INVESTIGATION OF SMALL MOLECULE MEDIATED REGULATION OF PROTEASOME ACTIVITY: COMPUTATIONALLY GUIDED DESIGN, SYNTHESIS, AND FORMULATION OF A THEORETICAL FRAMEWORK FOR ACTIVITY

By

Corey L. Jones

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Chemistry – Doctor of Philosophy

2019

ABSTRACT

INVESTIGATION OF SMALL MOLECULE MEDIATED REGULATION OF PROTEASOME ACTIVITY: COMPUTATIONALLY GUIDED DESIGN, SYNTHESIS, AND FORMULATION OF A THEORETICAL FRAMEWORK FOR ACTIVITY

By

Corey L. Jones

Proteins undergo constant proteolytic degradation to regulate intracellular processes and maintain biological homeostasis. One of the main intracellular proteolytic pathways involves the proteasome, which is responsible for the degradation of misfolded, oxidatively damaged, and redundant proteins. The age-related decline of the efficiency of this enzymatic pathway leads to accumulation of aberrant proteins which leads to aggregates and a host of amyloidosis disease states. Small molecule intervention has been suggested as a viable therapeutic strategy; however, enhancement of proteasome activity is not well understood. This work details efforts and evidence for mechanistic understanding of how small molecule action may enhance and bias proteasome proteolysis and provides a theoretical model for further advancement.

Dedicated to my parents Ronald and Sharmin Jones, and my sister Haley Jones.

ACKNOWLEDGMENTS

Grad school has been an arduous journey with many high and low points. I doubt such a journey could've been completed without the help of many wonderful people in my life. First, I would like to thank my family. My parents, Ronnie and Sharmin, for their support and encouragement, and my sister Haley who called regularly just to check and talk. You all were a constant source of support on days I would've gladly burned the building down rather than try one more reaction.

I could not have asked for better lab mates to journey with. Travis Bethel and Matthew Giletto. We had many great discussions about chemistry and many more about everything else in life, and I like to imagine the halls of the 5th floor will echo with our laughter for years to come. Nor shall I forget the loud and let's say "interesting" song choices Travis would play on weekends. I would also like to thank my other group mates for being good people and not making grad school any more difficult than it had to be. Evert Njomen, the bio-lab princess, for our discussions on the bio data were very helpful. Katarina and Shafaat, thank you for having the right try hard attitude. Helping Travis train you two was a rewarding experience in its own right and I look forward to seeing what you do in the future. Grace, Taylor, Sophie and Allison, I have enjoyed my slightly more limited time with you all. Keep working and know the journey does end eventually.

Beyond the training and education, I think the best thing I have received are lifelong friends who have forged the path with me. Tanner McDaniel, Tim Shannon, Kristen Gore (now Kristen Shannon), Yu-Ling Lien, Yukari Nishizawa-Brennan, Brennan Billow, Tyler Walter, Oscar Judd, Kristen Reese, Kelly Aldrich, and Olivia Chesniak. You all have been

an excellent source of information and lively discussion as well as much needed decompression from the stresses of lab work.

I would like to thank Dr. Dan Holmes for his NMR assistance and conversation over the years. I would also like to thank Marvey Olson, for her help in dealing with the outside world and for always being a delight to interact with, I hope you are enjoying retirement. I also want to thank Dawn Khun who moved in to fill your shoes and has done an excellent job. Also, Nancy Lavirik as my quick 10-minute updates would inevitably turn to hour long discussions and your incredible work coordinating the organic classes, they don't pay you nearly enough. I also want to thank Dr. Chrysoula Vasileiou who has been a great source of conversation and advice throughout the years. Another person MSU gets to keep for too little.

I would like to thank my committee, Dr. Wulff, Dr. Huang, and Dr. Odom, for many enlightening and fruitful conversations.

Finally, I would like to thank Dr. Jetze Tepe. I would like to thank him for the opportunity to work in his laboratory and the creative freedom he has encouraged. We have had many ups and downs between us but have managed to make it to the end and I look forward to seeing where the lab goes next.

I would also commend you if you have read this far. Good job!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

Arg	Arginine
Asp	Aspartamie
ATP	Adenosine Triphophate
Boc	Tert-butoxycarbonyl
BTZ	Bortezomib
CMA	Chaperone Mediated Autophagy
CP	Core Particle
CPZ	Chlropromazine
DIPA	Diisoproyl amine
DMF	Dimethyl Formamide
EC_{50}	Concentration producing 50% of maximal response
EDCI	N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide hydrochloride
HTS	High-throughput-screen
IBX	2-Iodoxybenzoic acid
IC_{50}	Concentration inhibiting 50% of maximal response
IDP	Intrinsically Disordered Proteins
IDPR	Intrinsically Disordered Protein Region
Lys	Lysine
MD	Molecular Dynamics
Ms	Mesyl
nBuLi	n-butyl lithium

NMR	Nuclear Magnetic Resonance
OAc	Acetate
PA	Protein Activator
Ph	Phenyl
PN	Proteostasis Network
PTM	Post Translational Modificaiton
QSAR	Quanitative Struture Activity
RP	Regulatory Particle
TBDPS	Tert-butyl dimethyl silyl
THF	Tetrahydrofuran
Thr	Threonine
Ub	Ubiquitin
Ub	Ubiquitin
UIPS	Ubiquitin Independent Proteasome System
UPS	Ubiquitin Proteasome System
UV	Ultra Violet

Chapter 1. The Proteasome and Related Systems

Multiple projects detailed herein concern the proteasome, and related systems, to varying degrees. To prevent repetitive introduction, this chapter will serve as a general primer to provide perspective, while other chapters will contain a more focused introduction to needed topics.

A. Proteostasis and Aging

Maintaining fidelity among the over 10,000 proteins[1], in mammalian cells, present at any given time is a herculean task. Cells maintain this protein homeostasis (proteostasis) via the intervention of the proteostasis network (PN): a highly coordinated and intricate system which acts to rectify proteome imbalance[2]. Exact definitions of what is and is not included in the PN vary between authors; however, the consensus appears to be that the PN incorporates all machinery directly involved in the synthesis (*ca.* 279 components[3]), folding and disaggregation (chaperones ~332 components[4]), or degradation (two canonical pathways ~1388 components[5]) of proteins. Regulators of posttranslational modifications (PTMs), structural components, and other such systems essential for PN function may be considered secondary to the PN depending on the focus of the work.

Due to the vast conformational space available to polypeptides, folding is inherently error prone[6]. Unhelpfully, the protein concentration of a cell is nearly 300g/L, an extremely crowded environment which significantly increases the tendency of proteins to aggregate (compared to dilute solutions)[7]. Compounding this problem is the decreased efficiency of the PN as the system ages[8].

Aging is a complex, and ultimately disheartening, aspect of living cells. A common feature of aged cells is the accumulation of non-native protein aggregates and general loss

of proteome fidelity. Recent reviews exhaustively cover proteostasis in aging[9], dysfunction of associated pathways[8], role in cardiac health[10], neurodegenerative diseases[11-13] and many other diseases that appear to be caused by dysregulation of the PN[14].

Many reasons for its decline are given[15] but in brief, age related decline of the PN may be explained by small mutations to the PN machinery over long time periods leading to decreased efficiency. This decreased efficiency leads acute stressors, such as misfolded or damaged proteins, becoming chronic stressors. Chronic dependence on the PN leads to an increase in the synthesis of chaperones and other machinery for proteostasis[16-18]. Increased synthesis leads to more errors and more stress[19-20]. A destructive cycle ensues where decreasing capacity leads to more stress, more stress leads to more mutations, more mutations leads to a greater decrease in capacity, which ultimately overcomes the PN resulting in a dysfunctional cellular state[20].

B. Systems of Degradation

Damage to proteins is natural and very common in cells and may be induced by a wide range of external stimuli like air pollution[21-22] and UV radiation[23]. Likewise, natural cellular processes[24-25] produce oxidants and reactive species that damages cellular proteins. These aberrant proteins must be dealt with quickly and selectively to prevent aggregate formation while simultaneously maintaining the proper distribution of necessary proteins (proteostasis). To achieve this flexibility and allow a wide range of half-lives, ranging from minutes to days, the eukaryotic PN relies upon two organelles: the lysosome and the 20S proteasome.

Lysosomes are closed organelles containing a multitude of digestive enzymes and proteases[26]. The membrane boundary prevents uncontrolled destruction of cellular contents

but requires additional pathways for uptake of proteins. These pathways are collectively termed autophagy and are discussed below.

C. Autophagy Lysosome System

Autophagy, a name meaning "self-devouring", has been seen in mainstream media due to the "intermittent fasting diet" trend and has been reviewed several times[27]: a special series published in *Nature Reviews Molecular Cell Biology* in July 2018 covering the topic exhaustively is also available. In general, all forms of autophagy proceed through the same steps: sequestering proteins, transport to the lysosome, translocation into the lysosome, and finally degradation by the lysosome. Autophagy appears to be largely non-selective but operates relatively slowly[28]. Often, autophagy is discussed in the context of three main forms: macroautophagy, microautophagy, and chaperone-mediated autophagy (CMA) depending on the transportation type.

Macroautophagy is primarily attributed with the destruction of damaged cell organelles, large protein aggregates, and anything else too large for proteasomal degradation[29]. This pathway has been implicated as having key roles in immunity and inflammation[30]; particularly in stress response during times of starvation[31]. The overall process proceeds through the sequestration of doomed species by a membrane termed a phagophore. The resulting double membrane autophagosome transports its cargo and fuses with the lysosome. The autophagosome and its contents are then degraded by lysosomal enzymes.

Microautophagy is characterized by the direct and non-selective degradation of cytoplasmic components[32]. Microautophagy complements the other two processes and may be induced through starvation. The general process is the direct invagination and vesicle

scission into the lumen, mediated by autophagic tubes and is believed to be involved in the maintenance of organelle size and membrane homeostasis.

Chaperone-mediated autophagy is, by comparison, less studied but is known to be highly selective[33]. The two main chaperones for CMA have been identified as hsc70 and hsp90 which direct selected proteins to and regulate translocation through the lysosomal membrane.

Other forms of autophagy are known, such as mitophagy[34] and lipophagy[35], which is the process of lysosomal degradation of mitochondria and lipids respectively, and more are likely to be discovered. However, the overall pathway is the same: isolation and transport of proteins to the lysosome for degradation.

While this chapter has only covered the basics of autophagy (*vide supra*) and its utilization by the cell, it is an important pathway to keep in mind as autophagy works in tandem with proteasome systems (*vide infra*)[36-37].

D. The Proteasome

The importance of the proteasome is difficult to overstate as the cell has evolved to utilize its proteolytic power in several specialized tasks including cell cycle regulation, differentiation, the inflammatory response, a large role in immune function and apoptosis[38]. The proteasome comprises between 1% and 2% of a healthy cell's proteome[39] and are found within all kingdoms of life. As detailed reviews of the proteasome's architecture[40], biological assembly[41], role in health[42], and other topics are available[37, 43-44], this section will merely outline the proteasome and associated pathways.

The human proteasome core particle (CP) is a large, *ca.* 750 kDa, multicatalytic protease comprised of four stacked rings in an αββα sequence[45-46]. Each ring is in turn

comprised of seven unique subunits. The interior β subunits form a hollow cavity and house two sets of three (6 total) distinct proteolytic active sites: chymotrypsin-like (β5), trypsin-like (β2), and caspase-like (β1)[47]. Access to the interior chamber is limited by the convergence of the N-terminal tails of the outer α subunits which meet to form a gate at either end of the core particle (Figure 1.1). The gate prevents non-selective degradation of cellular contents and possesses binding domains between adjacent α-rings termed intersubunit pockets[48]. These binding domains are utilized by a variety protein activators (PA) with a specific PA utilized for specific pathways (*vide infra*). A weak equilibrium exists between the open and closed forms of the proteasome, heavily favoring the closed form[49]. However, the gate does not open wide enough for properly folded globular proteins to gain access to the catalytic sites[50-51].

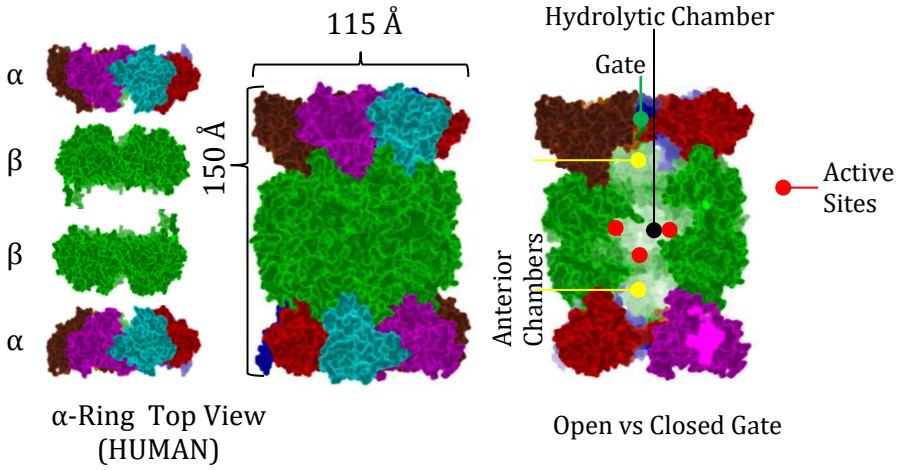


Figure 1.1: Overview of Proteasome Structure

E. The Ubiquitin-Proteasome System (UPS)

The 20S CP participates proteolytically in two distinct biological pathways distinguished by the presence or absence of ubiquitin. Ubiquitin (Ub) is a 76-residue protein unique to eukaryotic cells with high levels of conservation[52]. The addition of

ubiquitin to a substrate is termed ubiquitination or ubiquitylation. Protein modification using Ub produces a wide range of effects from tagging a protein for degradation[47] to signal transduction[53] among others[52, 54]. Ub conjugation is carried out in three steps utilizing three different enzyme types[55]. Gly76 is activated in an ATP dependent manner by a ubiquitin-activating enzyme (E1) forming a thioester linkage with the E1 cysteine. A second enzyme termed the ubiquitin-conjugating enzyme (E2) binds the Ub-E1 complex and transfers the Ub onto itself through trans(thio)esterification resulting in an E2-Ub complex. Finally, an E3 ubiquitin ligase transfers its bound target protein (most commonly through a lysine residue) to the E2 bound Ub and releases the Ub-protein complex back into the cell, ending the Ub cascade[56] (Figure 1.2).

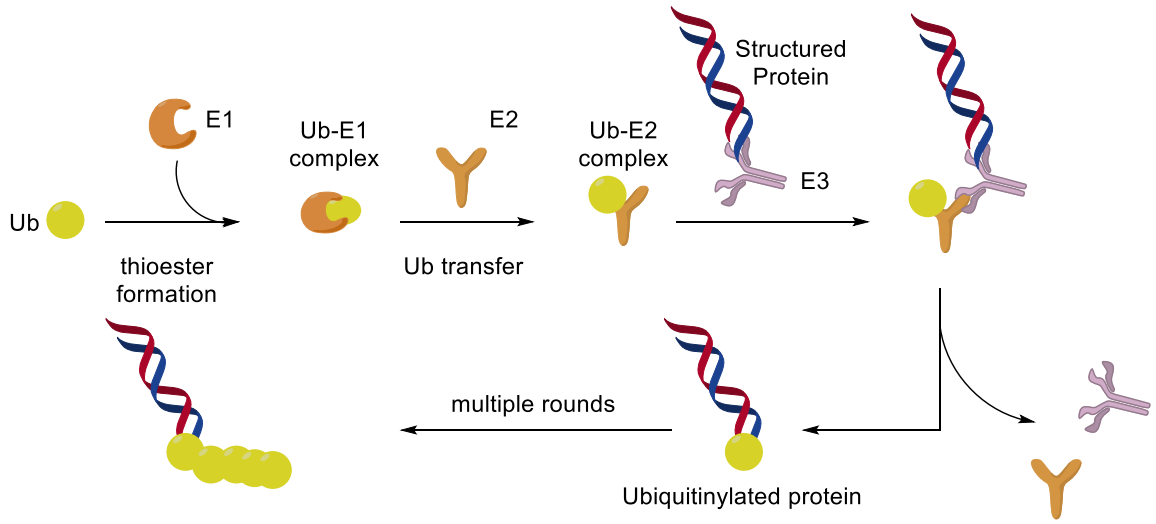


Figure 1.2: Cartoon of Ubiquitin (Ub) Cascade

The cascade is hierarchical with the two E1 enzymes[57] able to bind to multiple E2 enzymes (humans have 35)[58] and each E2 is able to bind to multiple E3 ligases (humans are estimated to have between 500-1000)[59]. This tiered system buried within a cascade allows tight regulation of Ub and ubiquitinylated systems. Originally it was thought that a multiubiquitin chain was required for protein recognition by the UPS; however, emerging

evidence suggests otherwise[60]. Likewise, a new level of regulation has been detailed in recent years illustrating the spatial arrangement and linkage specific conformations direct tagged protein outcome and is utilized by the cell for transient PTMs[60-61].

To participate in ubiquitin dependent proteolysis, the 20S CP requires endogenous activation[62]. This is achieved via the non-covalent addition of activator caps to special pockets formed between adjacent alpha subunits termed intersubunit pockets (see Section F). The most common activator is the 19S regulatory particle (RP) which is comprised of 19 subunits and split into two main parts (Figure 1.3): a base and a lid.

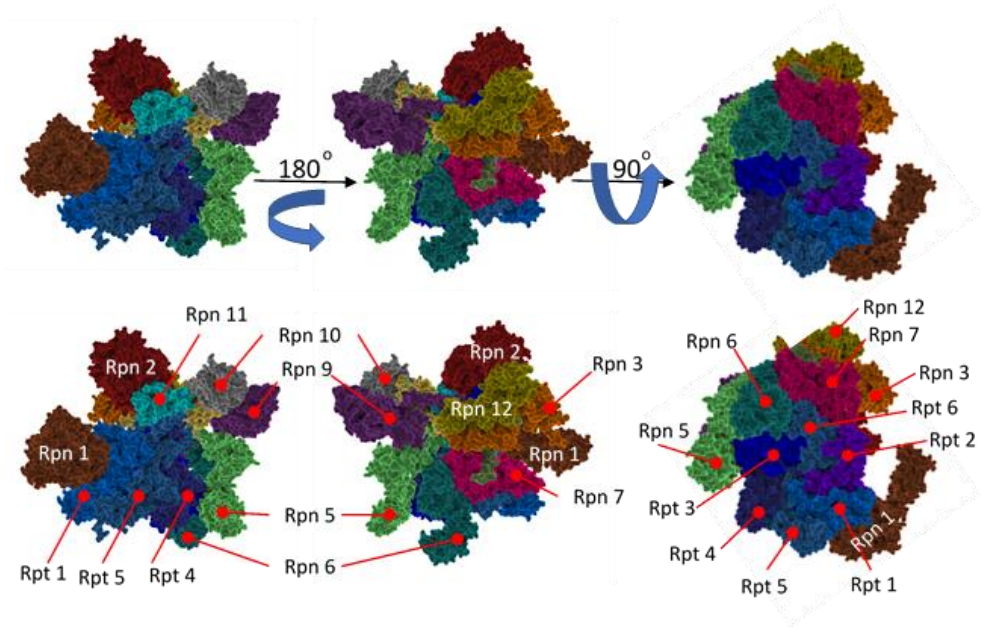


Figure1.3: Depiction of the 19S Subunits with Base (shades of blue) and Lid (multicolored)

The RP base binds directly to the CP while the lid extends outward. Once a ubiquitinylated protein is bound by the RP's lid many things may occur concomitantly. Deubiquitylating enzymes recycle the tag back to the cell while ATP hydrolysis, occurring on the lid, drives the unfolding and translocation of bound substrate into the proteolytic chamber of the CP. This process is regulated by the motor units that make up the base. Six

distinct subunits (Rpt1-6) from the AAA (ATPases associated with diverse cellular activities') family of ATPases regulate lid mediated engagement and unfolding[39, 63]. This process is conformationally complex as the holoenzyme has been reported to exits in 19 distinct states[50, 64-66]. Because of this complexity, the exact mechanism of how endogenous ligands open the proteasome gate is still poorly understood (See Section F).

Additionally, one or two 19S RPs may bind to a single 20S CP. The "most common" form encountered is a doubly bound CP with two identical 19S RPs generating the 26S proteasome. The 26S proteasome has been attributed with at least 80% of protein degradation in growing mammalian cells and widely assumed to automatically degrade any ubiquitinylated protein it encounters[67]; however, recent evidence suggests the 26S is tightly regulated and may not automatically degrade any bound protein[43]. Other endogenous caps exist, and the proteasome has been observed to exist as hybrid with two different caps[41, 63]. In any event, the coordinated action of the 26S proteasome and ubiquitin is widely acknowledged as the main proteolytic pathway for selective degradation of misfolded or redundant *structured* proteins[68].

F. Overview of Gating Mechanism

The proteasome gate is formed by the convergence of the N-termini of the seven alpha subunits (Figure 1.4).

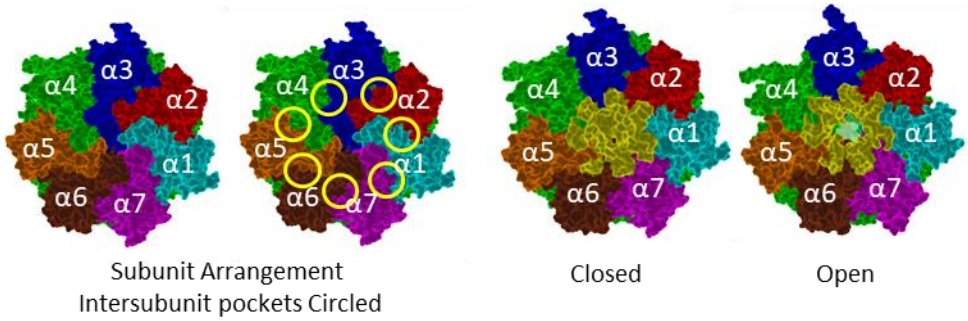


Figure 1.4: Top View of Proteasome CP Depicting the Gate (Yellow) and Intersubunit Pockets (circles)

Pockets formed between two adjacent subunits are used by endogenous protein activators (PA), like the 19S, for binding and gate opening[69]. PAs utilize a conserved three peptide hydrophobic-tyrosine-variable (HbYX) recognition unit, found at the C-terminus of a number of proteasome binding partners including assembly factors[70-71], to bind within these intersubunit pockets.

One of the first studies demonstrating the importance of the HbYX motif was done by Smith et. al. using PAN and archaeal 20S[71]. Systematic mutational studies demonstrated that no AA substitution was tolerated at the penultimate tyrosine. Additionally, they discovered that lys66 within the α-subunit was needed for PA-20S association. Based on this information the authors concluded that the penultimate tyrosine and a preceding hydrophobic residue were essential for gating, but the terminal AA was only required for PA association and played no role in inducing substrate hydrolysis. Forster et al[72] demonstrated gate opening with a different PA, the 11S, which lacks the HbYX motif. In that complex, C-terminal residues form main-chain to main-chain hydrogen bonding and a salt bridge between the C-terminal carboxylate and Lys 66. Interaction with the Pro17 reverse turn on the proteasome with an "activation loop" PA26 induces gate opening by small (0.5-3.5 Å) movements of each subunit. Four conserved residues were identified as crucial to binding and stabilization of the open form: Tyr8, Asp9, Pro17, and Tyr26. Severe reduction in model substrate degradation was observed in mutant archaeal proteasomes lacking any of these residues[72].

Rabl et al discovered a similar mechanism at play with the proteasomal ATPases in PAN (yeast version of the 19S)[73]. Using x-ray crystallography, Rabel et al illustrated that

PAN also induced the same radial and lateral displacement of the alpha ring reverse loop without contacting Pro17 and instead attributed the displacement to HbYX binding. They proposed association of PAN with the 20S through HbYX through contacts with Gly34, Lys66, and Leu81 induces the rotation of Pro17 away from the central pore and causes stabilization of the open-gate conformation. The importance of Lue81 was confirmed by mutagenesis[72]. Corroborated by several studies on peptide mimics possessing HbYX tails[71, 73], it was believed that the binding of the HbYX motif was sufficient to induce gate opening. However, recent Cryo-EM studies have revealed stably bound eukaryotic 26S structures with a *closed* gate[74].

In a more recent publication[75], *substrate engaged* human 26S has been described in 7 different states: termed E_{A1}, E_{A2}, E_B, E_{C1}, E_{C2}, E_{D1}, and E_{D2}. As the 26S progresses from E_A to E_D, the 19S is marked by major conformational changes as the Rpt units (Figure 1.3) insert more tails into the alpha rings (Figure 1.4). Intriguingly, the CP remains largely unchanged until the final gate opened form E_D.

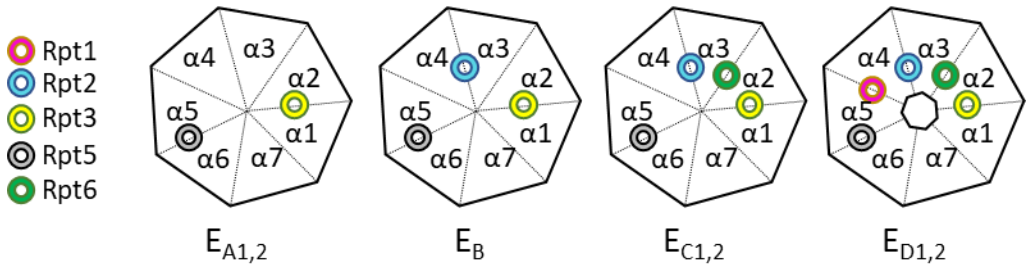


Figure 1.4: Rpt Tails' Insertion Points Through Each 26S Conformer[75]

Unfortunately, the authors did not comment extensively on what factors may be contributing to the opened form. Yet, their publication[75] does illustrate the mechanism of gate opening is far more complex than previously thought as 19S binding alone is

insufficient to induce an open-gate proteasome, a stark-contrast to previous reports (*vide supra*).

G. The Ubiquitin Independent Proteasome System (UIPS)

As mentioned above, the UPS is an ATP-dependent pathway; however, ATP is only required for driving the unfolding of structured proteins. Once a protein has entered into the proteolytic chamber, it is degraded passively. To pass unaided though the gate, a protein must already be unfolded. Two types of proteins fit this criterion: intrinsically disordered proteins (IDPs) and oxidatively damaged proteins (detail in the next section).

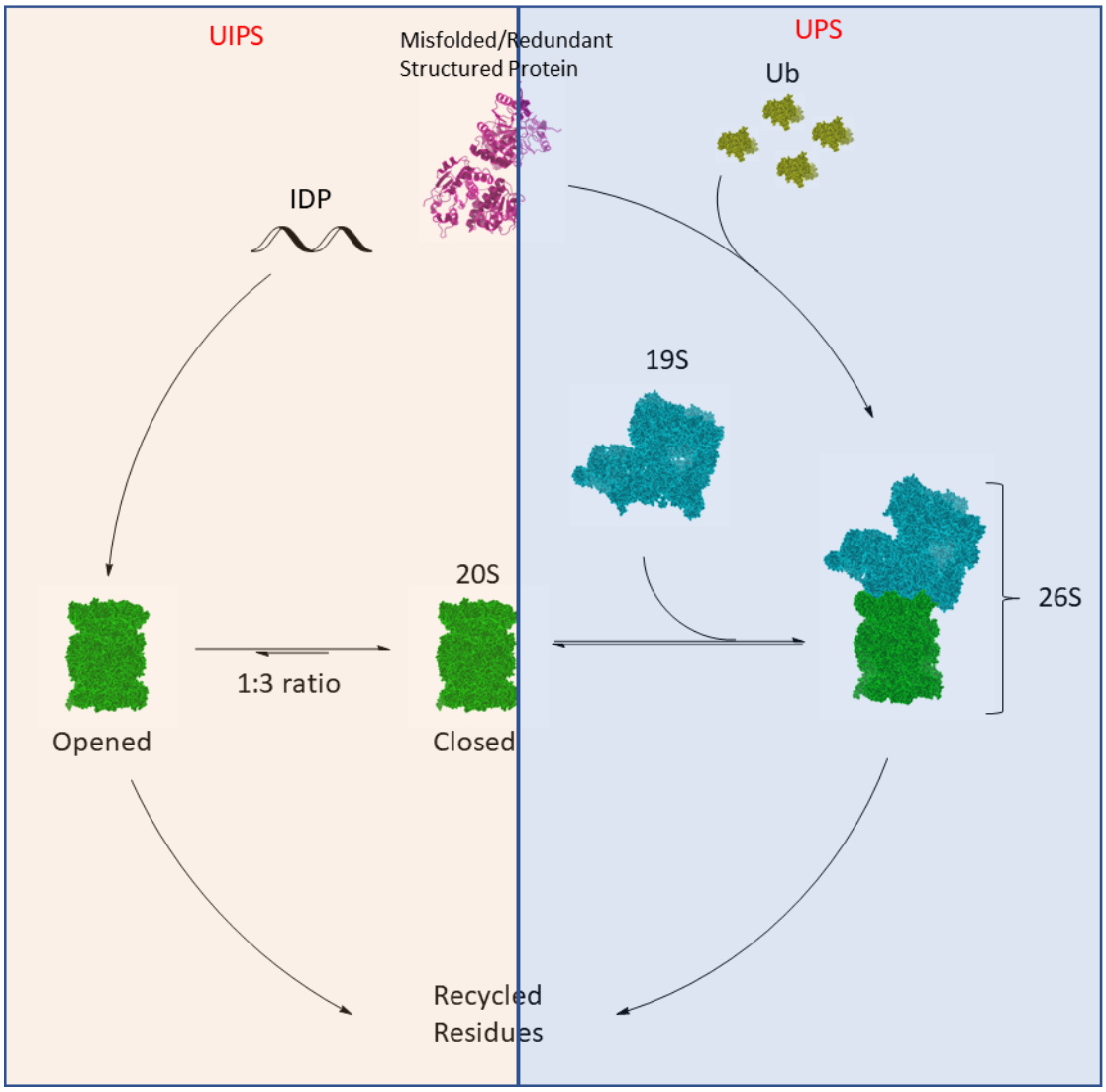


Figure 1.5: Simplified Outline of UPS vs UIPS

While IDPs and oxidatively damaged proteins are unstructured enough to pass through an open gate 20S CP, the open and closed gate exist in an equilibrium heavily favoring the closed form (Figure 1.5)[76]. Like the UPS pathway, the UIPS also contains specific PAs that bind and open the gate. The two main PAs for this pathway are the 11S (REG/PA28) and PA200. Both of these PAs bind in a similar manner to the 19S (i.e.

utilizing the intersubunit pockets) yet form complexes with open ends allowing suitable substrates access to the proteolytic chamber[69].

H. Intrinsically Disordered Proteins

Understanding IDPs as a broad class of functionally important proteins began in earnest in the mid-1990s with a bioinformatics study[77] on the emerging complete genome sequences. Analysis revealed that disordered regions were actually common in eukaryotic proteins[77-78], with some estimates proclaiming that 44% of human protein-coding genes have disordered segments of at least 30 residues[79]. Today it is known that IDPs play indispensable roles in numerous cellular processes like signaling, transcription /translation, and cell cycle progression[77-78, 80-82].

The recent awareness in IDP research publications would suggest these complexes are a recent discovery; yet, the reality is that IDPs have been reported periodically over the past 80 years[77, 83-84].

A plausible explanation of this confusion, proposed by Uversky[85], is a traditional lack of a unifying terminology. IDPs have been previously described as floppy, pliable, flexible, partially folded, natively denatured, and many more[86]. It has also been suggested that the term IDP is not ideal; however, it is the currently recognized umbrella term found in literature. Additionally, a special issue of *Chem Rev.* covers classification comprehensively[87]. The accepted definitions, derived from common use in the field, may be listed as[87-89]:

- Any functional protein or protein domain possessed of a unique 3D structure described by minimal fluctuation around their equilibrium Ramachandran angles is termed a "structured protein".

- Any functional protein or protein domain that exists as a dynamic ensemble lacking specific equilibrium Ramachandran angles with backbone atomic positions that naturally undertake non-cooperative conformational changes is termed an "intrinsically disordered protein or region" (IDP or IDPR).

IDP/IDPRs are characterized by low sequence complexity and biased amino acid composition (preference for highly charged and hydrophilic residues)[89]. The lack of hydrophobic/bulky residues results in a relatively flat energy surface and existence as a structural "ensemble" of interconverting conformational states[90]. Disordered regions leverage their high conformational freedom to maximize potential binding partners. Specificity is generally achieved per partner by multiple low affinity contact points, and the entropic loss keeps the binding interaction transient. As the disordered domain gets larger, the increase in surface area of binding begins to compensate for the loss of freedom and binding time increases in permanence. The main functional features described above exist on a continuum, though may be mutually exclusive. As such, a single protein can be comprised of multiple disordered regions that belong to different functional classes, offering immense conformational variability and adaptability. This is exploited by the cell to facilitate regulation through varied PTMs, and recruitment/localization of different binding partners. Recent reviews on IDP function[78, 86, 91-92], role in cellular signaling[80, 93-95], advantages in protein-protein interactions[82], regulation and disease[95-99], empirical studies using single molecule methods[100], NMR spectroscopy[101], methods of characterizing conformational ensembles[102], their identification[103] and specifics of IDP network interactions in great detail[104] are available and will not be discussed here.

Chapter 2. *In Silico* Investigation of Proteasome Regulators: Development of Theory of Action and Applications

A. Introduction

Previous work in the Tepe Lab detailed the diversity-oriented synthesis of imidazoline scaffolds as potent proteasome inhibitors[105]. However, Dr. Theresa Lansdell discovered, in a non-standard assay, that one of the more potent compounds could exhibit enhancement of proteasome activity (i.e. proteasome agonism). I found this dual ability to be both a proteasome inhibitor and activator, regardless of assay conditions, to be highly intriguing. I decided to investigate further with the aim of offering a plausible explanation of this odd behavior. For this, I turned to *in silico* methods, specifically molecular docking.

Molecular docking, hereafter called docking, generally refers to the computational effort of predicting the "best" intermolecular complex that may be formed between two species. Most often, the two species are a protein-ligand pair. A bit more formally, the docking "problem" may be stated as: Given the atomic coordinates of two species, predict their "correct" bound association. In principle, only the structural information of the two species is required for docking. In practice however, docking is often complemented with biological or other empirical information to aid refinement. One of the first practical suggestions for docking was posited by Crick in the early 1950s[106]. He proposed that complementarity in helical coiled coils could be modeled as knobs fitting holes. Yet, the first program written to represent a protein surface wasn't published until the late 1970s[107]. Connolly's development of a method for analytically calculating a smooth 3D contour about a molecule[108] was critical for the development of docking algorithms. The mid-1980s would see the field begin to flourish with advent of the first docking program by Kuntz and

coworkers called DOCK[109]. Today, computer-aided drug discovery/design (CADD) plays an increasingly central role in the quest for small molecule therapeutics[110]. CADD encompasses a number of powerful methods to aid hit-to-lead campaigns including molecular dynamics (MD), protein-ligand and protein-protein docking, homology modeling, quantitative structure-activity (QSAR), and many others[111]. Reviews discussing CADD's role in the drug discovery pipeline[110], methods of MD implementation[112], and detailed discussion docking software[113-114] and potential pitfalls[115] are available.

B. A Brief Overview of Molecular Docking

Many docking programs are available today (click2drug.org and Wikipedia maintain impressive lists). In all cases, docking attempts to predict the "correct" orientation and conformation of ligand (called a *pose*) complexed with its binding partner, most often a small molecule-protein complex. Generally, docking has two aims: accurate structural modelling and correct prediction of activity. Unfortunately, this far easier said than done. Identifying key molecular features responsible for biological recognition is a difficult task empirically, let alone predictively. As such, docking is routinely implemented in a multi-step process with each successive step adding complexity[116].

The first step is generation of a pose space. Dozens of implementations exist for this single step, but the idea is to cover as many conformations of the ligand of interest as possible in a reasonable amount of time. The search algorithm dictates how a ligand will be divided (if it is) and how the program will proceed in attempting to identify geometrically complementary surface interactions. Ligand-protein complexes with good complementarity are re-evaluated with a scoring function. The scoring function, depending on a particular implementation, will re-evaluate each pose on the basis of a more rigorous

geometric and other criteria. The "other criteria" are heavily dependent on program used and may be purely energetic (force-field based) or purely knowledge-based, wherein poses are evaluated with how well they match similar structures in data banks. Finally, top scoring poses are evaluated a final time with a ranking algorithm which is usually attempts to account for as many factors as possible (entropy, solvation/desolvation energies, rotational freedom, etc) to yield an accurate binding affinity.

A general summation of the preceding paragraph is thus: all docking programs follow the same basic process. First is make multiple conformations of the ligand, bind each to the protein of interest, evaluate each and rank them in order of best to worst. The details of this three-step process vary widely from program to program and even a brief description of each point would become a several page review of the field, which many are already provided[110, 114-118]. Further, as the field has progressed, and computational power has become more available, the once clear-cut denominations have blurred and mixed. Because of this complexity, interested parties are referred to the an excellent, and brief, basic overview to docking from Jurgen[119] or for more detailed discussion of components, the one from Haleprin[120]. Instead of recapitulating those publications, I will discuss my chosen program in greater detail and why I believe to be a good choice compared to other options in the next section.

C. Defining Boundary Conditions

At the outset of this project, compounds capable of enhancing proteasomal activity were sparse and an explanation for what an interaction may be occurring was non-existent. Additionally, a single paper reported transforming proteasome activators to inhibitors[121] though offered no mechanistic explanation. Imidazoline TCH-165 was unique as it

presented both of these activities without structural modification to the molecule. Intrigued, I wanted to determine how such a thing was possible and believed my best chance was docking. After a long literature investigation, I discovered that there is no "one-size fits all" docking program, and the success of such an endeavor would come down to accurate definition of boundary conditions and interpretation of results. Therefore, I defined the following boundary conditions.

Equipment/Funds. Practically speaking, I would begin docking on my laptop, a dual-core system with 8 GB of RAM. An excellent laptop for heavy student multitasking; however, it certainly was not meant for the computational complexity of molecular docking. The program would also have to be free as I personally could not afford the monetized options and without positive results it wouldn't make sense for the laboratory to invest in something expensive.

Available Crystal Structure. Fortuitously, the first human proteasome structure was published shortly after the start of this project[122]. From another project (see Chapter 4), I knew that species were different enough that accurate modelling would require a human crystal structure. Yet, the resolution was limited to 2.6 Å, solidly in the "high resolution" category for proteins allowing exact tracing of backbone position; however, this level of resolution is very near the 3 Å cutoff where only general protein contours can be made out. This meant that during the searching analysis, the amino acid side chains would be "fuzzy".

The lack of discrete side chain positions was problematic for two reasons. First, a crystal structure is a rigid image of a conformational ensemble. The proteasome in solution is undoubtedly transitioning between energetically near conformers, which are lost on a

single structure. Many docking programs attempt to obviate this limitation by allowing side chains of interest to flex and move around a bound ligand and allow SAR progress as sidechains of interest in the available form could be identified. However, a static structure of clearly defined secondary structure with "fuzzy" side-chains casts doubt on the calculation. Second, many docking programs weight hydrogen bonding quite heavily based on directionality[120]. Chemically this seems reasonable; however, I balked at this as without sub 1 Å resolution (very rare for proteins), positioning of hydrogens is pure guess-work and again, it is a non-dynamic system. I believed this would unfairly weight potentially unwanted poses on the basis of a randomly found and completely uncertain H-bonding interaction. I thought perhaps by making the ligand "fuzzy" accuracy could be regained, and, thankfully, I was not the only one who thought this way (see next section).

Finally, the sheer size of the proteasome rendered many server-based offerings used by other groups untenable and meant flexible side chain docking was unlikely due to computational expense. Also, the lab had numerous imidazoline scaffolds to compare with and docking each manually would take an incredible amount of time. Consequently, I also sought some way to manage a robust workflow.

D. Why Chose Vina?

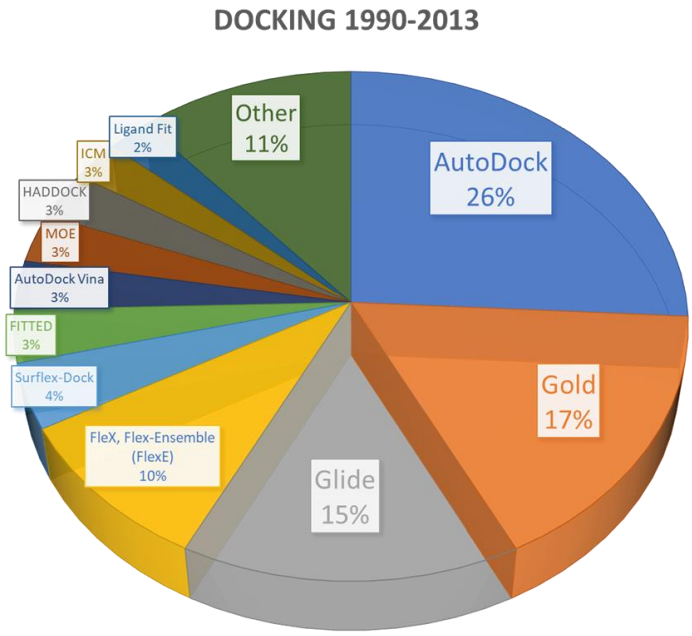


Figure 2.1: Published Utilization of Docking Software 1990-2013

Between 1990 and 2013, AutoDock was the most utilized docking program in the published literature. However, this lion share of utility is not correlated to performance[115]. AutoDock and GOLD are some of the earliest programs written and have had more time in the market to be assimilated by interested parties. Additionally, price is an unfortunate motivator for popularity as free options (AutoDock, GOLD, etc.) continue to be overrepresented in the scientific literature regardless of suitability to the problem at hand[115].

Despite this, I did look at AutoDock first. However, at the time, limitations on number of atoms, rotatable bonds, and grid map size were unacceptable limitations. Additionally, AutoDock employed a genetic search algorithm and a force-field based scoring method[116]. Neither of these are particularly troublesome in the scheme of docking as a whole; however, I had reservations due to my literature readings. In my readings, many purely calculation-based scoring methods were theoretically the best (i.e., theoretically this

should work!), but performance was inconsistent. My thoughts on this are that despite a probably excellent physics-based force field, the simplifying assumptions imposed for usability purposes (time vs accuracy) hobbles the utility of these systems at the moment. For example, AutoDock applies the same energy potential for all hydrogen bonds, as opposed to different terms for different types, and then weights the contribution based on directionality[123]. The scoring calculation is very physics-based, based on the AMBER force field (ff) used to predict protein folding. AMBER calculates the potential energy of the system by summing the contributions from van der Waals energy, geometry (sterics), torsional energy, and covalently bound elements atom by atom[124]. This means, AMBER will calculate each individual H-bond individually; however, AutoDock simply assigns a number to this interaction regardless of other factors, which seems counter-intuitive to the goal. This is just one example of a very non-physics-based simplification for the sake of computation time[123]. Other programs were considered as well though, I eventually decided to use AutoDock Vina.

AutoDock Vina (hereafter referred to as Vina) had many attractive features. Vina used a united atom type scheme wherein each atom is assigned a type with a corresponding set of symmetric interaction functions based on interatomic distance[123]. This greatly reduces computational expense as groups of atoms would be replaced with a dummy "atom" representing the group and makes the compound itself "fuzzy". For example, instead of a methyl group representing a carbon and three hydrogens, you would have methyl "atom" with requisite functions applied. In my opinion, this is conceptually brilliant as the author, Dr. Trott[123], has shifted the focus from energies (force field and the like) to chemical potentials. This is only a qualitative difference as force fields concern themselves with well

depth of a potential while a chemical potential is also concerned with the shape of the potential. This is a difficult thing to describe so observe Figure 2.2.

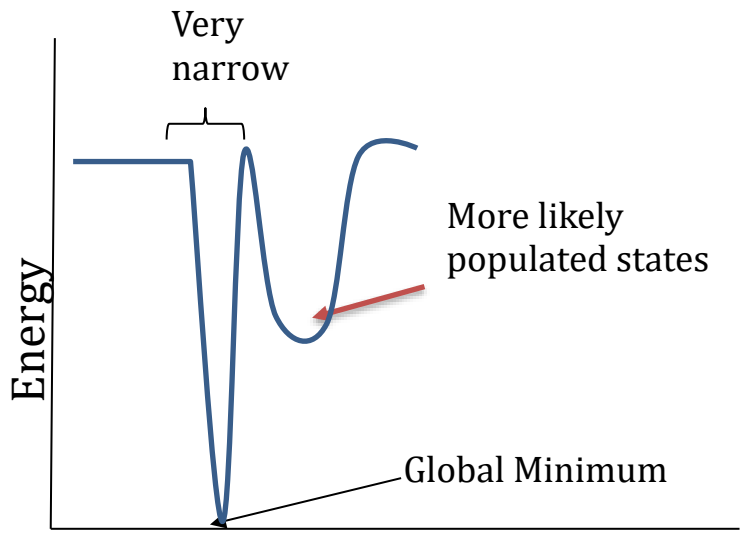


Figure 2.2: Conceptual Illustration of Energetically Preferred States

Docking methods that rely on energetic calculations for scoring are primarily concerned with finding the "Global Minimum" or the lowest energy binding potential possible. However, as illustrated above, this often neglects the shape of the potential well this global minimum occurs in. In the above case, the global minimum is within a well with very sheer and narrow sides, which is entropically very disfavored[125]. Entropic effects are often ignored until the final ranking procedure; however, ranking does not eliminate calculated poses, it simply ranks them. A binding complex so entropically unfavored should be culled during initial scoring which Vina does.

The scoring function utilized, in the author's own words, is a "machine learning" approach to the scoring function. Vina's scoring function is semi-empirical, and a full description has yet to be published. What has been published[123] explains the process as follows:

Initial population of a pose space is generated (i.e. the program makes lots of copies with different conformations). The set of conformations is then evaluated for clashes or other disfavored interactions and are culled, reducing the set of conformations required to be docked. After initial docking, the bound conformation is immediately locally optimized using a quasi-Newtonian method[126]. In other similar programs, the compound would be randomly mutated and rescored. The two scores would be compared, the better score kept, and reiterated until the program is satisfied no other conformation is better. Vina mutates a conformation in the bound state and looks at not only the scoring function but it's gradient (the derivative in each dimension). Chemically, the gradient of the scoring function is the total force acting on the ligand and so accounts for more than just the sum of good vs bad interactions, it accounts for which direction is the bad coming from and the random change is then made in an effort to optimize these interactions based on empirical considerations (i.e. a methyl group is more likely to be X distance than Y from group Z based on published structures). After a set of locally minimized structures are generated, Vina then ranks each and assigns binding affinity values.

A few final notes that need to be mentioned is that Vina does not explicitly utilize charged states nor hydrogens, though utilizes symmetric hydrogen bonding (i.e. directionality is ignored). Chemical intuition may initially balk at this; however, it is a surprisingly useful implementation choice for my system in particular. Many programs implicitly use a single protonation state of the protein in question, and it is a non-trivial task to account for changes in protonation state of every atom in a protein. The atom-type scheme appears to account for charged states during the assignment but doesn't use them explicitly. Hydrogens are often ignored in most cases anyway, and, once again, the atom-

type scheme seems to apply a “this can H-bond” parameter to hydrogen bonding-capable group. However, by treating them implicitly, directionality is no longer a factor, and the goal becomes minimizing the distance between an H-bonding group on the ligand with that on the protein. In recent years, Vina has become one of the most cited programs in use for its incredible ability to quickly identify empirical binding modes.

E. Results and Theory of Proteasome Activation

I eventually was able to set up Vina on my laptop. To begin, TCH-165 was drawn in ChemDraw™ and given 3D coordinates in ChemDraw3D™. The conformation was relaxed using the buil-in MM2 force field and the relaxed structure converted to PDB coordinates. The PDB coordinates were converted to pdqt file format (required for Vina). The proteasome crystal structure was accessed online from the PDB databank (code: 4R3O). Organics and waters were removed, and the protein was converted to pdbqt format.

Due to previously mentioned constraints, docking was done in stages (Figure 2.3). In the first stage, unbiased docking was achieved by including the whole proteasome in the search space and allowing unbiased conformational search.

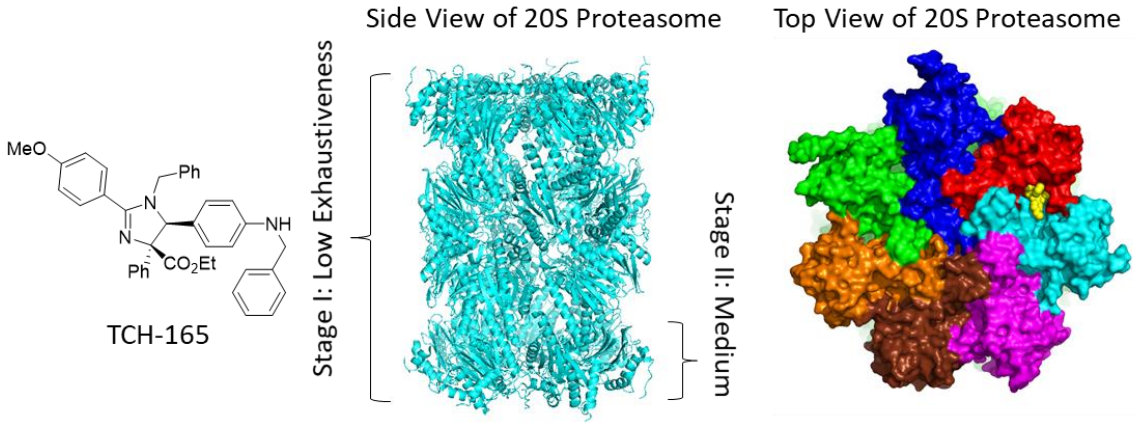


Figure 2.3: Graphical Overview of Docking Process and Result

Exhaustiveness, an arbitrary input for "how hard" the program looks for a solution, was set to 60, default is 8. In this case, 60 represents a "low" setting due to the size of the search space. After the run completed, I was amazed to find no binding modes were predicted exclusively on the alpha ring. I shrank the search space and resubmitted the docking at a higher exhaustiveness (80 was the highest achievable on my laptop due to memory constraints). The final predicted binding mode (shown in yellow in the above figure) rested within an intersubunit pocket. Other binding modes were predicted; however, manual inspection of each pose allowed me to put forth a small fraction of them as "most likely" forms.

The criteria for this identification was difficult to put together, though obvious now. The form of the proteasome crystal structure available was the close inactive form (see Ch. 1 for detail) and all the data suggested an active open form. As such, I was limited to an induced mechanism of action based on the available information and not on biochemical theory. Seeking an induced conformational change, I examined each predicted pose for potential interactions that would result in opening the gate (i.e. pulling this amino acid would pull this beta sheet, which would create space for the alpha helix to move back, which would affect gating residues). After settling on a subset of the predicted poses, I found that all of them resided in an intersubunit pocket formed by adjacent alpha subunits.

Binding in the same pocket would plausibly behave as a competitive inhibitor of the 19S cap resulting in an increase in ubiquitinylated proteins in cells, a published result[105], while simultaneously mimicking the 19S activity and inducing an open form of the proteasome.

Chapter 3. Phenothiazine Small Molecule Activators

A. Introduction to Phenothiazine Activators and Chlorpromazine

With a working theoretical frame-work in hand, the lab felt confident that proteasome activators of different scaffolds could be found. A high through put screen (HTS) was performed on the NIH Clinical Collection and Prestwick libraries. The compounds were ranked based on EC_{50} values, and we were able to successfully identify several hit candidates (Figure 3.1).

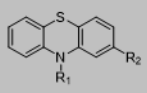
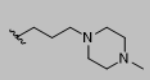
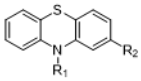
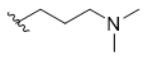
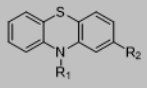
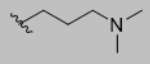
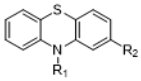
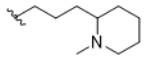
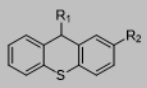
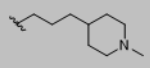
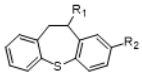
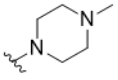
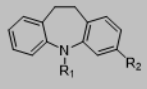
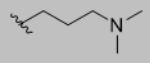
	Drug Name	*Structure*	*R_1*	*R_2*	*EC_{50} (μM)*
1	Thiethylperazine			SEt	1.7
2	Chlorpromazine			Cl	7.7
3	Triflupromazine			CF_3	7.7
4	Thioridazine			SMe	6.2
5	Metixene			H	11.7
6	Methiothepin			SMe	5.3
7	Clomipramine			Cl	13.0

Figure 3.1: Selected Results from HTS of Proteasome Activators

An EC_{50} value represents the concentration at which a molecule produces 50% of its maximal effect, in this case hydrolysis of peptides. The bioassay we use to evaluate is a kinetic assay where the rate hydrolytic release of a fluorescent molecule bound to idealized peptides is taken from the linear portion of the curve (see Section D for detail). Several recurring core structures stood out, phenothiazine chief among them, with activities in the low micromolar range. The phenothiazine core structure is quite old, originally prepared by Bernthsen in 1883 and is considered a privileged structure in medicinal chemistry as they have been found to have insecticidal, antiseptic[127-129], anthelmintic, anti-cancer[130], antiemetics[131], and antioxidants[132]. The identified phenothiazine compounds are all 1st generation antipsychotics. These compounds are believed to bind to D_2 receptors and prevent access by the endogenous ligand (dopamine)[132-134].

Chlorpromazine (CPZ) was selected to be the hit compound due to good reproducibility in bioassays. The goal then became removal of the natural dopamine antagonism with preservation, if not improvement of, proteasome activity. Structural features required for the dopamine antagonism were fairly well known in the literature[135-136] (Figure 3.2). Despite this, I was unwilling to purse a traditional SAR project for several reasons. First, phenothiazines have long been known in the literature with many synthetic routes to desired structures available in the literature[137]. Synthesis would be non-trivial; however, there was very little room to pursue a worthwhile methodology. Second, due to the phenothiazine's promiscuity[138], a group change could destroy both proteasome and dopamine activity while giving it some other unwanted activity. Next, random generation of analogues aimed at removal of dopamine activity while relying on bioassay evaluation

to determine if the change is good or bad, then making another *random change* equates to fumbling in the dark, and I would prefer a flashlight.

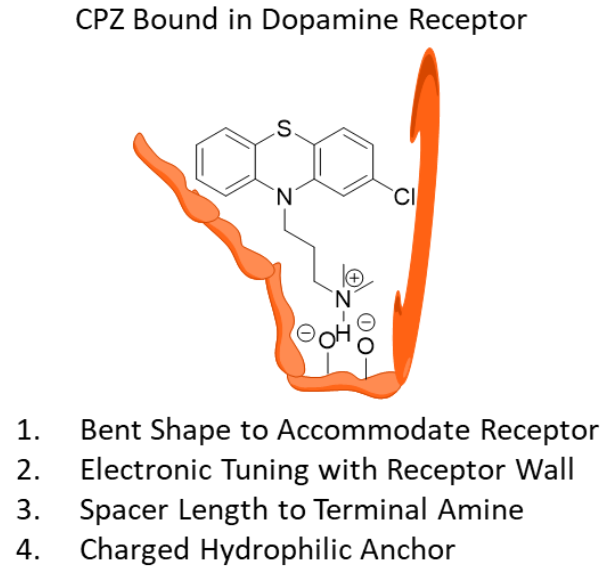


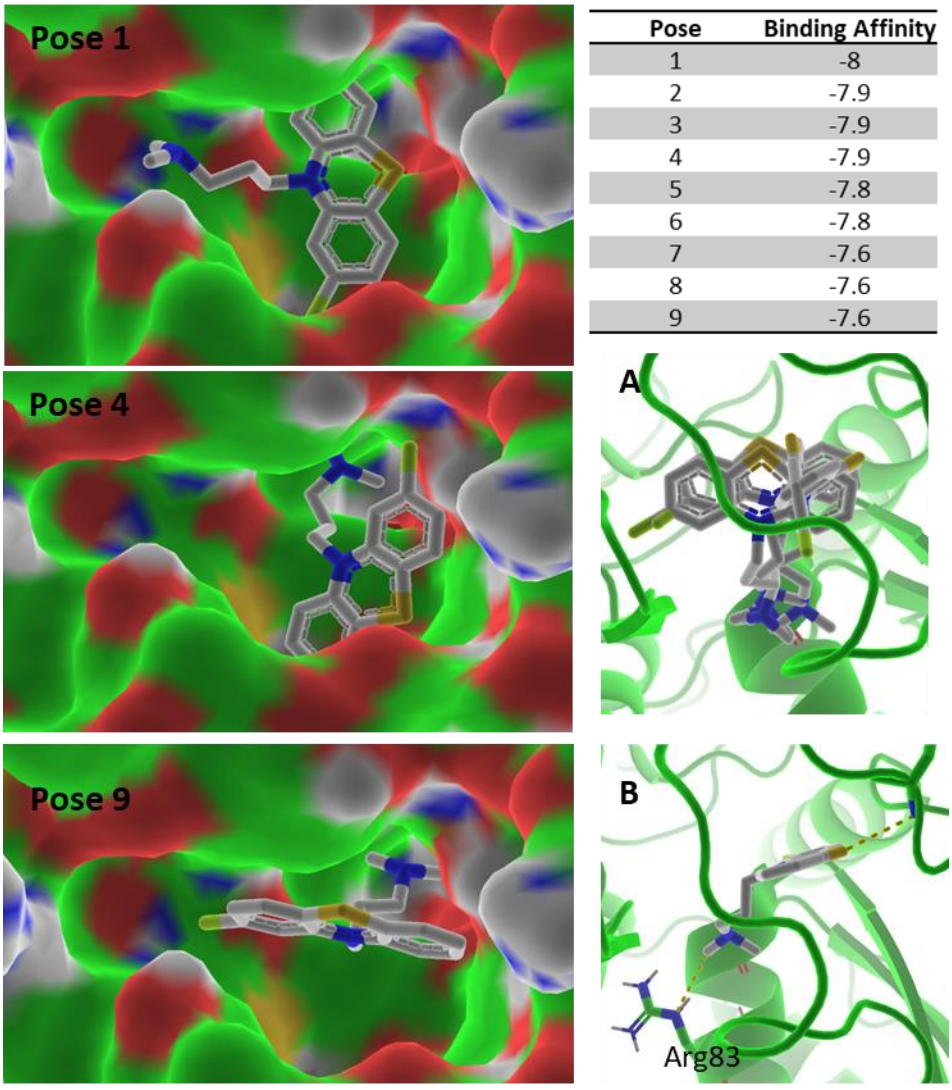
Figure 3.2: Graphic of Key Dopamine Receptor Interactions with CPZ

Finally, even a successful SAR campaign may end with a potent compound but no mechanistic insight as to how the effect is achieved. This could result in requiring a new SAR campaign if the compound has off-target effects that have to be removed. Instead, I decided to use this opportunity to validate the docking model empirically in hopes of gaining some mechanistic insight while meeting the requirements of the lab.

B. Validation of Docking Model by Repurposing of Neuroleptic Agent

CPZ was subjected to the same docking procedure detailed in Chapter 2. Likewise, it displayed an impressive preference for the alpha ring intersubunit pockets and found lodging within an intersubunit pocket, suggesting a similar mechanism of action to that of the imidazoline scaffolds. The binding affinities calculated by Vina (Figure 3.3, Table) had a very narrow range of values, despite some poses possessing very different orientations (Figure 3.3, Pose 1 vs Pose 2). Overlaying multiple binding poses (Figure 3.3A) revealed

preference for the tail to be oriented downward toward the center of the enzyme. Investigating this area revealed a proximate arginine residue (Arg83, Figure 3.3B) that was postulated to be the key residue anchoring the tail downward. As the terminal amine tail was known to be key in binding to the dopamine receptor (Figure 3.2), it was targeted for replacement first.



Pose	Binding Affinity
1	-8
2	-7.9
3	-7.9
4	-7.9
5	-7.8
6	-7.8
7	-7.6
8	-7.6
9	-7.6

Figure 3.3: Binding Affinity Values for Top 9 Poses (Table), Different Orientations with Similar Binding Affinity (Pose 1, 4, and 9), Overlay of Similar Poses (A) and Proposed Anchoring Residue Arg83 (B).

Two analogues were proposed based of the above model. A carbon analogue compound **3-1** (Figure 3.5) was proposed as a negative control as replacement of the terminal nitrogen with a carbon would be expected to remove dopamine activity as well as its interaction with Arg83 (Figure 3.3B). Replacing the terminal amine with a sulfonate tail (Figure 3.5, compound **3-2**) would instead be expected to make a powerful salt bridge with Arg83 while being unable to bind to the dopamine receptor. However, when both compounds were submitted to docking investigation both had no predicted poses within the intersubunit pocket. According to the current theory, both compounds would be expected to be inactive as proteasome agonists. Curious to know if the sulfonate was a bad choice or if a secondary factor was to blame, two additional analogues containing a longer chain of four (Figure 3.5, compound **3-3**) and five carbons (Figure 3.5, compound **3-5**) respectively docked against the proteasome. Intriguingly, only compound **3-3**, four carbon linker, was predicted to bind in the same pocket, and incredibly in the same pose as CPZ (Figure 3.4)!

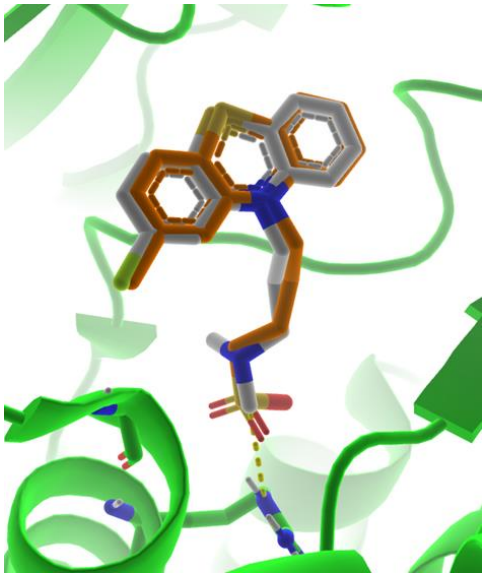


Figure 3.4: Overlay of Predicted Binding Modes of CPZ (white) and **3-3** (orange).

This was a fantastic result as the synthesis of all four analogues (Figure 3.5) and biological testing would allow swift determination of how trust-worthy the predictive power of docking could be in the absence of a known binding site.

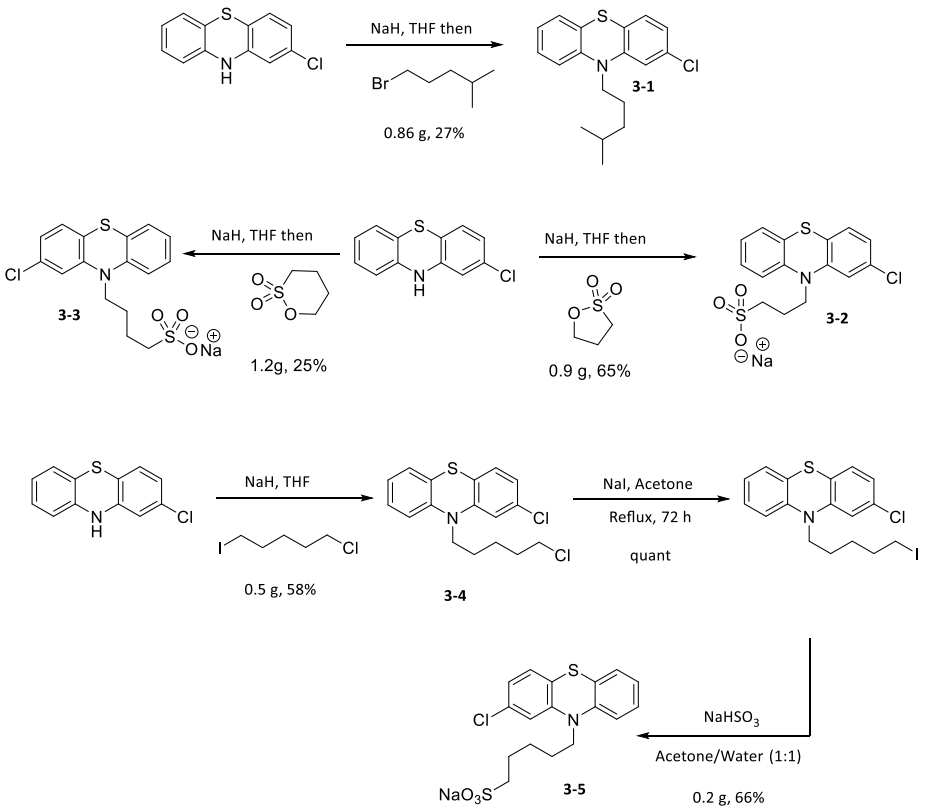


Figure 3.5: Synthesis of First-Generation Validation Compounds

Compound **3-1** was generated by treatment of the phenothiazine core with sodium hydride in tetrahydrofuran (THF) followed by addition of the alkyl halide. The reaction proceeded fairly smoothly; however, purification was unfortunately a challenge. The extreme similarity between product and starting material resulted in co-elution regardless of solvent polarity. Selective precipitation of the starting phenothiazine core with chloroform ultimately provided the alkylated core. Synthesis of the **3-2** and **3-3** proceeded smoothly with treatment of deprotonated phenothiazine with the appropriate sultone which precipitated upon cooling. Synthesis of **3-5** was a little more involved requiring alkylation

of the phenothiazine core to give **3-4** followed by Finkelstein conversion of the terminal halide to an alkyl iodide before displacement with sodium sulfite to attain **3-5**.

The results were outstanding[139]. As predicted, compounds **3-1,2**, and **5** were inactive in testing while compound **3-3** showed excellent dose-response up to 8-fold over proteasome alone (set to 100% in Figure 3.6A) and activating all three catalytic sites (Figure 3.6A, yellow, red and purple lines).

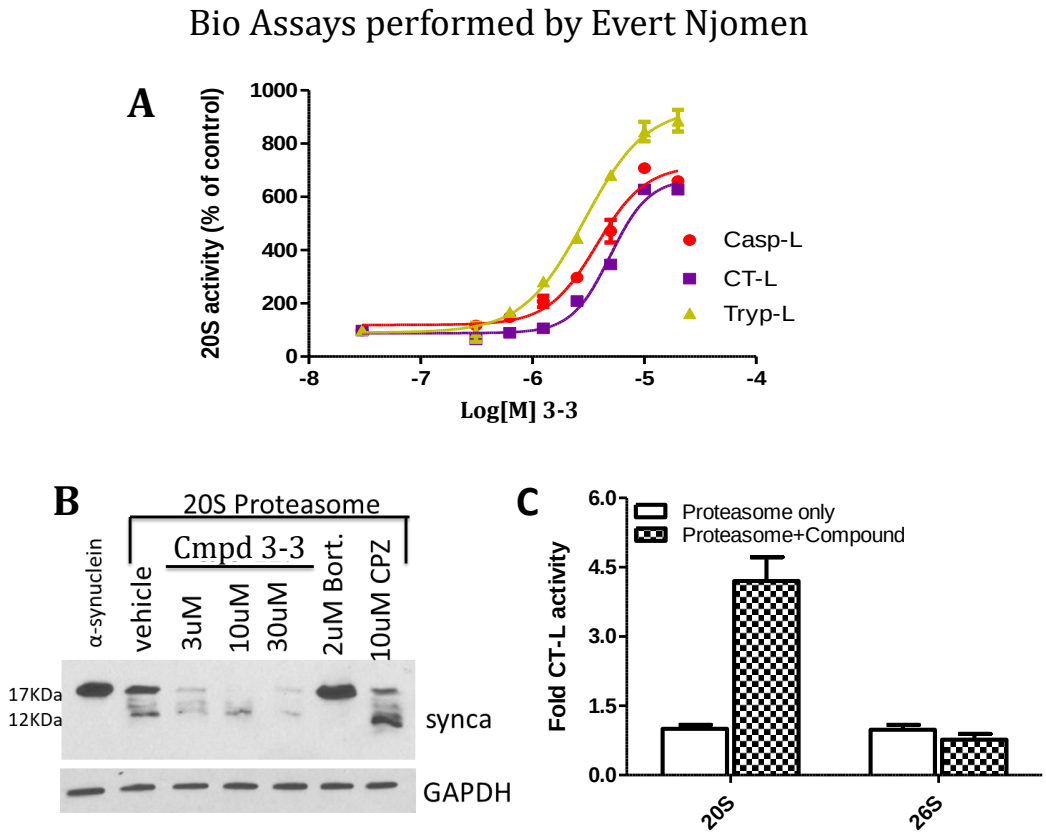


Figure 3.6: Dose Response of 3-3 (Top), IDP Degradation (B), and Complex Selectivity (C)

Compound **3-3** was also investigated for its ability to enhance degradation of an IDP (Figure 3.6B). From left to right in Figure 3.6B, lane one illustrates a control for where alpha synuclein (synca) resides on the gel. Lane two is the effect the proteasome alone on synca digestion (the smaller fragments below the synca level). Lane 3, 4, and 5 show

increasing digestion (absence of fragments) as the dose of **3-3** increases. Lane 6 contains a negative control, BTZ. BTZ is Bortezomib, a proteasome inhibitor which prevents the 20S from degrading any proteins and results in no change from the synca control in Lane 1. The final lane uses CPZ as a positive control to ensure validity of the assay and also demonstrates the increase in 20S proteolysis induced by compound **3-3** at the same dosage (lane 4 vs lane 7). GAPDH is an added protein used to ensure equal loading of sample. As each lane has an approximately equal amount of GAPDH, we can be confident that the differences in synca is due to proteasomal degradation and not simply different amount of protein. GAPDH, as a structured protein is not susceptible to degradation by the 20S CP (see Chapter 1 for details). Improvement in IDP proteolysis strongly supported the theory that proteasome stimulation could be a viable therapeutic strategy (see Chapter 1). **3-3** was also investigated for its effect on the 26S proteasome (Figure 3.6C). Excitingly, only the 20S core particle showed increased proteolysis in the presence of **3-3** but had no effect on the 19S (Figure 3.6); strongly supporting the docking model that binding is occurring utilizing the same binding site as the 19S.

Despite the inarguable success of compound **3-3**, it still had a few undesirable features. The permanently charged tail is unsuitable for cell permeability[140], at very high concentrations *inhibition* of the proteasome was observed (a property of detergents[141]). Additionally, polymorphism of the compound was suspected as different batches possessing identical spectral data performed differently under the same assay conditions. As such, a more drug-like molecule was sought to become a new lead structure to explore a docking guided SAR. Toward that end, a number of potential compounds were designed with the aim to further validate the docking model as well as yield a suitable lead compound.

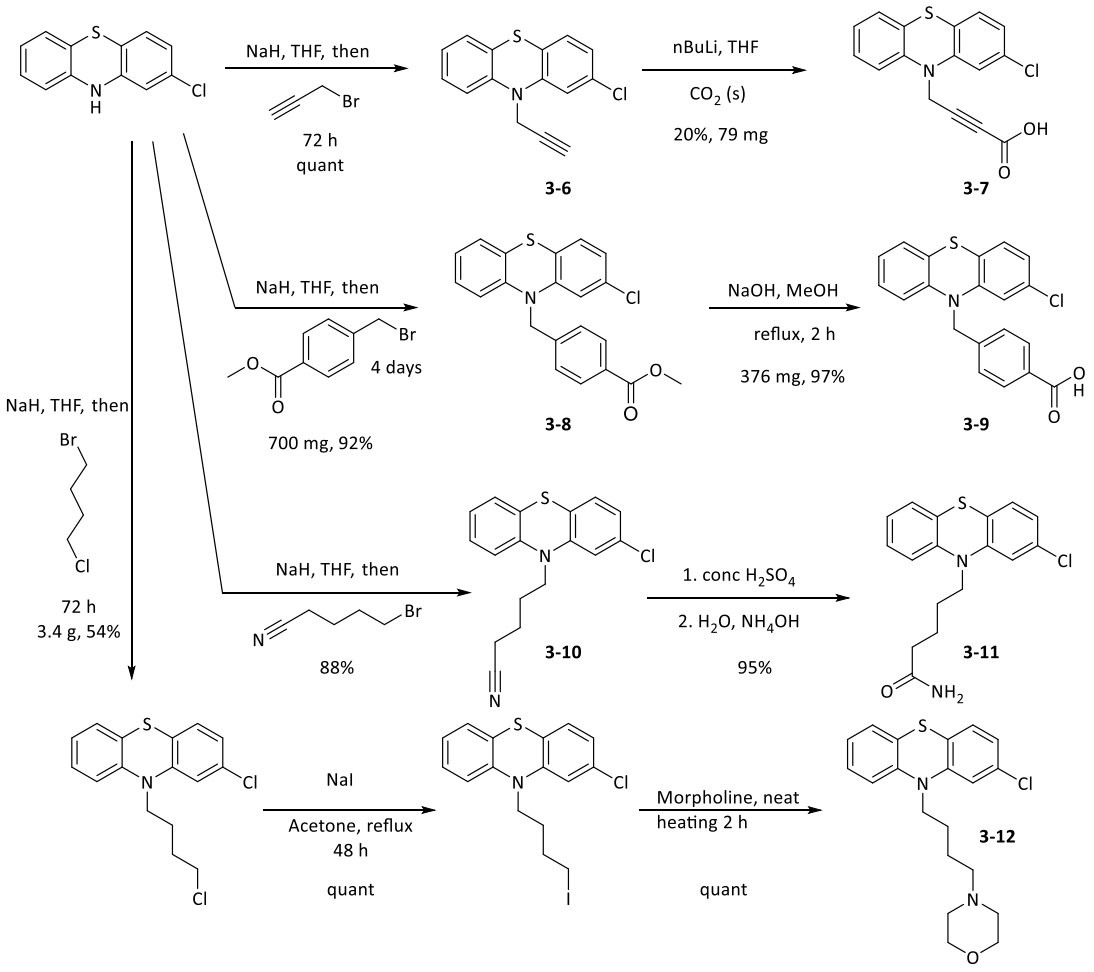


Figure 3.7: Synthesis of Lead Candidates

Compound **3-7** was proposed to investigate the flexibility of Arg83. Docking predicted the **3-7** would be inactive; however, as flexibility in protein side changes could not be simulated (see Chapter 2 for detail), such an induced change could not be checked computationally. Synthesis of **3-7** was therefore pursued via alkylating the core scaffold with propargyl bromide to yield **3-6**. The terminal alkyne was deprotonated using nBuLi, and the resultant anion used to trap CO_2 giving the desired product **3-7**.

From the activity differences between compounds **3-2,3,** and **5**, it was known that tail length had a strong effect on the ability to stimulate proteasome activity. A ligand with less conformational freedom often binds more effectively to a protein binding pocket due

to lower entropic penalty[125, 142]. As such, we sought to lower the number of rotatable bonds while preserving favorable interactions with Arg83. Compound **3-9** was expected to offer a more rigid system than **3-3** while preserving the ability to interact with Arg83. Synthesis was achieved through benzylation of the core structure followed by hydrolysis provided the desired compound with minimal issue.

We were also interested in exploring the tolerance of functionality at the end of the tail. As stated above, the conformational flexibility may be a detriment; however, the excellent activity of the sulfonate analogue made us unwilling to completely abandon the structure without an alternative. An analogue containing an amide tail (compound **3-11**) was generated by alkylating the phenothiazine core with valeronitrile (compound **3-10**) followed by acid catalyzed hydrolysis. A morpholine tail was also explored as a mimic for the active tail of several of the original active compounds (Figure 3.1, entry 1,4, and 5). Synthesis was pursed in a similar manner to **3-5** except displacement of the terminal iodide was done using morpholine to give **3-12**.

It is also worth noting that the flexible ester and carboxylic acid were obvious additions to this exploration. However, during the course of the synthesis of these compounds, it was found that this compound would be unsuitable. As previously mentioned, purification of these compounds was incredibly challenging/frustrating. Often, the resultant product mixture would be purified via column chromatography using gravity to isolate a mixture of phenothiazine starting material and product. This mixture would then have to be suspended in chloroform and chilled overnight to precipitate out phenothiazine to give a product. In the case of the flexible ester tail, the acidity of the silica gel resulted in some amount of hydrolyzed product. The carboxylate, in solution, then

began catalyzing the decomposition of the product into more phenothiazine starting material and a volatile side product (Figure 3.8).

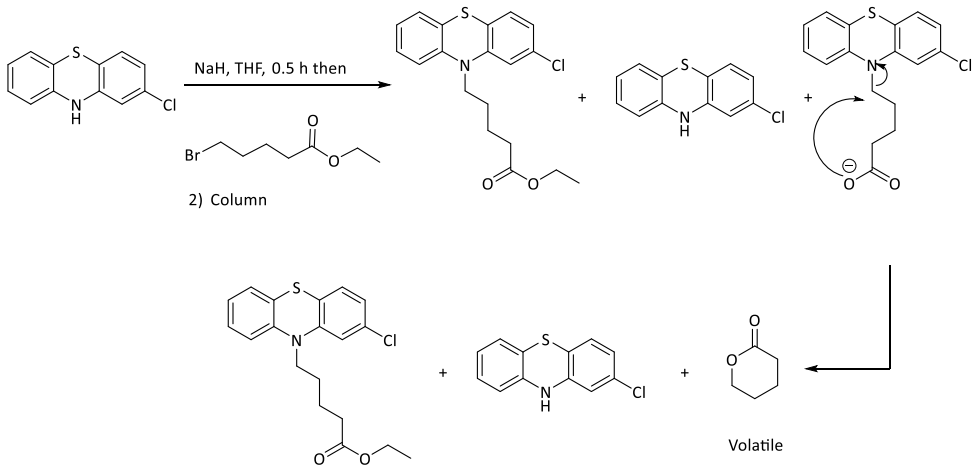
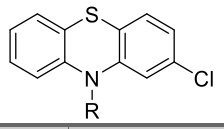
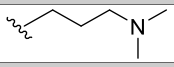
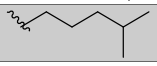
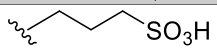
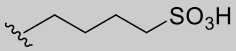
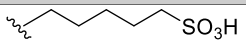
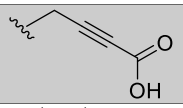
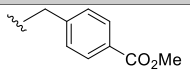
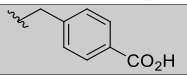
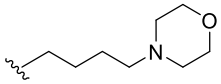


Figure 3.8: Synthesis of Flexible Ester and Depiction of Decomposition

This particular process was challenging to determine as after filtration and drying, it simply appeared as though the precipitation was incomplete. However, due to an impossible mass balance (I had isolated more phenothiazine from this process than should be possible based on crude mass), I investigated on a small scale with deuterated chloroform and was able to identify the lactone in the mixture. Synthetically, this is easily remedied. However, this decomposition pathway is likely to be far more prevalent under assay conditions, and even more so in a cellular assay. Because of this, it was agreed that pursuit of this compound should be abandoned due to it being unsuitable for biological testing.

Desired compounds in hand, they were submitted for biological evaluation (Table 1).



		20S		D2R	
Compound	R	EC_{50} (μM)	Max Fold	%Inhibition	Ki (μM)
CPZ		9.9	20	77.9	0.48
3-1		>25	--	0	>250
3-2		>25	--	--	--
3-3		6.3	8	-4.5	>250
3-5		>25	--	--	--
3-7		>25	--	-2	>250
3-8		15.6	10	2	>250
3-9		6.4	2-3	1	>250
3-12		8.9	4	74.5	2.97

20S Assays performed By Evert Njomen Dopamine Activity Assays Performed by Dr. Benita Sjogren

Table 3.1: Tabulated Depiction of Bioassay Results. [139]

Unfortunately, though not entirely surprising, **3-12** was a potent dopamine antagonist and limits functionality options for this position (i.e. no basic amine functions lest they regain unwanted off target effects). Compound **3-7** was completely inactive, suggesting Arg83 does not have a great deal of motion available to it as **3-9** was quite potent. The amide derivative **3-11** (Figure 3.7) was likewise inactive in the bio assay screen but was not tested for dopamine activity. Two active candidate compounds (**3-8** and **3-9**) were tested by Evert Njomen in a manner similar to **3-3** (protein assay Figure 3.6B) and **3-8** was chosen as the lead compound due to superior ability to degrade IDPs[139].

C. Attempts at SAR: Challenges of Irrational Data

The beginning of this project was an inarguable success. We had demonstrated the utility of a docking-based theory for identification of proteasome agonists (literature first), found a new scaffold capable of such agonism (the phenothiazines, literature first), illustrated its ability to remove IDPs implicated in common neurodegenerative pathologies in cells (offering a new strategy for targeting such pathologies), and demonstrated we could remove undesirable off target effects. Our interest then turned to making a more drug-like proteasome agonist. With little literature precedence, we arbitrarily decided that any compound unable to produce 2-fold activity of the lone 20S (i.e. double the 20S CP's hydrolytic activity) would be considered inactive as a rough guide to aid in optimization.

Many strategies were discussed; however, in the course of synthesizing the previously discussed compounds and other unpublished attempts (not detailed herein), the phenothiazine core was found to be intermittently thermo and photo-sensitive depending on group attachments, though no clear pattern emerged to predict this. Additionally, the nucleophilicity of the ring sulfur interfered with derivatization steps attempted on the sulfonate analogues. Many other points against continued use of the phenothiazine core can be given; however, the consensus was the first priority should be its removal.

As removal of the core structure was desired, investigation began there with the goal of logically transitioning to new chemical space. Using the candidate **3-8** ester tail, several analogues were generated via benzylation. A general procedure was discovered to give acceptable yield and ease of purification which was used in the generation of all benzylation reactions afterward. Using DMF as a solvent, phenothiazine would be deprotonated using sodium hydride to give a red to red/orange solution. The desired

electrophile would then be added as a single portion and stirred at room temperature wrapped in foil for 16 hours (Figure 3.9). The initial strategy was two-fold. First, the ring substituent was examined for its effect on activity.

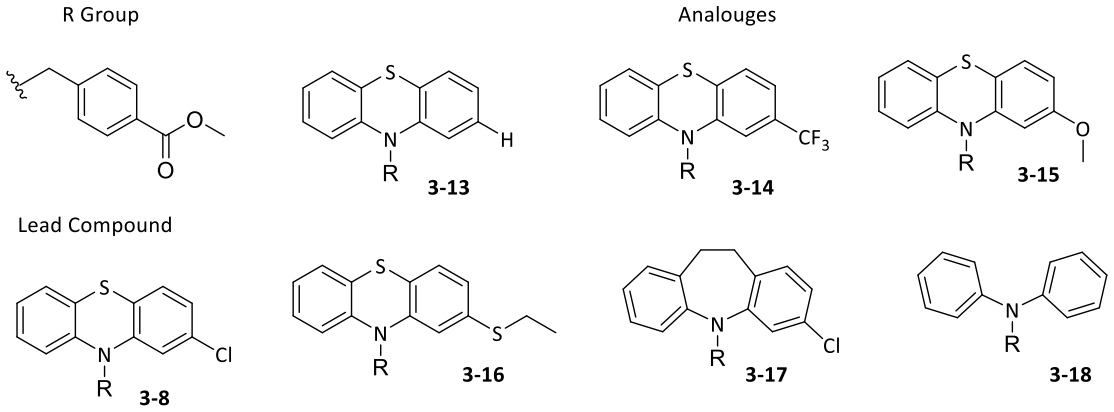


Figure 3.9: Analogues of Lead Compound **3-8**

The chloro-substituent was replaced with hydrogen to probe for a potential halogen bond (**3-13**). Trifluoromethyl was examined as it has literature precedence for increasing cell permeability and also acts as a powerful electron withdrawing group, though lacks π-donor ability (which the chlorine possesses)[143-144]. The thio-ether (**3-16**) was examined because it was the core structure that displayed greater activity in the original HTS (Figure 3.1 entry 1). The methoxy substituent (**3-15**) was desirable as it had been shown to greatly increase activity of the previously discussed imidazoline scaffolds. We also examined the requirement of the sulfur atom. The ethyl linkage (**3-17**) was chosen over a single methylene unit because two methylene allowed minor flexibility while preserving the overall geometry **3-8** without the sulfur atom. Excising the sulfur entirely (**3-18**) allows the phenyl rings to expand and greatly changes the shape of the molecule.

For the next discussion of bioactivity results, it is important to note here the *intentional* omission of collected bio data about to be discussed. It has been omitted for

narrative coherence and will be discussed in detail in the next section (Section D) in great detail. For now, activity results will be discussed in general terms.

Surprisingly, **3-15** and **3-18** were reported to be inactive while the remainder of the analogues had approximately the same potency. This was very difficult to rationalize using docking as the uncertainty in the binding mode meant several different poses could be the active one and each compound bound within an intersubunit pocket; although not all in the same one. The result was suggestive that the tail had greater effect in activating the proteasome than the attached core structure.

To evaluate this theory, both the phenothiazine (Figure 3.10, center left) and the iminodibenzyl (Figure 3.10, center right) scaffolds were used (Figure 3.10). However, early in this investigation it was clear something was wrong. Both compounds were performing approximately the same in *in vitro* tests. Also, equal potency was reported for the para-methyl ester (Figure 3.9, compounds **3-8** and **3-17**), the meta ester (Figure 3.10, compounds **3-21** and **3-22**), and even the unsubstituted benzene rings (Figure 3.10, compounds **3-19** and **3-20**) provoked 2-fold activities, albeit at high EC_{50} values (>25).

These results (all behaving basically the same) were both unexpected and frustrating. The absence of the ester moiety was expected to drastically decrease activity of the compounds and allow the determination of which core was more suitable to explore. The size of the intersubunit pockets was suspected to be a contributing factor as the amount of available binding surface made them oddly accommodating.

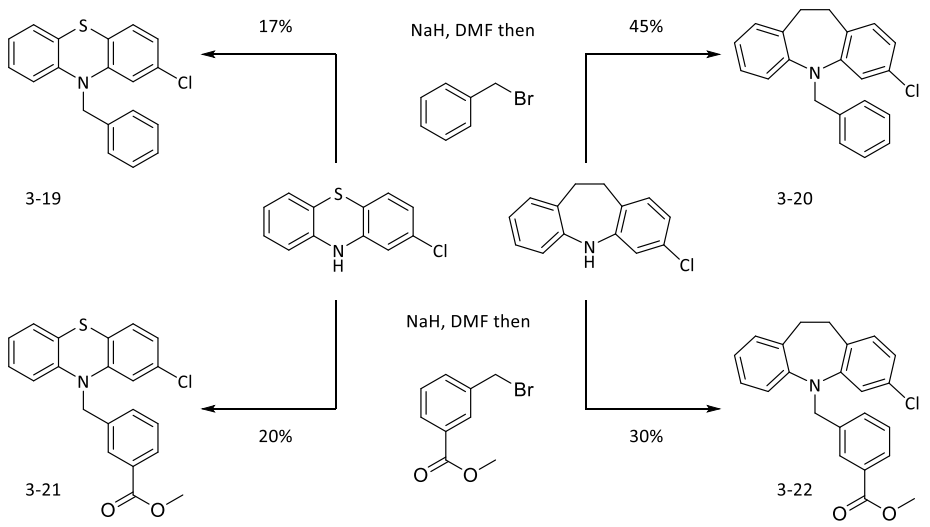


Figure 3.10: Synthesis of Tail Analogous

Expanding the structure outward was debated; however, the project goal of maintaining drug-like properties limited mass addition to approximately 120 $\frac{g}{mol}$ (as the average masses were ~380 and most sources claim a molecular weight of 500 $\frac{g}{mol}$ should be the goal[145-146]) which would not go far in expanding into the pocket. Further, the on/off nature of the analogues made docking refinement impossible. Nine predicted binding poses across 3 binding sites meant an argument could be made for any site. A distribution of activities was needed to make proper progress as on/off was simply too extreme to effectively build more from theory.

As I was unwilling to generate analogues through random synthesis of whatever was in the organic cabinets, I took inventory of what we had discovered and its relationship to the known literature. Despite SAR stagnation, a few key points could be gleaned:

- Core Structure requires some rigidity (compound **3-18** vs everything else)
- Tail Length is important but variable (compound **3-7** vs **3-9** and **3-3** vs **3-4/5**)

- Some measure of hydrophobicity is required along an edge of the molecule (performance of phenothiazine and iminodibenzyl core vs others)
- Ring substitution apparently does not matter (compound **3-13** vs **3-8**)

On top of my own work, literature at the time offered a few additional suggestions. Mutational studies on the C-terminal HbYX motifs (Chapter 1) demonstrated the necessity of a penultimate tyrosine residue and provided evidence that the C-terminus carboxylate was required for association but *not* activity of the 19S-20S complex[70]. This provides a possible explanation for the remarkable similarity of activity between ester, carboxylate, and the lack thereof should it be aiding in orientation but not effect.

I also went back to the original HTS list in an effort to see if a functional group suggestion could be found. Each potent compound possessed a tertiary amine tail (which everyone noticed); however, what now stood out was the pKa similarity between the tertiary amines, approximately 10 for the protonated analogue, and the required phenolic tyrosine pKa (9.6). Additionally, Trader and co-workers published the identification of two new proteasome stimulators AM-404 and MK-886 with strikingly similarity to our published compound and the back-bone tyrosine found in HbYX motifs (Figure 3.11)[147-148].

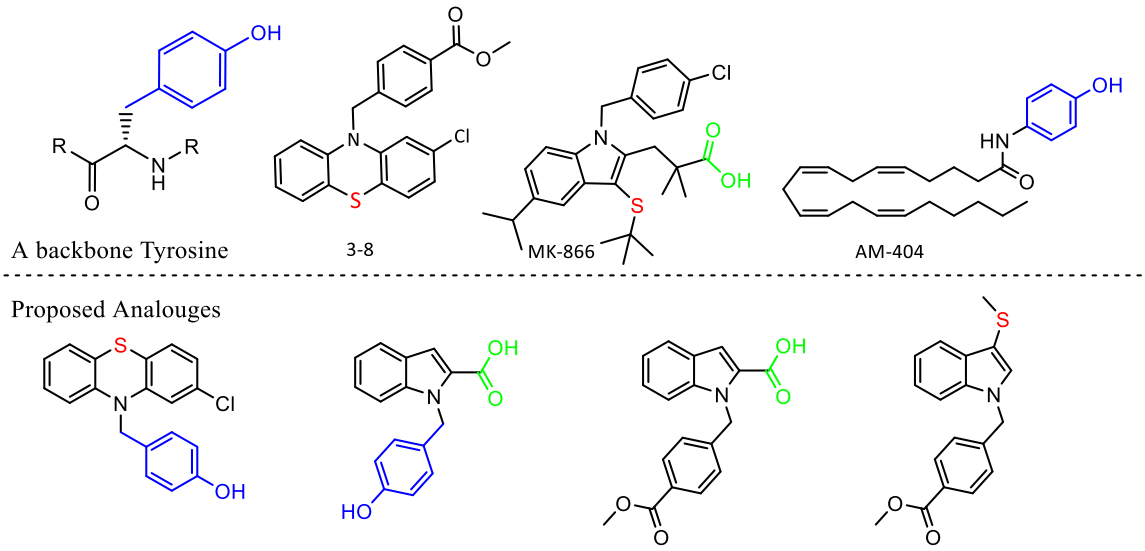


Figure 3.11: Proposed Activator Structures with Key Interaction Motifs Highlighted

As the phenolic tyrosine had been demonstrated to be necessary for activity[149], the obvious course of action seemed to be attachment to the phenothiazine core. By the same notion, attachment of the benzyl phenol at the nitrogen of an indole with a C2 carboxylate would make a very near mimic for the endogenous HbYX tails. As the methyl ester appeared to be providing some stimulatory activity to our previous compounds, I was curious to see its effect on an indole. Likewise, with a C2 carboxylate, but also with a thioether in the C3 position. This compound would be an effective mimic for **3-13** and allow more investigation of the phenothiazine core.

Synthesis of the indole analogues (Figure 3.12) was done in multiple steps. Nucleophilic attack on an activated disulfide gave the indole thioether **3-23**. After purification, this was benzylated with our general procedure (sodium hydride in DMF followed by electrophile) to yield the target compound **3-24**.

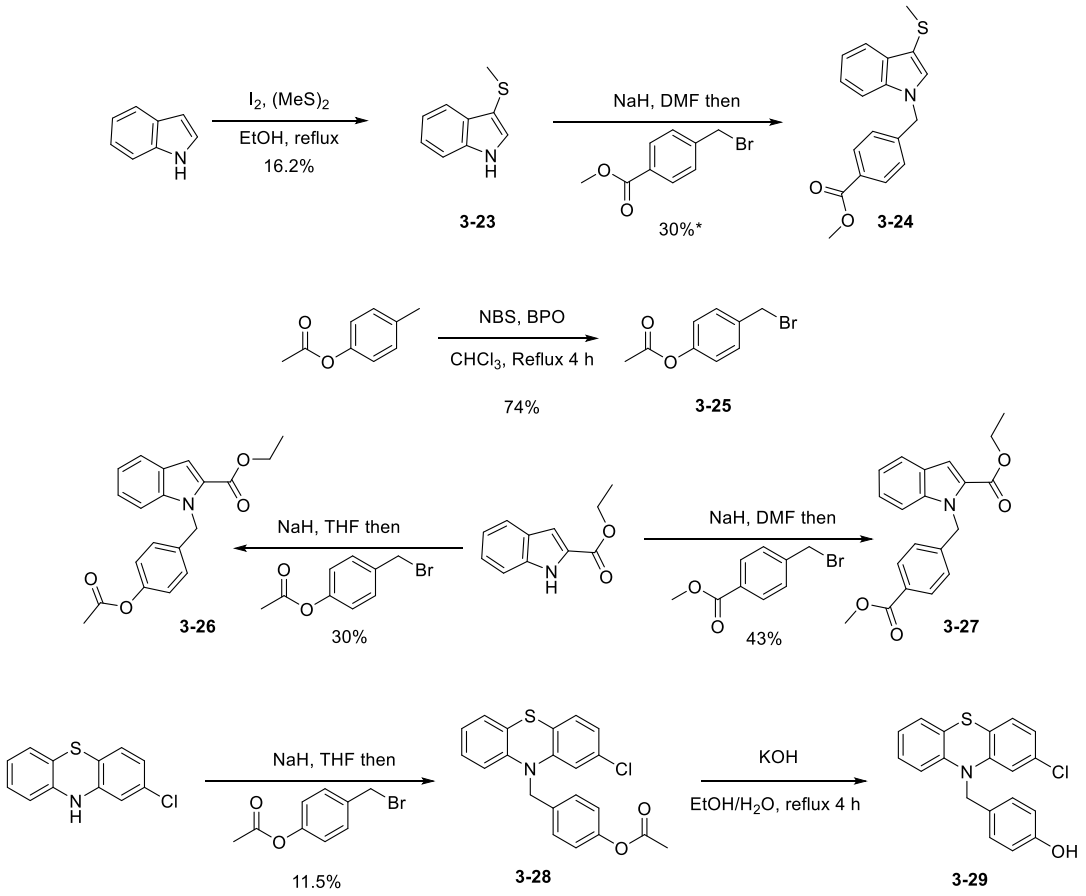


Figure 3.12: Synthesis of HbYX Small Molecule Mimics

The electrophile **3-25** was synthesized via benzylic bromination in freshly distilled chloroform and coupled with anionic phenothiazine to yield **3-28** which was hydrolyzed to yield the desired compound **3-29**. Instead of installing the C2 carboxylate, the commercially available C2 indole ester was benzylated to give the phenolic precursor **3-26**. The indole ester was also benzylated with the **3-8** tail to see if the diester species would be active. Unfortunately, only compound **3-28** was active though, as its hydrolyzed product was inactive, **3-28** is an unsuitable compound for further optimization as it would be expected to hydrolyze to the inactive compound under cellular conditions.

Concomitantly, with the above synthetic and computational hurdles, periodic complaints about the reproducibility of the compound's activity were raised. Some would

be minor variance expected in any biological system; however, others raised serious concerns as some compounds activity would drop below our 2x threshold. A scaffold that sometimes is “an activator” and other times “is not” would be disastrous. As accurate and reliable bioactivity was crucial to the success of my project, I transitioned into the bio-lab to investigate assay conditions.

D. Optimization of Proteasome Agonist Evaluation and Work Flow

The original protocol for the biological assay most commonly performed by our lab is as follows: Purified 20S proteasome is dissolved in a suitable buffer to maintain its biological activity. Drug is added, and the mixture is allowed to sit for 10-15 minutes. An idealized peptide bound to 7-aminomethylcoumarin (AMC) is then quickly added, and the rate of fluorescence measured over time at 37 °C for 1 hour. Rate is determined by taking the slope of the line of through the linear region of the relative fluorescent unit (RFU) vs time graph. By varying the concentration of drug applied and keeping all other concentrations constant, a dose-response curve of activity vs concentration can be drawn, and different drugs compared for their effect on the proteasome, most commonly based on their EC_{50} value. An EC_{50} value is the effective concentration (EC) that produces half of the maximum response (50%) and graphically is the inflection point of the dose-response curve.

After being taught the procedure and demonstrating competency, I began testing my own and other’s compounds. I too found intermittent performance to be an issue and tracked the cause to be an unstable vehicle (Figure 3.13, left). This discovery was distressing as the effect of a drug is now variable and a positive or negative determination was nearly random. For example, an example compound that adds 0.02 to the rate of

hydrolysis may measure <20% activity increase (Figure 3.13, Trial 2, blue block) in one trial while measuring double the activity in another (Figure 3.13, Trial 3, blue block).

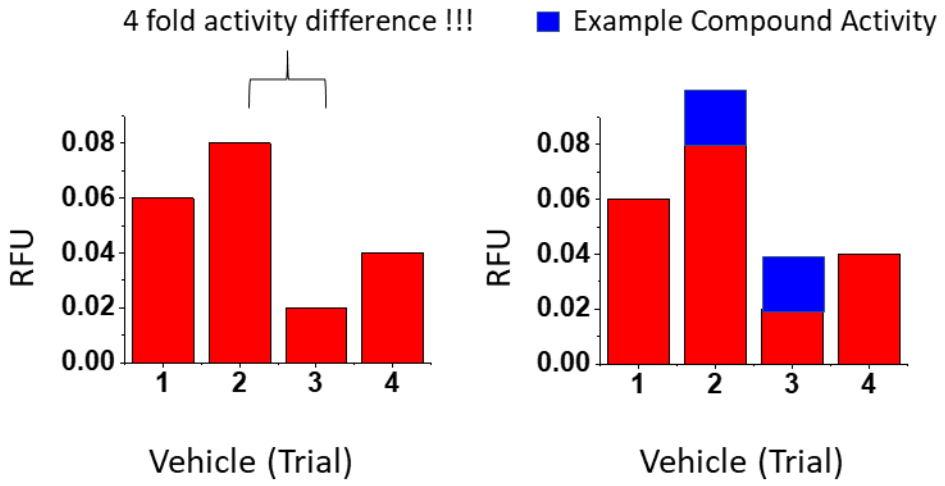


Figure 3.13: Illustration of Vehicle Stability Issue

A systematic investigation into the factors that can affect the proteasome background activity was undertaken. Determination of effects was a time-consuming process, as each change had to be subjected to the full protocol to determine outcome. Numerous factors such as salt concentration, cation source (i.e. potassium vs sodium), presence/absence of salt in buffer, as well as resistivity of the water were examined, and a summary of findings and corrective actions are depicted below (Table 2).

With greater understanding of factors effecting proteasome background activity, attention turned to modification of the protocol. Some changes made were by necessity. For example, many unusable data points are generated while the proteasome, drug, and substrate (fluorogenic peptide) cocktail warm to assay temperature due to the changing pH of the medium, and the fact kinetics of chemical reactions tend to increase with increasing temperature.

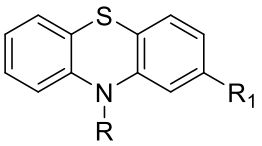
Error Source	Problem	Correction
Buffer	• Variable pH with Temperature • Instability of solvated Tris (~14 days) • Daily pH change	Buffer is to be made fresh and pH adjusted daily to a pH that will give 7.4 at 37 C. After 2 weeks buffer is disposed of, sooner if required.
Water	MiliQ water was not stably at high resistivity. Cascadingly effects pH, protein folding, etc. Accounted for >2-fold activity changes in the same batch of 20S proteasome	MiliQ water is obtained from a different lab with reproducible and stable water supply.
Salt effects	Addition or removal of salt greatly affects the activity of the proteasome and was sometimes added, sometimes not. This has wide ranging effects.	Using "good" miliQ water and sodium chloride gave the most consistence results across multiple assays and so was chosen to be in all *in vitro* assays
pH probe	pH probe was found to be inconsistent/faulty.	Probe replaced.
Enzyme Equilibrium	Immediate data collection gives wide variety of outputs due to the changing pH (Tris) and changing kinetics as the enzyme warms.	Incubate drug and Enzyme at 37 C before addition of substrate.

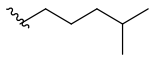
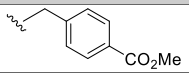
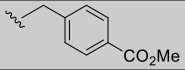
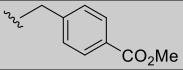
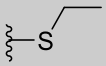
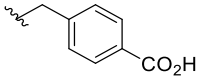
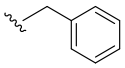
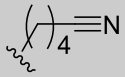
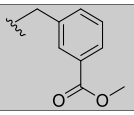
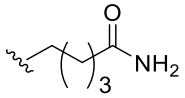
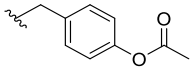
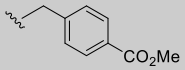
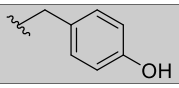
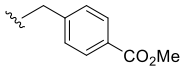
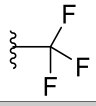
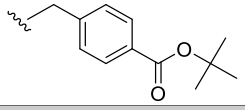
Table 3.2: Summary of Sources of Error and Corrective Actions

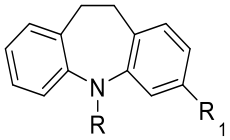
As such, incubation of drug and enzyme at assay temperature beforehand addition of substrate and immediate reading was an obvious modification. However, more changes needed to be made. The original protocol performed a full dose-response on every compound, limiting each plate to a maximum 4-6 compounds. As the synthetic section of the lab expanded and purification of compounds was becoming easier due to lab acquisition of a medium pressure liquid chromatography instrument (MPLC, colloquially referred to as an auto-column), analogue generation was rapidly progressing. This required a more efficient work flow to accommodate the newfound analogue backlog. To achieve this, a pre-screening protocol was generated. By checking each compound at three concentrations,

in duplicate, up to 13 compounds could be screened simultaneously. Further, discussion with a SpectraMax technician identified other variables that could optimize performance: lowering the filter cutoff to more than 5 nm below emission and volume in wells being key incorporations. Volumes were reduced to confident minimums allowing 50% reduction in enzyme usage. These changes allowed the same amount of enzyme to be utilized for twice as many assay plates with each plate containing >2x the usual number of compounds. A screening protocol was also established with: (1) a three-point screen graded on a pass/fail basis to determine active compounds (2) active compounds being titrated to determine EC_{50} points, and (3) most potent and drug-like compounds being introduced to a cellular assay as a final evaluation. Using this protocol, I re-tested all of my compounds (Table 2) as well as began testing compounds for other lab members.

Disappointingly, the HbYX mimics were inactive. Nevertheless, active compounds were carried forward to obtain EC_{50} values. Once again, despite good data the values were rather flat and still had the appearance of on/off regulation. Having now been so involved in the evaluation of compounds, I proposed a change to the goals of the project (*vide infra*). To this point, the assays looked exclusively at the chymotrypsin-like activity of the proteasome (one of the active site selectivity's, see Chapter 1 for detail). This was done as the majority of the literature supported the idea that this site was the most important to target[150-153]. However, that conclusion was drawn by looking specifically at inhibition, not activation, and the endogenous 20S CP would be utilizing all three sites at once. Shouldn't the other sites be examined for their effect? Also, with a stable assay came stable EC_{50} values and with multiple compounds to draw on, a glaring issue with the EC_{50} metric was brought to the fore.



Cmpd	R	R1	A/In	Cmpd	R	R1	A/In
3-1		Cl	In	3-15			A
3-8			A	3-16			A
3-9		Cl	In*	3-19		Cl	A
3-10		Cl	In	3-21		Cl	A
3-11		Cl	In	3-28		Cl	In
3-13		H	A	3-29		Cl	In
3-14			A	3-30		Cl	A



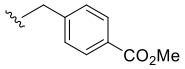
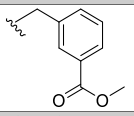
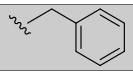
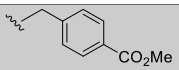
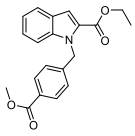
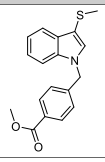
Cmpd	R	R1	A/In	Cmpd	R	R1	A/In
3-17		Cl	A	3-22		H	In
3-20		H	In	3-31		H	In
3-27			In	3-24			In

Table 3.3: Table of Screened Stimulator Compounds

As mentioned previously, an EC_{50} value is calculated from the dose-response curve and enables comparison of multiple compound's potency (Figure 3.17A). Notice, however, that the assumption is that all compounds involved (the three lines) are capable of producing a maximal effect.

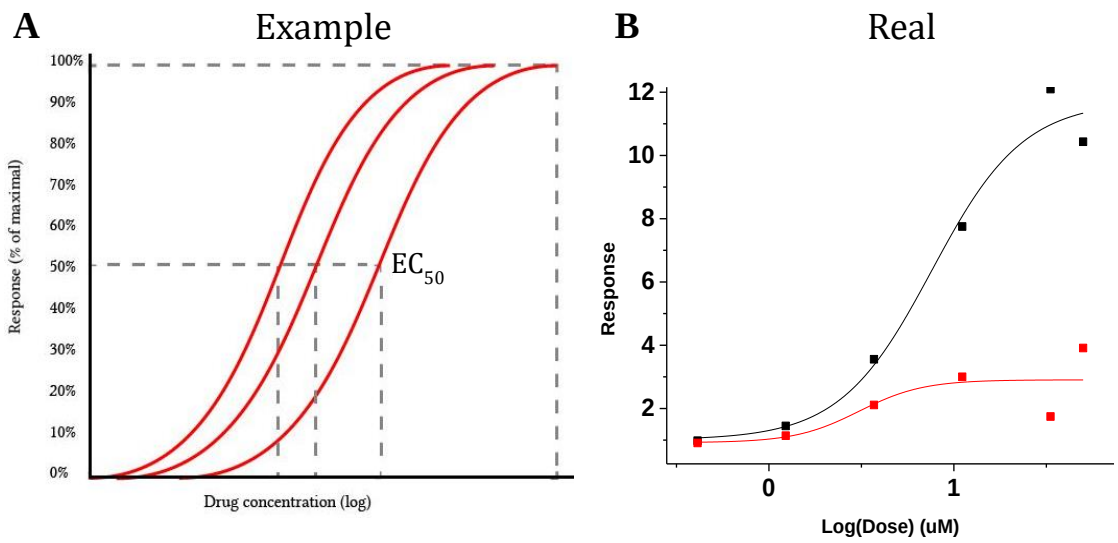


Figure 3.14: Illustration of EC_{50} Value Use (A) and Sample Project Data (B)

Unfortunately, project data (Figure 3.17B) demonstrates that this assumption can be misleading at times. According to the EC_{50} metric, the red line in the "Real" graph is superior to the black line. Methods exits to account for such partial agonism; however, they all rely on knowing what 100% agonism is. For proteasome activity, this is an unknown quantity. Over reliance on EC_{50} values was also beginning to show in the literature as compounds were beginning to be published with low micromolar EC_{50} values but producing less than 50% increase in proteasome activity[147-148, 154]. This gap in our knowledge base is the real barrier to improved compound activities, and we really needed to know if we should concern ourselves with the other active sites and what metrics translate to a biological system.

With so many unknowns, SAR could not realistically advance until more information on the "real world" effects could be gathered. As such, efforts were aimed at determining what factors translate to the protein level.

E. Construction of New Metrics for Agonism Effectiveness

First, I checked our lead compound and chlorpromazine for activity in each of the three catalytic sites and the mixture (Figure 3.18).

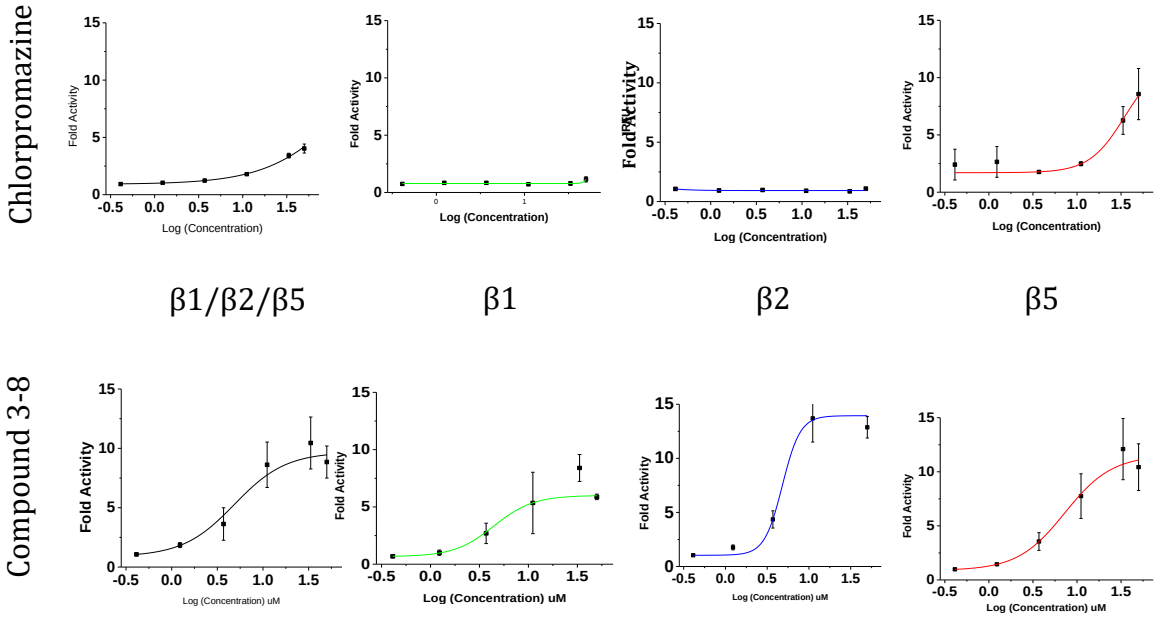
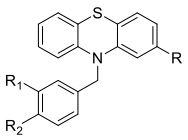


Figure 3.15: Comparison of Compound 3-8 and CPZ Activities per Active Site

The differences were stark. CPZ's activity is limited to only the chymotryptic-like site (β5) and shows no effect on any of the other sites (Note Evert Njomen tested CPZ for all three sites originally, and my data matches hers). Intriguingly, compound **3-8** produced the greatest effect on the trypsin-like site (β2) though had excellent activity in all sites. As the project goals had now changed to exploratory as opposed to performance oriented, only the phenothiazine compounds, along with some newly synthesized phenothiazine analogues, were subject to the screening protocol for each site and the mixture to give a pure data set (Table 3.4; note that inactive compounds are omitted). Additionally, as the lab was strongly opposed to interpret activity based on EC_{50} values, we decided to look at a different metric: the AC_{200}. AC_{200} is the concentration of drug that produces double the



				Caspase β1		Trypsin β2		Chymotrypsin β5		Combo 1:1:1	
Cmpd	R	R_1	R_2	AC_{200}	$Fold^M$	AC_{200}	$Fold^M$	AC_{200}	$Fold^M$	AC_{200}	$Fold^M$
3-8	Cl	H	CO_2CH_3	2.7	7	2.6	15	1.9	12	1.4	9
3-14	CF_3	H	CO_2CH_3	X	X	2.8	6	13.5	2	1.9	9
3-16	SEt	H	CO_2CH_3	X	X	2.2	3	X	X	1.4	3
3-13	H	H	CO_2CH_3	X	X	4.7	5	15.4	2	4.4	4
3-30	Cl	H	$CO_2(CH_3)_3$	X	X	X	X	X	X	X	X
3-21	Cl	CO_2CH_3	H	9.7	4	1.9	8	2.4	6	2.4	6
3-9	Cl	H	CO_2H	10	2	10	2	>30	2	10	2
3-31	Cl	CO_2H	H	X	X	X	X	X	X	X	X
3-32	Cl	NO_2	H	1.2	12	1.6	8	1.1	9	1.1	8
3-33	Cl	NH_2	H	3.8	10	9.2	10	6.5	9	3.6	8
3-34	Cl	OCH_3	H	1.3	4	5.4	2	X	X	1.5	4
3-19	Cl	H	H	X	X	3.1	2	2.8	3	1.5	2
CPZ	Chlorpromazine			X	X	X	X	7	>4	13.5	4

Table 3.4: Table of Individual and Combination Activities

activity of the enzyme. This artificially generates a "maximal" value to which each compound can be reliably compared. $Fold^M$ is another new metric that stands for maximal fold activity (i.e. the maximum effect this compound can produce is the listed fold over unactivated proteasome). Finally, significant differences in activity across multiple domains could be seen. Many different ranking orders can be offered by simple arranging compounds from best to least in a desired bracket (Figure 3.30). However, determining which one to follow would be decided by protein degradation studies. To that end, Evert Njomen took a selection of my compounds and used them to digest an IDP, α-synuclein (Figure 3.21).

	AC_{200}				Fold		
Combo	β1	β2	β5	Combo	β1	β2	β5
3-32	3-32	3-32	3-32	3-8	3-13	3-8	3-8
3-8	3-34	3-21	3-8	3-32	3-16	3-33	3-33
3-16	3-8	3-16	3-21	3-33	3-14	3-32	3-32
3-34	3-33	3-8	3-19	3-21	3-19	3-21	3-21
3-19	3-21	3-14	3-33	3-13	3-31	3-14	3-19
3-21	3-9	3-19	3-14	3-34	3-32	3-13	3-14
3-14	3-31	3-13	3-13	3-16	3-33	3-16	3-13
3-33	3-16	3-34	3-9	3-14	3-8	3-19	3-9
3-13	3-19	3-33	3-31	3-19	3-21	3-34	3-30
3-9	3-14	3-9	3-34	3-9	3-34	3-9	3-16
3-30	3-13	3-31	3-16	3-31	3-9	3-31	3-34
3-31	3-30	3-30	3-30	3-30	3-30	3-30	3-31

Table 3.5: All Top-to-Bottom Ranking Combinations

The results are quite interesting. From left to right in Figure 3.21, once again is synca control next to a vehicle control. We see that each active compound is indeed capable of enhancing synca degradation. The two standout observations are lanes 6, 7, and 8 holding compounds **3-13, 33,** and **32** respectively. Compounds **3-13** and **33** are able to completely remove synca while compound **3-32** appears to preferentially remove digested pieces first as they are absent in the 1 hour treatment (top image) and barely observable in the 2 hour treatment (bottom image), presumably this is because asyna has begun to be degraded again. Ranking these compounds on the basis of this degradation would likely be **3-33, 13, 8, 32/14** based only on the disappearance of synca. Which of course does not match any of the possibilities listed above (Figure 3.20). Note these results are bizarre as compound **3-13**, for example, is by every metric an inferior compound to compound **3-8**; however, it performs as well as if not better. The different site stimulation of each site compelled me to look deeper.

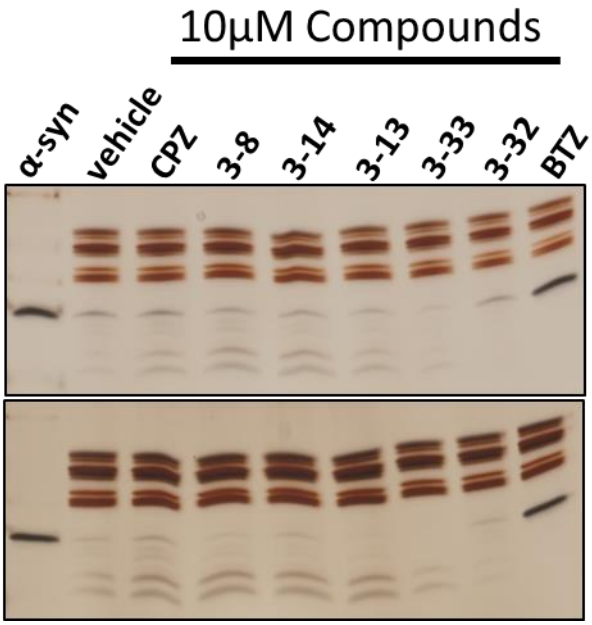
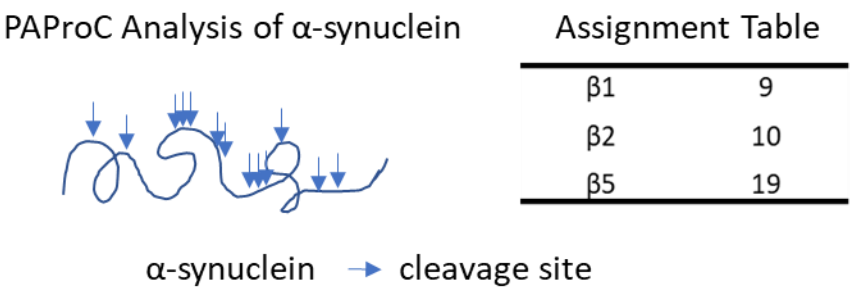


Figure 3.16: Selected Compounds Effect on Protein Degradation

Could the site selectivity's be translating to proteins and would that explain the somewhat bizarre results of the digestion? To investigate this, I took the gene sequence from the variant used in the above assay and subjected it to PAProC analysis. PAProC is an online server-based service which computationally predicts which sites will be cleaved by the human 20S proteasome. Inspecting these predicted sites, I assigned them a likely catalytic site. For example, prediction of cleavage at an acidic residue would be assigned β1, basic residues β2, and hydrophobic residues β5 in accordance with literature specificities[40] (Figure 3.22).

Intriguingly, there are very few β1 cleavage sites, only 9 total. The majority of cleavage sites are for the β5 site. Couple this to the known effect that site occupancy at the β1 inhibits β5 hydrolysis ability[155] and a compelling proposal can be made. If what has been presented so far is true, then to degrade synca (not its degradation products but the

protein itself) we would rank compounds with high β5 activity and low β1 activity as most useful (as high β1 activity would have an inverse effect on β5).



Assignment Table	
β1	9
β2	10
β5	19

Figure 3.18: Illustration of PAProC Analysis and Table of Assignents

This would explain the pronounced ability of **3-13** to degrade synca as it is unable to enhance β1 activity at any concentration (Figure 3.18). The resulting degradation products appear to be cleared best by overall activation of the proteasome, in which case the Fold^{M} activity of the combination (Figure 3.20, column 5) matches the best for clearance of these species.

F. Theory Update, Conclusion, and Future Outlook

Elated at the breakthrough we discovered in the previous section (i.e. how to effect change at the protein level), the obvious question remains, how can a small molecule produce such different effects?

The proteasome active sites contain a conserved catalytic triad of Asp17, Lys33, and Thr01 (Figure 3.19A). Proteolysis requires the concerted action of all three amino acids[156]. By examining recent crystallographic data[50, 75, 122, 157], I found that the catalytic sites undergo a range of changes varying from nearly identical to markedly different (Figure 3.19B). For the sake of illustration, assume the magenta form of the catalytic site shown is active and the yellow form is inactive.

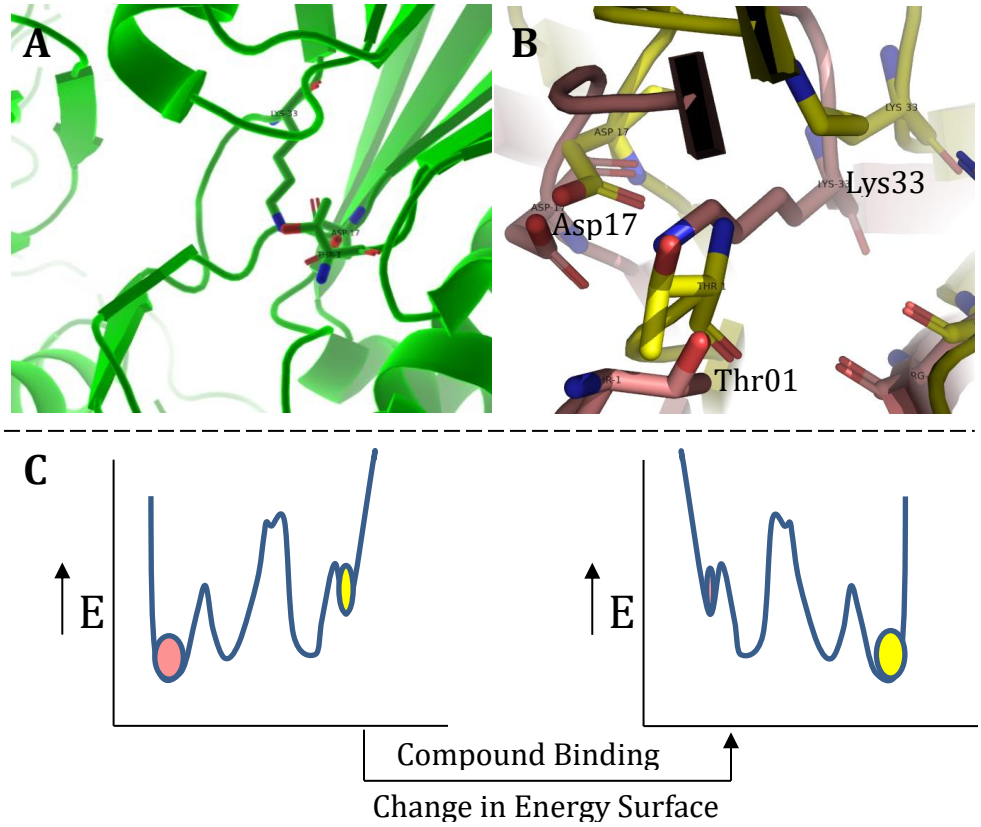


Figure 3.19: Illustration of Catalytic Site Conformers

These two forms likely exist in a dynamic equilibrium and our small molecules must be biasing/inducing these forms (i.e. active conformer of β5, inactive conformer of β2). Whether this mechanism is conformational trapping or induction is up for debate. Based on biochemical theory, conformational trapping is more likely. Fortuitously, solved crystal structures of the 20S CP possessing an open gate portion were recently published this year[75] and access to the powerful high-performance computational cluster (HPCC) here at MSU allowed high exhaustiveness docking to compare both the open and closed forms (Figure 3.20).

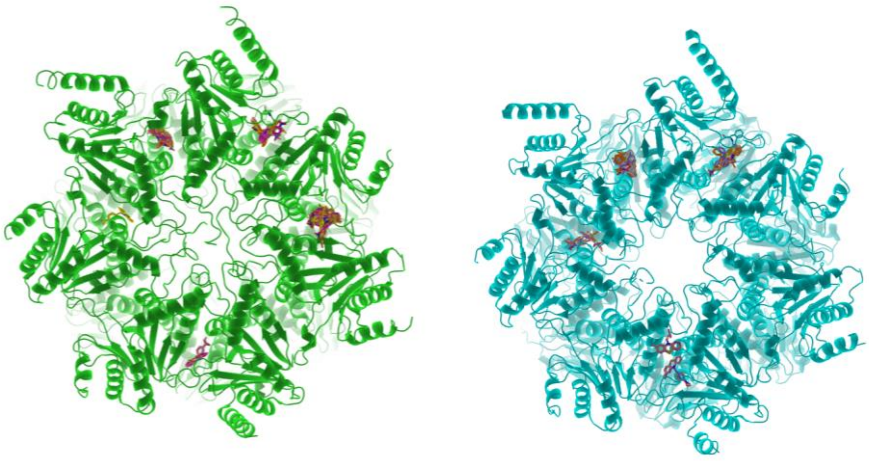


Figure 3.20: Pocket Preference Difference Among Different 20S Conformers

The remarkable change in binding pocket preference strongly suggests a mechanism of trapping an existing conformation and not binding followed by a conformational change (recall the original theory was pigeon-holed into this theory due to available data). Unfortunately, despite many forms of the crystal structures available with intersubunit pockets bound to Rpt tails (see Chapter 1), no investigation of key interactions has been done. Below I provide my own thoughts on what interactions may be present and suggestions for future directions of this project (Figure 3.21).

Excellent work published on 26S dynamics[41, 62, 67, 75, 158] allows for great understanding of substrate recognition, unfolding, and translocation events; however, due to pocket geometry, a variety of PA-CP interactions may be plausibly proposed resulting in limited mechanistic understanding of which residues are key to generate an open gate CP. Among the 11 states listed above only the C-termini of Rpt3 is generally unchanged whereas Rpt5 goes through minor conformational changes and Rpt1, 2, and 6 display high variability between states limiting us to only a few general observations.

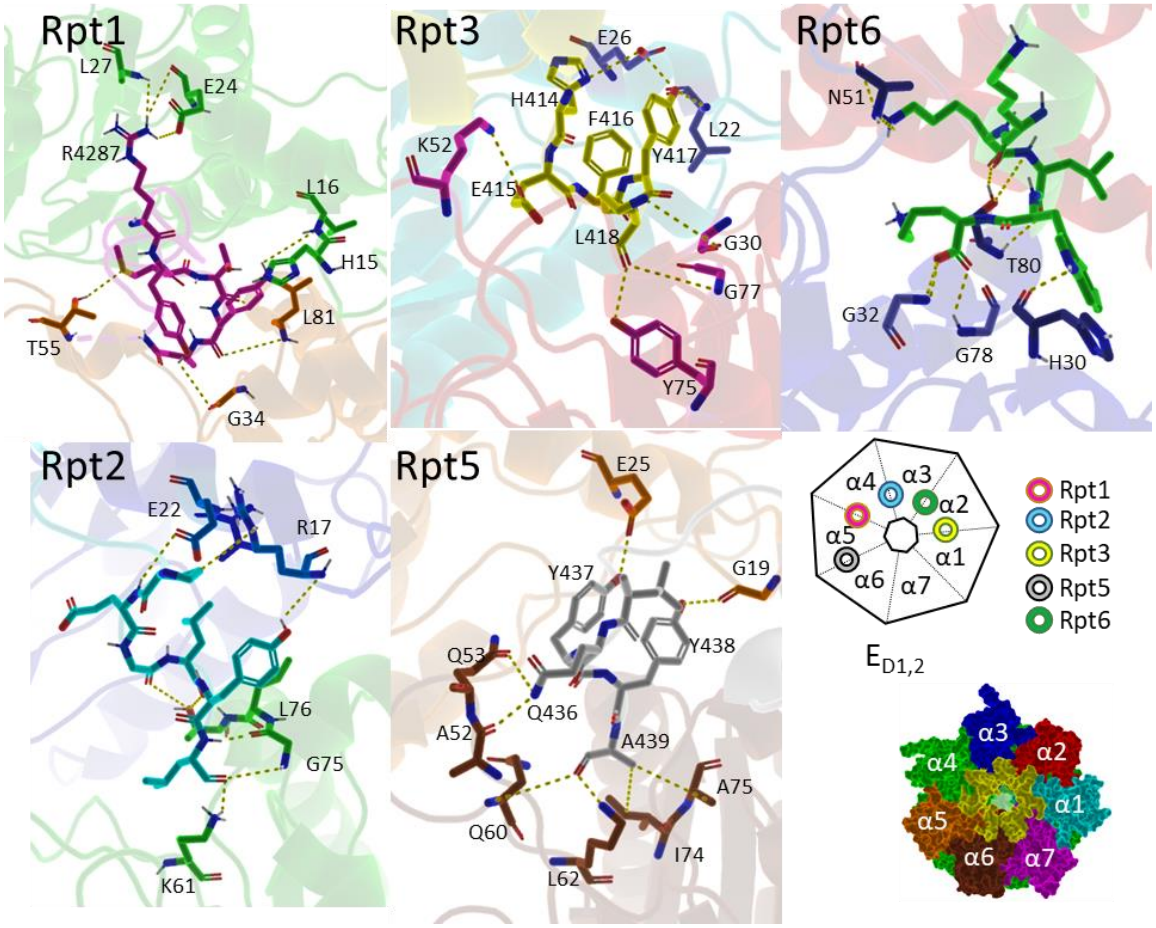


Figure 3.21: Intersubunit Pockets bound

Rpt3 appears to operate as an anchor for the 19S base as strong hydrogen bonding interactions are present at multiple points along the bottom of the α2 subunit and the front of the α6. These interactions are preserved across the majority of the other solved crystalline forms despite large conformational changes in the lid and minor changes in the alpha ring geometry. This binding dynamic is the same in the other HbYX containing Rpt2 and 5 (i.e. hydrogen bonding network along the back and bottom of one subunit and front of the adjacent subunit). Rpt5 displays minor conformational changes through the substrate processing process while Rpt2 possesses even more varied forms while preserving the number of contacts if not the same contacts.

However, Rpt1 and 6, lacking the HbYX motif, begin changing this commonality. Rpt1 binds most similarly to Rpt2, 3, and 5; however, it also bridges off to attack more centrally residing AAs within the α7 subunit by making a salt bridge with E26 and accepting a hydrogen bond from L27. In Rpt6, this multi-subunit binding breaks down completely as only minor hydrophobic interactions may be seen with the front of α2 as Rpt6 instead opts for numerous interactions with the α4 subunit. The α4 subunit is often cited as the most important subunit in gating as it possesses the greatest amount of electron density over the CP opening. Perhaps this extreme binding interaction with the Rpt6 tail is required to induce the conformational swing that moves the N-termini away from the CP opening; however, no clear indication of how this is achieved is currently available.

The above brief discussion provides a number of possible interactions and key residues utilized by numerous endogenous proteasome activators while also offering general comments about the apparent pocket geometries. I would suggest future endeavors utilize docking to provide a starting point for small molecule diversification, as was done at the beginning of this project. With the new metrics provided, reliable analogue differentiation can be achieved and progress on small molecule proteasome activators should increase. It is also strongly suggested that molecular dynamics be pursued as potential binding modes can be checked on their effect on the proteasome gate computationally. This would be a key contribution as intersubunit binding is most likely based on current information; however, this does not preclude the possibility of a β-ring/small molecule event and a computational simulation demonstrating stabilization of an open gate conformer would enable swift identification of necessary small molecule structural features.

This work has illustrated how complicated proteasome activation can be and the huge amount of work necessary in the future to elucidate more potent activators. Clearly, potency, in the case of activation, is extremely case dependent and a general activator may be untenable. Scientifically, this is the best-case scenario as it implies selective targeting of IDPs is possible based on their degradation sites and fragment products. Practically speaking, this is the worst-case scenario since knowledge of all metrics provided will be necessary for continued progress, as the current data suggests there is no-one-size-fit-all metric to gauge progress.

Chapter 4. Efforts Toward Small Molecule Binding Site Identification

Concomitantly with investigative efforts on factors that translate to the protein level, a new effort aimed at identification of the proteasome binding site/sites was initialized. Diazirine photoaffinity labels are often used for identification of protein-ligand binding pockets[159]. The diazirine is installed in an accommodating branch of the desired ligand[160]. Light of an appropriate wavelength (variable) irradiates the ligand-enzyme mixture generating a carbene in a bound pocket. The carbene then inserts into a proximate residue generating a covalently "labeled" system[161-162]. Since the benzyl tail of the phenothiazine compounds had thus far shown remarkable insensitivity to substitution, I decided to install the photo-labile group in the para position and replace the methyl ester (Figure 4.1,**4-A**).

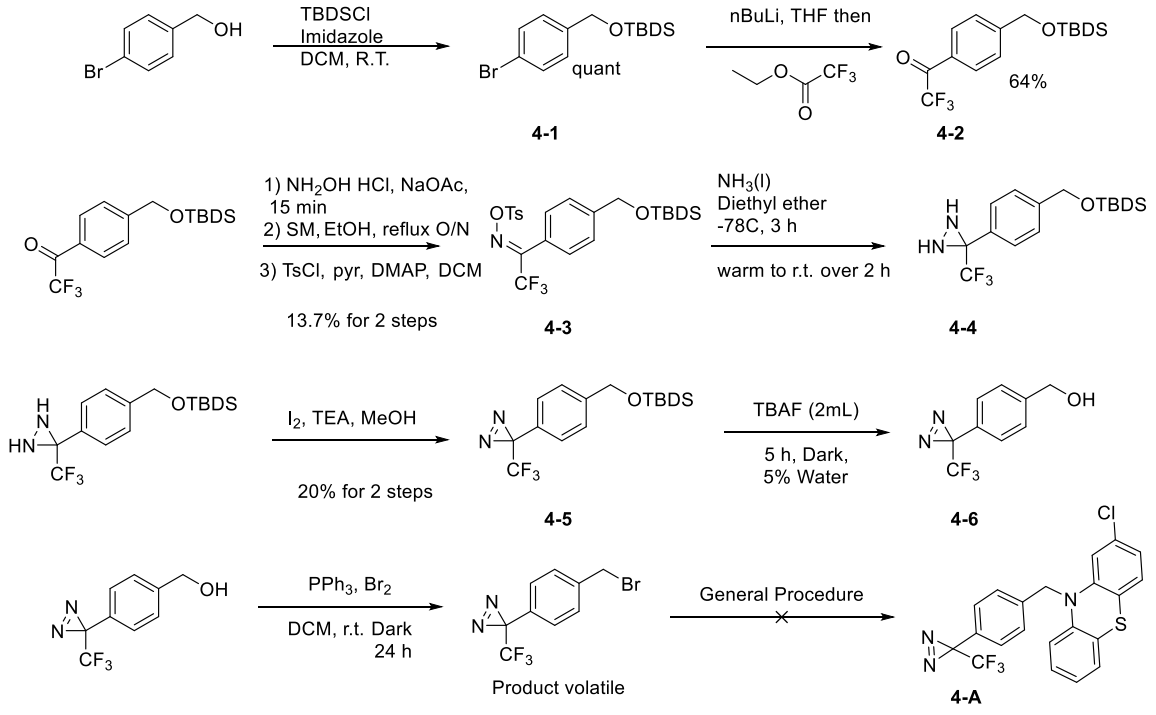


Figure 4.1: Synthesis of Phenothiazine Photoaffinity Label.

Synthesis of the benzyl tail proceeded following a literature precedence[162]. Due to a prior group member's experience with a similar pathway, the sequence was begun at 50

mmol scale with protection of the commercially available starting material to give protected alcohol **4-1**. Lithium/halogen exchange and quench with ethyl trifluoroacetate proceeded smoothly to afford **4-2**. The trifluromethy group offered a convenient handle through which to track reaction progress via fluorine NMR. Imine formation followed tosyl protection yielded the O-tosyl protected hydroxyl imine in moderate yield (**4-3**). Condensation of ammonia enabled nucleophilic displacement of the weakened N-O bond to give diaziridine in **4-4**. This compound was oxidized with iodide in methanol to yield the diazirine **4-5**. After this, great care was taken to ensure reactions would take place in darkness. TBDS deprotection with TBAF gave the alcohol **4-6** in excellent yield. The final two steps were performed in a one-pot two step procedure. Crude spectral data indicated the presence of desired end product. However, during purification, the diazirine decomposed (confirmed by FNMR). During repetition of the route, further research and discussion with the metabolomics cores on how identification of a binding site is done prompted abandonment of the diazirine target.

Binding site identification is ultimately done via computational search. A simulated fragmentation is performed generating a list of possible fragment products (Figure 4.2, red dots). The covalently bound enzyme (20S CP in this case) will be degraded into smaller fragments using a peptidase producing thousands of fragments (Figure 4.2, blue dots) and analyzed via mass spectrometry. The two fragment libraries generated are then compared to each other as a computer searches for fragmentation patterns that match the prediction. Clearly, as the size of the protein increases, the number of fragments that have to be checked also increases.

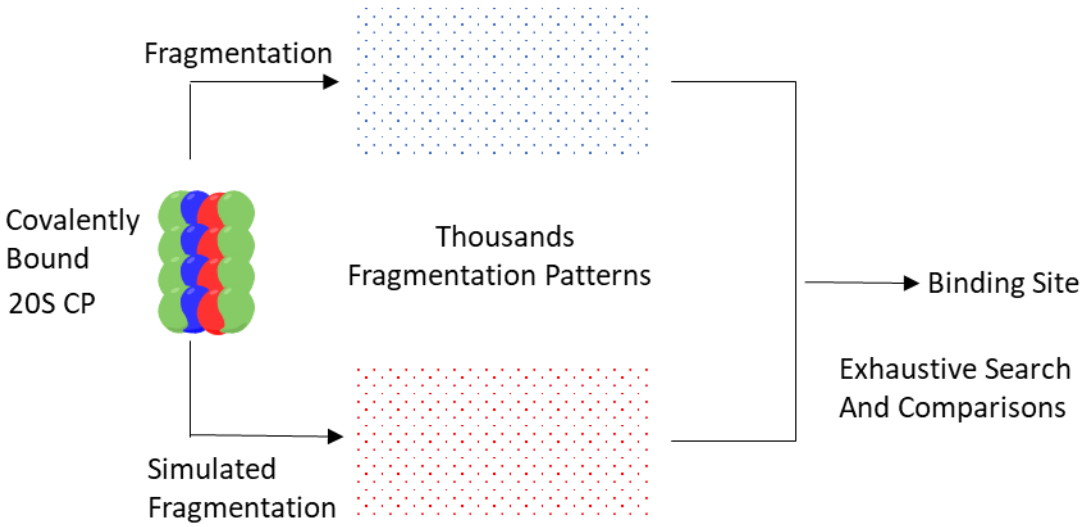


Figure 4.2: Diagram of the Binding Site Identification Process

However, what was non obvious is the nuance of fragment searching. A protein sequence, call it ABCD, would weigh the same regardless of which amino acid (AA) it is bound to. Identifying a covalent modification requires looking at the sequence *and* its fragments (Figure 4.3). Now a larger protein will have larger fragments (i.e. ABCDEFGHIJ, etc), requiring more patterns to be predicted and checked for. This problem is exacerbated if no knowledge is available to limit search parameters.

If Bound to A: 14 possible patterns:

A'BCD	BA'CD	CA'BD	DA'BC
A'BC	BA'C	CA'B	DA'B
A'B	BA'	CA'	DA'
	A'CD	A'BD	A'BC

- But…it could be bound to B or C or D.
- 56 total patterns for 4 AA only.
- The proteasome has ~6,286 AAs.
- Need to know *something* about the binding site.

Figure 4.3: Example of Combinatorial Explosion

In our example with ABCD, a covalent modification to a single AA requires comparison and identification of >56 possible patterns. However, if we know it is binding

to A only, this is reduced to 14 patterns. The 20S CP has >6000 amino acids and countless fragmentation possibilities. As a photoaffinity label can, in principle, bind to any AA, every possibility would have to be checked. Docking models can make suggested amino acid targets; however, we feel this to be a disingenuous protocol to follow. Despite good evidence we are binding in an intersubunit pocket, we still cannot rule out other allosteric binding sites and this would artificially limit possibilities. However, I was advised by the metabolomics core that, without some way to limit the possibilities, identifying the binding site would be extremely unlikely.

Several review articles discuss selective peptide modification[163-164] targeting different AAs in a variety ways. However, as docking did suggest binding in an intersubunit pocket, we elected to target a lysine residue. Lysine appears in every pocket at multiple points and we believed this would give us the best chance of covalently binding in a pocket. A brute force synthesis was undertaken to test the validity this idea (Figure 4.4).

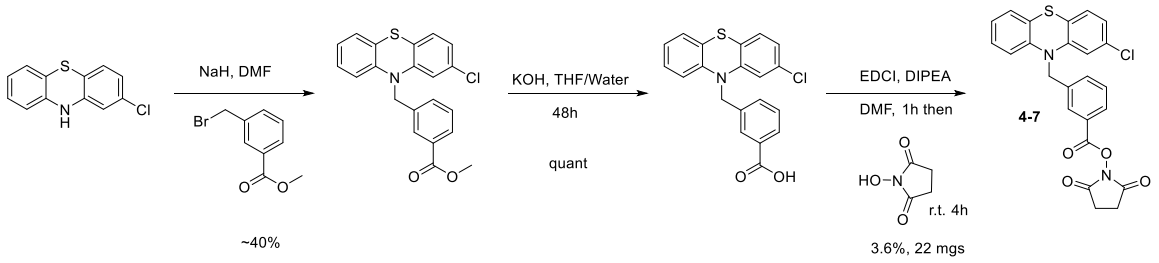


Figure 4.4: Synthesis of Lysine Targeting Substrate

The phenothiazine core was benzylated and hydrolyzed as detailed in Chapter 3. The N-hydroxysuccinamide (NHC) derivate was made using peptide-coupling reagent EDCI to give the lysine targeting compound **4-7**. Compound **4-7** was incubated with human proteasome for 1 hour at 37 °C before being frozen in a -80 °C freezer for an additional hour. This was taken to the metabolomics core where it was thawed, digested with trypsin

(i.e. the 20S complex was broken into many peptide fragments), and subjected to mass spectrometry and computational analysis (Figure 4.5).

Probability Legend:
over 95%
80% to 94%
50% to 79%
20% to 49%
0% to 19%

#	Visible?	Starred?	MS/MS View: 51 Proteins in 39 Clusters	Accession Number	Molecular Weight	Protein Grouping Ambiguity	Tepe ... Jones .temp (F012624)
1	☑	☆	Cluster of Keratin, type II cytoskeletal 1 OS=Homo sapiens GN=KR...	sp|P04264|K2C1_HU...	66 kDa	★	180
2	☑	☆	Proteasome subunit beta type-5 OS=Homo sapiens GN=PSMB5 PE...	sp|P28074|PSB5_HU...	28 kDa		139
3	☑	☆	Spectrin alpha chain, erythrocytic 1 OS=Homo sapiens GN=SPTA1 ...	sp|P02549|SPTA1_HU...	280 kDa		86
4	☑	☆	Proteasome subunit alpha type-6 OS=Homo sapiens GN=PSMA6 P...	sp|P60900|PSA6_HU...	27 kDa		84
5	☑	☆	Proteasome subunit alpha type-3 OS=Homo sapiens GN=PSMA3 P...	sp|P25788|PSA3_HU...	28 kDa		78
6	☑	☆	Proteasome subunit alpha type-1 OS=Homo sapiens GN=PSMA1 P...	sp|P25786|PSA1_HU...	30 kDa		72
7	☑	☆	Cluster of Keratin, type I cytoskeletal 10 OS=Homo sapiens GN=K...	sp|P13645|K1C10_HU...	59 kDa	★	71
8	☑	☆	Proteasome subunit alpha type-7 OS=Homo sapiens GN=PSMA7 P...	sp|O14818|PSA7_HU...	28 kDa		63
9	☑	☆	Proteasome subunit alpha type-2 OS=Homo sapiens GN=PSMA2 P...	sp|P25787|PSA2_HU...	26 kDa		45
10	☑	☆	sp|K1C9_HUMAN|	sp|K1C9_HUMAN| (+1)	62 kDa		47
11	☑	☆	Proteasome subunit beta type-1 OS=Homo sapiens GN=PSMB1 PE...	sp|P20618|PSB1_HU...	26 kDa		32
12	☑	☆	Proteasome subunit alpha type-5 OS=Homo sapiens GN=PSMA5 P...	sp|P28066|PSA5_HU...	26 kDa		31
13	☑	☆	Drebrin-like protein OS=Homo sapiens GN=DBNL PE=1 SV=1	sp|Q9UJU6|DBNL_HU...	48 kDa		31
14	☑	☆	sp|TRYP_PIG|	sp|TRYP_PIG|	24 kDa		32
15	☑	☆	Cluster of Spectrin beta chain, erythrocytic OS=Homo sapiens GN...	sp|P11277|SPTB1_HU...	246 kDa	★	30
16	☑	☆	Proteasome subunit beta type-6 OS=Homo sapiens GN=PSMB6 PE...	sp|P28072|PSB6_HU...	25 kDa		29
17	☑	☆	Proteasome subunit beta type-2 OS=Homo sapiens GN=PSMB2 PE...	sp|P49721|PSB2_HU...	23 kDa		27
18	☑	☆	Proteasome subunit beta type-4 OS=Homo sapiens GN=PSMB4 PE...	sp|P28070|PSB4_HU...	29 kDa		23
19	☑	☆	Cluster of Keratin, type I cytoskeletal 14 OS=Homo sapiens GN=K...	sp|P02533|K1C14_HU...	52 kDa	★	29
20	☑	☆	Proteasome subunit beta type-7 OS=Homo sapiens GN=PSMB7 PE...	sp|Q99436|PSB7_HU...	30 kDa		19
21	☑	☆	Proteasome subunit alpha type-4 OS=Homo sapiens GN=PSMA4 P...	sp|P25789|PSA4_HU...	29 kDa		15
22	☑	☆	Cluster of Heat shock protein HSP 90-alpha OS=Homo sapiens GN...	sp|P07900|HS90A_H...	85 kDa	★	11
23	☑	☆	Proteasome subunit beta type-3 OS=Homo sapiens GN=PSMB3 PE...	sp|P49720|PSB3_HU...	23 kDa		7
24	☑	☆	Proteasome subunit beta type-8 OS=Homo sapiens GN=PSMB8 PE...	sp|P28062|PSB8_HU...	30 kDa		6
25	☑	☆	Dermcidin OS=Homo sapiens GN=DCD PE=1 SV=2	sp|P81605|DCD_HUM...	11 kDa		5
26	☑	☆	Keratinocyte proline-rich protein OS=Homo sapiens GN=KPRP PE=...	sp|Q5T749|KPRP_HU...	64 kDa		5
27	☑	☆	Hornerin OS=Homo sapiens GN=HRNR PE=1 SV=2	sp|Q86YZ3|HORN_HU...	282 kDa		5
28	☑	☆	Proteasome subunit beta type-10 OS=Homo sapiens GN=PSMB10 ...	sp|P40306|PSB10_HU...	29 kDa		5
29	☑	☆	sp|HBB_HUMAN|	sp|HBB_HUMAN| (+14)	16 kDa		4
30	☑	☆	Junction plakoglobin OS=Homo sapiens GN=JUP PE=1 SV=3	sp|P14923|PLAK_HU...	82 kDa		4
31	☑	☆	Proteasome subunit beta type-9 OS=Homo sapiens GN=PSMB9 PE...	sp|P28065|PSB9_HU...	23 kDa		3
32	☑	☆	COP9 signalosome complex subunit 8 OS=Homo sapiens GN=COPS...	sp|Q99627|CSN8_HU...	23 kDa		3
33	☑	☆	Polyubiquitin-B OS=Homo sapiens GN=UBB PE=1 SV=1	sp|P0CG47|UBB_HUM...	26 kDa		3
34	☑	☆	sp|PRDX1_HUMAN|	sp|PRDX1_HUMAN| (...	22 kDa		3
35	☑	☆	COP9 signalosome complex subunit 2 OS=Homo sapiens GN=COPS...	sp|P61201|CSN2_HU...	52 kDa		2
36	☑	☆	TBC1 domain family member 17 OS=Homo sapiens GN=TBC1D17 P...	sp|Q9HA65|TBC17_H...	73 kDa		2
37	☑	☆	Desmoplakin OS=Homo sapiens GN=DSP PE=1 SV=3	sp|P15924|DESP_HU...	332 kDa		2

Table 4.1: List of Identified Fragments

Fifty-one proteins were identified with <1% chance of being misidentified, after deconvolution (see Experimental for Detail), and a bound fragment was discovered near the C-terminal end of α1 subunit attached to Lys242 (Figure 4.6)!

Valid	Weight	Sequence	Prob	Masc...	Masc...	Masc...	X! Ta...	...	Modifications
☑	1.0	(R)KAQPAQPADEPAEK(A)	100%	12.9	37.2	12.9	10.96	2	CJones_mod (+349)
☑	1.0	(R)KAQPAQPADEPAEKADEPMEH(-)	100%	25.0	33.6	25.0	6.92	2	CJones_mod (+349)

Table 4.2: Identified Bound Fragment

Our initial euphoria diminished after it was discovered this linkage is most likely spurious. The C-terminus of the α1 subunit is itself intrinsically disordered and juts out into the surrounding space, away from the proteasome. This area is so disordered, it does not appear in the proteasome crystal structures. We therefore believe this binding to be solution phase chemistry and not a potential binding mode. Additionally, the fragmentation of the proteasome resulted in a surprisingly low "coverage". A protein fragment is expected/predicted to fragment further a number of ways (i.e. from one-end to the other and *vice-versa*). Low coverage means a large number of these expected fragments were not found and could be one reason why only this spurious, exterior binding site was found. The low coverage most likely results from incomplete digestion by trypsin (according to Dr. Whitman).

At the time of writing, optimism is high as this is the first compound from the lab to be identified and investigation is ongoing on how a more likely binding site can be determined.

Chapter 5. Allosteric Inhibition of the 20S Proteasome

A. Design and Synthesis of an Allosteric Proteasome Inhibitor

Antineoplastic activity, through proteasome inhibition, has made the proteasome itself a high value target in the treatment of certain cancers[153]. Velcade™ (Figure 5.1A), an FDA approved drug, is used in the treatment of multiple myeloma (MM), mantel cell lymphoma, and acute allograft rejection[150]. Efficacy is achieved via the formation of a covalent bond of the electrophilic head group of Velcade™ to a threonine residue within the active site of the proteasome[150] (Figure 5.1B and C).

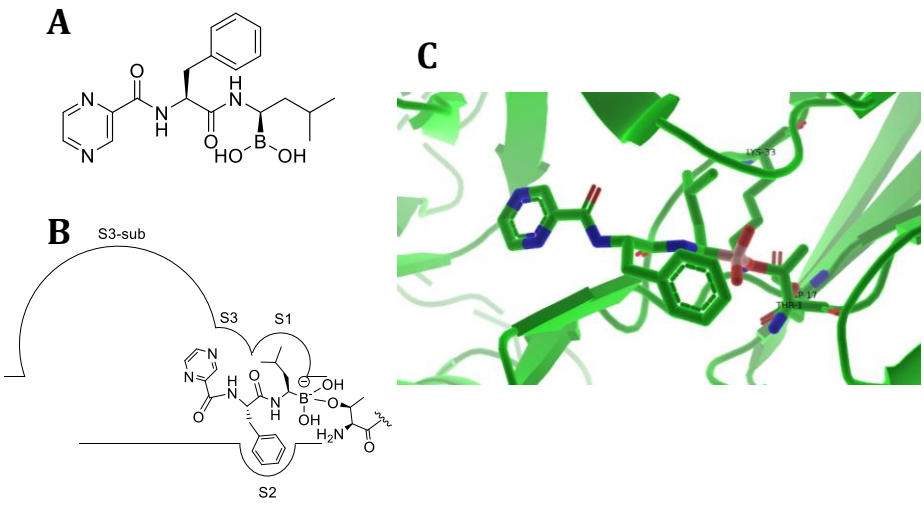


Figure 5.1: Velcade (A) Shown Bound in Catalytic Site Cartoon (B) and Crystal Structure (C)

This action halts the proteasome's catalytic activity, proteolytic degradation of proteins, required to maintain cellular homeostasis[165]. The excellent response of MM cells to this tactic is attributed to the prolific output of proteins by these cells, inflicting considerable stress on the degradation pathways within. Derailment of the proteasome degradation pathway results in apoptosis via the unfolded protein response[105]. The success of Velcade™ has launched many imitators with all current pipeline drugs operating by the

same competitive mechanism[166]. However, this binding mechanism has led to these agents exhibiting permanent abrogation of global protein degradation, lack of specificity, low systemic distribution, resistance, and severe off-target effects[150]. As a consequence, greater than 97% of patients become intolerant or resistant to treatment[167]. These severe side-effects and abysmal prognosis indicate the strong need for mechanistically distinct inhibitors.

Dr. Hewlett, of the Tepe lab, synthesized[168] a natural product analogue that was discovered[169] to inhibit the proteasome through an unprecedented binding mode[170]. This exclusive and non-covalent binding mode *broke with the dogma of proteasome inhibition*[171-172]. Various interactions are available; however, most striking is the hydrogen bonding network made possible through the specific orientation of the 5-6 fused ring[173] (Figure 5.2A). It has been shown that structurally related compounds of this family, preeminent among them Palau'amine, also inhibit the human proteasome[174]; implicating the 6-5 guanidine ring as a potentially useful scaffold for more potent inhibitors.

The lab was interested in expanding this S3-sub domain binding interaction to another part of the catalytic site. To establish the validity of such an approach, an existing inhibitor anchored in another domain would have to be modified for such a purpose. Fortuitously, Groll and co-workers reported an ideal inhibitor, unique in being anchored in the S1 domain of the β5 catalytic site while also evoking the S3 and S3-subdomain, though not exclusively[171]. The optimized structure (Figure 5.2B) is an alkynyl hydroxyurea mated to a 3-substituted phenyl ether, all features identified as crucial for potency[151].

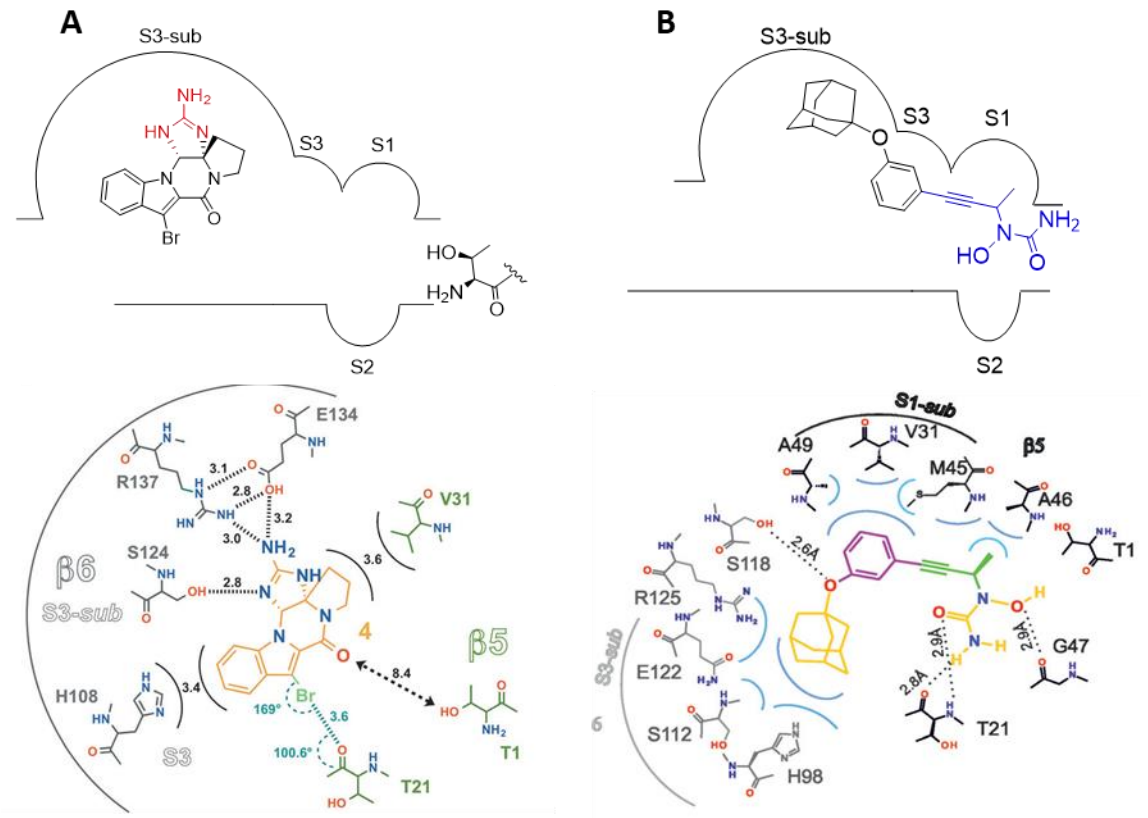


Figure 5.2: (A) Indolophakelin Bound in S3-Sub Domain (B) Hydroxyurea Inhibitor Bound in Active Site

Upon inspection, we believed replacement of the adamantyl group with a more hydrophilic motif would enable access to the hydrogen bonding network found in the S3-subdomain. Based upon PyMOL™ analysis (Figure 5.3A/B), a proof-of-concept structural analogue of the hydroxyurea inhibitor was proposed (Figure 5.3C). The goal would be to demonstrate the utility of interaction with the S3-sub domain as an anchor point from which a structurally distinct compound could be built (i.e. replacement of the hydroxurea alkyne motif).

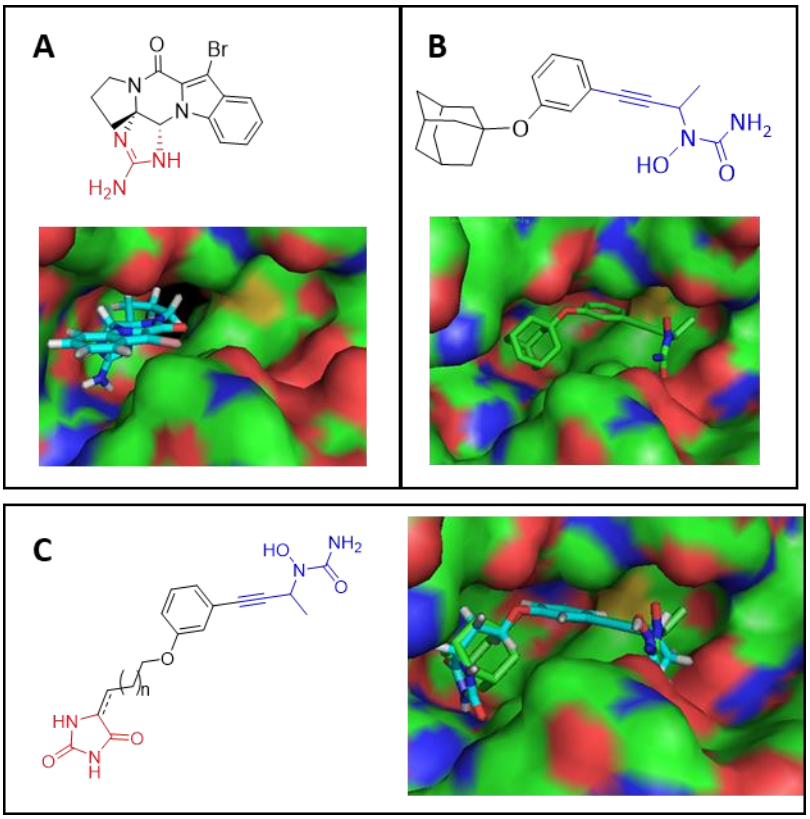


Figure 5.3: Indolophakelin (A) and Hydroxyurea (B) Bound in Catalytic Site. Overlay of Proposed Compound (C) with Hydroxyurea in Catalytic Site.

Retrosynthetic analysis is shown below (Figure 5.4). The general target is conveniently divided by Sonogashira coupling into two fragments. Fragment B is achievable through literature methods[151]. Fragment B could be obtained through either a Horner-Wadsworth Edmunds (HWE) condensation, if a double bond was preferred, or Bucherer-Bergs synthesis if the shorter carbon unit was required. Precursors to both are conveniently available through aldehyde **FB-1**. **FB-1** would be synthesized though oxidation of alcohol **FB-2** generated from alkylation of the commercially available m-iodophenol (**FB-3**).

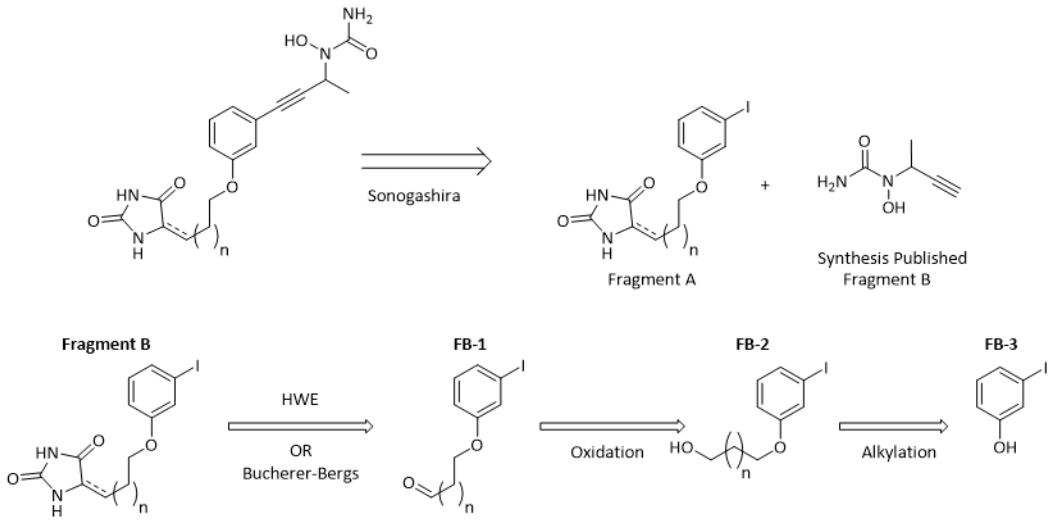


Figure 5.4: Generalized Retrosynthetic Analysis of Proposed Compound

In the forward case, synthesis of alkyne hydroxyurea **5-4** (Figure 5.5) went according to the literature method beginning with mesylation of propargyl alcohol (**5-1**). Displacement of the activated alcohol **5-2** with hydroxylamine followed by treatment with potassium cyanate and acid provided the desired fragment A (**5-4**) as a white solid.

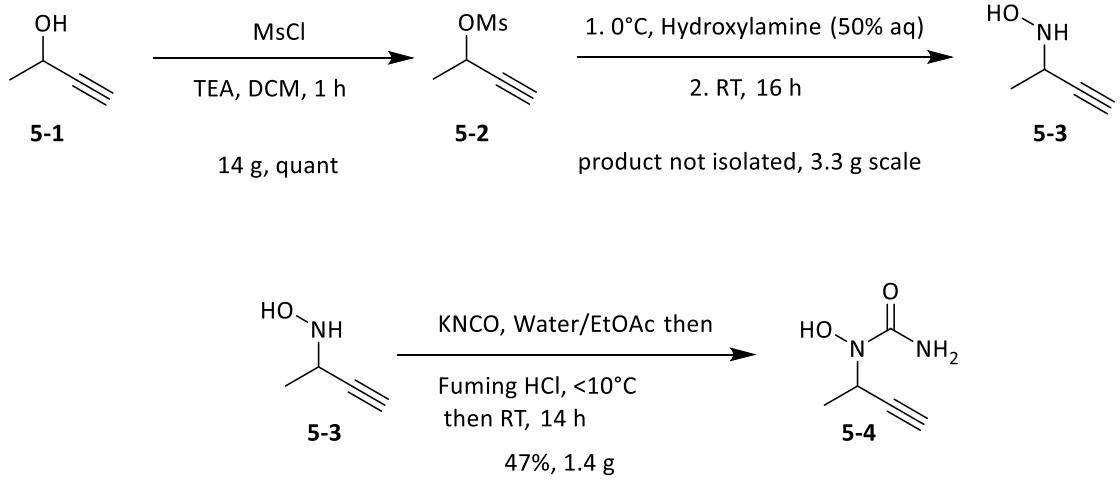


Figure 5.5: Synthesis of Fragment A (Compound **5-4**)

Synthesis of Fragment B and coupling to Fragment A is depicted below (Figure 5.6).

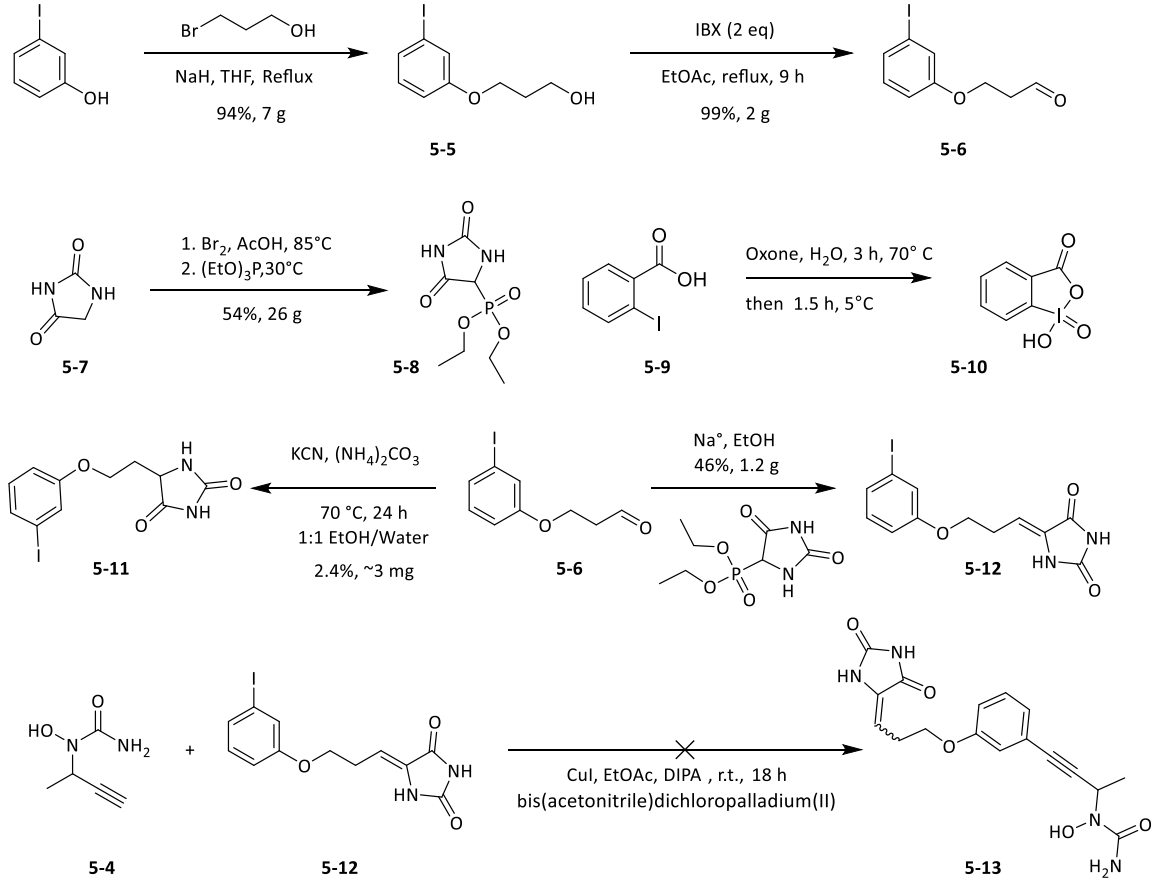


Figure 5.6: Synthesis of Allosteric Inhibitor **5-13**.

Alkylation of *m*-iodophenol with 3-bromo-propanol afforded the terminal alcohol **5-5** in good yield at gram scale. Oxidation to the aldehyde **5-6** was conveniently performed using 2-iodoxybenzoic acid (IBX, **5-10**) in refluxing ethyl acetate. IBX was generated using a literature procedure from 2-iodobenzoic acid **5-9**[175]. With the aldehyde **5-6** in hand, hydantoin **5-7** was converted to the HWE ylide precursor **5-8**. Compound **5-6** was then converted to the unsaturated coupling partner **5-12** following a literature HWE preparation and to the saturated hydantoin **5-11** using Bucherer-Bergs conditions. Unfortunately, the Bucherer-Bergs gave too little yield to carry forward. As such, only **5-12** was coupled with

hydroxyurea **5-4**. Unfortunately, the desired linker compound was not formed. Likely causes would be the high chelation potential of the hydantoin to either the copper or palladium species. A screen of different ligand sets would have to be undertaken to correct for this; however, the project was shortly hereafter ended.

Further screening was avoided because I began to look closer at the system. The crystal structures that the original analysis was based upon (Figure 5.2) were of yeast proteasome. The assay was performed with *human* proteasome. I copied the coordinates of the original hydroxurea inhibitor and copied them to the human proteasome map (Figure 5.7A, gold compound) for comparison purposes.

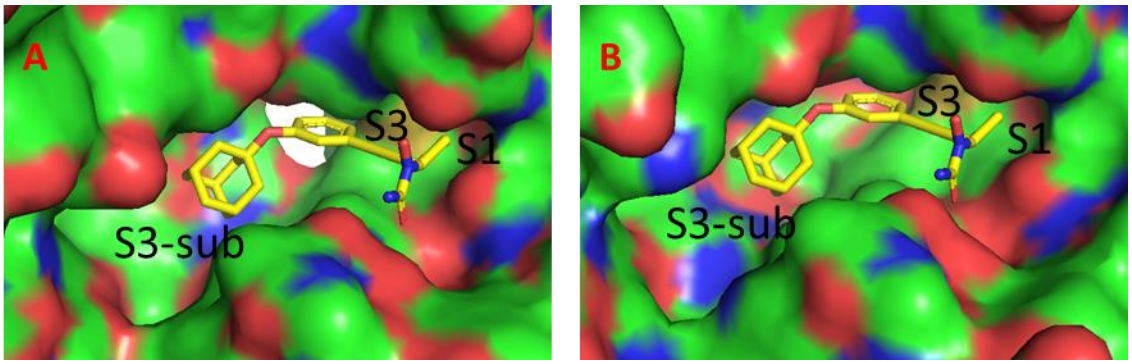


Figure 5.7: Electrostatic Potential Maps of β5 Catalytic Site in Human (A) and Yeast (B)

The differences in the two maps are most apparent in the S3-sub pocket. The yeast proteasome contains a generously positive surface (blue shading, Figure 5.8B), presumably from the all the basic amino acids our hydantoin was supposed to take advantage of. Unfortunately, this same area is more neutral (green shading) and even slightly negatively charged deeper in the pocket of the human proteasome (Figure 5.8A). By comparing their sequences against each other, I found there is very low sequence similarity between yeast and human proteasome, which would account for this binding difference and attribute

nearly all the activity to the right half of the molecule (note the remarkable similarity of the S1 pocket in Figure 5.7). If this were true, further effort would inevitable be in vain.

To answer this, I turned to computational docking methods. Once the docking workflow was established, I ran the proposed linker against human proteasome and found it had no predicted binding modes in the catalytic site and, shortly thereafter, used docking to form a framework for small molecule *activation* which pivoted my path away from proteasome inhibitors.

B. Computational Insight for New Allosteric Inhibitors

While I my personal projects moved away from the inhibition side of proteasome activities, I nonetheless still aided collaborators in their inhibitory endeavors. In collaboration with the Gaczynska group[176], which studies proteasome inhibition, a series of analogs based on SAR optimized motif (Figure 5.8, **B1**) was performed by Dr. Matthew Giletto (Tepe Group) to further their studies. Several of these published compounds[177] are depicted below (Figure 5.8). I was asked to rationalize a mechanism of action and explain the unusual behavior of the pipecolic acid enantiomers (Figure 5.8, **C, D,** and **E**). In most cases, the potency between two enantiomers is an order of magnitude (or more) or their individual activities sum to that of the racemate[178].

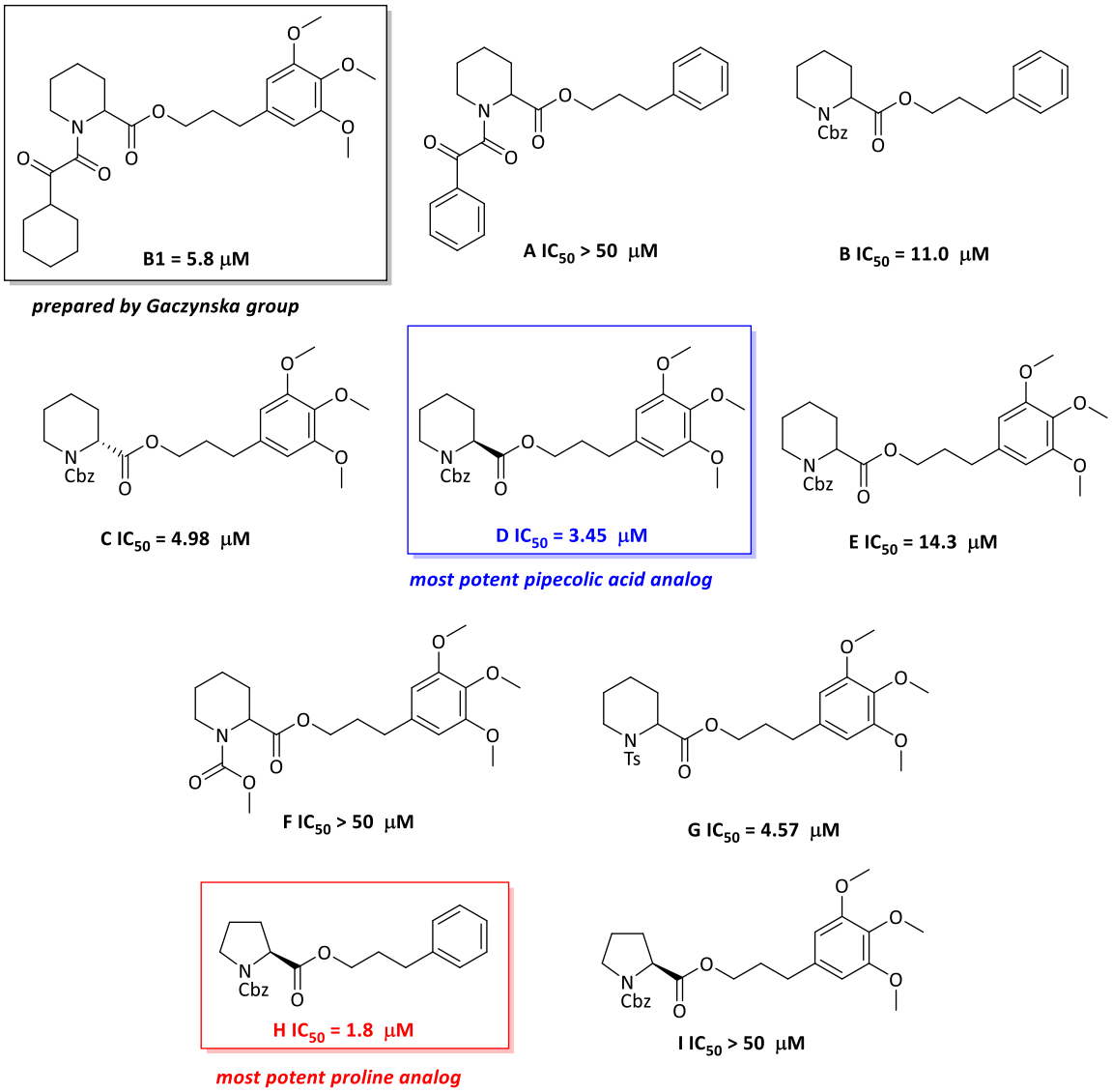


Figure 5.8: Published Structures Prepared by Dr. Giletto[177]

The large size of the human proteasome can make docking studies complicated due to the large search space and numerous possible binding pockets. Additionally, a completely exhaustive search is infeasible with the hardware available to the lab as the large number of rotatable bonds in the pipecolic ester structures increases the computational expense. To overcome these limitations, docking was undertaken in three parts:

1) Whole 20S proteasome

The entirety of the proteasome was searched with a low exhaustiveness (30-40) in multiple individual runs (on average ~5 times). This sampling revealed more binding modes within the alpha ring than the beta rings.

2) Alpha and Beta Rings

After (1), the search space was reduced, and exhaustiveness increased (~60). In two separate runs, each compound was docked against the beta rings, and then the alpha rings. The average binding affinity difference was ~1.2 kcal/mol suggesting a large preference for the alpha ring system.

3) Alpha Ring Site refinement

Finally, the search space was limited to the minimum space necessary to encompass just the alpha ring and as high an exhaustiveness as possible on our hardware:

Center: (135.9681, -38.0296, 65.3754)

Dimensions (Angstrom): X: 122.946 Y: 138.030 Z:53.184 Exhaustiveness: 80

Remarkably, the most potent compounds gathered in only a few of the intersubunit pockets (red circles, Figure 5.9). While, the less active compounds preferred either a different binding pocket or no preference.

This was an intriguing mechanism of action, intersubunit pocket binding for inhibition; however, after the crystallography data published this year[75] is accounted for, the result makes more sense[75].

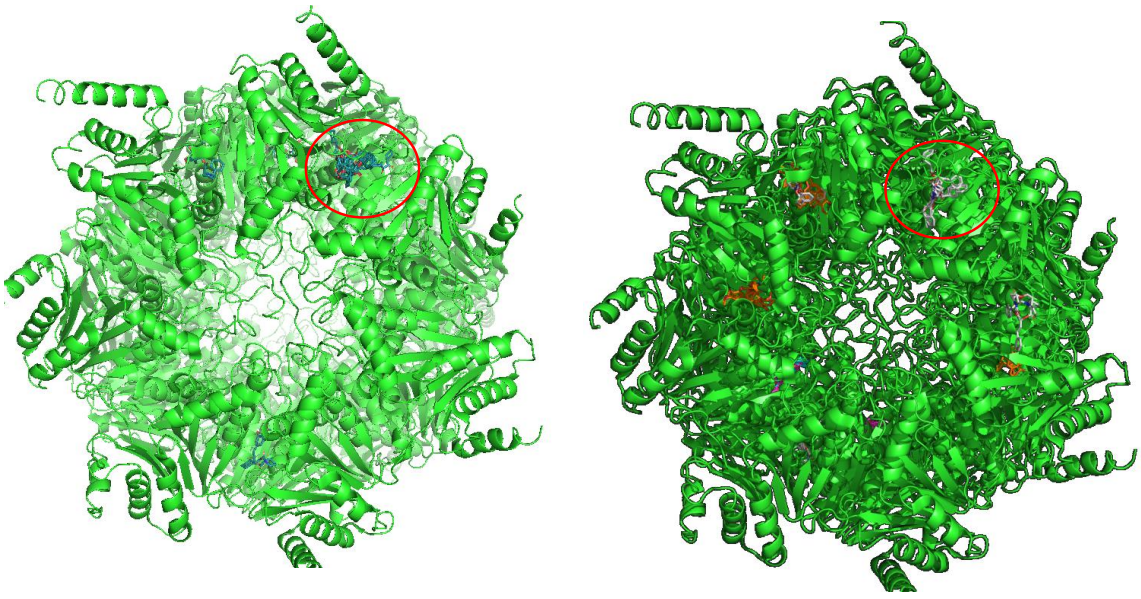


Figure 5.9: Binding Modes of Dr. Gilletto's Analogues.

The new data (discussed more in Chapter 1), demonstrated that 19S intersubunit binding could occur without inducing a gate opened form[75], implying other mechanisms of action for binders within this site. The deeply penetrating pipecolic analogous were therefore proposed to be inducing/trapping a non-active proteasome form through alpha ring conformer adjustment. The enantiomers bind to two different subunits of the alpha ring. An unusual result, but in the absence of other empirical data, a reasonable explanation for the observed results.

Experimental Section

A. Synthetic Methods

Reactions were carried out in flame-dried glassware under a nitrogen atmosphere. Reagents and solvents were purchased from commercial suppliers and used without further purification. Anhydrous THF was distilled over benzophenone and sodium immediately prior to use. All reactions were magnetically stirred. Yields refer to chromatographically and spectroscopically pure compounds unless otherwise noted. Infrared spectra, where applicable, were recorded on a JASCO Series 6600 FTIR spectrometer. ^{1}H and ^{13}C NMR spectra were recorded on a Varian Unity Plus-500 or 600 spectrometers, as noted in the experimental for each compound. Chemical shifts are reported relative to the residue peaks of the solvent ($CDCl_3$: 7.26 ppm for ^{1}H and 77.0 ppm for ^{13}C) (DMSO-d_6: 2.50 ppm for ^{1}H and 39.5 ppm for ^{13}C). The following abbreviations are used to denote the multiplicities: s = singlet, d = doublet, dd = doublet of doublets, t = triplet, and m = multiplet. Due to the hydrophobicity of a number of these compounds, small quantities of solvents were un-removable in some cases. These instances are labeled in accordance with literature precedent[179]. HRMS were obtained at the Mass Spectrometry Facility of Michigan State University with a Micromass Q-ToF Ultima API LC-MS/MS mass spectrometer. Purification of compounds was achieved in most cases using laboratory medium pressure liquid chromatogram (MPLC) on silica gel (20-40 microns). Standard method is gradient elution from 0-50% ethy acetate in hexanes over 45 minutes. Deviations from this will be noted per compound but were otherwise general. Attachment of substituted benzyl groups was performed through a general procedure (detailed below). Any deviations are listed where appropriate in the text of the compound.

General Benzylation Procedure:

Heterocyclic core (phenothiazine, iminodibenzyl, indole, etc) was dissolved in anhydrous DMF at room temperature under an inert atmosphere in a round bottom flask. Sodium hydride (1.1 eq) was added as a single portion, vigorous bubbling should be observed, and the mixture allowed to stir at room temperature for 0.5 h wrapped in foil. Substituted benzyl derivate (1.5 eq) is added as a single portion and allowed to react at room temperature in the dark (i.e. wrapped in foil) for 16 h. After 16 h, the reaction is diluted with ether (*ca.* 2 solvent volumes) and poured into separatory funnel containing a 10% wt/wt solution of LiBr in DI water. The ether layer is carefully washed 2x with LiBr (aq) and then Brine (1x solvent volume), dried over sodium sulfite and concentrated *in vacuo*. Crude material was then purified using MPLC standard methods.

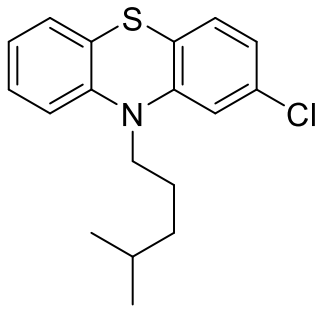


Figure E.1: Compound 3-1

2-Chloro-10-(4-methylpentyl)-10H-phenothiazine

A solution of 2-chloro-10H-phenothiazine (0.467 g, 2 mmol) in THF is added dropwise to a suspension of sodium hydride (60% wt/wt, 0.080 g, 2 mmol) at room temperature. The mixture is allowed to stir at room temp for 30 minutes. 1-Bromo-4-methyl pentane (0.146 mL, 1 mmol) was added neat, dropwise. After stirring for 2 hours, the solution was poured into saturated bicarbonate solution (*ca.* 50 mL) and extracted into ethyl acetate (3x 50 mL). The combined organic layers were washed with brine (*ca.* 50 mL) and dried over sodium sulfate and concentrated *in vacuo* to give a purple solid, which was purified via MPLC to give the final product as a white solid (85.1 mg, 26.8%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.20 – 7.10 (m, 2H), 7.02 (d, *J* = 8.2 Hz, 1H), 6.93 (td, *J* = 7.5, 1.1 Hz, 1H), 6.88 (dd, *J* = 8.1, 2.0 Hz, 1H), 6.85 (dd, *J* = 8.2, 1.1 Hz, 1H), 6.81 (d, *J* = 2.0 Hz, 1H), 3.78 (t, *J* = 7.2 Hz, 2H), 1.82-1.73 (m, 2H), 1.59-1.51 (m, 1H), 1.34-1.24 (m, 2H), 0.88 (d, *J* = 6.6 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 146.7, 144.7, 133.3, 128.0, 127.6, 127.5, 124.8, 123.5, 122.9, 122.2, 115.8, 115.8, 47.9, 36.2, 27.8, 24.8, 22.7. HRMS (ESI) *m/z*: $[M+H]^+$ Calcd for $C_{18}H_{22}ClNS$: 318.1083; Found 318.1082 ATIR: CH (2952 cm^{-1}), aromatic CH (3056 cm^{-1} and 3176 cm^{-1})

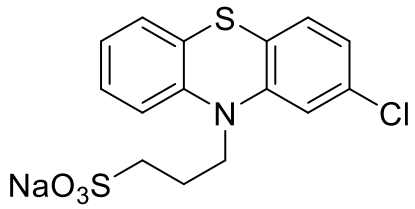


Figure E.2: Compound 3-2

3-(2-Chloro-10H-phenothiazin-10-yl)propane-1-sulfonate

2-chloro-10H-phenothiazine (0.981 g, 4.2 mmol) was added as a solution in anhydrous THF (10 mL), to a round bottom flask charged with sodium hydride (0.160 g, 4 mmol) and THF (15 mL). The solution was then heated to reflux for 1 hour to give a bright red solution, which was cooled to near room temperature before addition of 1,3-propane sultone (0.41 mL. 4 mmol). The solution immediately becomes yellow and forms a white precipitate. The solution was stirred for 1 hr at reflux. White solids formed upon cooling were washed with THF (100 mL) and diethyl ether (100 mL) before being left to dry in air overnight (985 mg, 65%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 7.21 (ddd, J = 8.6, 7.3, 1.6 Hz, 1H), 7.16 – 7.12 (m, 2H), 7.07 (dd, J = 8.3, 1.5 Hz, 2H), 7.01 – 6.93 (m, 2H), 3.99 (t, J = 7.2 Hz, 2H), 2.53 (t, J = 7.3 Hz, 2H), 1.96 (tt, J = 8.4, 6.5 Hz, 2H). ^{13}C NMR (126 MHz, (DMSO-d_6) δ 146.3, 144.0, 132.5, 128.1, 127.8, 127.2, 123.1, 122.9, 122.4, 122.1, 116.3, 115.7, 48.5, 45.7, 22.7. HRMS (ESI) *m/z*: $[M+H]^+$ Calcd for $C_{15}H_{115}ClNO_3S_2$ 356.0182; Found 356.0182. ATIR: Aromatic CH (3427 cm^{-1}), CH (2950 cm^{-1} , very weak) RSO_3^- (1049 cm^{-1})

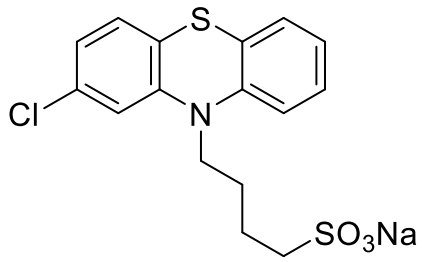


Figure E.3: Compound 3-3

4-(2-Chloro-10H-phenothiazin-10-yl)butane-1-sulfonate

2-chloro-10H-phenothiazine (3.5 g, 15 mmol) was added as a solution in anhydrous THF (10 mL), to a round bottom flask charged with sodium hydride (0.6 g, 15 mmol) and THF (15 mL). The solution was then heated to reflux for 1 hour to give a bright red solution, which was cooled to near room temperature and injected with 1,4-butane sultone (1.54 mL, 15 mmol). The solution was then refluxed for 24 hours. Upon cooling, title compound precipitated from solution as an off white solid (4.6 g, 78%) and may be used without further purification. Further purification can be achieved if desired by taking a portion of the compound and refluxing in benzene overnight (*ca.* 12 h) with a Dean-Stark trap. The benzene solution was then frozen, and the solvent sublimed off to give clean compound. ^{1}H NMR (500 MHz, (DMSO-d_6): δ 7.19 (t, J = 7.7 Hz, 1H), 7.15 – 7.10 (m, 2H), 7.06 – 7.01 (m, 2H), 6.99 – 6.93 (m, 2H), 3.84 (t, J = 6.7 Hz, 2H), 2.42 (t, J = 7.4 Hz, 2H), 1.74-1.62 (m, 4H). ^{13}C NMR (126 MHz, (DMSO-d_6): δ 146.4, 144.0, 132.5, 128.1, 127.8, 127.2, 123.2, 122.9, 122.5, 122.0, 116.3, 115.7, 50.9, 46.5, 25.5, 22.6. HRMS (ESI) *m/z*: $[M+H]^+$ Calcd for $C_{16}H_{17}ClNO_3S_2$ 370.0338; Found 370.0344. ATIR: Aromatic CH (3427 cm^{-1}), CH (2950 cm^{-1}, very weak), RSO_3^- (1049 cm^{-1})

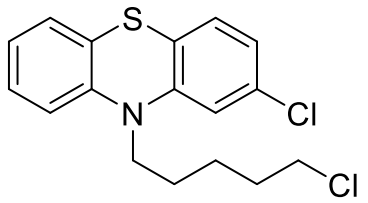


Figure E.4: Compound 3-4

2-Chloro-10-(5-chloropentyl)-10H-phenothiazine

To an oven dried round bottom flask charged with sodium hydride (600 mg, 15 mmol) was added 2-chloro phenothiazine (2.3 g, 10 mmol) as a solution in anhydrous THF (20 mL), dropwise at room temperature. An additional 20 mL of THF was added and the solution brought to reflux for one hour. The solution, now red to orange, was cooled to room temperature and 1-bromo-5-choropentane (1.58 mL, 12 mmol) was added in a single portion. The solution was allowed to stir at room temperature for 12 hours. The solution, now with brown solids, was poured into a separatory funnel containing an equivalent volume of sodium bicarbonate and extracted with diethyl ether (2x 50 mL). The organic layers were dried over sodium sulfate and concentrated to dryness before being applied to the MPLC for purification to give a dark oil as the final product (1.44 g, 42%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.24 - 7.13 (m, 2H), 7.04 (d, *J* = 8.2 Hz, 1H), 6.98 (td, *J* = 7.5, 1.1 Hz, 1H), 6.91 (dd, *J* = 8.2, 2.1 Hz, 1H), 6.88 - 6.83 (m, 2H), 3.82 (t, *J* = 7.0 Hz, 2H), 3.52 (t, *J* = 6.6 Hz, 2H), 1.88 - 1.72 (m, 4H), 1.66 - 1.49 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 146.5, 144.5, 133.2, 128.0, 127.6, 127.5, 125.0, 123.7, 123.0, 122.3, 115.9, 115.8, 47.2, 44.9, 32.2, 26.1, 24.3. This compound (18.5 mmol) was dissolved in acetone (150 mL). Finely ground sodium iodide (*ca.* 50 g, 333mmol) was added and the mixture vigorously stirred. The mixture was refluxed for 3 days. The mixture was then placed in a -20°C freezer

for 4 hours and then filtered through a medium frit. Solids were washed with acetone (2x100 mL) and the filtrate concentrated to dryness to give the product as a waxy brown solid in quantitative yield (7.6 g). The iodo-products were used without further purification for the preparation of compound **3-5**.

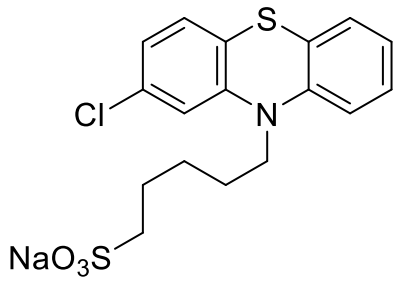


Figure E.5: Compound 3-5

5-(2-Chloro-10H-phenothiazin-10-yl)pentane-1-sulfonic acid

2-Chloro-10-(5-iodopentyl)-10H-phenothiazine (343 mg, 0.8 mmol) was added to a mixture of acetone and water (1:1) and sodium sulfite (403.2 mg, 3.2 mmol) in a round bottom flask giving a milky white solution. Mixture was allowed to stir overnight (ca. 12 hr) to give a clear yellow solution with a white precipitate. The solution is concentrated to dryness in vacuo and residue swirled with dichloromethane (30 mL) and solids collected on a medium frit. Solids washed with acetone (30 mL) to give the product as an off white solid. Solid is collected and dried overnight on a high vacuum line to give the title compound (207 mg, 66%). ^{13}C NMR (126 MHz, DMSO-d6) δ 146.4, 144.0, 132.5, 128.1, 127.8, 127.2, 123.3, 123.0, 122.6, 122.1, 116.3, 115.7, 51.4, 46.5, 26.1, 25.6, 24.9. ^{1}H NMR (500 MHz, DMSO-d6) δ 7.25 – 7.17 (m, 1H), 7.17 – 7.09 (m, 2H), 7.04 – 7.00 (m, 2H), 6.98 – 6.93 (m, 2H), 3.83 (t, J = 7.0 Hz, 2H), 2.42 – 2.26 (m, 2H), 1.65-1.60 (m, 2H), 1.58 – 1.46 (m, 2H), 1.42-1.36 (m, 2H). ATIR: OH (from hydrate, 3428 cm-1, broad), Aromatic CH (2930 cm-1), CH (2850 cm-1) HRMS (ESI) m/z: [M-H]- Calc'd for $C_{17}H_{17}ClNO_3S_2$ 382.0338; found: 382.0336.

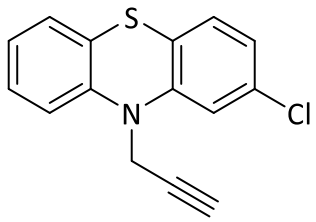


Figure E.6: Compound 3-6

2-Chloro-10-(prop-2-yn-1-yl)-10H-phenothiazine

2-Chloro-10H-phenothiazine (0.467 g, 2 mmol) was added as a solution in anhydrous THF (*ca.* 10 mL), to a round bottom flask charged with sodium hydride (0.076 g, 1.9 mmol). The solution was stirred one hour to give a reddish-brown solution before the addition of propargyl bromide (80% in Toluene, 0.24 mL, 2.2 mmol) in a single portion. The solution was stirred for 12 hours and then concentrated to dryness *in vacuo* before being placed on a high vacuum line for approximately 4 hours. Crude material was suspended in chloroform and cooled to -20°C overnight. The starting material precipitates out and was removed by filtration. The filtrate was concentrated to dryness to give thick, dark oil as product (0.300 g, 58.1%) which was used in the next step without further purification. ^{13}C NMR (126 MHz, cdcl$_3$) δ 145.4, 143.7, 133.5, 127.8, 127.6, 127.2, 123.5, 123.2, 122.9, 122.1, 115.3, 115.2, 78.6, 75.1, 38.8. ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.23−7.18 (m, 3H), 7.13 (dd, J = 7.6, 1.4 Hz, 1H), 7.02 (d, J = 8.2 Hz, 1H), 6.98 (ddd, J = 7.8, 6.7, 2.0 Hz, 1H), 6.93 (dd, J = 8.2, 2.0 Hz, 1H), 4.48 (d, J = 2.4 Hz, 2H), 2.51 (t, J = 2.4 Hz, 1H). HRMS (ESI) *m/z*: [M]$^{\cdot}$- Calc’ d for $C_{15}H_{10}ClNS$ 271.0222; Found 271.0228.

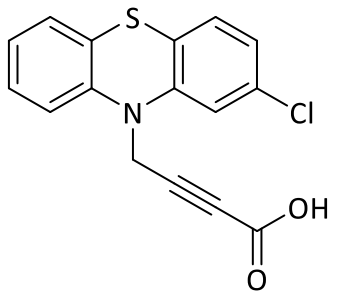


Figure E.7: Compound 3-7

4-(2-Chloro-10H-phenothiazin-10-yl) but-2-ynoic acid

A solution of 2-chloro-10-(prop-2-yn-1-yl)-10Hphenothiazine (0.338 g, 1.25 mmol) in THF was cooled to -78°C in an acetone/dry ice bath and allowed to stand for one hour in a round bottom flask. A solution of nBuLi (2.5 M in hexanes, 0.52 mL, 1.31 mmol) was added dropwise and allowed to stir for ca. 20 min. An excess of solid carbon dioxide was added and the round bottom flask sealed. The reaction was allowed to stir for 2 hours before warming to 10°C. Solution was then poured into a small beaker containing 10% HCl solution (pH ~2) and extracted into ether. The organic layer was adjusted with 10% NaOH to a pH of 11, organic layer discarded, and the aqueous layer acidified to 2 by the addition of 10% HCl. The aqueous layer was then extracted into ether, washed with brine, dried using sodium sulfate, and concentrated *in vacuo* to give the product as a light brown solid (0.078g, 20%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.16 (d, J = 7.8 Hz, 1H), 7.09 (d, J = 7.6 Hz, 1H), 7.05 (d, J = 8.1 Hz, 1H), 7.00 (d, J = 7.9 Hz, 2H), 6.95 (t, J = 7.6 Hz, 1H), 6.91 (dd, J = 8.1, 1.9 Hz, 1H), 4.60 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 156.0, 144.9, 143.1, 133.6, 127.9, 127.8, 127.4, 123.9, 123.3, 123.3, 122.2, 115.0, 115.0, 83.9, 77.8, 38.7. HRMS (ESI) *m/z*: [M+H]+ Calcd for $C_{16}H_{11}ClNO_2S$ 316.0189; Found 316.0201. ATIR CO_2H (br, 3400 cm^{-1} – 2700 cm^{-1}), -C=C-CO2 (2236 cm^{-1})

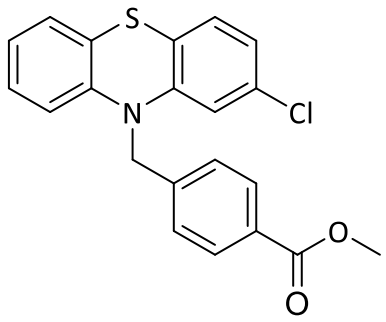


Figure E.8: Compound 3-8

Methyl 4-((2-chloro-10H-phenothiazin-10-yl)methyl)benzoate

2-chloro-10H-phenothiazine (0.583 g, 2.5 mmol) was added as a solution in anhydrous THF (ca. 15), to a round bottom flask charged with sodium hydride (0.090 g, 2.25 mmol). The mixture was stirred at room temperature for 1 hour followed by addition of 4-bromomethyl benzoate (0.458 g, 2 mmol). Upon addition, the reaction becomes a brown-orange solution and was covered in foil before stirring for 4 days, after which the solution was green. The solution was poured into a seperatory funnel containing diethyl ether and turned purple. Saturated sodium bicarbonate was added and the aqueous layer (brown in color) was discarded. The organic layer was washed with brine, dried over sodium sulfate and concentrated *in vacuo* to give the crude product. The crude was suspended in chloroform and placed in a -20°C freezer overnight to precipitate out unreacted starting material. The solution was decanted, and the chloroform concentrated *in vacuo* to give the pure product as a white solid (0.700 g, 92%). A sample was applied to an MPLC (Hexane: Ethyl Acetate gradient) for analytical purity. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.01 (d, *J* = 8.3 Hz, 2H), 7.42 – 7.35 (m, 2H), 7.09 (dd, *J* = 7.6, 1.6 Hz, 1H), 7.00 (d, *J* = 8.2 Hz, 1H), 6.89 (td, *J* = 7.5, 1.2 Hz, 1H), 6.85 (dd, *J* = 8.2, 2.0 Hz, 1H), 6.60 – 6.54 (m, 3H), 5.09 (s, 2H), 3.91 (s, 3H) ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.9, 145.7, 143.7, 141.6, 133.3, 130.3,

130.3, 129.5, 127.6, 127.6, 127.2, 126.8, 123.3, 122.7, 115.8, 115.7, 52.6, 52.3.. HRMS (ESI) *m/z*: $[M+H]^+$ Calcd for $C_{21}H_{17}ClNO_2S$ 382.0669; Found 382.0670 ATIR Aromatic CH (3100 cm^{-1}), CH(2949 cm^{-1}, 2922 cm^{-1}) -CO_2R (st, sharp, 1716 cm^{-1})

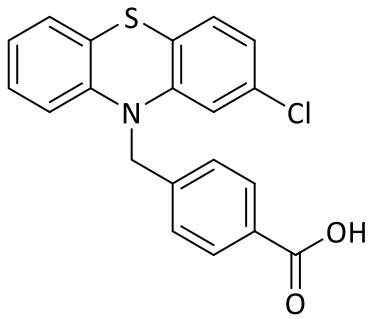


Figure E.9: Compound 3-9

4-((2-Chloro-10H-phenothiazin-10-yl)methyl)benzoic acid

Methyl 4-((2-chloro-10H-phenothiazin-10-yl)methyl)benzoate (0.400 g, 1.05 mmol) was added to a solution of 10% NaOH and methanol (1:1) and refluxed for 2 hours. The reaction was extracted with ether and the organic layer discarded. The aqueous layer was acidified with 10% HCl to pH 2 and extracted into ether (50 mL), washed with brine (50 mL), dried over sodium sulfate. The resulting solution was concentrated *in vacuo* to give a white solid (0.376 g, 97%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.09 (d, *J* = 8.2 Hz, 2H), 7.43 (d, *J* = 8.1 Hz, 2H), 7.10 (dd, *J* = 7.6, 1.5 Hz, 1H), 7.06 – 6.95 (m, 2H), 6.93 – 6.81 (m, 2H), 6.62 – 6.50 (m, 2H), 5.12 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 171.1, 145.7, 143.7, 142.7, 133.3, 131.0, 128.5, 127.7, 127.6, 127.3, 126.9, 123.5, 123.4, 122.8, 122.3, 115.8, 115.7, 52.6. HRMS (ESI) *m/z*: $[M]^-$ Calcd for $C_{20}H_{13}ClNO_2S^-$ 366.0356; Found 366.0356 ATIR CH Aromatic (2995 cm-1), CH (2912 cm^{-1}), CO2H (w, br, 3433 cm^{-1}), (w, 1610 cm^{-1})

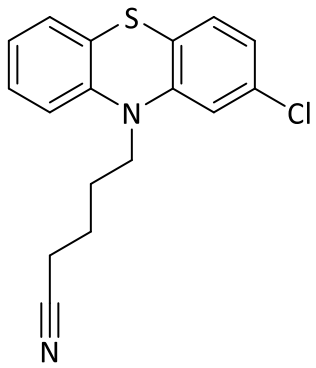


Figure E.10: Compound 3-10

5-(2-chloro-10H-phenothiazin-10-yl)pentanenitrile To a suspension of sodium hydride (200 mg, 5 mmol) in anhydrous THF, was added a solution of 2-chloro-10H-phenothiazine dropwise at room temperature in a round bottom flask. Solution was stirred for 0.5 hours before addition of valeronitrile (0.58 mL, 5 mmol) in a single portion. The mixture was allowed to stir for 48 hours before being concentration to dryness, extracted into ether, washed with brine (3 x 50 mL), concentrated *in vacuo* and dissolved in chloroform to remove as much starting material as possible, filtered and concentrated to give the final product that was used without further purification (21%, 131 mg). *Special Note: The above procedure was repeated after the lab obtained a medium pressure liquid chromatogram (MPLC). However, after brine wash, the crude material was concentrated *in vacuo* then applied to the MPLC using a very slow gradient of hexane:ethyl acetate giving superior yield (1.3g, 88%). The NMR spectra attached is from this run and not the original. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.22 – 7.12 (m, 2H), 7.05 (d, *J* = 8.2 Hz, 1H), 6.97 (td, *J* = 7.5, 1.2 Hz, 1H), 6.91 (dd, *J* = 8.2, 2.1 Hz, 1H), 6.86 (dd, *J* = 8.1, 1.2 Hz, 1H), 6.82 (d, *J* = 2.1 Hz, 1H), 3.88 (t, *J* = 6.4 Hz, 2H), 2.32 (t, *J* = 7.1 Hz, 2H), 1.96 – 1.88 (m, 2H), 1.80 – 1.71 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 146.4, 144.3, 133.3, 128.2, 127.8,

127.6, 125.5, 124.3, 123.3, 122.7, 119.4, 116.0, 46.2, 25.6, 22.7, 16.9. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{17}H_{16}ClN_2S$ 315.0717; Found 315.0720.

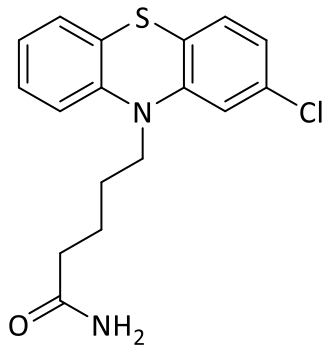


Figure E.11: Compound 3-11

5-(2-Chloro-10H-phenothiazin-10-yl)pentanamide

5-(2-chloro-10H-phenothiazin-10-yl)pentanenitrile was dissolved in concentrated sulfuric acid and stirred for 2.5 hours in a round bottom flask. The solution was then poured into an ice-water mixture and basified with concentrated ammonium hydroxide. Mixture was then extracted with EtOAc (3x 100 mL), washed with brine (1 x 100 mL), dried over sodium sulfite, and concentrated *in vacuo* to give the title compound (95%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.19 – 7.10 (m, 2H), 7.02 (d, *J* = 8.1 Hz, 1H), 6.93 (td, *J* = 7.5, 1.1 Hz, 1H), 6.88 (dd, *J* = 8.2, 2.0 Hz, 1H), 6.85 (dd, *J* = 8.2, 1.1 Hz, 1H), 6.81 (d, *J* = 2.0 Hz, 1H), 3.84 (t, *J* = 6.6 Hz, 2H), 2.21 (t, *J* = 7.3 Hz, 2H), 1.88 – 1.78 (m, 2H), 1.78 – 1.70 (m, 2H).^{13}C NMR (126 MHz, $CDCl_3$) δ 175.23, 146.61, 144.57, 133.36, 128.10, 127.70, 127.60, 124.98, 123.75, 123.11, 122.46, 115.97, 115.95, 47.05, 35.29, 26.13, 22.91. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd $C_{17}H_{18}ClN_2OS$ 333.0823; Found 358.1387.

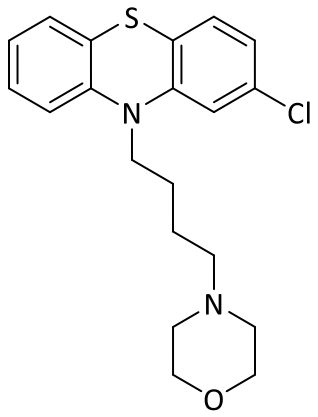


Figure E.12: Compound 3-12

4-(4-(2-Chloro-10H-phenothiazin-10-yl) butyl) morpholine (5)

2-Chloro-10-(4-iodobutyl)-10H-phenothiazine was added to a neat solution of morpholine and gently refluxed for 2 hours. The solution was poured into separatory funnel containing 0.5M HCl solution. The aqueous layer was extracted with EtOAc (2x 100 mL), and the combined organic layers washed with brine (100 mL), dried over sodium sulfite and concentrated *in vacuo* to give the title compound in quantitative yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.15 (ddd, *J* = 8.1, 7.4, 1.6 Hz, 1H), 7.12 (dd, *J* = 7.7, 1.5 Hz, 1H), 7.02 (d, *J* = 8.1 Hz, 1H), 6.93 (td, *J* = 7.5, 1.2 Hz, 1H), 6.88 (dt, *J* = 8.2, 1.8 Hz, 2H), 6.83 (d, *J* = 2.0 Hz, 1H), 3.86 (t, *J* = 6.9 Hz, 2H), 3.66 (t, *J* = 4.6 Hz, 4H), 2.36 (dd, *J* = 14.9, 7.7 Hz, 6H), 1.85 (tt, *J*= 7.7, 6.2 Hz, 2H), 1.62 (p, *J* = 7.3 Hz, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 146.6, 144.7, 133.3, 128.0, 127.7, 127.5, 124.9, 123.7, 123.0, 122.6, 115.9, 115.9, 67.1, 58.3, 53.8, 47.3, 24.4, 23.6. HRMS (ESI) *m/z*: [M+H]+ Calcd for $C_{20}H_{24}ClN_2OS$ 375.1298; found: 375.1306. ATIR: Aromatic CH (3100 cm-1), CH (2945 cm-1), CH (2846 cm-1).

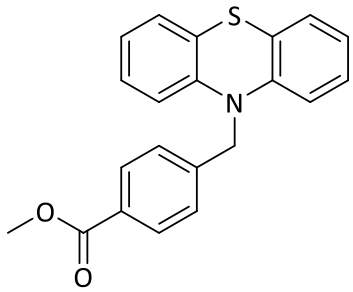


Figure E.13: Compound 3-13

Methyl 4-((10H-phenothiazin-10-yl)methyl)benzoate

General procedure. 101.3 mg, 15% yield as a waxy off-white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.02 (d, *J* = 8.3 Hz, 2H), 7.41 (d, *J* = 8.6 Hz, 2H), 7.12 (dd, *J* = 7.6, 1.6 Hz, 2H), 7.04 – 6.93 (m, 2H), 6.89 (t, *J* = 7.5 Hz, 2H), 6.60 (dd, *J* = 8.2, 1.1 Hz, 2H), 5.12 (s, 2H), 3.91 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.8, 144.3, 142.3, 130.1, 129.1, 127.3, 127.0, 126.8, 123.5, 122.8, 115.4, 52.5, 52.1. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{21}H_{18}NO_2S$: 348.1058 Found: 348.1099 ATIR: CH_2 (2843, 2882 cm^{-1}), aromatic CH (3004, 3059 cm^{-1}), carbonyl (s, 1713 cm^{-1})

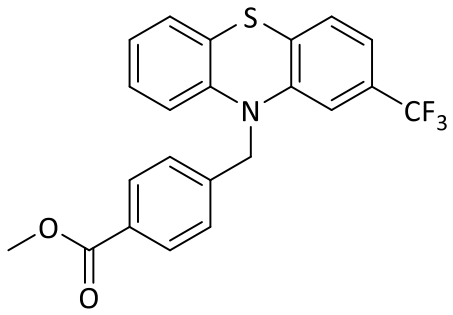


Figure E.14: Compound 3-14

Methyl 4-((2-(trifluoromethyl)-10H-phenothiazin-10-yl)methyl)benzoate

General procedure. 257.1 mg, 31% yield as a sticky white semi-solid that discolors on standing to green-white. ^{1}H NMR (500 MHz**,** $CDCl_3$) δ 8.02 (d, *J* = 8.3 Hz, 2H), 7.42 (d, *J* = 8.4 Hz, 2H), 7.19 (dd, *J* = 8.0, 0.9 Hz, 1H), 7.14 – 7.10 (m, 2H), 7.02 (ddd, *J* = 8.1, 7.4, 1.6 Hz, 1H), 6.92 (td, *J* = 7.5, 1.2 Hz, 1H), 6.80 (d, *J* = 1.7 Hz, 1H), 6.64 (dd, *J* = 8.2, 1.1 Hz, 1H), 5.14 (s, 2H), 3.91 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.8, 144.9, 143.7, 141.4, 130.3, 129.6 (q, $^{2}J_{CF}$ = 32.1 Hz, 1C), 129.5, 127.8, 127.3, 127.2, 126.8, 124.0 (q, $^{1}J_{CF}$ = 274.2 Hz, 1C), 123.5, 122.9, 119.4 (q, $^{3}J_{CF}$ = 3.9 Hz, 1C), 115.9, 111.7 (q, $^{3}J_{CF}$ = 3.7 Hz, 1C), 52.5, 52.2. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{22}H_{17}F_3NO_2S^+$ 416.0932; Found 416.0835 ATIR: CH (2957 cm^{-1}), aromatic CH (3009 cm^{-1}), carbonyl (s, 1713 cm^{-1})

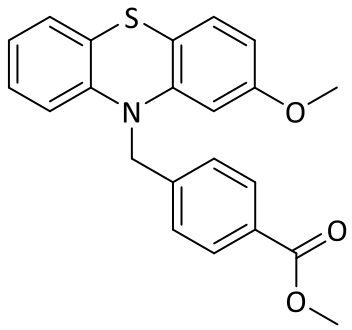


Figure E.15: Compound 3-15

Methyl 4-((2-methoxy-10H-phenothiazin-10-yl)methyl)benzoate

General Procedure. 5%, 37.7 mg as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.99 (d, *J* = 8.0 Hz, 2H), 7.40 (d, *J* = 8.0 Hz, 2H), 7.16 – 7.06 (m, 1H), 7.02 – 6.93 (m, 2H), 6.88 (t, *J* = 7.5 Hz, 1H), 6.59 (d, *J* = 8.5 Hz, 1H), 6.45 (dd, *J* = 8.4, 2.4 Hz, 1H), 6.20 (d, *J* = 2.4 Hz, 1H), 5.11 (s, 2H), 3.90 (s, 3H), 3.63 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.0, 159.7, 145.7, 144.2, 142.3, 130.2, 129.2, 127.4, 127.2, 127.1, 126.9, 122.9, 115.6, 107.0, 103.6, 55.5, 52.7, 52.3. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{22}H_{20}NO_3S^+$ 378.1158; Found 378.1161. ATIR: CH (2957 cm^{-1}), aromatic CH (3009 cm^{-1}), carbonyl (s, 1713 cm^{-1})

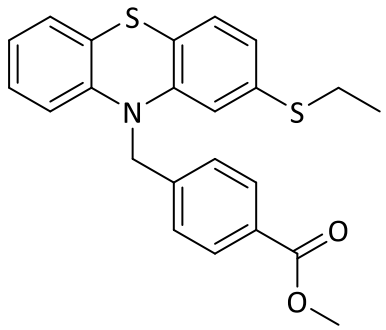


Figure E.16: Compound 3-16

Methyl 4-((2-(ethylthio)-10H-phenothiazin-10-yl)methyl)benzoate

Benzylation achieved with the general procedure to give product (95.1 mg, 12%) as a thick dark yellow oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.00 (d, *J* = 8.4 Hz, 2H), 7.44 – 7.35 (m, 2H), 7.10 (dd, *J* = 7.6, 1.5 Hz, 1H), 7.00 (d, *J* = 7.9 Hz, 2H), 6.91 – 6.72 (m, 2H), 6.61 (dd, *J* = 8.2, 1.2 Hz, 1H), 6.56 (d, *J* = 1.7 Hz, 1H), 5.10 (s, 2H), 3.90 (s, 3H), 2.67 (q, *J* = 7.3 Hz, 2H) ,1.09 (t, *J* = 7.3 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.8, 144.6, 144.1, 142.1, 135.5, 130.1, 129.2, 127.4, 127.1, 127.1, 126.8, 123.7, 123.6, 122.9, 121.4, 116.5, 115.5, 52.4, 52.2, 28.0, 14.3. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{23}H_{22}NO_2S_2$ 408.1086; Found 408.1084. ATIR: CH (2957 cm^{-1}), aromatic CH (3012 cm^{-1}), carbonyl (s, 1714 cm^{-1})

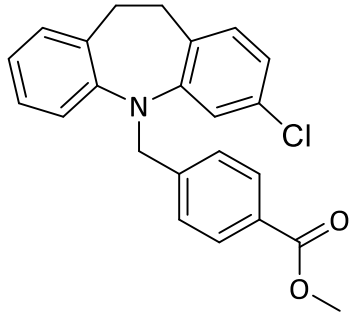


Figure E.17: Compound 3-17

Methyl 4-((3-chloro-10,11-dihydro-5H-dibenzo[b,f]azepin-5-yl)methyl)benzoate

Synthesized by general procedure to give the title compound as a white solid (400 mg, 34%). ^{1}H NMR (500 MHz**,** $CDCl_3$**)** δ 7.92 (d, *J* = 8.4 Hz, 2H), 7.44 (d, *J* = 8.4 Hz, 2H), 7.10 (d, *J* = 7.1 Hz, 1H), 7.08 – 7.06 (m, 2H), 7.03 (d, *J* = 2.0 Hz, 1H), 6.95 – 6.90 (m, 2H), 6.84 (dd, *J* = 8.1, 2.1 Hz, 1H), 4.97 (s, 2H), 3.87 (s, 3H), 3.29 – 3.09 (m, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.9, 148.5, 147.5, 143.2, 134.7, 131.6, 131.5, 131.2, 129.7, 129.6, 129.0, 128.0, 126.6, 123.5, 122.4, 120.4, 120.0, 55.4, 52.0, 32.3, 31.8. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{23}H_{21}NClO_2^+$ 378.1261; Found 378.1257 ATIR: CH (w, 2958 cm^{-1}), aromatic CH (w, 3006 cm^{-1}), carbonyl (s, 1711 cm^{-1})

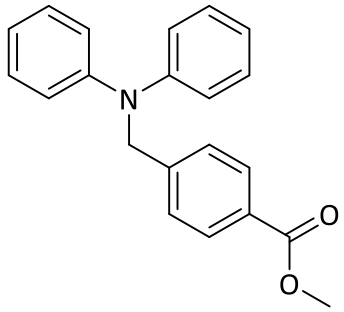


Figure E.18: Compound 3-18

Methyl 4-(((3-chlorophenyl)(phenyl)amino)methyl)benzoate

General Procedure gave the product as a white solid, which decays over a long period of time to a purple solid, 283 mg, 29.8%. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.07 (d, *J* = 8.4 Hz, 2H), 7.49 (d, *J* = 8.5 Hz, 2H), 7.31 (dd, *J* = 8.7, 7.4 Hz, 4H), 7.13 (dd, *J* = 8.8, 1.1 Hz, 4H), 7.05 – 6.98 (m, 2H), 5.09 (s, 2H), 3.94 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.9, 147.9, 144.9, 130.0, 129.5, 128.9, 126.6, 121.8, 120.7, 56.3, 52.1. HRMS (ESI+) *m/z*: $[M+H]^+$ Calc'd for $C_{21}H_{20}NO_2^+$ 318.1494; Found 318.1498 ATIR: CH (2957 cm^{-1}), aromatic CH (3065, 3045, 3009 cm^{-1}), carbonyl (s, 1711 cm^{-1})

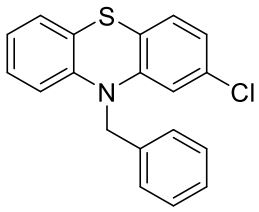


Figure E.19: Compound 3-19

10-Benzyl-2-chloro-10H-phenothiazine

General Procedure provided the product as a colorless oil (200 mg, 17%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.39 – 7.26 (m, 5H), 7.12 (dd, *J* = 7.6, 1.6 Hz, 1H), 7.04 – 6.97 (m, 2H), 6.94 – 6.84 (m, 2H), 6.71 – 6.64 (m, 2H), 5.06 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 145.8, 143.9, 136.0, 133.1, 128.9, 127.5, 127.4, 127.3, 127.0, 126.6, 123.1, 123.0, 122.4, 121.9, 115.9, 115.7, 52.7. HRMS (ESI+) *m/z*: $[M+H]^+$ Calc'd for $C_{19}H_{15}ClNS^+$ 324.0608; Found 324.0611 ATIR: CH (2924, 2851 cm^{-1}), aromatic CH (3112, 3060, 3021 cm^{-1})

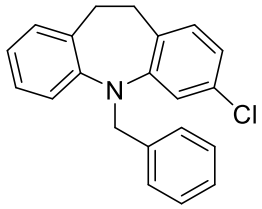


Figure E.20: Compound 3-20

5-Benzyl-10,11-dihydro-5H-dibenzo[b,f]azepine

Synthesized by the general procedure to yield the product as a brown solid (2.83 mg, 45%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.49 (d, *J* = 8.1, 2H), 7.36 (t, *J* = 6.5 Hz, 2H), 7.29 – 7.19 (m, 5H), 7.08 – 7.04 (m, 2H), 6.96 (dd, *J* = 8.1, 2.1 Hz, 1H), 5.01 (s, 2H), 3.35 – 3.32 (m, 2H), 3.3 – 3.25 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 148.8, 147.8, 137.9, 134.7, 131.6, 131.5, 131.2, 129.5, 128.4, 128.1, 127.1, 126.5, 123.3, 122.2, 120.7, 120.2, 55.7, 32.4, 31.9. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{21}H_{19}ClN^+$ 320.1200; Found 320.1207 Found 378.1257 ATIR: CH (w, 2958 cm^{-1}), aromatic CH (w, 3006 cm^{-1})

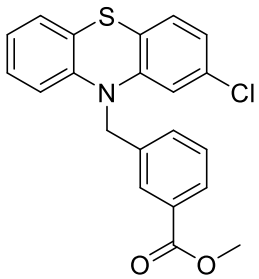


Figure E.21: Compound 3-21

Methyl 3-((2-chloro-10H-phenothiazin-10-yl)methyl)benzoate

General Procedure generated title compound as a white solid (535 mg, 20%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.05 (s, 1H), 7.94 (d, *J* = 7.7 Hz, 1H), 7.47 (d, *J* = 7.7 Hz, 1H), 7.37 (t, *J* = 7.7 Hz, 1H), 7.09 (dd, *J* = 7.6, 1.6 Hz, 1H), 7.00 – 6.98 (m, 2H), 6.88 (td, *J* = 7.5, 1.2 Hz, 1H), 6.84 (dd, *J* = 8.2, 2.0 Hz, 1H), 6.64 – 6.53 (m, 2H), 5.06 (s, 2H), 3.91 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.8, 145.7, 143.6, 136.7, 133.1, 131.1, 130.8, 129.0, 128.6, 127.8, 127.5, 127.5, 127.1, 123.4, 123.2, 122.6, 122.3, 115.8, 115.6, 52.3, 52.3. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{21}H_{17}ClNO_2S^+$ 382.0669; Found: 382.0670 ATIR Aromatic CH (3100 cm^{-1}), CH(2949 cm^{-1}, 2922 cm^{-1}) , (s, 1713 cm^{-1})

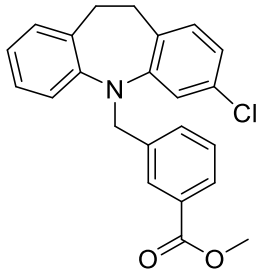


Figure E.22: Compound 3-22

Methyl 3-((3-chloro-10,11-dihydro-5H-dibenzo[b,f]azepin-5-yl)methyl)benzoate

General procedure provided the product as a pale, sticky oil (18 mg, 30%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.08 (s, 1H), 7.84 (dt, *J* = 7.7, 1.5 Hz, 1H), 7.56 (d, *J* = 7.7 Hz, 1H), 7.32 (t, *J* = 7.7 Hz, 1H), 7.18 – 7.03 (m, 4H), 7.00 (d, *J* = 8.0 Hz, 1H), 6.97 – 6.90 (m, 1H), 6.85 (dd, *J* = 8.1, 2.1 Hz, 1H), 4.97 (s, 2H), 3.91 (s, 3H), 3.29 – 3.21 (m, 2H), 3.22 – 3.16 (m, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 167.0, 148.6, 147.5, 138.4, 134.8, 132.5, 131.6, 131.5, 131.2, 130.3, 129.6, 129.3, 128.5, 128.3, 126.6, 123.4, 122.4, 120.5, 120.0, 55.2, 52.1, 32.3, 31.8. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{23}H_{21}NClO_2^+$ 378.1261; Found 378.1273 ATIR: CH (2848 cm^{-1}), aromatic CH (3089, 3062 cm^{-1}), carbonyl (m, 1719 cm^{-1})

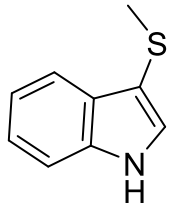


Figure E.23: Compound 3-23

3-(Methylthio)-1H-indole

Indole (585 mg, 5 mmol), iodine (634 mg, 2.5 mmol) and dimethyl disulfide (0.24 mL, 2.75 mL) were combined in a round bottom flask charged with ethanol and refluxed for 12 hours. After which, the solution was cooled to room temperature, washed with sodium thiosulfate (reaction volume x 2), extracted into EtOAc (reaction volume x2), dried over sodium sulfite, and purified via MPLC (132.4 mg, 16.2%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.12 (bs, 1NH), 7.91 – 7.76 (m, 1H), 7.37 (dd, *J* = 7.0, 1.5 Hz, 1H), 7.33 – 7.27 (m, 2H), 7.26 (d, *J* = 2.6 Hz, 1H), 2.43 (d, *J* = 0.9 Hz, 3H).^{13}C NMR (126 MHz, $CDCl_3$) δ 136.3, 128.7, 128.0, 122.8, 120.4, 119.2, 111.7, 107.9, 20.3. NMR spectra of the compound matched literature values[180].

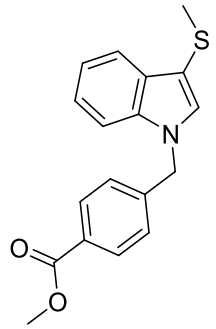


Figure E.24: Compound 3-24

Methyl 4-((3-(methylthio)-1H-indol-1-yl)methyl)benzoate

Benzylation via general procedure to give the title compound as a yellow solid (59 mg, 30%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.98 (d, *J* = 8.4 Hz, 2H), 7.83 – 7.76 (m, 1H), 7.33 – 7.02 (m, 6H), 5.33 (s, 2H), 3.90 (s, 3H), 2.40 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.7, 142.1, 136.9, 131.6, 130.2, 129.8, 129.6, 126.8, 122.8, 120.4, 119.7, 110.1, 107.6, 52.3, 50.0, 20.5. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{18}H_{18}NO_2S^+$ 312.1058; Found 312.0369. ATIR: CH (2982, 2915 cm^{-1}), aromatic CH (3108, 3060 cm^{-1}), carbonyl (s, 1715 cm^{-1})

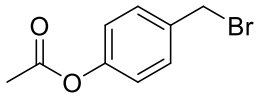


Figure E.25: Compound 3-25

4-(Bromomethyl)phenyl acetate

Powdered NBS (3.1 g, 17.5 mmol) and benzoyl peroxide (0.8 g, 3.3 mmol) were added to a stirred solution of p-tolyl-acetate in freshly distilled chlroform (20 mL). Solution was refluxed for 4 hours. After cooling to room temperature, solids were removed by filtration before concentrating *in vacuo.* Residue was then dissolved in DCM, washed with DI water and dried over sodium sulfate. Compound was applied to an MPLC to give the title compound (2.8 g, 74%) with identical NMR spectra to literature[181]. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.42 (d, *J* = 6.8 Hz, 2H), 7.08 (d, *J* = 6.8 Hz, 2H), 4.43 (s, 2H), 2.31 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.3, 150.5, 130.2, 127.9, 122.0, 32.7, 21.1.

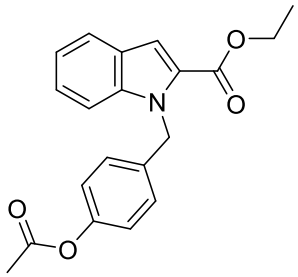


Figure E.26: Compound 3-26

Ethyl 1-(4-acetoxybenzyl)-1H-indole-2-carboxylate

Benzylation by general procedure gave the title compound as a clear colorless oil (303 mg, 30%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.76 (d, *J* = 8.0 Hz, 1H), 7.46 (s, 1H), 7.42 – 7.32 (m, 2H), 7.22 (t, *J* = 7.9 Hz, 1H), 7.12 (d, *J* = 8.6 Hz, 2H), 7.05 – 6.99 (d, J = 8.6 Hz, 2H), 5.86 (s, 2H), 4.38 (q, *J* = 7.1 Hz, 2H), 2.28 (s, 3H), 1.40 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.4, 161.9, 149.7, 139.4, 135.9, 127.6, 127.4, 126.1, 125.4, 122.7, 121.7, 120.9, 111.1, 110.8, 60.6, 47.3, 21.1, 14.3. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{20}H_{20}NO_4^+$ 338.1392; Found 388.1388 ATIR: CH (2985, 2930 cm^{-1}), aromatic CH (3108, 3061 cm^{-1}), carbonyl (s, 1704 cm^{-1}), carbonyl (s, 1762 cm^{-1})

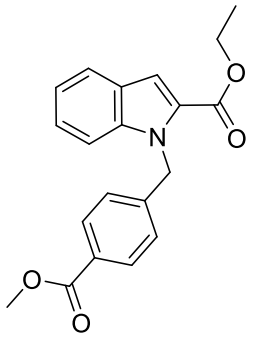


Figure E.27: Compound 3-27

Ethyl 1-(4-(methoxycarbonyl)benzyl)-1H-indole-2-carboxylate

Benzylation by general procedure gave the title compound as a beige solid (444 mg, 43%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.99 – 7.93 (m, 2H), 7.76 (d, J = 8.0 Hz, 1H), 7.47 (s, 1H), 7.36 – 7.25 (m, 2H), 7.23 – 7.13 (m, 1H), 7.11 (d, J = 8.1 Hz, 2H), 5.87 (s, 2H), 4.35 (q, J = 7.1 Hz, 2H), 3.87 (s, 3H), 1.37 (t, J = 7.1 Hz, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 166.7, 161.9, 143.7, 139.5, 130.0, 129.1, 127.6, 126.2, 125.5, 122.8, 121.1, 120.5, 111.3, 110.6, 60.7, 52.0, 47.7, 14.3. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{20}H_{20}NO_4^+$ 338.1392; Found 388.1372 ATIR: CH (2983, 2853, 2924 cm^{-1}), aromatic CH (3309, 3062 cm^{-1}), carbonyl (s, 1717 cm^{-1}), carbonyl (s, 1700 cm^{-1})

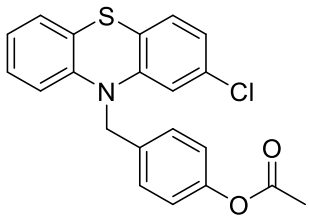


Figure E.28: Compound 3-28

4-((2-Chloro-10H-phenothiazin-10-yl)methyl)phenyl acetate

Benzylation by general procedure to give an off-white solid (12%, 127 mg). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32 (d, *J* = 8.6 Hz, 2H), 7.09 (m, 3H), 6.99 (d, *J* = 8.2 Hz, 2H), 6.90 (td, *J* = 7.5, 1.2 Hz, 1H), 6.86 (dd, *J* = 8.2, 2.0 Hz, 1H), 6.68 – 6.61 (m, 2H), 5.02 (s, 2H), 2.31 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.4, 149.8, 145.7, 143.7, 133.5, 133.2, 127.6, 127.5, 127.5, 127.0, 123.1, 123.1, 122.5, 122.0, 122.0, 115.9, 115.6, 52.2, 21.2. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{21}H_{17}ClNO_2S^+$ 382.0663; Found 382.0669.

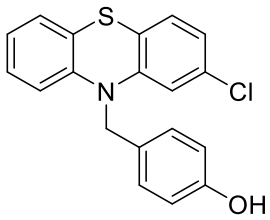


Figure E.29: Compound 3-29

4-((2-Chloro-10H-phenothiazin-10-yl)methyl)phenol

4-((2-chloro-10H-phenothiazin-10-yl)methyl)phenyl acetate (88 mg, 0.23 mmol) was added to a stirred solution of KOH (excess) in a mixture of ethanol/water (1:1) and refluxed for 4 hours. After cooling, mixture was concentrated to half volume and extracted with ether. Organic layer was discarded, and the aqueous layer acidified to pH 2 and extracted into ether. Crude material was then concentrated and applied to the MPLC for purification to give a white solid that rapidly turns purple (131 mg, 21%) ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.24 (d, *J* = 7.9 Hz, 1H), 7.16 (d, *J* = 8.4 Hz, 2H), 7.07 (dd, *J* = 7.8, 1.4 Hz, 1H), 7.02 – 6.94 (m, 2H), 6.91 – 6.82 (m, 2H), 7.70 (d, *J* = 9.3 Hz, 2H), 6.68 – 6.59 (m, 2H), 4.98 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 154.73, 145.91, 144.01, 133.20, 129.78, 128.05, 128.01, 127.51, 127.45, 127.03, 123.18, 123.07, 122.43, 122.02, 115.99, 115.88, 115.84, 115.34, 52.23. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{19}H_{15}ClNOS^+$ 340.0536; Found 340.0474 ATIR: CH (2962 cm^{-1}), aromatic CH (3062 cm^{-1}), broad OH (3321 cm^{-1})

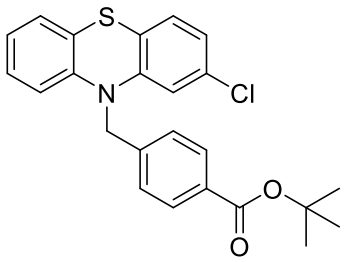


Figure E.30: Compound 3-30

Tert-butyl 4-((2-chloro-10H-phenothiazin-10-yl)methyl)benzoate

Benzylation using general procedure gave the desired compound as a white solid (64 mg, 7.6%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.98 (d, *J* = 8.2 Hz, 2H), 7.36 (d, *J* = 8.1 Hz, 2H), 7.09 (dd, *J* = 7.6, 1.5 Hz, 1H), 7.04 – 6.93 (m, 2H), 6.92 – 6.88 (m, 2H), 6.61 – 6.55 (m, 2H), 5.07 (s, 2H), 1.60 (s, 9H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 165.4, 145.6, 143.6, 140.8, 133.1, 131.2, 130.1, 127.5, 127.4, 127.1, 126.5, 123.3, 123.2, 122.6, 122.1, 115.8, 115.6, 81.1, 52.4, 28.2. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{24}H_{23}ClNO_2S^+$ 424.1133; Found 424.1168. ATIR Aromatic CH (3100 cm^{-1}), CH(2949 cm^{-1}, 2922 cm^{-1}) , (s, 1713 cm^{-1})

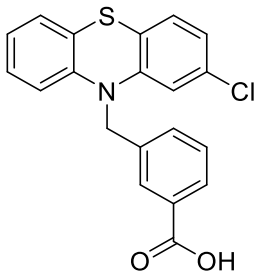


Figure E.31: Compound 3-31

3-((2-Chloro-10H-phenothiazin-10-yl)methyl)benzoic acid

Methyl 3-((2-chloro-10H-phenothiazin-10-yl)methyl)benzoate (0.400 g, 1.05 mmol) was added to a solution of 10% KOH and ethanol (1:1) and refluxed for 2 hours in a round bottom flask. The reaction was extracted with ether and the organic layer discarded. The aqueous layer was acidified with 10% HCl to pH 2 and extracted into ether (50 mL), washed with brine (50 mL) and dried over sodium sulfate. The resulting solution was concentrated *in vacuo* to give a white solid (376 mg, 97%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.11 (s, 1H), 8.01 (d, *J* = 7.7 Hz, 1H), 7.54 (d, *J* = 1.5 Hz, 1H), 7.42 (t, *J* = 7.7 Hz, 1H), 7.10 (d, *J* = 7.6, 1H), 7.03 – 6.95 (m, 2H), 6.94 – 6.83 (m, 2H), 6.69 – 6.56 (m, 2H), 5.10 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 178.0, 145.6, 143.6, 136.9, 133.2, 132.0, 130.0, 129.3, 129.2, 128.4, 127.6, 127.5, 127.1, 123.5, 123.2, 122.6, 122.4, 115.8, 115.6, 52.2. HRMS (ESI) *m/z*: $[M]^-$ Calcd for $C_{20}H_{13}ClNO_2S^-$ 366.0356; Found 366.0356 ATIR CH Aromatic (2995 cm^-1), CH (2912 cm^{-1}), CO2H (w, br, 3433 cm^{-1}), (w, 1610 cm^{-1})

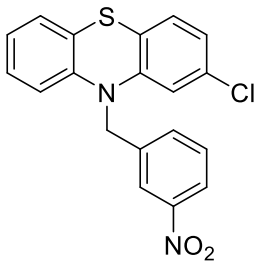


Figure E.32: Compound 3-32

2-Chloro-10-(3-nitrobenzyl)-10H-phenothiazine

Benzylation followed the general procedure with one addendum. After addition of 3-(bromomethyl)-nitrobenzene, potassium iodide (0.166g, 1 mmol) was added and the reaction was stirred while covered in foil at room temperature overnight. Workup is as written in the general procedure to give the title compound as a bright yellow solid (1.16 g, 31 %). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.24 (s, 1H), 8.13 (d, *J* = 7.2 Hz, 1H), 7.66 (d, *J* = 6.8 Hz, 1H), 7.50 (t, *J* = 7.9 Hz, 1H), 7.13 (dd, *J* = 7.6, 1.5 Hz, 1H), 7.08 – 6.99 (m, 2H), 6.95 – 6.86 (m, 2H), 6.65 – 6.50 (m, 2H), 5.14 (s, 2H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 148.7, 145.4, 143.4, 138.6, 133.2, 132.8, 129.9, 127.8, 127.5, 127.4, 124.0, 123.5, 122.9, 122.8, 122.6, 121.8, 115.7, 115.6, 51.8. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{19}H_{14}ClN_2O_2S^+$ 369.0464; Found 369.0511 ATIR: CH (weak, 2924, 2855 cm^{-1}), aromatic CH (3309, 3062 cm^{-1}), NO (1459, 1404 cm^{-1})

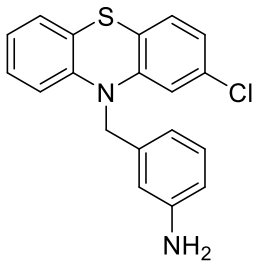


Figure E.33: Compound 3-33

3-((2-Chloro-10H-phenothiazin-10-yl)methyl)aniline

2-chloro-10-(3-nitrobenzyl)-10H-phenothiazine (1.16 g, 3.1 mmol) was dissolved in EtOAc. Tin(II) chloride dihydrate (10 eq) was added and the mixture refluxed for 2 h. The mixture was cooled to room temperature and poured into ice, and the pH was adjusted to 10. The resulting slurry was extracted with EtOAc (3 x 50 mL) and the combined organic fractions washed with brine (1 x 50 mL), dried over sodium sulfite and concentrated *in vacuo* to give the title compound in quantitative yield as a white solid (1.17 g). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.12 (t, *J* = 7.6 Hz, 1H), 7.06 (d, *J* = 7.5, 1H), 7.00 – 6.92 (m, 2H), 6.91 – 6.79 (m, 2H), 6.74 – 6.61 (m, 3H), 6.58 – 6.53 (m, 2H), 4.93 (s, 2H), 3.59 (bs, 2NH). ^{13}C NMR (126 MHz, $CDCl_3$) δ 146.9, 145.8, 143.8, 137.2, 133.1, 129.9, 127.5, 127.2, 126.8, 123.0, 122.6, 122.3, 121.5, 116.6, 115.9, 115.6, 114.0, 112.8, 52.8. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{19}H_{16}ClN_2S^+$ 339.0723; Found 339.0822 ATIR: NH (3446, 3365 cm^{-1}), CH (2924, 2855 cm^{-1}), aromatic CH (3205, 3011 cm^{-1}), (s, 1717 cm^{-1}), NH bend (1616, 1590 cm^{-1}).

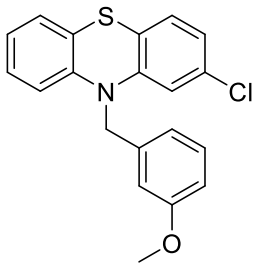


Figure E.34: Compound 3-34

2-Chloro-10-(3-methoxybenzyl)-10H-phenothiazine

Benzylation was done using the general procedure to isolate product as a yellow oil (99 mg, 28%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.21 (d, *J* = 8.5 Hz, 2H), 7.08 (dd, *J* = 7.6, 1.6 Hz, 1H), 7.04 – 6.95 (m, 2H), 6.93 – 6.79 (m, 4H), 6.71 – 6.59 (m, 2H), 4.99 (s, 2H), 3.79 (s, 3H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 158.7, 145.8, 143.9, 133.1, 129.8, 127.7, 127.7, 127.4, 127.3, 126.9, 123.1, 123.0, 122.3, 121.9, 115.9, 115.7, 114.3, 113.9, 55.3, 52.1. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{20}H_{17}ClNOS^+$ 353.0641; Found 353.0672 ATIR: CH (2924, 2851 cm^{-1}), aromatic CH (3112, 3060, 3021 cm^{-1})

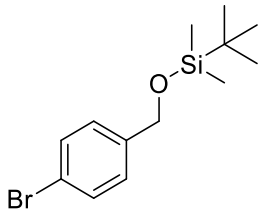


Figure E.35: Compound 4-1

((4-Bromobenzyl)oxy)(tert-butyl)dimethylsilane

p-Bromo benzyl alcohol (10 g, 53 mmol) was added to a stirring solution of *tert*-butyl dimethylsilyl chloride (8.8g, 58.3 mmol) and imidazole (7.9 g, 116.6 mmol) in DCM. Solution was stirred at room temperature for three days and then quenched with aqueous ammonia chloride. The aqueous layer was extracted with DCM (3x 100 mL) and the combined organic layers were washed with brine (1x 300 mL) and dried over sodium sulfite. Solution was concentrated *in vacuo* and purified via hexane:ethyl acetate gradient on the MPLC (16.2 g, 99%). Spectroscopic data matches that reported for this compound[162]. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.66 – 7.35 (m, 2H), 7.35 – 7.11 (m, 2H), 4.72 (s, 2H), 0.98 (s, 9H), 0.27 (s, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 140.5, 131.3, 127.7, 120.6, 64.3, 26.0, 18.4, -5.2.

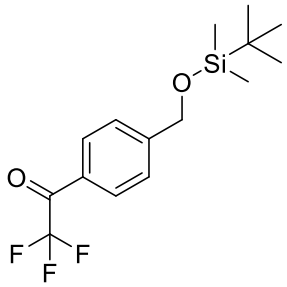


Figure E.36: Compound 4-2

1-(4-(((Tert-butyldimethylsilyl)oxy)methyl)phenyl)-2,2,2-trifluoroethan-1-one

((4-Bromobenzyl)oxy)(tert-butyl)dimethylsilane (16.2 g, 53 mmol) was dissolved in anhydrous THF and stirred for 15 minutes at -78 °C (acetone dry ice bath). nBuLi (25.4 mL, 63.6 mmol) was added dropwise under constant temperature over 1 hour. Solution was left to stir at -78 °C for 1.5 hours before dropwise addition of ethyl trifluoroacetate (8.9 mL, 74.2 mmol). Temperature and stirring was maintained for another 1.5 hours. To quench, saturated aqueous ammonium chloride (reaction volume) was added at -78 °C and reaction was allowed to warm to room temperature before being diluted with diethyl ether. Organic layer was washed with saturated aqueous ammonium chloride (2x 100 mL), brine (2x 100 mL). Organic layer was dried over sodium sulfate and concentrated *in vacuo* to give the title compound (10.9 g, 64%) as a colorless liquid. Product was used without further purification and matched literature[162]. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.07 (d, J = 8.0 Hz, 2H), 7.52 (d, J = 8.1 Hz, 2H), 4.85 (s, 2H), 0.98 (s, 9H), 0.14 (s, 6H).^{13}C NMR (126 MHz, $CDCl_3$) δ 180.3, 180.2 (q, $^{2}J_{CF}$ = 38.6 Hz, 1C), 150.1, 130.2 (q, $^{4}J_{CF}$ = 1 Hz , 1C) 128.6, 128.2, 126.9, 126.1, 126.0, 117.9, 116.7 (q, $^{1}J_{CF}$ = 292.5 Hz, 1C), 115.6, 64.2, 25.9, -5.4.

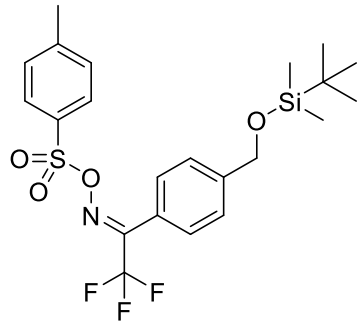


Figure E.37: Compound 4-3

(E)-1-(4-(((Tert-butyldimethylsilyl)oxy)methyl)phenyl)-2,2,2-trifluoroethan-1-one O-tosyl oxime

Hydroxylamine hydrochloride salt (7.3 g, 105 mmol) and sodium acetate (11.5 g, 140 mmol) were combined in ethanol stirred for 15 minutes at room temperature. Solids were filtered and the supernatant was added to a stirring solution of 1-(4-(((tert-butyldimethylsilyl)oxy)methyl)phenyl)-2,2,2-trifluoroethan-1-one (10.9 g, 35 mmol) and excess sodium sulfate. Mixture was refluxed overnight. After cooling to room temperature, solution is filtered and concentrated *in vacuo*. The crude material is backfilled with argon and charged with anhydrous DCM (5 mL/2 mmol of product calculated from crude mass), pyridine (5.6 mL, 70 mmol), DMAP (3.8 g, 31.5 mmol) and cooled to zero centigrade. In batches, p-toluene sulfonyl chloride (7.3 g, 38.5 mmol) was added and mixture stirred for 0.5 h at zero then 2 hours at room temperature. Reaction was then quenched with DI water and extracted with DCM (3x 100 mL). Combined organic layers were dried over sodium sulfate, concentrated *in vacuo* and purified on silica gel with 10% EtOAc in hexanes (2 g, 13.7%) as a pale green oil. Compound exists as a mixture of E/Z isomers and matches spectra from the literature[162].

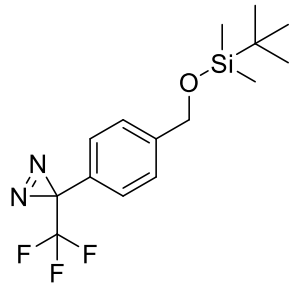


Figure E.38: Compound 4-5

3-(4-(((Tert-butyldimethylsilyl)oxy)methyl)phenyl)-3-(trifluoromethyl)-3H-diazirine

(E)-1-(4-(((tert-butyldimethylsilyl)oxy)methyl)phenyl)-2,2,2-trifluoroethan-1-one O-tosyl oxime (2 g, 4.7 mmol) was dissolved in diethyl ether (anhydrous 20 mL) and cooled to -78 °C. Ammonia was bubbled through until ~60 mL had condensed and reaction was stirred for 3 hours at -78 °C, then allowed to warm to room temperature over 2 hours. After evaporation of ammonia, the reaction mixture was filtered and concentrated to give a translucent paste that was used in subsequent transformations without further purification. The paste was dissolved in anhydrous MeOH and TEA (excess) in a darkened fume hood. Molecular iodine (0.6 g, 2.4 mmol) was added in portions until a red/orange color persists. After stirring for 20 minutes at room temperature, the reaction was quenched with a few drops of 5% sodium metabisulphite and neutralized with 10% citric acid. Mixture was poured into ether (150 mL), dried over sodium sulfate, and concentrated *in vacuo*. Crude material was purified via MPLC in the dark using 2:1 hexane/DCM to give the title compound as a light yellow oil (303 mg, 20%). Spectral data matches literature report[162]. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.41 – 7.36 (m, 2H), 7.19 (d, *J* = 8.1 Hz, 2H), 4.77 (s, 2H), 0.97 (d, *J* = 1.0 Hz, 9H), 0.13 (d, *J* = 1.0 Hz, 6H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 143.3, 127.6, 126.4, 126.4, 126.2, 125.5, 123.3, 121.1, 118.9, 64.2, 25.9, 18.4, -5.4.

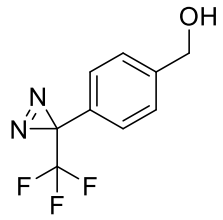


Figure E.39: Compound 4-6

(4-(3-(Trifluoromethyl)-3H-diazirin-3-yl)phenyl)methanol

3-(4-(((Tert-butyldimethylsilyl)oxy)methyl)phenyl)-3-(trifluoromethyl)-3H-diazirine (303 mg, 0.9 mmol) was dissolved in 1M TBAF in THF (2 mL) with 5% water (0.1 mL) in a round bottom flask wrapped in foil. Reaction was stirred for 5 hours at room temperature before being diluted with diethyl ether (40 mL), washed with DI water (3 x12 mL). Organic layers were combined, dried over sodium sulfate, and concentrated *in vacuo*. Purify on silica gel with 100% DCM to give the title compound (quant, 0.9 mmol) as a pale-yellow oil. Spectral data matches literature report[162]. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.35 (d, J = 8.0 Hz, 2H), 7.17 (d, J = 8.0 Hz, 2H), 4.66 (d, J = 5.5 Hz, 2H), 2.3 (bs, 1H).^{13}C NMR (126 MHz, $CDCl_3$) δ 142.5, 128.3, 127.1, 126.6, 126.6, 122.1 (q, $^{1}J_{CF}$ = 275.4 Hz, 1C), 64.3.

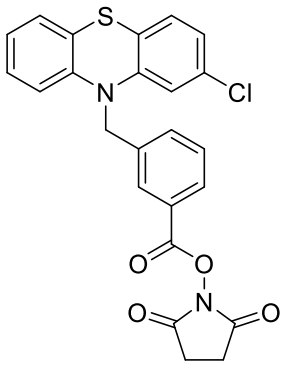


Figure E.40: Compound 4-7

2,5-Dioxopyrrolidin-1-yl 3-((2-chloro-10H-phenothiazin-10-yl)methyl)benzoate

Compound **3-31** (460 mg, 1.3 mmol), EDCI (255 mg, 2 mmol), and DIPEA (0.7 mL, 2 mmol) were combined in anhydrous DMF and stirred at room temperature. After 1 hour, N-hydroxysuccinamide was added and allowed to stir for 4 hours at room temperature. Solution was then diluted with EtOAc (~ 100 mL), washed with saturated bicarbonate solution (2x 100 mL), DI water (3x 100 mL), and brine (1x 100 mL). Organic layer was dried over sodium sulfite and concentrated *in vacuo* before being purified via standard MPLC conditions to give the title compound (22 mg, 3.6%) as a white solid. ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.16 (s, 1H), 8.05 (d, J = 8.7 Hz, 1H), 7.59 (d, J = 7.6 Hz, 1H), 7.45 (t, *J* = 7.8 Hz, 1H), 7.10 (d, *J* = 7.6 Hz, 1H), 7.05 – 6.97 (m, 2H), 6.95 – 6.80 (m, 2H), 6.65 – 6.52 (m, 2H), 5.11 (s, 2H), 2.91 (s, 4H). ^{13}C NMR (126 MHz, $CDCl_3$) δ 169.2, 161.6, 145.6, 143.4, 137.4, 133.3, 133.1, 129.6, 129.5, 128.8, 127.7, 127.5, 127.2, 125.8, 123.6, 123.3, 122.7, 122.4, 115.8, 115.5, 52.0, 25.7. HRMS (APCI) *m/z*: $[M+H]^+$ Calc'd for $C_{24}H_{18}ClN_2O_2S^+$ 465.0676; Found 465.0693

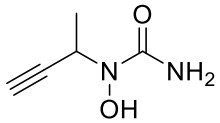


Figure E.41: Compound 5-4

1-(But-3-yn-2-yl)-1-hydroxyurea

But-3-yn-2-ol (5 mL, 63 mmol) was dissolved in dry DCM (ca. 110 mL) to give a colorless solution which was cooled to 0°C before freshly distilled TEA (11.67 mL, 83 mmol) was added as a single portion. This solution was allowed to stir for 10 min at 0° C before methane sulfonyl chloride (5.96 mL, 76.7 mmol) was added dropwise while keeping the temperature below 10°C. The reaction was stirred below 10°C until TLC showed complete conversion of starting material. An equal volume of 0.5 M HCl was added and the layers were separated. The aqueous layer was further extracted with DCM (2x30 mL). Organic layers were combined, washed with brine (2x70 mL), dried over $MgSO_4$, filtered and dried in vacuo before dissolving the residue in MeOH and cooling to 0°C. Hydroxylamine (50% aq. solution, 31.3 mmol) was added and the mixture was allowed to warm to room temperature with stirring. After stirring for 16 hr, the solution was concentrated in vacuo and pH adjusted to 9 with sodium hydroxide pellets. The mixture was then extracted with ethyl acetate (4x 100 mL) and concentrated to ca. 30 mL. To this pale yellow solution, potassium cyanate (13.8g, 170 mmol) in ca. 100 mL water was added in a single portion and allowed to sitr for 20 minutes. Afterwards, fuming HCl (15.45 mL, 502 mmol) was added dropwise via addition funnel while keeping the temperature below 10°C with an ice bath. After addition was complete, the solution was warmed to room temperature and stirred for an additional 14 hr. The layers were then separated and the aqueous layer further extracted with EtOAc (5 x 200 mL). The organic layers were combined into two fractions

[due to limitation of available glassware size], both washed with brine (3x 100mL) and dried over MgSO4. The solvent was removed under reduced pressure to give ca. 30mL. This was diluted under vigorous stirring with heptane (100 mL, white precipitate forms) and concentrated under reduced pressure. The residue was recrystallized from EtOAc/diethyl ether 1:1 and the resulting white solid was filtered off and the mother liquour concentrated under reduced pressure without heat to precipitate out additional product. The flask was placed into a fridge overnight and the resulting white solid was also filtered and combined with the previous fraction and dried under reduced pressure to yield the title compound (2.8 g, 22 %) as a white free flowing powder. Spectral data match literature report[151]. ^{1}HNMR (500 MHz, DMSO-d_6): δ 9.23 (s, 1H), 6.50 (s, 2H), 4.84 (q, J = 7.0 Hz, 1H), 3.03 (d, J = 2.2 Hz, 1H), 1.24 (d, J = 7.1 Hz, 3H).^{13}CNMR (126 MHz, DMSO-d_6): δ 162.0, 84.3, 73.3, 45.6, 18.9.

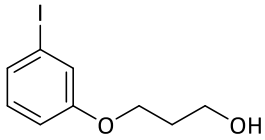


Figure E.42: Compound 5-5

3-(3-Iodophenoxy)propan-1-ol

A round bottom flask was charged with THF and sodium hydride (1.05 g, 26.25 mmol) under an inert atmosphere. M-iodophenol (6g, 27.5 mmol) was carefully added in 3 portions and stirred for 15 minutes at room temperature. 3-Bromo-propanol (2.3 mL, 25 mmol) was then added dropwise and the solution was stirred for 3 days. Solution was then concentrated to dryness *in vacuo* then diluted with DI water before being extracted into diethyl ether. Organic layer was washed with Brine (2 x 150 mL), dried over sodium sulfite and concentrated *in vacuo*. Crude material was purified on a large silica gel column eluted with 100% hexanes to give the title compound (6.58 g, 94%). Spectral data matched literature[182]. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32 – 7.22 (m, 2H), 6.99 (t, *J* = 7.8 Hz, 1H), 6.87 (ddd, *J* = 8.4, 2.5, 1.0 Hz, 1H), 4.09 (t, *J* = 6.0 Hz, 2H), 3.85 (t, *J* = 5.9 Hz, 2H), 2.03 (p, *J* = 6.0 Hz, 2H).^{13}C NMR (126 MHz, $CDCl_3$) δ 159.3, 130.8, 130.0, 123.6, 114.2, 94.4, 65.7, 60.2, 31.9.

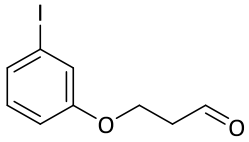


Figure E.43: Compound 5-6

3-(3-Iodophenoxy)propanal

3-(3-Iodophenoxy)propan-1-ol (2g, 7.2 mmol) and IBX (4.03 g, 14.4 mmol) were combined in a round bottom flask charged with EtOAc and brought to reflux for 9 hours. After cooling, the solution was placed in a refrigerator for 1 hour and solids filtered off. Organic liquid was concentrated *in vacuo* to give the title compound as an oil which was used without further purification (2.1g, 99%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 9.75 (t, *J* = 1.5 Hz, 1H), 7.27 – 7.16 (m, 2H), 6.96 – 6.85 (m, 1H), 6.80 (ddd, *J* = 8.4, 2.5, 1.0 Hz, 1H), 4.16 (t, *J* = 6.1 Hz, 2H), 2.80 (td, *J* = 6.0, 1.6 Hz, 2H).^{13}C NMR (126 MHz, $CDCl_3$) δ 200.1, 159.0, 131.0, 130.2, 123.6, 114.2, 94.6, 61.7, 43.0.

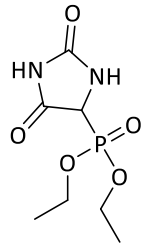


Figure E.44: Compound 5-7

Diethyl (2,5-dioxoimidazolidin-4-yl)phosphonate

Imidazolidine-2,4-dione (hydantoin) (20 g, 200 mmol) and acetic acid (80 mL) was heated to 85 °C in an oil bath. An addition funnel was charged with bromine (11.2 mL, 220 mmol) and a small amount of bromine (~2 mL) was introduced into the reaction mixture with vigorous stirring. Once the orange color had disappeared, the remainder of the bromine was added rapidly dropwise to afford a clear solution. After being stirred at 85 °C for 30 min, the reaction mixture was cooled to 30 °C in an ice bath and triethyl phosphite (47.9 mL, 280 mmol) introduced at such a rate that the internal temperature was maintained between 40-45 °C. After the addition was completed, the mixture was stirred at room temperature for 90 min. The solvent was removed *in vacuo* and the residue diluted with diethyl ether (80 mL) with stirring to induce precipitation of a white solid. The mixture was poured onto diethyl ether (200 mL) with vigorous stirring. After 30 min, filtration afforded the title compound (25.6 g, 54.2%) as a white solid, which matched literature reports[183] and was used without further purification.

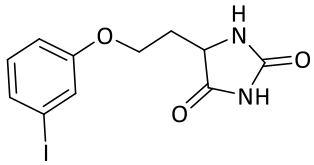


Figure E.44: Compound 5-11

5-(2-(3-Iodophenoxy)ethyl)imidazolidine-2,4-dione

To a solution of 1-iodo-3-(prop-2-yn-1-yloxy)benzene (1 g, 3.9 mmol) in dry DCM (*ca.* 35 mL) was added borane dimethyl sulfide solution (2.75 mL, 4.5 mmol) at room temperature in the dark under inert atmosphere in one portion. The reaction was stirred vigorously for 0.5 hr, after which 1 mL of MeOH was added carefully, followed by 10% NaOH (4 mL) and 30% peroxide solution (4 mL). This solution was left to stir for 20 hours before extraction with DCM. The organic layers were combined and washed with 10% HCl, brine, and dried over $MgSO_4$ before being concentrated in vacuo. The crude is combined with potassium cyanide (0.507 g, 7.8 mmol), and ammonium carbonante (1.5 g, 15.6 mmol) in a round bottom flask in ethanol/water (1:1) mixture and sealed with a rubber septum. The slurry is heated to 70°C and stirred 24 hours. After cooling to room temperature, the solution was concentrated to half volume and acidified to pH <6 with concentrated hydrochloric acid. Solution is extracted with diethyl ether and concentrated *in vacuo*. DCM is added to the crude yellow oil and filtered to collect the white precipitate. The precipitate is dried *in vacuo* to give a white solid (0.030 g, 2.4%). ^{1}HNMR (DMSO-d_6, 500 MHz): δ 10.65 (s, 1NH), δ 8.02 (s, 1H), δ 7.26-7.27 (m, 2H), δ 7.06 (t, J= 7.8 Hz, 1H), δ 6.94 (d, J=7.8 Hz, 1H), δ 4.06 (t, J=5.6 Hz, 2H), δ 2.11-2.07 (m, 1H), δ 1.95-1.90 (m, 1H). ^{13}CNMR (DMSO-d_6, 125 MHz): δ 176.4, 159.5, 157.9, 131.8, 130.0, 123.6, 114.8, 95.5, 64.3, 55.2, 31.3. MS (ESI+): *m/z* calculated $[M+H]^+$ $C_{11}H_{12}IN_2O_3^+$ 346.99, found 347.00

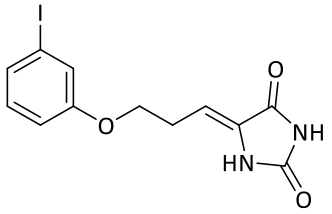


Figure E.45: Compound 5-12

(E/Z)-5-(3-(3-Iodophenoxy)propylidene)imidazolidine-2,4-dione

Diethyl (2,5-dioxoimidazolidin-4-yl)phosphonate (2.13g, 9 mmol) is dissolved in ethanol under an inert atmosphere. Sodium metal (216 mg, 9 mmol) is added and mixture stirred for 20 minutes at room temperature. Then, 3-(3-iodophenoxy)propanal (2 g, 7.2 mmol) is added as a solution in ethanol and reaction stirred for 20 minutes. After which, the mixture is poured into 10% HCl solution to precipitate the title compound as a mixture of E/Z isomers (1.18g, 45.7%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 11.00 (s, 1NH), 10.27 (s, 1NH), 7.28 (bs, 2H), 7.17 – 7.00 (m, 1H), 7.00 – 6.89 (m, 1H), 5.55 (t, J = 7.7 Hz, 1H), 4.03 (dt, J = 13.8, 6.5 Hz, 2H), 2.98 (q, J = 6.8 Hz, 1H), 2.58 (q, J = 6.8 Hz, 1H). ^{13}C NMR (126 MHz, DMSO-d_6) δ 165.1, 164.7, 159.6, 159.6, 155.3, 154.4, 132.2, 131.8, 131.0, 130.0, 128.1, 124.4, 123.4, 123.4, 115.4, 114.9, 114.9, 112.5, 107.7, 95.5, 67.5, 66.8, 26.7, 25.8.

B. Proteomic Methods

Determination of the proteasome binding site was performed as described. Compound **4-7** was incubated with human proteasome in HEPES Buffer (pH 7.4, 100 mM NaCl) with a drop of DMSO (to solubilize **4-7** in the aqueous media) for 1 hour at 37 °C before being frozen in a -80 °C freezer for an additional hour. This was taken to the metabolomics core where it was thawed, digested with trypsin and subjected to mass spectrometry and computational analysis. Analysis constraints were as follows:

DATABASE SEARCHING-- Tandem mass spectra were extracted by [unknown] version [unknown]. Charge state deconvolution and deisotoping were not performed. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.6.0) and X! Tandem (The GPM, thegpm.org; version X! Tandem Alanine (2017.2.1.4)). Mascot was set up to search the UP_human_crap_20171102 database (unknown version, 160685 entries) assuming the digestion enzyme stricttrypsin. X! Tandem was set up to search a reverse concatenated subset of the UP_human_crap_20171102 database (unknown version, 1404 entries) also assuming strict trypsin. Mascot and X! Tandem were searched with a fragment ion mass tolerance of 0.020 Da and a parent ion tolerance of 10.0 PPM. Carbamidomethyl of cysteine was specified in Mascot and X! Tandem as a fixed modification. Glu->pyro-Glu of the n-terminus, ammonia-loss of the n-terminus, gln->pyro-Glu of the n-terminus, oxidation of methionine and CJones_mod of cysteine and lysine were specified in X! Tandem as variable modifications. Oxidation of methionine and CJones_mod of cysteine and lysine were specified in Mascot as variable modifications.

CRITERIA FOR PROTEIN IDENTIFICATION-- Scaffold (version Scaffold_4.8.9, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 17.0% probability by the Scaffold Local FDR algorithm. Protein identifications were accepted if they could be established at greater than 80.0% probability to achieve an FDR less than 1.0% and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii, Al et al Anal. Chem. 2003;75(17):4646-58). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters.

Tepe - Jones Mod Data 20190122, Publication report created on 04/12/2019"

Experiment: Tepe - Jones Mod Data 20190122
- Peak List Generator: unknown
  - Version: unknown
  - Charge States Calculated: unknown
  - Deisotoped: unknown
  - Textual Annotation: unknown
- Database Set: 2 Databases
  - Database Name: a reverse concatenated subset of the UP_human_crap_20171102 database
    - Version: unknown
    - Taxonomy: All Entries
    - Number of Proteins: 1404
  - Database Name: the UP_human_crap_20171102 database
    - Version: unknown
    - Taxonomy: All Entries
    - Number of Proteins: 160685
  - Does database contain common contaminants?: unknown
- Search Engine Set: 2 Search Engines
  - Search Engine: Mascot
    - Version: 2.6.0
    - Samples: All Samples
    - Fragment Tolerance: 0.020 Da (Monoisotopic)

Parent Tolerance: 10.0 PPM (Monoisotopic)
Fixed Modifications: +57 on C (Carbamidomethyl)
"Variable Modifications: +16 on M (Oxidation), +349 on CK (CJones_mod)"
"Database: the UP_human_crap_20171102 database (unknown version, 160685 entries)"
Digestion Enzyme: stricttrypsin
Max Missed Cleavages: 2
Probability Model:
"Jones .temp (F012624): LFDR Model, Classifier data: Bayes, Good (78%) m:22.9/s:24.9 m:56.1/s:25.3 m:NA m:NA m:NA, Bad (22%) m:-26.1/s:7.66 m:6.64/s:7.02 m:NA m:NA m:NA [all charge states]"
Search Engine: X! Tandem
Version: X! Tandem Alanine (2017.2.1.4)
Samples: All Samples
Fragment Tolerance: 0.020 Da (Monoisotopic)
Parent Tolerance: 10.0 PPM (Monoisotopic)
Fixed Modifications: +57 on C (Carbamidomethyl)
"Variable Modifications: -18 on Peptide N-Terminal (Glu->pyro-Glu), -17 on Peptide N-Terminal (Ammonia-loss), -17 on Peptide N-Terminal (Gln->pyro-Glu), +16 on M (Oxidation), +349 on CK (CJones_mod)"
"Database: a reverse concatenated subset of the UP_human_crap_20171102 database (unknown version, 1404 entries)"
Digestion Enzyme: stricttrypsin
Max Missed Cleavages: 2
Probability Model:
"Jones .temp (F012624): LFDR Model, No Classifier [all charge states]"
Scaffold: Version: Scaffold_4.8.9
Modification Metadata Set: 2334 modifications
Source: C:\Program Files\Scaffold 4\parameters\unimod.xml
Comment:
Protein Grouping Strategy: Experiment-wide grouping with protein cluster analysis
Peptide Thresholds: 17.0% minimum
Protein Thresholds: 80.0% minimum and 2 peptides minimum
Peptide FDR: 0.0% (Decoy)
Protein FDR: 0.0% (Decoy)
GO Annotation Source(s):
Alternate ID Source(s):

C. Proteasome Activity Assays

The activity assays were carried out in a 100 μL reaction volume. Different concentrations of test compounds were added to a black flat/clear bottom 96-well plate containing 1 nM of human constitutive 20S proteasome, in 50 mM Tris-HCl at pH 7.8 and allowed to sit for 15 min at 37 °C. Fluorogenic substrates were then added and the enzymatic activity measured at 37 °C on a SpectraMax M5e spectrometer by measuring the change in fluorescence unit per 5 minutes for 1 h at 380−460 nm. The fluorescence units for the vehicle control were set at a 100%, and the ratio of drug-treated sample set to that of vehicle control was used to calculate the fold change in enzymatic activity. Fold activity was plotted as a function of drug concentration, using Origin Pro 9. The fluorogenic substrates used were Suc-LLVY-AMC (CT-L activity, 10 μM), Z-LLE-AMC (Casp-L activity, 10 μM), and Boc-LRR-AMC (T-L activity, 10 μM).

Combination

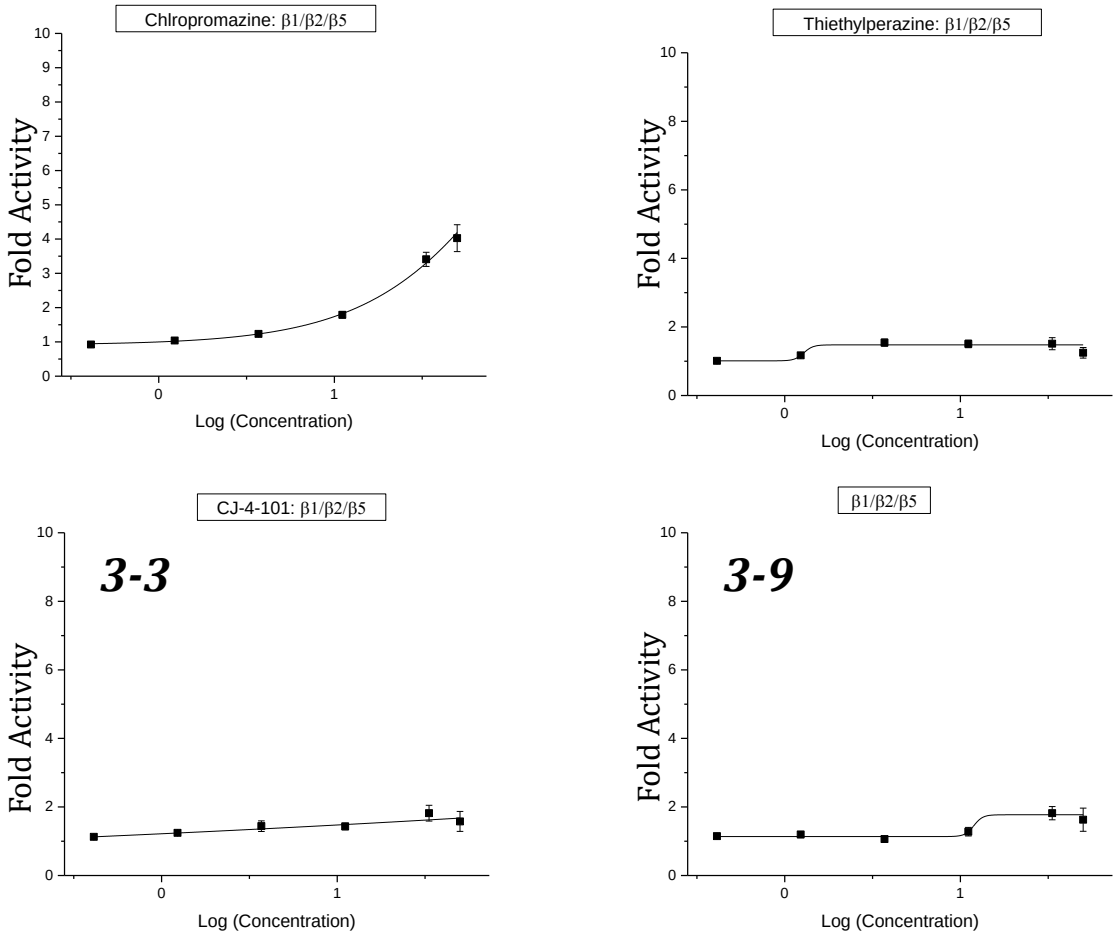


Figure E.46: Dose Response of Combination Set 1

Combination

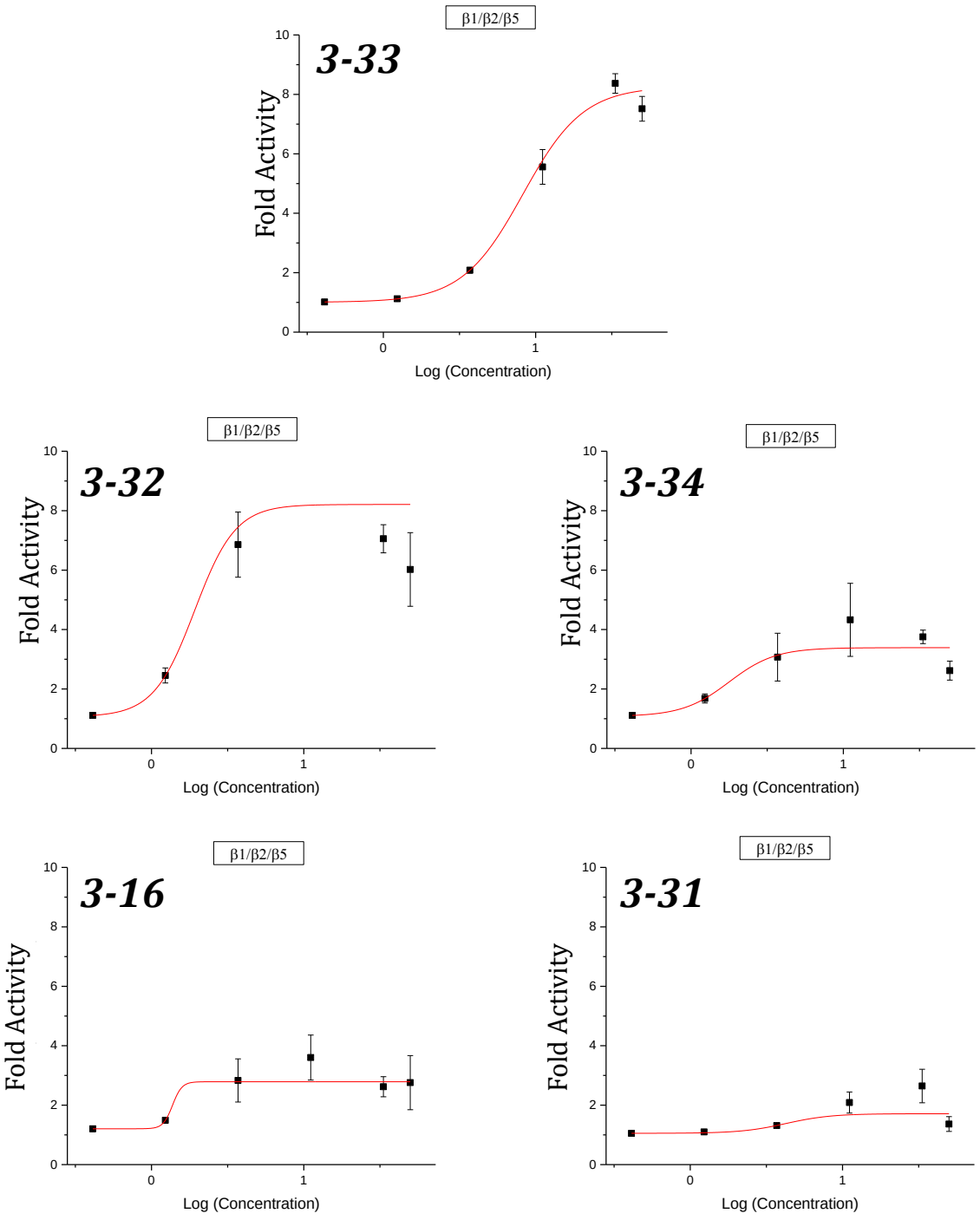


Figure E.47: Dose Response of Combination Set 2

Combination

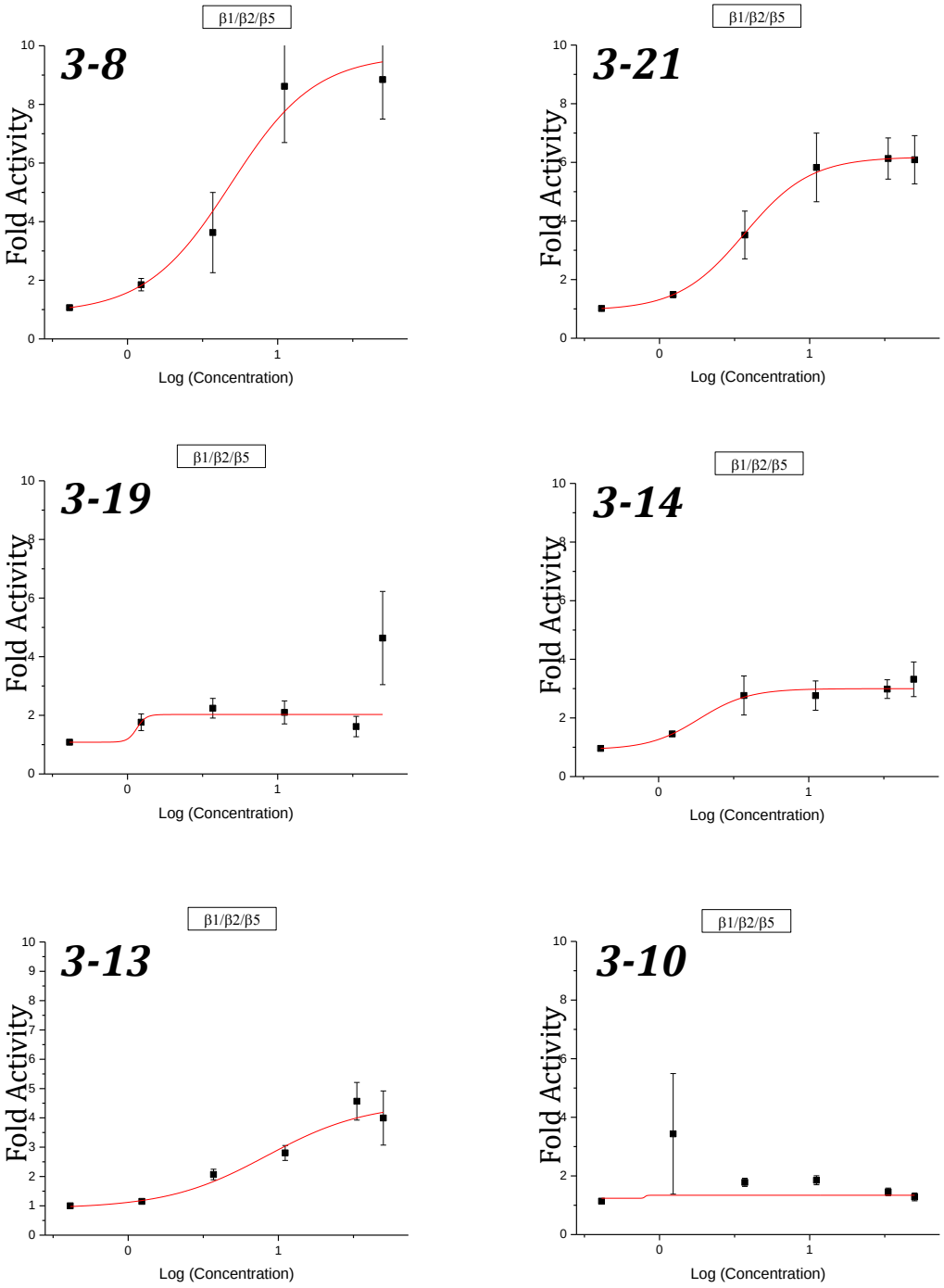


Figure E.48: Dose Response of Combination Set 3

β1-Site Activities

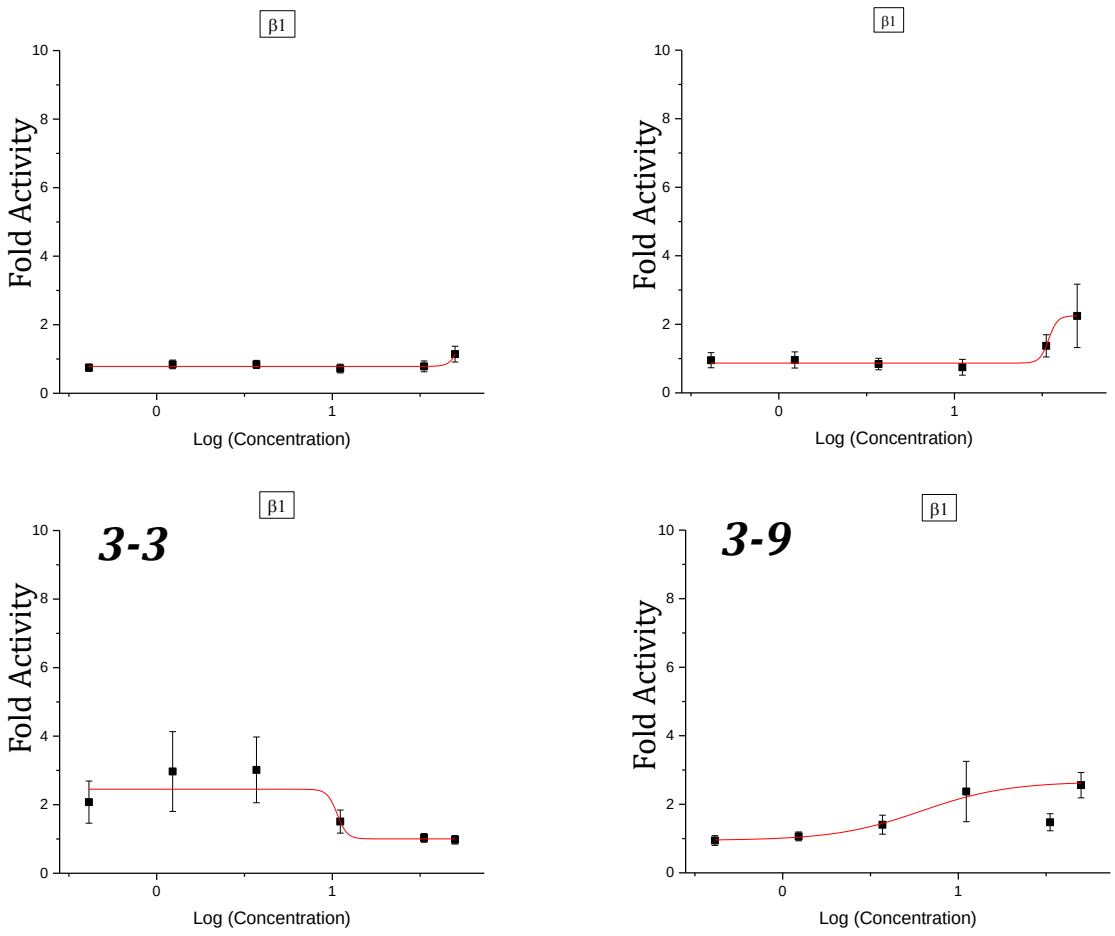


Figure E.49: Dose Response of B1 Set 1

β1-Site Activities

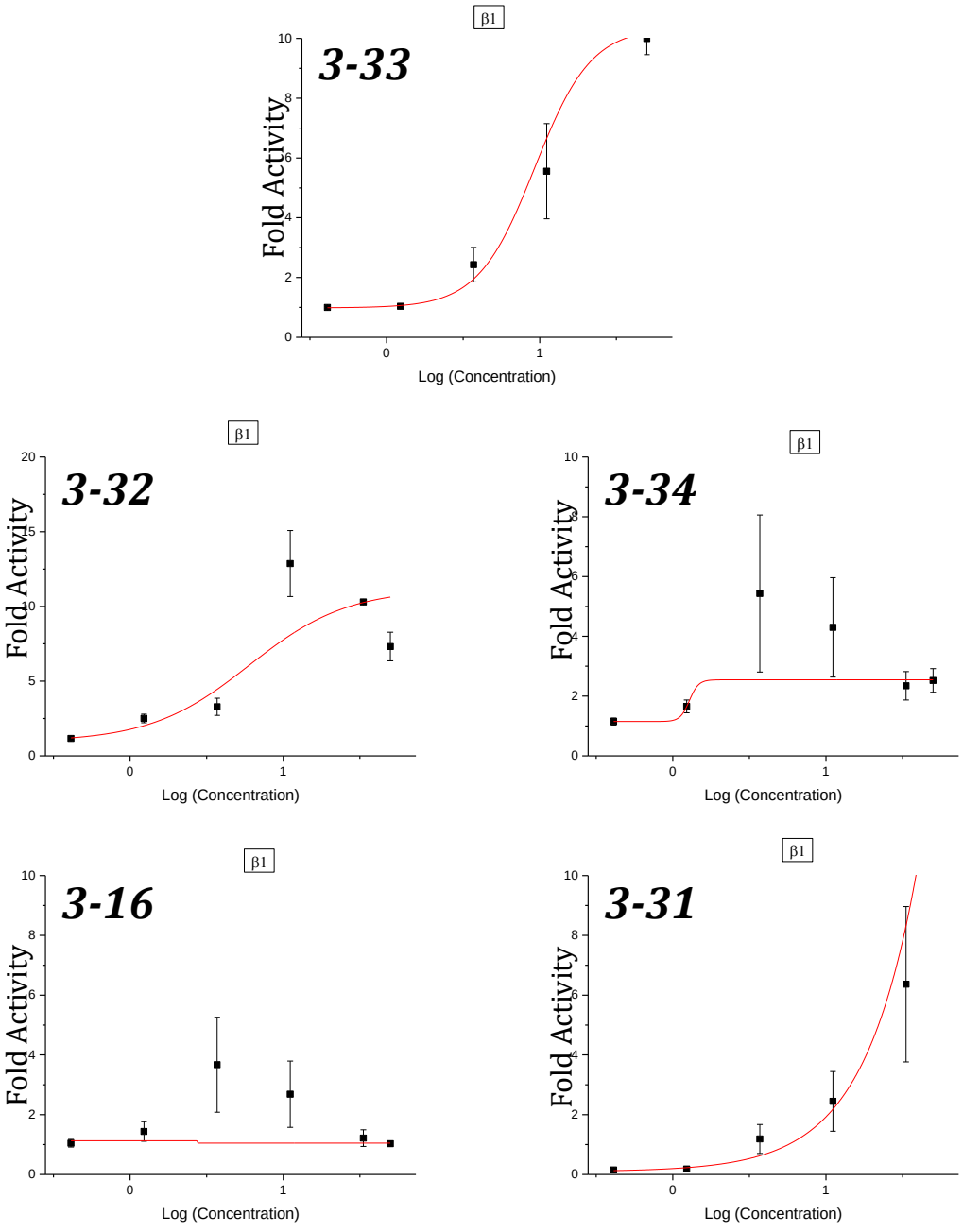


Figure E.50: Dose Response of B1 Set 2

β1-Site Activities

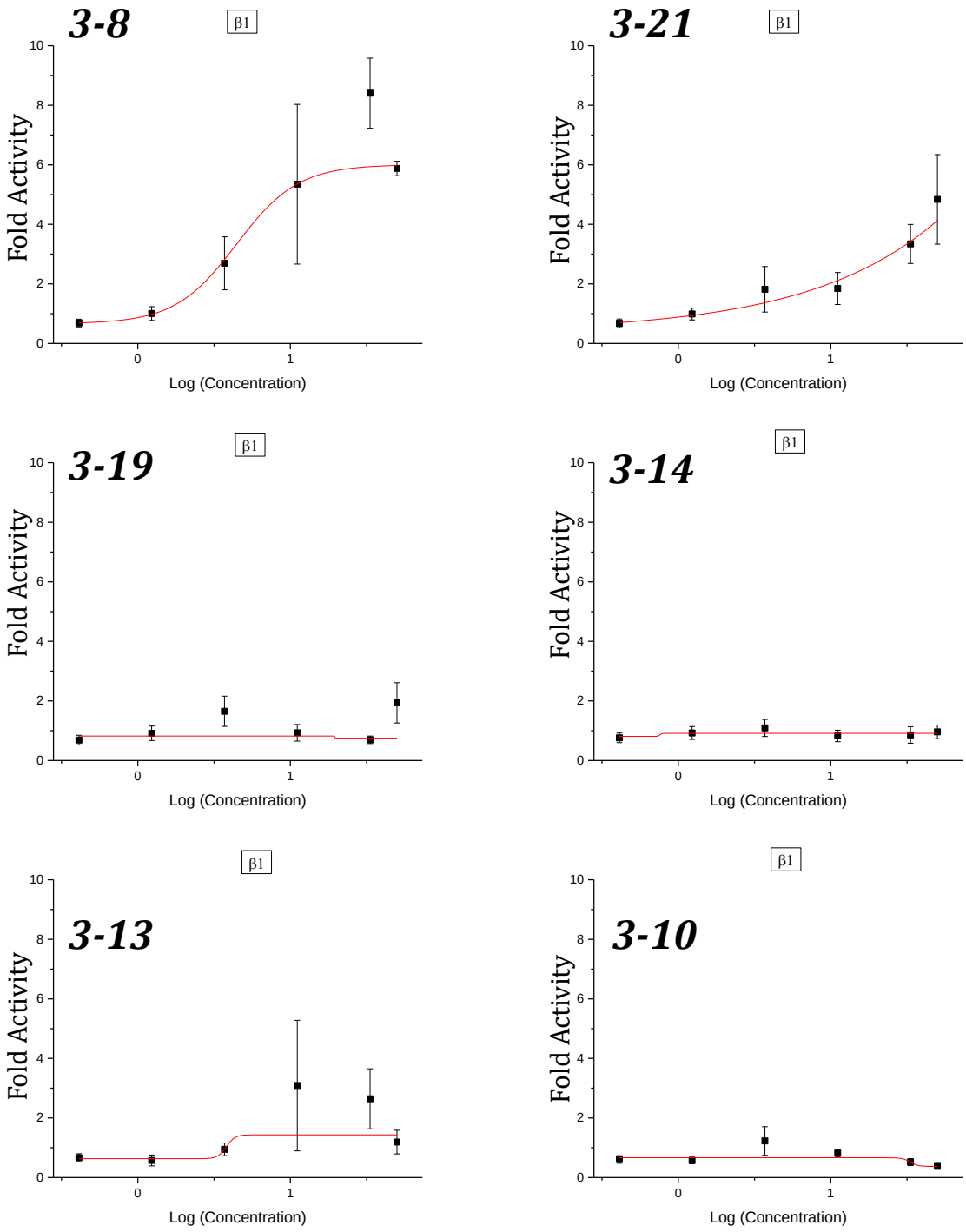


Figure E.51: Dose Response of B1 Set 3

β2-Site Activities

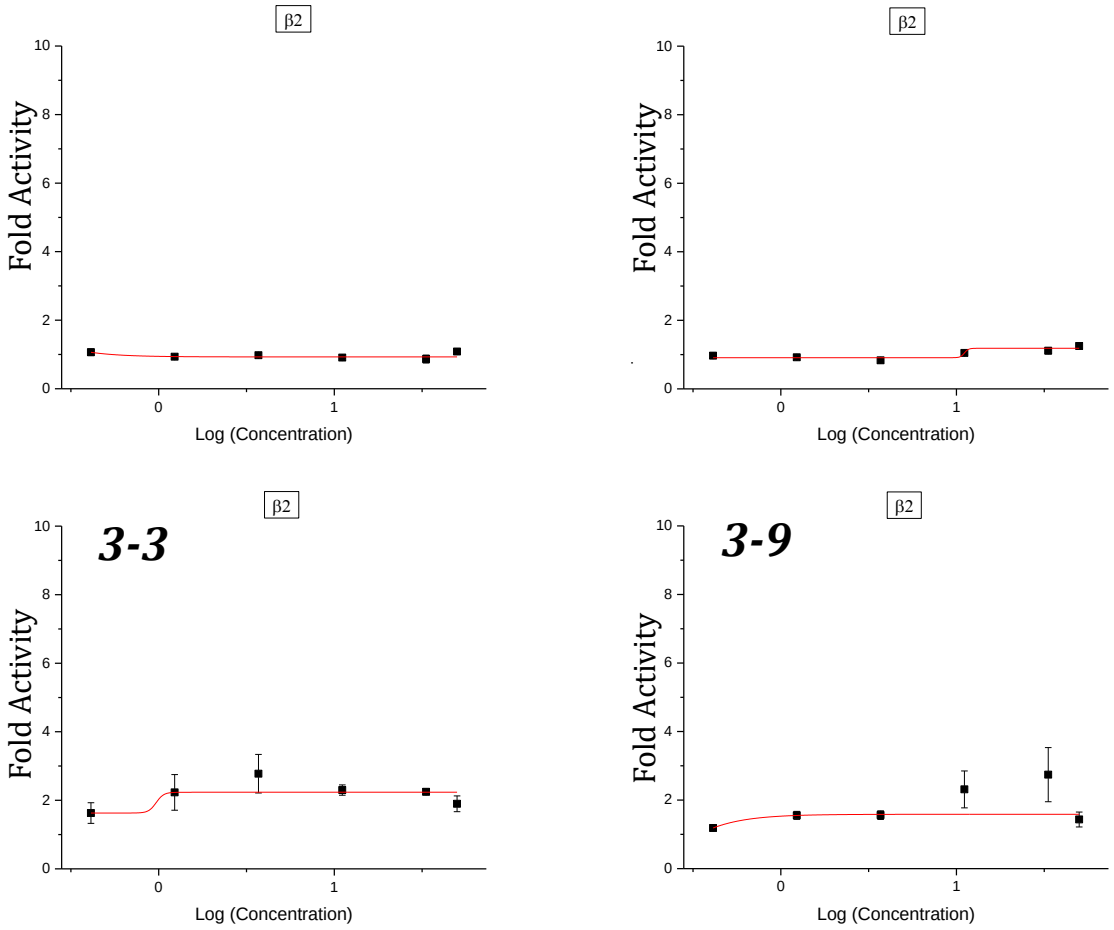


Figure E.52: Dose Response of B2 Set 1

β2-Site Activities

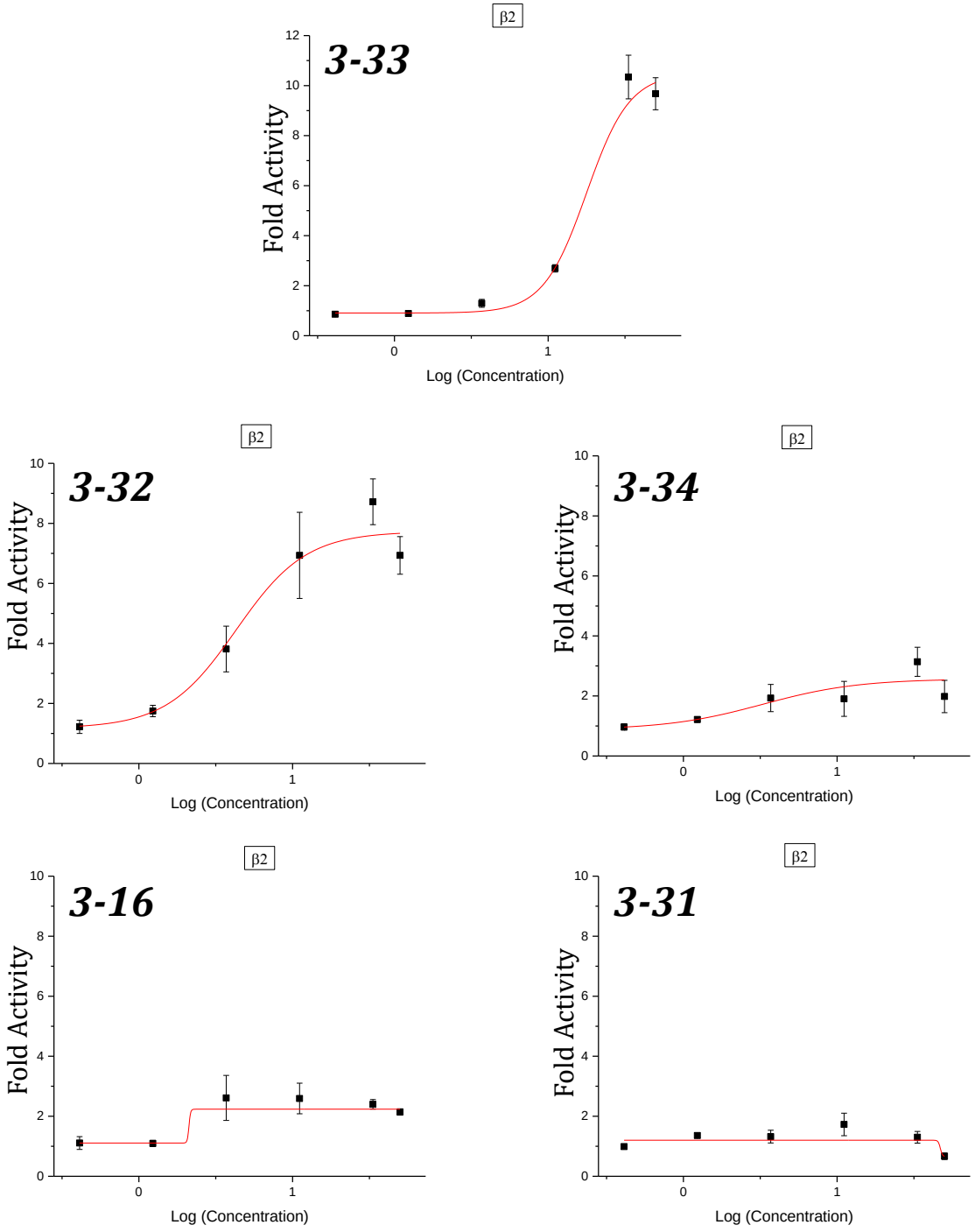


Figure E.53: Dose Response of B2 Set 2

β2-Site Activities

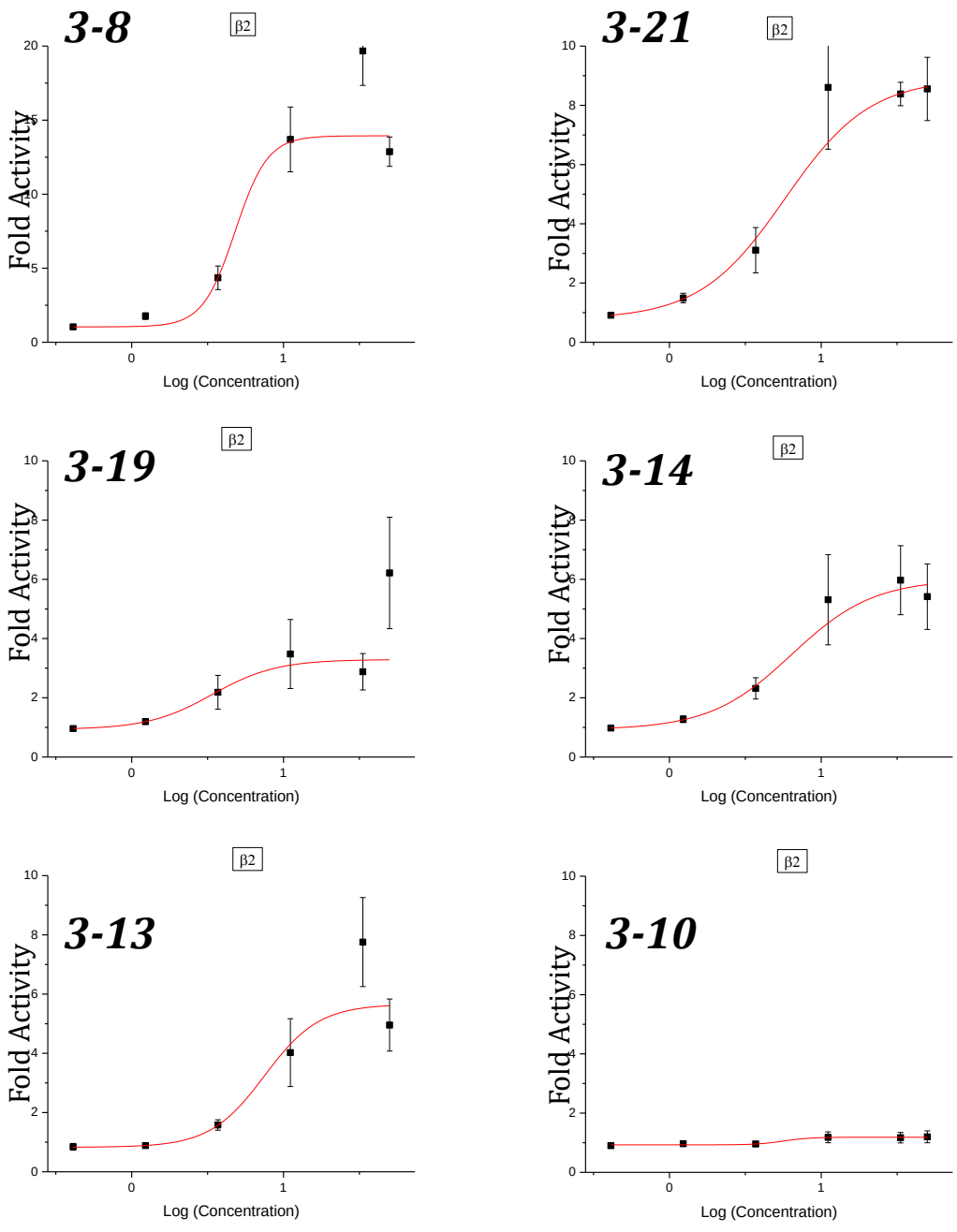


Figure E.54: Dose Response of B2 Set 3

β5-Site Activities

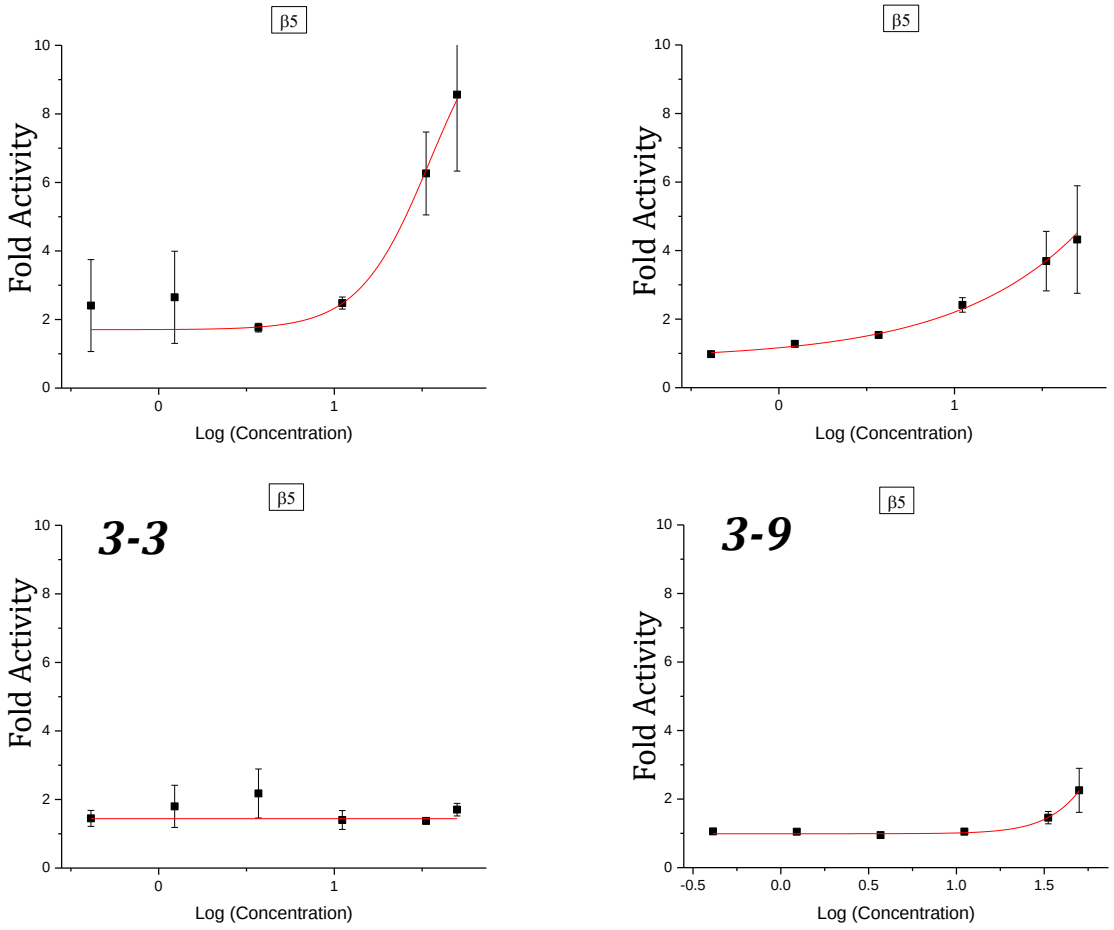


Figure E.55: Dose Response of B5 Set 1

β5-Site Activities

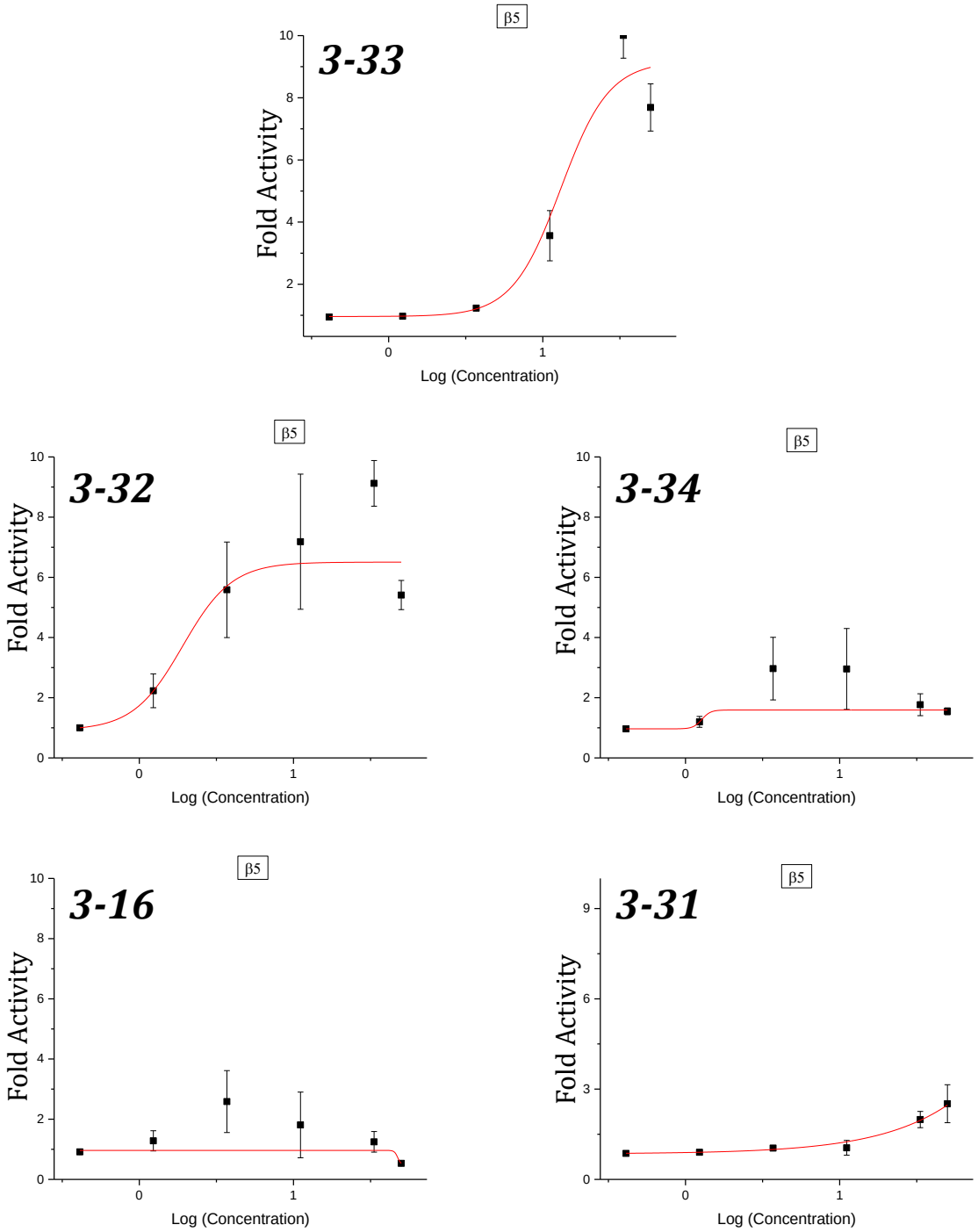


Figure E.56: Dose Response of B5 Set 2

β5-Site Activities

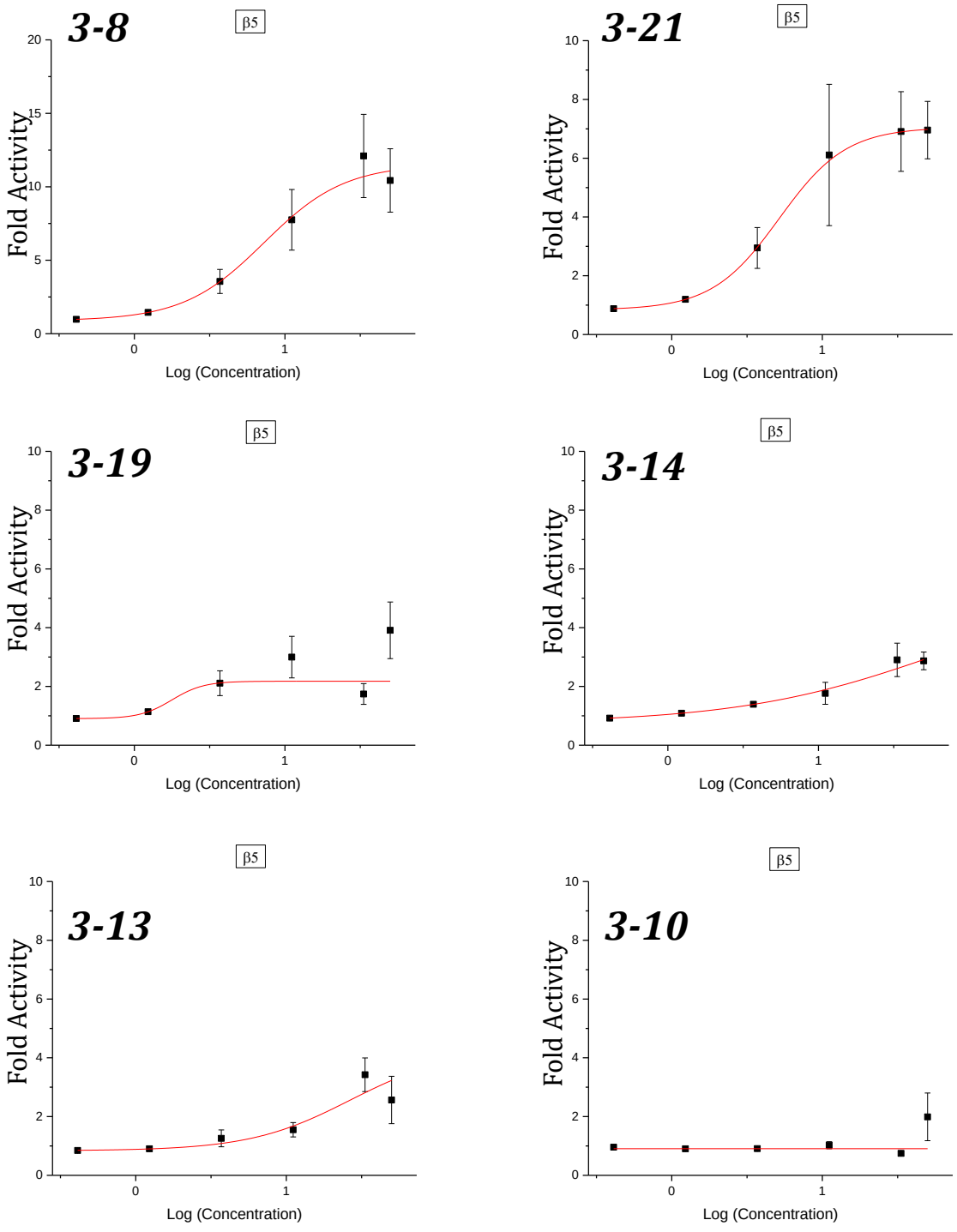


Figure E.57: Dose Response of B5 Set 3

APPENDIX

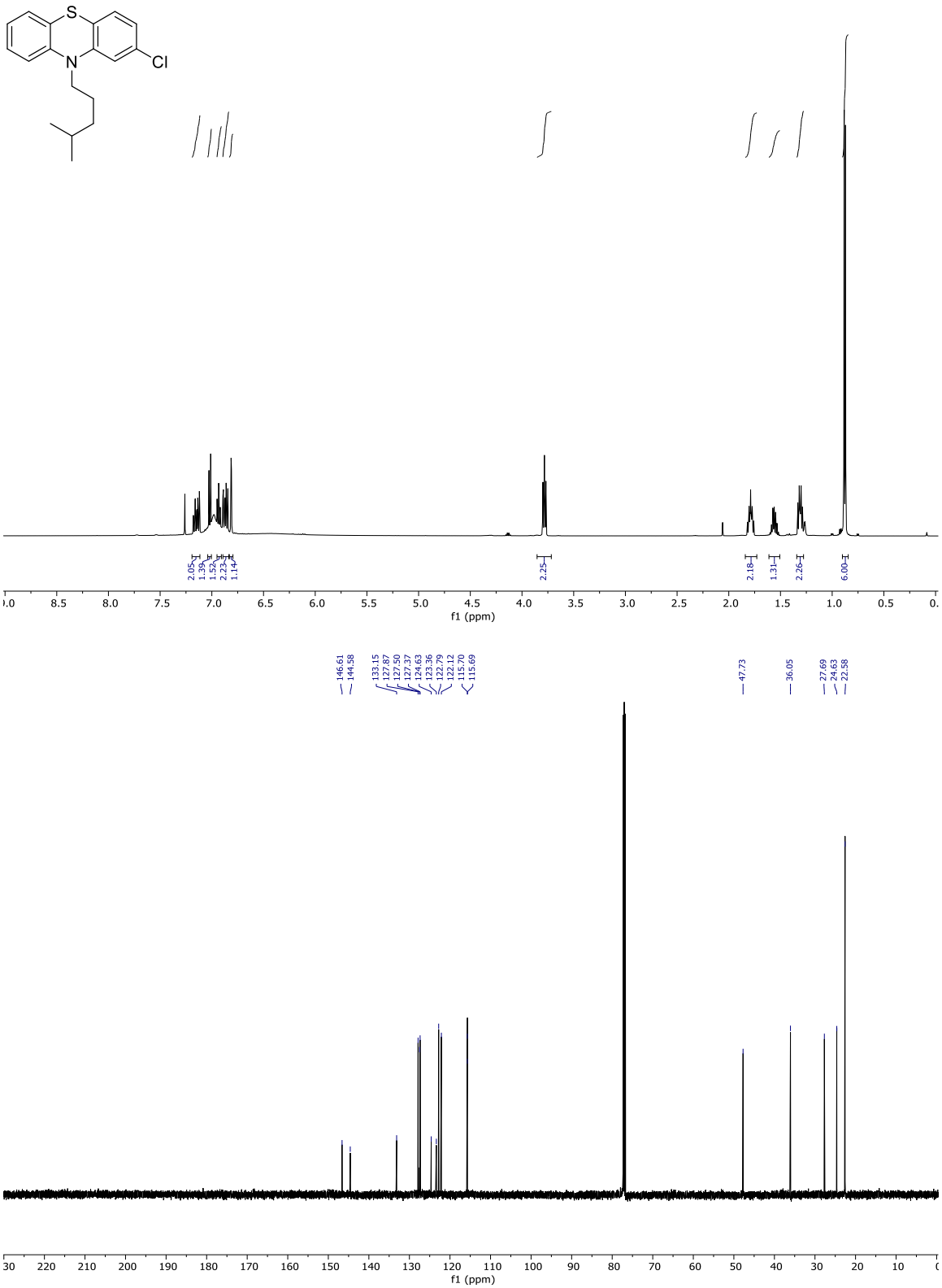


Figure A.1: ^{1}H and ^{13}C NMR Spectra of Compound 3-1

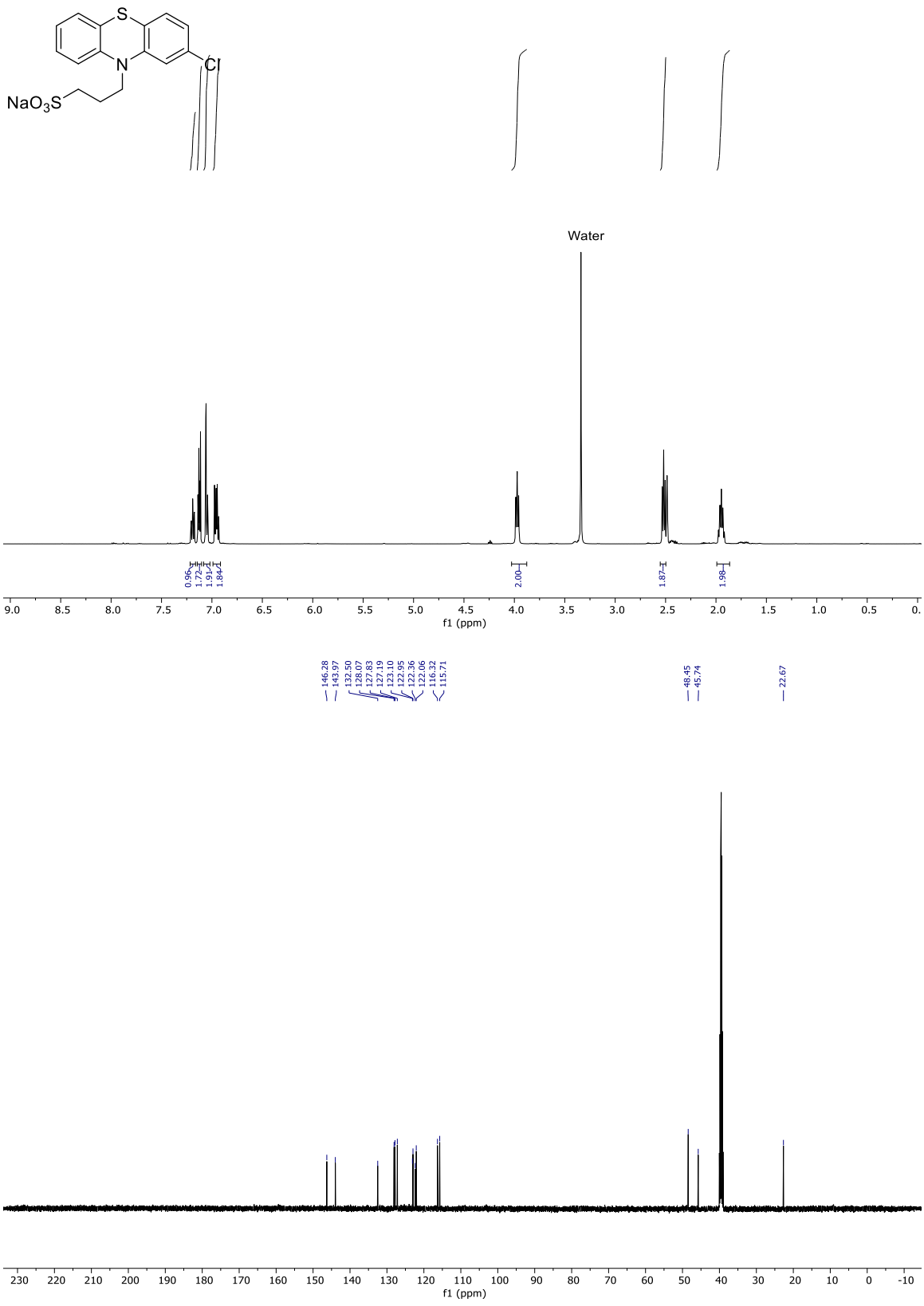


Figure A.2: ^{1}H and ^{13}C NMR Spectra of Compound 3-2

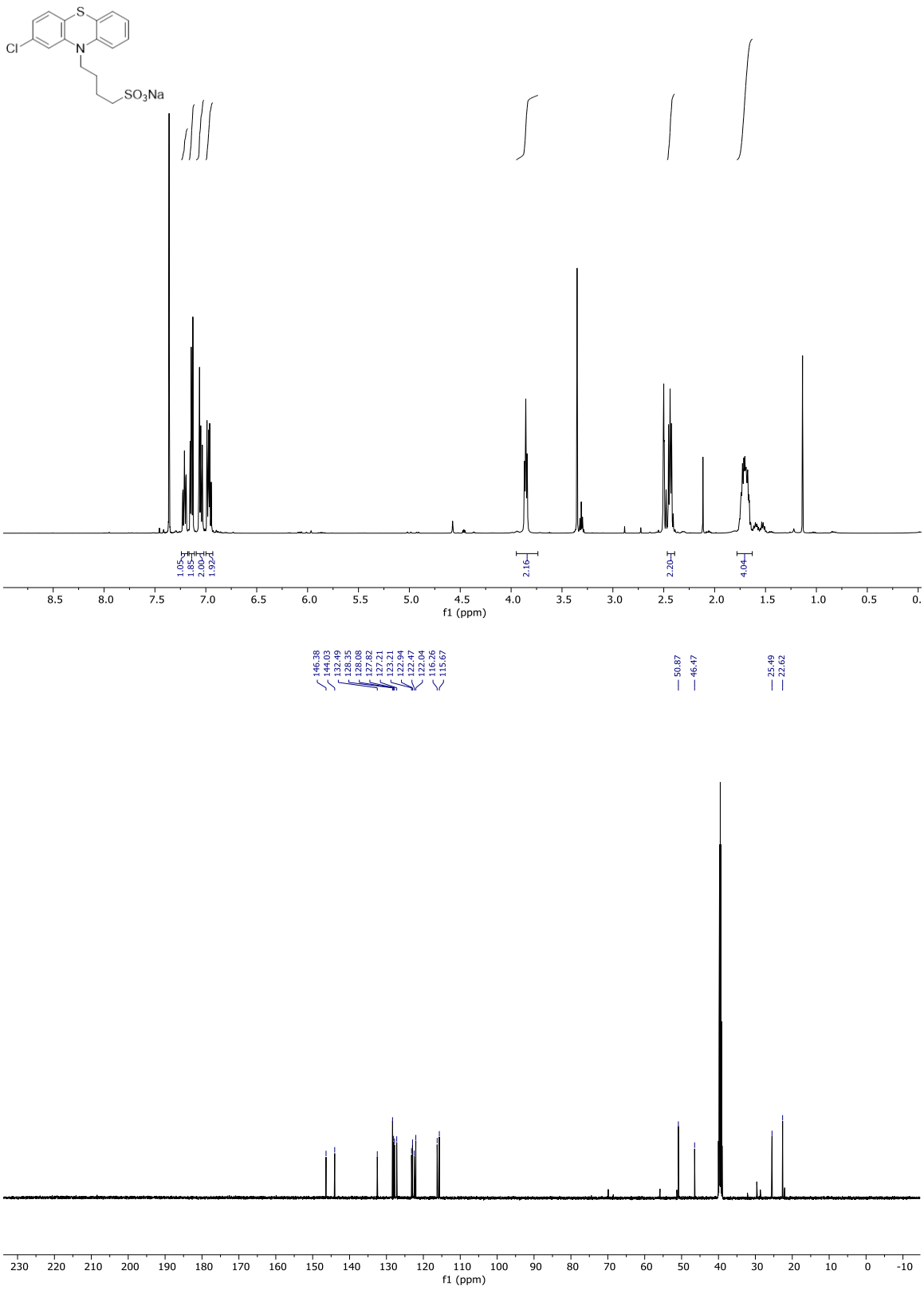


Figure A.3: ^{1}H and ^{13}C NMR Spectra of Compound 3-3

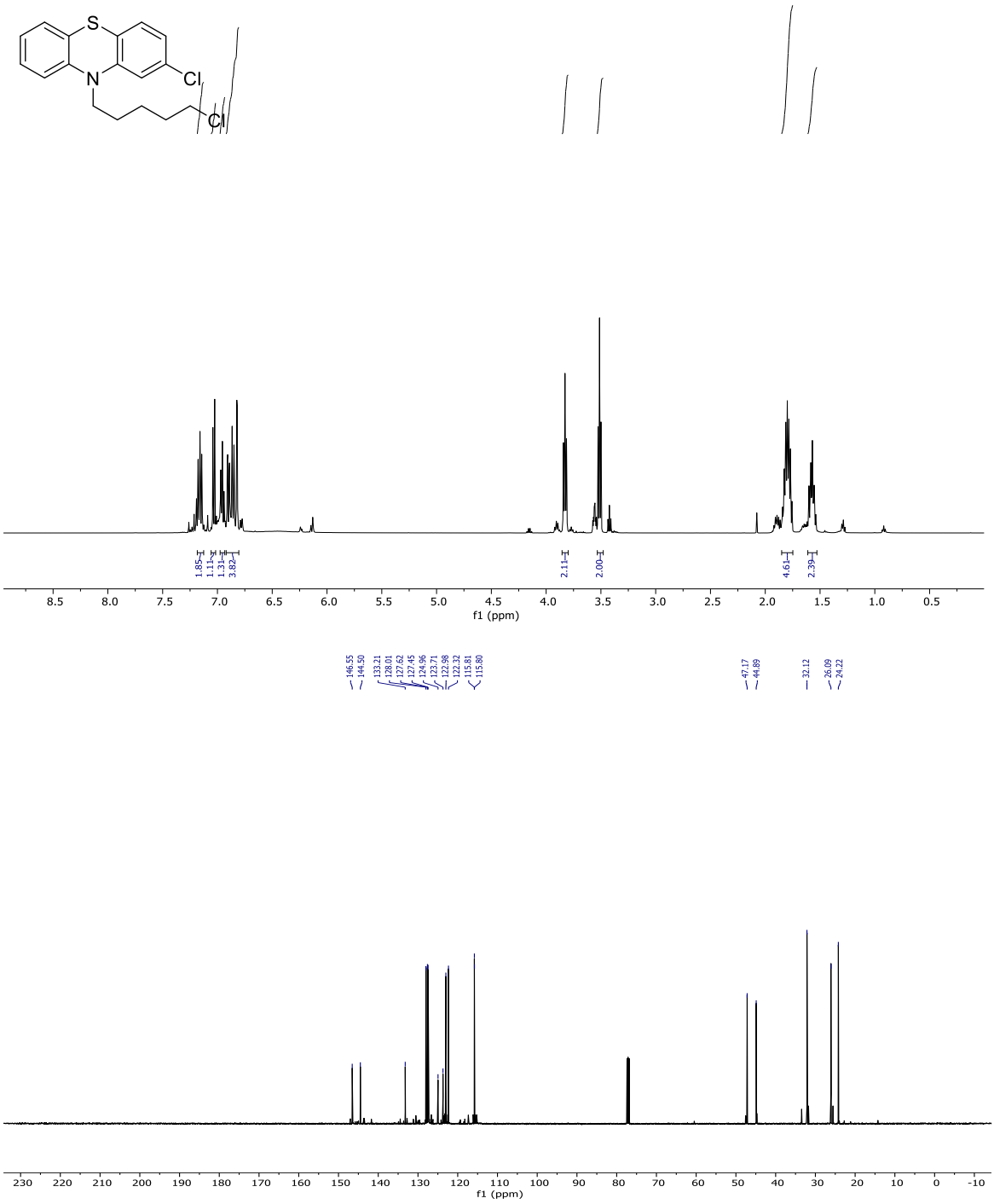


Figure A.4: ^{1}H and ^{13}C NMR Spectra of Compound 3-4

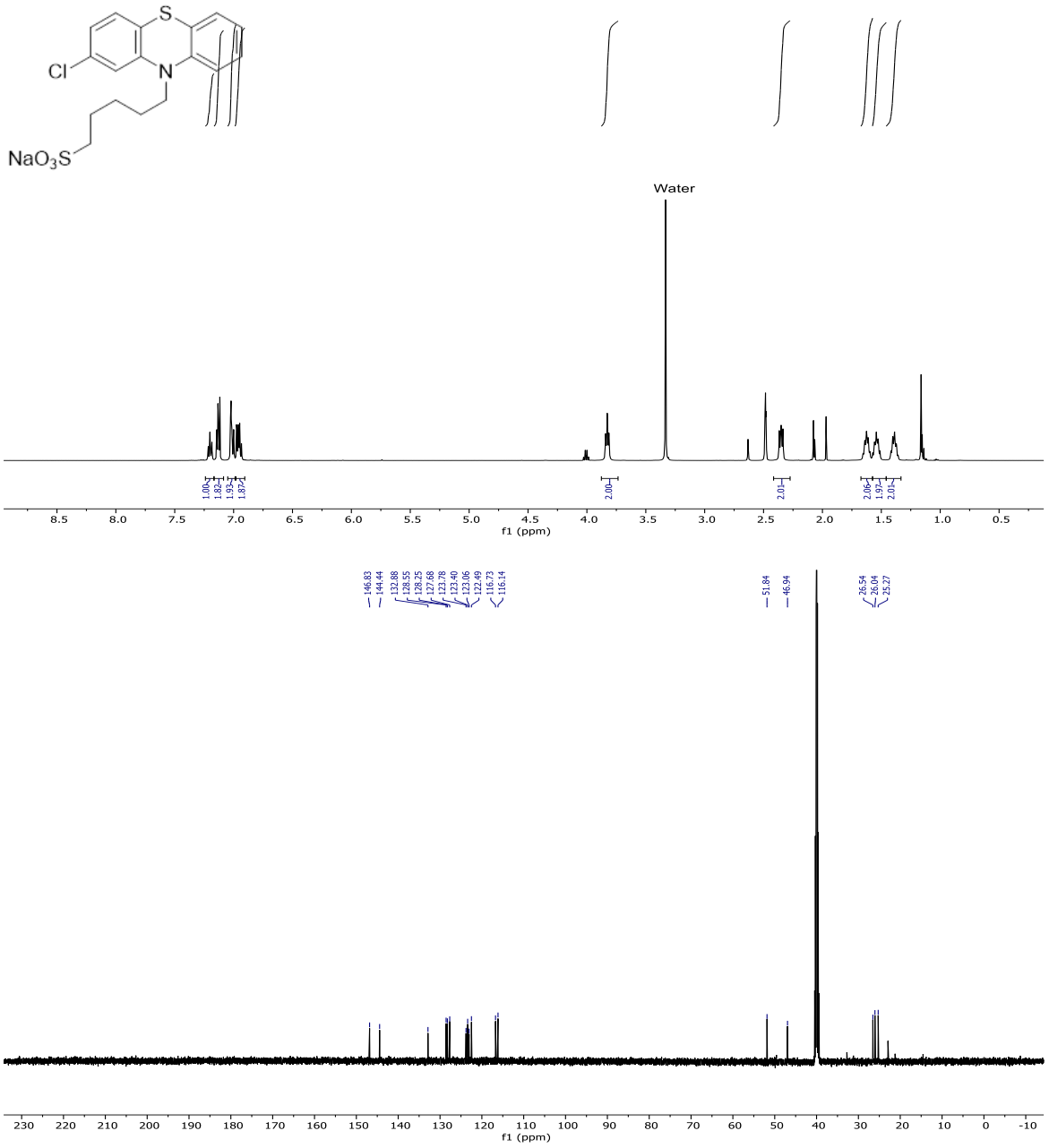


Figure A.5: ^{1}H and ^{13}C NMR Spectra of Compound 3-5

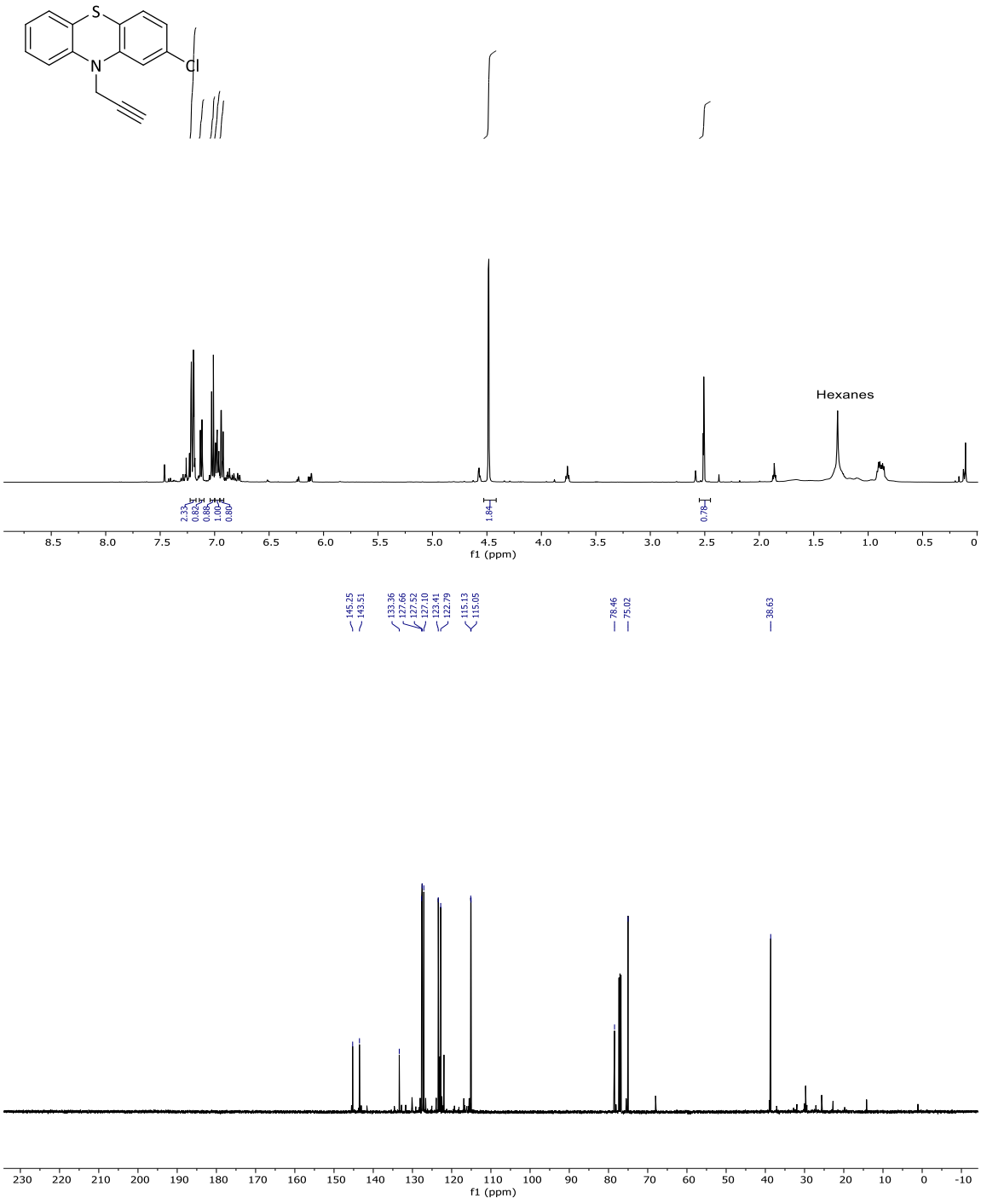


Figure A.6: ^{1}H and ^{13}C NMR Spectra of Compound 3-6

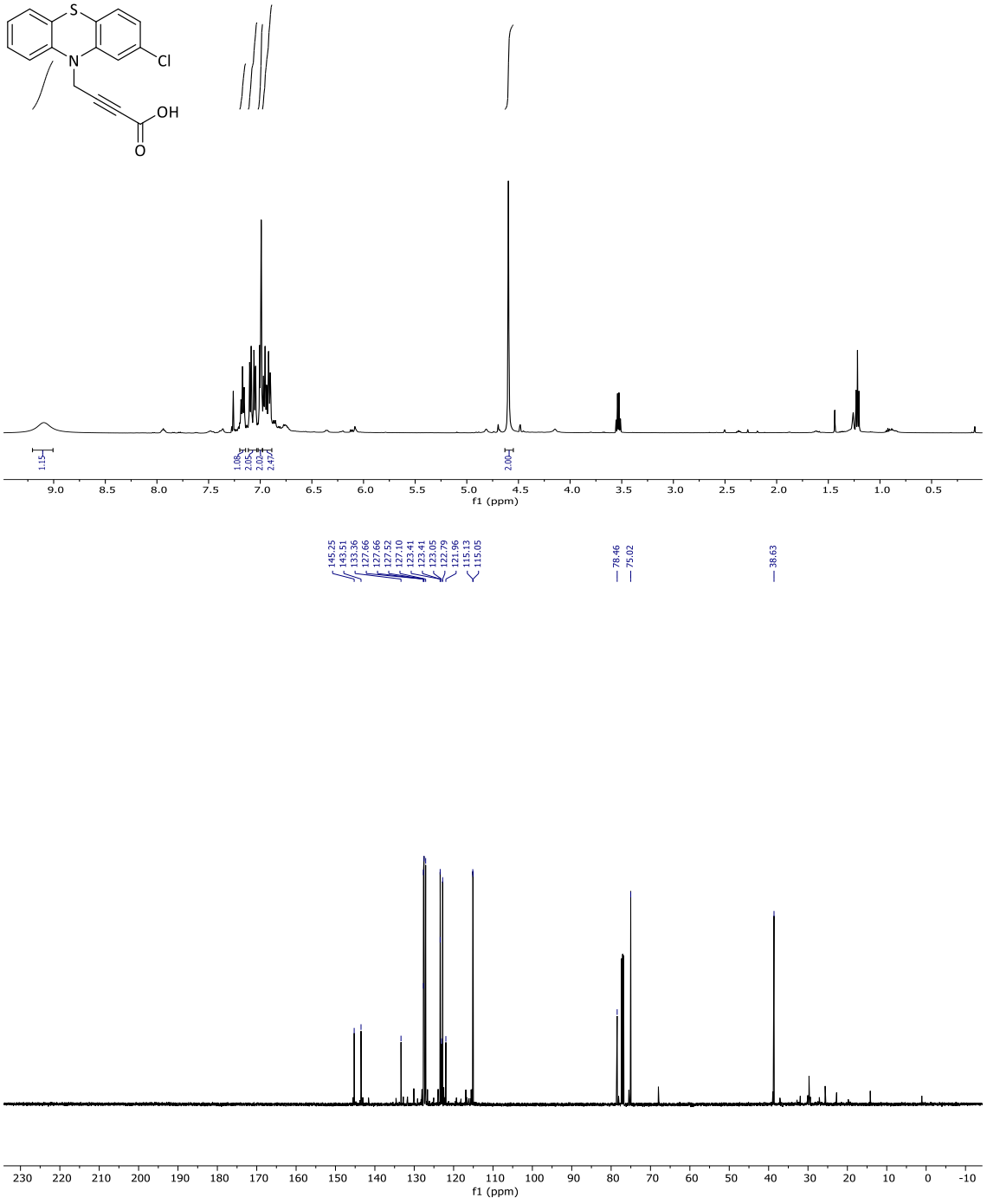


Figure A.7: ^{1}H and ^{13}C NMR Spectra of Compound 3-7

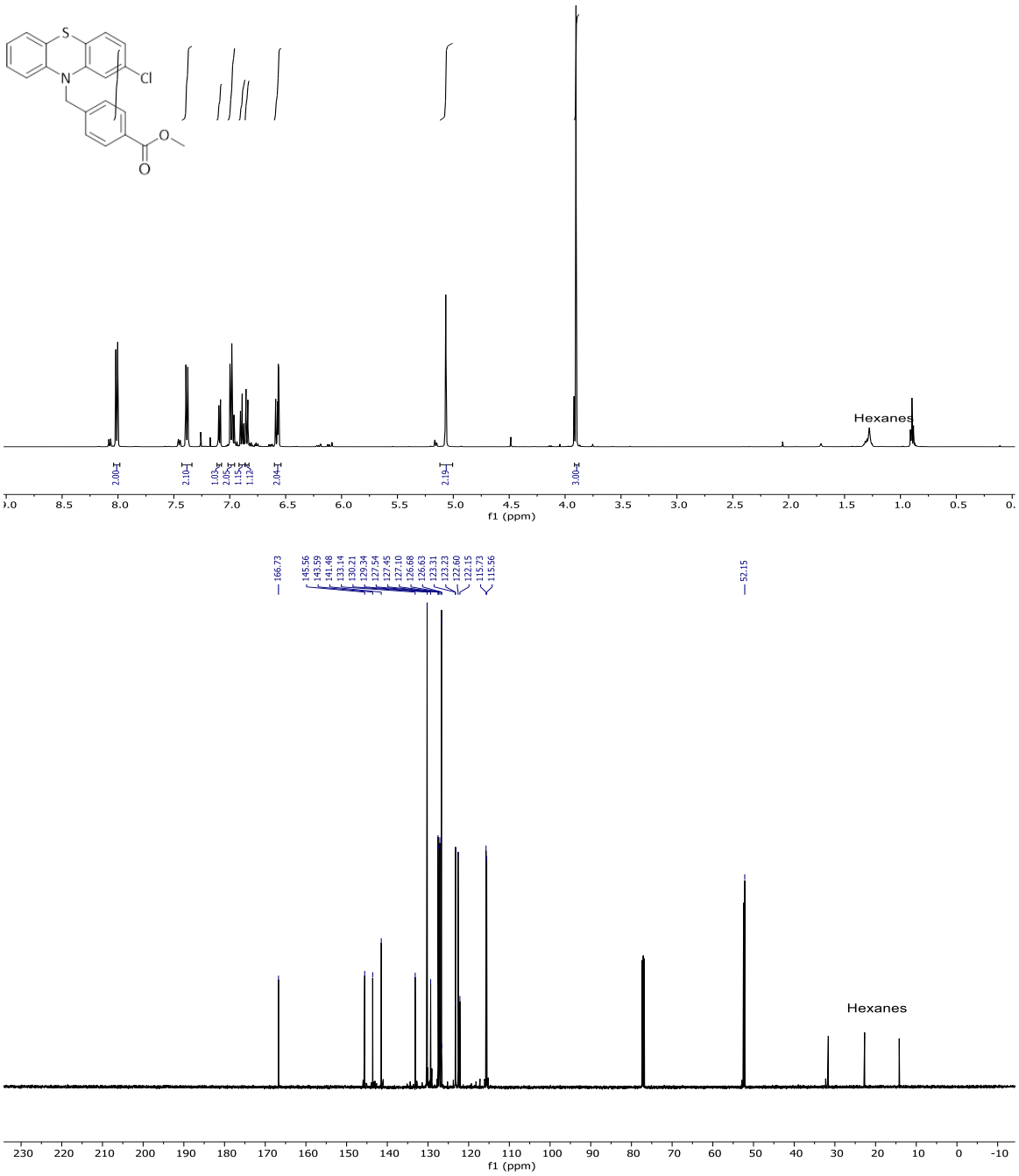


Figure A.8: ^{1}H and ^{13}C NMR Spectra of Compound 3-8

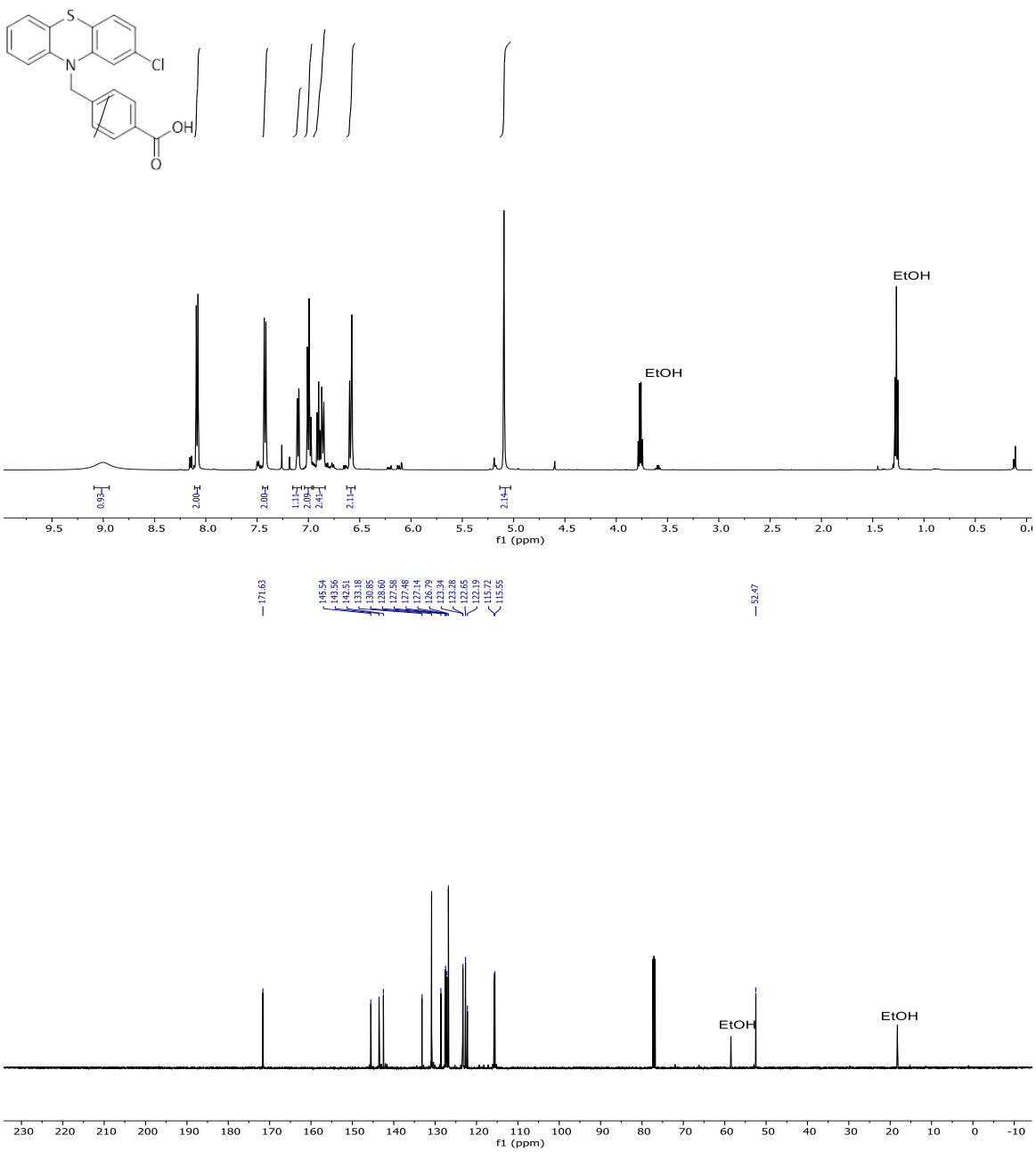


Figure A.9: ^{1}H and ^{13}C NMR Spectra of Compound 3-9

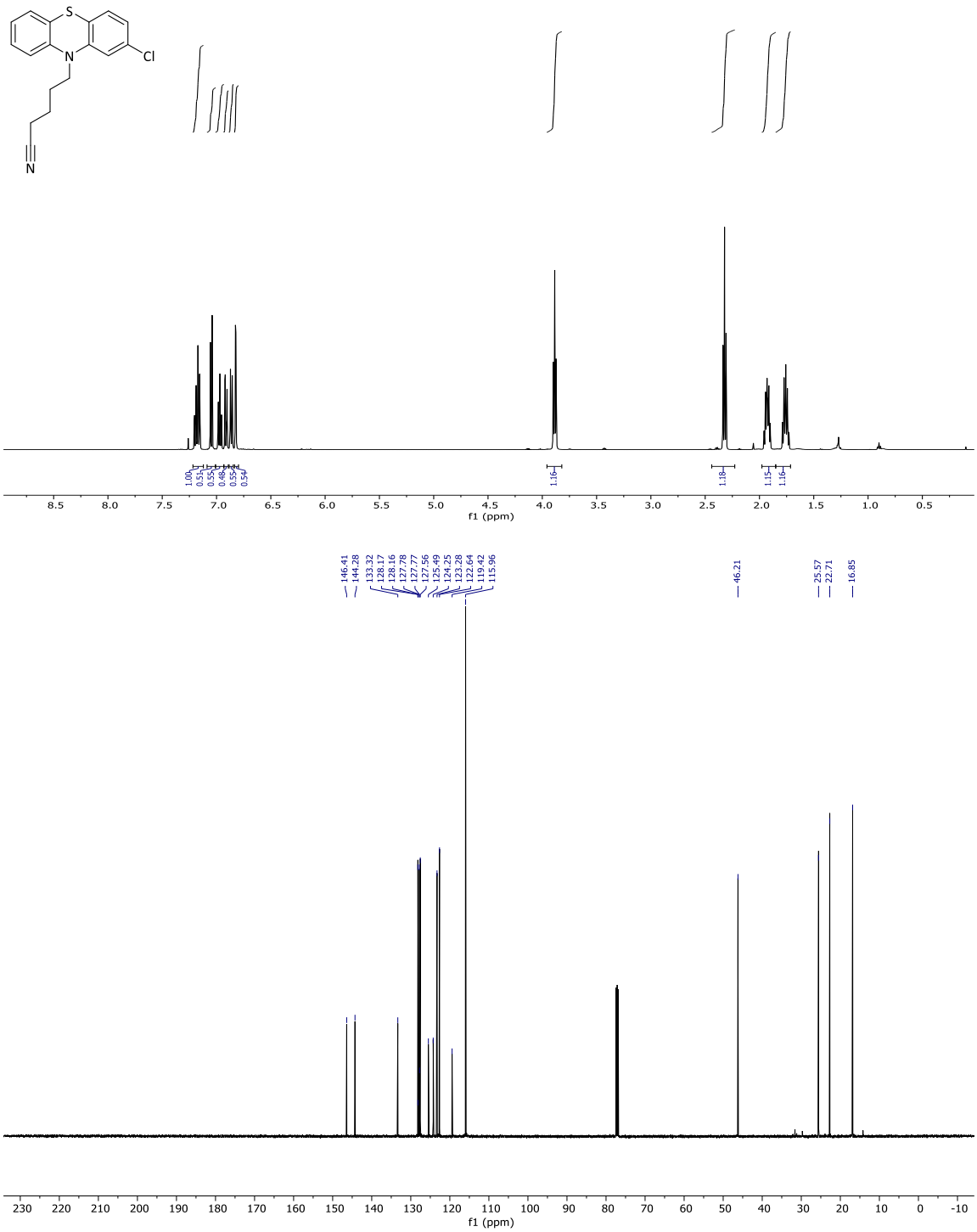


Figure A.10: ^{1}H and ^{13}C NMR Spectra of Compound 3-10

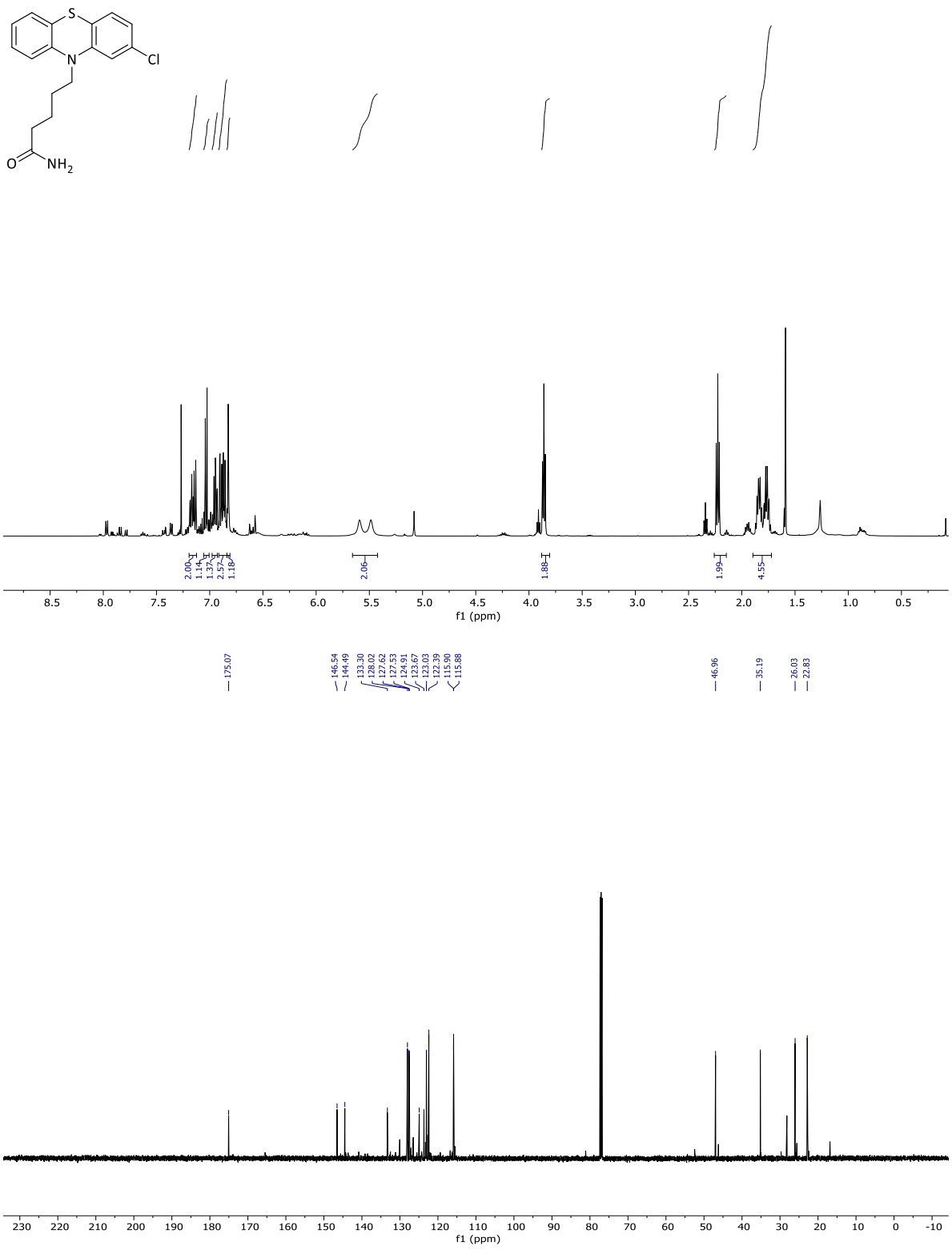


Figure A.11: ^{1}H and ^{13}C NMR Spectra of Compound 3-11

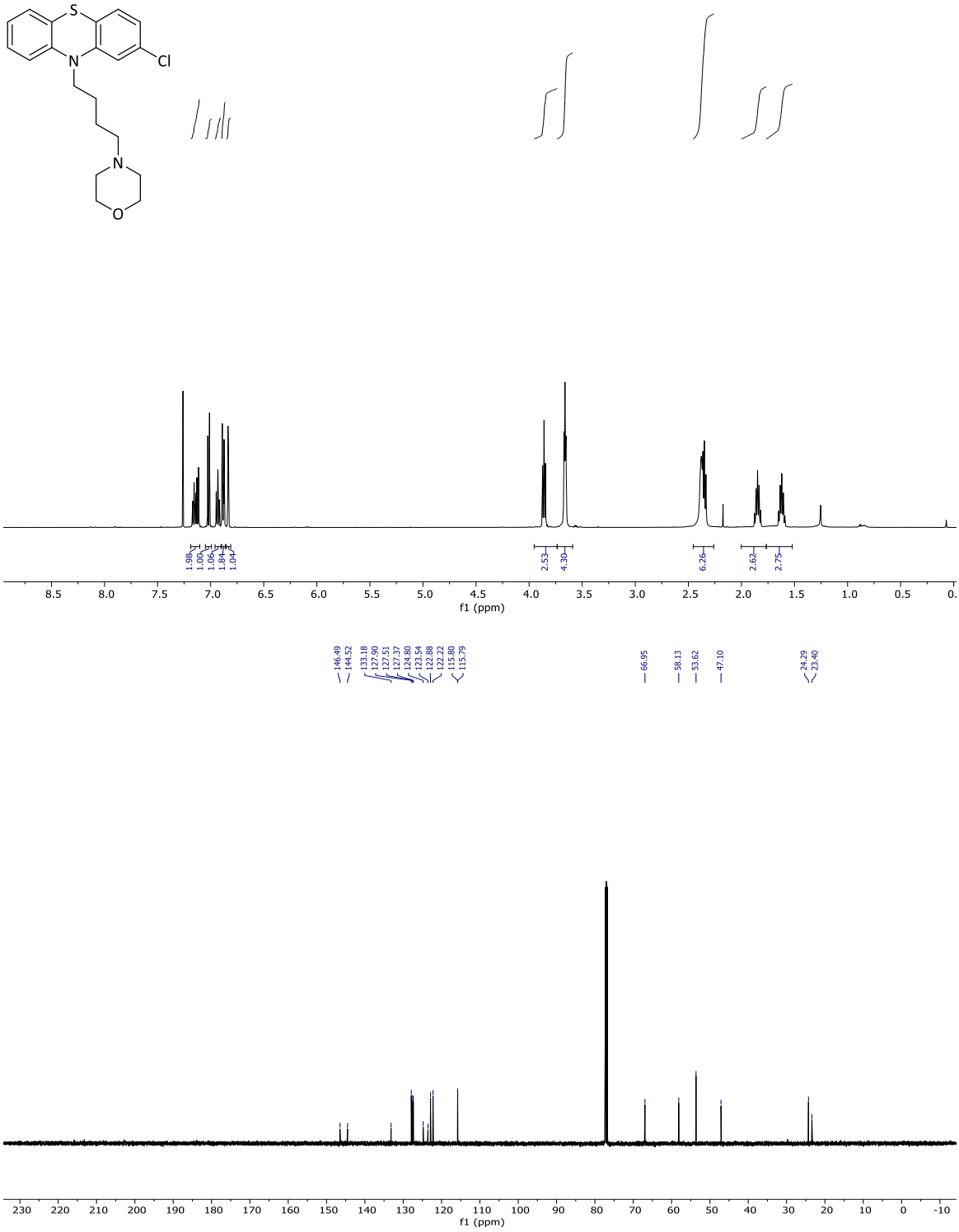


Figure A.12: ^{1}H and ^{13}C NMR Spectra of Compound 3-12

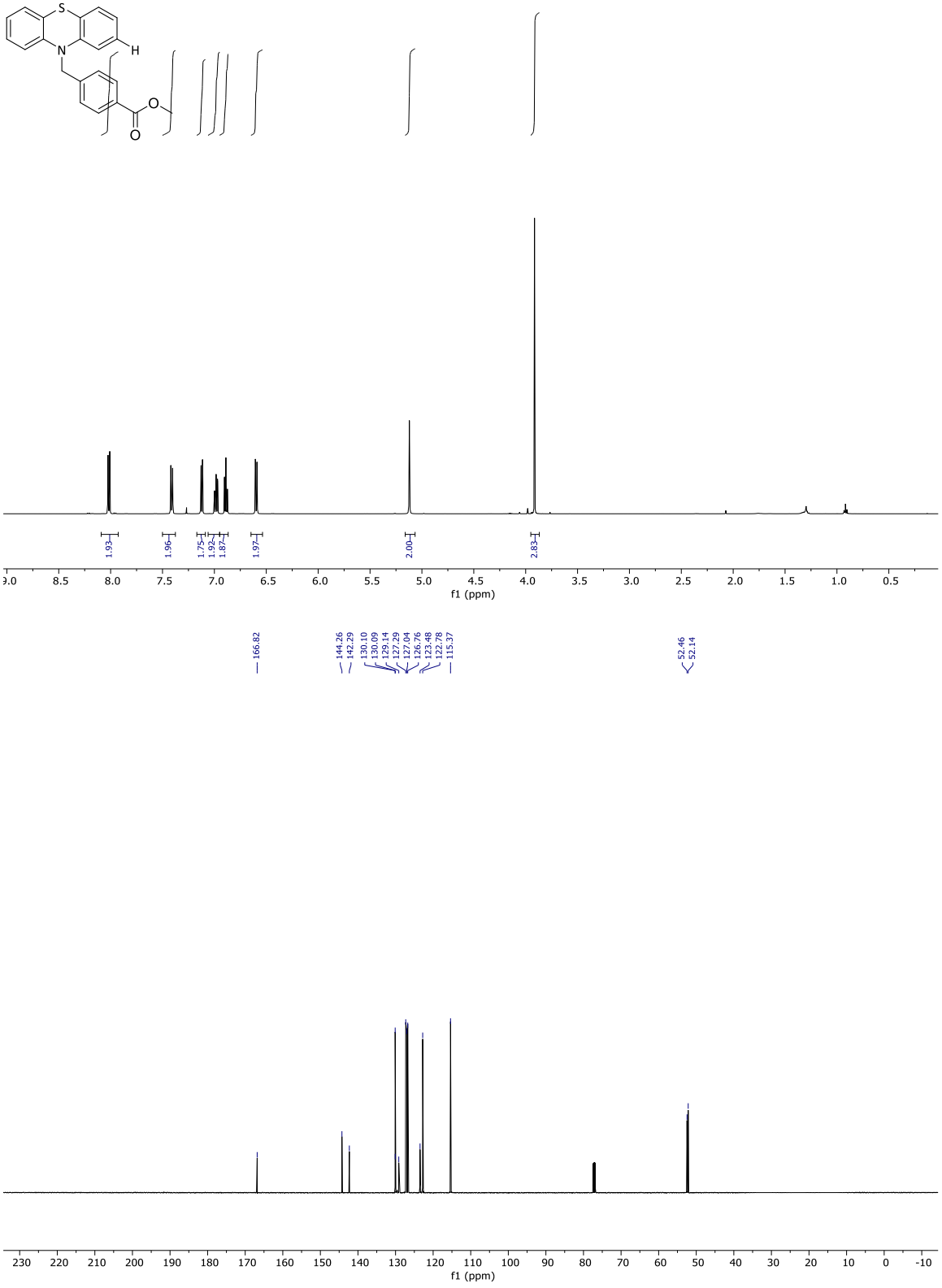


Figure A.13: ^{1}H and ^{13}C NMR Spectra of Compound 3-13

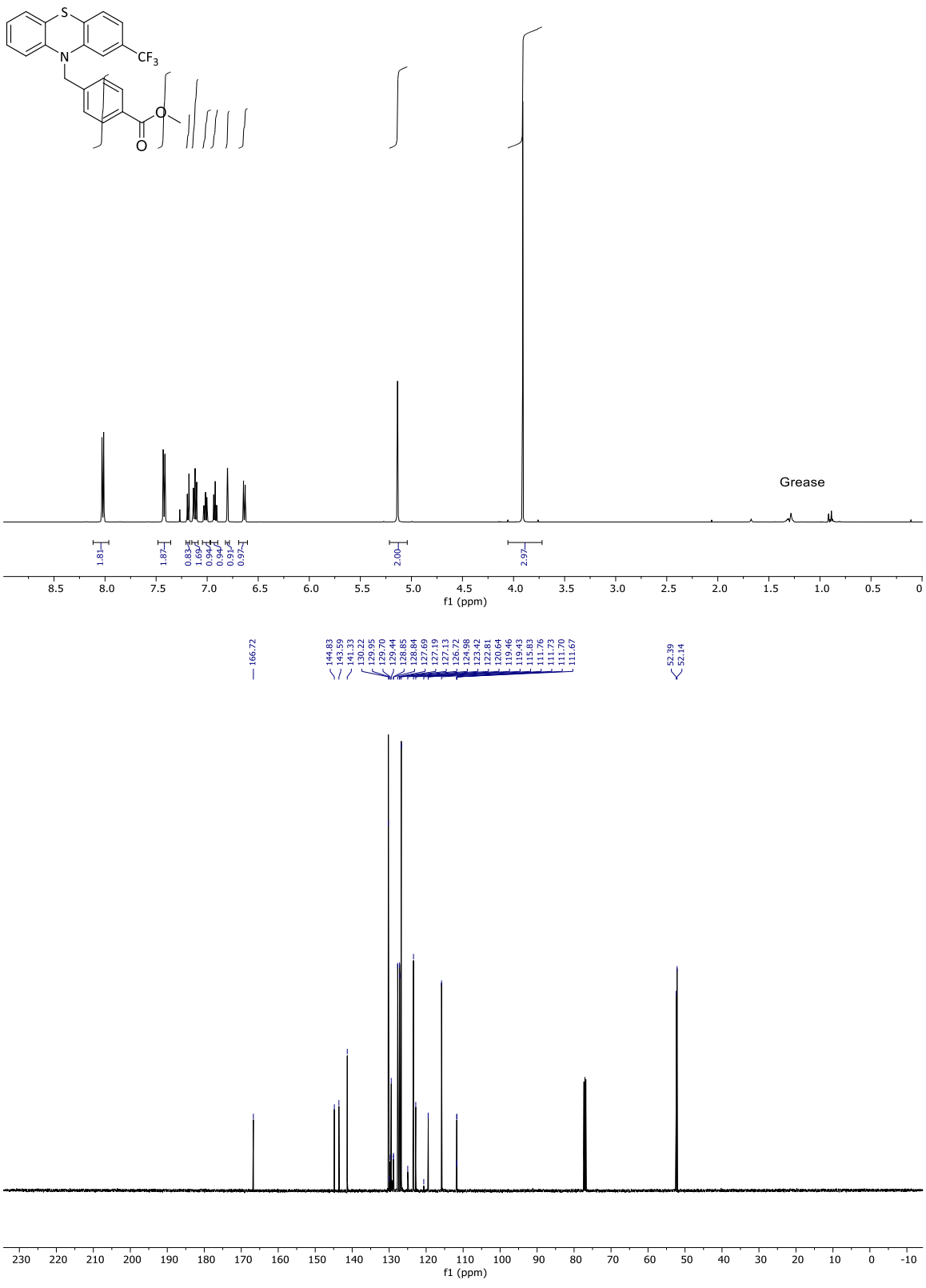


Figure A.14: ^{1}H and ^{13}C NMR Spectra of Compound 3-14

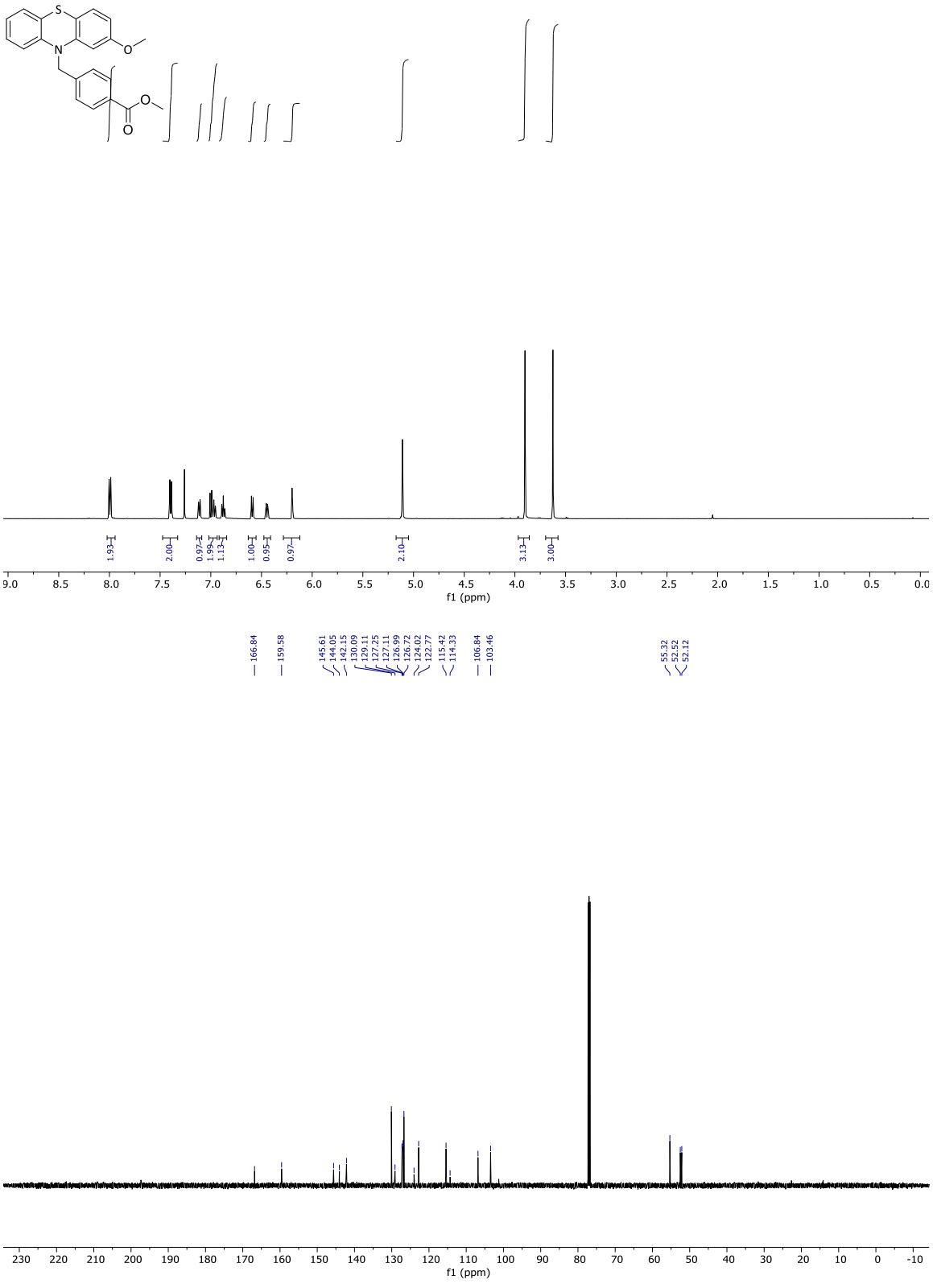


Figure A.15: ^{1}H and ^{13}C NMR Spectra of Compound 3-15

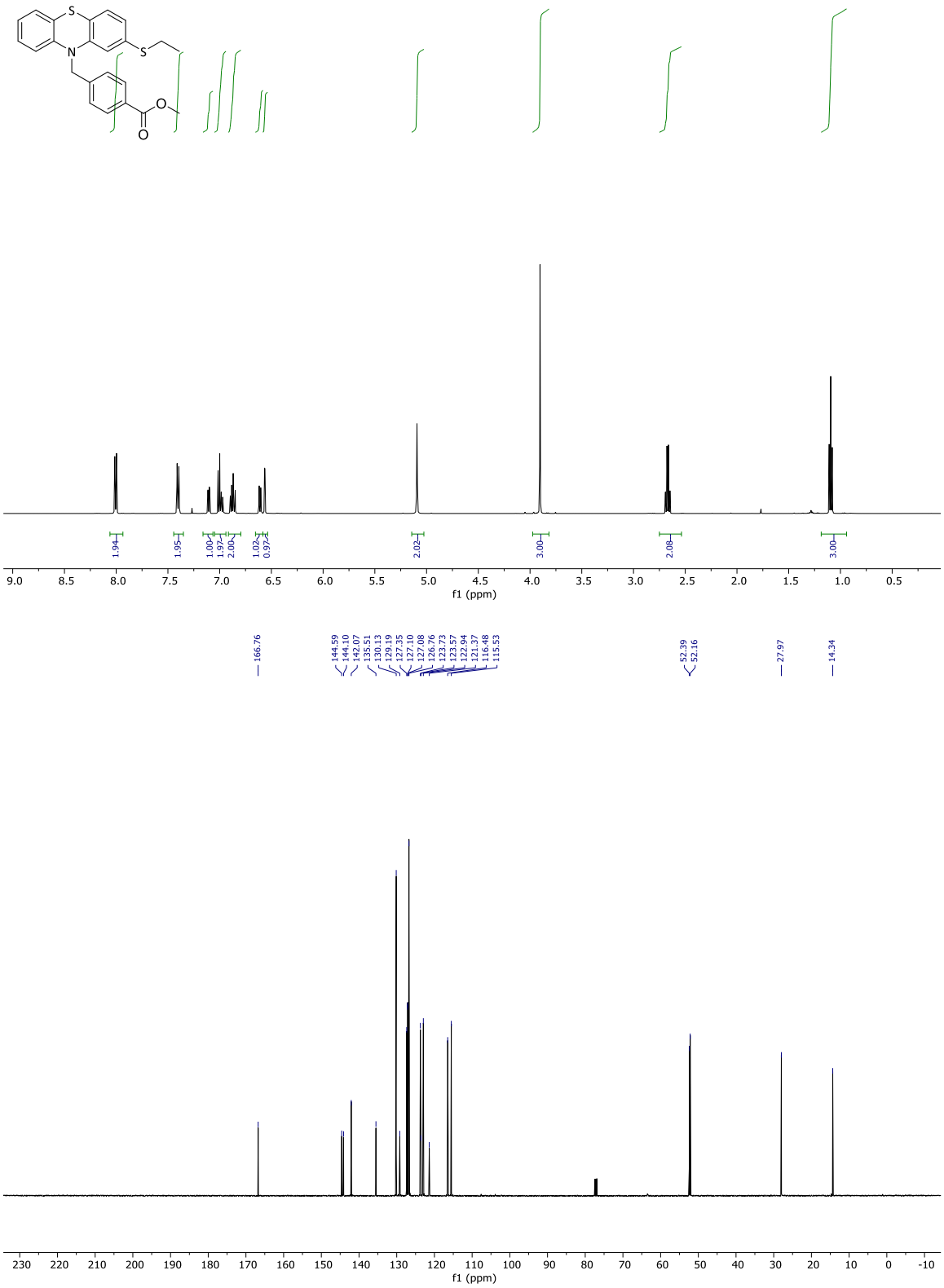


Figure A.16: ^{1}H and ^{13}C NMR Spectra of Compound 3-16

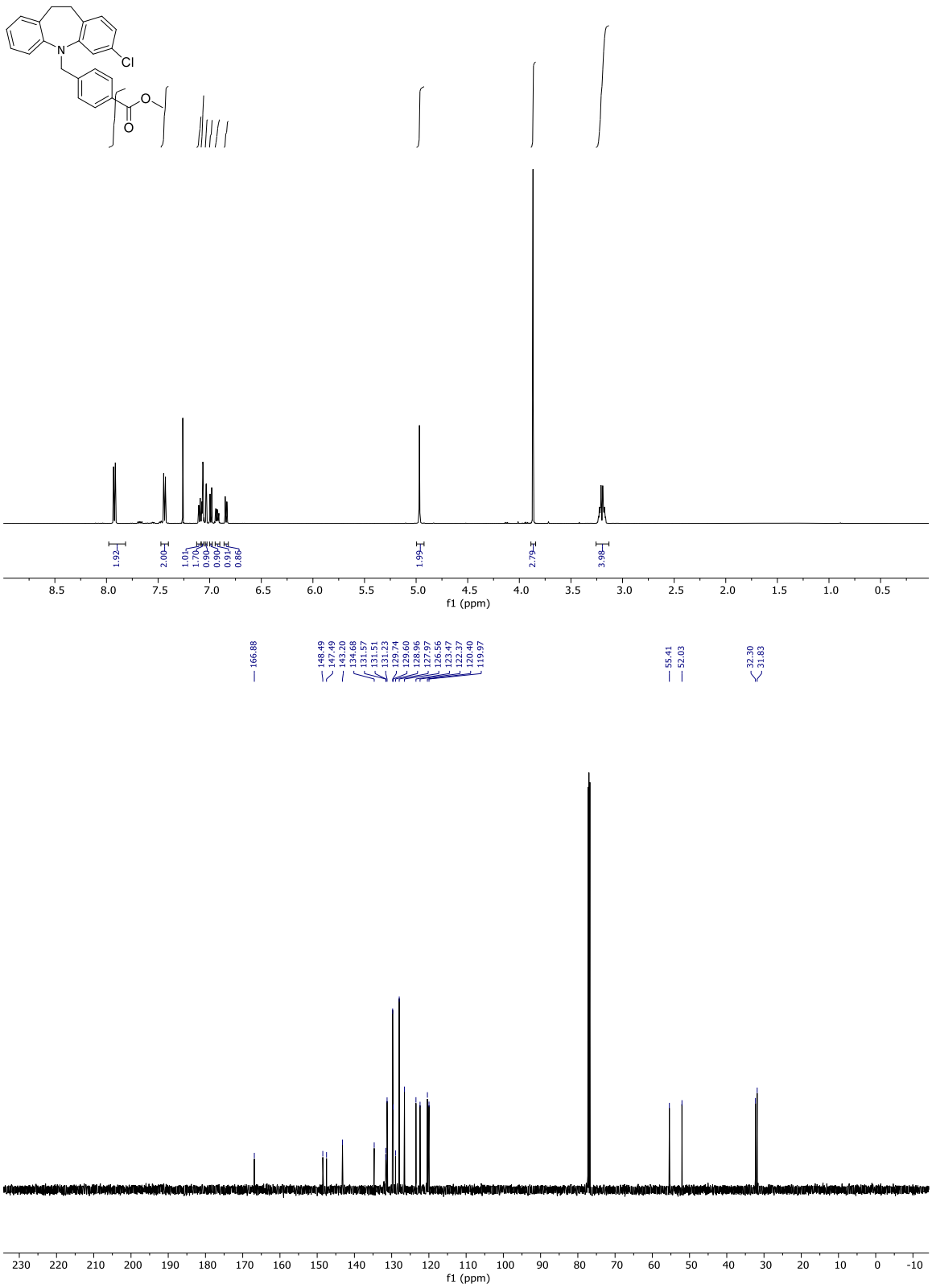


Figure A.17: ^{1}H and ^{13}C NMR Spectra of Compound 3-17

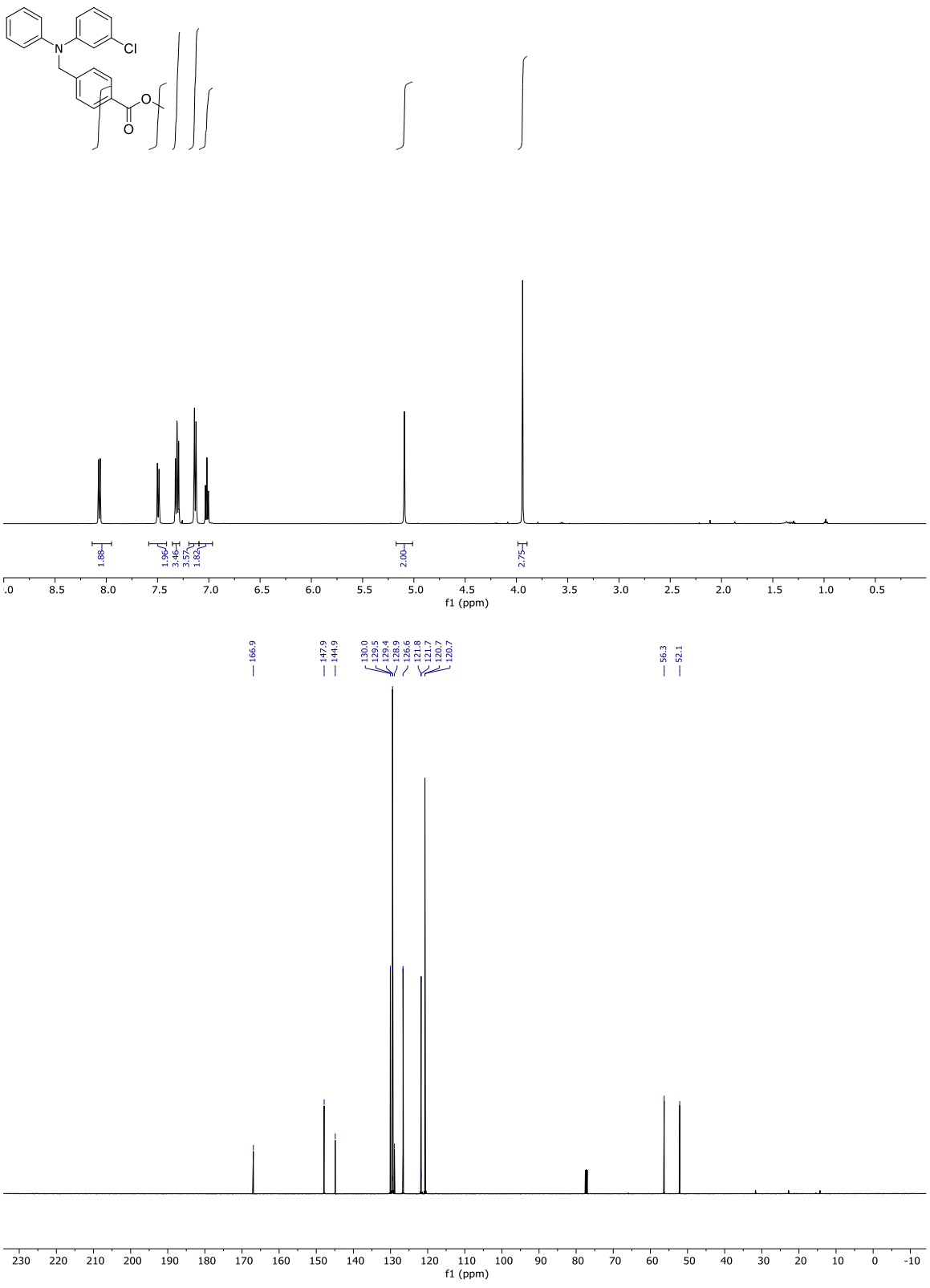


Figure A.18: ^{1}H and ^{13}C NMR Spectra of Compound 3-18

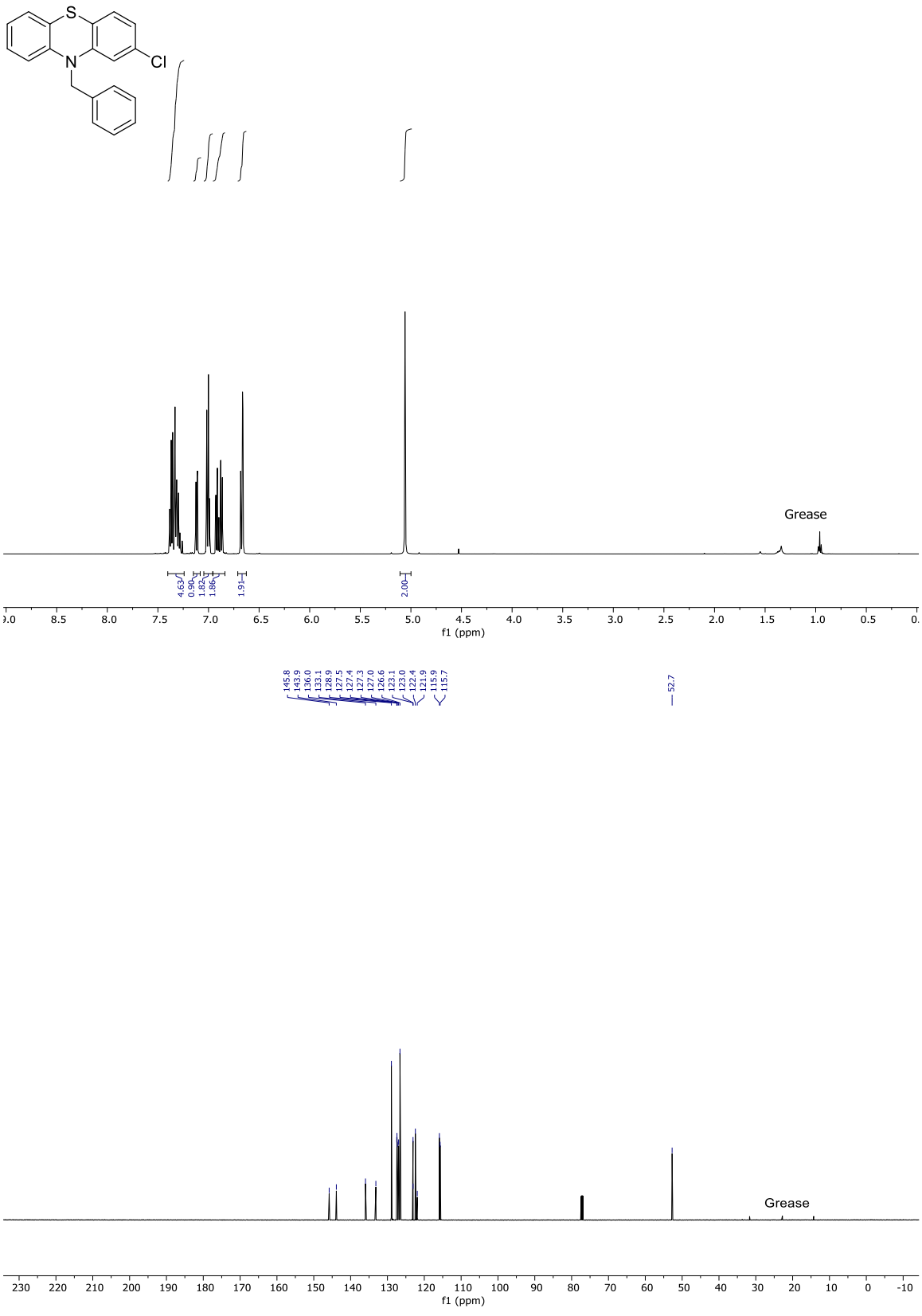


Figure A.19: ^{1}H and ^{13}C NMR Spectra of Compound 3-19

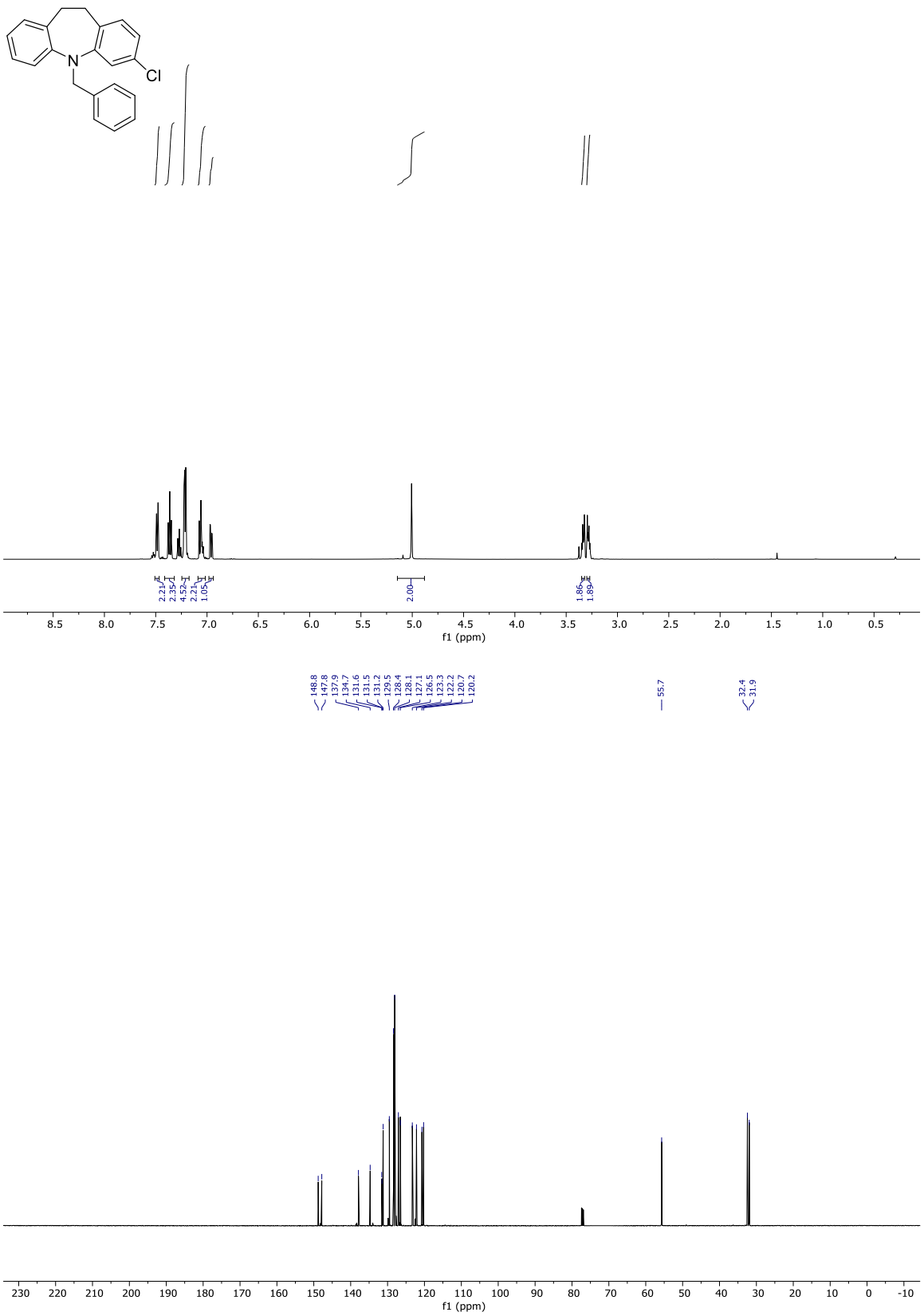


Figure A.20: ^{1}H and ^{13}C NMR Spectra of Compound 3-20

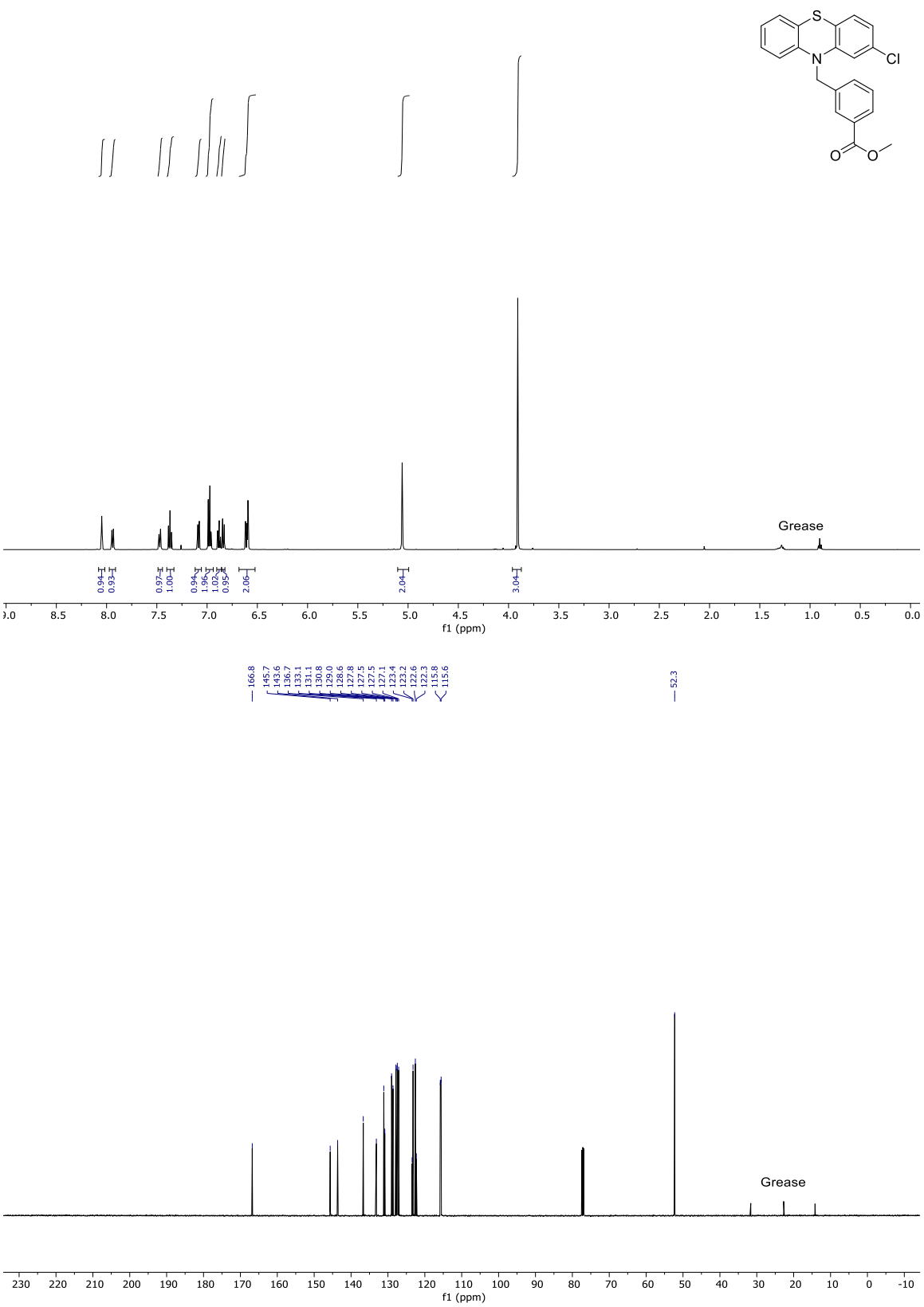


Figure A.21: ^{1}H and ^{13}C NMR Spectra of Compound 3-21

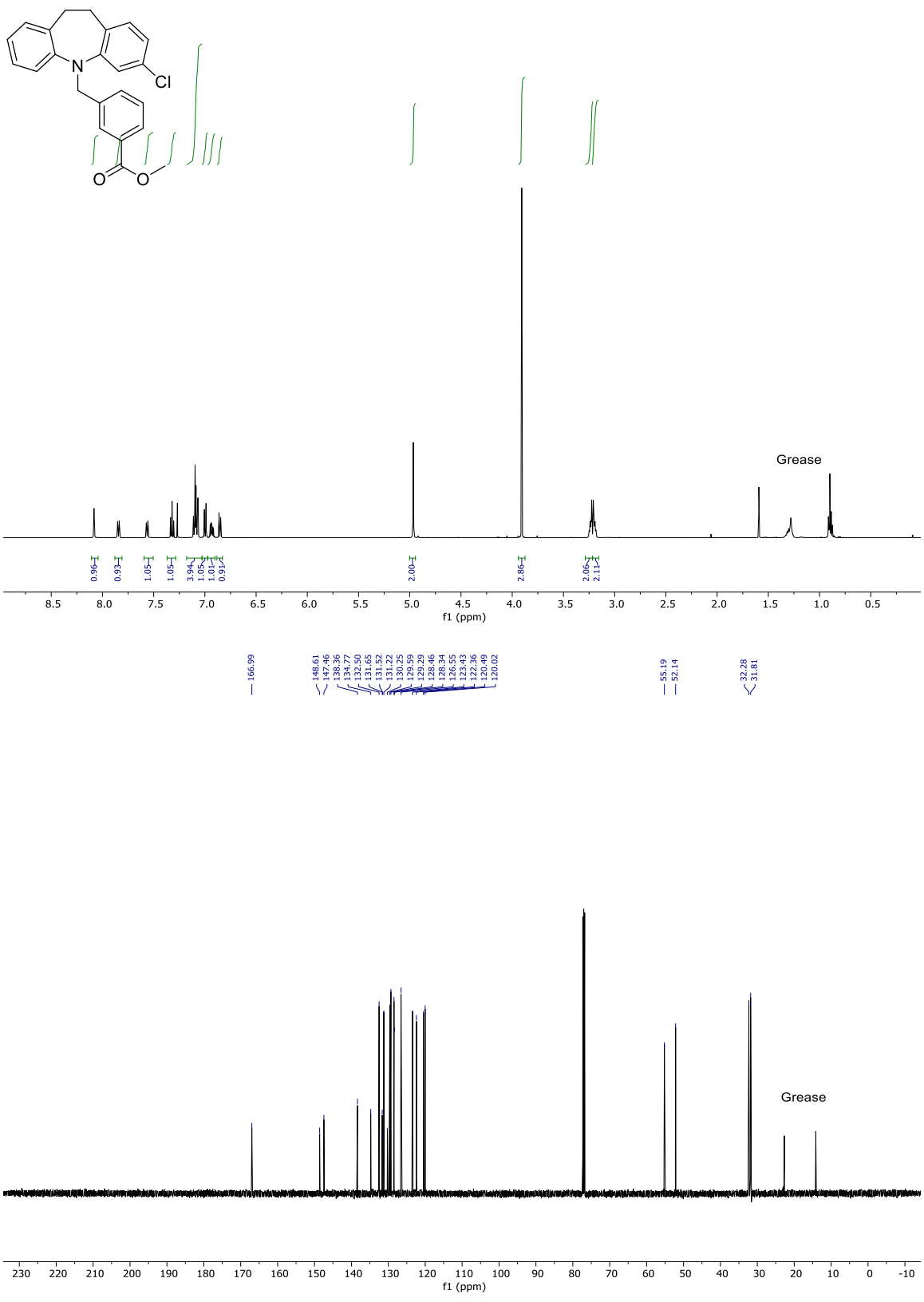


Figure A.22: ^{1}H and ^{13}C NMR Spectra of Compound 3-22

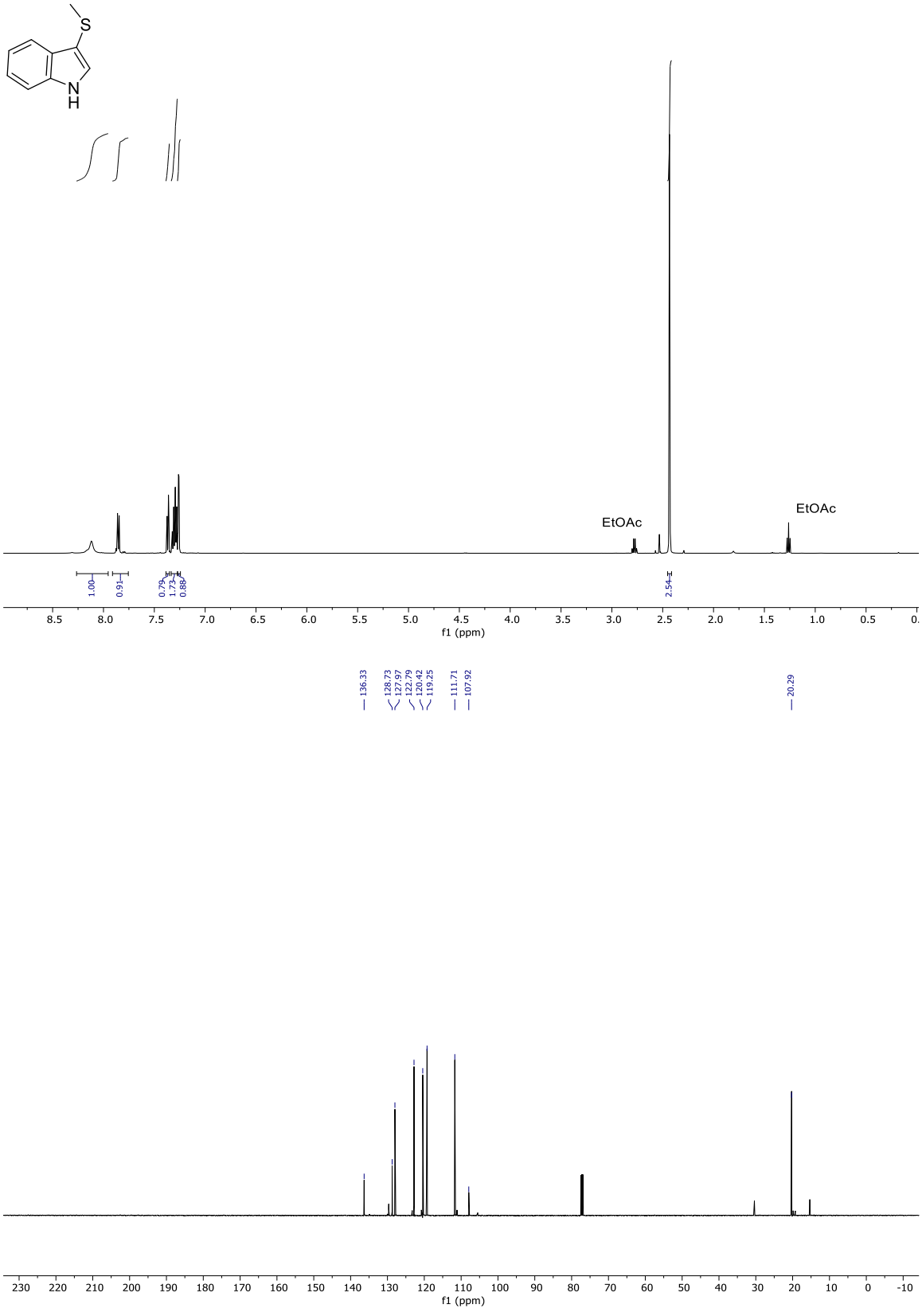


Figure A.23: ^{1}H and ^{13}C NMR Spectra of Compound 3-23

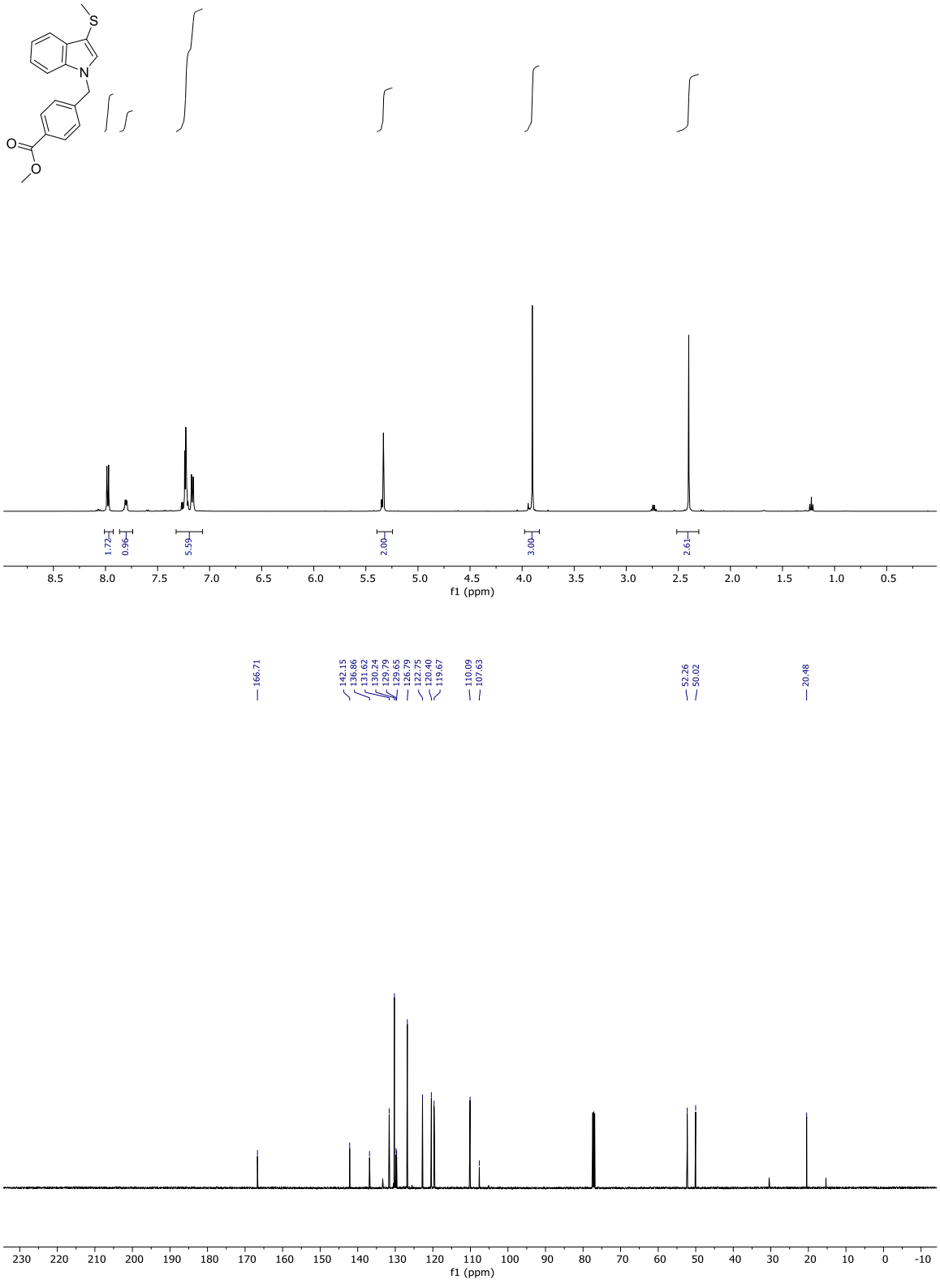


Figure A.24: ^{1}H and ^{13}C NMR Spectra of Compound 3-24

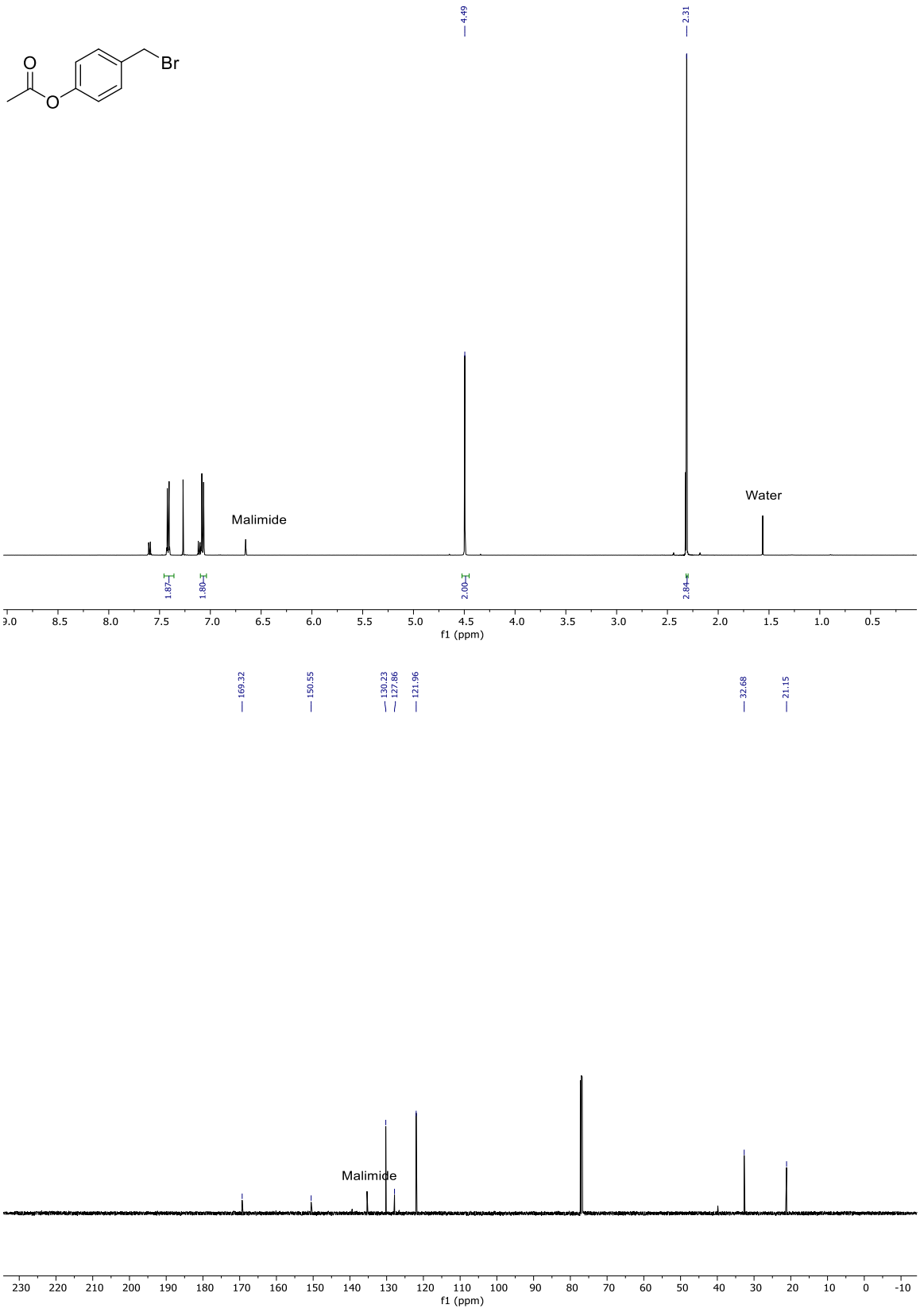


Figure A.25: ^{1}H and ^{13}C NMR Spectra of Compound 3-25

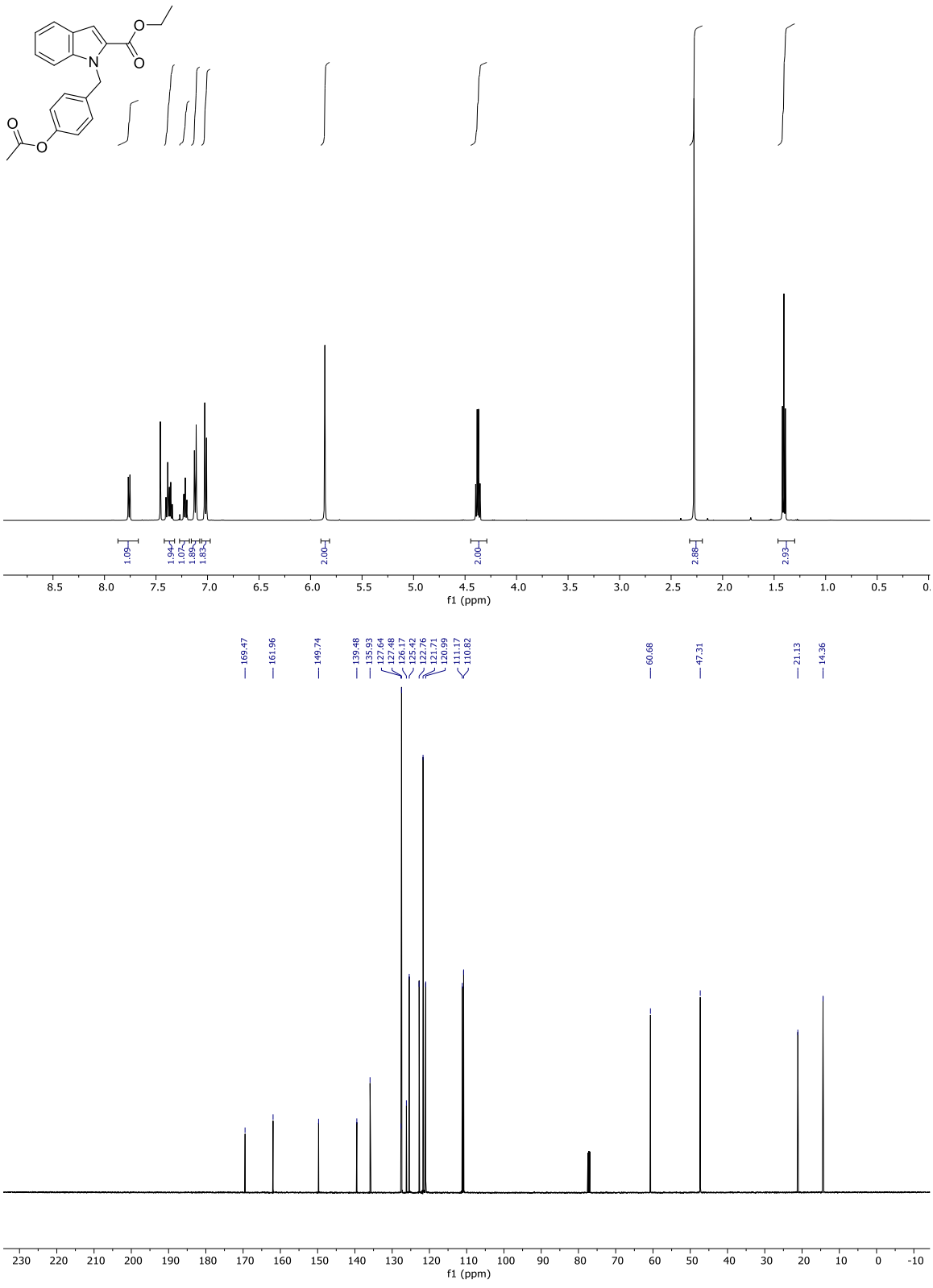


Figure A.26: ^{1}H and ^{13}C NMR Spectra of Compound 3-26

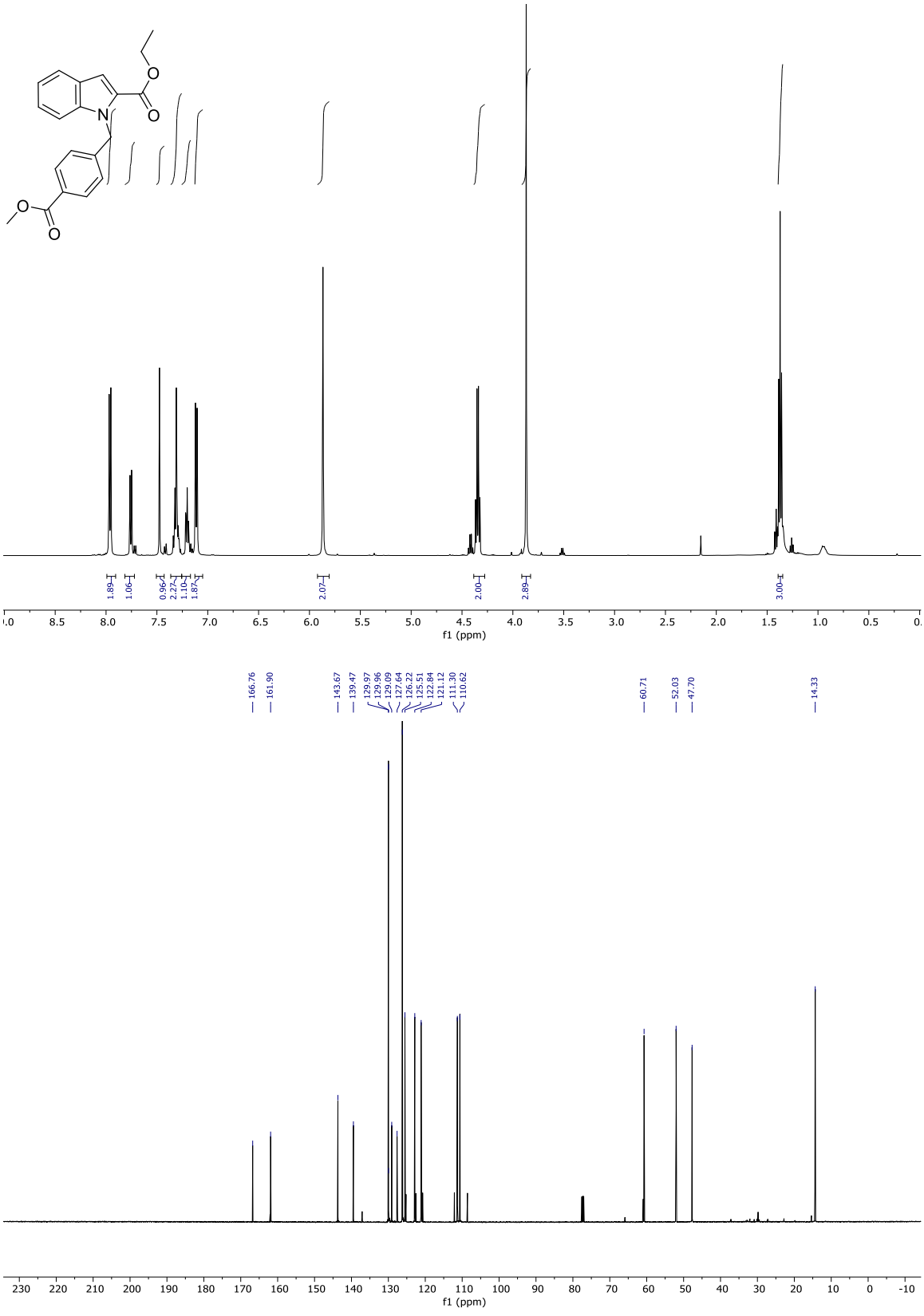


Figure A.27: ^{1}H and ^{13}C NMR Spectra of Compound 3-27

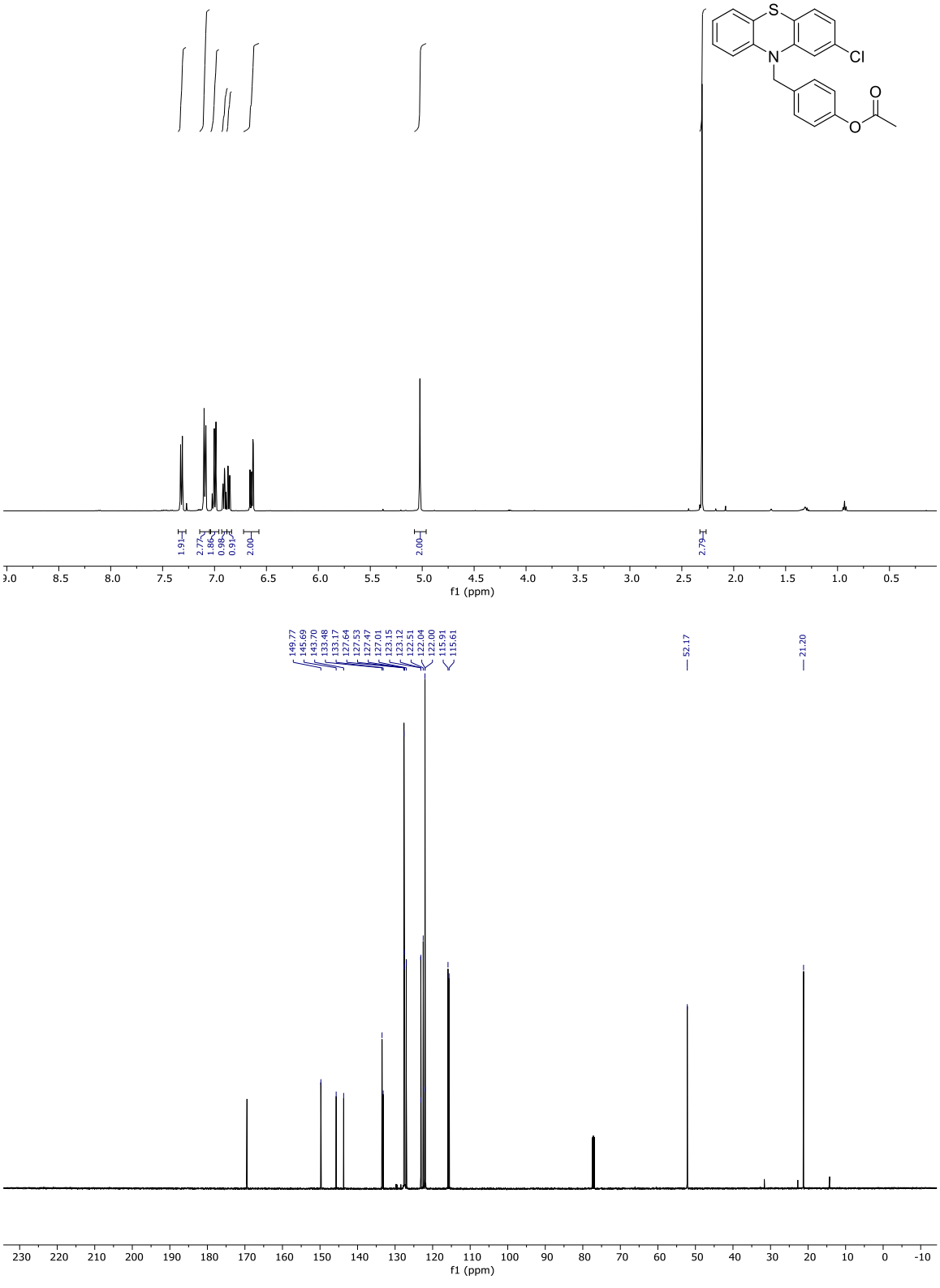


Figure A.28: ^{1}H and ^{13}C NMR Spectra of Compound 3-28

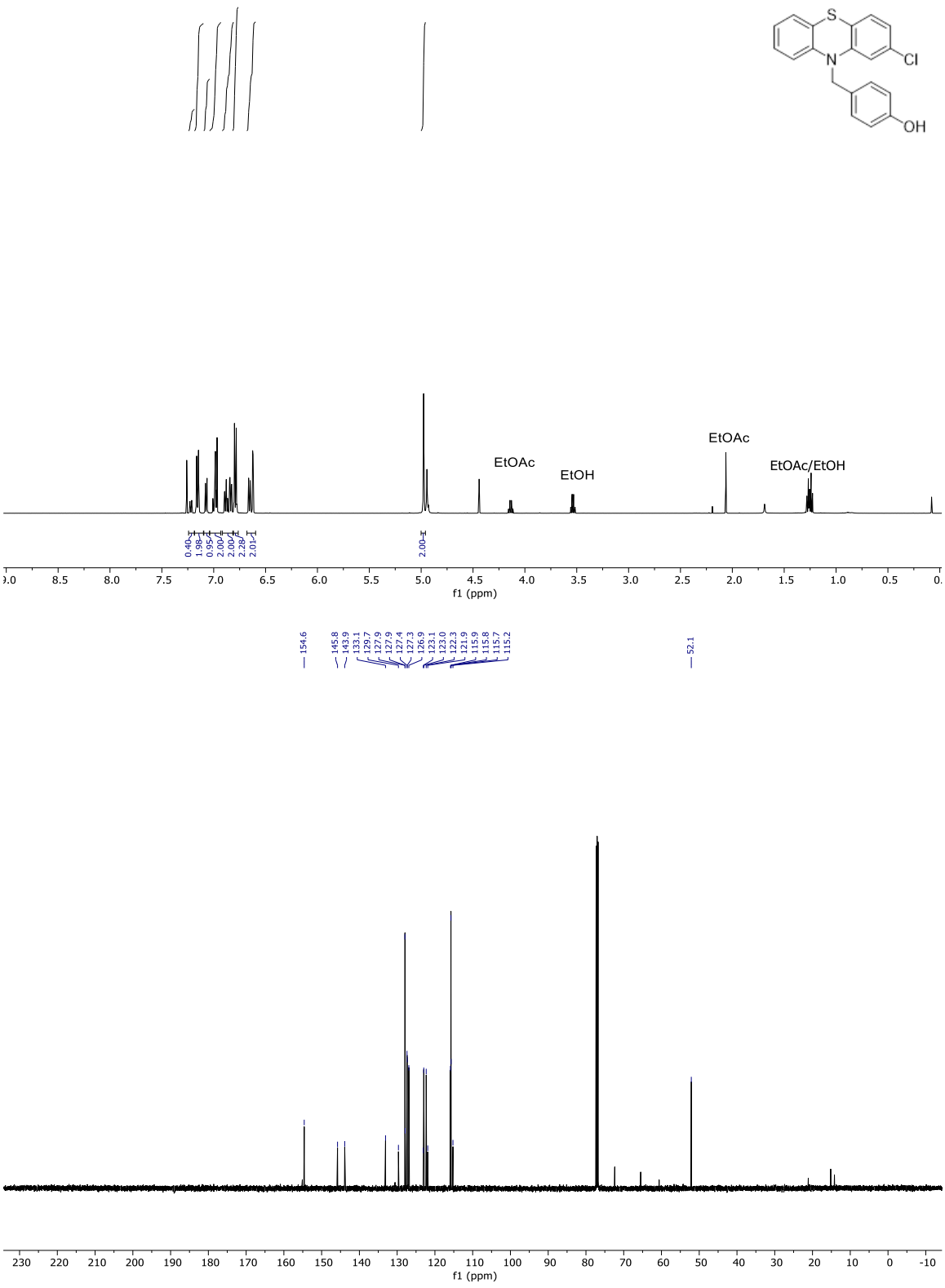


Figure A.29: ^{1}H and ^{13}C NMR Spectra of Compound 3-29

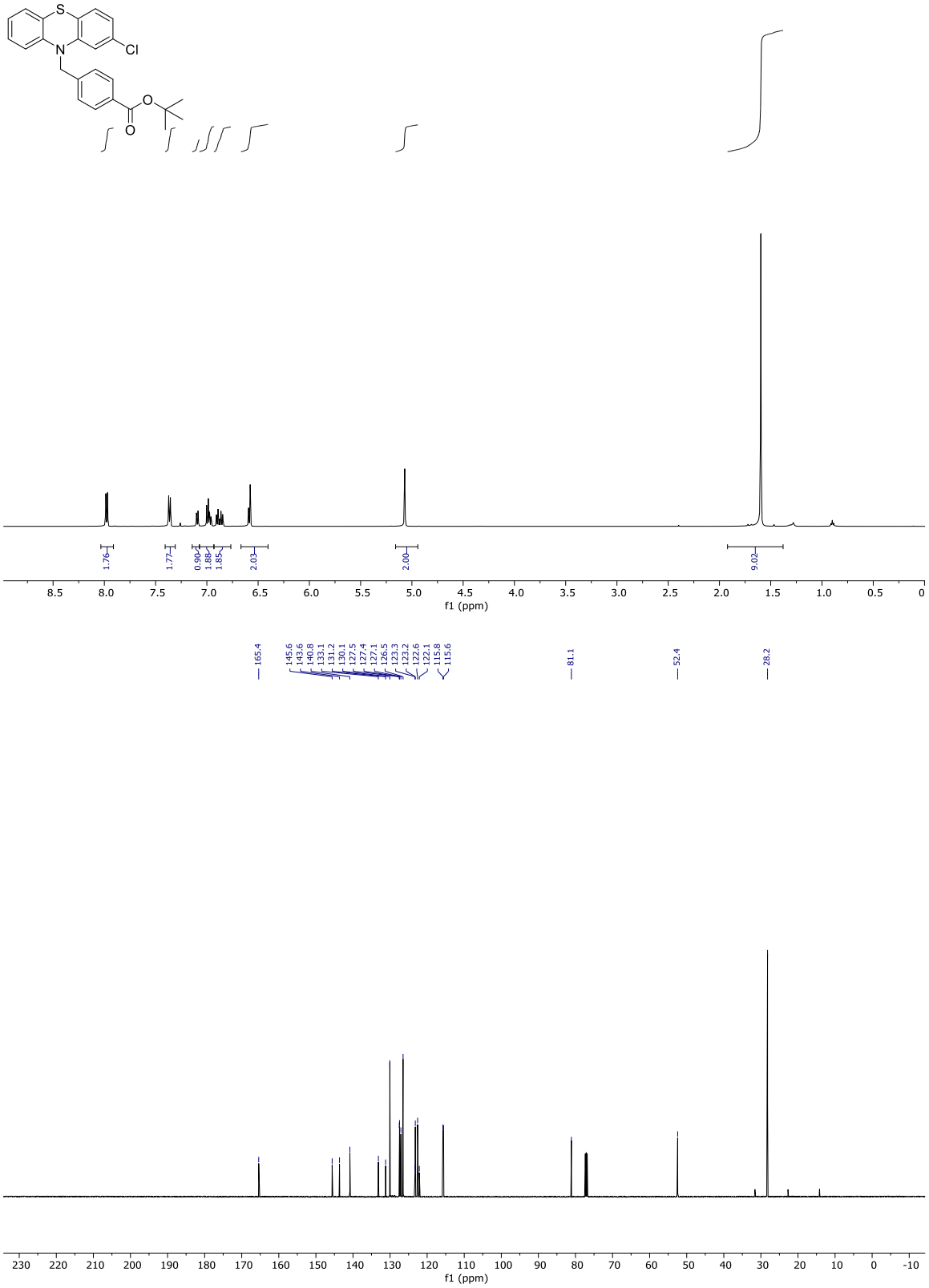


Figure A.30: ^{1}H and ^{13}C NMR Spectra of Compound 3-30

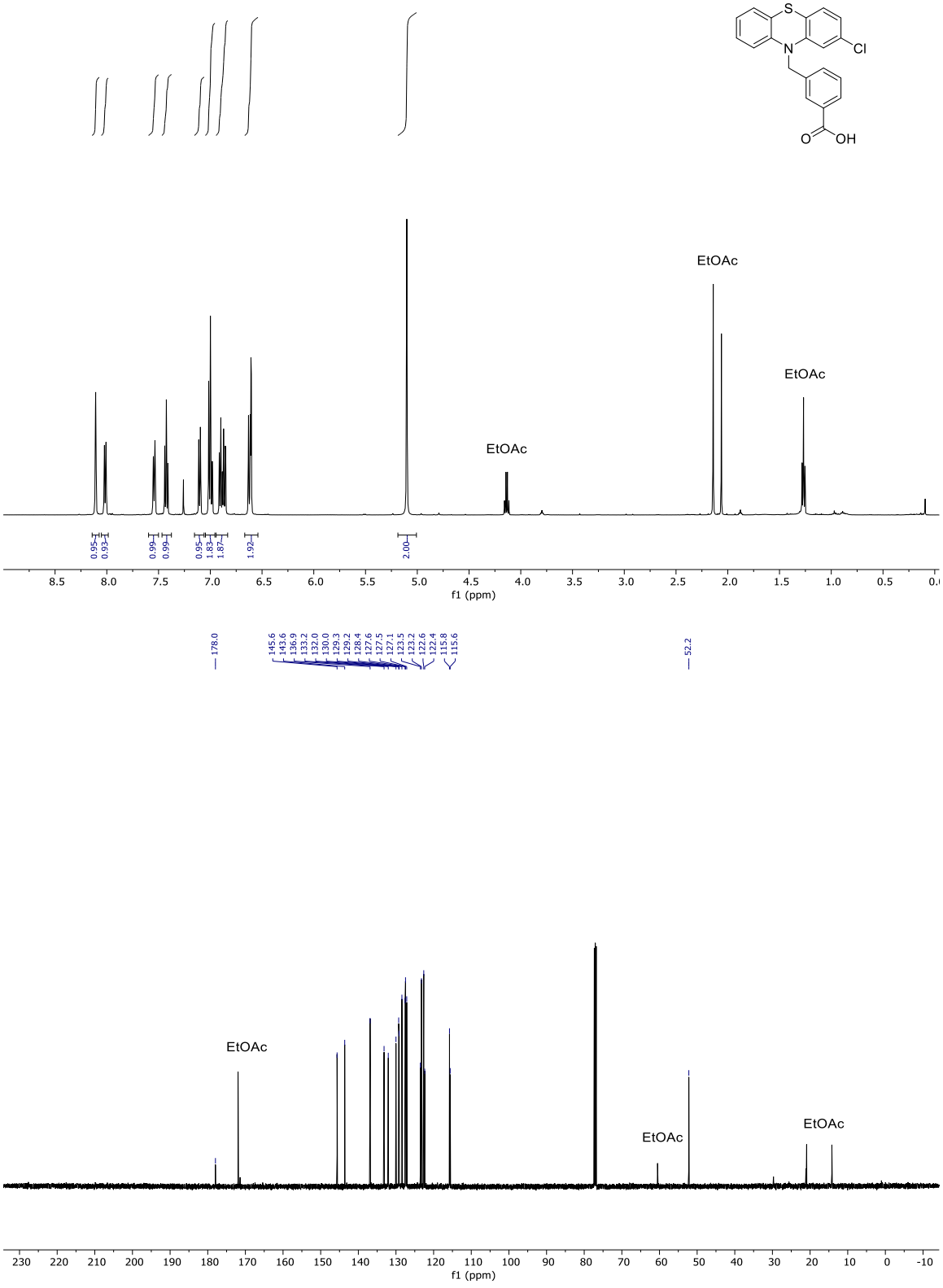


Figure A.31: ^{1}H and ^{13}C NMR Spectra of Compound 3-31

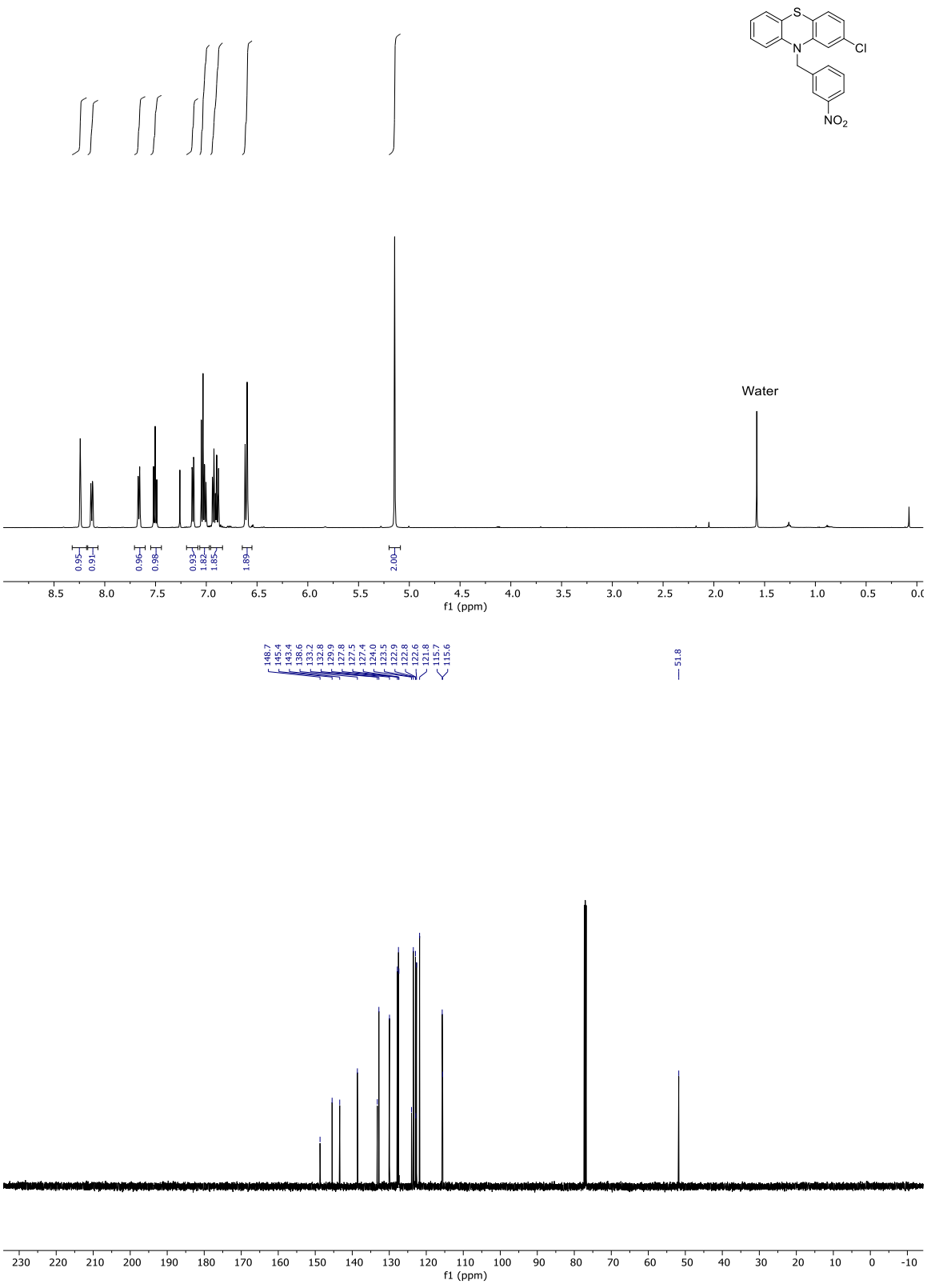


Figure A.32: ^{1}H and ^{13}C NMR Spectra of Compound 3-32

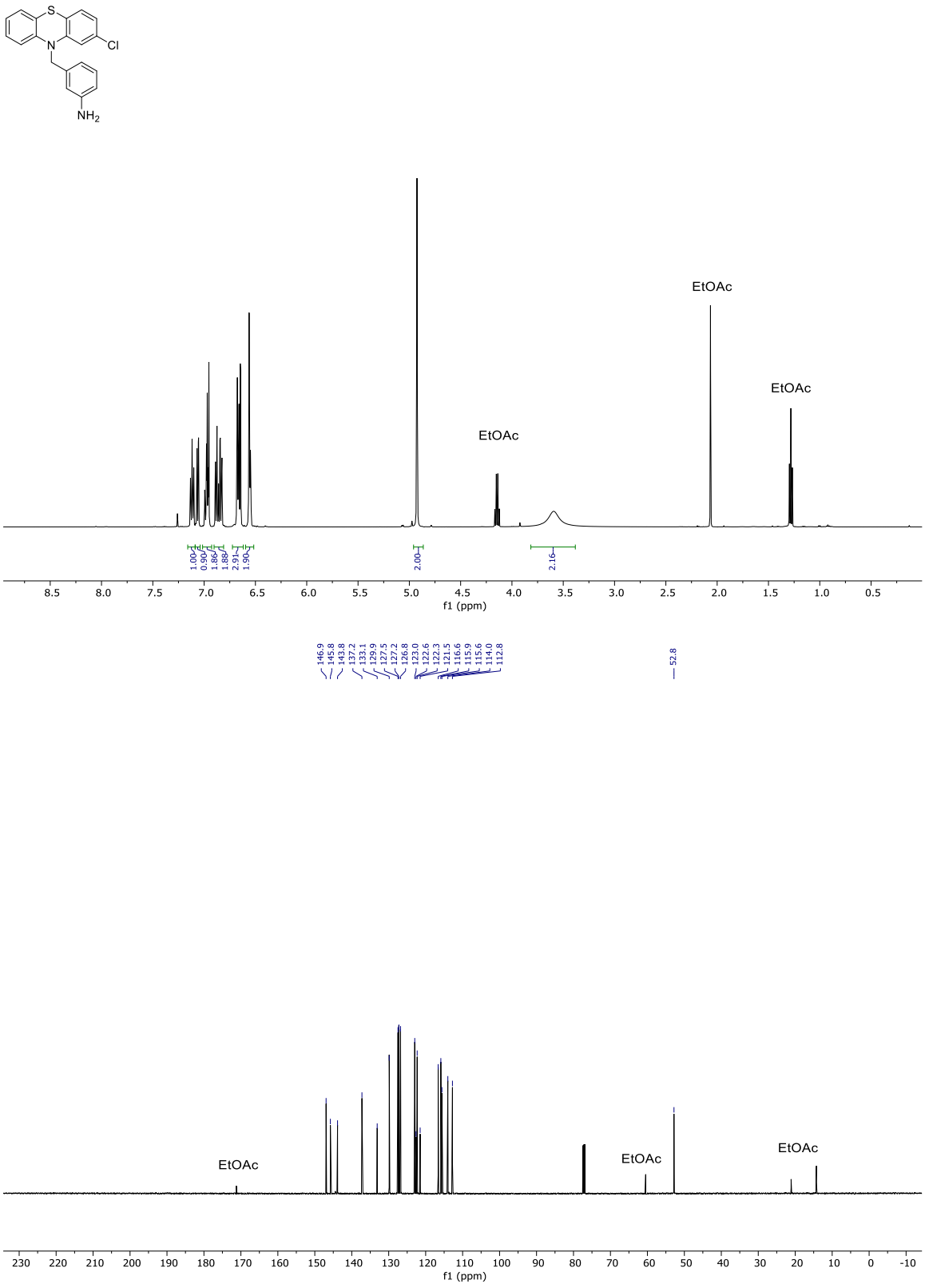


Figure A.33: ^{1}H and ^{13}C NMR Spectra of Compound 3-33

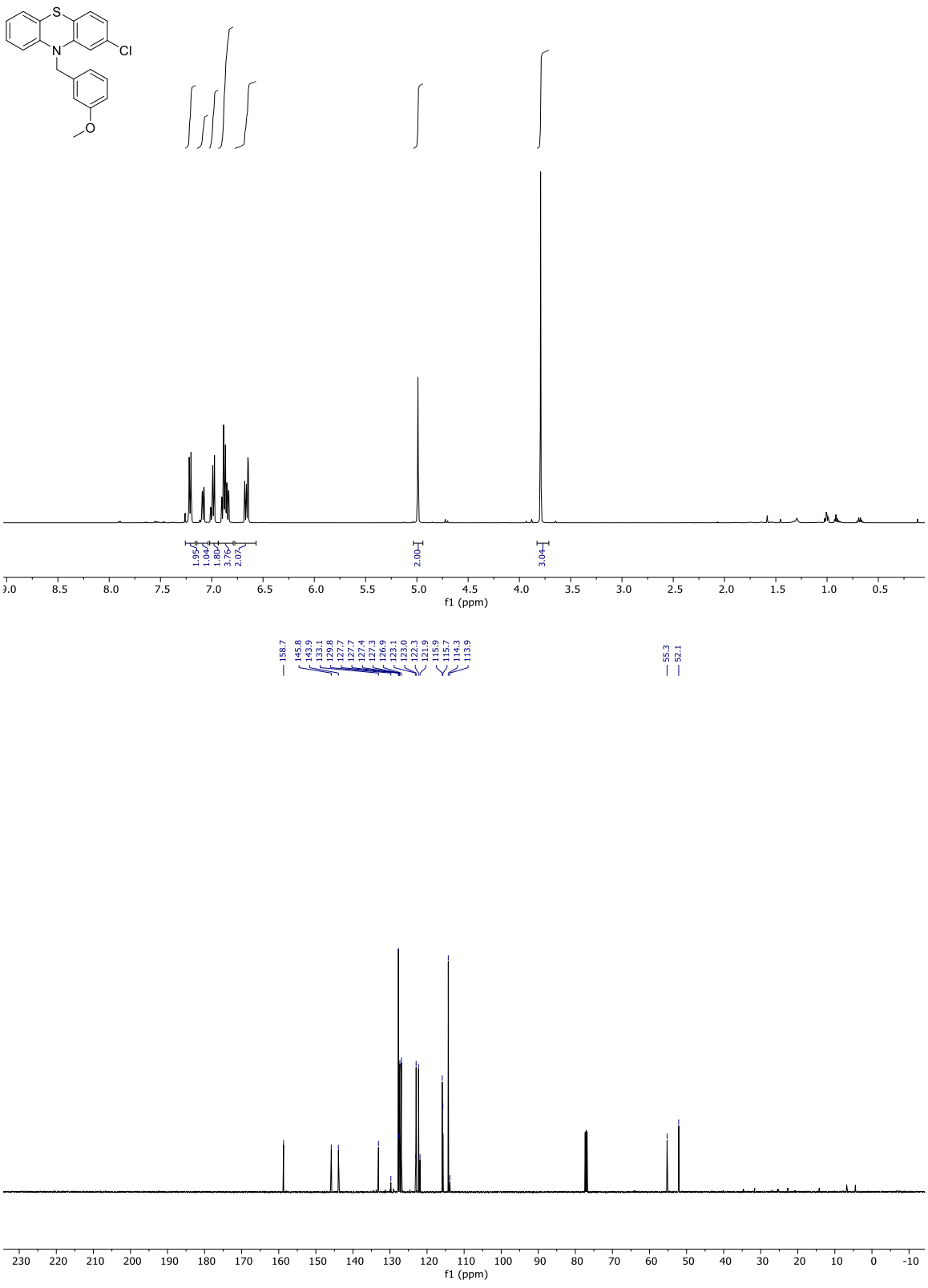


Figure A.34: ^{1}H and ^{13}C NMR Spectra of Compound 3-34

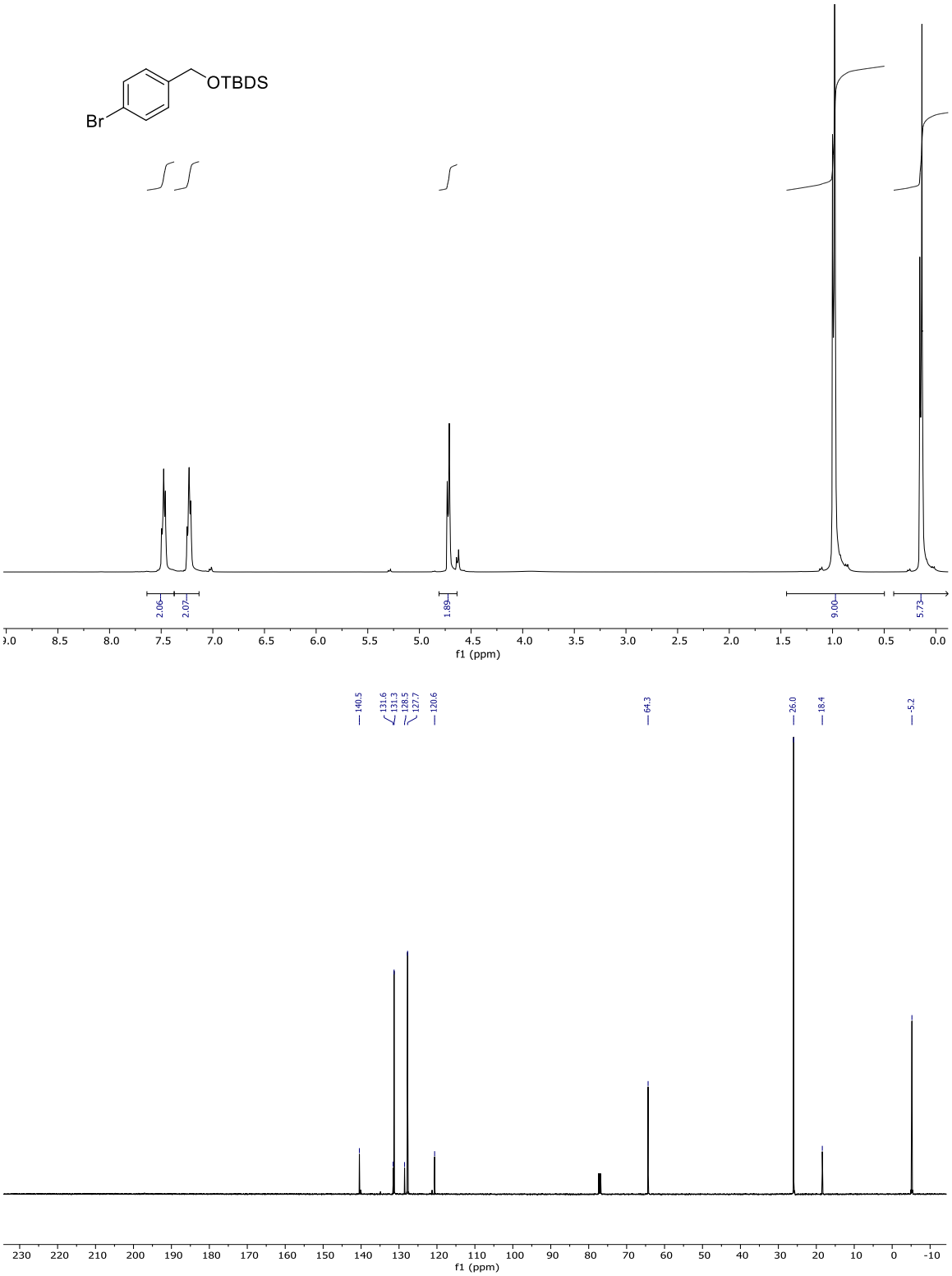


Figure A.35: ^{1}H and ^{13}C NMR Spectra of Compound 4-1

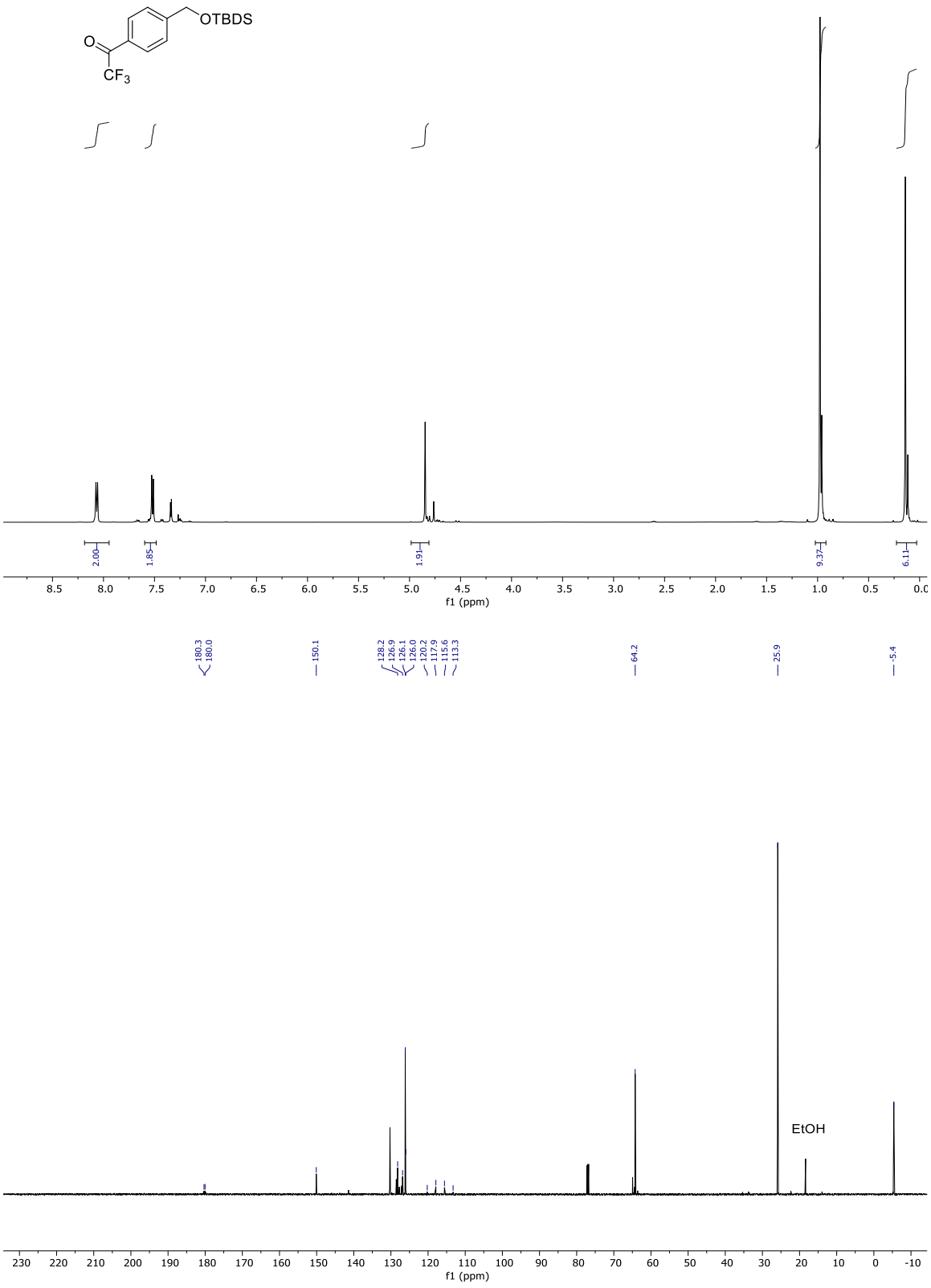


Figure A.36: ^{1}H and ^{13}C NMR Spectra of Compound 4-2

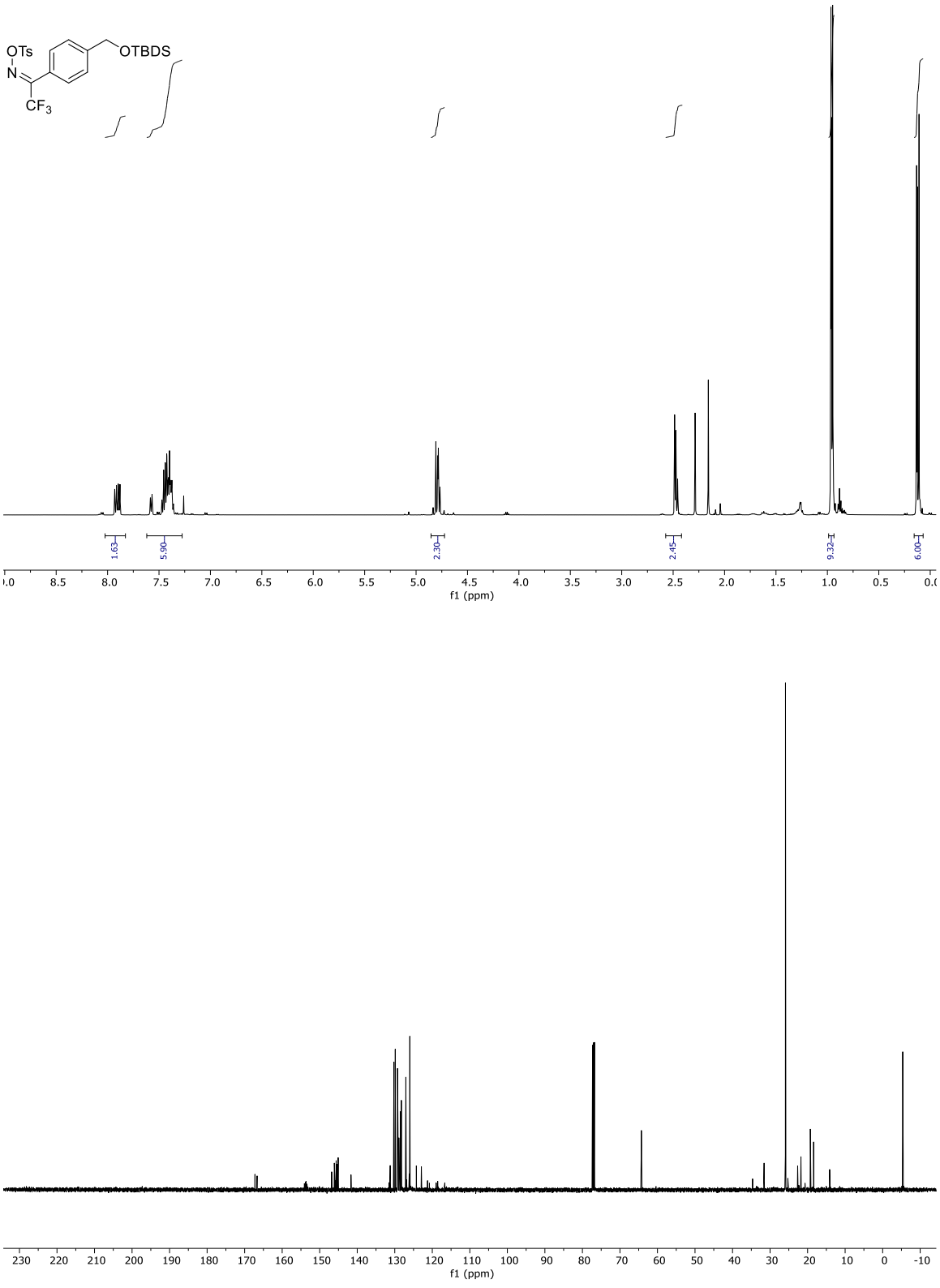


Figure A.37: ^{1}H and ^{13}C NMR Spectra of Compound 4-3

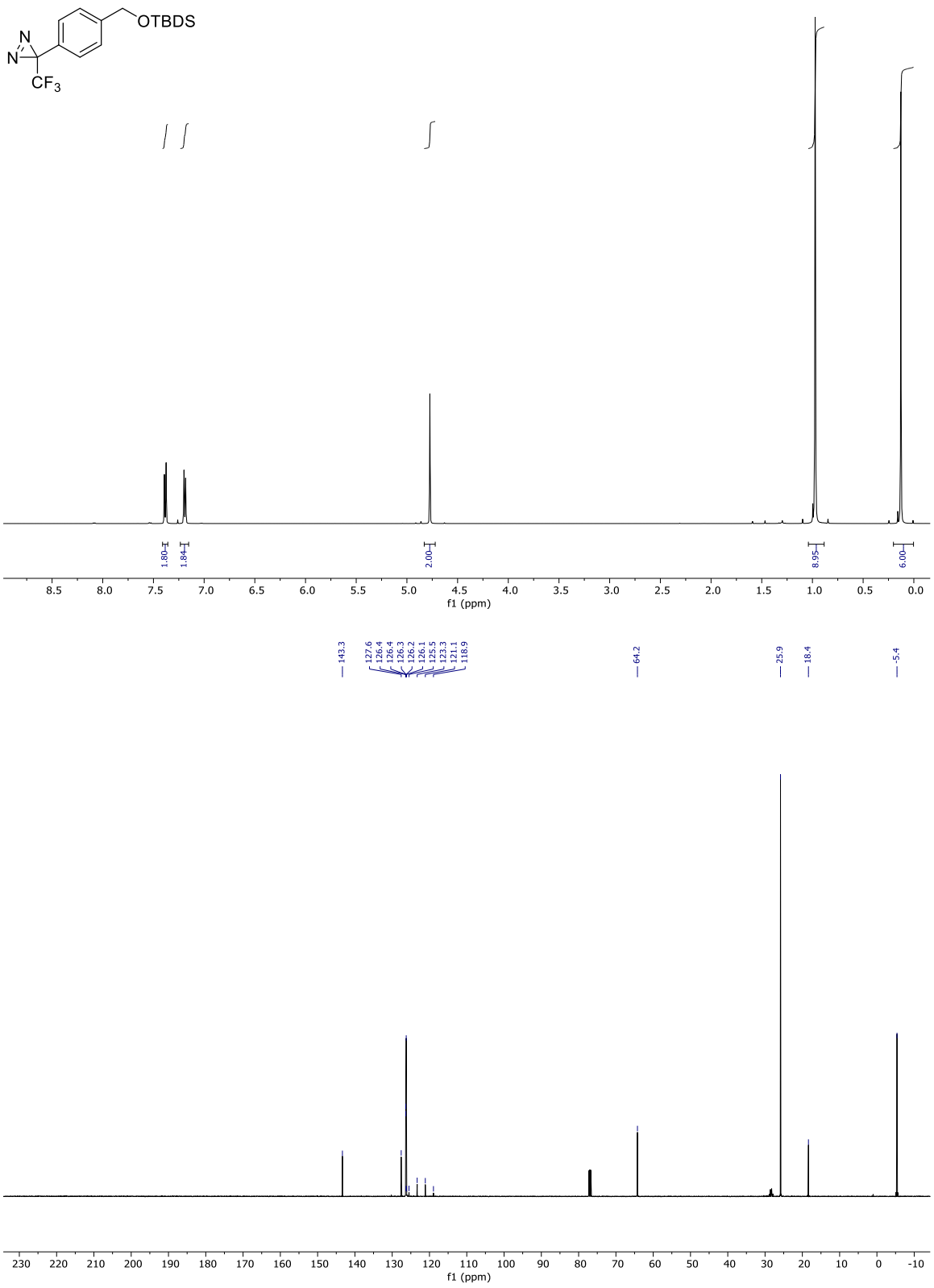


Figure A.38: ^{1}H and ^{13}C NMR Spectra of Compound 4-4

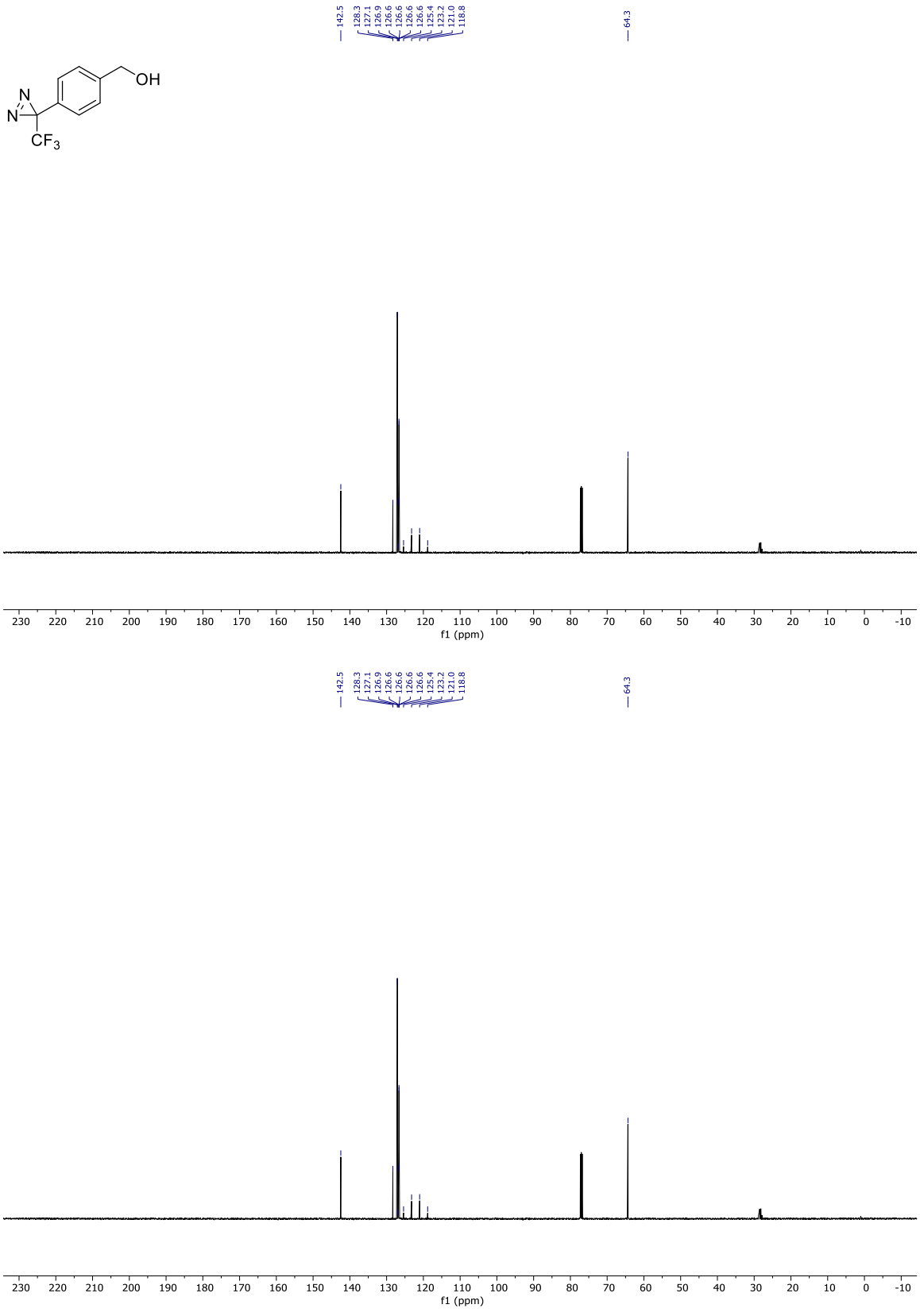


Figure A.39: ^{1}H and ^{13}C NMR Spectra of Compound 4-5

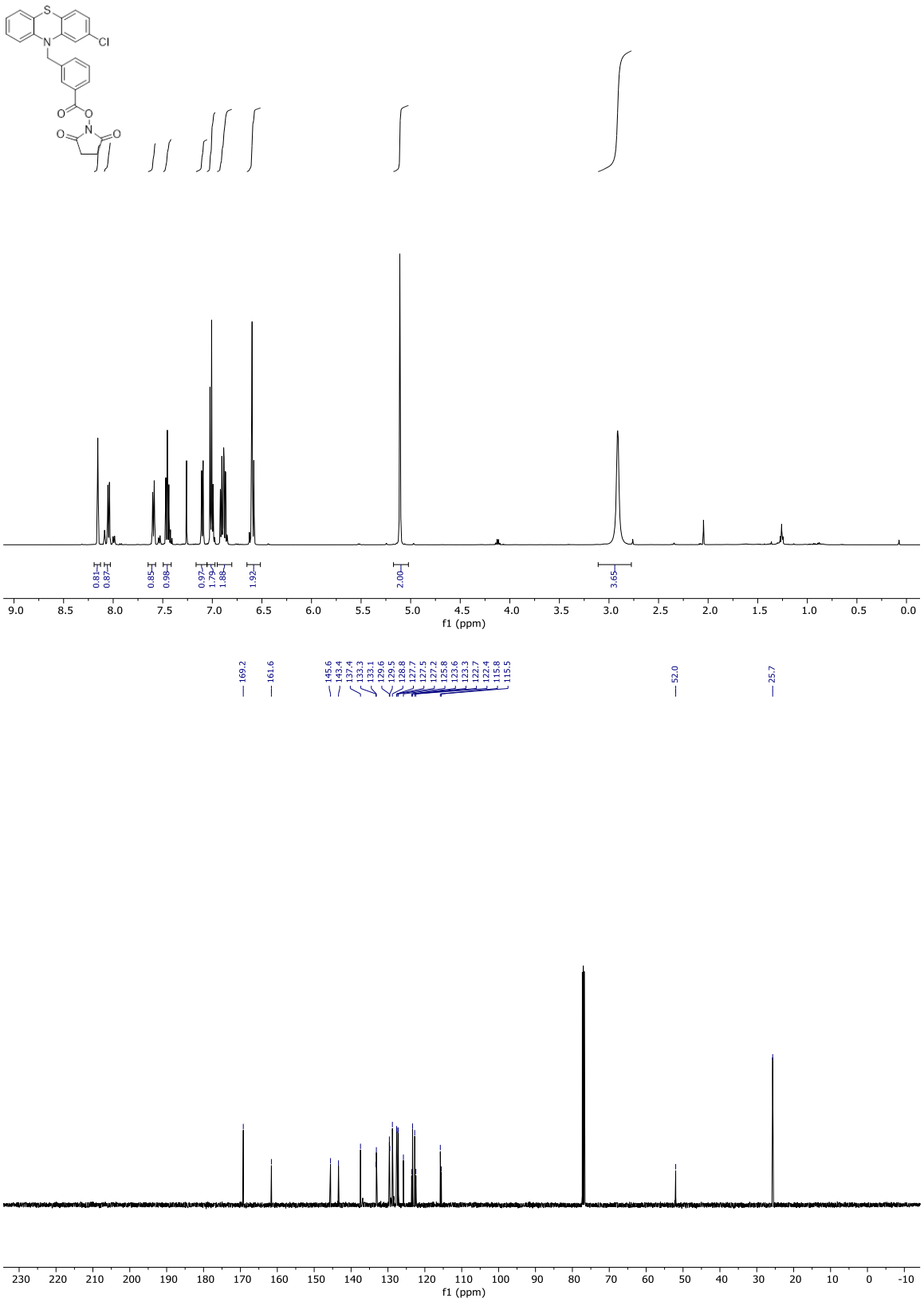


Figure A.40: ^{1}H and ^{13}C NMR Spectra of Compound 4-7

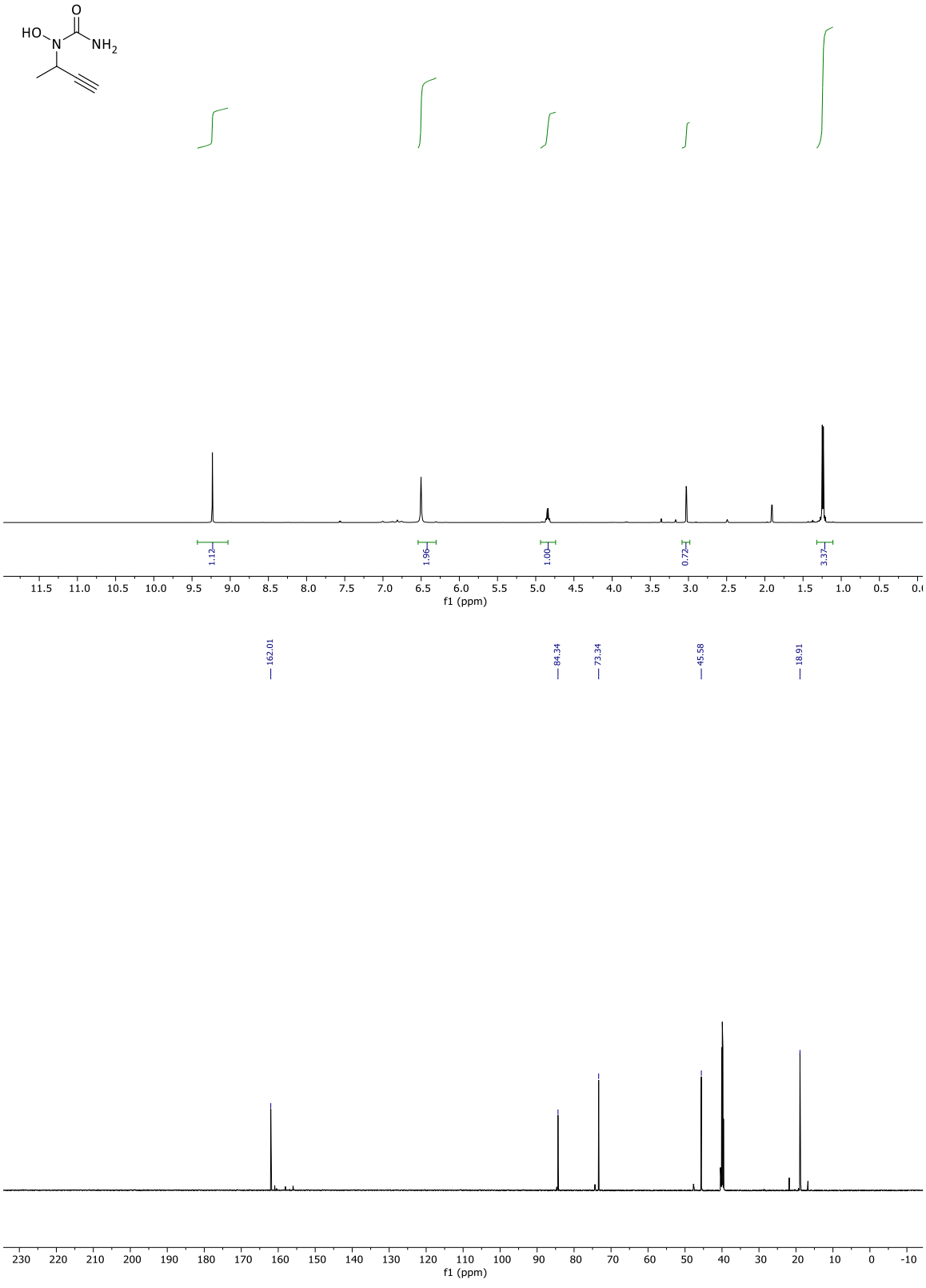


Figure A.41: ^{1}H and ^{13}C NMR Spectra of Compound 5-4

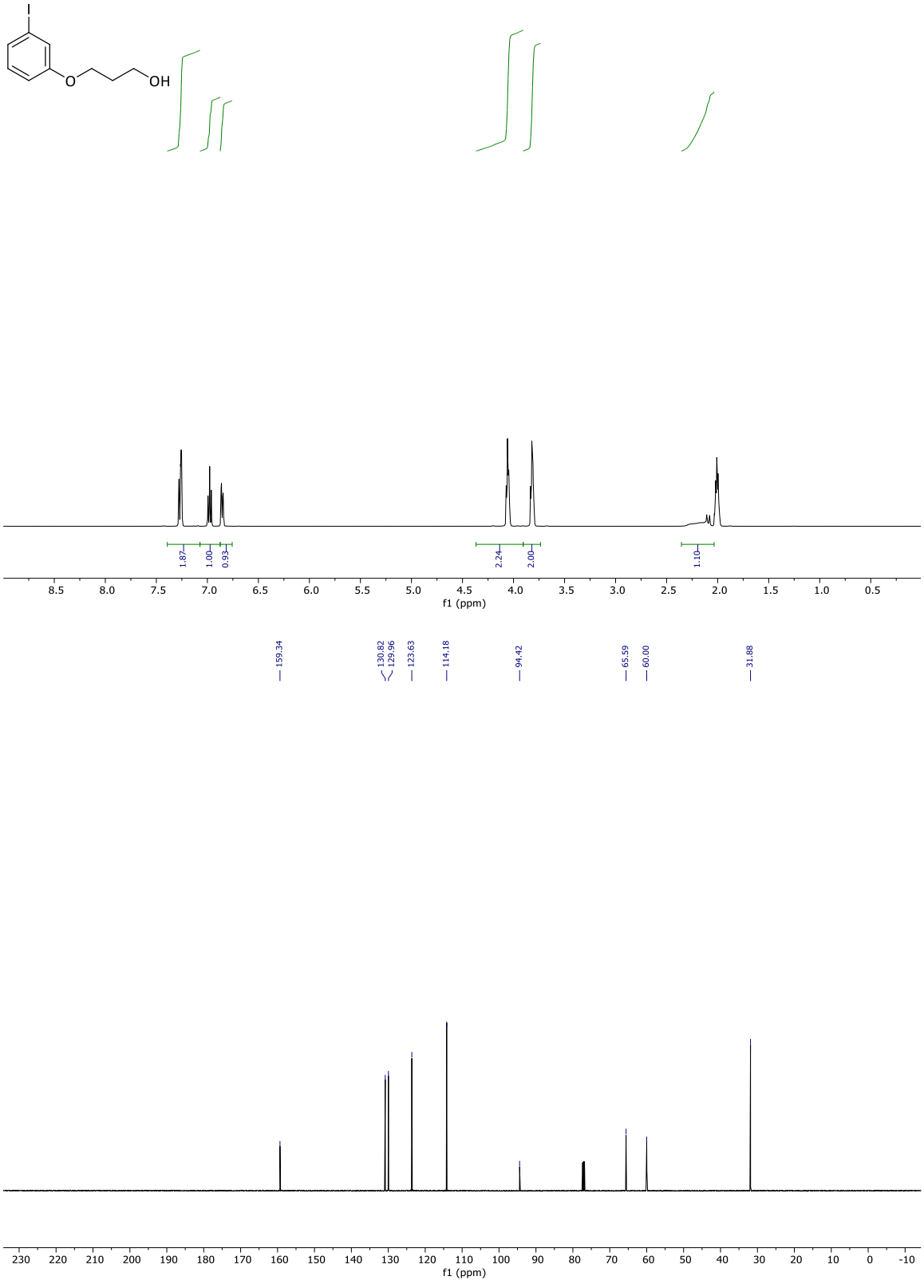


Figure A.42: ^{1}H and ^{13}C NMR Spectra of Compound 5-5

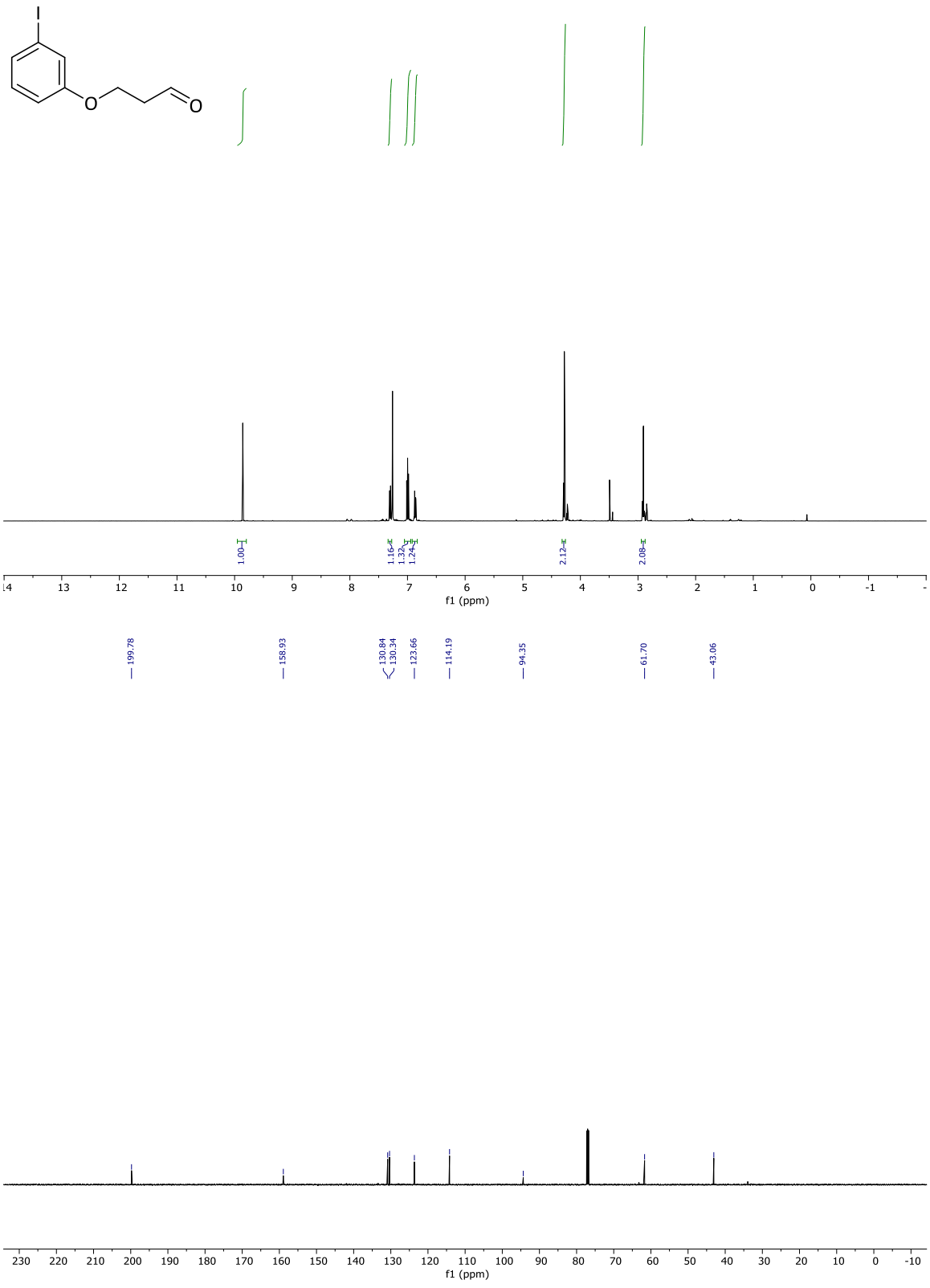


Figure A.43: ^{1}H and ^{13}C NMR Spectra of Compound 5-6

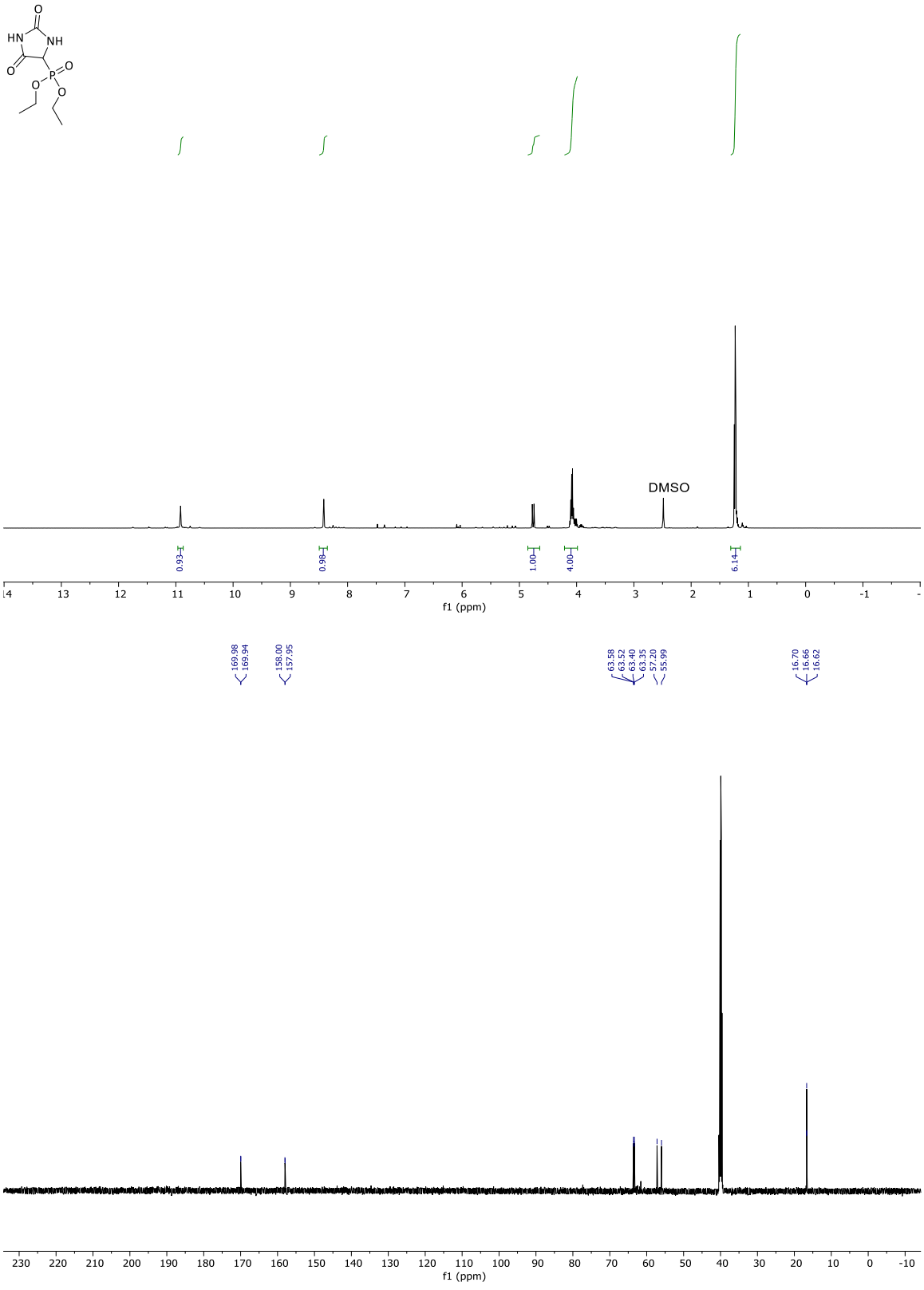


Figure A.44: ^{1}H and ^{13}C NMR Spectra of Compound 5-7

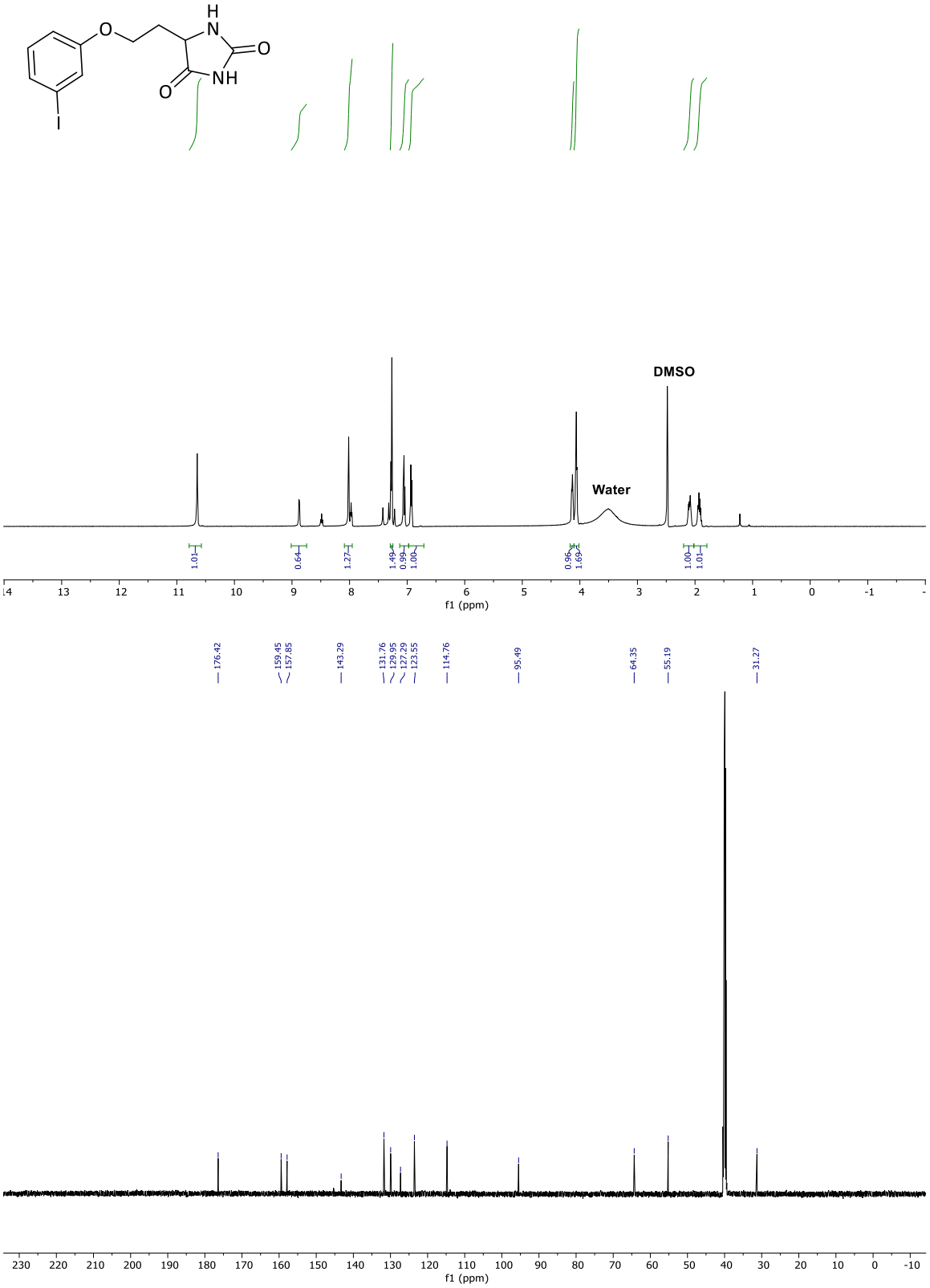


Figure A.45: ^{1}H and ^{13}C NMR Spectra of Compound 5-11

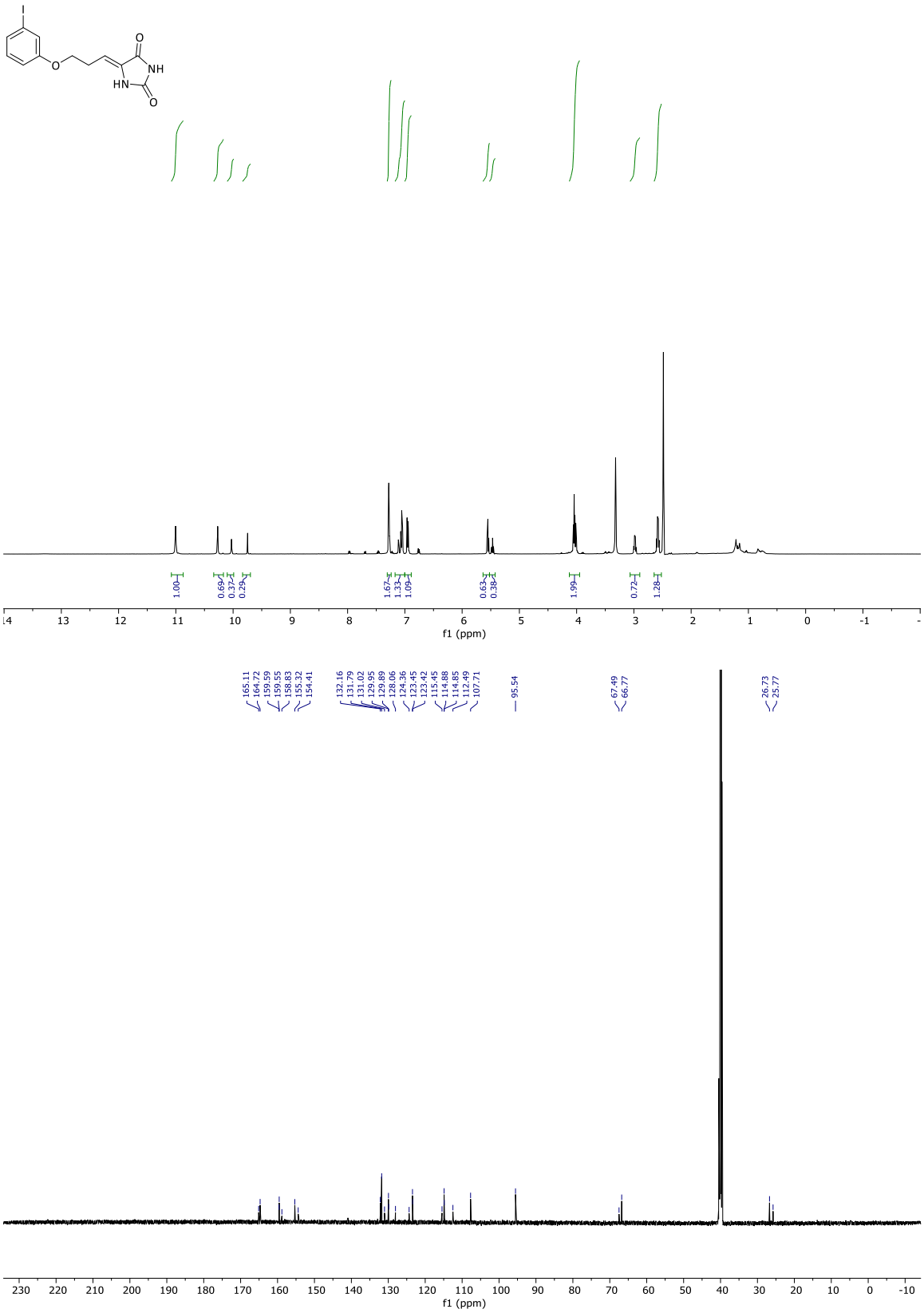


Figure A.46: ^{1}H and ^{13}C NMR Spectra of Compound 5-12

REFERENCES

REFERENCES

1. Kulak, N. A.; Geyer, P. E.; Mann, M., Loss-less Nano-fractionator for High Sensitivity, High Coverage Proteomics. *Molecular & Cellular Proteomics* **2017,** *16* (4), 694-705.

2. Hetz, C.; Glimcher, L. H., Protein homeostasis networks in physiology and disease. *Current opinion in cell biology* **2011,** *23* (2), 123-125.

3. Rouillard, A. D.; Gundersen, G. W.; Fernandez, N. F.; Wang, Z.; Monteiro, C. D.; McDermott, M. G.; Ma'ayan, A., The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. *Database* **2016,** *2016*, baw100-baw100.

4. Brehme, M.; Voisine, C.; Rolland, T.; Wachi, S.; Soper, James H.; Zhu, Y.; Orton, K.; Villella, A.; Garza, D.; Vidal, M.; Ge, H.; Morimoto, Richard I., A Chaperome Subnetwork Safeguards Proteostasis in Aging and Neurodegenerative Disease. *Cell Reports* **2014,** *9* (3), 1135-1150.

5. García-Prat, L.; Martínez-Vicente, M.; Perdiguero, E.; Ortet, L.; Rodríguez-Ubreva, J.; Rebollo, E.; Ruiz-Bonilla, V.; Gutarra, S.; Ballestar, E.; Serrano, A. L.; Sandri, M.; Muñoz-Cánoves, P., Autophagy maintains stemness by preventing senescence. *Nature* **2016,** *529*, 37.

6. Bartlett, A. I.; Radford, S. E., An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nature Structural &Amp; Molecular Biology* **2009,** *16*, 582.

7. Ellis, R. J.; Minton Allen, P., Protein aggregation in crowded environments. In *Biological Chemistry*, 2006; Vol. 387, p 485.

8. Klaips, C. L.; Jayaraj, G. G.; Hartl, F. U., Pathways of cellular proteostasis in aging and disease. *The Journal of Cell Biology* **2017**.

9. Kaushik, S.; Cuervo, A. M., Proteostasis and aging. *Nature Medicine* **2015,** *21*, 1406.

10. Henning, R. H.; Brundel, B. J. J. M., Proteostasis in cardiac health and disease. *Nature Reviews Cardiology* **2017,** *14*, 637.

11. Tanaka, K.; Matsuda, N., Proteostasis and neurodegeneration: The roles of proteasomal degradation and autophagy. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2014,** *1843* (1), 197-204.

12. Labbadia, J.; Morimoto, R. I., The Biology of Proteostasis in Aging and Disease. *Annual Review of Biochemistry* **2015,** *84* (1), 435-464.

13. Medinas, D. B.; Valenzuela, V.; Hetz, C., Proteostasis disturbance in amyotrophic lateral sclerosis. *Human Molecular Genetics* **2017,** *26* (R2), R91-R104.

14. Powers, E. T.; Morimoto, R. I.; Dillin, A.; Kelly, J. W.; Balch, W. E., Biological and Chemical Approaches to Diseases of Proteostasis Deficiency. *Annual Review of Biochemistry* **2009,** *78* (1), 959-991.

15. Korovila, I.; Hugo, M.; Castro, José P.; Weber, D.; Höhn, A.; Grune, T.; Jung, T., Proteostasis, oxidative stress and aging. *Redox Biology* **2017,** *13*, 550-567.

16. Morley, J. F.; Brignull, H. R.; Weyers, J. J.; Morimoto, R. I., The threshold for polyglutamine-expansion protein aggregation and cellular toxicity is dynamic and influenced by aging in Caenorhabditis elegans. *Proc Natl Acad Sci U S A* **2002,** *99* (16), 10417-22.

17. David, D. C.; Ollikainen, N.; Trinidad, J. C.; Cary, M. P.; Burlingame, A. L.; Kenyon, C., Widespread protein aggregation as an inherent part of aging in C. elegans. *PLoS biology* **2010,** *8* (8), e1000450.

18. Ben-Zvi, A.; Miller, E. A.; Morimoto, R. I., Collapse of proteostasis represents an early molecular event in Caenorhabditis elegans aging. *Proc Natl Acad Sci U S A* **2009,** *106* (35), 14914-9.

19. Taylor, R. C.; Berendzen, K. M.; Dillin, A., Systemic stress signalling: understanding the cell non-autonomous control of proteostasis. *Nature reviews. Molecular cell biology* **2014,** *15* (3), 211-217.

20. Hipp, M. S.; Park, S.-H.; Hartl, F. U., Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends in Cell Biology* **2014,** *24* (9), 506-514.

21. Halliwell, B.; Hu, M.-L.; Louie, S.; Duvall, T. R.; Tarkington, B. K.; Motchnik, P.; Cross, C. E., Interaction of nitrogen dioxide with human plasma Antioxidant depletion and oxidative damage. *FEBS Letters* **1992,** *313* (1), 62-66.

22. Menzel, D. B., The toxicity of air pollution in experimental animals and humans: the role of oxidative stress. *Toxicology Letters* **1994,** *72* (1), 269-277.

23. Hu, M.-L.; Tappel, A. L., POTENTIATION OF OXIDATIVE DAMAGE TO PROTEINS BY ULTRAVIOLET-A AND PROTECTION BY ANTIOXIDANTS. *Photochemistry and Photobiology* **1992,** *56* (3), 357-363.

24. Kappus, H., Oxidative stress in chemical toxicity. *Archives of Toxicology* **1987,** *60* (1), 144-149.

25. Davies, K. J. A., Oxidative stress: the paradox of aerobic life. *Biochemical Society Symposium* **1995,** *61*, 1-31.

26. Appelqvist, H.; Kågedal, K.; Wäster, P.; Öllinger, K., The lysosome: from waste bag to potential therapeutic target. *Journal of Molecular Cell Biology* **2013,** *5* (4), 214-226.

27. Glick, D.; Barth, S.; Macleod, K. F., Autophagy: cellular and molecular mechanisms. *The Journal of pathology* **2010,** *221* (1), 3-12.

28. Hayat, M. A., Chapter 1 - Overview of Autophagy. In *Autophagy: Cancer, Other Pathologies, Inflammation, Immunity, Infection, and Aging*, Hayat, M. A., Ed. Academic Press: 2017; pp 3-90.

29. Feng, Y.; He, D.; Yao, Z.; Klionsky, D. J., The machinery of macroautophagy. *Cell Research* **2013,** *24*, 24.

30. Levine, B.; Mizushima, N.; Virgin, H. W., Autophagy in immunity and inflammation. *Nature* **2011,** *469*, 323.

31. Shang, L.; Chen, S.; Du, F.; Li, S.; Zhao, L.; Wang, X., Nutrient starvation elicits an acute autophagic response mediated by Ulk1 dephosphorylation and its subsequent dissociation from AMPK. *Proceedings of the National Academy of Sciences of the United States of America* **2011,** *108* (12), 4788-4793.

32. Li, W.-w.; Li, J.; Bao, J.-k., Microautophagy: lesser-known self-eating. *Cellular and Molecular Life Sciences* **2012,** *69* (7), 1125-1136.

33. Bandyopadhyay, U.; Kaushik, S.; Varticovski, L.; Cuervo, A. M., The Chaperone-Mediated Autophagy Receptor Organizes in Dynamic Protein Complexes at the Lysosomal Membrane. *Molecular and Cellular Biology* **2008,** *28* (18), 5747-5763.

34. Youle, R. J.; Narendra, D. P., Mechanisms of mitophagy. *Nature Reviews Molecular Cell Biology* **2010,** *12*, 9.

35. Liu, K.; Czaja, M. J., Regulation of lipid stores and metabolism by lipophagy. *Cell Death And Differentiation* **2012,** *20*, 3.

36. Dikic, I., Proteasomal and Autophagic Degradation Systems. *Annual Review of Biochemistry* **2017,** *86* (1), 193-224.

37. Lilienbaum, A., Relationship between the proteasomal system and autophagy. *International Journal of Biochemistry and Molecular Biology* **2013,** *4* (1), 1-26.

38. Goldberg, A. L., Functions of the proteasome: from protein degradation and immune surveillance to cancer therapy. *Biochemical Society Transactions* **2007,** *35* (1), 12-17.

39. Smith, D. M.; Benaroudj, N.; Goldberg, A., Proteasomes and their associated ATPases: A destructive combination. *Journal of Structural Biology* **2006,** *156* (1), 72-83.

40. Finley, D.; Chen, X.; Walters, K. J., Gates, Channels, and Switches: Elements of the Proteasome Machine. *Trends in Biochemical Sciences* **2016,** *41* (1), 77-93.

41. Budenholzer, L.; Cheng, C. L.; Li, Y.; Hochstrasser, M., Proteasome Structure and Assembly. *J Mol Biol* **2017**.

42. Schmidt, M.; Finley, D., Regulation of proteasome activity in health and disease. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2014,** *1843* (1), 13-25.

43. Collins, G. A.; Goldberg, A. L., The Logic of the 26S Proteasome. *Cell* **2017,** *169* (5), 792-806.

44. Pickering, A. M.; Davies, K. J. A., Chapter 6 - Degradation of Damaged Proteins: The Main Function of the 20S Proteasome. In *Progress in Molecular Biology and Translational Science*, Grune, T., Ed. Academic Press: 2012; Vol. 109, pp 227-248.

45. Ronald Hough, G. P., and Martin Rechsteiner, Purification of Two High Molecular Weight Proteasomes from Rabbit Reticulocyte Lysate. *The Journal of Biological Chemistry* **1987,** (262).

46. Lowe, J.; Stock, D.; Jap, B.; Zwickl, P.; Baumeister, W.; Huber, R., Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. *Science* **1995,** *268* (5210), 533-539.

47. Glickman, M. H.; Ciechanover, A., The Ubiquitin-Proteasome Proteolytic Pathway: Destruction for the Sake of Construction. *Physiological Reviews* **2002,** *82* (2), 373-428.

48. Ciechanover, A., Proteolysis: from the lysosome to ubiquitin and the proteasome. *Nature Reviews Molecular Cell Biology* **2005,** *6*, 79.

49. Tanaka, K., The proteasome: overview of structure and functions. *Proceedings of the Japan Academy. Series B, Physical and biological sciences* **2009,** *85* (1), 12-36.

50. Bard, J. A. M.; Goodall, E. A.; Greene, E. R.; Jonsson, E.; Dong, K. C.; Martin, A., Structure and Function of the 26S Proteasome. *Annual Review of Biochemistry* **2018,** *87* (1), 697-724.

51. Choi, W. H.; de Poot, S. A. H.; Lee, J. H.; Kim, J. H.; Han, D. H.; Kim, Y. K.; Finley, D.; Lee, M. J., Open-gate mutants of the mammalian proteasome show enhanced ubiquitin-conjugate degradation. *Nature Communications* **2016,** *7*, 10963.

52. Pickart, C. M.; Eddins, M. J., Ubiquitin: structures, functions, mechanisms. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2004,** *1695* (1), 55-72.

53. Mukhopadhyay, D.; Riezman, H., Proteasome-Independent Functions of Ubiquitin in Endocytosis and Signaling. *Science* **2007,** *315* (5809), 201-205.

54. Schnell, J. D.; Hicke, L., Non-traditional Functions of Ubiquitin and Ubiquitin-binding Proteins. *Journal of Biological Chemistry* **2003,** *278* (38), 35857-35860.

55. Finley, D., Recognition and Processing of Ubiquitin-Protein Conjugates by the Proteasome. *Annual Review of Biochemistry* **2009,** *78* (1), 477-513.

56. Pickart, C. M., Mechanisms Underlying Ubiquitination. *Annual Review of Biochemistry* **2001,** *70* (1), 503-533.

57. Groettrup, M.; Pelzer, C.; Schmidtke, G.; Hofmann, K., Activating the ubiquitin family: UBA6 challenges the field. *Trends in Biochemical Sciences* **2008,** *33* (5), 230-237.

58. Wijk, S. J. L. v.; Timmers, H. T. M., The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. *The FASEB Journal* **2010,** *24* (4), 981-993.

59. Nakayama, K. I.; Nakayama, K., Ubiquitin ligases: cell-cycle control and cancer. *Nature Reviews Cancer* **2006,** *6*, 369.

60. Kwon, Y. T.; Ciechanover, A., The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends in Biochemical Sciences* **2017,** *42* (11), 873-886.

61. Komander, D.; Rape, M., The Ubiquitin Code. *Annual Review of Biochemistry* **2012,** *81* (1), 203-229.

62. Strickland, E.; Hakala, K.; Thomas, P. J.; DeMartino, G. N., Recognition of misfolding proteins by PA700, the regulatory subcomplex of the 26 S proteasome. *J Biol Chem* **2000,** *275* (8), 5565-72.

63. Livneh, I.; Cohen-Kaplan, V.; Cohen-Rosenzweig, C.; Avni, N.; Ciechanover, A., The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. *Cell research* **2016,** *26* (8), 869-885.

64. Zhu, Y.; Wang, W. L.; Yu, D.; Ouyang, Q.; Lu, Y.; Mao, Y., Structural mechanism for nucleotide-driven remodeling of the AAA-ATPase unfoldase in the activated human 26S proteasome. *Nature Communications* **2018,** *9* (1), 1360.

65. Eisele, M. R.; Reed, R. G.; Rudack, T.; Schweitzer, A.; Beck, F.; Nagy, I.; Pfeifer, G.; Plitzko, J. M.; Baumeister, W.; Tomko, R. J.; Sakata, E., Expanded Coverage of the 26S Proteasome Conformational Landscape Reveals Mechanisms of Peptidase Gating. *Cell Reports* **2018,** *24* (5), 1301-1315.e5.

66. Wehmer, M.; Rudack, T.; Beck, F.; Aufderheide, A.; Pfeifer, G.; Plitzko, J. M.; Förster, F.; Schulten, K.; Baumeister, W.; Sakata, E., Structural insights into the functional

cycle of the ATPase module of the 26S proteasome. *Proceedings of the National Academy of Sciences* **2017,** *114* (6), 1305-1310.

67. Enenkel, C., Proteasome dynamics. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **2014,** *1843* (1), 39-46.

68. Nandi, D.; Tahiliani, P.; Kumar, A.; Chandu, D., The ubiquitin-proteasome system. *Journal of Biosciences* **2006,** *31* (1), 137-155.

69. Sadre-Bazzaz, K.; Whitby, F. G.; Robinson, H.; Formosa, T.; Hill, C. P., Structure of a Blm10 Complex Reveals Common Mechanisms for Proteasome Binding and Gate Opening. *Molecular Cell* **2010,** *37* (5), 728-735.

70. Kusmierczyk, A. R.; Kunjappu, M. J.; Kim, R. Y.; Hochstrasser, M., A conserved 20S proteasome assembly factor requires a C-terminal HbYX motif for proteasomal precursor binding. *Nature structural & molecular biology* **2011,** *18* (5), 622-9.

71. Smith, D. M.; Chang, S.-C.; Park, S.; Finley, D.; Cheng, Y.; Goldberg, A. L., Docking of the Proteasomal ATPases' Carboxyl Termini in the 20S Proteasome's α Ring Opens the Gate for Substrate Entry. *Molecular Cell* **2007,** *27* (5), 731-744.

72. Förster, A.; Masters, E. I.; Whitby, F. G.; Robinson, H.; Hill, C. P., The 1.9 Å Structure of a Proteasome-11S Activator Complex and Implications for Proteasome-PAN/PA700 Interactions. *Molecular Cell* **2005,** *18* (5), 589-599.

73. Rabl, J.; Smith, D. M.; Yu, Y.; Chang, S.-C.; Goldberg, A. L.; Cheng, Y., Mechanism of Gate Opening in the 20S Proteasome by the Proteasomal ATPases. *Molecular Cell* **2008,** *30* (3), 360-368.

74. Chen, S.; Wu, J.; Lu, Y.; Ma, Y.-B.; Lee, B.-H.; Yu, Z.; Ouyang, Q.; Finley, D. J.; Kirschner, M. W.; Mao, Y., Structural basis for dynamic regulation of the human 26S proteasome. *Proceedings of the National Academy of Sciences* **2016,** *113* (46), 12991-12996.

75. Dong, Y.; Zhang, S.; Wu, Z.; Li, X.; Wang, W. L.; Zhu, Y.; Stoilova-McPhie, S.; Lu, Y.; Finley, D.; Mao, Y., Cryo-EM structures and dynamics of substrate-engaged human 26S proteasome. *Nature* **2019,** *565* (7737), 49-55.

76. Njomen, E.; Osmulski, P. A.; Jones, C. L.; Gaczynska, M.; Tepe, J. J., Small Molecule Modulation of Proteasome Assembly. *Biochemistry* **2018,** *57* (28), 4214-4224.

77. Wright, P. E.; Dyson, H. J., Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *Journal of Molecular Biology* **1999,** *293* (2), 321-331.

78. Dunker, A. K.; Brown, C. J.; Lawson, J. D.; Iakoucheva, L. M.; Obradović, Z., Intrinsic Disorder and Protein Function. *Biochemistry* **2002,** *41* (21), 6573-6582.

79. Oates, M. E.; Romero, P.; Ishida, T.; Ghalwash, M.; Mizianty, M. J.; Xue, B.; Dosztányi, Z.; Uversky, V. N.; Obradovic, Z.; Kurgan, L.; Dunker, A. K.; Gough, J., D(2)P(2): database of disordered protein predictions. *Nucleic Acids Research* **2013,** *41* (Database issue), D508-D516.

80. Iakoucheva, L. M.; Brown, C. J.; Lawson, J. D.; Obradović, Z.; Dunker, A. K., Intrinsic Disorder in Cell-signaling and Cancer-associated Proteins. *Journal of Molecular Biology* **2002,** *323* (3), 573-584.

81. Galea, C. A.; Wang, Y.; Sivakolundu, S. G.; Kriwacki, R. W., Regulation of Cell Division by Intrinsically Unstructured Proteins: Intrinsic Flexibility, Modularity, and Signaling Conduits. *Biochemistry* **2008,** *47* (29), 7598-7609.

82. Perkins, J. R.; Diboun, I.; Dessailly, B. H.; Lees, J. G.; Orengo, C., Transient Protein-Protein Interactions: Structural, Functional, and Network Properties. *Structure* **2010,** *18* (10), 1233-1243.

83. Tompa, P., Intrinsically disordered proteins: a 10-year recap. *Trends in Biochemical Sciences* **2012,** *37* (12), 509-516.

84. Tompa, P., Intrinsically unstructured proteins. *Trends in Biochemical Sciences* **2002,** *27* (10), 527-533.

85. Uversky, V. N., Introduction to Intrinsically Disordered Proteins (IDPs). *Chemical Reviews* **2014,** *114* (13), 6557-6560.

86. Dunker, A. K.; Babu, M. M.; Barbar, E.; Blackledge, M.; Bondos, S. E.; Dosztányi, Z.; Dyson, H. J.; Forman-Kay, J.; Fuxreiter, M.; Gsponer, J.; Han, K.-H.; Jones, D. T.; Longhi, S.; Metallo, S. J.; Nishikawa, K.; Nussinov, R.; Obradovic, Z.; Pappu, R. V.; Rost, B.; Selenko, P.; Subramaniam, V.; Sussman, J. L.; Tompa, P.; Uversky, V. N., What's in a name? Why these proteins are intrinsically disordered: Why these proteins are intrinsically disordered. *Intrinsically Disordered Proteins* **2013,** *1* (1), e24157.

87. van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M., Classification of Intrinsically Disordered Regions and Proteins. *Chemical Reviews* **2014,** *114* (13), 6589-6631.

88. Habchi, J.; Tompa, P.; Longhi, S.; Uversky, V. N., Introducing Protein Intrinsic Disorder. *Chemical Reviews* **2014,** *114* (13), 6561-6588.

89. Dyson, H J., Making Sense of Intrinsically Disordered Proteins. *Biophysical Journal* **2016,** *110* (5), 1013-1016.

90. Drake, J. A.; Pettitt, B. M., Thermodynamics of Conformational Transitions in a Disordered Protein Backbone Model. *Biophysical Journal* **2018,** *114* (12), 2799-2810.

91. Dyson, H. J.; Wright, P. E., Intrinsically unstructured proteins and their functions. *Nature Reviews Molecular Cell Biology* **2005,** *6*, 197.

92. Uversky, V. N., Functional roles of transiently and intrinsically disordered regions within proteins. *The FEBS Journal* **2015,** *282* (7), 1182-1189.

93. Wright, P. E.; Dyson, H. J., Intrinsically disordered proteins in cellular signalling and regulation. *Nature Reviews Molecular Cell Biology* **2014,** *16*, 18.

94. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K., Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *Journal of Molecular Recognition* **2005,** *18* (5), 343-384.

95. Uversky, V. N., Flexible Nets of Malleable Guardians: Intrinsically Disordered Chaperones in Neurodegenerative Diseases. *Chemical Reviews* **2011,** *111* (2), 1134-1166.

96. Babu, M. M.; van der Lee, R.; de Groot, N. S.; Gsponer, J., Intrinsically disordered proteins: regulation and disease. *Current Opinion in Structural Biology* **2011,** *21* (3), 432-440.

97. Uversky, V. N., Wrecked regulation of intrinsically disordered proteins in diseases: pathogenicity of deregulated regulators. *Frontiers in Molecular Biosciences* **2014,** *1*, 6.

98. Uversky, V. N.; Oldfield, C. J.; Dunker, A. K., Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept. *Annual Review of Biophysics* **2008,** *37* (1), 215-246.

99. Darling, A. L.; Uversky, V. N., Intrinsic Disorder and Posttranslational Modifications: The Darker Side of the Biological Dark Matter. *Frontiers in Genetics* **2018,** *9*, 158.

100. Brucale, M.; Schuler, B.; Samorì, B., Single-Molecule Studies of Intrinsically Disordered Proteins. *Chemical Reviews* **2014,** *114* (6), 3281-3317.

101. Jensen, M. R.; Zweckstetter, M.; Huang, J.-r.; Blackledge, M., Exploring Free-Energy Landscapes of Intrinsically Disordered Proteins at Atomic Resolution Using NMR Spectroscopy. *Chemical Reviews* **2014,** *114* (13), 6632-6660.

102. Wei, G.; Xi, W.; Nussinov, R.; Ma, B., Protein Ensembles: How Does Nature Harness Thermodynamic Fluctuations for Life? The Diverse Functional Roles of Conformational Ensembles in the Cell. *Chemical Reviews* **2016,** *116* (11), 6516-6551.

103. Uversky, V. N., Targeting intrinsically disordered proteins in neurodegenerative and protein dysfunction diseases: another illustration of the D(2) concept. *Expert Review of Proteomics* **2010,** *7* (4), 543-564.

104. Csizmok, V.; Follis, A. V.; Kriwacki, R. W.; Forman-Kay, J. D., Dynamic Protein Interaction Networks and New Structural Paradigms in Signaling. *Chemical Reviews* **2016,** *116* (11), 6424-6462.

105. Azevedo, L. M.; Lansdell, T. A.; Ludwig, J. R.; Mosey, R. A.; Woloch, D. K.; Cogan, D. P.; Patten, G. P.; Kuszpit, M. R.; Fisk, J. S.; Tepe, J. J., Inhibition of the Human Proteasome by Imidazoline Scaffolds. *Journal of Medicinal Chemistry* **2013,** *56* (14), 5974-5978.

106. Crick, F., The packing of [alpha]-helices: simple coiled-coils. *Acta Crystallographica* **1953,** *6* (8-9), 689-697.

107. Lee, B.; Richards, F. M., The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology* **1971,** *55* (3), 379-IN4.

108. Connolly, M., Analytical molecular surface calculation. *Journal of Applied Crystallography* **1983,** *16* (5), 548-558.

109. Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E., A geometric approach to macromolecule-ligand interactions. *J Mol Biol* **1982,** *161* (2), 269-88.

110. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W., Jr., Computational methods in drug discovery. *Pharmacological reviews* **2014,** *66* (1), 334-395.

111. Macalino, S. J.; Gosu, V.; Hong, S.; Choi, S., Role of computer-aided drug design in modern drug discovery. *Archives of pharmacal research* **2015,** *38* (9), 1686-701.

112. Adcock, S. A.; McCammon, J. A., Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chemical Reviews* **2006,** *106* (5), 1589-1615.

113. Usha, T.; Shanmugarajan, D.; Goyal, A. K.; Kumar, C. S.; Middha, S. K., Recent Updates on Computer-aided Drug Discovery: Time for a Paradigm Shift. *Current topics in medicinal chemistry* **2017,** *17* (30), 3296-3307.

114. Pagadala, N. S.; Syed, K.; Tuszynski, J., Software for molecular docking: a review. *Biophysical reviews* **2017,** *9* (2), 91-102.

115. Chen, Y. C., Beware of docking! *Trends in pharmacological sciences* **2015,** *36* (2), 78-95.

116. Brooijmans, N.; Kuntz, I. D., Molecular Recognition and Docking Algorithms. *Annual Review of Biophysics and Biomolecular Structure* **2003,** *32* (1), 335-373.

117. Barril, X., Ligand discovery: Docking points. *Nat Chem* **2014,** *6* (7), 560-561.

118. Meng, X., Zhang, H., Mezei, M., Cui, M., Molecular Docking: A powerful approach for structure-based drug discovery. *Curr Comput Aided Drug Des.* **2011,** *7* (2), 146-157.

119. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J., Docking and scoring in virtual screening for drug discovery: methods and applications. *Nature Reviews Drug Discovery* **2004,** *3*, 935.

120. Halperin, I.; Ma, B.; Wolfson, H.; Nussinov, R., Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics* **2002,** *47* (4), 409-443.

121. Huang, L.; Ho, P.; Chen, C.-H., Activation and inhibition of the proteasome by betulinic acid and its derivatives. *FEBS letters* **2007,** *581* (25), 4955-4959.

122. Huang, X.; Luan, B.; Wu, J.; Shi, Y., An atomic structure of the human 26S proteasome. *Nature Structural &Amp; Molecular Biology* **2016,** *23*, 778.

123. Trott, O.; Olson, A. J., AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of computational chemistry* **2010,** *31* (2), 455-461.

124. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J Comput Chem* **2004,** *25* (9), 1157-74.

125. Chang, C.-e. A.; Chen, W.; Gilson, M. K., Ligand configurational entropy and protein binding. *Proceedings of the National Academy of Sciences* **2007,** *104* (5), 1534-1539.

126. Haelterman, R.; Van Eester, D.; Verleyen, D., Accelerating the solution of a physics model inside a tokamak using the (Inverse) Column Updating Method. *Journal of Computational and Applied Mathematics* **2015,** *279*, 133-144.

127. Kristiansen, J. E.; Mortensen, I., Antibacterial effect of four phenothiazines. *Pharmacology & toxicology* **1987,** *60* (2), 100-3.

128. Dastidar, S. G.; Debnath, S.; Mazumdar, K.; Ganguly, K.; Chakrabarty, A. N., Triflupromazine: a microbicide non-antibiotic compound. *Acta microbiologica et immunologica Hungarica* **2004,** *51* (1-2), 75-83.

129. Mazumder, R.; Ganguly, K.; Dastidar, S. G.; Chakrabarty, A. N., Trifluoperazine: a broad spectrum bactericide especially active on staphylococci and vibrios. *International journal of antimicrobial agents* **2001,** *18* (4), 403-6.

130. VARGA, B.; CSONKA, Á.; CSONKA, A.; MOLNÁR, J.; AMARAL, L.; SPENGLER, G., Possible Biological and Clinical Applications of Phenothiazines. *Anticancer Research* **2017,** *37* (11), 5983-5993.

131. Dronca, R. S.; Loprinzi, C., Chapter 18 - Nausea and Vomiting. In *Management of Cancer in the Older Patient*, Naeim, A.; Reuben, D. B.; Ganz, P. A., Eds. W.B. Saunders: Philadelphia, 2012; pp 171-181.

132. Massie, S. P., The Chemistry of Phenothiazine. *Chemical Reviews* **1954,** *54* (5), 797-833.

133. Lester, P. A.; Moore, R. M.; Shuster, K. A.; Myers, D. D., Chapter 2 - Anesthesia and Analgesia. In *The Laboratory Rabbit, Guinea Pig, Hamster, and Other Rodents*, Suckow, M. A.; Stevens, K. A.; Wilson, R. P., Eds. Academic Press: Boston, 2012; pp 33-56.

134. Mocko, J. B.; Kern, A.; Moosmann, B.; Behl, C.; Hajieva, P., Phenothiazines interfere with dopaminergic neurodegeneration in Caenorhabditis elegans models of Parkinson's disease. *Neurobiol Dis* **2010,** *40* (1), 120-9.

135. Liemburg, E. J.; Knegtering, H.; Klein, H. C.; Kortekaas, R.; Aleman, A., Antipsychotic medication and prefrontal cortex activation: a review of neuroimaging findings. *European neuropsychopharmacology : the journal of the European College of Neuropsychopharmacology* **2012,** *22* (6), 387-400.

136. Pickar, D.; Litman, R. E.; Konicki, P. E.; Wolkowitz, O. M.; Breier, A., Neurochemical and neural mechanisms of positive and negative symptoms in schizophrenia. *Modern problems of pharmacopsychiatry* **1990,** *24*, 124-51.

137. Silberg, I. A.; Cormos, G.; Oniciu, D. C., Retrosynthetic Approach to the Synthesis of Phenothiazines. In *Advances in Heterocyclic Chemistry*, Katritzky, A. R., Ed. Academic Press: 2006; Vol. 90, pp 205-237.

138. Mitchell, S. C., Phenothiazine: the parent molecule. *Current drug targets* **2006,** *7* (9), 1181-9.

139. Jones, C. L.; Njomen, E.; Sjögren, B.; Dexheimer, T. S.; Tepe, J. J., Small Molecule Enhancement of 20S Proteasome Activity Targets Intrinsically Disordered Proteins. *ACS Chemical Biology* **2017,** *12* (9), 2240-2247.

140. Di, L.; Kerns, E. H.; Carter, G. T., Drug-like property concepts in pharmaceutical design. *Curr Pharm Des* **2009,** *15* (19), 2184-94.

141. Shibatani, T.; Ward, W. F., Sodium Dodecyl Sulfate (SDS) Activation of the 20S Proteasome in Rat Liver. *Archives of Biochemistry and Biophysics* **1995,** *321* (1), 160-166.

142. Mobley, D. L.; Dill, K. A., Binding of small-molecule ligands to proteins: "what you see" is not always "what you get". *Structure (London, England : 1993)* **2009,** *17* (4), 489-498.

143. Yale, H. L., The Trifluoromethyl Group in Medical Chemistry. *Journal of Medicinal and Pharmaceutical Chemistry* **1959,** *1* (2), 121-133.

144. Betageri, R.; Zhang, Y.; Zindell, R. M.; Kuzmich, D.; Kirrane, T. M.; Bentzien, J.; Cardozo, M.; Capolino, A. J.; Fadra, T. N.; Nelson, R. M.; Paw, Z.; Shih, D. T.; Shih, C. K.; Zuvela-Jelaska, L.; Nabozny, G.; Thomson, D. S., Trifluoromethyl group as a

pharmacophore: effect of replacing a CF3 group on binding and agonist activity of a glucocorticoid receptor ligand. *Bioorg Med Chem Lett* **2005,** *15* (21), 4761-9.

145. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews* **2001,** *46* (1-3), 3-26.

146. Ghose, A. K.; Viswanadhan, V. N.; Wendoloski, J. J., A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *Journal of combinatorial chemistry* **1999,** *1* (1), 55-68.

147. Trader, D. J.; Simanski, S.; Dickson, P.; Kodadek, T., Establishment of a suite of assays that support the discovery of proteasome stimulators. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2017,** *1861* (4), 892-899.

148. Coleman, R. A.; Trader, D. J., Development and Application of a Sensitive Peptide Reporter to Discover 20S Proteasome Stimulators. *ACS Combinatorial Science* **2018,** *20* (5), 269-276.

149. Kumar, B.; Kim, Y.-C.; DeMartino, G. N., The C Terminus of Rpt3, an ATPase Subunit of PA700 (19 S) Regulatory Complex, Is Essential for 26 S Proteasome Assembly but Not for Activation. *Journal of Biological Chemistry* **2010,** *285* (50), 39523-39535.

150. Kisselev, Alexei F.; van der Linden, W. A.; Overkleeft, Herman S., Proteasome Inhibitors: An Expanding Army Attacking a Unique Target. *Chemistry & Biology* **2012,** *19* (1), 99-115.

151. Gallastegui, N.; Beck, P.; Arciniega, M.; Huber, R.; Hillebrand, S.; Groll, M., Hydroxyureas as Noncovalent Proteasome Inhibitors. *Angewandte Chemie International Edition* **2012,** *51* (1), 247-249.

152. Manasanch, E. E.; Orlowski, R. Z., Proteasome inhibitors in cancer therapy. *Nature Reviews Clinical Oncology* **2017,** *14*, 417.

153. Park, J. E.; Miller, Z.; Jun, Y.; Lee, W.; Kim, K. B., Next-generation proteasome inhibitors for cancer therapy. *Translational Research* **2018,** *198*, 1-16.

154. Coleman, R. A.; Muli, C. S.; Zhao, Y.; Bhardwaj, A.; Newhouse, T. R.; Trader, D. J., Analysis of chain length, substitution patterns, and unsaturation of AM-404 derivatives as 20S proteasome stimulators. *Bioorganic & Medicinal Chemistry Letters* **2019,** *29* (3), 420-423.

155. Kisselev, A. F.; Garcia-Calvo, M.; Overkleeft, H. S.; Peterson, E.; Pennington, M. W.; Ploegh, H. L.; Thornberry, N. A.; Goldberg, A. L., The caspase-like sites of proteasomes, their substrate specificity, new inhibitors and substrates, and allosteric interactions with the trypsin-like sites. *J Biol Chem* **2003,** *278* (38), 35869-77.

156. Huber, E. M.; Heinemeyer, W.; Li, X.; Arendt, C. S.; Hochstrasser, M.; Groll, M., A unified mechanism for proteolysis and autocatalytic activation in the 20S proteasome. *Nature Communications* **2016,** *7*, 10900.

157. Harshbarger, W.; Miller, C.; Diedrich, C.; Sacchettini, J., Crystal structure of the human 20S proteasome in complex with carfilzomib. *Structure* **2015,** *23* (2), 418-24.

158. Fabre, B.; Lambour, T.; Garrigues, L.; Ducoux-Petit, M.; Amalric, F.; Monsarrat, B.; Burlet-Schiltz, O.; Bousquet-Dubouch, M.-P., Label-Free Quantitative Proteomics Reveals the Dynamics of Proteasome Complexes Composition and Stoichiometry in a Wide Range of Human Cell Lines. *Journal of Proteome Research* **2014,** *13* (6), 3027-3037.

159. Dubinsky, L.; Krom, B. P.; Meijler, M. M., Diazirine based photoaffinity labeling. *Bioorganic & Medicinal Chemistry* **2012,** *20* (2), 554-570.

160. Hill, J. R.; Robertson, A. A. B., Fishing for Drug Targets: A Focus on Diazirine Photoaffinity Probe Synthesis. *Journal of Medicinal Chemistry* **2018,** *61* (16), 6945-6963.

161. Kumar, A. B.; Tipton, J. D.; Manetsch, R., 3-Trifluoromethyl-3-aryldiazirine photolabels with enhanced ambient light stability. *Chemical Communications* **2016,** *52* (13), 2729-2732.

162. Smith, D. P.; Anderson, J.; Plante, J.; Ashcroft, A. E.; Radford, S. E.; Wilson, A. J.; Parker, M. J., Trifluoromethyldiazirine: an effective photo-induced cross-linking probe for exploring amyloid formation. *Chemical Communications* **2008,** (44), 5728-5730.

163. deGruyter, J. N.; Malins, L. R.; Baran, P. S., Residue-Specific Peptide Modification: A Chemist's Guide. *Biochemistry* **2017,** *56* (30), 3863-3873.

164. Chen, X.; Wu, Y.-W., Selective chemical labeling of proteins. *Organic & Biomolecular Chemistry* **2016,** *14* (24), 5417-5439.

165. Murata, S.; Yashiroda, H.; Tanaka, K., Molecular mechanisms of proteasome assembly. *Nat Rev Mol Cell Biol* **2009,** *10* (2), 104-115.

166. Schrader, J.; Henneberg, F.; Mata, R. A.; Tittmann, K.; Schneider, T. R.; Stark, H.; Bourenkov, G.; Chari, A., The inhibition mechanism of human 20S proteasomes enables next-generation inhibitor design. *Science* **2016,** *353* (6299), 594-8.

167. Gerecke, C.; Fuhrmann, S.; Strifler, S.; Schmidt-Hieber, M.; Einsele, H.; Knop, S., The Diagnosis and Treatment of Multiple Myeloma. *Deutsches Arzteblatt international* **2016,** *113* (27-28), 470-476.

168. Hewlett, N. M.; Tepe, J. J., Total Synthesis of the Natural Product (±)-Dibromophakellin and Analogues. *Organic Letters* **2011,** *13* (17), 4550-4553.

169. Lansdell, T. A.; Hewlett, N. M.; Skoumbourdis, A. P.; Fodor, M. D.; Seiple, I. B.; Su, S.; Baran, P. S.; Feldman, K. S.; Tepe, J. J., Palau'amine and Related Oroidin Alkaloids

Dibromophakellin and Dibromophakellstatin Inhibit the Human 20S Proteasome. *Journal of Natural Products* **2012,** *75* (5), 980-985.

170. Beck, P.; Lansdell, T. A.; Hewlett, N. M.; Tepe, J. J.; Groll, M., Indolo-Phakellins as β5-Specific Noncovalent Proteasome Inhibitors. *Angewandte Chemie International Edition* **2015,** *54* (9), 2830-2833.

171. Gallastegui, N.; al., e., *Angew. Chem., Int. Ed. Engl.* **2012,** (51), 247-249.

172. Kisselev, A. F.; van der Linden, W. A.; Overkleeft, H. S., *Chem. Bio.* **2012,** *19*, 99-115.

173. Beck, P.; Lansdell, T. A.; Hewlett, N. M.; Tepe, J. J.; Groll, M., *Angew. Chem., Int. Ed. Engl.* **2015,** (54), 2830-2833.

174. Lansdell, T. A.; al, e., *J. Nat. Prod.* **2012,** *75*, 980-985.

175. Frigerio, M.; Santagostino, M.; Sputore, S., A User-Friendly Entry to 2-Iodoxybenzoic Acid (IBX). *The Journal of Organic Chemistry* **1999,** *64* (12), 4537-4538.

176. Gaczynska, M.; Osmulski, P. A.; Gao, Y.; Post, M. J.; Simons, M., Proline- and arginine-rich peptides constitute a novel class of allosteric inhibitors of proteasome activity. *Biochemistry* **2003,** *42* (29), 8663-70.

177. Giletto, M. B.; Osmulski, P. A.; Jones, C. L.; Gaczynska, M. E.; Tepe, J. J., Pipecolic esters as minimized templates for proteasome inhibition. *Organic & Biomolecular Chemistry* **2019**.

178. Nguyen, L. A.; He, H.; Pham-Huy, C., Chiral drugs: an overview. *International journal of biomedical science : IJBS* **2006,** *2* (2), 85-100.

179. Gottlieb, H. E.; Kotlyar, V.; Nudelman, A., NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *The Journal of Organic Chemistry* **1997,** *62* (21), 7512-7515.

180. Chen, S.-Q.; Wang, Q.-M.; Xu, P.-C.; Ge, S.-P.; Zhong, P.; Zhang, X.-H., Iodine-promoted selective 3-selanylation and 3-sulfenylation of indoles with dichalcogenides under mild conditions. *Phosphorus, Sulfur, and Silicon and the Related Elements* **2016,** *191* (1), 100-103.

181. Nakata, E.; Yukimachi, Y.; Nazumi, Y.; Uto, Y.; Maezawa, H.; Hashimoto, T.; Okamoto, Y.; Hori, H., A newly designed cell-permeable SNARF derivative as an effective intracellular pH indicator. *Chemical Communications* **2010,** *46* (20), 3526-3528.

182. Kim, C.; Wallace, J. U.; Chen, S. H.; Merkel, P. B., Effects of Dilution, Polarization Ratio, and Energy Transfer on Photoalignment of Liquid Crystals Using Coumarin-Containing Polymer Films. *Macromolecules* **2008,** *41* (9), 3075-3080.

183. Meanwell, N. A.; Roth, H. R.; Smith, E. C. R.; Wedding, D. L.; Wright, J. J. K., Diethyl 2,4-dioxoimidazolidine-5-phosphonates: Horner-Wadsworth-Emmons reagents for the mild and efficient preparation of C-5 unsaturated hydantoin derivatives. *The Journal of Organic Chemistry* **1991,** *56* (24), 6897-6904.