃) δ 171.8, 160.9, 160.4, 156.9, 138.4, 128.8, 128.0, 127.6, 43.7, 35.0, 25.2, 22.1, 11.4. HRMS (ESI⁺) *m/z* calculated for C₁₅H₁₇NO₂Na [M+Na]⁺ 266.1157, found 266.1149.

2.6 References

- [1]. Rosetto, O.; Pirazzini, M.; Montecucco, C. *Nat. Rev. Microbiol.* **2014**, *12*, 535–549.
- [2]. Nelson, E. J.; Harris, J. B.; Morris, J. G., Jr. *Nat. Rev. Microbiol.* **2009**, *7*, 693–702.
- [3]. Sen, R.; Nayak, L.; De, R. K. *Eur. J. Microbiol. Infec. Dis.* **2016**, *35*, 1581–1599.
- [4]. Cornelis, G. R. *Nat. Rev. Microbiol.* **2006**, *4*, 811–825.

- [5]. Kusch, H.; Engelmann, S. *Int. J. Med. Microbiol.* **2014**, *304*, 133–141.
- [6]. Zheng, J.; Ren, X.; Wei, C.; Yang, J.; Hu, Y.; Liu, L.; Xu, X.; Wang, J.; Jin, Q. *Mol. Cell. Proteomics* **2013**, *12*, 2081–2095.
- [7]. Prescher, J. A.; Bertozzi, C. R. *Nat. Chem. Biol.* **2005**, *1*, 13–21.
- [8]. Best, M. D.; Rowland, M. M.; Bostic, H. E. *Acc. Chem. Res.* **2011**, *44*, 686–698.
- [9]. Laughlin, S. T.; Baskin, J. M.; Amacher, S. L.; Bertozzi, C. R. *Science* **2008**, *320*, 664–667.
- [10]. Patterson, D. M.; Jones, K. A.; Prescher, J. A. *Mol. BioSyst.* **2014**, *10*, 1693–1697.
- [11]. Merkel, M.; Peewasan, K.; Arndt, S.; Ploschik, D.; Wagenknecht, H.-A. *ChemBioChem* **2015**, *16*, 1541–1553.
- [12]. Dieterich, D. C.; Link, A. J.; Graumann, J.; Tirrell, D. A.; Schuman, E. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9482–9487.
- [13]. Grammel, M.; Dossa, P. D.; Taylor-Salmon, E.; Hang, H. C. *Chem. Commun.* **2012**, *48*, 1473–1474.
- [14]. Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 19–24.
- [15]. Mahdavi, A.; Szychowski, J.; Ngo, J. T.; Sweredoski, M. J.; Graham, R. L. J.; Hess, S.; Schneewind, O.; Mazmanian, S. K.; Tirrell, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 433–438.
- [16]. Link, A. J.; Vink, M. K. S.; Tirrell, D. A. *Nat. Protoc.* **2007**, *2*, 1879–1883.
- [17]. Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G. *Bioconjug. Chem.* **2010**, *21*, 1912–1916.
- [18]. Tian, H.; Sakmar, T. P.; Huber, T. *Chem. Commun.* **2016**, *52*, 5451–5454.

- [19]. Fairbanks, B. D.; Sims, E. A.; Anseth, K. S.; Bowman, C. N. *Macromolecules* **2010**, *43*, 4113–4119.
- [20]. Patterson, D. M.; Nazarova, L. A.; Prescher, J. A. *ACS Chem. Biol.* **2014**, *9*, 592–605.
- [21]. Kamber, D. N.; Liang, Y.; Blizzard, R. J.; Liu, F.; Mehl, R. A.; Houk, K. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2015**, *137*, 8388–8391.
- [22]. Kamber, D. N.; Nguyen, S. S.; Liu, F.; Briggs, J. S.; Shih, H.-W.; Row, R. D.; Long, Z. G.; Houk, K. N.; Liang, Y.; Prescher, J. A. *Chem. Sci.* **2019**, *10*, 9109–9114.
- [23]. Šečkutė, J.; Devaraj, N. K. *Curr. Opin. Chem. Biol.* **2013**, *17*, 761–767.
- [24]. Yoshikawa, E.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *Macromolecules* **1994**, *27*, 5471–5475.
- [25]. Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A. *ChemBioChem* **2002**, *3*, 235–237.
- [26]. Tu, J.; Svatunek, D.; Parvez, S.; Liu, A. C.; Levandowski, B. J.; Eckvahl, H. J.; Peterson, R. T.; Houk, K. N.; Franzini, R. M. *Angew. Chem. Int. Ed.* **2019**, *58*, 9043–9048.
- [27]. Hu, Y.; Roberts, J. M.; Kilgore, H. R.; Mat Lani, A. S.; Raines, R. T.; Schomaker, J. M. *J. Am. Chem. Soc.* **2020**, *142*, 18826–18835.
- [28]. Landgraf, P.; Antileo, E. R.; Schuman, E. M.; Dieterich, D. C. *Methods Mol. Biol.* **2015**, *1266*, 199–215.
- [29]. Mermershtain, I.; Finarov, I.; Klipcan, L.; Kessler, N.; Rozenberg, H.; Safro, M. G. *Protein Sci.* **2011**, *20*, 160–167.
- [30]. Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455–461.

- [31]. Tabanella, S.; Valancogne, I.; Jackson, R. F. W. *Org. Biomol. Chem.* **2003**, *1*, 4254–4261.
- [32]. Perez-Gonzalez, M.; Jackson, R. F. W. *Org. Synth.* **2005**, *81*, 77–87.
- [33]. Ross, A. J.; Lang, H. L.; Jackson, R. F. W. *J. Org. Chem.* **2010**, *75*, 245–248.
- [34]. Datta, D.; Wang, P.; Carrico, I. S.; Mayo, S. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2002**, *124*, 5652–5653.
- [35]. Yuet, K. P.; Doma, M. K.; Ngo, J. T.; Sweredoski, M. J.; Graham, R. L. J.; Moradian, A.; Hess, S.; Schuman, E. M.; Sternberg, P. W.; Tirrell, D. A. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112*, 2705–2710.
- [36]. Yang, N. J.; Hinner, M. J. *J. Methods Mol. Biol.* **2015**, *1266*, 29–53.
- [37]. Wang, Z.; Matthews, H. *RSC Adv.* **2020**, *10*, 11013–11023.
- [38]. Kamber, D. N.; Liang, Y.; Blizzard, R. J.; Liu, F.; Mehl, R. A.; Houk, K. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2015**, *137*, 8388–8391.
- [39]. Liang, Y.; Mackey, J. L.; Lopez, S. A.; Liu, F.; Houk, K. N. *J. Am. Chem. Soc.* **2012**, *134*, 17904–17907.
- [40]. Yang, J.; Liang, Y.; Šečková, J.; Houk, K. N.; Devaraj, N. K. *Chem. Eur. J.* **2014**, *20*, 3365–3375.
- [41]. Liu, F.; Liang, Y.; Houk, K. N. *J. Am. Chem. Soc.* **2014**, *136*, 11483–11493.
- [42]. Yang, Y.-F.; Liang, Y.; Liu, F.; Houk, K. N. *J. Am. Chem. Soc.* **2016**, *138*, 1660–1667.
- [43]. Bickelhaupt, F. M.; Houk, K. N. *Angew. Chem. Int. Ed.* **2017**, *56*, 10070–10086.
- [44]. Narayanam, M. K.; Liang, Y.; Houk, K. N.; Murphy, J. M. *Chem. Sci.* **2016**, *7*, 1257–61.

- [45]. Liu, F.; Liang, Y.; Houk, K. N. *Acc. Chem. Res.* **2017**, *50*, 2297–2308.
- [46]. Engelen, W.; van de Wiel, K. M.; Meijer, L. H. H.; Saha, B.; Merkx, M. *Chem. Commun.* **2017**, *53*, 2862–2865.
- [47]. Elliot, T. S.; Townsley, F. M.; Bianco, A.; Ernst, R. J.; Sachdeva, A.; Elsässer, S. J.; Davis, L.; Lang, K.; Pisa, R.; Greiss, S.; Lilley, K. S.; Chin, J. W. *Nat. Biotechnol.* **2014**, *32*, 465–472.
- [48]. Italia, J. S.; Addy, P. S.; Erickson, S. B.; Peeler, J. C.; Weerapana, E.; Chatterjee, A. *J. Am. Chem. Soc.* **2019**, *141*, 6204–6212.
- [49]. Patterson, D. M.; Nazarova, L. A.; Xie, B.; Kamber, D. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2012**, *134*, 18638–18643.
- [50]. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785–2791.
- [51]. Hammill, J. T.; Miyake-Stoner, S.; Hazen, J. L.; Jackson, J. C.; Mehl, R. A. *Nat. Protoc.* **2007**, *2*, 2601–2607.
- [52]. Pédelacq, J.-D.; Cabantous, S.; Tran, T.; Terwilliger, T. C.; Waldo, G. S. *Nat. Biotechnol.* **2006**, *24*, 79–88.
- [53]. de Almeida, G.; Sletten, E. M.; Nakamura, H.; Palaniappan, K. K.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2012**, *51*, 2443–2447.
- [54]. Ernd, M.; Heuchmann, M.; Zipse, H. *Helv. Chim. Acta* **2005**, *88*, 1491–1518.
- [55]. Kamber, D. N.; Nazarova, L. A.; Liang, Y.; Lopez, S. A.; Patterson, D. M.; Shih, H.-W.; Houk, K. N.; Prescher, J. A. *J. Am. Chem. Soc.* **2013**, *135*, 13680–13683.
- [56]. Shao, H.; Lockman, J. W.; Parquette, J. R. *J. Am. Chem. Soc.* **2007**, *129*, 1884–1885.

- [57]. Atmuri, P. N. D.; Lubell, W. D. *Org. Synth.* **2015**, *92*, 103–116.
- [58]. Yang, J.; Karver, M. R.; Li, W.; Sahu, S.; Devaraj, N. K. *Angew. Chem. Int. Ed.* **2012**, *51*, 5222–5225.
- [59]. Darwish, T. A.; Evans, R. A.; James, M.; Malic, N.; Triani, G.; Hanley, T. L. *J. Am. Chem. Soc.* **2010**, *132*, 10748–10755.
- [60]. Gröst, C.; Berg, T. *Org. Biomol. Chem.* **2015**, *13*, 3866–3870.
- [61]. Bertozzi, C. R.; Agard, N. J.; Prescher, J. A.; Baskin, J. M.; Sletten, E. M. Compositions and methods for modification of biomolecules. U.S. Patent 9,260,371, February 16, 2016.
- [62]. Zhuo, J.; Zhang, C.; Xu, M.; Qian, D. Q.; Yao, W.; Jalluri, R. K. Triazolotriazines as kinase inhibitors. U.S. Patent 7,862,060, March 23, 2010.
- [63]. Row, R. D.; Prescher, J. A. *Org. Lett.* **2018**, *20*, 5614–5617.

Chapter 3: Chemically triggered crosslinking with bioorthogonal cyclopropenones

Adapted from published work, with permission from the Royal Society of Chemistry: Row, R. D.; Nguyen, S. S.; Ferreira, A. J.; Prescher, J. A. *Chem. Commun.* **2020**, 56, 10883–10886.

3.1 Abstract

The bioorthogonal ligation between cyclopropenones (CpOs) and phosphines features a key ketene-ylide intermediate. In the first iteration of this chemistry, the Prescher lab demonstrated that this electrophile could be trapped in an intermolecular fashion with exogenous amines. The reaction only proceeded when large concentrations of the nucleophilic trap were present. While this was initially seen as a limitation, we began to wonder whether this requirement for intermolecular trapping could be exploited for capturing biomolecule interactions. The high concentrations of trapping nucleophile required for effective ligation could be achieved locally between two interacting partners. In this Chapter, I highlight how CpOs were optimized as chemically triggered crosslinkers to capture protein-protein interactions.

3.2 Introduction

Covalent crosslinkers are valuable tools for examining biomolecule interactions in physiologically relevant environments. Such tools typically fall into two categories: photocrosslinkers [1] and chemical crosslinkers. Popular photocrosslinkers include diazirines [2-3], aryl azides [4-5], benzophenones [6-7], and aryl carboxytetrazoles [8]. These groups are routinely used *in vitro* to monitor interactions between proteins and other biomolecules [9-15]. When exposed to intense UV light, the motifs photolyze to

provide high-energy intermediates; these species can be trapped by neighboring biomolecules to forge covalent adducts. UV light can be delivered to samples on demand, enabling both spatially and temporally controlled crosslinking. Furthermore, the small size of most photocrosslinkers ensures that they are compatible with a variety of cellular targets and pathways for installation [9, 11-12, 16-18].

While widely employed, photocrosslinkers also have limitations. Many suffer from high levels of background signal due to off-target labeling [19]. Additionally, the requisite UV light precludes applications in thick tissues and live organisms—environments that are either refractory to light delivery or sensitive to irradiation [20-21]. These issues can be avoided using chemical crosslinkers, such as α -haloacetamides or other chemical warheads for covalent trapping [22-28]. However, such electrophiles are permanently “on” and not responsive to external stimuli, contributing to non-specific labeling. Efforts to tune the specificity of these probes have come at the expense of versatility, as less potent electrophiles react with a more limited set of nucleophiles [26, 29].

To develop more general and triggerable crosslinkers, we investigated cyclopropanones (CpOs) as *chemically* activated motifs. These scaffolds undergo bioorthogonal reactions with functionalized phosphines [30-31]. The reaction proceeds through a ketene-ylide intermediate, which can be trapped by a variety of nucleophiles to produce covalent adducts (Figure 3-1). Nucleophile trapping is typically accomplished in an *intramolecular* fashion using ortho-substituted phosphines. We hypothesized that *intermolecular* trapping would afford biomolecule crosslinks. Upon phosphine treatment, CpOs positioned near targets of interest would be activated for covalent adduct formation.

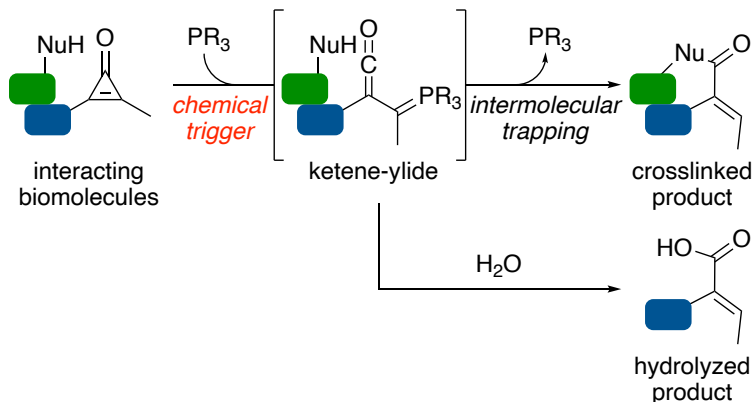


Figure 3-1. Cyclopropanones as chemically triggered crosslinkers. Cyclopropanones generate ketene-yldes upon treatment with bioorthogonal phosphines. These intermediates can be trapped by neighboring nucleophiles to form covalent crosslinks or hydrolyze to give innocuous byproducts.

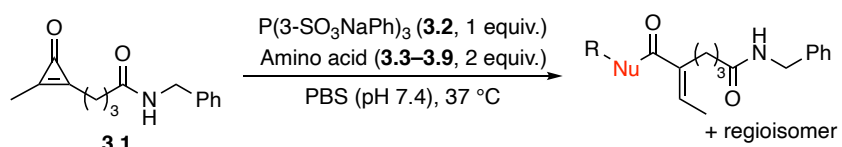
We previously showed that ketene-yldes form upon phosphine incubation with CpOs, and that such intermediates can be trapped with alkylamines [30]. Large concentrations (>10 mM) of nucleophile were required to outcompete ketene hydrolysis (Figure 3-1). We hypothesized that such features would be ideal for biomolecule crosslinking. High local concentrations of trapping residues would only be achieved with interacting biomolecules. In all other cases, hydrolysis would dominate, generating an innocuous byproduct and minimizing non-specific crosslinking. Most protein binding interfaces also comprise residues amenable to trapping electrophiles and thus affording stable crosslinks [32].




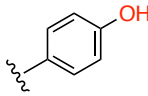

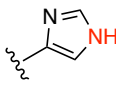
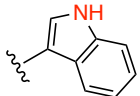
3.3 Results and Discussion

To test whether a broad range of residues could suffice for *intermolecular* trapping, Dr. Drew Ferreira from the group incubated a model cyclopropanone (CpO **3.1**, Table 3-1) with a panel of amino acids. CpO **3.1** was activated for trapping via treatment with a water-soluble triarylphosphine (TPPTS, **3.2**), and the reactions were monitored by LCMS.

Crosslinked products were observed in the presence of Ser, Cys, Lys, and Tyr (Figures 3-2–3-5), but only when excess amino acid was used (Table 3-1). The hydrolyzed product was the major species formed in each case. These results suggested that CpO-mediated crosslinks could be forged at protein interfaces, where local concentrations of amino acids can exceed 100 mM [33].

Table 3-1. Phosphine-Mediated Crosslinking of CpO with Amino Acids.



Entry	Amino acid	Nucleophile (Nu)	% crosslink	
			500 μ M	4 mM
1	Fmoc-L-Ser-OH 3.3		<1%	1%
2	Fmoc-L-Cys-OH 3.4		5% ^a	8% ^a
3	Fmoc-L-Lys-OH 3.5		<1%	2%
4	Fmoc-L-Tyr-OH 3.6		0%	1%
5	Fmoc-L-Ala-OH 3.7		0%	N.D.
6	Fmoc-L-His-OH 3.8		0%	N.D.
7	Fmoc-L-Trp-OH 3.9		0%	N.D.

^a Disulfide formation observed over course of reaction

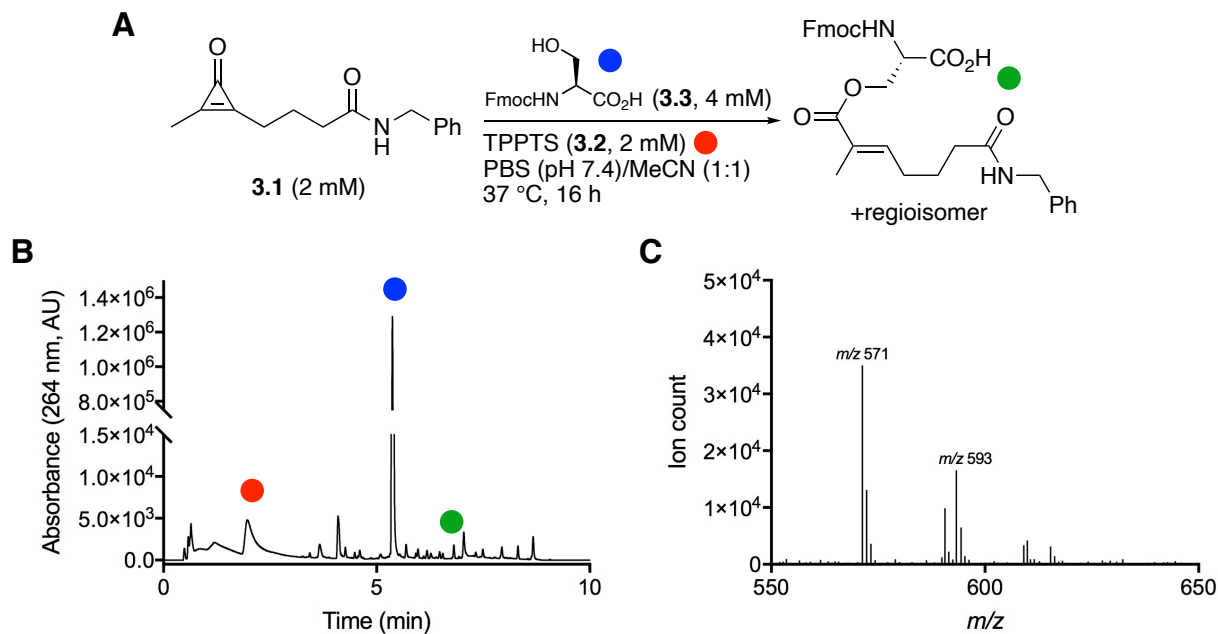


Figure 3-2. Chemically triggered CpO trapping with serine. (A) CpO **3.1** was incubated with phosphine **3.2** and Fmoc-protected serine **3.3**. Conversion to the crosslinked product was monitored via (B) UV absorption spectroscopy and (C) mass spectrometry.

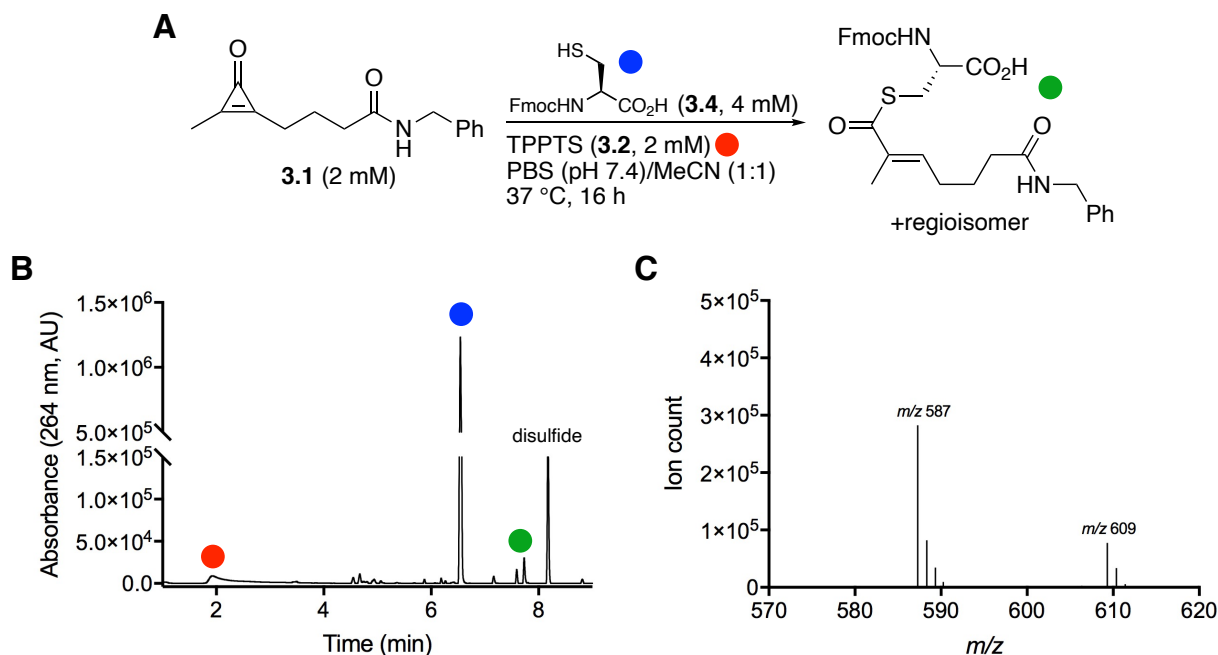


Figure 3-3. Chemically triggered CpO trapping with cysteine. (A) CpO **3.1** was incubated with phosphine **3.2** and Fmoc-protected cysteine **3.4**. Conversion to the crosslinked product was monitored via (B) UV absorption spectroscopy and (C) mass spectrometry.

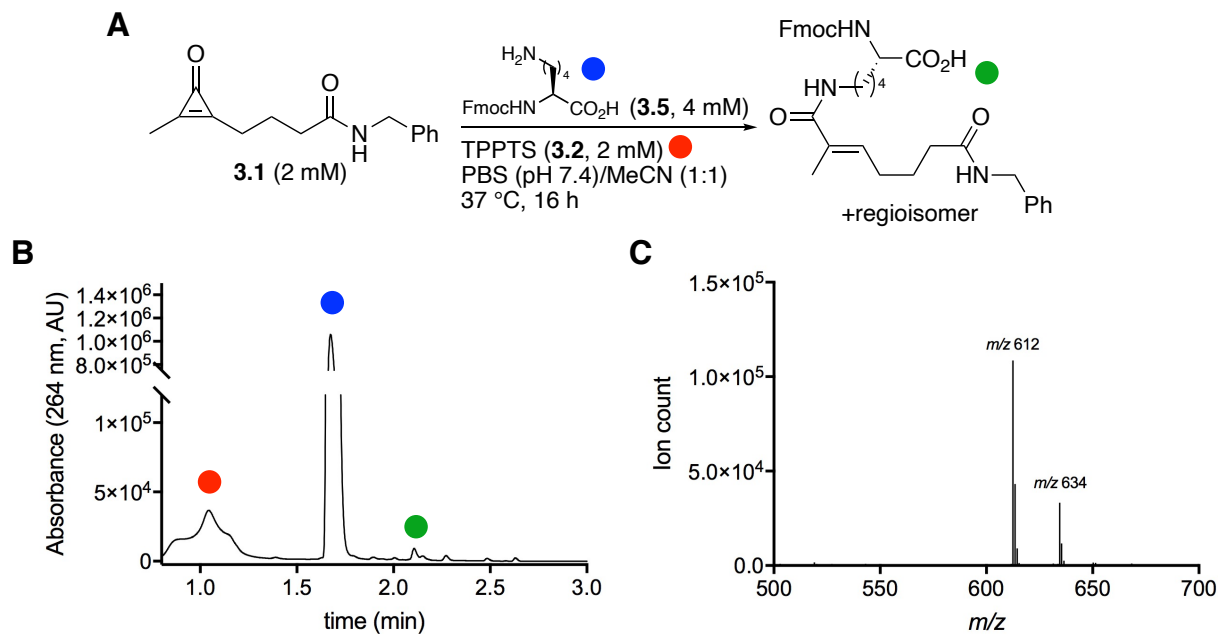


Figure 3-4. Chemically triggered CpO trapping with lysine. (A) CpO **3.1** was incubated with phosphine **3.2** and Fmoc-protected lysine **3.5**. Conversion to the crosslinked product was monitored via (B) UV absorption spectroscopy and (C) mass spectrometry.

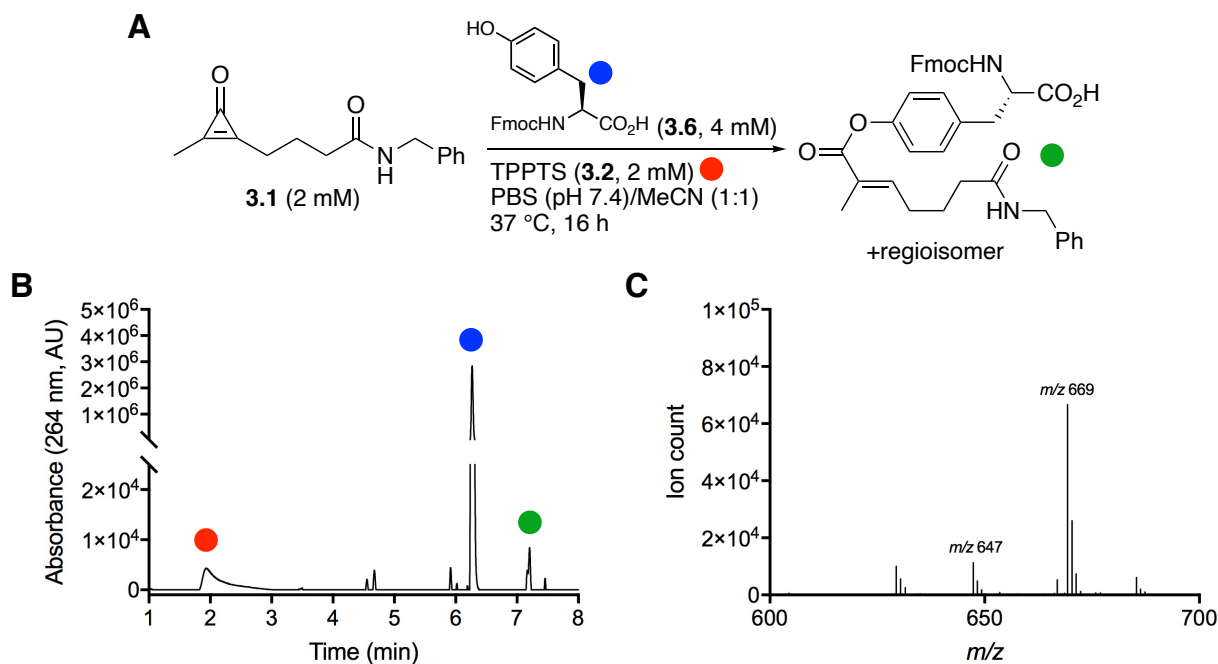


Figure 3-5. Chemically triggered CpO trapping with tyrosine. (A) CpO **3.1** was incubated with phosphine **3.2** and Fmoc-protected tyrosine **3.6**. Conversion to the crosslinked product was monitored via (B) UV absorption spectroscopy and (C) mass spectrometry.

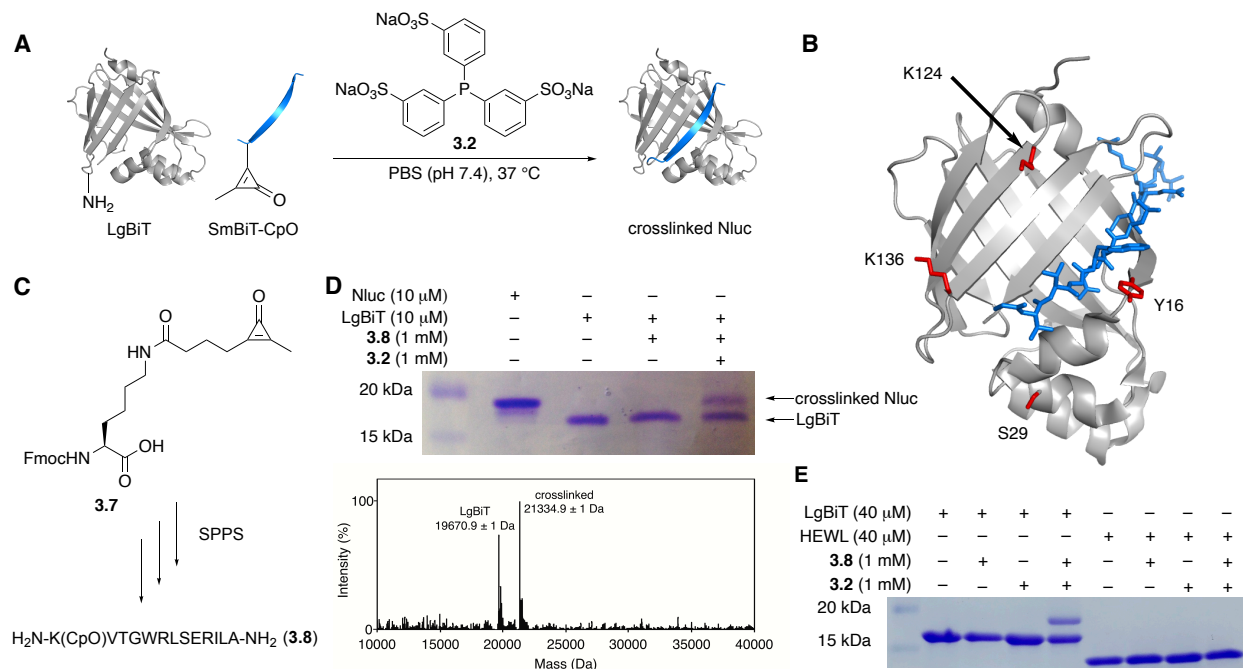


Figure 3-6. Chemically triggered crosslinking of split luciferase probes. (A) Bioorthogonal crosslinking of LgBiT and SmBiT to form Nanoluciferase (Nluc). (B) Crystal structure of Nluc (PDB 5IBO), with SmBiT highlighted in blue. Potential trapping residues are shown in red. (C) CpO-functionalized SmBiT-CpO **3.8** was prepared via solid-phase peptide synthesis. (D) Covalent crosslinking of LgBiT and SmBiT-CpO **3.8** with phosphine **3.2**. Crosslinked Nluc was observed by SDS-PAGE analysis (top) and mass spectrometry (bottom). The crosslinking yield (33%) was determined by ImageJ analysis. (E) SmBiT-CpO **3.8** does not form covalent adducts with off-target proteins. LgBiT (40 μ M) or HEWL (40 μ M) and SmBiT-CpO **3.8** (1 mM) were incubated in PBS (pH 7.4) in the presence of phosphine **3.2** (1 mM) at 37 $^{\circ}$ C for 10 h. Samples were analyzed by SDS-PAGE. The crosslinking yield (31%) was determined by ImageJ analysis.

We proceeded to examine CpO crosslinking using a model protein interaction: split Nanoluciferase (Nluc, Fig. 3-6A) [34]. Split Nluc comprises a short C-terminal peptide (SmBiT) and a larger engineered fragment (LgBiT). LgBiT and SmBiT bind readily and can reconstitute the full-length, light-emitting enzyme in aqueous buffers. Examining the Nluc crystal structure, we hypothesized that CpO appended to the N-terminus of SmBiT could access a handful of potential trapping residues when bound to LgBiT (Figure 3-6B) [35]. To prepare the functionalized peptide, alumnus Dr. David Row first synthesized CpO amino acid **3.7**. This Fmoc-protected variant could be prepared in gram-scale quantities, and was amenable to solid-phase peptide synthesis. CpO **3.7** was installed at the N-terminal position of SmBiT to provide the desired conjugate (SmBiT-CpO **3.8**,

Figure 3-6C). Importantly, SmBiT-CpO **3.8** was still capable of binding LgBiT to produce a light-emitting enzyme (Figure 3-7) [34].

Dr. David Row then demonstrated that the functionalized peptide could form a covalent crosslink with LgBiT. LgBiT and peptide **3.8** were incubated in phosphate buffered saline (PBS, pH 7.4) and treated with phosphine **3.2**. SDS-PAGE analysis revealed a higher molecular weight band, indicating that crosslinked Nluc was formed. The identity of the product was verified via mass spectrometry (Figure 3-6D). Hydrolysis products were also observed after 3 h, along with cyclized SmBiT (due to *intramolecular* trapping). Unreacted SmBiT-CpO was still present at this point, though, and available for crosslinking.

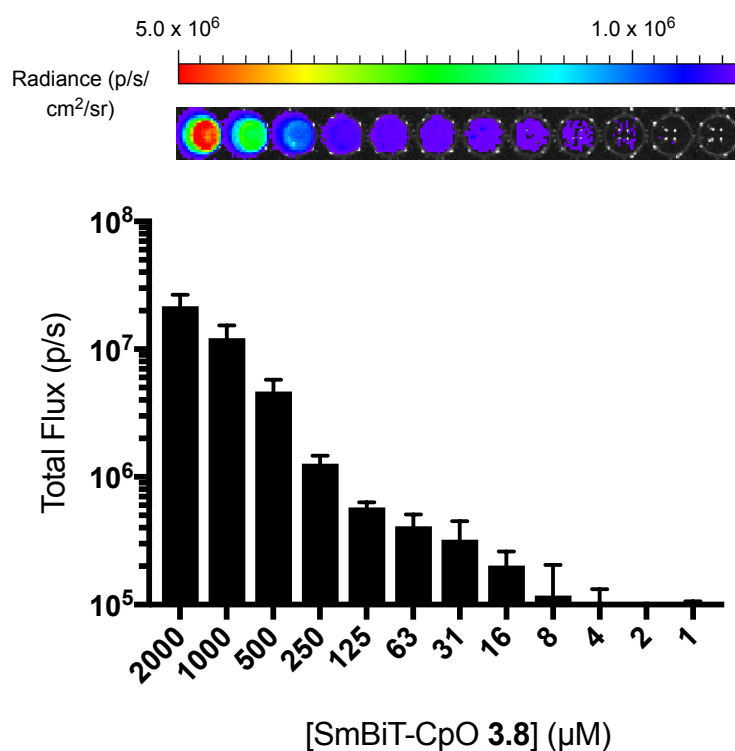


Figure 3-7. SmBiT-CpO **3.8** and LgBiT associate to form a functional, light-emitting enzyme. Furimazine (50X dilution, Promega) was added to wells containing SmBiT-CpO **3.8** (1–2000 μM) and LgBiT (40 μM). Light emission was quantified and error bars represent the standard error of the mean for $n = 3$ experiments. Sample images are shown.

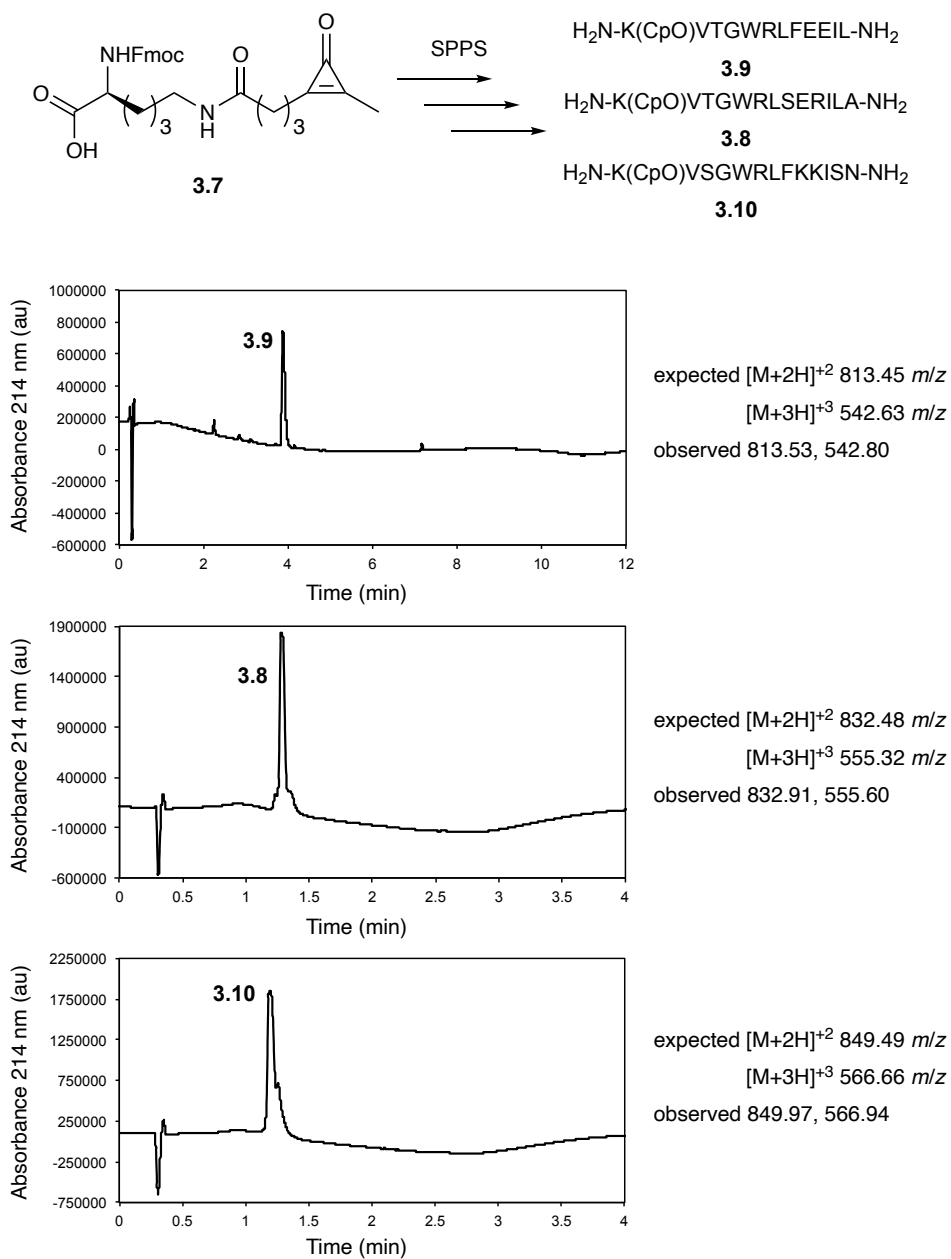


Figure 3-8. LC-MS characterization of SmBiT-CpO peptides **3.9**, **3.8**, and **3.10**. The peptides were synthesized on solid phase and purified via HPLC.

To test whether crosslinking was specific to interacting fragments, we performed the reaction in the presence of a non-target protein, hen egg-white lysozyme (HEWL). HEWL is similar in size to LgBiT. Both proteins also comprise >20 surface-exposed nucleophiles that could potentially trap the activated cyclopropanone [36]. When SmBiT-

CpO **3.8** was incubated with HEWL, though, no crosslinked products were observed upon phosphine treatment (Figure 3-6E). These data suggest that the crosslinking strategy can provide a selective readout on productive interactions.

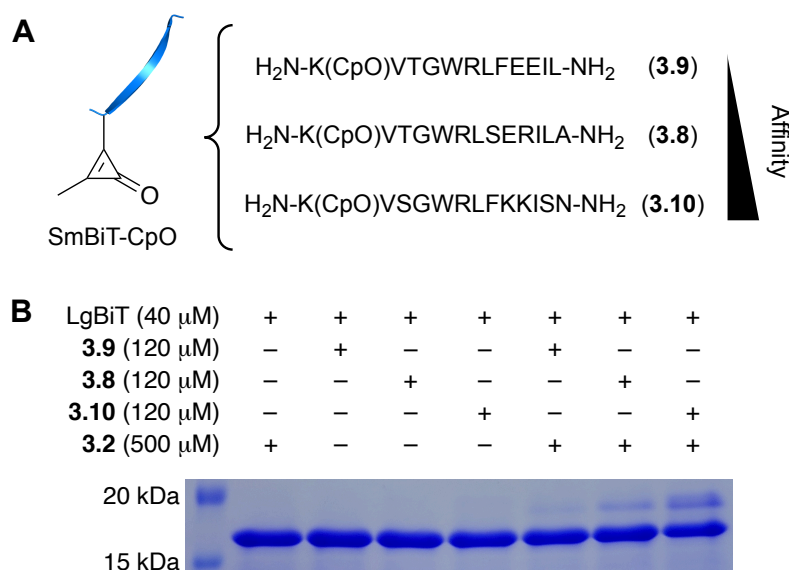


Figure 3-9. Phosphine-triggered crosslinking tracks with affinity. (A) SmBiT-CpO peptides (covering a range of LgBiT affinities [34]) were accessed via SPPS. (B) Crosslinking is most pronounced with high affinity interactions. LgBiT (40 μM) and SmBiT-CpO peptides (120 μM) were incubated with phosphine **3.2** (500 μM) at 37 $^\circ\text{C}$ for 4 h. Samples were analyzed by SDS-PAGE. The crosslinking yields (3-20%) were determined by ImageJ analysis.

We further examined whether the approach could distinguish among interactions of variable affinity. We synthesized two additional SmBiT-CpO peptides (**3.9** and **3.10**, Figure 3-8) comprising sequences known to exhibit different degrees of binding to LgBiT [34]. We hypothesized that the range of binding affinities would manifest in variable degrees of crosslinking. The SmBiT peptides were mixed with LgBiT, and the CpO residues were activated with phosphine **3.2**. SDS-PAGE analysis revealed a clear trend in crosslinking efficiency (Figure 3-9). The highest affinity peptide (**3.10**) provided the largest number of LgBiT crosslinks, while no adducts were observed with the weakest

binder (**3.9**). Importantly, SmBiT-CpO **3.10** (i.e., the highest affinity peptide) also did not form covalent crosslinks with off-target proteins (Figure 3-10).

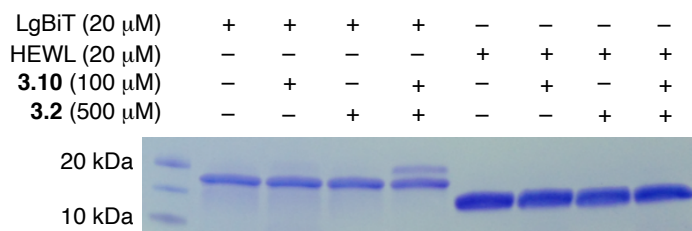


Figure 3-10. SmBiT-CpO **3.10** does not form covalent adducts with HEWL. LgBiT (20 μ M) or HEWL (20 μ M), SmBiT-CpO **3.10** (100 μ M), and phosphine **3.2** (500 μ M) were added to PBS (pH 7.4). Samples were incubated at 37 $^{\circ}$ C for 16 h, then denatured with urea (1.6 M final concentration) and analyzed by SDS-PAGE. The crosslinking yield (20%) was determined by ImageJ analysis.

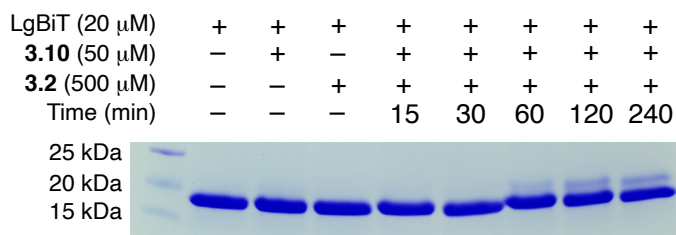


Figure 3-11. Crosslinking is time-dependent with phosphine **3.2**. LgBiT (20 μ M) and SmBiT-CpO **3.10** (50 μ M) were combined in PBS (pH 7.4). Phosphine **3.2** (500 μ M) was added, and the reactions were incubated at 37 $^{\circ}$ C for 15–240 min. Reactions were then quenched with H₂O₂ (1% final concentration) prior to SDS-PAGE analysis. Crosslinking yields (1–4%) were determined by ImageJ analysis.

While split luciferase adducts were captured upon CpO activation, maximal crosslinking was not achieved until 4 h post-phosphine addition (Figure 3-11). We hypothesized that crosslinking speed could be improved by tuning the phosphine trigger. Alkyl substituents are well known to increase the nucleophilicity of phosphines, but can render the probes susceptible to oxidation [37]. However, monocyclohexyl diaryl phosphines are bench stable and have been used in biological environments [31].

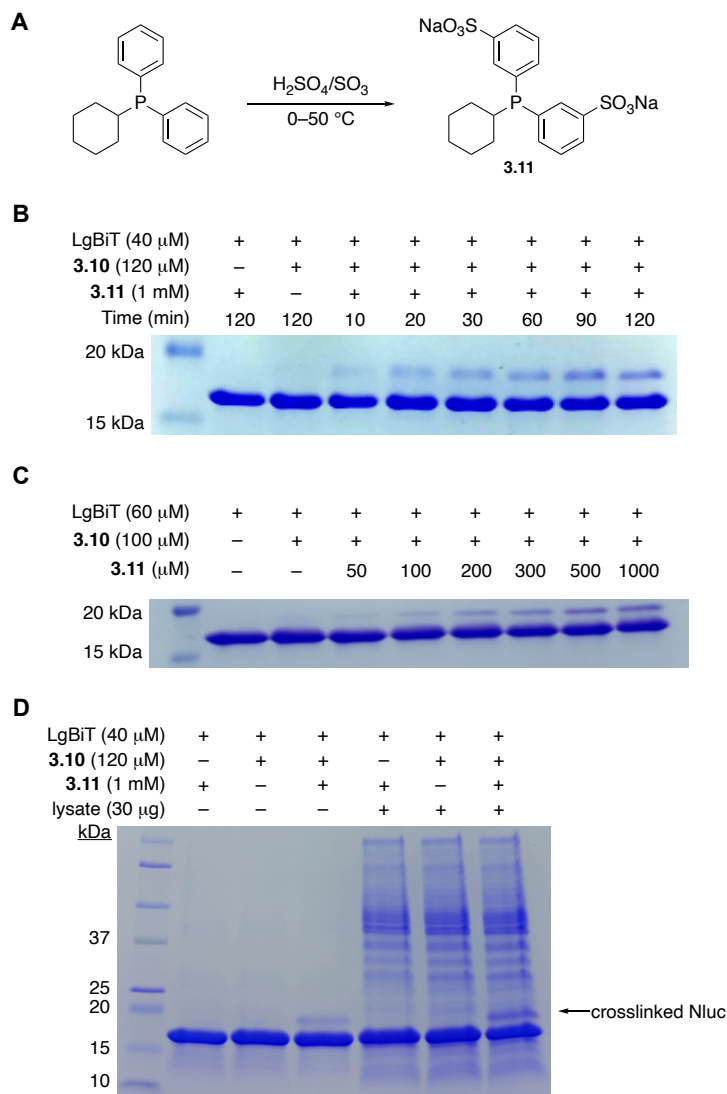


Figure 3-12. Improved crosslinking with an optimized trigger. (A) Phosphine **3.11** was designed as a more nucleophilic, water-soluble probe. (B) Rapid crosslinking was observed with optimized phosphine. LgBiT (40 μ M) and SmBiT-CpO **3.10** (120 μ M) were incubated with phosphine **3.11** (1 mM). Samples were analyzed by SDS-PAGE. The crosslinking yields (5-16%) were determined by ImageJ analysis. (C) Crosslinking reactions were dose-dependent with phosphine **3.11**. (D) Crosslinking in cell lysate. LgBiT (40 μ M) and SmBiT-CpO **3.10** (120 μ M) were incubated in PBS (pH 7.4) containing bacterial lysate (30 μ g) and phosphine **3.11** (1 mM). Samples were incubated at 37 $^{\circ}$ C for 30 min, then analyzed by SDS-PAGE. The crosslinking yields in the presence (13%) and absence (5%) of lysate were determined by ImageJ analysis.

Thus, we reasoned that phosphine **3.11** would strike a balance between increased reactivity and stability toward oxidation, and also be water-soluble (Figure 3-12A) [38-40].

Phosphine **3.11** was easily accessed from commercially available starting materials using standard sulfonation chemistries. When **3.11** was used to activate SmBiT-CpO **3.10** in

the presence of LgBiT, crosslinked products were observed in as little as 10 minutes (Figure 3-12B). The probe could also activate CpO motifs for crosslinking in a dose-dependent manner (Figure 3-12C). No adducts were observed with triarylphosphine **3.2** over the same time period, and prolonged reaction times were required for robust crosslinking signal (Figure 3-13A).

The crosslinking approach was further examined in bacterial cell lysate. In this complex environment, many proteins could potentially outcompete LgBiT for trapping activated CpOs. LgBiT and SmBiT-CpO **3.10** were added to freshly prepared lysate, then treated with phosphine **3.11**. The expected crosslinked product was observed after 30 min of incubation at 37 °C (Fig. 3-12D). Interestingly, SDS-PAGE analysis suggested that crosslinking was more efficient in lysate than in buffer alone. A similar result was observed with the less nucleophilic phosphine **3.2** (Figure 3-13B). Macromolecular crowding likely facilitates biomolecular association—and thus more effective crosslinking—in heterogeneous environments [41].

Having demonstrated successful crosslinking with split luciferase probes, we were curious about the nature of the covalent linkage. We initially suspected LgBiT residues K124 or K136 were involved in CpO trapping (Figure 3-6B). However, when these sites were mutated to alanine, no decrease in crosslinking efficiency was observed (Figure 3-14). Mass spectrometry analyses of tryptic digests revealed LgBiT fragment (residues 16–64) involved in the trap (Figure 3-15). Further attempts to localize the crosslink (via MS/MS analysis) were inconclusive, though, possibly due to non-specific cleavage of the crosslink. Within the LgBiT fragment, Tyr16, Ser28, and Ser29 were the most likely traps based on predicted proximity to bound SmBiT (Fig. 3-6B). We prepared LgBiT mutants

Y16F, S28A, and S29A to probe their participation in CpO trapping. Only mutant Y16F exhibited reduced crosslinking efficiency not attributed to impaired LgBiT binding (Figures 3-16 and 3-17). Few chemical crosslinkers target tyrosine or other less potent nucleophiles, highlighting the versatility of our approach.

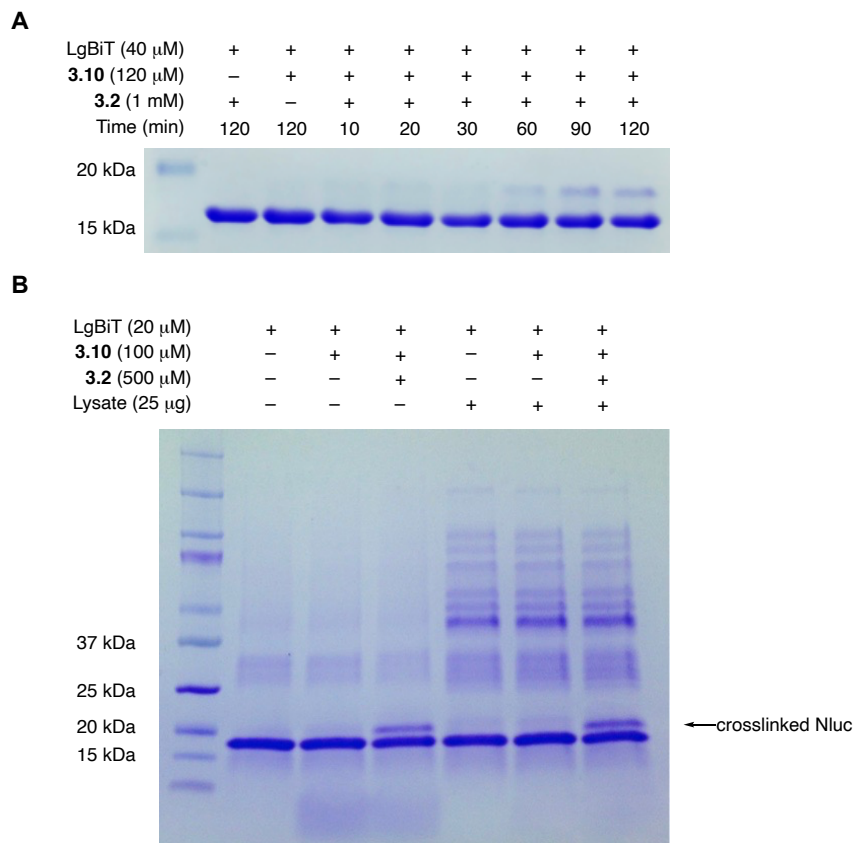


Figure 3-13. (A) Slower crosslinking reactions were observed with less nucleophilic phosphine **3.2**. LgBiT (40 μ M) and SmBiT-CpO **3.10** (120 μ M) were combined in PBS (pH 7.4). Phosphine **3.2** (1 mM) was added, and the reactions were incubated at 37 °C for 10–120 min. Reaction samples were denatured with urea (1.6 M final concentration) prior to SDS-PAGE analysis. (B) Crosslinking observed in bacterial lysate using phosphine **3.2**. LgBiT (20 μ M) and SmBiT-CpO **3.10** (100 μ M) were incubated in the presence or absence of bacterial lysate (25 μ g) and phosphine **3.2** (500 μ M). Reactions were incubated at 37 °C for 4 h, then denatured with urea (1.6 M final concentration) and analyzed by SDS-PAGE. Crosslinking yields in the presence (20%) and absence (19%) of lysate were determined by ImageJ analysis.

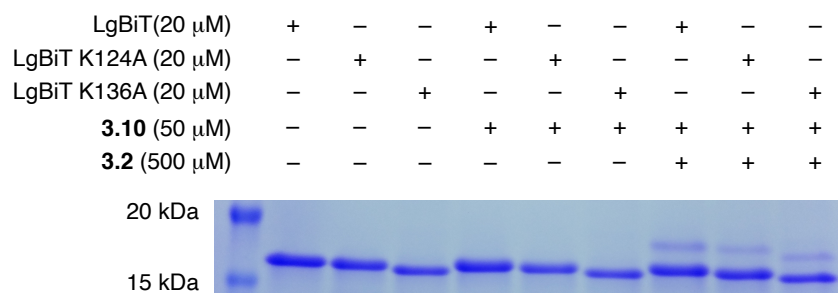


Figure 3-14. LgBiT residues K124 and K136 are not necessary for crosslinking. WT LgBiT, or LgBiT mutants K124A or K136A (20 μ M) were combined with SmBiT-CpO **3.10** (50 μ M) and phosphine **3.2** (500 μ M). Samples were incubated at 37 $^{\circ}$ C for 4 h, then denatured with urea (1.6 M final concentration) and analyzed by SDS-PAGE. Crosslinked products were observed with both mutants. The crosslinking yields for LgBiT (15%) and the K124A (11%) and K136A (14%) mutants were determined by ImageJ analysis.

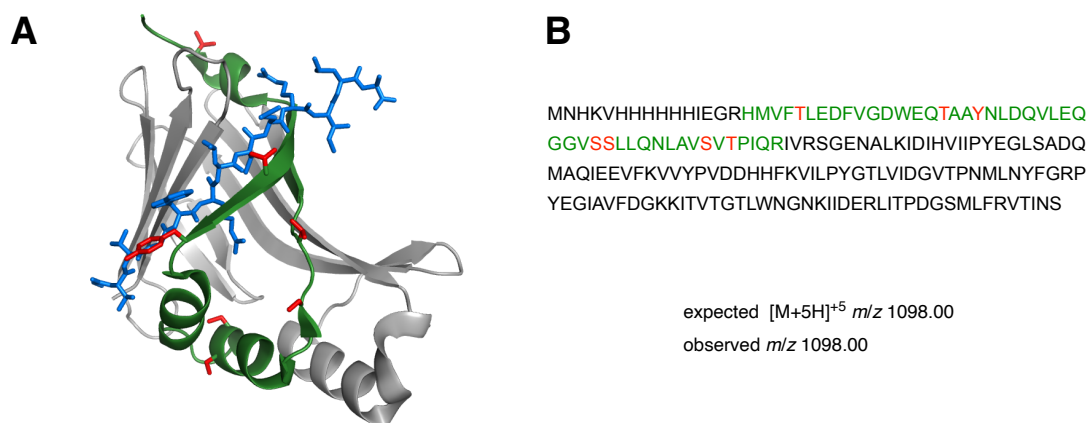


Figure 3-15. Site of Nluc crosslink analyzed via tryptic digest. (A) A peptide matching residues 1–43 (green) showed a mass shift corresponding to the addition of the first 6 residues of SmBiT-CpO peptide **3.10** (blue). Residues bearing potential trapping sidechains are highlighted in red. (B) Primary sequence of Nluc in the pCOLD construct. The residues 16–64 are highlighted in green. Potential nucleophilic side chains within the sequence are highlighted in red.

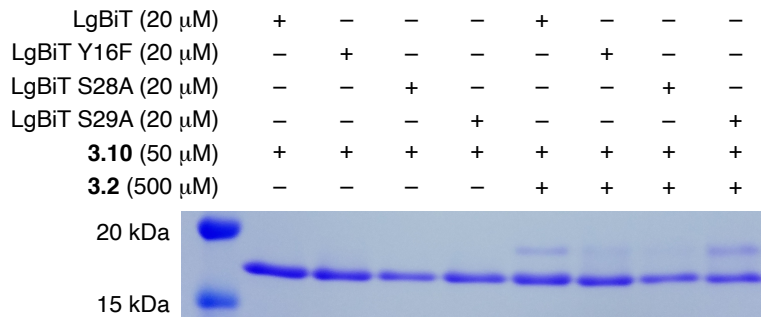


Figure 3-16. Diminished crosslinking is observed with LgBiT mutants Y16F and S28A. WT LgBiT, or LgBiT mutants comprising Y16F, S28A, or S29A (20 μ M) were combined with SmBiT-CpO **3.10** (50 μ M) and phosphine **3.2** (500 μ M), then incubated at 37 $^{\circ}$ C for 4 h. Samples were denatured with urea (1.6 M final concentration), then analyzed by SDS-PAGE. The crosslinking yields for LgBiT (12%) and the Y16F (6%), S28A (5%), and S29A (22%) mutants were determined by ImageJ analysis.

The CpO crosslinking strategy was also compared to a common photocrosslinking approach with diazirine (Dz) motifs. Upon UV irradiation, diazirines form carbenes that can react with a variety of protein residues to afford covalent adducts. A diazirine-functionalized SmBiT peptide (SmBiT-Dz, **3.12**) was synthesized. The Dz linker was one carbon shorter than the corresponding CpO probe, but was pursued based on the synthetic tractability of a precursor amino acid (**S-3.1**). SmBiT-Dz (**3.12**) was incubated with LgBiT, and the sample was irradiated with UV light (312 nm). Crosslinked adducts were observed via SDS- PAGE analysis (Figure 3-18A). Prolonged irradiation times resulted in more crosslinked adducts, but also promoted LgBiT degradation. The crosslinking efficiency of SmBiT-Dz was also examined alongside SmBiT-CpO (**3.10**). Each peptide was incubated with LgBiT. Samples containing SmBiT-CpO were exposed to phosphine **3.11** for chemical crosslinking, while samples containing SmBiT-Dz were irradiated with UV light for photocrosslinking. Crosslinked products were observed in both cases (Figure 3-18B), with the chemically triggered reaction providing more adducts.

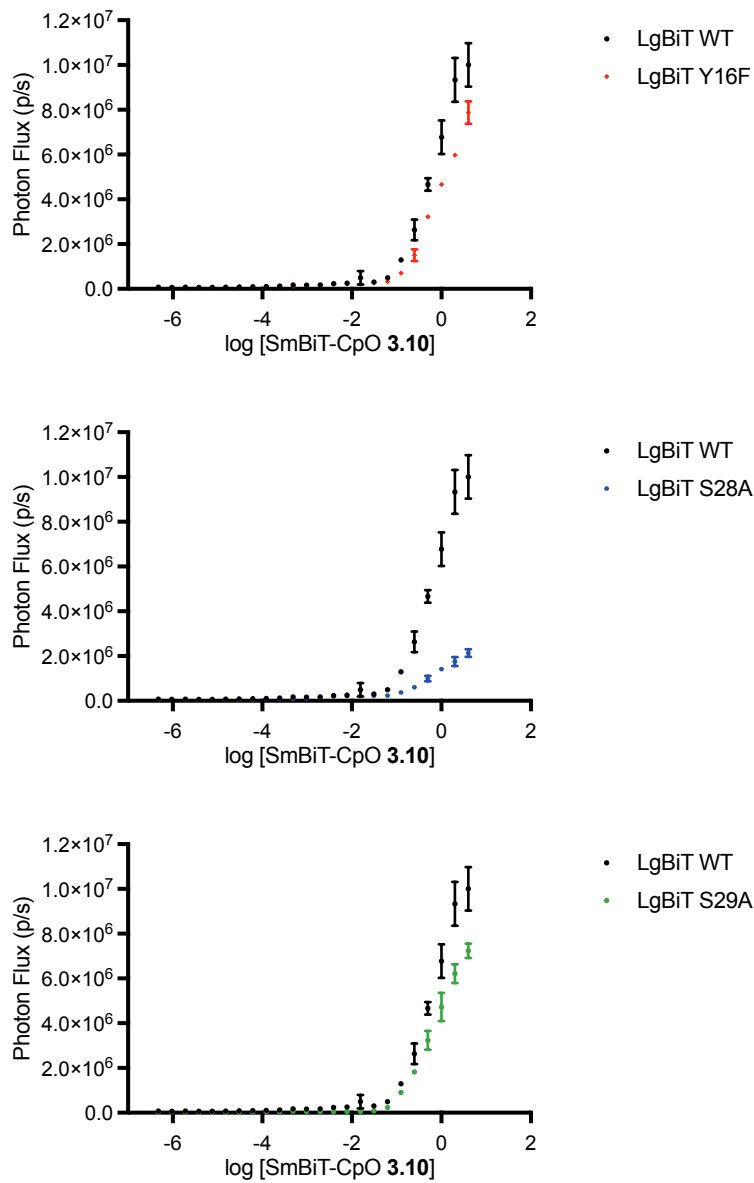


Figure 3-17. Dissociation constants for SmBiT-CpO **3.10** and either native (WT) or mutant LgBiTs. Apparent K_D values for mutants Y16F and S29A are comparable to WT. Error bars represent standard error of the mean for $n = 3$ experiments.

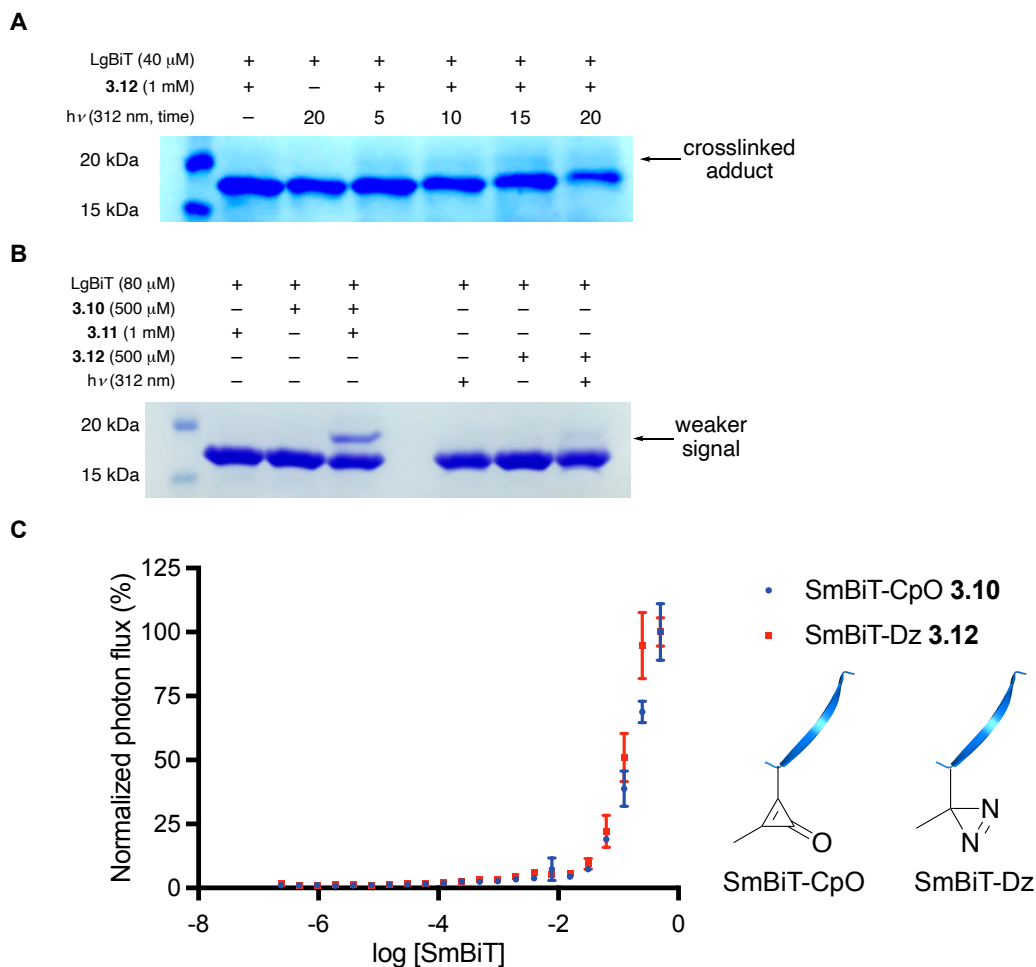


Figure 3-18. (A) Prolonged UV irradiation promotes photocrosslinking. LgBiT (40 μ M) was combined with SmBiT-Dz **3.12** (1 mM) in PBS (pH 7.4). The samples were spun down and left to equilibrate (10 min) at room temperature. The samples were then placed on a transilluminator (312 nm) for the indicated time. Crosslinking was further confirmed via mass spec analysis. LgBiT degradation was observed upon prolonged UV exposure. Crosslinking yields ranged from 1–5%, as determined by ImageJ analysis. (B) CpO chemical crosslinking is more efficient than diazirine photocrosslinking. LgBiT (80 μ M) was combined with either SmBiT-CpO **3.10** (500 μ M) or SmBiT-Dz **3.12** (500 μ M) in PBS (pH 7.4). For chemical crosslinking samples (lanes 2–4), phosphine **3.11** (1 mM) was added, and the reactions were incubated at 37 $^{\circ}$ C for 30 min. For photocrosslinking reactions (lanes 6–8), samples were irradiated with UV light (312 nm) for 10 min. All reaction samples were denatured with urea (1.6 M final concentration) prior to SDS-PAGE analysis. The crosslinking yields for SmBiT-CpO (22%) and SmBiT-Dz (1%) were determined by ImageJ analysis. (C) Dissociation constants for native (WT) LgBiT and either SmBiT-CpO **3.10** or SmBiT-Dz **3.12**. Apparent K_D values for both SmBiT peptides are comparable. Error bars represent standard error of the mean for $n = 3$ experiments.

While quantitative comparisons cannot be made due to differences in probe structure, SmBiT-Dz and SmBiT-CpO exhibited similar affinities for LgBiT (Figure 3-18C), suggesting that the observed differences in crosslinking were not simply due to altered binding interactions. They could instead be due to differences in probe reactivity. Carbenes are formed irreversibly upon diazirine photolysis. These high-energy intermediates exhibit relatively short half-lives (ns– μ s) and can insert into a variety of different bonds [2]. As a consequence, carbenes can trap even weakly associated off-target molecules, leading to false positives. Ketene-ylides, by contrast, are formed reversibly from cyclopropanones and phosphines. They have relatively long half-lives and react with a smaller subset of biological functional groups [30, 42]. Such differences likely influence the degree and specificity of crosslinking observed.

In principle, the CpO crosslinking strategy can be extended to capture other biomolecule interactions (Figure 3-19A). Nucleic acids, for example, have been implicated in regulating many biological processes through interactions with protein partners. Misregulation of these events are at the core of many human diseases [43]. To explore the possibility of capturing RNA-mediated interactions, we required a method to install CpOs into the transcriptome of the cell. We drew inspiration from recent work from Spitale and coworkers, who demonstrated that azide-containing adenosine nucleotides were enzymatically incorporated into nascent RNA *in cellulo* [44]. We reasoned that CpO analogues could also undergo metabolism into the cell's RNA based on the small size of the CpO scaffold (Figure 3-19B).

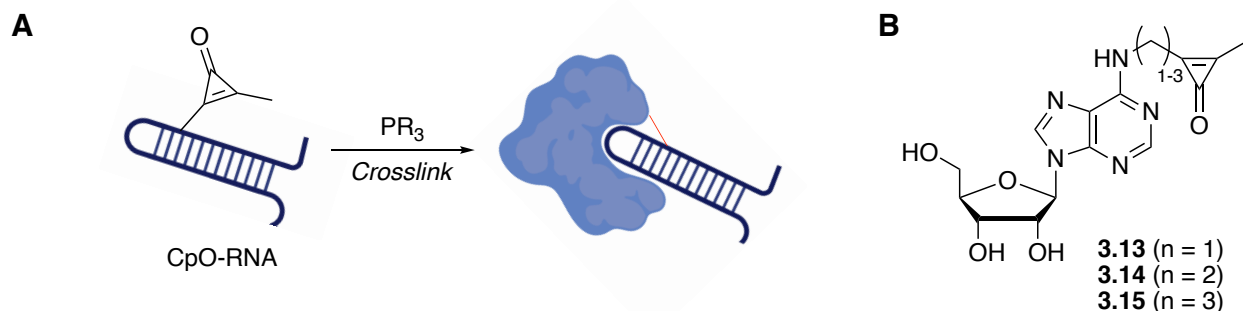
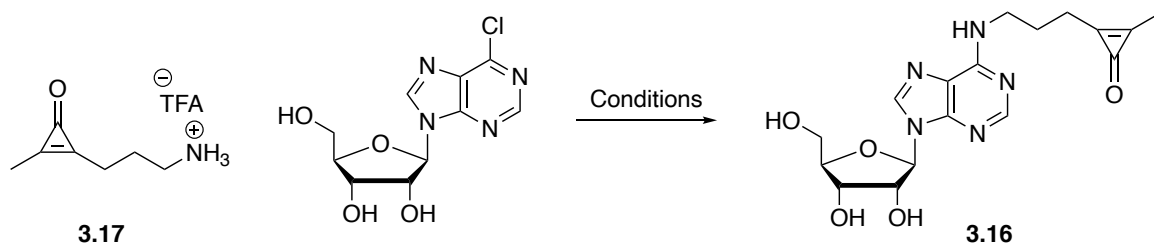


Figure 3-19. (A) CpO crosslinkers can be applied to capturing RNA-mediated interactions. (B) CpO-adenosine analogues with various linker lengths were targeted. Similar probes containing azide handles at the N^6 position were incorporated at the transcriptome level *in cellulo*.

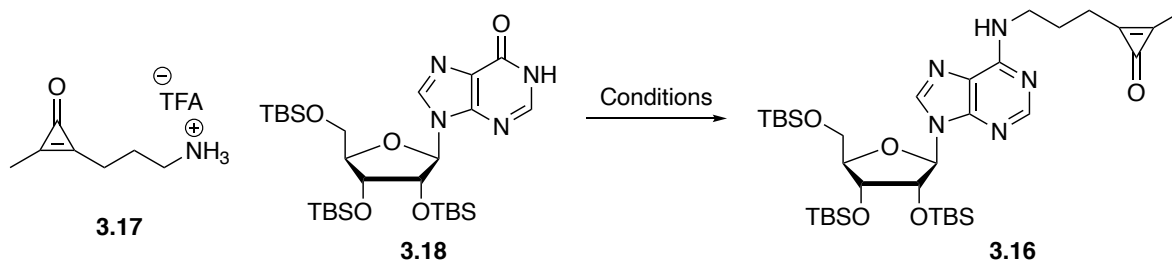
Adenosine analogues containing modifications at the N^6 position are typically accessed through S_NAr transformations with 6-chloroadenosine [44-46] (Table 3-2). Under conditions employed by Spitale *et al.*, no conversion to the desired analog was observed (entry 1). Elevated temperatures led to decomposition of CpO-amine **3.17** (entry 2). The use of inorganic base resulted in no conversion to analog **3.16**, but did not destroy CpO **3.17** (entry 3). Alternative solvent and base were evaluated, but were also unsuccessful (entry 4).

Table 3-2. Synthesis of CpO-adenosine analogues via S_NAr Transformations.



entry	base	solvent	temperature	outcome
1	Et ₃ N	EtOH	23 °C	No conversion by TLC
2	Et ₃ N	EtOH	80 °C	No CpO peaks in crude NMR
3	CaCO ₃	<i>t</i> -BuOH	40 °C	No reaction. CpO intact by crude NMR
4	iPrNEt ₂	Dioxane/H ₂ O	50 °C	Little conversion of SM after 2 d

Table 3-3. Alternative strategy to access CpO-adenosine analogues.



entry	Coupling Agent	solvent	temperature	outcome
1	BOP	DMF	23 °C	Dimethylamino-adenosine as major pdt
2	BOP	THF	23 °C	Dimethylamino-adenosine as major pdt
3	HATU	DMF	23 °C	Mostly SM by TLC. MS hit for pdt, but strong MS hit for SM
4	HBTU	DMF	23 °C	Mostly SM by TLC. MS hit for pdt, weakest MS hit for SM
5	HCTU	DMF	23 °C	Mostly SM by TLC. MS hit for pdt, 2nd weakest MS hit for SM

Adenosine derivatives modified at N^6 have also been accessed from inosine precursors such as **3.18**. Such transformations typically involve activation of O^6 using coupling agents containing 1-hydroxybenzotriazole motifs [47-48]. Literature conditions using BOP as a coupling agent (Table 3-3, entries 1 and 2) resulted in the formation of 6-dimethylamino adenosine as the major product. The dimethylamine likely resulted from partial hydrolysis of the BOP reagent. To circumvent this off-target pathway, alternative coupling agents were screened (entries 3-5). In all cases, little conversion to product **3.16** was observed. The identity of the coupling agent appeared to only affect the amount of starting amine **3.17** that was recovered.

3.4 Conclusions

In conclusion, we demonstrated that cyclopropenones can be chemically activated by phosphines to forge biomolecule crosslinks. Both the phosphine and cyclopropenone reagents are readily accessible, biocompatible, and easy to handle. Using a model split reporter, we showed that the triggered crosslinking reaction is specific and proceeds readily in physiological buffers and in the presence of cell lysate. CpO probes are complementary to photocrosslinkers and add to the growing arsenal of tools for capturing protein contacts. Phosphine accessibility to CpO probes will dictate the full scope of targets, and is the subject of ongoing work. The ability to robustly and site-specifically encode CpO motifs will also provide additional opportunities for interrogating protein binding in cells [49-50]. These experiments, as well as exploration into RNA-protein interactions are underway [51].

3.5 Materials and Methods

3.5.1 General information

All reagents and solvents were used as received, unless otherwise specified. Anhydrous organic solvents were prepared by degassing with argon and passing through two 4 x 36 in. columns of anhydrous neutral A2 (8 x 12 mesh; LaRoche Chemicals; activated at 350 °C for 12 h under a flow of argon). Column chromatography was carried out using SiliCycle 60 Å (230–400 mesh) silica gel. Thin-layer chromatography (TLC) was carried out with Agela Technologies 200 mm silica gel MF254 plates, and plates were visualized using UV light or KMnO₄ stain. Organic solutions were concentrated under reduced pressure using a Büchi rotary evaporator. HPLC purifications were performed on

an Agilent Technologies 1260 Infinity II equipped with a multiple wavelength detector, using an Agilent Eclipse XDB-C18 semiprep column (9.4 x 250 mm, 5 μ m) and a 3 mL/min flow rate.

^1H , ^{13}C , and ^{31}P NMR spectra were obtained using either a Bruker DRX400 or a Bruker AVANCE600 instrument equipped with a cryo probe. ^1H NMR spectra were acquired at 600 MHz, ^{13}C NMR spectra were acquired at 150 MHz, and ^{31}P NMR spectra were acquired at 162 MHz. Spectra were internally referenced to residual solvent signals (CDCl_3 was referenced to 7.27 ppm for ^1H and 77.16 ppm for ^{13}C , $(\text{CD}_3)_2\text{SO}$ was referenced to 2.50 ppm for ^1H and 39.52 ppm for ^{13}C , D_2O was referenced to 4.79 ppm for ^1H). All spectra were acquired at 298 K. Chemical shifts are reported in ppm, and coupling constants (J) are reported in Hz. Mass spectra were acquired at the University of California, Irvine Mass Spectrometry Facility.

3.5.2 LC-MS Analysis of CpO Trapping with Amino Acids

Trapping of phosphine-activated cyclopropenone (CpO) **3.1** was performed with amino acids **3.3–3.9**. Conversion to crosslinked products was analyzed by LC-MS. Each amino acid (10 μL of a 5 mM stock in 50% MeCN/PBS, 0.5 mM final concentration) was combined with a solution CpO **3.1** (2.5 μL of a 10 mM stock in 50% MeCN/PBS, 0.25 mM final concentration) and phosphine **3.2** (2.5 μL of a 10 mM stock in 50% MeCN/PBS, 0.25 mM final concentration). The resulting solution was diluted to 100 μL with PBS (12 mM, pH 7.4). Reactions were incubated at 37 $^\circ\text{C}$ overnight. Samples were then analyzed on a Waters ACQUITY UPLC with an ACQUITY QDa Mass Detector, using a C18 column and

eluting with a gradient of 10–90% MeCN/H₂O (containing 0.1% formic acid) over 5 min with a flow rate of 0.5 mL/min.

Trapping reactions with amino acids **3.3–3.6** were also examined at higher concentrations. Amino acids **3.3–3.6** (200 μL of 10 mM stock solutions in 50% MeCN/PBS, 4 mM final concentration) were combined with CpO **3.1** (100 μL of a 10 mM stock in 50% MeCN/PBS, 2 mM final concentration) and phosphine **3.2** (100 μL of a 10 mM stock in 50% MeCN/PBS, 2 mM final concentration). The resulting solutions were then diluted to 500 μL with PBS (12 mM, pH 7.4). Reactions were incubated at 37 °C overnight. Samples were analyzed on a Waters ACQUITY UPLC with an ACQUITY QDa Mass Detector, using a C18 column and eluting with a gradient of 10–90% MeCN/H₂O (containing 0.1% formic acid) over 15 min with a flow rate of 0.5 mL/min.

Percent crosslinking was calculated by comparing the integrations of the Fmoc-protected amino acids and product using UV absorbance at 264 nm (λ_{max} for product). The ratio was scaled to account for the stoichiometric ratio of the amino acid (2 equivalents).

$$\% \text{ crosslink} = \frac{I_P}{I_{aa} + I_P} \times 100 \times 2$$

I_P = product integration at 264 nm

I_{aa} = Fmoc-protected amino acid integration at 264 nm

3.5.3 Molecular Cloning and Plasmid Construction

LgBiT point mutants (Y16F, S28A, S29A, K124A, and K136A) were constructed using overlapping circular polymerase extension cloning (CPEC). [52] Two PCR reactions (PCR1 and PCR2) were used to assemble the insert containing the mutation of interest, using primers listed in Table 3-4. The insert was then integrated into a linearized pCOLD vector using primers outlined in Table 3-5. The resulting plasmid was transformed into BL21 *E. coli* for protein expression and purification.

Forward primers for each mutant and Insert-For1 (for Y16F, S28A, and S29A) or Insert-For2 (for K124A and K136A) (PCR1), along with reverse primers for each mutant and Insert-Rev1 (for Y16F and S29A), Insert-Rev2 (for S28A), or Insert-Rev3 (for K124A and K136A) (PCR2), were used to construct individual halves of each mutant gene product. The PCR1 and PCR2 reactions were assembled using appropriate Insert-For and Insert-Rev to give the fully constructed mutant gene product (Tables 3-4 and 3-5). PCR products from all insert reactions were analyzed using 1% agarose gels and GelRed™ (Biotium, Inc.) staining. The bands of interest were excised, and DNA was isolated following incubation at 65 °C for 20 min in ADB buffer (400 µL, Zymo Research Company). DNA products were purified using a ZymoPURE Plasmid Miniprep Kit (Zymo Research Company), eluting with Nanopure water (30 µL).

The constructed mutant gene products were ultimately inserted into a pCOLD vector backbone. The pCOLD vector was linearized by amplification using primers pCOLD-vector-For1 and pCOLD-vector-Rev (for Y16F and S29A), pCOLD-vector-For2 and pCOLD-vector-Rev (for S28A), or pCOLD-vector-For3 and pCOLD-vector-Rev2 (for

K124A and K136A, Table 3-5). Q5[®] High-Fidelity DNA Polymerase (New England BioLabs) was used in the amplifications. The linearization reaction was analyzed and purified as above. The resulting linear vector was combined in a 4:1 volumetric ratio with the constructed mutants described above. The insert and vector were combined via CPEC using the following conditions: 1X Q5[®] High-Fidelity DNA Polymerase reaction buffer, 1X Q5[®] High GC Enhancer buffer, dNTPs (0.8 mM), and Q5[®] High-Fidelity DNA Polymerase (1 U) in a total volume of 50 μ L. The following thermal cycling conditions were used: initial denaturation at 95 °C for 60 s, then 24 cycles of denaturation (95 °C, 30 s), annealing (60 °C for insert construction, 68 °C for vector linearization and full vector amplification), and extension (72 °C, 60 s for insert construction, 72 °C, 180 s for vector linearization and full vector amplification). The fully constructed plasmid was purified using a ZymoPURE Plasmid Miniprep Kit (Zymo Research Company) and eluted with Nanopure water (10 μ L). The full elution volume was used for transformation into TOP10 *E. coli* cells via electroporation.

3.5.4 Primer Lists

All primers were purchased from Integrated DNA Technologies, Inc. (San Diego, CA) and are written in the 5'→3' direction. Bases highlighted in red denote the target position for site-directed mutagenesis.

Table 3-4. Forward and reverse primers used to generate mutant LgBiT plasmids.

Forward Primers	
Y16F	GAACAGACAGCCGCCTTTAACCTGGACCAAG
S28A	CTTGAACAGGGAGGTGTGGCGAGTTTGCTGCAGAATC
S29A	CTTGAACAGGGAGGTGTGTCCGCGTTGCTGCAGAATC
K124A	CGTGTTTCGACGGCAAAGCGATCACTGTAACAGGGACC
K136A	CTGTGGAACGGCAACGCGATTATCGACGAGCG
Reverse Primers	
Y16F	CTTGGTCCAGGTTAAAGGCGGCTGTCTGTTC
S28A	GATTCTGCAGCAAACCTCGCCACACCTCCCTGTTCAAG
S29A	GATTCTGCAGCAACCGCGGACACACCTCCCTGTTCAAG
K124A	GGTCCCTGTTACAGTGATCGCTTTTGCCGTCGAACACG
K136A	CGCTCGTCGATAATCGCGTTGCCGTTCCACAG

Table 3-5. Primers used to generate gene inserts and vector backbones.

Insert Construction and Amplification Primers	
Insert-For1	CTTCCACTTTTTCCCGCGTTTTTCGCAGAAAC
Insert Rev1	GGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAA AAAACCTTGATTTG
Insert Rev2	CAAATAAAAAAATCCCCGCCAAATGGCAGGGATCTTAG ATTC
Insert-For2	CCAGTGTAAGGCAAGTCCCTTCAAGAGTTATCGTT GATACCCCTCGTAG

Insert-Rev3	GGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAA AAAACCTTGATTTG
Vector Linearization	
pCOLD-vector-For1	CCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC GATTTAGAG
pCOLD-vector-Rev1	GAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACAT TC
pCOLD-vector-For2	GGCGGGGATTTTTTTATTTGTTTTTCAGGAAATAAATAAT CGATC
pCOLD-vector-For3	CCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCC GATTTAGAG
pCOLD-vector-Rev2	TGAAGGGACTTGCCTTACTACACTGGATATGCGCTAGC ACATCAAATTG

3.5.5 Protein Expression and Purification

E. coli TOP10 cells expressing plasmids pZER09 (pCOLD-WT_LgBiT), pDR15 (pCOLD-LgBiT K136A), pDR16 (pCOLD-LgBiT K124A), pSN09 (pCOLD-LgBiT Y16F), pSN10 (pCOLD-LgBiT S28A), or pSN11 (pCOLD-LgBiT S29A) were used to inoculate 50 mL of Luria-Bertani broth (LB) containing ampicillin (100 µg/mL). After overnight incubation at 37 °C with shaking (225 rpm), 15 mL of each starter culture was used to inoculate 1 L of LB containing ampicillin (100 µg/mL). The cultures were incubated at 37 °C with shaking until an OD₆₀₀ of ~0.6 was achieved. The cultures were then chilled in an ice bath for 20 min. IPTG was added (1 mM final concentration) to induce protein expression, and cultures were incubated at 16 °C with shaking (225 rpm) for 20 h.

Cells were collected via centrifugation at 4000 rpm for 20 min (4 °C). The supernatant was decanted and the pellets were suspended in 25 mL of 50 mM phosphate buffer (pH 7.8). Protease inhibitor cocktail (Sigma-Aldrich) and phenylmethylsulfonyl fluoride (PMSF, Gold Biotechnology, 500 µM final concentration) were added to each suspension. Each sample was sonicated, and the lysate was centrifuged at 10000 rpm for 40 min (4 °C). The clarified lysates were filtered using 0.45 µm filters (Olympus). Profinity™ IMAC resin (BioRad, 5 mL bed volume) was added to each clarified lysate, and the mixtures were rocked at 4 °C for 30 min. For each sample, the resin was rinsed with wash buffer (50 mM phosphate, 20 mM imidazole, pH 7.8, 50 mL), and proteins were isolated with elution buffer (50 mM phosphate, 250 mM imidazole, pH 7.8, 3 x 5 mL). The fractions were analyzed via SDS-PAGE, and those containing the desired protein were combined and concentrated via spin filter centrifugation (3 kDa MW cutoff). Protein concentrations were determined using a Pierce™ BCA Protein Assay Kit (Thermo Fisher) or by absorbance at 280 nm using a JASCO V730 UV-Vis spectrophotometer and an extinction coefficient of 19940 M⁻¹ cm⁻¹ (calculated using ExPASy ProtParam).

3.5.6 Peptide K_D Measurements

All bioluminescence imaging experiments were performed in black 96-well plates (Greiner Bio-One). Plates containing luminescent reagents were imaged in a light-proof chamber with an IVIS Lumina (Xenogen) CCD camera chilled to -90 °C. The stage was kept at 37 °C during the imaging session, and the camera was controlled using Living Image software. The exposure time was set to 30 s, with data binning levels set to

medium. Regions of interest were selected for quantification and total flux values were analyzed using Living Image software.

Peptide affinities for LgBiT were measured as previously reported. [34] SmBiT-CpO **3.8** (1–2000 μ M) or **3.10** (0.5 nM–2 mM) was combined with native LgBiT (WT, 40 μ M) or mutant LgBiT (40 μ M) and samples were incubated at room temperature for 30 min. Furimazine (Promega, 100 μ L of a 50X dilution of the commercial stock) was then added, and bioluminescence output was measured.

3.5.7 Crosslinking Reactions with Split Nluc

All crosslinking experiments were carried out in PBS (pH 7.4) at 37 °C. For chemical crosslinking experiments, LgBiT (10–80 μ M final concentration) or hen egg-white lysozyme (HEWL, 20 μ M) was combined with SmBiT-CpO peptides **3.8**, **3.9**, and **3.10** (50 μ M–1 mM). Phosphine **3.2** (500 μ M–1 mM) or **3.11** (1 mM) was then added, and the reactions were incubated at 37 °C for 10 min–16 h.

For photocrosslinking reactions, LgBiT (80 μ M final concentration) was combined with SmBiT-Dz peptide **3.12** (500 μ M) and incubated at room temperature for 10 min. The samples were irradiated with UV light (312 nm for 10 min) using an FBTIV-816 Transilluminator (Fisher Scientific).

All samples were denatured with urea (1.6 M final concentration) for 30 min at 25 °C, then analyzed by SDS-PAGE (BioRad 4–20% gels, 130 V, 4 °C, 70–80 min). Gel staining was accomplished using Coomassie Blue. In Figure S9, excess phosphine was

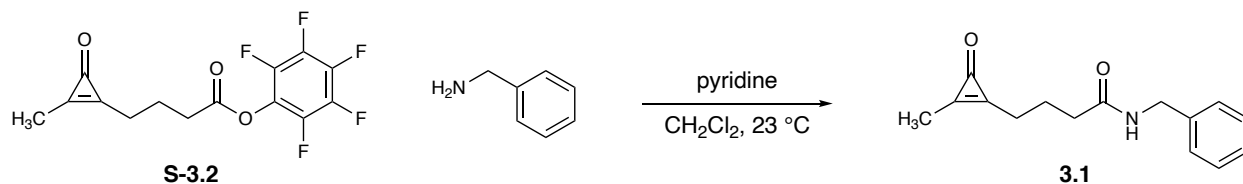
quenched with H₂O₂ (1% final concentration) prior to SDS-PAGE analysis. The percentages of crosslinked products were determined via gel scanning and calculating the area of protein bands using ImageJ software.

3.5.8 Crosslink Analysis via Trypsin Digestion

LgBiT (20 μM) was incubated with SmBiT-CpO **3.10** (100 μM) and phosphine **3.2** (500 μM) for 4 h at 37 °C. Samples were then analyzed via SDS-PAGE. Bands corresponding to the crosslinked product were excised and destained (50 mM NH₄HCO₃ in 50% MeCN, 1 mL). The destain buffer was removed, and the sample was washed with MeCN (1 x 1 mL). After drying, the samples were incubated with Trypsin Gold, Mass Spectrometry Grade (Promega, 50 ng/μL, 1 μL) at 37 °C overnight. Peptides were eluted in 0.1% formic acid and analyzed via LC-MS.

3.5.9 Synthetic Procedures

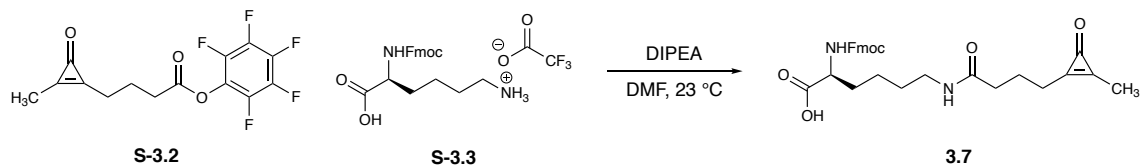
Compounds **3.18** [48], **S-3.2** [53], **S-3.3** [54] and **S-3.4** [55] were synthesized as previously described.



***N*-Benzyl-4-(2-methyl-3-oxocycloprop-1-en-1-yl)butanamide (3.1)**

To a flame-dried round-bottom flask under a nitrogen atmosphere was added **S-3.2** (192 mg, 0.600 mmol), followed by anhydrous CH_2Cl_2 (6 mL), anhydrous pyridine (72 μL , 0.89 mmol), and benzylamine (33 μL , 0.30 mmol). The solution was stirred at 23 $^\circ\text{C}$ for 16 h, then diluted with H_2O (25 mL) and washed with 1 M HCl (5 mL). The layers were separated, and the aqueous layer was extracted with CH_2Cl_2 (3 x 10 mL). The organic layers were combined, dried with MgSO_4 , filtered, and concentrated *in vacuo*. The crude residue was passed through a silica plug (eluting with 0–100% acetone in EtOAc), then further purified by HPLC (eluting with 10–90% MeCN + 0.1% TFA in H_2O + 0.1% TFA over 15 min). The desired fractions were collected and lyophilized to give **3.1** as a colorless oil (18 mg, 25%).

^1H NMR (CDCl_3 , 600 MHz) δ 7.31–7.29 (m, 2H), 7.26–7.23 (m, 3H), 6.74 (br s, 1H), 4.41 (d, $J = 5.8$ Hz, 2H), 2.62 (t, $J = 6.9$ Hz, 2H), 2.38 (t, $J = 7.0$ Hz, 2H), 2.24 (s, 3H), 2.02 (quint, $J = 7.0$ Hz, 2H). ^{13}C NMR (CDCl_3 , 150 MHz) δ 171.8, 160.9, 160.4, 156.9, 138.4, 128.8, 128.0, 127.6, 43.7, 35.0, 25.2, 22.1, 11.4. HRMS (ESI $^+$) calculated for $\text{C}_{15}\text{H}_{17}\text{NO}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ m/z 266.1157, found 266.1149.

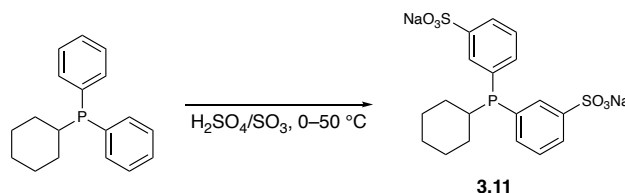


Cyclopropenone amino acid (**3.7**)

To a flame-dried round-bottom flask under a nitrogen atmosphere was added a solution of **S-3.2** (298 mg, 0.932 mmol) in anhydrous DMF (2 mL), followed by DIPEA (0.97 mL, 5.6 mmol). A solution of **S-3.3** (886 mg, 1.84 mmol) in anhydrous DMF (2 mL) was added dropwise, and the reaction was stirred at room temperature overnight. The solution was transferred to a separatory funnel containing 1 M HCl (~100 mL) and extracted into CH₂Cl₂ (3 x 20 mL). The combined organic layers were dried over MgSO₄, filtered, and concentrated *in vacuo*. The residue was diluted with toluene (100 mL) and concentrated *in vacuo* to remove residual DMF. This dilution-concentration sequence was repeated twice. The resulting residue was purified by flash column chromatography (eluting with a gradient of 2–10% MeOH in CH₂Cl₂) to give **3.7** as an off-white solid (0.27 g, 0.54 mmol). This material was used in solid-phase peptide synthesis without further purification. A small portion was further purified by HPLC (40–90% MeCN + 0.1% TFA in H₂O + 0.1% TFA over 15 min) for characterization. The desired fractions were collected and lyophilized.

¹H NMR (600 MHz, (CD₃)₂SO) δ 7.89 (d, *J* = 7.5 Hz, 2H), 7.83 (t, *J* = 5.3 Hz, 1H), 7.72 (d, *J* = 7.2 Hz, 2H), 7.61 (d, *J* = 8.0 Hz, 1H), 7.42 (t, *J* = 7.4 Hz, 2H), 7.33 (t, *J* = 7.4 Hz, 2H), 4.27 (d, *J* = 7.1 Hz, 2H), 4.22 (t, *J* = 7.1 Hz, 1H), 3.90 (td, *J* = 9.2, 4.7 Hz, 1H), 3.04–3.00 (m, 2H), 2.56 (t, *J* = 7.1 Hz, 2H), 2.22 (s, 3H), 2.14 (t, *J* = 7.4 Hz, 2H), 1.83 (quint, *J* = 7.2 Hz, 2H), 1.72–1.67 (m, 1H), 1.62–1.56 (m, 1H), 1.43–1.27 (m, 4H). ¹³C NMR (150

MHz, (CD₃)₂SO) δ 174.0, 171.1, 160.6, 158.1, 157.3, 156.2, 143.8, 140.7, 127.6, 127.1, 125.3, 120.1, 65.6, 53.7, 46.7, 38.2, 34.4, 30.4, 28.7, 25.3, 23.1, 21.9, 11.1. HRMS (ESI⁺) calculated for C₂₉H₃₂N₂O₆Na [M+Na]⁺ 527.2158 *m/z*, found 527.2162.

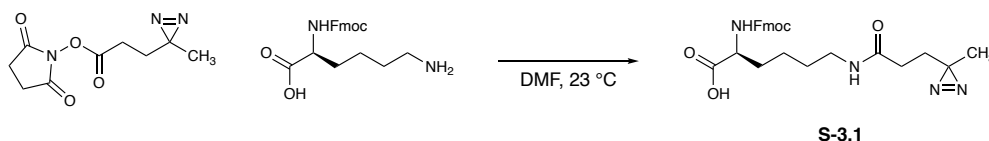


Sodium 3,3'-(cyclohexylphosphanediy) dibenzenesulfonate (3.11)

Compound **3.11** was prepared using a previously reported procedure [40], with some modifications. To a round-bottom flask under a nitrogen atmosphere was added oleum (18–24% free SO₃, 1.25 mL) and the solution was cooled to 0 °C. Cyclohexyldiphenylphosphine (268 mg, 1.00 mmol) was added in portions over 5 min, and the resulting solution was allowed to warm to room temperature. The solution was then heated at 50 °C for 4 hours until starting material was consumed. The reaction was diluted with cold water (70 mL), then triisooctylamine (10 mL) and toluene (10 mL) were added. The resulting biphasic mixture was stirred for 1 h at room temperature. In order to separate the product from impurities, the crude mixture was extracted at different pH levels. To the mixture was added a deoxygenated 5% NaOH solution until pH 3 was reached. The solution was transferred to a separatory funnel and the aqueous layer was discarded. Additional deoxygenated 5% NaOH solution was added to the separatory funnel containing the organic layer with periodic agitation until the aqueous layer reached pH 6. The aqueous layer was then discarded. To the separatory funnel containing the

organic layer was carefully added a deoxygenated 5% NaOH solution with periodic agitation until the aqueous layer reached pH 8. The organic layer was then discarded and the aqueous layer was collected and concentrated *in vacuo*. The conversion was measured by ^{31}P NMR (86% desired product, 14% corresponding oxide). This material was further purified as needed by HPLC (10–50% MeCN in H_2O over 15 min) to obtain the desired product as a white crystalline solid (trace amounts of phosphine oxide present in ^1H NMR).

^1H NMR (D_2O , 600 MHz) δ 7.98 (d, $J = 7.4$ Hz, 2H), 7.82 (d, $J = 7.7$ Hz, 2H), 7.70 (t, $J = 7.4$ Hz, 2H), 7.54 (t, $J = 7.7$ Hz, 2H), 2.50–2.45 (m, 1H), 1.74–1.66 (m, 5H), 1.38–1.32 (m, 2H), 1.26–1.15 (m, 3H). ^{31}P NMR (162 MHz, D_2O) δ -2.2. ^{13}C NMR (D_2O , 150 MHz) δ 142.7 (d, $J = 7.4$ Hz), 136.9 (d, $J = 11.2$ Hz), 136.5 (d, $J = 16.5$ Hz), 130.2 (d, $J = 21$ Hz), 129.4 (d, $J = 6.7$ Hz), 126.3, 34.2 (d, $J = 4.5$ Hz), 29.0 (d, $J = 13.6$ Hz), 26.1 (d, $J = 11.8$ Hz), 25.9. HRMS (ESI $^-$) calculated for $\text{C}_{18}\text{H}_{20}\text{NaO}_6\text{PS}_2$ $[\text{M}+\text{Na}]^-$ m/z 449.0275, found 449.0277.

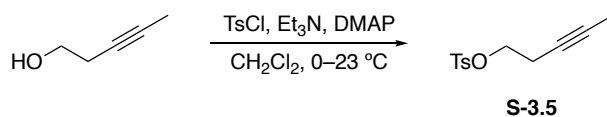


Diazirine amino acid (S-3.1)

To a flame-dried round-bottom flask under an argon atmosphere was added succinimidyl 4,4'-azipentanoate (38 mg, 0.17 mmol) and Fmoc-Lys-OH (93 mg, 0.25 mmol). Anhydrous DMF (0.56 mL) was added, and the solution was stirred at ambient

temperature overnight. The reaction was diluted with H₂O (20 mL) and extracted with Et₂O (4 x 20 mL). The organic layers were combined and washed with H₂O (2 x 20 mL) and sat. LiCl (1 x 20 mL). The organic layers were then dried with MgSO₄, filtered, and concentrated *in vacuo* to give compound **S-3.1** as a white solid (61 mg, 75%).

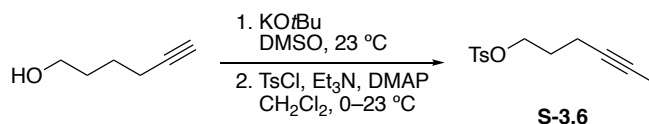
¹H NMR (CDCl₃, 500 MHz) δ 7.76 (d, *J* = 7.5 Hz, 2H), 7.60 (t, *J* = 7.4 Hz, 1.5 H), 7.52–7.56 (m, 0.5 H, rotamer), 7.39 (t, *J* = 7.4 Hz, 2H), 7.30 (t, *J* = 7.4 Hz, 2H), 5.83 (br s, 0.4 H, rotamer) 5.77–5.72 (m, 1H), 5.65 (d, *J* = 7.8 Hz, 0.7 H), 4.51 (br s, 0.4 H, rotamer) 4.44–4.34 (m, 3H), 4.21 (t, *J* = 6.9 Hz, 1H), 3.30–3.22 (m, 2H), 1.97 (t, *J* = 7.6 Hz, 2H), 1.94–1.87 (m, 1H), 1.81–1.77 (m, 1 H), 1.75 (t, *J* = 7.6 Hz, 2 H) 1.59–1.51 (m, 2H), 1.49–1.36 (m, 2H), 0.98 (s, 3H). ¹³C NMR (CDCl₃, 150 MHz) δ 174.8, 172.3, 156.5, 144.0 (rotamer), 143.8, 141.5, 127.9, 127.3, 125.3, 120.2, 67.3, 53.6, 47.3, 39.3, 31.9, 30.8, 29.0, 25.6, 22.3, 20.0. HRMS (ESI⁺) calculated for C₂₆H₃₀N₄O₅Na [M+Na]⁺ 501.2114 *m/z*, found 501.2117.



Pent-3-yn-1-yl 4-methylbenzenesulfonate (**S-3.5**)

To a flame-dried round-bottom flask was added a solution of 3-pentyn-1-ol (0.20 mL, 2.2 mmol, 1.0 equiv.) in anhydrous CH₂Cl₂ (12 mL). 4-Dimethylaminopyridine (DMAP, 27.3 mg, 0.223 mmol, 0.103 equiv.) and anhydrous triethylamine (Et₃N, 0.36 mL, 2.6 mmol, 1.2 equiv.) were added, and the resulting solution was cooled to 0 °C. 4-

Toluenesulfonyl chloride (TsCl, 0.658 g, 3.45 mmol, 1.59 equiv.) was added, and the reaction was slowly warmed to ambient temperature and stirred overnight. The solution was diluted with saturated NH₄Cl (40 mL) and extracted with CH₂Cl₂ (3 x 20 mL). The organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The resulting crude residue dry-loaded on silica, then purified by flash column chromatography (eluting with 10% EtOAc in hexanes) to give **S-3.5** (0.47 g, 91%) as a colorless oil. NMR spectra matched those previously reported [56].

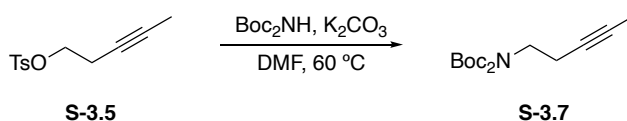


Hex-4-yn-1-yl 4-methylbenzenesulfonate (**S-3.6**)

To a flame-dried round-bottom flask was added anhydrous DMSO (5.4 mL) and 5-hexyn-1-ol (0.50 mL, 4.5 mmol, 1.0 equiv.). Potassium *tert*-butoxide (KOtBu, 1.064 g, 9.478 mmol, 2.090 equiv.) was added, and the solution was stirred overnight. The resulting yellow solution was diluted with 1 M HCl (40 mL) and extracted with Et₂O (3 x 40 mL). The organic layers were combined, and washed with brine (2 x 40 mL), then dried with MgSO₄, filtered, and concentrated under partial vacuum on an ice-water bath. The resulting volatile alcohol was used immediately without further purification.

To a flame-dried round-bottom flask was added a solution of crude alcohol in anhydrous CH₂Cl₂ (30 mL). DMAP (55.4 mg, 0.454 mmol, 0.100 equiv.) and anhydrous Et₃N (0.76 mL, 5.4 mmol, 1.2 equiv.) were added, and the resulting solution was cooled to 0 °C. TsCl (1.723 g, 9.072 mmol, 2.000 equiv.) was added, and the reaction was slowly

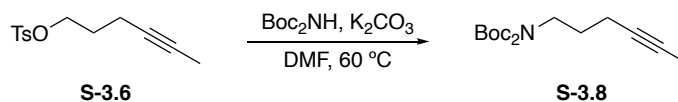
warmed to ambient temperature and stirred overnight. The solution was diluted with saturated NH_4Cl (40 mL) and extracted with CH_2Cl_2 (2 x 40 mL). The organic layers were combined, dried with MgSO_4 , filtered, and concentrated *in vacuo*. The resulting crude residue dry-loaded on silica, then purified by flash column chromatography (eluting with 10% EtOAc in hexanes) to give **S-3.6** (0.53 g, 46% over two steps) as a light yellow oil. NMR spectra matched those previously reported [57].



Di-*tert*-butyl pent-3-yn-1-yliminodicarbonate (**S-3.7**)

To a flame-dried round-bottom flask was added di-*tert*-butyl-iminodicarboxylate (Boc_2NH , 0.995 g, 4.58 mmol, 1.50 equiv.) and potassium carbonate (K_2CO_3 , 0.850 g, 6.15 mmol, 2.01 equiv.). A solution of **S-3.5** (0.728 g, 3.05 mmol, 1.00 equiv.) in anhydrous DMF (3.0 mL) was added, and the resulting slurry was stirred at 60 °C overnight. The mixture was diluted with H_2O (40 mL) and extracted with Et_2O (3 x 30 mL). The organic layers were combined and washed with brine (2 x 40 mL), then dried with MgSO_4 , filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (eluting with 5% EtOAc in hexanes) to give **S-3.7** (0.54 g, 62%) as a white solid.

^1H NMR (600 MHz, CDCl_3) δ 3.70 (t, $J = 7.3$ Hz, 2H), 2.43–2.37 (m, 2H), 1.75 (t, $J = 2.5$ Hz, 3H), 1.51 (s, 18H). ^{13}C NMR (151 MHz, CDCl_3) δ 152.5, 82.5, 77.1, 75.9, 45.5, 28.2, 19.3, 3.7. HRMS (ESI $^+$) calcd. for $\text{C}_{15}\text{H}_{25}\text{NO}_4\text{Na}$ $[\text{M}+\text{Na}]^+$ 306.1681 m/z , found 306.1684.



Di-*tert*-butyl hex-4-yn-1-yliminodicarbonate (**S-3.8**)

To a flame-dried round-bottom flask was added Boc_2NH (0.673 g, 3.10 mmol, 1.49 equiv.) and K_2CO_3 (0.577 g, 4.18 mmol, 2.01 equiv.). A solution of **S-3.6** (0.524 g, 2.08 mmol, 1.00 equiv.) in anhydrous DMF (2.0 mL) was added, and the resulting slurry was stirred at 60 °C overnight. The mixture was diluted with H_2O (30 mL) and extracted with Et_2O (2 x 40 mL). The organic layers were combined and washed with brine (2 x 30 mL), then dried with MgSO_4 , filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (eluting with 5% EtOAc in hexanes) to give **S-3.8** (0.55 g, 88%) as a pale yellow oil.

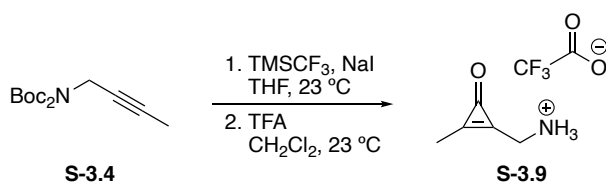
^1H NMR (600 MHz, CDCl_3) δ 3.64 (t, $J = 7.4$ Hz, 2H), 2.16–2.13 (m, 2H), 1.76–1.72 (m, 5H), 1.50 (s, 18H). ^{13}C NMR (151 MHz, CDCl_3) δ 152.6, 82.3, 78.4, 75.9, 46.0, 28.6, 28.2, 16.5, 3.6. HRMS (ESI⁺) calcd. for $\text{C}_{16}\text{H}_{27}\text{NO}_4\text{Na}$ [$\text{M}+\text{Na}$]⁺ 320.1838 m/z , found 320.1843.

General procedure A for the synthesis of CpO-ammonium compounds

To a flame-dried Schlenk tube was added sodium iodide (NaI, 1.8–2.5 equiv.). The reagent was gently flame-dried under vacuum, then resuspended in a solution of alkyne (1 equiv.) in anhydrous THF (0.25 M alkyne). Trifluoromethyltrimethylsilane (TMSCF_3 , 2–2.2 equiv.) was added, and the vessel was sealed and stirred vigorously at ambient temperature for 2 d. The slurry was diluted with water (30 mL) and extracted with CH_2Cl_2 (3 x 20 mL). The organic layers were combined, dried with MgSO_4 , filtered, and

concentrated *in vacuo*. The resulting crude difluorocyclopropene was used immediately without further purification.

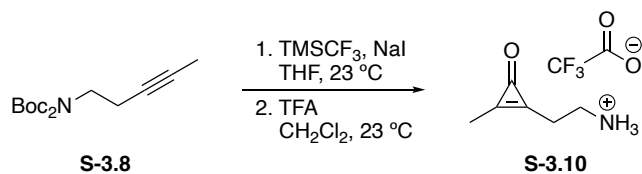
To a flame-dried round-bottom flask was added a solution of crude difluorocyclopropene in anhydrous CH₂Cl₂ (~ 0.25 M w.r.t. alkyne), followed by trifluoroacetic acid (TFA, 5–8.6 equiv.). The solution was stirred overnight, then concentrated *in vacuo*. The crude residue was resuspended in H₂O (5 mL) and washed with Et₂O (3 x 10 mL). The aqueous layer was collected and lyophilized to give the desired CpO-ammonium as the TFA salt.



(2-Methyl-3-oxocycloprop-1-en-1-yl)methanaminium trifluoroacetate (**S-3.9**)

General procedure A was used with the following reagents: **S-3.4** (0.221 g, 0.821 mmol, 1.00 equiv), NaI (0.277 g, 1.84 mmol, 2.25 equiv.), TMSCF₃ (0.24 mL, 1.6 mmol, 2.0 equiv.), anhydrous THF (3.3 mL); TFA (0.54 mL, 7.1 mmol, 8.6 equiv.), anhydrous CH₂Cl₂ (4.0 mL). Compound **S-3.9** (91 mg, 52% over two steps) was isolated as a yellow residue.

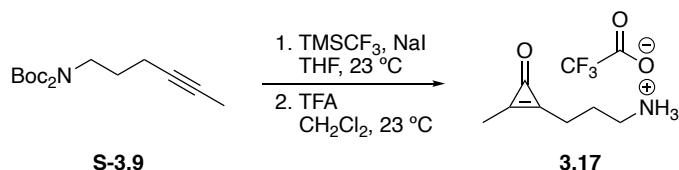
¹H NMR (400 MHz, (CD₃)₂SO) δ 8.46 (br s, 3H), 4.20 (s, 2H), 2.31 (s, 3H). ¹⁹F NMR (565 MHz, (CD₃)₂SO) δ –73.5. ¹³C NMR (151 MHz, (CD₃)₂SO) δ 158.2, 155.2, 155.0, 36.5, 11.4. HRMS could not be acquired for this compound.



2-(2-Methyl-3-oxocycloprop-1-en-1-yl)ethan-1-aminium trifluoroacetate (**S-3.10**)

General procedure A was used with the following reagents: **S-3.8** (0.319 g, 1.13 mmol, 1.00 equiv), NaI (0.383 g, 2.56 mmol, 2.27 equiv.), TMSCF₃ (0.33 mL, 2.3 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); TFA (0.43 mL, 5.6 mmol, 5.0 equiv.), anhydrous CH₂Cl₂ (4.0 mL). Compound **S-3.10** (0.17 g, 68% over two steps) was isolated as a reddish-brown residue.

¹H NMR (400 MHz, (CD₃)₂SO) δ 7.74 (br s, 3H), 3.17 (t, *J* = 7.3 Hz, 2H), 2.90 (t, *J* = 7.4 Hz, 2H), 2.27 (s, 3H). ¹⁹F NMR (565 MHz, (CD₃)₂SO) δ -73.5. ¹³C NMR (151 MHz, (CD₃)₂SO) δ 159.0, 157.3, 157.2, 36.0, 24.8, 11.4. HRMS could not be acquired for this compound.



3-(2-Methyl-3-oxocycloprop-1-en-1-yl)propan-1-aminium trifluoroacetate (**3.17**)

General procedure A was used with the following reagents: **S-3.9** (0.144 g, 0.483 mmol, 1.00 equiv), NaI (0.170 g, 1.13 mmol, 2.36 equiv.), TMSCF₃ (0.14 mL, 0.97 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); TFA (0.43 mL, 5.6 mmol, 5.0 equiv.), anhydrous

CH₂Cl₂ (4.0 mL). Compound **3.17** (0.080 g, 69% over two steps) was isolated as a yellow residue.

¹H NMR (400 MHz, (CD₃)₂SO) δ 7.72 (br s, 3H), 2.87 (t, *J* = 7.4 Hz, 2H), 2.69 (t, *J* = 7.1 Hz, 2H), 2.25 (s, 3H), 1.90 (quint, *J* = 7.4 Hz, 2H). ¹⁹F NMR (565 MHz, (CD₃)₂SO) δ –73.5. ¹³C NMR (151 MHz, (CD₃)₂SO) δ 159.9, 158.0, 157.4, 38.3, 23.7, 23.0, 11.1. HRMS could not be acquired for this compound.

SmBiT peptides: H₂N-K(CpO)VTGWRLSERILA-NH₂ (3.8), H₂N-K(CpO)VTGWRLFEEIL-NH₂ (3.9), H₂N-K(CpO)VSGWRLFKKISN-NH₂ (3.10), and H₂N-K(Dz)VSGWRLFKKISN-NH₂ (3.12).

To a fritted glass reservoir was added NovaPEG Rink Amide resin (0.46 mmol/g, 110–220 mg, 0.05–0.10 mmol respectively). Throughout the synthesis, the resin was agitated by bubbling N₂ through the reservoir. The resin was initially swelled and washed with DMF (2 x 15 mL, 15 min per cycle), then treated with 20% 4-methylpiperidine/DMF (10 mL, 20 min) for Fmoc removal. After the first deprotection step, the resin was washed with DMF (3 x 10 mL, 5 min). Protected amino acids were then coupled in the following manner: Fmoc-protected amino acid (0.30 mmol), HCTU (120 mg, 0.30 mmol), and diisopropylethylamine (160 μL, 0.89 mmol) were dissolved in 3.0 mL of DMF. The solution was mixed thoroughly, incubated at room temperature for ~3 min, and then added to the deprotected resin. Each coupling step was performed at room temperature for a minimum of 45 min. The resin was then washed with DMF (3 x 10 mL, 5 min). Subsequent Fmoc deprotection was achieved by incubating the resin with 10 mL of 20% 4-

methylpiperidine/DMF (20 min). The resin was then washed with DMF (3 x 10 mL, 5 min) prior to the next coupling step. Repeated cycles of amino acid coupling and deprotection were used to achieve the desired peptides. After the final cycle, the resin was washed with CH₂Cl₂ (3 x 10 mL, 5 min). Peptides were then isolated by incubating the resin with 95:2.5:2.5 trifluoroacetic acid (TFA):triisopropylsilane:H₂O (2 x 5 mL, 30 min). The cleavage solutions were combined, concentrated *in vacuo*, and treated with 40 mL cold diethyl ether. The precipitate was then collected via centrifugation (3000 x *g*, 5 min) and dried under vacuum. The peptides were purified by HPLC (eluting with 40–90% MeCN containing 0.1% TFA over 15 min), and the purified products were lyophilized. Product identity and purity were assessed via LC-MS. Peptide stocks were prepared in PBS (pH 7.4), and concentrations were determined using a Pierce™ BCA Protein Assay Kit (Thermo Fisher) or by absorbance at 280 nm using an extinction coefficient of 5500 M⁻¹ cm⁻¹ (calculated using ExPASy ProtParam).

3.6 References

- [1]. Tanaka, Y.; Bond, M. R.; Kohler, J. J. *Mol. BioSyst.* **2008**, *4*, 473–480.
- [2]. Das, J. *Chem. Rev.* **2011**, *111*, 4405–4417.
- [3]. Lin, S.; He, D.; Long, T.; Zhang, S.; Meng, R.; Chen, P. R. *J. Am. Chem. Soc.* **2014**, *136*, 11860–11863.
- [4]. Preston, G. W.; Wilson, A. J. *Chem. Soc. Rev.* **2013**, *42*, 3289–3301.
- [5]. Coin, I.; Katritch, V.; Sun, T.; Xiang, Z.; Siu, F. Y.; Beyermann, M.; Stevens, R. C.; Wang, L. *Cell* **2013**, *155*, 1258–1269.

- [6]. Dormán, G.; Nakamura, H.; Pulsipher, A.; Prestwich, G. D. *Chem. Rev.* **2016**, *116*, 15284–15398.
- [7]. Joiner, C. M.; Breen, M. E.; Clayton, J.; Mapp, A. K. *ChemBioChem* **2017**, *18*, 181–184.
- [8]. Tian, Y.; Jacinto, M. P.; Zeng, Y.; Yu, Z.; Qu, J.; Liu, W. R.; Lin, Q. *J. Am. Chem. Soc.* **2017**, *139*, 6078–6081.
- [9]. Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 9026–9027.
- [10]. Tippmann, E. M.; Liu, W.; Summerer, D.; Mack, A. V.; Schultz, P. G. *ChemBioChem* **2007**, *8*, 2210–2214.
- [11]. Chin, J. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11020–11024.
- [12]. Qiu, Z.; Lu, L.; Jian, X.; He, C. *J. Am. Chem. Soc.* **2008**, *130*, 14398–14399.
- [13]. Buchmueller, K. L.; Hill, B. T.; Platz, M. S.; Weeks, K. M. *J. Am. Chem. Soc.* **2003**, *125*, 10850–10861.
- [14]. Parker, C. G.; Galmozzi, A.; Wang, Y.; Correia, B. E.; Sasaki, K.; Joslyn, C. M.; Kim, A. S.; Cavallaro, C. L.; Lawrence, R. M.; Johnson, S. R.; Narvaiza, I.; Saez, E.; Cravatt, B. F. *Cell* **2017**, *168*, 527–541.
- [15]. Niphakis, M. J.; Lum, K. M.; Cognetta, A. B., III; Correia, B. E.; Ichu, T.-A.; Olucha, J.; Brown, S. J.; Kundu, S.; Piscitelli, F.; Rosen, H.; Cravatt, B. F. *Cell* **2015**, *161*, 1668–1680.
- [16]. Yu, S.-H.; Boyce, M.; Wands, A. M.; Bond, M. R.; Bertozzi, C. R.; Kohler, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 4834–4839.

- [17]. Han, S.; Collins, B. E.; Bengston, P.; Paulson, J. C. *Nat. Chem. Biol.* **2005**, *1*, 93–97.
- [18]. Janz, J. M.; Ren, Y.; Looby, R.; Kazmi, M. A.; Sachdev, P.; Grunbeck, A.; Haggis, L.; Chinnapen, D.; Lin, A. Y.; Seibert, C.; McMurry, T.; Carlson, K. E.; Muir, T. W.; Hunt, S., III; Sakmar, T. P. *J. Am. Chem. Soc.* **2011**, *133*, 15878–15881.
- [19]. Kleiner, P.; Heydenreuter, W.; Stahl, M.; Korotkov, V. S.; Sieber, S. A. *Angew. Chem. Int. Ed.* **2017**, *56*, 1396–1401.
- [20]. Pham, N. D.; Parker, R. B.; Kohler, J. J. *Curr. Opin. Chem. Biol.* **2013**, *17*, 90–101.
- [21]. Rice, B. W.; Cable, M. D.; Nelson, M. B. *J. Biomed. Opt.* **2001**, *6*, 432–440.
- [22]. Cigler, M.; Müller, T. G.; Horn-Ghetko, D.; von Wrisberg, M.-K.; Fottner, M.; Goody, R. S.; Itzen, A.; Müller, M. P.; Lang, K. *Angew. Chem. Int. Ed.* **2017**, *56*, 15737–15741.
- [23]. Xuan, W.; Li, J.; Luo, X.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2016**, *55*, 10065–10068.
- [24]. Xuan, W.; Shao, S.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2017**, *56*, 5096–5100.
- [25]. Dong, J.; Krasnova, L.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2014**, *53*, 9430–9448.
- [26]. Yang, B.; Wu, H.; Schnier, P. D.; Liu, Y.; Liu, J.; Wang, N.; DeGrado, W. F.; Wang, L. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 11162–11167.
- [27]. Martín-Gago, P.; Olsen, C. A. *Angew. Chem. Int. Ed.* **2019**, *58*, 957–966.
- [28]. Niessen, S.; Dix, M. M.; Barbas, S.; Potter, Z. E.; Lu, S.; Brodsky, O.; Planken, S.; Behenna, D.; Almaden, C.; Gajiwala, K. S.; Ryan, K.; Ferre, R.; Lazear, M. R.; Hayward, M. M.; Kath, J. C.; Cravatt, B. F. *Cell Chem. Biol.* **2017**, *24*, 1388–1400.

- [29]. Shindo, N.; Fuchida, H.; Sato, M.; Watari, K.; Shibata, T.; Kuwata, K.; Miura, C.; Okamoto, K.; Hatsuyama, Y.; Tokunaga, K.; Sakamoto, S.; Morimoto, S.; Abe, Y.; Shiroishi, M.; Caaveiro, J. M. M.; Ueda, T.; Tamura, T.; Matsunaga, N.; Nakao, T.; Koyanagi, S.; Ohdo, S.; Yamaguchi, Y.; Hamachi, I.; Ono, M.; Ojida, A. *Nat. Chem. Biol.* **2019**, *15*, 250–258.
- [30]. Shih, H.-W.; Prescher, J. A. *J. Am. Chem. Soc.* **2015**, *137*, 10036–10039.
- [31]. Row, R. D.; Shih, H.-W.; Alexander, A. T.; Mehl, R. A.; Prescher, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 7370–7375.
- [32]. Yan, C.; Wu, F.; Jernigan, R. L.; Dobbs, D.; Honavar, V. *Protein J.* **2008**, *27*, 59–70.
- [33]. Nguyen, T.-A.; Cigler, M.; Lang, K. *Angew. Chem. Int. Ed.* **2018**, *57*, 14350–14361.
- [34]. Dixon, A. S.; Schwinn, M. K.; Hall, M. P.; Zimmerman, K.; Otto, P.; Lubben, T. H.; Butler, B. L.; Binkowski, B. F.; Machleidt, T.; Kirkland, T. A.; Wood, M. G.; Eggers, C. T.; Encell, L. P.; Wood, K. V. *ACS Chem. Biol.* **2016**, *11*, 400–408.
- [35]. Tomabechi, Y.; Hosoya, T.; Ehara, H.; Sekine, S.-I.; Shirouzu, M.; Inouye, S. *Biochem. Biophys. Res. Commun.* **2016**, *470*, 88–93.
- [36]. Weiss, M. S.; Palm, G. J.; Hilgenfeld, R. *Acta Cryst.* **2000**, *D56*, 952–958.
- [37]. Ni, H.; Chan, W.-L.; Lu, Y. *Chem. Rev.* **2018**, *118*, 9344–9411.
- [38]. Barder, T. E.; Buchwald, S. L. *J. Am. Chem. Soc.* **2007**, *129*, 5096–5101.
- [39]. Caiazza, A.; Dalili, S.; Yudin, A. K. *Org. Lett.* **2002**, *4*, 2597–2600.
- [40]. Mika, L. T.; Orha, L.; Farkas, N.; Horváth, I. T. *Organometallics* **2009**, *28*, 1593–1596.
- [41]. Aguilar, X.; Weise, C. F.; Sparrman, T.; Wolf-Watz, M.; Wittung-Stafshede, P. *Biochemistry* **2011**, *50*, 3034–3044.
- [42]. Hamada, A.; Takizawa, T. *Tetrahedron Lett.* **1972**, *13*, 1849–1850.

- [43]. Khalil, A. M.; Rinn, J. L. *Semin. Cell Dev. Biol.* **2011**, *22*, 359–365.
- [44]. Nainar, S.; Beasley, S.; Fazio, M.; Kubota, M.; Dai, N.; Corrêa, I. R., Jr.; Spitale, R. C. *ChemBioChem* **2016**, *17*, 2149–2152.
- [45]. Du, X.; Li, J.; Gao, Y.; Kuang, Y.; Xu, B. *Chem. Commun.* **2012**, *48*, 2098–2100.
- [46]. Barlow, N.; Baker, S. P.; Scammells, P. J. *ChemMedChem* **2013**, *8*, 2036–2046.
- [47]. Wan, Z.-K.; Binnun, E.; Wilson, D. P.; Lee, J. *Org. Lett.* **2005**, *7*, 5877–5880.
- [48]. Bae, S.; Lakshman, M. K. *J. Am. Chem. Soc.* **2007**, *129*, 782–789.
- [49]. Smith, E.; Collins, I. *Future Med. Chem.* **2015**, *7*, 159–183.
- [50]. Wright, M. H.; Sieber, S. A. *Nat. Prod. Rep.* **2016**, *33*, 681–708.
- [51]. Ayele, T. M.; Loya, T.; Valdez-Sinon, A. N.; Bassell, G. J.; Heemstra, J. M. *BioRxiv* **2020**, 942482.
- [52]. Quan, J.; Tian, J. *Nat. Protoc.* **2011**, *6*, 242–251.
- [53]. Row, R. D.; Prescher, J. A. *Org. Lett.* **2018**, *20*, 5614–5617.
- [54]. Papeo, G.; Giordano, P.; Brasca, M. G.; Buzzo, F.; Caronni, D.; Ciprandi, F.; Mongelli, N.; Veronesi, M.; Vulpetti, A.; Dalvit, C. *J. Am. Chem. Soc.* **2007**, *129*, 5665–5672.
- [55]. Anderson, J. C.; Flaherty, A.; Swarbrick, M. E. *J. Org. Chem.* **2000**, *65*, 9152–9156.
- [56]. Fang, F.; Vogel, M.; Hines, J. V.; Bergmeier, S. C. *Org. Biomol. Chem.* **2012**, *10*, 3080–3091.
- [57]. Cruz, F. A.; Chen, Z.; Kurtoic, S. I.; Dong, V. M. *Chem. Commun.* **2016**, *52*, 5836–5839.

Chapter 4: Butenolide synthesis from functionalized cyclopropenones

Adapted from published work, with permission from the American Chemical Society: Nguyen, S. S.; Ferreira, A. J.; Long, Z. G.; Heiss, T. K.; Dorn, R. S.; Row, R. D.; Prescher, J. A. *Org. Lett.* **2019**, *21*, 8695–8699.

4.1 Abstract

The stringent requirements placed on bioorthogonal reagents – high chemoselectivity, broad functional group tolerance, generalizability – are also traits highly sought after in methodology development. We envisaged that the cyclopropenone (CpO)-phosphine ligation could be optimized to access a class of cyclic lactones known as butenolides. Retrosynthetically, these scaffolds could be accessed from the CpO-phosphine ligation via intramolecular trapping of the ketene-ylide intermediate. Whereas previous renditions of this reaction featured an intramolecular trap on the phosphine probe, this strategy would require the intramolecular trap on the CpO reagent. In this Chapter, I discuss the development of a method to synthesize highly substituted butenolides from bioorthogonal CpOs.

4.2 Introduction

Butenolides are found in a variety of natural product scaffolds and possess desirable bioactive properties [1]. For example, linderalactone (Figure 4-1A) can protect hepatocytes from oxidative damage [2]. Other butenolides, including (–)-incrustoporin, are potent inhibitors of pathogenic fungi [3]. These and related scaffolds [4] have inspired chemists to develop efficient syntheses of α,β -unsaturated lactones. While several

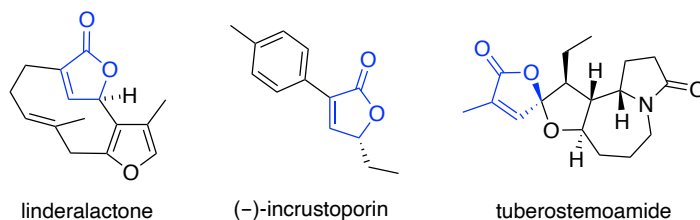
methods now exist [5-9], most are limited in terms of substitution pattern, functional group tolerance, or starting material accessibility. More general methods to build butenolides are therefore needed.

Cyclopropenones (CpOs) are attractive synthons for butenolide formation. These microcycles map readily onto the target structures and are easily accessible from alkyne precursors [10]. CpOs have only recently been exploited for lactone synthesis, though [11-13]. The groups of Lin and Sun developed methods to convert symmetric CpOs to functionalized butenolides (Figures 4-1B and 4-1C). Both transformations proceed via intermolecular 1,2-addition of carbonyls into the CpO scaffold, followed by ring opening. The resulting vinyl anion intermediates subsequently cyclize to deliver the desired lactones. While robust, both methods require electron-rich carbonyl fragments. The vinyl anion intermediates are also not compatible with a range of functional groups, limiting the scope of these chemistries. Furthermore, the intermolecular nature of the reactions can present regioselectivity challenges when unsymmetrical CpOs are employed.

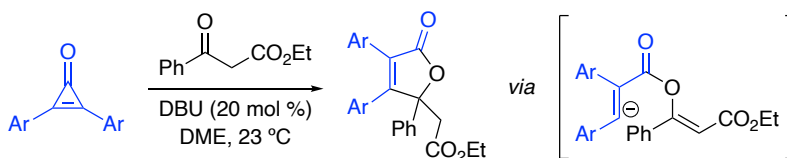
We hypothesized that an *intramolecular* reaction could convert CpOs to functionalized butenolides and potentially broaden the scope of accessible products. Toward this end, we drew inspiration from our previous work on bioorthogonal CpOs [14-15]. These motifs undergo conjugate addition reactions with phosphines, producing ketene-ylide intermediates upon ring fragmentation [16]. The electrophiles can be readily trapped by pendant nucleophiles on the phosphine (Figure 4-2A) to afford covalent adducts. We surmised that pendant nucleophiles on the CpO – rather than the phosphine – could also trap the ketene (Figure 4-2B). If the nucleophile was part of a hydroxymethyl

tether, the products would comprise lactones. Subsequent ylide protonation and phosphine elimination could ultimately deliver functionalized butenolides. Since the proposed method involves bioorthogonal reagents, it would likely be compatible with a variety of functional groups. The proximity of the hydroxy group on the CpO would also promote intramolecular cyclization, outcompeting any exogenous nucleophile in the trapping step [15].

A) Bioactive natural products comprising butenolides



B) Lin (2017)



C) Sun (2018)

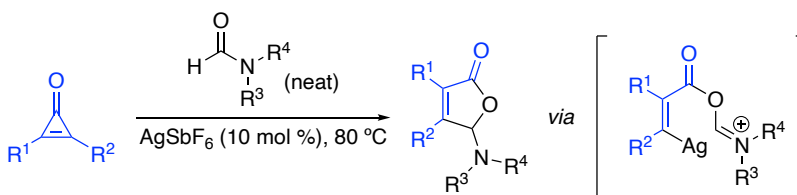
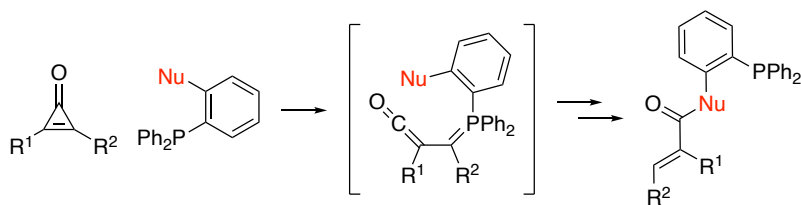


Figure 4-1. (A) Butenolides (blue) are common motifs in bioactive natural products. Recent work by (B) Lin [11] and (C) Sun [12] featured intermolecular reactions between CpOs and carbonyls to generate functionalized butenolides.

A) Phosphine-triggered CpO ligation



B) This work

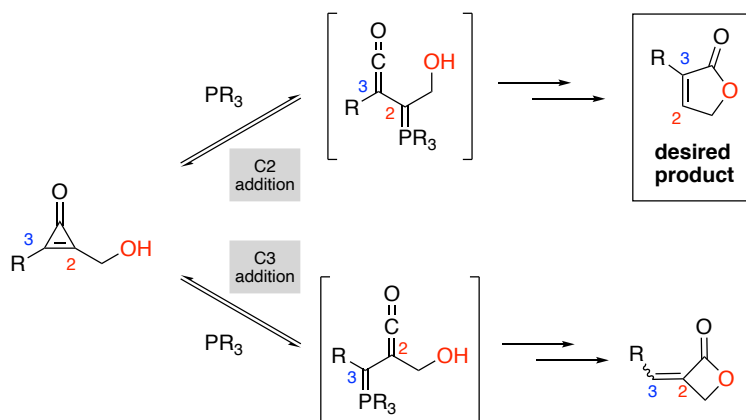


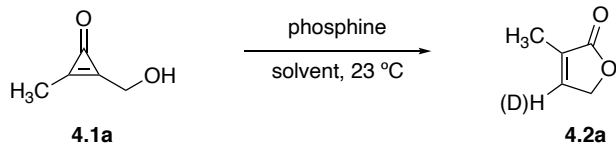
Figure 4-2. CpOs react with bioorthogonal phosphines to reveal ketene ylides. These intermediates can be trapped with A) pendant nucleophiles on phosphines or B) hydroxy group nucleophiles on CpO scaffolds. In the latter case, phosphine addition at C2 affords butenolides (top). Phosphine addition at C3 could provide undesired β -lactones (bottom).

The regioselectivity of the proposed reaction was further considered. Phosphine addition at C2 (i.e., the carbon bearing the hydroxymethyl tether) would likely provide the desired products. However, phosphine addition at C3 could give undesired β -lactones. Cyclization en route to the β -lactone (4-exo-dig) is disfavored according to Baldwin's rules, but notable examples exist [17-19]. We reasoned that butenolide formation could still predominate in the reaction, even if the phosphine attacked C3. Ketene-ylide formation is reversible in the absence of trapping nucleophiles [14,20], and cyclization en route to the β -lactone would likely be slower than the reverse reaction to reform the CpO. Thus, the reaction could funnel to a butenolide product, regardless of the initial site of phosphine addition.

4.3 Results and Discussion

To examine the overall strategy, model CpO **4.1a** was synthesized and treated with a panel of phosphines. The reactions were performed in various solvents and monitored using ^1H NMR spectroscopy (Table 4-1, See Appendix B). 1,3,5-Triaza-7-phosphaadamantane (PTA) was initially selected due to its potent reactivity and stability in protic solvents. When combined stoichiometrically with **4.1a**, PTA afforded rapid conversion to the desired butenolide (entry 1). The reaction slowed significantly when the catalyst loading was reduced to 10 mol % (entry 2). Notably, no intermolecular solvent trapping was observed even at the longer reaction times. Sluggish reactivity was not general to all alkyl-substituted phosphines. When **4.1a** was treated with cyclohexyldiphenylphosphine (CyDPP), rapid formation of **4.2a** was observed even at low catalyst loadings (entry 3).

Anticipating that alkyl phosphines would be prone to oxidation and thus less general, we further investigated triarylphosphines. We were particularly drawn to tri(*o*-tolyl)phosphine, as this reagent would likely be sufficiently nucleophilic to add into the CpO core, but be reasonably air stable. When CpO **4.1a** was treated with tri(*o*-tolyl)phosphine, though, no conversion to butenolide **4.2a** was observed (entry 4). No butenolide was formed even with longer reaction times or elevated temperatures (See Appendix B). The increased steric bulk surrounding the phosphine likely precluded efficient conjugate addition. Indeed, this phosphine is rarely used in Michael-type reactions and is primarily employed as a metal ligand [21-23].

Table 4-1. Optimization of butenolide cyclization.

entry	phosphine (mol %)	solvent	time	conversion (%) ^b
1	PTA (100)	CD ₃ OD	10 min	> 95
2	PTA (10)	CD ₃ OD	2 h	67
3	CyDPP (5)	CD ₃ OD	2.5 h	90
4	P(<i>o</i> -tolyl) ₃ (5)	CD ₃ OD	2.5 h	0
5	PPh ₃ (100)	CD ₃ OD	10 min	> 95
6	PPh ₃ (100)	DMSO- <i>d</i> ₆	22 h	> 95
7	PPh ₃ (5)	DMSO- <i>d</i> ₆	24 h	16
8	PPh ₃ (100)	C ₆ D ₆	1 h	> 95
9	PPh ₃ (5)	C ₆ D ₆	24 h	94
10	PPh ₃ (10)	CD ₃ OD	1.5 h	> 95
11	PPh₃ (5)	CD₃OD	2.5 h	> 95
12	PTA (1)	CD ₃ OD	2 h	9
13	PTA (5)	D ₂ O	2.5 h	52
14	P(<i>o</i> -tolyl) ₃ (5)	CD ₃ OD	21 h	0
15 ^c	P(<i>o</i> -tolyl) ₃ (5)	CH ₃ OH	18 h	0

^aReaction conditions: CpO (15 μmol), TMS-acetylene (3 μmol), solvent (600 μL)

^bNMR conversion, calculated from integral ratios between starting CpO and butenolide product

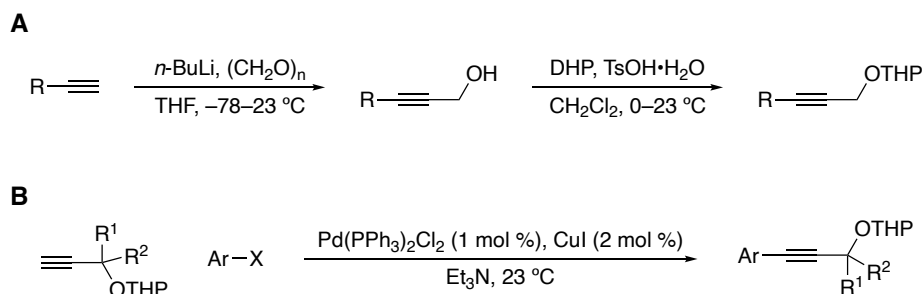
^cReaction conditions: CpO (0.36 mmol), solvent (14.4 mL), 60 °C

We next tested a less sterically encumbered reagent, triphenylphosphine (PPh₃). PPh₃ is commercially available, inexpensive, and bench stable, making it attractive for methods development. When **4.1a** was treated with stoichiometric amounts of PPh₃ in CD₃OD, rapid conversion to butenolide **4.2a** was observed (entry 5). Efficient cyclization occurred even at reduced catalyst loadings (entries 10–11) with no evidence of intermolecular solvent trapping. Additionally, in the realm of phosphine organocatalysis, few reactions feature catalyst loadings below 10 mol %, and even fewer use air-stable, commercially available reagents [24]. We were also surprised that low catalyst loadings

afforded rapid butenolide formation, considering the diminished nucleophilicity of PPh₃. We hypothesized that the polar protic solvent accelerated the reaction via hydrogen-bond activation of the starting CpO. Similar observations were made when CpOs were tuned for bioorthogonal ligation [15]. When the cyclizations were performed in DMSO-*d*₆ or C₆D₆, longer reaction times or stoichiometric amounts of PPh₃ were required for full conversion (entries 6–9).

We aimed to test the optimized reaction conditions with a variety of hydroxymethyl CpOs. Such probes can be readily accessed via appropriately functionalized alkynes [10]. We thus prepared a variety of alkyl- and aryl-substituted alkynes via acetylide addition to formaldehyde or Sonogashira cross-coupling reactions, respectively (Scheme 4-1). Each alkyne also comprised a THP-protected hydroxymethyl tether. The alkyne products were subjected to difluorocarbene, a reagent generated *in situ* via the conditions of Olah [25] (Scheme 4-2). The resulting difluorocyclopropenes were then hydrolyzed and deprotected to furnish the desired hydroxymethyl CpOs (Scheme 4-2). The carbene insertion and hydrolysis sequence was not compatible with alkynes bearing nitrogen heterocycles, and attempts to isolate the corresponding CpOs resulted in decomposition. These results are in stark contrast to heterocyclic alkenes, which undergo robust difluorocarbene insertion [26-27].

Scheme 4-1. General synthesis of functionalized alkynes with masked hydroxymethyl tethers. (A) Alkyl-functionalized alkynes were accessed via acetylide addition to formaldehyde, followed by THP protection. (B) Aryl-functionalized alkynes were prepared via Sonogashira cross coupling reactions with THP-protected propargyl alcohols and aryl halides.

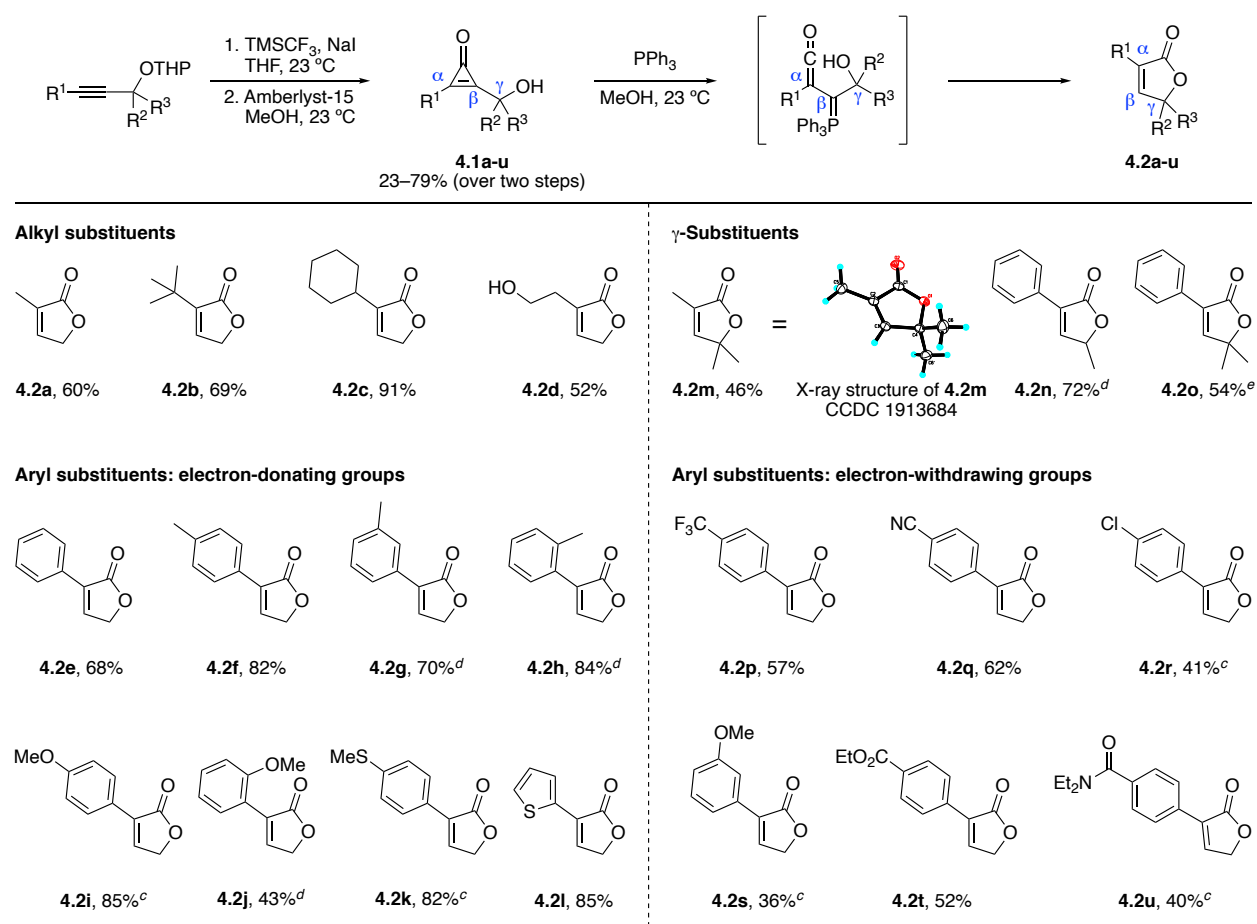


The panel of hydroxymethyl-tethered CpOs was subjected to butenolide formation conditions (Scheme 4-2). As anticipated, the cyclization was tolerant of a broad range of functional groups. α -Alkyl-substituted CpOs (**4.2a-d**), including branched (**4.2b**), and cyclic (**4.2c**) substrates, were efficiently converted. The reaction also proceeded in the presence of competing nucleophiles (**4.2d**), albeit with lower yields. α -Aryl substituted CpOs also generated butenolides in the presence of PPh_3 . Both electron-donating (**4.2e-l**) and electron-withdrawing (**4.2p-u**) groups were examined, along with thiophene heterocycles (**4.2l**). Notably, robust product formation was observed even in the presence of electrophilic esters (**4.2t**) and cyano groups (**4.2q**). The reaction was also tolerant of different substitution patterns on the aryl ring (**4.2f-j**, **4.2s**).

The CpO-phosphine reaction further enabled access to more highly substituted butenolides. As noted above, the cyclization is tolerant of numerous α -substituents (Scheme 4-2). These groups are positioned away from the ketene, and thus minimally interfere with trapping. Additional substituents on the hydroxymethyl tether provided access to α,γ -disubstituted butenolides (**4.2m-o**). Compound **4.2n**, in particular,

comprises the α,γ -substitution pattern present in incrustoporin and related natural products (Figure 4-1A). While the CpO reaction cannot provide β -substituted butenolides, such scaffolds are readily accessible post-cyclization. Cycloadditions [28-29] and Heck couplings [30] can be used in this regard, along with several other methods [31-32]. Notably, functionalized butenolides can also serve as gateways to butyrolactones [33-36] and other interesting scaffolds [37-38].

Scheme 4-2. Diverse butenolides were synthesized from substituted CpOs^{a,b}



^aReaction conditions: CpO (1 equiv), PPh₃ (5 mol %), methanol (0.25 M). ^bIsolated yields. ^c10 mol % PPh₃. ^d20 mol % PPh₃. ^eC₆H₆ (0.25 M) was used.

Our collective results demonstrated that butenolides are favored in the hydroxymethyl CpO-phosphine reaction. All but one of the cyclizations proceeded with no competing β -lactone formation. Such side products were observed only when CpO **4.1m** was subjected to low catalyst loadings in methanol (Figures 4-3A). The *gem*-dimethyl groups on **4.1m** likely promoted phosphine addition to the more accessible C3 position and accelerated β -lactone cyclization to provide **4.3a-b** [39]. An appreciable amount of product **4.2m** was still formed under these challenging conditions, though, supporting the original hypothesis that butenolide formation can predominate, despite the potential for competing pathways.

We carried out additional experiments to examine the propensity for hydroxymethyl CpOs to form butenolides. The observed product distributions likely reflect the faster rate of five- versus four-membered ring cyclization following ketene formation (Figure 4-2B). An alternative explanation is that phosphine attack is favored at C2. To investigate these possibilities, we devised a competitive trapping experiment with diol **4.1d**. This compound comprises two hydroxy group tethers, and is capable of cyclizing to five- or six-membered rings, depending on the ketene formed. We anticipated that **4.1d** would provide a mixture of γ - and δ -lactones upon phosphine treatment, as the steric environments around C2 and C3 are similar and unlikely to bias phosphine addition. Additionally, five- or six-membered ring formation (post-ketene generation) should be similarly facile. When diol **4.1d** was treated with PPh_3 , a mixture of γ -lactones was observed, in addition to butenolide **4.2d** (See Appendix B). No δ -lactone (**4.5**) formed, though, to our surprise (Figure 4-3B). The ketene en route to **4.5** was capable of forming and cyclizing, as demonstrated with a CpO probe outfitted with a single hydroxyethyl tether (See Appendix B). These data suggest

that phosphine addition to C2 could be preferred when CpOs are functionalized with hydroxymethyl appendages. The exact mechanism is the subject of ongoing work, but the tether could promote internal hydrogen-bond activation, preorganizing the CpO for C2 attack.

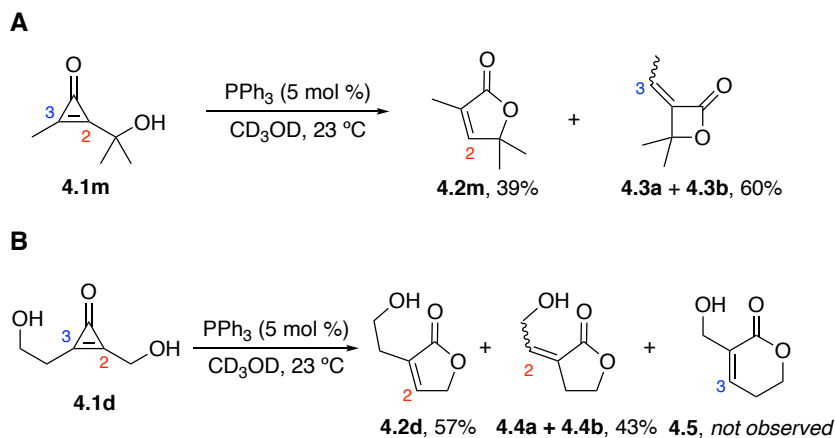


Figure 4-3. Mechanistic studies involving butenolide formation. (A) Upon phosphine treatment, CpO **4.1m** formed β -lactone products **4.3a-b**. The additional steric bulk at C2 likely disfavored phosphine addition. (B) δ -Lactone **4.5** was not observed when diol-CpO **4.1d** was treated with PPh_3 . This reaction produced a mixture of γ -lactones (**4.4a-b**), in addition to the desired butenolide **4.2d**. For (A)-(B), percent conversion values (from NMR analyses) are reported.

4.4 Conclusions

In conclusion, we developed a method to prepare substituted butenolides using mild and bioorthogonal reagents: hydroxymethyl-tethered cyclopropenones and aryl phosphines. This method features mild reaction conditions and low catalyst loadings, and can produce a variety of targets. The reaction exhibits wide functional group tolerance, typical of biocompatible reagents. The reported transformation is complementary to existing methods that furnish α,γ -substituted butenolides [40-49], but is potentially more generalizable. Importantly, the requisite hydroxymethyl CpOs can be derived from propargyl alcohols, widely used and available materials in organic synthesis. We further

anticipate that CpOs and other bioorthogonal reagents will continue to inspire the development of useful methodologies.

4.5 Methods and Materials

4.5.1 General Information

Reactions were run at ambient temperature under a nitrogen atmosphere, unless otherwise indicated. Tetrahydrofuran (THF), diethyl ether (Et₂O), dichloromethane (CH₂Cl₂), dimethylformamide (DMF), and acetonitrile (MeCN) were degassed with argon and run through two 4 x 36 inch columns of anhydrous neutral A-2 (8 x 14 mesh; LaRoche Chemicals; activated under a flow of argon at 350 °C for 12 h). Thin-layer chromatography was performed using Silica Gel 60 F₂₅₄-coated glass plates (0.25 mm thickness), and visualization was performed with KMnO₄ stain and/or UV irradiation. Chromatography was performed with 60 Å (240–400 mesh) silica gel, commercially available from Sorbent Technologies. Organic solutions were concentrated under reduced pressure using a Büchi rotary evaporator. NMR spectra were collected on a Bruker DRX400 instrument (400 MHz ¹H, 376 MHz ¹⁹F, 162 MHz ³¹P), a Bruker DRX500 instrument equipped with a cryo probe (500 MHz ¹H), or an AVANCE600 instrument equipped with a cryo probe (600 MHz ¹H, 151 MHz ¹³C). Spectra were internally referenced to residual solvent signals (CDCl₃ was referenced to 7.26 ppm for ¹H and 77.16 ppm for ¹³C, CD₃OD was referenced to 3.31 ppm for ¹H and 49.0 ppm for ¹³C, C₆D₆ was referenced to 7.16 ppm for ¹H and 128.06 ppm for ¹³C, (CD₃)₂SO was referenced to 2.50 ppm for ¹H, D₂O was referenced to 4.79 ppm for ¹H). ¹⁹F and ³¹P NMR spectra were referenced by indirect absolute chemical shift to residual protio solvent signals. All spectra were collected at 298 K unless stated otherwise. High resolution mass spectrometry (HRMS) was performed by the University

of California, Irvine Mass Spectrometry Facility. Crystallographic data were acquired and processed by the University of California, Irvine X-Ray Crystallography Facility.

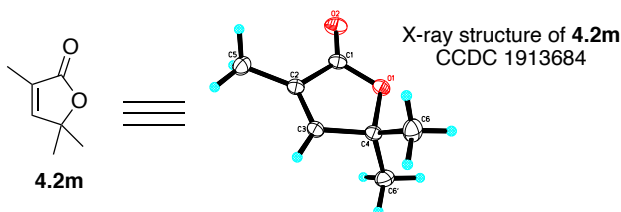
4.5.2 Optimization of butenolide cyclization conditions

Reaction conditions were analyzed using ^1H NMR spectroscopy. Cyclopropenone **4.1a** (25 mM) and phosphine (1–100 mol %) were incubated in the presence of air at ambient temperature. NMR spectra were acquired over 24 h, or until full conversion was observed.

4.5.3 Crystallization of compound **4.2m**

A scintillation vial containing compound **4.2m** (~60 mg) was dissolved in Et_2O and concentrated *in vacuo* using a Büchi rotary evaporator. The resulting residue was placed under reduced pressure (~0.1 mm Hg). After 2 h, crystals were serendipitously found deposited onto the needle. Crystals were transferred to a new scintillation vial, parafilm, and stored at $-20\text{ }^\circ\text{C}$ until diffraction.

X-ray crystal structure of compound **4.2m**



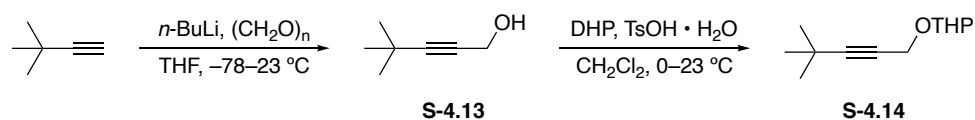
A colorless crystal (0.169 x 0.306 x 0.408 mm) was mounted on a glass fiber and transferred to a Bruker SMART APEX II diffractometer. The APEX2 [50] program package was used to determine the unit-cell parameters and for data collection (15 sec/frame scan time for a sphere of diffraction data). The raw frame data was processed using SAINT [51] and SADABS [52] to yield the reflection data file. Subsequent calculations were carried out using the SHELXTL [53] program. The diffraction symmetry was $2/m$ and the systematic absences were consistent with the monoclinic space groups $C2$, Cm and $C2/m$. It was later determined that space group $C2/m$ was correct.

The structure was solved by direct methods and refined on F^2 by full-matrix least-squares techniques. The analytical scattering factors [54] for neutral atoms were used throughout the analysis. Hydrogen atoms were located from a difference-Fourier map and refined (x, y, z and U_{iso}). The molecule was located on a mirror plane.

Least-squares analysis yielded $wR2 = 0.0922$ and $Goof = 1.062$ for 74 variables refined against 920 data (0.74 Å), $R1 = 0.0341$ for those 864 data with $I > 2.0\sigma(I)$.

4.5.4 Synthetic Procedures

Compounds **S-4.1** [15], **S-4.5** [55], **S-4.6** [56], **S-4.7** [57], **S-4.8** [58], **S-4.9** [59], **S-4.10** [60], **S-4.11** [61], and **S-4.12** [62] were synthesized as previously described. All other reagents were obtained from commercial sources and used without further purification.



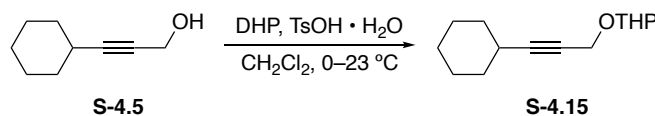
2-((4,4-Dimethylpent-2-yn-1-yl)oxy)tetrahydro-2H-pyran (**S-4.14**)

To a flame-dried round-bottom flask was added 3,3-dimethylbut-1-yne (0.50 mL, 4.1 mmol, 1.0 equiv.) in anhydrous THF (8.0 mL). The solution was cooled to $-78\text{ }^\circ\text{C}$, and *n*-butyllithium (2.5 M in hexanes, 1.8 mL, 4.5 mmol, 1.1 equiv.) was added dropwise. The solution was stirred at $-78\text{ }^\circ\text{C}$ for 1 h. Paraformaldehyde (0.145 g, 4.83 mmol, 1.20 equiv.) was added in one portion, and the reaction was stirred at ambient temperature. When full consumption of the alkyne was observed (as determined by TLC), the solution was quenched with sat. NH_4Cl (90 mL) and extracted with EtOAc (3 x 30 mL). The organic layers were combined, dried with MgSO_4 and filtered. The solvent was concentrated *in vacuo*, and the resulting crude oil (**S-4.13**) was used in the next step without further purification.

To a flame-dried round-bottom flask was added 4-toluenesulfonic acid monohydrate ($\text{TsOH} \cdot \text{H}_2\text{O}$, 5.2 mg, 0.040 mmol, 1 mol %) and a solution of **S-4.13** (0.798 g, 4.07 mmol, 1.00 equiv.) in anhydrous CH_2Cl_2 (8.0 mL). After cooling to $0\text{ }^\circ\text{C}$, 3,4-dihydro-2H-pyran (DHP, 0.39 mL, 4.5 mmol, 1.1 equiv.) was added dropwise, and the

solution was slowly warmed to ambient temperature and stirred overnight. The resulting dark solution was diluted with CH₂Cl₂ (30 mL) and washed with sat. NaHCO₃ (1 x 50 mL). The aqueous layer was extracted with CH₂Cl₂ (2 x 20 mL), and the organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (eluting with 5% EtOAc in hexanes) to give compound **S-4.14** (0.65 g, 81% over two steps) as a clear, viscous oil.

¹H NMR (600 MHz, CDCl₃) δ 4.86 (t, *J* = 3.4 Hz, 1H), 4.31 (d, *J* = 15.4 Hz, 1H), 4.25 (d, *J* = 15.2 Hz, 1H), 3.89 (ddd, *J* = 11.5, 9.3, 2.9 Hz, 1H), 3.52–3.49 (m, 1H), 1.86–1.79 (m, 1H), 1.75–1.70 (m, 1H), 1.64–1.56 (m, 2H), 1.55–1.48 (m, 2H). ¹³C NMR (151 MHz, CDCl₃) δ 96.4, 94.8, 74.2, 62.0, 54.6, 31.0, 30.4, 27.4, 25.4, 19.2. HRMS (ESI⁺) calcd. for C₁₂H₂₀O₂Na [M+Na]⁺ 219.1361 *m/z*, found 219.1360.

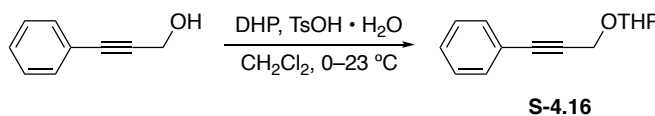


2-((3-Cyclohexylprop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (**S-4.15**)

To a flame-dried round-bottom flask was added TsOH · H₂O (5.1 mg, 0.027 mmol, 1 mol %) and a solution of **S-4.5** (0.354 g, 2.56 mmol, 1.00 equiv.) in anhydrous CH₂Cl₂ (3 mL). After cooling to 0 °C, DHP (0.24 mL, 2.8 mmol, 1.1 equiv.) was added dropwise, and the solution was slowly warmed to ambient temperature and stirred overnight. The resulting dark solution was diluted with CH₂Cl₂ (30 mL) and washed with sat. NaHCO₃ (1 x 50 mL). The aqueous layer was extracted with CH₂Cl₂ (2 x 20 mL), and the organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude

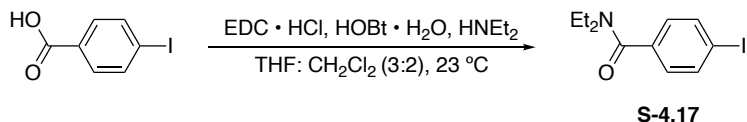
residue was purified by flash column chromatography (eluting with 5% Et₂O in hexanes) to give compound **S-4.15** (0.46 g, 81%) as a colorless oil.

¹H NMR (500 MHz, CDCl₃) δ 4.83 (t, *J* = 3.4 Hz, 1H), 4.30 (dd, *J* = 15.3, 1.9 Hz, 1H), 4.23 (dd, *J* = 15.3, 1.8 Hz, 1H), 3.85 (ddd, *J* = 11.6, 9.4, 2.8 Hz, 1H), 2.39 (m, 1H), 1.87–1.48 (m, 12H), 1.46–1.40 (m, 2H), 1.34–1.24 (m, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 96.7, 90.9, 75.7, 62.1, 54.8, 32.8, 30.5, 29.3, 26.0, 25.6, 25.0, 19.3. HRMS (ESI⁺) calcd. for C₁₄H₂₂O₂Na [M+Na]⁺ 245.1517 *m/z*, found 245.1512.



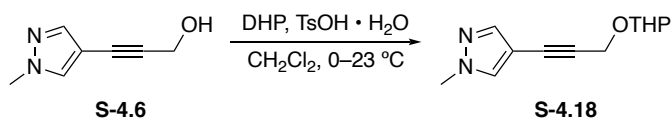
2-((3-Phenylprop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (**S-4.16**)

To a flame-dried round-bottom flask was added TsOH • H₂O (28.8 mg, 0.151 mmol, 1 mol %), anhydrous CH₂Cl₂ (30 mL), and 3-phenyl-2-propyn-1-ol (1.86 mL, 15.1 mmol, 1 equiv.). After cooling to 0 °C, DHP (1.28 mL, 15.1 mmol, 1.00 equiv.) was added dropwise, and the solution was slowly warmed to ambient temperature and stirred overnight. The resulting dark solution was diluted with CH₂Cl₂ (30 mL) and washed with sat. NaHCO₃ (1 x 100 mL). The aqueous layer was extracted with CH₂Cl₂ (3 x 50 mL), and the organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (eluting with 0–5% EtOAc in hexanes) to give compound **S-4.16** (2.80 g, 85%) as a pale yellow oil. NMR spectra matched those previously reported [63].



***N,N*-Diethyl-4-iodobenzamide (S-4.17)**

To a flame-dried round-bottom flask was added 4-iodobenzoic acid (0.496 g, 2.00 mmol, 1.00 equiv.), EDC · HCl (0.461 g, 2.40 mmol, 1.20 equiv.), and HOBT · H₂O (0.367 g, 2.40 mmol, 1.20 equiv.). Anhydrous THF (3.0 mL), anhydrous CH₂Cl₂ (2.0 mL), and diethylamine (0.62 mL, 6.0 mmol, 3.0 equiv.) were added, and the solution was stirred at ambient temperature overnight. The reaction was concentrated *in vacuo* and the crude residue was purified by flash column chromatography (eluting with 0–2% MeOH in CH₂Cl₂) to give compound **S-4.17** (0.58 g, 96%) as a white powder. NMR spectra matched those previously reported [64].

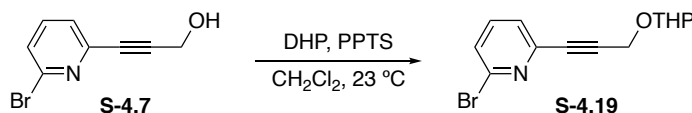


1-Methyl-4-(3-((tetrahydro-2*H*-pyran-2-yl)oxy)prop-1-yn-1-yl)-1*H*-pyrazole (S-4.18)

To a flame-dried round-bottom flask was added TsOH · H₂O (12.2 mg, 0.0641 mmol, 6 mol %) and a solution of **S-4.6** (0.152 g, 1.12 mmol, 1.00 equiv.) in anhydrous CH₂Cl₂ (3.0 mL). After cooling to 0 °C, DHP (0.18 mL, 2.0 mmol, 1.8 equiv.) was added dropwise, and the solution was slowly warmed to ambient temperature and stirred overnight. The resulting solution was diluted with CH₂Cl₂ (30 mL) and washed with sat. NaHCO₃ (1 x 50 mL). The aqueous layer was extracted with CH₂Cl₂ (2 x 20 mL), and the

organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (eluting with 20% EtOAc in hexanes) to give compound **S-4.18** (0.18 g, 74%) as a yellow oil.

¹H NMR (400 MHz, C₆D₆) δ 7.64 (s, 1H), 6.81 (s, 1H), 4.97 (t, *J* = 3.3 Hz, 1H), 4.52 (d, *J* = 15.7 Hz, 1H), 4.46 (d, *J* = 15.7 Hz, 1H), 3.75 (ddd, *J* = 13.5, 11.0, 2.9 Hz, 1H), 3.41–3.36 (m, 1H), 3.02 (s, 3H), 1.80–1.67 (m, 1H), 1.66–1.53 (m, 2H), 1.41–1.31 (m, 1H), 1.30–1.18 (m, 2H). ¹³C NMR (151 MHz, C₆D₆) δ 142.2, 132.8, 103.2, 96.6, 86.7, 78.0, 61.5, 54.9, 38.2, 30.6, 25.8, 19.2. HRMS (ESI⁺) calcd. for C₁₂H₁₆N₂O₂Na [M+Na]⁺ 243.1109 *m/z*, found 243.1106.



2-Bromo-6-(3-((tetrahydro-2H-pyran-2-yl)oxy)prop-1-yn-1-yl)pyridine (**S-4.19**)

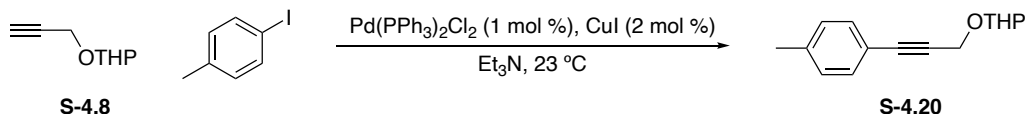
Compound **S-4.19** was synthesized following the general procedure of Duffey, *et al* [56]. To a flame-dried round-bottom flask was added a solution of **S7** (0.201 g, 0.947 mmol, 1.00 equiv.) in anhydrous CH₂Cl₂ (3.0 mL). DHP (0.17 mL, 1.9 mmol, 2.0 equiv.) was added dropwise, followed by pyridinium *p*-toluenesulfonate (PPTS, 14.2 mg, 0.0577 mmol, 6 mol %). The solution was stirred at ambient temperature overnight. The reaction was poured into a separatory funnel containing sat. NaHCO₃ (40 mL), then extracted with CH₂Cl₂ (2 x 20 mL). The combined organic layers were dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography

(eluting with 10% EtOAc in hexanes) to give compound **S-4.19** (0.21 g, 76%) as a yellow oil.

^1H NMR (600 MHz, CDCl_3) δ 7.50 (t, $J = 7.8$ Hz, 1H), 7.43 (d, $J = 7.6$ Hz, 1H), 7.39 (d, $J = 7.4$ Hz, 1H), 4.87 (t, $J = 3.2$ Hz, 1H), 4.52 (d, $J = 16.1$ Hz, 1H), 4.46 (d, $J = 16.1$ Hz, 1H), 3.86 (ddd, $J = 11.5, 9.3, 2.9$ Hz, 1H), 3.57–3.54 (m, 1H), 1.88–1.80 (m, 1H), 1.78–1.73 (m, 1H), 1.67–1.59 (m, 2H), 1.56–1.52 (m, 2H). ^{13}C NMR (151 MHz, CDCl_3) δ 143.4, 141.8, 138.4, 127.9, 126.2, 97.2, 87.3, 83.9, 62.2, 54.6, 30.3, 25.5, 19.1. HRMS (ESI $^+$) calcd. for $\text{C}_{13}\text{H}_{14}\text{BrNO}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 318.0106 m/z , found 318.0094.

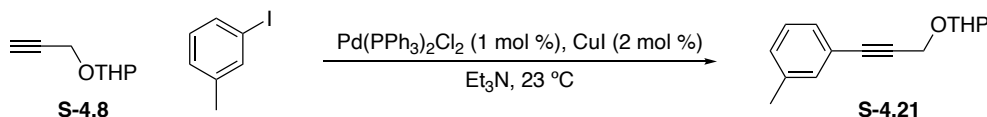
4.4.4a General procedure A for Sonogashira Coupling

To a flame-dried round-bottom flask was added $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (1 mol %), copper (I) iodide (CuI , 2–3 mol %), anhydrous triethylamine (Et_3N , 0.25–0.31 M aryl halide), aryl halide (1.0 equiv.), and alkyne (1.0–1.1 equiv.). The mixture was stirred at ambient temperature overnight. The resultant slurry was dissolved in CH_2Cl_2 (20 mL) and washed with sat. NH_4Cl (1 x 30 mL). The organic layer was dried with MgSO_4 , filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (eluting with hexanes/EtOAc) to afford pure product.



2-((3-(*p*-Tolyl)prop-2-yn-1-yl)oxy)tetrahydro-2*H*-pyran (S-4.20)

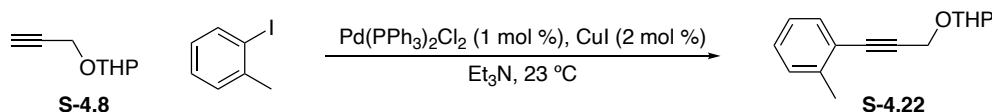
General procedure A was used with the following reagents: 4-iodotoluene (0.654 g, 3.00 mmol, 1.00 equiv.), **S-4.8** (0.42 mL, 3.0 mmol, 1.0 equiv.), Pd(PPh₃)₂Cl₂ (21.0 mg, 0.0299 mmol, 1 mol %), CuI (12.0 mg, 0.0630 mmol, 2 mol %), anhydrous Et₃N (12 mL). The crude residue was purified by flash column chromatography (eluting with 1–8% EtOAc in hexanes) to give compound **S-4.20** (0.67 g, 96%) as a yellow oil. NMR spectra matched those previously reported [63].



2-((3-(*m*-Tolyl)prop-2-yn-1-yl)oxy)tetrahydro-2*H*-pyran (S-4.21)

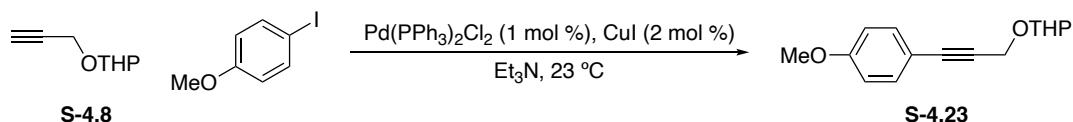
General procedure A was used with the following reagents: 3-iodotoluene (0.13 mL, 1.0 mmol, 1.0 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₂Cl₂ (7.2 mg, 0.010 mmol, 1 mol %), CuI (4.3 mg, 0.023 mmol, 2 mol %), anhydrous Et₃N (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 1–3% EtOAc in hexanes) to give compound **S-4.21** (0.19 g, 80%) as a light brown oil.

^1H NMR (400 MHz, CDCl_3) δ 7.28–7.24 (m, 2H), 7.19 (t, $J = 7.5$ Hz, 1H), 7.12 (d, $J = 7.6$ Hz, 1H), 4.91 (t, $J = 3.4$ Hz, 1H), 4.51 (d, $J = 15.7$ Hz, 1H), 4.45 (d, $J = 15.7$ Hz, 1H), 3.89 (ddd, $J = 11.9, 9.0, 3.1$ Hz, 1H), 3.59–3.54 (m, 1H), 2.32 (s, 3H), 1.92–1.73 (m, 2H), 1.70–1.54 (m, 4H). ^{13}C NMR (151 MHz, CDCl_3) δ 138.0, 132.5, 129.4, 129.0, 128.3, 122.7, 96.9, 86.1, 84.9, 62.1, 54.9, 30.4, 25.5, 21.3, 19.2. HRMS (ESI $^+$) calcd. for $\text{C}_{15}\text{H}_{18}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 253.1205 m/z , found 253.1195.



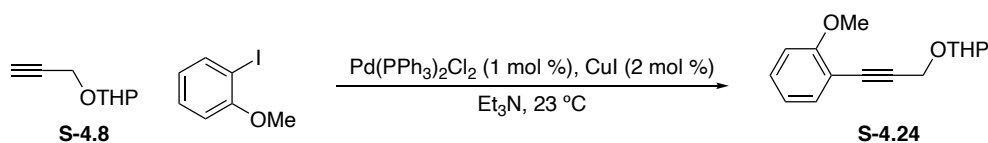
2-((3-(*o*-Tolyl)prop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (**S-4.22**)

General procedure A was used with the following reagents: 2-iodotoluene (0.13 mL, 1.0 mmol, 1.0 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (8.1 mg, 0.012 mmol, 1 mol %), CuI (4.2 mg, 0.022 mmol, 2 mol %), anhydrous Et_3N (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 1% EtOAc in hexanes) to give compound **S-4.22** (0.18 g, 78%) as a yellow oil. NMR spectra matched those previously reported [63].



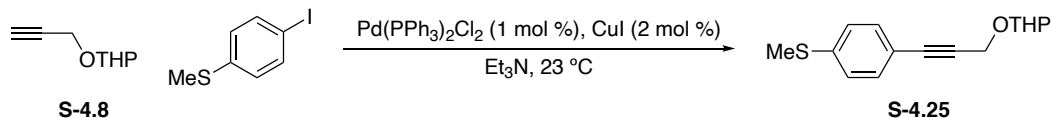
2-((3-(4-Methoxyphenyl)prop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (**S-4.23**)

General procedure A was used with the following reagents: 4-iodoanisole (0.235 g, 1.00 mmol, 1.00 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₂Cl₂ (7.2 mg, 0.010 mmol, 1 mol %), Cul (4.1 mg, 0.022 mmol, 2 mol %), anhydrous Et₃N (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 1–3% EtOAc in hexanes) to give compound **S-4.23** (0.22 g, 89%) as a yellow oil. NMR spectra matched those previously reported [65].



2-((3-(2-methoxyphenyl)prop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (**S-4.24**)

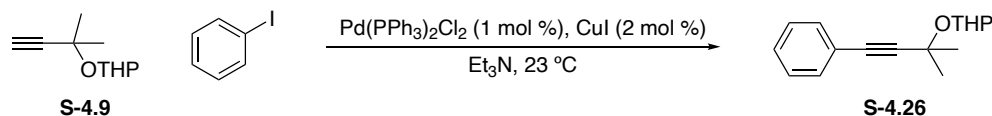
General procedure A was used with the following reagents: 2-iodoanisole (0.26 mL, 2.0 mmol, 1.0 equiv.), **S-4.8** (0.31 mL, 2.2 mmol, 1.1 equiv.), Pd(PPh₃)₂Cl₂ (14.1 mg, 0.0199 mmol, 1 mol %), Cul (7.5 mg, 0.039 mmol, 2 mol %), anhydrous Et₃N (8.0 mL). The crude residue was purified by flash column chromatography (eluting with 2–3% EtOAc in hexanes) to give compound **S-4.24** (0.44 g, 89%) as a brown oil. NMR spectra matched those previously reported [63].



2-((3-(4-(Methylthio)phenyl)prop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (S-4.25)

General procedure A was used with the following reagents: 4-iodoanisole (0.250 g, 1.00 mmol, 1.00 equiv.), **S-4.8** (0.14 mL, 1.0 mmol, 1.0 equiv.), Pd(PPh₃)₂Cl₂ (7.0 mg, 0.010 mmol, 1 mol %), CuI (3.8 mg, 0.020 mmol, 2 mol %), anhydrous Et₃N (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 99:0:1–93:6:1 hexanes:EtOAc:C₆H₆) to give compound **S-4.25** (0.24 g, 92%) as an off-white solid.

¹H NMR (400 MHz, CDCl₃) δ 7.36–7.30 (m, 2H), 7.16–7.09 (m, 2H), 4.87 (t, *J* = 3.6 Hz, 1H), 4.50 (d, *J* = 16.0 Hz, 1H), 4.42 (d, *J* = 16.0 Hz, 1H), 3.86 (ddd, *J* = 12.0, 8.8, 3.2 Hz, 1H), 3.57–3.50 (m, 1H), 2.44 (s, 3H), 1.90–1.46 (m, 6H). ¹³C NMR (151 MHz, CDCl₃) δ 139.5, 132.1, 125.7, 119.0, 96.8, 85.6, 85.2, 62.0, 54.8, 30.3, 25.4, 19.1, 15.3. HRMS (ESI⁺) calcd. for C₁₅H₁₈O₂S [M+Na]⁺ 285.0925 *m/z*, found 285.0925.

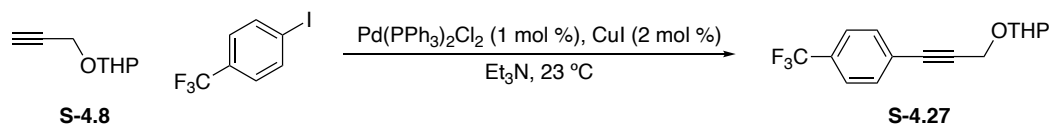


2-((2-Methyl-4-phenylbut-3-yn-2-yl)oxy)tetrahydro-2H-pyran (S-4.26)

General procedure A was used with the following reagents: iodobenzene (0.41 mL, 3.7 mmol, 1.0 equiv.), **S-4.9** (0.683 g, 4.06 mmol, 1.10 equiv.), Pd(PPh₃)₂Cl₂ (25.2 mg, 0.0359 mmol, 1 mol %), CuI (16.7 mg, 0.0878 mmol, 2 mol %), anhydrous Et₃N (12 mL).

The crude residue was purified by flash column chromatography (eluting with 1% EtOAc in hexanes) to give compound **S-4.26** (0.65 g, 72%) as a yellow oil.

^1H NMR (400 MHz, CDCl_3) δ 7.43–7.41 (m, 2H), 7.32–7.28 (m, 3H), 5.16–5.13 (m, 1H), 4.02–3.97 (m, 1H), 3.56–3.50 (m, 1H), 1.92–1.82 (m, 1H), 1.79–1.71 (m, 1H), 1.63 (s, 3H), 1.59 (s, 3H), 1.57–1.52 (m, 4H). ^{13}C NMR (151 MHz, CDCl_3) δ 131.8, 128.4, 128.3, 123.1, 96.5, 91.8, 84.1, 71.7, 63.6, 32.2, 30.9, 30.1, 25.6, 20.8. HRMS (ESI⁺) calcd. for $\text{C}_{16}\text{H}_{20}\text{O}_2\text{Na}$ [M+Na]⁺ 267.1361 m/z , found 267.1360.

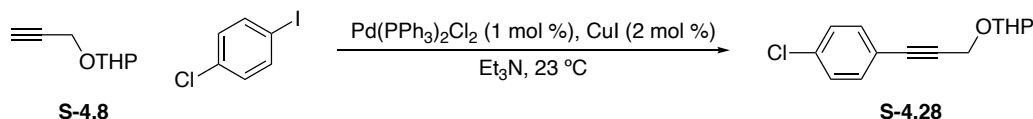


2-((3-(4-(Trifluoromethyl)phenyl)prop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (**S-4.27**)

General procedure A was used with the following reagents: 4-iodobenzotrifluoride (0.15 mL, 1.0 mmol, 1.0 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (7.6 mg, 0.010 mmol, 1 mol %), CuI (5.5 mg, 0.029 mmol, 3 mol %), anhydrous Et_3N (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 1–2% EtOAc in hexanes) to give compound **S-4.27** (0.18 g, 63%) as a pale yellow oil.

^1H NMR (500 MHz, CDCl_3) δ 7.56–7.52 (m, 4H), 4.88 (t, $J = 3.4$ Hz, 1H), 4.53 (d, $J = 15.9$ Hz, 1H), 4.46 (d, $J = 16.0$ Hz, 1H), 3.88 (ddd, $J = 11.8, 9.3, 3.0$ Hz, 1H), 3.59–3.54 (m, 1H), 1.89–1.53 (m, 6H). ^{19}F NMR (376 MHz, CDCl_3) δ -63.1. ^{13}C NMR (151 MHz, CDCl_3) δ 132.1, 130.1 (q, $J = 32.7$ Hz), 126.7 (d_{app}, $J = 1.3$ Hz), 125.3 (q, $J = 3.7$ Hz), 124.0 (q, J

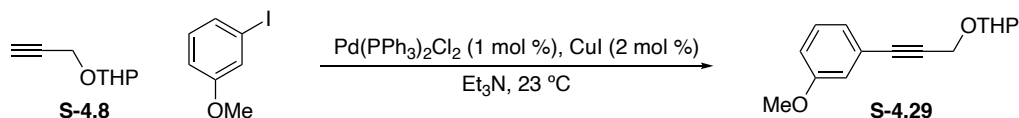
= 272.2 Hz), 97.1, 87.9, 84.5, 62.1, 54.7, 30.4, 25.5, 19.1. HRMS (ESI⁺) calcd. for C₁₅H₁₅F₃O₂Na [M+Na]⁺ 307.0922, found 307.0920.



2-((3-(4-Chlorophenyl)prop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (**S-4.28**)

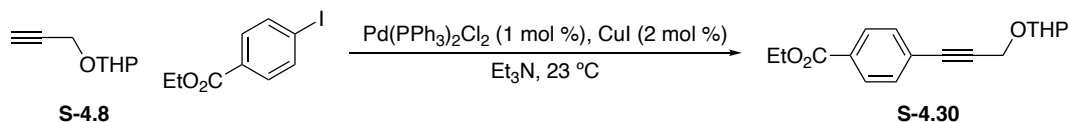
General procedure A was used with the following reagents: 1-chloro-4-iodobenzene (0.242 g, 1.00 mmol, 1.00 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₂Cl₂ (7.2 mg, 0.010 mmol, 1 mol %), CuI (6.2 mg, 0.033 mmol, 3 mol %), anhydrous Et₃N (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 2–5% EtOAc in hexanes) to give compound **S-4.28** (0.19 g, 76%) as a yellow oil.

¹H NMR (400 MHz, CDCl₃) δ 7.37 (d, *J* = 8.4 Hz, 2H), 7.27 (d, *J* = 8.6 Hz, 2H), 4.88 (t, *J* = 2.8 Hz, 1H), 4.51 (d, *J* = 15.8 Hz, 1H), 4.44 (d, *J* = 15.8 Hz, 1H), 3.89 (ddd, *J* = 12.0, 9.4, 3.0 Hz, 1H), 3.59–3.54 (m, 1H), 1.89–1.56 (m, 6H). ¹³C NMR (151 MHz, CDCl₃) δ 134.6, 132.2, 128.8, 121.5, 97.2, 86.4, 84.8, 62.2, 54.9, 30.5, 25.6, 19.2. HRMS (ESI⁺) calcd. for C₁₄H₁₅ClO₂Na [M+Na]⁺ 273.0658, found 273.0660.



2-((3-(3-Methoxyphenyl)prop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (S-4.29)

General procedure A was used with the following reagents: 3-iodoanisole (0.12 mL, 1.0 mmol, 1.0 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₂Cl₂ (7.1 mg, 0.010 mmol, 1 mol %), CuI (3.8 mg, 0.020 mmol, 2 mol %), anhydrous Et₃N (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 1–5% EtOAc in hexanes) to give compound **S-4.29** (0.18 g, 71%) as a yellow oil. NMR spectra matched those previously reported [63].

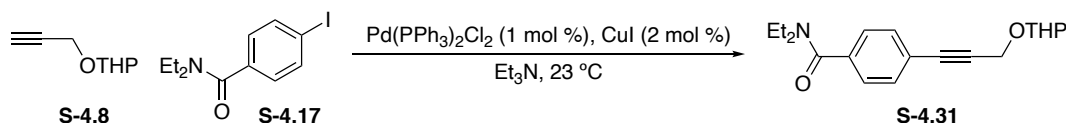


Ethyl 4-(3-((tetrahydro-2H-pyran-2-yl)oxy)prop-1-yn-1-yl)benzoate (S-4.30)

General procedure A was used with the following reagents: ethyl 4-iodobenzoate (0.26 mL, 1.0 mmol, 1.0 equiv.), **S-4.8** (0.14 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₂Cl₂ (7.0 mg, 0.010 mmol, 1 mol %), CuI (4.0 mg, 0.021 mmol, 2 mol %), anhydrous Et₃N (4 mL). The crude residue was purified by flash column chromatography (eluting with 98:1:1 – 84:15:1 hexanes:EtOAc:C₆H₆) to give compound **S-4.30** (0.24 g, 82%) as a yellow oil.

¹H NMR (400 MHz, CDCl₃) δ 8.00–7.93 (m, 2H), 7.52–7.45 (m, 2H), 4.88 (t, *J* = 3.6 Hz, 1H), 4.53 (d, *J* = 15.6 Hz, 1H), 4.46 (d, *J* = 15.6 Hz, 1H), 4.36 (q, *J* = 7.2 Hz, 2H), 3.88 (ddd, *J* = 12.0, 8.8, 3.2 Hz, 1H), 3.61–3.53 (m, 1H), 1.93–1.46 (m, 7H), 1.38 (t, *J* = 7.2

Hz, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 166.0, 131.7, 130.0, 129.4, 127.3, 97.0, 88.2, 85.1, 62.0, 61.1, 54.7, 30.3, 25.4, 19.0, 14.3. HRMS (ESI⁺) calcd. for $\text{C}_{17}\text{H}_{20}\text{O}_4\text{Na}$ $[\text{M}+\text{Na}]^+$ 311.1259 m/z , found 311.1264.



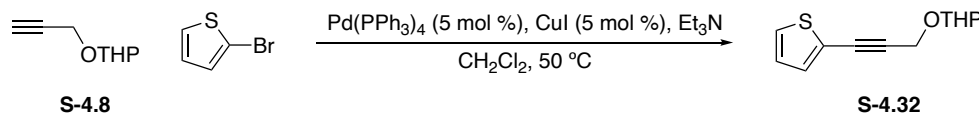
***N,N*-Diethyl-4-(3-((tetrahydro-2*H*-pyran-2-yl)oxy)prop-1-yn-1-yl)benzamide (**S-4.31**)**

General procedure A was used with the following reagents: **S-4.17** (0.542 g, 1.8 mmol, 1.00 equiv.), **S-4.8** (0.28 mL, 2.0 mmol, 1.1 equiv.), $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (13.2 mg, 0.019 mmol, 1 mol %), CuI (8.5 mg, 0.045 mmol, 2 mol %), anhydrous Et_3N (8.0 mL). The crude residue was purified by flash column chromatography (eluting with 50% EtOAc in hexanes) to give compound **S-4.31** (0.56 g, >99%, 97% purity).

^1H NMR (600 MHz, CDCl_3) δ 7.47 (d, $J = 8.2$ Hz, 2H), 7.31 (d, $J = 8.2$ Hz, 2H), 4.89 (t, $J = 3.3$ Hz, 1H), 4.53 (d, $J = 15.9$ Hz, 1H), 4.46 (d, $J = 15.9$ Hz, 1H), 3.89 (ddd, $J = 11.6, 9.5, 3.0$ Hz, 1H), 3.59–3.53 (m, 3H), 3.23 (bs, 2H), 1.89–1.81 (m, 1H), 1.80–1.74 (m, 1H), 1.70–1.62 (m, 2H), 1.59–1.53 (m, 2H), 1.23 (bs, 3H), 1.10 (bs, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 170.7, 137.2, 132.0, 126.5, 123.8, 97.1, 86.5, 85.3, 62.2, 54.9, 43.4, 39.5, 30.4, 25.5, 19.2, 14.3, 13.0. HRMS (ESI⁺) calcd. for $\text{C}_{19}\text{H}_{25}\text{NO}_3\text{Na}$ $[\text{M}+\text{Na}]^+$ 338.1732 m/z , found 338.1734.

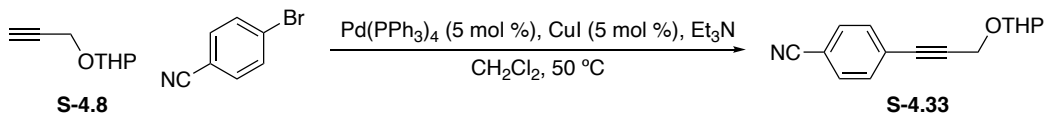
4.4.4b General procedure B for Sonogashira Coupling

To a flame-dried Schlenk tube was added Pd(PPh₃)₄ (5 mol %), copper (I) iodide (CuI, 5–8 mol %), a solution of aryl halide (1.0 equiv.) in anhydrous CH₂Cl₂ (0.25 M aryl halide), anhydrous Et₃N (3.0–6.0 equiv.), and alkyne (1.1 equiv.). The tube was sealed, and the solution was stirred at 50 °C overnight). After cooling to ambient temperature, the reaction was diluted with CH₂Cl₂ (20 mL) and washed with sat. NH₄Cl (1 x 30 mL). The organic layer was dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (eluting with hexanes/EtOAc) to afford pure product.



2-((3-(Thiophen-2-yl)prop-2-yn-1-yl)oxy)tetrahydro-2H-pyran (S-4.32)

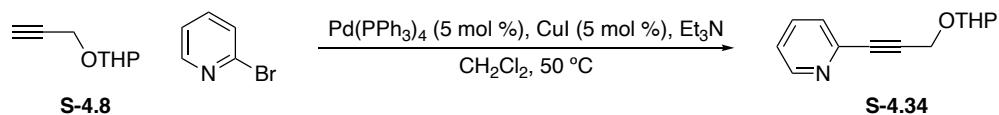
General procedure B was used with following reagents: 2-bromothiophene (0.10 mL, 1.0 mmol, 1.0 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₄ (56.8 mg, 0.0492 mmol, 5 mol %), CuI (15.6 mg, 0.0819 mmol, 8 mol %), anhydrous Et₃N (0.44 mL, 3.1 mmol, 3.0 equiv.), anhydrous CH₂Cl₂ (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 2% EtOAc in hexanes) to give compound **S-4.32** (0.13 g, 57%) as a yellow oil. NMR spectra matched those previously reported [66].



4-(3-((Tetrahydro-2H-pyran-2-yl)oxy)prop-1-yn-1-yl)benzonitrile (S-4.33)

General procedure B was used with the following reagents: 4-bromobenzonitrile (0.203 g, 1.10 mmol, 1.00 equiv.), **S-4.8** (0.17 mL, 1.2 mmol, 1.1 equiv.), Pd(PPh₃)₄ (64.5 mg, 0.0558 mmol, 5 mol %), CuI (11.1 mg, 0.0583 mmol, 5 mol %), anhydrous Et₃N (0.47 mL, 3.3 mmol, 3.0 equiv.), anhydrous CH₂Cl₂ (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 3–5% Et₂O in hexanes) to give compound **S-4.33** (0.22 g, 82%) as an amber oil.

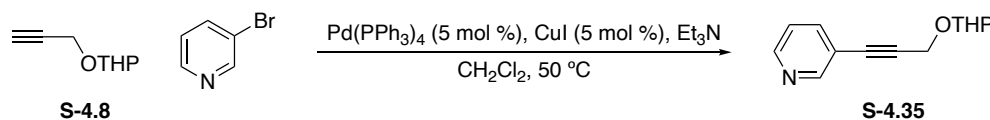
¹H NMR (400 MHz, CDCl₃) δ 7.60 (d, *J* = 8.4 Hz, 2H), 7.52 (d, *J* = 8.4 Hz, 2H), 4.87 (t, *J* = 3.3 Hz, 1H), 4.54 (d, *J* = 16.0 Hz, 1H), 4.47 (d, *J* = 16.0 Hz, 1H), 3.88 (ddd, *J* = 12.0, 9.4, 3.0 Hz, 1H), 3.60–3.54 (m, 1H), 1.91–1.74 (m, 2H), 1.71–1.53 (m, 4H). ¹³C NMR (151 MHz, CDCl₃) δ 132.5, 132.1, 127.8, 118.6, 112.0, 97.2, 90.0, 84.3, 62.2, 54.8, 30.4, 25.5, 19.1. HRMS (FI⁺) calcd. for C₁₁H₁₅NO₂ [M]⁺ 241.1103 *m/z*, found 241.1106.



2-(3-((Tetrahydro-2H-pyran-2-yl)oxy)prop-1-yn-1-yl)pyridine (S-4.34)

General procedure B was used with the following reagents: 2-bromopyridine (0.10 mL, 1.0 mmol, 1.0 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₄ (60.0 mg, 0.0519 mmol, 5 mol %), CuI (11.0 mg, 0.0578 mmol, 5 mol %), anhydrous Et₃N (0.44 mL,

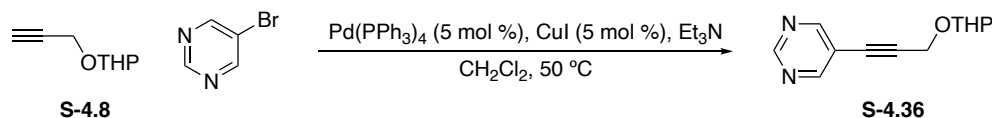
3.2 mmol, 3.0 equiv.), anhydrous CH₂Cl₂ (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 10–50% EtOAc in hexanes) to give compound **S-4.34** (0.13 g, 56%) as a brown oil. NMR spectra matched those previously reported [67].



3-(3-((Tetrahydro-2H-pyran-2-yl)oxy)prop-1-yn-1-yl)pyridine (S-4.35)

General procedure B was used with the following reagents: 3-bromopyridine (0.10 mL, 1.0 mmol, 1.0 equiv.), **S-4.8** (0.16 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₄ (60.4 mg, 0.0523 mmol, 5 mol %), CuI (10.5 mg, 0.0551 mmol, 5 mol %), anhydrous Et₃N (0.88 mL, 6.3 mmol, 6.3 equiv.), anhydrous CH₂Cl₂ (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 2–5% EtOAc in hexanes) to give compound **S-4.35** (0.12 g, 51%) as a yellow oil.

¹H NMR (500 MHz, CDCl₃) δ 8.68 (dd, *J* = 1.9, 0.6 Hz, 1H), 8.53 (dd, *J* = 4.9, 1.7 Hz, 1H), 7.72 (dt, *J* = 7.8, 1.9 Hz, 1H), 7.23 (ddd, *J* = 7.9, 4.9, 0.8 Hz, 1H), 4.88 (t, *J* = 3.4 Hz, 1H), 4.54 (d, *J* = 15.9 Hz, 1H), 4.46 (d, *J* = 15.9 Hz, 1H), 3.88 (ddd, *J* = 11.8, 9.3, 3.0 Hz, 1H), 3.59–3.55 (m, 1H), 1.90–1.81 (m, 1H), 1.80–1.74 (m, 1H), 1.70–1.52 (m, 4H). ¹³C NMR (151 MHz, CDCl₃) δ 152.6, 148.9, 138.9, 123.1, 120.0, 97.2, 88.8, 82.6, 62.2, 54.8, 30.4, 25.5, 19.1. HRMS (ESI⁺) calcd. for C₁₃H₁₆NO₂ [M+H]⁺ 218.1181 *m/z*, found 218.1181.



5-(3-((Tetrahydro-2H-pyran-2-yl)oxy)prop-1-yn-1-yl)pyrimidine (S-4.36)

General procedure B was used with the following reagents: 5-bromopyrimidine (0.159 g, 1.00 mmol, 1.00 equiv.), **S-4.8** (0.15 mL, 1.1 mmol, 1.1 equiv.), Pd(PPh₃)₄ (59.2 mg, 0.0512 mmol, 5 mol %), CuI (15.8 mg, 0.0830 mmol, 8 mol %), anhydrous Et₃N (0.42 mL, 3.0 mmol, 3.0 equiv.), anhydrous CH₂Cl₂ (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 90:9:1 – 85:14:1 hexanes:EtOAc:PhMe) to give compound **S-4.36** (0.15 g, 68%) as a light yellow oil.

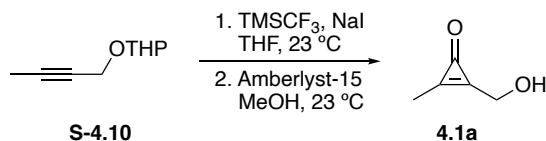
¹H NMR (600 MHz, CDCl₃) δ 9.11 (s, 1H), 8.77 (s, 2H), 4.86 (t, *J* = 3.4 Hz, 1H), 4.54 (d, *J* = 16.0 Hz, 1H), 4.47 (d, *J* = 16.1 Hz, 1H), 3.87 (ddd, *J* = 11.5, 9.5, 3.0 Hz, 1H), 3.58–3.55 (m, 1H), 1.88–1.81 (m, 1H), 1.79–1.73 (m, 1H), 1.69–1.54 (m, 4H). ¹³C NMR (151 MHz, CDCl₃) δ 159.1, 157.1, 119.4, 97.3, 92.8, 79.1, 62.2, 54.6, 30.3, 25.4, 19.0. HRMS (ESI⁺) calcd. for C₁₂H₁₅N₂O₂ [M+H]⁺ 219.1134 *m/z*, found 219.1131.

4.4.4c General procedure C for substituted hydroxymethylcyclopropenones

To a flame-dried Schlenk tube was added sodium iodide (NaI, 1.8–2.5 equiv.). The reagent was gently flame-dried under vacuum, then resuspended in a solution of alkyne (1 equiv.) in anhydrous THF (0.25 M alkyne). Trifluoromethyltrimethylsilane (TMSCF₃, 2–2.2 equiv.) was added, and the vessel was sealed and stirred vigorously at ambient temperature for 2 d. The slurry was diluted with water (30 mL) and extracted with CH₂Cl₂

(3 x 20 mL). The organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The resulting crude difluorocyclopropene was used immediately without further purification.

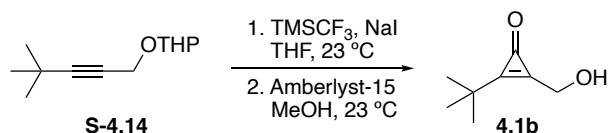
To a flame-dried round-bottom flask was added a solution of crude difluorocyclopropene in anhydrous MeOH (~ 0.5 M w.r.t. alkyne), followed by Amberlyst-15 resin (~70–80 mg mmol⁻¹ alkyne). The mixture was stirred until starting material was fully consumed (as observed by TLC, 1:1 CH₂Cl₂:EtOAc). The mixture was filtered to remove the Amberlyst-15 resin, then concentrated *in vacuo*. If necessary, the crude residue was purified by flash column chromatography to afford pure product.



2-(Hydroxymethyl)-3-methylcycloprop-2-en-1-one (4.1a)

General procedure C was used with following reagents: **S-4.10** (1.08 g, 7.00 mmol, 1.00 equiv.), NaI (2.31 g, 15.4 mmol, 2.20 equiv.), TMSCF₃ (2.1 mL, 15 mmol, 2.2 equiv.), anhydrous THF (28 mL); Amberlyst-15 (0.150 g), anhydrous MeOH (5 mL). The crude residue was purified by flash column chromatography (eluting with 0–50% acetone in EtOAc) to give compound **4.1a** (0.44 g, 64% over two steps) as a brown oil.

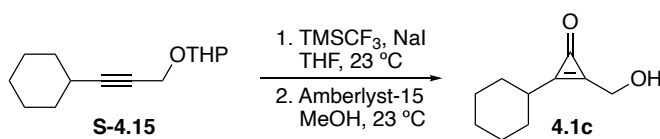
¹H NMR (400 MHz, CD₃OD) δ 4.67 (s, 2H), 2.34 (s, 3H). ¹³C NMR (151 MHz, CD₃OD) δ 162.4, 160.5, 157.6, 58.5, 10.4. HRMS (ESI⁺) calcd. for C₅H₆O₂Na [M+Na]⁺ 121.0266 *m/z*, found 121.0264.



2-(*tert*-Butyl)-3-(hydroxymethyl)cycloprop-2-en-1-one (4.1b)

General procedure C was used with the following reagents: **S-4.14** (0.647 g, 3.30 mmol, 1.00 equiv.), NaI (0.937 g, 6.25 mmol, 1.90 equiv.), TMSCF₃ (1.1 mL, 7.3 mmol, 2.2 equiv.), anhydrous THF (8.0 mL); Amberlyst-15 (0.200 g), anhydrous MeOH (6.0 mL). The crude residue was purified by flash column chromatography (eluting with 25% acetone in CH₂Cl₂) to give compound **4.1b** (0.25 g, 55% over two steps) as a pale yellow solid.

¹H NMR (600 MHz, CDCl₃) δ 4.69 (s, 2H), 4.49 (bs, 1H), 1.31 (s, 9H). ¹³C NMR (151 MHz, CDCl₃) δ 166.1, 158.5, 158.3, 57.8, 33.8, 28.0. HRMS (ESI⁺) calcd. for C₈H₁₂O₂Na [M+Na]⁺ 163.0735 *m/z*, found 163.0733.

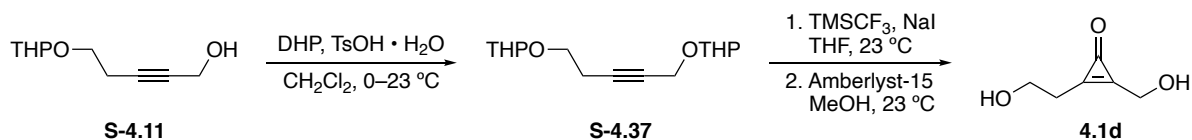


2-Cyclohexyl-3-(hydroxymethyl)cycloprop-2-en-1-one (4.1c)

General procedure C was used with the following reagents: **S-4.15** (0.218 g, 0.980 mmol, 1.00 equiv.), NaI (0.324 g, 2.16 mmol, 2.20 equiv.), TMSCF₃ (0.29 mL, 2.0 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (83.4 mg), anhydrous MeOH (2.0 mL).

The crude residue was purified by flash column chromatography (eluting with 25–50% EtOAc in CH₂Cl₂) to give compound **4.1c** (0.11 g, 65% over two steps) as a yellow oil.

¹H NMR (500 MHz, CD₃OD) δ 4.66 (s, 2H), 2.90–2.84 (m, 1H), 2.00–1.95 (m, 2H), 1.75–1.68 (m, 2H), 1.66–1.59 (m, 3H), 1.50–1.35 (m, 3H). ¹³C NMR (151 MHz, CD₃OD) δ 163.7, 161.2, 160.8, 58.1, 36.9, 30.9, 26.7, 26.0. HRMS (ESI⁺) calcd. for C₁₀H₁₄O₂Na [M+Na]⁺ 189.0892 *m/z*, found 189.0898.



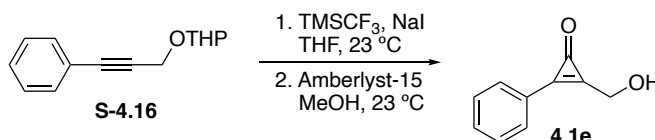
2-(2-Hydroxyethyl)-3-(hydroxymethyl)cycloprop-2-en-1-one (**4.1d**)

To a flame-dried round-bottom flask was added TsOH • H₂O (1.1 mg, 0.0086 mmol, 1 mol %) and a solution of **S-4.11** (0.160 g, 0.868 mmol, 1.00 equiv.) in anhydrous CH₂Cl₂ (2.0 mL). After cooling to 0 °C, DHP (0.083 mL, 0.95 mmol, 1.1 equiv.) was added dropwise, and the solution was slowly warmed to ambient temperature and stirred overnight. The resulting dark solution was diluted with CH₂Cl₂ (30 mL) and washed with sat. NaHCO₃ (1 x 50 mL). The aqueous layer was extracted with CH₂Cl₂ (2 x 20 mL), and the organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude residue passed through a plug of silica (eluting with 30% EtOAc in hexanes) to give **S-4.37**, which was used immediately without further purification.

General procedure C was used with the following reagents: **S-4.37** (0.892 g, 3.33 mmol, 1.00 equiv.), NaI (0.945 g, 6.30 mmol, 1.90 equiv.), TMSCF₃ (1.1 mL, 7.3 mmol,

2.2 equiv.), anhydrous THF (7.0 mL); Amberlyst-15 (0.300 g), anhydrous MeOH (7.0 mL). The crude residue was purified by flash column chromatography (eluting with 10% MeOH in CH₂Cl₂) to give compound **4.1d** (0.27 g, 62% over three steps) as an orange oil.

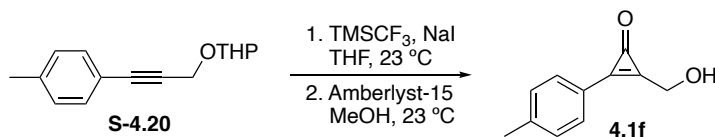
¹H NMR (400 MHz, CDCl₃) δ 4.78 (s, 2H), 3.97 (t, *J* = 5.3 Hz, 2H), 2.93 (t, *J* = 5.3 Hz, 2H). ¹³C NMR (151 MHz, CDCl₃) δ 160.5, 157.7, 157.5, 58.7, 58.3, 29.7. HRMS (ESI⁺) calcd. for C₆H₈O₃Na [M+Na]⁺ 151.0371 *m/z*, found 151.0371.



2-(Hydroxymethyl)-3-phenylcycloprop-2-en-1-one (**4.1e**)

General procedure C was used with the following reagents: **S-4.16** (2.68 g, 12.4 mmol, 1.00 equiv.), NaI (4.10 g, 27.3 mmol, 2.20 equiv.), TMSCF₃ (3.7 mL, 25 mmol, 2.0 equiv.), anhydrous THF (52 mL); Amberlyst-15 (0.750 g), anhydrous MeOH (40 mL). The crude residue was purified by flash column chromatography (eluting with 0–100% acetone in EtOAc) to give compound **4.1e** (1.5 g, 74% over two steps) as a peach solid.

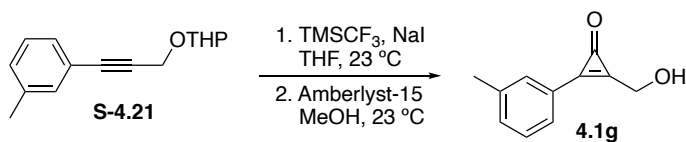
¹H NMR (500 MHz, CDCl₃) δ 7.95 (d, *J* = 8.0 Hz, 2H), 7.53 (m, 3H), 4.93 (s, 2H). ¹³C NMR (125 MHz, CDCl₃) δ 155.5, 154.4, 154.0, 133.1, 132.4, 129.4, 123.0, 58.9. HRMS (ESI⁺) calcd. for C₁₀H₈O₂ [M+Na]⁺ 183.0422 *m/z*, found 183.0422.



2-(Hydroxymethyl)-3-(*p*-tolyl)cycloprop-2-en-1-one (4.1f)

General procedure C was used with the following reagents: **S-4.20** (0.194 g, 0.842 mmol, 1.00 equiv.), NaI (0.278 g, 1.85 mmol, 2.20 equiv.), TMSCF_3 (0.25 mL, 1.7 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (0.145 g), anhydrous MeOH (3.0 mL). After filtration, the resulting solution was concentrated *in vacuo* to give compound **4.1f** (83 mg, 57% over two steps) as a dark brown solid.

^1H NMR (400 MHz, CDCl_3) δ 7.81 (d, $J = 8.0$ Hz, 2H), 7.26 (d, $J = 8.0$ Hz, 2H), 4.86 (s, 2H), 2.40 (s, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 155.9, 153.5, 153.0, 144.1, 132.5, 130.0, 120.2, 58.6, 22.0. HRMS (ESI⁺) calcd. for $\text{C}_{11}\text{H}_{10}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 197.0578 m/z , found 197.0570.

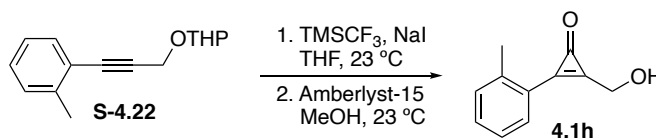


2-(Hydroxymethyl)-3-(*m*-tolyl)cycloprop-2-en-1-one (4.1g)

General procedure C was used with the following reagents: **S-4.21** (0.183 g, 0.794 mmol, 1.00 equiv.), NaI (0.273 g, 1.82 mmol, 2.29 equiv.), TMSCF_3 (0.23 mL, 1.6 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (80.8 mg), anhydrous MeOH (2.0 mL).

After filtration, the solution was concentrated *in vacuo* to give **4.1g** (91 mg, 66% over two steps) as a brown solid.

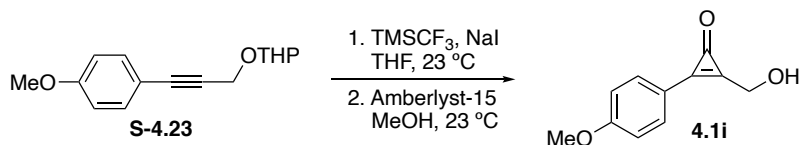
^1H NMR (400 MHz, CD_3OD) δ 7.84 (s, 1H), 7.82–7.80 (m, 1H), 7.47–7.46 (m, 2H), 4.85 (s, 2H), 2.43 (s, 3H). ^{13}C NMR (151 MHz, CD_3OD) δ 157.9, 156.3, 154.2, 140.6, 135.0, 133.8, 130.6, 130.3, 123.8, 58.4, 21.2. HRMS (ESI⁺) calcd. for $\text{C}_{11}\text{H}_{10}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 197.0578 m/z , found 197.0577.



2-(Hydroxymethyl)-3-(o-tolyl)cycloprop-2-en-1-one (**4.1h**)

General procedure C was used with the following reagents: **S-4.22** (0.183 g, 0.796 mmol, 1.00 equiv.), NaI (0.265 g, 1.77 mmol, 2.22 equiv.), TMSCF_3 (0.24 mL, 1.6 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (77.0 mg), anhydrous MeOH (2.0 mL). After filtration, the solution was concentrated *in vacuo* to give compound **4.1h** (92 mg, 67% over two steps) as a peach solid.

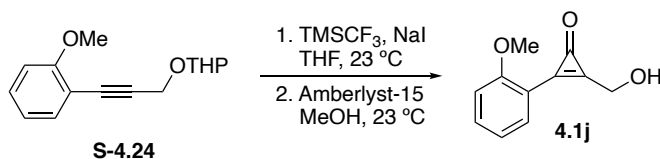
^1H NMR (400 MHz, CD_3OD) δ 8.08 (d, $J = 7.7$ Hz, 1H), 7.49 (td, $J = 7.6, 1.3$ Hz, 1H), 7.40–7.33 (m, 2H), 4.88 (s, 2H), 2.64 (s, 3H). ^{13}C NMR (151 MHz, CD_3OD) δ 157.7, 155.5, 153.5, 142.9, 134.5, 134.1, 131.7, 127.6, 124.1, 58.4, 20.6. HRMS (ESI⁺) calcd. for $\text{C}_{11}\text{H}_{10}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 197.0578 m/z , found 197.0587.



2-(Hydroxymethyl)-3-(4-methoxyphenyl)cycloprop-2-en-1-one (4.1i)

General procedure C was used with the following reagents: **S-4.23** (0.219 g, 0.890 mmol, 1.00 equiv.), NaI (0.299 g, 1.99 mmol, 2.24 equiv.), TMSCF₃ (0.26 mL, 1.8 mmol, 2.0 equiv.), anhydrous THF (4.0 mL). After the difluorocarbene insertion, the residue was passed through a plug of silica (eluting with 15–50% EtOAc in CH₂Cl₂) and used in the next step without further purification. Amberlyst-15 (71.2 mg), anhydrous MeOH (2.0 mL). After filtration, the solution was concentrated *in vacuo* to give compound **4.1i** (55 mg, 33% over two steps) as a tan solid.

¹H NMR (500 MHz, CDCl₃) δ 7.88 (d, *J* = 8.8 Hz, 2H), 6.94 (d, *J* = 8.8 Hz, 2H), 4.85 (s, 2H), 3.85 (s, 3H), 3.66 (br s, 1H). ¹³C NMR (151 MHz, CDCl₃) δ 163.4, 155.5, 153.0, 150.4, 134.5, 115.7, 114.7, 58.5, 55.6. HRMS (ESI⁺) calcd. for C₁₁H₁₁O₃ [M+H]⁺ 191.0708 *m/z*, found 191.0704.

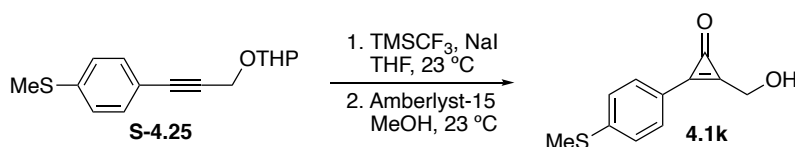


2-(Hydroxymethyl)-3-(2-methoxyphenyl)cycloprop-2-en-1-one (4.1j)

General procedure C was used with the following reagents: **S-4.24** (0.435 g, 1.77 mmol, 1.00 equiv.), NaI (0.538 g, 3.59 mmol, 2.03 equiv.), TMSCF₃ (0.52 mL, 3.5 mmol,

2.0 equiv.), anhydrous THF (7.0 mL); Amberlyst-15 (0.136 g), anhydrous MeOH (3.5 mL). The crude residue was purified by flash column chromatography (eluting with 5% MeOH in CH₂Cl₂) to give compound **4.1j** (0.22 g, 66% over two steps) as an off-white solid.

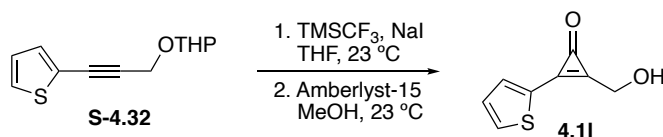
¹H NMR (600 MHz, CDCl₃) δ 7.90 (d, *J* = 7.5 Hz, 1H), 7.62 (t, *J* = 7.9 Hz, 1H), 7.18 (d, *J* = 8.5 Hz, 1H), 7.10 (t, *J* = 7.5 Hz, 1H), 4.84 (s, 2H), 3.98 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 161.0, 157.9, 154.4, 150.1, 136.9, 113.1, 112.7, 59.0, 56.5. HRMS (ESI⁺) calcd. for C₁₁H₁₀O₃Na [M+Na]⁺ 213.0528 *m/z*, found 213.0530.



2-(Hydroxymethyl)-3-(4-(methylthio)phenyl)cycloprop-2-en-1-one (**4.1k**)

General procedure C was used with the following reagents: **S-4.25** (0.239 g, 0.912 mmol, 1.00 equiv.), NaI (0.308 g, 2.05 mmol, 2.25 equiv.), TMSCF₃ (0.27 mL, 1.8 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (72.6 mg), anhydrous MeOH (2.0 mL). After filtration, the solution was concentrated *in vacuo* to give compound **4.1k** (0.15 g, 79% over two steps) as a light beige solid.

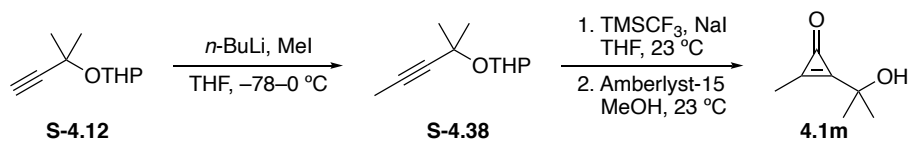
¹H NMR (500 MHz, CD₃OD) δ 7.91 (d, *J* = 8.4 Hz, 2H), 7.42 (d, *J* = 8.4 Hz, 2H), 4.82 (s, 2H), 2.55 (s, 3H). ¹³C NMR (151 MHz, CD₃OD) δ 157.4, 154.3, 153.5, 148.4, 133.7, 126.6, 119.7, 58.3, 14.5. HRMS (ESI⁺) calcd. for C₁₁H₁₀O₂SNa [M+Na]⁺ 229.0299 *m/z*, found 229.0291.



2-(Hydroxymethyl)-3-(thiophen-2-yl)cycloprop-2-en-1-one (4.11)

General procedure C was used with the following reagents: **S-4.32** (0.132 g, 0.601 mmol, 1.00 equiv.), NaI (0.221 g, 1.47 mmol, 2.45 equiv.), TMSCF₃ (0.19 mL, 1.3 mmol, 2.2 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (80.9 mg), anhydrous MeOH (2.0 mL). The crude residue was purified by flash column chromatography (eluting with 25–50% EtOAc in CH₂Cl₂) to give **4.11** (61 mg, 61% over two steps) as a brown solid.

¹H NMR (400 MHz, CD₃OD) δ 8.00 (dd, *J* = 5.1, 1.1 Hz, 1H), 7.88 (dd, *J* = 3.7, 1.1 Hz, 1H), 7.30 (dd, *J* = 5.1, 3.8 Hz, 1H), 4.77 (s, 2H). ¹³C NMR (151 MHz, CD₃OD) δ 154.5, 149.1, 146.9, 138.1, 137.0, 129.9, 124.9, 58.2. HRMS (ESI⁺) calcd. for C₈H₆O₂SNa [M+Na]⁺ 188.9986 *m/z*, found 188.9989.



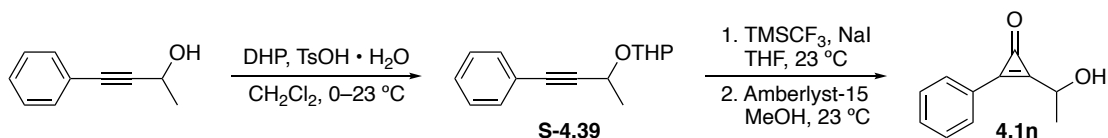
2-(2-Hydroxypropan-2-yl)-3-methylcycloprop-2-en-1-one (4.1m)

To a flame-dried round-bottom flask was added a solution of **S-4.12** (1.00 g, 5.96 mmol, 1.00 equiv.) in anhydrous THF (12 mL). The solution was cooled to –78 °C, and *n*-butyllithium was added dropwise (2.5 M in hexanes, 2.6 mL, 6.5 mmol, 1.1 equiv.). The solution was stirred at –78 °C for 30 min, then methyl iodide was added dropwise (0.45

mL, 7.2 mmol, 1.2 equiv.). The solution was slowly warmed to ambient temperature and stirred overnight. The solution was quenched with sat. NH_4Cl (90 mL), then extracted with Et_2O (2 x 50 mL). The combined organic layers were dried with MgSO_4 , filtered, and concentrated *in vacuo*. The crude oil **S-4.38** was used in the next step without further purification.

General procedure C was employed using the following amounts of reagents: **S-4.38** (1.05 g, 5.76 mmol, 1.00 equiv.), NaI (1.91 g, 12.7 mmol, 2.21 equiv.), TMSCF_3 (1.7 mL, 12 mmol, 2.0 equiv.), anhydrous THF (12 mL); Amberlyst-15 (0.385 g), anhydrous MeOH (10 mL). The crude residue was purified by flash column chromatography (eluting with 50% EtOAc in CH_2Cl_2) to give compound **4.1m** (0.39 g, 52% over three steps) as a yellow oil.

^1H NMR (600 MHz, CD_3OD) δ 2.33 (s, 3H), 1.50 (s, 6H). ^{13}C NMR (151 MHz, CD_3OD) δ 166.0, 160.2, 156.3, 71.2, 28.1, 10.1. HRMS (ESI⁺) calcd. for $\text{C}_7\text{H}_{10}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 149.0578 *m/z*, found 149.0577.



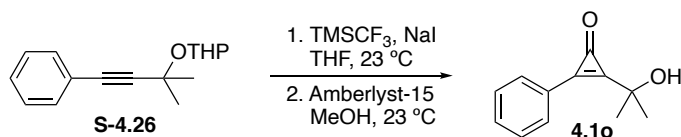
2-(1-Hydroxyethyl)-3-phenylcycloprop-2-en-1-one (**4.1n**)

To a flame-dried round-bottom flask was added $\text{TsOH} \cdot \text{H}_2\text{O}$ (3.0 mg, 0.016 mmol, 0.6 mol %) and a solution of 4-phenyl-3-butyn-2-ol (0.369 g, 2.52 mmol, 1.00 equiv.) in anhydrous CH_2Cl_2 (5.0 mL). After cooling to 0 °C, DHP (0.24 mL, 2.8 mmol, 1.1 equiv.)

was added dropwise, and the solution was slowly warmed to ambient temperature and stirred overnight. The resulting dark solution was diluted with CH₂Cl₂ (30 mL) and washed with sat. NaHCO₃ (1 x 50 mL). The aqueous layer was extracted with CH₂Cl₂ (2 x 20 mL), and the organic layers were combined, dried with MgSO₄, filtered, and concentrated *in vacuo*. The crude residue was purified by flash column chromatography (eluting with 5–10% EtOAc in hexanes) to give compound **S-4.39** (0.56 g, 97%) as a colorless oil. The mixture of diastereomers was used in the next step without further purification.

General procedure C was employed using the following amounts of reagents: **S-4.39** (0.460 g, 2.0 mmol, 1.00 equiv.), NaI (0.568 g, 3.78 mmol, 1.89 equiv.), TMSCF₃ (0.69 mL, 4.4 mmol, 2.2 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (0.100 g), anhydrous MeOH (5.0 mL). The crude residue was purified by flash column chromatography (eluting with 5–20% acetone in CH₂Cl₂) to give compound **4.1n** (0.21 g, 59% over two steps) as a tan solid.

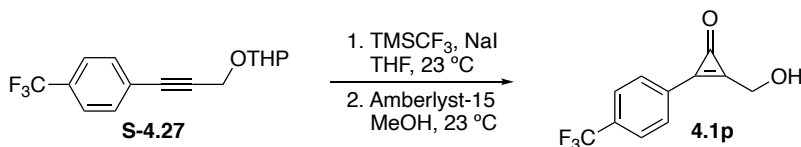
¹H NMR (600 MHz, CDCl₃) δ 7.97–7.94 (m, 2H), 7.58–7.49 (m, 3H), 5.14 (qd, *J* = 6.9, 4.9 Hz, 1H), 2.75 (d, *J* = 5.0 Hz, 1H), 1.64 (d, *J* = 7.0 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 156.9, 155.9, 153.1, 133.0, 132.5, 129.3, 123.0, 65.0, 21.4. HRMS (ESI⁺) calcd. for C₁₁H₁₀O₂Na [M+Na]⁺ 197.0578 *m/z*, found 197.0570.



2-(2-Hydroxypropan-2-yl)-3-phenylcycloprop-2-en-1-one (4.1o)

General procedure C was used with the following reagents: **S-4.26** (0.620 g, 2.54 mmol, 1.00 equiv.), NaI (0.853 g, 5.69 mmol, 2.24 equiv.), TMSCF₃ (0.75 mL, 5.1 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (0.173 g), anhydrous MeOH (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 25–50% EtOAc in CH₂Cl₂) to give compound **4.1o** (0.35 g, 74% over two steps) as a light brown solid.

¹H NMR (400 MHz, CD₃OD) δ 8.03–8.00 (m, 2H), 7.67–7.62 (m, 1H), 7.60–7.56 (m, 2H), 1.61 (s, 6H). ¹³C NMR (151 MHz, CD₃OD) δ 160.4, 157.9, 152.6, 134.4, 133.7, 130.5, 123.9, 71.3, 28.4. HRMS (ESI⁺) calcd. for C₁₂H₁₂O₂Na [M+Na]⁺ 211.0735 *m/z*, found 211.0728.

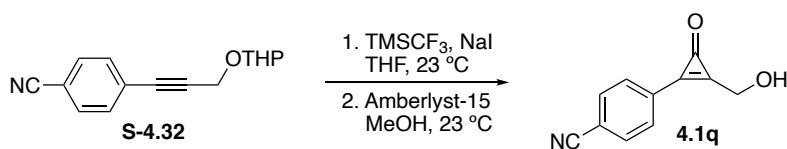


2-(Hydroxymethyl)-3-(4-(trifluoromethyl)phenyl)cycloprop-2-en-1-one (4.1p)

General procedure C was used with the following reagents: **S-4.27** (0.183 g, 0.644 mmol, 1.00 equiv.), NaI (0.229 g, 1.53 mmol, 2.37 equiv.), TMSCF₃ (0.19 mL, 1.9 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (78.6 mg), anhydrous MeOH (2.0 mL).

The crude residue was purified by flash column chromatography (eluting with 25–50% EtOAc in CH₂Cl₂) to give compound **4.1p** (45 mg, 31% over two steps) as an off-white solid.

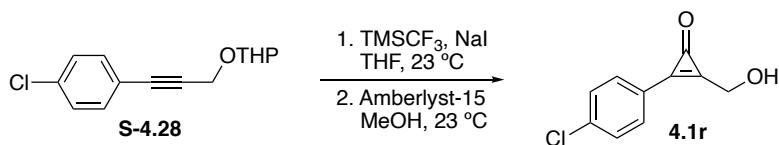
¹H NMR (400 MHz, CD₃OD) δ 8.20 (d, *J* = 8.0 Hz, 2H), 7.89 (d, *J* = 8.1 Hz, 2H), 4.90 (s, 2H). ¹⁹F NMR (376 MHz, CD₃OD) δ -65.0. ¹³C NMR (151 MHz, CD₃OD) δ 160.2, 157.2, 152.9, 135.0 (q, *J* = 32.7 Hz), 133.9, 127.3 (q, *J* = 3.8 Hz), 127.2 (d_{app}, *J* = 0.9 Hz), 125.1 (q, *J* = 271.9 Hz), 58.7. HRMS (ESI⁺) calcd. for C₁₁H₈F₃O₃ [M+H]⁺ 229.0476, found 229.0472.



4-(2-(Hydroxymethyl)-3-oxocycloprop-1-en-1-yl)benzonitrile (**4.1q**)

General procedure C was used with the following reagents: **S-4.32** (0.217 g, 0.901 mmol, 1.00 equiv.), NaI (0.303 g, 2.02 mmol, 2.24 equiv.), TMSCF₃ (0.27 mL, 1.8 mmol, 2.0 equiv.), anhydrous THF (4.0 mL); Amberlyst-15 (75.2 mg), anhydrous MeOH (2.0 mL). The crude residue was purified by flash column chromatography (eluting with 50–100% EtOAc in CH₂Cl₂) to give compound **4.1q** (45 mg, 27% over two steps) as a light tan solid.

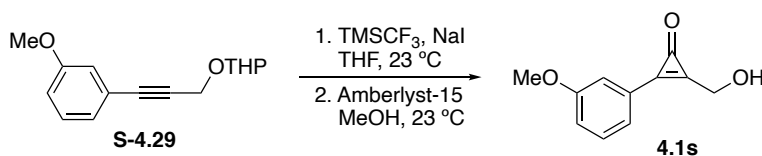
¹H NMR (400 MHz, (CD₃)₃SO) δ 8.07 (s, 4H), 6.02 (t, *J* = 5.4 Hz, 1H), 4.83 (d, *J* = 5.4 Hz, 2H). ¹³C NMR (151 MHz, (CD₃)₃SO) δ 161.7, 153.5, 151.0, 133.4, 132.0, 126.3, 118.1, 114.4, 57.8. HRMS (ESI⁺) calcd. for C₁₁H₇NO₂Na [M+Na]⁺ 208.0374 *m/z*, found 208.0374.



2-(4-Chlorophenyl)-3-(hydroxymethyl)cycloprop-2-en-1-one (4.1r)

General procedure C was used with the following reagents: **S-4.28** (0.194 g, 0.773 mmol, 1.00 equiv.), NaI (0.258 g, 1.72 mmol, 2.22 equiv.), TMSCF_3 (0.23 mL, 1.5 mmol, 2.0 equiv.), anhydrous THF (4.0 mL). The crude residue was passed through a plug of silica (eluting with 10–20% EtOAc in CH_2Cl_2) and used in the next step without further purification. Amberlyst-15 (77.0 mg), anhydrous MeOH (2.0 mL). After filtration, the solution was concentrated *in vacuo* to give compound **4.1r** (35 mg, 23% over two steps) as an off-white solid.

^1H NMR (600 MHz, CD_3OD) δ 8.01 (m, 2H), 7.61 (m, 2H), 4.85 (s, 2H). ^{13}C NMR (151 MHz, CD_3OD) δ 157.3, 157.2, 153.0, 140.5, 134.9, 130.8, 122.6, 58.5. HRMS (ESI⁺) calcd. for $\text{C}_{10}\text{H}_7\text{ClO}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 217.0032 m/z , found 217.0031.

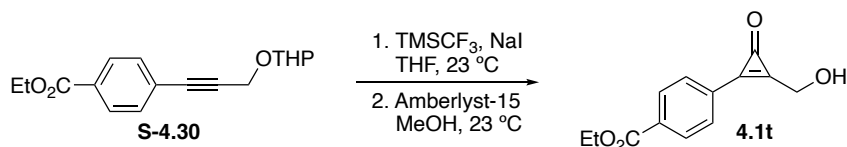


2-(Hydroxymethyl)-3-(3-methoxyphenyl)cycloprop-2-ene-1-one (4.1s)

General procedure C was used with the following reagents: **S-4.29** (0.176 g, 0.714 mmol, 1.00 equiv.), NaI (0.236 g, 1.57 mmol, 2.21 equiv.), TMSCF_3 (0.21 mL, 1.4 mmol, 2.0 equiv.), anhydrous THF (2.9 mL); Amberlyst-15 (60.1 mg), anhydrous MeOH (2.0 mL).

After filtration, the solution was concentrated to give compound **4.1s** (83 mg, 66% over two steps) as a dark yellow solid.

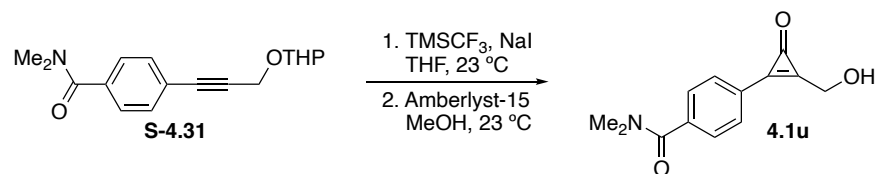
^1H NMR (600 MHz, CDCl_3) δ 7.53 (d, $J = 7.4$ Hz, 1H) 7.47 (s, 1H) 7.39 (t, $J = 7.7$ Hz, 1H), 7.09 (d, $J = 8.3$ Hz, 1H), 4.91 (s, 2H), 3.83 (s, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 160.0, 155.7, 154.7, 153.7, 124.8, 123.7, 119.7, 116.6, 58.73, 55.6. HRMS (ESI $^+$) calcd. for $\text{C}_{11}\text{H}_{10}\text{O}_3\text{Na}$ [M+Na] $^+$ 213.0528 m/z , found 213.0529.



Ethyl 4-(2-(hydroxymethyl)-3-oxocycloprop-1-en-1-yl)benzoate (**4.1t**)

General procedure C was used with the following amounts of reagents and modifications: **S-4.30** (0.224 g, 0.777 mmol, 1.00 equiv.), NaI (0.257 g, 1.71 mmol, 2.20 equiv.), TMSCF_3 (0.23 mL, 1.6 mmol, 2.0 equiv.), anhydrous THF (3.0 mL). The reaction was sealed and stirred at 85 °C overnight. Amberlyst-15 (75.0 mg), anhydrous MeOH (4.0 mL). The crude residue was purified by flash column chromatography (eluting with 0–5% acetone in EtOAc) to give compound **4.1t** (0.11 g, 46% over two steps) as a light brown solid.

^1H NMR (400 MHz, CD_3OD) δ 8.20 (m, 2H), 8.11 (m, 2H), 4.89 (s, 2H), 4.41 (q, $J = 7.1$ Hz, 2H), 1.41 (t, $J = 7.1$ Hz, 3H). ^{13}C NMR (151 MHz, CD_3OD) δ 166.8, 159.6, 157.4, 153.2, 135.3, 133.3, 131.2, 127.5, 62.7, 58.7, 14.5. HRMS (ESI $^+$) calcd. for $\text{C}_{13}\text{H}_{12}\text{O}_4\text{Na}$ [M+Na] $^+$ 255.0633 m/z , found 255.0632.



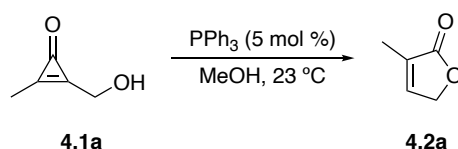
***N,N*-Diethyl-4-(2-(hydroxymethyl)-3-oxocycloprop-1-en-1-yl)benzamide (4.1u)**

General procedure C was used with the following reagents: **S-4.31** (0.564 g, 1.79 mmol, 1.00 equiv.), NaI (0.508 g, 3.39 mmol, 1.89 equiv.), TMSCF₃ (0.61 mL, 3.9 mmol, 2.2 equiv.), anhydrous THF (3.6 mL); Amberlyst-15 (0.200 g), anhydrous MeOH (3.6 mL). The crude residue was purified by flash column chromatography (eluting with 10% MeOH in CH₂Cl₂) to give **4.1u** (0.18 g, 38% over two steps) as a clear orange oil.

¹H NMR (600 MHz, CDCl₃) δ 7.86 (d, *J* = 8.1 Hz, 2H), 7.32 (d, *J* = 8.1 Hz, 2H), 5.41 (t, *J* = 5.8 Hz, 1H), 4.74 (d, *J* = 5.8 Hz, 2H), 3.52 (q, *J* = 7.0 Hz, 2H), 3.16 (q, *J* = 7.0 Hz, 2H), 1.23 (t, *J* = 7.0 Hz, 3H), 1.06 (t, *J* = 7.0 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 170.4, 157.3, 155.3, 152.7, 140.5, 132.6, 126.7, 123.5, 58.4, 43.5, 39.7, 14.1, 12.8. HRMS (ESI⁺) calculated for C₁₅H₁₇NO₃Na [M+Na]⁺ 282.1106 *m/z*, found 282.1107.

4.4.4d General procedure D for substituted butenolides

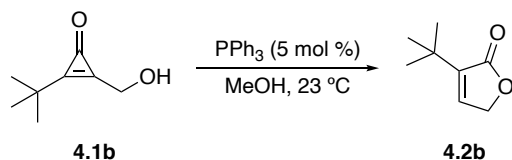
To a scintillation vial containing substituted 2-hydroxymethyl-cyclopropenone (1.0 equiv.) was added anhydrous MeOH (0.25 M) or anhydrous C₆H₆ (0.25 M), followed by triphenylphosphine (PPh₃, 5 mol %, unless otherwise stated). The reaction was stirred at ambient temperature, and monitored until full consumption of starting material was observed (as determined by TLC, 1:1 CH₂Cl₂:EtOAc). The mixture was concentrated *in vacuo*, then purified by flash column chromatography to afford pure product.



3-Methylfuran-2(5H)-one (4.2a)

General procedure D was employed using the following reagents: **4.1a** (48 mg, 0.49 mmol, 1.0 equiv.), triphenylphosphine (PPh₃, 6.2 mg, 0.024 mmol, 5 mol %), anhydrous MeOH (2.0 mL). The crude residue was purified by flash column chromatography (eluting with 50% Et₂O in *n*-pentane) to give compound **4.2a** (20 mg, 62%) as a clear oil.

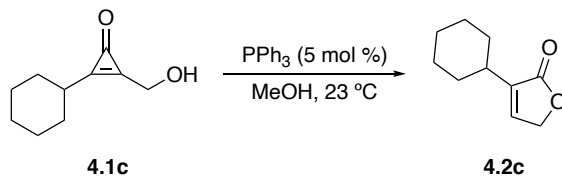
¹H NMR (600 MHz, CDCl₃) δ 7.13 (s, 1H), 4.73 (s, 2H), 1.90 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 174.9, 145.1, 130.0, 70.1, 10.8. HRMS (ESI⁺) calcd. for C₅H₆O₂Na [M+Na]⁺ 121.0266 *m/z*, found 121.0266.



3-(*tert*-Butyl)furan-2(5*H*)-one (4.2b)

General procedure D was used with the following reagents: **4.1b** (0.177 g, 1.26 mmol, 1.00 equiv.), triphenylphosphine (16.5 mg, 0.0629 mmol, 5 mol %), anhydrous MeOH (5.0 mL). The crude residue was purified by flash column chromatography (eluting with 10% EtOAc in hexanes) to give compound **4.2b** (0.12 g, 69%) as a clear oil.

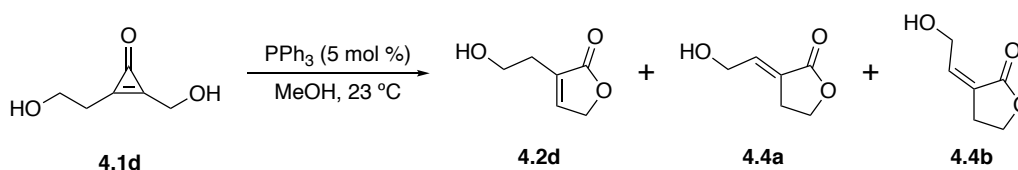
$^1\text{H NMR}$ (600 MHz, CDCl_3) δ 7.05 (s, 1H), 4.71 (s, 2H), 1.26 (s, 9H). $^{13}\text{C NMR}$ (151 MHz, CDCl_3) δ 172.7, 142.6, 142.5, 69.3, 31.7, 28.2. HRMS (ESI⁺) calcd. for $\text{C}_{12}\text{H}_{12}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 163.0735 m/z , found 163.0731.



3-Cyclohexylfuran-2(5*H*)-one (4.2c)

General procedure D was used with the following reagents: **4.1c** (0.105 g, 0.633 mmol, 1.00 equiv.), PPh_3 (8.7 mg, 0.033 mmol, 5 mol %), anhydrous MeOH (2.5 mL). The crude residue was purified by flash column chromatography (eluting with 20% Et_2O in hexanes) to give compound **4.2c** (95 mg, 91%) as a pale yellow oil.

^1H NMR (400 MHz, CDCl_3) δ 7.03 (q_{app}, J = 1.6 Hz, 1H), 4.75 (t, J = 1.7 Hz, 2H), 2.35 (toct_{app}, J = 11.6, 1.5 Hz, 1H), 1.97–1.92 (m, 2H), 1.81–1.68 (m, 3H), 1.43–1.30 (m, 2H), 1.26–1.15 (m, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 174.1, 142.7, 139.5, 70.2, 35.0, 31.6, 26.2, 26.1. HRMS (ESI⁺) calcd. for $\text{C}_{10}\text{H}_{14}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 189.0892 m/z , found 189.0893.



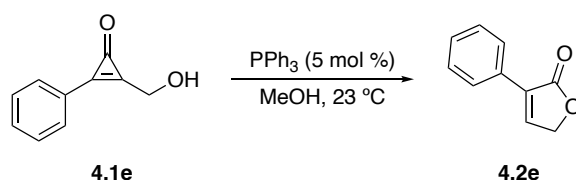
3-(2-Hydroxyethyl)furan-2(5H)-one (4.2d)

General procedure D was used with the following reagents: **4.1d** (0.102 g, 0.790 mmol, 1.00 equiv.), PPh_3 (42.0 mg, 0.160 mmol, 5 mol %), anhydrous MeOH (3.2 mL). The crude residue was purified by flash column chromatography (eluting with 20% EtOAc in hexanes) to give compound **4.2d** (50 mg, 52%) as a clear oil.

^1H NMR (600 MHz, CDCl_3) δ 7.30 (quint, J = 1.6 Hz, 1H) 4.83 (q, J = 1.8 Hz, 2H), 3.85 (q, J = 5.8 Hz, 2H), 2.60 (tq, J = 5.8, 1.6 Hz, 2H), 2.08 (t, J = 5.8 Hz, 1H). ^{13}C NMR (151 MHz, CDCl_3) δ 175.0, 146.8, 131.8, 70.7, 60.6, 29.1. HRMS (ESI⁺) calcd. for $\text{C}_6\text{H}_8\text{O}_3\text{Na}$ $[\text{M}+\text{Na}]^+$ 151.0371 m/z , found 151.0370.

Compound **4.4a** (11 mg), was isolated as a colorless oil [68]. Further attempts at purification led to decomposition. ^1H NMR (400 MHz, CDCl_3) δ 6.80 (tt_{app}, J = 5.4, 2.9 Hz, 1H) 4.40–4.36 (m, 4H), 2.98 (dddd, J = 9.8, 5.0, 4.3, 2.9 Hz, 2H). ^{13}C NMR (151 MHz, CDCl_3) δ 171.4, 138.0, 126.1, 65.8, 61.0, 25.5. HRMS (ESI⁺) calcd. for $\text{C}_6\text{H}_8\text{O}_3\text{Na}$ $[\text{M}+\text{Na}]^+$ 151.0371 m/z , found 151.0367.

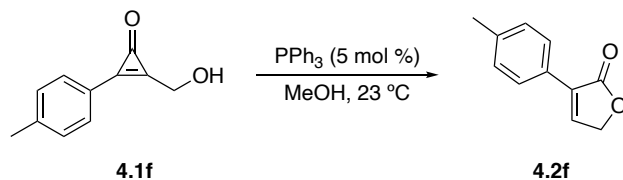
Compound **4.4b** (8.0 mg), was isolated as a colorless oil [68]. Further attempts at purification led to decomposition. ^1H NMR (400 MHz, CDCl_3) δ 6.46 (tt_{app}, J = 5.7, 2.4 Hz, 1H) 4.60 (ddd, J = 4.8, 3.2, 2.3 Hz, 2H), 4.42 (t, J = 7.4 Hz, 2H), 2.99 (dddd, J = 9.7, 5.1, 4.6, 2.3 Hz, 2H). ^{13}C NMR (151 MHz, CDCl_3) δ 171.4, 142.4, 126.3, 66.5, 59.1, 28.9. HRMS (ESI⁺) calcd. for $\text{C}_6\text{H}_8\text{O}_3\text{Na}$ $[\text{M}+\text{Na}]^+$ 151.0371 m/z , found 151.0368.



3-Phenylfuran-2(5H)-one (4.2e)

General procedure D was used with the following reagents: **4.1e** (15.5 mg, 0.0966 mmol, 1.00 equiv.), PPh_3 (1.3 mg, 0.0050 mmol, 5 mol %), anhydrous MeOH (0.4 mL). The crude residue was purified by flash column chromatography (eluting with 20% Et_2O in hexanes) to give compound **4.2e** (10 mg, 68%) as a pale yellow solid.

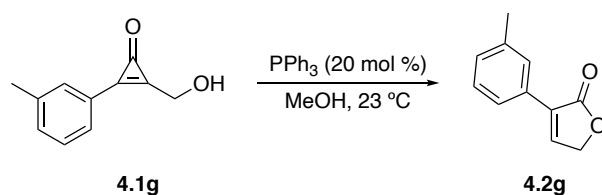
^1H NMR (500 MHz, CDCl_3) δ 7.87–7.83 (m, 2H), 7.64 (t, J = 2.0 Hz, 1H), 7.44–7.36 (m, 3H), 4.92 (d, J = 2.0 Hz, 2H). ^{13}C NMR (151 MHz, CDCl_3) δ 172.3, 144.4, 131.9, 129.7, 129.5, 128.8, 127.1, 69.7. HRMS (ESI⁺) calcd. for $\text{C}_{10}\text{H}_8\text{O}_2$ $[\text{M}+\text{Na}]^+$ 183.0422 m/z , found 183.0421.



3-(*p*-Tolyl)furan-2(5*H*)-one (4.2f)

General procedure D was used with the following reagents: **4.1f** (79.6 mg, 0.457 mmol, 1.00 equiv.), PPh₃ (12.0 mg, 0.0458 mmol, 10 mol %), anhydrous MeOH (1.8 mL). The crude residue was purified by flash column chromatography (eluting with 20% Et₂O in hexanes) to give compound **4.2f** (65 mg, 82%) as a white solid.

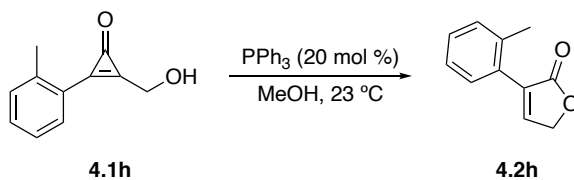
¹H NMR (400 MHz, CDCl₃) δ 7.75 (d, *J* = 8.4 Hz, 2H), 7.58 (t, *J* = 2.0 Hz, 1H), 7.22 (d, *J* = 8.0 Hz, 2H), 4.90 (d, *J* = 2.0 Hz, 2H), 2.38 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 172.5, 143.4, 139.6, 131.6, 129.5, 126.9, 126.8, 69.6, 21.5. HRMS (ESI⁺) calcd. for [M+Na]⁺ 197.0578 *m/z*, found 197.0585.



3-(*m*-Tolyl)furan-2(5*H*)-one (4.2g)

General procedure D was used with the following reagents: **4.1g** (86.8 mg, 0.498 mmol, 1.00 equiv.), PPh₃ (26.3 mg, 0.100 mmol, 20 mol %), anhydrous MeOH (2.0 mL). The crude residue was purified by flash column chromatography (eluting with 20% Et₂O in hexanes) to give compound **4.2g** (61 mg, 70%) as an off-white solid.

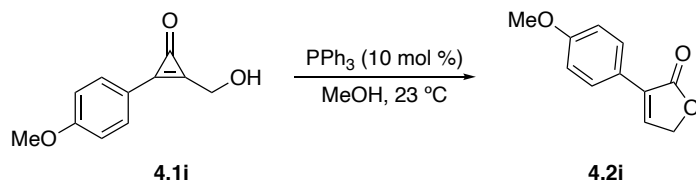
^1H NMR (400 MHz, CDCl_3) δ 7.67 (br s, 1H), 7.64–7.61 (m, 2H), 7.31 (t, $J = 7.7$ Hz, 1H), 7.20 (d, $J = 7.6$ Hz, 1H), 4.92 (d, $J = 2.0$ Hz, 2H), 2.39 (s, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 172.4, 144.2, 138.5, 132.0, 130.3, 129.6, 128.7, 127.7, 124.2, 69.6, 21.6. HRMS (ESI $^+$) calcd. for $\text{C}_{11}\text{H}_{10}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 197.0578 m/z , found 197.0572.



3-(*o*-Tolyl)furan-2(5*H*)-one (**4.2h**)

General procedure D was used with the following reagents: **4.1h** (86.6 mg, 0.497 mmol, 1.00 equiv.), PPh_3 (26 mg, 0.099 mmol, 20 mol %), anhydrous MeOH (2.0 mL). The crude residue was purified by flash column chromatography (eluting with 20% Et_2O in hexanes) to give compound **4.2h** (73 mg, 84%) as a clear oil.

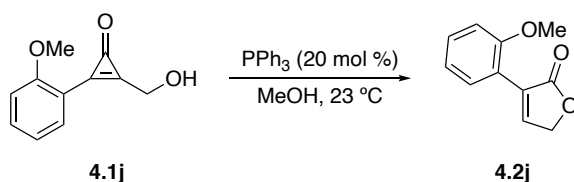
^1H NMR (400 MHz, CDCl_3) δ 7.45 (t, $J = 1.8$ Hz, 1H), 7.35 (d, $J = 7.7$ Hz, 1H), 7.32–7.27 (m, 2H), 7.25–7.22 (m, 1H), 4.98 (d, $J = 1.8$ Hz, 2H), 2.35 (s, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 172.7, 147.8, 136.7, 133.6, 130.7, 129.64, 129.60, 129.2, 126.0, 70.1, 20.5. HRMS (ESI $^+$) calcd. for $\text{C}_{11}\text{H}_{10}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 197.0578 m/z , found 197.0570.



3-(4-Methoxyphenyl)furan-2(5H)-one (4.2i)

Using Procedure D. The following amounts of reagents were used: **4.1i** (54.6 mg, 0.287 mmol, 1.00 equiv.), PPh₃ (7.4 mg, 0.028 mmol, 10 mol %), anhydrous MeOH (1.2 mL). The crude residue was purified by flash column chromatography (eluting with 20% Et₂O in hexanes) to give compound **4.2i** (47 mg, 85%) as a pale-yellow solid.

¹H NMR (400 MHz, CDCl₃) δ 7.83 (d, *J* = 8.9 Hz, 2H), 7.52 (t, *J* = 2.0 Hz, 1H), 6.94 (d, *J* = 8.9 Hz, 2H), 4.90 (d, *J* = 2.0 Hz, 2H), 3.84 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 172.6, 160.6, 142.0, 131.2, 128.5, 122.3, 114.2, 69.6, 55.5. HRMS (ESI⁺) calcd. for C₁₁H₁₀O₃Na [M+Na]⁺ 213.0528 *m/z*, found 213.0520.

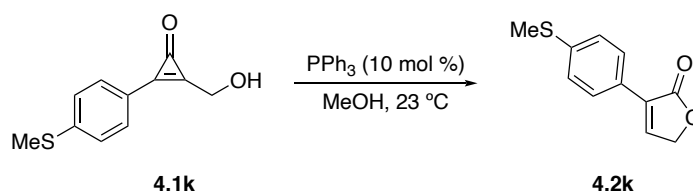


3-(2-Methoxyphenyl)furan-2(5H)-one (4.2j)

General procedure D was used with the following reagents: **4.1j** (20.2 mg, 0.106 mmol, 1.00 equiv.), PPh₃ (5.6 mg, 0.021 mmol, 20 mol %), anhydrous MeOH (0.42 mL). The crude residue was purified by flash column chromatography using deactivated silica

(deactivated with 1% Et₃N in CH₂Cl₂, then eluting with 15% EtOAc in hexanes) to give compound **4.2j** (8.6 mg, 43%) as a white solid.

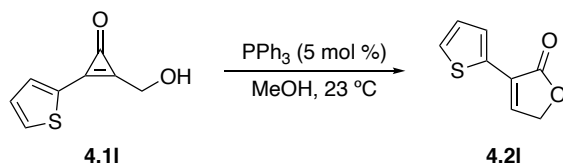
¹H NMR (600 MHz, CDCl₃) δ 8.15 (d, *J* = 7.7 Hz, 1H), 7.98 (s, 1H), 7.35 (t, *J* = 7.8 Hz, 1H), 7.04 (t, *J* = 7.6 Hz, 1H), 6.97 (d, *J* = 8.3 Hz, 1H), 4.93 (d, *J* = 1.4 Hz, 2H), 3.89 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 173.2, 157.8, 148.4, 130.3, 130.8, 129.6, 126.9, 120.8, 118.7, 110.9, 69.8, 55.6. HRMS (ESI⁺) calcd. for C₁₁H₁₀O₃Na [M+Na]⁺ 213.0528 *m/z*, found 213.0527.



3-(4-(Methylthio)phenyl)furan-2(5H)-one (**4.2k**)

General procedure D was used with the following reagents: **4.1k** (0.142 g, 0.689 mmol, 1.00 equiv.), PPh₃ (18.2 mg, 0.0694 mmol, 10 mol %), anhydrous MeOH (2.8 mL). The crude residue was purified by flash column chromatography (eluting with 20% Et₂O in hexanes) to give compound **4.2k** (0.12 g, 82%) as a yellow solid.

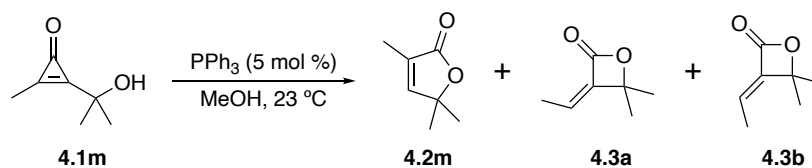
¹H NMR (400 MHz, CDCl₃) δ 7.80 (d, *J* = 8.5 Hz, 2H), 7.60 (t, *J* = 2.0 Hz, 1H), 7.27 (d, *J* = 8.6 Hz, 2H), 4.92 (d, *J* = 2.0 Hz, 2H), 2.50 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 172.3, 143.3, 140.6, 131.2, 127.4, 126.2, 126.16, 69.6, 15.5. HRMS (ESI⁺) calcd. for C₁₁H₁₀O₂SNa [M+Na]⁺ 229.0299 *m/z*, found 229.0305.



3-(Thiophen-2-yl)furan-2(5H)-one (4.2I)

General procedure D was used with the following reagents: **4.1I** (58.5 mg, 0.352 mmol, 1.00 equiv.), PPh₃ (5.0 mg, 0.019 mmol, 5 mol %), anhydrous MeOH (2.8 mL). The crude residue was purified by flash column chromatography (eluting with 20% Et₂O in hexanes) to give compound **4.2I** (50 mg, 85%) as an off-white solid.

¹H NMR (400 MHz, CDCl₃) δ 7.77 (d, *J* = 3.4 Hz, 1H), 7.47 (t, *J* = 2.1 Hz, 1H), 7.38 (d, *J* = 5.0 Hz, 1H), 7.09 (dd, *J* = 5.0, 3.7 Hz, 1H), 4.94 (d, *J* = 2.1 Hz, 2H). ¹³C NMR (151 MHz, CDCl₃) δ 171.5, 140.5, 131.6, 127.9, 127.5, 127.3, 126.7, 70.1. HRMS (ESI⁺) calcd. for C₈H₆O₂SNa [M+Na]⁺ 188.9986 *m/z*, found 188.9989.



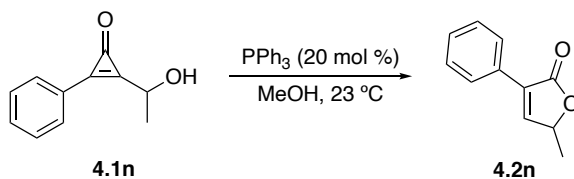
3,5,5-Trimethylfuran-2(5H)-one (4.2m)

General procedure D was used with the following reagents: **4.1m** (0.101 g, 0.800 mmol, 1.00 equiv.), PPh₃ (10.6 mg, 0.0404 mmol, 5 mol %), anhydrous MeOH (3.2 mL). The crude residue was purified by flash column chromatography (eluting with 3–4 Et₂O in hexanes) to give compound **4.2m** (47 mg, 46%) as a white solid.

^1H NMR (400 MHz, CDCl_3) δ 6.98–6.97 (m, 1H), 1.88 (d, $J = 1.4$ Hz, 3H), 1.43 (s, 6H). ^{13}C NMR (151 MHz, CDCl_3) δ 173.8, 154.1, 128.3, 84.3, 25.8, 10.6. HRMS (Cl^+) calcd. for $\text{C}_7\text{H}_{10}\text{O}_2$ $[\text{M}]^+$ 126.0681 m/z , found 126.0675.

Compound **4.3a** (5.4 mg, 5%) was isolated as a volatile, colorless oil. [69] ^1H NMR (400 MHz, CDCl_3) δ 5.83 (q, $J = 7.2$ Hz, 1H), 2.03 (d, $J = 7.2$ Hz, 3H), 1.58 (s, 6H). ^{13}C NMR (151 MHz, CDCl_3) δ 164.1, 142.5, 129.2, 83.8, 25.6, 14.7. HRMS (Cl^+) calcd. for $\text{C}_7\text{H}_{11}\text{O}_2$ $[\text{M}+\text{H}]^+$ 127.0759 m/z , found 127.0760.

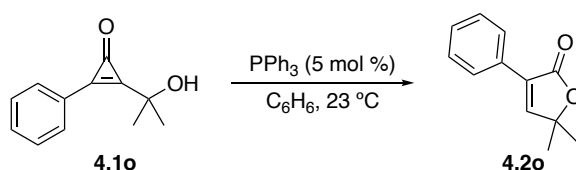
Compound **4.3b** (5.0 mg, 5%) was isolated as a volatile, colorless oil. [69] ^1H NMR (400 MHz, CDCl_3) δ 6.30 (q, $J = 7.3$ Hz, 1H), 1.80 (d, $J = 7.3$ Hz, 3H), 1.66 (s, 6H). ^{13}C NMR (151 MHz, CDCl_3) δ 163.8, 142.8, 127.2, 84.0, 25.0, 13.3. HRMS (Cl^+) calcd. for $\text{C}_7\text{H}_{11}\text{O}_2$ $[\text{M}+\text{H}]^+$ 127.0759 m/z , found 127.0761.



5-Methyl-3-phenylfuran-2(5H)-one (4.2n)

General procedure D was used with the following reagents: **4.1n** (7.0 mg, 0.039 mmol, 1.0 equiv.), PPh_3 (2.0 mg, 0.0076 mmol, 20 mol %), anhydrous MeOH (0.15 mL). The crude residue was purified by flash column chromatography (eluting with 20% EtOAc in hexanes) to give compound **4.2n** (4.8 mg, 72%) as a clear oil.

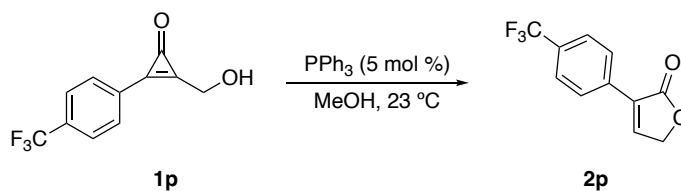
^1H NMR (600 MHz, CDCl_3) δ 7.85 (d, $J = 7.0$ Hz, 2H), 7.54 (d, $J = 1.4$ Hz, 1H) 7.43–7.37 (m, 3H), 5.15 (qd, $J = 6.8, 1.4$ Hz, 1H), 1.52 (d, $J = 6.8$ Hz, 3H). ^{13}C NMR (151 MHz, CDCl_3) δ 171.8, 149.1, 131.6, 129.7, 129.5, 128.8, 127.2, 76.8, 19.3. HRMS (ESI $^+$) calcd. for $\text{C}_{11}\text{H}_{10}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 197.0578 m/z , found 197.0579.



5,5-Dimethyl-3-phenylfuran-2(5H)-one (**4.2o**)

General procedure D was used with the following reagents and modifications: **4.1o** (50.2 mg, 0.267 mmol, 1.00 equiv.), PPh_3 (3.6 mg, 0.014 mmol, 5 mol %), anhydrous C_6H_6 (1.1 mL). The crude residue was purified by flash column chromatography (eluting with 5–10% Et_2O in hexanes) to give compound **4.2o** (28 mg, 55%) as a white solid.

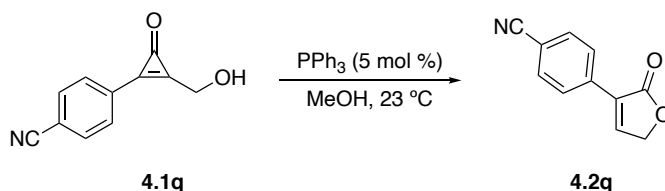
^1H NMR (600 MHz, CDCl_3) δ 7.85 (d, $J = 7.1$ Hz, 2H), 7.50 (s, 1H), 7.42–7.36 (m, 3H), 1.56 (s, 6H). ^{13}C NMR (151 MHz, CDCl_3) δ 171.2, 153.1, 130.2, 129.7, 129.4, 128.8, 127.2, 83.5, 25.9. HRMS (ESI $^+$) calcd. for $\text{C}_{12}\text{H}_{12}\text{O}_2\text{Na}$ $[\text{M}+\text{Na}]^+$ 211.0735 m/z , found 211.0736.



3-(4-(Trifluoromethyl)phenyl)furan-2(5H)-one (4.2p)

General procedure D was used with the following reagents: **4.1p** (44.8 mg, 0.196 mmol, 1.00 equiv.), PPh₃ (2.6 mg, 0.0099 mmol, 5 mol %), anhydrous MeOH (0.78 mL). The crude residue was purified by flash column chromatography (eluting with 50% EtOAc in hexanes) to give compound **4.2p** (28 mg, 61%) as a white solid.

¹H NMR (600 MHz, CDCl₃) δ 8.00 (d, *J* = 8.2 Hz, 2H), 7.78 (t, *J* = 1.9 Hz, 1H), 7.69 (d, *J* = 8.3 Hz, 2H), 4.99 (d, *J* = 1.9 Hz, 2H). ¹⁹F NMR (376 MHz, CDCl₃) δ -63.1. ¹³C NMR (151 MHz, CDCl₃) δ 171.8, 146.3, 133.0 (q, *J* = 1.2 Hz), 131.4 (q, *J* = 32.7 Hz), 130.9, 125.8 (q, *J* = 3.8 Hz), 124.0 (q, *J* = 272.2 Hz), 69.8. HRMS (ESI⁻) calcd. for C₁₁H₆F₃O₂ [M-H]⁻ 227.0320 *m/z*, found 227.0323.

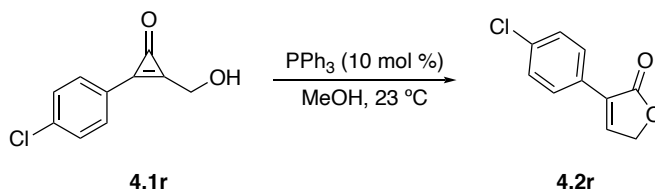


4-(2-Oxo-2,5-dihydrofuran-3-yl)benzonitrile (4.2q)

General procedure D was used with the following reagents: **4.1q** (26.8 mg, 0.145 mmol, 1.00 equiv.), PPh₃ (1.9 mg, 0.0072 mmol, 5 mol %), anhydrous MeOH (0.60 mL).

The crude residue was purified by flash column chromatography (eluting with 75% Et₂O in hexanes) to give compound **4.2q** (17 mg, 62%) as a pale yellow solid.

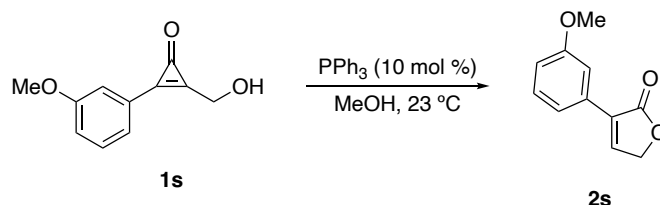
¹H NMR (400 MHz, CDCl₃) δ 8.00 (d, *J* = 8.2 Hz, 2H), 7.81 (s, 1H), 7.72 (d, *J* = 8.3 Hz, 2H), 5.00 (s, 2H). ¹³C NMR (151 MHz, CDCl₃) δ 171.4, 147.1, 133.8, 132.6, 130.5, 127.7, 118.5, 113.1, 69.8. HRMS (ESI⁻) calcd. for C₁₁H₆NO₂ [M-H]⁻ 184.0399 *m/z*, found 184.0397.



3-(4-Chlorophenyl)furan-2(5H)-one (**4.2r**)

General procedure D was used with the following reagents: **4.1r** (29.9 mg, 0.154 mmol, 1.00 equiv.), PPh₃ (4.0 mg, 0.015 mmol, 10 mol %), anhydrous MeOH (0.60 mL). The crude residue was purified by flash column chromatography (eluting with 50% Et₂O in CH₂Cl₂) to give compound **4.2r** (12 mg, 41%) as a white solid.

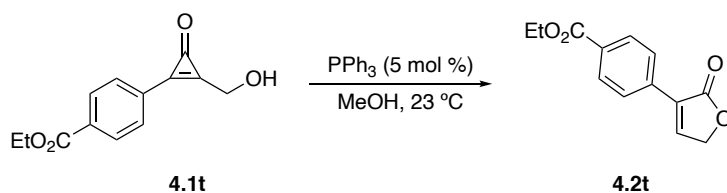
¹H NMR (400 MHz, CDCl₃) δ 7.87–7.79 (m, 2H), 7.66 (t, *J* = 2.0 Hz, 1H), 7.44–7.37 (m, 2H), 4.94 (d, *J* = 2.0 Hz, 2H). ¹³C NMR (151 MHz, CDCl₃) δ 172.1, 144.6, 135.6, 130.8, 129.1, 128.4, 128.1, 69.7. HRMS (ESI⁻) calcd. for C₁₀H₉O₂Cl [M-H]⁻ 193.0056 *m/z*, found 193.0064.



3-(3-Methoxyphenyl)furan-2(5H)-one (**4.2s**)

General procedure D was used with the following reagents: **4.1s** (83.0 mg, 0.436 mmol, 1.00 equiv.), PPh₃ (12.0 mg, 0.0458 mmol, 10 mol %), anhydrous MeOH (1.8 mL). The crude residue was purified by flash column chromatography (eluting with 50% EtOAc in hexanes) to give compound **4.2s** (30 mg, 36%) as a yellow oil.

¹H NMR (600 MHz, CDCl₃) δ 7.64 (t, *J* = 1.8 Hz, 1H) 7.44 (t, *J* = 1.8 Hz, 1H), 7.40 (d, *J* = 7.7 Hz, 1H), 7.32 (t, *J* = 8.0 Hz, 1H), 6.93 (dd, *J* = 8.2, 2.3 Hz, 1H), 4.90 (d, *J* = 1.9 Hz, 2H), 3.83 (s, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 172.2, 159.8, 144.8, 131.4, 130.8, 129.7, 119.4, 115.1, 112.4, 69.5, 55.4. HRMS (ESI⁺) calcd. for C₁₁H₁₀O₃Na [M+Na]⁺ 213.0528 *m/z*, found 213.0524.

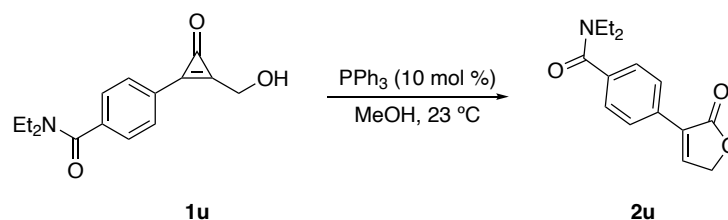


Ethyl 4-(2-oxo-2,5-dihydrofuran-3-yl)benzoate (**4.2t**)

General procedure D was used with the following reagents: **4.1t** (27.0 mg, 0.116 mmol, 1.00 equiv.), PPh₃ (1.5 mg, 0.0057 mmol, 5 mol %), anhydrous MeOH (0.46 mL).

The crude residue was purified by flash column chromatography (eluting with 20–50% Et₂O in hexanes) to give compound **4.2t** (14 mg, 52%) as a white solid.

¹H NMR (400 MHz, CDCl₃) δ 8.12–8.06 (m, 2H), 7.99–7.92 (m, 2H), 7.77 (t, *J* = 2.0 Hz, 1H), 4.97 (d, *J* = 2.0 Hz, 1H), 4.40 (q, *J* = 7.1 Hz, 2H), 1.41 (t, *J* = 7.1 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 171.9, 166.2, 146.2, 133.7, 131.2, 131.1, 130.0, 127.0, 69.8, 61.3, 14.5. HRMS (ESI⁺) calcd. for C₁₃H₁₂O₄Na [M+Na]⁺ 255.0633 *m/z*, found 255.0635.



***N,N*-Diethyl-4-(2-(hydroxymethyl)-3-oxocycloprop-1-en-1-yl)benzamide (4.2u)**

General procedure D was used with the following reagents: **4.1u** (0.110 g, 0.424 mmol, 1.00 equiv.), PPh₃ (11.0 mg, 0.0419 mmol, 10 mol %), anhydrous MeOH (1.7 mL). The crude residue was purified by flash column chromatography (eluting with 10% acetone in CH₂Cl₂) to give compound **4.2u** (50 mg, 45%) as a white solid.

¹H NMR (600 MHz, CDCl₃) δ 7.89 (d, *J* = 8.3 Hz, 2H), 7.69 (t, *J* = 1.9 Hz, 1H), 7.43 (d, *J* = 8.3 Hz, 2H), 4.96 (d, *J* = 1.9 Hz, 2H), 3.55 (bs, 2H), 3.26 (bs, 2H), 1.25 (bs, 3H), 1.12 (bs, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 172.1, 170.7, 145.3, 138.2, 132.2, 132.1, 131.1, 130.3, 128.64, 128.56, 127.1, 126.8, 69.7, 43.4, 39.4, 14.4, 13.0. HRMS (ESI⁺) calcd. for C₁₅H₁₇NO₃Na [M+Na]⁺ 282.1106 *m/z*, found 282.1107.

4.6 References

- [1]. Karuppiyah, V.; Sun, W.; Li, Z. *Stud. Nat. Prod. Chem.* **2016**, *48*, 417–446.
- [2]. Gan, L.-S.; Zheng, Y.-L.; Mo, J.-X.; Liu, X.; Li, X.-H.; Zhou, C.-X. *J. Nat. Prod.* **2009**, *72*, 1497–1501.
- [3]. Lu, A.; Wang, J.; Liu, T.; Han, J.; Li, Y.; Su, M.; Chen, J.; Zhang, H.; Wang, L.; Wang, Q. *J. Agric. Food Chem.* **2014**, *62*, 8799–8807.
- [4]. Hou, Y.; Shi, T.; Yang, Y.; Fan, X.; Chen, J.; Cao, F.; Wang, Z. *Org. Lett.* **2019**, *21*, 2952–2956.
- [5]. Rao, Y. S. *Chem. Rev.* **1964**, *64*, 353–388.
- [6]. Rao, Y. S. *Chem. Rev.* **1976**, *76*, 625–694.
- [7]. Knight, D. W. *Contemp. Org. Synth.* **1994**, *1*, 287–315.
- [8]. Carter, N. B.; Nadany, A. E.; Sweeney, J. B. *J. Chem. Soc., Perkin Trans. 1* **2002**, 2324–2342.
- [9]. Mao, B.; Fañanás-Mastral, M.; Feringa, B. L. *Chem. Rev.* **2017**, *117*, 10502–10566.
- [10]. Komatsu, K.; Kitagawa, T. *Chem. Rev.* **2003**, *103*, 1371–1427.
- [11]. Li, X.; Han, C.; Yao, H.; Lin, A. *Org. Lett.* **2017**, *19*, 778–781.
- [12]. Ren, J.-T.; Wang, J.-X.; Tian, H.; Xu, J.-L.; Hu, H.; Aslam, M.; Sun, M. *Org. Lett.* **2018**, *20*, 6636–6639.
- [13]. Xu, J.; Cao, J.; Fang, C.; Lu, T.; Du, D. *Org. Chem. Front.* **2017**, *4*, 560–564.
- [14]. Shih, H.-W.; Prescher, J. A. *J. Am. Chem. Soc.* **2015**, *137*, 10036–10039.
- [15]. Row, R. D.; Shih, H.-W.; Alexander, A. T.; Mehl, R. A.; Prescher, J. A. *J. Am. Chem. Soc.* **2017**, *139*, 7370–7375.
- [16]. Hamada, A.; Takizawa, T. *Tetrahedron Lett.* **1972**, *13*, 1849–1850.

- [17]. Cooke, M. P., Jr. *J. Org. Chem.* **1993**, *58*, 6833–6837.
- [18]. Cooke, M. P., Jr. *J. Org. Chem.* **1994**, *59*, 2930–2931.
- [19]. Alabugin, I. V.; Gilmore, K.; Manoharan, M. *J. Am. Chem. Soc.* **2011**, *133*, 12608–12623.
- [20]. Wei, Y.; Zhao, W.-T.; Yang, Y.-L.; Zhang, Z.; Shi, M. *ChemCatChem* **2015**, *7*, 3340–3349.
- [21]. Ziegler, C. B., Jr.; Heck, R. F. *J. Org. Chem.* **1978**, *43*, 2941–2946.
- [22]. Hartwig, J. F. *Angew. Chem. Int. Ed.* **1998**, *37*, 2046–2067.
- [23]. Wolfe, J. P.; Wagaw, S.; Marcoux, J.-F.; Buchwald, S. L. *Acc. Chem. Res.* **1998**, *31*, 805–818.
- [24]. Guo, H.; Fan, Y. C.; Sun, Z.; Wu, Y.; Kwon, O. *Chem. Rev.* **2018**, *118*, 10049–10293.
- [25]. Wang, F.; Luo, T.; Hu, J.; Wang, Y.; Krishnan, H. S.; Jog, P. V.; Ganesh, S. K.; Prakash, G. K. S.; Olah, G. A. *Angew. Chem. Int. Ed.* **2011**, *50*, 7153–7157.
- [26]. Zhou, J.; Campbell-Conroy, E. L.; Silina, A.; Uy, J.; Pierre, F.; Hurley, D. J.; Hilgraf, N.; Frieman, B. A.; DeNinno, M. P. *J. Org. Chem.* **2015**, *80*, 70–79.
- [27]. Nosik, P. S.; Ryabukhin, S. V.; Pashko, M. O.; Grabchuk, G. P.; Grygorenko, O. O.; Volochnyuk, D. M. *J. Fluorine Chem.* **2019**, *217*, 80–89.
- [28]. Hanessian, S.; Murray, P. J. *Tetrahedron* **1987**, *43*, 5055–5072.
- [29]. Boukouvalas, J.; Maltais, F.; Lachance, N. *Tetrahedron Lett.* **1994**, *35*, 7897–7900.
- [30]. Khoobi, M.; Alipour, M.; Zarei, S.; Jafarpour, F.; Shafiee, A. *Chem. Commun.* **2012**, *48*, 2985–2987.
- [31]. Almirante, N.; Cerri, A. *J. Org. Chem.* **1997**, *62*, 3402–3404.

- [32]. Mukai, C.; Moharram, S. M.; Azukizawa, S.; Hanaoka, M. *J. Org. Chem.* **1997**, *62*, 8095–8103.
- [33]. Feringa, B. L.; de Lange, B. *Tetrahedron* **1988**, *44*, 7213–7222.
- [34]. de Lange, B.; van Bolhuis, F.; Feringa, B. L. *Tetrahedron* **1989**, *45*, 6799–6818.
- [35]. Sweidan, A.; Chollet-Krugler, M.; van de Weghe, P.; Chokr, A.; Tomasi, S.; Bonnaure-Mallet, M.; Bousarghin, L. *Bioorg. Med. Chem.* **2016**, *24*, 5823–5833.
- [36]. Goldfogel, M. J.; Roberts, C. C.; Manan, R. S.; Meek, S. J. *Org. Lett.* **2017**, *19*, 90–93.
- [37]. Chu, C. K.; Beach, J. W.; Ullas, G. V.; Kosugi, Y. *Tetrahedron Lett.* **1988**, *29*, 5349–5352.
- [38]. Okabe, M.; Sun, R.-C.; Tam, S. Y.-K.; Todaro, L. J.; Coffen, D. L. *J. Org. Chem.* **1988**, *53*, 4780–4786.
- [39]. Beesley, R. M.; Ingold, C. K.; Thorpe, J. F. *J. Chem. Soc., Trans.* **1915**, *107*, 1080–1106.
- [40]. Cowell, A.; Stille, J. K. *J. Am. Chem. Soc.* **1980**, *102*, 4193–4198.
- [41]. Buchwald, S. L.; Fang, Q.; King, S. M. *Tetrahedron Lett.* **1988**, *29*, 3445–3448.
- [42]. DeShong, P.; Sidler, D. R.; Rybczynski, P. J.; Slough, G. A.; Rheingold, A. L. *J. Am. Chem. Soc.* **1988**, *110*, 2575–2585.
- [43]. Tiecco, M.; Testaferri, L.; Tingoli, M.; Bagnoli, L.; Santi, C. *Synlett* **1993**, 798–800.
- [44]. Yoneda, E.; Zhang, S.-W.; Zhou, D.-Y.; Onitsuka, K.; Takahashi, S. *J. Org. Chem.* **2003**, *68*, 8571–8576.
- [45]. Oh, C. H.; Park, S. J.; Ryu, J. H.; Gupta, A. K. *Tetrahedron Lett.* **2004**, *45*, 7039–7042.

- [46]. Kang, J.-E.; Lee, E.-S.; Park, S.-I.; Shin, S. *Tetrahedron Lett.* **2005**, *46*, 7431–7433.
- [47]. Takii, K.; Kanbayashi, N.; Onitsuka, K. *Chem. Commun.* **2012**, *48*, 3872–3874.
- [48]. Matsushita, K.; Komori, T.; Oi, S.; Inoue, Y. *Tetrahedron Lett.* **1994**, *35*, 5889–5890.
- [49]. Yu, W.-Y.; Alper, H. *J. Org. Chem.* **1997**, *62*, 5684–5687.
- [50]. APEX2 Version 2014.11-0, Bruker AXS, Inc.; Madison, WI 2014.
- [51]. SAINT Version 8.34a, Bruker AXS, Inc.; Madison, WI 2013.
- [52]. Sheldrick, G. M. SADABS, Version 2014/5, Bruker AXS, Inc.; Madison, WI 2014.
- [53]. Sheldrick, G. M. SHELXTL, Version 2014/7, Bruker AXS, Inc.; Madison, WI 2014.
- [54]. International Tables for Crystallography 1992, Vol. C., Dordrecht: Kluwer Academic Publishers.
- [55]. Corkum, E. G.; Hass, M. J.; Sullivan, A. D.; Bergens, S. H. *Org. Lett.* **2011**, *13*, 3522–3525.
- [56]. Duffey, M., O.; England, D., B.; Hu, Z.; Iyo, M.; Langston, S., P.; McIntyre, C.; Mizutani, H.; Xu, H. Heteroaryl Inhibitors of SUMO Activating Enzyme. Int. Patent Application WO2015002994A2, 2015.
- [57]. Crowley, J. D.; Leigh, D. A.; Lusby, P. J.; McBurney, R. T.; Perret-Aebi, L.-E.; Petzold, C.; Slawin, A. M. Z.; Symes, M. D. *J. Am. Chem. Soc.* **2007**, *129*, 15085–15090.
- [58]. Suzuki, H.; Mori, M.; Shibakami, M. *Synlett* **2003**, 2163–2166.
- [59]. Salomon, R. G.; Basu, B.; Roy, S.; Sachinvala, N. D. *J. Am. Chem. Soc.* **1991**, *113*, 3096–3106.
- [60]. Deagostino, A.; Prandi, C.; Toppino, A.; Venturello, P. *Tetrahedron* **2008**, *64*, 10344–10349.
- [61]. Kong, A.; Andreansky, E. S.; Blakey, S. B. *J. Org. Chem.* **2017**, *82*, 4477–4483.

- [62]. Ginn, J. D.; Sorcek, R. J.; Turner, M. R.; Young, E. R. R. Substituted 3-amino-thieno[2,3-b]pyridine-2-carboxylic acid amide compounds and processes for preparing and their uses. U.S. Patent 7592353, 2007.
- [63]. Thakur, K. G.; Sekar, G. *Synthesis* **2009**, 2785–2789.
- [64]. Wallinder, C.; Botros, M.; Rosenström, U.; Guimond, M.-O.; Beaudry, H.; Nyberg, F.; Gallo-Payet, N.; Hallberg, A.; Alterman, M. *Bioorg. Med. Chem.* **2008**, *16*, 6841–6849.
- [65]. Li, J.-H.; Li, J.-L.; Wang, D.-P.; Pi, S.-F.; Xie, Y.-X.; Zhang, M.-B.; Hu, X.-C. *J. Org. Chem.* **2007**, *72*, 2053–2057.
- [66]. Van den Hoven, B. G.; Alper, H. *J. Org. Chem.* **1999**, *64*, 9640–9645.
- [67]. Charpe, V. P.; Hande, A. A.; Sagadevan, A.; Hwang, K. C. *Green Chem.* **2018**, *20*, 4859–4864.
- [68]. Merritt, A. T.; Pouwer, R. H.; Williams, D. J.; Williams, C. M.; Ley, S. V. *Org. Biomol. Chem.* **2011**, *9*, 4745–4747.
- [69]. Danheiser, R. L.; Choi, Y. M.; Menichincheri, M.; Stoner, E. J. *J. Org. Chem.* **1993**, *58*, 322–327.

Appendix A: NMR Spectra for Chapters 2 and 3

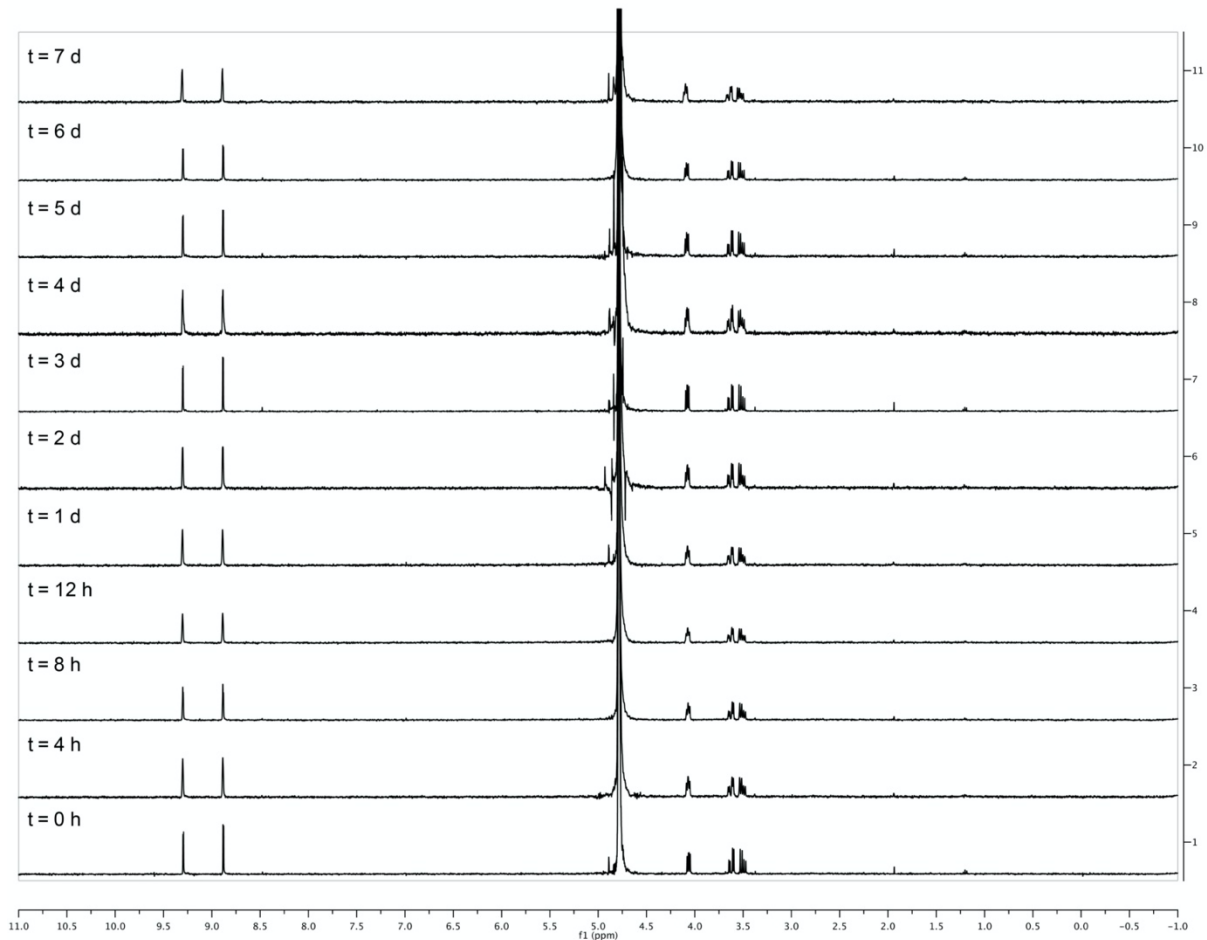
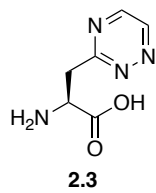


Figure S2-1. Triazine amino acid **2.3** is stable in aqueous solution over 7 d. Compound **2.3** (5 mM) was incubated in *d*-PBS (pH 7.4) at 37 °C and monitored via ¹H NMR spectroscopy.

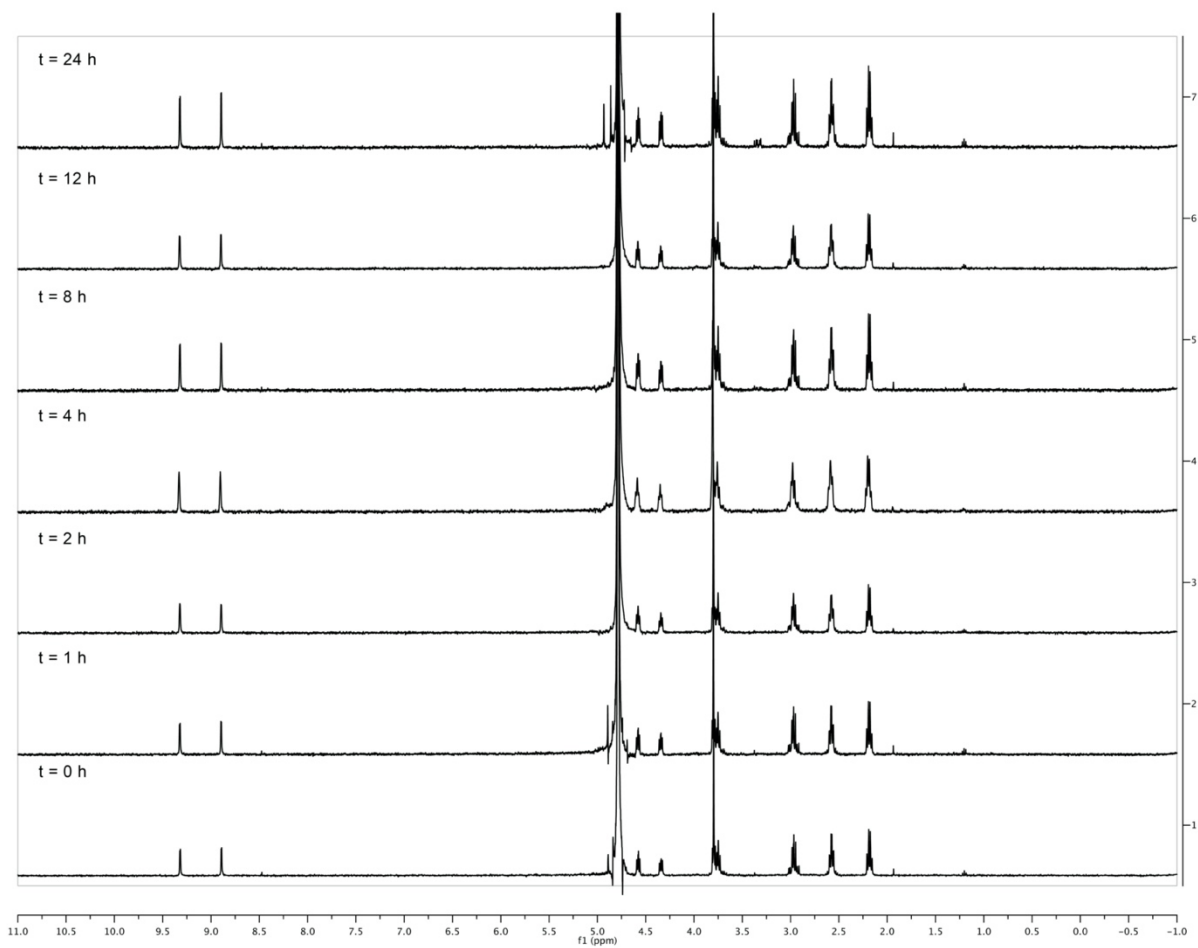
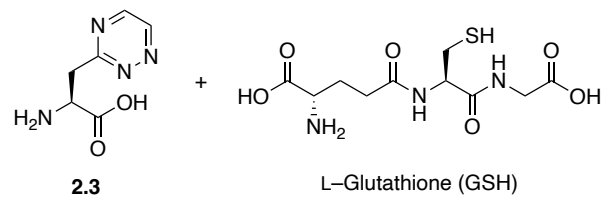


Figure S2-2. Triazine amino acid **2.3** is stable to L-glutathione over 24 h. Compound **2.3** (5 mM) and L-glutathione (5 mM) were incubated in *d*-PBS (pH 7.4) at 37 °C and monitored via ¹H NMR spectroscopy.

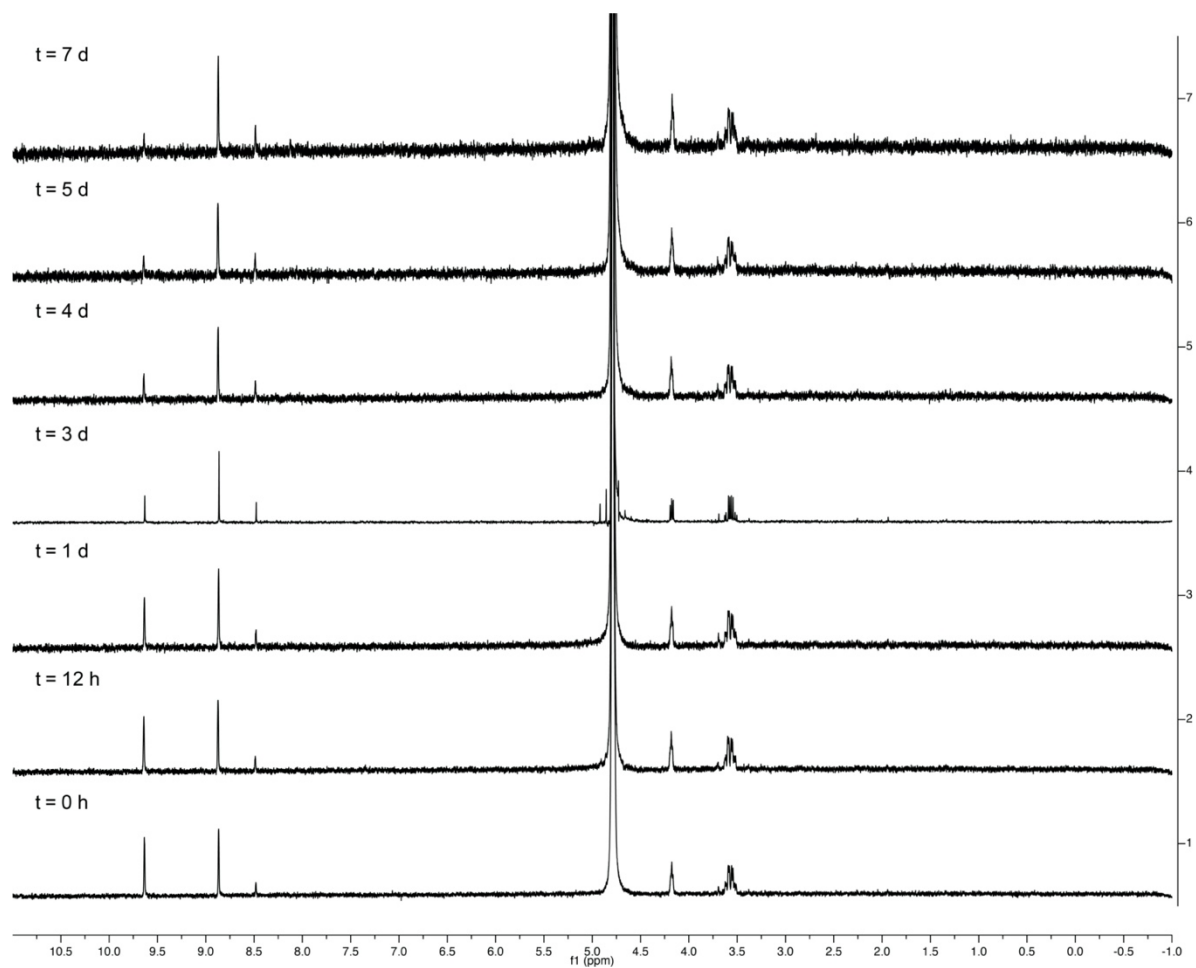
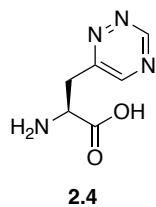
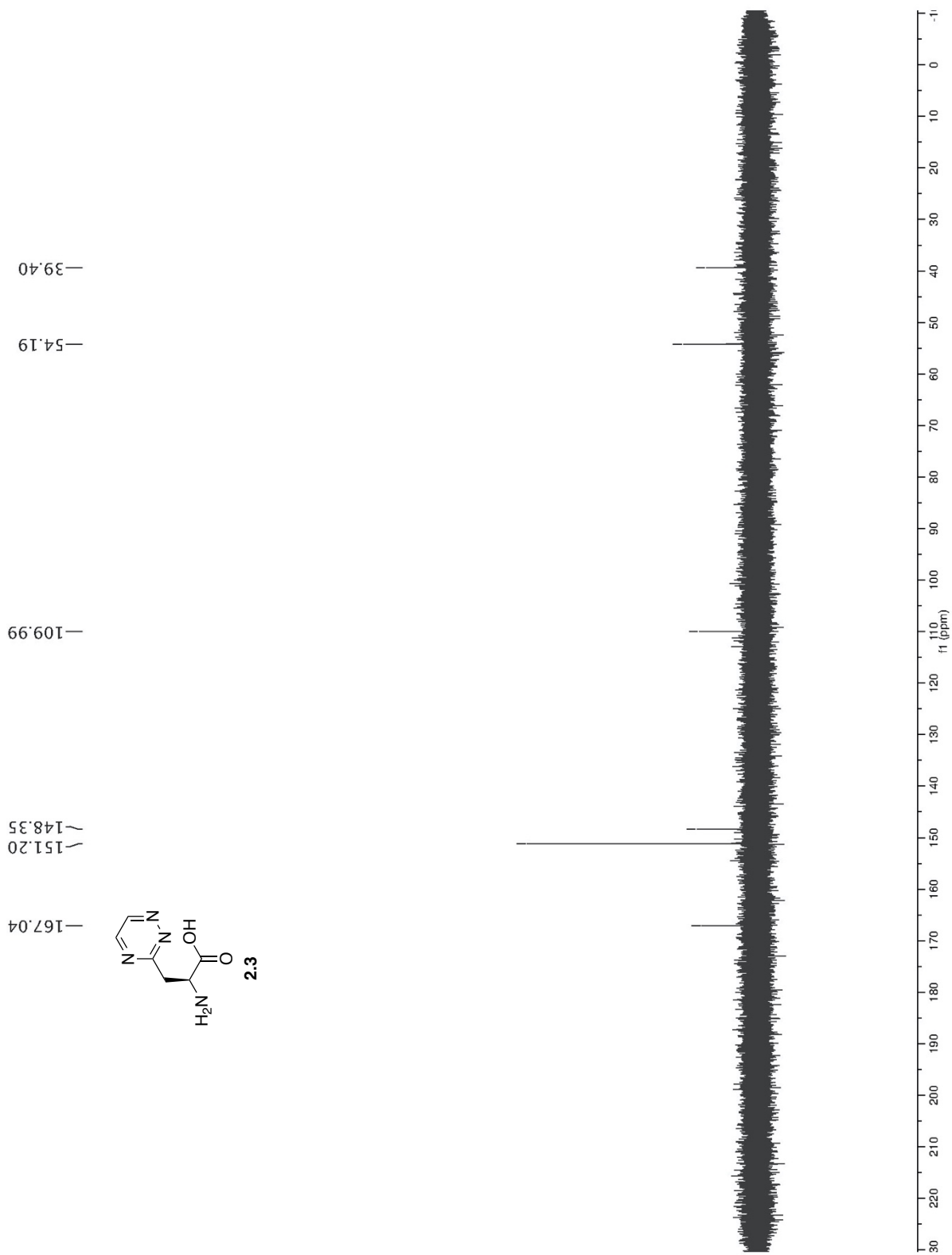
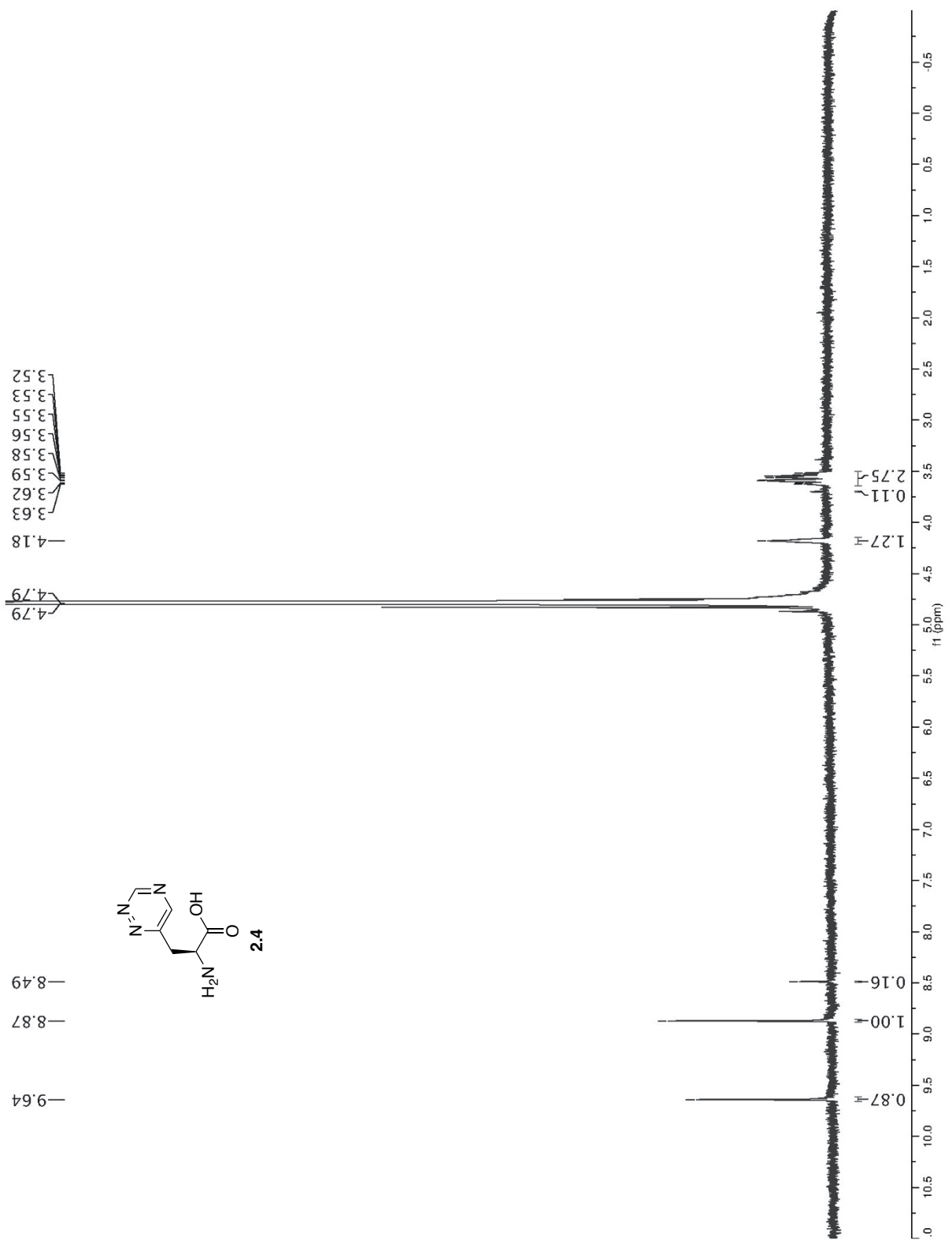
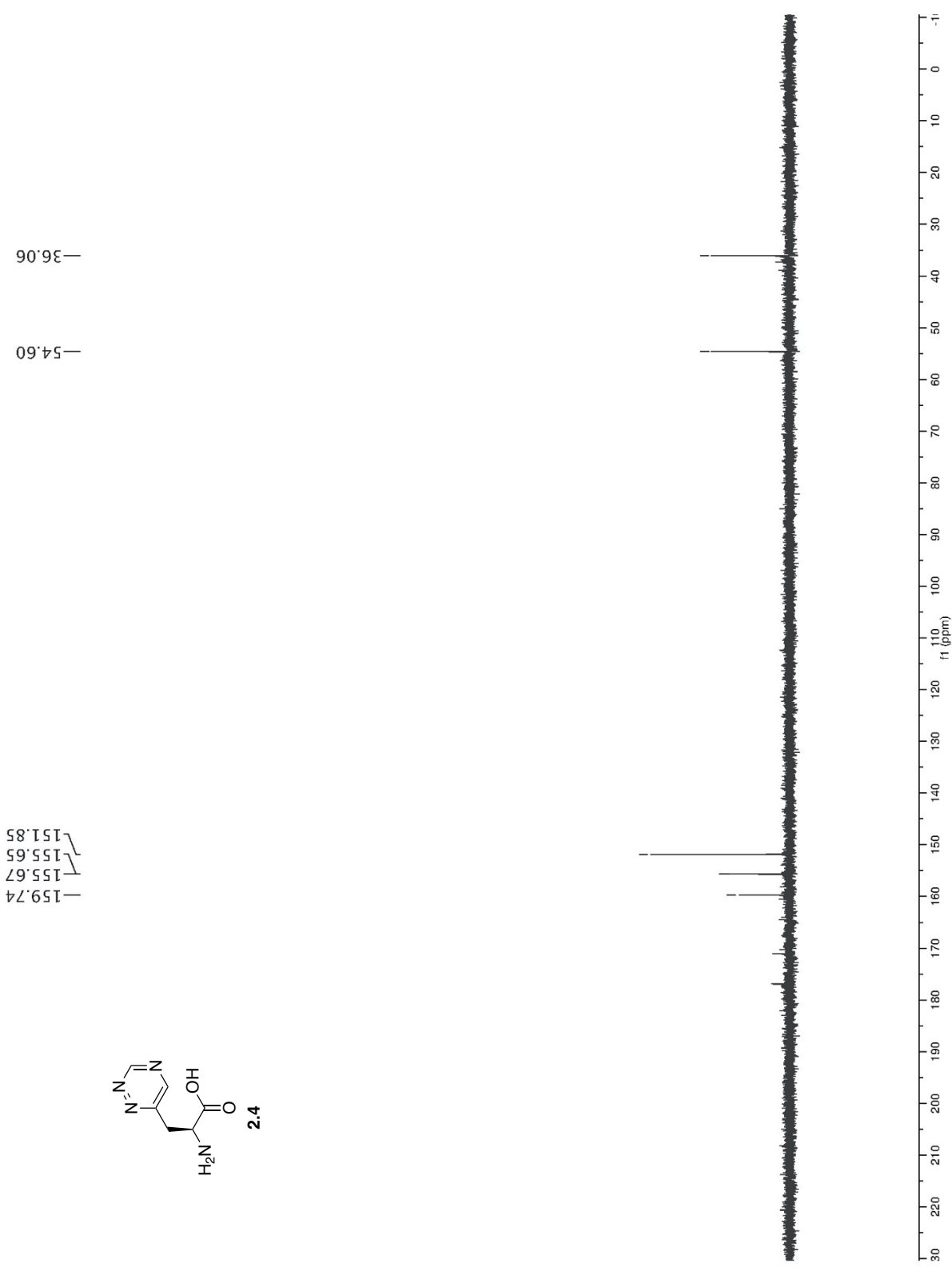


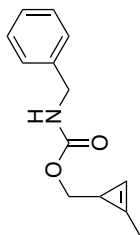
Figure S2-3. Triazine amino acid **2.4** is stable in aqueous solution over 7 d. Compound **2.4** (5 mM) was incubated in *d*-PBS (pH 7.4) at 37 °C and monitored via ¹H NMR spectroscopy. Deuterium exchange was observed over time.



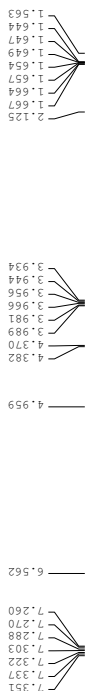




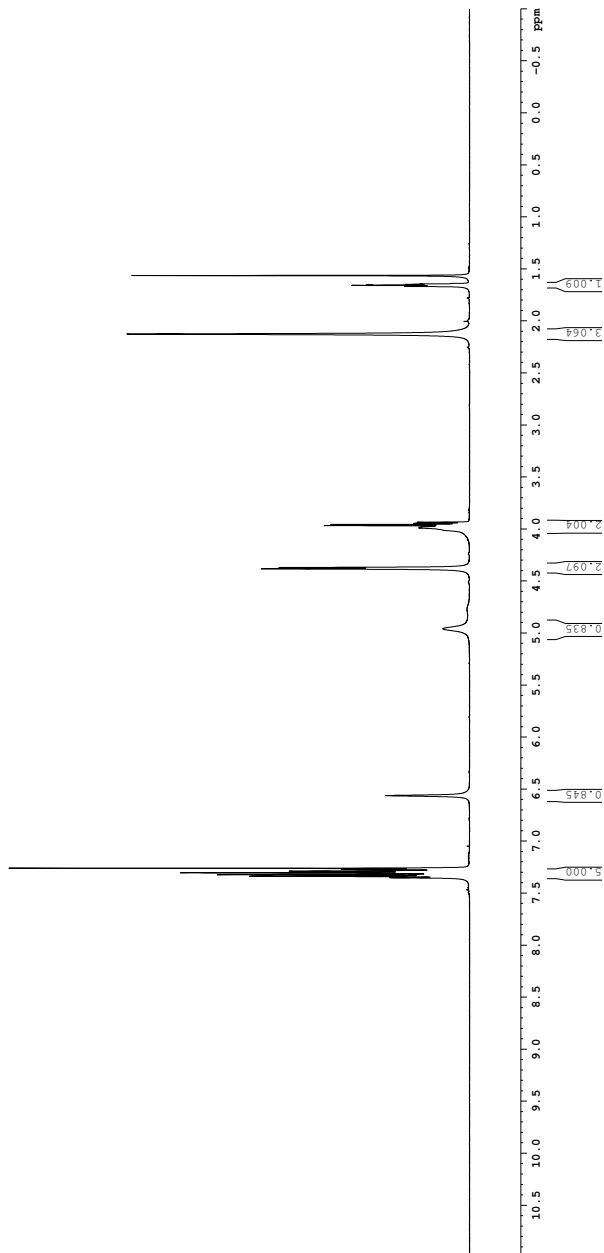
1H spectrum



2.18

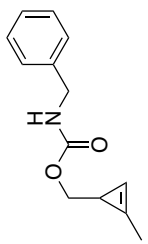
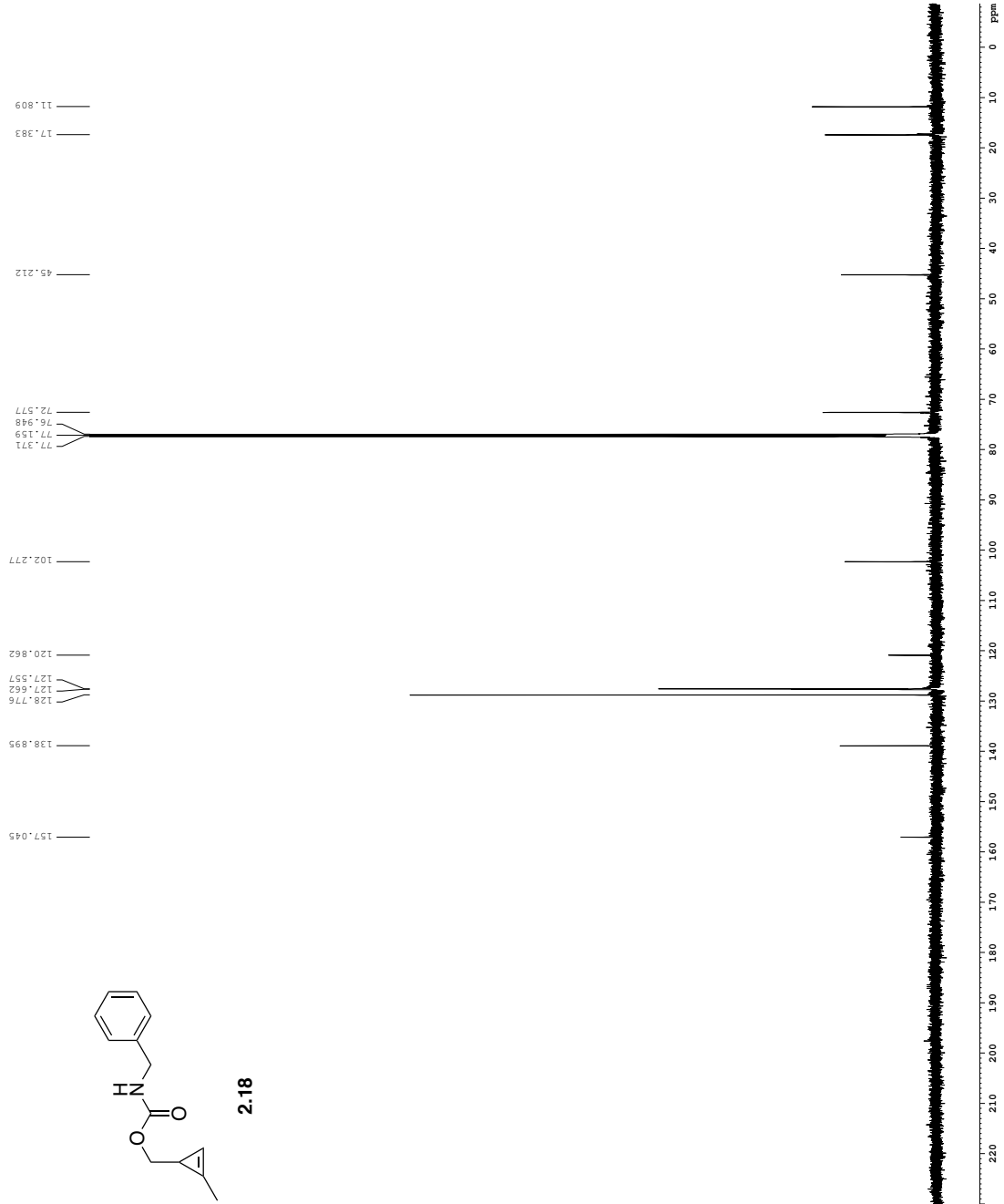


Current Data Parameters
NAME: 212-225
EXPNO: 3
PROCNO: 1
DATA1: /v/data/asanin/nmr
Date_: 20181004
INSTRUM: cryo500
PROBHD: 5 mm CPTCI 1H-
TD: 65536
F2 - Acquisition Parameters
SOLVENT: CDCl3
DS: 2
SWH: 8012.820 Hz
FIDRES: 0.190000 Hz
AQ: 5.0998273 sec
RG: 7.1
DE: 62.00 usec
TE: 300.2 K
D1: 0.1000000 sec
MCREST: 0 sec
MCWRK: 0.0500000 sec
===== CHANNEL f1 =====
NUC1: 13C
P1: 7.50 usec
PL1: 0.00 dB
SFO1: 500.220322 MHz
F2 - Processing parameters
ACQ: EM
SF: 500.220322 MHz
WDW: EM
SSB: 0
LB: 0.30 Hz
GB: 0
PC: 1.00



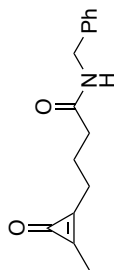
```

Current Data Parameters
EXPNO 2
PROCNO 2
DATAF1CH1 //data/summit/mr
F2 - Acquisition Parameters
Date_ 20181004
Time 11:50:00
INSTRUM spect
PROBHD 5 mm CPBBO-BB
PULPROG zgpg30
TD 65536
SFO1 150.1494000 MHz
AQ 0.4000000 sec
RG 1024
SC 1
SI 1
SF 150.1494000 MHz
NUC1 13C
NUC2 13C
PC 1.00
===== CHANNEL F2 =====
NAME waltz16
NUC1 13C
NUC2 13C
PC 1.00
===== CHANNEL F2 =====
CPDPRG2 waltz16
PC 1.00
PULPROG zgpg30
TD 65536
SFO1 150.1494000 MHz
SFO2 150.1494000 MHz
SF 150.1494000 MHz
===== Processing parameters =====
SI 1
SF 150.1494000 MHz
WDW EM
SSB 0
GB 0
PC 1.00
  
```



2.18

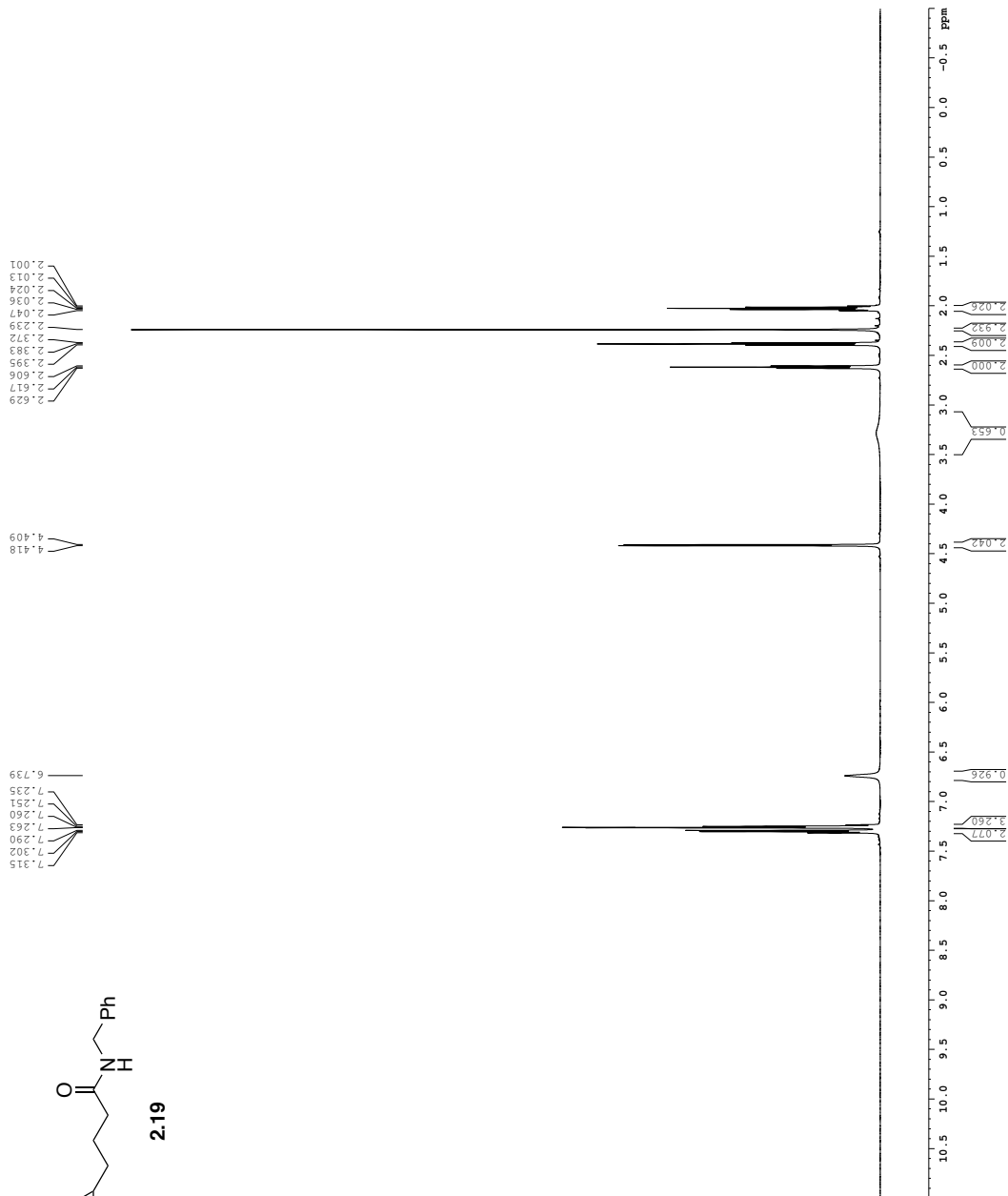
13C spectrum with 1H decoupling



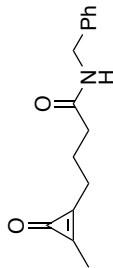
2.19

```

Current Data Parameters
NAME          DMS-123
EXPNO        2
PROCNO       1
DATAFATH     /v/data/asan/nf/mr
F2 - Acquisition Parameters
Date_        20180913
Time         11.40
INSTRUM      av600
PROBHD      5 mm CPBBO-BB-
PULPROG     zgpg30
TD          65536
SOLVENT     CDCl3
DS           2
SFR         9615.385 Hz
AQ          5.0998478 sec
RG          32.00
DE          11.81 MHz
TE          300.2 K
D1          0.10000000 sec
TD0         1
===== CHANNEL f1 =====
SFO1        600.134209 MHz
NUC1        13C
P1          11.50 usec
PL1         20.0000000 W
F2 - Processing Parameters
SI          65536
SF          600.1300929 MHz
RG          32.00
DE          11.81 MHz
TE          300.2 K
D1          0.10 Hz
GB          0
PC          1.00
    
```



¹³C spectrum with ¹H decoupling

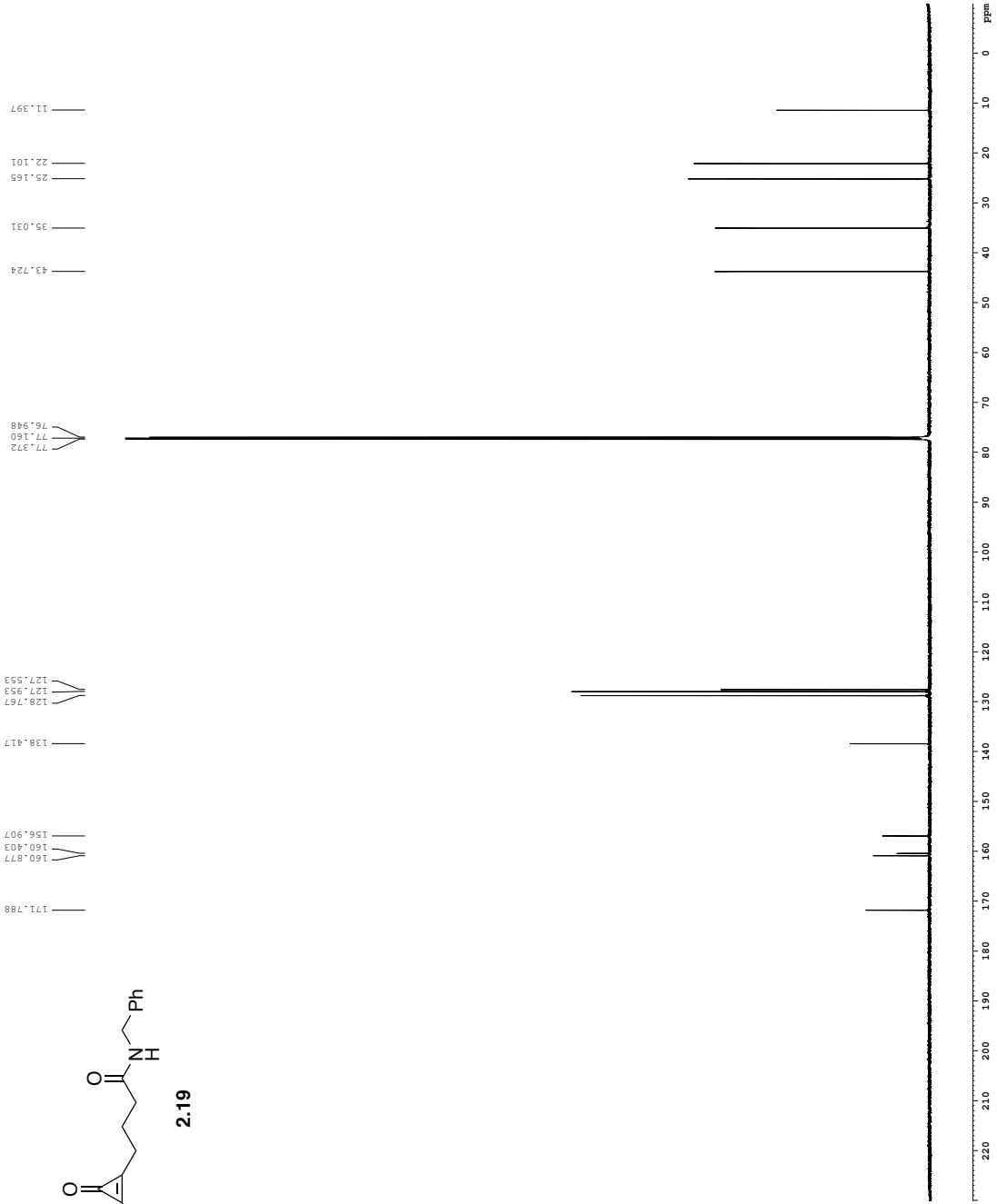


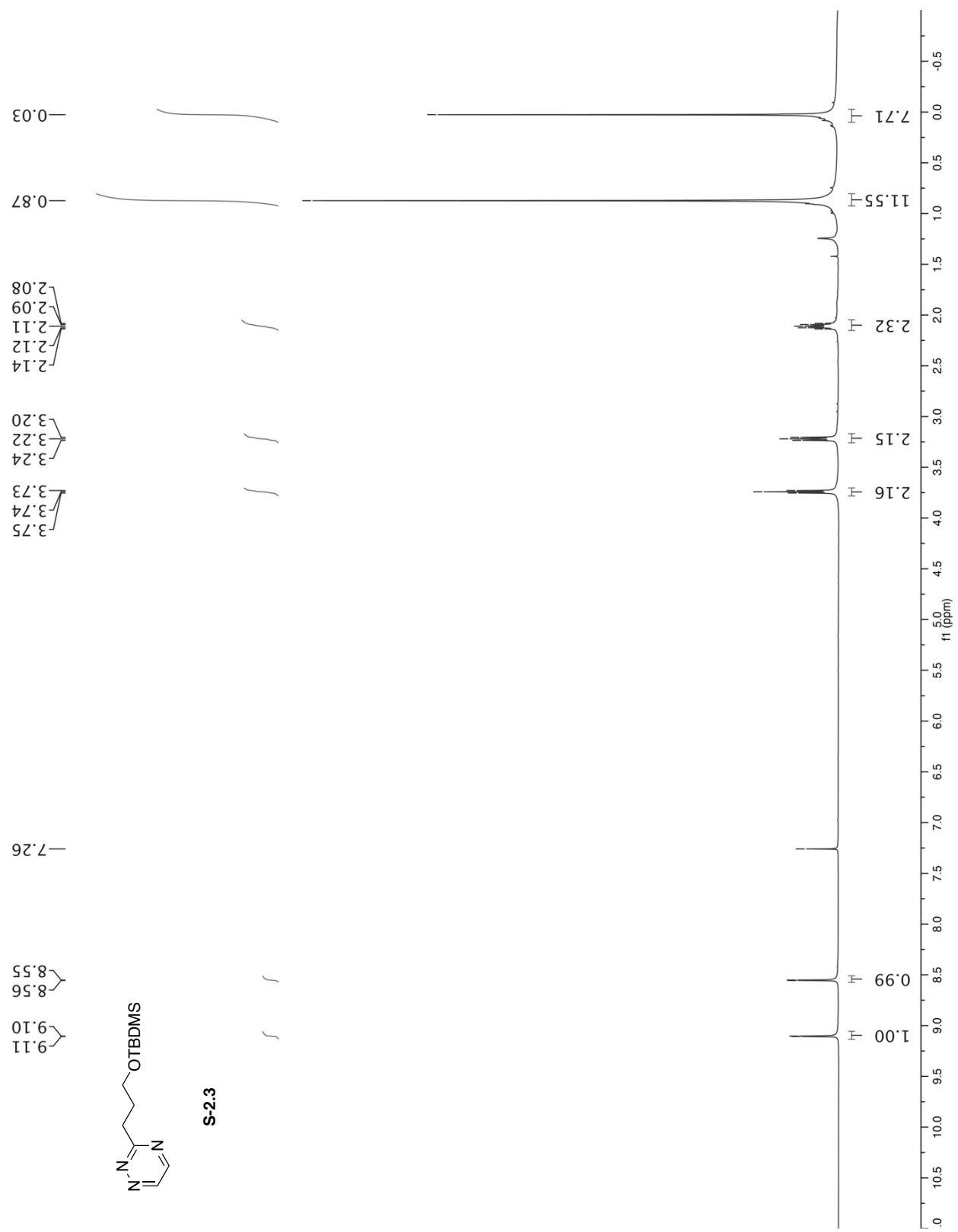
```

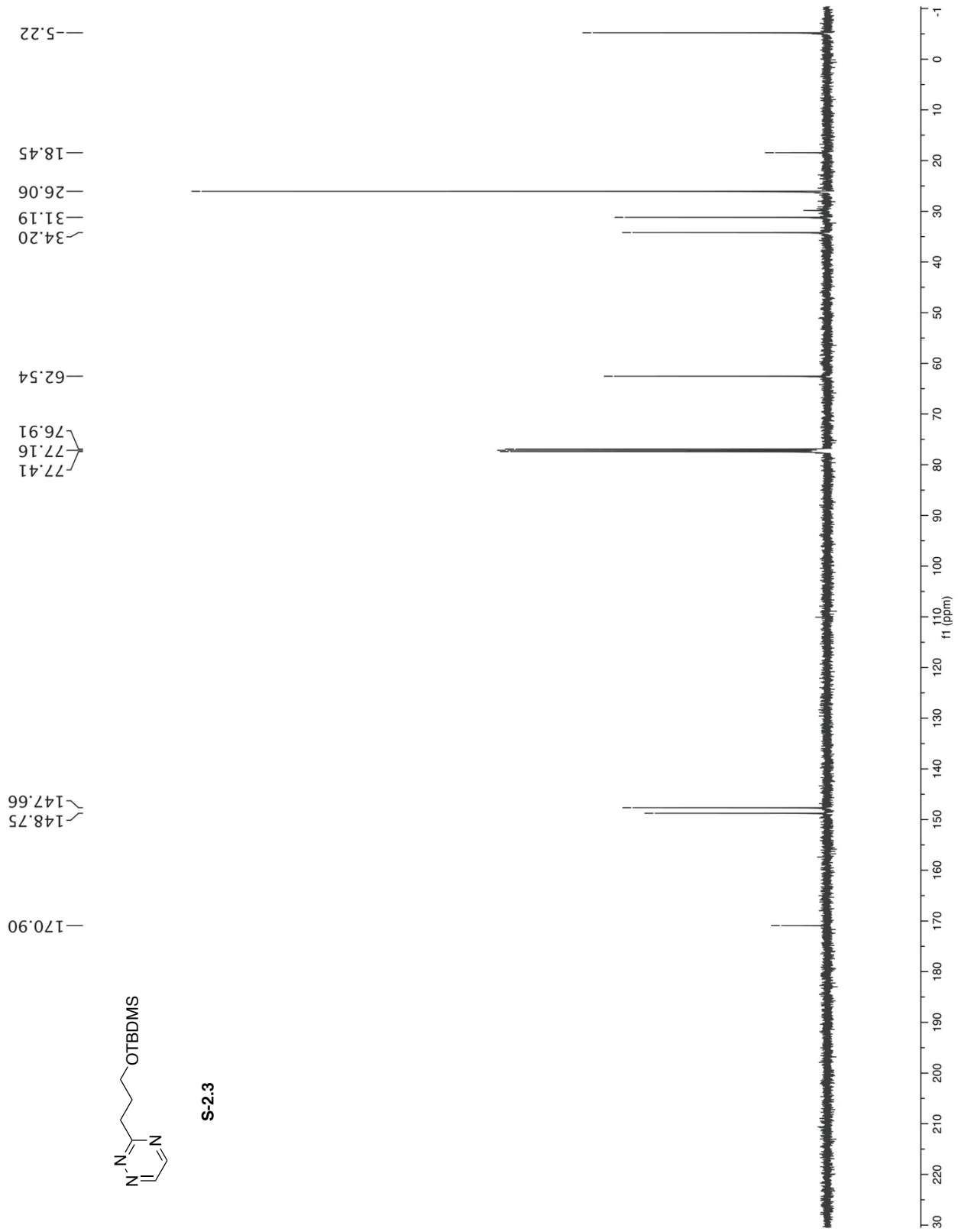
Current Data Parameters
EXPNO 1
PROCNO 1
PROBHD 5 mm CPBBO-BB
PULPROG zgpg30
TD 65536
SFO1 125.7614000 MHz
SF 125.7614000 MHz
WDW EM
SSB 0
GB 0
PC 1.00

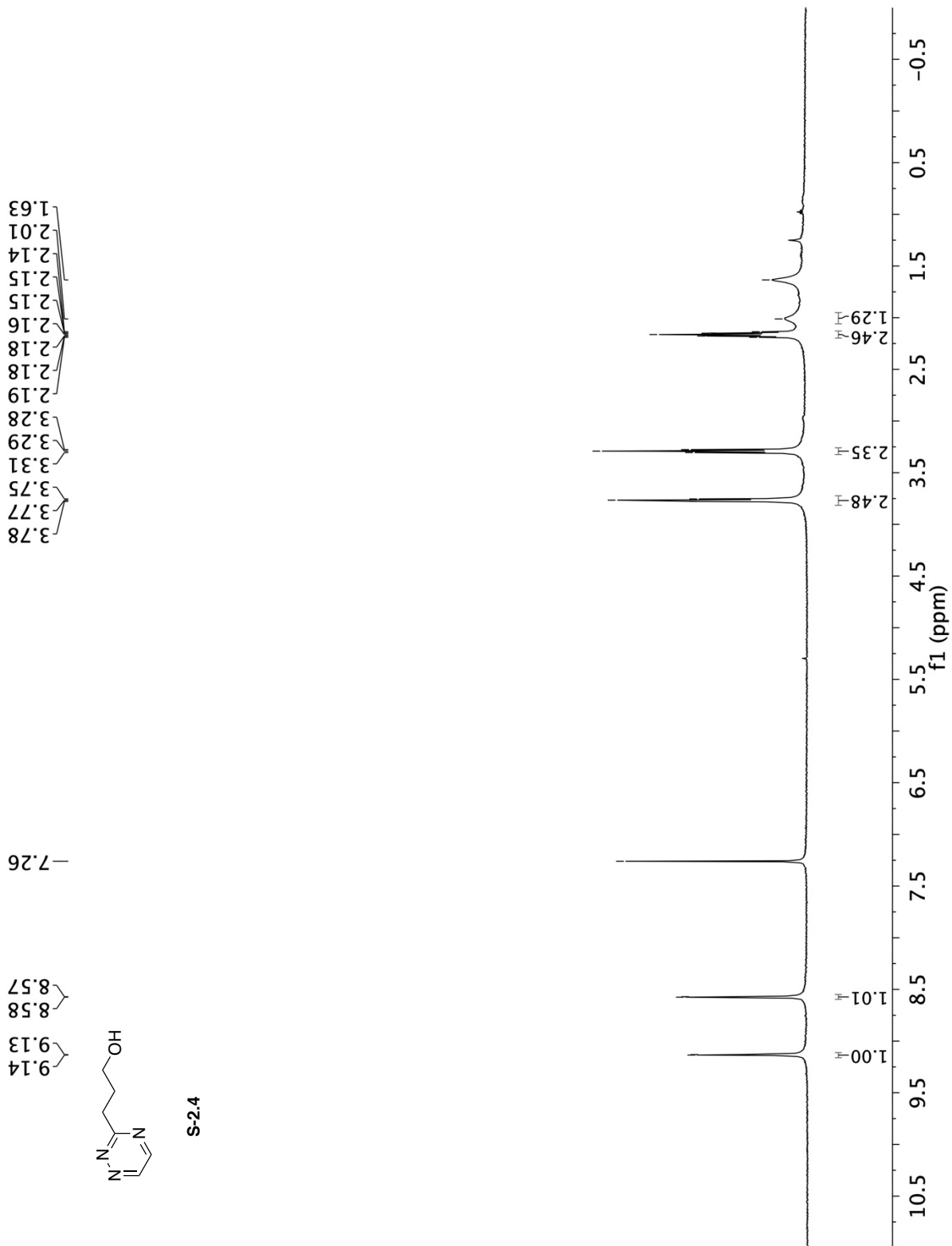
===== CHANNEL f1 =====
NUC1 13C
P1 1.20
PL1 0.00000000 MHz
F1 64.00000000 MHz
===== CHANNEL f2 =====
NUC2 13C
P2 1.20
PL2 0.00000000 MHz
F2 64.00000000 MHz
===== CHANNEL f3 =====
CPDPRG2 waltz16
NUC3 1H
P3 1.20
PL3 0.00000000 MHz
F3 500.1364000 MHz
===== CHANNEL f4 =====
CPDPRG4 zgpg30
NUC4 1H
P4 1.20
PL4 0.00000000 MHz
F4 500.1364000 MHz

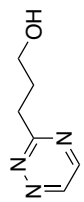
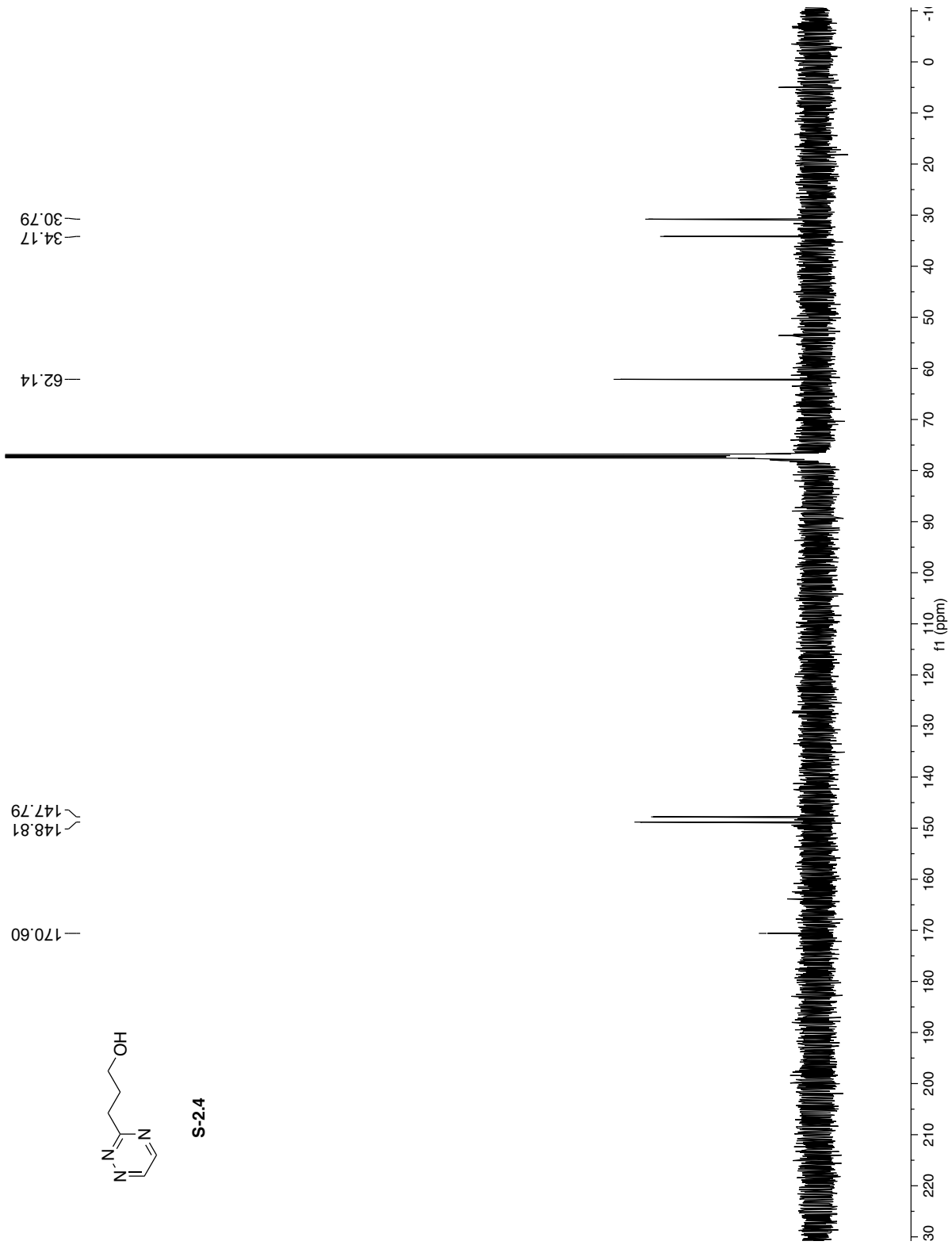
===== Processing parameters =====
SI 32768
SF 125.7614000 MHz
WDW EM
SSB 0
GB 0
PC 1.00
  
```



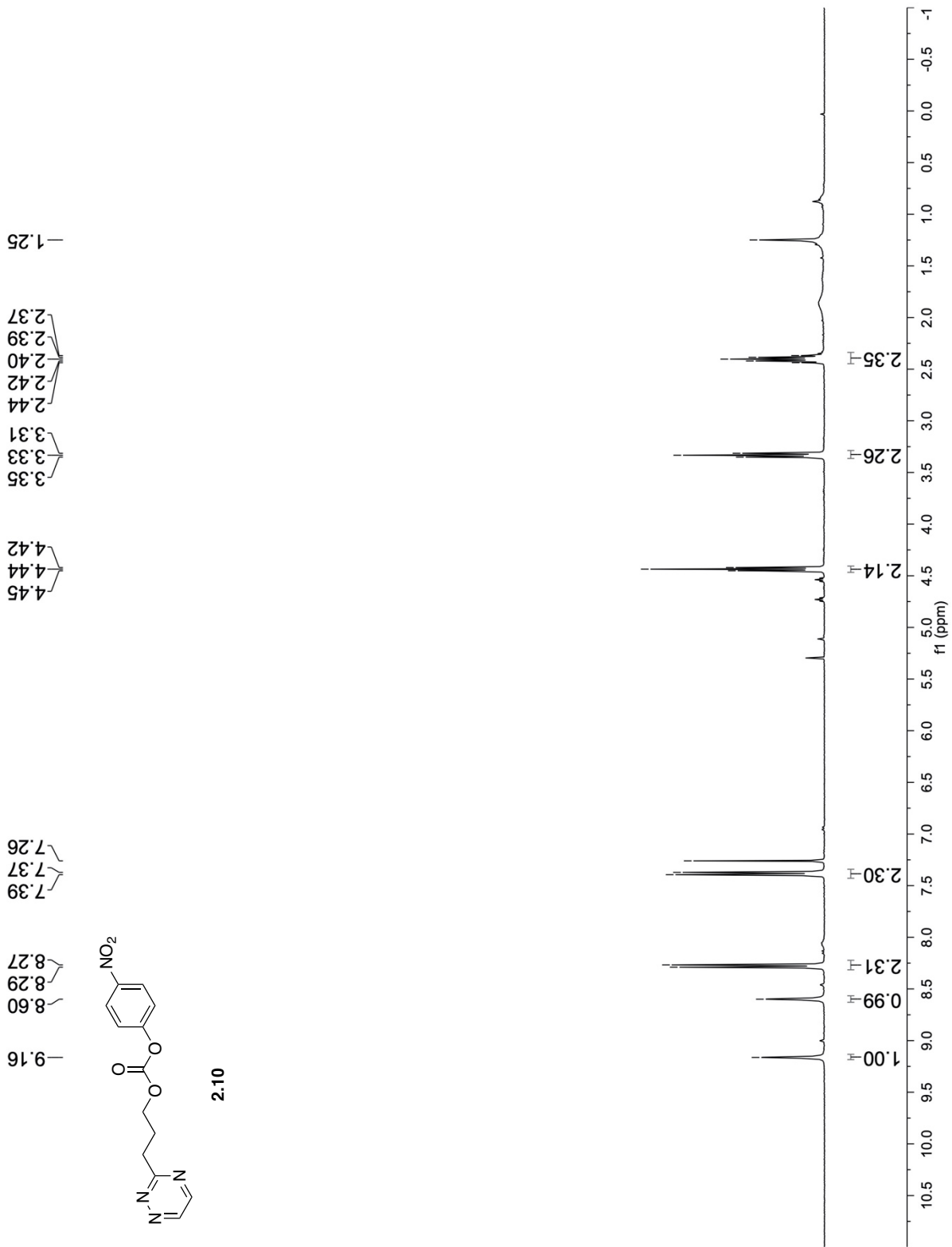


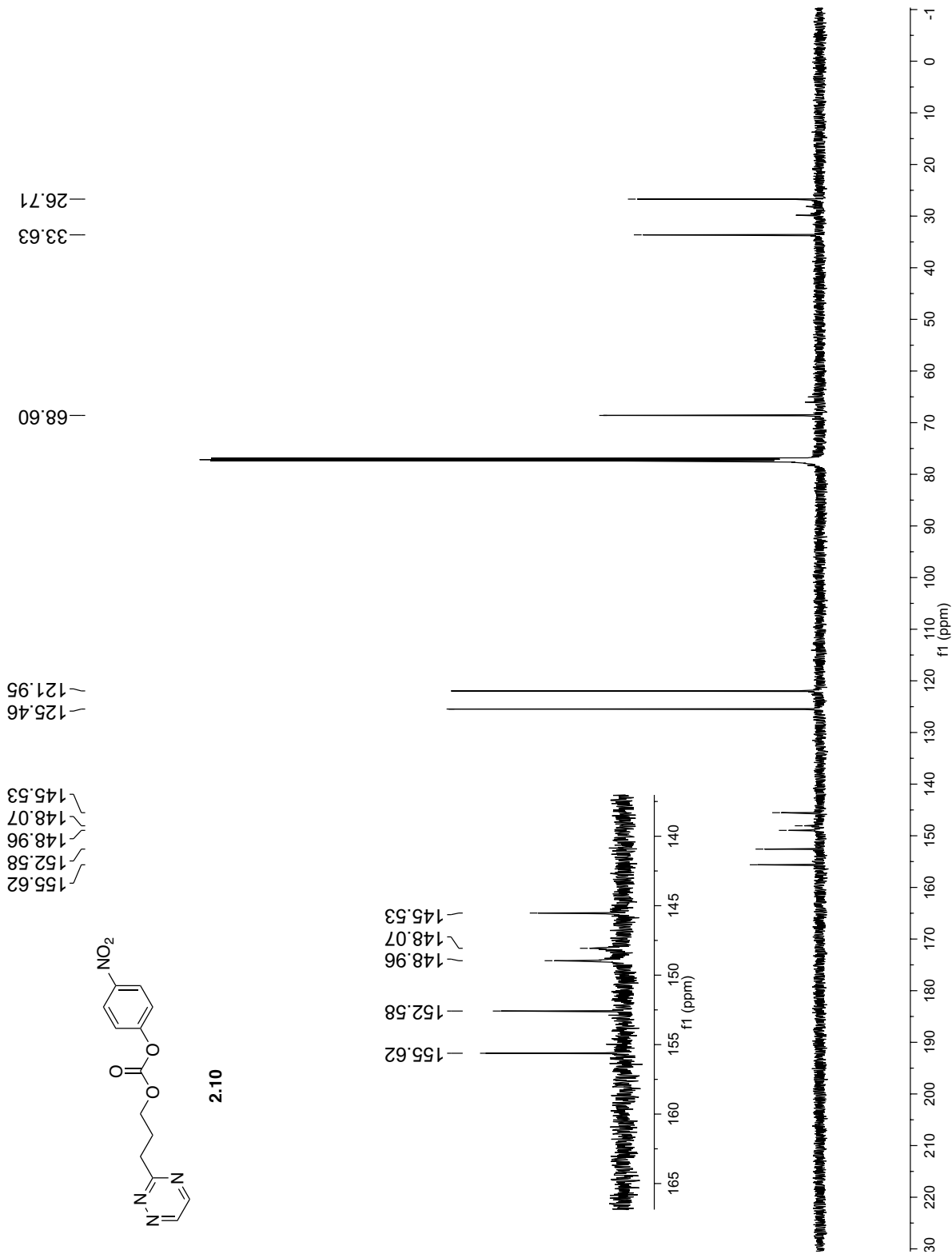






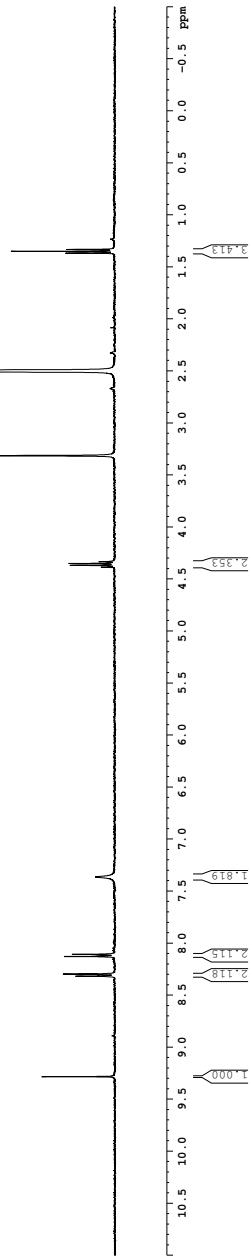
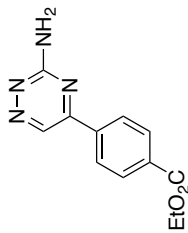
S-2.4



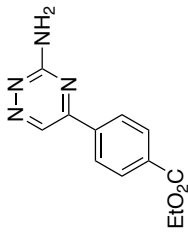


1H spectrum

Current Data Parameters
NAME: S-2.6
PROCNO: 1
Date Acquired: 20170608
Time: 15.02
PROBHD: 5 mm Multichannel
PULPROG: zgpg30
SOLVENT: CD3OD
NS: 8
DS: 4
SWH: 6410.256 Hz
FIDRES: 0.097813 Hz
RG: 51.18246 sec
AQ: 78.000 usec
TE: 298.2 K
MCREST: 0 sec
MCWRR: 0.0150000 sec
===== CHANNEL F1 =====
NUC1: 1H
P1: 12.00 usec
PL1: -1.00 dB
SFO1: 400.132809 MHz
F2 - Processing Parameters
SI: 65536
WDW: EM
SSB: 0
GB: 0
PC: 2.00

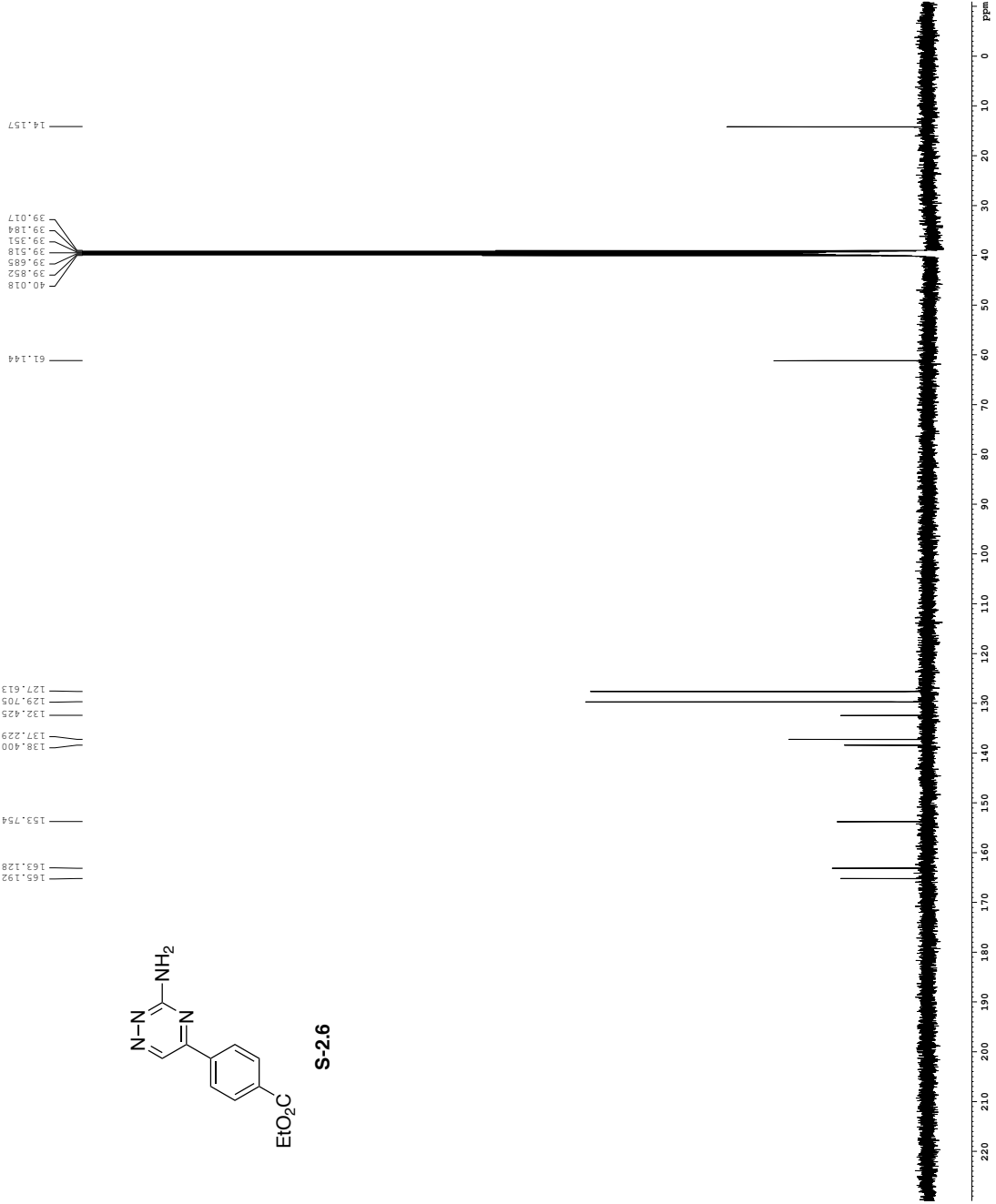


Z-restored spin-echo 13C spectrum with 1H decoupling

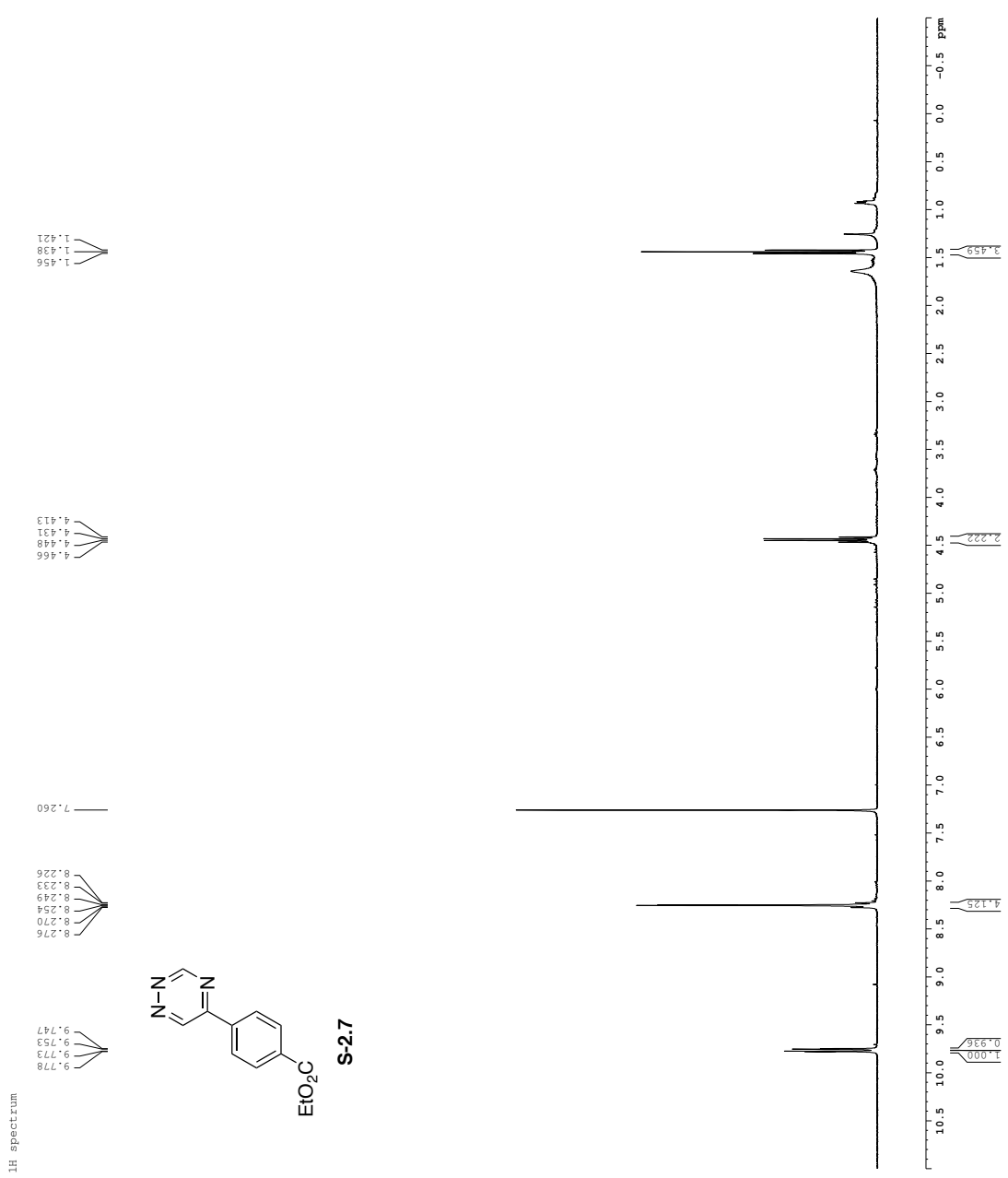


```

Current Data Parameters
EXPNO  1
PROCNO 1
Date_ Acquired  07.16.05
Date_  07.16.05
Time  15.09
INSTRUM spect
PROBHD  5 mm CPXI 1H-
PULPROG zgpg30
TD  65536
SOLVENT  CDCl3
NS  4096
DS  4
AQ  0.0390625 Hz
F2  300.136 MHz
F1  0.462388 Hz
RG  1.02384 sec
AQ2  0.0390625 Hz
DE  16.500 usec
TE  300.2 K
D1  0.20000000 sec
d11  0.03000000 sec
d12  0.03000000 sec
d13  0.03000000 sec
d14  0.03000000 sec
d15  0.03000000 sec
d16  0.03000000 sec
d17  0.03000000 sec
d18  0.03000000 sec
d19  0.03000000 sec
d20  0.03000000 sec
d21  0.03000000 sec
d22  0.03000000 sec
d23  0.03000000 sec
d24  0.03000000 sec
d25  0.03000000 sec
d26  0.03000000 sec
d27  0.03000000 sec
d28  0.03000000 sec
d29  0.03000000 sec
d30  0.03000000 sec
d31  0.03000000 sec
d32  0.03000000 sec
d33  0.03000000 sec
d34  0.03000000 sec
d35  0.03000000 sec
d36  0.03000000 sec
d37  0.03000000 sec
d38  0.03000000 sec
d39  0.03000000 sec
d40  0.03000000 sec
d41  0.03000000 sec
d42  0.03000000 sec
d43  0.03000000 sec
d44  0.03000000 sec
d45  0.03000000 sec
d46  0.03000000 sec
d47  0.03000000 sec
d48  0.03000000 sec
d49  0.03000000 sec
d50  0.03000000 sec
d51  0.03000000 sec
d52  0.03000000 sec
d53  0.03000000 sec
d54  0.03000000 sec
d55  0.03000000 sec
d56  0.03000000 sec
d57  0.03000000 sec
d58  0.03000000 sec
d59  0.03000000 sec
d60  0.03000000 sec
d61  0.03000000 sec
d62  0.03000000 sec
d63  0.03000000 sec
d64  0.03000000 sec
d65  0.03000000 sec
d66  0.03000000 sec
d67  0.03000000 sec
d68  0.03000000 sec
d69  0.03000000 sec
d70  0.03000000 sec
d71  0.03000000 sec
d72  0.03000000 sec
d73  0.03000000 sec
d74  0.03000000 sec
d75  0.03000000 sec
d76  0.03000000 sec
d77  0.03000000 sec
d78  0.03000000 sec
d79  0.03000000 sec
d80  0.03000000 sec
d81  0.03000000 sec
d82  0.03000000 sec
d83  0.03000000 sec
d84  0.03000000 sec
d85  0.03000000 sec
d86  0.03000000 sec
d87  0.03000000 sec
d88  0.03000000 sec
d89  0.03000000 sec
d90  0.03000000 sec
d91  0.03000000 sec
d92  0.03000000 sec
d93  0.03000000 sec
d94  0.03000000 sec
d95  0.03000000 sec
d96  0.03000000 sec
d97  0.03000000 sec
d98  0.03000000 sec
d99  0.03000000 sec
d100 0.03000000 sec
===== CHANNEL f1 =====
NUC1  13C
P1  14.55 usec
PL1  0.00 dB
PC1  191.10 MHz
===== CHANNEL f2 =====
NUC2  1H
P2  12.00 usec
PL2  19.00 dB
PC2  500.136091 MHz
===== GRABENT CHANNEL =====
GRAB1(1) SINE.100
GRAB1(2) SINE.100
GRAB1  0 %
GRAB2  0 %
GRAB3  0 %
GRAB4  0 %
GRAB5  0 %
GRAB6  0 %
GRAB7  0 %
GRAB8  0 %
GRAB9  0 %
GRAB10 0 %
===== Processing Parameters =====
SI  65536
SF  125.760369 MHz
WDW  EM
SSB  0
LB  1.00 Hz
GB  0
PC  2.00
  
```



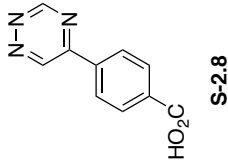
Current Data Parameters
 NAME SN3-103
 PROCNO 1
 F2 - Acquisition Parameters
 Date_ Time 20170703 10:25
 PROBNM 103
 PULPROG 5 mm zgpg30
 PULPROG 6930
 SOLVENT DMSO
 NS 9
 SRH 6410.256 Hz
 FIDRES 0.09033 Hz
 RG 327.12302
 DW 78.000 usec
 DE 238.0 K
 TE 0.10000000 sec
 MCREST 0 sec
 MCWRE 0.01500000 sec
 ===== CHANNEL f1 =====
 NUC1 1H
 P1 11.00 usec
 PL1 -1.10 dB
 SFO1 400.132809 MHz
 F2 - Processing parameters
 SI 6536
 SF 400.132809 MHz
 EM
 XNM 0
 XBN 0 0.30 Hz
 GB 0
 PC 2.00



Current Data Parameters
 NAME S13-107
 POCNO 1
 Date Acquisition Parameters
 Time 15.57
 PULPROG 5 mm MULL101
 SOLVENT CDCl3
 NS 8
 SNH 6410.256 Hz
 FIDRES 0.097813 Hz
 RG 32.18246 sec
 DW 78.000 usec
 TE 298.1 K
 MCREST 0 sec 0.1000000 sec
 MCHWK 0.0150000 sec
 ===== CHANNEL f1 =====
 NUCL 1H
 P1 11.00 usec
 PL1 -1.10 dB
 SFO1 400.132809 MHz
 F2 - Processing Parameters
 SI 6536 MHz
 KW 0 EM
 LB 0 0.30 Hz
 GB 0
 PC 2.00



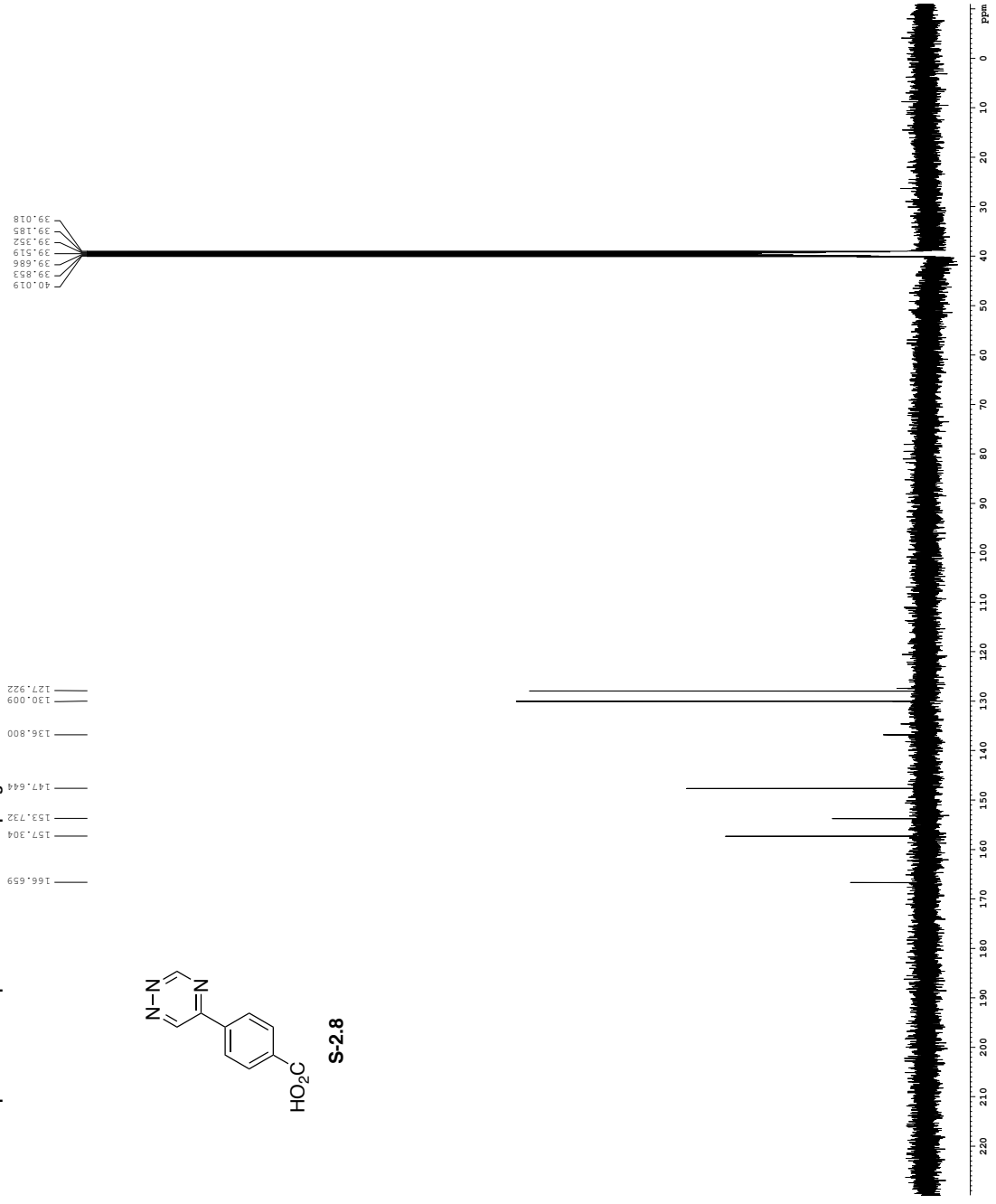
Z-restored spin-echo 13C spectrum with 1H decoupling



```

Current Data Parameters
EXPNO 3
PROCNO 1
-----
F2 - Acquisition Parameters
Date_ 20110616
Time 16:33
Operator jk
PROBHD 5 mm CPXI 1H
PULPROG zgpg30
TD 65536
SOLVENT CDCl3
DS 4
DE 0.00116
FIDRES 0.000166
AQ 1.081194
RG 16.500
DM 16.500
SFO1 125.762549
TE 300.2
D1 0.001000
d11 0.001000
d15 0.001000
MCSHFT 0
PC 33.10
-----
CHANNEL f1
NUC1 13C
P1 14.130
P2 500.000
F1 500.000
F2 125.762
SFO1 125.762549
SFO2 2.76
SFO3 2.76
SFO4 2.76
SFO5 2.76
SFO6 2.76
SFO7 2.76
SFO8 2.76
SFO9 2.76
SFO10 2.76
SFO11 2.76
SFO12 2.76
SFO13 2.76
SFO14 2.76
SFO15 2.76
SFO16 2.76
SFO17 2.76
SFO18 2.76
SFO19 2.76
SFO20 2.76
SFO21 2.76
SFO22 2.76
SFO23 2.76
SFO24 2.76
SFO25 2.76
SFO26 2.76
SFO27 2.76
SFO28 2.76
SFO29 2.76
SFO30 2.76
SFO31 2.76
SFO32 2.76
SFO33 2.76
SFO34 2.76
SFO35 2.76
SFO36 2.76
SFO37 2.76
SFO38 2.76
SFO39 2.76
SFO40 2.76
SFO41 2.76
SFO42 2.76
SFO43 2.76
SFO44 2.76
SFO45 2.76
SFO46 2.76
SFO47 2.76
SFO48 2.76
SFO49 2.76
SFO50 2.76
SFO51 2.76
SFO52 2.76
SFO53 2.76
SFO54 2.76
SFO55 2.76
SFO56 2.76
SFO57 2.76
SFO58 2.76
SFO59 2.76
SFO60 2.76
SFO61 2.76
SFO62 2.76
SFO63 2.76
SFO64 2.76
SFO65 2.76
SFO66 2.76
SFO67 2.76
SFO68 2.76
SFO69 2.76
SFO70 2.76
SFO71 2.76
SFO72 2.76
SFO73 2.76
SFO74 2.76
SFO75 2.76
SFO76 2.76
SFO77 2.76
SFO78 2.76
SFO79 2.76
SFO80 2.76
SFO81 2.76
SFO82 2.76
SFO83 2.76
SFO84 2.76
SFO85 2.76
SFO86 2.76
SFO87 2.76
SFO88 2.76
SFO89 2.76
SFO90 2.76
SFO91 2.76
SFO92 2.76
SFO93 2.76
SFO94 2.76
SFO95 2.76
SFO96 2.76
SFO97 2.76
SFO98 2.76
SFO99 2.76
SFO100 2.76
-----
CHANNEL f2
NUC2 13C
P1 14.130
P2 500.000
F1 500.000
F2 125.762
SFO1 125.762549
SFO2 2.76
SFO3 2.76
SFO4 2.76
SFO5 2.76
SFO6 2.76
SFO7 2.76
SFO8 2.76
SFO9 2.76
SFO10 2.76
SFO11 2.76
SFO12 2.76
SFO13 2.76
SFO14 2.76
SFO15 2.76
SFO16 2.76
SFO17 2.76
SFO18 2.76
SFO19 2.76
SFO20 2.76
SFO21 2.76
SFO22 2.76
SFO23 2.76
SFO24 2.76
SFO25 2.76
SFO26 2.76
SFO27 2.76
SFO28 2.76
SFO29 2.76
SFO30 2.76
SFO31 2.76
SFO32 2.76
SFO33 2.76
SFO34 2.76
SFO35 2.76
SFO36 2.76
SFO37 2.76
SFO38 2.76
SFO39 2.76
SFO40 2.76
SFO41 2.76
SFO42 2.76
SFO43 2.76
SFO44 2.76
SFO45 2.76
SFO46 2.76
SFO47 2.76
SFO48 2.76
SFO49 2.76
SFO50 2.76
SFO51 2.76
SFO52 2.76
SFO53 2.76
SFO54 2.76
SFO55 2.76
SFO56 2.76
SFO57 2.76
SFO58 2.76
SFO59 2.76
SFO60 2.76
SFO61 2.76
SFO62 2.76
SFO63 2.76
SFO64 2.76
SFO65 2.76
SFO66 2.76
SFO67 2.76
SFO68 2.76
SFO69 2.76
SFO70 2.76
SFO71 2.76
SFO72 2.76
SFO73 2.76
SFO74 2.76
SFO75 2.76
SFO76 2.76
SFO77 2.76
SFO78 2.76
SFO79 2.76
SFO80 2.76
SFO81 2.76
SFO82 2.76
SFO83 2.76
SFO84 2.76
SFO85 2.76
SFO86 2.76
SFO87 2.76
SFO88 2.76
SFO89 2.76
SFO90 2.76
SFO91 2.76
SFO92 2.76
SFO93 2.76
SFO94 2.76
SFO95 2.76
SFO96 2.76
SFO97 2.76
SFO98 2.76
SFO99 2.76
SFO100 2.76
-----
CONSISTENT CHANNELS
GPM1(1) SINE100
GPM1(2) SINE100
GPM2(1) SINE100
GPM2(2) SINE100
GPM3(1) SINE100
GPM3(2) SINE100
GPM4(1) SINE100
GPM4(2) SINE100
GPM5(1) SINE100
GPM5(2) SINE100
GPM6(1) SINE100
GPM6(2) SINE100
GPM7(1) SINE100
GPM7(2) SINE100
GPM8(1) SINE100
GPM8(2) SINE100
GPM9(1) SINE100
GPM9(2) SINE100
GPM10(1) SINE100
GPM10(2) SINE100
GPM11(1) SINE100
GPM11(2) SINE100
GPM12(1) SINE100
GPM12(2) SINE100
GPM13(1) SINE100
GPM13(2) SINE100
GPM14(1) SINE100
GPM14(2) SINE100
GPM15(1) SINE100
GPM15(2) SINE100
GPM16(1) SINE100
GPM16(2) SINE100
GPM17(1) SINE100
GPM17(2) SINE100
GPM18(1) SINE100
GPM18(2) SINE100
GPM19(1) SINE100
GPM19(2) SINE100
GPM20(1) SINE100
GPM20(2) SINE100
GPM21(1) SINE100
GPM21(2) SINE100
GPM22(1) SINE100
GPM22(2) SINE100
GPM23(1) SINE100
GPM23(2) SINE100
GPM24(1) SINE100
GPM24(2) SINE100
GPM25(1) SINE100
GPM25(2) SINE100
GPM26(1) SINE100
GPM26(2) SINE100
GPM27(1) SINE100
GPM27(2) SINE100
GPM28(1) SINE100
GPM28(2) SINE100
GPM29(1) SINE100
GPM29(2) SINE100
GPM30(1) SINE100
GPM30(2) SINE100
GPM31(1) SINE100
GPM31(2) SINE100
GPM32(1) SINE100
GPM32(2) SINE100
GPM33(1) SINE100
GPM33(2) SINE100
GPM34(1) SINE100
GPM34(2) SINE100
GPM35(1) SINE100
GPM35(2) SINE100
GPM36(1) SINE100
GPM36(2) SINE100
GPM37(1) SINE100
GPM37(2) SINE100
GPM38(1) SINE100
GPM38(2) SINE100
GPM39(1) SINE100
GPM39(2) SINE100
GPM40(1) SINE100
GPM40(2) SINE100
GPM41(1) SINE100
GPM41(2) SINE100
GPM42(1) SINE100
GPM42(2) SINE100
GPM43(1) SINE100
GPM43(2) SINE100
GPM44(1) SINE100
GPM44(2) SINE100
GPM45(1) SINE100
GPM45(2) SINE100
GPM46(1) SINE100
GPM46(2) SINE100
GPM47(1) SINE100
GPM47(2) SINE100
GPM48(1) SINE100
GPM48(2) SINE100
GPM49(1) SINE100
GPM49(2) SINE100
GPM50(1) SINE100
GPM50(2) SINE100
GPM51(1) SINE100
GPM51(2) SINE100
GPM52(1) SINE100
GPM52(2) SINE100
GPM53(1) SINE100
GPM53(2) SINE100
GPM54(1) SINE100
GPM54(2) SINE100
GPM55(1) SINE100
GPM55(2) SINE100
GPM56(1) SINE100
GPM56(2) SINE100
GPM57(1) SINE100
GPM57(2) SINE100
GPM58(1) SINE100
GPM58(2) SINE100
GPM59(1) SINE100
GPM59(2) SINE100
GPM60(1) SINE100
GPM60(2) SINE100
GPM61(1) SINE100
GPM61(2) SINE100
GPM62(1) SINE100
GPM62(2) SINE100
GPM63(1) SINE100
GPM63(2) SINE100
GPM64(1) SINE100
GPM64(2) SINE100
GPM65(1) SINE100
GPM65(2) SINE100
GPM66(1) SINE100
GPM66(2) SINE100
GPM67(1) SINE100
GPM67(2) SINE100
GPM68(1) SINE100
GPM68(2) SINE100
GPM69(1) SINE100
GPM69(2) SINE100
GPM70(1) SINE100
GPM70(2) SINE100
GPM71(1) SINE100
GPM71(2) SINE100
GPM72(1) SINE100
GPM72(2) SINE100
GPM73(1) SINE100
GPM73(2) SINE100
GPM74(1) SINE100
GPM74(2) SINE100
GPM75(1) SINE100
GPM75(2) SINE100
GPM76(1) SINE100
GPM76(2) SINE100
GPM77(1) SINE100
GPM77(2) SINE100
GPM78(1) SINE100
GPM78(2) SINE100
GPM79(1) SINE100
GPM79(2) SINE100
GPM80(1) SINE100
GPM80(2) SINE100
GPM81(1) SINE100
GPM81(2) SINE100
GPM82(1) SINE100
GPM82(2) SINE100
GPM83(1) SINE100
GPM83(2) SINE100
GPM84(1) SINE100
GPM84(2) SINE100
GPM85(1) SINE100
GPM85(2) SINE100
GPM86(1) SINE100
GPM86(2) SINE100
GPM87(1) SINE100
GPM87(2) SINE100
GPM88(1) SINE100
GPM88(2) SINE100
GPM89(1) SINE100
GPM89(2) SINE100
GPM90(1) SINE100
GPM90(2) SINE100
GPM91(1) SINE100
GPM91(2) SINE100
GPM92(1) SINE100
GPM92(2) SINE100
GPM93(1) SINE100
GPM93(2) SINE100
GPM94(1) SINE100
GPM94(2) SINE100
GPM95(1) SINE100
GPM95(2) SINE100
GPM96(1) SINE100
GPM96(2) SINE100
GPM97(1) SINE100
GPM97(2) SINE100
GPM98(1) SINE100
GPM98(2) SINE100
GPM99(1) SINE100
GPM99(2) SINE100
GPM100(1) SINE100
GPM100(2) SINE100
-----
F2 - Processing Parameters
SI 65536
SF 125.762549
WDW EM
SSB 0
GB 0
PC 2.00

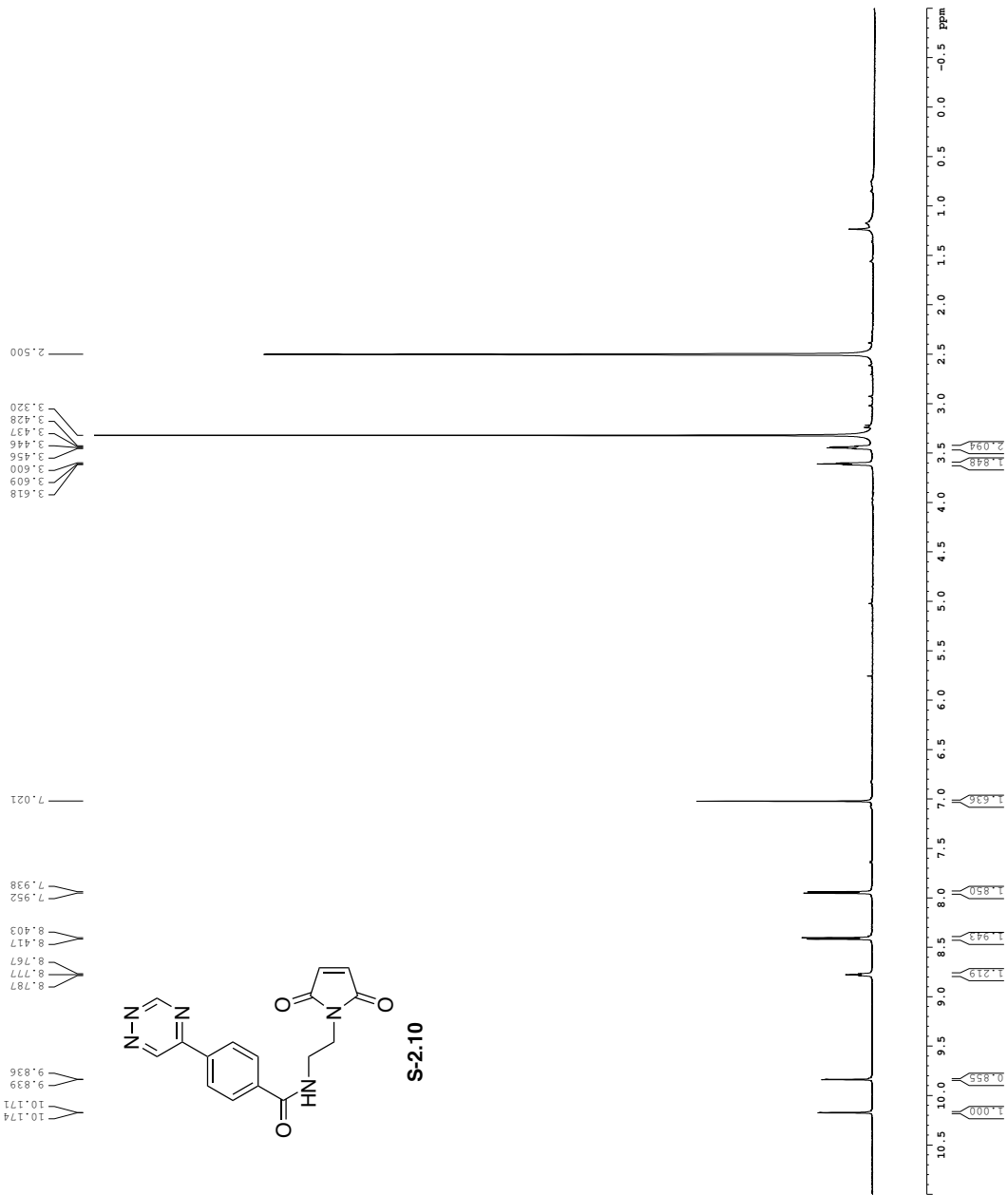
```

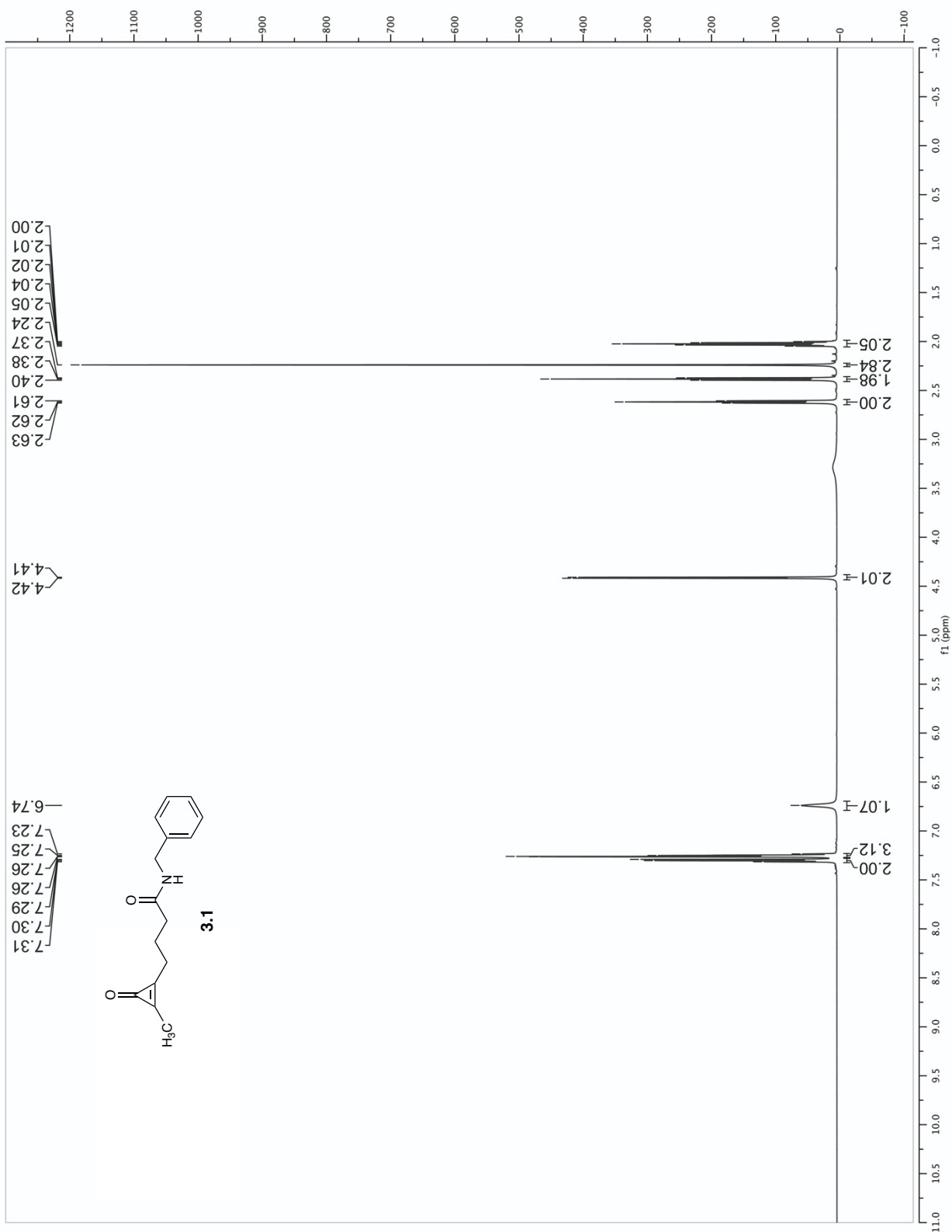


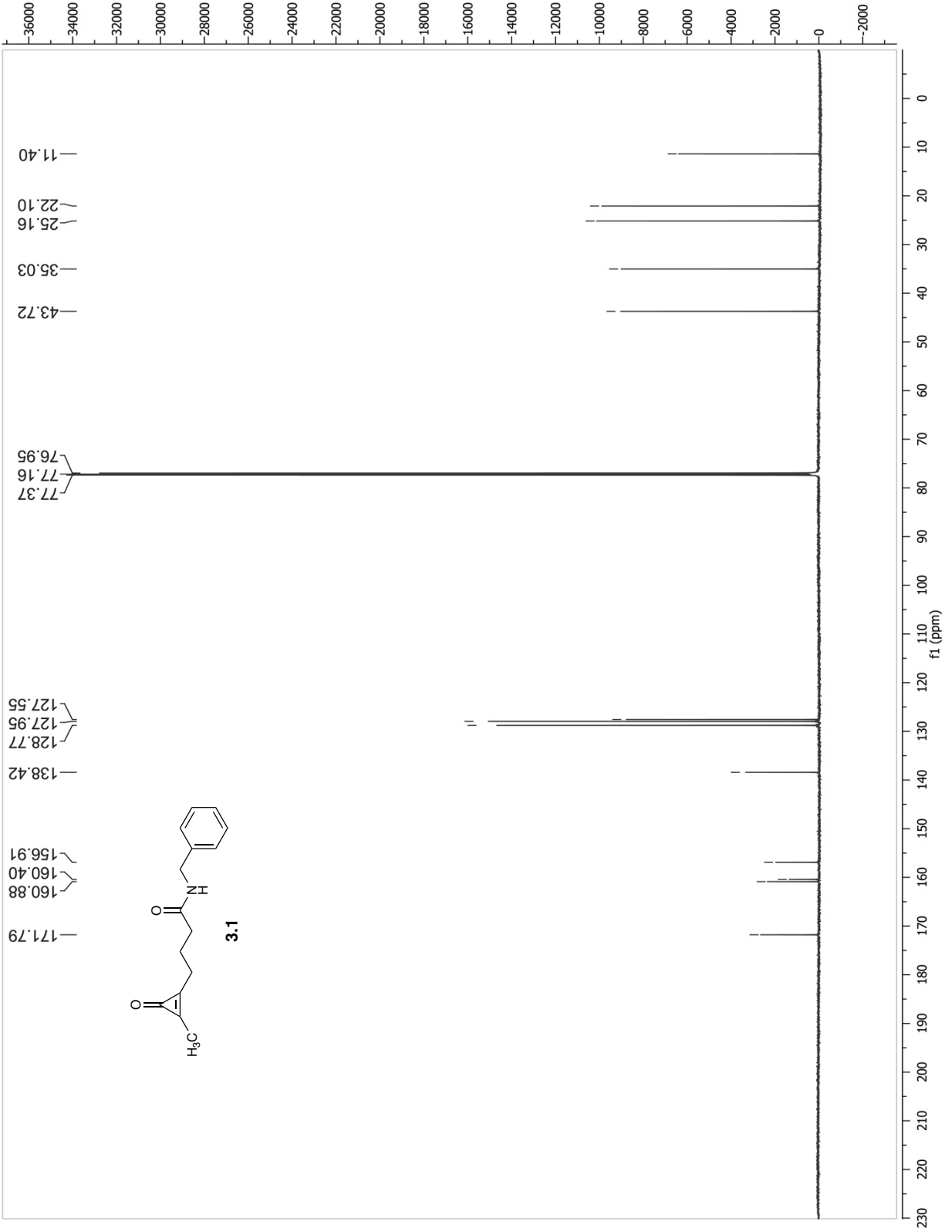
```

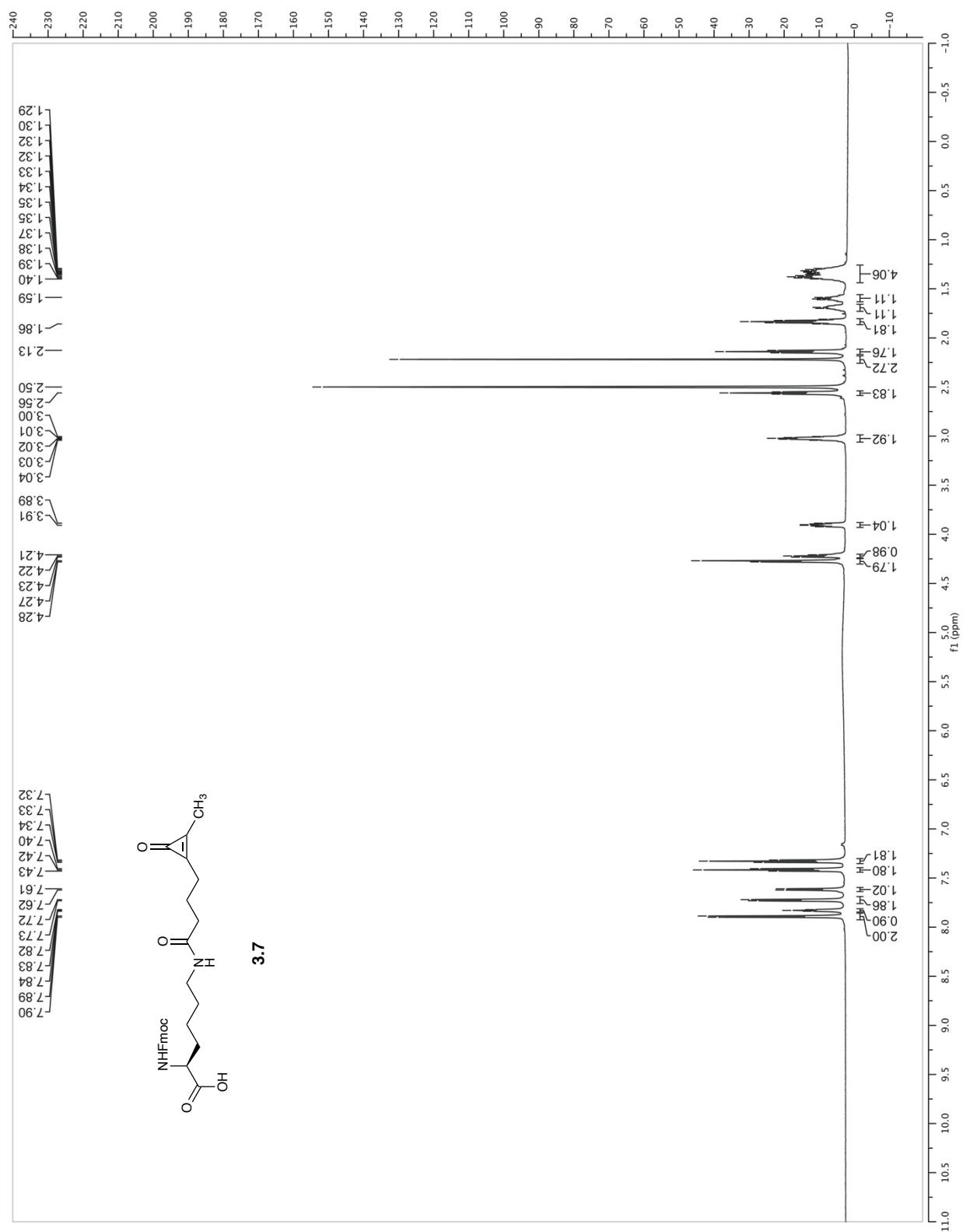
Current Data Parameters
=====
NAME          S13-110
EXPNO         1
PROCNO        1
Date_         20170707
Time          05:00
INSTRUM       5 mm CPBBO-BB-
PROBHD        5 mm
PULPROG       zgpg30
TD            65536
SOLVENT       DMSO
DS            2
SS            9615.385 Hz
SM          5.0000000 W
AQ          5.0998478 sec
RG          52.000 Hz
DE          13.70 usec
TE          300.00 K
D1          0.1000000 sec
TD0         1

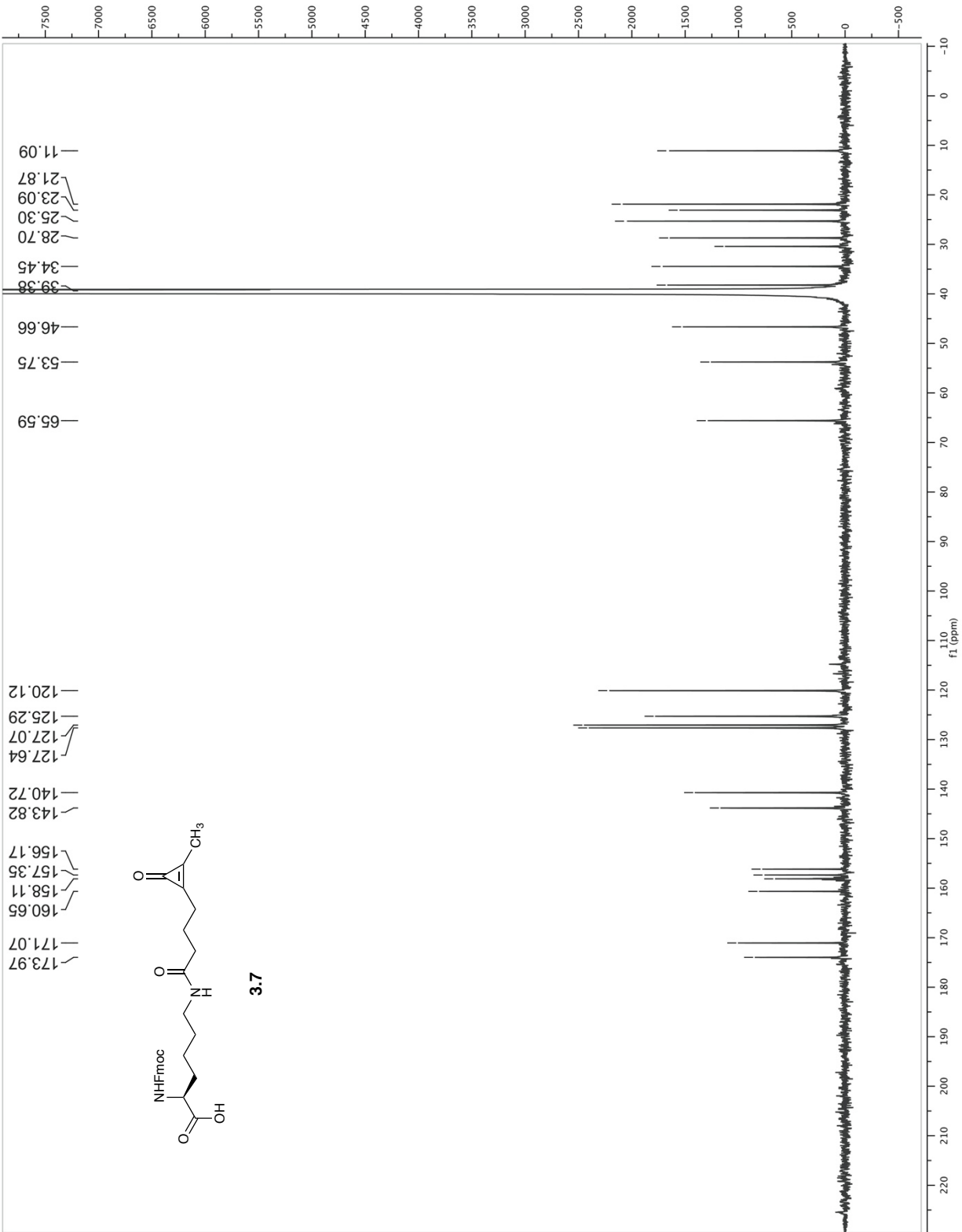
===== CHANNEL f1 =====
SFO1        600.1342019 MHz
NUC1         13C
PUL1        12.00 usec
PL1         20.0000000 W
F2 - Processing Parameters
SI          65536
RG          52.000 Hz
DE          13.70 usec
TE          300.00 K
D1          0.1000000 sec
D10         1.00
  
```

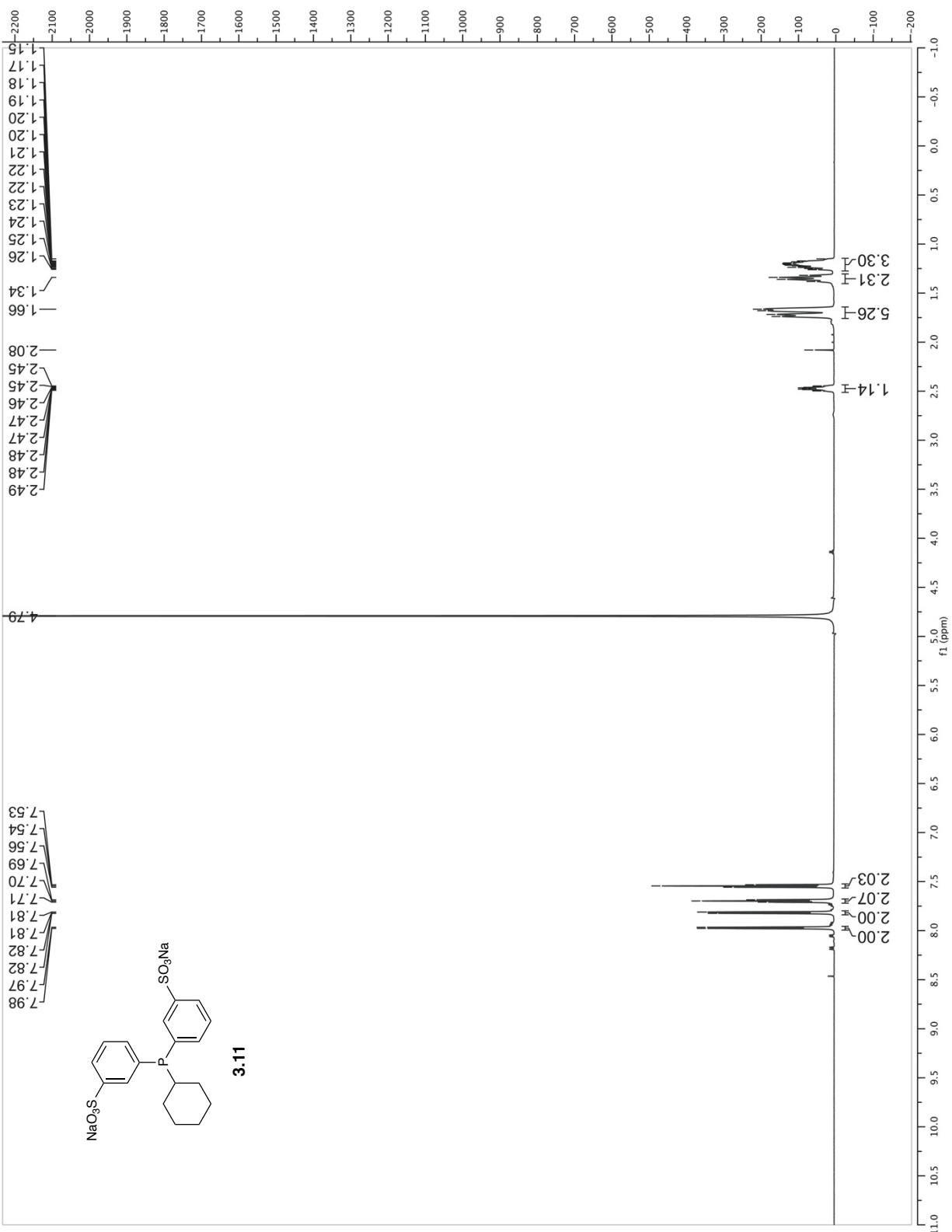


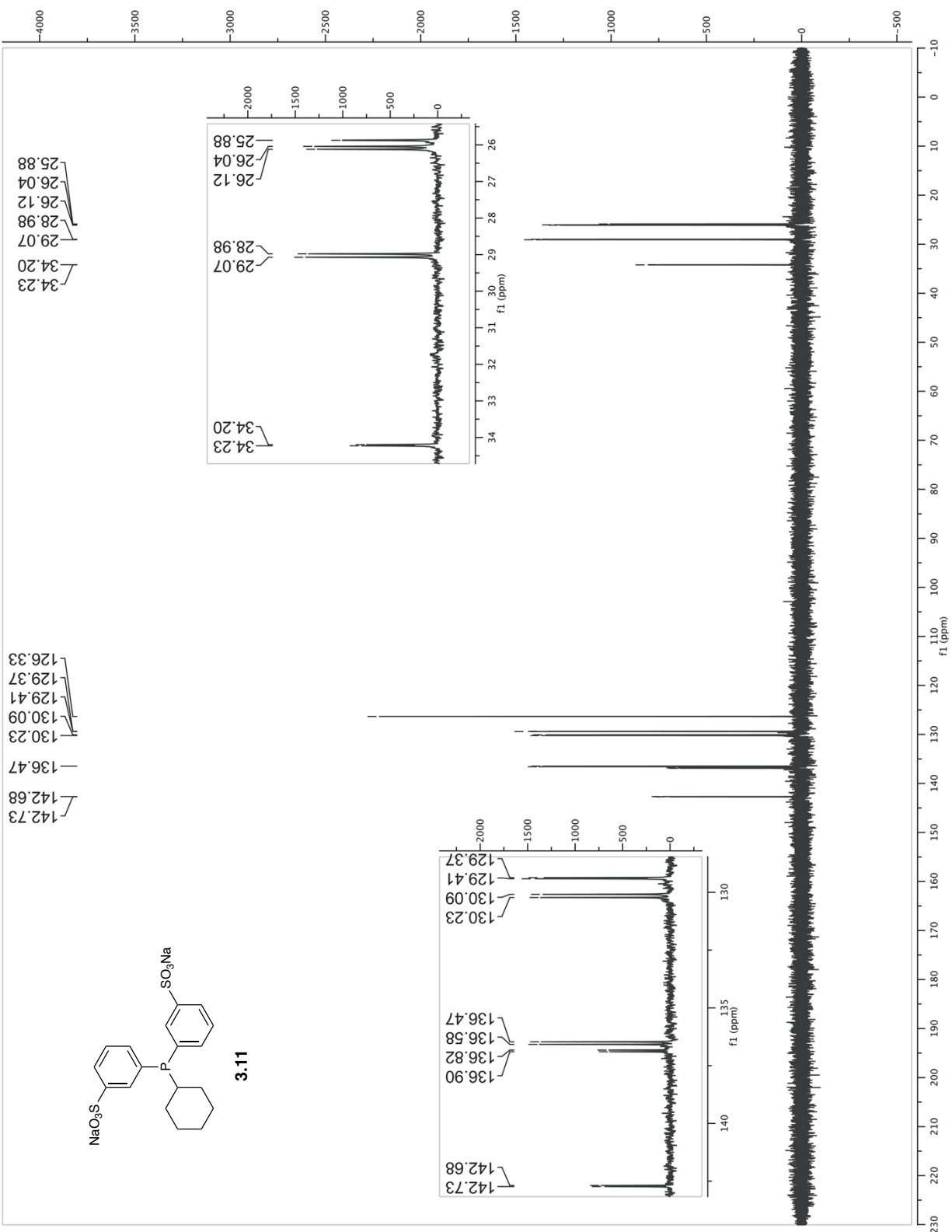


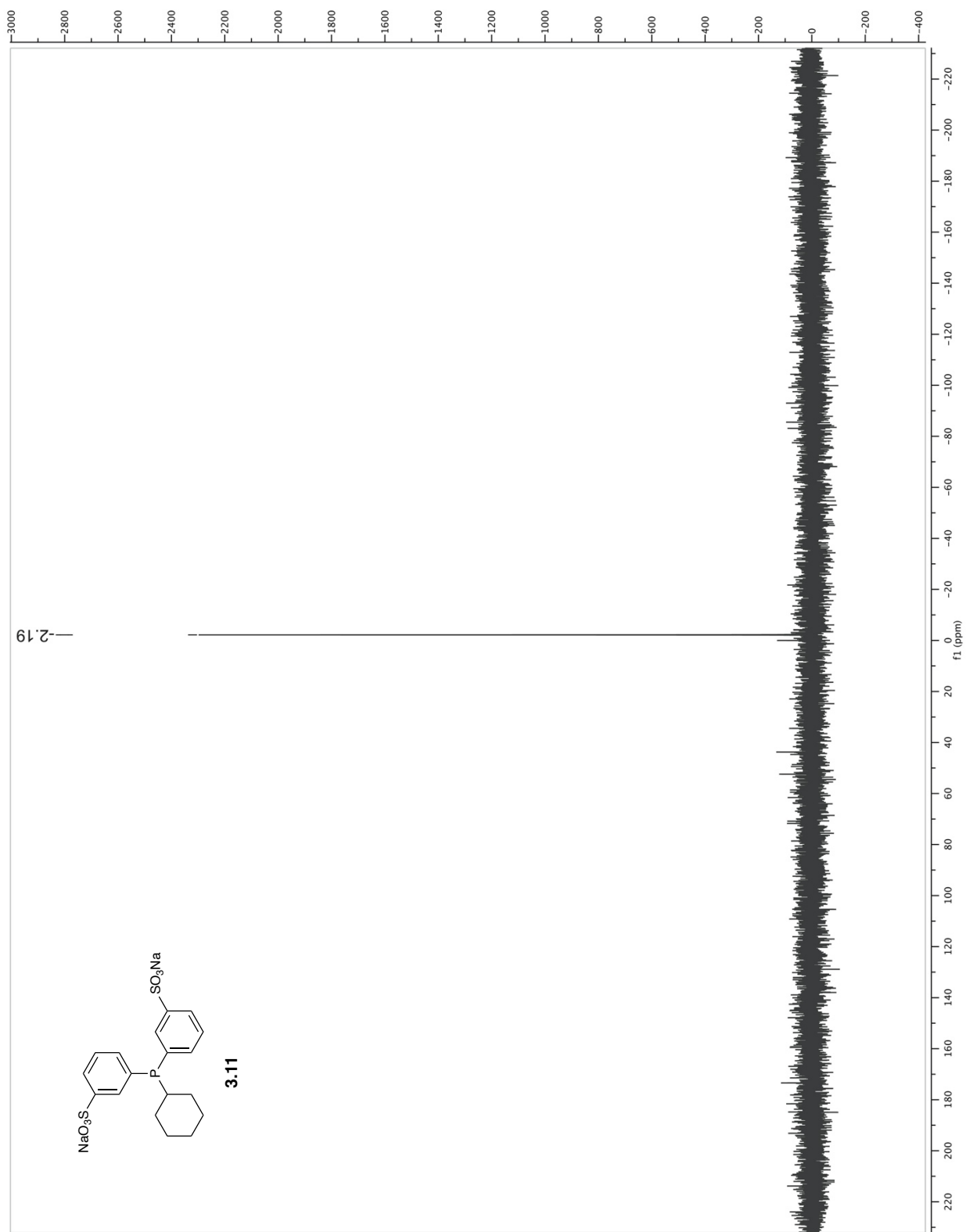


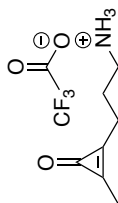
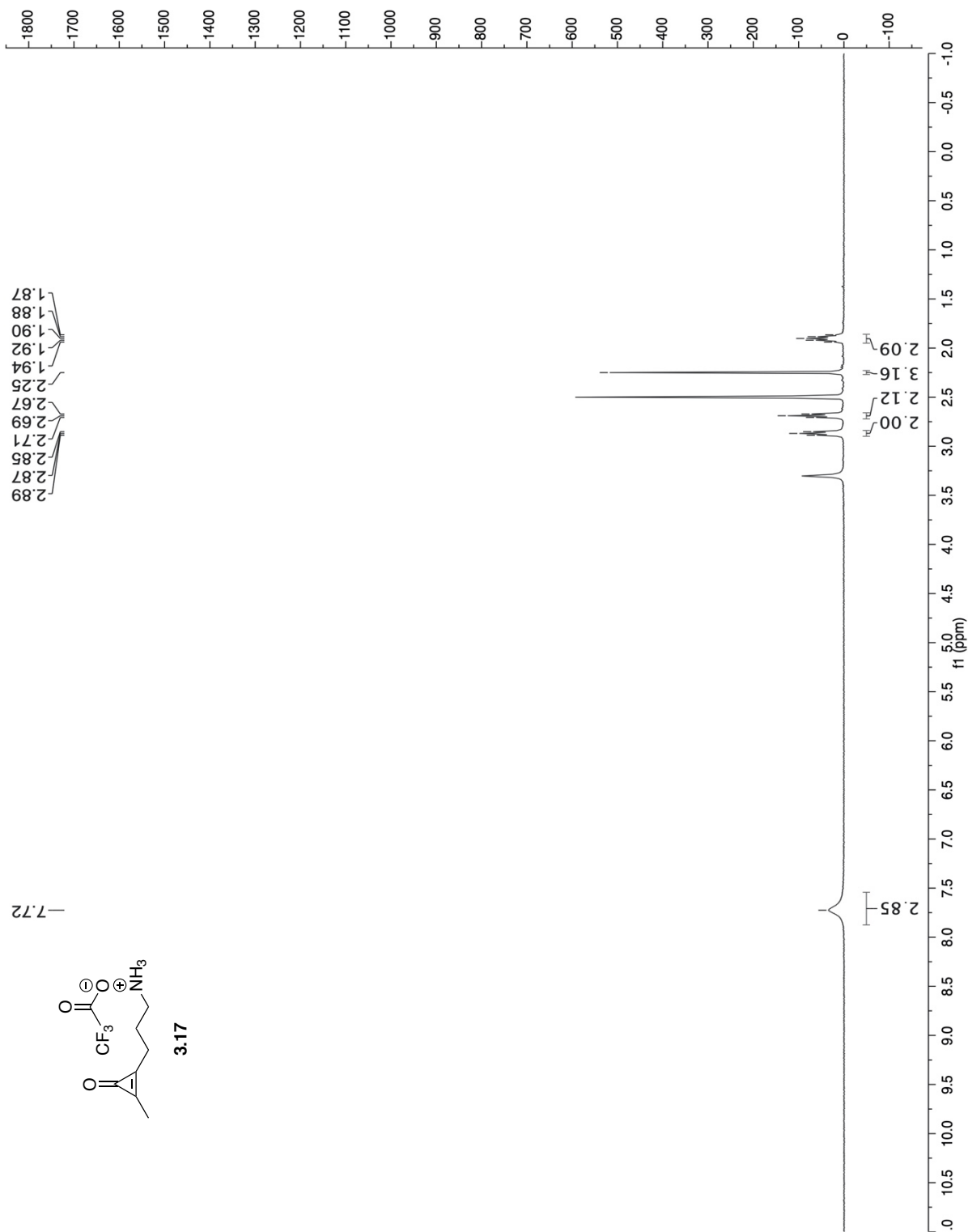




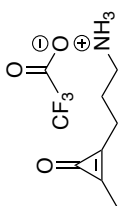
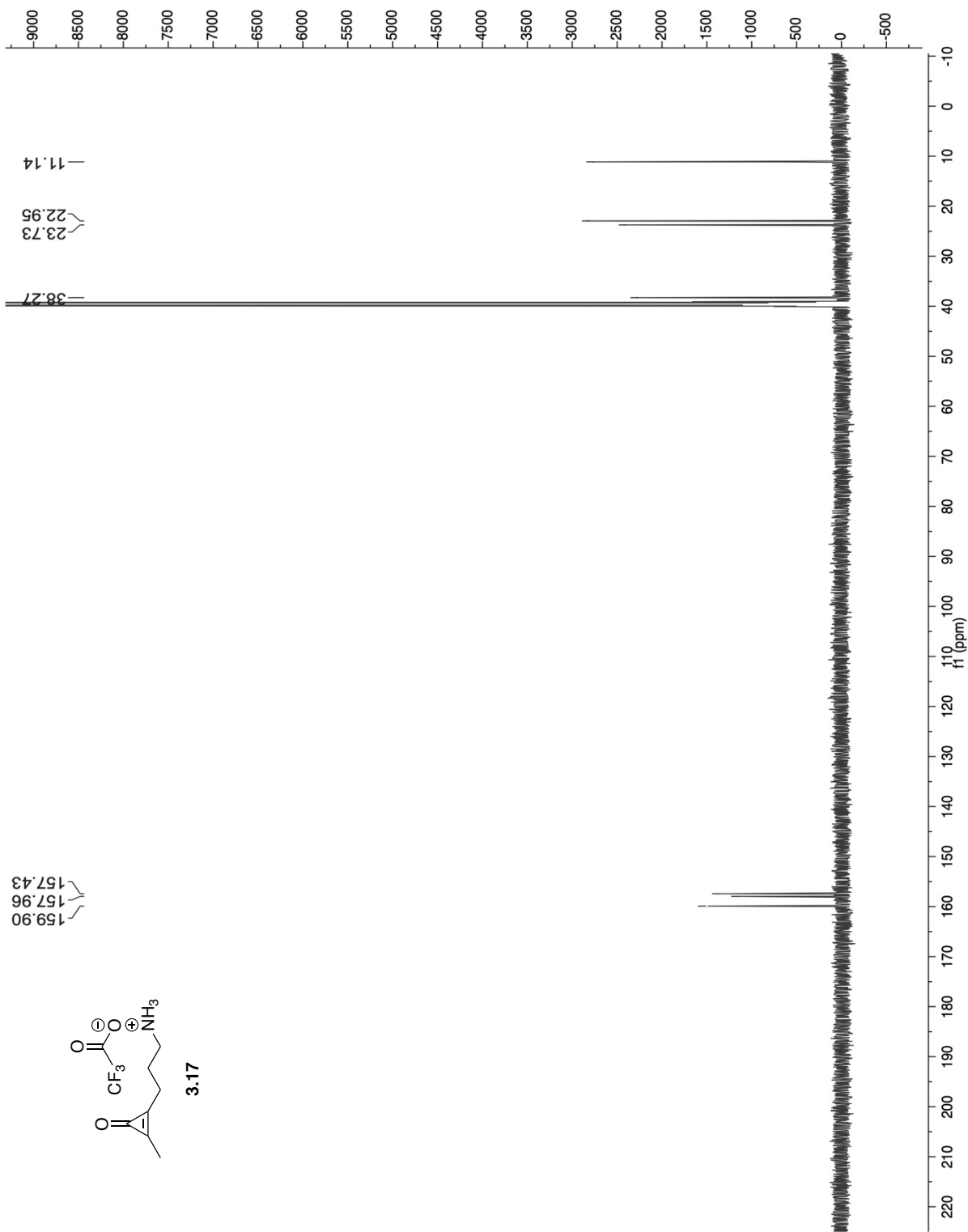




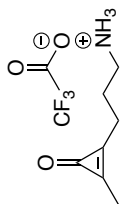




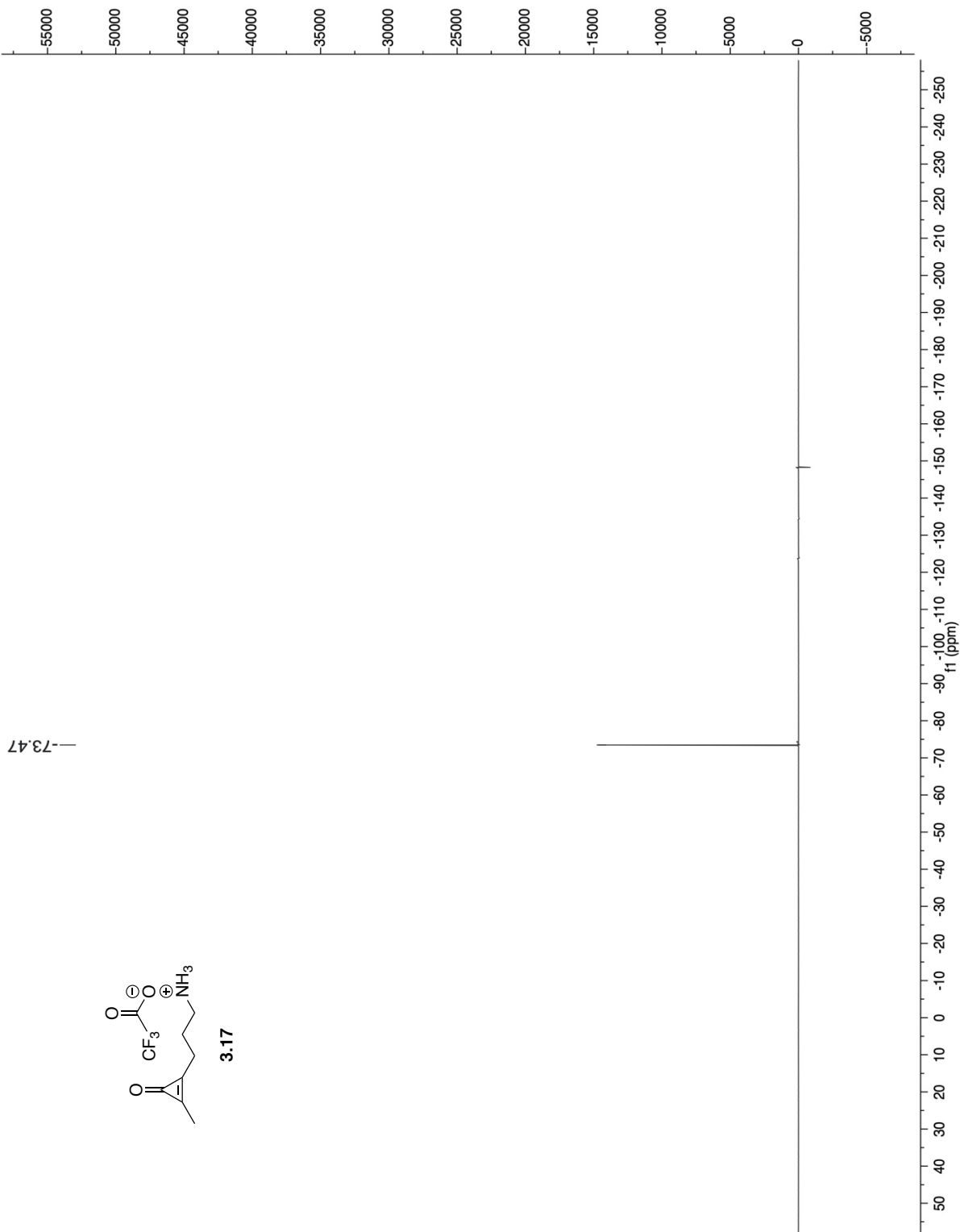
3.17

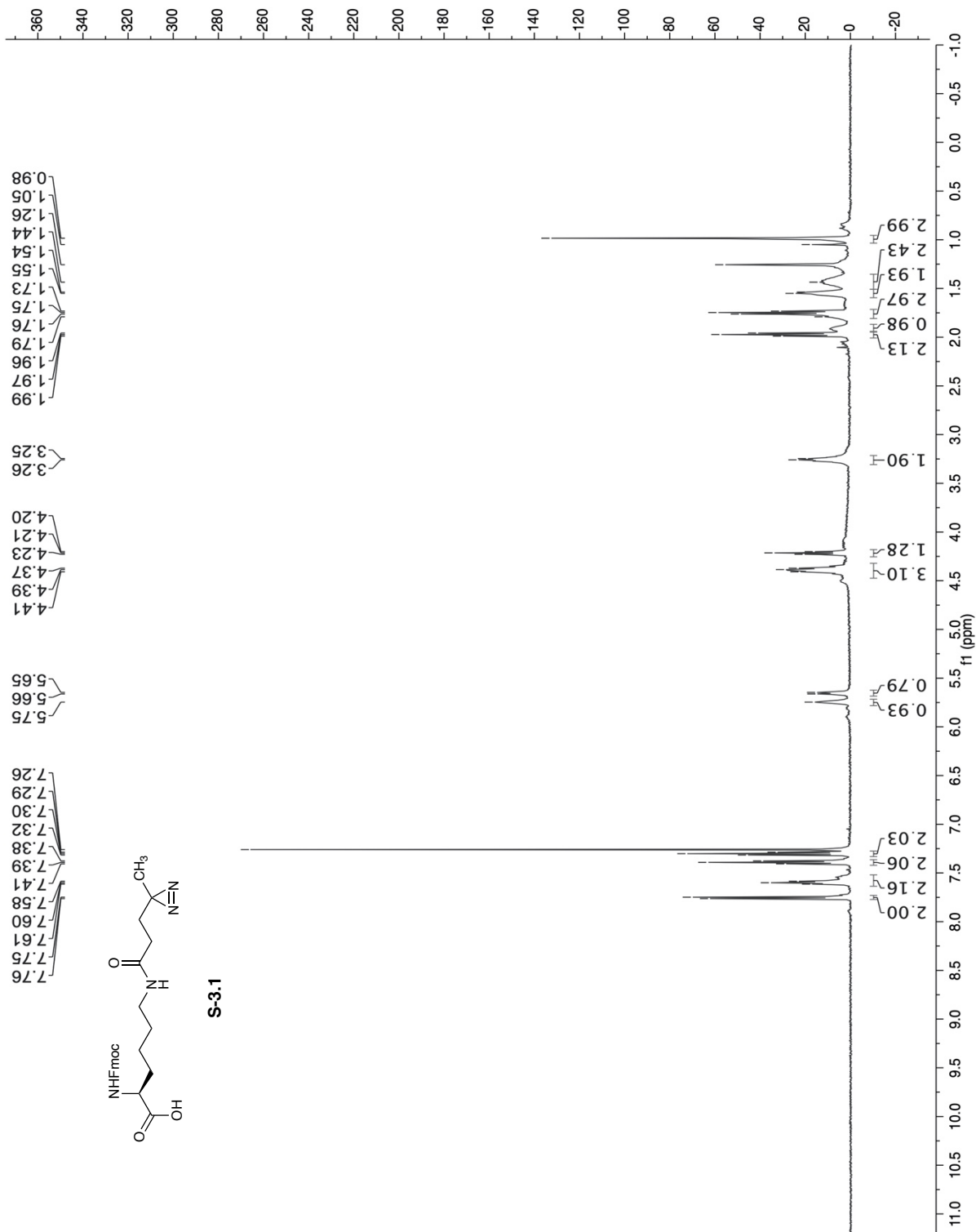


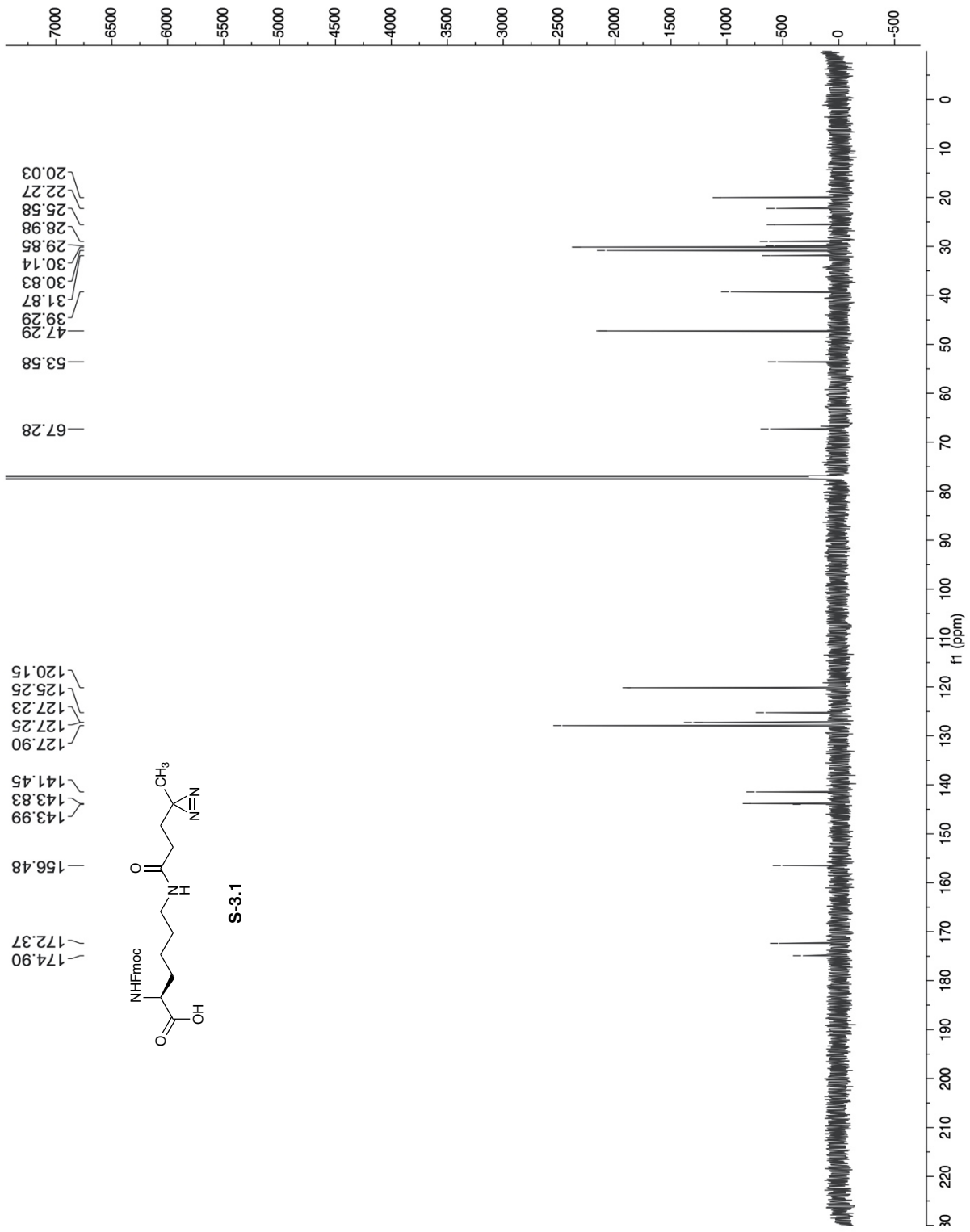
3.17

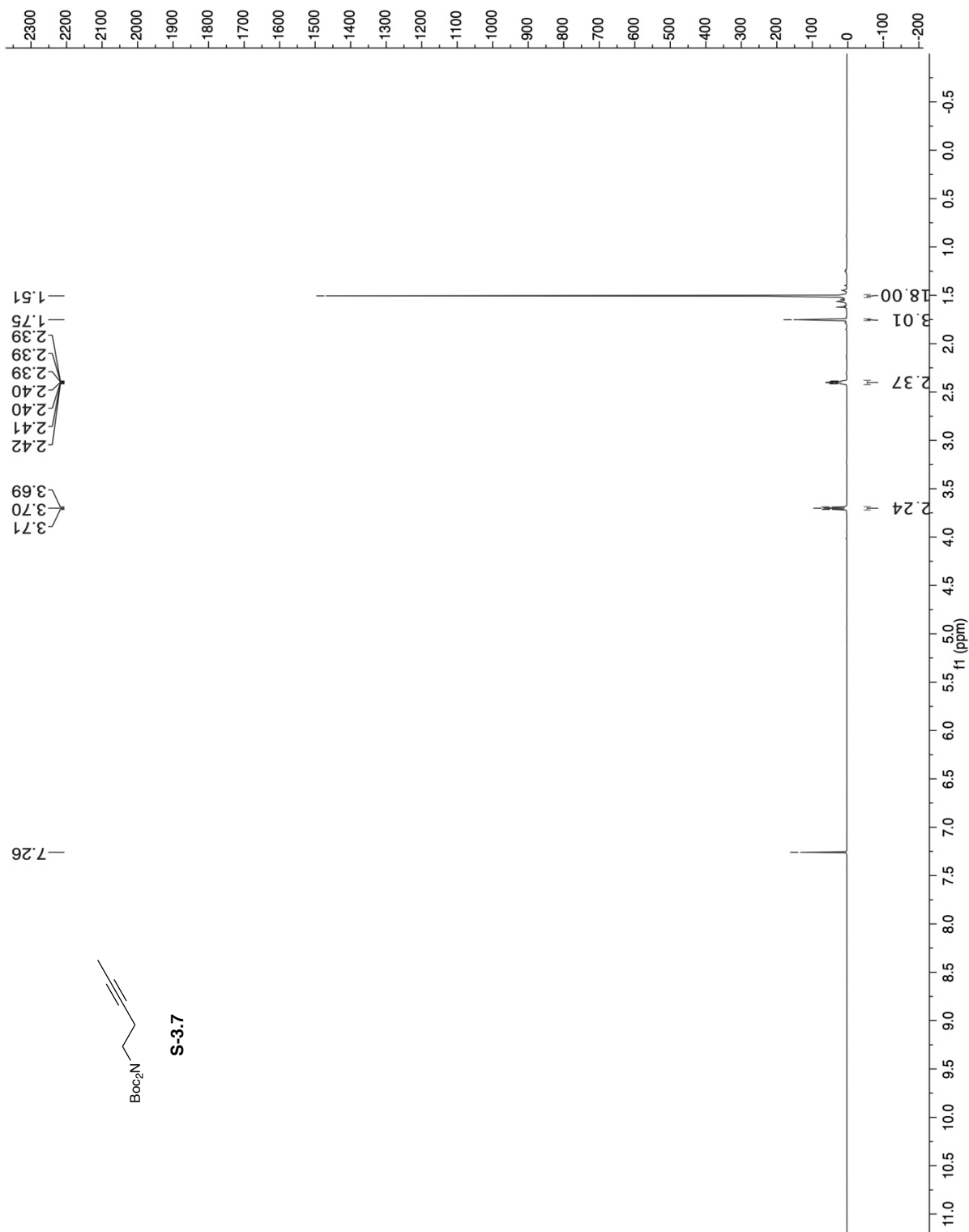


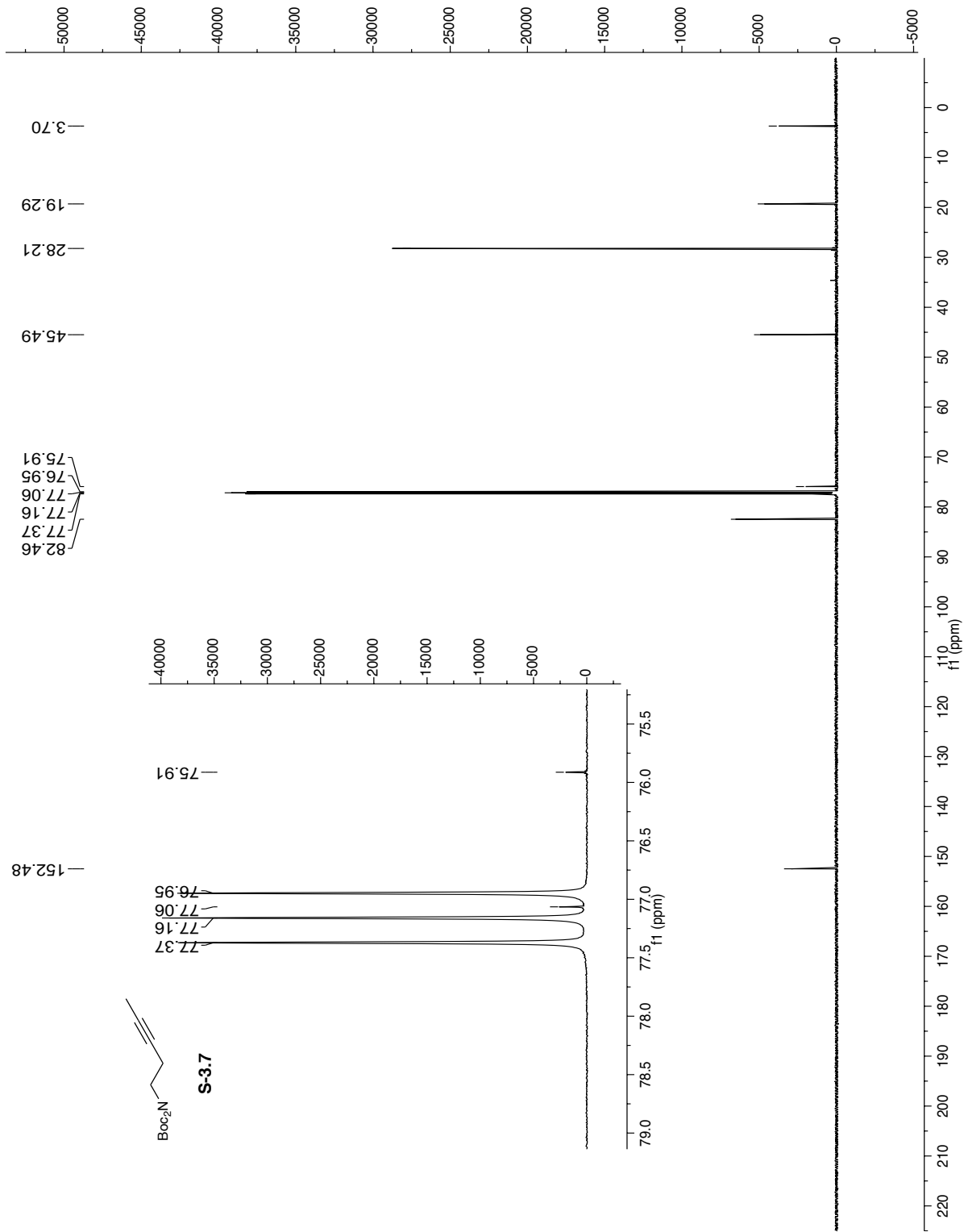
3.17

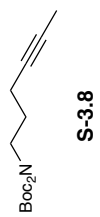
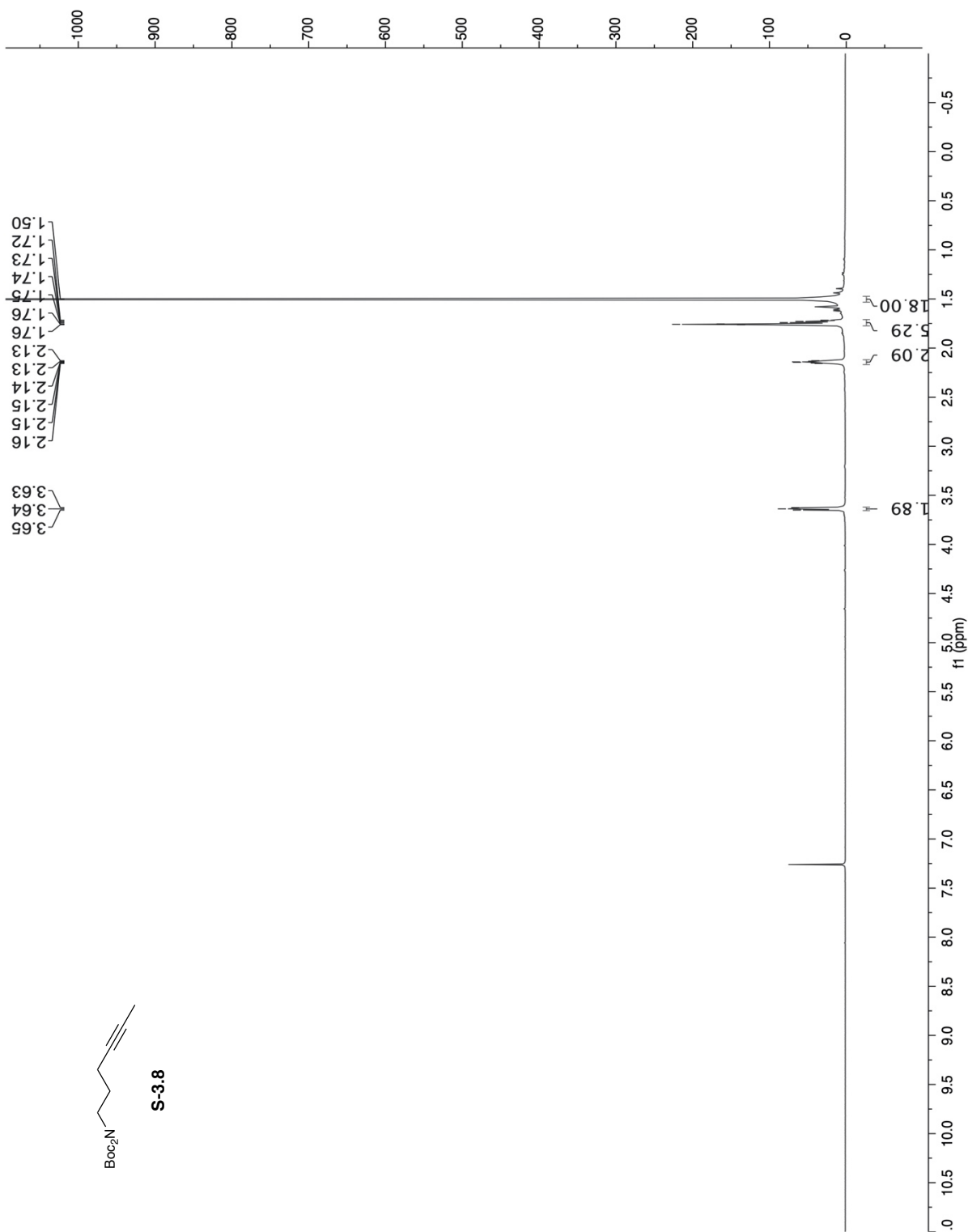


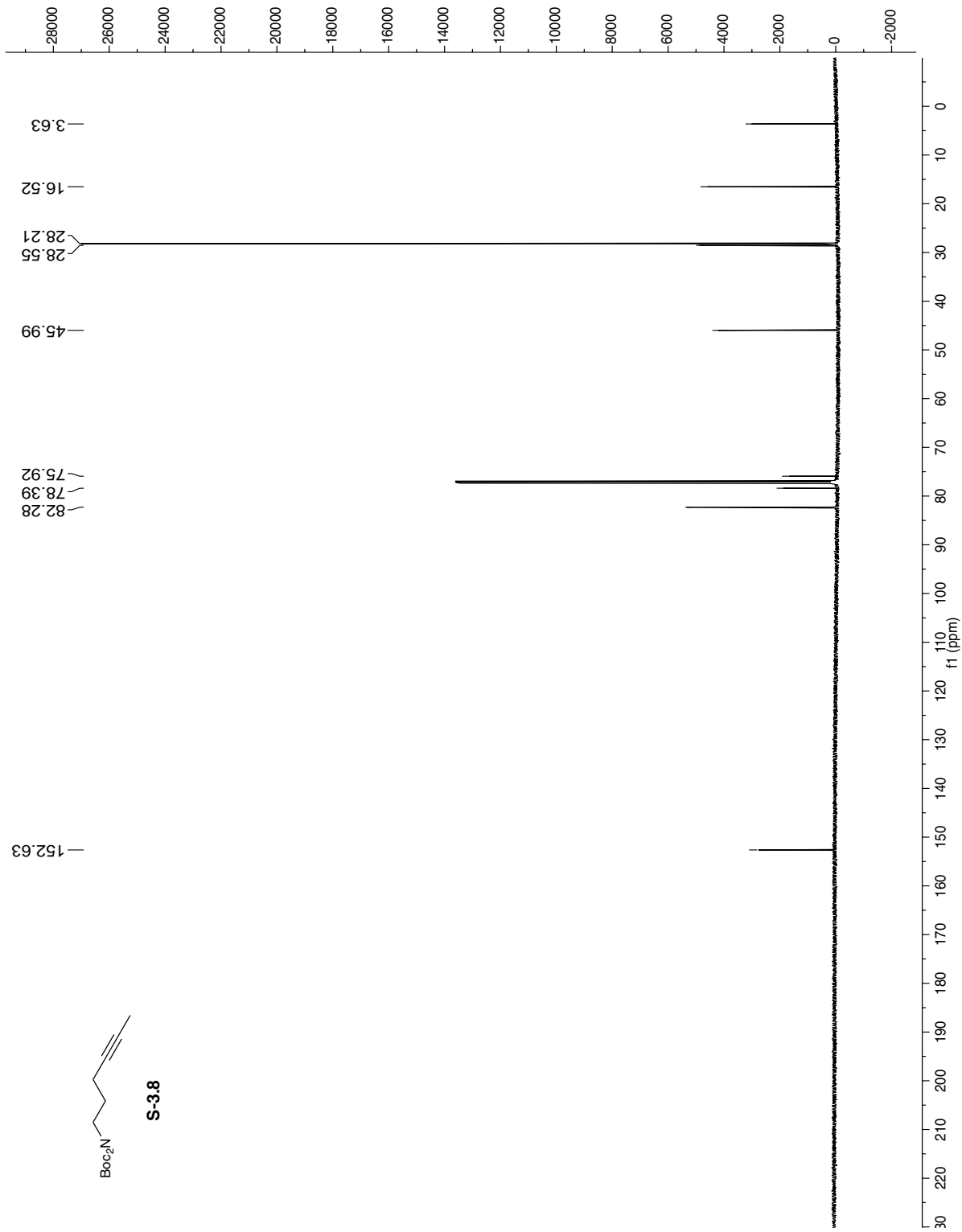


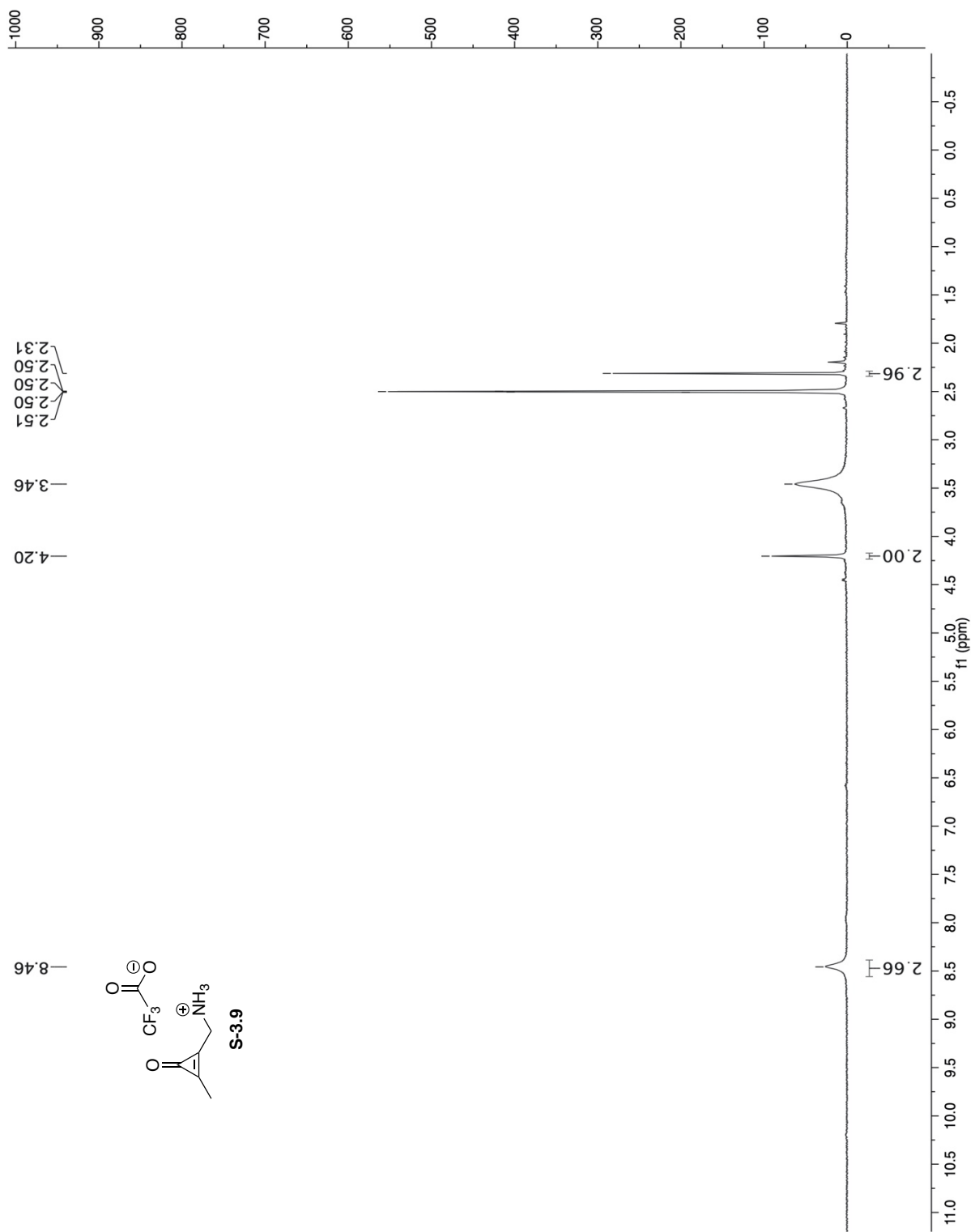


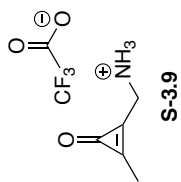
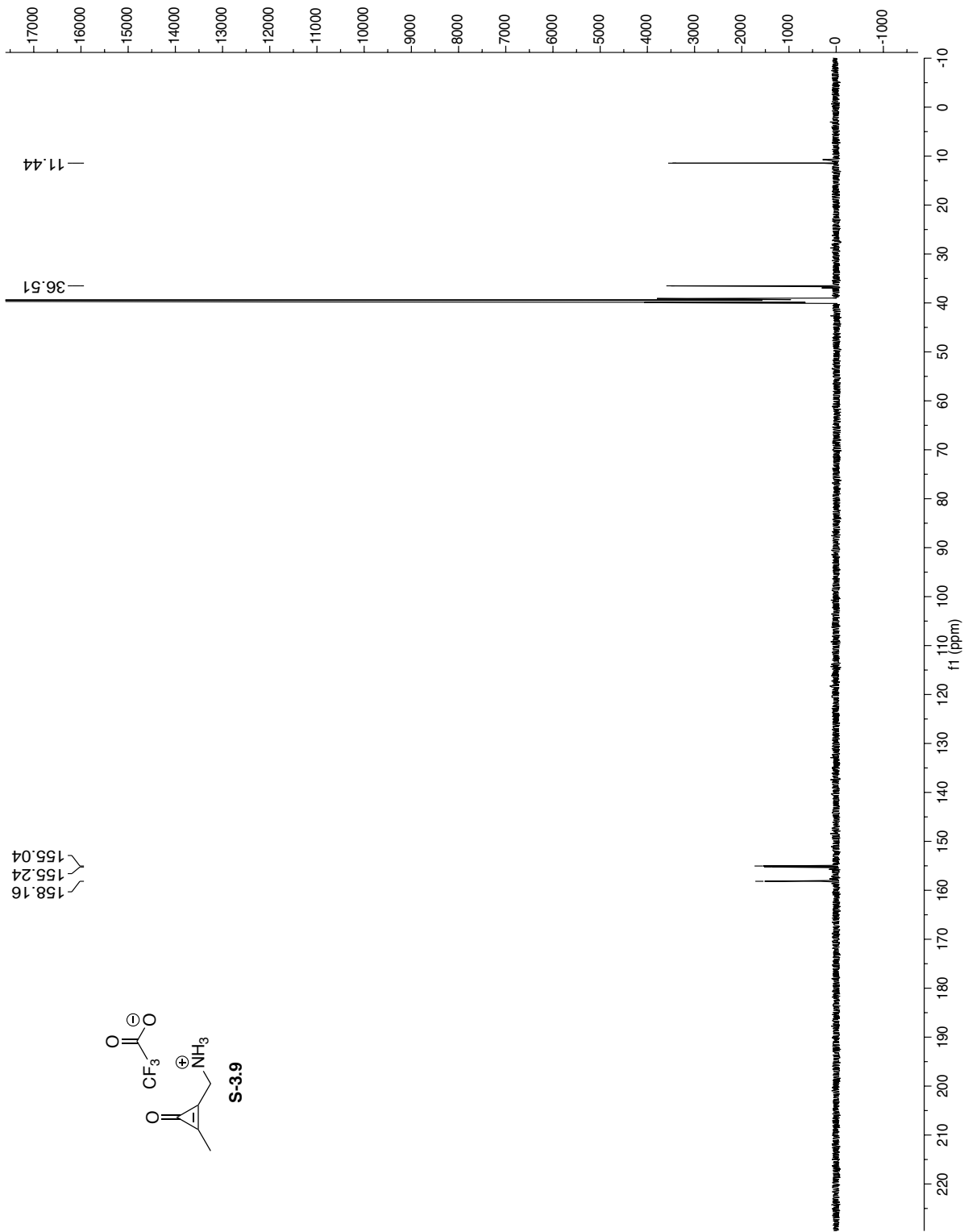


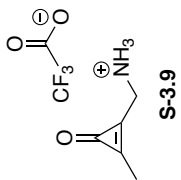
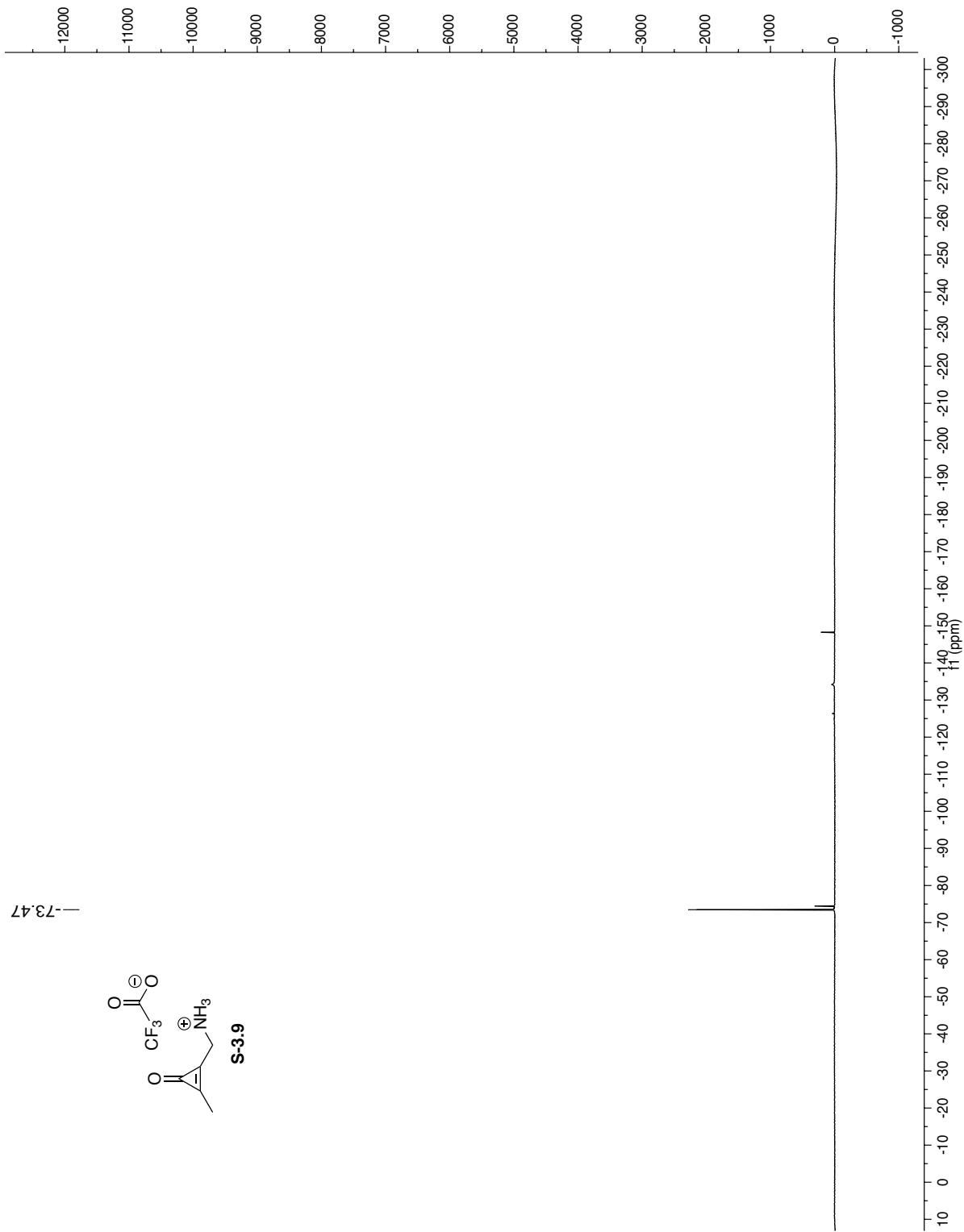


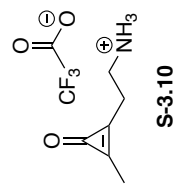
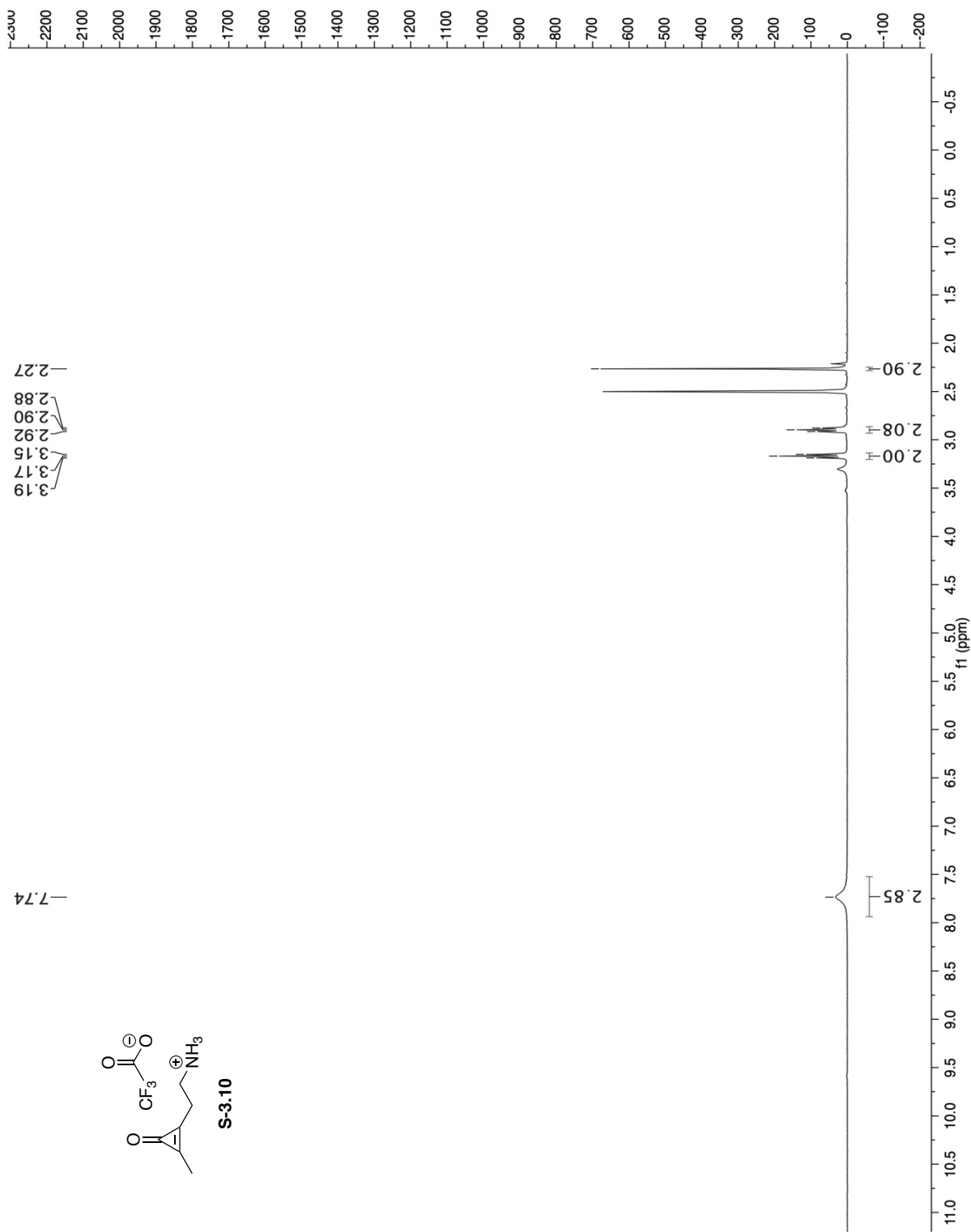


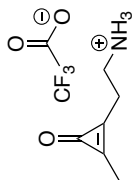
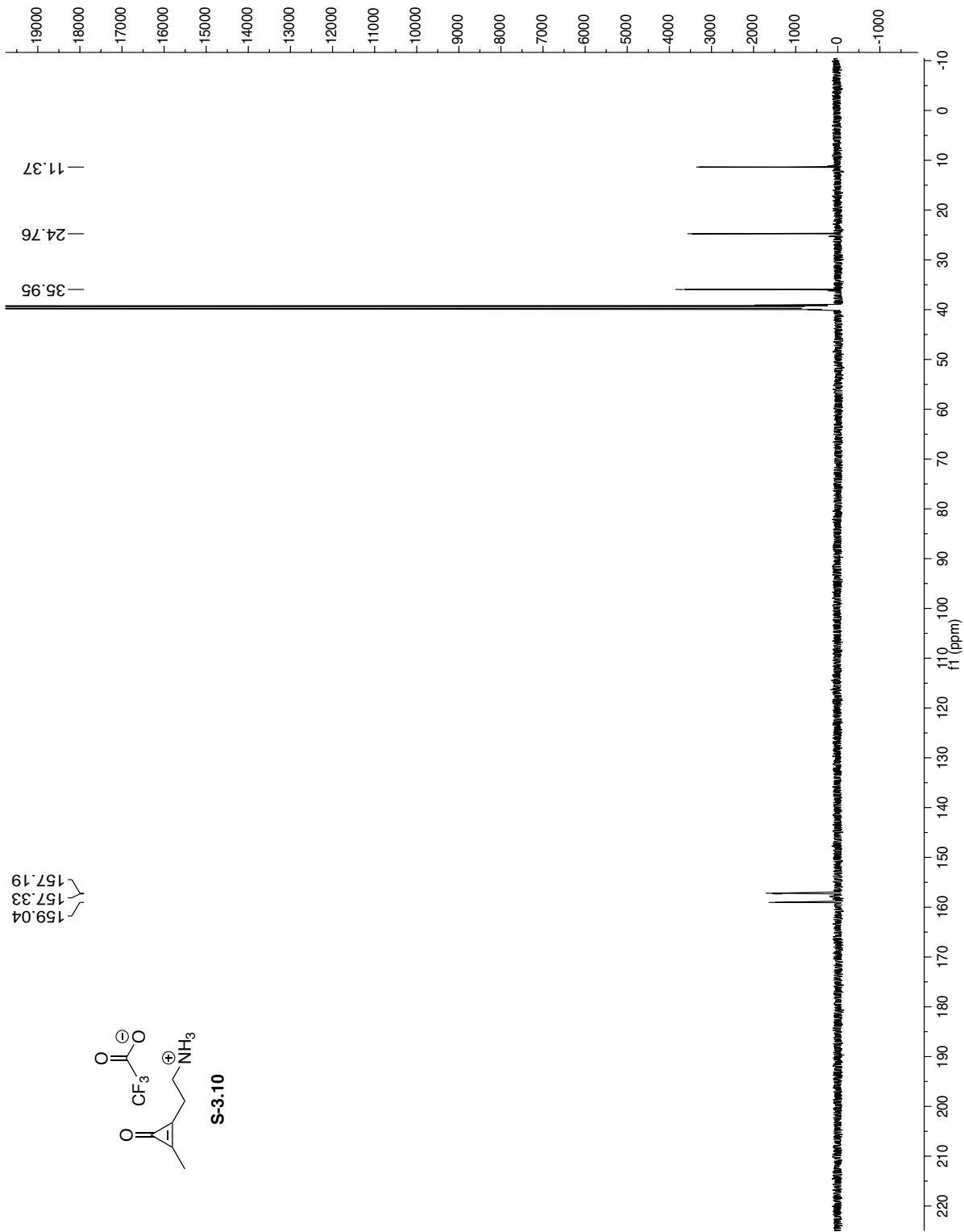




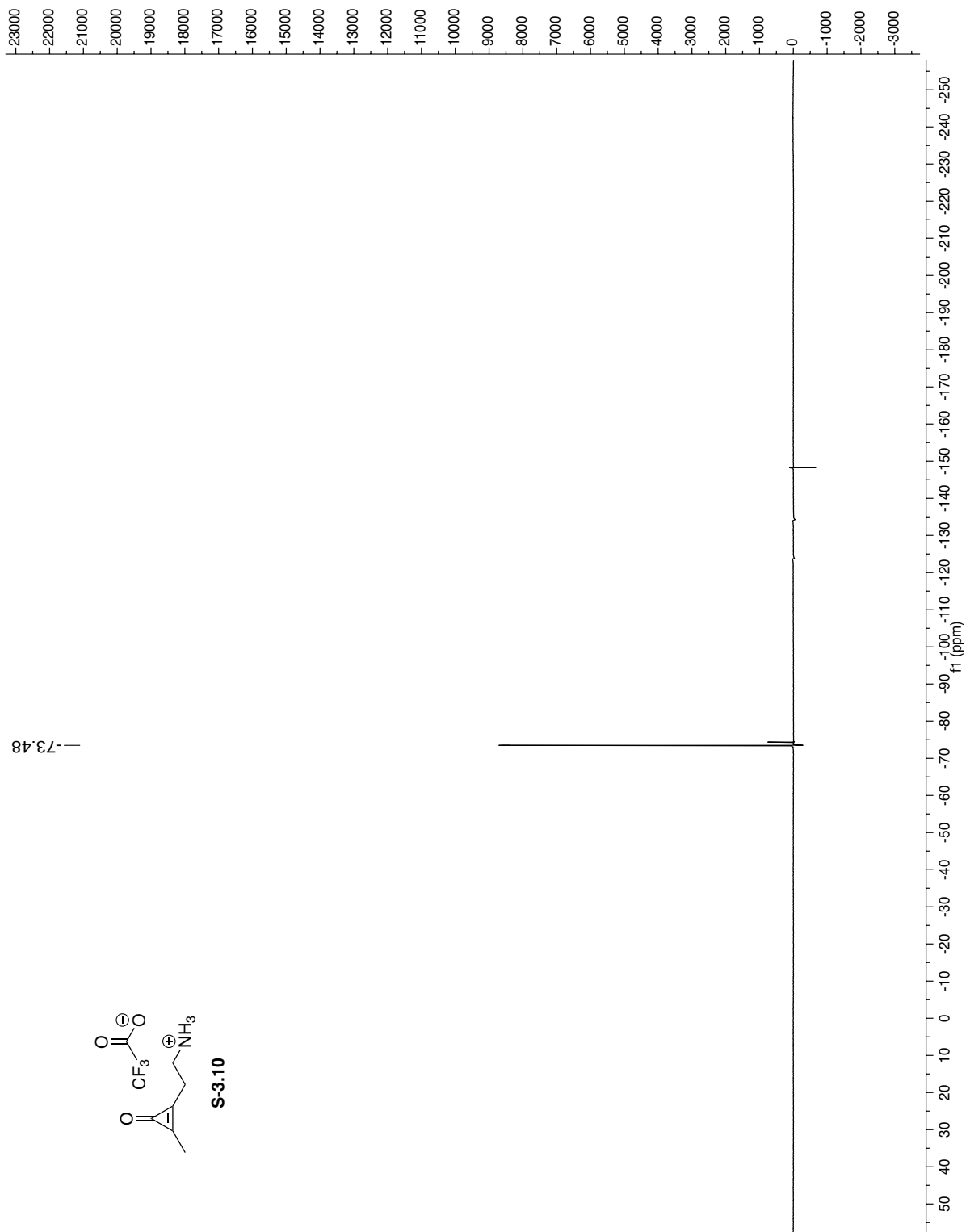








S-3.10



Appendix B: NMR Spectra for Chapter 4

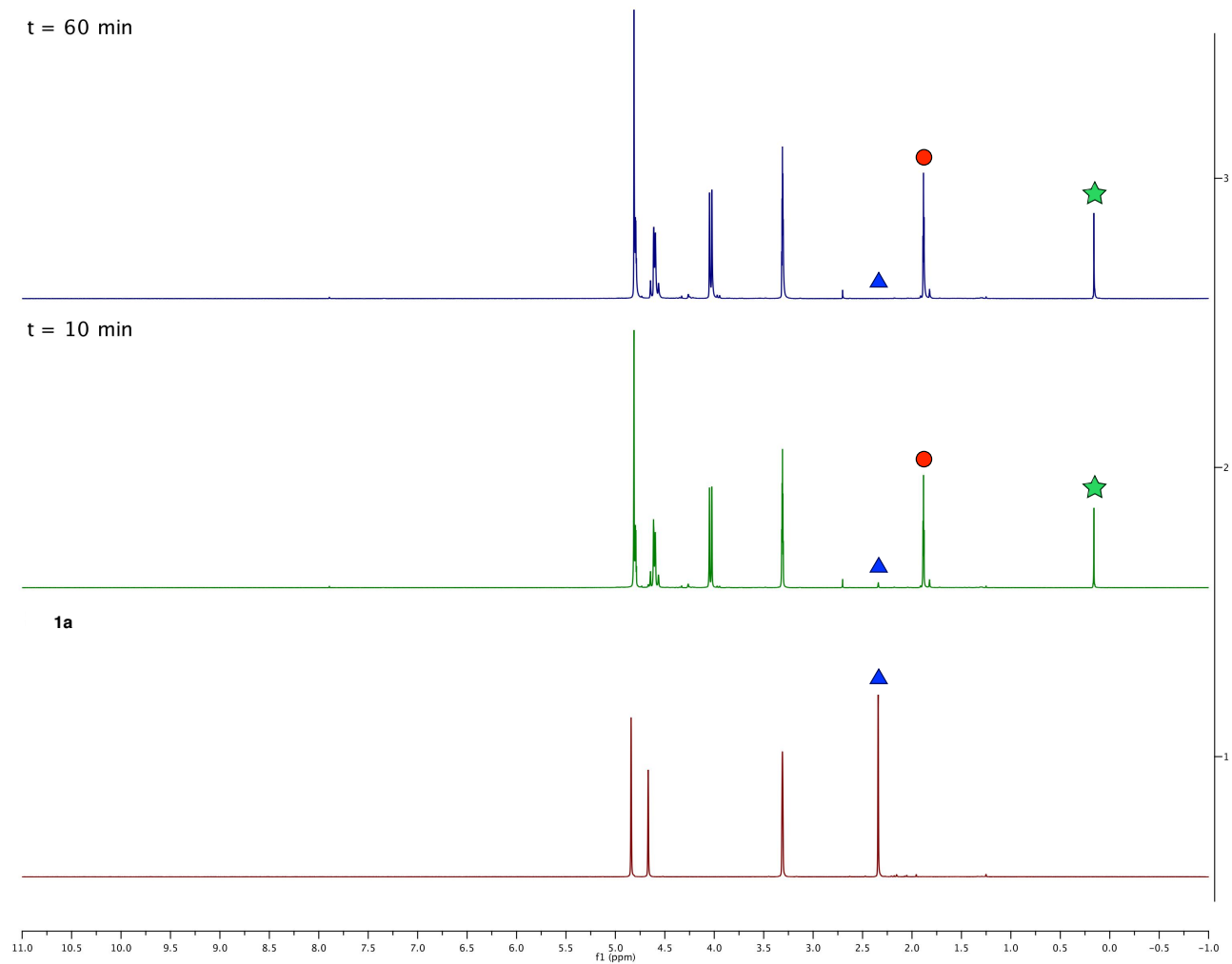
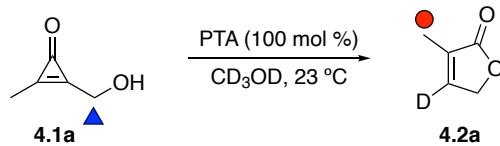


Figure S4-1. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PTA (100 mol %) in CD₃OD (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

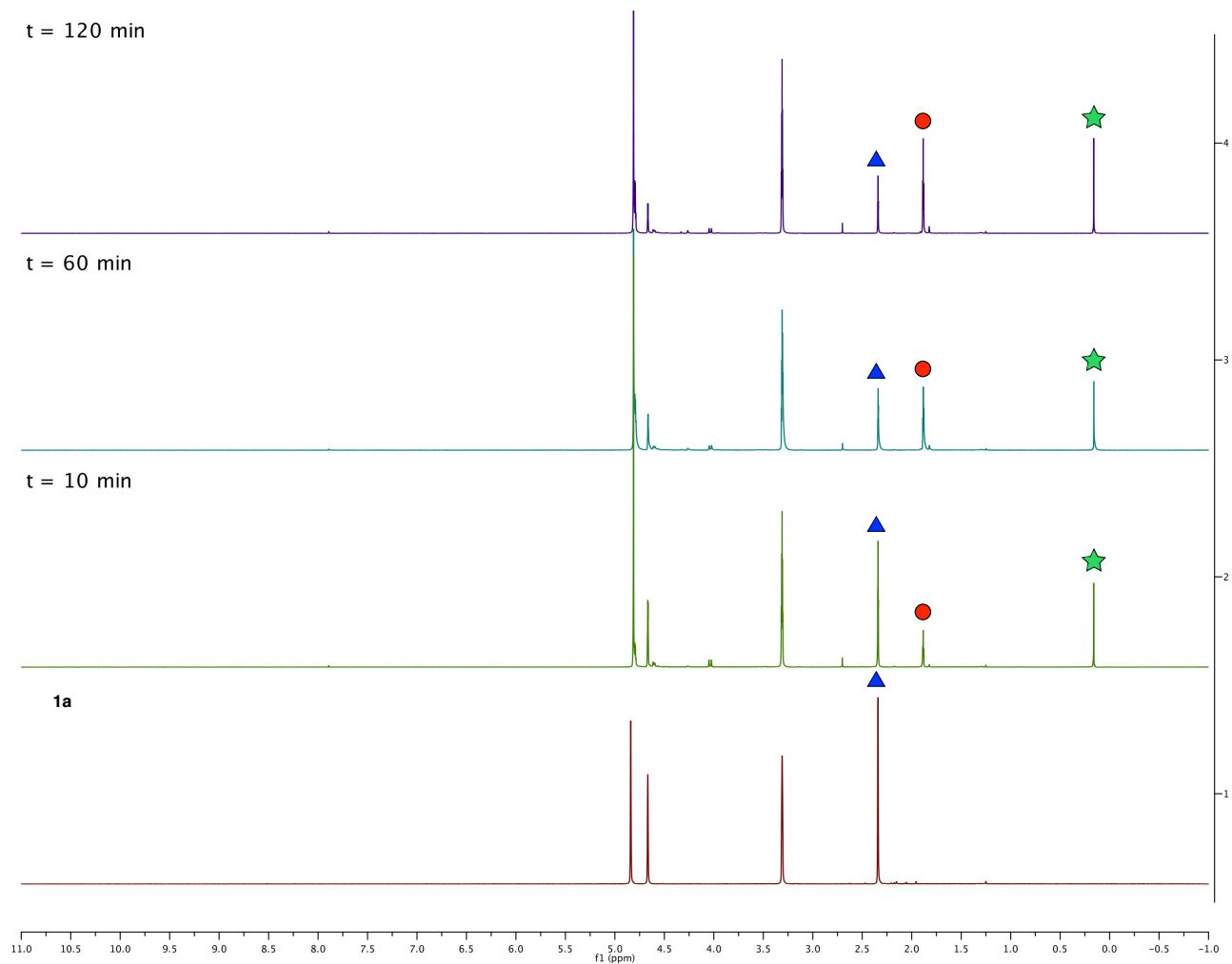
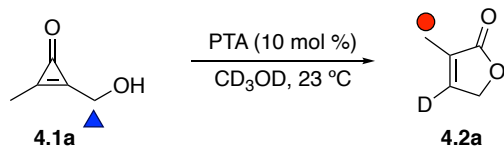


Figure S4-2. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PTA (10 mol %) in CD₃OD (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

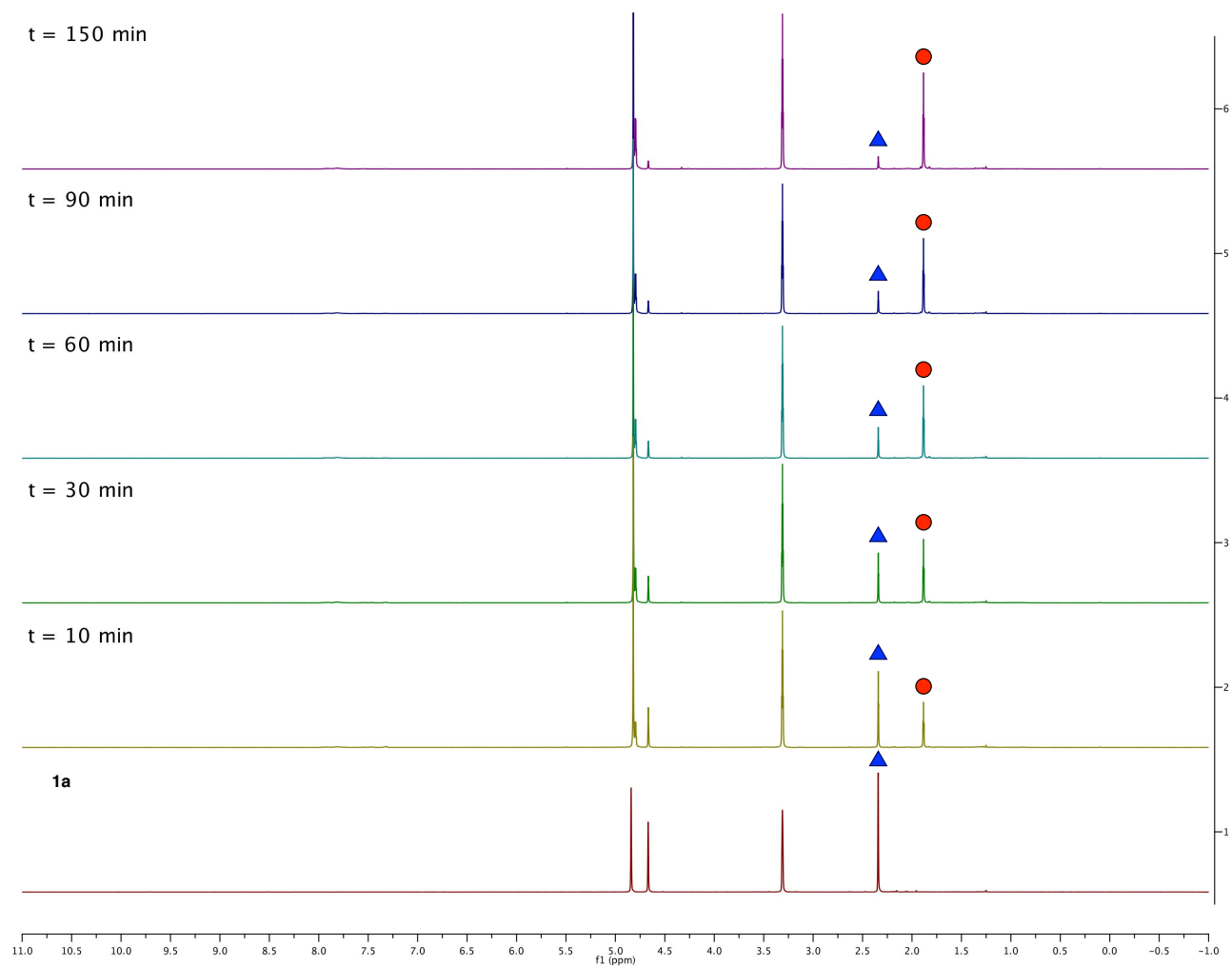
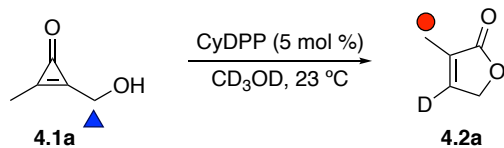


Figure S4-3. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with cyclohexyldiphenylphosphine (CyDPP, 5 mol %) in CD₃OD (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle).

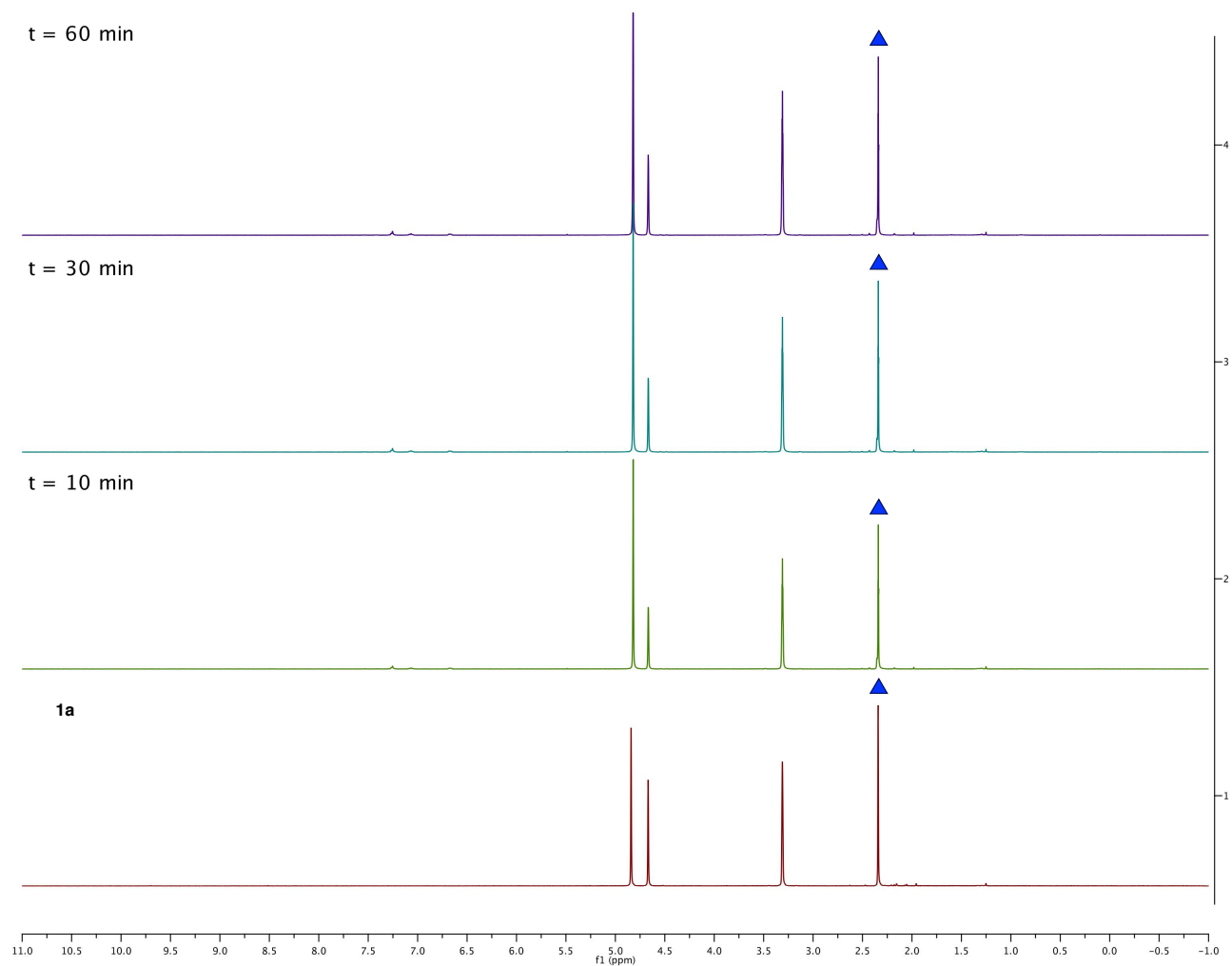
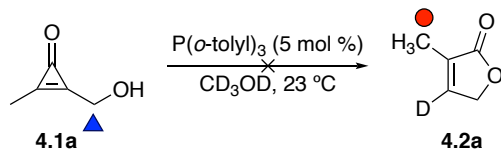


Figure S4-4. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with P(*o*-tolyl)₃ (5 mol %) in CD₃OD (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy. No conversion to butenolide **4.2a** was observed.

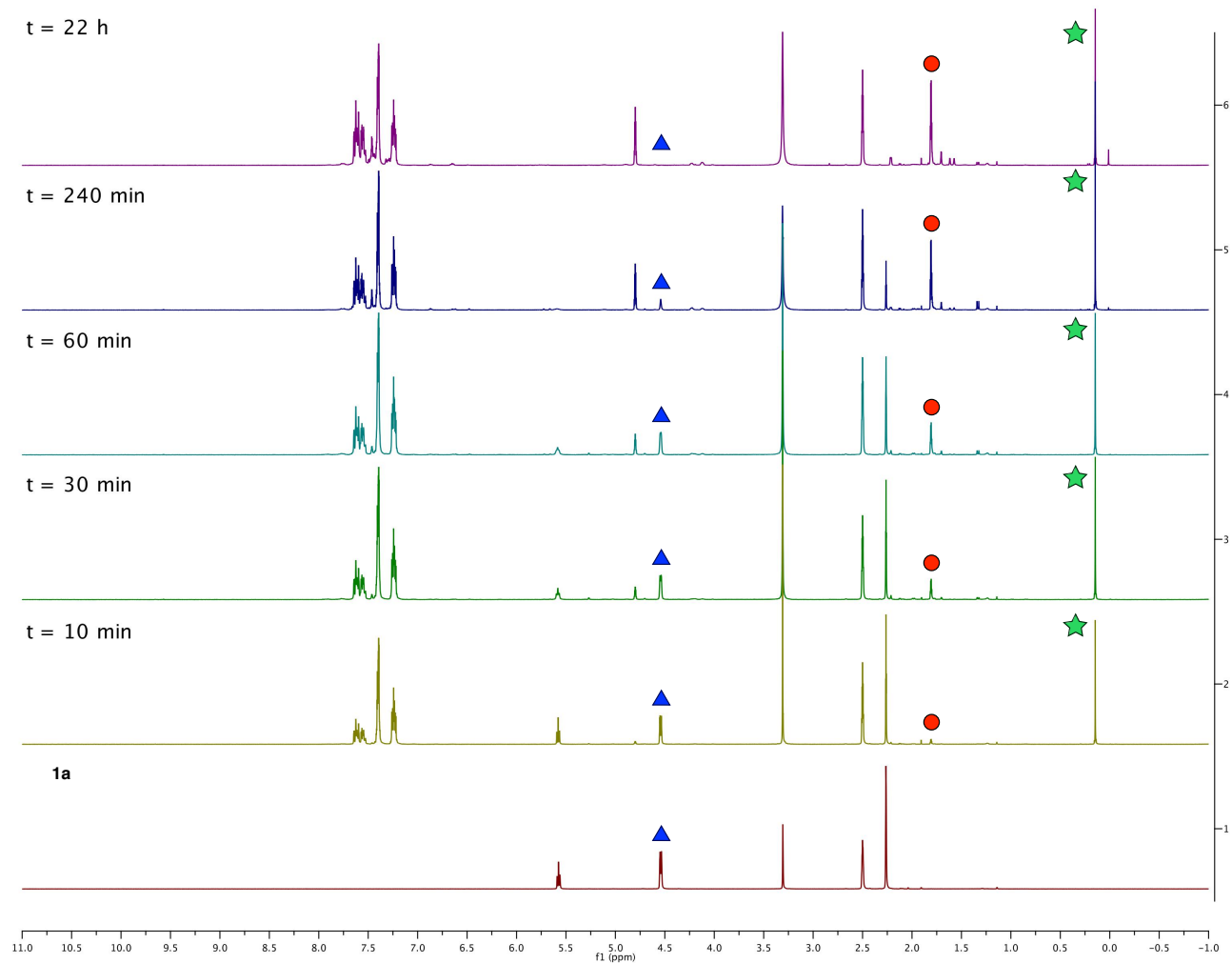
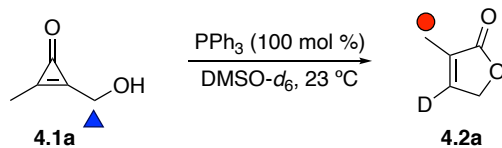


Figure S4-5. Solvent screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PPh₃ (100 mol %) in DMSO-*d*₆ (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

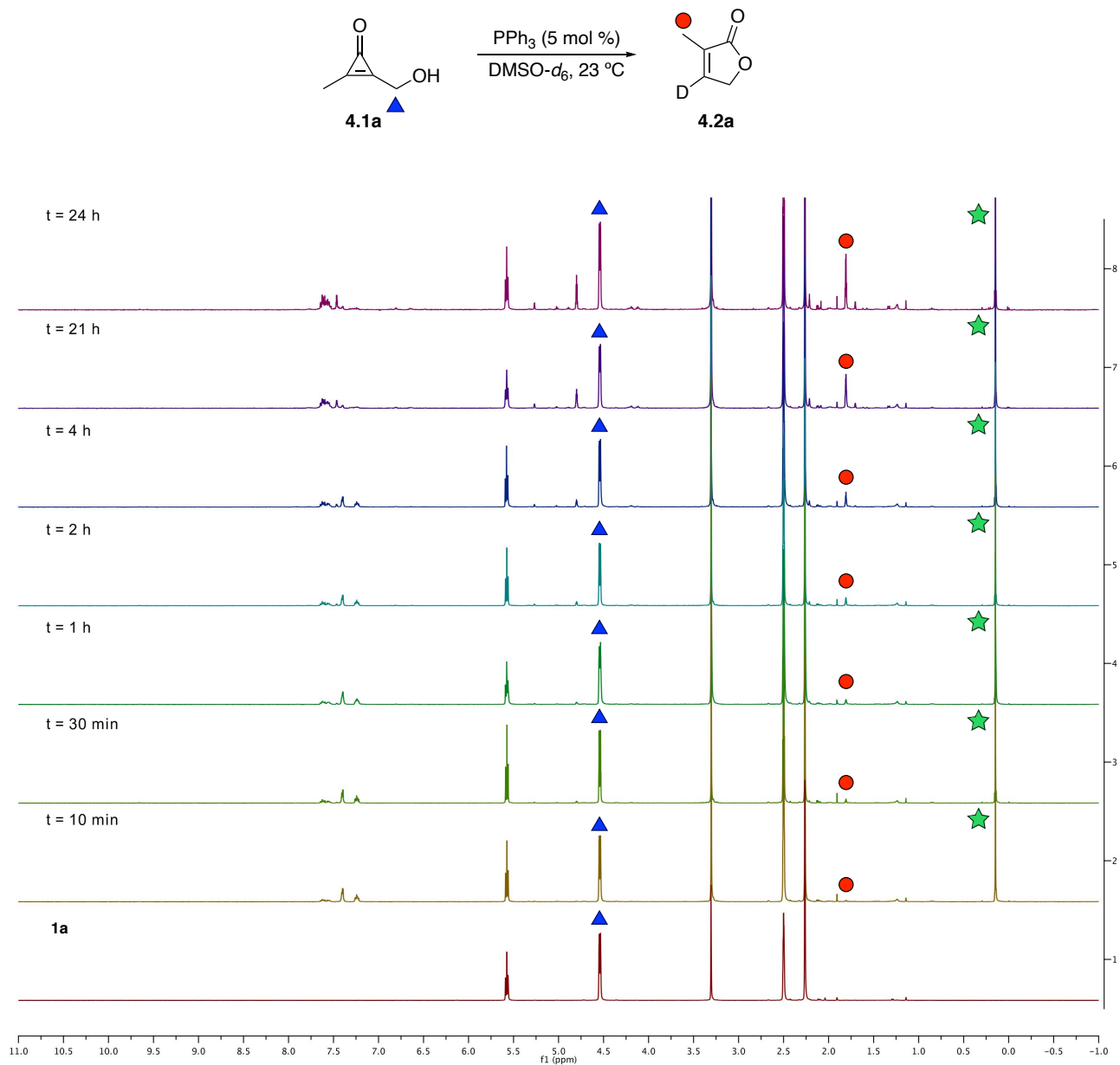


Figure S4-6. Solvent screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PPh_3 (5 mol %) in $\text{DMSO-}d_6$ (600 μL) at ambient temperature. The reaction was monitored periodically by ^1H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

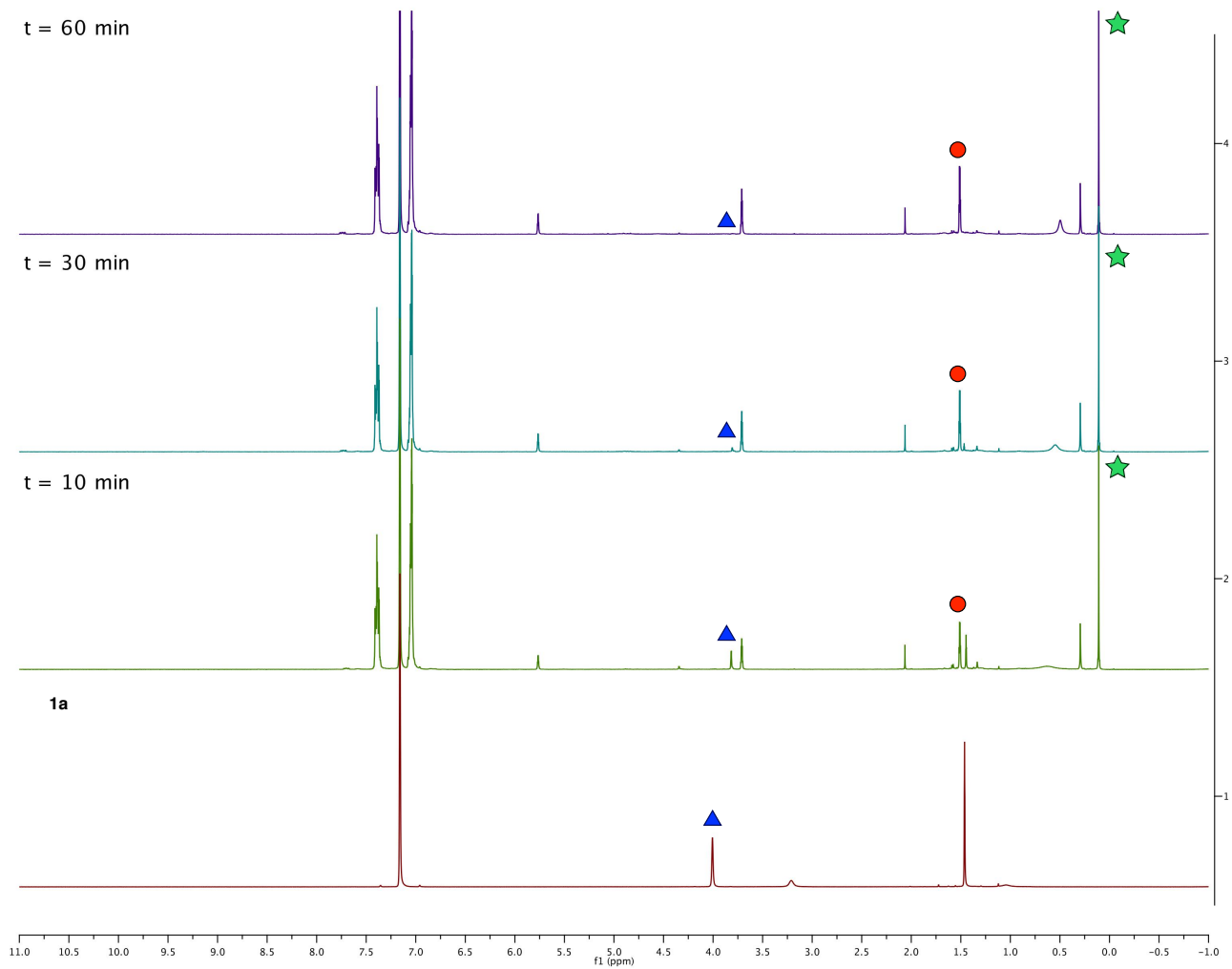
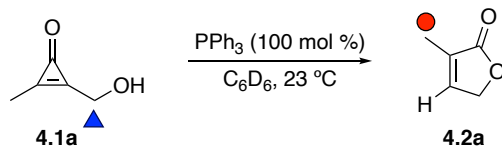


Figure S4-7. Solvent screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PPh₃ (100 mol %) in C₆D₆ (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

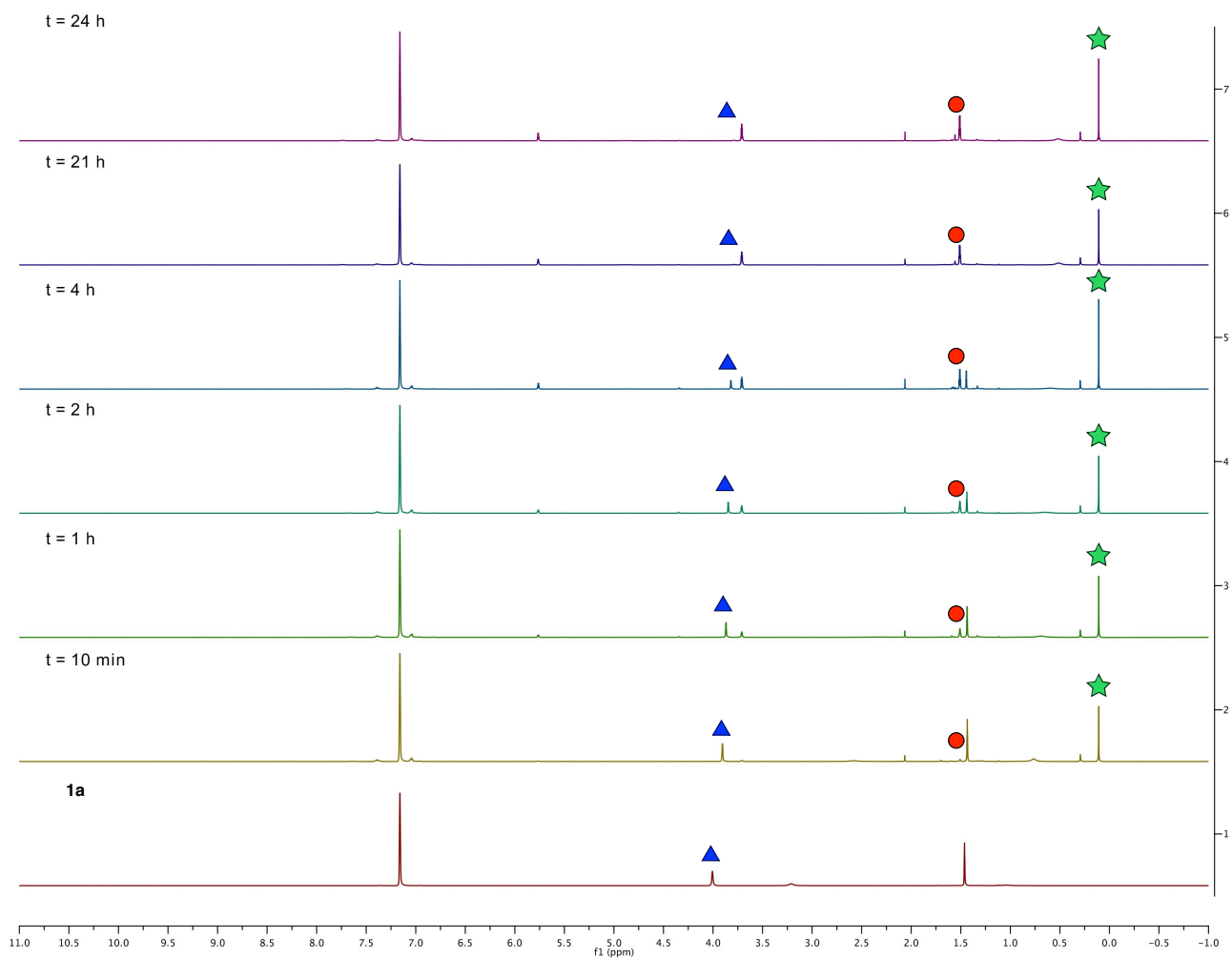
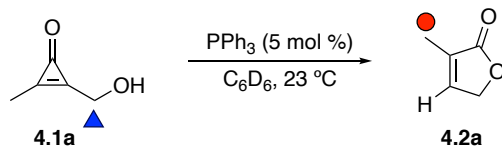


Figure S4-8. Solvent screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PPh₃ (5 mol %) in C₆D₆ (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

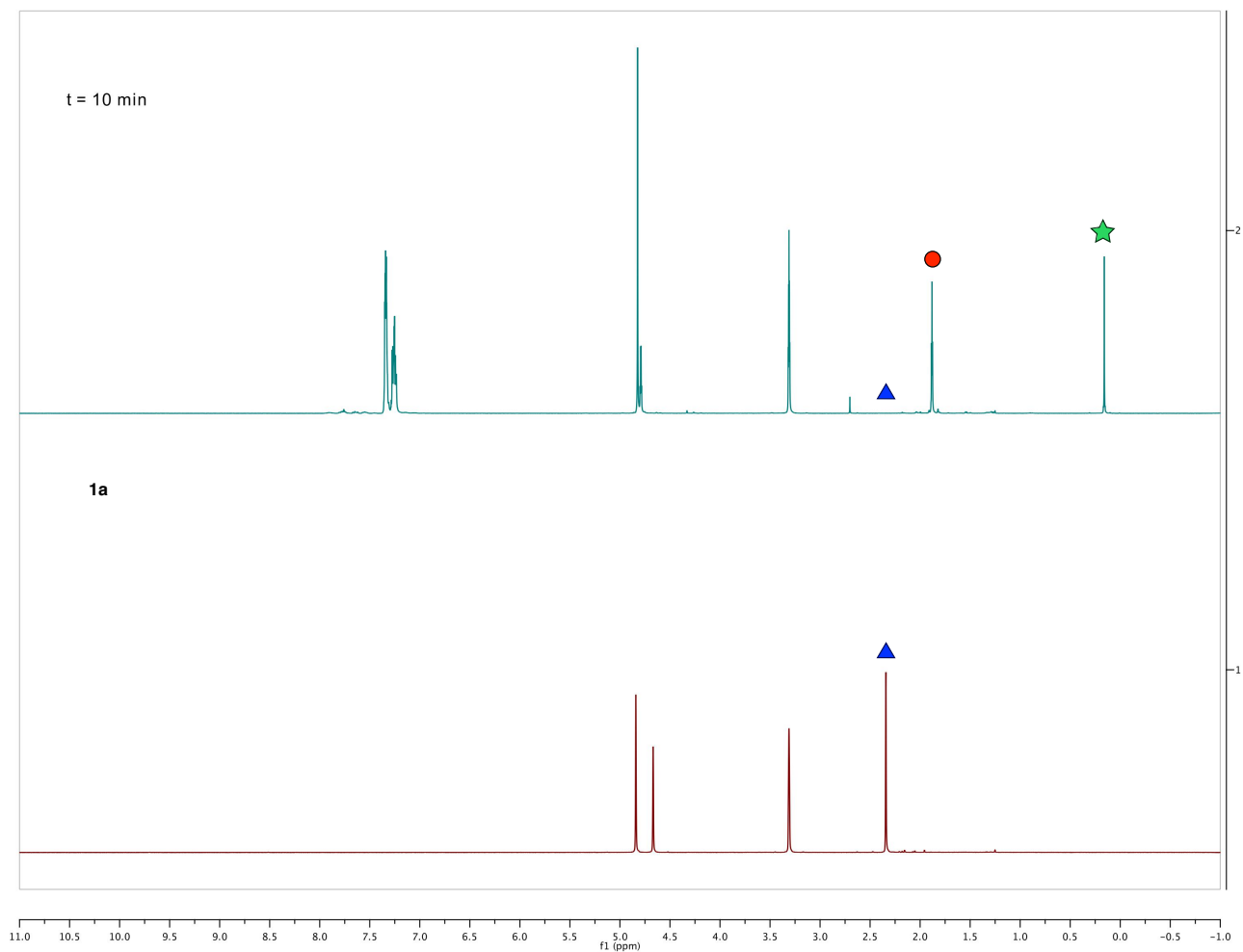
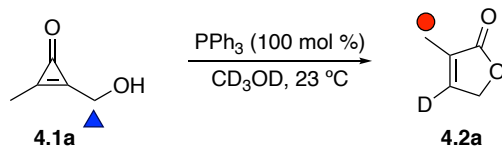


Figure S4-9. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PPh₃ (100 mol %) in CD₃OD (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

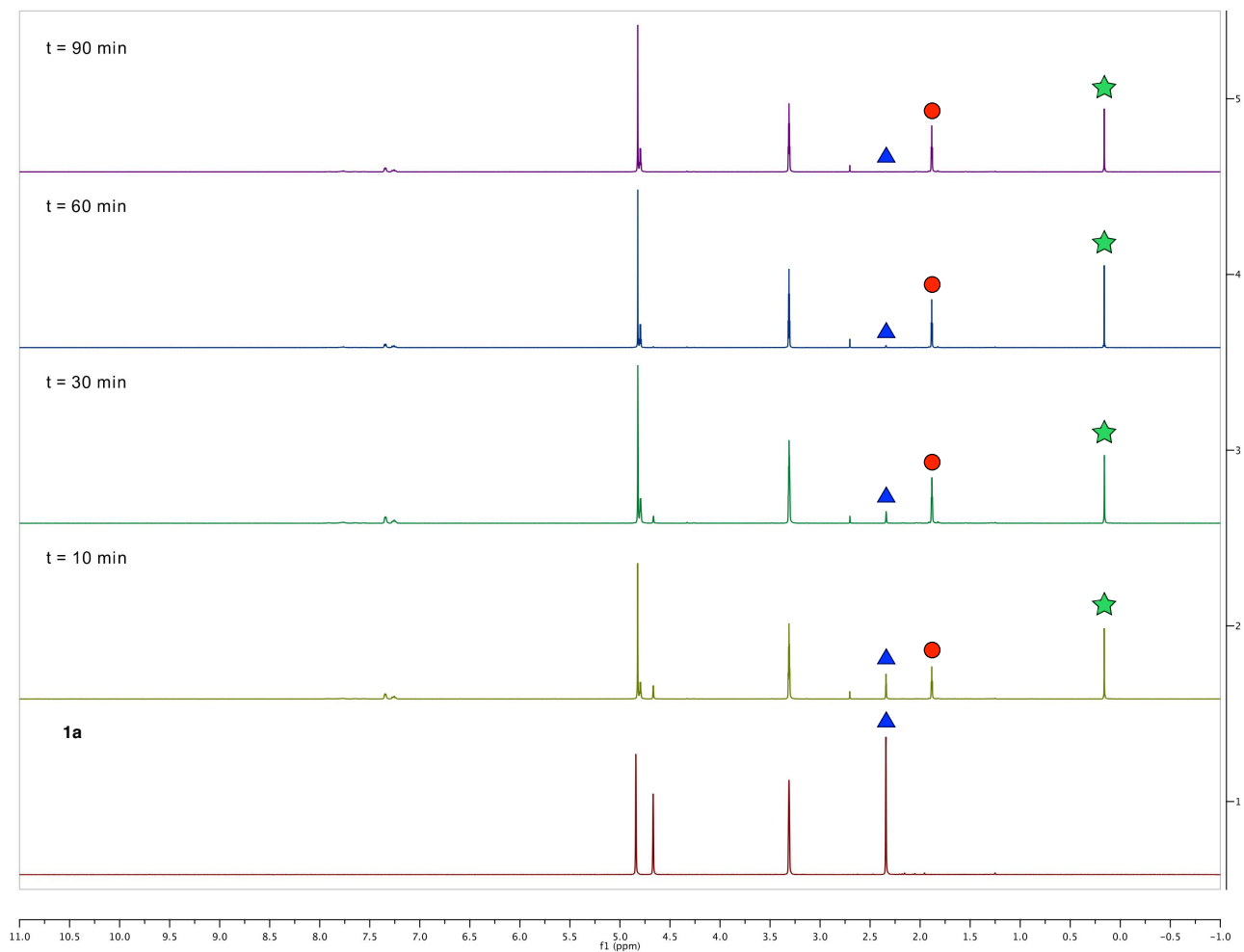
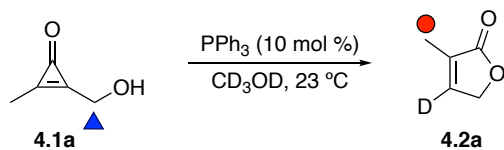


Figure S4-10. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PPh₃ (10 mol %) in CD₃OD (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

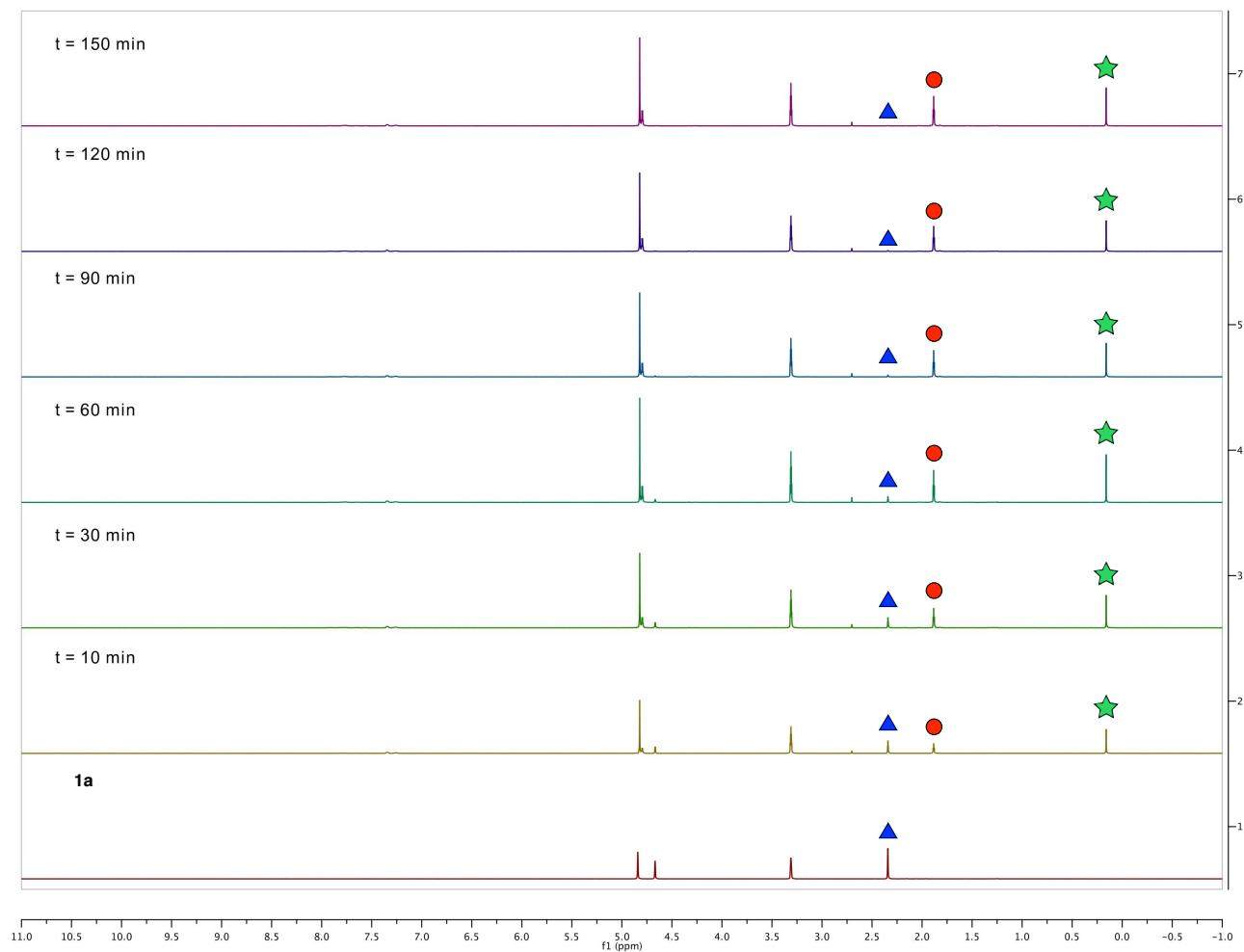
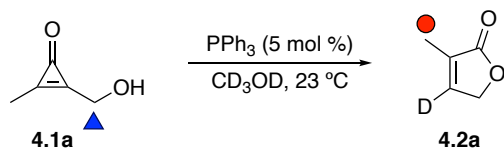


Figure S4-11. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PPh_3 (5 mol %) in CD_3OD (600 μL) at ambient temperature. The reaction was monitored periodically by ^1H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

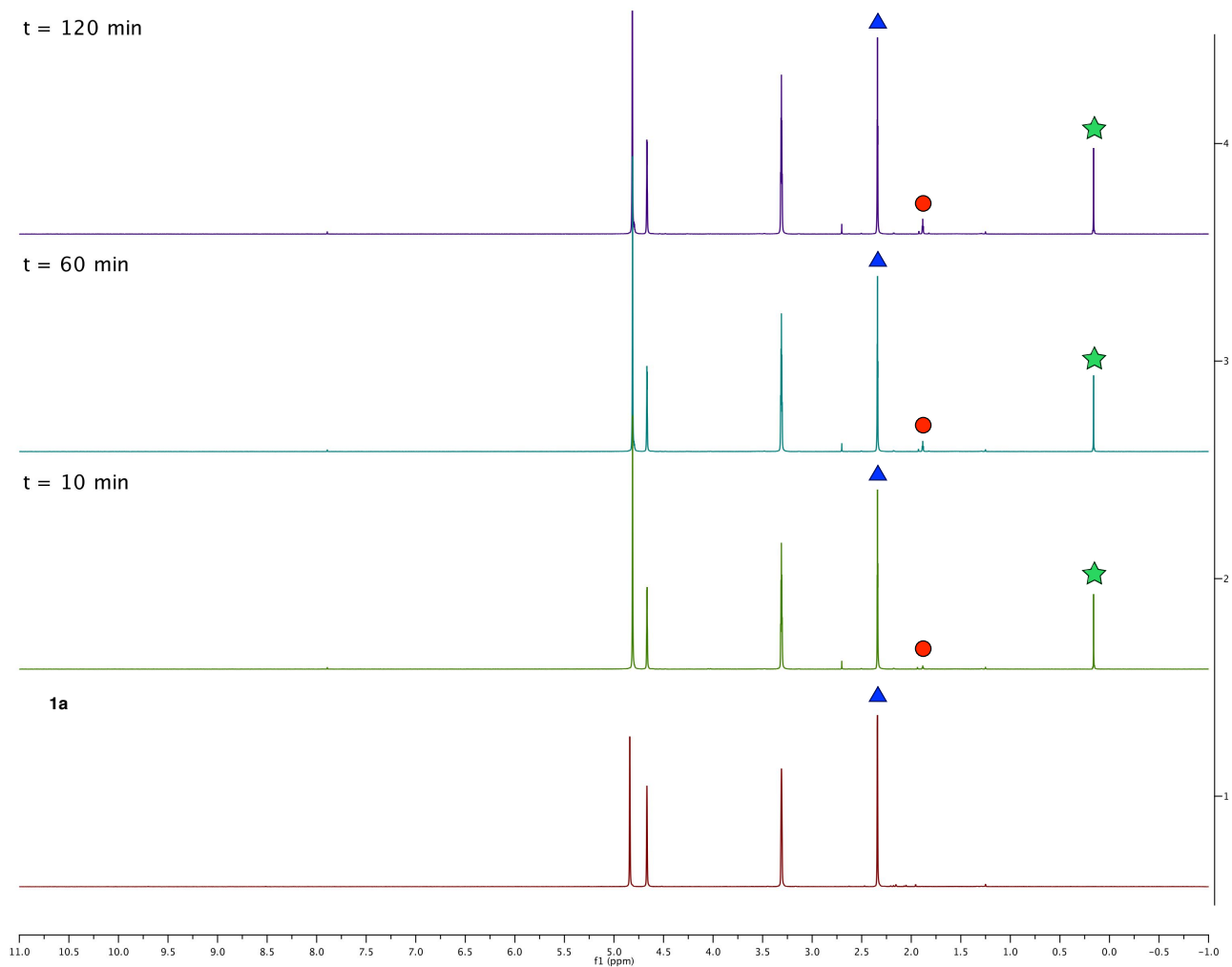
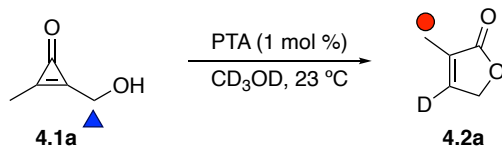


Figure S4-12. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PTA (1 mol %) in CD_3OD (600 μL) at ambient temperature. The reaction was monitored periodically by ^1H NMR spectroscopy for the formation of butenolide **4.2a** (red circle). TMS-acetylene (5 mM, green star) was added as a reference.

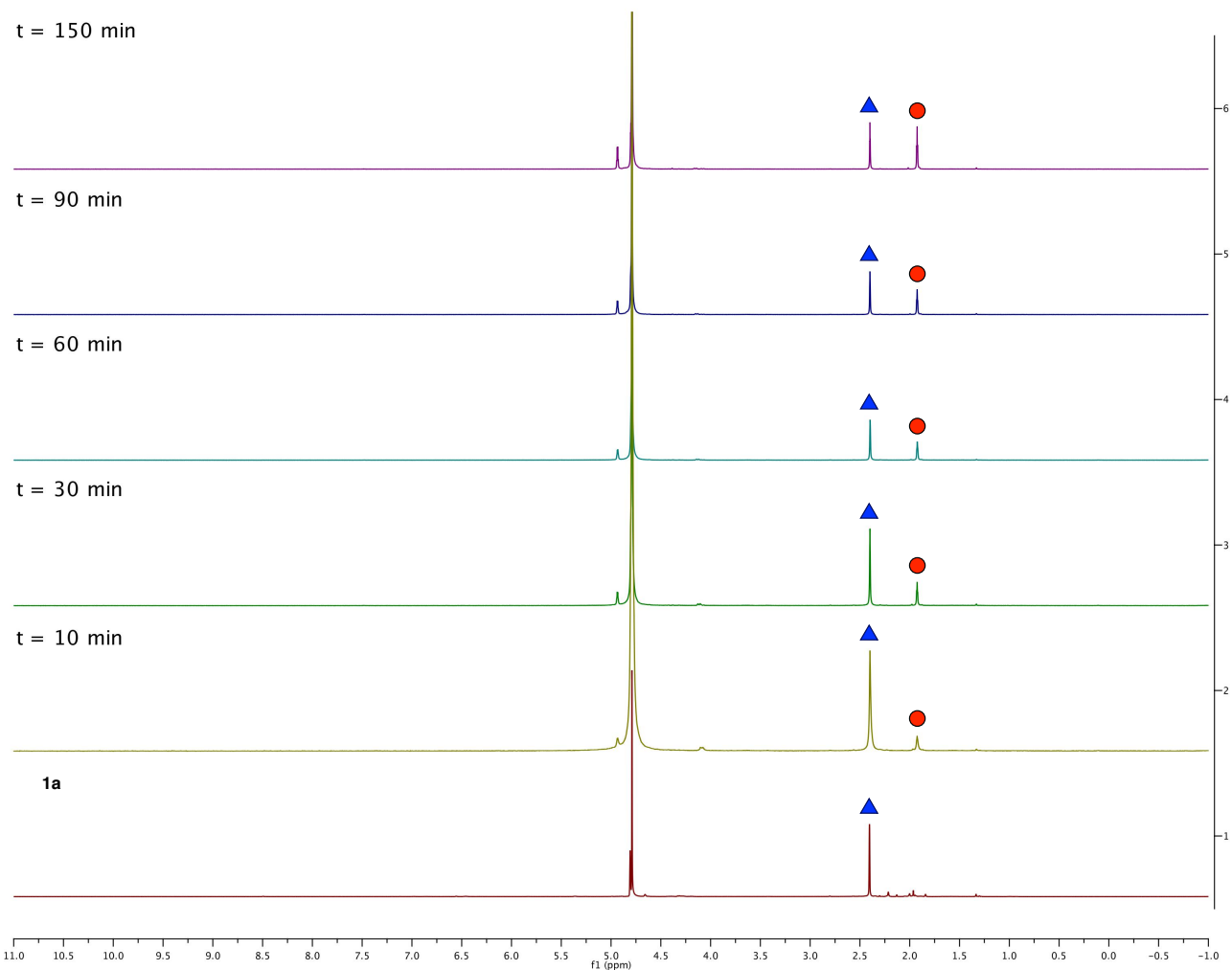
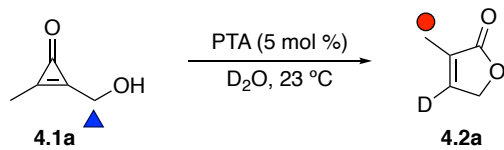


Figure S4-13. Solvent screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with PTA (5 mol %) in D₂O (600 μL) at ambient temperature. The reaction was monitored periodically by ¹H NMR spectroscopy for the formation of butenolide **4.2a** (red circle).

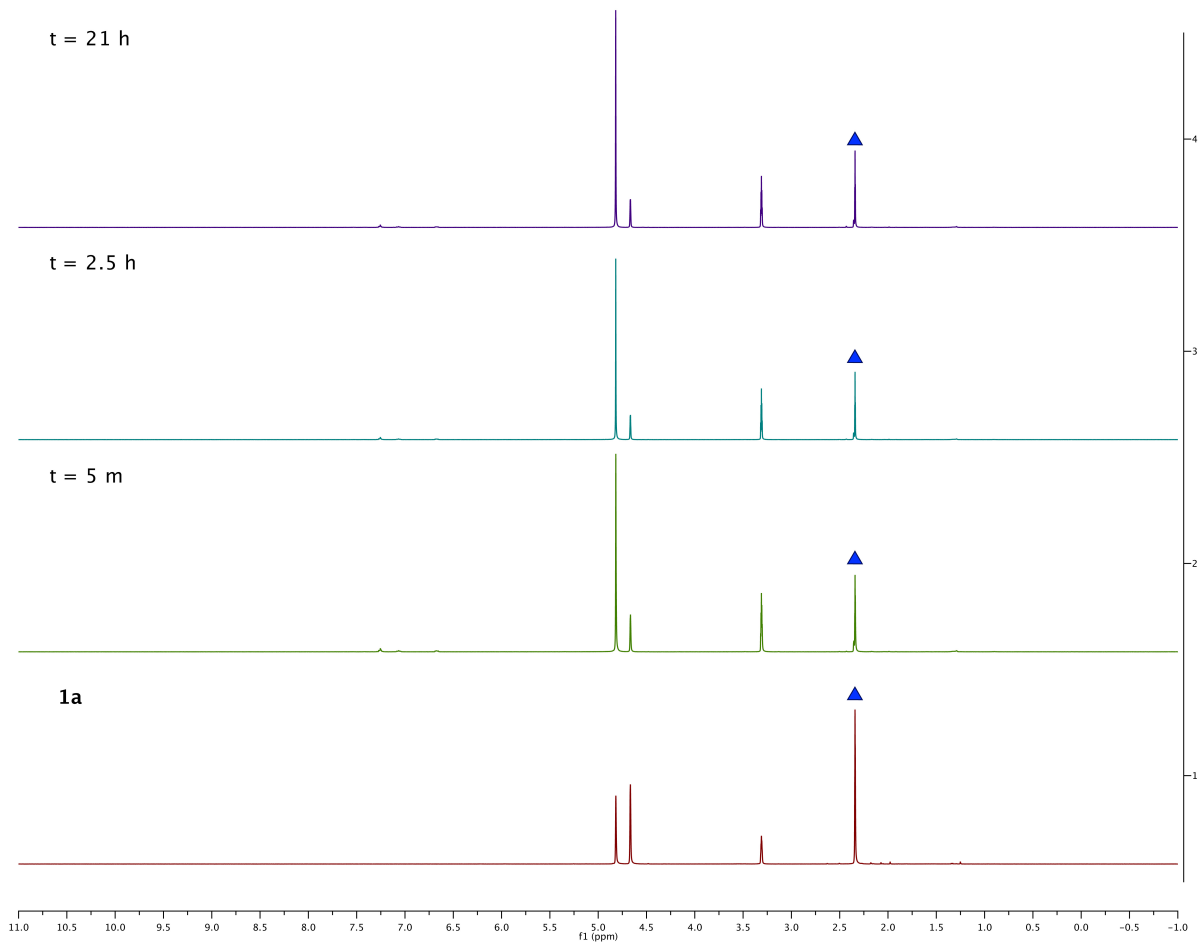
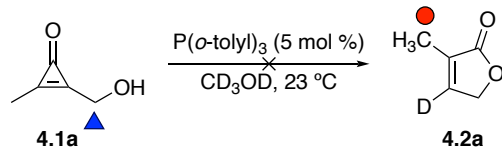


Figure S4-14. Phosphine screen for butenolide formation. CpO **4.1a** (25 mM, blue triangle) was incubated with $P(o\text{-tolyl})_3$ (5 mol %) in CD_3OD (600 μL) at ambient temperature. The reaction was monitored periodically by ^1H NMR spectroscopy. No conversion to butenolide **4.2a** was observed.

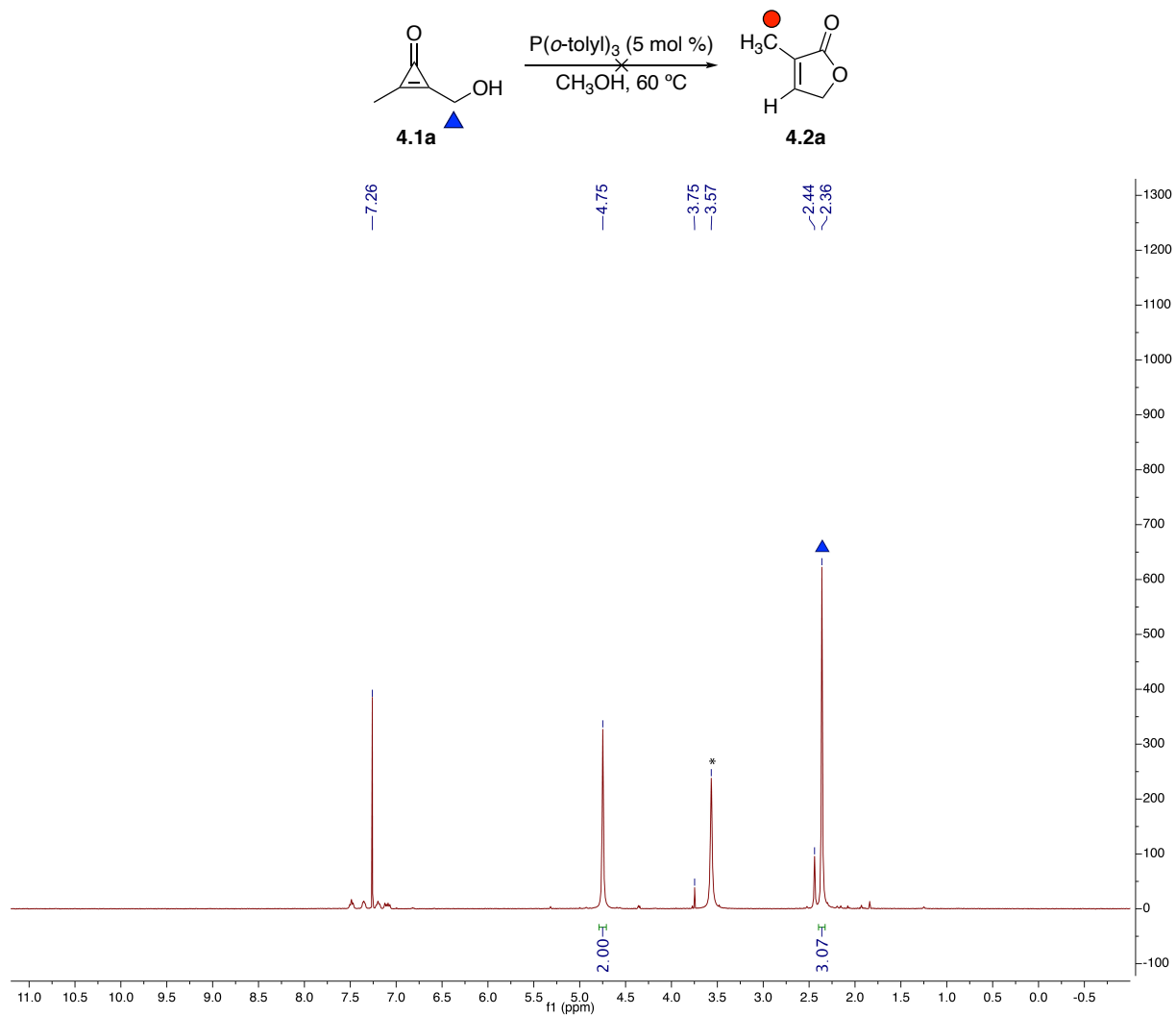


Figure S4-15. Phosphine screen for butenolide formation. CpO **4.1a** (35.5 mg, 0.361 mmol, 1.00 equiv., blue triangle) was incubated with $P(o\text{-tolyl})_3$ (5.5 mg, 0.021 mmol, 5 mol %) in CH_3OH (14.4 mL) at reflux. After 18 h, the reaction was cooled to ambient temperature and concentrated. The crude mixture was analyzed (in CDCl_3) by ^1H NMR spectroscopy. No conversion to butenolide **4.2a** was observed (* residual methanol).

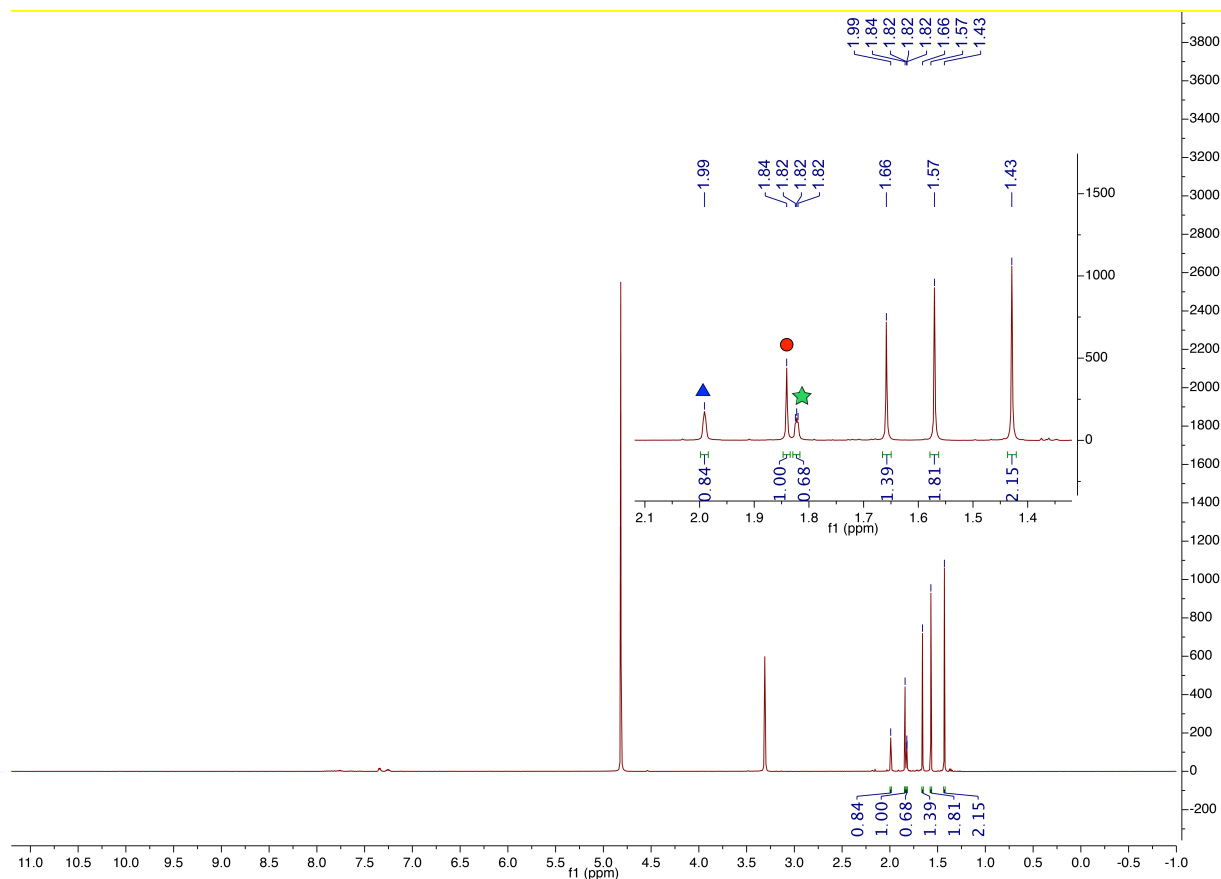
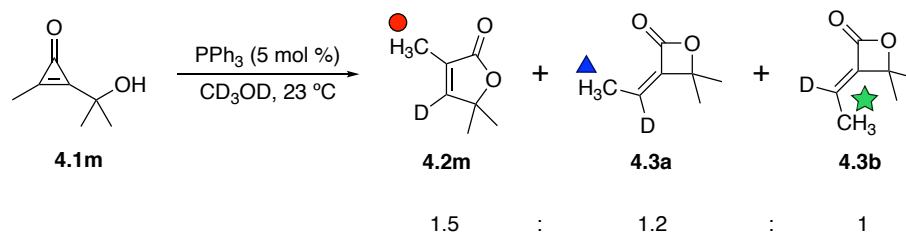


Figure S4-16. β -Lactone products were only observed with **4.1m**. CpO **4.1m** (25 mM) was incubated with PPh_3 (5 mol %) in CD_3OD at ambient temperature. The reaction was monitored by ^1H NMR spectroscopy until full consumption of starting material was observed. CpO **4.1m** gave a mixture of butenolide **4.2m** (red circle) and β -lactones **4.3a** [1] (blue triangle) and **4.3b** [1] (green star) in an approximate 1.5 : 1.2 : 1.0 molar ratio. Splitting of the β -lactone methyl peaks was attributed to vinylic coupling to deuterium.

[1]. Danheiser, R. L.; Choi, Y. M.; Menichincheri, M.; Stoner, E. J. *J. Org. Chem.* **1993**, *58*, 322–327.

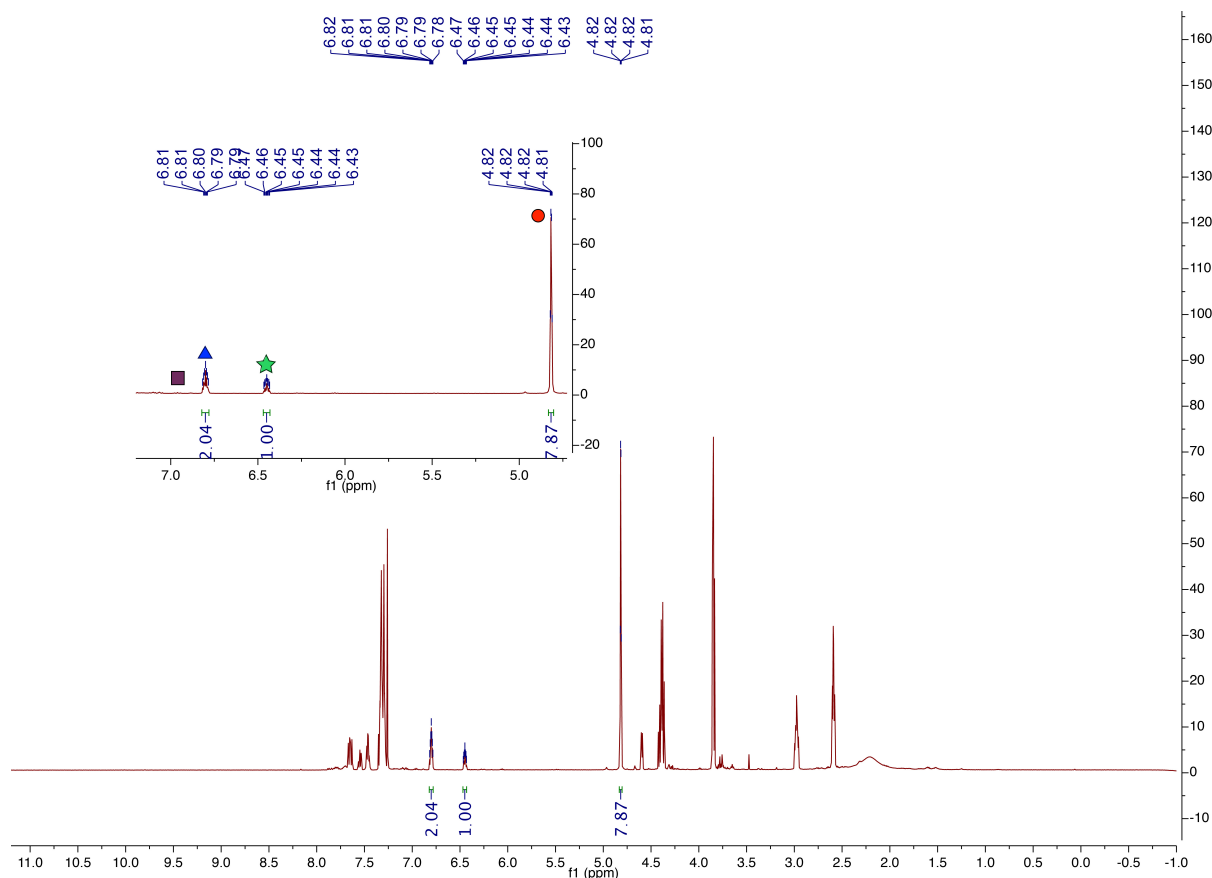
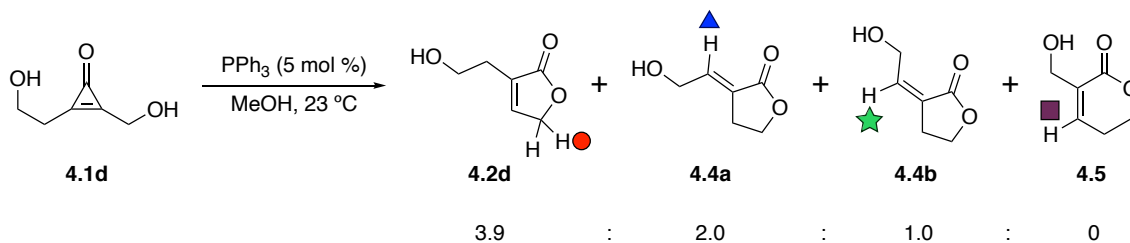


Figure S4-17. Mechanistic study on butenolide cyclization. Diol-CpO **4.1d** was incubated with PPh₃ (5 mol %) in CH₃OH. The reaction was concentrated, and the crude mixture was analyzed by ¹H NMR spectroscopy. A mixture of lactones **4.2d** (red circle), **4.4a** (blue triangle) [2], and **4.4b** (green star) [2] was formed in an approximate 3.9 : 2.0 : 1.0 ratio. No δ -lactone **4.5** (purple square) [3] was observed.

[2]. Merritt, A. T.; Powner, R. H.; Williams, D. J.; Williams, C. M.; Ley, S. V. *Org. Biomol. Chem.* **2011**, *9*, 4745–4747.

[3]. Fumiyama, H.; Sadayuki, T.; Osada, Y.; Goto, Y.; Nakao, Y.; Hosokawa, S. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 4355–4357.

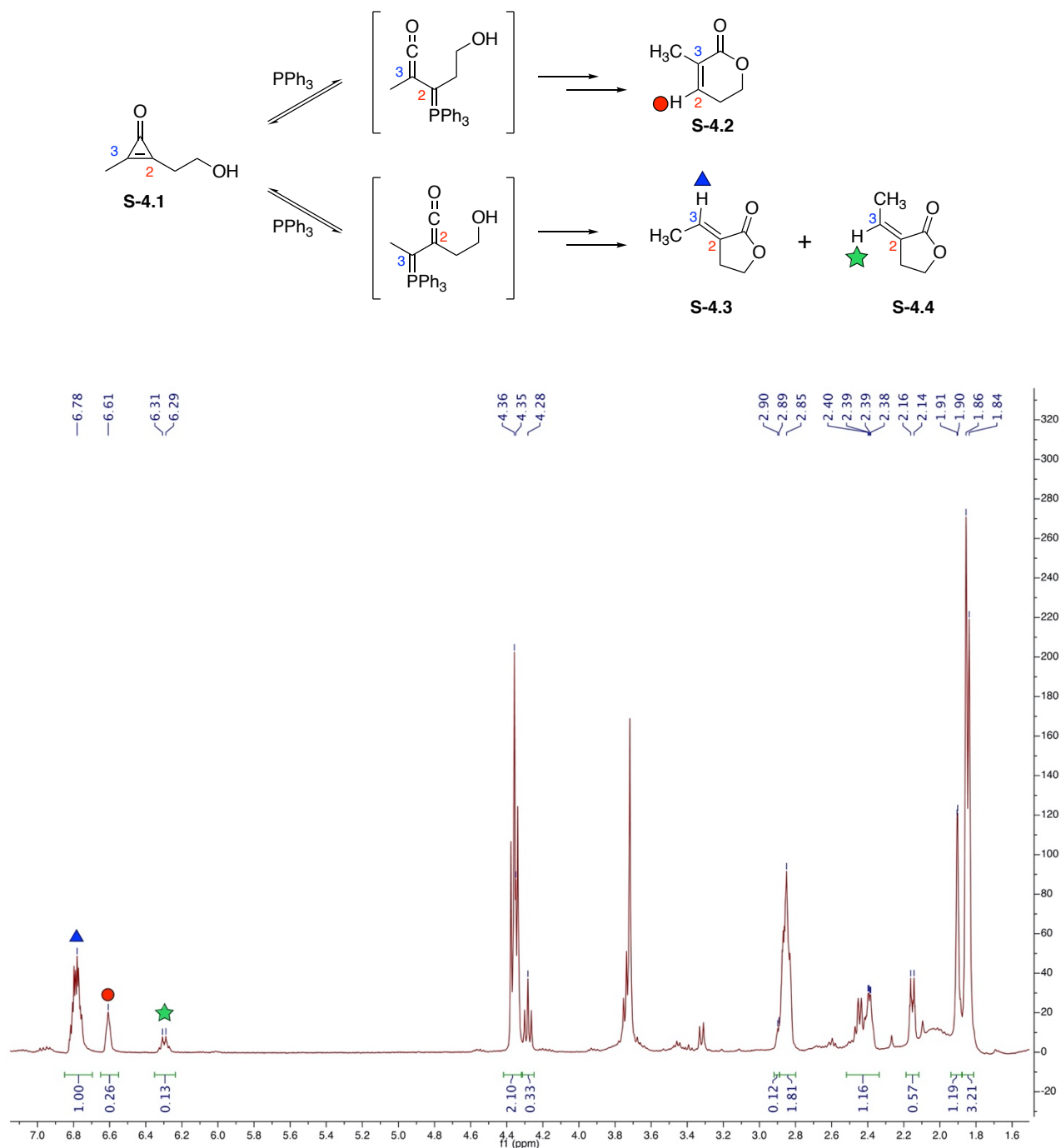
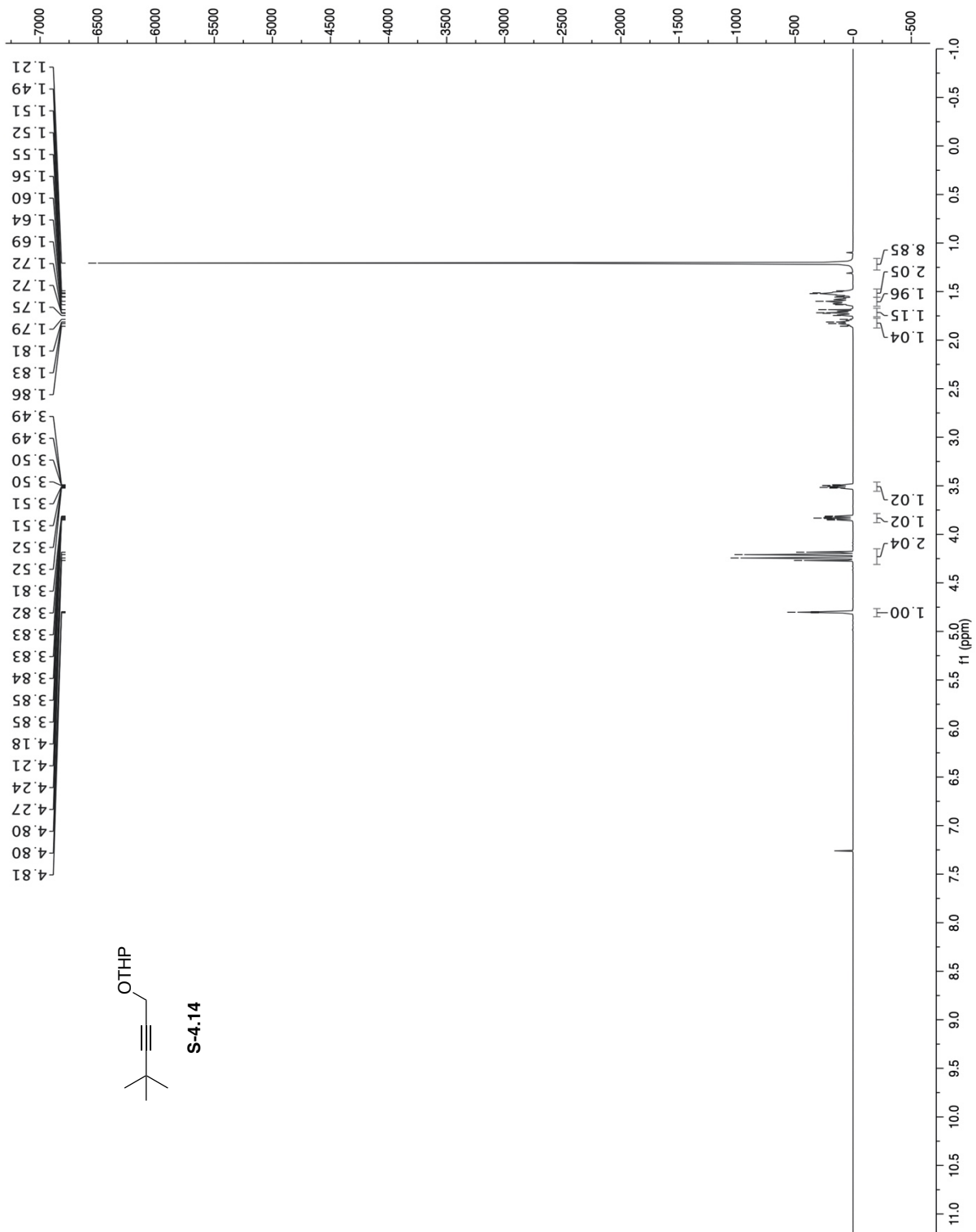
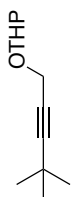


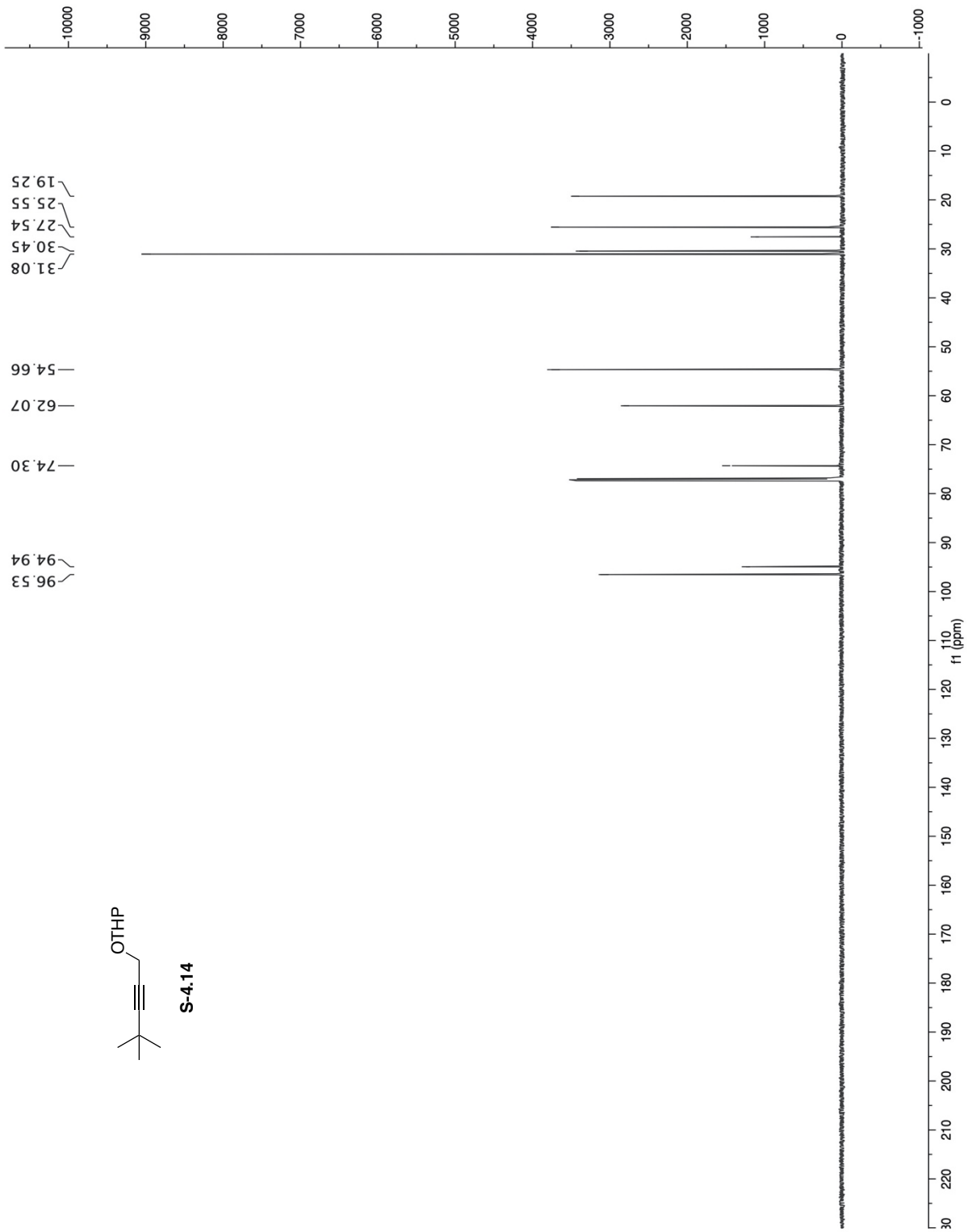
Figure S4-18. Mechanistic study on ketene-ylide formation. CpO **S-4.1** (25 mM) was incubated with PPh_3 (5 mol %) in CH_3OH . The reaction was concentrated, and the crude mixture was analyzed by ^1H NMR spectroscopy. A mixture of lactones **S-4.2** (red circle) [4], **S-4.3** (blue triangle) [2], and **S-4.4** (green star) [2] was observed. This experiment demonstrates that both ketene-ylides can be formed and trapped.

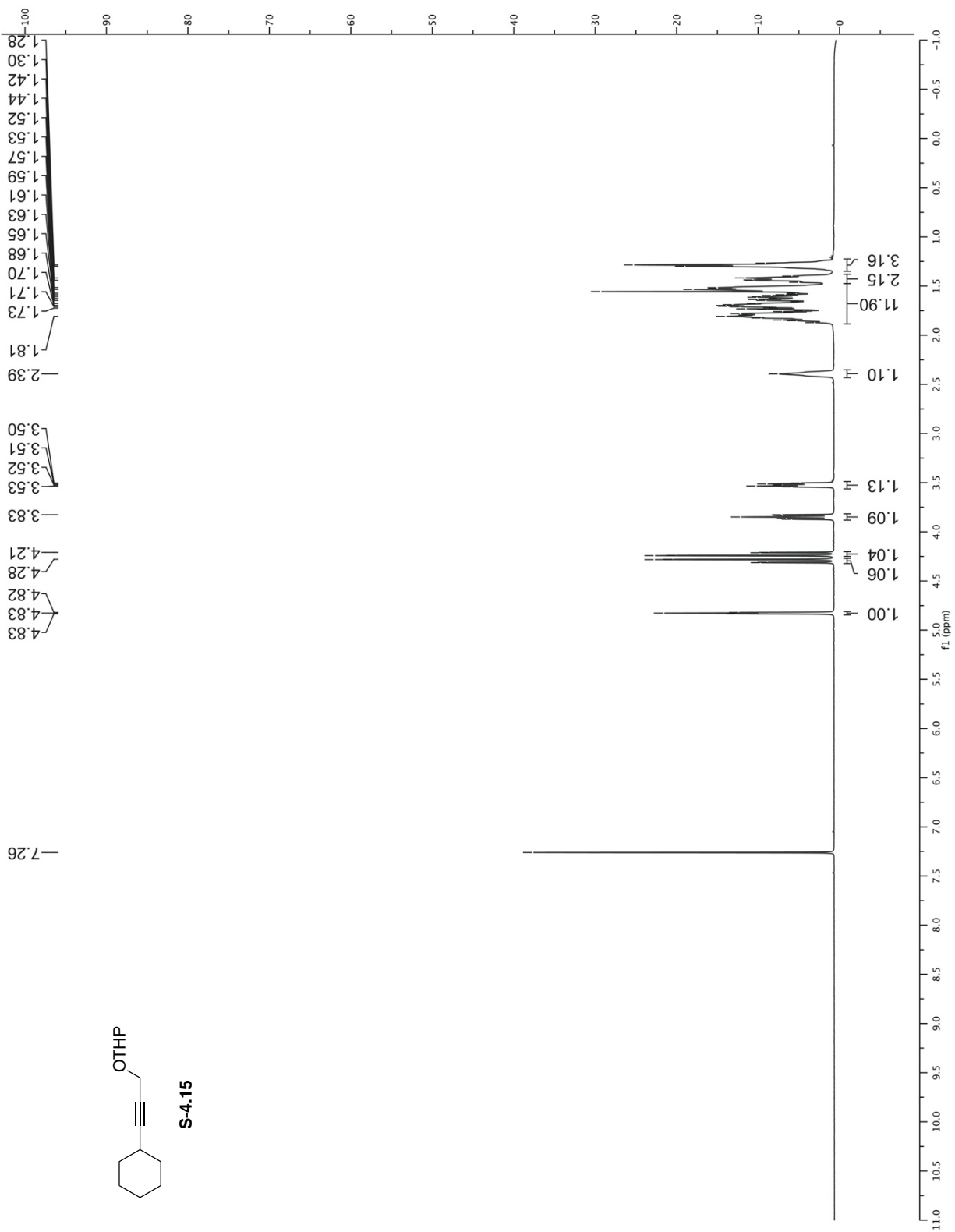
[4]. D'Annibale, A.; Ciaralli, L.; Bassetti, M.; Pasquini, C. *J. Org. Chem.* **2007**, *72*, 6067–6074.

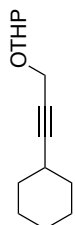




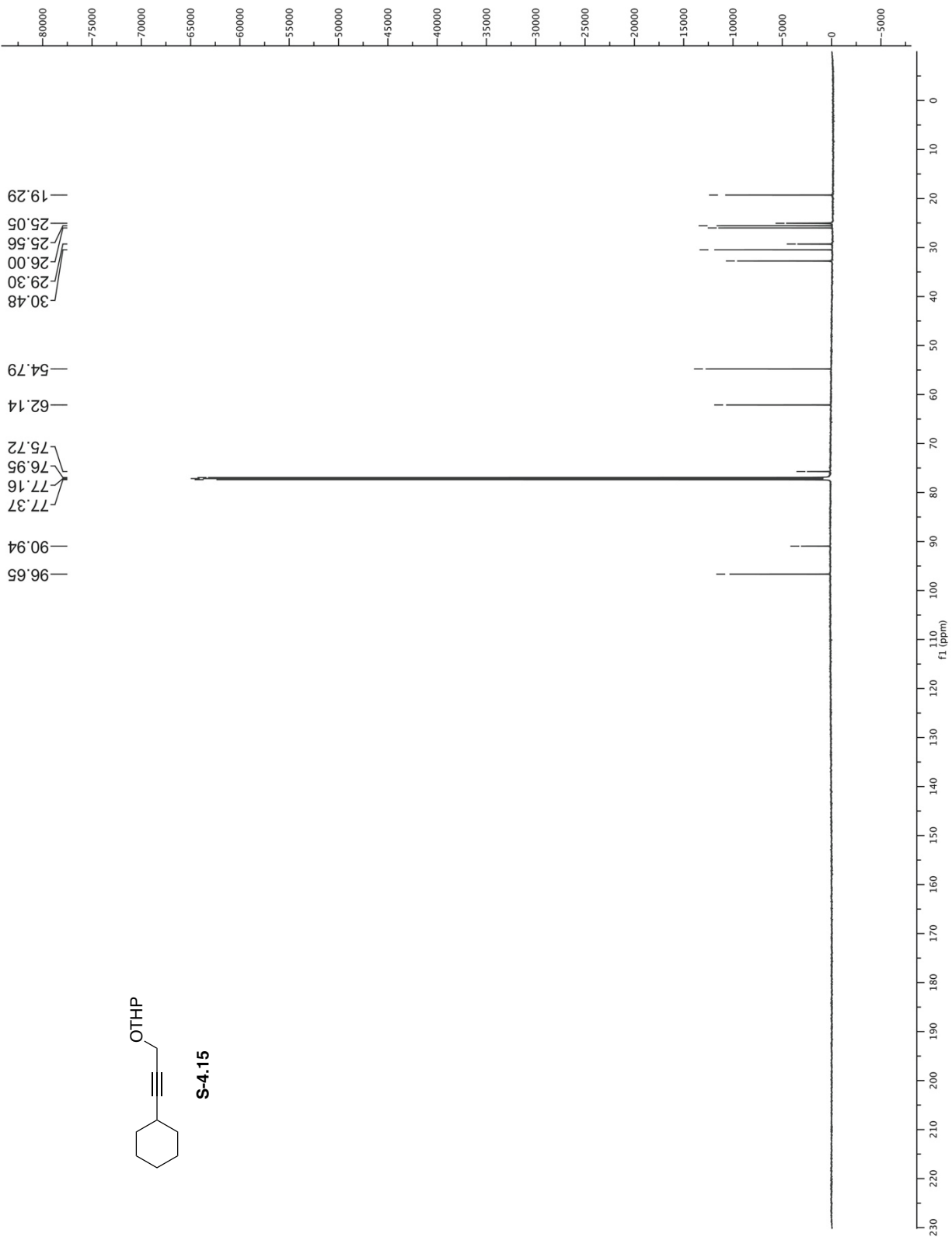
S-4.14

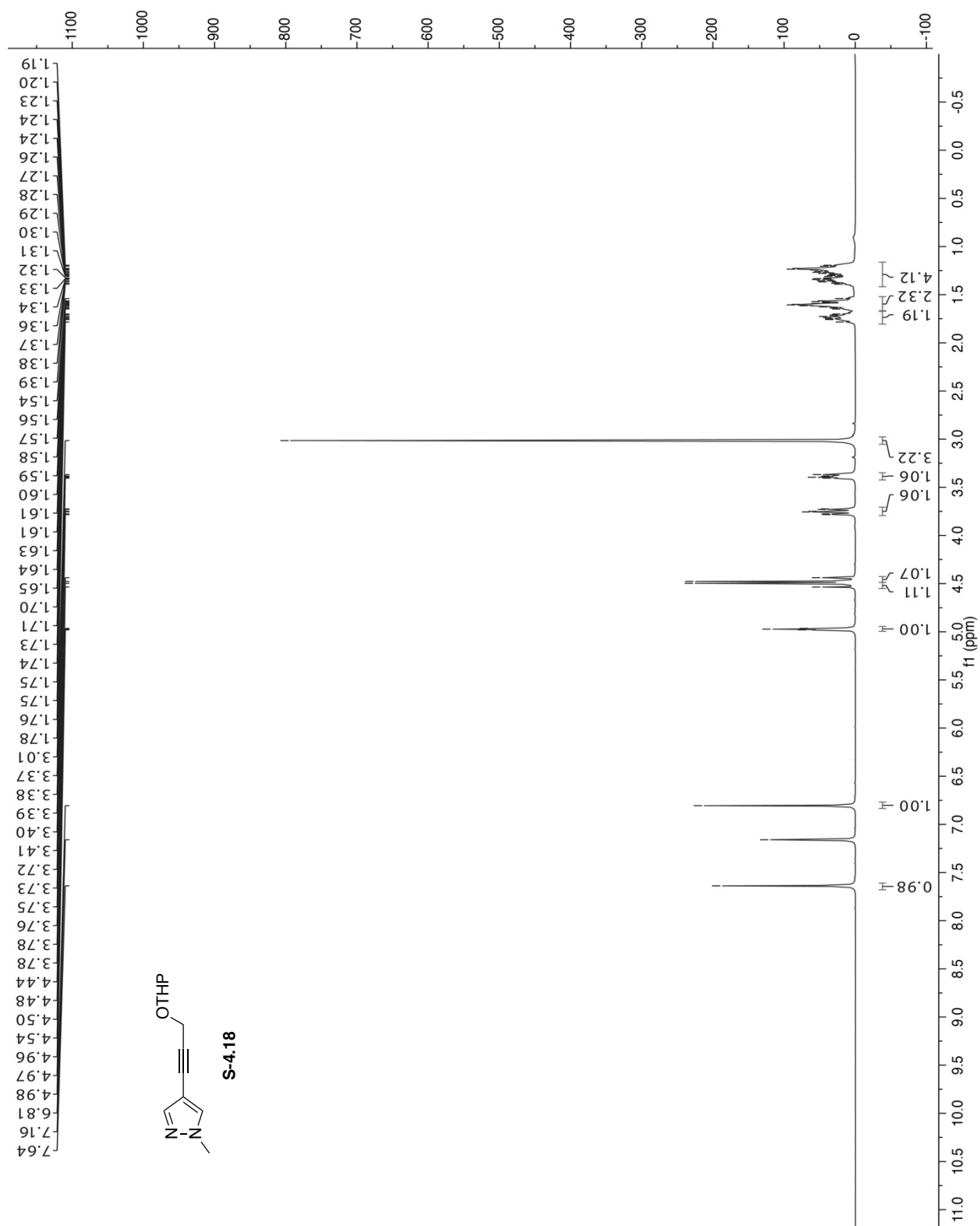


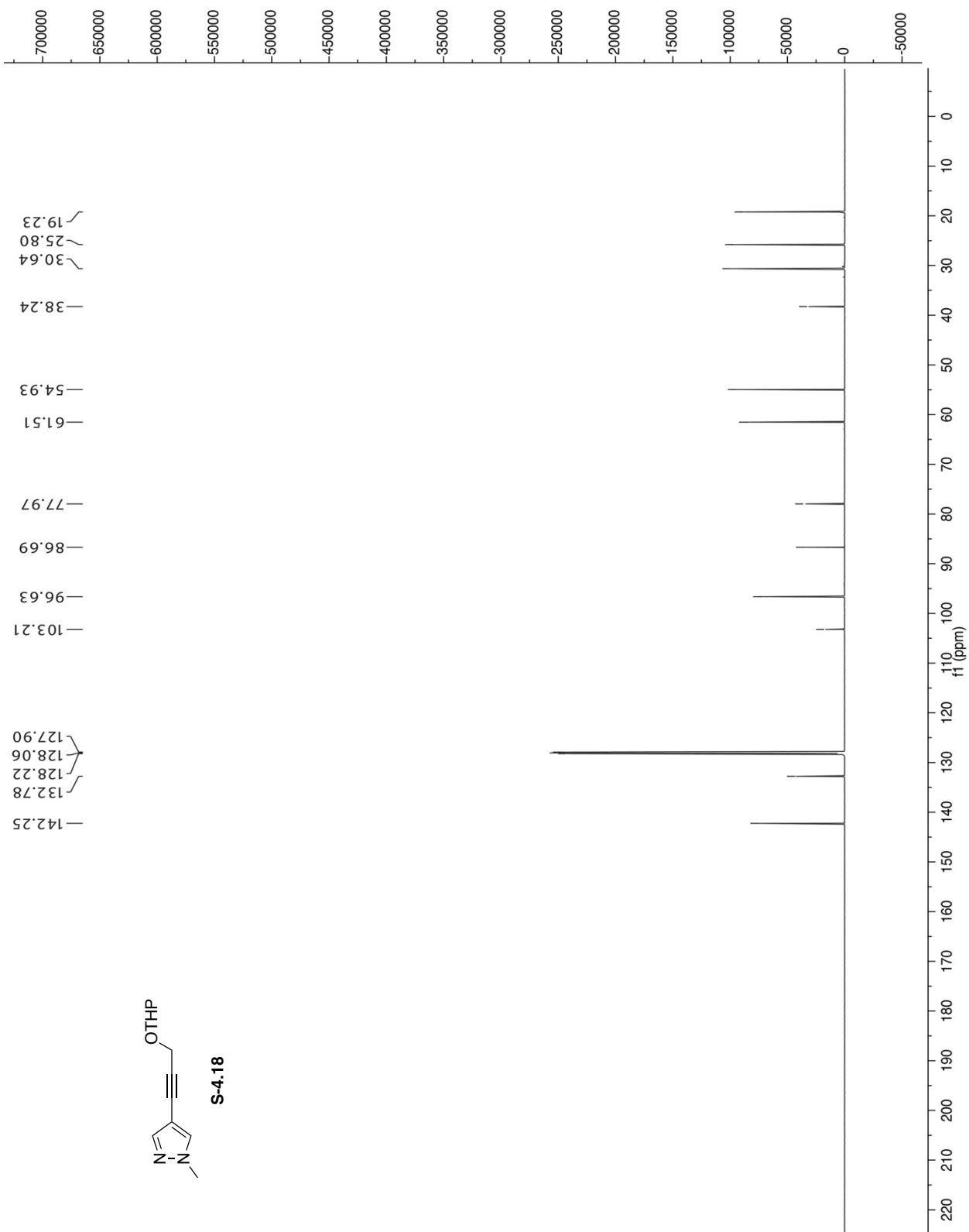


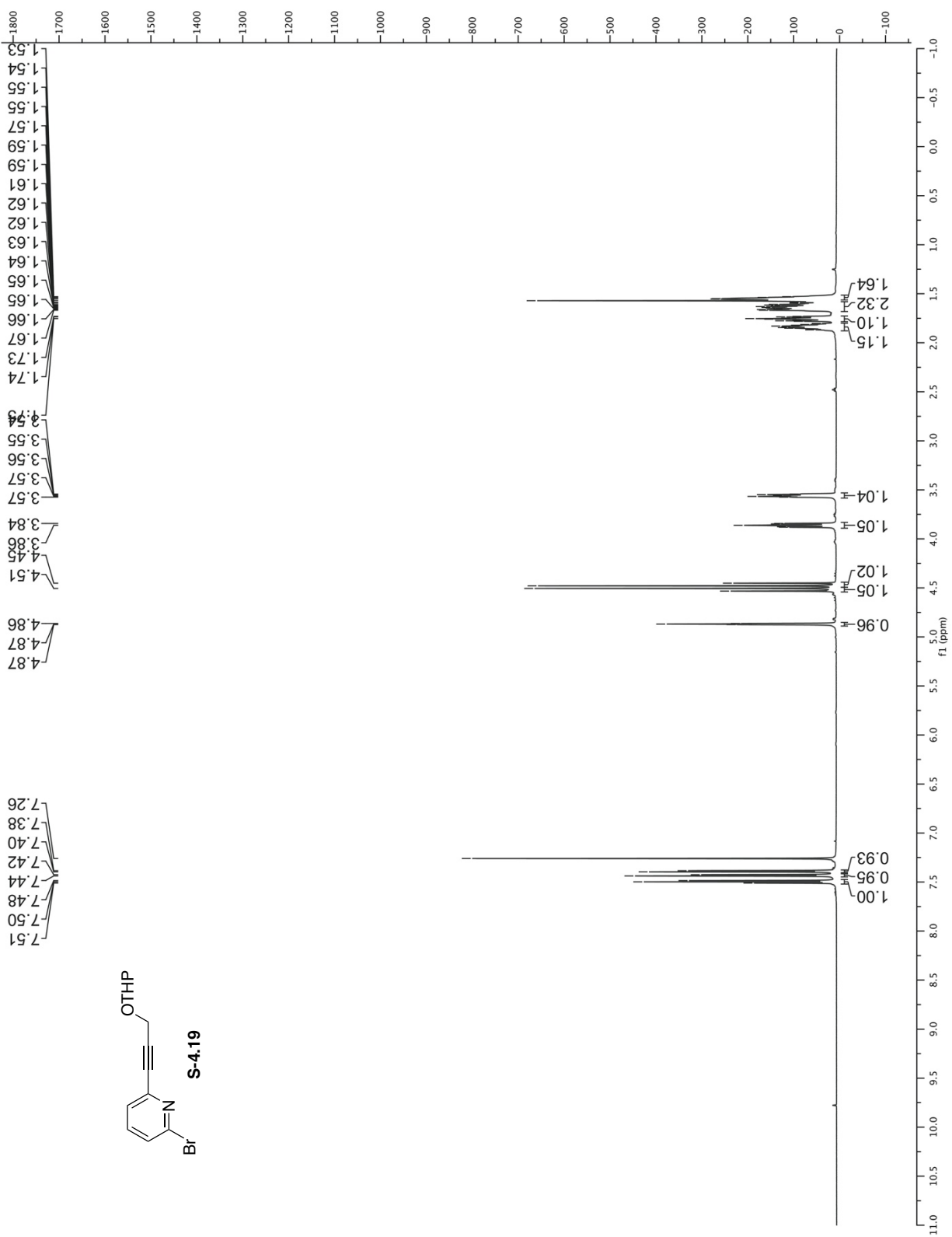


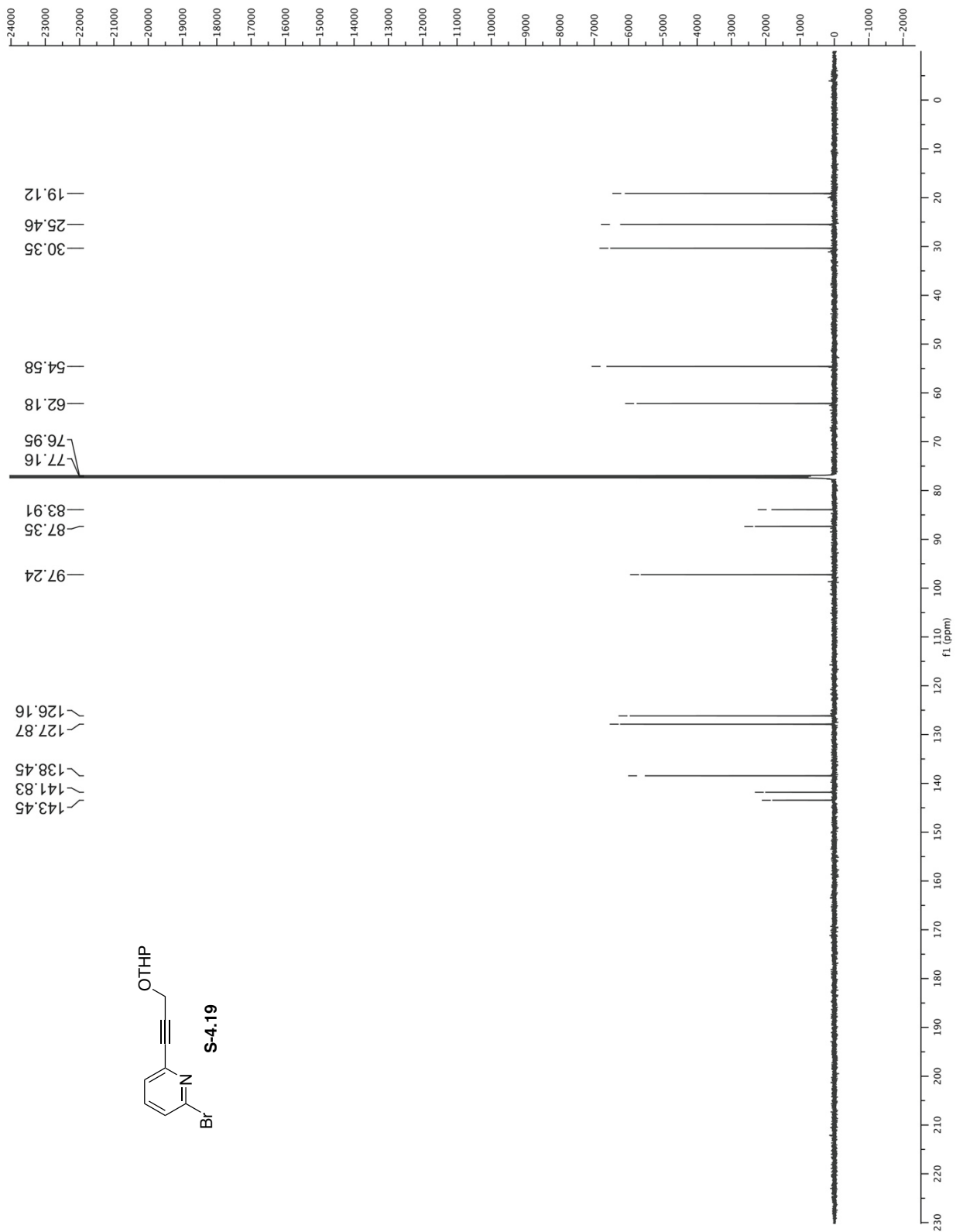
S-4.15

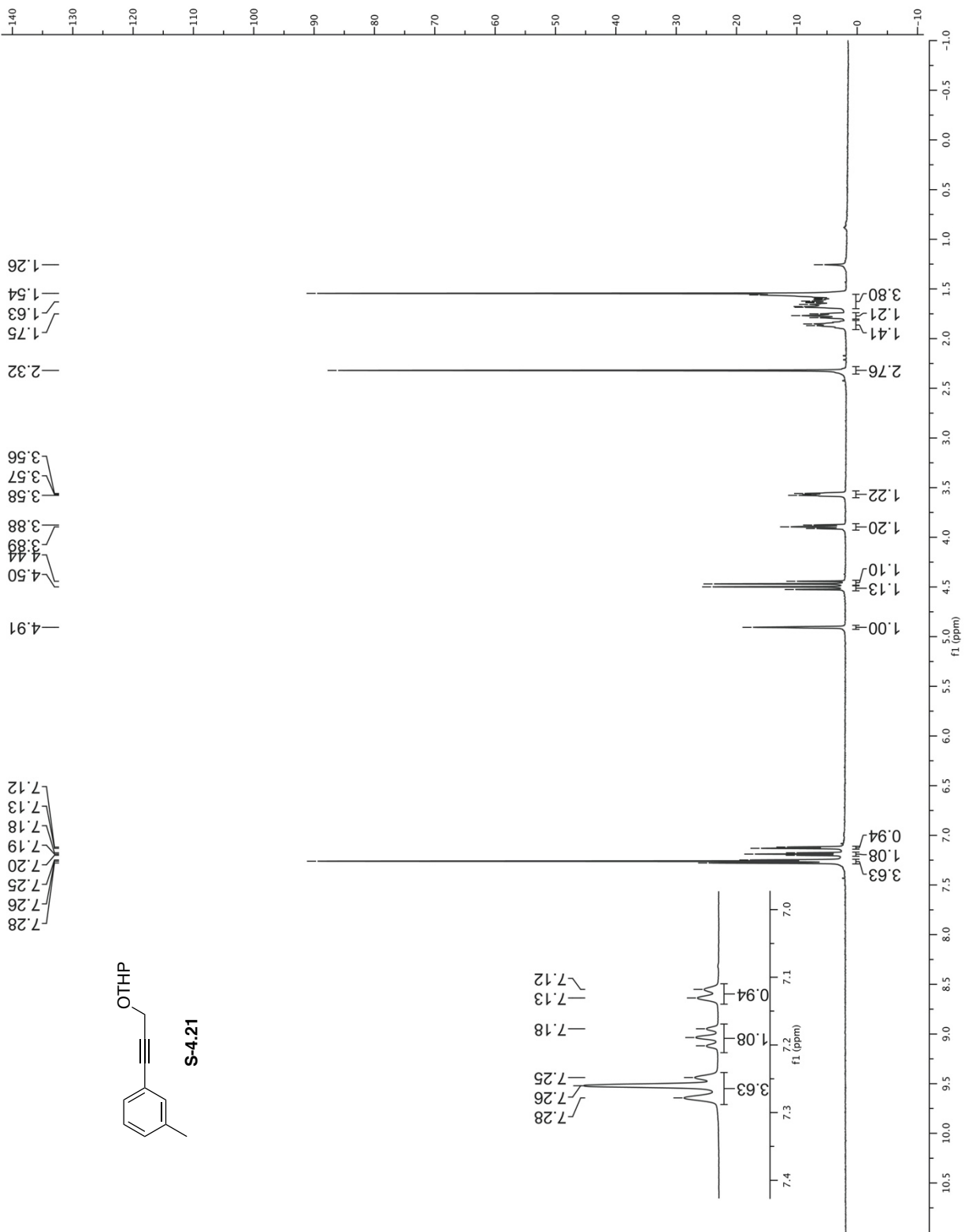


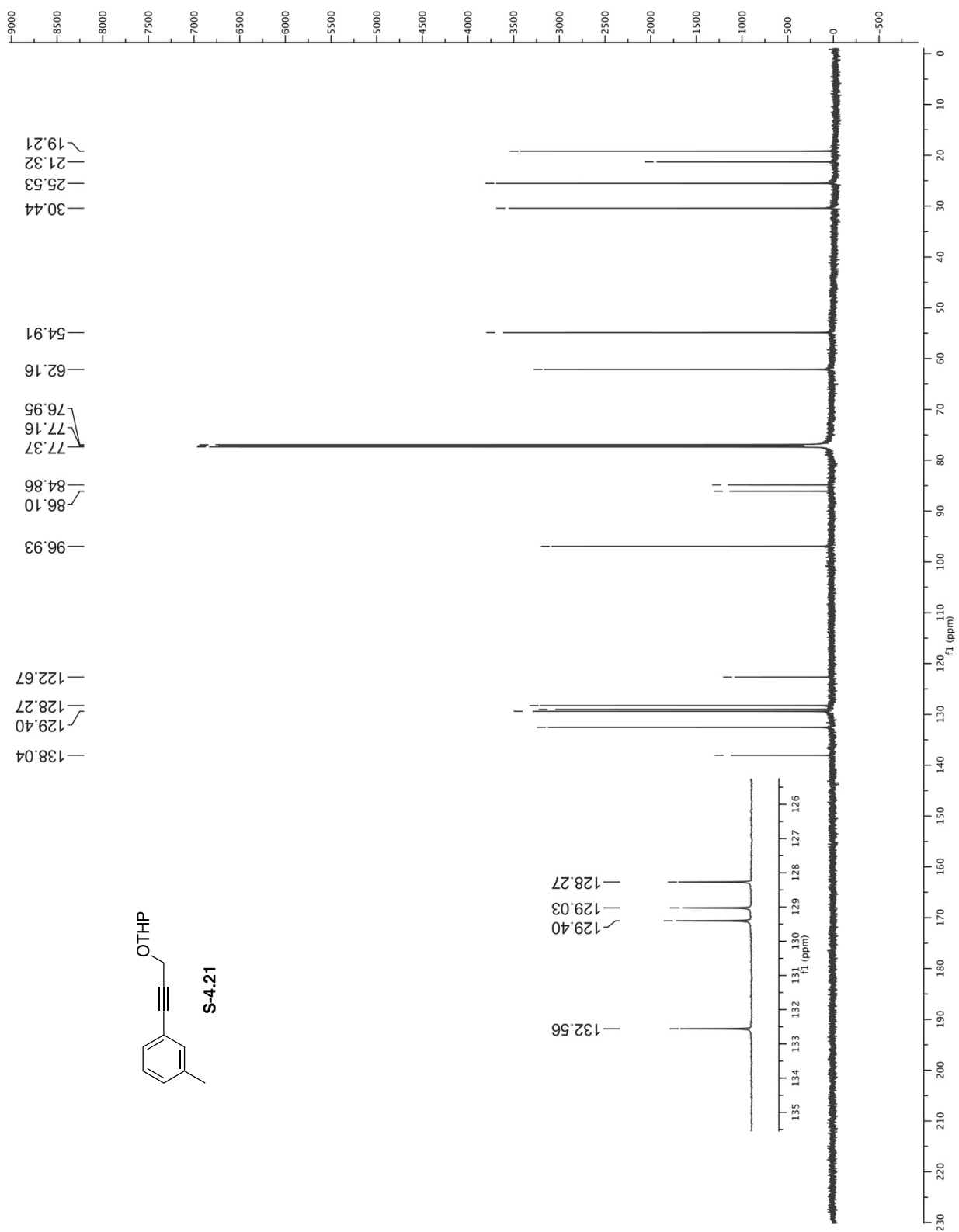


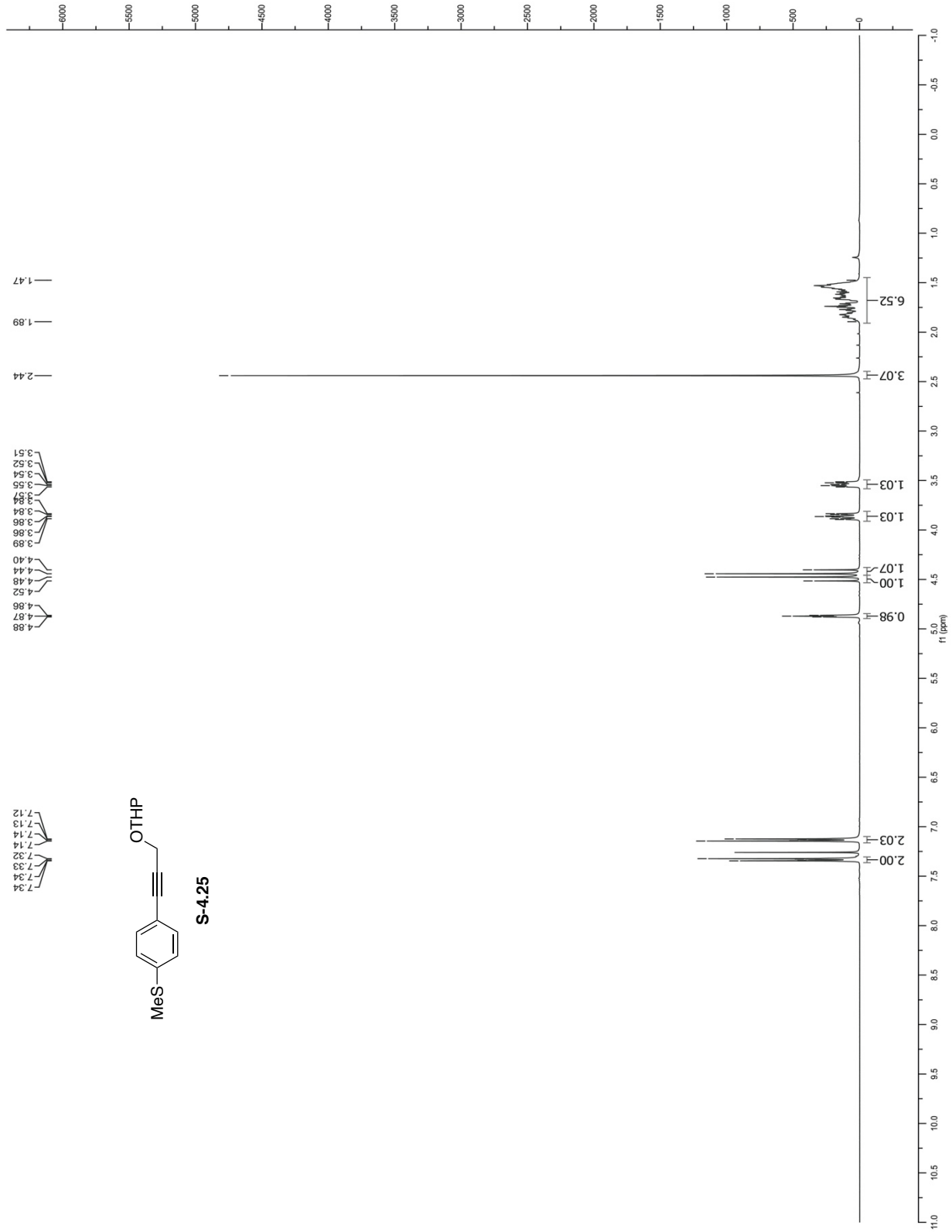


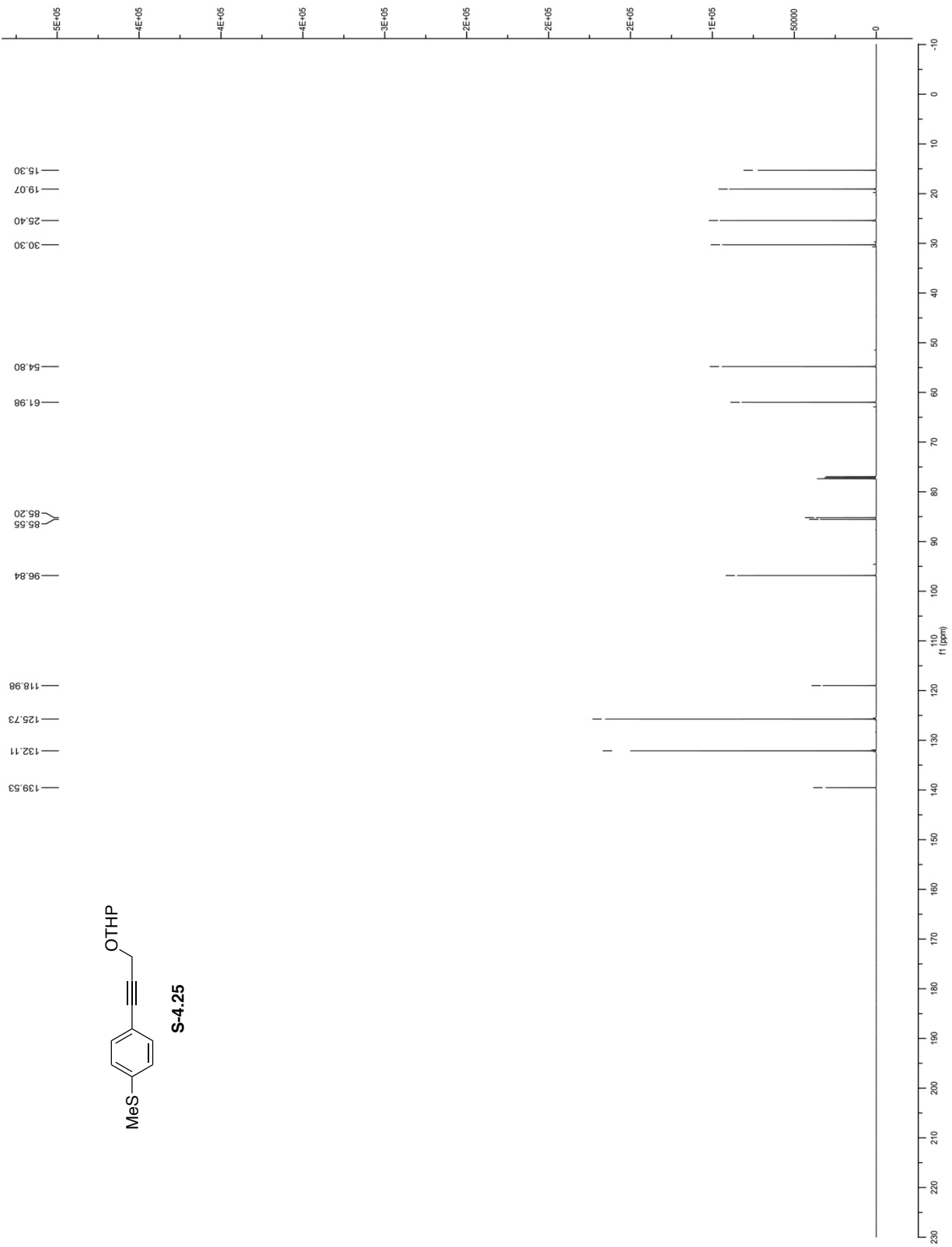
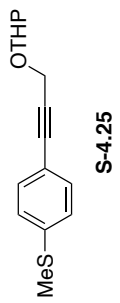


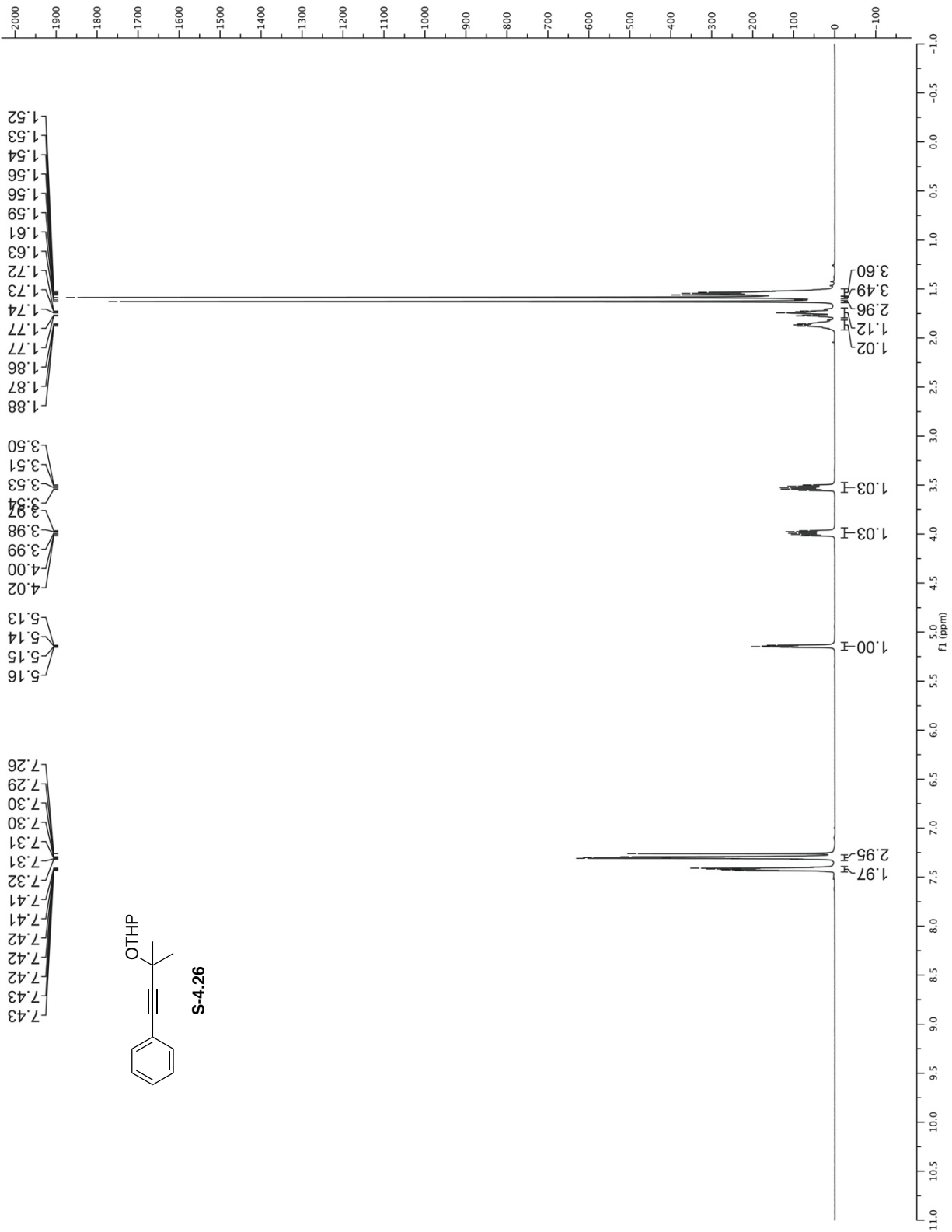


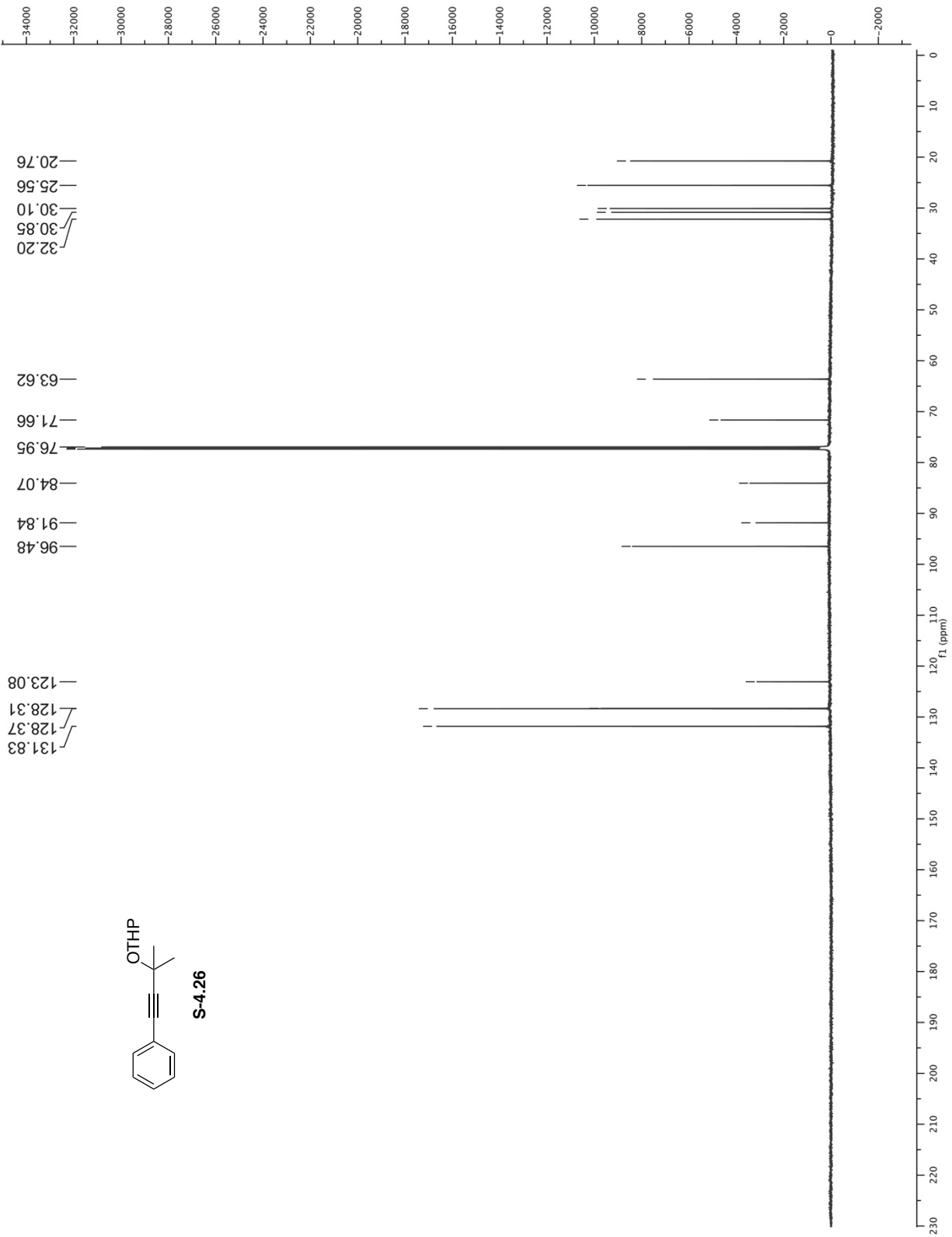


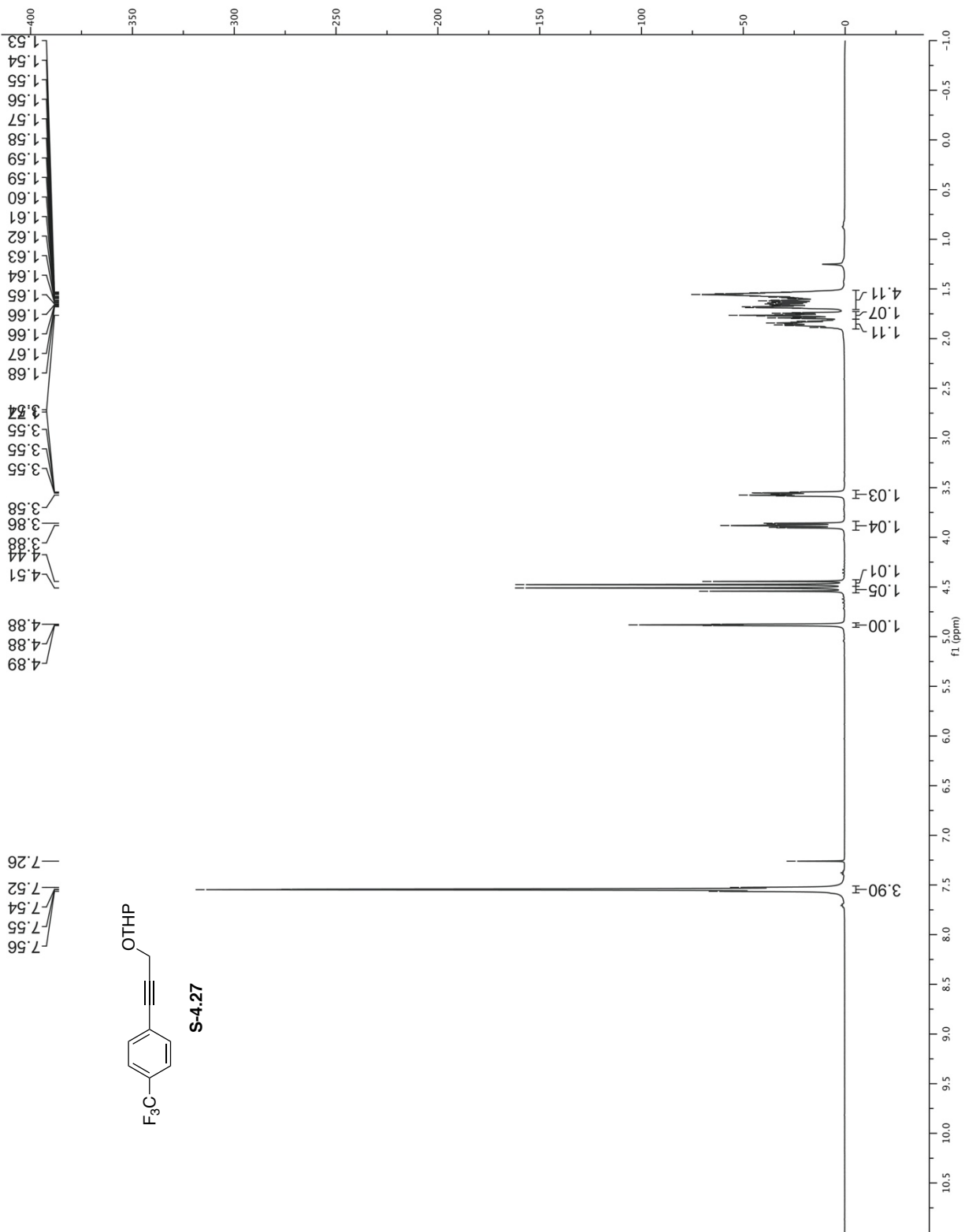


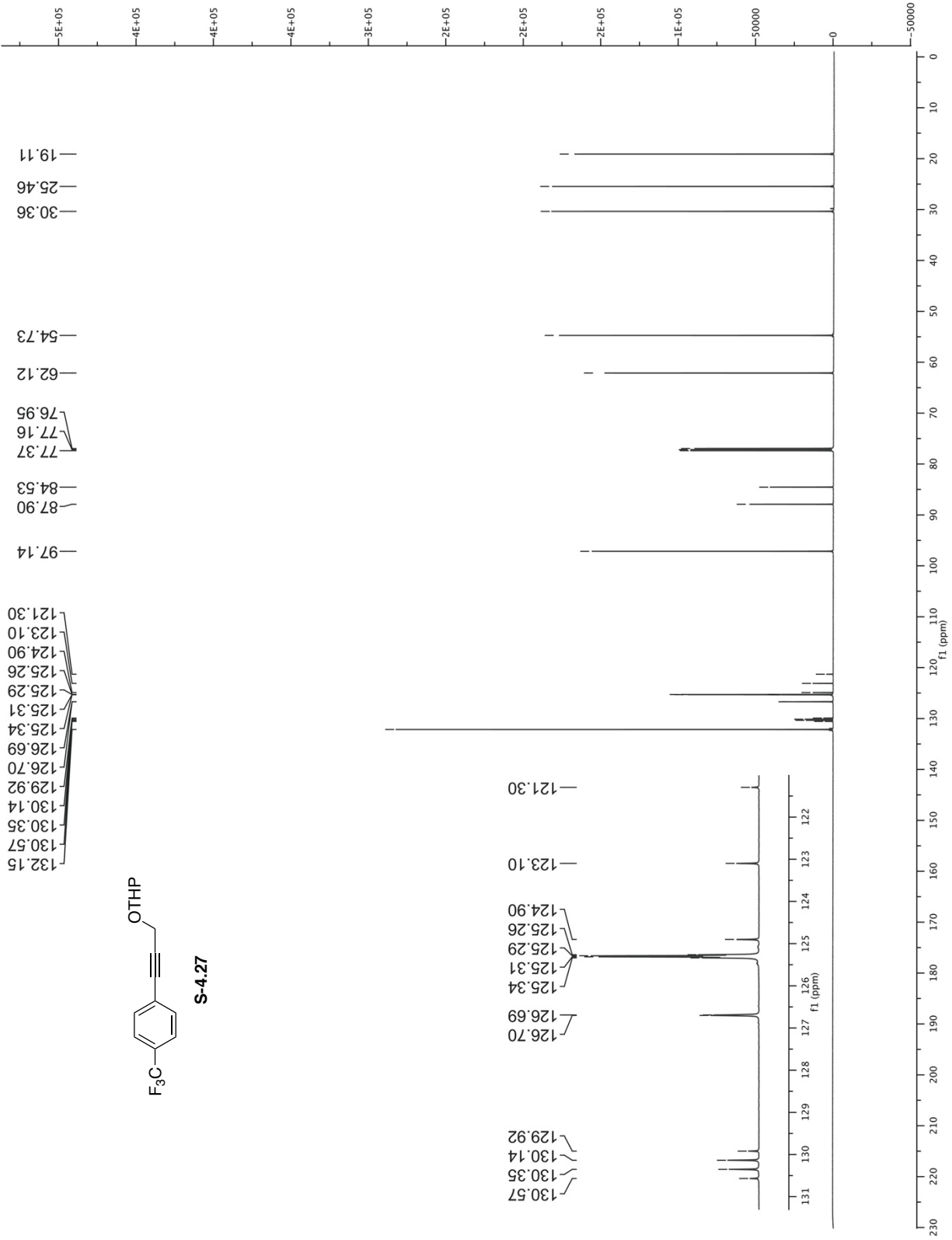


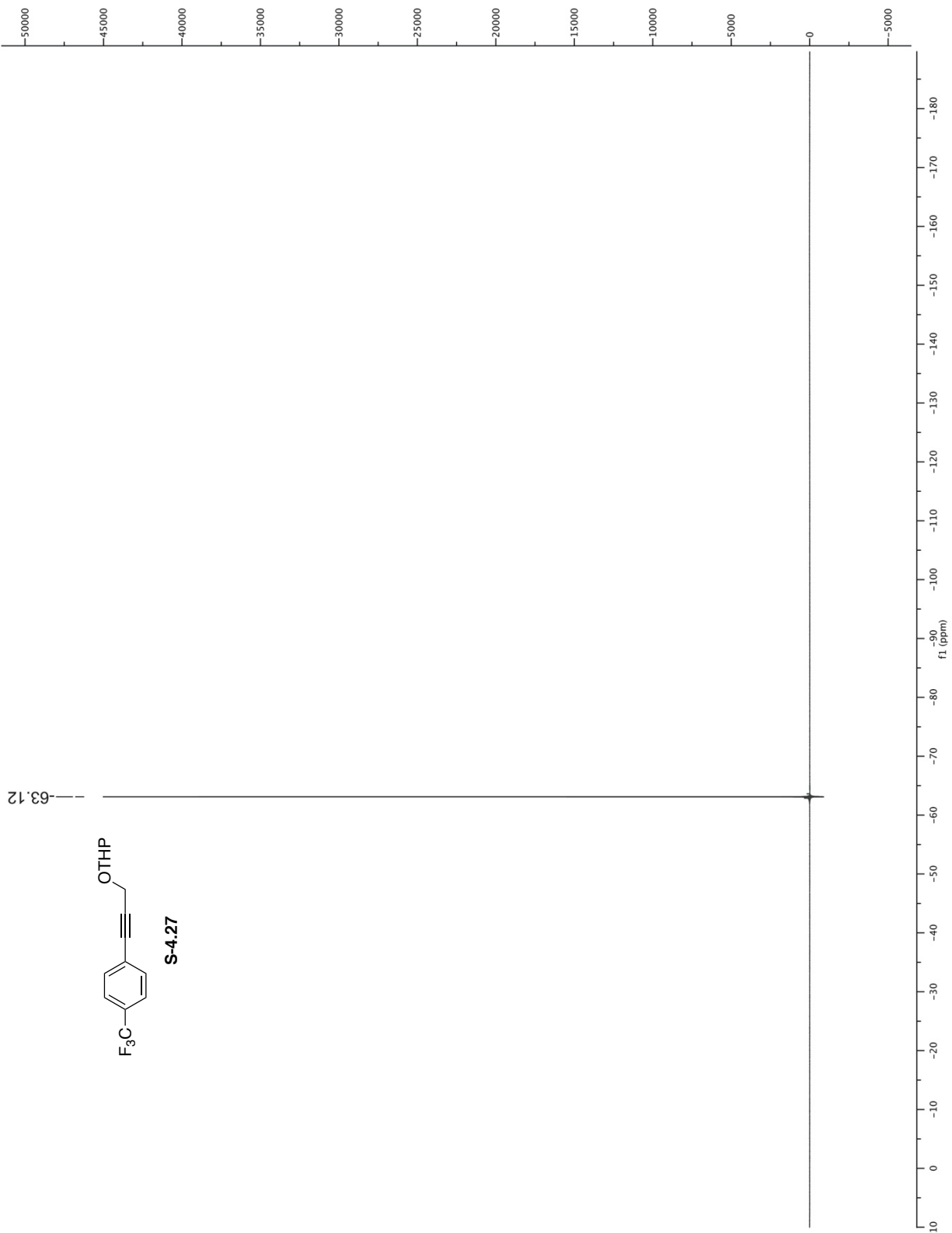


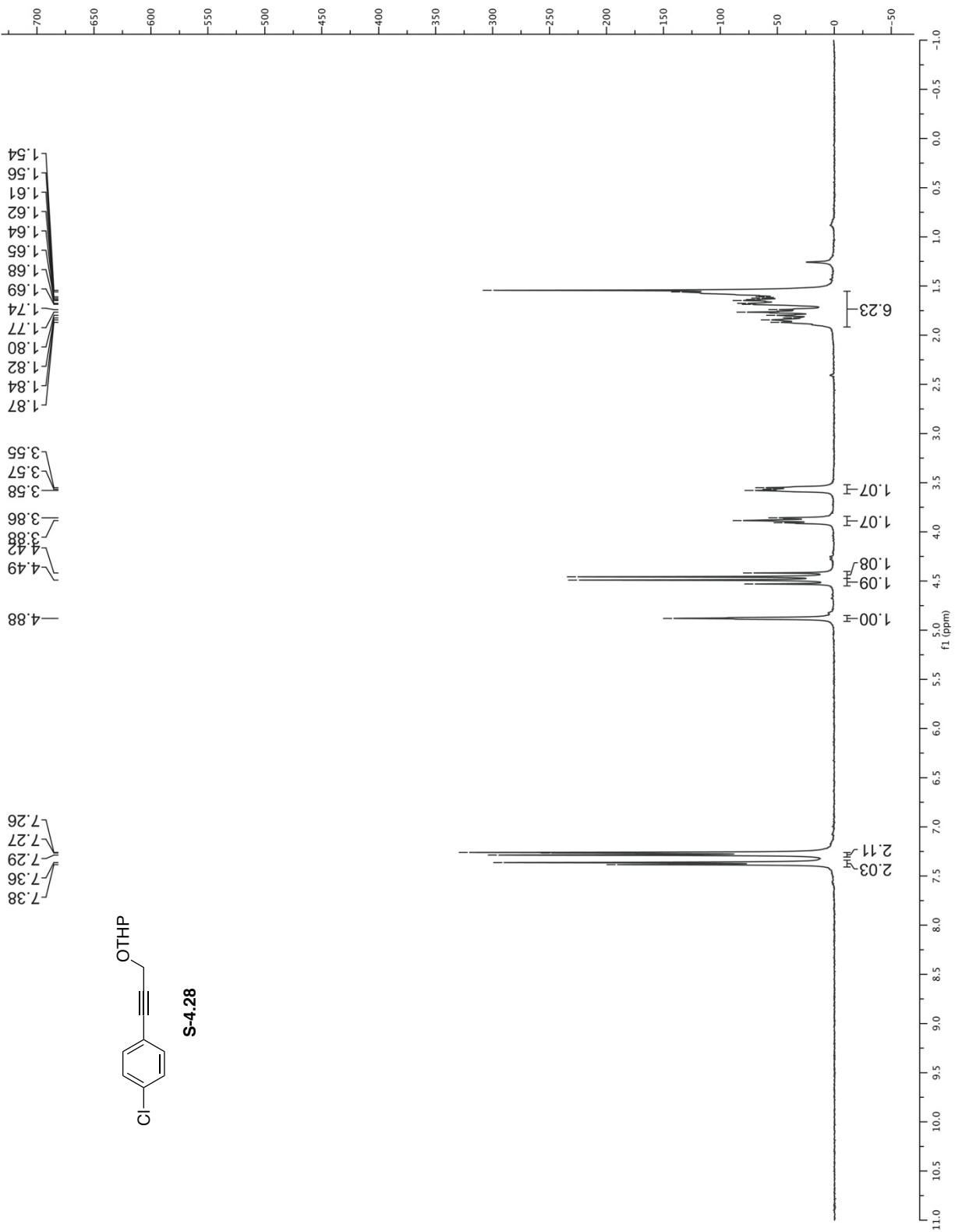


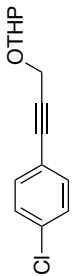
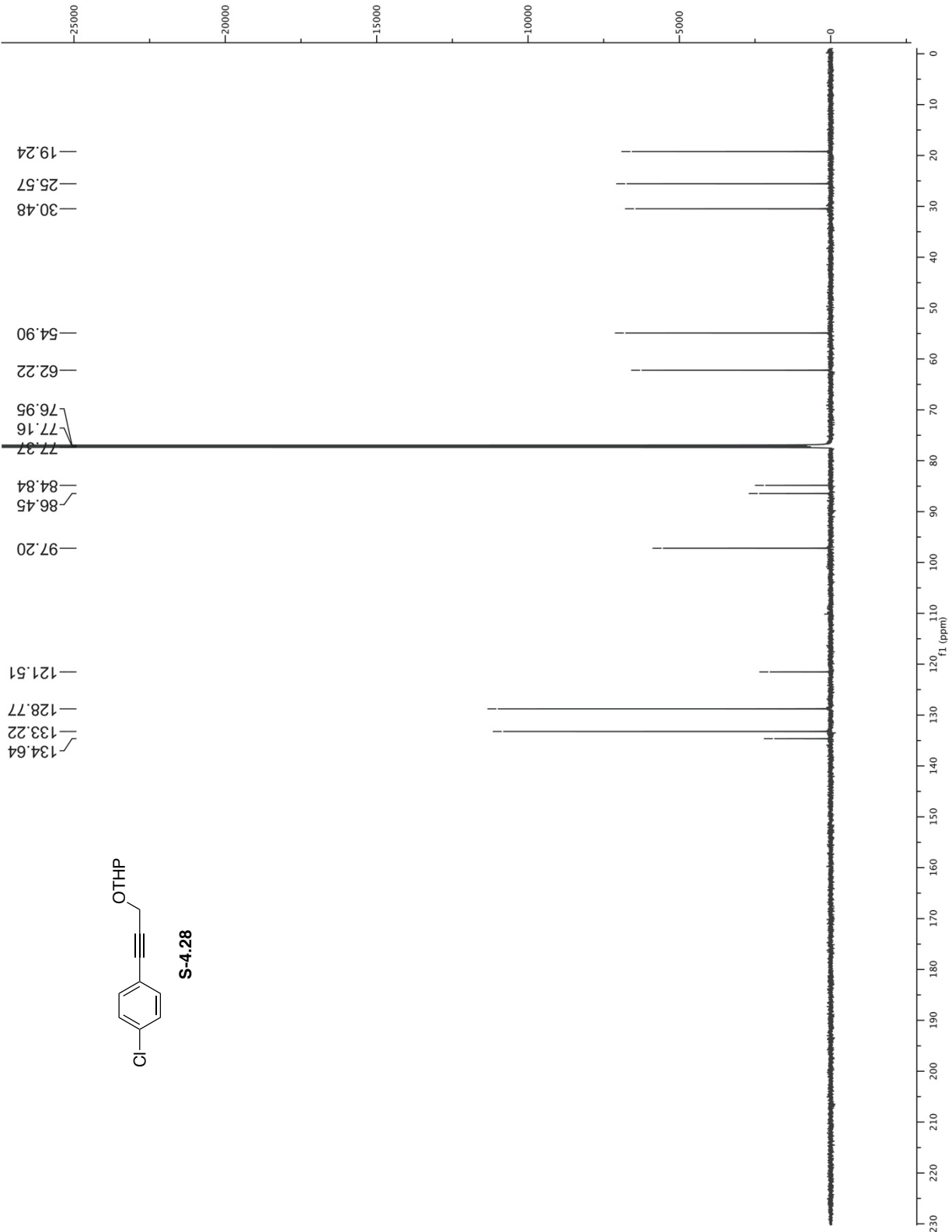




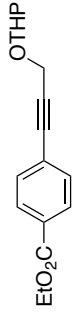
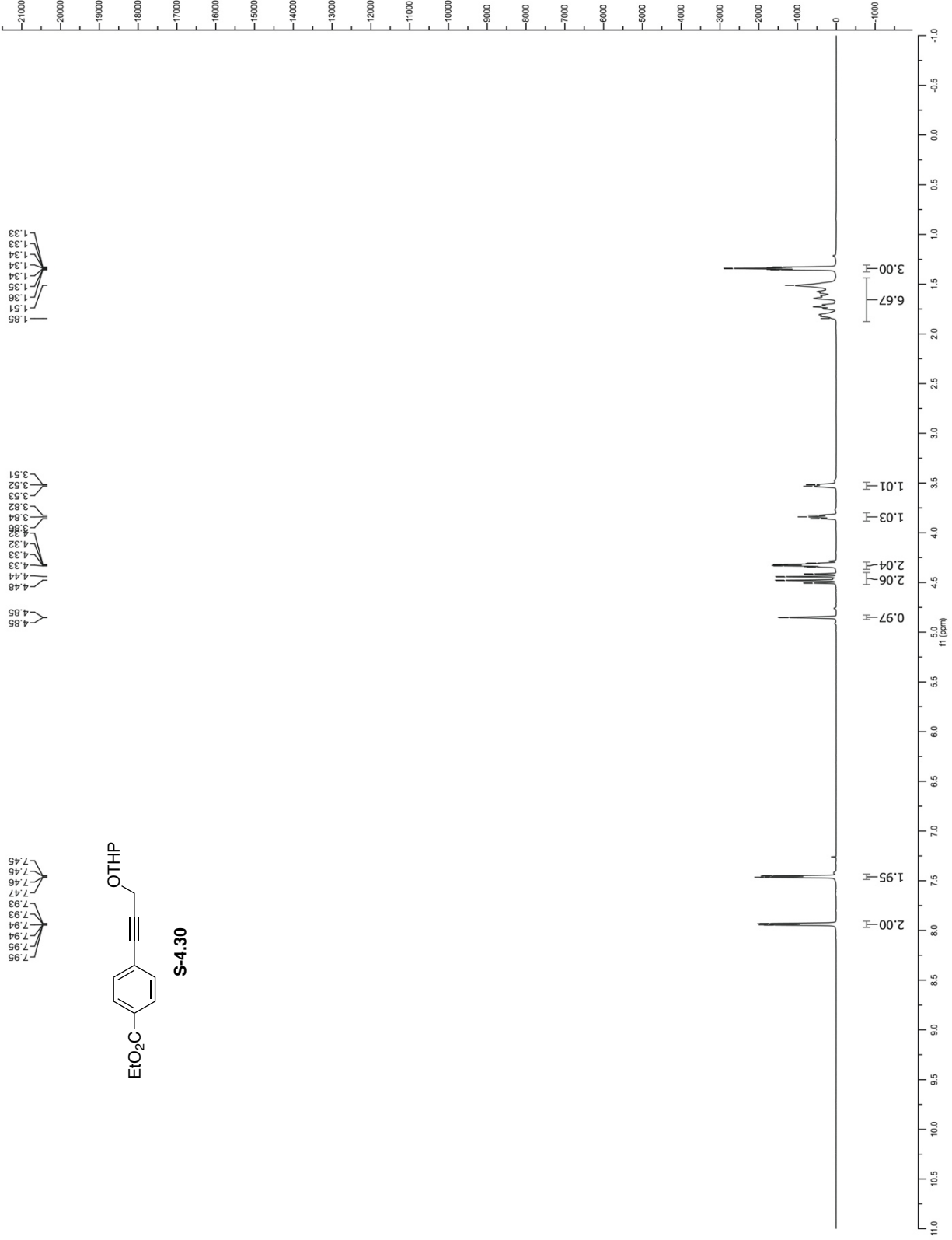




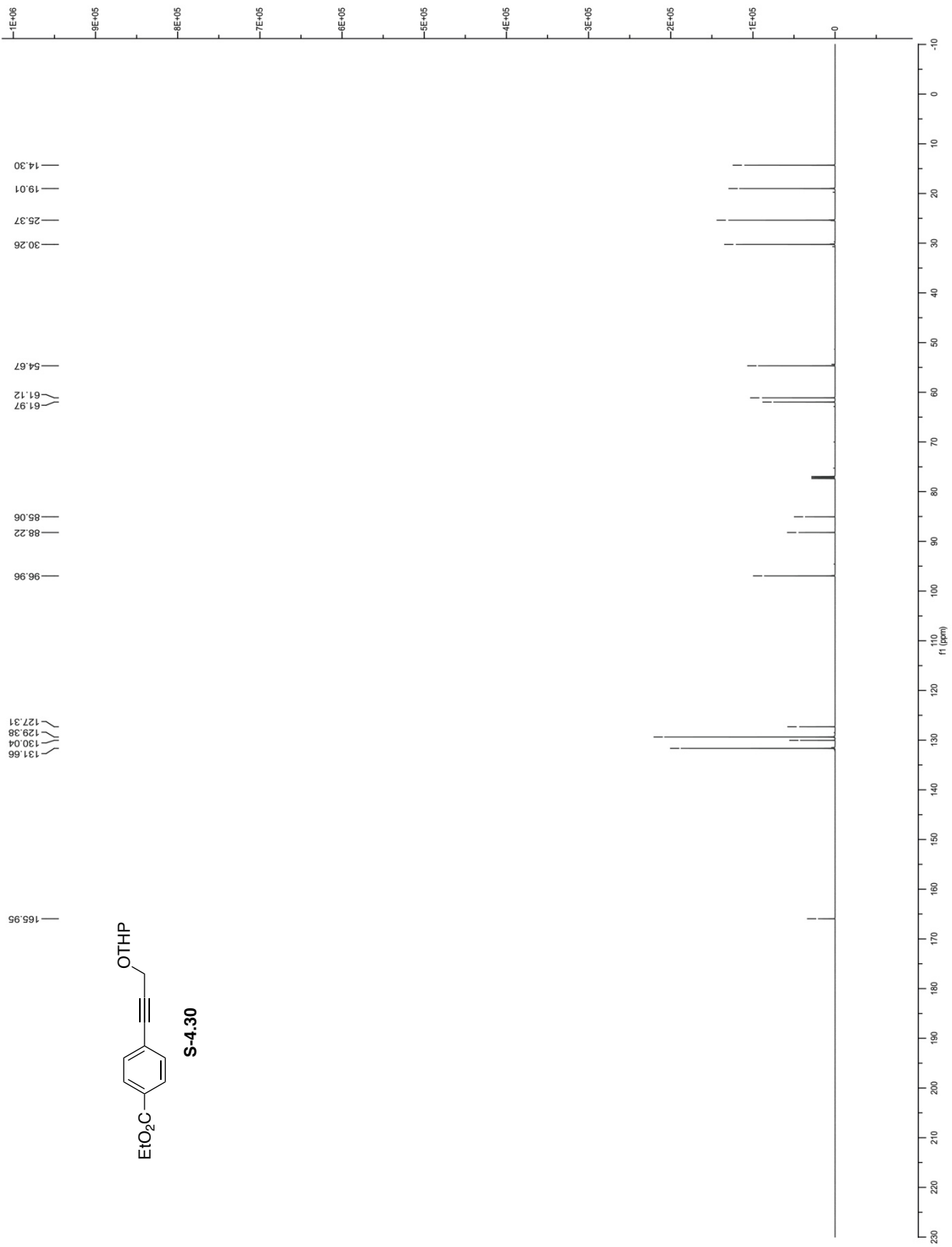


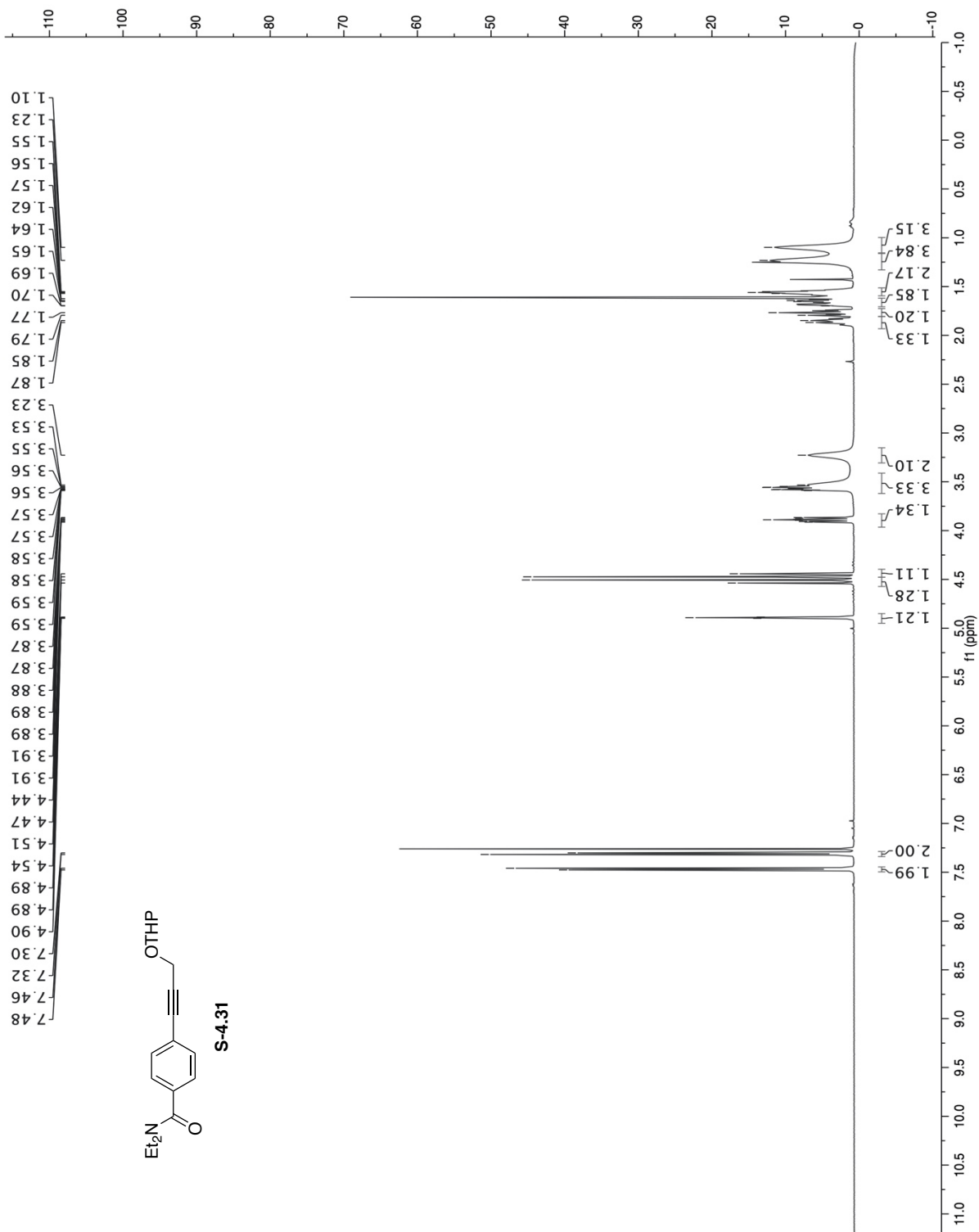


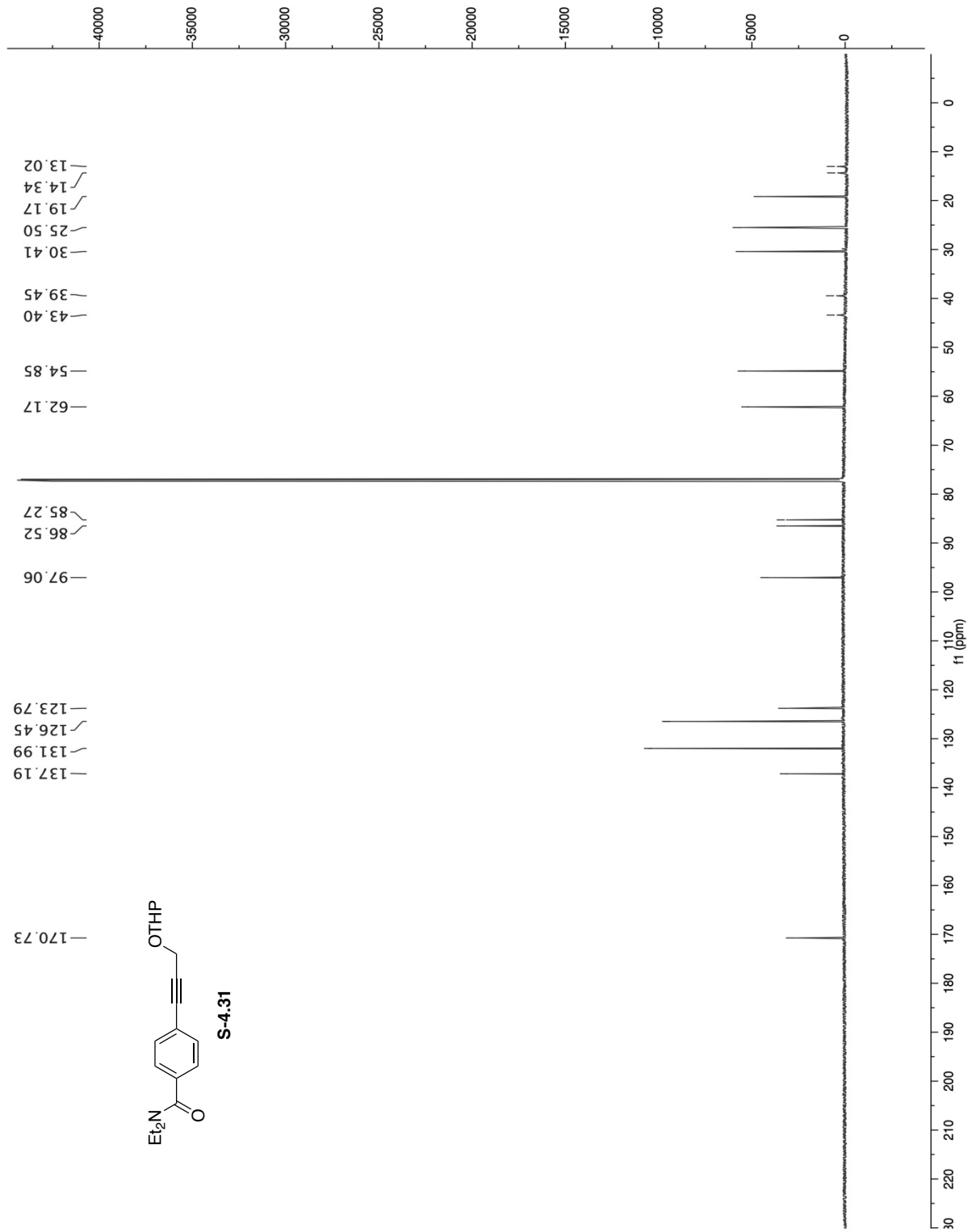
S-4.28

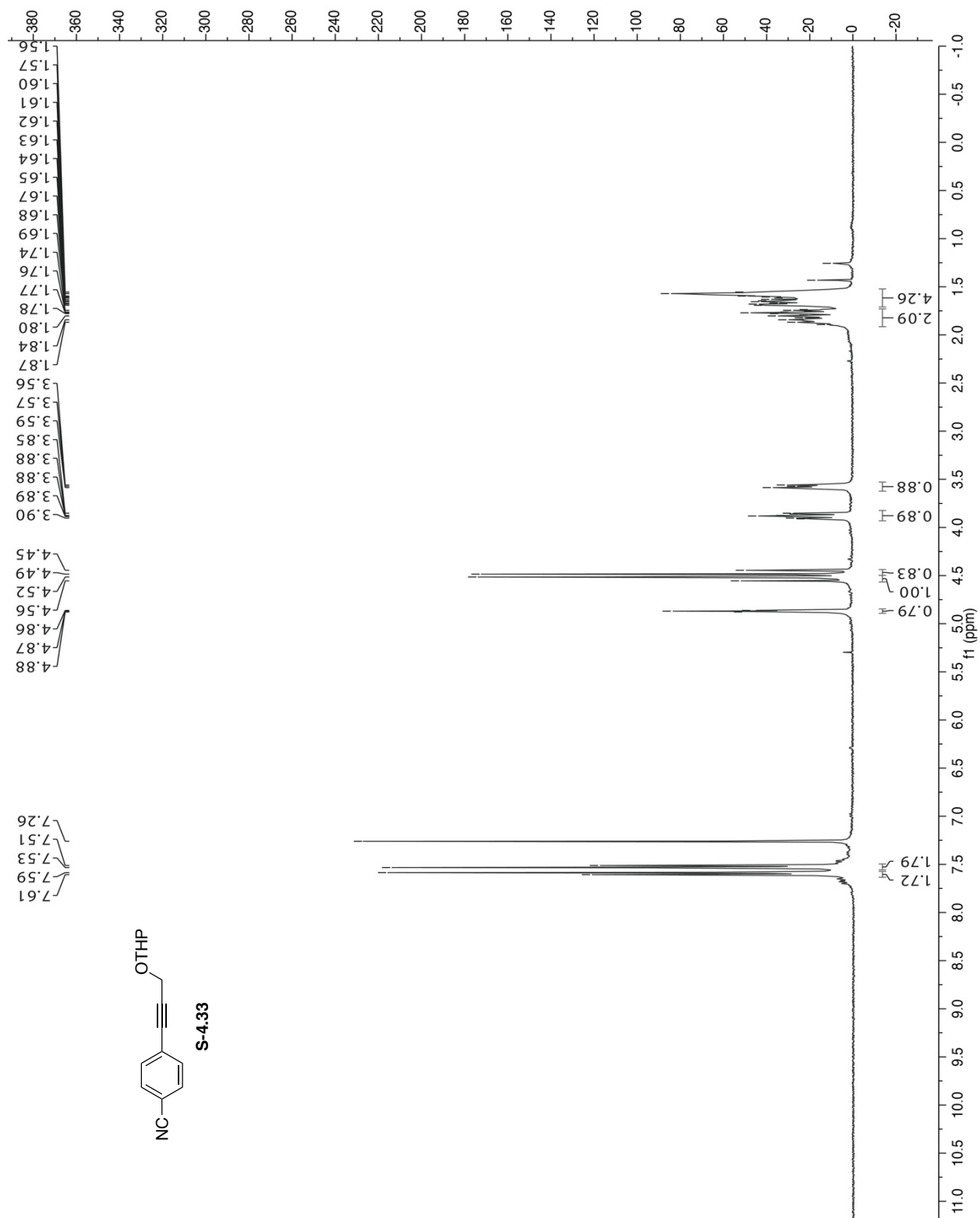


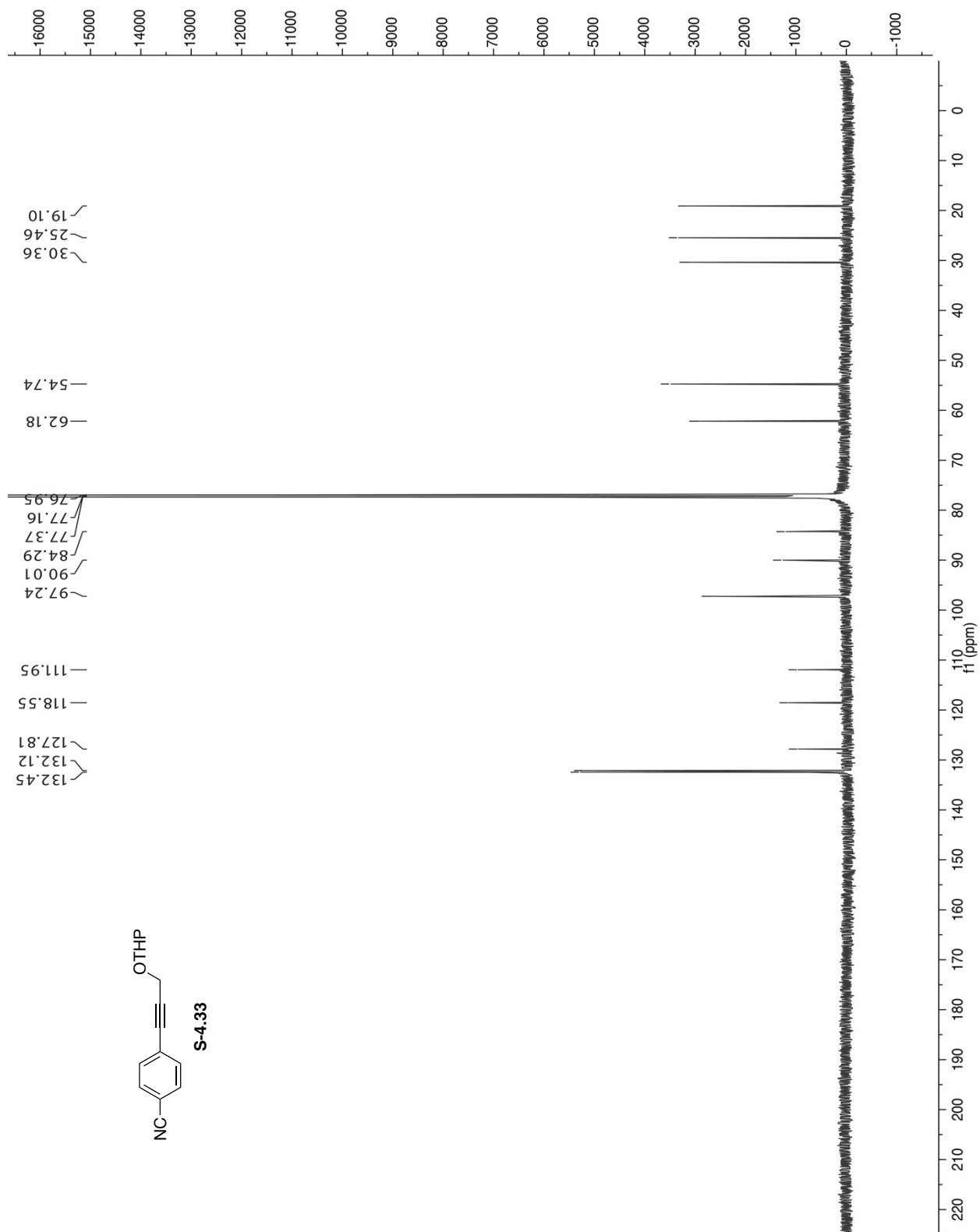
S-4.30

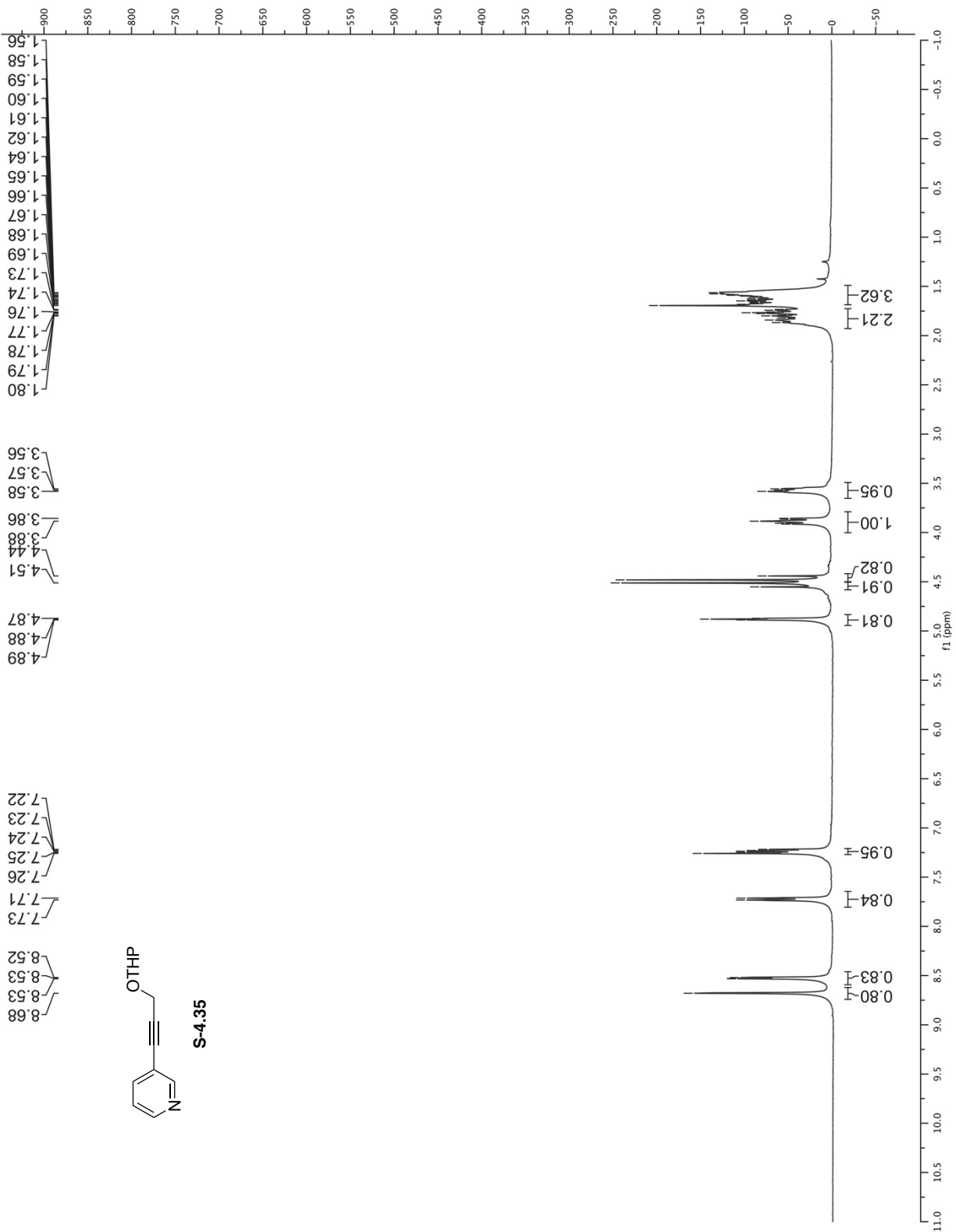


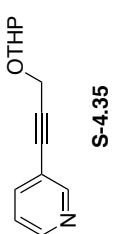
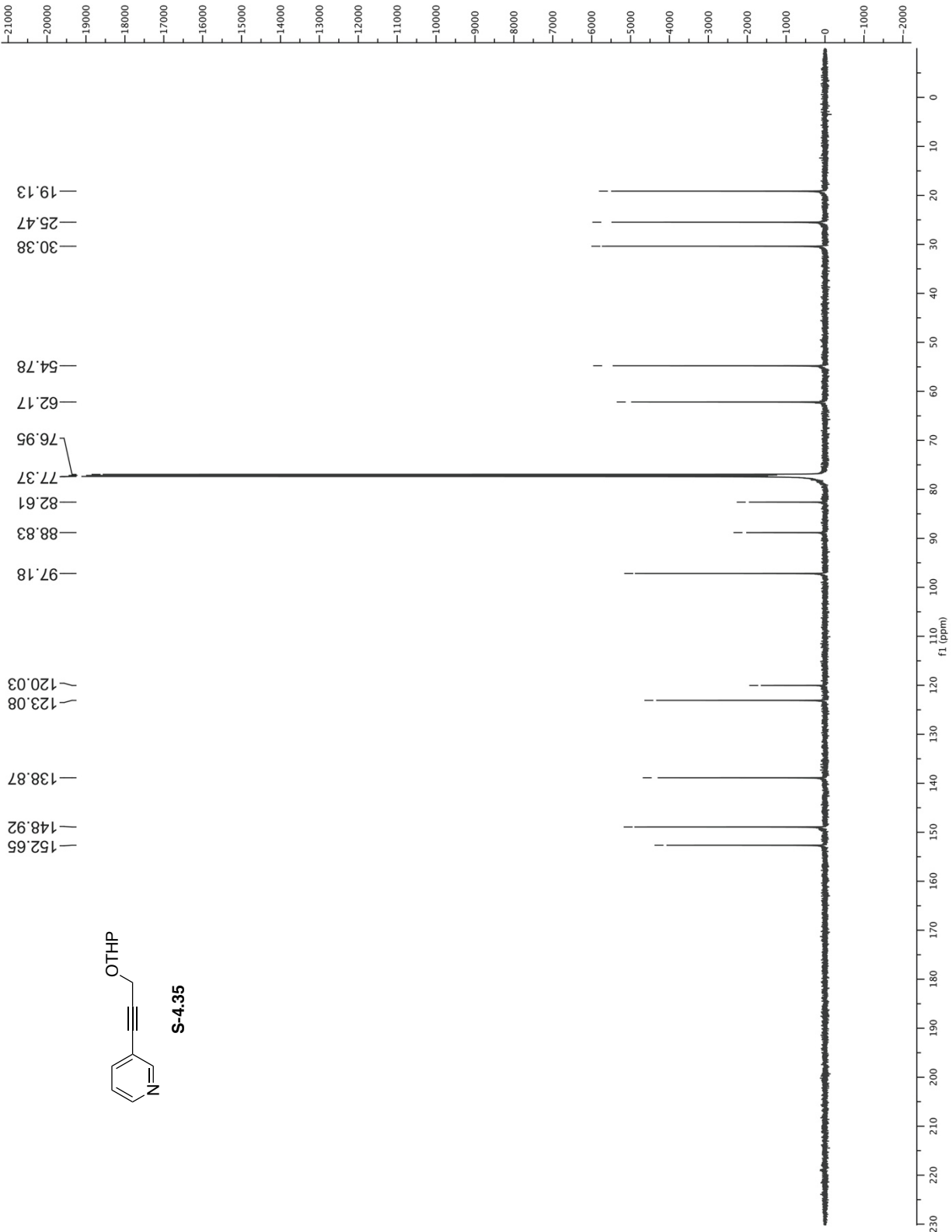


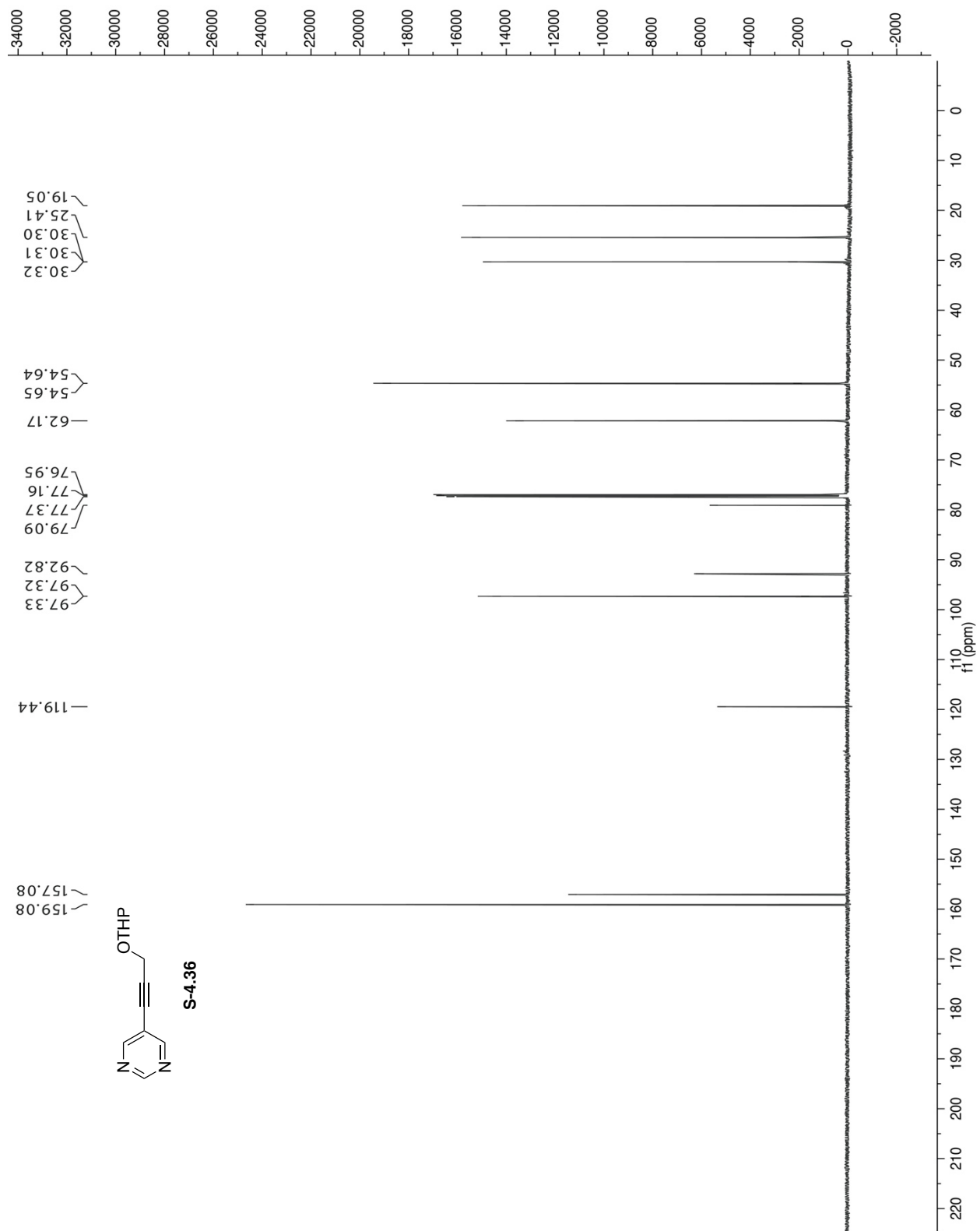


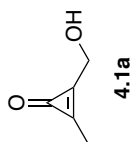
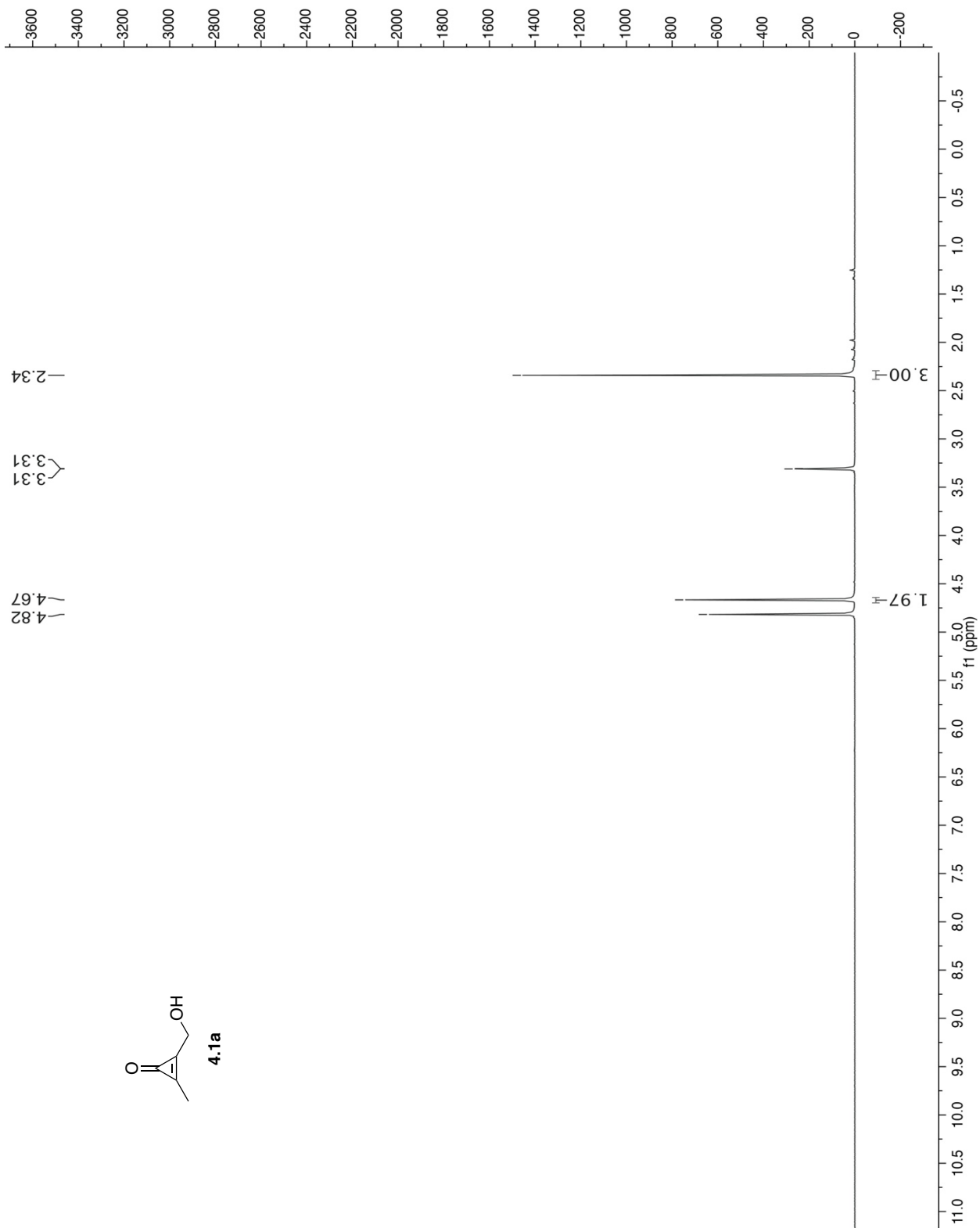


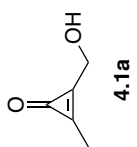
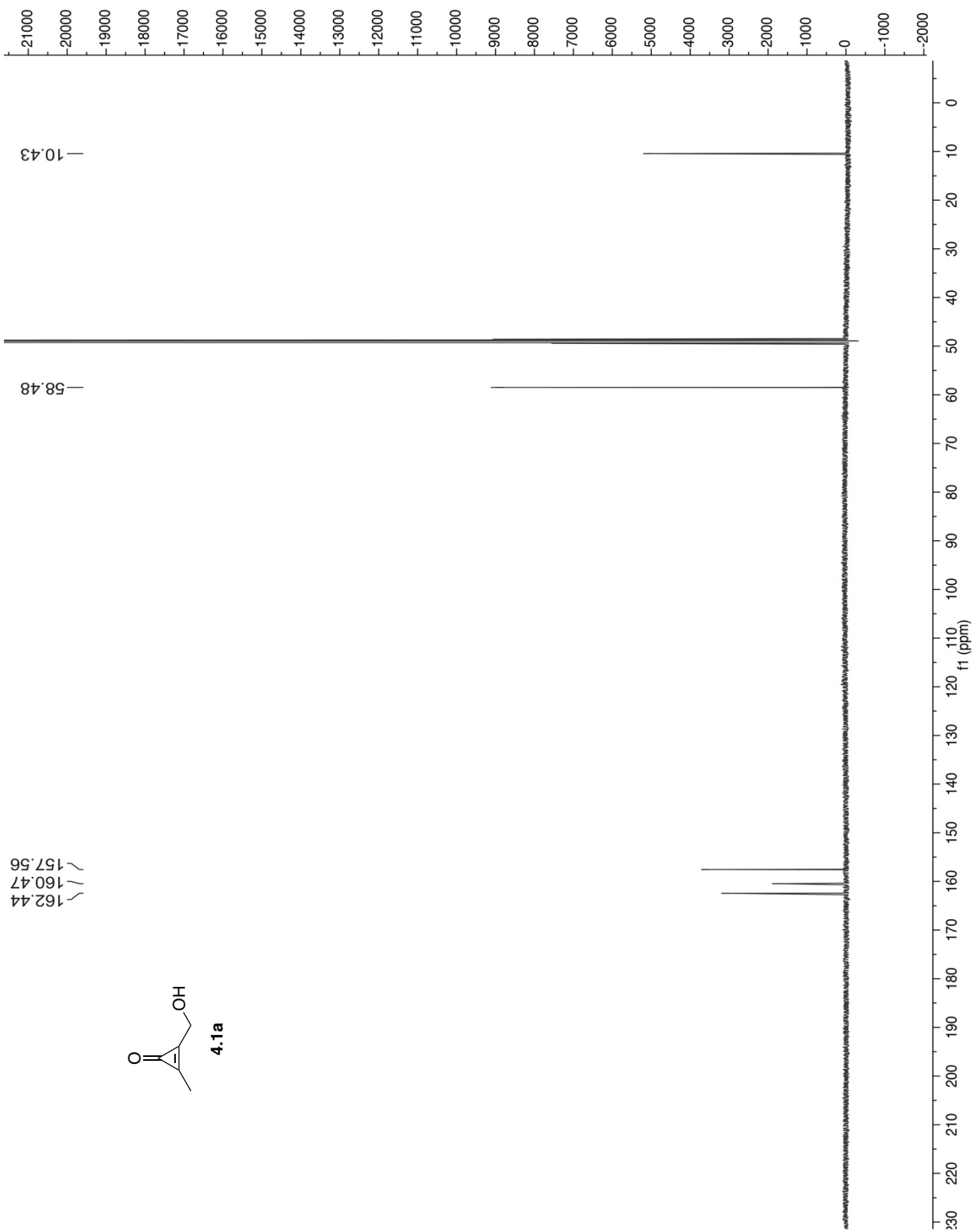


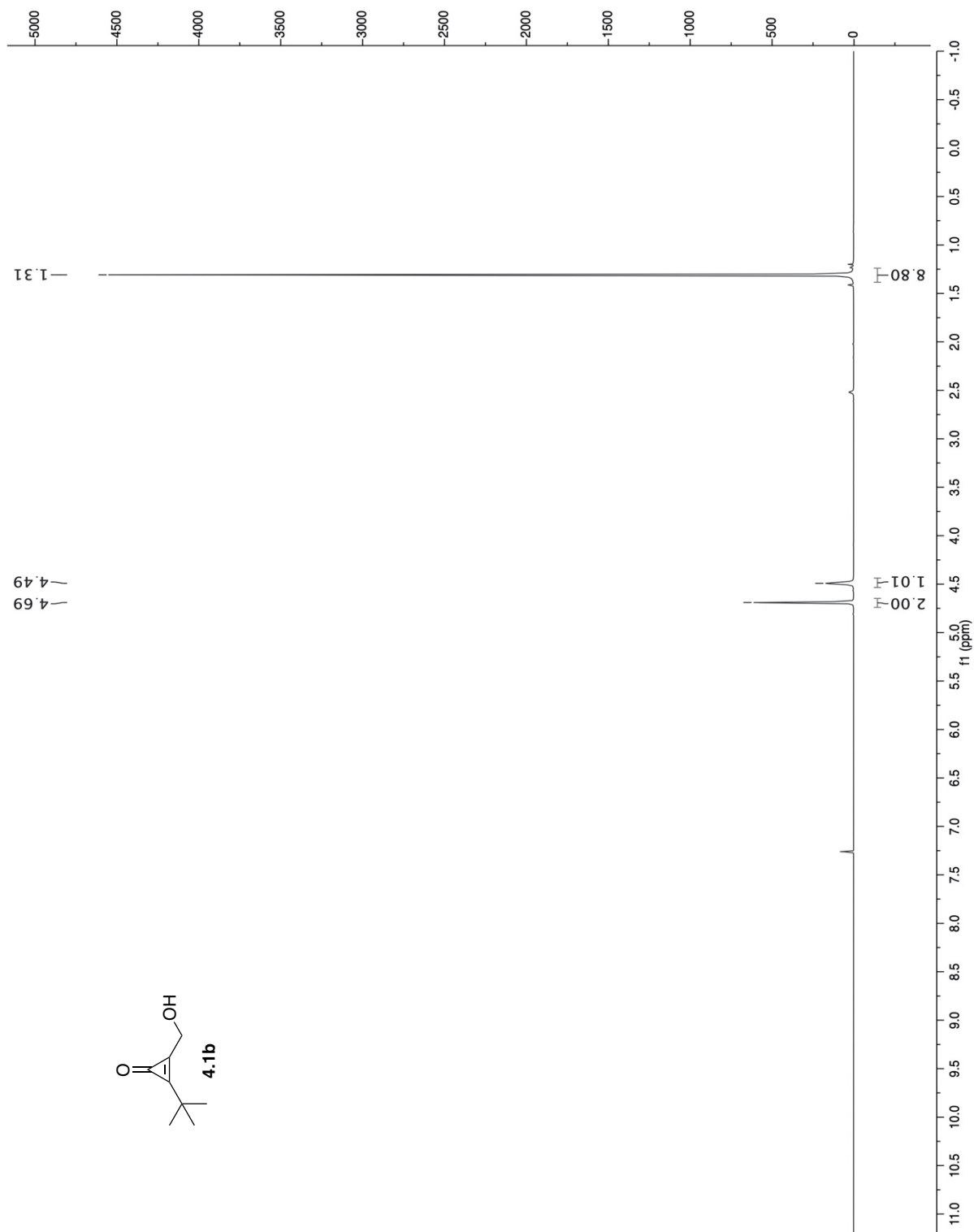


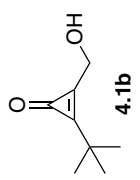
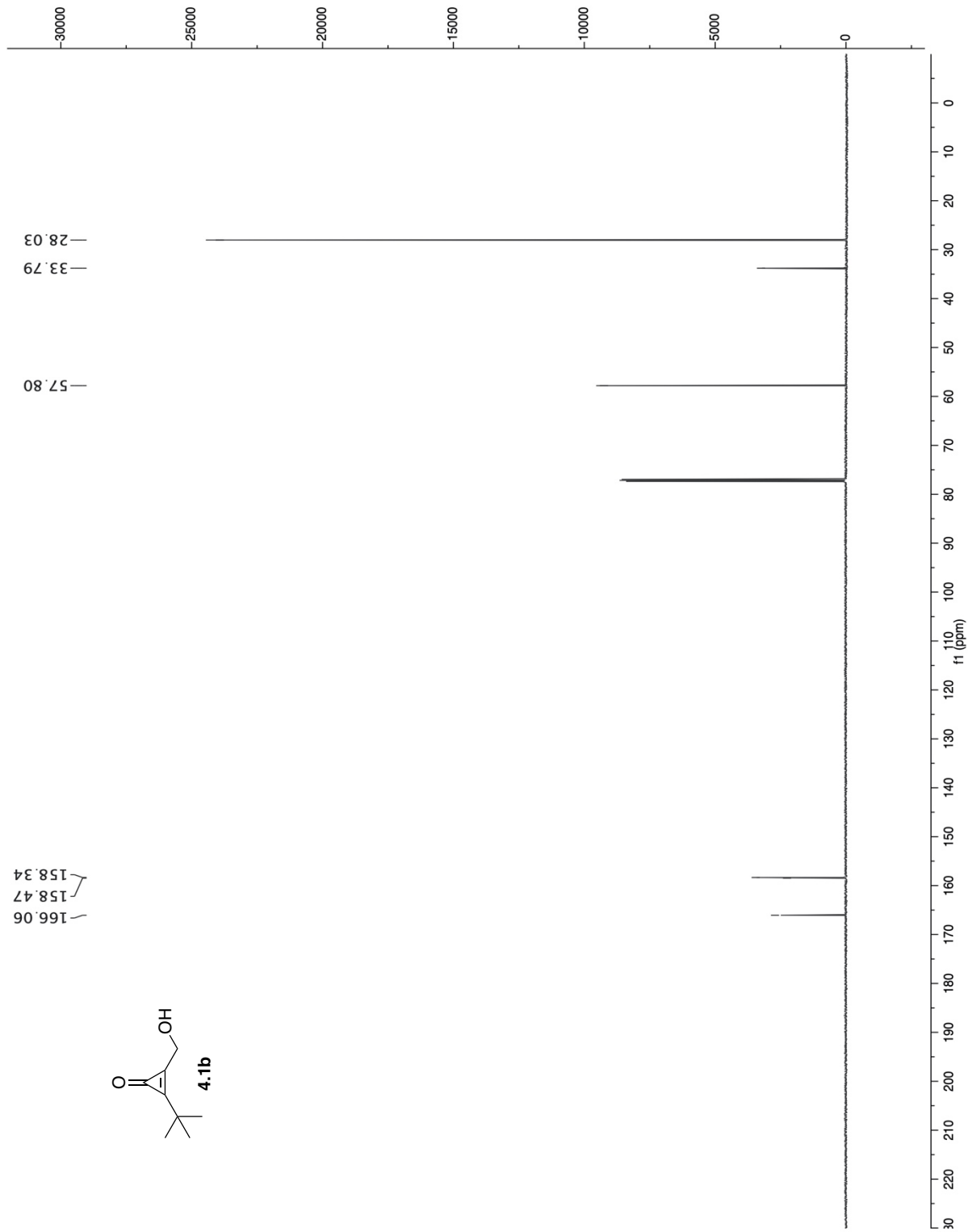


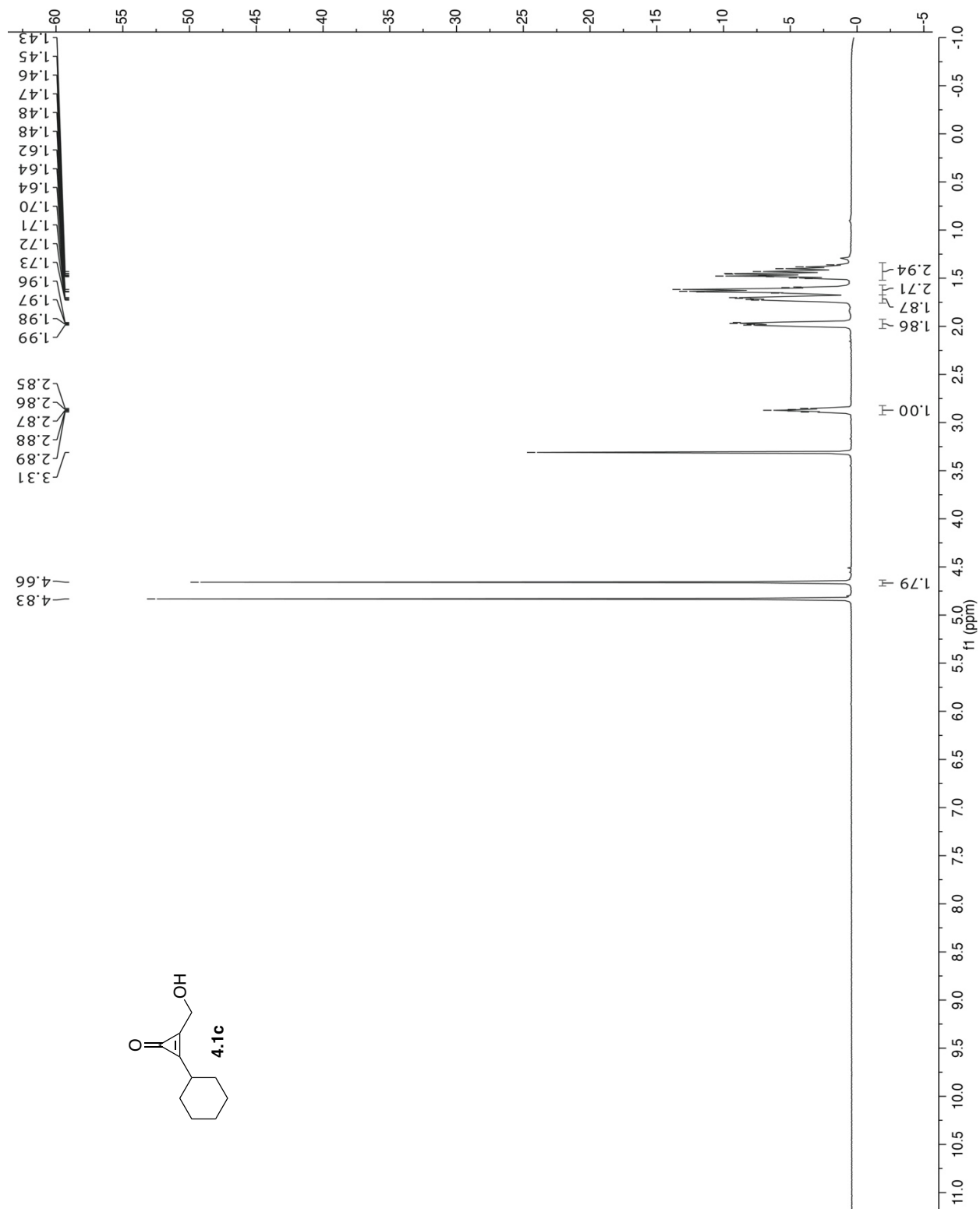


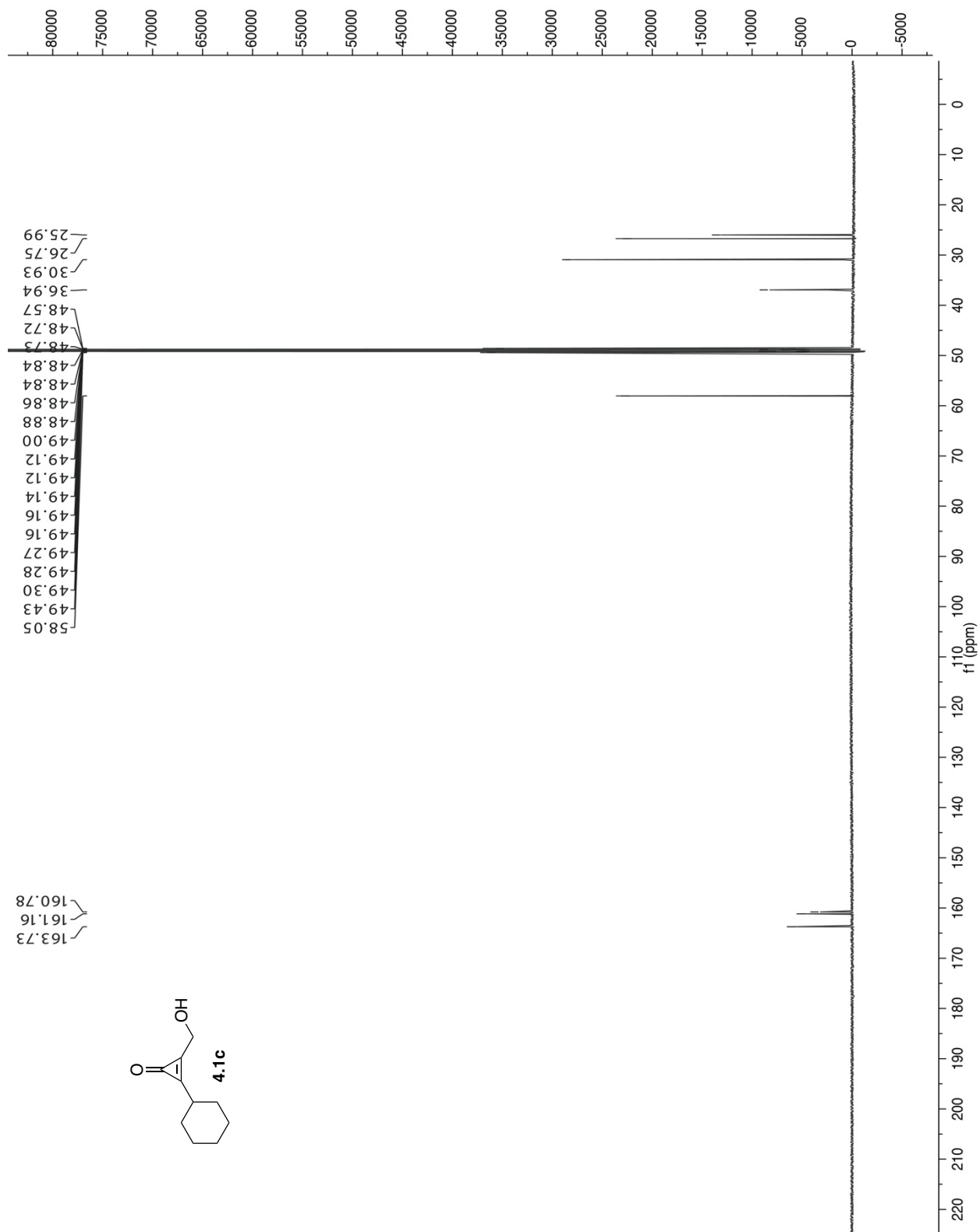


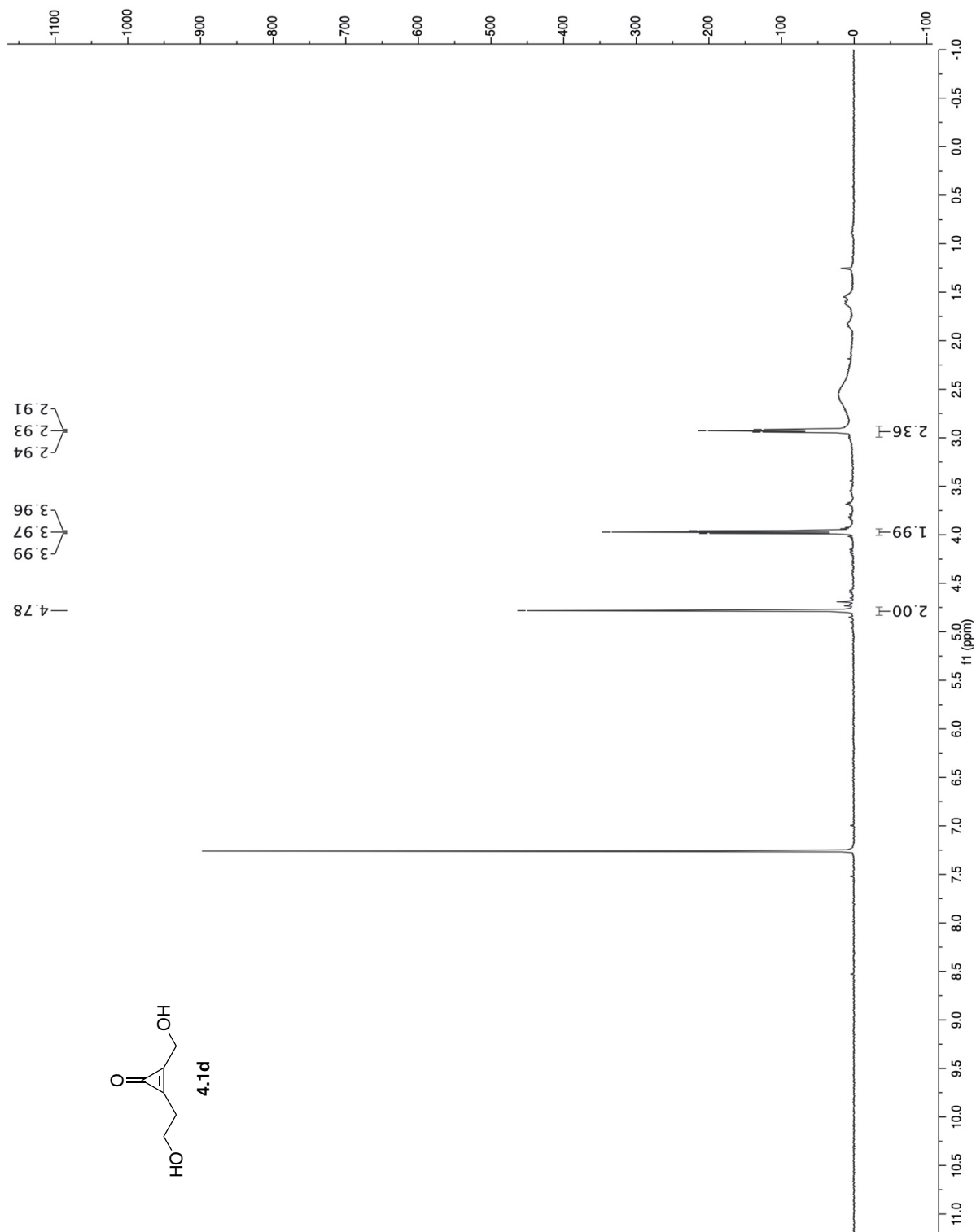


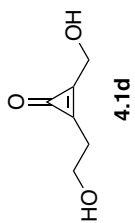
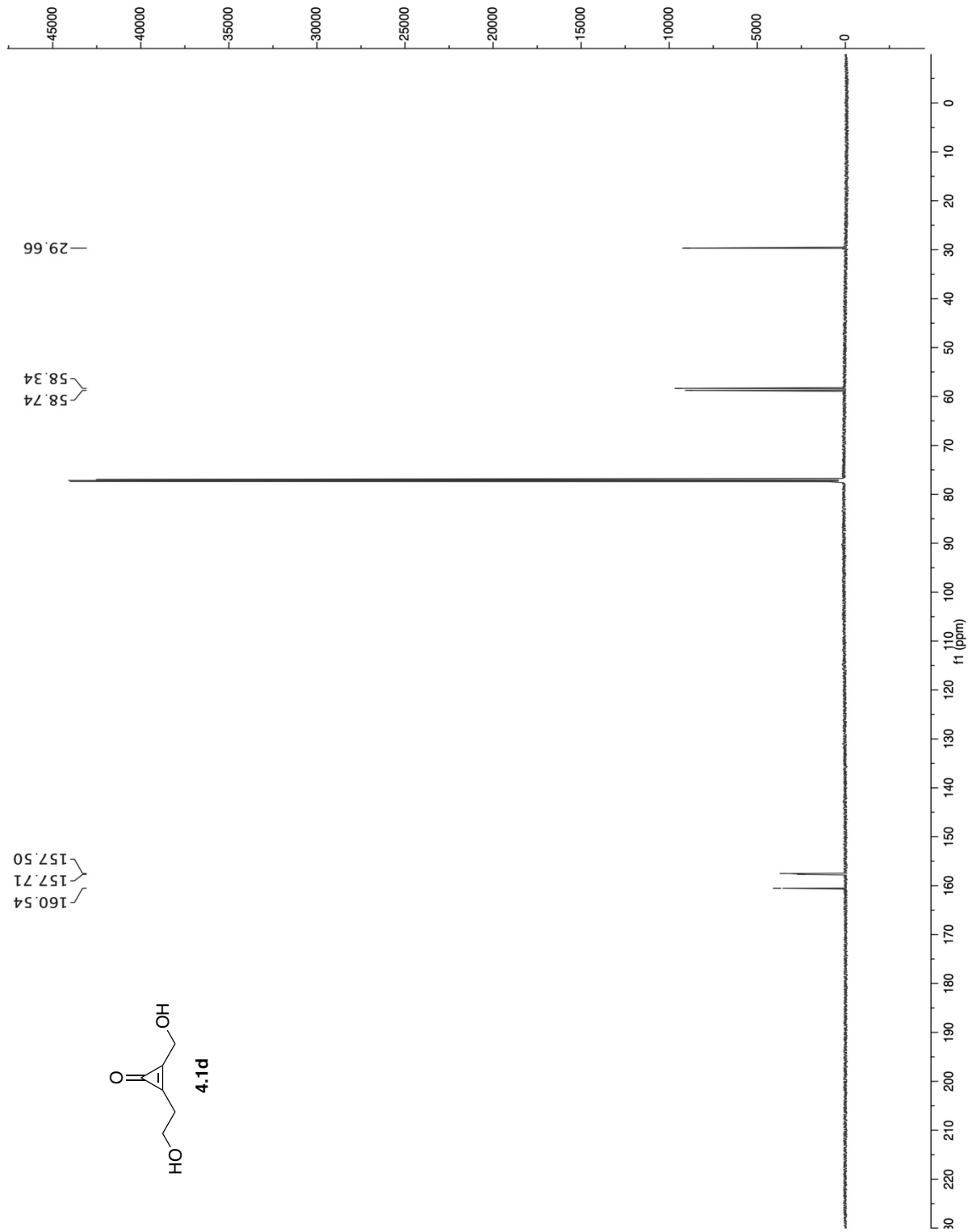


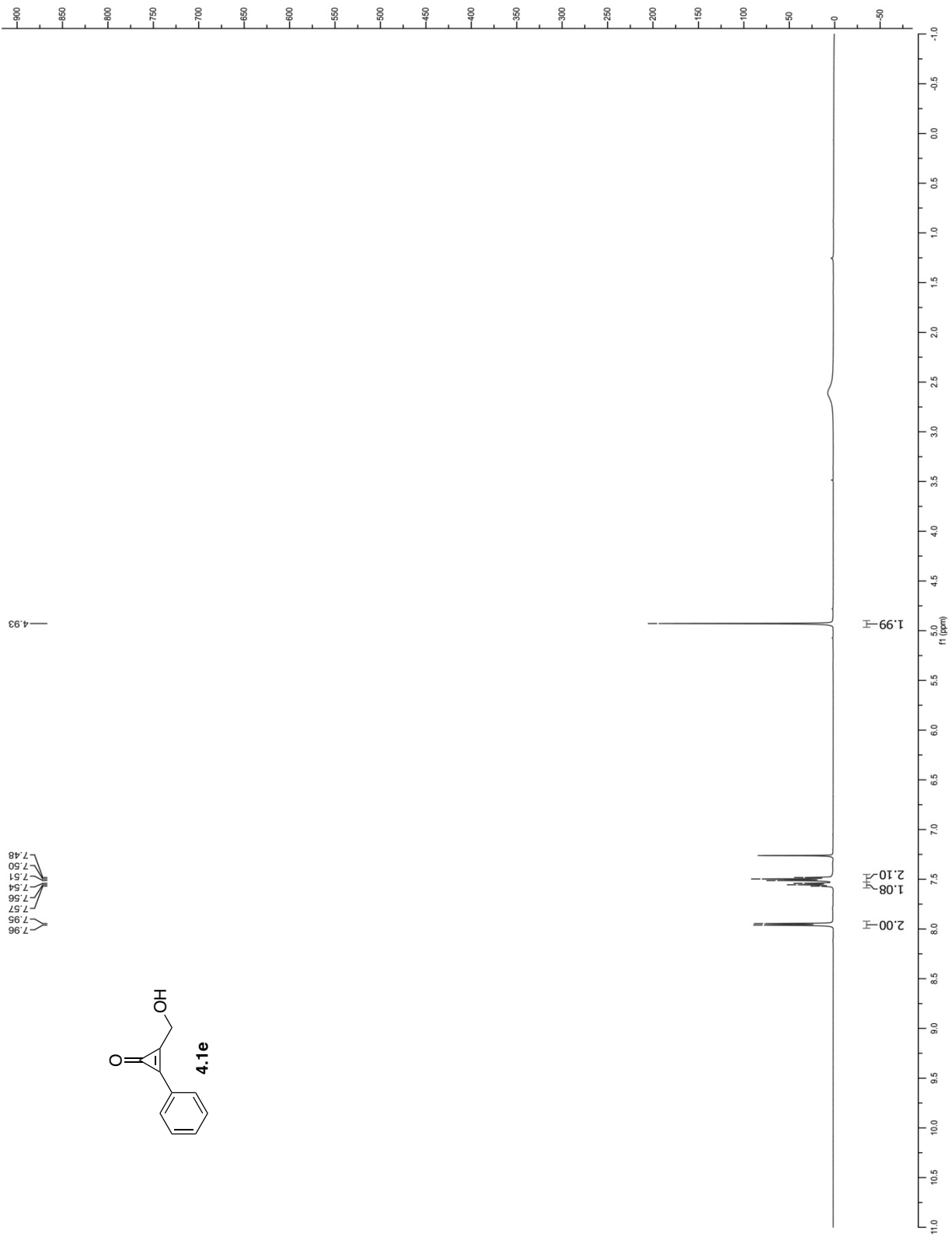


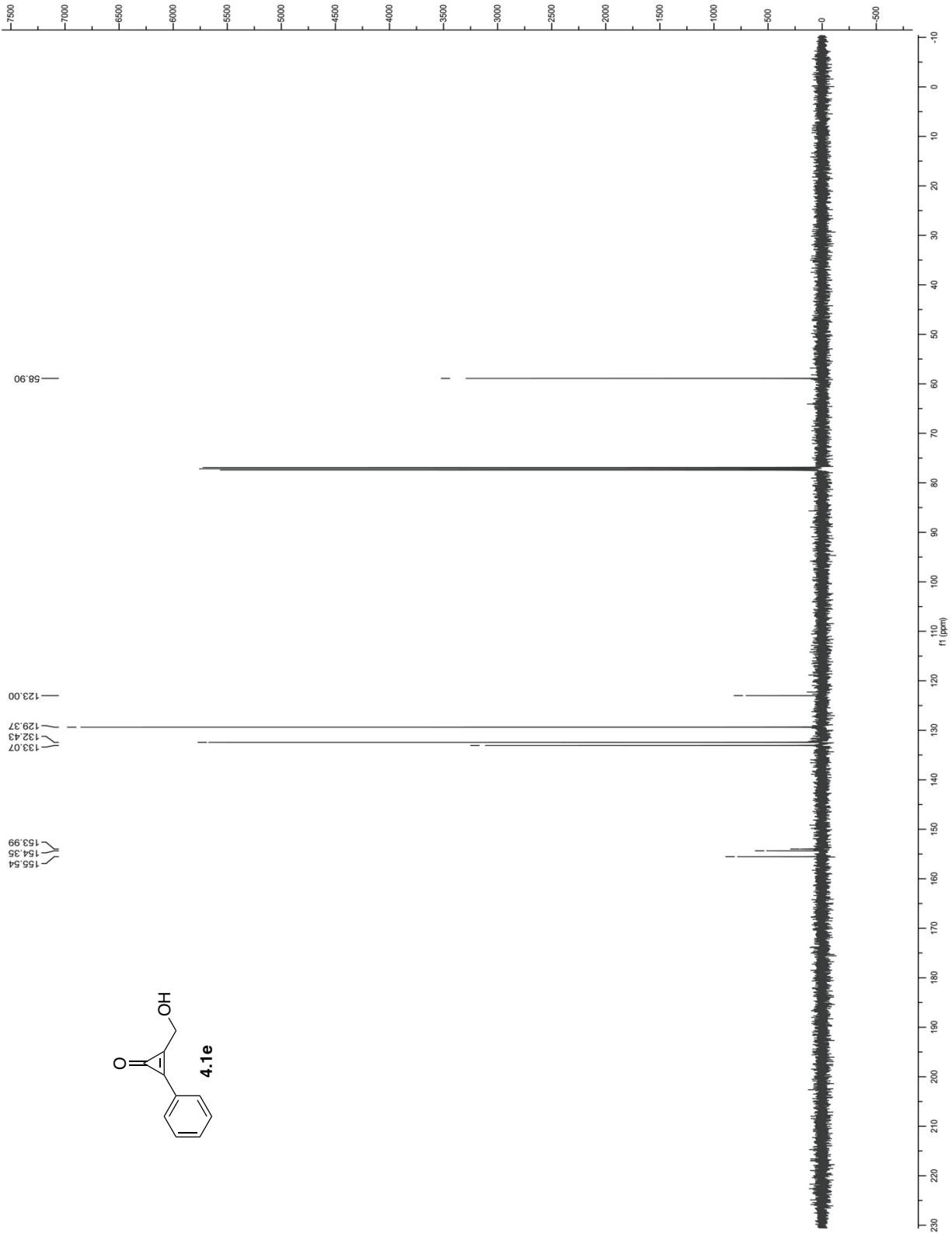


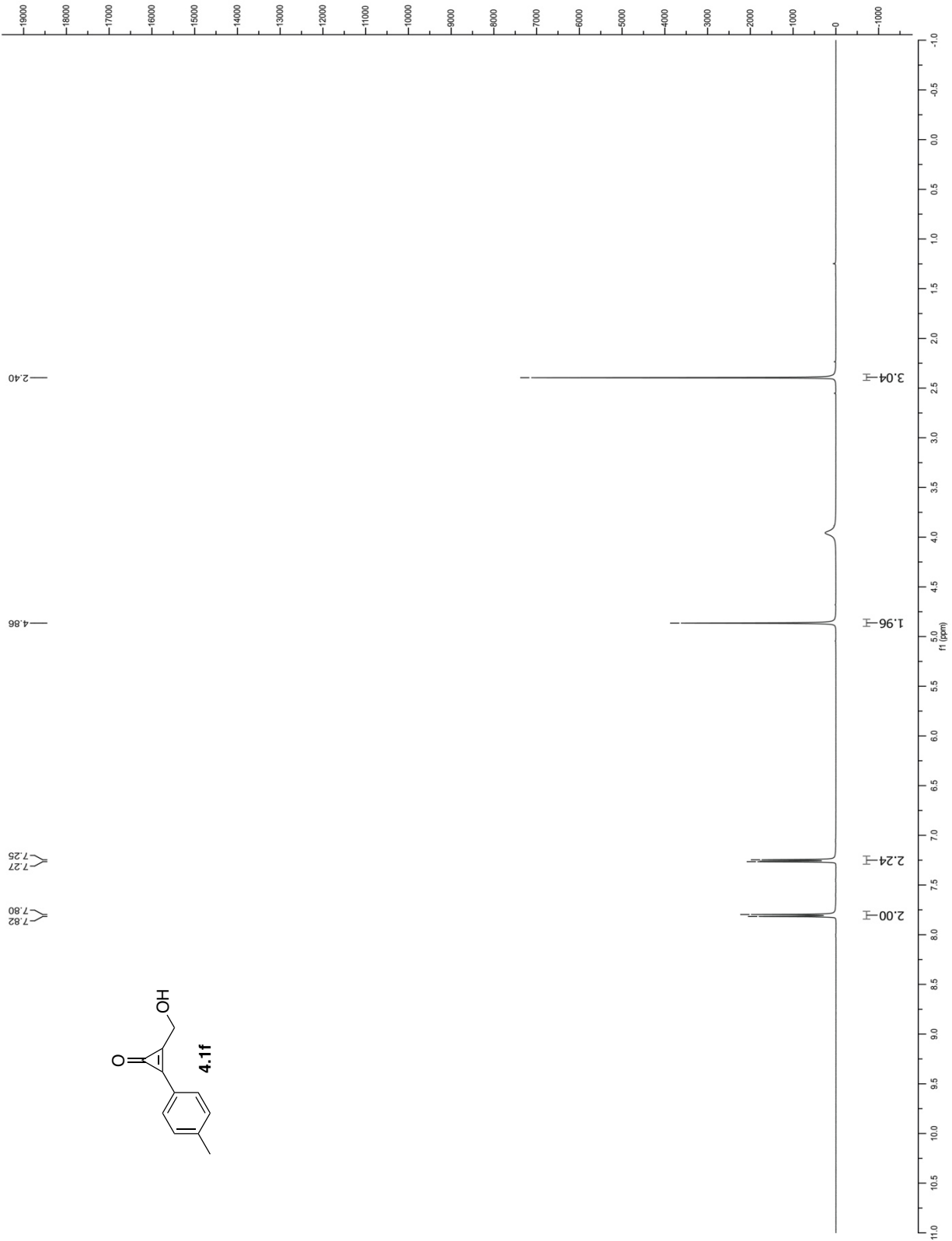


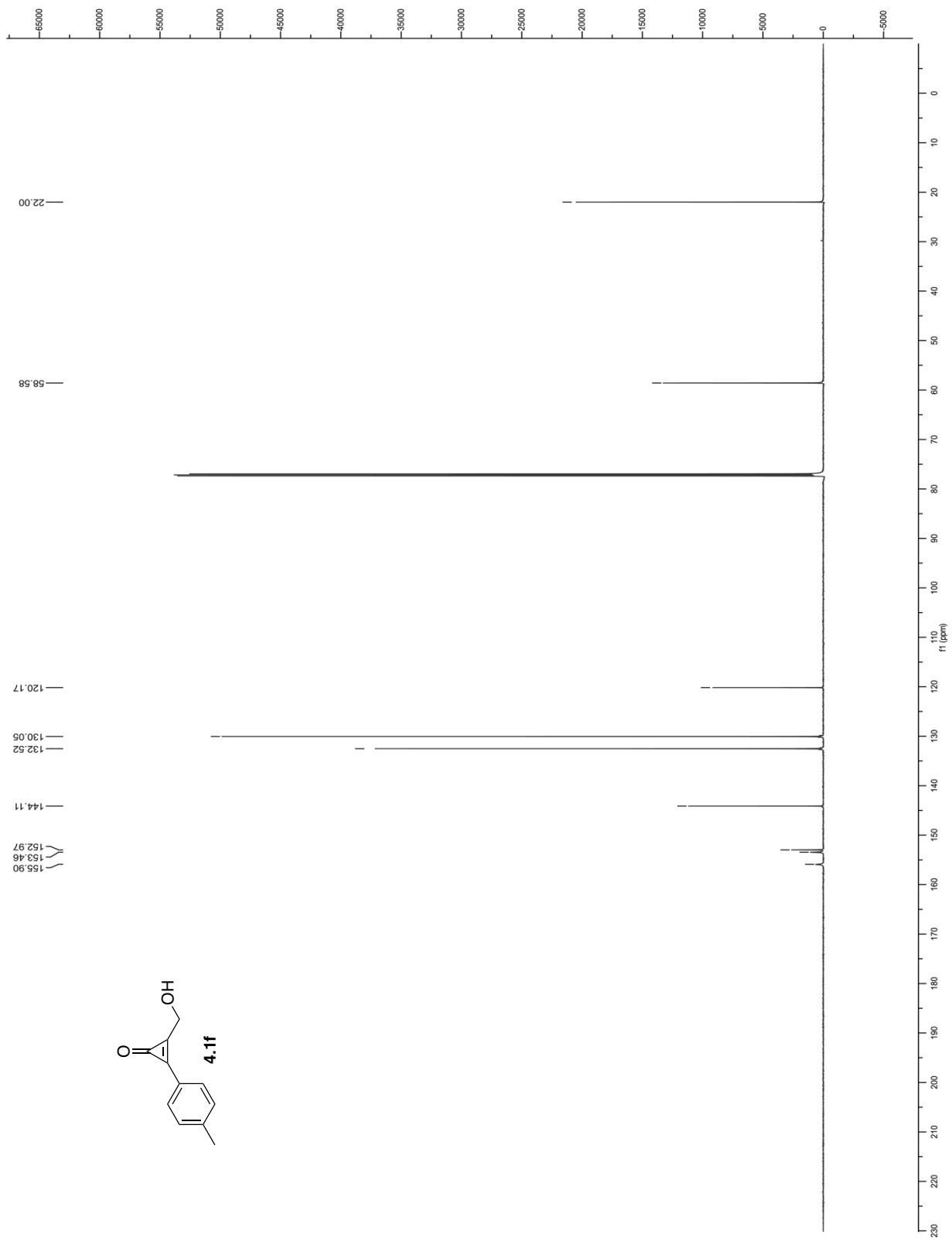


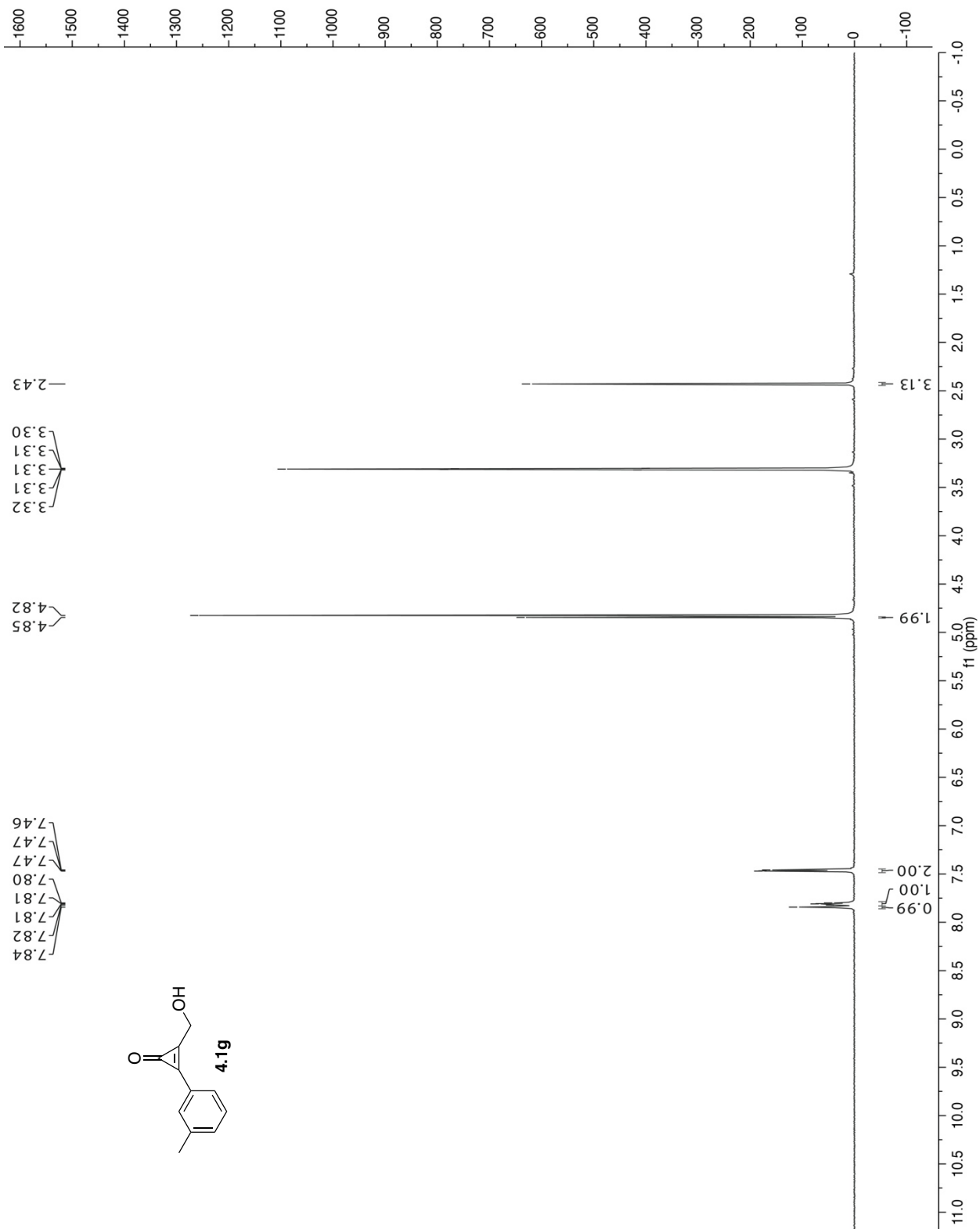


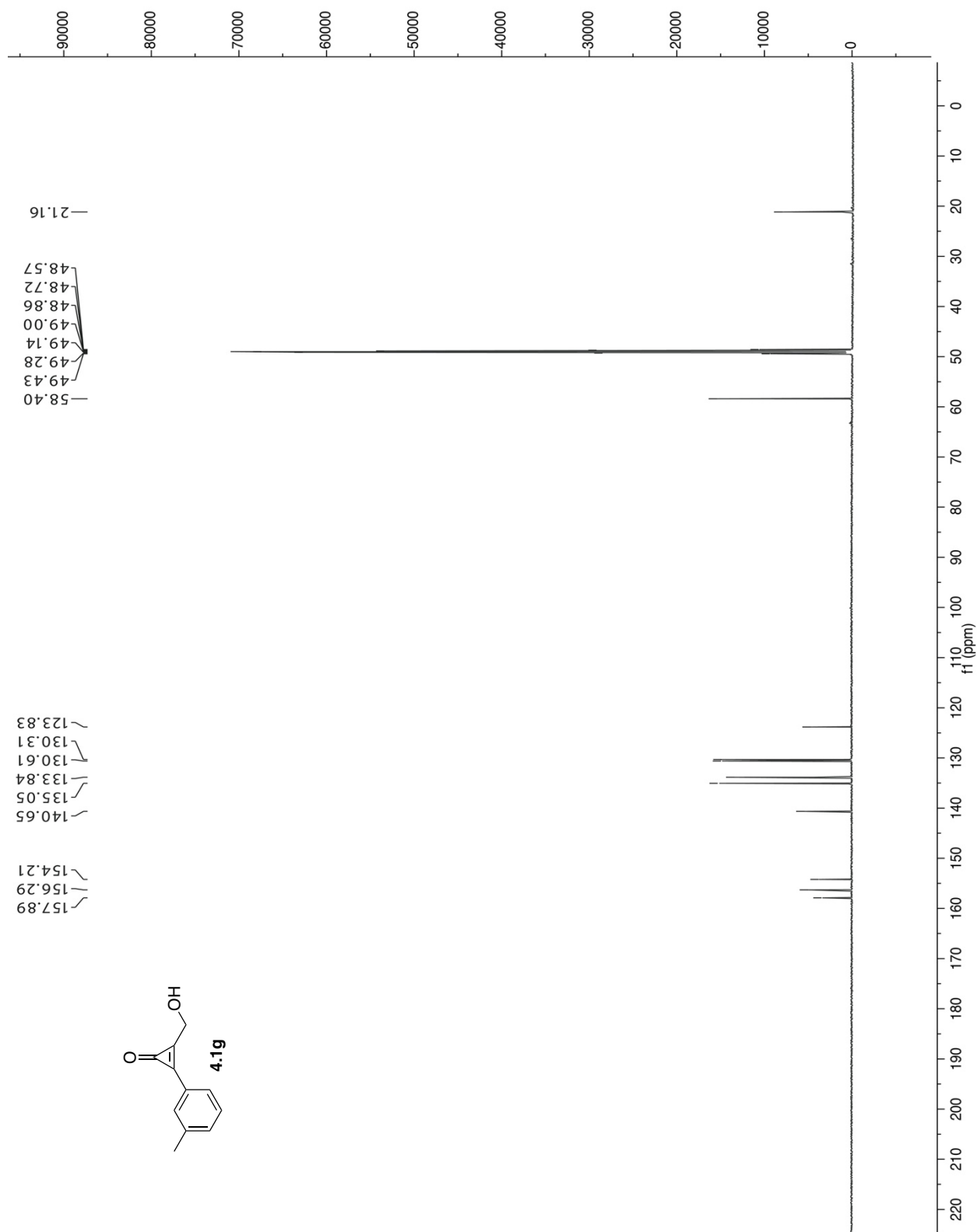


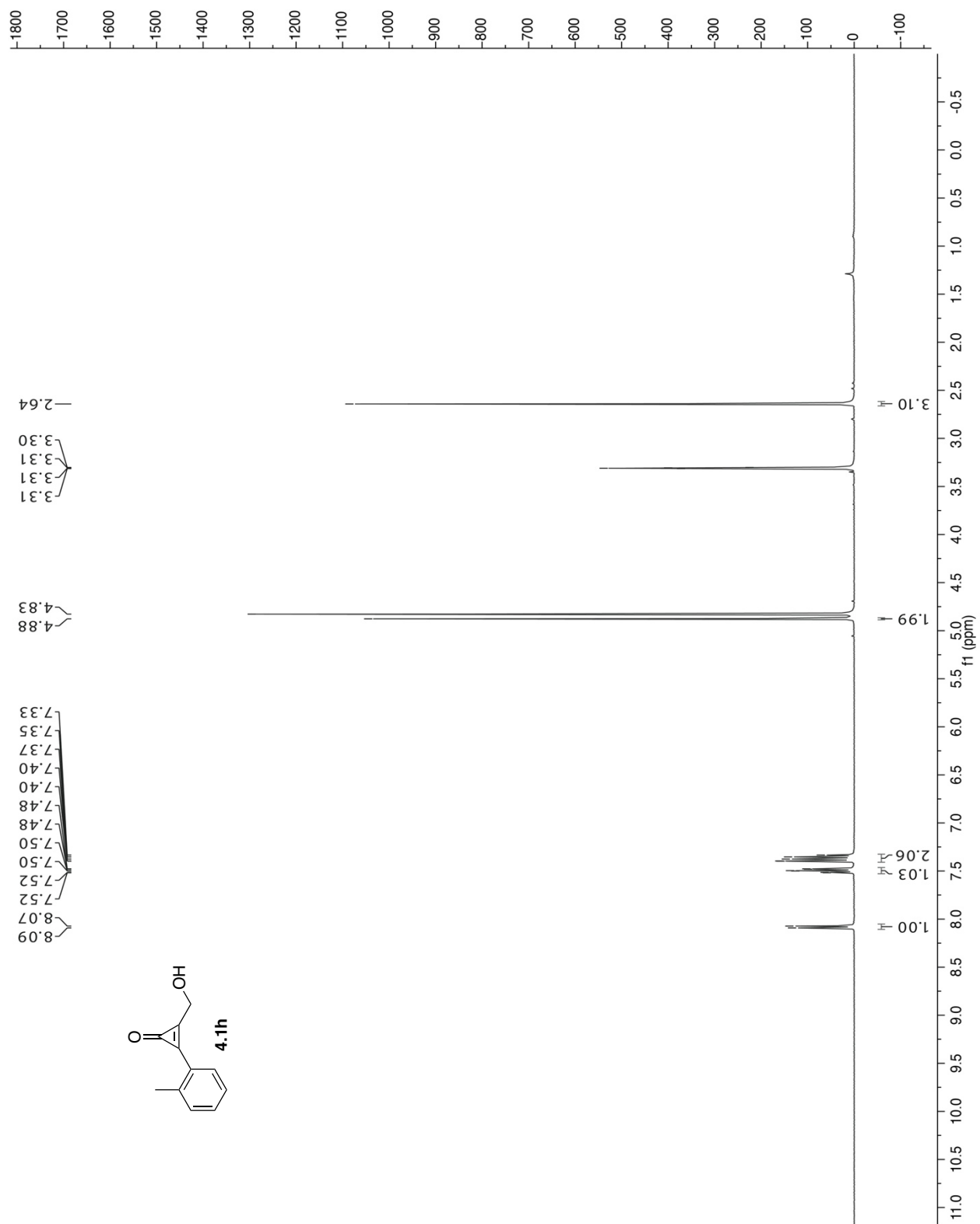


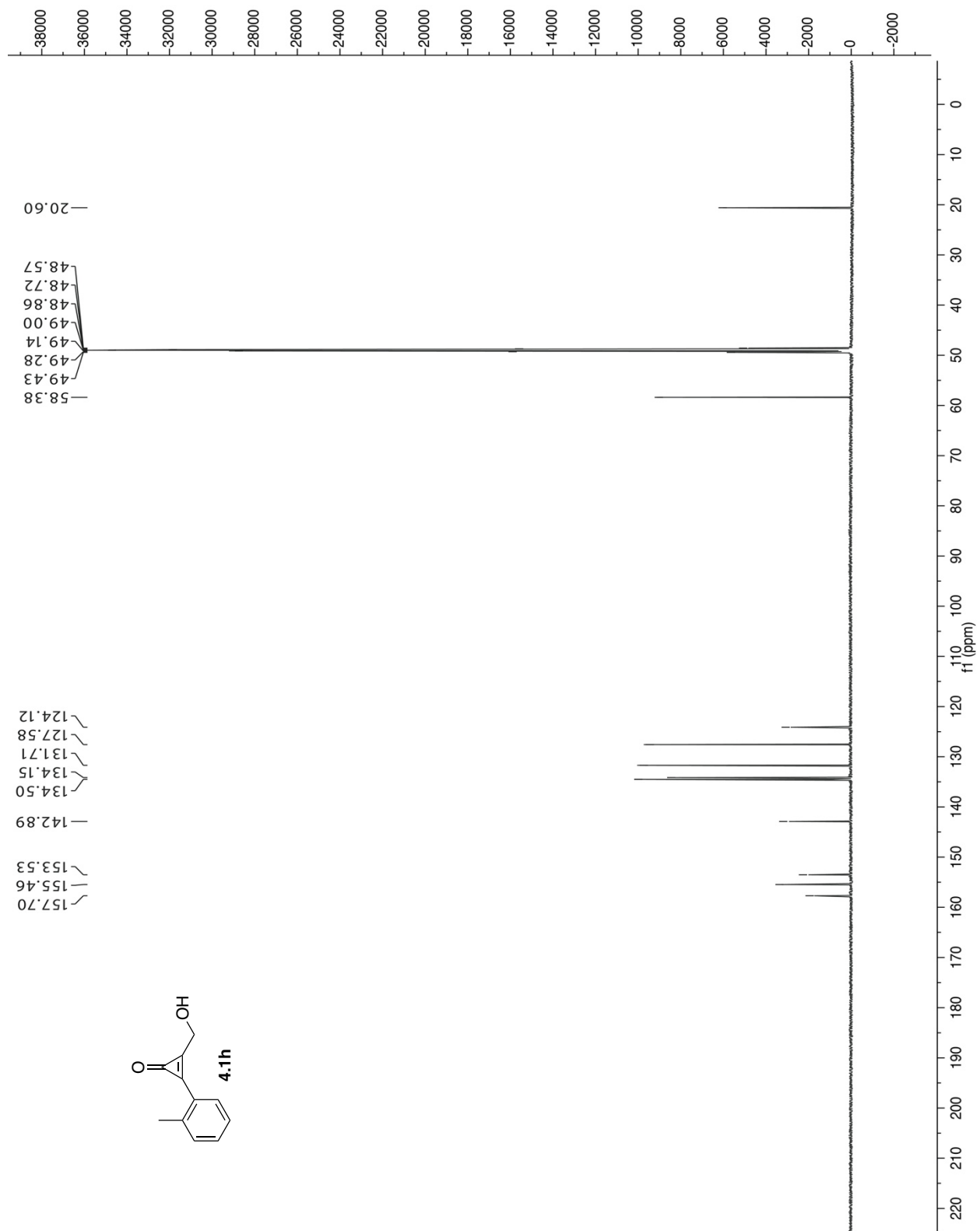


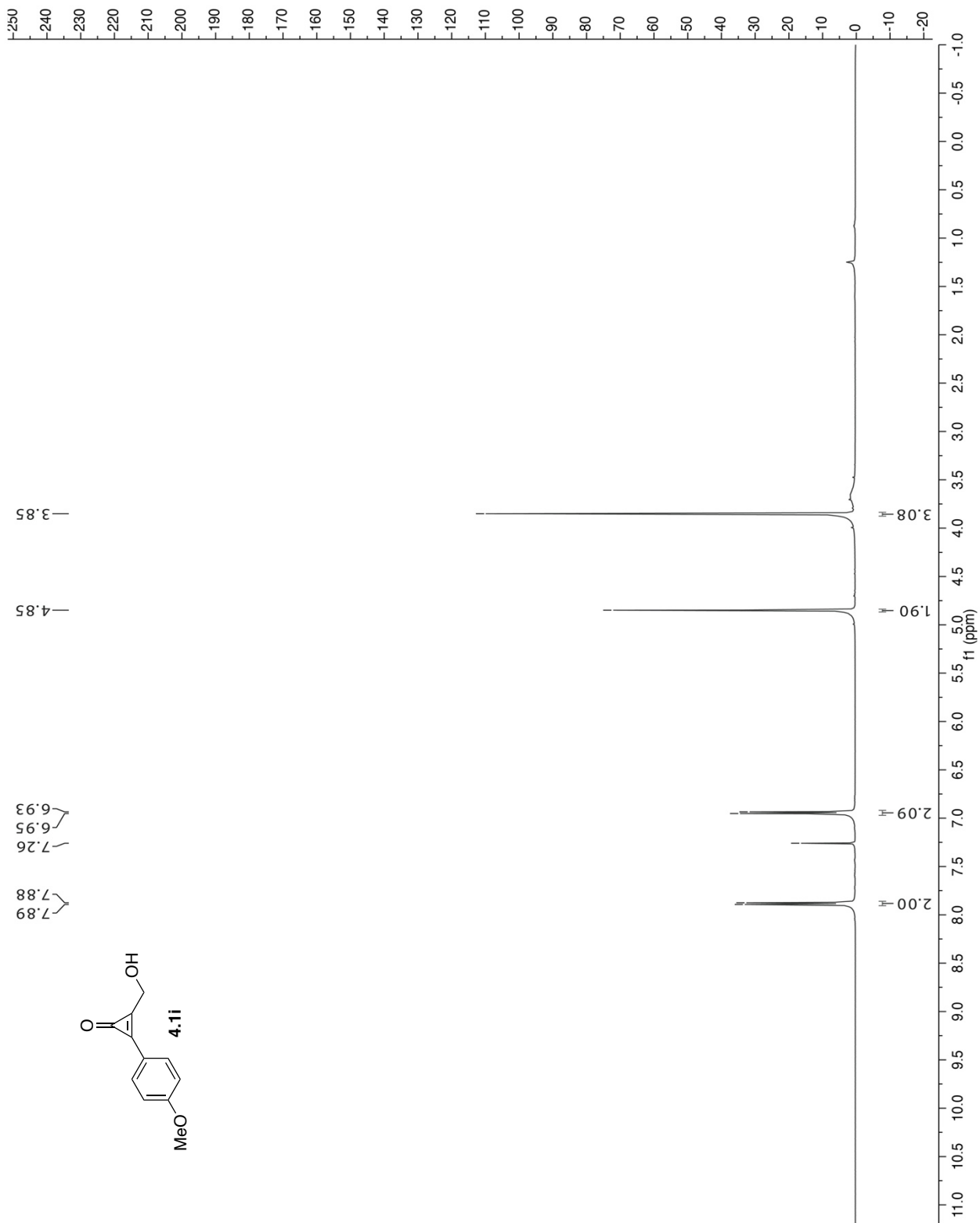


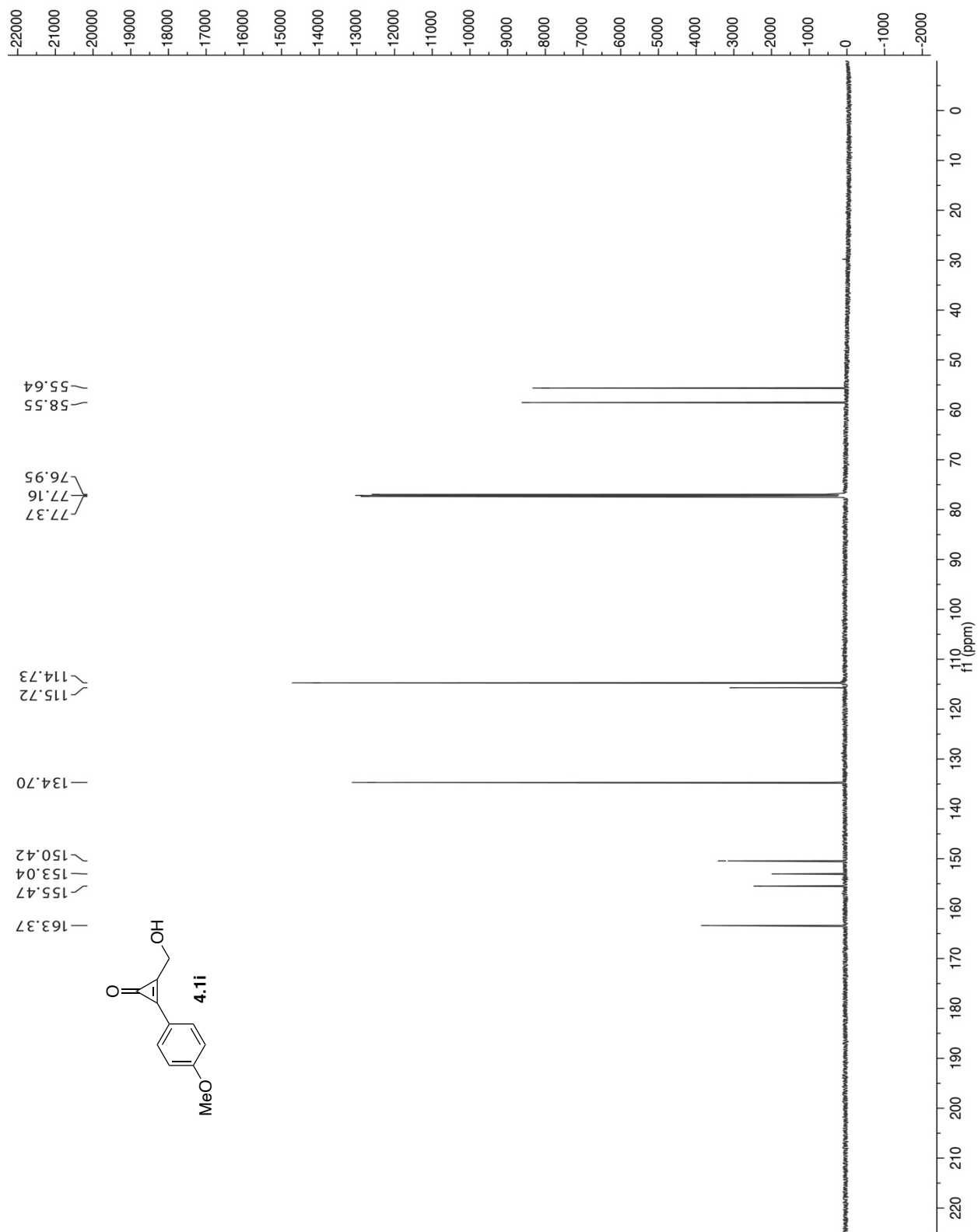


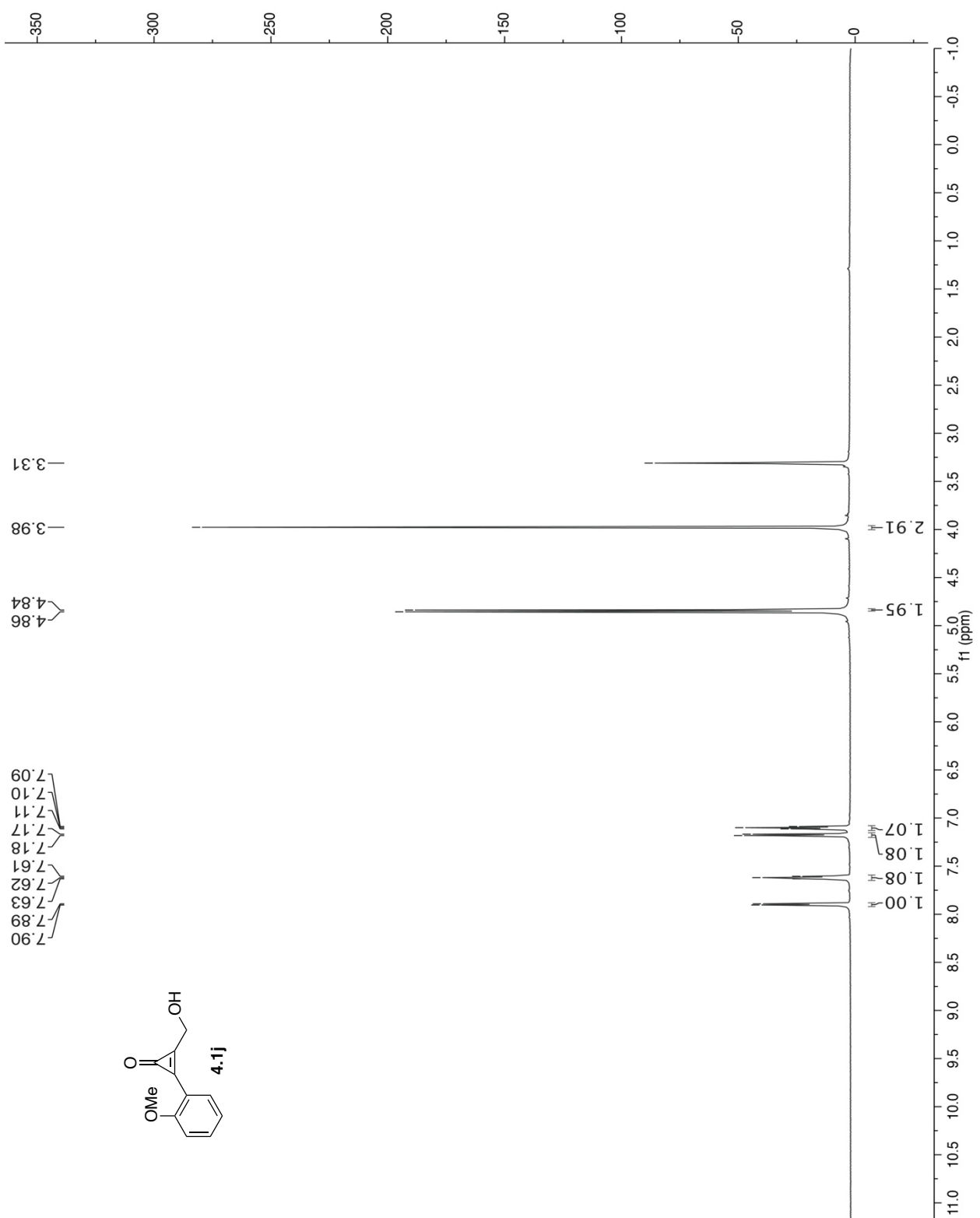


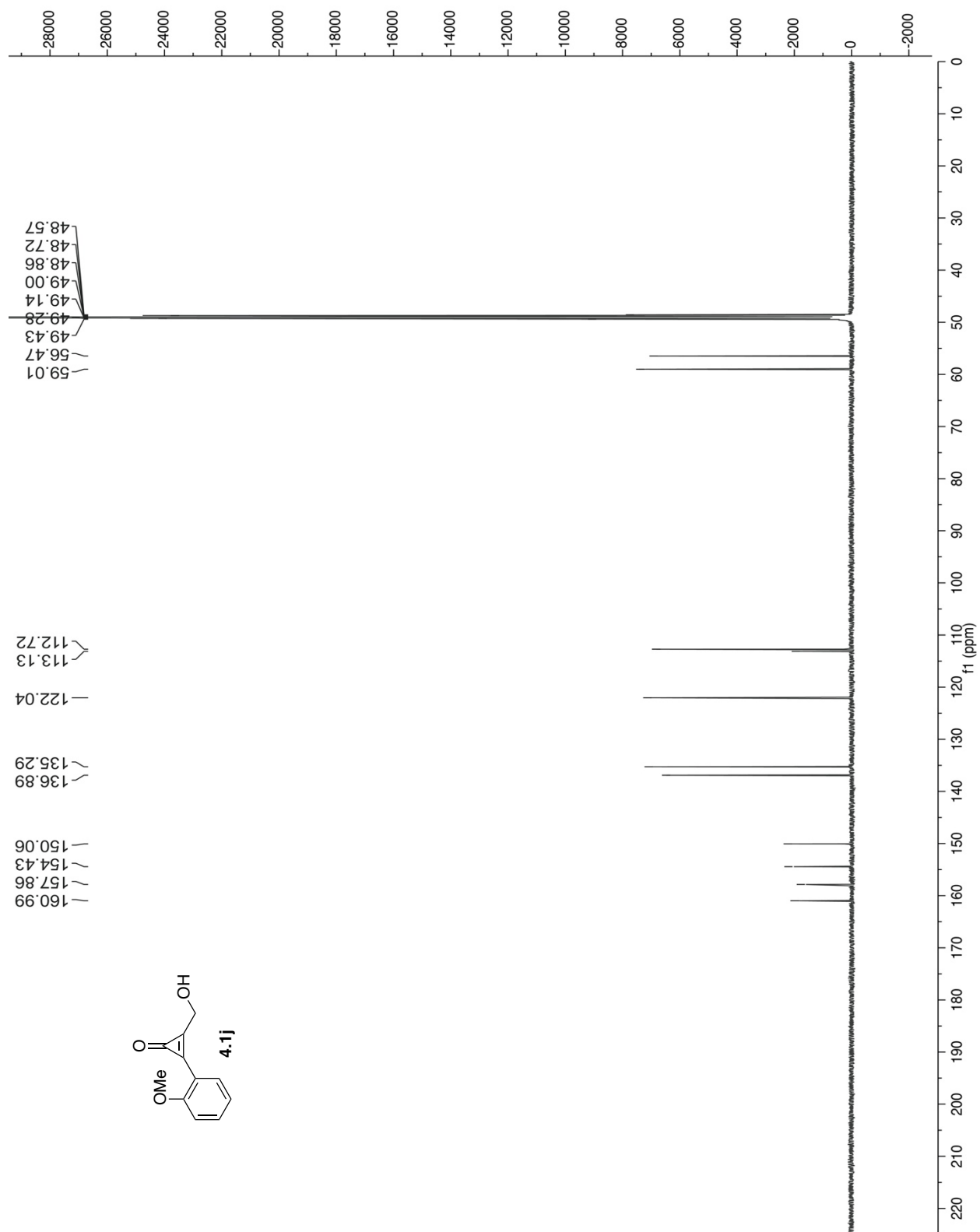


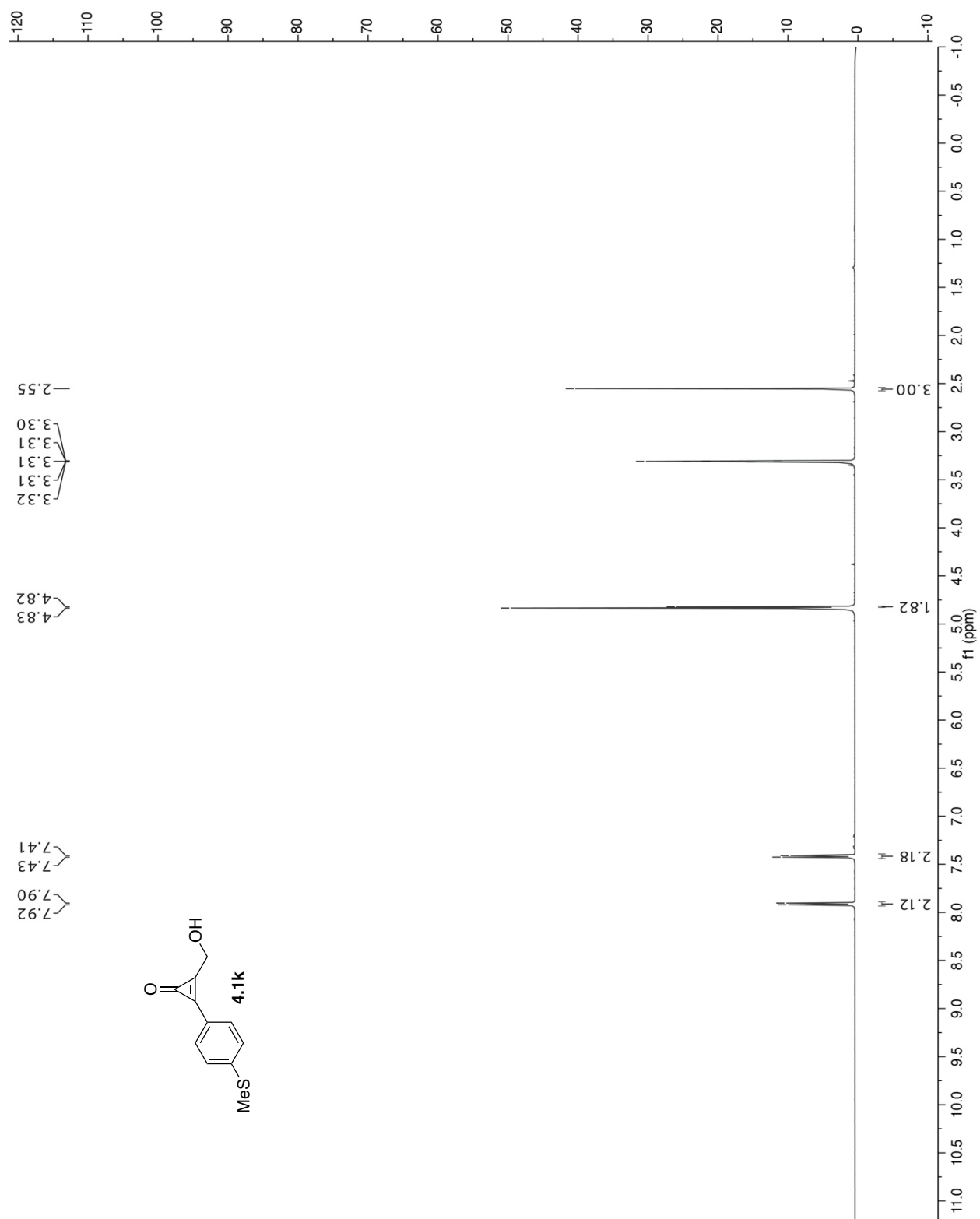


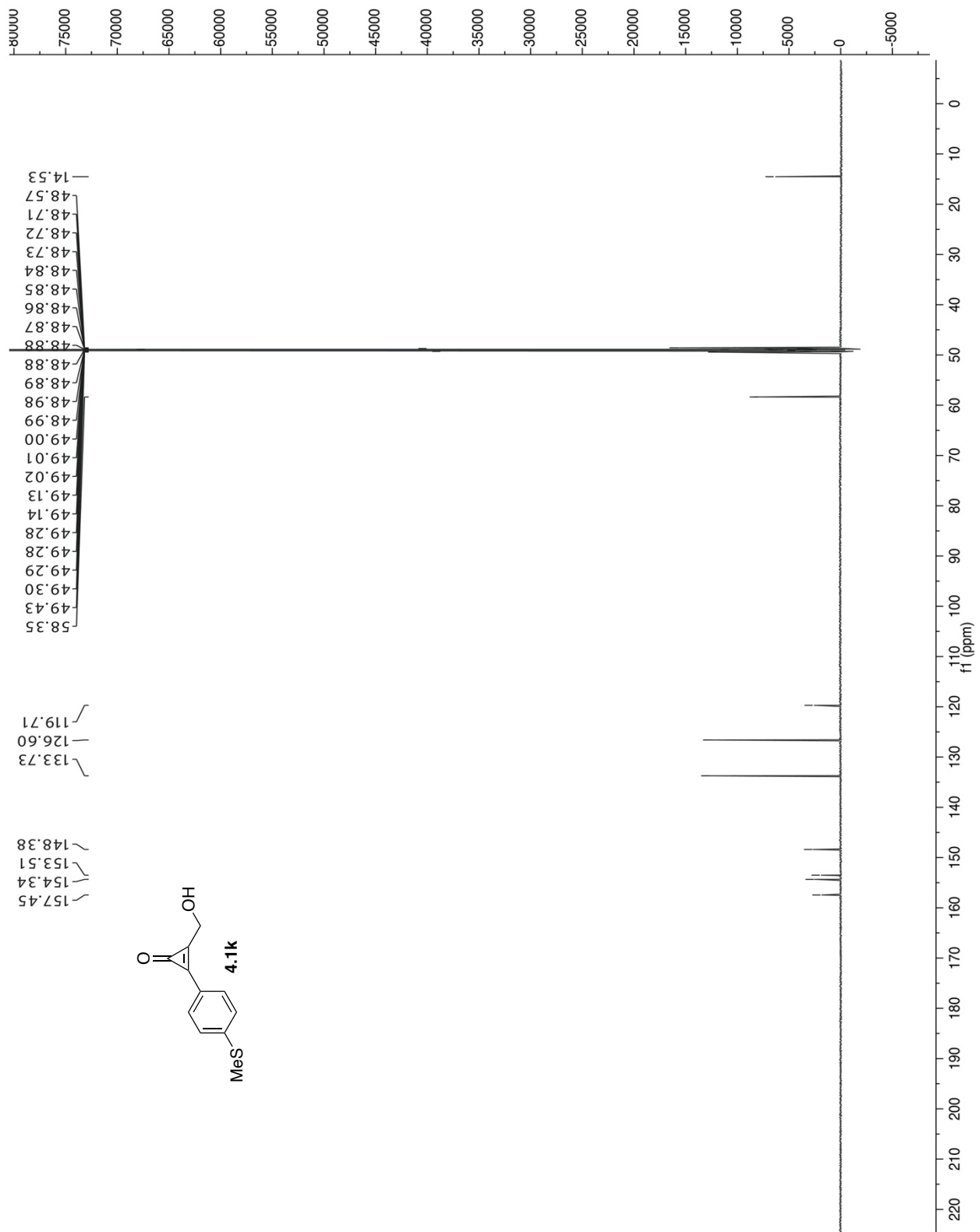


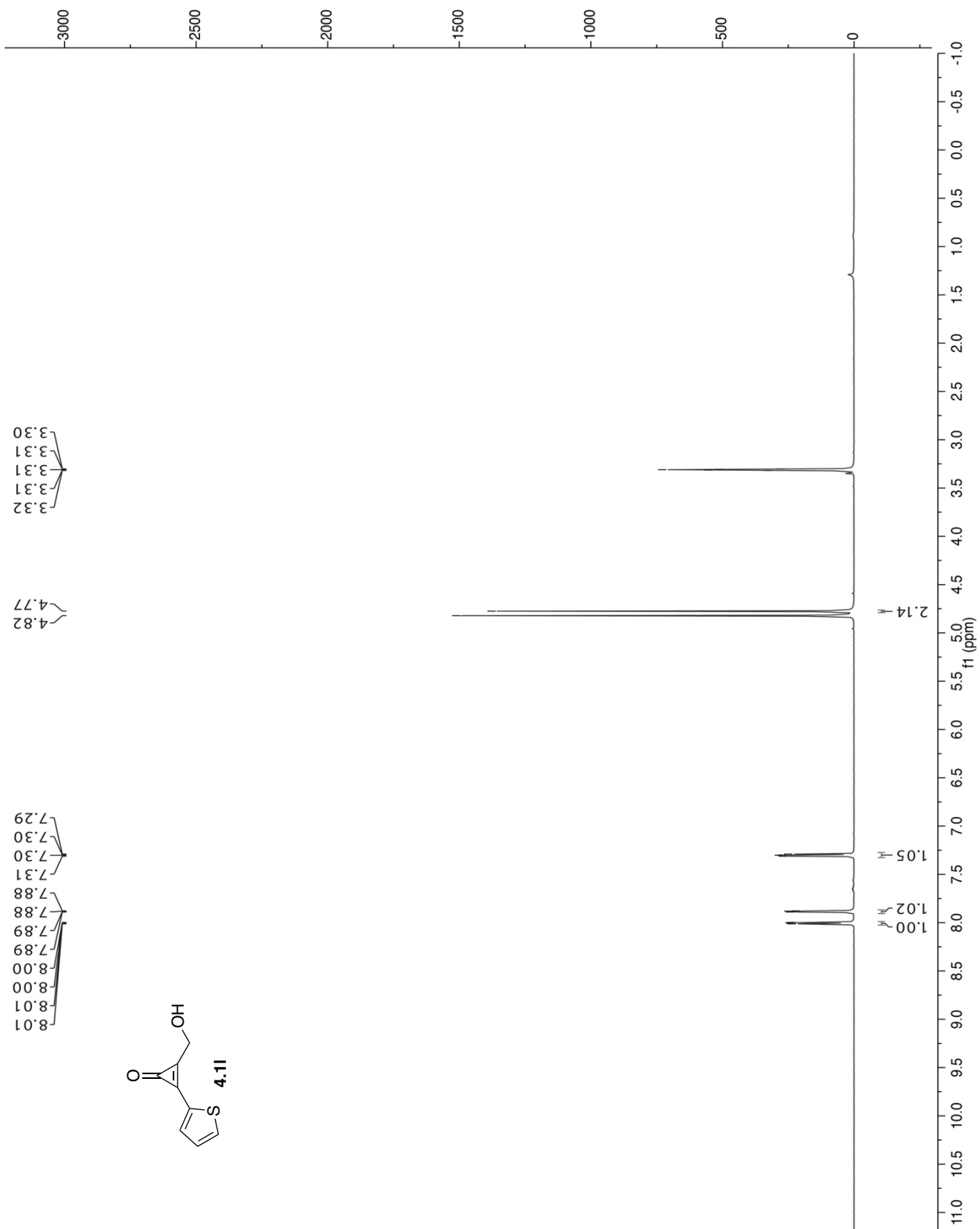


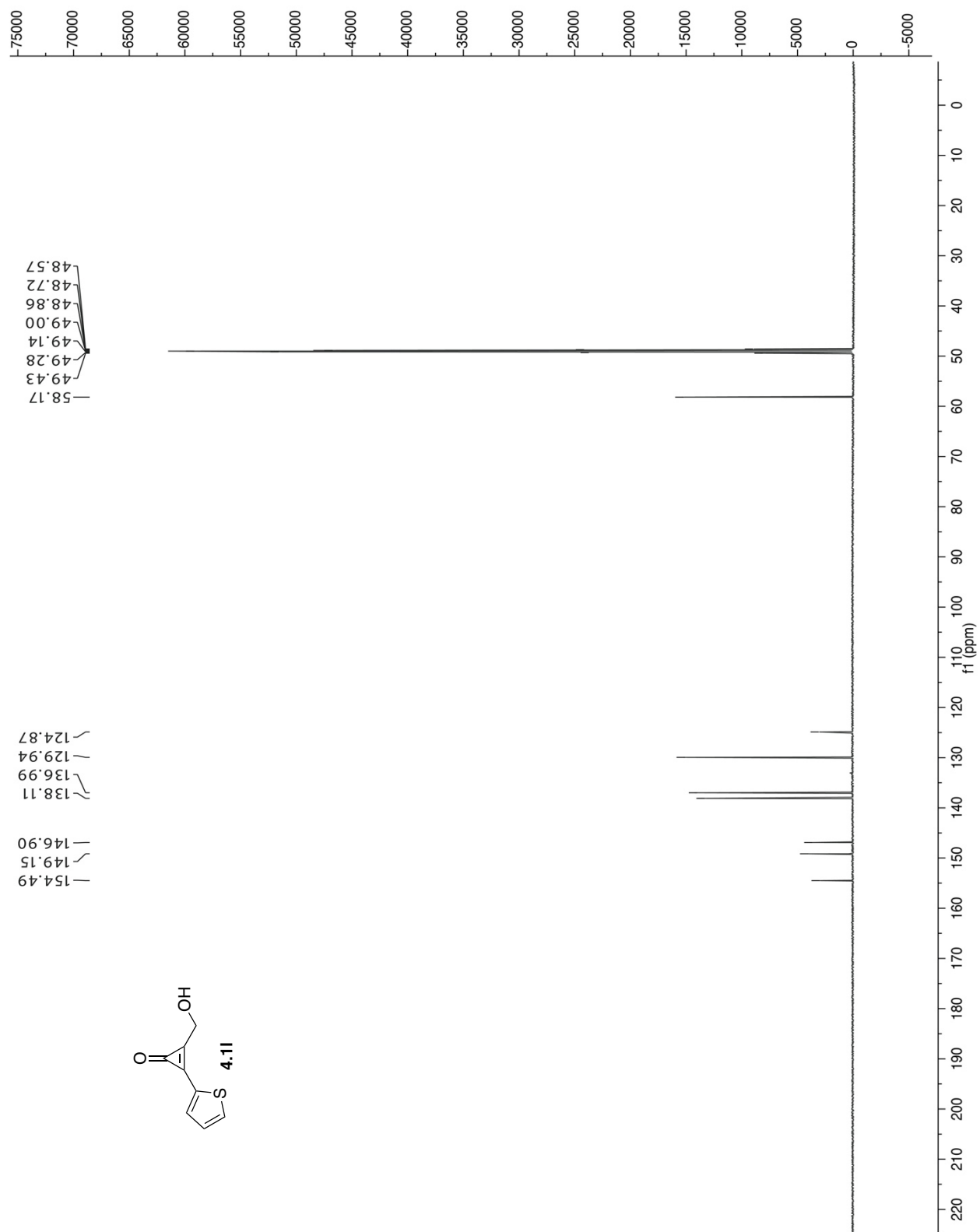


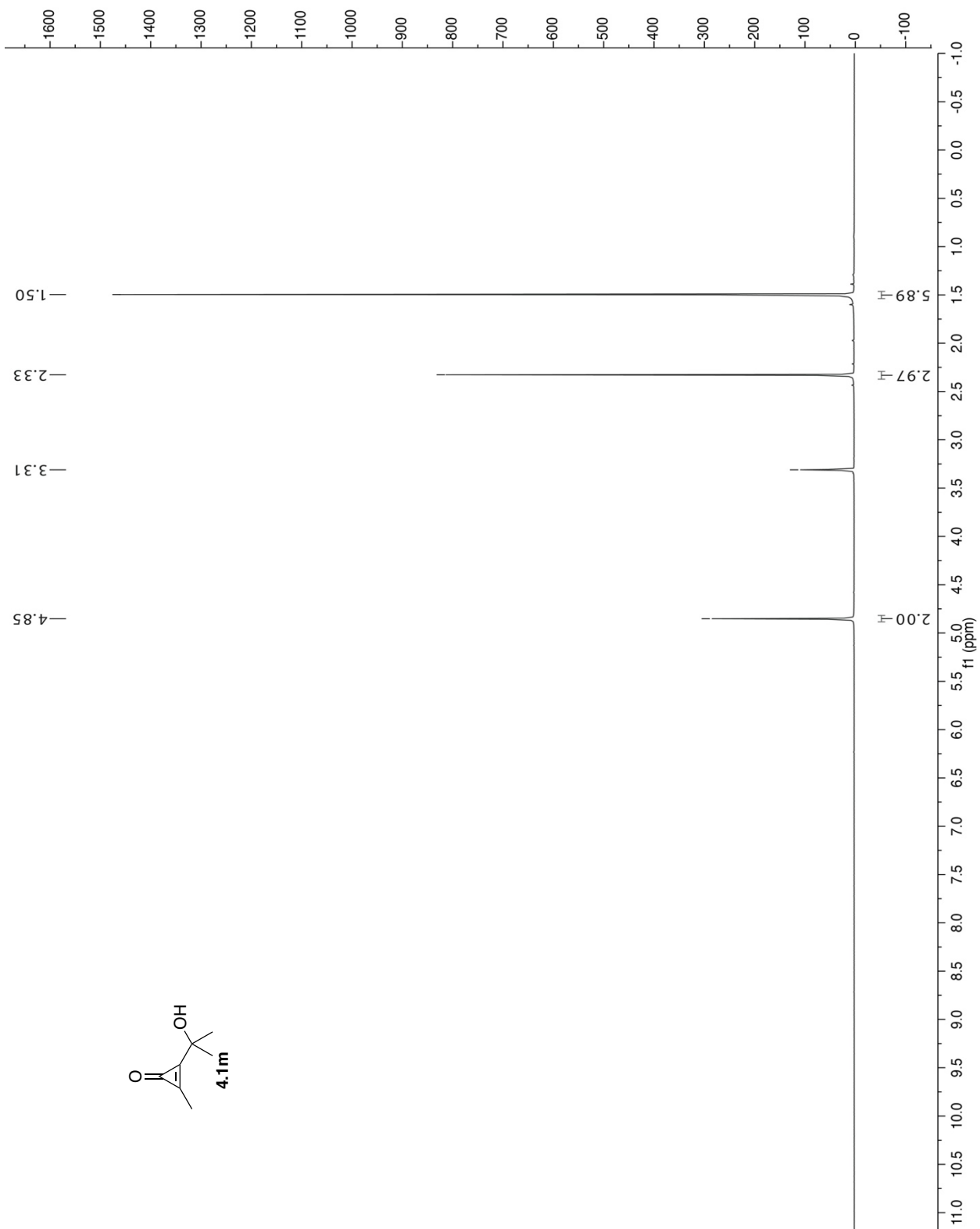


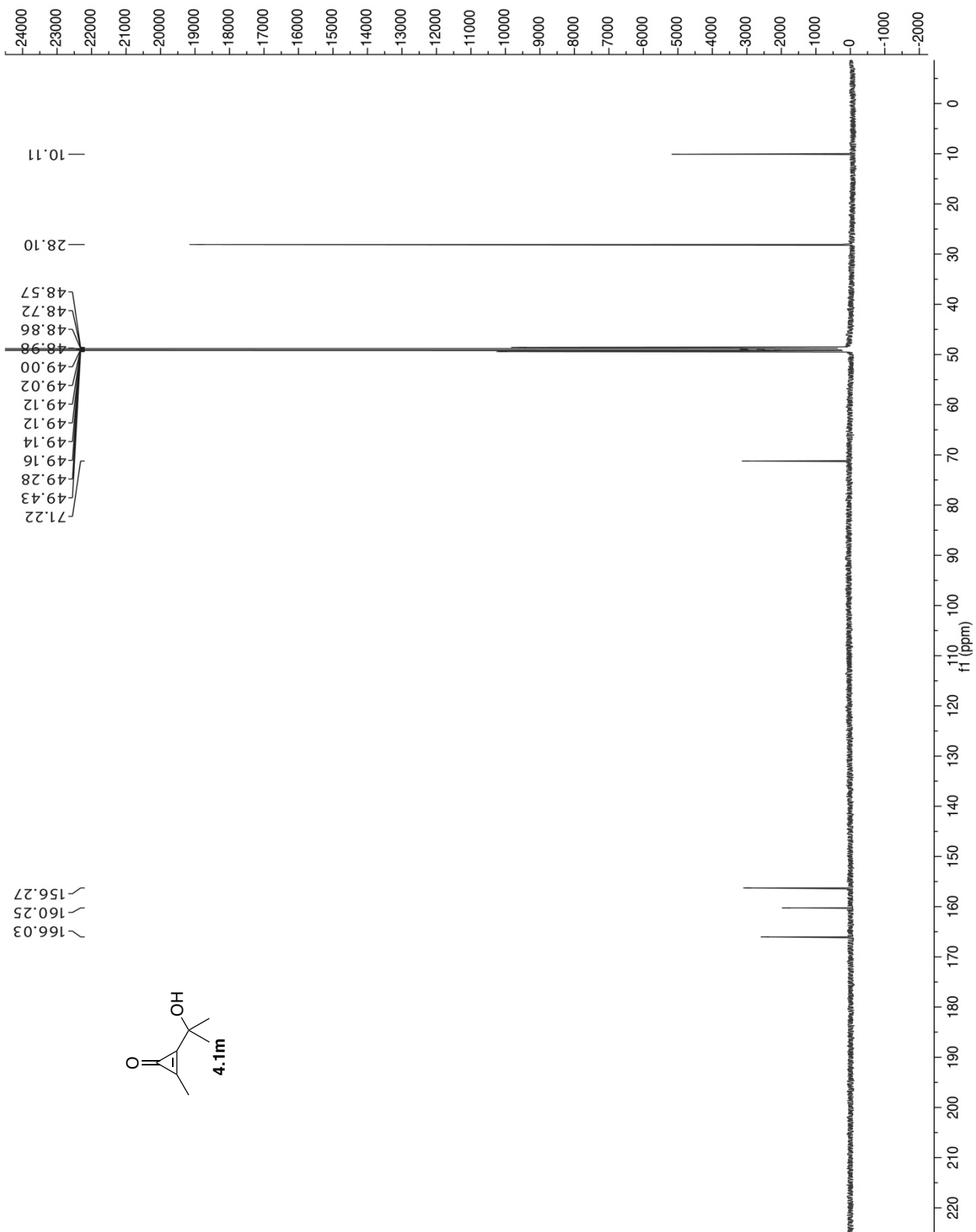


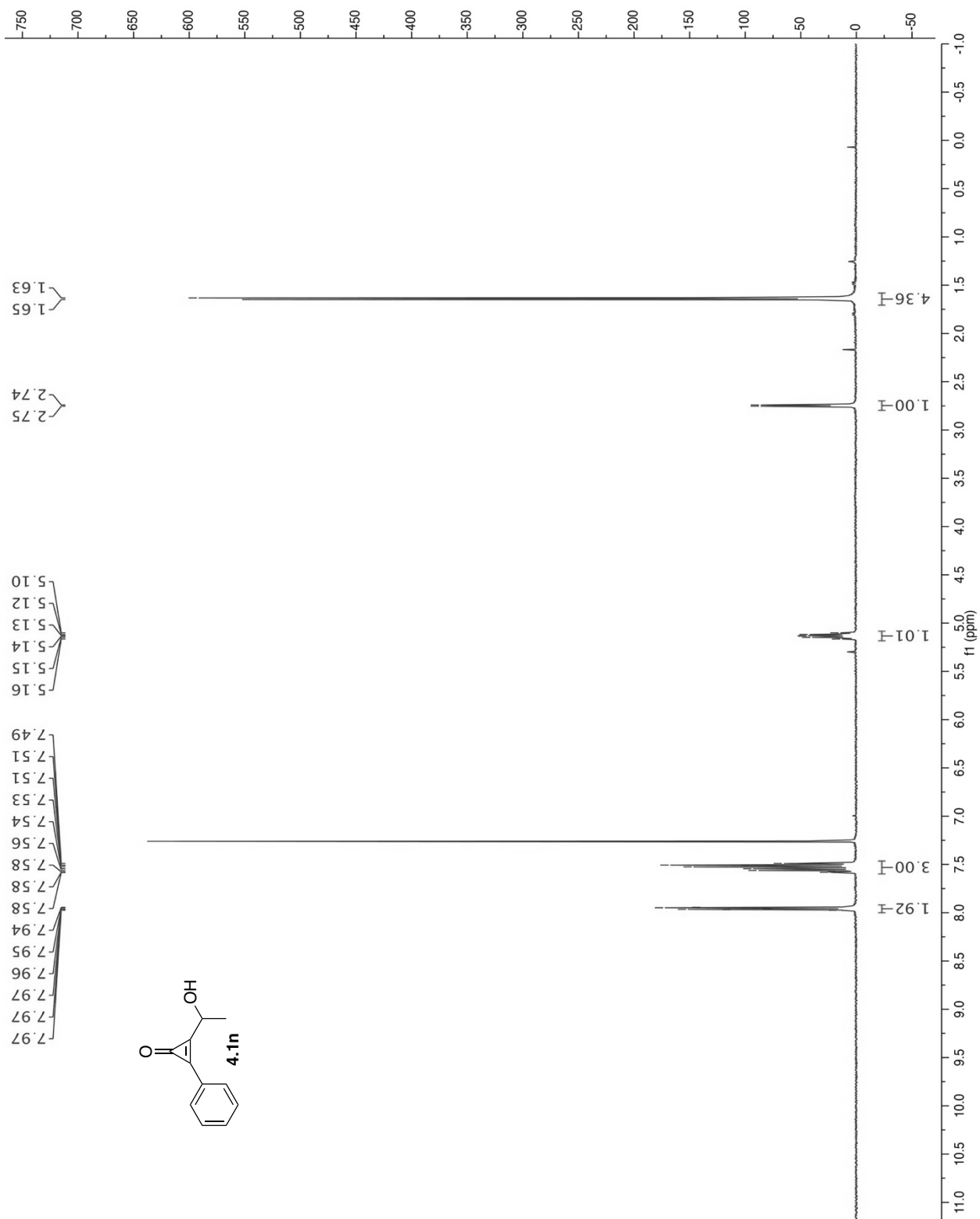


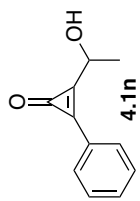
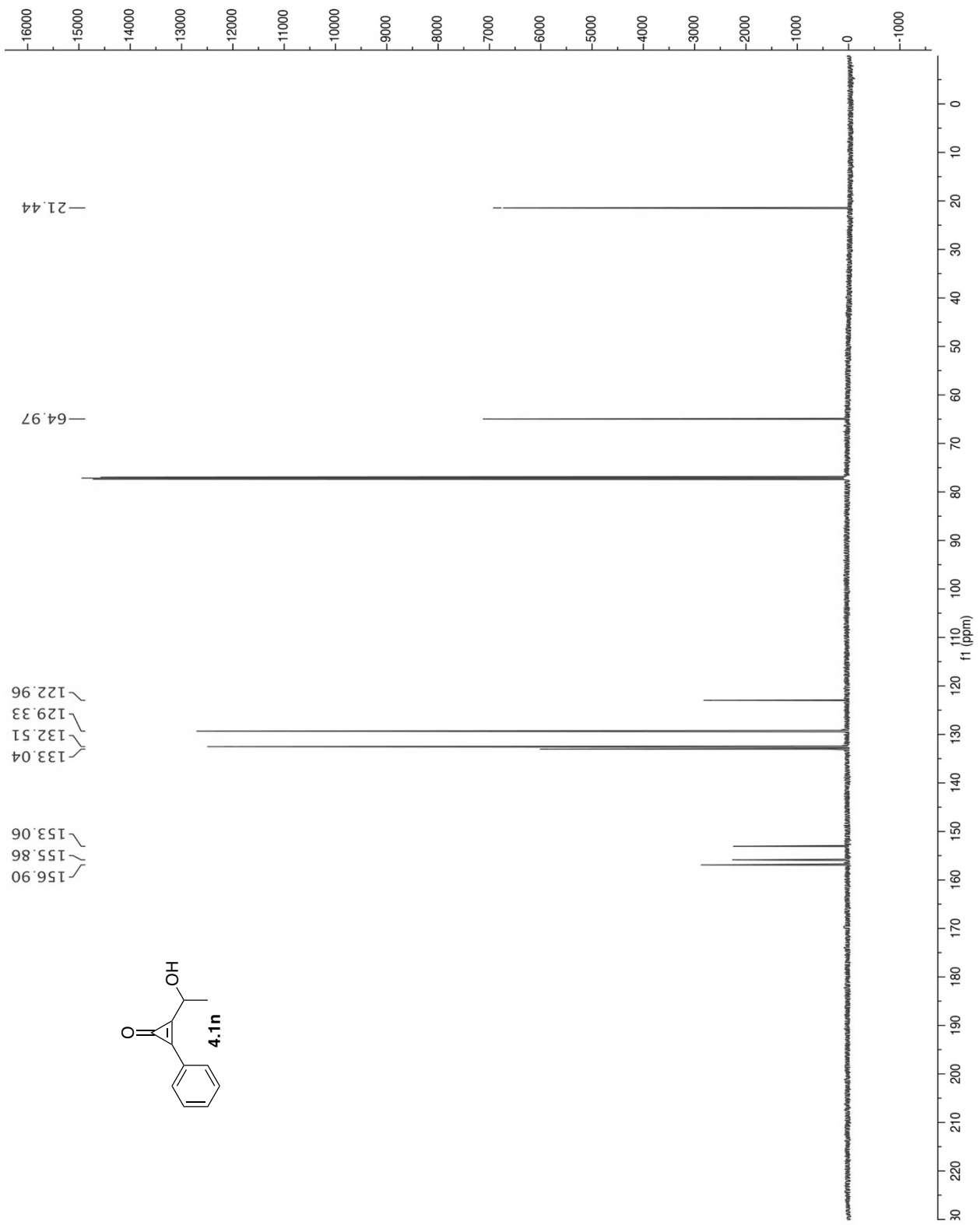


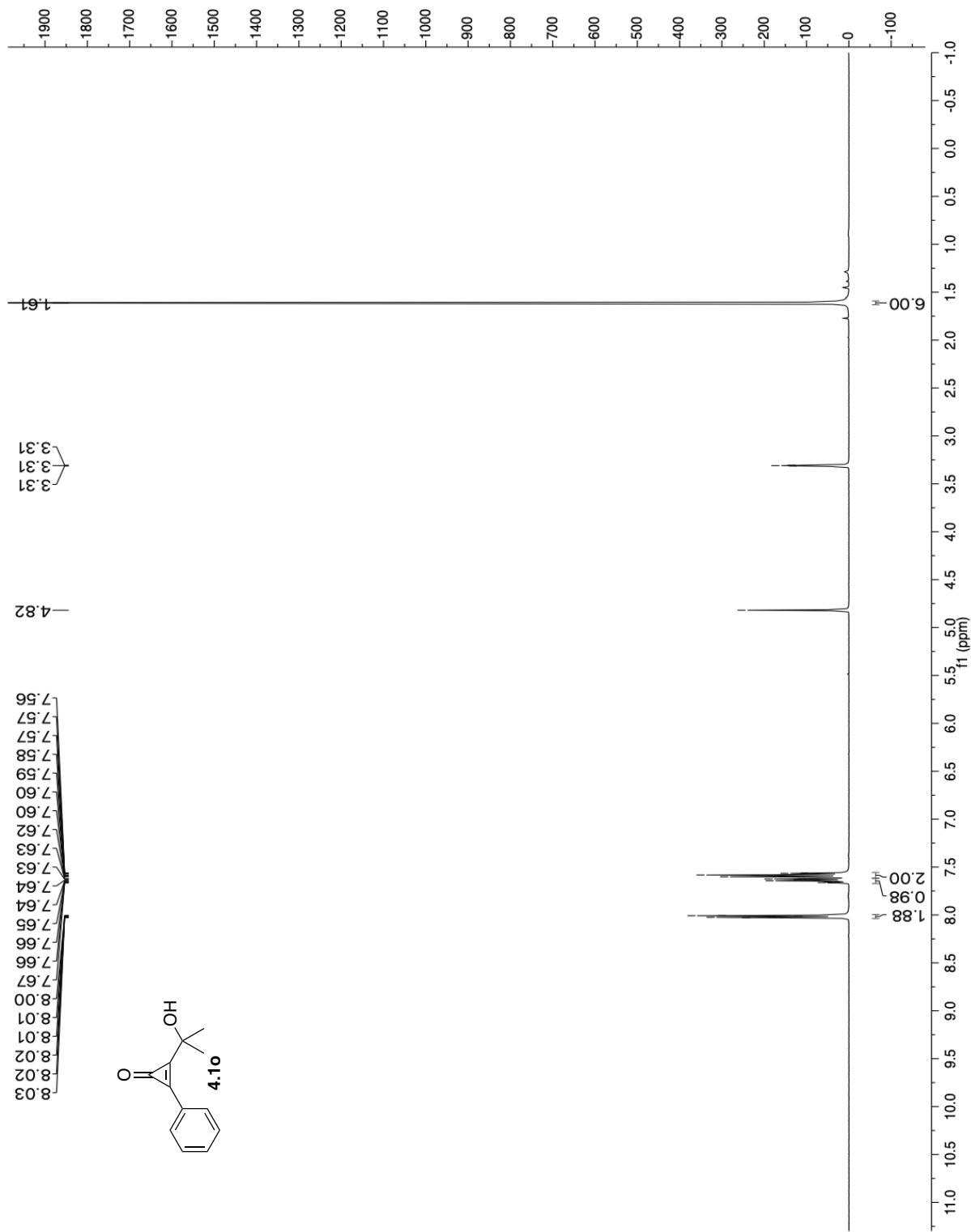


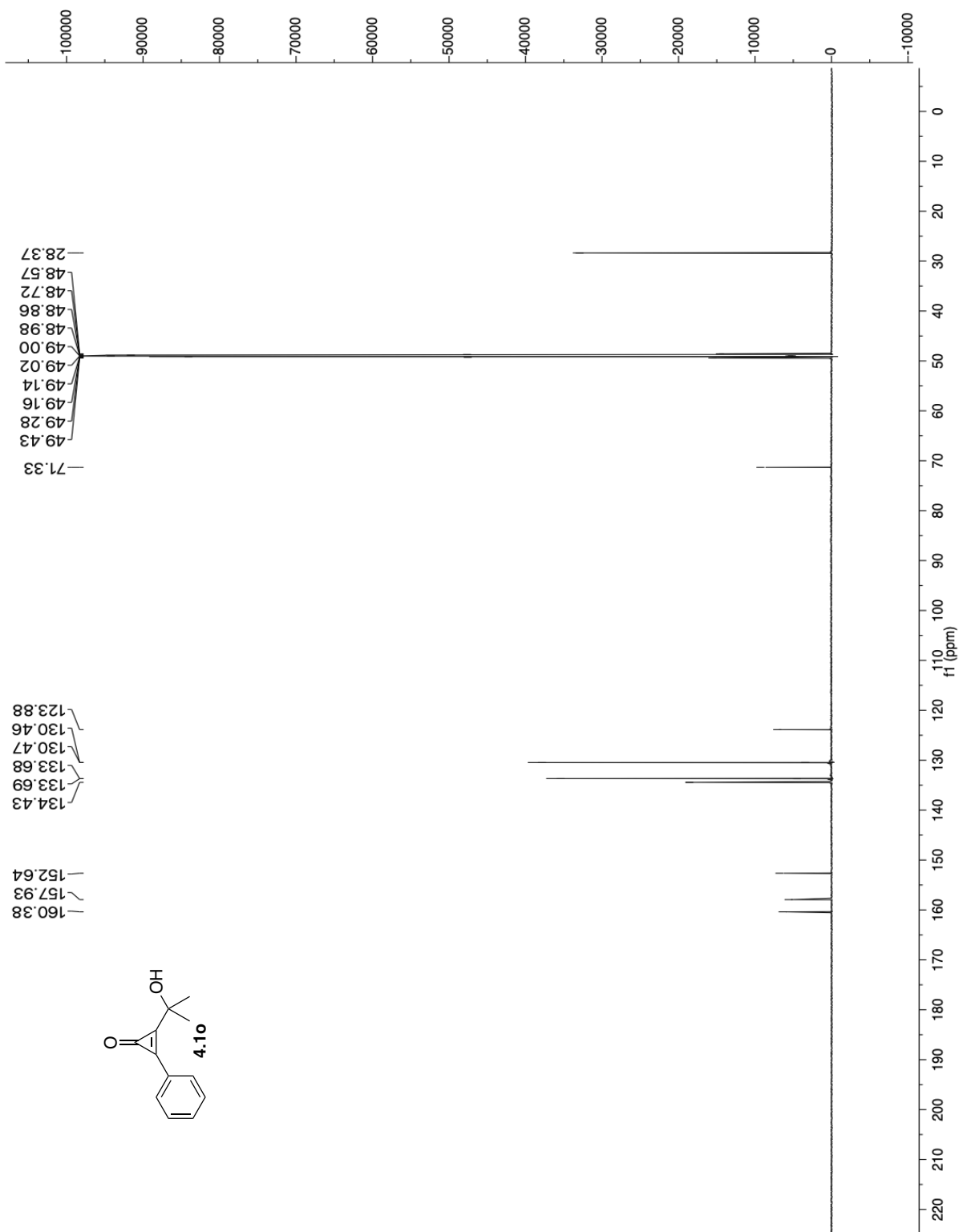


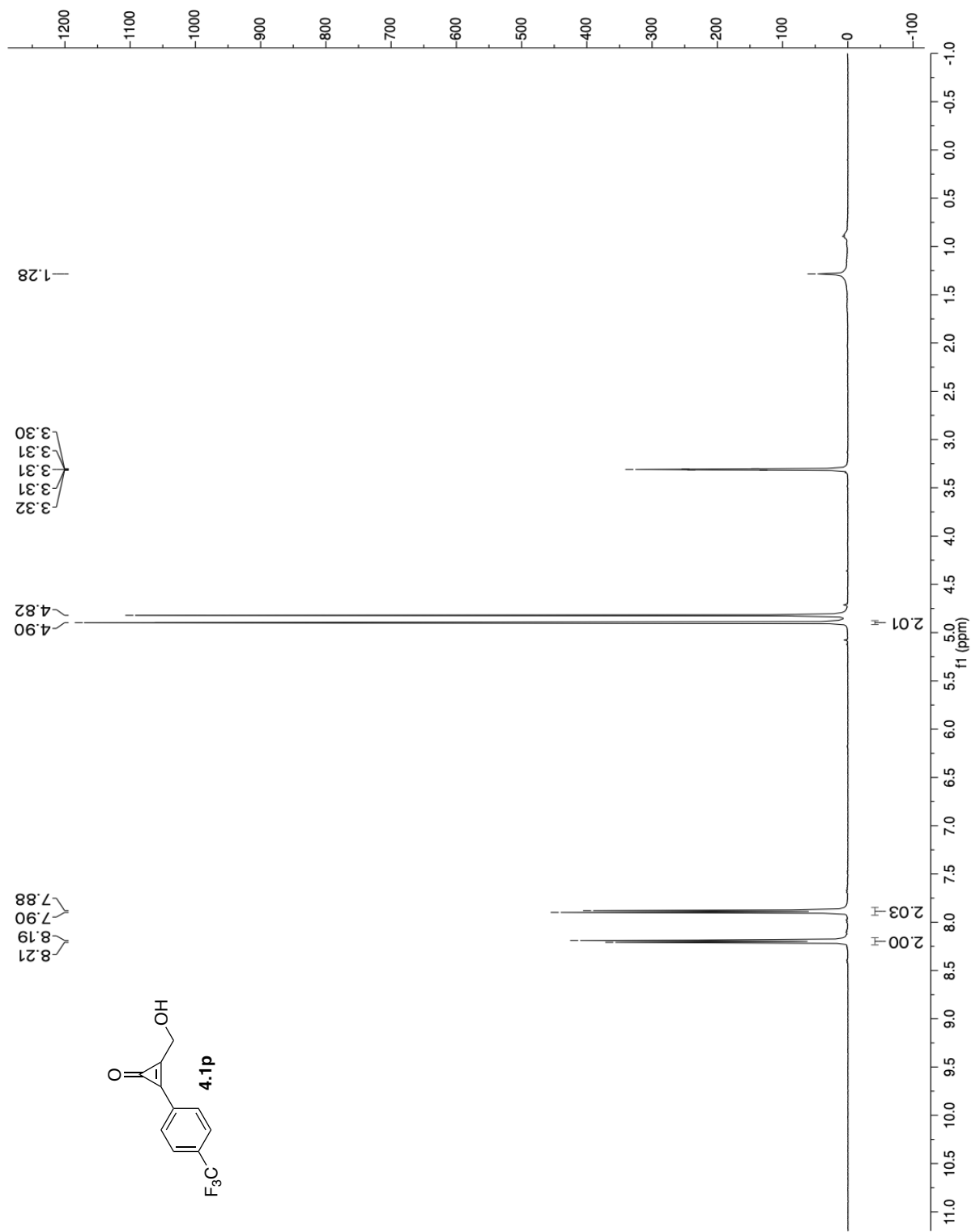


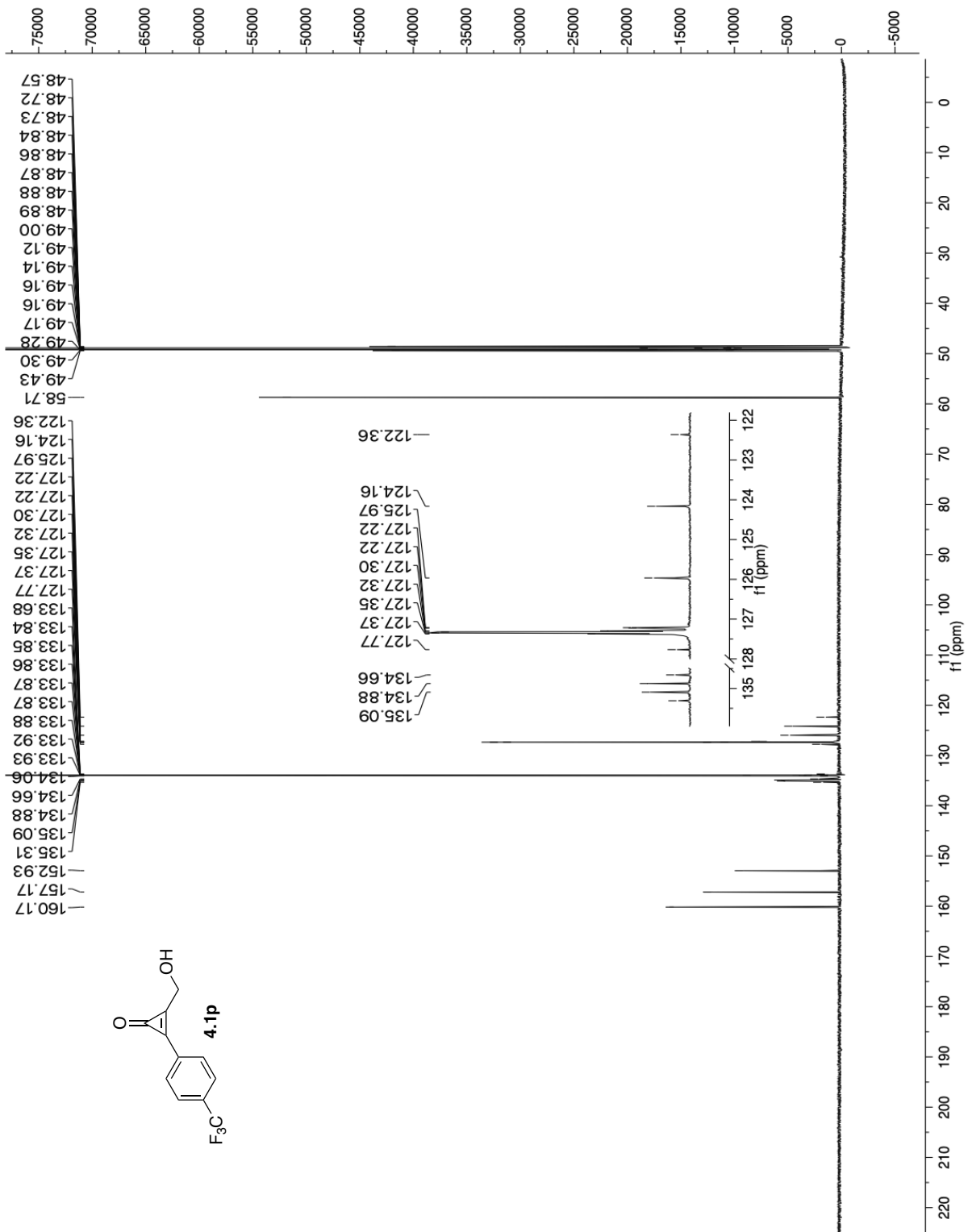


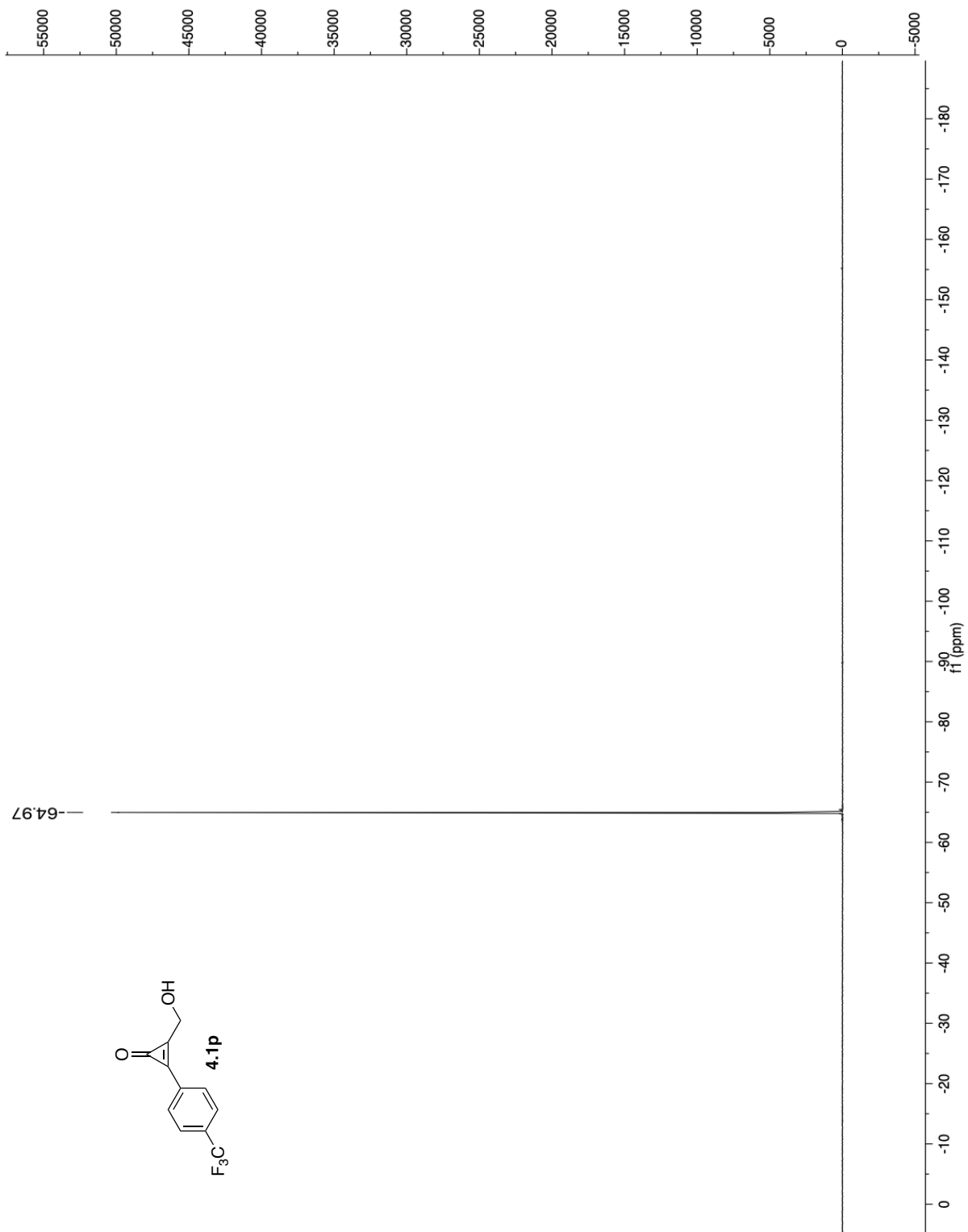


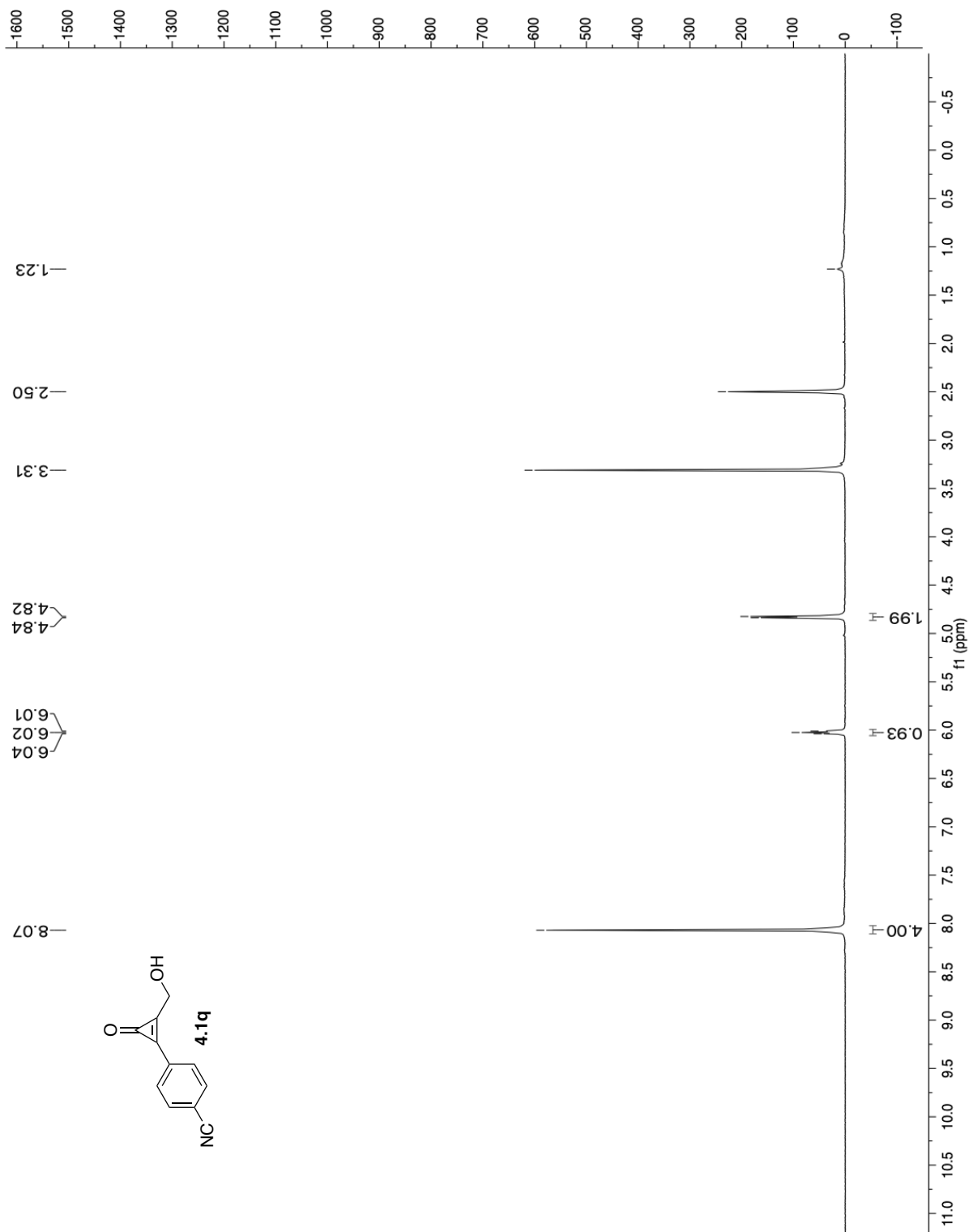


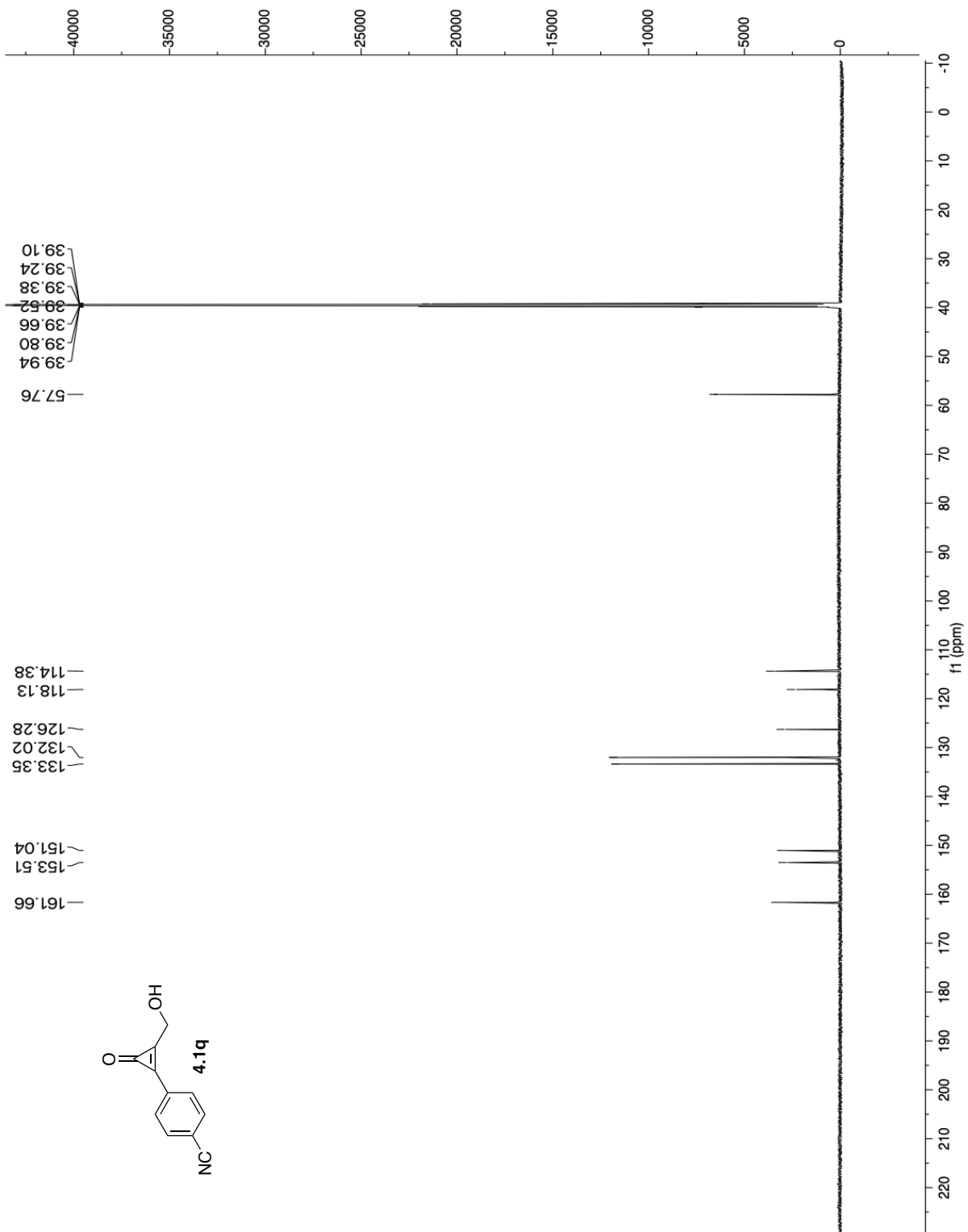


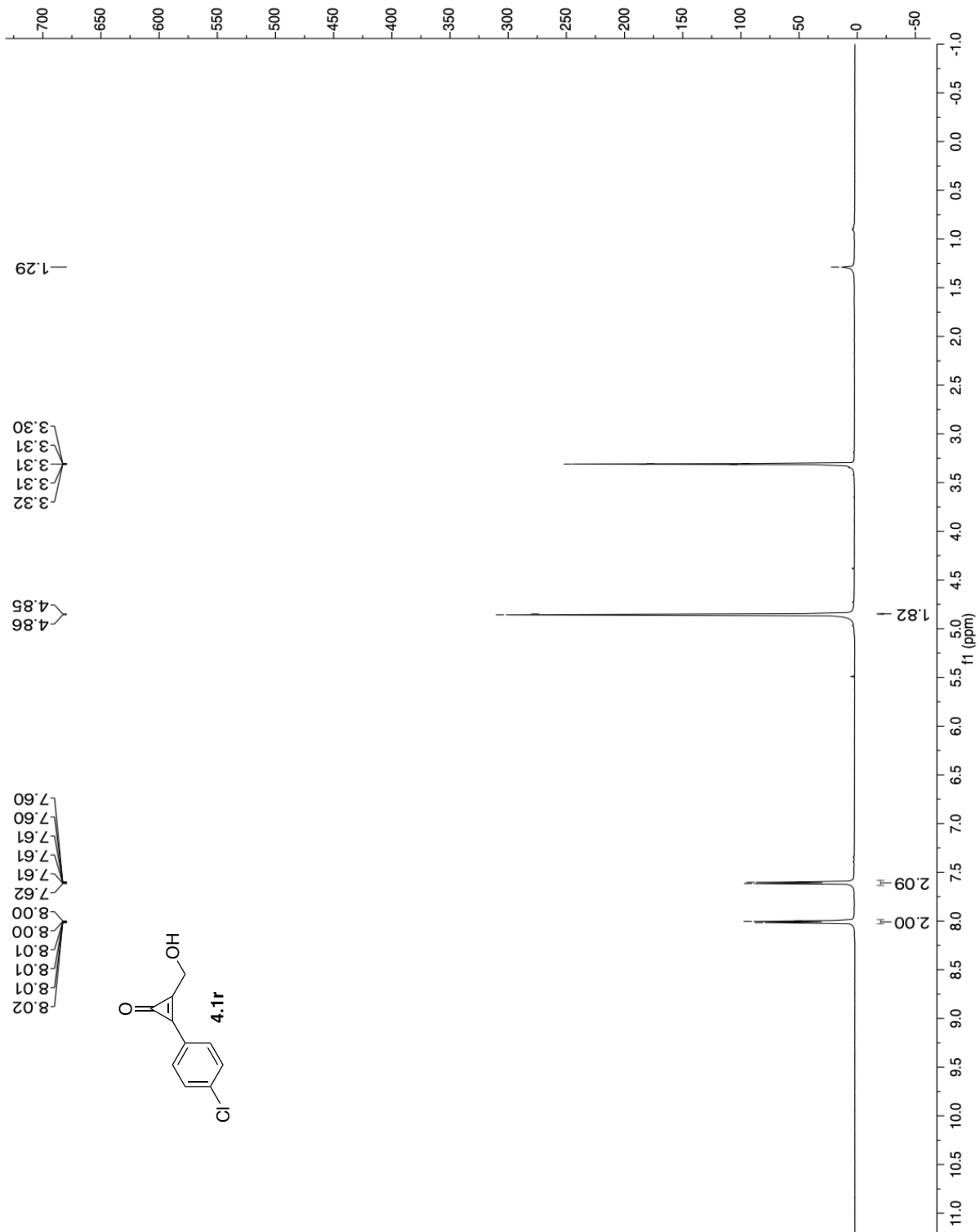


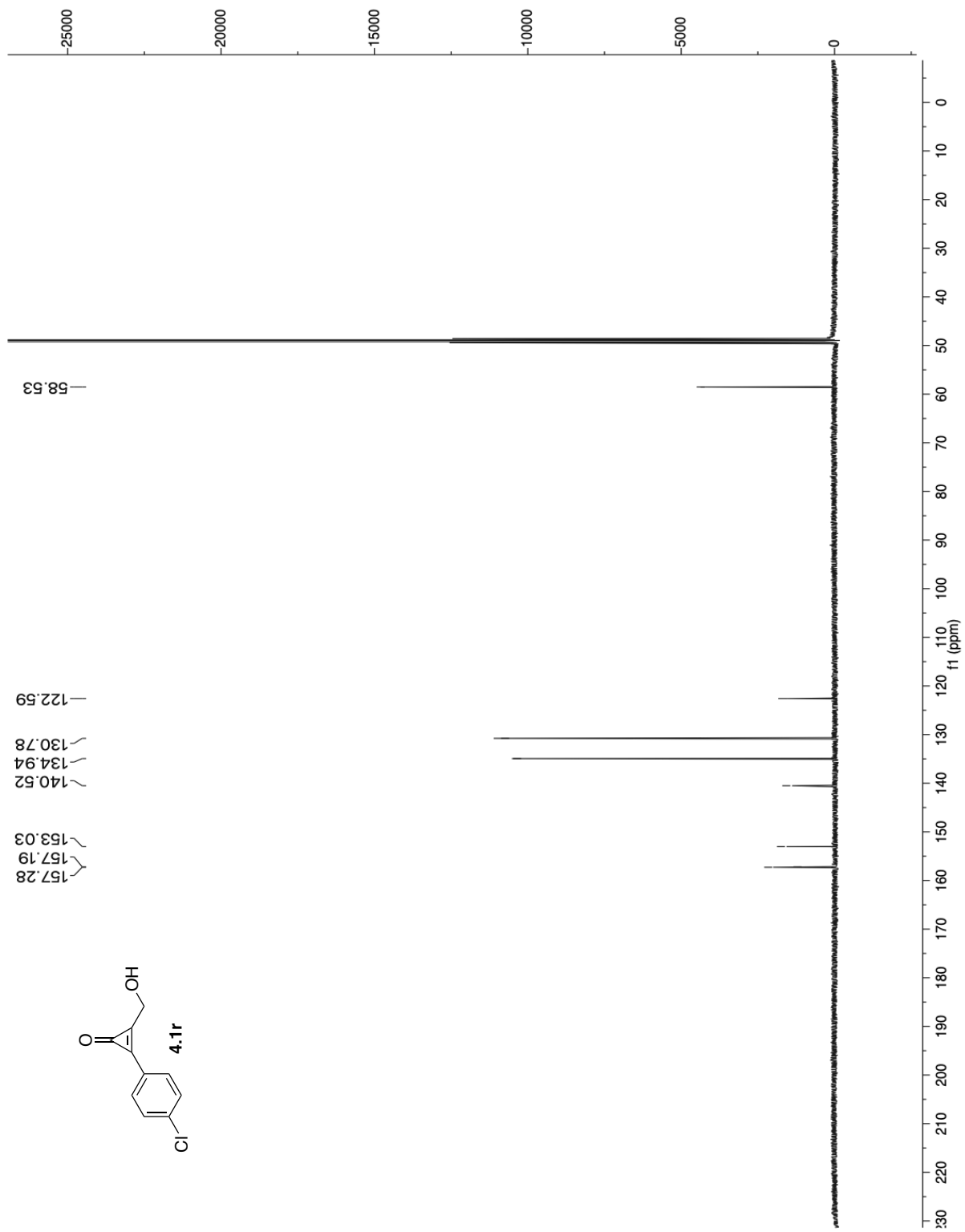


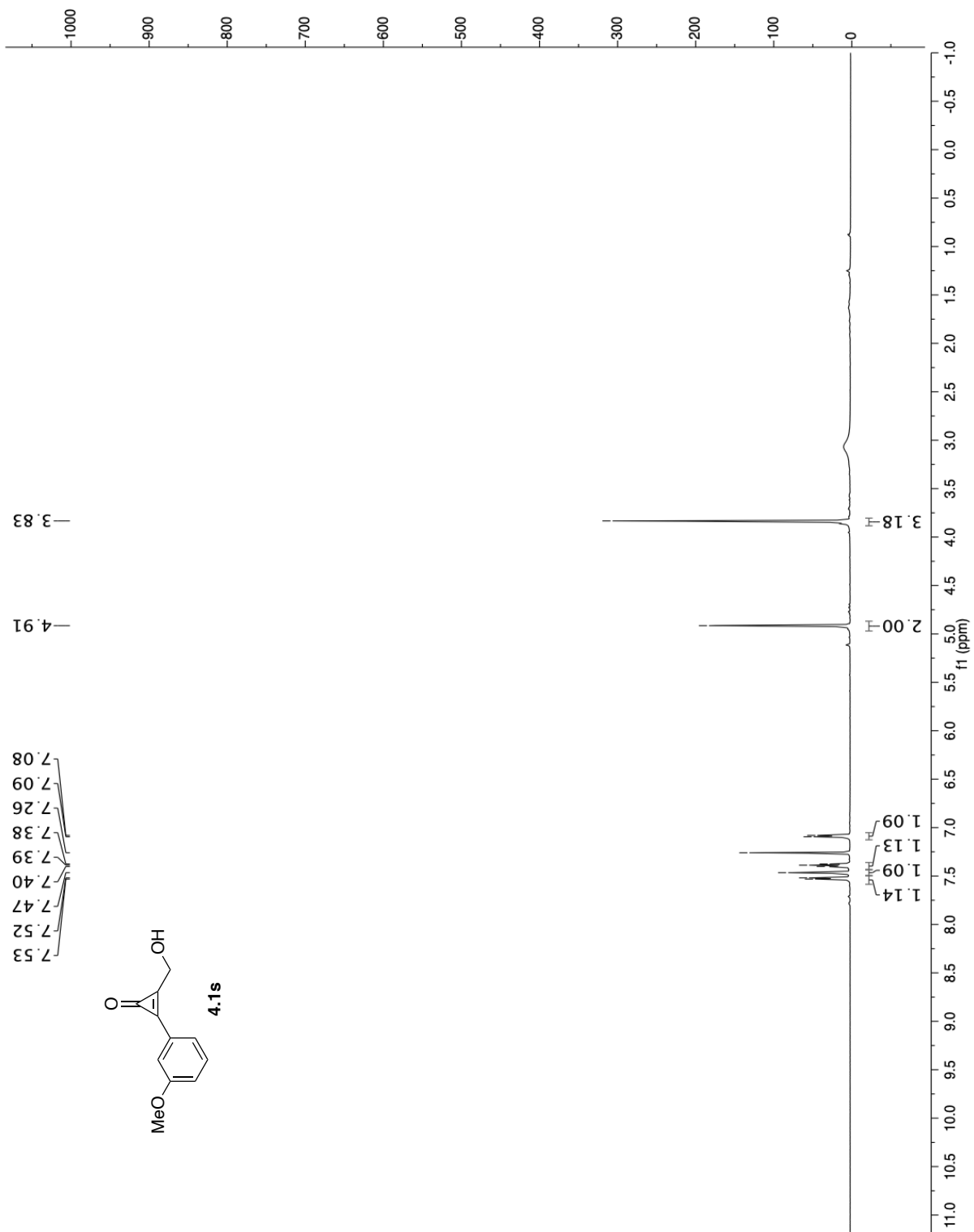


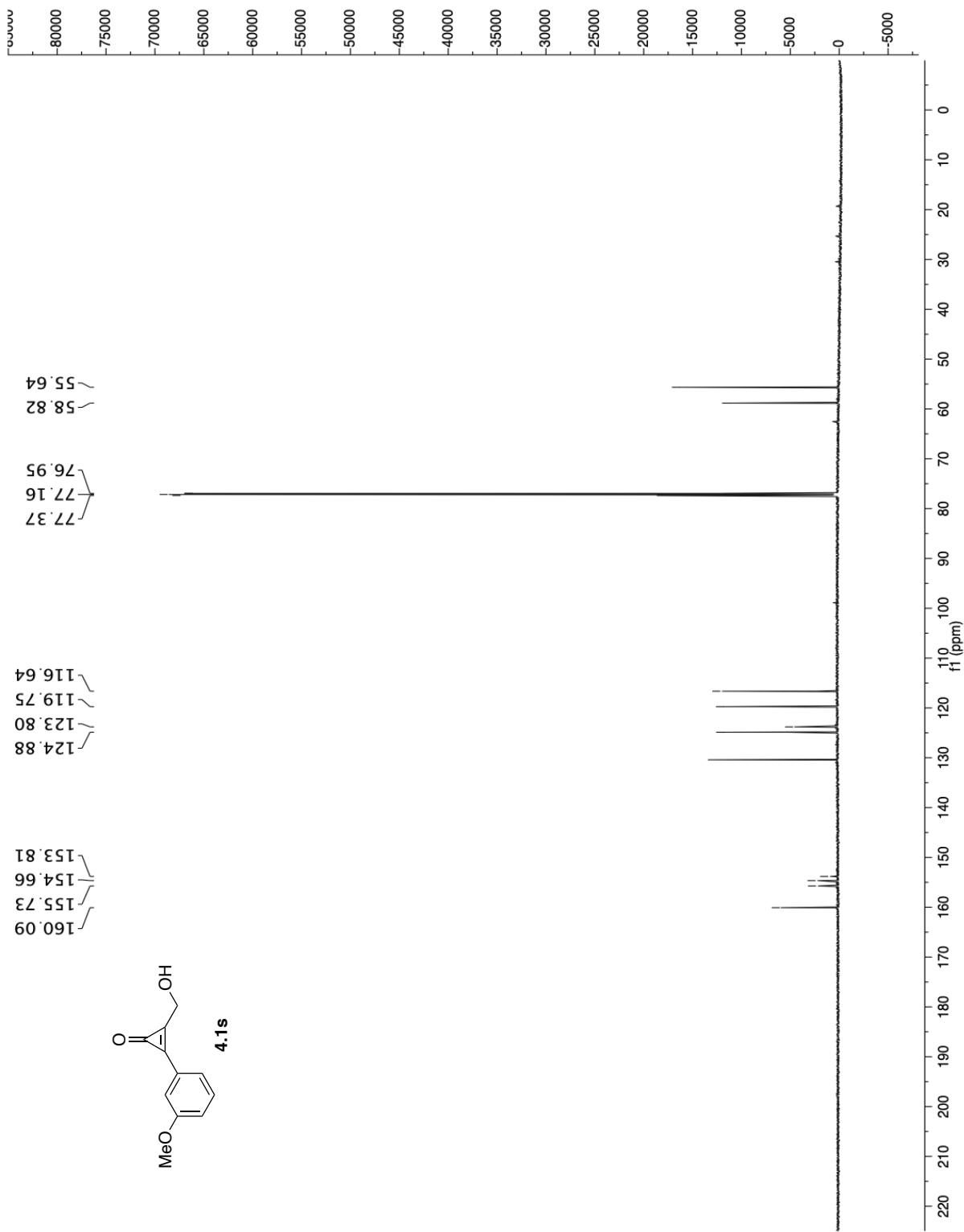


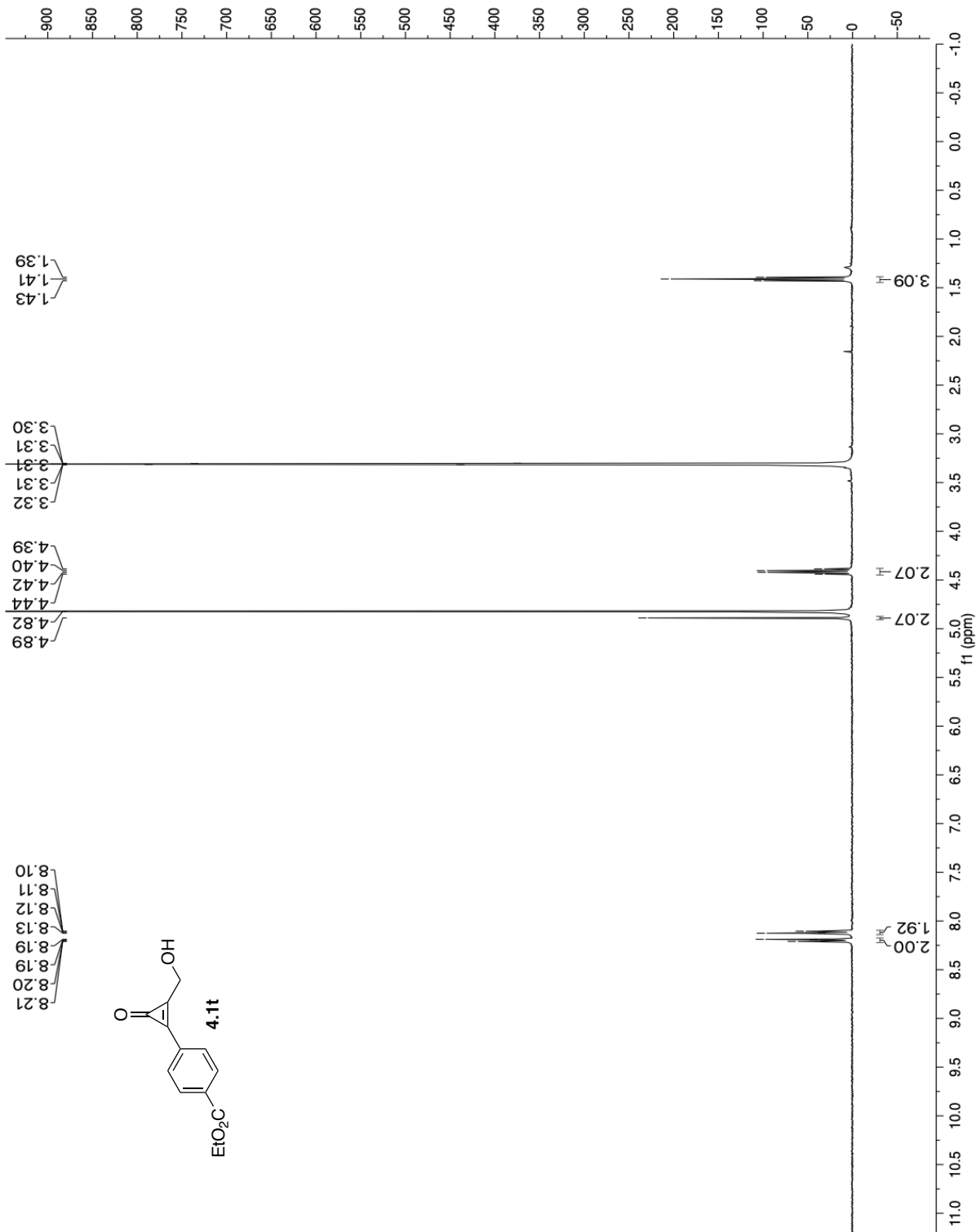


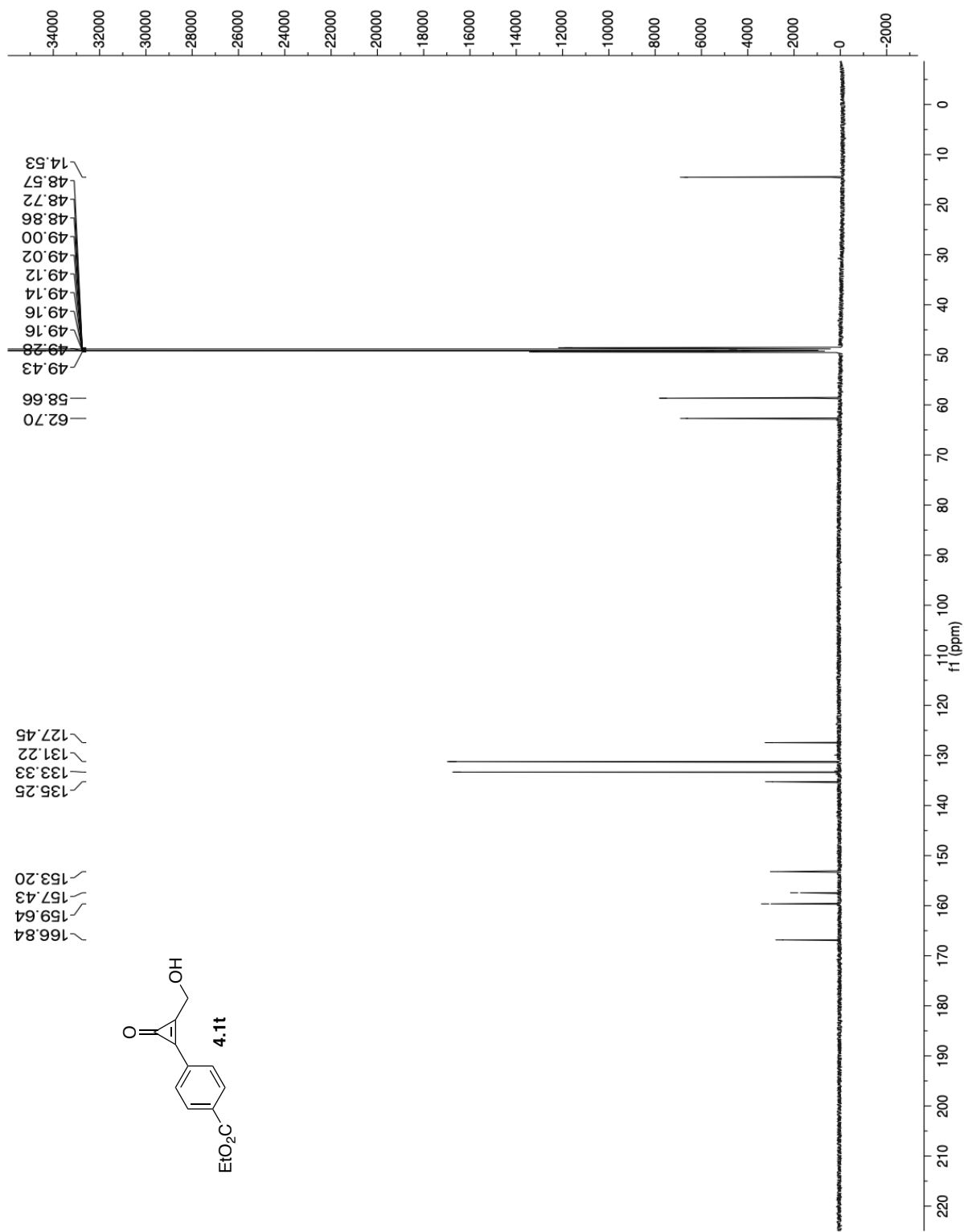


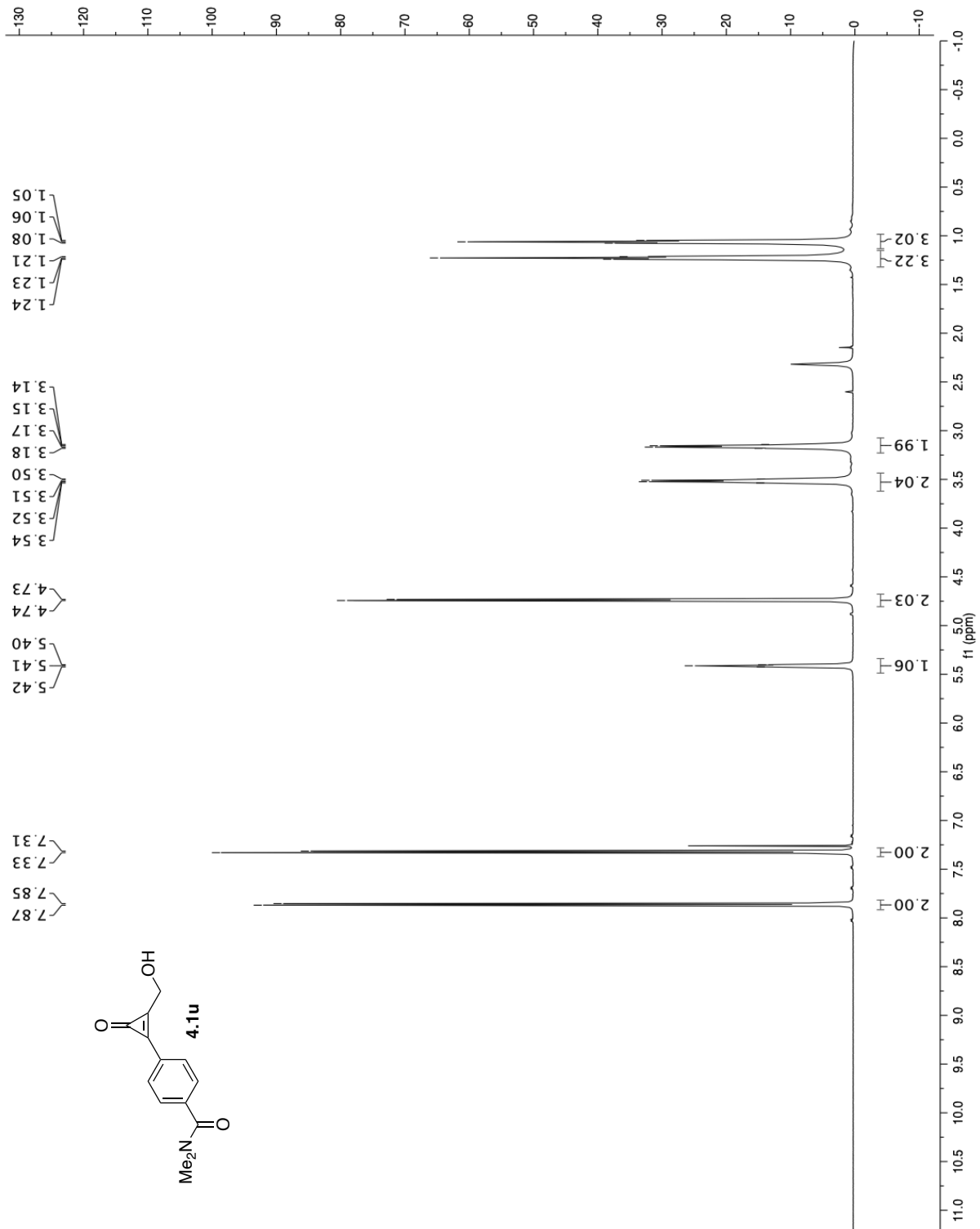


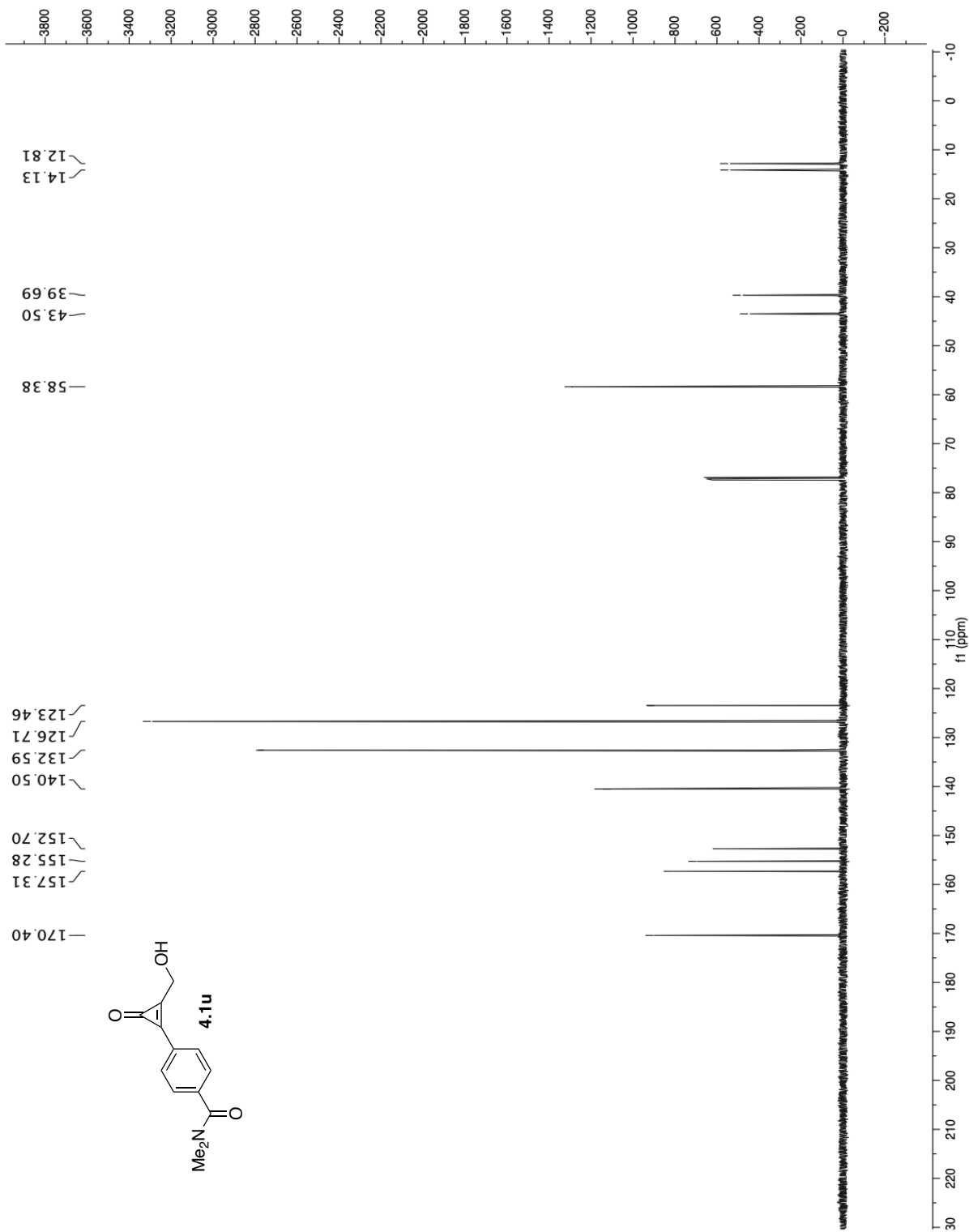


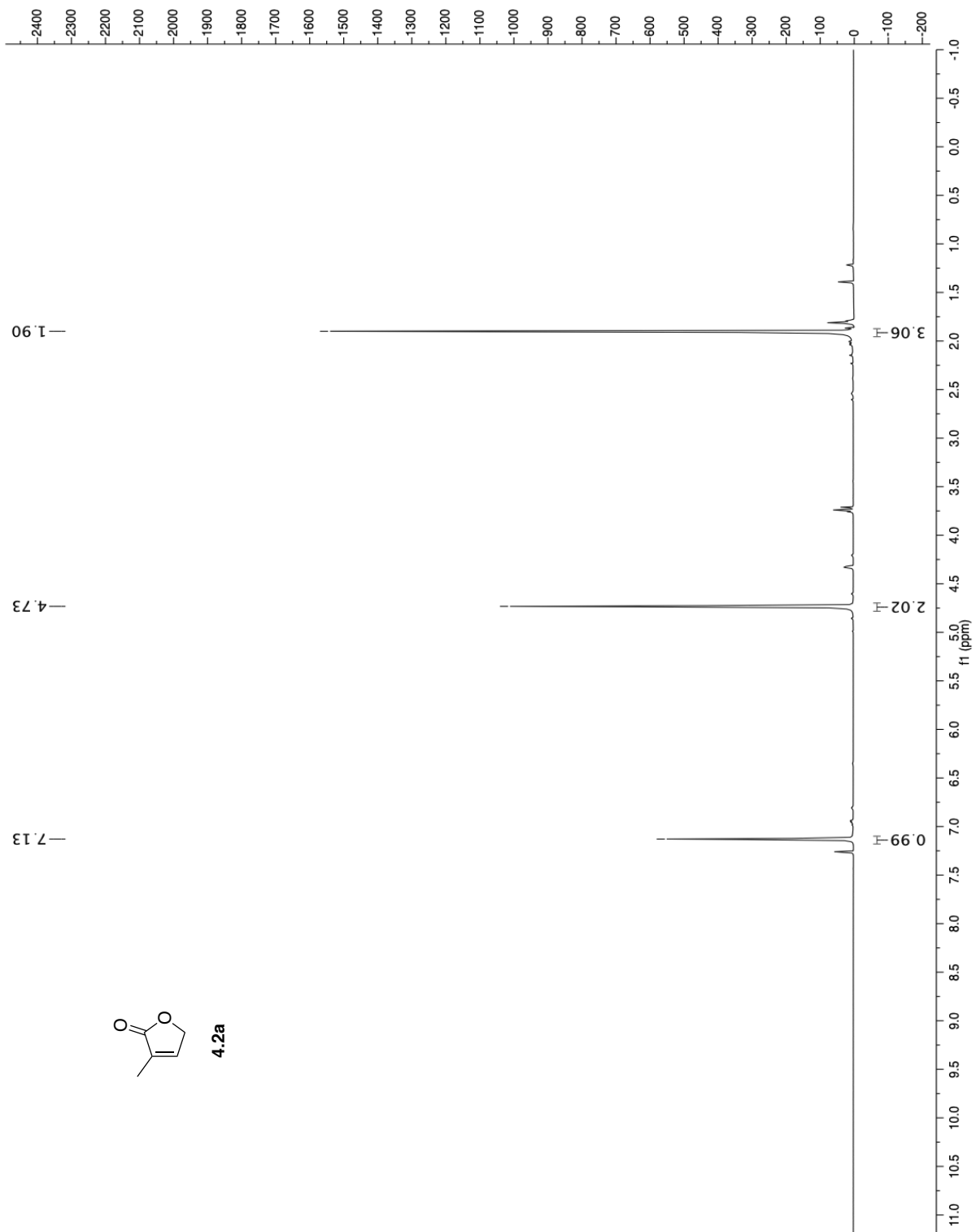




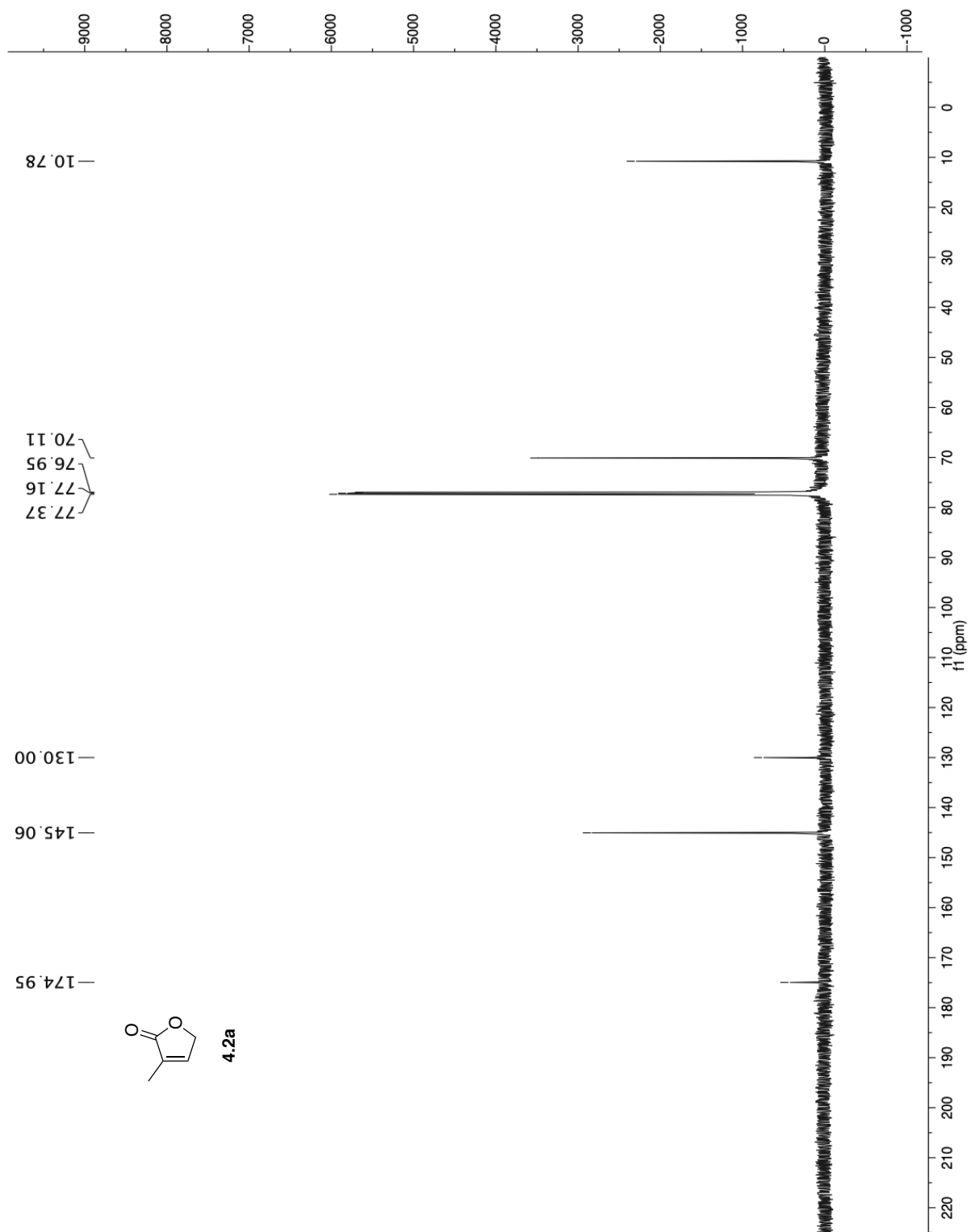


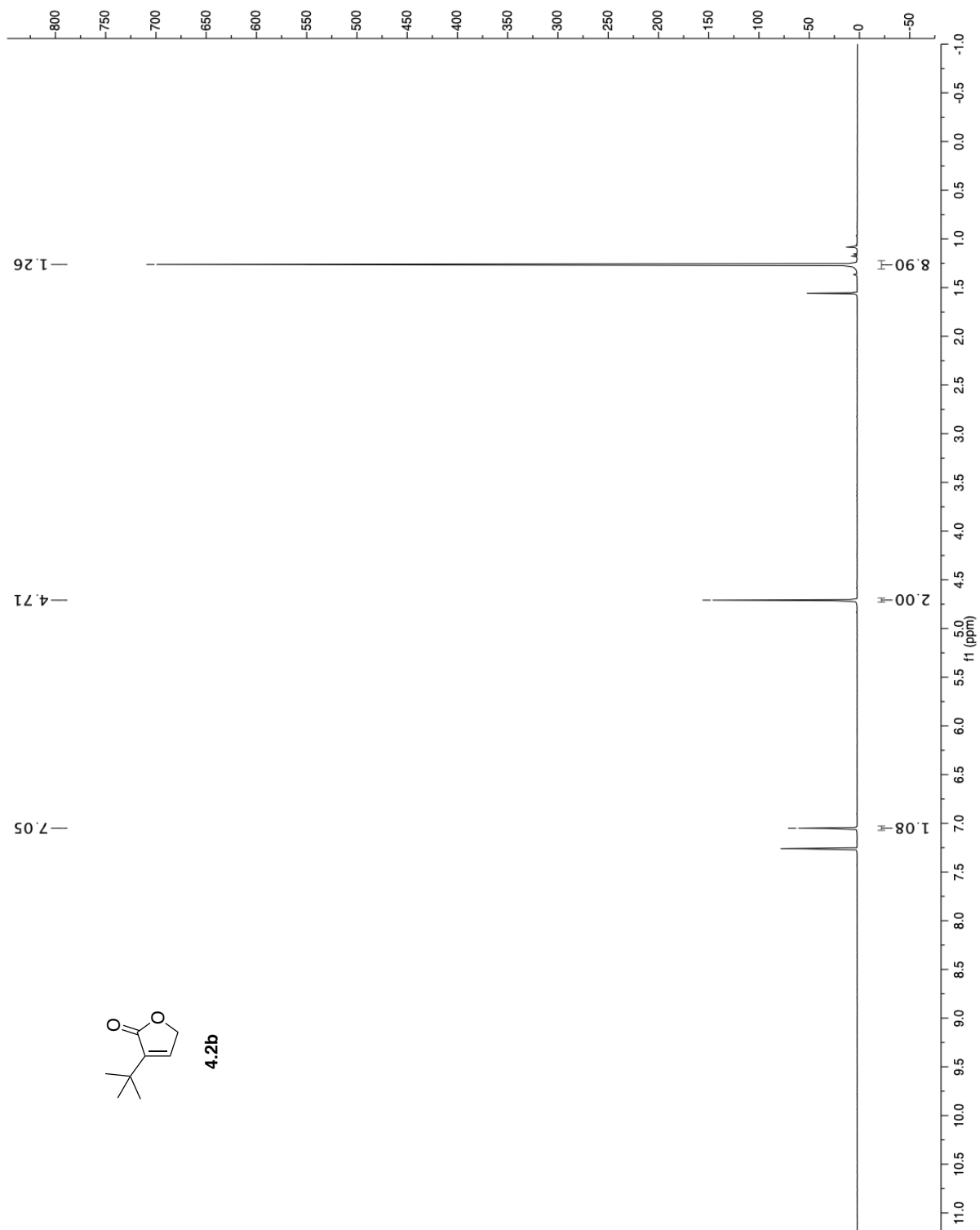


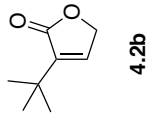
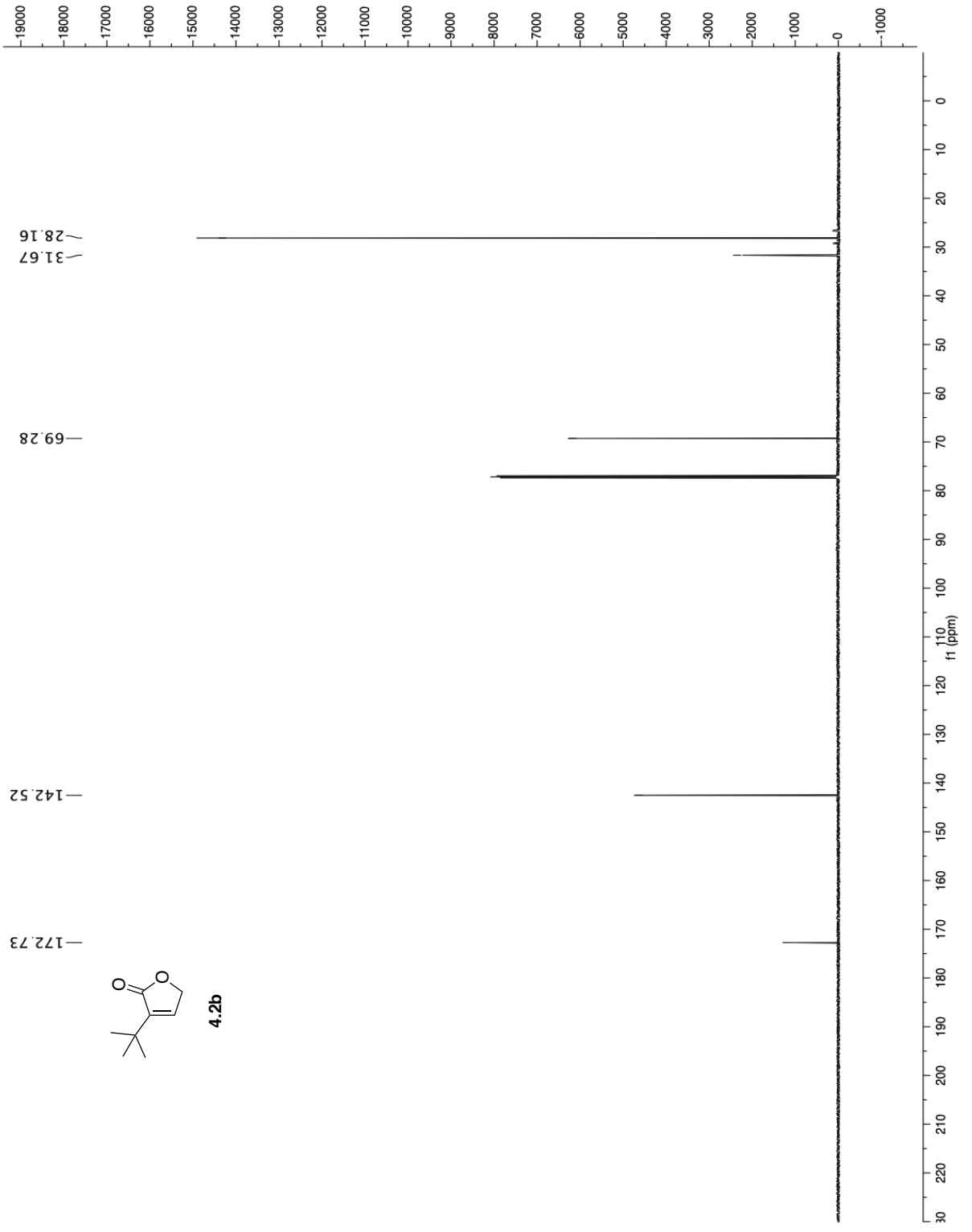


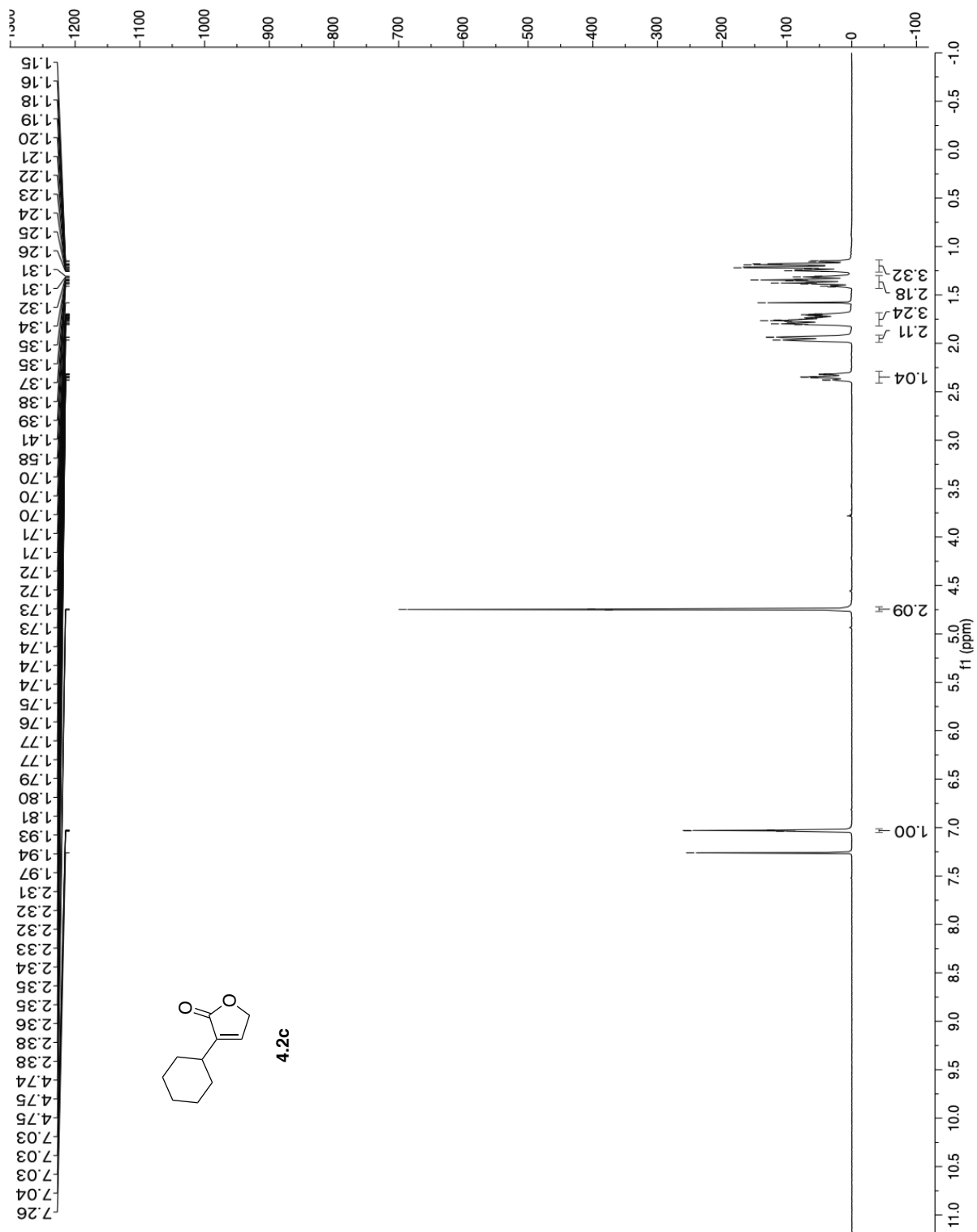


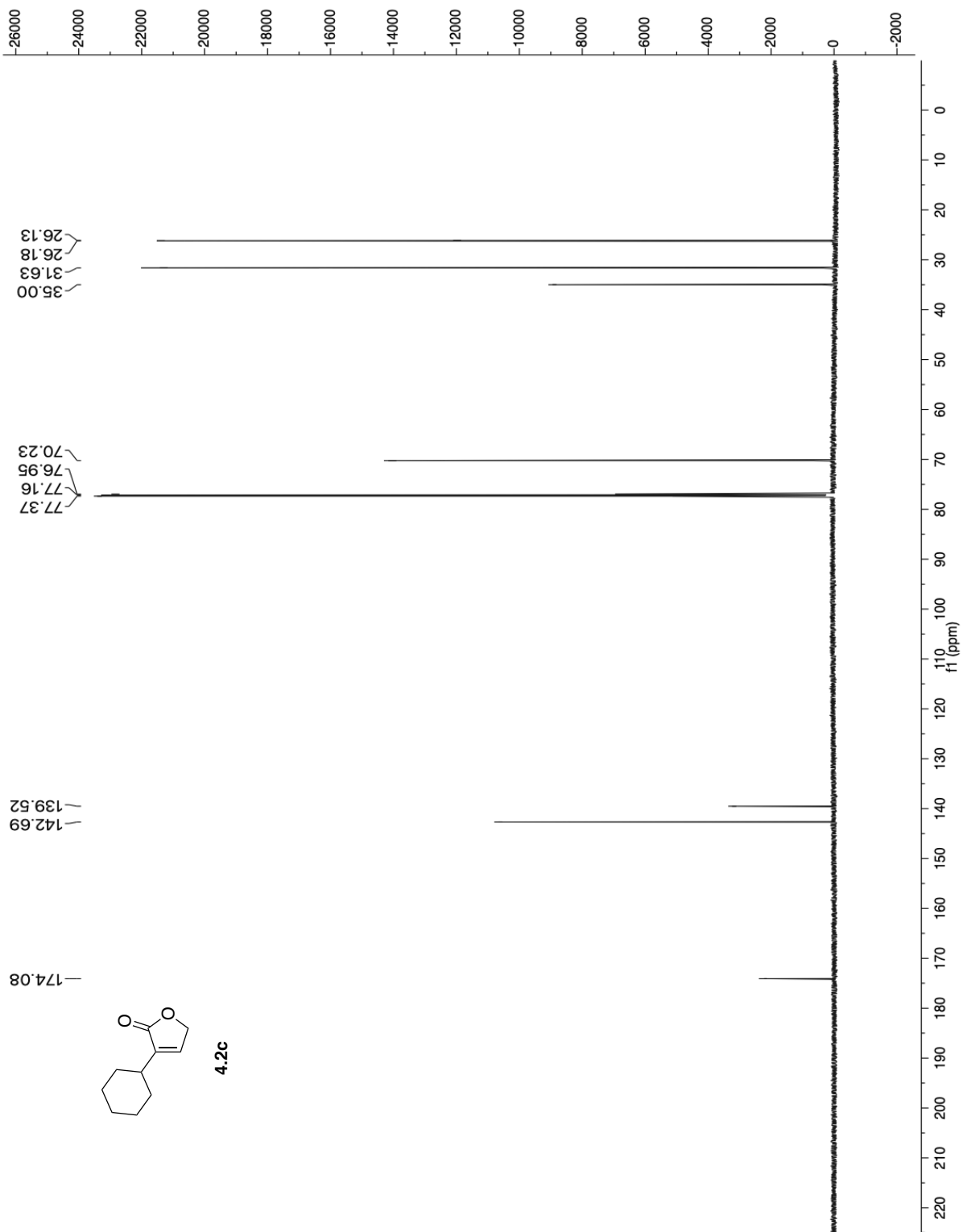
4.2a

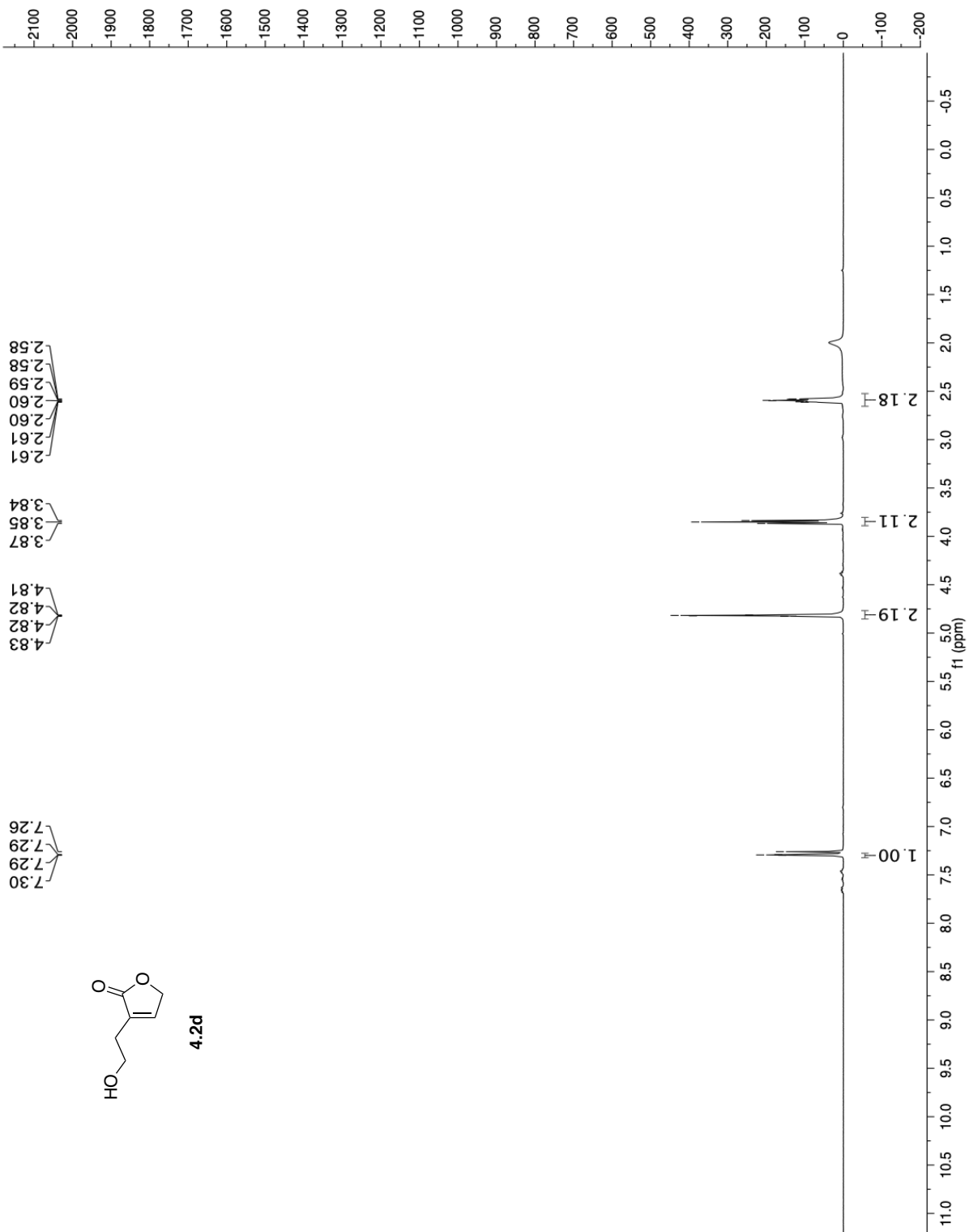


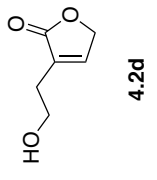
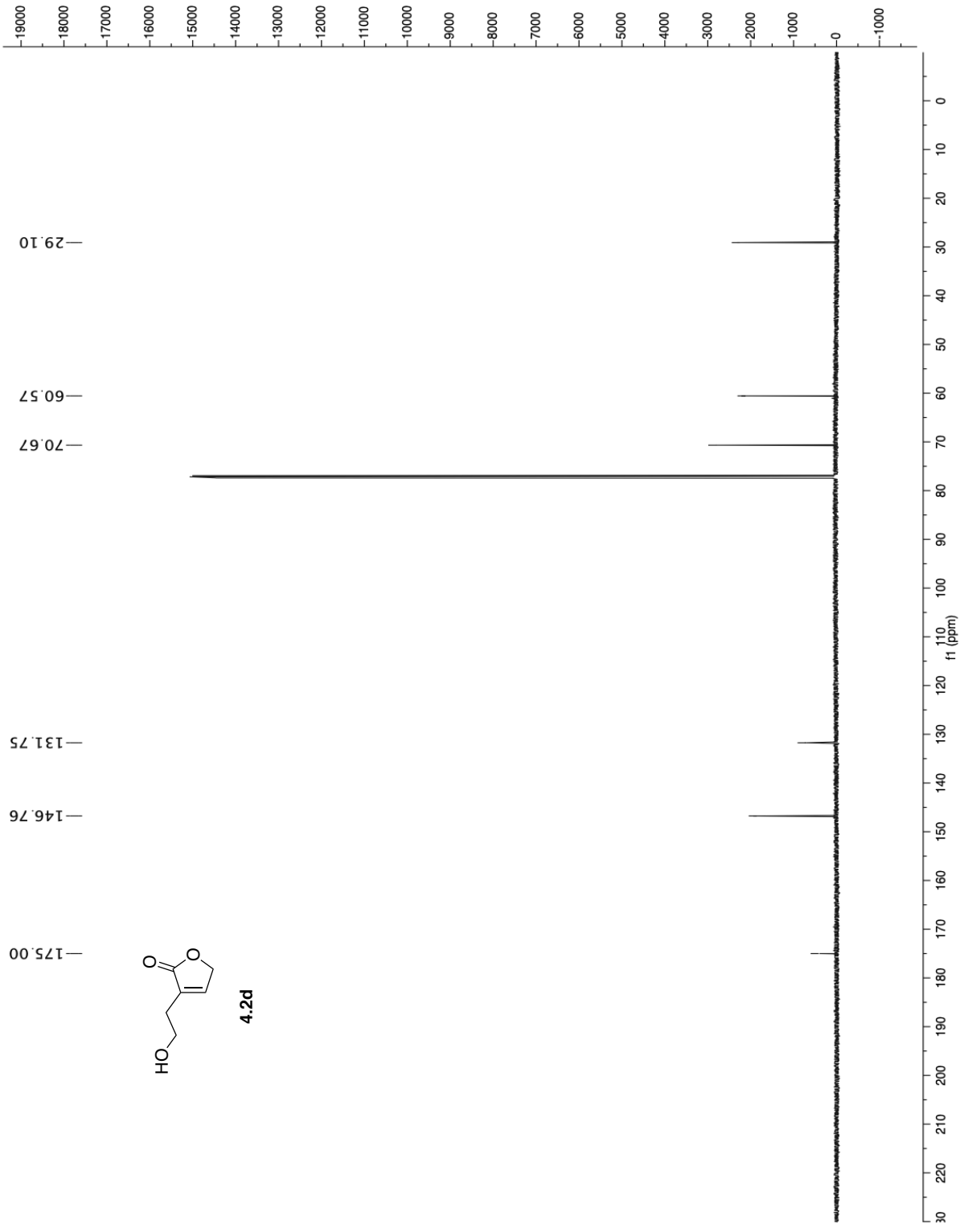


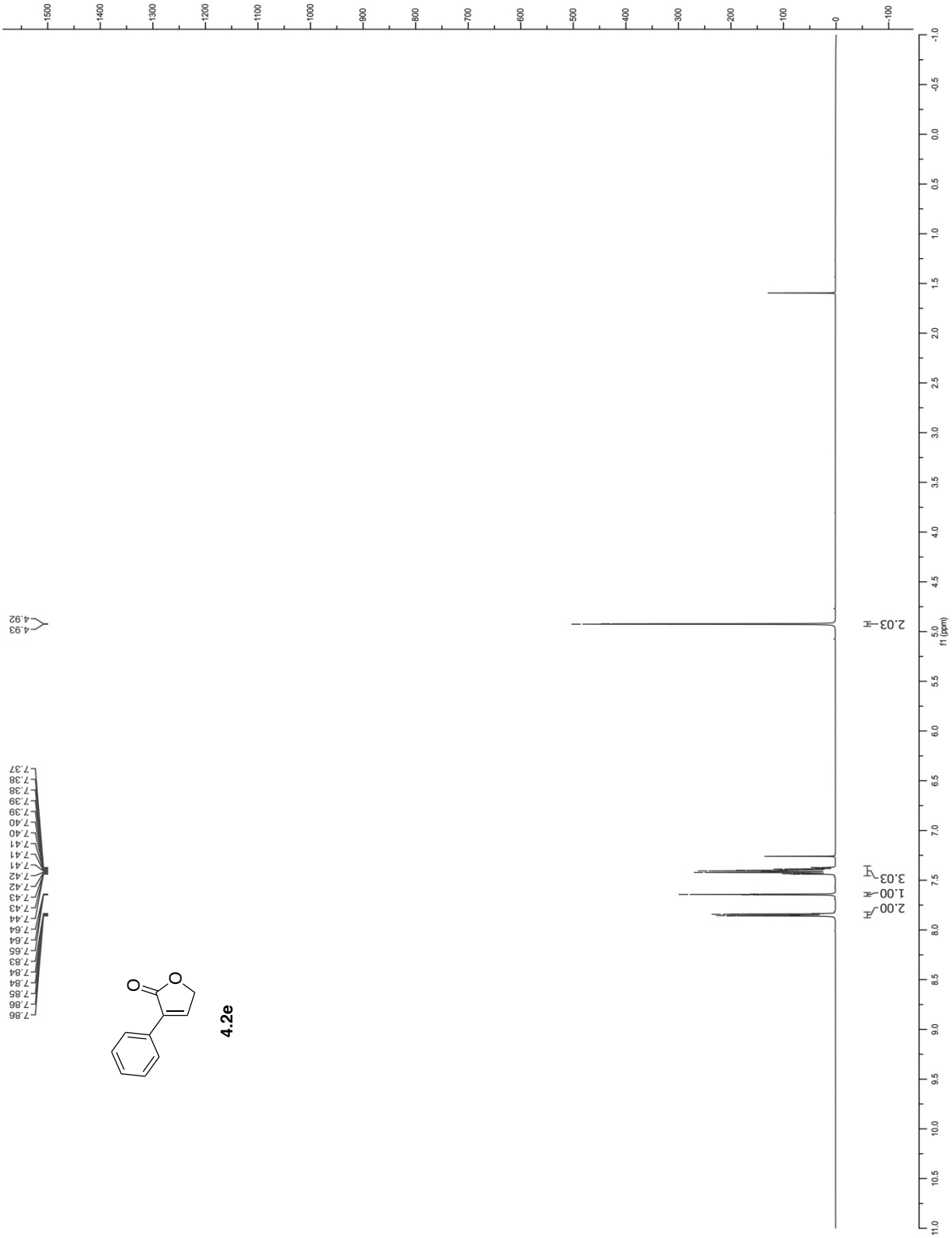


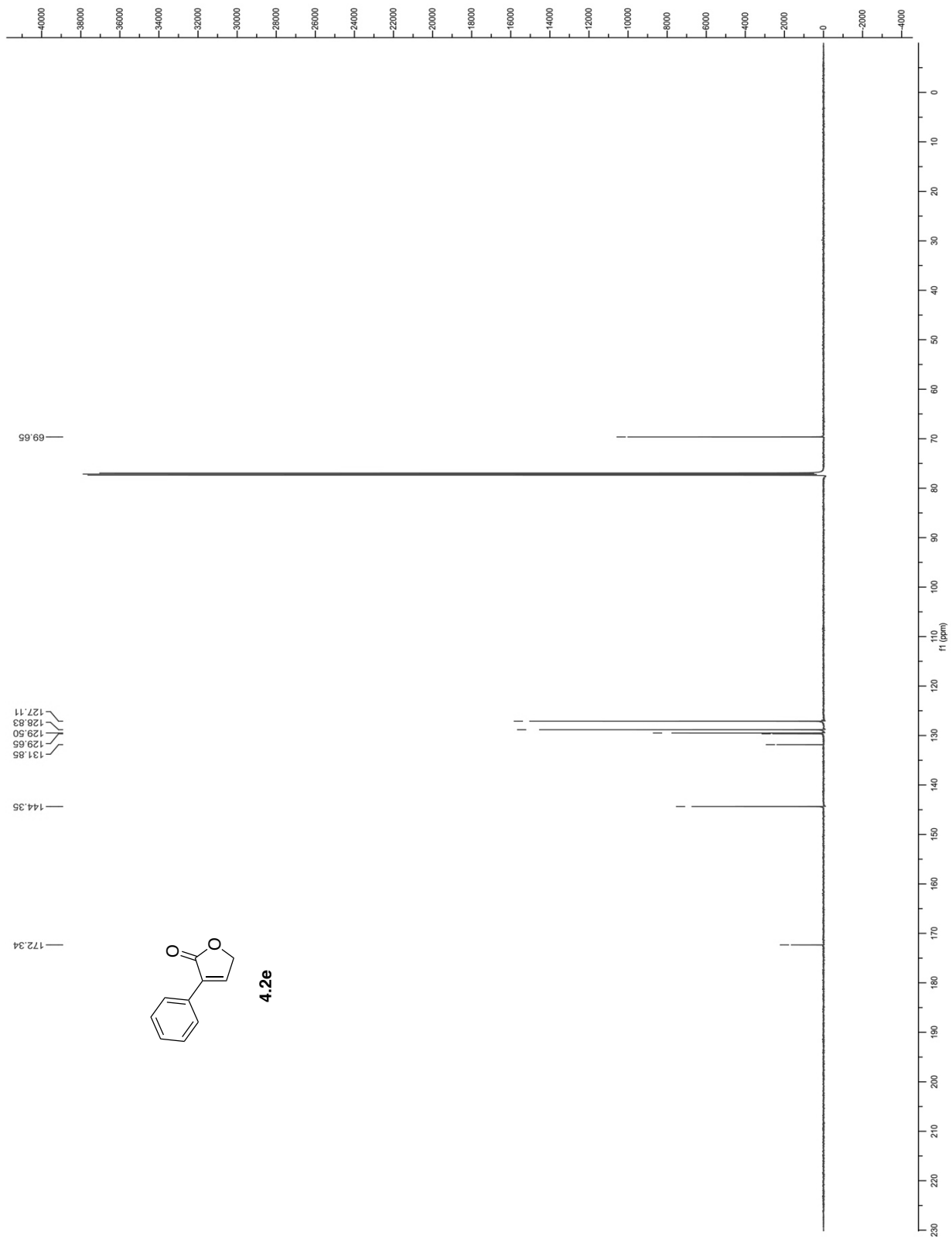


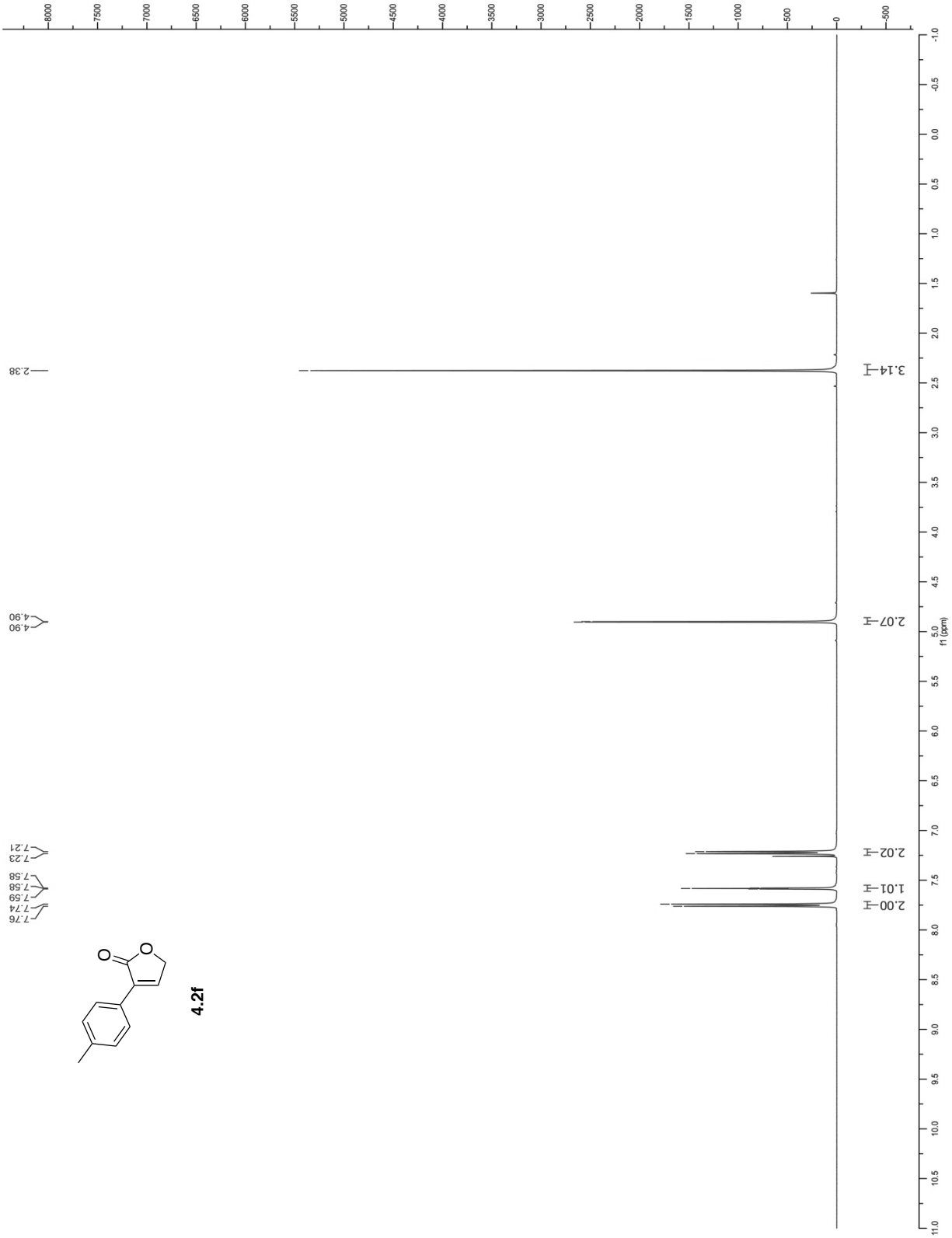


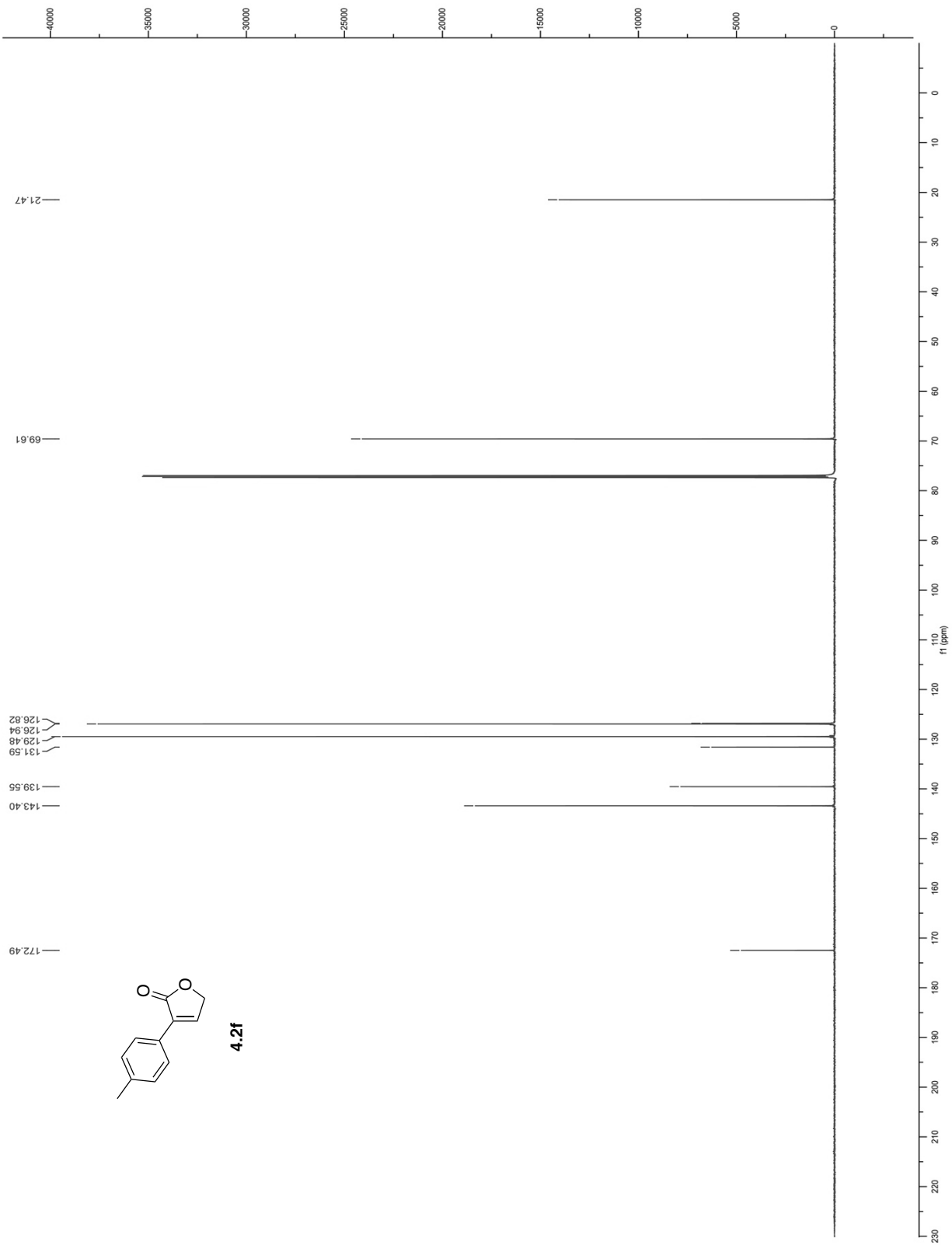


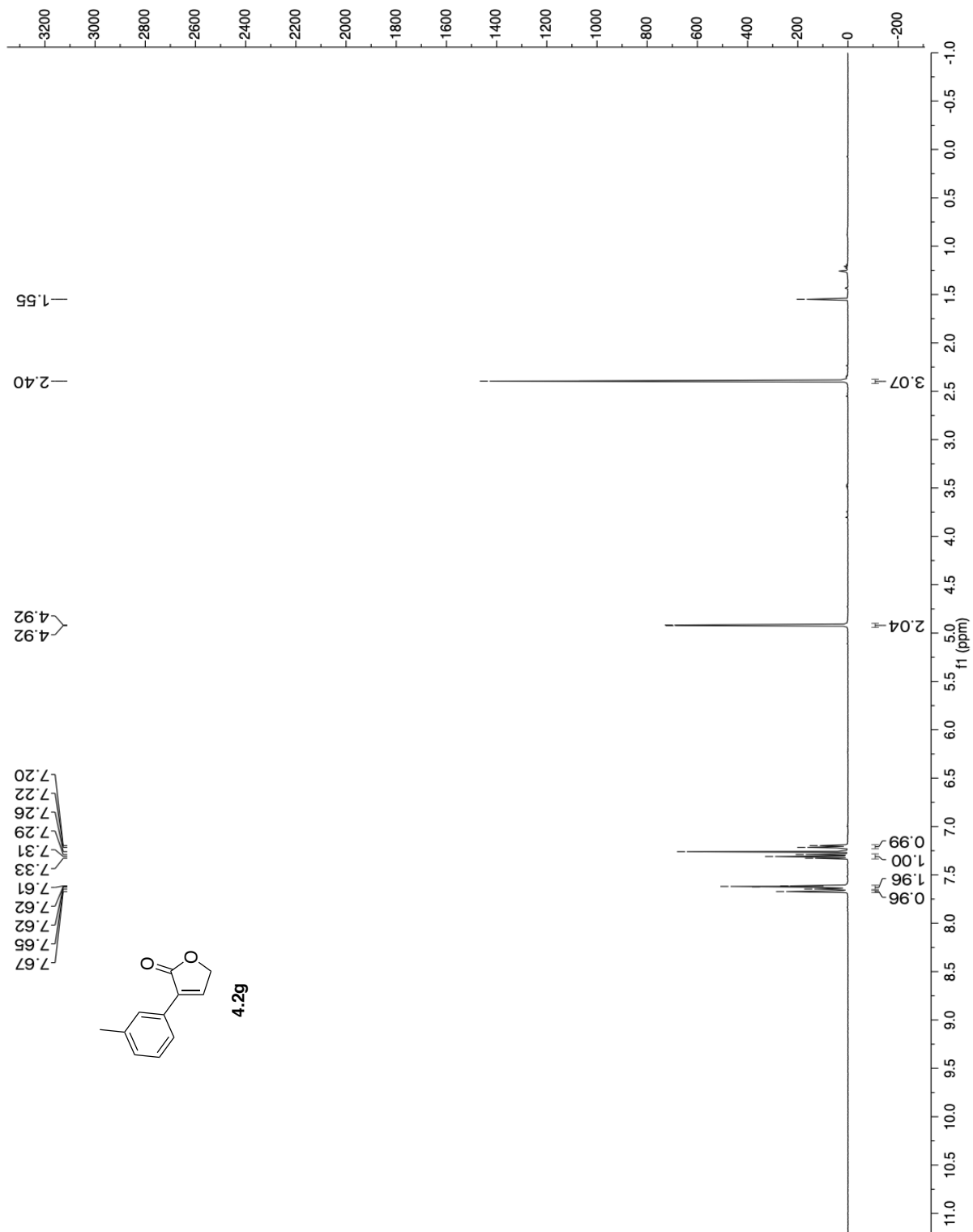


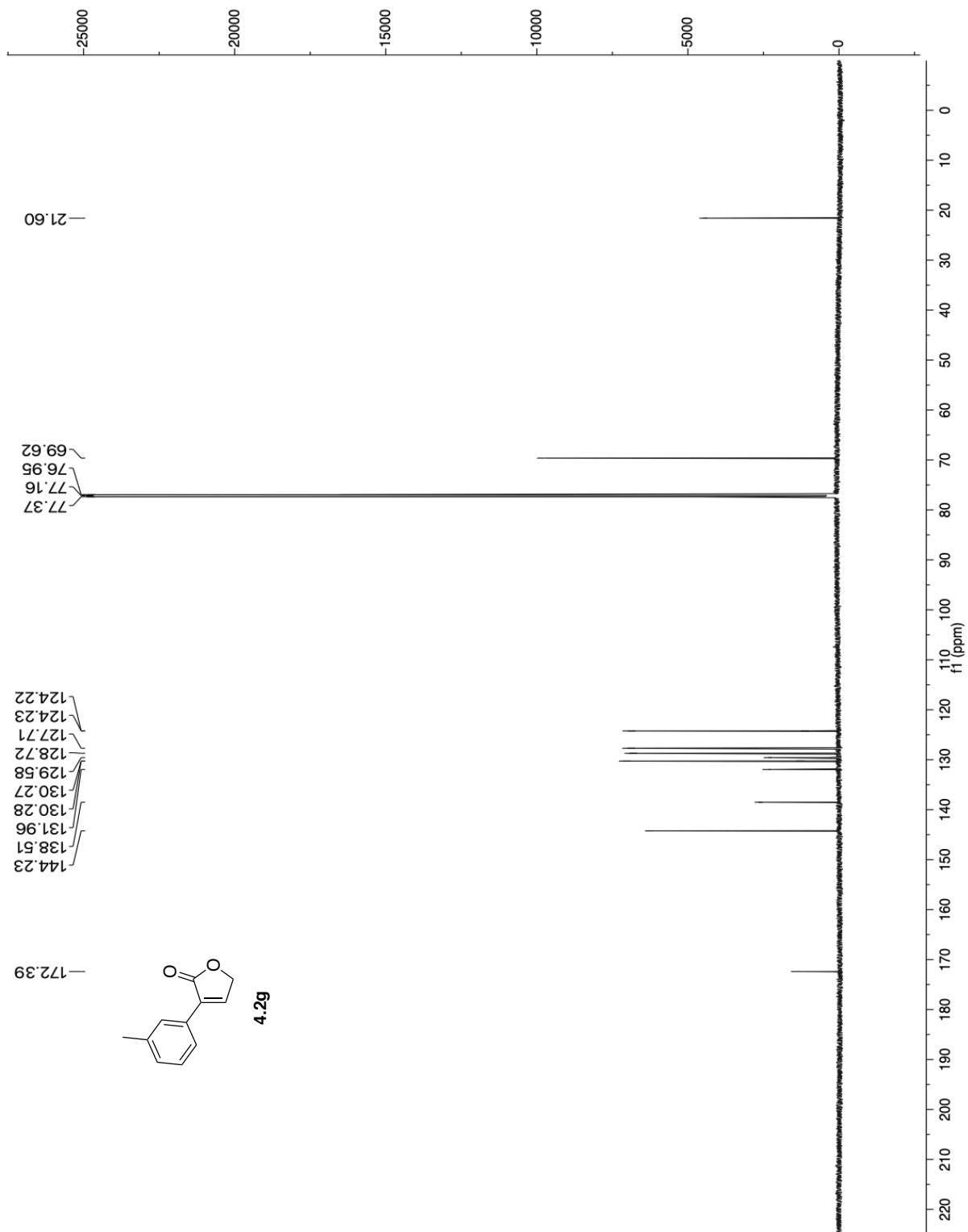


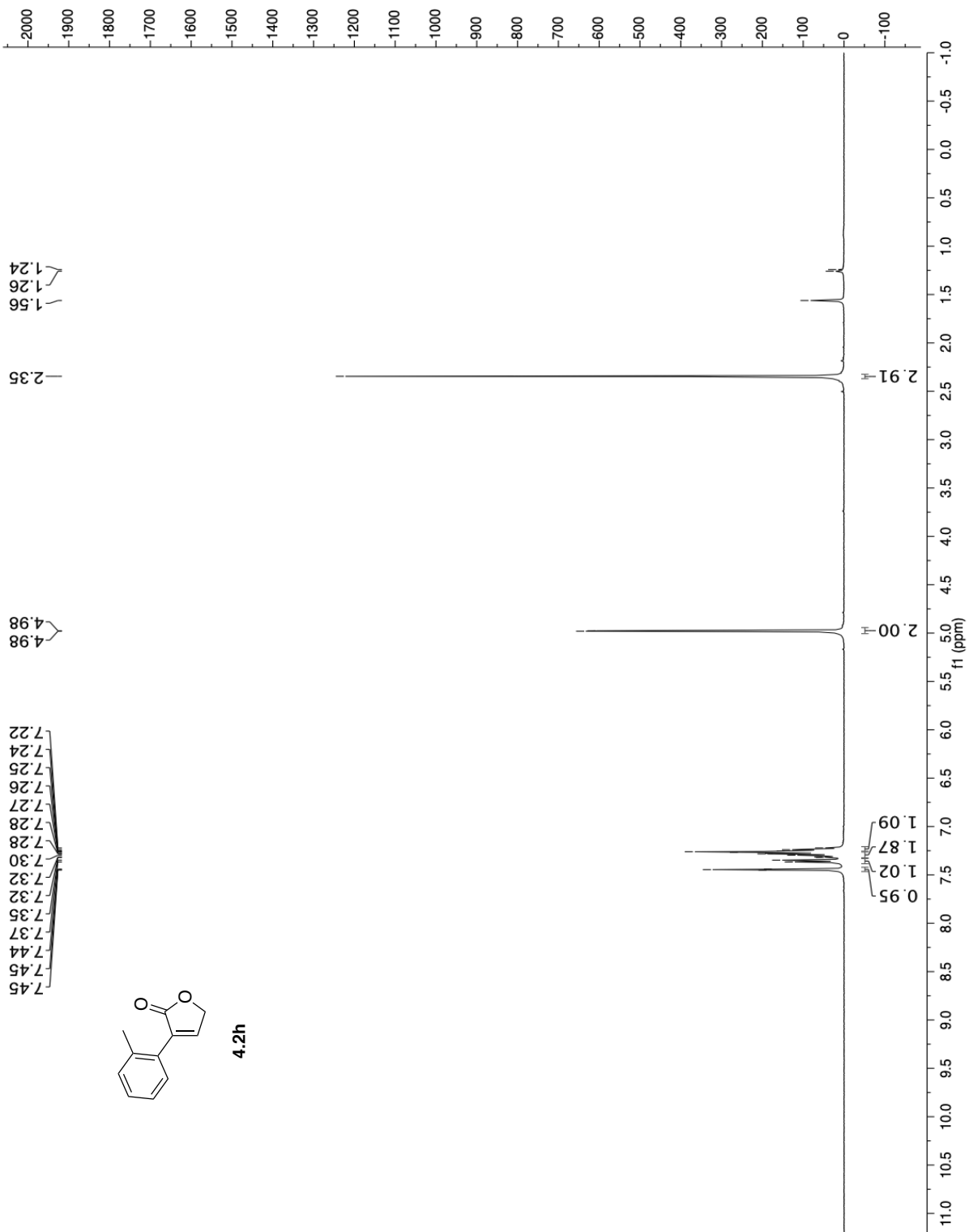


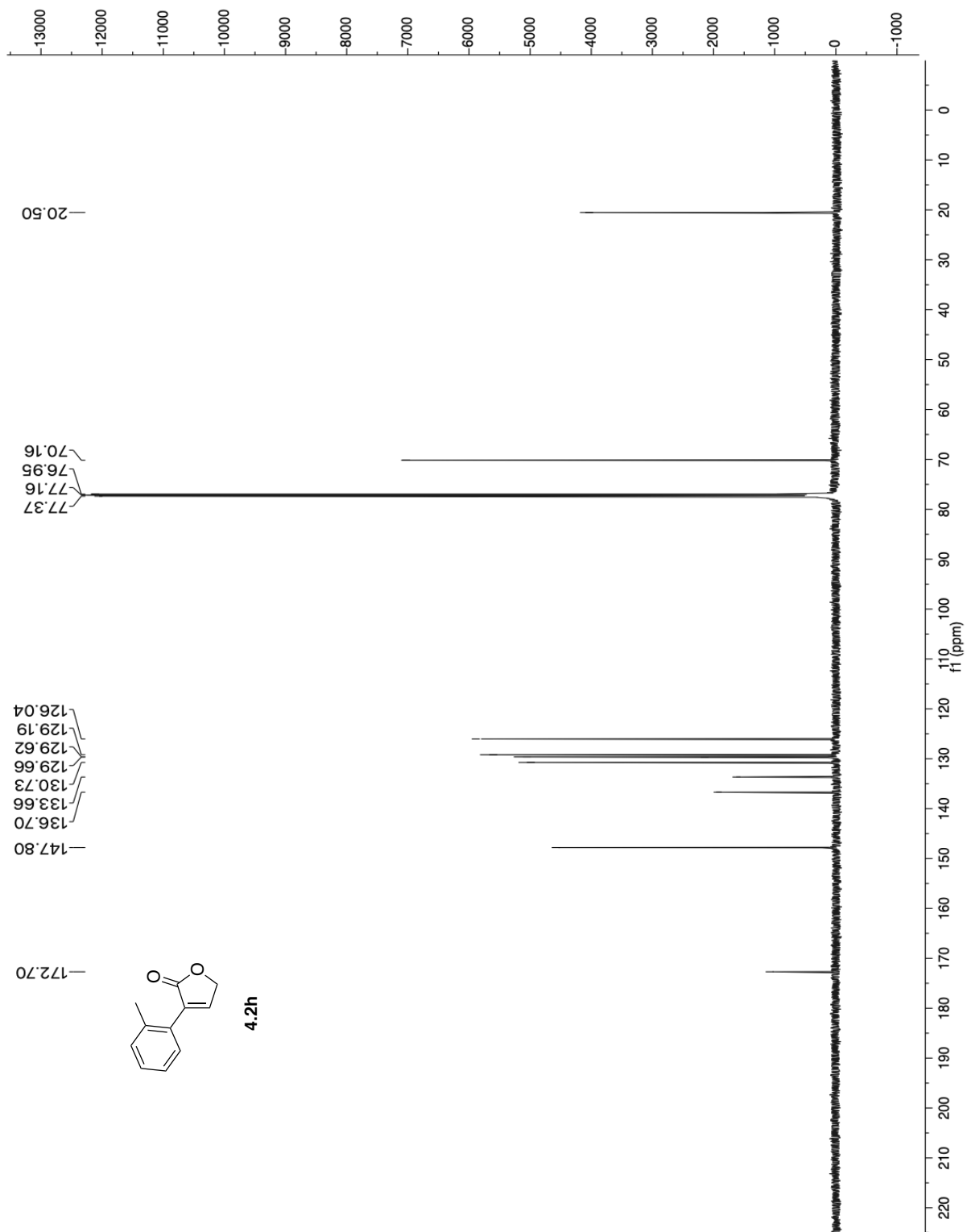


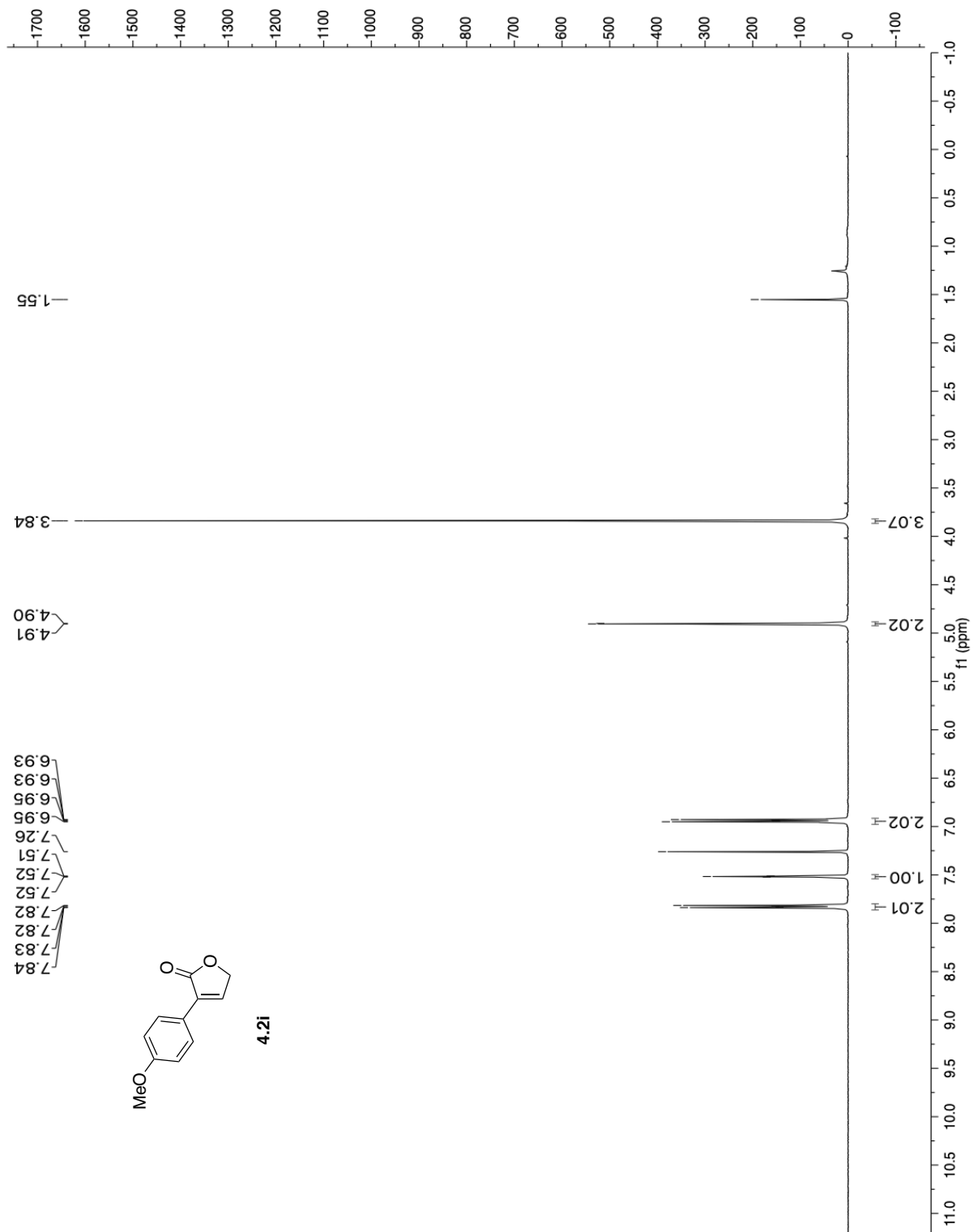


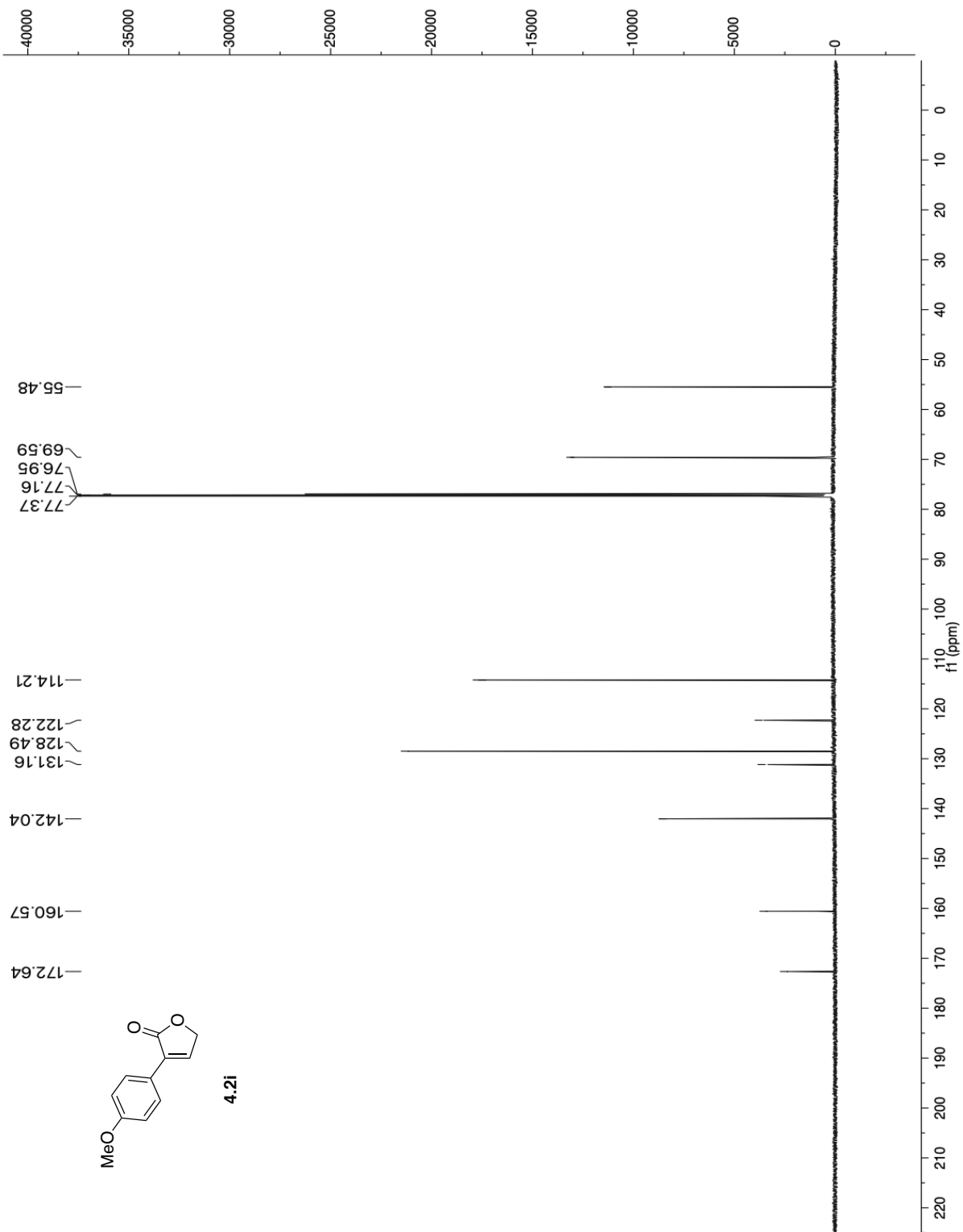


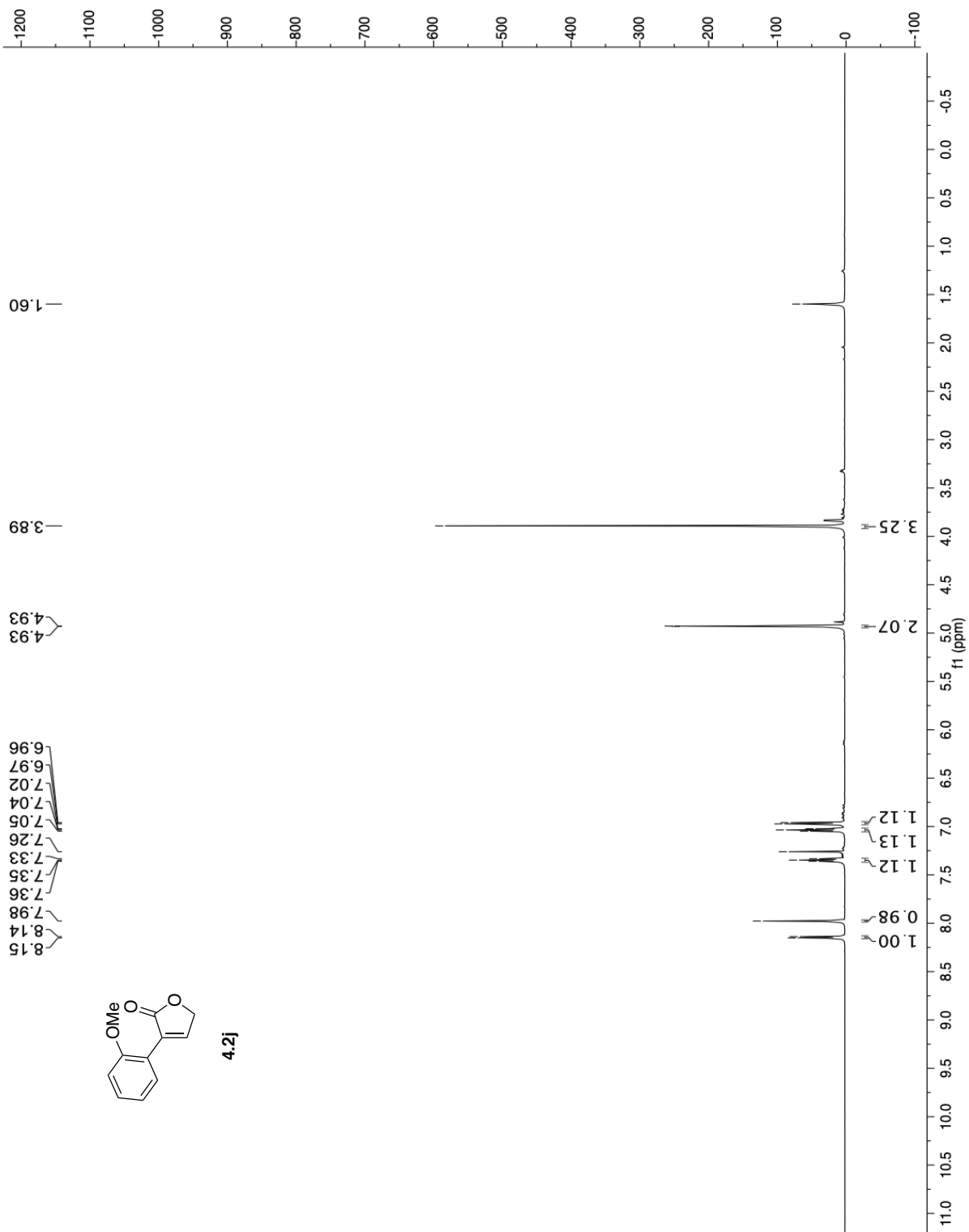


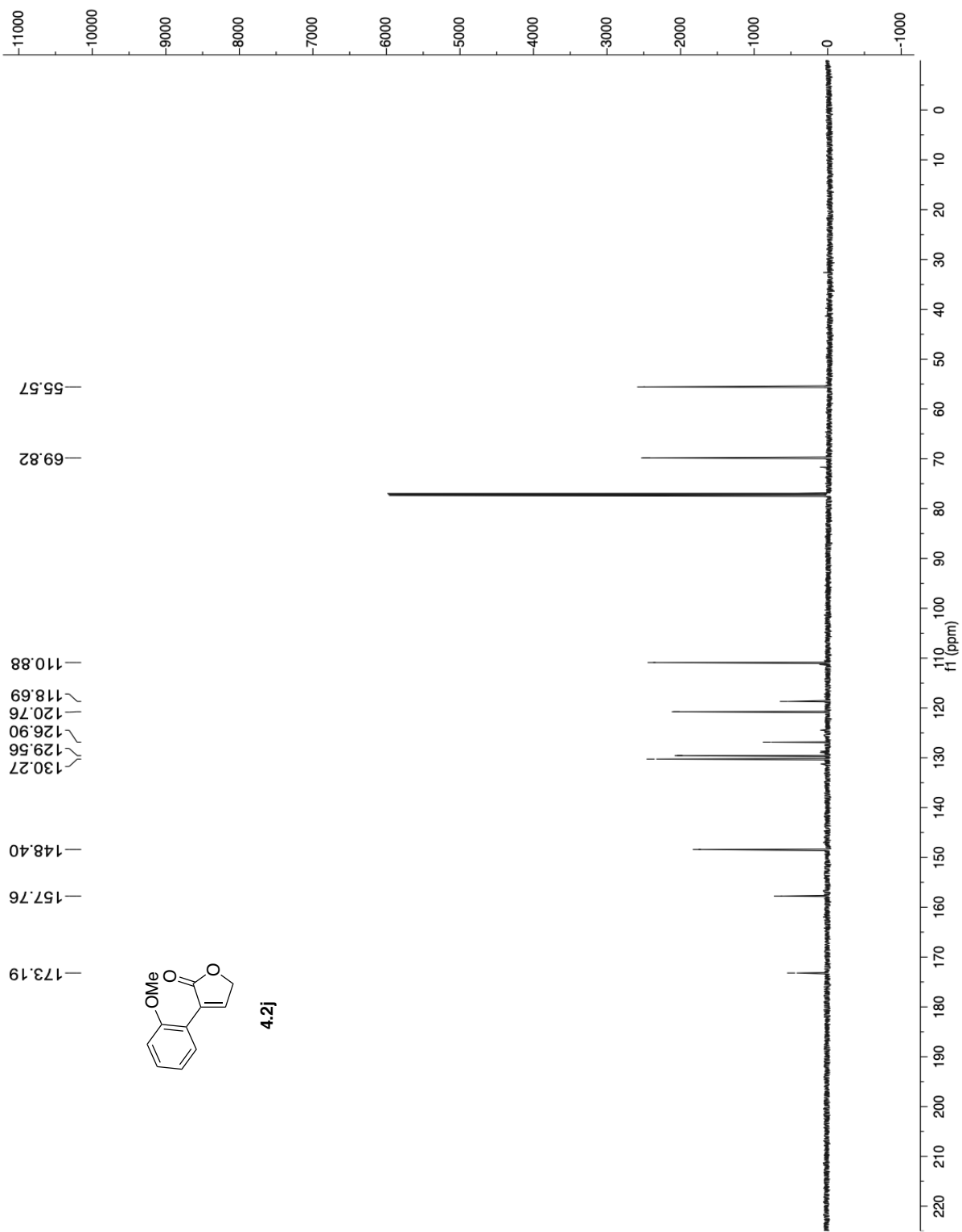


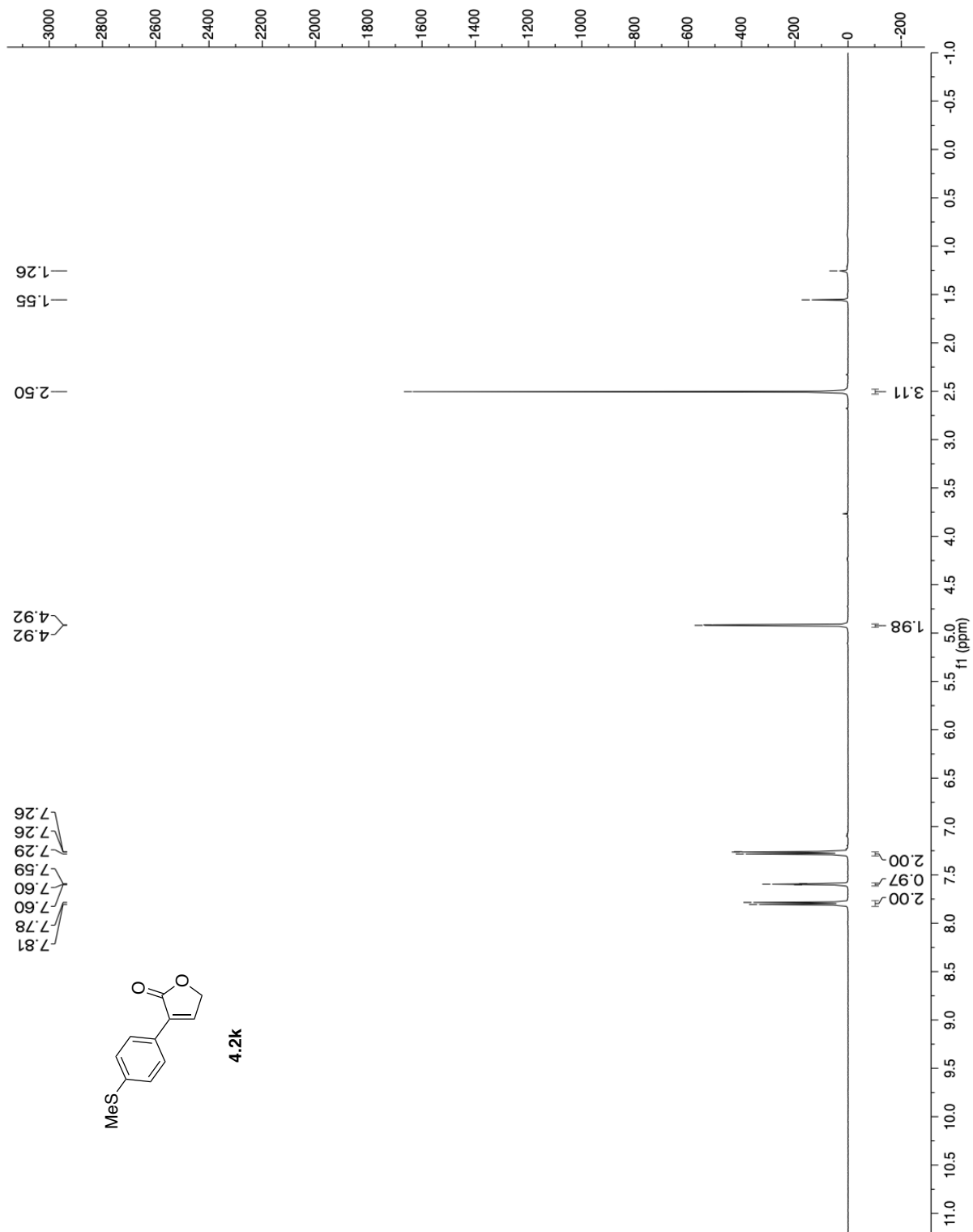


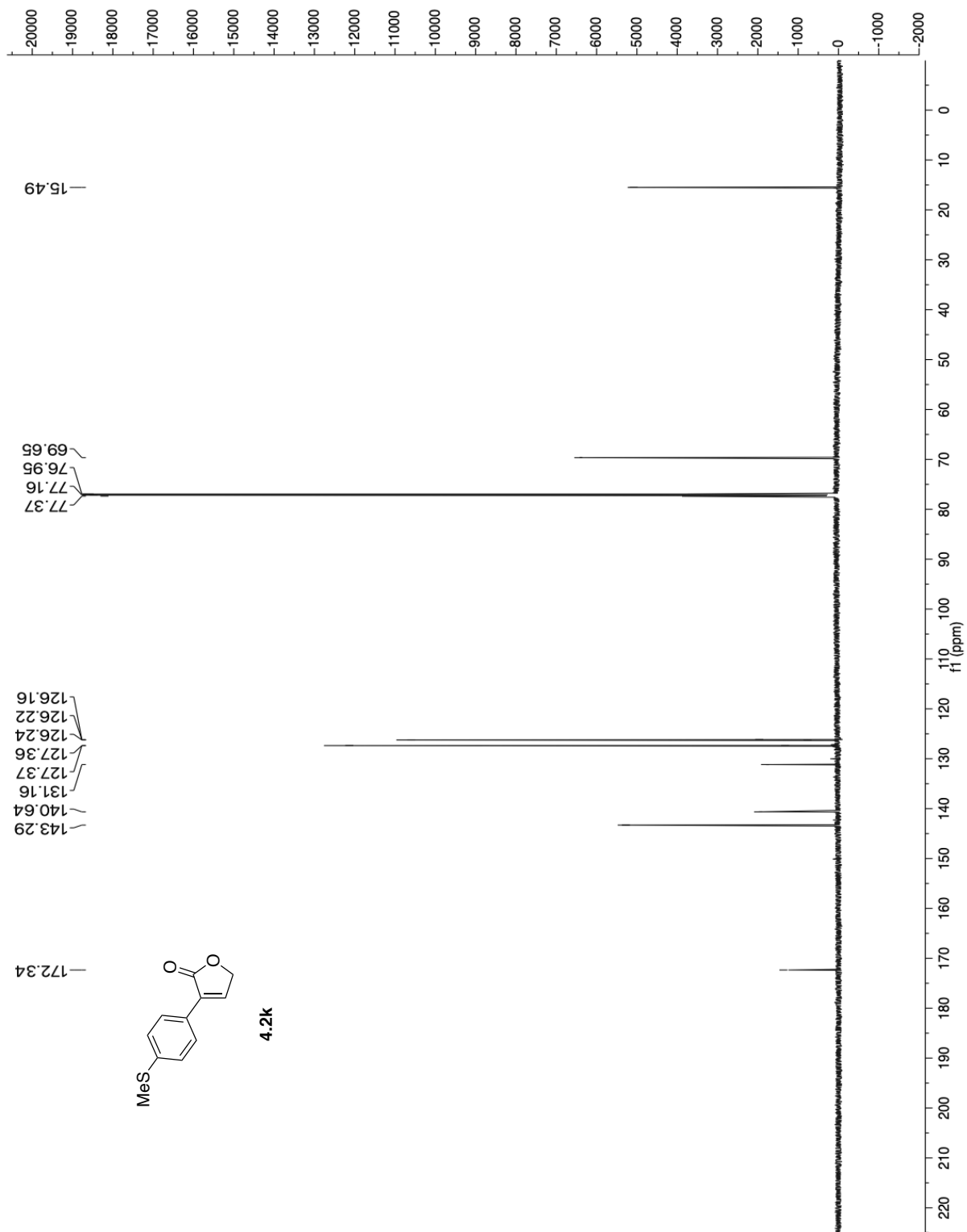


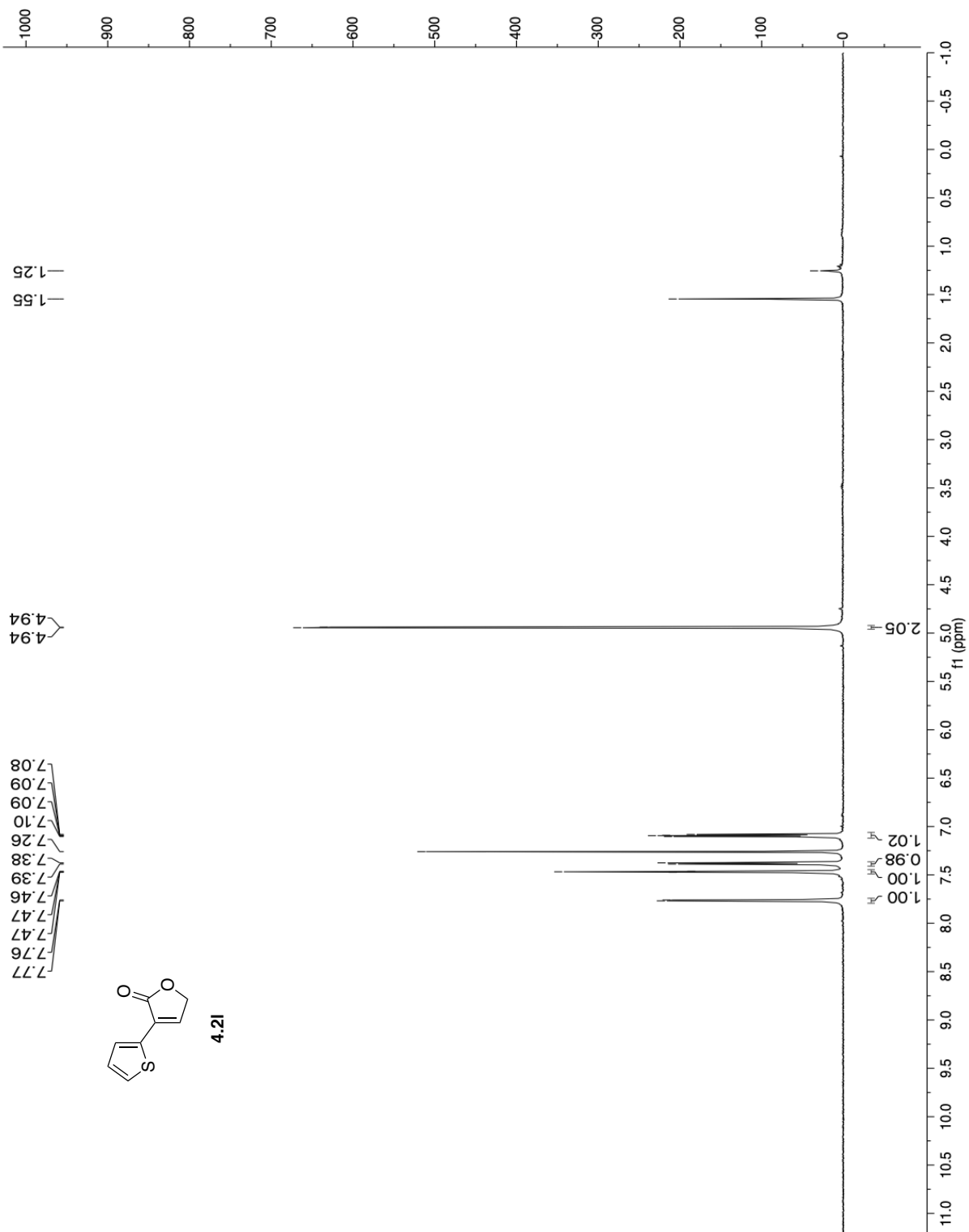


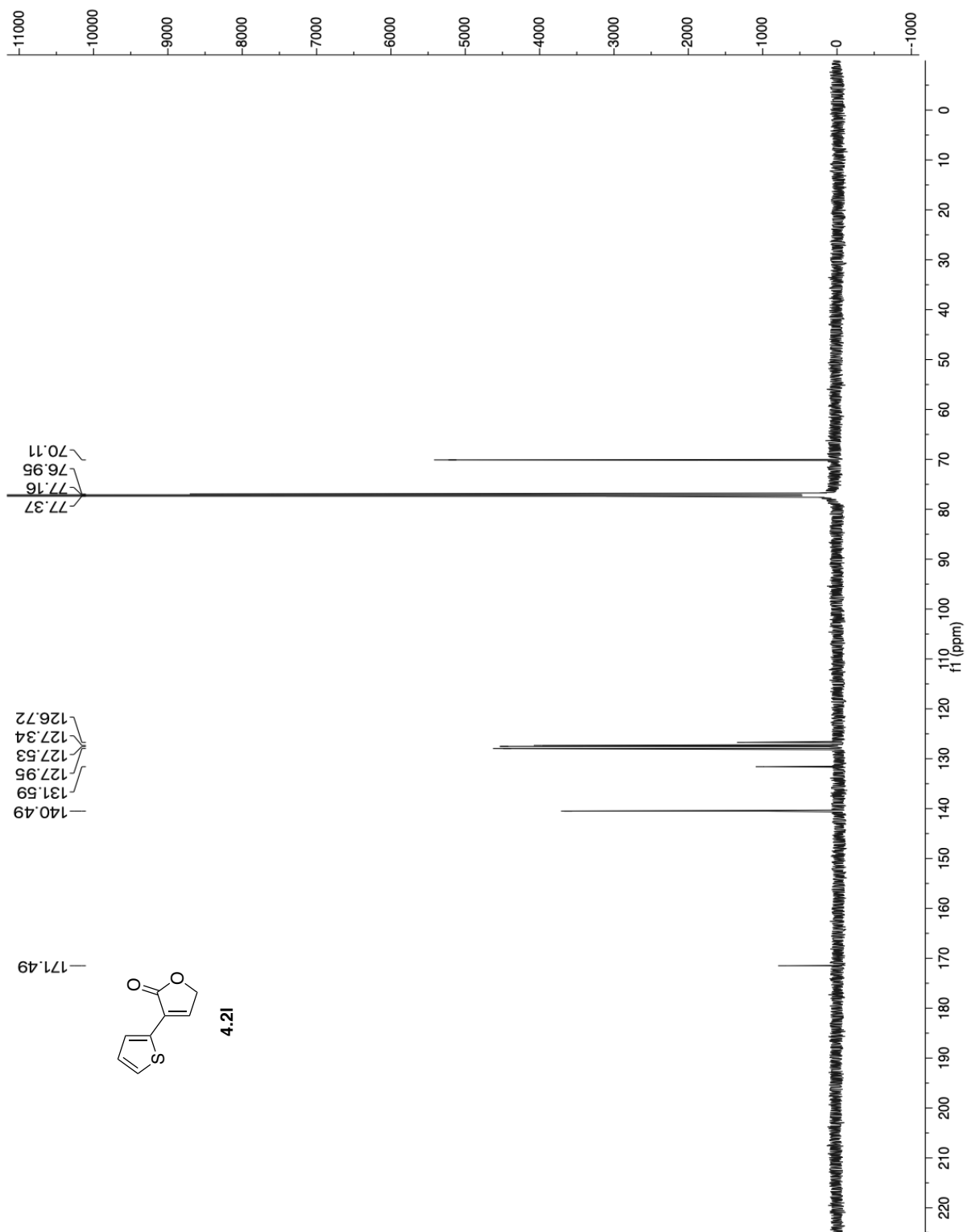


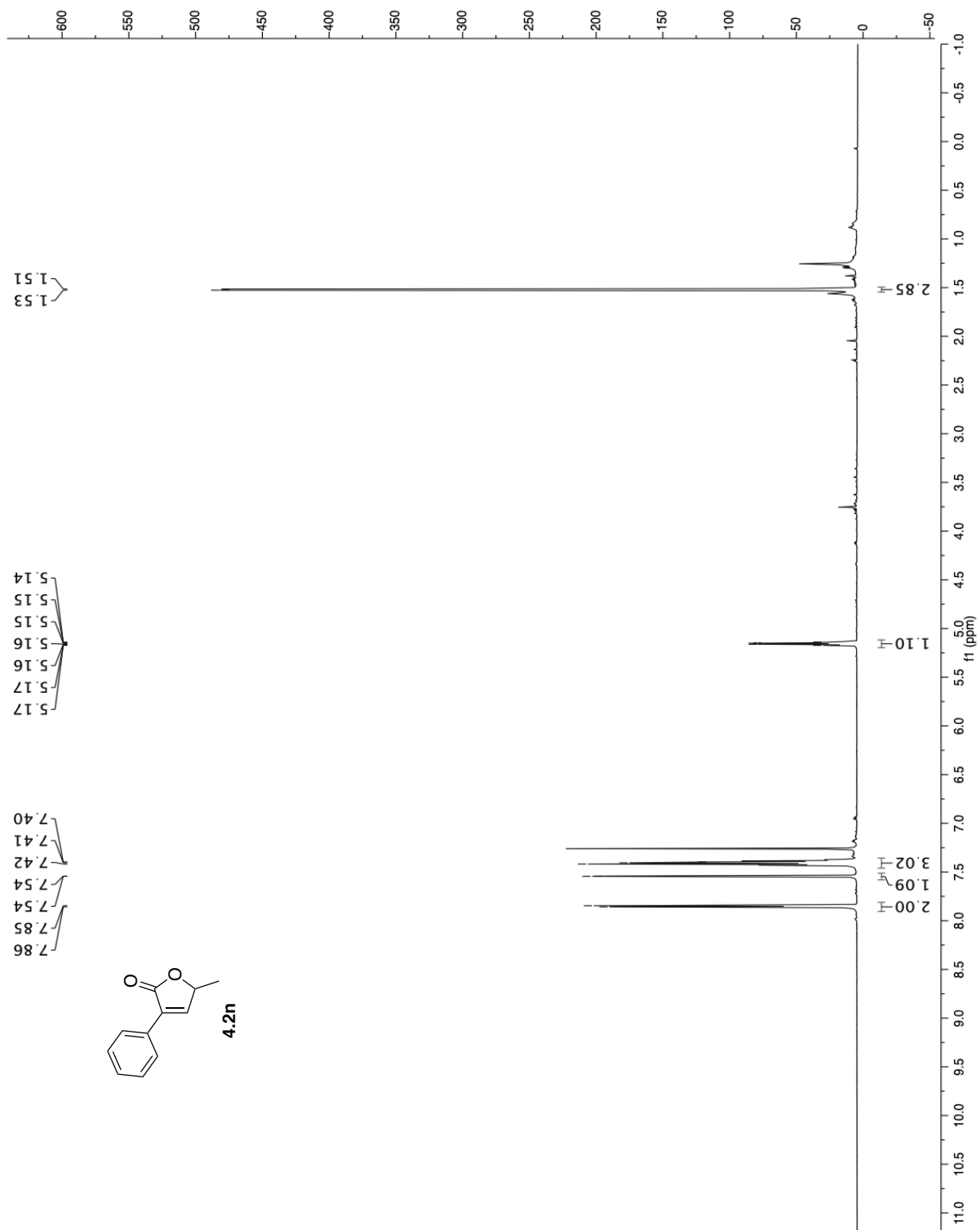


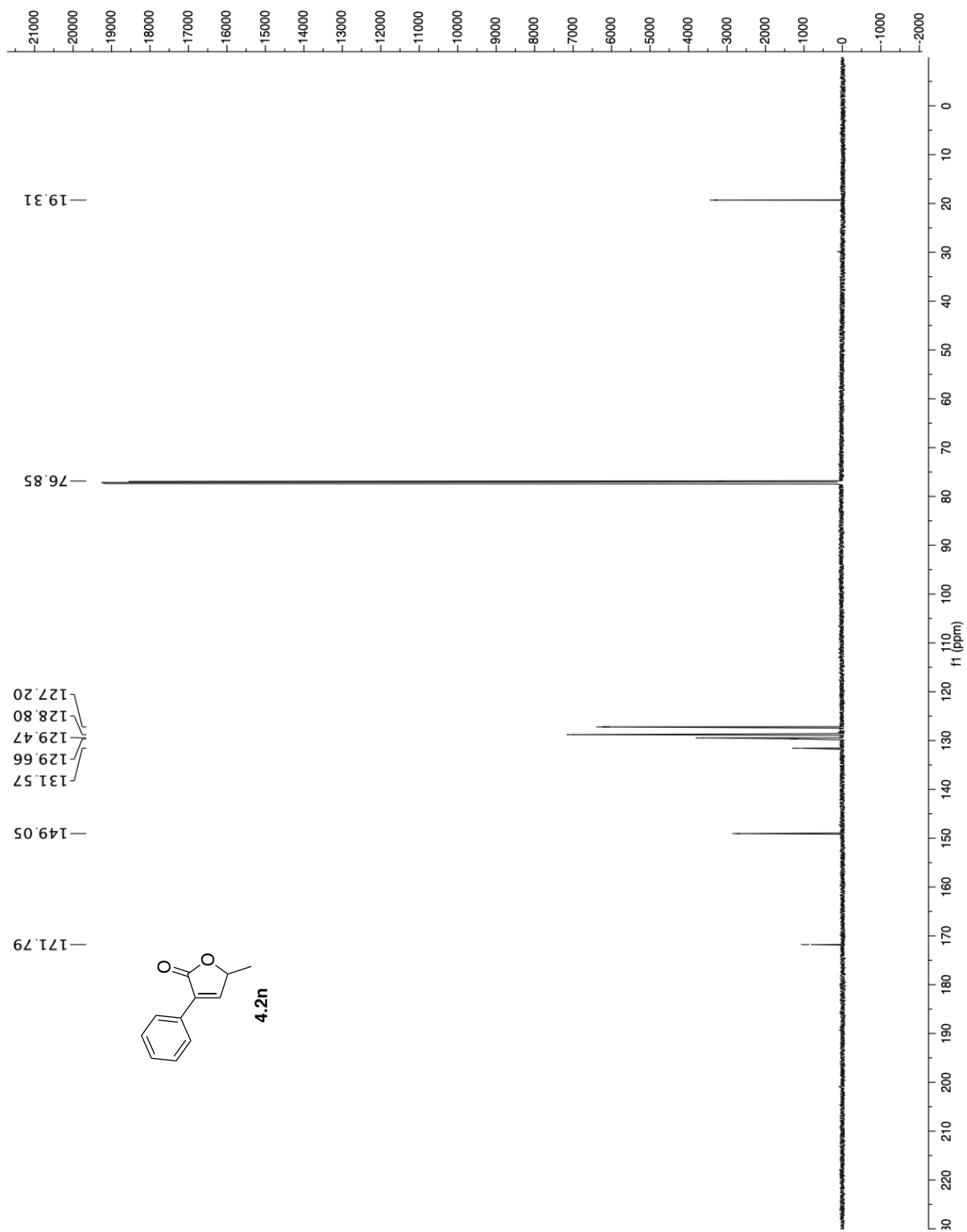


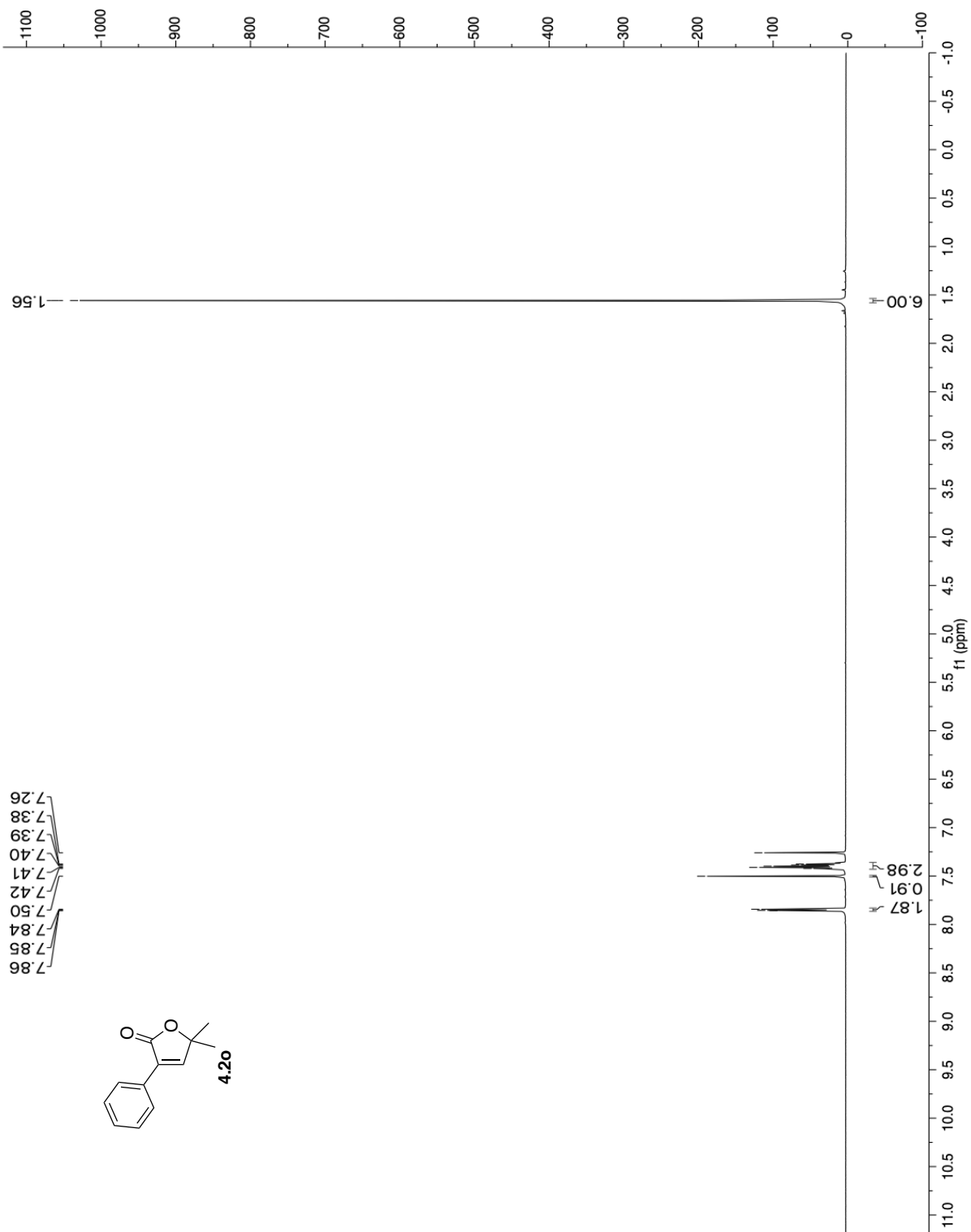


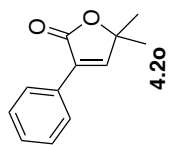
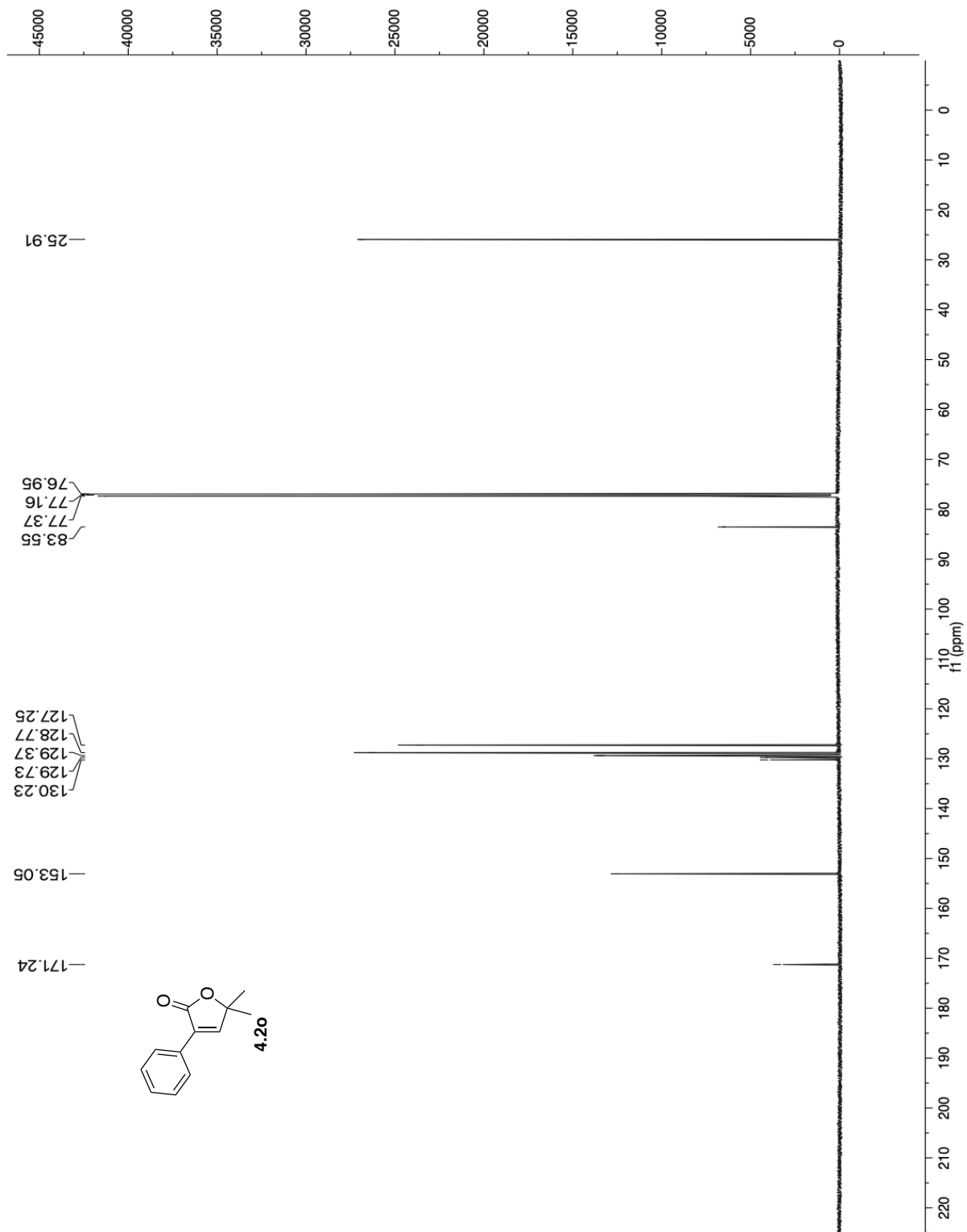


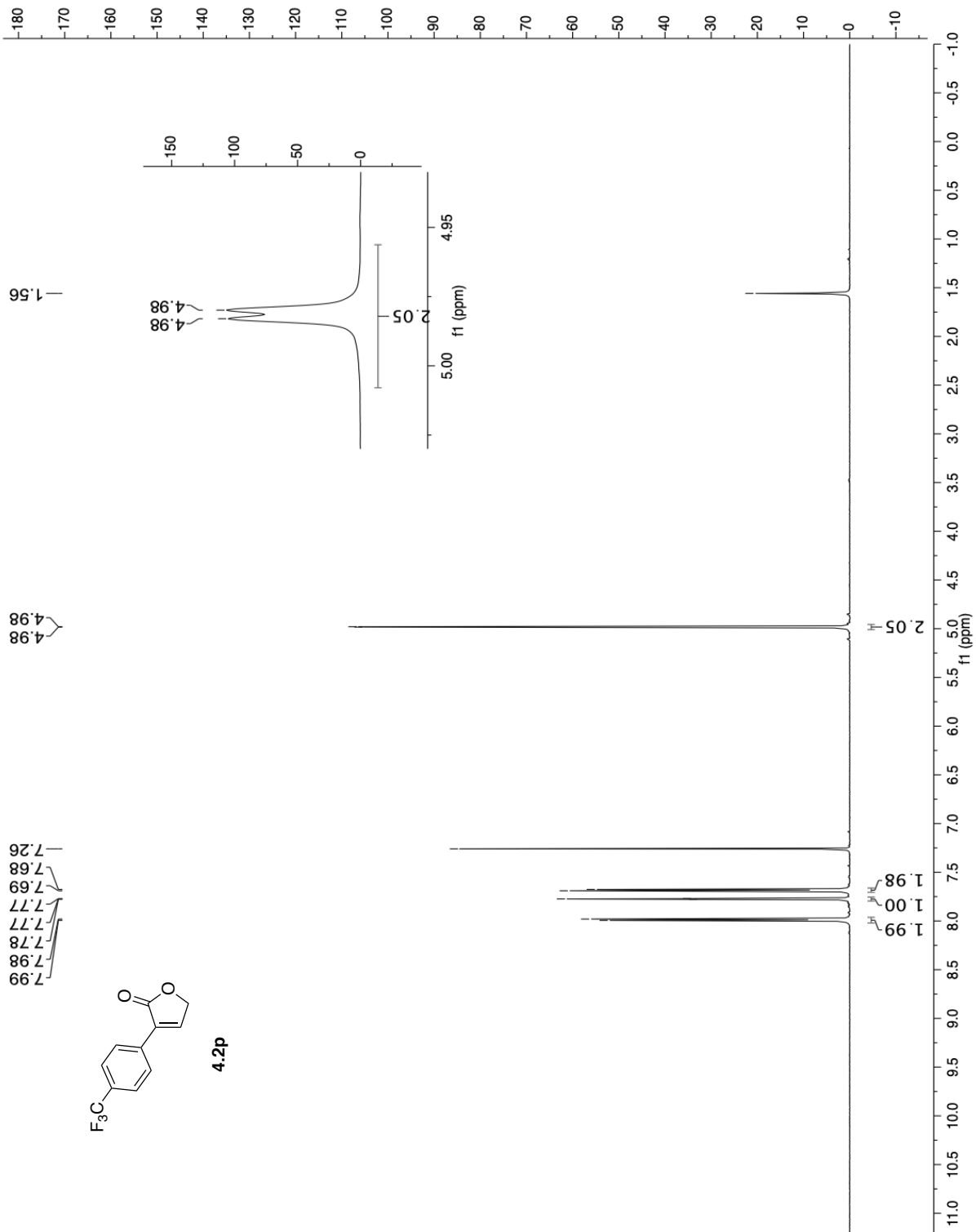


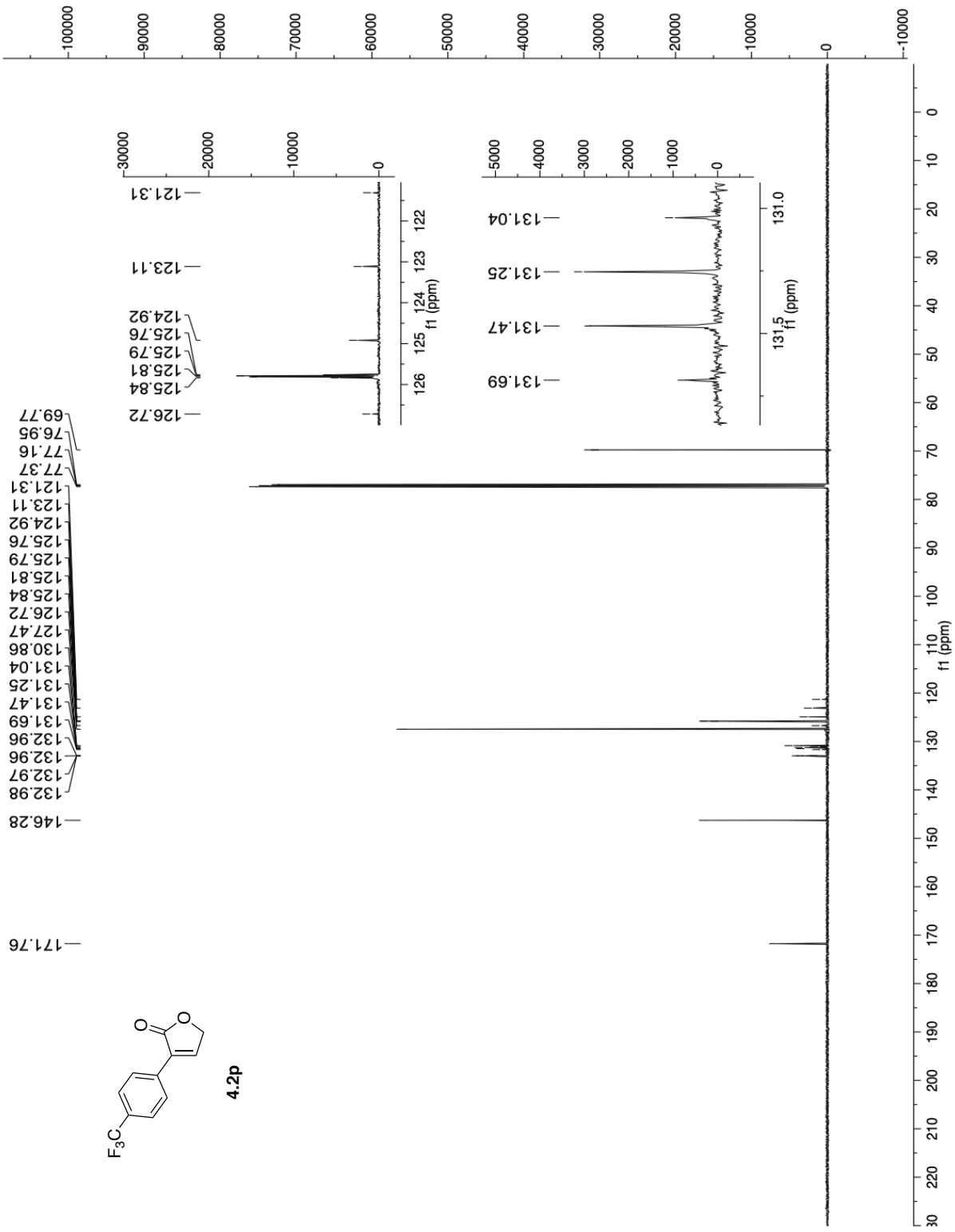


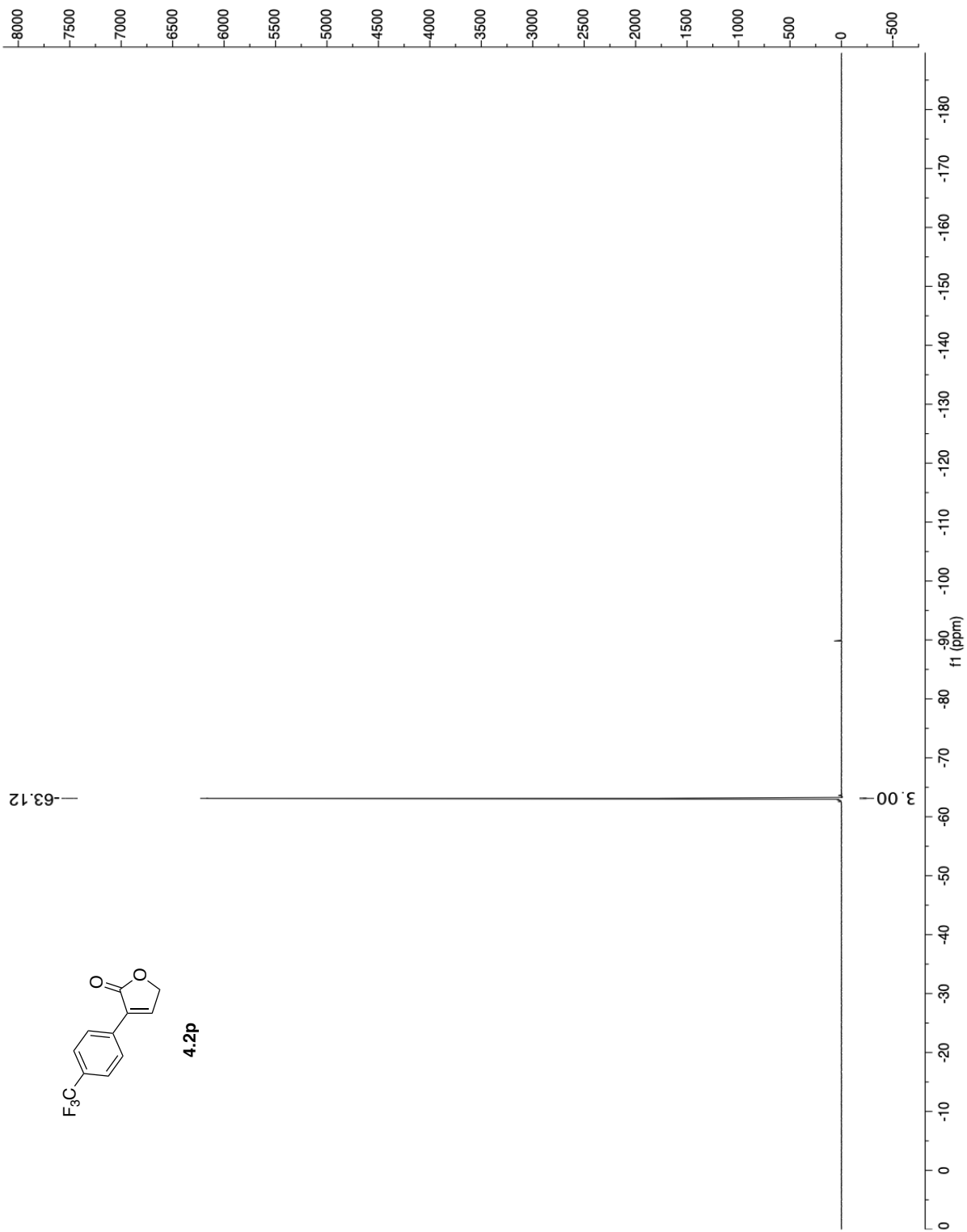


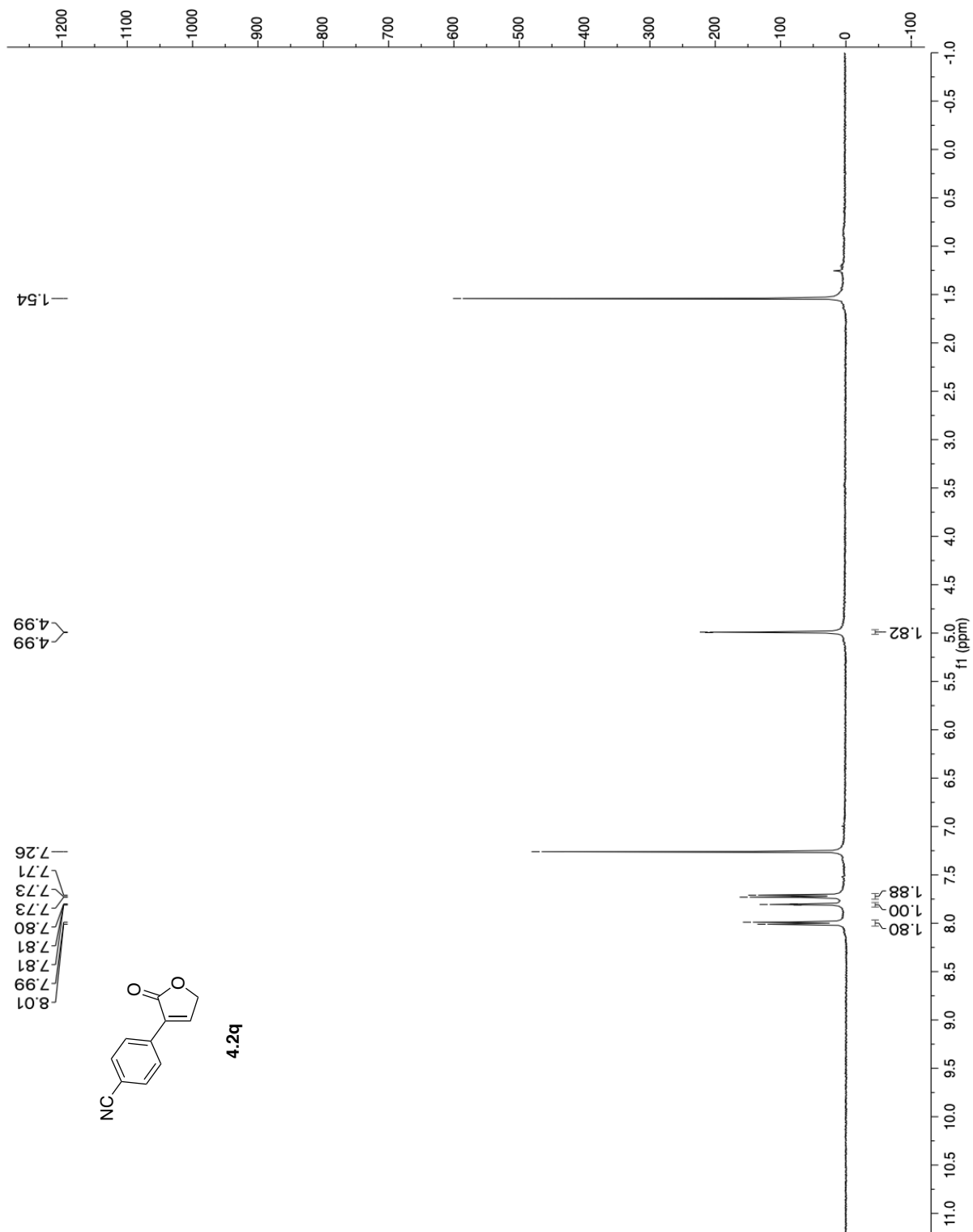


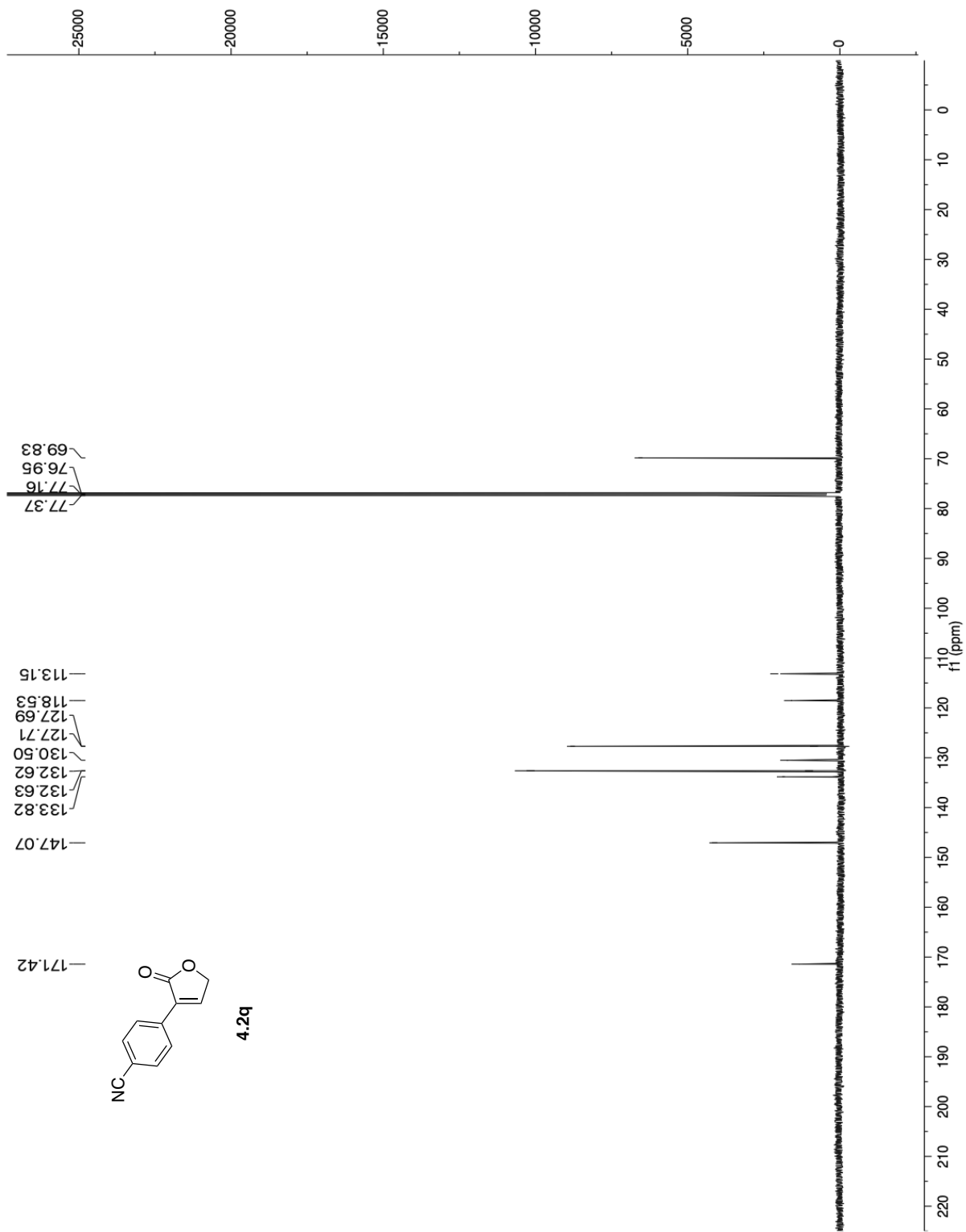


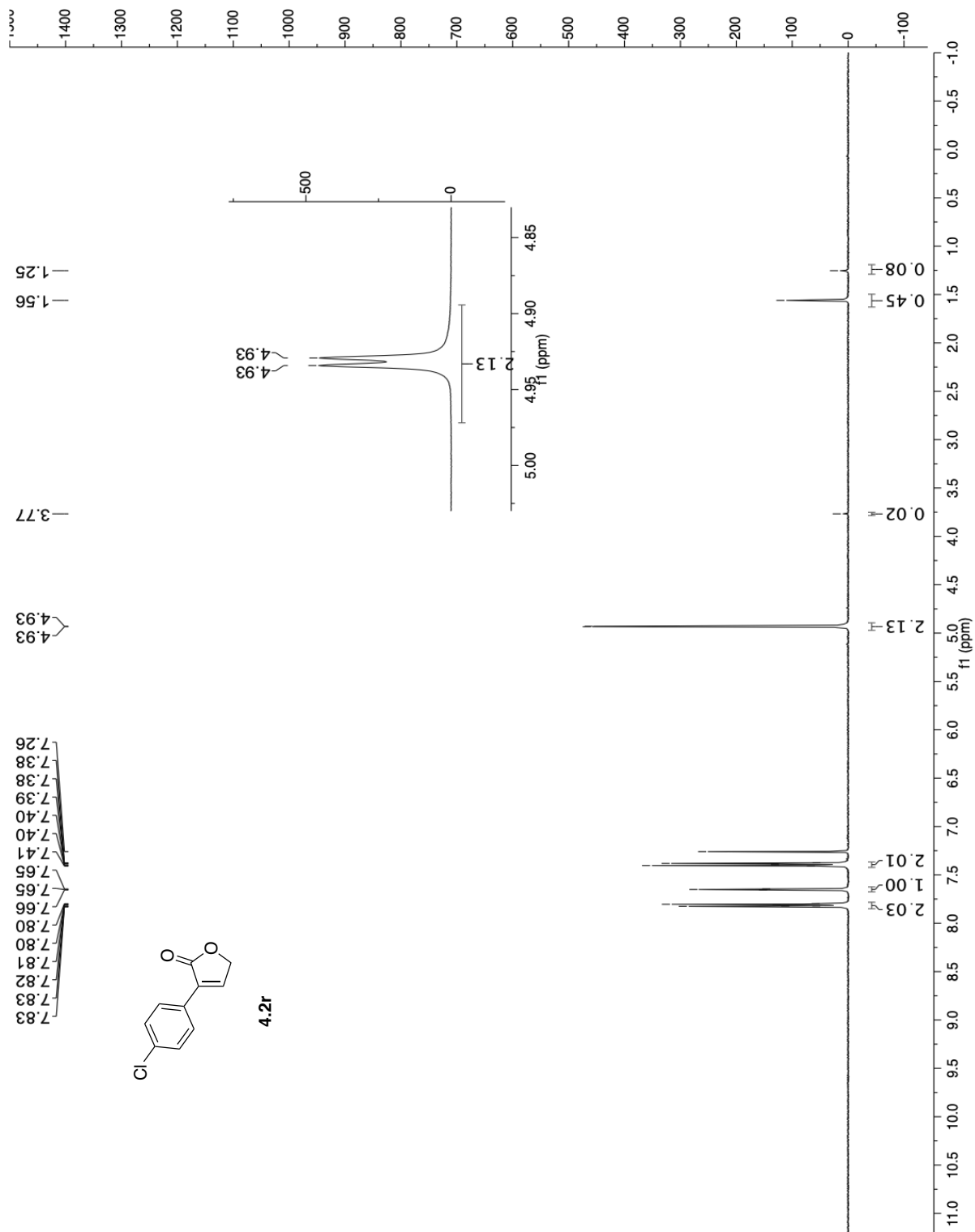


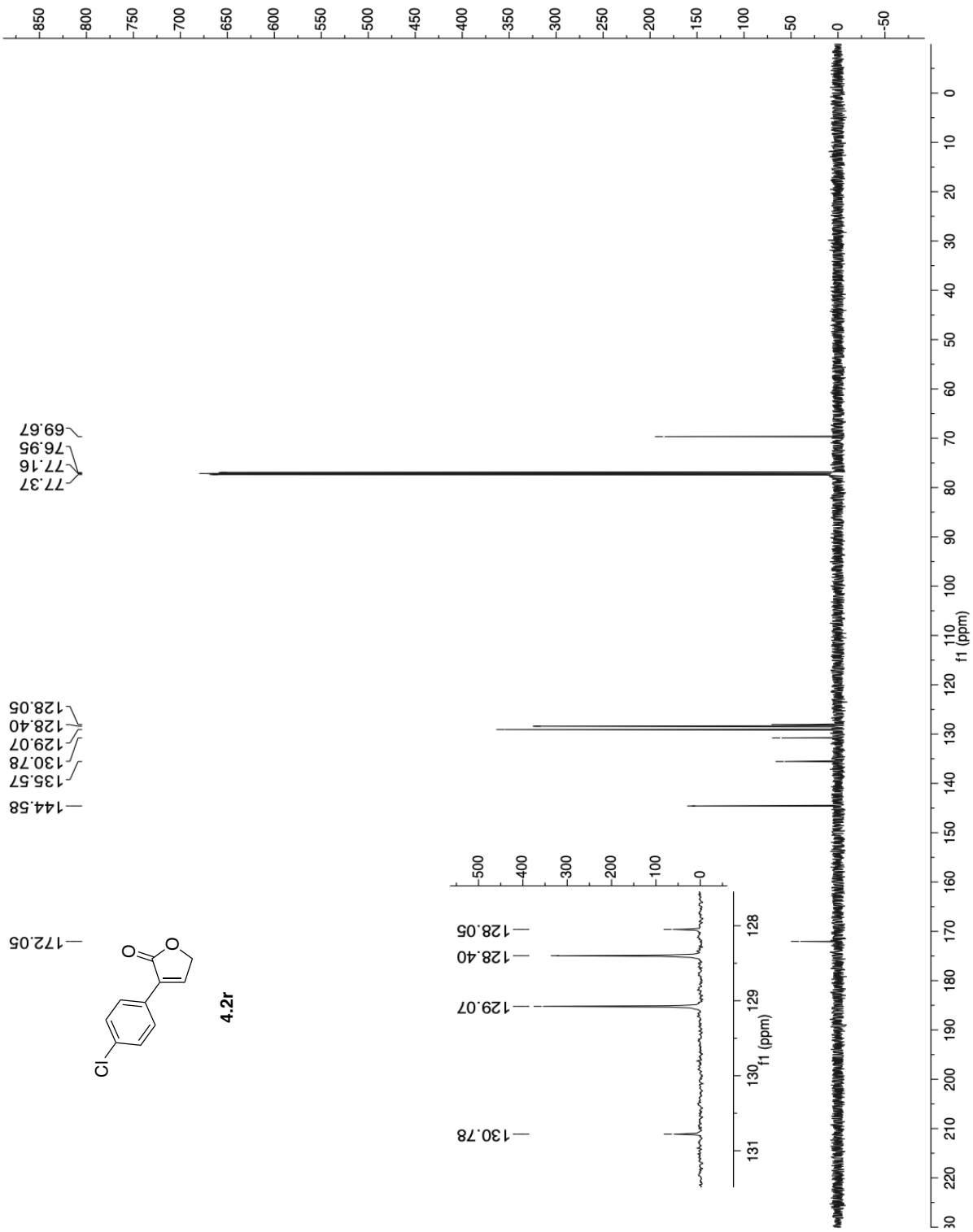


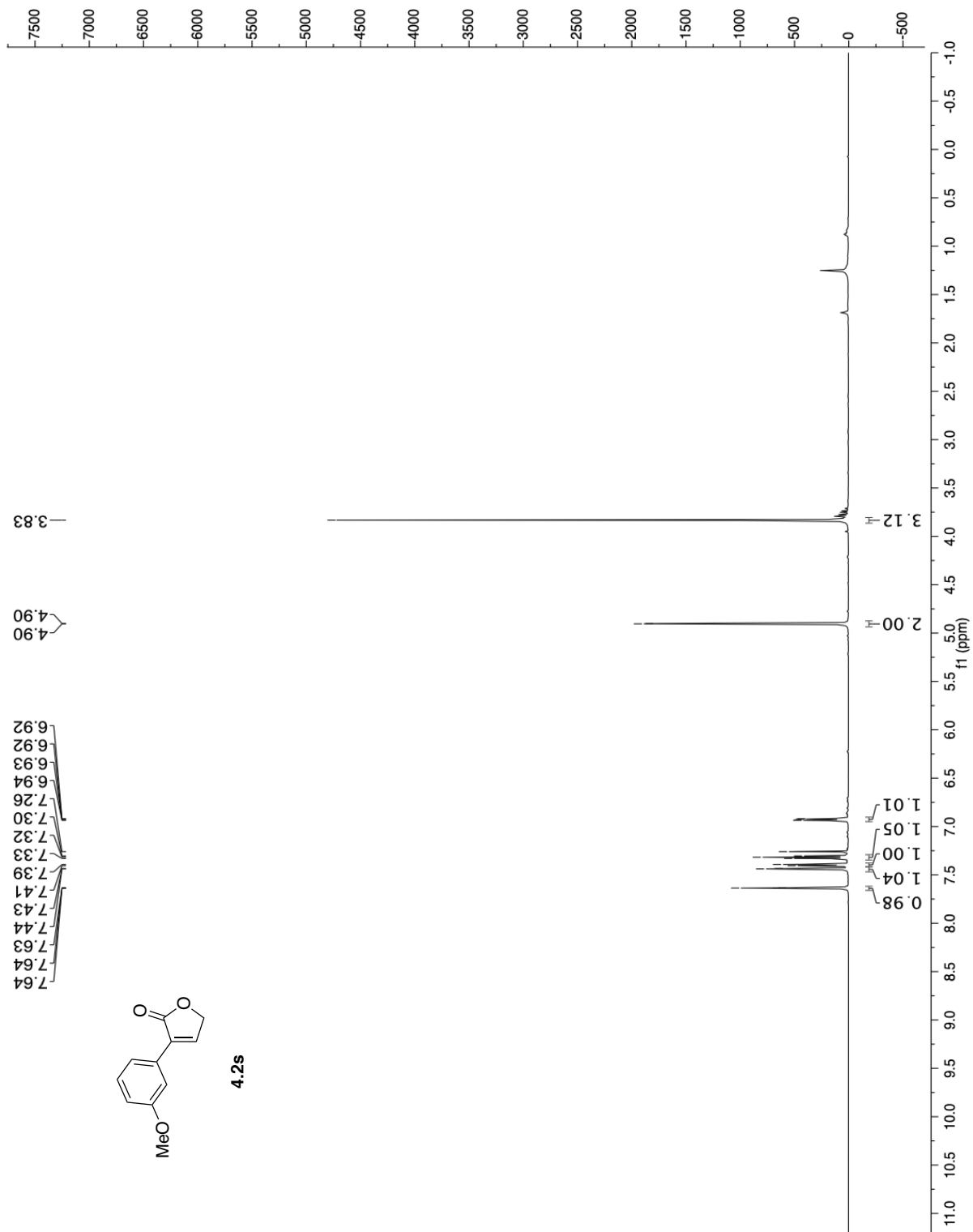


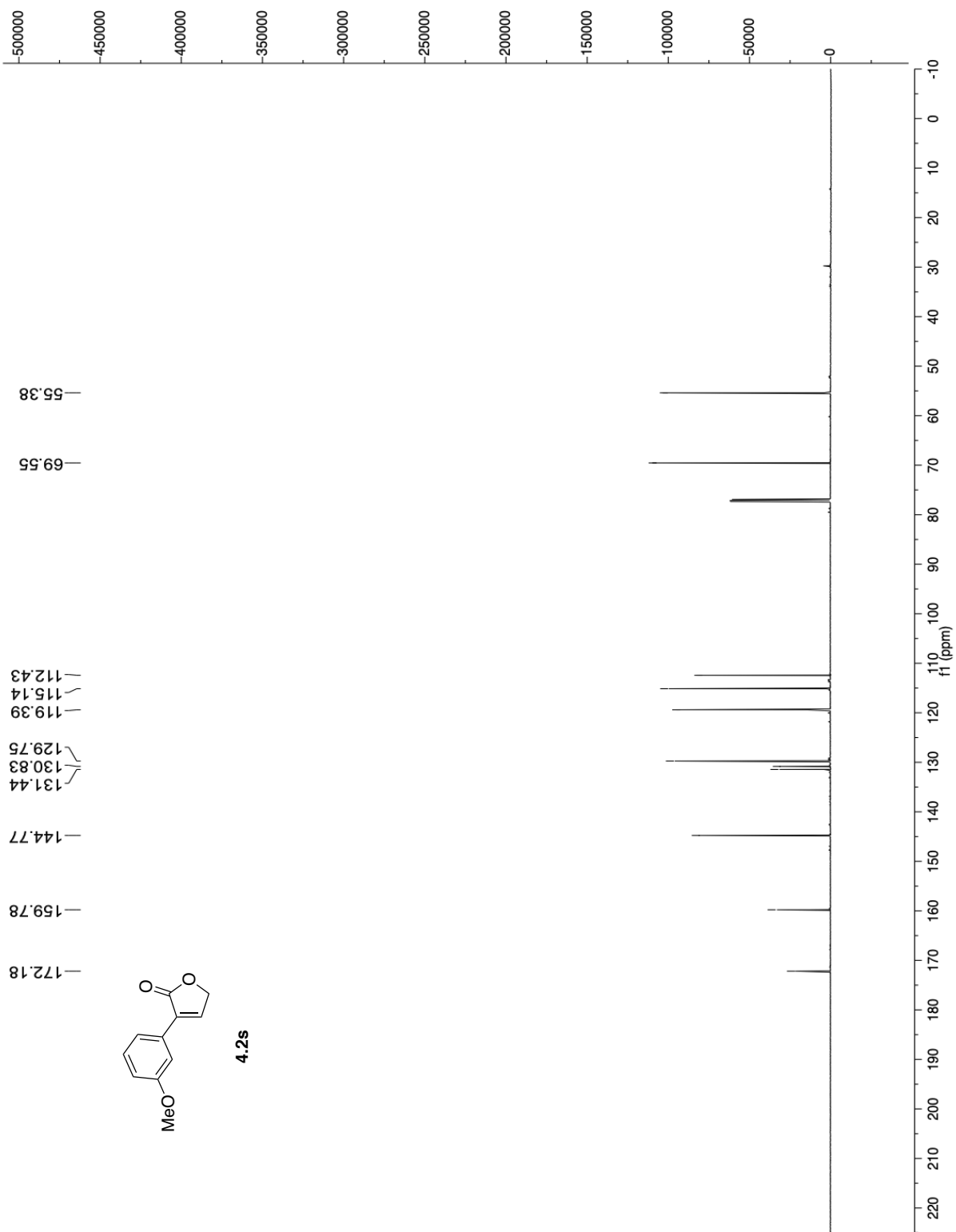


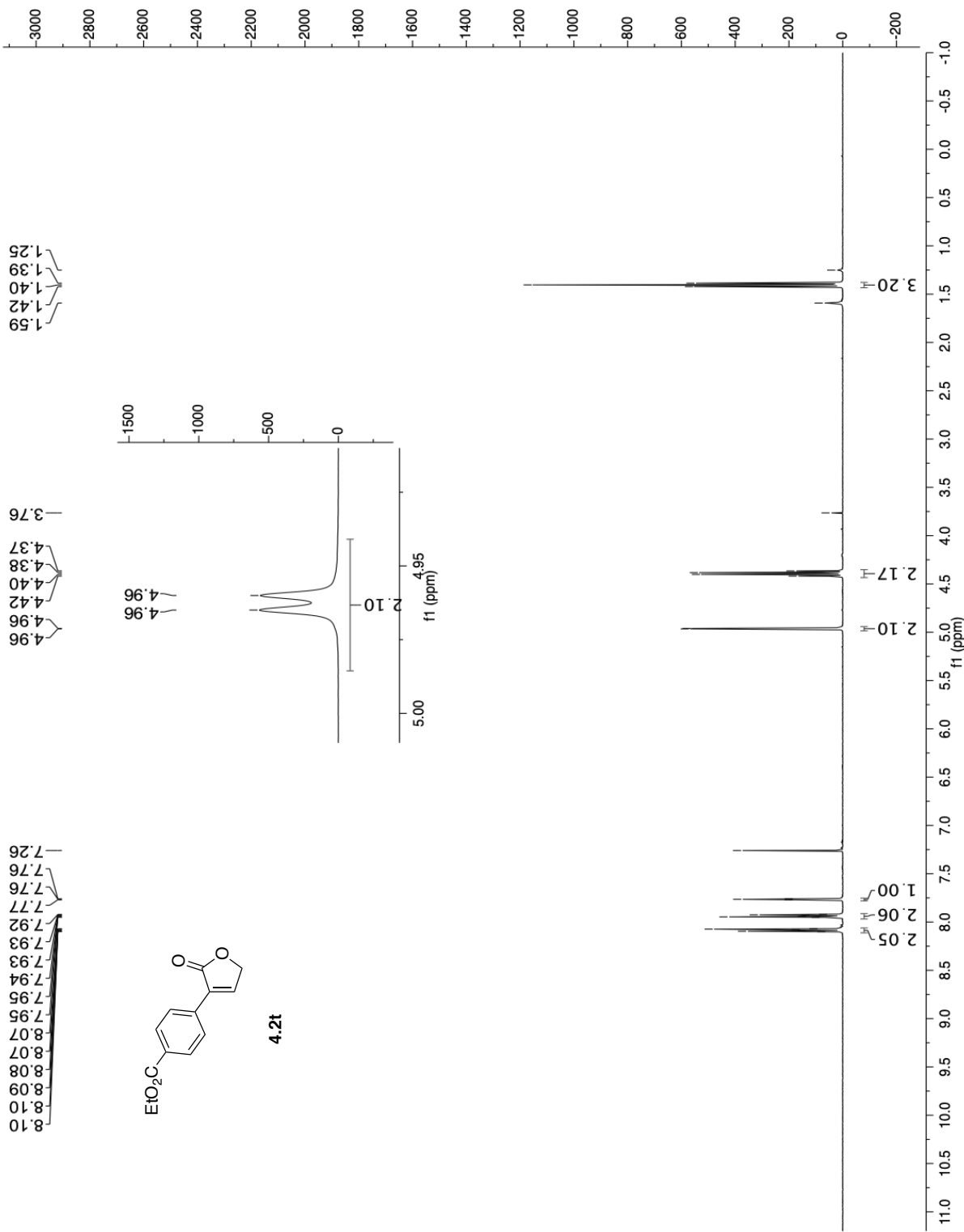


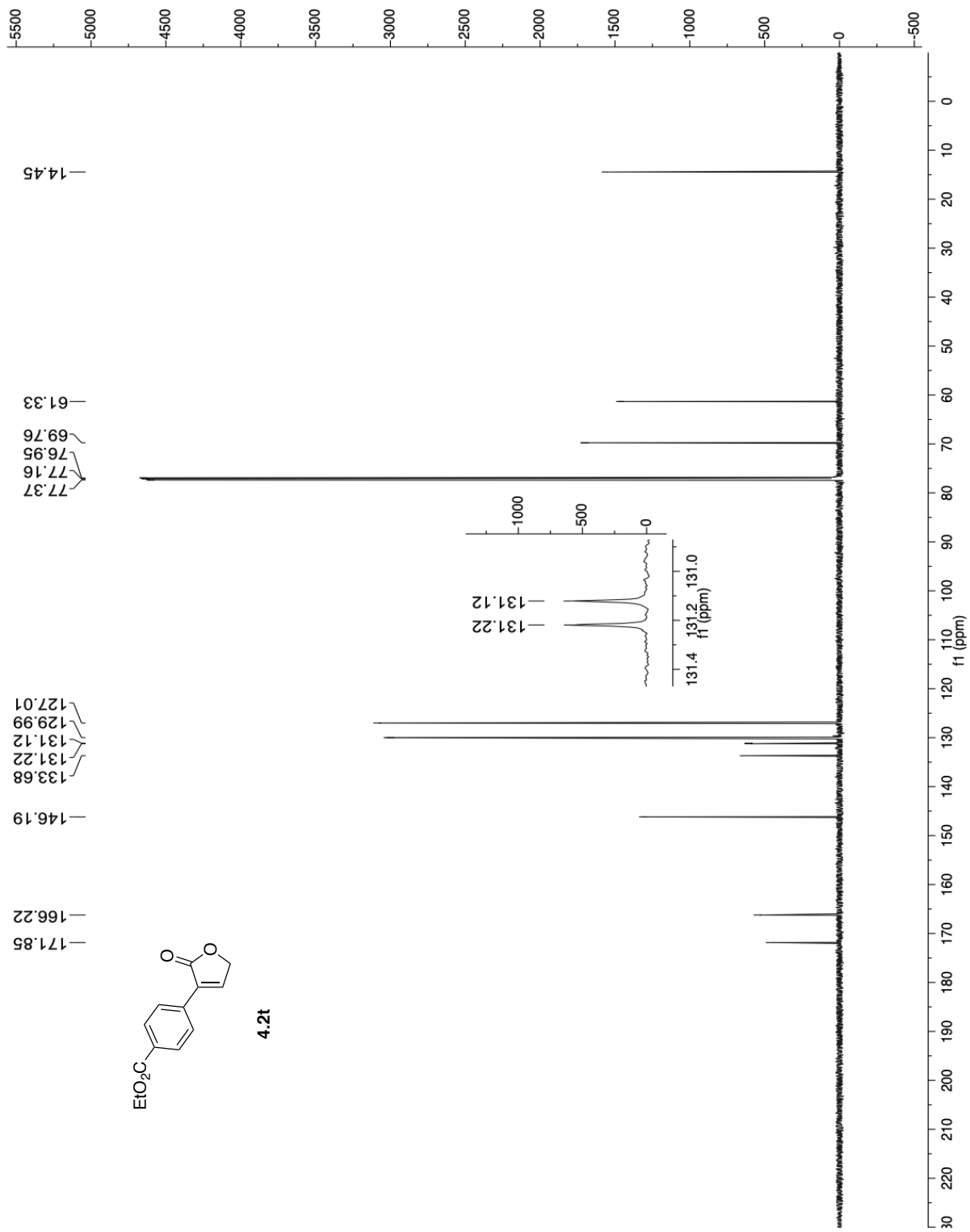


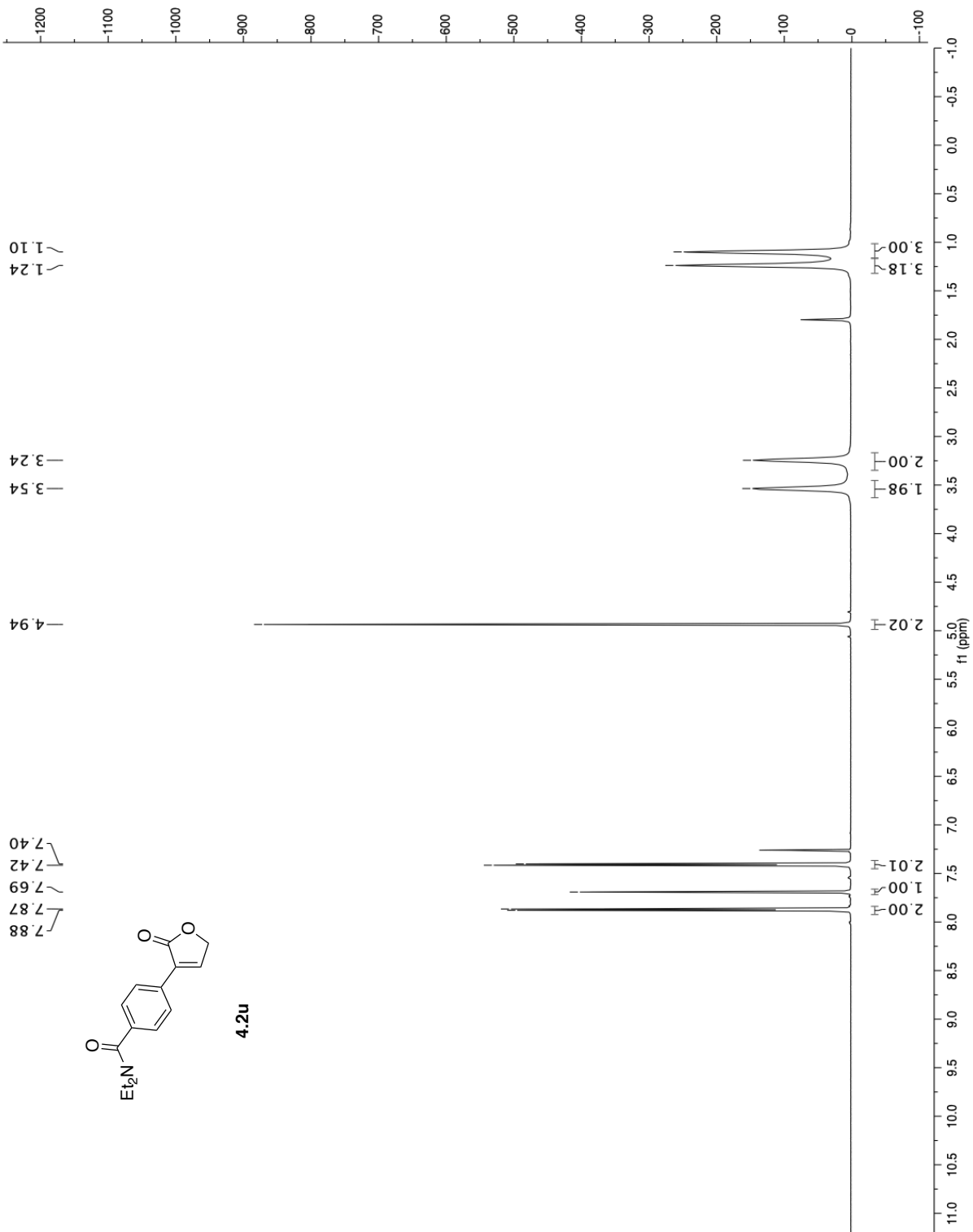


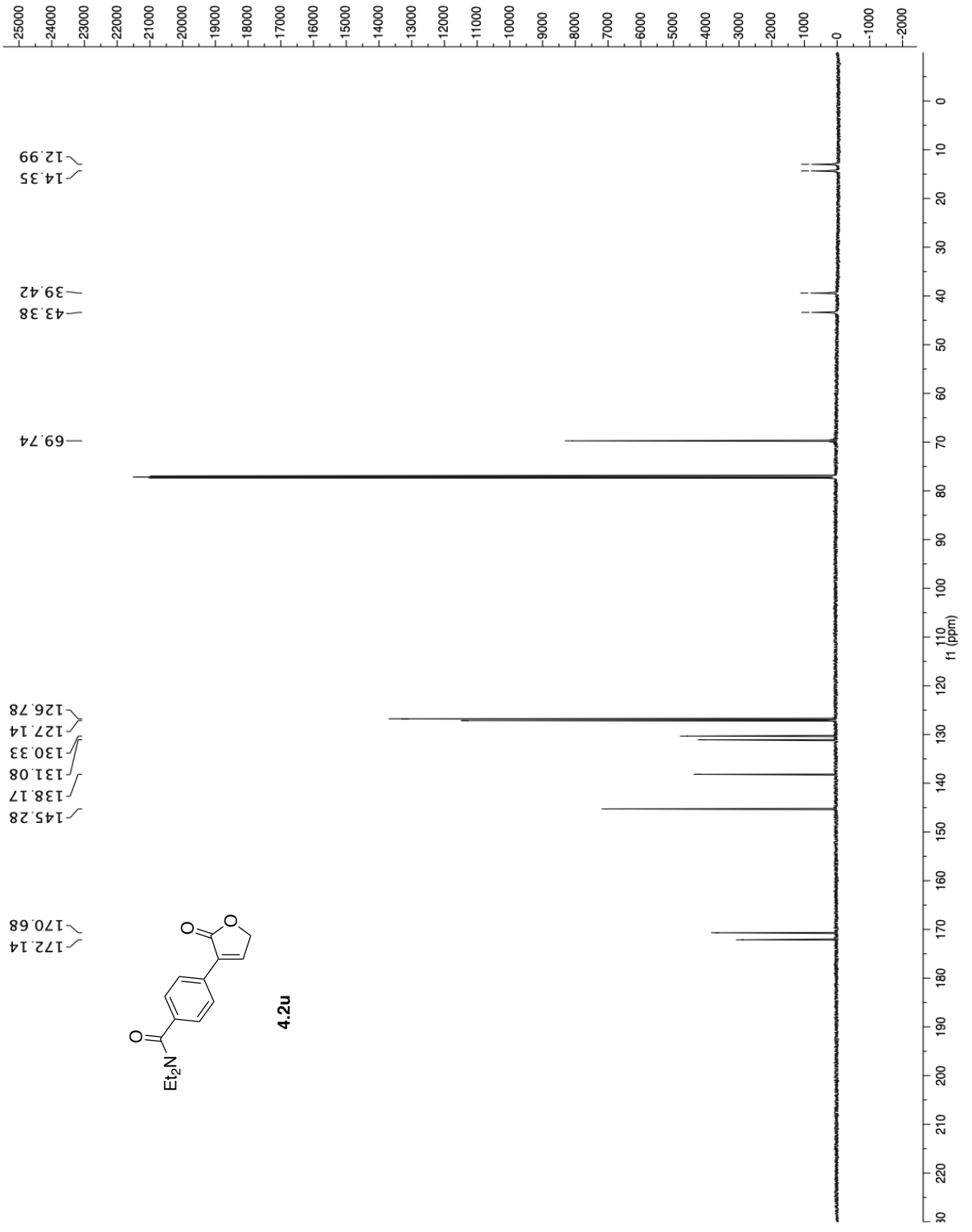


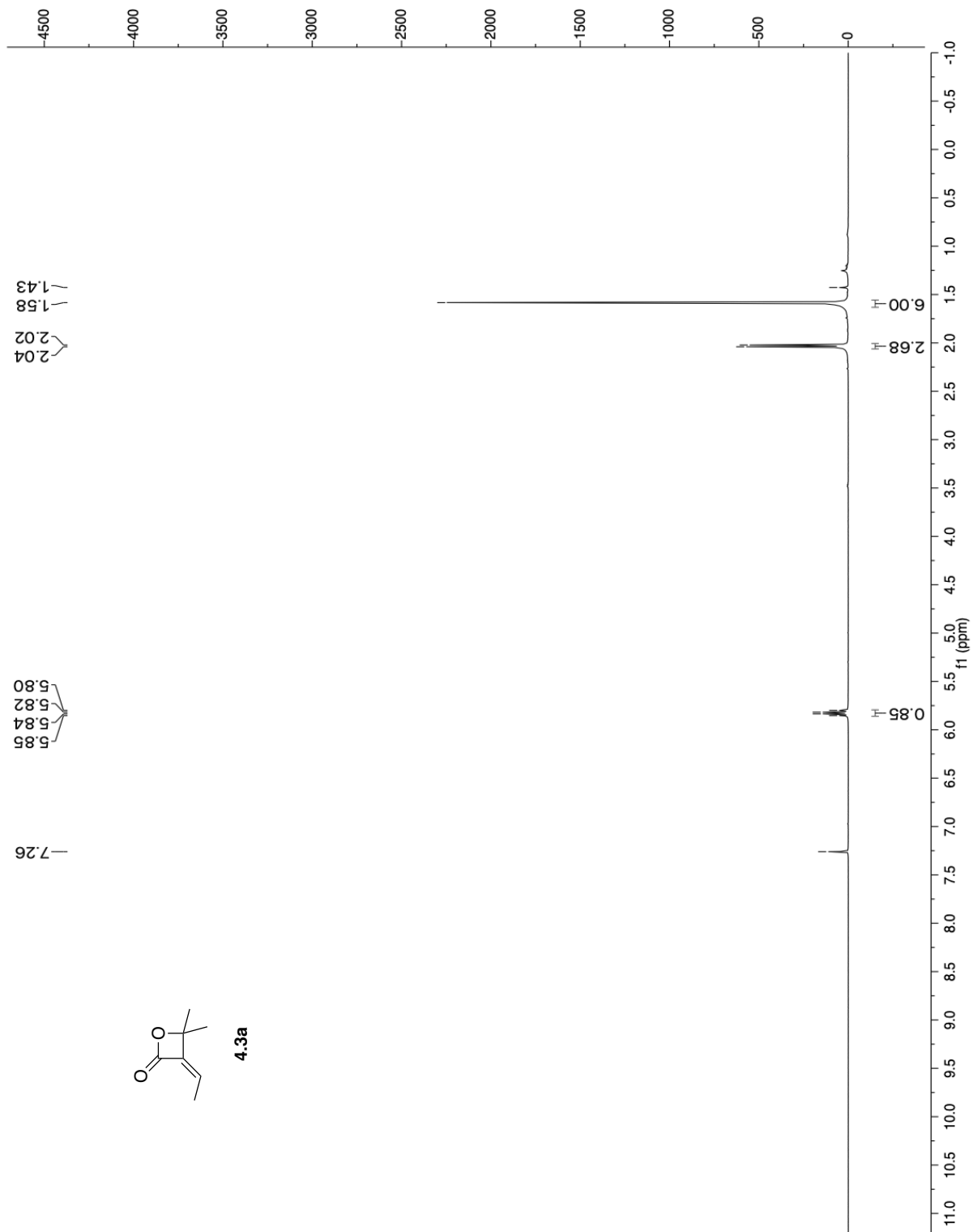


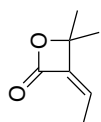
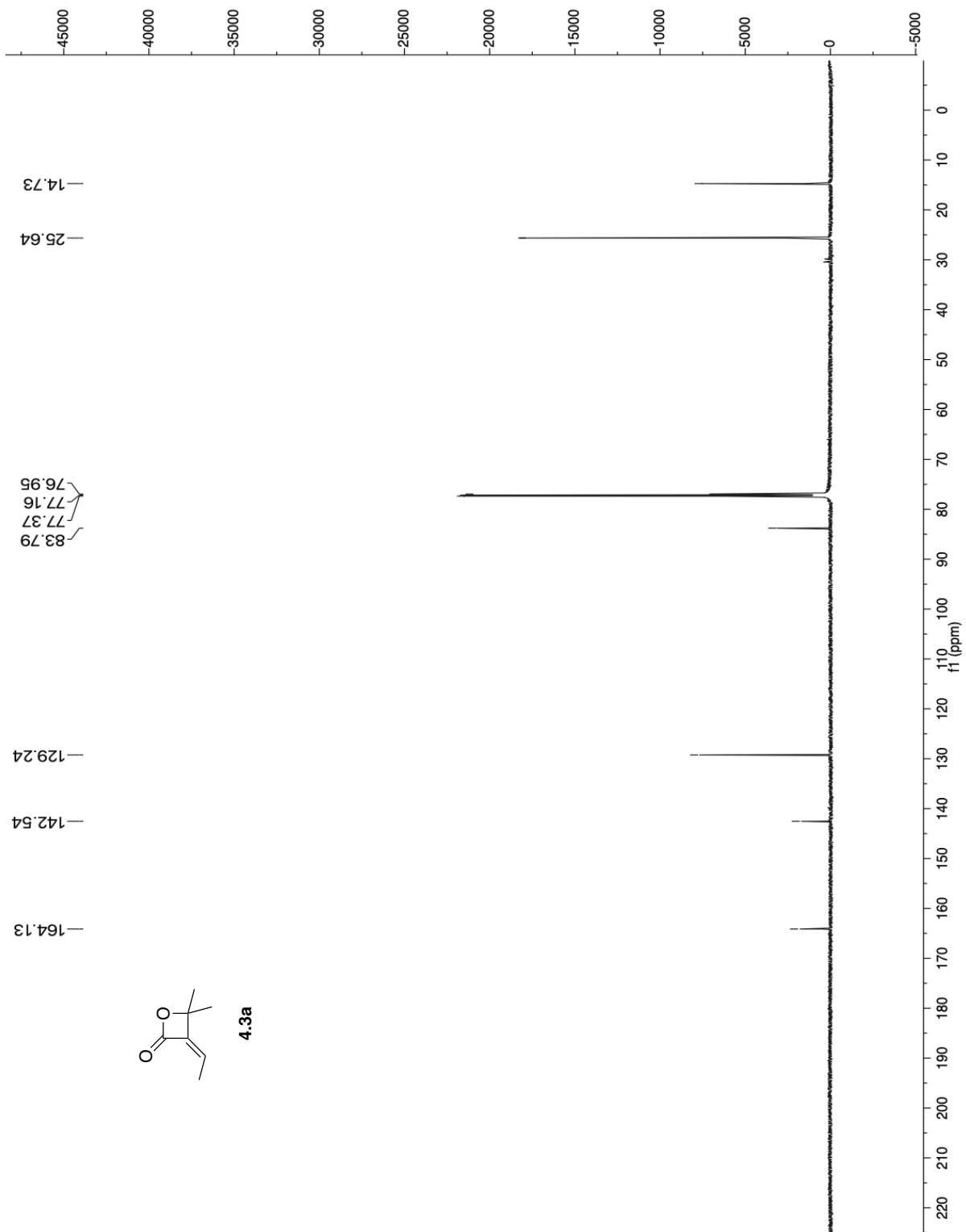












4.3a

