

**Evaluation of Biological Signaling and Transport Molecules Using
Synthetic, Computational, and Mass Spectrometric Techniques.**

**A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

Brandon James Kennedy

May 2019

Copyright © 2019 by Brandon J. Kennedy

All rights reserved.

DEDICATION

For the work presented herein - I would like to dedicate to my family, most notably my sister, Shana, and my parents, Bridget and Douglas Kennedy, whom have been a paragon of support and encouragement. I believe I would not have been successful in this endeavor if not for their contributions to raising me as a well-rounded, conscientious, and indefatigable individual.

I would also like to dedicate this dissertation in part to my friends and co-workers whom have made this experience enjoyable, educational, and forever memorable. I wish them well in their futures and I hope I have influenced their lives and education as much as they have enriched my own.

Thanks everybody!

:3

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. Shawn R. Campagna, for providing numerous research projects, intellectual stimulation, and experimental freedom to hone my skills as a synthetic organic chemist. I have grown as a scientist, critical thinker, and problem solver under his tutelage and am grateful for the education I received during my time in his research group. I would also like to thank my Ph.D. committee members Dr. Michael D. Best, Dr. Tessa R. Calhoun, and Dr. Gladys Alexandre for their time spent reading my work and guiding my rhetoric toward the successful completion of a philosophical doctorate in chemistry.

I would also like to thank the numerous collaborators I have worked with over the years that have provided me with research opportunities in several disciplines; I have learned so much from each research project I have been involved with. My education would have been less fruitful without collaborations with the labs of Dr. Jason Collier, Dr. Steven Wilhelm, Dr. Sarah Lebeis, Dr. Larry Shimkets, Dr. Elizabeth Fozo, Dr. Abbas Shilabin, Dr. Ralph Lydic, Dr. Helen Baghdoyan, Dr. Candida Rebello, and Dr. Zhihong Xu. I enjoyed addressing your research concerns with a chemical approach and hope that the work I have completed was as valuable to you as it was to me.

I have had the opportunity to build friendships with many colleagues during my appointment at UTK: Dr. Shawn R. Campagna, Dr. Hector Castro-Gonzales, Dr. Amanda L. May, Dr. Stephen P. Dearth, Dr. Carson W. Prevatte, Dr. Abigail T. Farmer, Dr. Allen K. Bourdon, Dr. Alexander R. Fisch, Jamie Rash, Maggie S. Lookadoo, Jordon Rodgerson, Dr. Eric D. Tague, Dr. Caleb M. Gibson, Josh B. Powers, Sarah M. Howard, Ashley M. Lato, Katharina M. Hölland, and Katarina A. Jones. As I had mentioned in the dedication, good luck “y’all”, and I hope I have at least in some way contributed to your education and growth as scientists.

ABSTRACT

There are many unanswered questions at the molecular level of biology, and a chemical approach is most often necessary to answer those questions. A chemical approach in this instance is a combination of synthetic organic chemistry, analytical chemistry, and computational chemistry. Three unrelated projects will be discussed in this document, which utilize a combination of chemical tools to make sense of what is happening at the molecular level in these systems. Each of the projects described in this dissertation rely on my abilities as an analytical and organic chemist to propose substantiated hypotheses on molecular feature, structure, and biological function. I believe that through this work, progress has been made to better understand the chemical knowledge of these systems for future investigations.

Chapter 1 will address drug-induced diabetes, which is a condition that is usually brought on from treatment of inflammation with glucocorticoids; Although these drugs are powerful anti-inflammatory agents, they have a propensity to hinder the secretion of insulin from pancreatic β [beta]-cells. To address this issue, a library of glucocorticoid analogues exhibiting a non-steroidal scaffold were synthesized, characterized, computationally modeled, and tested *in-vitro* as anti-inflammatory drug candidates with little to no effect on insulin production.

Chapter 2 focuses on an isotope tracing experiment using synthetic lipids to differentiate straight-chain lipid species from isobaric congeners containing a terminal isopropyl moiety. This was devised to investigate plausible semiochemicals in *Myxococcus xanthus* towards a developmental phenotype initiated by an unknown lipid potentially containing this motif. This method has several implications in lipidomic and metabolomic analyses of biological systems, being that these isobars are impossible to distinguish from one another without analytical standards.

Chapter 3 is an investigation of microbial iron-sequestration, which is essential for nearly all life forms but is of particular interest in marine microbes. Many microbes use specialized molecules known as siderophores to capture iron from the environment to use for several life-sustaining functions. This final chapter spotlights the use of chromatographic and mass spectrometric techniques to investigate uncharacterized siderophores isolated from the marine bacterium *S. loktanella* sp. SE62.

TABLE OF CONTENTS

INTRODUCTION.....	1
CHAPTER I <i>Evaluation of the anti-inflammatory and insulin-suppressive effects of a series of novel non-steroidal 1-(4-substituted phenyl)pyrazole-based glucocorticoid analogues.....</i>	6
1.1 – Abstract	7
1.2 – Introduction	11
1.2.1 – Glucocorticoids as anti-inflammatory drugs	11
1.2.2 – Synthetic glucocorticoids with specific bioactivity	13
1.2.3 – Molecular modeling to inform structure-activity relationships	14
1.3 – Synthesis and purification of glucocorticoid analogues	17
1.3.1 – Synthetic route and characterization	17
1.3.2 – Optimization of flash chromatography	20
1.3.3 – Synthetic methods	23
1.4 – Molecular modeling.....	41
1.4.1 – Structure activity relationships	41
1.4.2 – Modeling and visualization software	42
1.5 – <i>In-vitro</i> studies with synthetic glucocorticoids	47
1.5.1 – Assessing therapeutic profile	47
1.5.2 – Cell culture and assay parameters	49
1.6 – Results	50
1.6.1 – Inflammatory action of synthetic glucocorticoids.....	50
1.6.2 – Effect on insulin secretion in top-performing GC analogues	51

1.6.3 – Structure-activity relationship of dissociated therapeutic profile	52
1.7 – Conclusion.....	53
1.8 – Chapter I appendix	57
1.8.1 – CCL2 Inflammatory assay data	57
1.8.2 – 3xGRE Transcription assay data	61
1.8.3 – Glucose-stimulated insulin secretion (GSIS) assay data.....	65
1.8.4 – Full characterization data of novel and previously reported compounds	66
1.8.5 – ¹ H-NMR and ¹³ C-NMR spectra.....	80
1.8.6 – 2D/3D representations of ligand-receptor interactions	115
1.8.7 – Tables of ligand orientations in 5NFP pocket.....	125
1.8.8 – EICs and LCMS method for analysis of crude product 13ag.....	130
CHAPTER II <i>Synthetic lipids as Isotope tracers to differentiate branched-chain lipids</i>	
<i>from their straight-chain isobars in LCMS analyses.</i>	133
2.1 – Abstract	134
2.2 – Introduction	138
2.2.1 – Lipidomics and lipid semiochemicals	138
2.2.2 – Differentiating isobars in LCMS analyses	141
2.2.3 – Lipid biosynthesis and isotope tracing	144
2.2.4 – Isolipid model organism: <i>Myxococcus xanthus</i>	146
2.3 – Results and discussion	147
2.3.1 – Analysis of <i>M. xanthus</i> lipids	147
2.3.2 – Isotope tracing in <i>M. xanthus</i>	150
2.3.3 – Synthesis of isotope tracers	153
2.4 – Methods.....	156
2.4.1 – Synthesis of isotopically labeled tracing lipids	156
2.4.2 – Cell culturing and biological incorporation of tracing lipids.....	167
2.4.3 – Lipid extraction, preparative TLC, and sample preparation	167

2.4.4 – LC-HRMS method	169
2.4.5 – Data analysis	170
2.5 – Conclusion.....	171
2.6 – Chapter II appendix.....	174
2.6.1 – Percent-labeling for detected lipids in <i>M. xanthus</i> DK1622.....	183
2.6.2 – Percent-labeling for detected lipids in <i>M. xanthus</i> LS1191.....	191
2.6.3 – Extracted ion chromatograms of homospermidine lipids and labeled congeners detected in <i>M. xanthus</i> DK1161 and LS1191.	198
2.6.4 – ¹ H-NMR of synthetic intermediates and isolipid tracers iC ₁₅ -d ₆ and iC ₅ -d ₃	207
<i>CHAPTER III Isolation and preliminary structural elucidation of unknown siderophore from the marine bacterium Sulfitobacter loktanella sp. SE62.</i>	215
3.1 – Abstract	216
3.2 – Introduction	220
3.2.1 – Siderophores	220
3.2.2 – <i>Sulfitobacter loktanella</i> sp. SE62.....	222
3.2.3 – Natural product isolation	223
3.2.4 – Technologies for small molecule structure elucidation	224
3.3 – Results and discussion	225
3.3.1 – First-pass screening of culture extracts by UHPLC-UV-MS	225
3.3.2 – Colorimetric analysis and preparative thin-layer chromatography	226
3.3.3 – UHPLC-HRMS analysis of Fe-binding TLC fractions	232
3.3.4 – HRMS data to calculate potential molecular formula	233
3.3.5 – LC-MS-MS analysis of putative siderophore	235
3.3.6 – Scale-up purification and NMR analysis of unknown siderophore	236
3.3.7 – Data analysis and construction of a potential siderophore structure.....	237
3.4 – Methods.....	245
3.4.1 – Iron-limited cell culturing.....	245

3.4.2 – Isolation of siderophores	246
3.4.3 – UHPLC-MS and UHPLC-MS-MS analysis of preparative TLC samples.....	248
3.4.4 – Flash chromatography and subsequent NMR analysis	249
3.4.5 – Data analysis	250
3.5 – Conclusion.....	250
 <i>CONCLUSION</i>	<i>253</i>
 <i>REFERENCES</i>.....	<i>257</i>
 <i>VITA</i>	<i>269</i>

LIST OF TABLES

Table 2.6.1 – Table of Bode lipidomics data (DK1622 vs LS1191)	174
Table 2.6.2 – Table of lipidomics data showing incorporation of iC5-d ₃	178

LIST OF FIGURES

Figure 1.1.1 – Human glucocorticoid receptor 5NFP.....	8
Figure 1.1.2 – Top dissociated GC therapy candidates from this study	10
Figure 1.2.1 – Natural vs. synthetic GC ring system	11
Figure 1.2.2 – Synthetic route toward 1-(4-substitued phenyl)pyrazolyl glucocorticoid analogues	15
Figure 1.3.1 - Synthetic modifications to previously published work	18
Figure 1.3.2 – Extracted ion chromatograms of compounds 11ag and 13ag.....	21
Figure 1.4.1 – Ligand-receptor docking	43
Figure 1.4.2 – Ligand orientation within 5NFP binding pocket	44
Figure 1.4.3 – Ligand-residue interaction map and therapeutic index scores	45
Figure 2.2.1 – Lipids expressed during <i>M. Xanthus</i> stringent-response.....	140
Figure 2.2.2 – Isobaric compounds indistinguishable by LC-ESI-MS analysis.....	141
Figure 2.2.3 – Marfey’s reagent for the separation of amino acid enantiomers	143
Figure 2.2.4 – Branched-chain and straight-chain lipid biosynthesis.....	145
Figure 2.3.1 – Lipid composition between <i>M. xanthus</i> strains DK1622 and LS1191.....	149
Figure 2.3.2 – PLSDA and VIP score analysis of unknown spectral features between <i>M. xanthus</i> strains DK1622 and LS1191	151
Figure 2.3.3 – Synthetic route toward $iC_{15}-d_6$	154
Figure 2.3.4 – Attempted synthetic routes toward isovaleric acid	155
Figure 2.3.5 – Synthetic route toward iC_5-d_3	157
Figure 3.1.1 – TLC analysis of crude siderophore extract from Vib1, DSS3 and SE62	217
Figure 3.1.2 – Chrome azurol sulfate assay of iron-binding activity	218
Figure 3.2.1 – Diversity of structure and functionality in siderophores.....	221
Figure 3.3.1 – Total ion chromatograms from the UHPLC-MS analysis of crude siderophore extracts.....	227

Figure 3.3.2 – Select PDA chromatogram from UHPLC-MS analysis of crude siderophore extract	228
Figure 3.3.3 – Thin layer chromatographic analysis of crude siderophore extracts	230
Figure 3.3.4 – Mass spectrum of putative siderophore in <i>Sulfitobacter loktanella</i> sp. SE62	234
Figure 3.3.5 – ^1H -NMR of partially purified SE62 siderophore	238
Figure 3.3.6 – ^{13}C -NMR of partially purified SE62 siderophore	239
Figure 3.3.7 – ^1H - ^1H homonuclear correlation spectrum of partially purified SE62 siderophore.....	240
Figure 3.3.8 – Tandem mass spectrometric analysis of putative SE62 siderophore	241
Figure 3.3.9 – Possible SE62 siderophore structure from MS data	243
Figure 3.3.10 – Predicted ^1H -NMR spectrum of the putative SE62 siderophore structure	244

INTRODUCTION

Chemistry is a fascinating science that is becoming more interdisciplinary as time moves forward, especially in the dependence of biological research on the disciplines of analytical and organic chemistry. The development of new analytical tools and the days of developing novel molecular transformations is certainly not past, but a large focus of this field of physical science has adapted to be more directly influential in probing and analyzing living systems. Analytical chemists are called upon to give detailed information of the molecular makeup of things, from the composition of biological tissues to the rates of enzymatic reactions in living organisms. There is increasing need for analytical technologies and approaches to determine the biochemical mechanisms of disease pathology, chemically resistant microbes, and the understanding of normal function in any number of biological systems. Organic chemists use known techniques and reactions to develop medicines, investigate structure-activity relationships, synthesize commercially unavailable standards to aid analytical chemists in their work, or create isotopically labeled molecular analogues to probe metabolism in microbes or higher-order organisms. Organic chemists are constantly improving on existing molecular scaffolds to improve the action and outcome of medicinal therapies and treatments. The creation of novel drugs with precise biological activities is paramount in the battle against disease and dysfunction in all parts of the animal kingdom. This type of work is highly valued; not only in the medical field, but also in others such as agriculture, sustainable energies, cosmetics, and even the flavor and fragrance industry. The applications of chemical technologies to answer questions about the nature of life seems to be unending and ever-transforming, which is an intellectually invigorating notion.

Medicinal chemistry is an area of research that is constantly evolving, to achieve newer medicines with greater activities and fewer side effects. This can be said for practically every available medicinal treatment that humans have discovered or developed. Even now with the advent of targeted immunotherapies for cancers and other

maladies, previous treatments such as chemotherapy agents that at one time had revolutionized the prognosis of a cancer diagnosis are becoming obsolete in their comparable efficacy, yet combinatorial studies of these therapies are finding promising outcomes for patients.¹⁻² Side effects from effective drug therapies are another aspect of medicinal chemistry that is constantly being improved upon. One such class of molecules that have a huge impact on medicine are anti-inflammatory drugs, which are used eponymously to alleviate inflammation and the downstream conditions caused by an inflammatory response. A major draw-back to the use of these medications is their undesired effects on the immune system, as well as their effect on other signaling cascades that result in damage to by-standing systems such as the insulin producing β -cells of the pancreas.³ It is becoming well-established that diabetes is an inflammatory condition beginning with inflammation of pancreatic β -cells and the subsequent immune response that destroys the β -cells, leading to a diabetic disease-state.³⁻⁶ The reason that anti-inflammatory drugs, specifically glucocorticoids, are causing this immune-mediated β -cell death is that although these medications repress inflammation, they also cause cells to produce pro-inflammatory cytokines that turn the immune system against the same cells the therapy is intended to target.⁷

Glucocorticoid-induced diabetes is a condition that could be eradicated by the implementation of an appropriate anti-inflammatory agent that addresses the initial inflammation of the β -cells but does not lead to an immune response that invariably destroys the insulin producing ability of the patient undergoing treatment. Chapter 1 of this dissertation focuses on a series of molecules that have implications as being important in treating inflammation in patients that have a predisposition to diabetes being induced by typical glucocorticoid therapies. Structure-activity relationships were assessed with molecular docking experiments and density functional theory (DFT) calculations to support further research into optimizing a structure with the desired therapeutic profile. Though the molecules synthesized and evaluated for this study are by

no means perfect candidates for replacing dexamethasone or hydrocortisone as anti-inflammatory agents, this work will provide strong evidence that a therapy of this specificity is well within our capabilities as scientists to develop and implement.

Mass spectrometry (MS) has become invaluable in the analysis of a wide variety of samples and liquid chromatography coupled mass spectrometry (LCMS) has become a powerful tool in the analysis of biological systems. The inquiries of recent past, present, and future are focused on unknown aspects of life at the molecular level. Metabolomics is the final stage of the ‘-omics’ cascade, which seeks to identify and quantify small molecules (metabolites) in a biological system and relate them back to phenotype – being that metabolic mass space is a direct consequence of the proteomic, transcriptomic, and genomic capability of a system taking post-translational modification into account. Lipidomics is a subset of metabolomics that seeks to identify and quantify the hydrophobic small molecules present in a biological sample. Chemists have the instrumentation to directly analyze the large and complex pool of molecules such as the metabolome and lipidome but can also utilize isotopically labeled molecules to probe how an organism might metabolize biogenic and xenobiotic compounds into useful things. This gives an analytical chemist the ability to analyze the small molecule differences between mutant strains of microbes, or genetically similar organisms grown in different environmental conditions to see what metabolites are responsible for observable changes in phenotype. Lipids have been thought of to be important in cell-cell communication and other signaling pathways since the mid-twentieth century, but only recently have become routinely analyzed by LCMS to investigate the known and unknown lipids in a sample.

Obtaining exact and accurate mass information on metabolites and lipids is currently a somewhat trivial task, determining the structure of those mass features is much more difficult, especially by the very commonly employed method of electrospray ionization mass spectrometry (ESI-MS). ESI is a soft ionization source that does not

typically induce fragmentation of analytes being introduced to the mass spectrometer. This of benefit in LC-ESI-MS experiments since the identity of the intact analytes are detected and identified by their exact mass and retention time against a library of analytical standards subjected to the same chromatographic method. This however does not allow the analyst to differentiate between compounds of the same exact mass and retention profile that have slight structural differences from one another. This issue can be resolved by using a hard-ionization source such as electron impact (EI) to produce structurally characteristic fragments.; however, EI is not a compatible ionization source under atmospheric pressures due to the negative effect on the source filaments. It would be useful if there was a way to shift the mass of structural isomers from each other so that they may be detected independently. In the second chapter of this dissertation, an isotope tracing experiment was devised to do just this in microbial systems. This was accomplished by feeding microbes a deuterated isolipid building block, isovaleric acid, which would be catabolized by the organisms to produce isotopically labeled compounds featuring a terminal isopropyl unit; thusly permitting simultaneous detection of structurally isomeric lipids with the isolipids shifted higher in mass than their straight-chain isomers.

The analysis of biological systems through mass spectrometry has allowed for scientists to make substantiated conclusions about the influence of small molecules on the overall function of an organism. There are many examples of biologically derived molecules that have potent activities toward pathogens and dysfunctional cells, such as the antibiotic compound penicillin and the anti-cancer compound taxol. These molecules were discovered serendipitously to have beneficial properties in treating disease, but their isolation and identification is what allowed for the chemical production of these compounds for use in medicine. Natural product chemistry is a field in which molecules with interesting properties are isolated and characterized to not only inform the scientific community of their structure and function, but to potentially be synthetic targets for

chemists to produce in order to exploit their useful properties. There is not always a need for biologically derived molecules to be isolated and characterized to find a purpose outside the realm of scientific knowledge. One such case is in the understanding of nutrient uptake and cycling in marine microbes, where low concentrations of organic and inorganic compounds present in the environment pressure the organisms to develop chemical mechanisms to sustain life and thrive in these situations. In particular, the low abundance and bioavailability of iron in these environments coerces the production of small organic molecules with a high-affinity for iron, called siderophores. These iron-binding compounds are not produced by all marine microbes but can be found in most marine environments including surface waters, as dissolved organic compounds, on water-suspended particles, and on the ocean floor.⁸ The preliminary isolation and mass spectrometric analyses of an unknown siderophore produced by a poorly studied *Roseobacter* isolate is described in the third and final chapter of this dissertation. Determining the structure of this unknown iron-binding compound would not only contribute to the scientific literature but depending on potential antimicrobial activity could be used to treat antibiotic resistant infections.

CHAPTER I

EVALUATION OF THE ANTI-INFLAMMATORY AND INSULIN-SUPPRESSIVE EFFECTS OF A SERIES OF NOVEL NON-STEROIDAL 1-(4-SUBSTITUTED PHENYL)PYRAZOLE-BASED GLUCOCORTICOID ANALOGUES.

This chapter contains material that is to be part of a future publication titled:

Evaluation of the anti-inflammatory and insulin-suppressive effects of a series of synthetic non-steroidal glucocorticoid analogues.

Proposed Authors:

Brandon J. Kennedy, Alexander R. Fisch, Susan J. Burke, Ashley M. Lato, Carson W. Prevatte, Justin K. Kirkland, Lee E. Dulap, Russell T. Smith,, Konstantinos Vogiatzis, J. Jason Collier, and Shawn R. Campagna

BJK wrote the paper, synthesized, purified, and characterized the drug candidates, performed all molecular modeling, analyzed the data, generated conclusions, and organized the supplemental information document. ARF assisted in the synthesis, purification, and characterization of the drug candidates. SJB performed the *in-vitro* testing of the compounds synthesized and provided graphical bio-assay data. AML synthesized, purified, and fully characterized specific reaction biproducts and submitted them for biological testing. CWP synthesized the entire series of molecules and submitted them for biological testing, optimized reaction conditions for troublesome synthetic steps. JKK performed all DFT calculations and provided ligand-residue interactions from those simulations. LED as an undergraduate research assistant helped in synthesis, purification, and gathering of characterization data. RTS as an undergraduate research assistant aided in the optimization of reaction conditions and purification methods. JJC and SRC provided funding and conceptualization of this project.

1.1 – Abstract

Glucocorticoids (GCs) have been a mainstay in medicine for the treatment of autoimmune-mediated conditions, certain cancers, and to manage immune response in organ transplantation. GCs are steroid molecules that have many modes of action, usually acting as transcription factors that regulate pro-inflammatory gene expression. The glucocorticoid receptor (GR) (**Figure 1.1.1**) is ubiquitously expressed

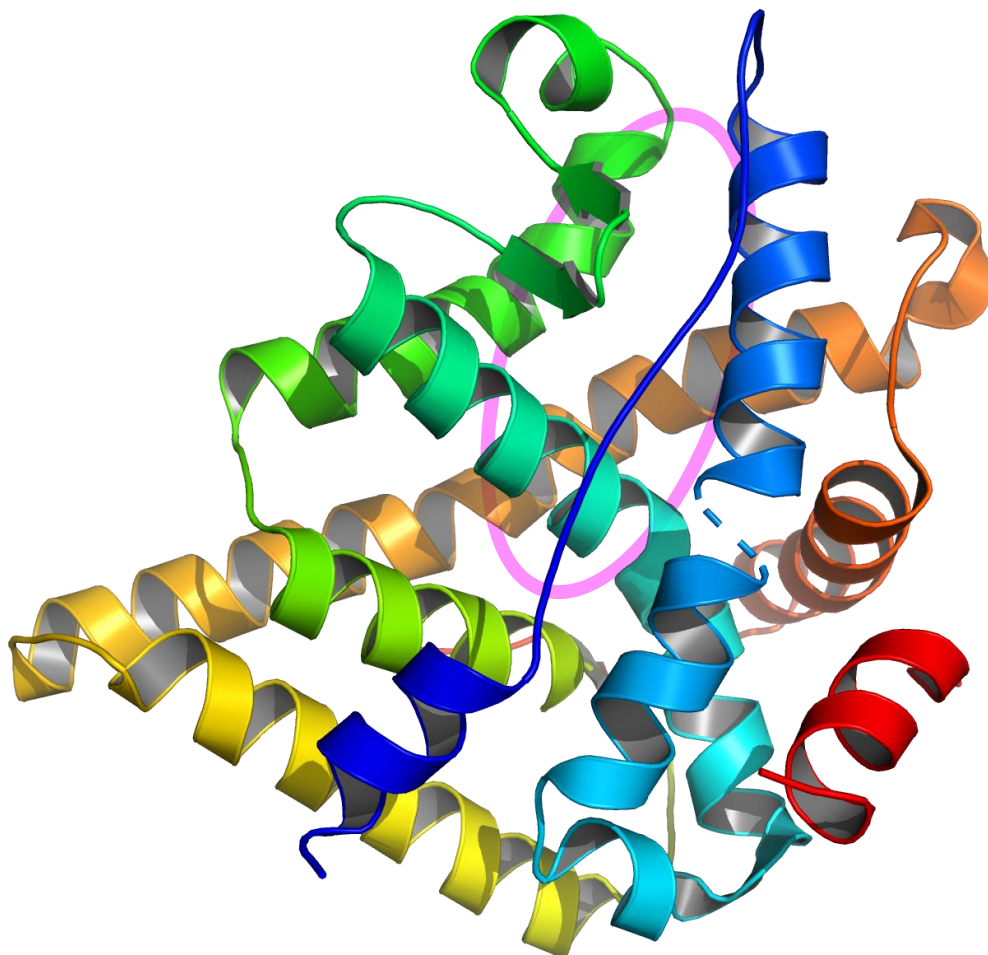


Figure 1.1.1 – Human glucocorticoid receptor 5NFP

Crystal structure of the ligand binding domain of the human GR complex with ligand (budesonide) and protein crystallization additives removed. The ligand binding pocket is circled in pink.

in mammalian tissues, rendering agonistic molecules as valuable therapeutic agents for a variety of conditions requiring the regulation of inflammation in some capacity. Previously reported 1-(4-fluorophenyl)pyrazole-based GC analogues have been shown to reduce interleukin-6 (IL-6) induced inflammation (transrepression), but were accompanied by an undesired upregulation of pro-inflammatory genes associated with GC-induced diabetes, muscle wasting, and osteoporosis (transactivation).⁹ Development of an anti-inflammatory GC with a therapeutic profile that minimizes transactivation could greatly influence the prevention and treatment of inflammation induced type-1 diabetes mellitus (T1DM).¹⁰ In this study, an unexplored series of 1-(4-substituted phenyl)pyrazole-based GC analogues were synthesized and screened for structure-activity relationships that lead to the desired therapeutic profile. Compounds **11aa** and **11ab** (**Figure 1.1.2**) were found to have improved properties, including anti-inflammatory efficacy comparable to dexamethasone, as well as reduced ability to suppress pancreatic β -cell insulin secretion. Repression of IL-1 β -induced inflammation by molecules **11aa** and **11ab** at 1 μ M were comparable to dexamethasone at 10 nM, while also having similar GRE activation. Compound **11aa** reduced insulin secretion at levels compared to the control, dimethylsulfoxide (DMSO), while dexamethasone reduced insulin secretion ~6-fold. Results from the biological testing of these compounds suggest this molecular scaffold to have therapeutic value in treating inflammation in patients whom require unabated insulin function.

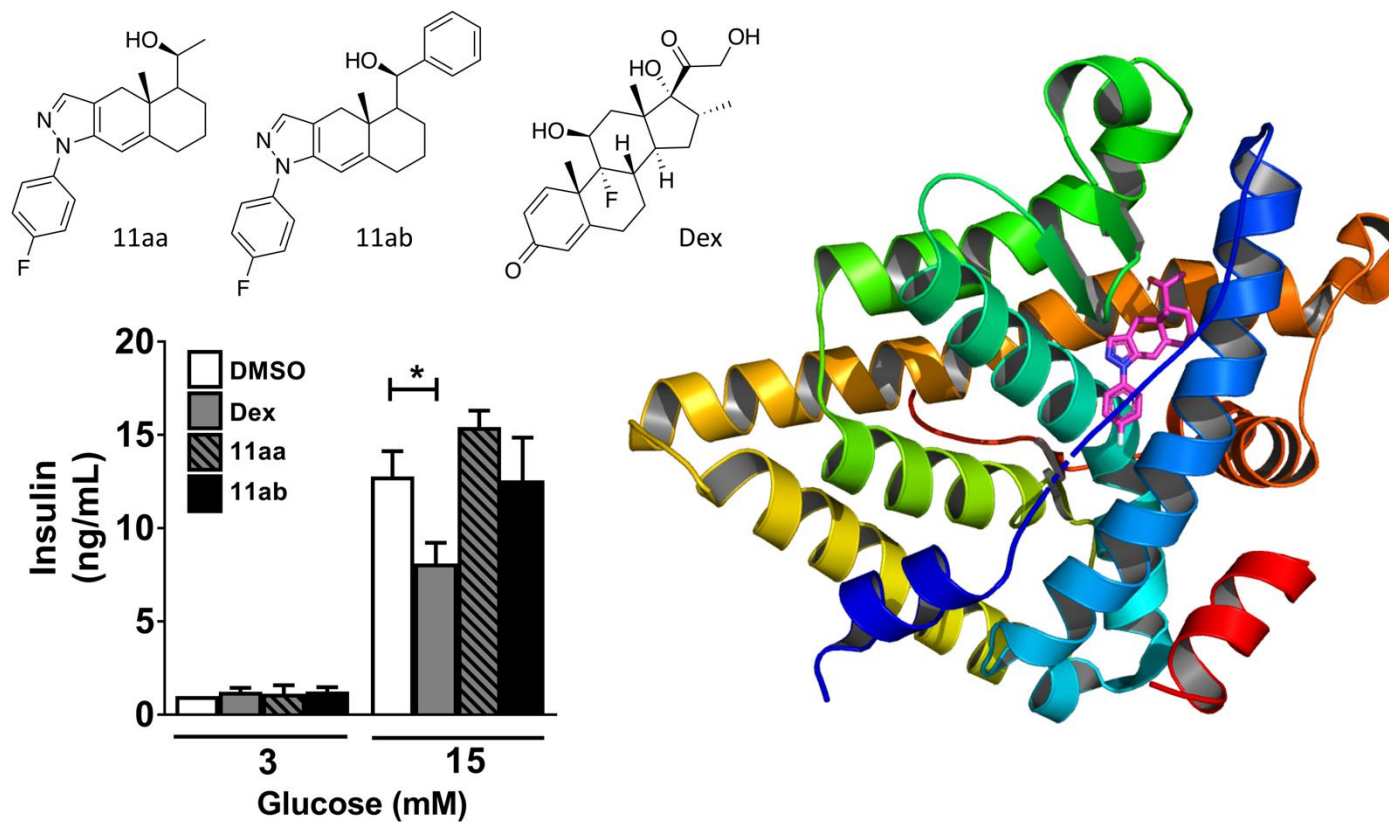


Figure 1.1.2 – Top dissociated GC therapy candidates from this study

Top candidate shown in binding pocket of 5NFP (right) along with glucose stimulated insulin secretion data (bottom left)

1.2 – Introduction

1.2.1 – Glucocorticoids as anti-inflammatory drugs

Non-steroidal drugs such as salicylic acid have been used since ancient times, with records recalling the use of extracts of myrtle and willow tree bark for the management of inflammation.¹¹ It was not until the late 19th century that chemists finally made modifications to these salicylate-containing remedies with the synthesis of acetylsalicylic acid (aspirin) by chemists Felix Hoffmann and Arthur Eichengrün.¹² During the time period between 1930 and 1950, corticosteroids were isolated and structure-activity relationships for their anti-inflammatory action were investigated. From these studies on cortisol and other steroid molecules it was found that steroid drugs bearing a hydroxyl-group on C11 (**Figure 1.2.1**) imparted an effect on immune-regulated inflammatory response but also gluconeogenesis.¹³

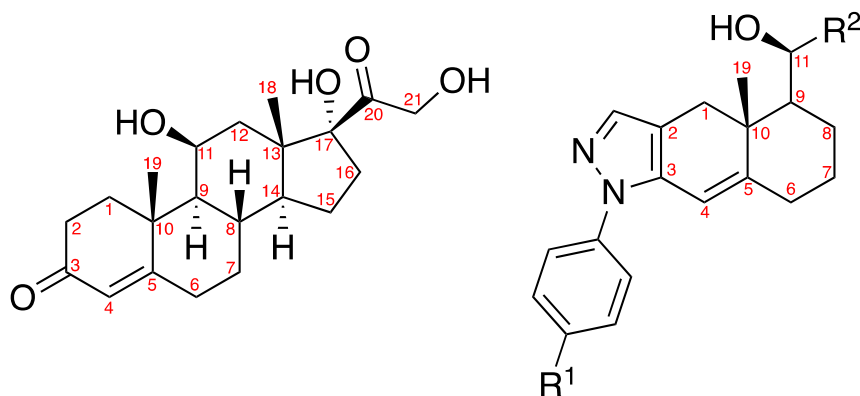


Figure 1.2.1 – Natural vs. synthetic GC ring system

The positioning of the C11 hydroxyl group in traditional steroid molecules is important for glucocorticoid receptor specificity.¹⁴ Carbon atom numbering kept consistent to show structural similarities in terms of scaffold functionalization.

The mechanism of action for aspirin eluded scientists until 1971, finding that it is a potent inhibitor of a family of enzymes called cyclooxygenases (COX-1 and COX-2), which are essential for the formation of prostaglandins that have a variety of downstream effects in vivo.¹⁵ Notably, certain prostaglandins stimulate the production of mucous that protects our gastrointestinal system from the actions of pepsin and acids produced in the stomach. During treatment with aspirin, this mucous production is halted, and can lead to the formation of ulcers in the stomach and intestines and lead to internal bleeding.¹⁶ Indicating a need to not rely on these therapies for all instances of inflammation, especially in chronic inflammatory conditions. Throughout the following years synthetic anti-inflammatory drugs with greater specificities were developed, but to this day there has been little effort towards developing a non-steroidal anti-inflammatory drug (NSAID) that does not affect the production of insulin.

GCs are steroid hormones, possessing both anti-inflammatory and metabolism altering effects, synthesized endogenously in the adrenal glands. GCs operate by acting as a ligand for the glucocorticoid receptor (GR), located in the cellular cytoplasm bound to chaperone proteins. The GR is expressed by nearly all cell types, including pancreatic β -cells, and is the product of a single gene, NR3C1, that is post-translationally modified to produce functionally distinct subtypes of the GR.⁵ These functionally differentiable GR subtypes are one of the greatest contributors to the numerous side-effects of NSAID therapy and why developing drug candidates with specific activities is so challenging. Post-translational modifications to proteins are something that is not encoded by the genome and are very difficult to characterize the various modified forms of the GR, and thusly their individual influence on transcription. The GR has at least six modes of activation at the cellular level. The ligand-bound GR (GR-GC) can translocate to the nucleus, whereby it can bind to a glucocorticoid-response element (GRE) to (1) induce or (2) suppress gene expression; additionally, while in the nucleus the GR-GC can (3) facilitate or (4) hinder the actions of other transcription factors. Alternatively, the GR-GC

can remain in the cytoplasm and (5) increase, or (6) decrease, the activity of proximal proteins.¹⁷ Due to the anti-inflammatory and metabolic-altering effects of GCs, along with the ubiquity of the GR, GC-based therapies have been successfully applied to treating a wide range of inflammatory, auto-immune disorders, and cancers as well aiding in organ and tissue transplantation.¹⁷⁻¹⁸

1.2.2 – Synthetic glucocorticoids with specific bioactivity

Application of synthetic GCs as a possible anti-inflammatory drug for type 1 diabetes mellitus (T1DM) and immunosuppressive agent for organ transplants is attractive for a variety of reasons. Diabetes is the seventh leading cause of death in the United States, affecting at least 7.2% of the U.S. population in 2015; 3.3% of these cases are juvenile type 1 diabetes mellitus (T1DM) in children and adolescents ≤ 20 years of age.¹⁹ Early onset T1DM is now thought to be partially caused by inflammation.^{6, 20} Application of synthetic GCs as possible therapeutics for the inflammatory component of T1DM and as immunosuppressive agents for organ transplants is attractive; however, these compounds, pose a long list of serious side effects including Cushing's disease, metabolic syndrome, osteoporosis, and the most common cause of drug-induced diabetes is clinical administration of GCs.³ T1DM is an autoimmune disorder characterized by inflammation and subsequent autoimmune-mediated destruction of the insulin-producing β -cells.²¹ One manner of autoimmune β -cell death commonly associated with T1DM occurs through insulinitis, a process by which β -cell damage and death occur via infiltration by macrophages and T-cells due to an accumulation of pro-inflammatory cytokines such as IL-1 β , IFN- γ , and TNF- α .^{4, 22-23} Cytokine-mediated inflammation of β -cells causes the expression of chemokines CXCL9 and CXCL10, which through binding to integral membrane protein CXCR3 signal the recruitment of T-cells followed by infiltration and cell death.^{7, 22, 24} The ideal GC-based T1DM therapeutic would suppress the autoimmune response, which ultimately destroys the β -cells. Moreover, synthetic GCs

are designed to exhibit a "dissociated profile", possessing enhanced anti-inflammatory activity with reduced side effects.

Two of the six modes of activation for the GR are often focused on as being the major contributors of the deleterious and salutary effects. One being the translocation of the GR-GC to the nucleus to increase transcription of pro-inflammatory genes, termed transactivation. The metabolic dysregulation caused by the direct binding of the GR-GC to GREs is thought to cause several side effects including those already mentioned, as well as hypertension and hyperglycemia. The other significant mode of GR activation results in translocation of the GR-GC to the nucleus to inhibit the action of pro-inflammatory cytokines, such as IL-1 β , NF- κ B, or AP-1, termed transrepression. The repression of various pro-inflammatory genes through the direct binding of the GR-GC to transcription factors are thought to be responsible for the observed anti-inflammatory effects of GCs scaffolds resembling 11, 12, and 13 (**Figure 1.2.2**).²⁵⁻²⁷

Different GREs differentially affect GR-GC binding conformation which implies that a GR-GC that cannot bind one GRE may be able to bind another GRE based on the specific target gene. Furthermore, different types and ratios of transcriptional cofactors are dependent on the cell or tissue type, and interactions with the GR-GC and other transcription factors contribute to the activity and dissociative properties of the GR-GC in terms of tissue type.^{18, 28-30} Synthetic GCs have been shown to improve β -cell development and suppress inflammation induced by IL-1 β ; However, they have also been shown to reduce adult β -cell mass.³¹⁻³² In this study, a series of novel non-steroidal 1-(4-substituted phenyl)pyrazole-based GC analogues were synthesized and subjected to *ex-vivo* assays measuring transactivation and transrepression. The top candidates from these assays were further assayed to determine how insulin production was affected.

1.2.3 – Molecular modeling to inform structure-activity relationships

When developing a drug therapy, biological assays are used to determine their efficacy. The assays chosen are typically targeted for a specific response the therapy is

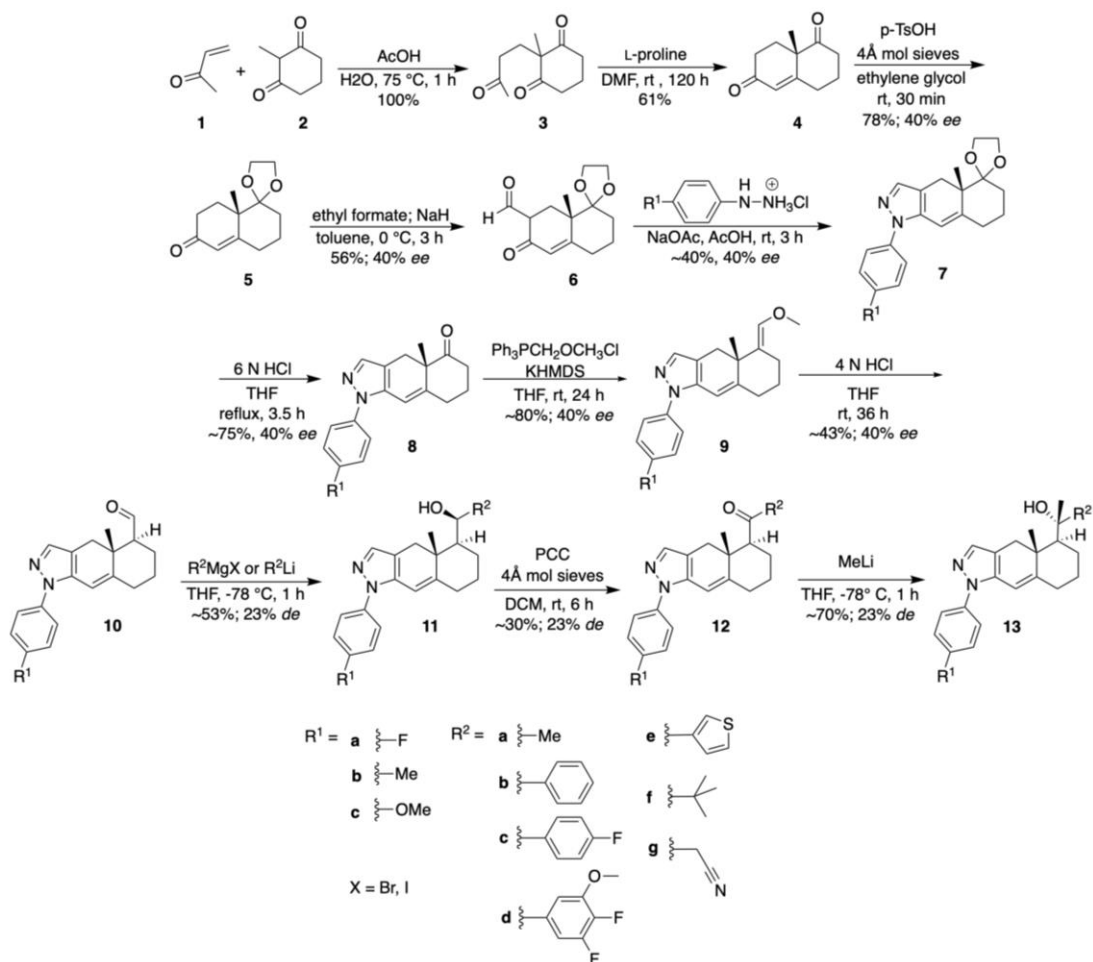


Figure 1.2.2 – Synthetic route toward 1-(4-substituted phenyl)pyrazolyl glucocorticoid analogues

Synthetic procedure adapted from the work of Ali and coworkers.⁹

intended to affect. Knowing the structures of the molecules used as therapeutic agents and the results from these assays provide key information as to how the structures affect the observed activity, but this usually does not describe what interactions are occurring within the ligand-receptor complex, which are contributing to the change in activity. Changes in ligand size, electron accepting/donating capability, hydrogen-bonding, π - π interactions, and dipolar interactions all play a part in how a molecule binds to the target receptor. Computational methods, known collectively as molecular modeling, assist chemists in understanding the physical interactions between a drug and its target receptor. Modeling these interactions inform the chemist of what structural features enhance or diminish the desired bioactivity and guide the logic of creating more effective medications.

The most typically employed methods of molecular modeling are docking experiments between a ligand and a rigid receptor.³³ In these calculations, electrostatic interactions between two bodies are evaluated in a number of physical orientations until the lowest-energy system is achieved, which would be representative of the most likely scenario that the ligand would bind to the receptor. The interacting bodies in this case are a ligand with all allowable rotations (torsions) and a non-moving (rigid) receptor. Receptor proteins are typically modeled in the conformation determined by x-ray crystallography and do not permit allowable torsions within the receptor. It is possible to produce a dynamic protein model; However, modeling an entire protein as a flexible molecule would be computationally exhaustive. These types of calculations need either a powerful processor or a lot of time, usually requiring a super-computer to perform these calculations in a reasonable timeframe. Even calculating the real-time molecular dynamics of a single small protein over the course of 10 microseconds (μ s) took 3 months of supercomputing time for Dr. Kalus Schulten's computational biophysics laboratory at the University of Illinois at Urbana-Champaign.³⁴ Although this 11-year old example exaggerates the time requirements for these calculations on modern computing systems,

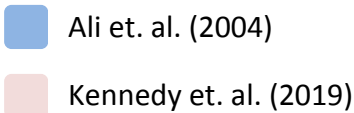
the computational cost of these molecular dynamic calculations for large-body protein systems are still very costly in terms of time and processing power.

There are higher levels of modeling methods than simple docking calculations that go beyond electrostatic interactions, such as DFT and symmetry adapted perturbation theory (SAPT) calculations, which take quantum-level electron density probabilities into account. These methods are also more computationally cumbersome, so simple docking experiments are typically acceptable when determining what residues are involved in the binding of the drug candidates to the target receptor, and how those interactions are playing a role in the bioactivity of the tested compounds.

1.3 – Synthesis and purification of glucocorticoid analogues

1.3.1 – Synthetic route and characterization

The total synthesis of all 1-(4-substituted phenyl)pyrazole-based GC analogues tested in this study were prepared by a previously described synthesis for a series of GC analogues⁹ with a few modifications made to synthetic procedure and substrate scope from Merck's published method (**Figure 1.3.1**).⁹ A simple change of reagent is employed in the conversion of secondary alcohol **11** to ketone **12**, accomplished using a standard PCC (pyridinium chlorochromate) oxidation in place of the costly oxidation with tetrapropylammonium perruthenate (TPAP) using *N*-methyldmorpholine-*N*-oxide (NMO) as a sacrificial oxidant. The molecules investigated by Ali and coworkers (Merck 2004) focused on alterations to the R² moiety while leaving the 4-fluorophenyl R¹ group as a constant. Changing the identity of R¹ was realized through the addition of the appropriate phenylhydrazine to β -ketoaldehyde **6**, thereby installing the desired R¹-arylpyrazole moiety (termed "headgroup"). Additionally, to further enhance the scope of functionalities examined in our biological evaluation was included a cyanomethylene unit as an R² moiety (termed "sidegroup") that was previously unexamined.



Above are shown our synthetic modifications to the previously published synthetic route toward 1-(4-fluorophenyl)pyrazole-based glucocorticoid analogues. Additionally, is shown the scope of target structures between the two studies.⁹

Briefly, the total synthesis of the steroid core begins with a Robinson annulation of compounds **1** and **2** to form intermediate trione **3**, which upon asymmetric intramolecular cyclization gives Wieland-Miescher ketone **4**. Selective ethylene glycol protection of compound **4** leaves the α - β -unsaturated ketone available for enolate formation and subsequent addition to ethyl formate, giving β -ketoaldehyde **6**. Addition of the appropriate phenylhydrazine to **6** completes the core structure of the 1-(4-substituted phenyl)pyrazole-based GCs as compounds **7**. The acid-catalyzed deprotection of acetal **7** provided synthon **8**, which was formylated via Wittig reaction to afford aldehyde **10**, the essential intermediate for the divergent synthesis of all analogues presented in this study. Separation of starting material from the product of the one-pot Wittig/hydrolysis sequence from ketone **8** to aldehyde **10** was futile. In order to obtain aldehyde **10** in high purity, the intermediate enol-ether **9** was isolated from the Wittig reaction, purified via flash chromatography to remove any unreacted starting material, and subsequently subjected to hydrolysis conditions to afford compound **10** in sufficient purity to proceed with the next reaction.

Insertion of the sidegroup (R^2) into the GC scaffold was achieved through generation of the appropriate alkyl lithium reagent followed by addition of aldehyde **10**, resulting in a series of 2° alcohols (**11**) that were subjected to biological testing. Intermediate ketone **12** was obtained through a PCC oxidation of **11**, and after purification was treated with MeLi to afford a series of 3° alcohols (**13**), which were also assayed. Synthesis of undesired byproducts **11ag** and **13ag** were attempted using the reaction conditions for generating compounds **11**, **12**, and **13**. Persistent ^1H -NMR impurities of purified target compounds **11ad**, **11ae**, and **11af** prompted the need to synthesize a *t*-butyl analogue of compound **11**. Addition of *t*-BuLi to aldehyde **10a** was achieved using reaction conditions described in generating the GC analogues in question - with consideration to the overall polarity of the reaction solvent system. The reaction to make **11ag** would not proceed if the volume of polar organic solvent was $\leq 50\%$ of the total

solvent volume. Even though *t*-BuLi is generally a poor nucleophile for additions to carbonyl compounds, the reaction of *t*-BuLi with aldehyde **10** produced a pure yield ~25% suggesting a non-negligible amount of **11ag** is being produced in the reactions affording **11ad**, **11ae**, and **11af**. Compound **13ag** could not be synthesized under the reaction conditions used for generating the other compound **13** analogues, ketone **12ag** was recovered from a reaction with 10 equivalents of MeLi with minimal evidence of **13ag** being synthesized – this was verified by performing an HPLC analysis of crude recovered material where **12ag** was detected as the major component, and **13ag** was essentially detected as instrumental noise (**Figure 1.3.2**; **appendix 1.8.8**).

1.3.2 – Optimization of flash chromatography

Throughout the synthetic procedure are purification steps requiring silica-based flash chromatography. TLC analysis was primarily done using a 25% solution of EtOAc in hexanes. Visualization of resolved compounds using short-wave UV light reveals separations that do not seem challenging, but when applied to a flash column by routine methods result in poor separation. The best chromatographic resolution was obtained using a silica stationary phase and a low-polarity mixture of EtOAc in hexanes as the mobile phase. Other mobile phases such as PetEt/Et₂O, Et₂O/pentane, MeOH/DCM, or THF/hexanes proved to be less effective toward the isolation of the target compounds. Dissolution of the compounds in halogenated solvents allow for wet-loading concentrated solutions to the top of the column but will create unwanted band-broadening. All the compounds purified in this synthesis would dissolve mostly in polar solvents that would obscure the chromatography and result in poor separation. It was found that the best chromatographic resolution can be obtained by loading compounds onto silica, wet-packing the silica column, carefully placing a thin layer of sand to maintain column integrity during compound loading, and finally loading the vacuum-dried compound-adsorbed silica on top of the sand layer. This procedure mimics that of a typical TLC analysis where the compound is loaded onto a small section of silica

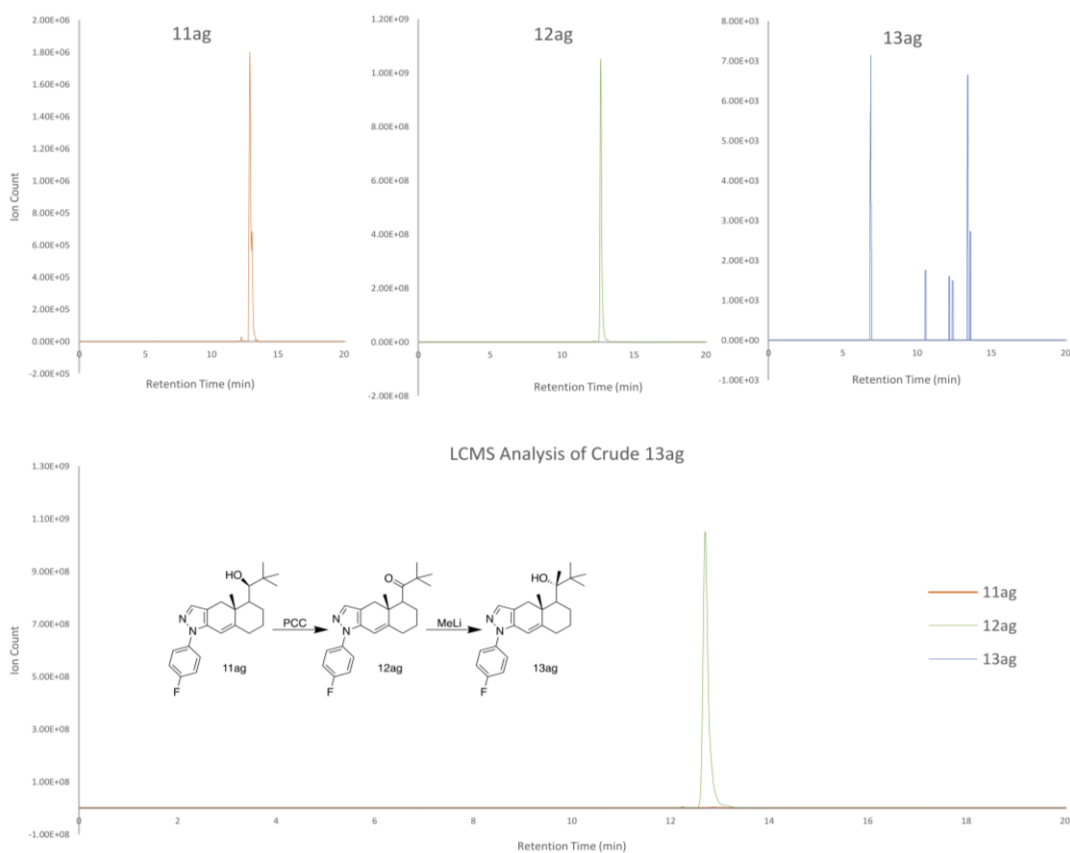


Figure 1.3.2 – Extracted ion chromatograms of compounds 11ag and 13ag

LCMS analysis of crude 3°-OH **13ag** after oxidation of 2°-OH **11ag** and attempted methylation of ketone **12ag** compared to a sample of compound **11ag**. Virtually no **13ag** was detected, and a small population (<0.002% of **13ag**) of 2°-OH **11ag** survived PCC oxidation.

the solvent is allowed to evaporate, and then the whole chromatographic system is subjected to the desired mobile phase composition.

¹H-NMR analysis of the purified intermediates revealed the presence of low intensity and low ppm “impurities” that were attributable to the presence of structural isomers of the desired intermediates and final products. This issue was likely avoided in the previous work of Ali et al. who used preparative chiral HPLC to achieve enantiomeric purity at each intermediate step of the synthesis of their series of *N*-arylpyrazolo[3,2-*c*]-based glucocorticoid analogues.⁹ Preparative HPLC was not an accessible technology in the synthesis of the target compounds, and therefore isomers of intermediates were carried through the entire synthesis of the glucocorticoid analogues reported in this study. It should be noted that these spectral anomalies were fairly consistent across all tested compounds and should not be a significant factor in the responses observed in the biological assays reported.

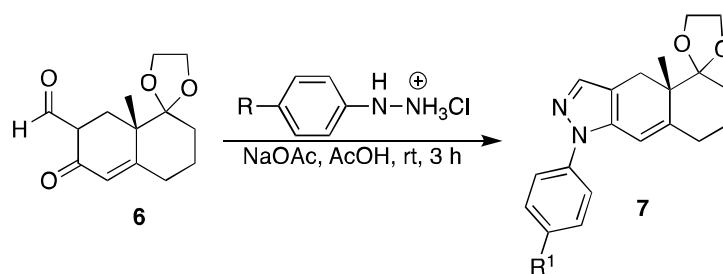
The compounds subjected to biological testing in this preliminary study have a purity ~80% by ¹H-NMR analysis, which is due in part to subtle differential intramolecular relaxation effects, as well as the presence of diastereomers of the final compounds. For compounds bearing R² = 2,3-difluoroanisoyl, 3-thiophenyl, and cyanomethyl - a single byproduct could not be removed by the chromatographic methods employed throughout the synthesis of the target compounds. The byproduct in question originates from the conversion of **10** to **11**, where freshly generated organolithium species using *t*-BuLi are added to aldehydes **10**. Being that lithium-halogen exchange reactions are an equilibrium process based on the comparative stability between the exchanging species, these reactions will generate a small amount of the undesired “*t*-butyl adduct”, which was synthesized and subjected to the same bioassays as all other compounds to determine how the presence of this undesired byproduct affected the observed activities of the screened GC analogues. All intermediates were only carried forward if TLC analysis revealed the purified compound as a single spot by UV visualization and/or TLC staining.

Although the synthetic GCs subjected to *ex-vivo* biological testing were not pure single molecules, the fact that these “impurities” were found in nearly all the compounds assayed should indicate that the observed activities for the compounds tested are directly related to the desired synthetic GC analogue. The isolation of pure compounds was quite challenging, and since this study is a first-pass screening of biological activity, further efforts in the absolute isolation of the desired compounds was deemed impractical. For future explorations of compounds similar in structure and synthesis – it should be noted that these molecules require rigorous purification strategies for silica-based flash chromatography such as low polarity mobile phase (5-10% EtOAc in hexanes) and several large columns (column dimensions ~ 50 x 4 cm) for single-gram quantities

1.3.3 – Synthetic methods

Procedures for obtaining compound **6** were followed from a previously reported synthesis of *N*-arylpyrazolo[3,2-*c*]-based GC derivatives. The same general procedure can be followed for obtaining both tolyl- and anisoly- ($R^1 = 1(4\text{-methylphenyl})$; $1(4\text{-methoxyphenyl})$) GC analogue series starting from compound **6**, the point of synthetic originality for this study. Synthetic procedures, chromatographic methods, and characterization data for intermediates and previously reported GC analogues will not be shown (namely the fluorophenyl series of GC analogues) except for GC analogues which were subjected to biological testing. Chemical reagents were purchased commercially and used without further purification, except for methyl vinyl ketone, which was purified by distillation immediately prior to use. Dry ethereal solvents were obtained by distillation from a potassium-benzophenone ketyl still, all other dry solvents were obtained via distillation over CaH₂(s). Thin-layer chromatography (TLC) was carried out using Sorbent Technologies silica G w/UV254 TLC plates. All purified compounds of interest were visualized as single spots using short-wave UV light; more permanent staining of TLC spots was achieved using vanillin, *p*-anisaldehyde, or ceric ammonium molybdate (Hannesian’s) TLC staining solutions.

^1H -NMR spectra were obtained on an Oxford 300 Mhz or Varian InNova 500 MHz instruments in solutions of CDCl_3 . All ^{13}C -NMR spectra were recorded on a Varian InNova 500 MHz instrument in solutions of CDCl_3 . High resolution mass spectrometry (HRMS) was performed using either a JEOL direct analysis in real-time (DART) source in positive ion mode coupled to a JEOL AccuTOF JMS-T100LC mass spectrometer, or an Applied Biosystems QStar Elite mass spectrometer equipped with a Sciex Turboionspray electrospray ionization (ESI) source in positive ionization mode. Optical rotation was measured using a Perkin-Elmer 241 polarimeter equipped with a sodium source lamp. IR spectra were collected using a Nicolet iS5 spectrophotometer as a dry film on single NaCl(s) plates. LCMS analysis was conducted on a Dionex Ultimate3000 UHPLC system coupled to a Thermo Exactive Plus™ Orbitrap mass spectrometer with an ESI source operating in positive ionization mode (Thermo Scientific; Pittsburgh, PA). All annotated characterization data as well as tabulated bioassay data, NMR spectra of final compounds, and data/methods for the single HPLC analysis can be found in the appendix for the first chapter.



General procedure for GC headgroup installation - compound series 7: To a solution of **6** (5.0 g, 20 mmol) in glacial acetic acid (50 mL) was added the appropriate phenylhydrazine HCl salt (23 mmol) and sodium acetate (1.88 g, 23 mmol) and allowed to stir at room

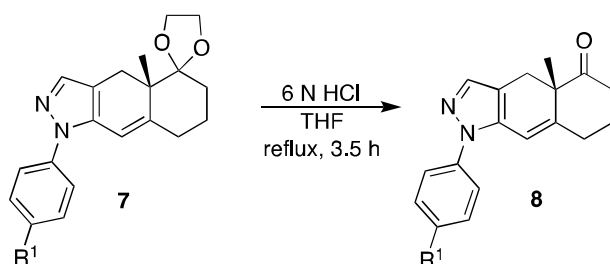
temperature for 2 h. The reaction mixture was then slowly and carefully poured into a cold (0-5 °C) and actively stirred solution of saturated aqueous NaHCO₃. The neutralized solution was then extracted with EtOAc (5x100 mL), washed with brine, dried with anhydrous MgSO₄, filtered, and concentrated in vacuo. Purification of the crude product was accomplished using flash chromatography (20% EtOAc in hexanes) to give a viscous orange oil.

4a-methyl-1-(p-tolyl)-1,4,4a,6,7,8-hexahydrospiro-[benzo[f]indazole-5,2'-

[1,3]dioxolane] (7b): From **6** using p-tolylphenylhydrazine HCl. Yield = 44%. TLC R_f = 0.28 (30% EtOAc/hexanes). ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.21 (s, 3H), 1.65-1.83 (m, 4H), 2.28 (dt, 2H), 2.4 (s, 3H), 2.51 (d, 1H), 3.16 (d, 1H), 3.99-4.08 (m, 4H), 6.23 (s, 1H), 7.24-7.26 (d, 2H), 7.35-7.37 (d, 2H), 7.43 (s, 1H).

1-(4-methoxyphenyl)-4a-methyl-1,4,4a,6,7,8-hexahydrospiro[benzo[f]indazole-5,2'-

[1,3]dioxolane] (7c): From **6** using p-methoxyphenylhydrazine HCl. Yield = 38%. TLC R_f = 0.27 (50% EtOAc/hexanes). ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.22 (s, 3H), 1.68-1.83 (m, 3H), 2.30-2.31 (d, 1H), 2.51-2.54 (d, 2H), 3.15-3.18 (d, 2H), 3.85 (s, 3H), 4.01-4.13 (m, 4H), 6.17 (s, 1H), 6.97-7.41 (m, 5H).

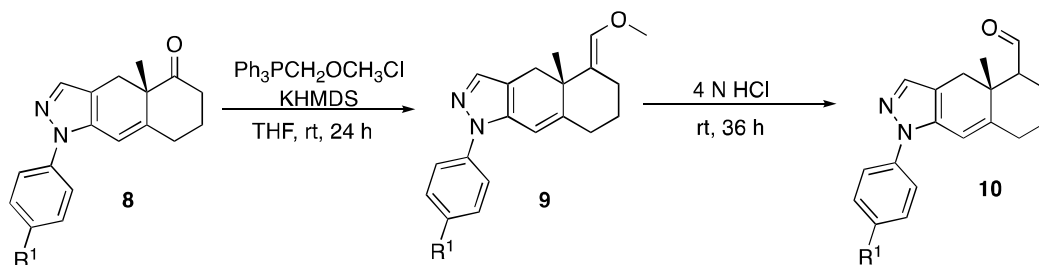


Acetal deprotection for preparation of ketone 8: To a solution of **7** (5.9 mmol) in THF (50 mL) was added 6 M HCl (3.9 mL, 23.6 mmol) and subjected to reflux for 3.5 h. The reaction

mixture was then neutralized by the slow addition of saturated aqueous NaHCO_3 and extracted with ethyl acetate (4x100 mL). The pooled organic layer was washed with brine, dried with anhydrous MgSO_4 , and concentrated in vacuo. Purification of the crude product was accomplished using flash chromatography (20% EtOAc in hexanes) to give a dull orange solid.

4a-methyl-1-(p-tolyl)-1,4,4a,6,7,8-hexahydro-5H-benzo[f]indazol-5-one (8b): From **7b**. Flash column mobile phase - 15% EtOAc/hexanes. Yield = 72%. TLC R_f = 0.34 (30% EtOAc/hexanes). $^1\text{H-NMR}$ (CDCl_3 , 499.73 MHz) δ 1.25 (s, 3H), 1.62-1.71 (m, 1H), 2.06-2.11 (m, 1H), 2.40 (s, 3H), 2.51-2.57 (m, 2H), 2.60-2.67 (m, 2H), 2.90 (s, 2H), 6.33 (s, 1H), 7.26-7.28 (d, 2H), 7.35-7.38 (d, 2H), 7.48 (s, 1H).

1-(4-methoxyphenyl)-4a-methyl-1,4,4a,6,7,8-hexahydro-5H-benzo[f]indazol-5-one (8c): From **7c**. Flash column mobile phase - 20% EtOAc/hexanes. Yield = 77%. TLC R_f = 0.28 (40% EtOAc/hexanes). $^1\text{H-NMR}$ (CDCl_3 , 499.73 MHz) δ 1.24 (s, 3H), 1.60-1.73 (m, 1H), 2.04-2.11 (m, 1H), 2.49-2.69 (m, 4H), 2.93 (d, 2H), 3.85 (s, 3H), 6.26 (s, 1H), 6.95-7.47 (m, 5H).



General Procedure for the synthesis of aldehydes 10: A commercially available solution of potassium bis(trimethylsilyl)amide in toluene (0.5 M, 33.12 mL, 16.5 mmol) was added dropwise with a syringe to an actively stirred suspension of methoxymethyltriphenylphosphonium chloride (6.81 g, 19.8 mmol) in THF (65 mL) cooled

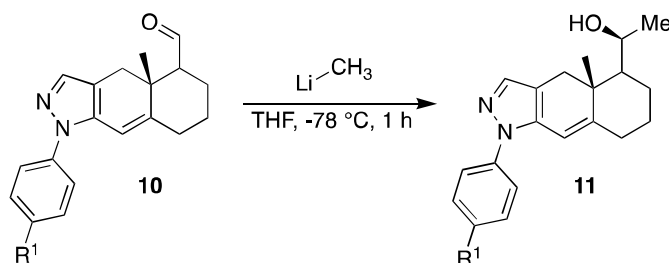
to -40 °C. The solution was warmed to 0 °C and stirred at that temperature for 15 min and a color change from pale yellow to dark red was observed. A solution of the appropriate ketone precursor **8** (6.6 mmol) in THF (16 mL) was transferred dropwise via canula to the chilled ylide. After stirring for 24 h, the reaction was quenched by the addition of 1:1 THF/MeOH (16 mL), diluted with three reaction volumes of EtOAc, washed with water and brine, dried over MgSO₄, filtered, concentrated in vacuo and subjected to flash chromatography (15% EtOAc/hexanes). Fractions containing enol either **9** were collected, concentrated in a round bottom flask, and dissolved in THF. To the prepared solution of **9** in THF was added 4 M HCl and was allowed to stir at room temperature for 36 h. The reaction mixture was afterward diluted with EtOAc - washed with water, saturated aqueous NaHCO₃, and brine. The organic layer was dried over MgSO₄, filtered, and concentrated in vacuo to give crude aldehyde **10** as a beige solid. Purification of the crude product was accomplished using flash chromatography (15% EtOAc/hexanes).

4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazole-5-carbaldehyde

(10b): Yield = 40.0 %. TLC R_f = 0.31 (25% EtOAc/hexanes). ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.12 (s, 3H), 1.30-1.40 (m, 2H) 1.70-1.78 (m, 2H), 1.90-1.96 (m, 1H), 2.32-2.44 (m, 2H), 2.40 (s, 3H), 2.91 (d, 1H), 3.10 (d, 1H), 6.21 (s, 1H), 7.26-7.28 (d, 2H), 7.35-7.37 (d, 2H), 7.44 (s, 1H), 9.90 (s, 1H)

1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazole-5-

carbaldehyde (10c): Yield = 45.0 %. TLC R_f = 0.28 (25% EtOAc/hexanes). ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.11 (s, 3H), 1.39-1.43 (m, 1H), 1.72-1.75 (m, 1H), 1.90-1.96 (m, 2H), 2.33-2.44 (m, 3H), 2.89-2.92 (d, 1H), 3.08-3.11 (d, 1H), 3.85-3.86 (s, 3H), 6.15 (s, 1H), 6.95-7.42 (m, 5H), 9.90 (s, 1H).



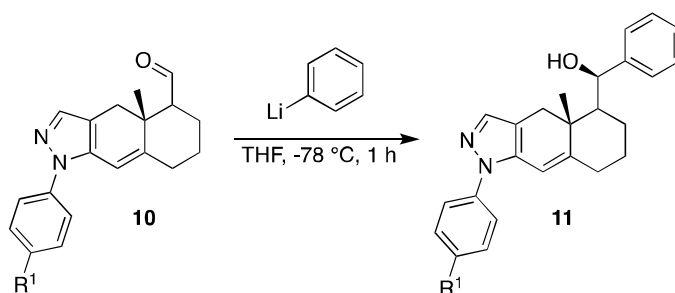
General Procedure for R² = methyl 2°OH analogues: A commercially available solution of 1.6 M MeLi in Et₂O (312 μ L, 0.5 mmol) was transferred via syringe to a round bottom flask containing diethyl ether (5 mL) and cooled to -78 °C. A solution of the appropriate aldehyde **10** (0.05 mmol) in THF (5 mL) was added dropwise via canula, and the reaction was stirred at -78 °C for 45 min. The reaction was quenched with isopropyl alcohol (IPA), poured into a separatory funnel containing saturated aqueous NH₄Cl, and extracted three times with EtOAc. The organic phase was washed with once each with water and brine, dried over anhydrous MgSO₄, filtered, and concentrated in vacuo. Purification of the crude product was accomplished using flash chromatography (15% EtOAc/hexanes) to give a white solid

1-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (11aa): ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.14 (s, 3H), 1.26-1.28 (d, 3H), 1.41-.147 (dd, 2H), 1.63-1.68 (dd, 1H), 1.73-1.80 (m, 1H), 1.90-1.97 (m, 1H), 2.31-2.43 (m, 2H), 2.46-2.52 (d, 1H), 2.98-3.03 (d, 1H), 4.25 (s, 1H), 6.10 (s, 1H), 7.11-7.19 (m, 2H), 7.41-7.47 (m, 3H).

1-(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (11ba): From **10b** using MeLi (1.6 M solution in Et₂O). Yield = 50%. Flash column mobile phase – 15% EtOAc in hexanes. TLC R_f = 0.16 (30% EtOAc/hexanes). HRMS (DART-RTOF)

m/z : $[M+H]^+$ Calcd for $C_{21}H_{27}N_2O$, 323.21179; found 323.21001. 1H -NMR ($CDCl_3$, 499.73 MHz): δ 1.13 (s, 3H), 1.23-1.29 (m, 3H), 1.34-1.45 (m, 2H), 1.56-1.62 (m, 2H), 1.72-1.78 (m, 1H), 1.90-1.94 (m, 1H), 2.27-2.30 (m, 1H), 2.32-2.37 (m, 4H) 2.47-2.50 (d, 1H), 2.97-3.00 (d, 1H), 6.14 (s, 1H), 7.23-7.26 (d, 2H), 7.34-7.36 (d, 2H), 7.40 (s, 1H)

1-(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (11ca): From **10c** using MeLi (1.6 M solution in Et_2O). Yield = 58%. Flash column mobile phase – 30% EtOAc/hexanes. TLC R_f = 0.11 (25% EtOAc/hexanes). HRMS (DART-RTOF) m/z : $[M+H]^+$ Calcd for $C_{21}H_{27}N_2O_2$, 339.20670; found 339.20344. 1H -NMR ($CDCl_3$, 499.73 MHz): δ 1.13 (s, 3H), 1.23-1.29 (m, 4H), 1.34-1.45 (m, 3H), 1.74-1.78 (d, 1H), 1.88-1.95 (d, 1H), 2.28-2.31 (d, 1H), 2.35-2.39 (m, 1H), 2.47-2.51 (d, 1H), 2.98-3.01 (d, 1H), 3.84-3.87 (s, 3H), 6.10 (s, 1H), 6.96-7.40 (m, 5H).



General Procedure for R^2 = phenyl 2°OH analogues: A commercially available solution of 1.9 M PhLi in di-n-butyl ether (3.42 mL, 6.5 mmol) was transferred to a round bottom flask containing a small portion of dry Et_2O and was cooled to $-78\text{ }^\circ\text{C}$. A solution of **10** (0.65 mmol) in THF (2 mL) was added dropwise via canula, and the reaction was stirred at $-78\text{ }^\circ\text{C}$ for 45 min. The reaction was quenched with IPA, poured into a separatory funnel containing saturated aqueous NH_4Cl , and extracted three times with EtOAc. The organic

phase was washed with once each with water and brine, dried over anhydrous MgSO_4 , filtered, and concentrated in vacuo. Purification of the crude product was accomplished using flash chromatography.

(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-

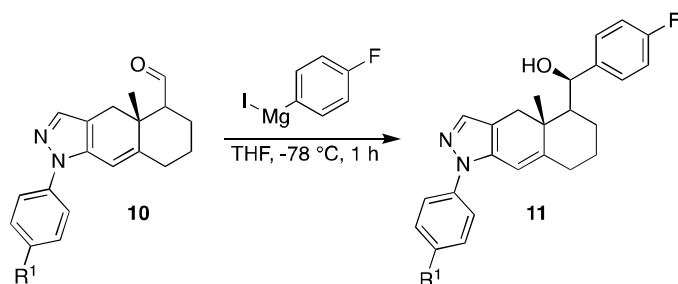
yl)(phenyl)methanol (11ab): $^1\text{H-NMR}$ (CDCl_3 , 499.73 MHz) δ 1.18 (s, 3H), 1.52-1.58 (m, 2H), 1.60-1.63 (dd, 1H), 1.70-1.74 (m, 2H), 2.17-2.21 (m, 1H), 2.28-2.36 (m, 1H), 2.66-2.69 (d, 1H), 3.08-3.11 (d, 1H), 5.10 (d, 1H), 6.02 (s, 1H), 7.04-7.08 (m, 2H), 7.16-7.18 (m, 1H), 7.27-7.29 (m, 4H), 7.35-7.38 (m, 3H).

(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-

yl)(phenyl)methanol (11bb): From **10b** using PhLi (1.9 M solution in di-n-butyl ether). Flash column mobile phase – 15% EtOAc/hexanes, to give a white solid. Yield = 28%. TLC R_f = 0.20 (30% EtOAc/hexanes). HRMS (DART-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{26}\text{H}_{29}\text{N}_2\text{O}$, 385.22744; found 385.22612. $^1\text{H-NMR}$ (CDCl_3 , 499.73 MHz): δ 1.26 (s, 3H), 1.60-1.81 (m, 5H), 2.23-2.28 (d, 1H), 2.36-2.43 (m, 4H), 2.74-2.77 (d, 1H) 3.15-3.18 (d, 1H), 5.17 (s, 1H), 6.15 (s, 1H), 7.23-7.27 (d, 3H), 7.34-7.38 (m, 6H), 7.41 (s, 1H).

(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-

yl)(phenyl)methanol (11cb): From **10c** using PhLi (1.9 M solution in di-n-butyl ether). Flash column mobile phase – 20% EtOAc/hexanes, to give a beige solid. Yield = 73 %. TLC R_f = 0.11 (30% EtOAc/hexanes). HRMS (DART-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{26}\text{H}_{29}\text{N}_2\text{O}_2$, 401.22235; found 401.21952. $^1\text{H-NMR}$ (CDCl_3 , 499.73 MHz): δ 1.27 (s, 3H), 1.56-1.81 (m, 6H), 2.24-2.27 (d, 1H), 2.36-2.44 (m, 1H), 2.75-2.78 (d, 1H), 3.17-3.20 (d, 1H), 3.85 (s, 3H), 5.20 (s, 1H), 6.11 (s, 1H), 6.96-7.44 (m, 9H).



General Procedure for R² = 4-fluorophenyl 2°OH analogues: A commercially available solution of 2.0 M 4-fluorophenylmagnesium bromide in diethyl ether was transferred to a round bottom flask and cooled to -78 °C. A solution of 8 R¹ = Me in THF was added dropwise via canula, and the reaction was stirred at -78 °C for 45 min. The reaction was quenched with IPA, poured into a separatory funnel containing saturated aqueous NH₄Cl, and extracted three times with EtOAc. The organic phase was washed with once each with water and brine, dried over anhydrous MgSO₄, filtered, and concentrated in vacuo. Purification of the crude product was accomplished using flash chromatography (10% EtOAc in to give a white solid.

(4-fluorophenyl)(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-

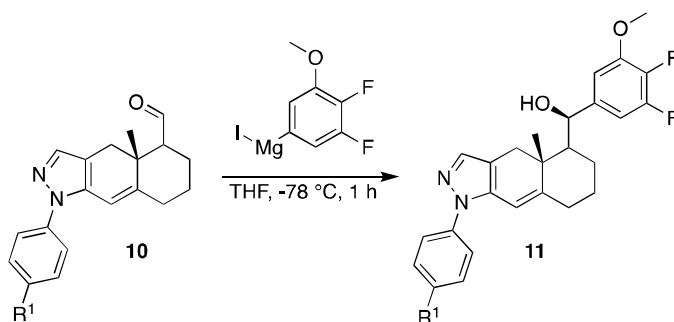
benzo[f]indazol-5-yl)methanol (11ac): ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.26 (s, 3H), 1.65-1.69 (m, 3H), 1.77-1.86 (m, 1H), 2.26-2.41 (m, 2H), 2.72-2.77 (d, 1H), 3.17-3.22 (d, 1H), 5.19-5.22 (s, 1H), 6.12 (s, 1H), 7.02-7.08 (t, 2H), 7.13-7.19 (t, 2H), 7.30-7.34 (dd, 2H), 7.44-7.48 (m, 3H).

(4-fluorophenyl)(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-

yl)methanol (11bc): From **10b** using 4-fluorophenylmagnesium bromide (2.0 M in Et₂O). Flash column mobile phase – 10% EtOAc/hexanes. Yield = 11.5 %. TLC R_f = 0.24 (25%

EtOAc/hexanes). HRMS (DART-TOF) m/z : $[M+H]^+$ Calcd for $C_{26}H_{28}FN_2O$, 403.21802; found 403.21649. 1H -NMR ($CDCl_3$, 499.73 MHz): δ 1.26 (s, 3H), 1.50-1.61 (m, 3H), 1.67-1.71 (m, 2H), 1.79-1.82 (d, 1H), 2.25-2.28 (d, 1H), 2.39 (s, 3H), 2.74-2.77 (d, 1H), 3.16-3.20 (d, 1H), 5.19 (s, 1H), 6.16 (s, 1H), 7.02-7.46 (m, 9).

(4-fluorophenyl)(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)methanol (11cc): From **10c** using 4-fluorophenylmagnesiumbromide (2.0M in Et_2O). Flash column mobile phase - 15% EtOAc/hexanes. Yield = 69%. TLC R_f = 0.26 (30% EtOAc/hexanes). HRMS (DART-TOF) m/z : $[M+H]^+$ Calcd for $C_{26}H_{28}FN_2O_2$, 419.21293; found 419.20660. 1H -NMR ($CDCl_3$, 499.73 MHz): δ 1.26 (s, 3H), 1.56-1.58 (m, 2H), 1.66-1.70 (m, 2H), 1.79-1.82 (m, 1H), 2.25-2.28 (m, 1H), 2.37-2.40 (m, 1H), 2.73-2.76 (d, 1H), 3.17-3.20 (d, 1H), 3.85 (s, 3H), 5.20 (d, 1H), 6.11 (s, 1H), 6.96-7.06 (m, 4H), 7.30-7.44 (m, 5H).



General Procedure for R^2 = 2,3-difluoroanisyl 2°OH analogues: A solution of 5-Bromo-2,3-difluoroanisole (4.6 mmol) in THF (10 mL) was cooled to $-78\text{ }^{\circ}C$, to which a commercially available solution of 1.9 M t -BuLi in pentanes (4.6 mmol, 2.9 mL) was added dropwise via syringe. The reaction was stirred for 20 min at $-78\text{ }^{\circ}C$, raised to $0\text{ }^{\circ}C$ for 5-10 min until the solution transitioned from pale yellow to dark yellow-green, and then was

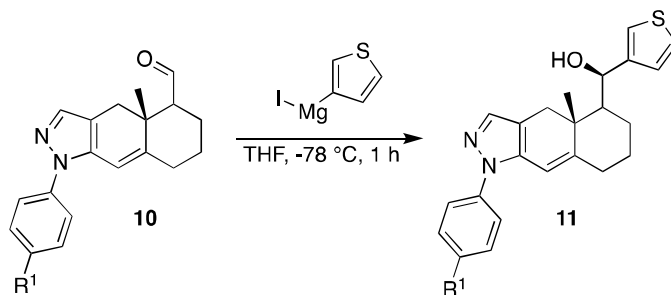
immediately cooled back to -78 °C for 5 min. A solution of **10** (0.46 mmol) in THF (2 mL) was added dropwise via canula, and the reaction was stirred at -78 °C for 1 h. The reaction was quenched with IPA, poured into a separatory funnel containing saturated aqueous NH₄Cl, and extracted three times with EtOAc. The organic phase was washed with water and brine, dried over anhydrous MgSO₄, filtered, and concentrated in vacuo. Purification of the crude product was accomplished using flash chromatography to give a white solid.

(3,4-difluoro-5-methoxyphenyl)(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)methanol (11ad): From **10a** using commercially available 5-bromo-2,3-difluoroanisole. Flash column mobile phase – 30% EtOAc/hexanes. Yield = 61.9 %. TLC R_f = 0.26 (25% EtOAc/hexanes). ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.26 (s, 3H), 1.54-1.57 (m, 1H), 1.64-1.72 (m, 1H), 1.82-1.86 (m, 2H), 2.28-2.46 (m, 3H), 2.72-2.77 (d, 1H), 3.17-3.22 (d, 1H), 3.92 (s, 3H), 5.12 (s, 1H), 6.12 (s, 1H), 6.72-6.80 (m, 2H), 7.15-7.21 (m, 2H), 7.45-7.54 (m, 3H).

(3,4-difluoro-5-methoxyphenyl)(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)methanol (11bd): From **10b** using commercially available 5-bromo-2,3-difluoroanisole. Flash column mobile phase - 10% EtOAc/hexanes. Yield = 76.7 %. TLC R_f = 0.22 (25% EtOAc/hexanes). HRMS (DART-RTOF) *m/z*: [M+H]⁺ Calcd for C₂₇H₂₉F₂N₂O₂, 451.21916; found 451.21790. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.25 (m, 3H), 1.37-1.50 (m, 2H), 1.54-1.65 (m, 2H), 1.76-1.79 (d, 1H), 2.04 (s, 3H), 2.26-2.29 (d, 1H), 2.72-2.75 (d, 1H), 3.14-3.17 (d, 1H), 3.88 (s, 3H), 5.10 (s, 1H), 6.16 (s, 1H), 6.72-6.78 (m, 2H), 7.24-7.46 (m, 6H).

(3,4-difluoro-5-methoxyphenyl)(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)methanol (11cd): From **10c** using commercially available 5-bromo-2,3-difluoroanisole. Flash column mobile phase - 15% EtOAc/hexanes. Yield = 71.4 %. TLC R_f = 0.16 (30% EtOAc/hexanes). HRMS (DART-RTOF) *m/z*: [M+H]⁺ Calcd for C₂₇H₂₉F₂N₂O₃, 467.21408; found 467.21488. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.26 (s, 3H), 1.52-1.56 (m, 2H), 1.66-1.69 (m, 2H), 1.80-1.84 (m, 2H), 2.26-2.29 (m, 1H), 2.73-2.76

(d, 1H), 3.15-3.18 (d, 1H), 3.85 (s, 3H), 3.91 (s, 3H), 5.12 (s, 1H), 6.11 (s, 1H), 6.72-6.79 (m, 2H), 6.96-6.99 (m, 2H), 7.37-7.44 (m, 3H).



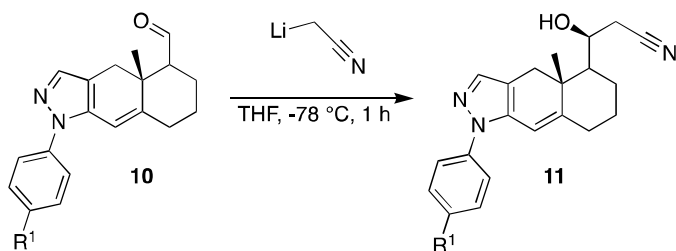
General Procedure for R² = thiophenyl 2°OH analogues: A solution of 3-bromothiophene (0.6 mL, 6.4 mmol) in THF (10 mL) was cooled to -78 °C, to which a commercially available solution of 1.9 M *t*-BuLi in pentanes (6.8 mL, 12.8 mmol) was added dropwise via syringe. After the reaction stirred for 30 min at -78 °C, a solution of **10** (6.4 mmol) in THF (2 mL) was added dropwise via canula, and the reaction was allowed to gradually warm to room temperature over the course of 1.5 h. After returning the reaction to -78 °C, the reaction was quenched with IPA, poured into a separatory funnel containing saturated aqueous NH₄Cl, and extracted three times with EtOAc. The organic phase was washed with water and brine, dried over anhydrous MgSO₄, filtered, and concentrated in vacuo. Purification of the crude product was accomplished using flash chromatography (10% EtOAc/hexanes) to give a white solid.

(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(thiophen-3-yl)methanol (11ae): From **10a** using commercially available 3-bromothiophene. Flash column mobile phase – 10% EtOAc/hexanes. Yield = 35.8%. TLC R_f = 0.26 (25% EtOAc/hexanes). ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.24 (s, 3H), 1.65-1.93 (m,

5H), 2.27-2.48 (m, 2H), 2.67-2.72 (d, 1H), 3.08-3.15 (d, 1H), 5.24 (s, 1H), 6.12 (s, 1H), 6.94-6.99 (m, 1H), 7.12-7.24 (m, 3H), 7.30-7.33 (m, 1H), 7.43-7.48 (m, 3H).

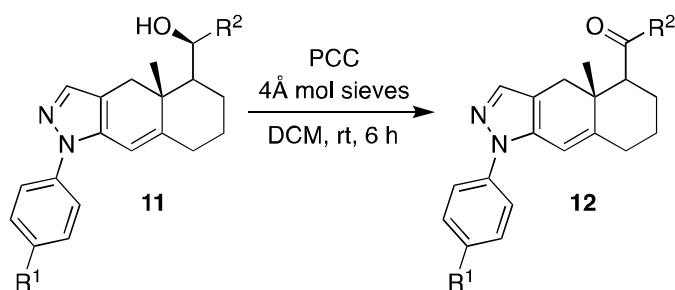
(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(thiophen-3-yl)methanol (11be): From **10b** using commercially available 3-bromothiophene. Flash column mobile phase – 10% EtOAc/hexanes. Yield = 35.8%. TLC R_f = 0.26 (25% EtOAc/hexanes). HRMS (DART-TOF) m/z : $[M+H]^+$ Calcd for $C_{24}H_{27}N_2OS$, 391.18386; found 391.13976. 1H -NMR ($CDCl_3$, 499.73 MHz) δ 1.23 (s, 3H), 1.62-1.91 (m, 6H), 2.26-2.31 (m, 1H), 2.40 (s, 3H), 2.67-2.72 (m, 1H), 3.06-3.14 (m, 1H), 5.39 (s, 1H), 6.16 (s, 1H), 6.95-7.43 (m, 8H).

(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(thiophen-3-yl)methanol (11ce): From **10c** using commercially available 3-bromothiophene. Flash column mobile phase – 15% EtOAc/hexanes. Yield = 79.5%. TLC R_f = 0.17 (30% EtOAc/hexanes). HRMS (ESI-TOF) m/z : $[M+H]^+$ Calcd for $C_{24}H_{27}N_2O_2S$, 407.1787; found 407.1720. 1H -NMR ($CDCl_3$, 499.73 MHz) δ 1.26 (m, 4H), 1.62-1.72 (m, 2H), 1.77-1.85 (m, 2H), 2.27-2.41 (m, 2H), 2.69-2.72 (d, 1H), 3.12-3.15 (d, 1H), 3.85 (s, 3H), 5.25 (s, 1H), 6.12 (s, 1H), 6.96-7.40 (m, 8H).



3-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-3-hydroxypropanenitrile (11af): A commercially available solution of 1.7 M *t*-BuLi in

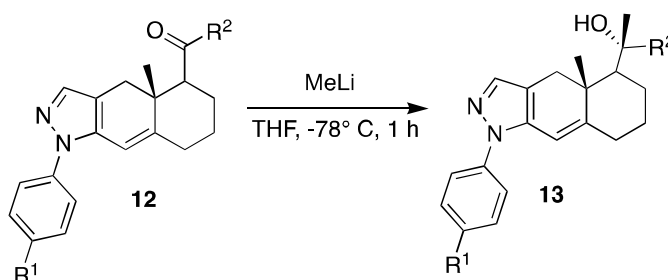
pentanes (0.9 mL, 1.53 mmol) was transferred to a round bottom flask containing THF (2 mL) and cooled to -78 °C. A solution of ACN (80 µL) in THF (2 mL) was added dropwise at -78 °C. Following addition, a solution of **10a** (400 mg, 1.3 mmol) in THF (2 mL) was added dropwise via canula, and the reaction was stirred at -78 °C for 1 h. The reaction was quenched with saturated aqueous NH₄Cl, acidified to a pH of 7 via dropwise addition of 0.5 M HCl, and extracted three times with diethyl ether. The organic phase was washed with water and brine, dried over anhydrous MgSO₄, filtered, and concentrated in vacuo. Purification of the crude product was accomplished using flash chromatography (25% EtOAc/hexanes) to give a white solid. Yield = 45.5%. TLC R_f = 0.14 (25% EtOAc/hexanes). HRMS (DART-TOF) *m/z*: [M+H]⁺ Calcd for C₂₁H₂₃FN₃O, 352.18197; found 352.18293. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.04-1.13 (s, 3H), 1.25-1.46 (m, 2H), 1.61-1.71 (m, 2H), 1.78-1.97 (m, 2H), 2.30-2.42 (m, 2H), 2.47-2.72 (m, 3H), 2.93-3.21 (d, 1H), 4.16-4.40 (m, 1H), 6.12 (s, 1H), 7.14-7.44 (m, 5H).



General procedure for the synthesis of intermediate 12: A solution of the appropriate 2°-OH precursor (**11**) in dichloromethane was transferred dropwise via canula to a suspension of pyridinium chlorochromate (1.5 molar equivalents to compound being oxidized) and freshly activated 4 Å molecular sieves (an equivalent weight to compound

being oxidized) in dichloromethane. The reaction was stirred at room temperature until complete conversion was observed by TLC analysis (reaction times varied from 1-7 h). Upon complete conversion, the reaction mixture was immediately loaded onto a prepared florisil column (packed DCM slurry) and subjected to purification via flash chromatography (100% DCM flush) to give ketone **12** in sufficient purity to move forward with the synthesis.

3-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-3-hydroxybutanenitrile (12af**):** White solid obtained from **11af** with PCC as the oxidizing agent. Yield = 16%. TLC R_f = 0.11 (30% EtOAc/hexanes). HRMS (DART-TOF) m/z : $[M+H]^+$ Calcd for $C_{21}H_{21}FN_3O$, 350.16632; found 350.16529. 1H -NMR ($CDCl_3$, 499.73 MHz) δ 1.20 (s, 3H), 1.45-1.49 (m, 2H), 1.82-1.87 (t, 2H), 1.93-1.97 (d, 1H), 2.32-2.44 (m, 3H), 2.74-2.77 (d, 1H), 2.82-2.88 (m, 2H), 3.53-3.61 (s, 2H), 6.15 (s, 1H), 7.14-7.44 (m, 5H).



General procedure for the synthesis of compound 13 analogues: The appropriate ketone precursor (**12**) was dissolved in diethyl ether and cooled to -40 °C, and a commercially available solution of MeLi (1.9 M, 10 equivalents to ketone starting material) was added dropwise via syringe. The reaction was stirred under argon at -40° for 45 min. The reaction was quenched with IPA (10 equivalents to MeLi added), poured into a separatory funnel

containing saturated aqueous NH_4Cl , and extracted with EtOAc. The organic phase was washed with water and brine, dried over anhydrous MgSO_4 , filtered, and concentrated in vacuo. Purification was accomplished using flash chromatography (see each entry for respective column conditions).

2-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)propan-2-ol (13aa): $^1\text{H-NMR}$ (CDCl_3 , 300.08 MHz) δ 1.14 (s, 3H), 1.26-1.28 (d, 3H), 1.35-1.47 (m, 2H), 1.58 (s, 3H), 1.63-1.68 (dd, 1H), 1.73-1.80 (m, 1H), 1.89-1.98 (m, 1H), 2.26-2.42 (m, 2H), 2.46-2.52 (d, 1H), 2.98-3.03 (d, 1H), 6.10 (s, 1H), 7.12-7.18 (m, 2H), 7.41-7.47 (m, 3H).

2-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)propan-2-ol (13ba): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et_2O) to give a white solid in 55% yield. Flash column mobile phase - 20% EtOAc/hexanes. TLC R_f = 0.17 (25% EtOAc/hexanes). HRMS (ESI-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{22}\text{H}_{29}\text{N}_2\text{O}$, 337.2274; found 337.2200. $^1\text{H-NMR}$ (CDCl_3 , 499.73 MHz) δ 1.19 (s, 3H), 1.33-1.37 (s, 8H), 1.38-1.39 (d, 1H), 1.66-1.69 (d, 1H), 1.77-1.80 (dt, 1H), 1.84-1.87 (dt, 1H), 2.31 (m, 1H), 2.39 (s, 3H), 2.73-2.76 (d, 1H), 3.53-3.57 (d, 1H), 6.16 (s, 1H) 7.24-7.25 (d, 2H), 7.35-7.38 (d, 2H), 7.40 (s, 1H).

1-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-phenylethan-1-ol (13bb): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et_2O) to give a white solid in 82% yield. Flash column mobile phase - 15% EtOAc/hexanes. TLC R_f = 0.27 (30% EtOAc/hexanes). HRMS (ESI-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{27}\text{H}_{31}\text{N}_2\text{O}$, 399.2430; found 399.2368.); $^1\text{H-NMR}$ (CDCl_3 , 499.73 MHz) δ 1.27 (s, 3H), 1.62-1.66 (m, 2H), 1.68 (s, 3H), 1.74-1.79 (m, 2H), 2.04-2.07 (d, 1H), 2.24-2.28 (d, 1H), 2.33-2.40 (m, 4H), 2.60-2.63 (d, 1H), 3.10-3.13 (d, 1H), 6.14 (s, 1H), 7.23-7.29 (m, 4H), 7.33-7.36 (m, 4H), 7.49-7.52 (m, 2H).

1-(4-fluorophenyl)-1-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (13bc): From appropriate ketone **12** using commercially

available MeLi (1.9 M in Et₂O) to give a white solid in 80% yield. Flash column mobile phase - 15% EtOAc/hexanes. TLC R_f = 0.25 (30% EtOAc/hexanes). HRMS (DART-TOF) m/z : $[M+H]^+$ Calcd for C₂₇H₃₀FN₂O, 417.23367; found 417.23232. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.26 (s, 3H), 1.60-1.63 (m, 3H), 1.66 (s, 3H), 1.75-1.79 (d, 1H), 1.99-2.02 (d, 1H), 2.24-2.27 (d, 1H), 2.39 (s, 3H), 2.60-2.63 (d, 1H), 3.12-3.15 (d, 1H), 5.30 (s, 1H), 6.14 (s, 1H), 7.01-7.48 (m, 9H).

1-(3,4-difluoro-5-methoxyphenyl)-1-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (13bd): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et₂O) to give a white solid in 94% yield. Flash column mobile phase - 10% EtOAc/hexanes. TLC R_f = 0.19 (30% EtOAc/hexanes). HRMS (ESI-TOF) m/z : $[M+H]^+$ Calcd for C₂₈H₃₁F₂N₂O₂, 465.2348; found 465.2283. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.22-1.35 (m, 6H), 1.58-1.62 (m, 4H), 1.79-1.83 (d, 1H), 1.96-1.99 (d, 1H), 2.25-2.28 (d, 1H), 2.34-2.44 (s, 3H), 2.57-2.61 (d, 1H), 3.05-3.08 (d, 1H), 3.91-3.96 (s, 3H), 6.15 (s, 1H), 6.87-6.92 (m, 2H), 7.23-7.30 (m, 4H), 7.36 (s, 1H).

1-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-(thiophen-2-yl)ethan-1-ol (13be): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et₂O) to give a white solid in 76% yield. Flash column mobile phase - 10% EtOAc/hexanes. TLC R_f = 0.28 (25% EtOAc/hexanes). HRMS (ESI-TOF) m/z : $[M+H]^+$ Calcd for C₂₅H₂₉N₂OS, 405.1995; found 405.1899. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.25 (s, 3H), 1.40-1.41 (d, 1H), 1.57-1.59 (d, 1H), 1.65 (s, 3H), 1.69-1.70 (d, 2H), 1.76-1.78 (d, 1H), 2.03-2.06 (d, 1H), 2.26-2.28 (d, 1H), 2.39 (s, 3H), 2.66-2.69 (d, 1H), 3.22-3.25 (d, 1H), 6.15 (s, 1H), 7.12-7.37 (m, 8H).

2-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)propan-2-ol (13ca): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et₂O) to give a white solid in 58% yield. Flash column mobile phase - 25% EtOAc/hexanes. TLC R_f = 0.08 (30% EtOAc/hexanes). HRMS (ESI-TOF) m/z : $[M+H]^+$ Calcd for C₂₂H₂₉N₂O₂, 353.2223; found 353.2184. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.19 (s, 3H)

1.30-1.38 (d, 6H), 1.48-1.55 (m, 1H), 1.66-1.69 (d, 1H), 1.76-1.87 (dd, 2H), 2.27-2.37 (m, 3H), 2.73-2.76 (d, 1H), 3.53-3.57 (d, 1H), 3.85 (s, 3H), 6.11 (s, 1H) 6.95-7.41 (m, 5H).

1-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-phenylethan-1-ol (13cb): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et₂O) to give a beige solid in 100% yield. Flash column mobile phase - 10% EtOAc/hexanes. TLC R_f = 0.20 (30% EtOAc/hexanes). HRMS (DART-TOF) *m/z*: [M+H]⁺ Calcd for C₂₇H₃₁N₂O₂, 415.23800; found 415.23680. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.27 (s, 3H), 1.55-1.63 (m, 3H), 1.68 (s, 3H), 1.75-1.77 (m, 1H), 2.04-2.07 (m, 1H), 2.23-2.26 (m, 1H), 2.33-2.38 (m, 1H), 2.60-2.63 (d, 1H), 3.10-3.13 (d, 1H), 3.84 (s, 3H), 6.08 (s, 1H), 6.95-7.51 (m, 10H).

1-(4-fluorophenyl)-1-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (13cc): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et₂O) to give a white solid in 49% yield. Flash column mobile phase - 20% EtOAc/hexanes. TLC R_f = 0.28 (30% EtOAc/hexanes). HRMS (DART-TOF) *m/z*: [M+H]⁺ Calcd for C₂₇H₃₀FN₂O₂, 433.22858; found 433.22897. ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.25 (s, 3H), 1.53-1.62 (m, 2H), 1.67 (s, 3H), 1.73-1.78 (m, 1H), 1.99-2.02 (d, 1H), 2.24-2.36 (m, 3H), 2.60-2.63 (d, 1H), 3.12-3.15 (d, 1H), 3.85 (s, 3H), 6.09 (s, 1H), 6.95-7.49 (m, 9H).

1-(3,4-difluoro-5-methoxyphenyl)-1-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (13cd): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et₂O) to give a white solid in 65% yield. Flash column mobile phase - 15% EtOAc/hexanes. TLC R_f = 0.13 (30% EtOAc/hexanes). HRMS (DART-TOF) *m/z*: [M+H]⁺ Calcd for C₂₇H₃₀FN₂O₂, 433.22858; found 433.22897. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.26 (s, 3H), 1.60-1.67 (m, 6H) 1.79-1.82(d, 1H), 1.96-1.99 (d, 1H), 2.24-2.39 (m, 2H), 2.57-2.60 (d, 1H), 3.05-3.08 (d, 1H), 3.85 (s, 3H), 3.91 (s, 3H), 6.10 (s, 1H), 6.87-7.01 (m, 4H), 7.32-7.45 (m, 3H).

1-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-(thiophen-2-yl)ethan-1-ol (13ce): From appropriate ketone **12** using commercially available MeLi (1.9 M in Et₂O) to give a white solid in 48% yield. Flash column mobile phase - 15% EtOAc/hexanes. TLC R_f = 0.16 (30% EtOAc/hexanes). HRMS (ESI-TOF) *m/z*: [M+H]⁺ Calcd for C₂₅H₂₉N₂O₂S, 421.1944; found 421.1907. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.25 (s, 3H), 1.59-1.65 (m, 2H), 1.68 (s, 3H), 1.73-1.78 (m, 2H), 2.03-2.06 (d, 1H), 2.25-2.36 (m, 2H), 2.65-2.69 (d, 1H), 3.22-3.25 (d, 1H), 3.84 (s, 3H), 6.10 (s, 1H), 6.94-7.40 (m, 8H).

1.4 – Molecular modeling

1.4.1 – Structure activity relationships

Structure-activity relationships are considered when designing drug candidates for specific biological activities. It is well understood that with most drug therapies, there is a target receptor and/or tissue the drug is designed to interact with. Receptors in the biological and chemical sense are pockets within an individual protein or a large multi-protein complex. These receptors have a specific shape, and a complex and dynamic electrostatic surface that can be targeted for doing a number of biologically important functions. Some of these functions include up/down-regulating the production of transcription factors that alter cellular production of proteins and/or enzymes, enhancing/blocking the action of cellular machinery such as proton pumps or cation channels, or simply by altering the circulating levels of hormones or other signaling molecules. Modeling the interactions between a receptors and ligands with known biological response can help tailor the design of molecules with more precise biological activity.

To better understand the results from the data obtained from bioassays, ligand-receptor interactions were modeled between the synthesized molecules subjected to *ex-vivo* testing and the human glucocorticoid receptor (PDB ID: 5NFP).³⁵ Ligand models were prepared in Chem3D (PerkinElmer) to generate a structure in the lowest-energy

conformation to be used in the docking experiment. Docking was performed between the rigid 5NFP GR and the flexible ligands generated from Chem3D using AutoDock Vina (The Scripps Research Institute).³⁶ The resulting docking conformations were viewed in 3D with PyMol (Schrödinger)³⁷ and 2D ligand-residue interaction figures were generated with LigPlot+ (European Molecular Biology Laboratory)(**Figure 1.4.1; appendix 1.8.9**).³⁸ Four characteristic binding orientations were observed in most AutoDock Vina output files (**Figure 1.4.2**). Orientation within the pocket may be an important factor to contributing to the observed effects of these compounds. 2D and 3D representations of the modeled interactions as well as calculated binding affinities (AutoDock Vina) can be found in the appendix of chapter 1 (**section 1.8.7**). Ligand-residue interactions obtained from analysis using LigPlot+ were tabulated and further processed in R to create an interaction map (**Figure 1.4.3**). Most of the top binding orientations for the modeled compounds are in “orientation 1” with six exceptions: dexamethasone, **11ab**, **11ae**, **11af**, **11be**, and **13ca**.

1.4.2 – Modeling and visualization software

There are a number of programs available for free to visualize molecular structures for modeling purposes in 3D. AutoDock Tools is a visualization program from The Scripps Research Institute that allows you not only to view structures but also lets you manage the physical parameters of the structures to prepare them for docking experiments. These parameters - such as allowable bond rotations, defining an area of a protein docking will be performed within, or bonding capabilities of atoms – are set prior to a docking experiment that is performed with a related program, AutoDock Vina. A traditional docking experiment calculates the lowest energy orientation that a submitted ligand can exist as within a predefined volume of space in a protein crystal structure. The output of these calculations is a structure file with the (X, Y, Z) coordinates of each atom in Cartesian space relative to the spatial coordinates of the protein the ligand was modeled with. After calculations, these files can be visualized and manipulated in another

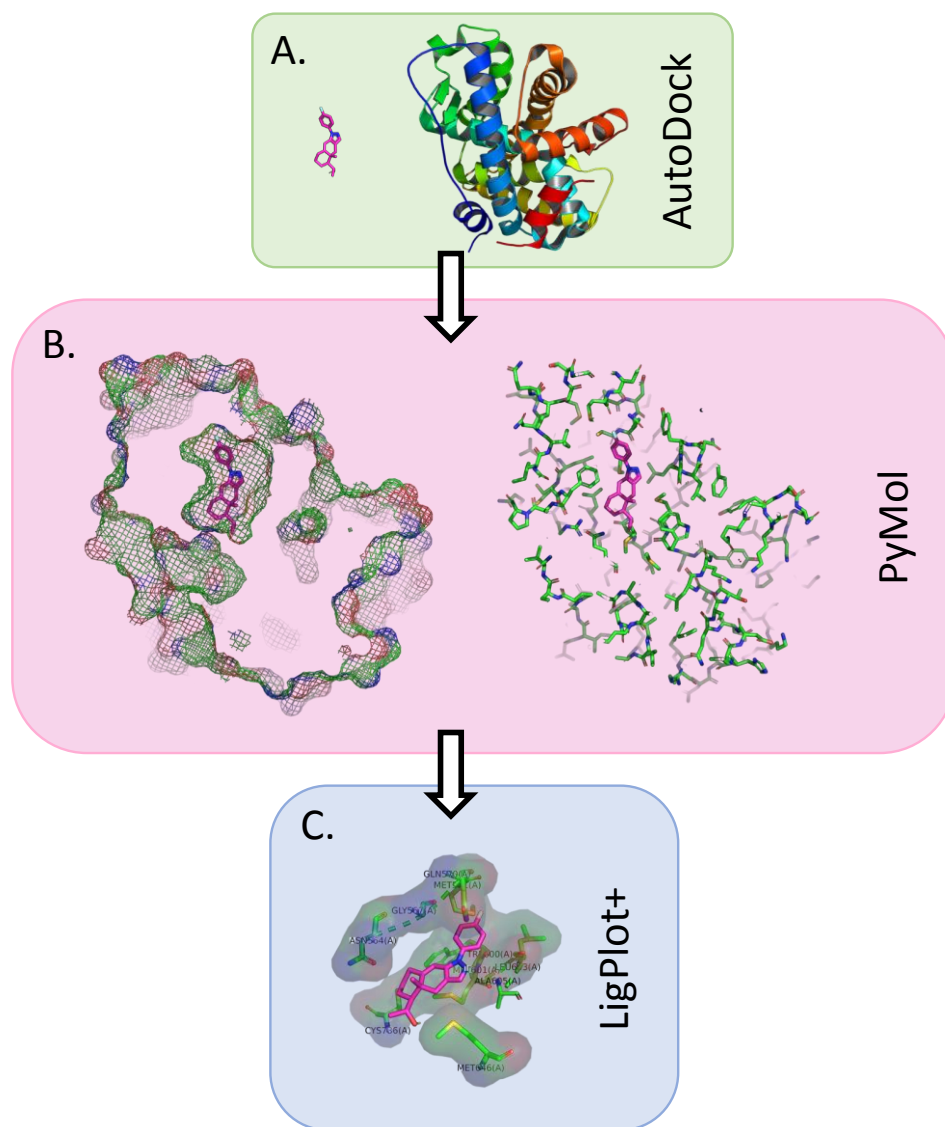


Figure 1.4.1 – Ligand-receptor docking

A truncated process map where first, the flexible ligand and target receptor are prepared in AutoDock for docking simulations (A). Once calculations are performed, the docking conformations can be viewed in PyMol in a number of ways (B). Secondary programs like LigPlot+ can then map out proximal interactions of the ligand within the receptor (C).

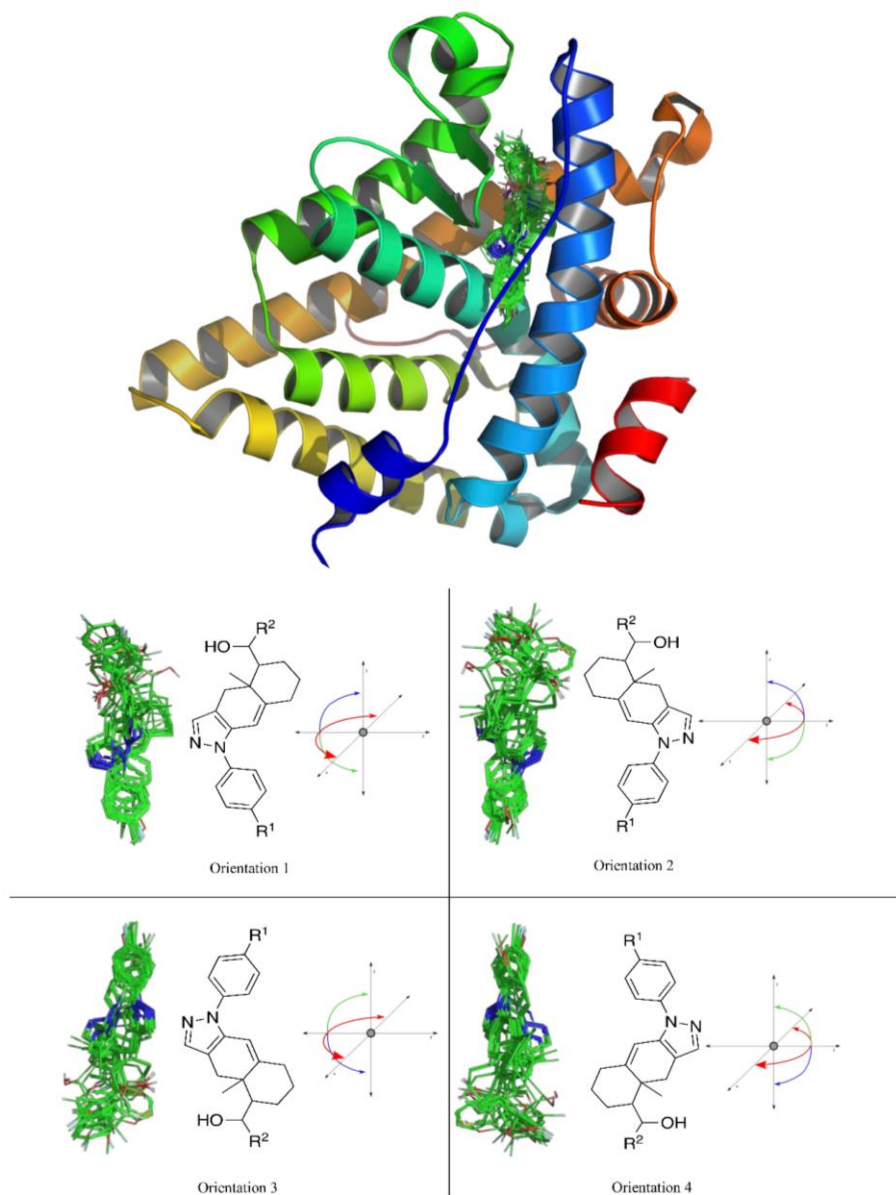


Figure 1.4.2 – Ligand orientation within 5NFP binding pocket

Autodock calculations revealed that there are four relative orientations for ligands within the GC receptor pocket, shown above. All molecules are shown stacked upon each other in the 5NFP binding pocket (top), as well as in each of the representative orientations (orientations 1,2,3, and 4).

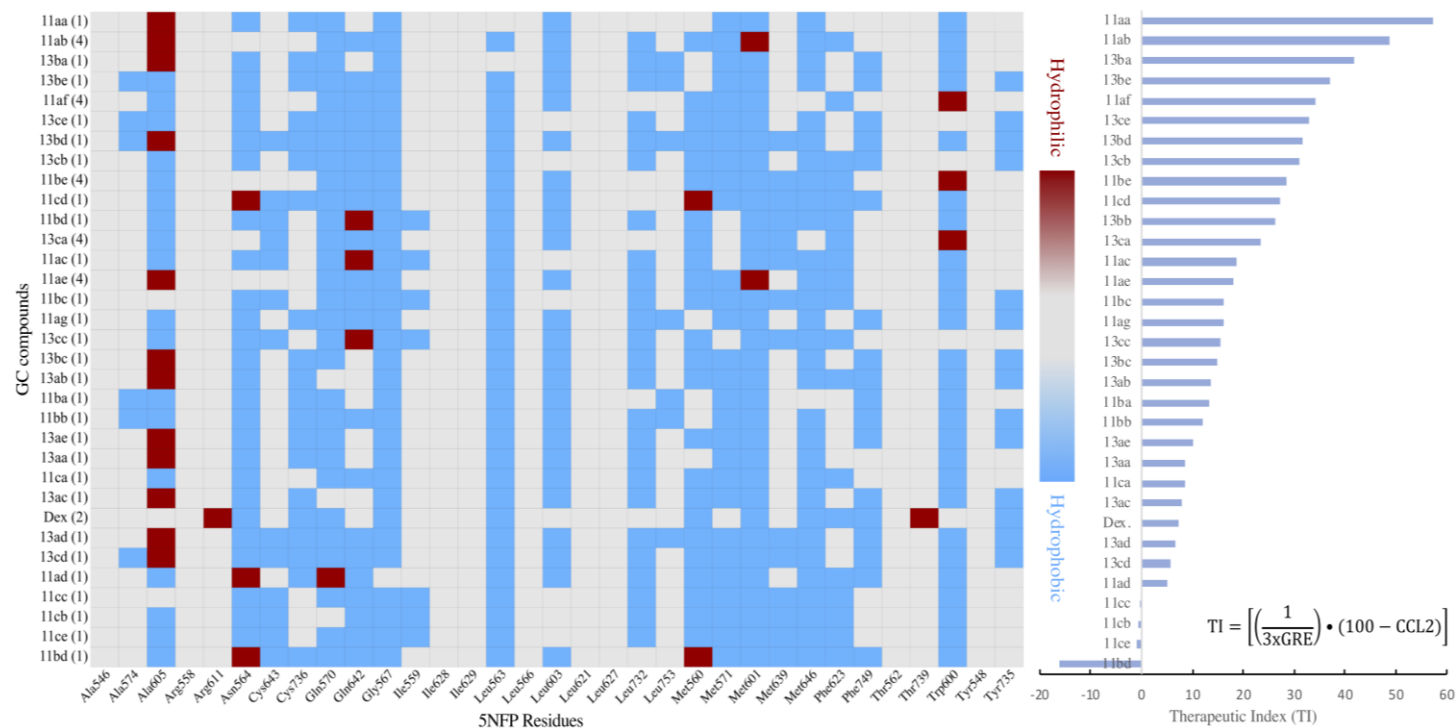


Figure 1.4.3 – Ligand-residue interaction map and therapeutic index scores

Ligand-residue interaction map (left) displaying hydrophilic (red) and hydrophobic (blue) interactions, as well as non-interacting residues (grey) with the respective compound in the human glucocorticoid receptor (5NFP) binding pocket. Compounds in interaction map are organized by therapeutic index (right). Residue interactions shown are of the ligands in their highest binding-affinity orientations (denoted by parenthetical numbers beside compound codes).

free software package called PyMol, which allows the user to label atoms, set color schemes, and view protein residues and/or tertiary structure.

With PyMol you can move the structures around to see what residues are close in space to the ligand and make educated assumptions about the most significant molecular interactions that are contributing to the simulated binding. There is another cost-free software for visualizing the interactions that are most significant called LigPlot+ out of the European Molecular Biology Laboratory that takes the output from AutoDock Vina and maps the interactions between ligand and protein in 2D, differentiating hydrophobic from hydrophilic interactions. This program also calculates donor-acceptor distances for the hydrophilic interactions, which is a nice feature. By using the combination of PyMol and LigPlot+, one can begin to make conclusions about the specific interactions that may be large contributors to the ligands observed bioactivity. The combined visualizations of ligand-residue interactions for all the tested compounds from this study can be found in chapter 1 appendix **section 1.8.6**.

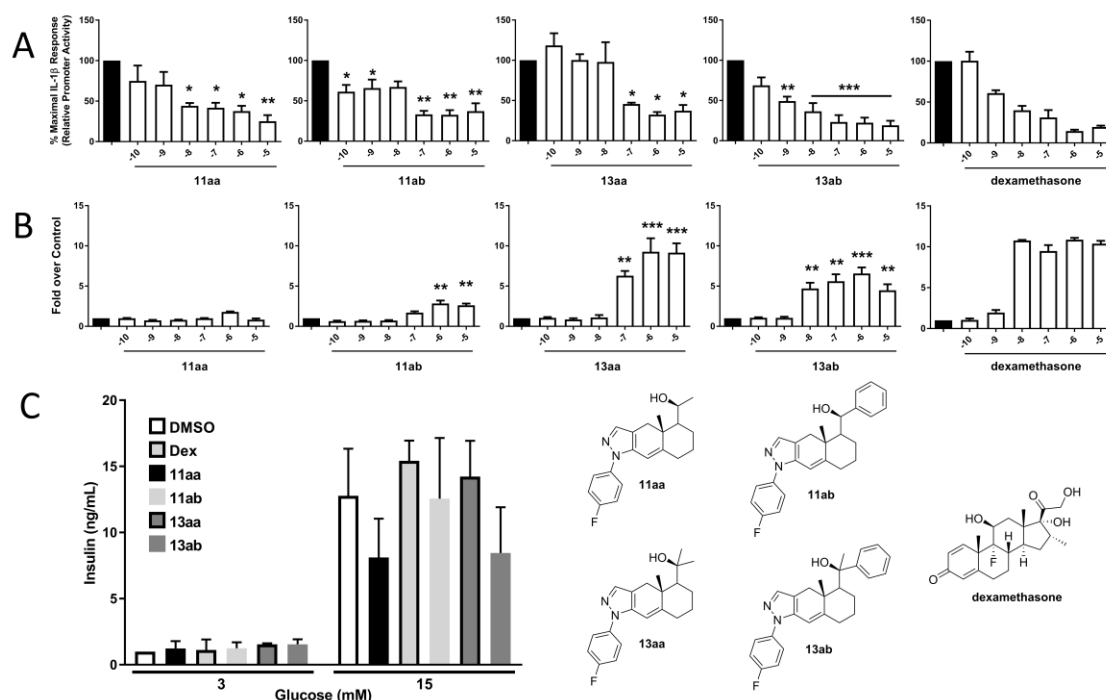
For density functional theory-based calculations (DFT) performed by the research group of Dr. Konstantinos Vogiatzis at the University of Tennessee, Knoxville; the following procedures were followed for ligand-receptor interaction calculations on a truncated receptor pocket: Calculation were performed on all synthetic GCs subjected to biological testing, which molecular models were prepared by the author. This was done by generating a minimized energy model of the ligands in Chem3D (Perkin Elmer) and converting this file to the appropriate .pdbqt file-type by AutoDock Tools. The receptor pocket was modeled by including all main group atoms within 6 Å of the co-crystallized ligand (budesonide) within the 5NFP structure obtained from PDB with solvent molecules removed. All valences were capped with appropriate atoms based on geometry of the included fragments. Calculations were performed with the TURBOMOLE 7.2 quantum chemical software package using DFT.³⁹ Constrained geometry optimizations were performed on docking output orientations by relaxing all hydrogen atoms and the full

glucocorticoid molecule within the pocket with atoms of the pocket frozen during convergence. The BP86 density functional, comprised of the Becke 1988 exchange and the Perdew 1986 correlation functionals, was employed along with Grimme's D3 dispersion correction with the Becke-Johnson damping function (D3(BJ)) and the def2-SVP basis set for all atoms.⁴⁰⁻⁴⁴ The resolution of identity (ri) approximation was used for rapid integral evaluation.⁴⁵ In addition to the optimized structures at the riBP86-D3(BJ)/def2-SVP level, single-point calculations were performed using the riB3LYP-D3(BJ)/def2-TZVPP level for higher accuracy.⁴⁶⁻⁴⁸ Interaction energies between the GCs and the pocket will be calculated by subtracting the energy of the empty pocket and the GC from the pocket-GC supersystem.

1.5 – *In-vitro* studies with synthetic glucocorticoids

1.5.1 – *Assessing therapeutic profile*

Inflammation can be assessed by a number of assays measuring the expression of inflammatory cytokines such as IL-1 β , NF- κ B, or AP-1. For this study, the %-maximal IL-1 β response is used to determine the extent of transrepression in the *ex-vivo* assays. The amount of IL-1 β transcribed was measured as a percentage compared to the IL-1 β response toward the drug vehicle, DMSO, which was quantified by a luminometer as in a typical luciferase assay. For the quantitative assessment of the production of this pro-inflammatory cytokine, modified cells with a luciferase-reporter inserted before the gene responsible for the production of chemokine C-C motif ligand 2 (CCL2)(**Figure 1.5.1A**). In this way, as IL-1 β was transcribed, so was an equivalent amount of luciferase. Upon treating the culture with luciferin, the resulting fluorescence is measured and is quantitatively representative of the production of IL-1 β . The second inflammatory assay (3xGRE) that was used in this study was another luciferase assay measuring transactivation, or the transcription of pro-inflammatory chemokines by activation of the glucocorticoid response element (GRE), which is partially responsible for the recruitment



1.5.1 – Truncated results from assays measuring inflammation and insulin secretion

Bioassay data for compounds exhibiting desired directionality of therapy. (A) CCL2-promoter-luciferase-promoter activity assay; x-axis: log concentration (molar); y-axis: % maximal IL-1 β response (relative promoter activity). (B) 3xGRE-promoter-luciferase-promoter activity assay; x-axis: log concentration (molar); y-axis: fold over control. (C) Glucose stimulated insulin secretion (GSIS). All assays were performed *ex vivo*, using 832/13 cells.

of immune cells, subsequent inflammation, and autoimmune-mediated damage to pancreatic β -cells. This assay was a means of measuring the amount of transactivation, which minimizing was a goal of this study with respect to the synthetic GCs employed (**Figure 1.5.1B**). Once the synthetic GCs were assayed for their therapeutic profile in terms of anti-inflammatory action, the top performing drug candidates were subjected to an assay measuring glucose-stimulated insulin secretion (GSIS). One of the primary goals of investigating these 1-(4-substituted phenyl)arylpyrazolyl GCs is to see how they compare to established GC therapeutics in terms of their influence on insulin production *in vivo*. Saturating concentrations (100 nM) of synthetic GCs were dissolved in DMSO and added to isolated human islets (pancreatic β -cells) to which known concentrations of glucose were added to stimulate the production of insulin. An enzyme-linked immunosorbent assay (ELISA) was performed on the GSIS samples to determine the concentration of insulin produced.

1.5.2 – Cell culture and assay parameters

Methods for the culture of 832/13 cells and measurements of insulin secretion have been described previously.⁴⁹ Six-point dose-response curves were used to generate EC50 and saturating concentrations of each steroid compound. For glucose-stimulated insulin secretion (GSIS), saturating concentrations (100 nM) of each commercial steroid or 1-(4-substituted phenyl)arylpyrazolyl based GC analogues were used (**Figure 1.5.1C**). Insulin secretion was measured with ELISA as previously described, using a kit from Mercodia (Uppsala, Sweden). Measurements for the 3xGRE luciferase assay were performed as described in the literature.⁷ Briefly, 832/13 cells were grown in 24 well plates to 50% confluence and then transiently transfected with 25 ng of indicated plasmid per well using TransFectin Lipid Reagent (Bio-Rad; Hercules, CA) following the manufacturer's directions. Cell lysis, luciferase assays, and normalization to total protein content were also carried out as described previously.⁷ Quantification of CCL2 secreted into the cell

media was carried out using a rat Quantikine ELISA kit from R&D Systems, Inc (Minneapolis, MN).

1.6 – Results

1.6.1 – *Inflammatory action of synthetic glucocorticoids*

Synthetic GC analogues were assayed for their ability to reduced inflammation through modulating transrepression and transactivation. Three different sets of GC analogues with respect to the installed headgroup (R^1) were tested for biological activity, and the 4-fluorophenyl series (compounds **11a**, **12a**, and **13a**) was observed to have the greatest dissociated profile. Levels of transrepression were measured by a CCL2-luciferase-reporter assay determining the %-maximal response to pro-inflammatory cytokine IL-1 β . Levels of transactivation was monitored through a 3xGRE-luciferase-reporter assay measuring the transcriptional activation toward the production of pro-inflammatory cytokines. The desired dissociated profile of the administered drug would reduce transrepression while also minimizing the transcription of cytokines known to be involved with an inflammatory response (transactivation). It was apparent from the CCL2 and 3xGRE assay data that the R^1 = 4-fluorophenyl series gave more desirable bioactivity than the R^1 = 4-methoxyphenyl or R^1 = 4-methylphenyl compounds. This may be a function of the bioavailability of the fluorinated analogues, which are known to be more permeable toward cell membranes and stable towards metabolism of the drug itself.⁵⁰

Compounds **11aa**, **11ab**, **13aa**, and **13ab** were found to have the desired reduction in IL-1 β response. Transrepression commenced at nM and μ M concentrations with ~70% reduction in relative promoter activity for the higher concentrations of fluorophenyl GCs (~1 μ M). Compounds with the greatest observed transrepression were those whose sidegroup (R^2) possesses a delocalized system of π -electrons (R^2 = phenyl, 4-fluorophenyl, 1,2-difluoroanisyl, and 3-thiophenyl). The observed activity is likely a function of hydrophobic interactions between GC and bound GR. GCs with R^2 moieties of smaller

(methyl-), and larger (cyanomethyl-) steric bulk that do not bear any delocalized π -electrons show a fairly pronounced reduction in desired activity, suggesting that size of the R^2 group has little influence on biological response with respect to transrepression – but rather that electronic interactions between GC and GR are more impactful.

The observed trend with transactivation was an increase in transcription of pro-inflammatory chemokine CCL2 with analogues bearing an R^2 moiety having delocalized system of π -electrons. Transactivation is observed at GC concentrations ranging from nM to μ M, where there is differential activity between molecule sets **11** and **13** having the same sidegroup. The additional steric bulk around the alcohol introduced by a methyl group in compound series **13** seems to enhance transrepression but also has a significant upregulating influence on transactivation, which makes series **13** less therapeutically valuable than the series **11** analogues. The intermediate ketone series **12aa-12ag** was tested for biological activity, but was found to enhance transactivation more than transrepression, indicating the alcohol group of series **11** and **13** analogues being important in influencing transrepressive activity of these molecules. It seems that coordination of the alcohol and/or an electron rich aromatic R^2 group in the GR binding pocket elicit a dissociated profile, unlike with the ketone intermediate; Therefore, intermediate series **12** (excluding analogues **12aa-12ag**) were not subjected to biological testing.

1.6.2 – Effect on insulin secretion in top-performing GC analogues

Glucose-stimulated insulin secretion was measured for the synthetic GCs that conferred the most dissociated therapeutic profile, meaning that the drug suppressed inflammatory response and did not also incite the production of immunostimulating chemokines. Compounds **11aa**, **11ab**, **13aa**, and **13bb** were chosen as the molecules with the most desirable combination of bioactivities and were subjected to an immunoassay to quantify the amount of insulin produced in response to a concentration of glucose, as indicated in the figure legend (**Figure 1.4B**). It was found that the 1-(4-substituted

phenyl)arylpyrazolyl based GC analogues had a range of effects on insulin production with **11aa** (R^2 = methyl) being a compound that did not hinder the production of insulin and compound **11ab** (R^2 = phenyl) , which had only a slight impact on insulin secretion.

There are not enough data points in terms of GSIS assays to make assumptions of what effect the identity of the R^2 moiety had on insulin production, but from what was performed, one could assume that larger and more electron-dense sidegroups may have a positively correlating factor in this assay. One other feature of the synthetic glucocorticoids that alludes to steric bulk around the R^2 moiety is that the compound **13** analogues **13aa** and **13ab** both had a significant impact on insulin secretion. This made compound **13** analogue therapeutically less valuable than their compound **11** series counterparts. From the data describing the ligand-residue interactions, there was no apparent residue interaction that was consistently present or not present in compounds **13** versus compounds **11**, suggesting that binding features were not the major drivers of the activity, but rather that the conformation of the GC-GR complex acting as a transcription factor played a more pivotal role. This will have to be investigated in the future in order to pin-point the underlying mechanisms associated with the downregulation of insulin production in response to these compounds.

1.6.3 – Structure-activity relationship of dissociated therapeutic profile

Considering the results from the CCL2, 3xGRE, and GSIS assays and comparing the results from the molecular docking experiment, there is not much to say about the simulated ligand binding scenarios. The binding profiles for all the compounds subjected to biological testing were all very similar, with only a few indications of hypothetically important residues involved in binding to have an influence on the observed activity. Hydrophilic interactions with residue Ala605 seemed to be a primary indicator of beneficial activity in conjunction with missing interactions toward Met639 and Cys643, but compounds with a much lower therapeutic index also featured this binding pattern. If anything should be taken from the ligand-residue interaction docking experiments, it's

that these molecules have similar binding features, but range in structure and electronic properties that may be affecting the structure of the GC-GR complex, which must have a powerful effect on DNA transcription specificity. Although this is something that could be modeled, it would require a computer with huge processing power to calculate the resulting GC-GR complex structure from a flexible ligand/receptor simulation and then compute interactions between the GC-GR complex against relevant DNA loci toward insulin function and inflammatory response. The data that is currently being collected from the higher-order DFT calculations on a truncated receptor pocket with the flexible ligands does point to a trend where orientation of the GC within the pocket leads to the desired therapeutic profile. There are four orientations that were considered where the GC was either flipped horizontally, vertically, and both horizontally and vertically with respect to the original orientation. Drug-residue interactions are currently being investigated by Justin K. Kirkland whom is a graduate research student in the computational chemistry research group of Dr. Konstatinos Vogiatzis at the University of Tennessee, Knoxville.

Without taking the docking experiments into consideration, the apparent effectors of biological activity of the synthetic GCs is their range of structural, chemical, and electronic properties that confer the observed bioactivities. As stated before, compounds with an R^2 group that can participate in π - π interactions or could act as a hydrogen bond-acceptor seemed to increase the incidence of transactivation, but also enhanced transrepression. From this, it could be stated that smaller more electronically incapable sidegroups such as R^2 = methyl would be moving toward a scaffold with the desired activity.

1.7 – Conclusion

Synthetic GCs were tested for their ability to reduce IL-1 β -induced inflammation (CCL2 assay - transrepression), and their propensity for enhancing transcription of a

synthetic GRE-containing promoter luciferase reporter gene (3xGRE assay - transactivation). Compounds providing the desired dissociated profile reduce inflammation with reduced transactivation compared to dexamethasone. Unfortunately for most of the compounds that exhibit transrepression, they also enhance transactivation presumably through dimerization of the GR-GC complex. It has been previously shown that dimerization of the GR-GC complex prior to binding to the nuclear GREs enhance the side effects of these drug therapies by enhancing the transcription of pro-inflammatory cytokines. At a SAR-level, this increase in transactivation seems to be influenced by GCs whose sidegroup can participate in π - π interactions with a bound GR. Only two of the compounds synthesized and screened for their therapeutic profile showed the desired activity (**Figure 1.5.1 A/B**) and were further assayed to determine their effect on insulin secretion. The GSIS assay for these compounds concluded that the 2'-OH compound **11** series did not hinder insulin secretion. A simple change in series **11** to the 3'-OH compound **13** series proved to begin impacting the function of β -cells and were thusly determined to be less therapeutically valuable in this study. CCL2 and 3xGRE assay data for all tested compounds can be found in appendix for chapter 1.

Both 1-(4-methoxyphenyl) and 1-(4-methylphenyl) headgroup analogues did not show a significant decrease in transrepression or an increase in transactivation, except for compounds whose R² moiety contains delocalized π -electrons. For those compounds, appreciable transrepression was observed at concentrations in the μ M-range, and an increase in transactivation was also observed at high GC concentrations ($\geq \mu$ M). This difference in activity compared to the 4-fluorophenyl headgroup series may be attributable to differential solubility between the analogues, being that fluorinated drugs are known for their increased lipid-solubility making them more easily transported throughout the cell.⁵¹ Data from the inflammatory assays portend that compared to the R¹ = 1-(4-methoxyphenyl)- / 1-(4-methylphenyl)-GCs, the R¹ = 1-(4-fluorophenyl) series would appear to have a greater affinity for the hGR, or may just be more easily transported

throughout the cell due to solubility properties compared to the other GC analogues. Binding affinity calculations (**Chapter I Appendix 1.8.7**) for all tested compounds reveal similar affinities to the hGR (5NFP) with an average binding affinity of -9.5 ± 1.75 kcal/mol. These calculations are not a very precise view of the true binding affinities for the compounds within the GR binding pocket, as the calculations were performed on a rigid receptor model. Ongoing work is being finished for the final publication of this study that features DFT calculations of a dynamic truncated GR-pocket and the interactions of the synthetic GCs within this pocket. As previously stated, orientation seems to be an important feature in the bioactivity of these compounds. This has been corroborated by preliminary data obtained from these DFT calculations as three of four possible orientations within the pocket had a positive correlation with the observed therapeutic index of these compounds as a function of increasing binding affinity. The remaining orientation displays a positive correlation of therapeutic index as a function of decreasing binding affinity. So future work with designing compounds with the desired dissociated therapeutic profile will have to take these orientation considerations into account for targeting specific drug-residue interactions that dissuade the negative effects that GCs can have.

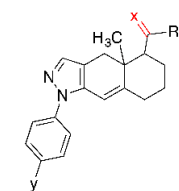
An increase in transactivation is an undesired effect of these drug candidates, and this activity is enhanced in compounds containing an aromatic sidegroup (R^2); Although, the modeled binding patterns within the hGR between the R^2 analogues are similar, the aromatic sidegroups may generate strong π - π interactions not accounted for with a non-flexible receptor model, which may be an important factor in enhancing transactivation. These aromatic sidegroups may be decreasing the binding affinity for the compounds in the correct orientation, or rather, they may be increasing the affinity for binding the receptor pocket in the orientation that decreases the therapeutic index score as a function increasing binding affinity. There is also an observable trend in which series **13** compounds have a higher average affinity for the hGR than series **11**. This increase in binding affinity

may be strongly associated with the transactivating orientation, but these observations have not yet been made with the DFT studies. Compounds **11** may have more freedom in the ways they may bind in the receptor pocket due to a greater degree of accessibility, which may explain why compounds **12** and **13** exhibit stronger transactivating activity when compared to 2°OH series **11**. The differential activity observed between compounds bearing different R¹ headgroups as suggested before may be attributable to differences in bioavailability, but when considering the binding features in the R¹ = 1-(4-fluorophenyl) series, there is a hydrophilic interaction with residue Ala605 in nearly all instances, especially among the molecules conferring the desired therapeutic profile. This may suggest some of the observed differences in biological activity for the R¹ = 1-(4-fluorophenyl) series as being a function of bioavailability, or more likely the resultant shape of the GR-GC complex specific to each GC compound. Determining the precise structure of these complexes would either require crystallizing each GR-GC complex and determining a crystal structure to be compared with DNA binding domains, or by powerful computational calculations, both of which would be exhausting bodies of work – especially the former.

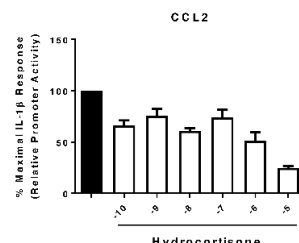
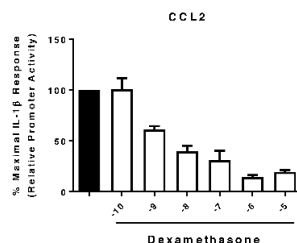
The work presented here on synthetic GCs with the desired anti-inflammatory profile without disturbing insulin production shows promise for future work in tailoring a structure with optimal activity. This study focuses on the structure-activity relationship of these synthetic compounds on inflammation with the added observation that of the compounds assayed for their effects on insulin secretion show little to no ability to diminish β -cell function. Compounds **11aa** and **11ab** were found to be most efficacious in this regard. Their *ex vivo* efficacy has been demonstrated, but their efficacy *in vivo* will need to be determined to validate these findings.

1.8 – Chapter I appendix

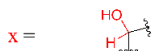
1.8.1 – CCL2 Inflammatory assay data



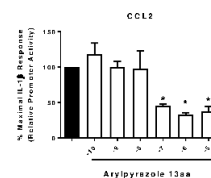
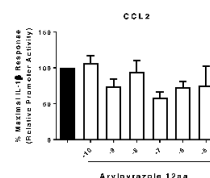
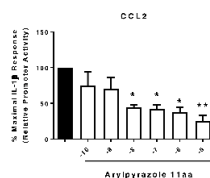
General Structure of arylpyrazole-based GC scaffold



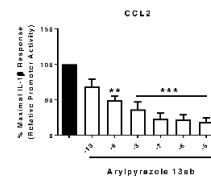
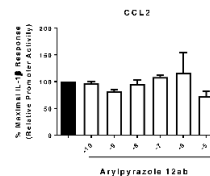
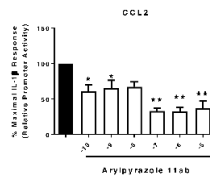
GC Analog



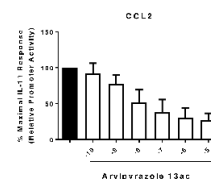
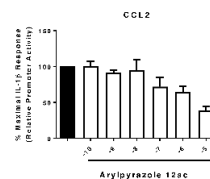
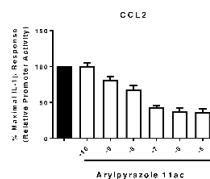
y = fluoro
R = methyl



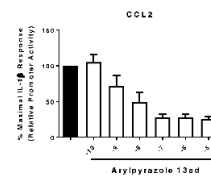
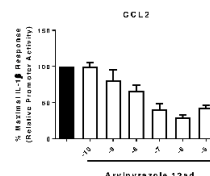
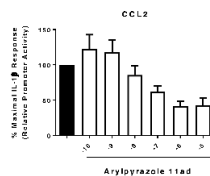
y = fluoro
R = phenyl



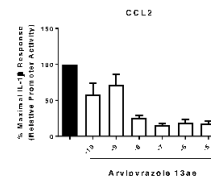
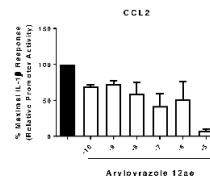
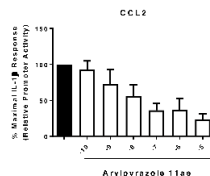
y = fluoro
R = 4-fluorophenyl



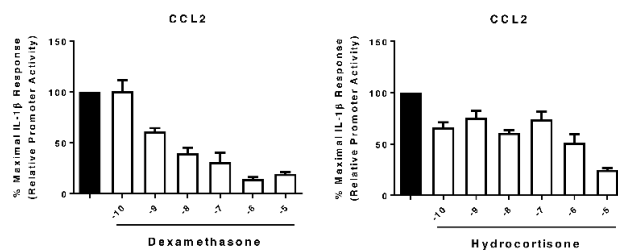
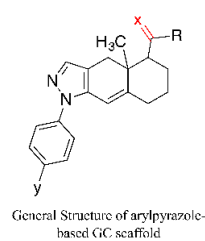
y = fluoro
R = 2,3-difluoroanisoyl



y = fluoro
R = 3-thiophenyl



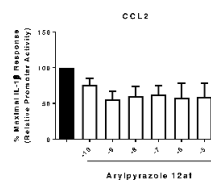
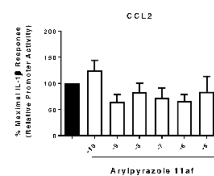
1.8.1 – CCL2 inflammatory assay data (continued)



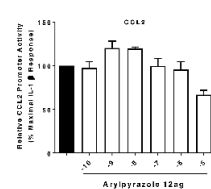
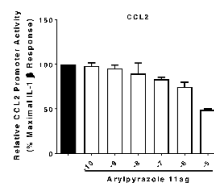
GC Analog



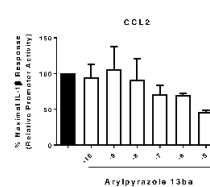
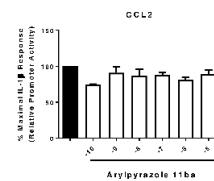
y = fluoro
R = cyanomethyl



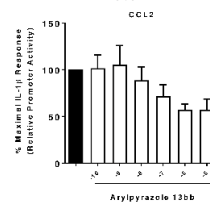
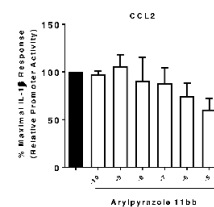
y = fluoro
R = t-butyl



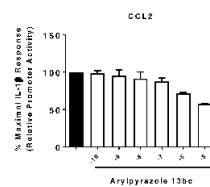
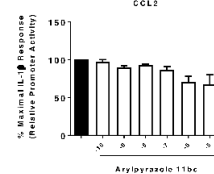
y = methyl
R = methyl



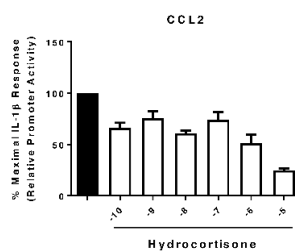
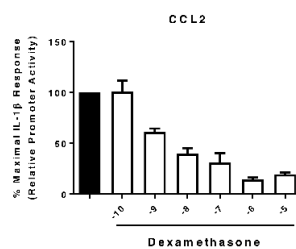
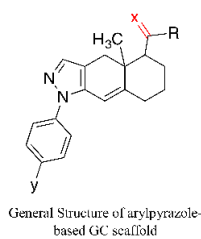
y = methyl
R = phenyl



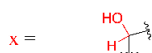
y = methyl
R = 4-fluorophenyl



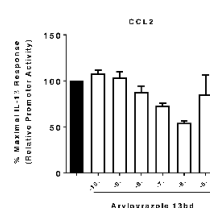
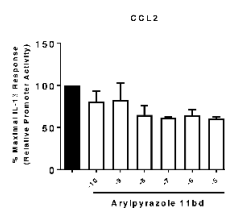
1.8.1 – CCL2 inflammatory assay data (continued)



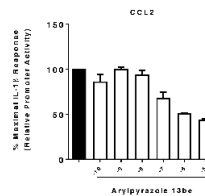
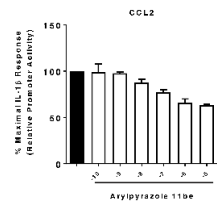
GC Analog



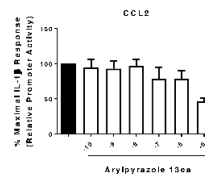
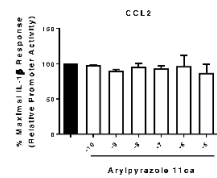
y = methyl
R = 2,3-difluoroanisoyl



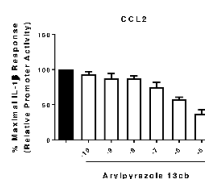
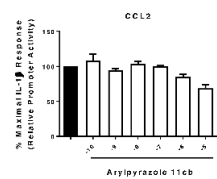
y = methyl
R = 3-thiophenyl



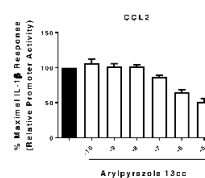
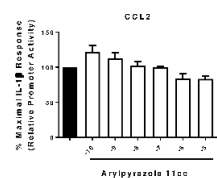
y = methoxy
R = methyl



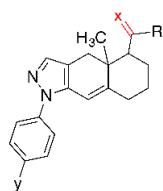
y = methoxy
R = phenyl



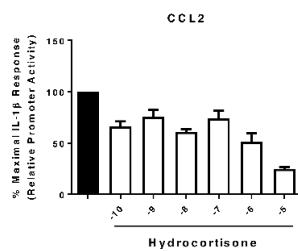
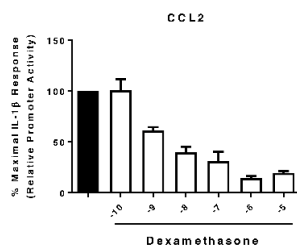
y = methoxy
R = 4-fluorophenyl



1.8.1 – CCL2 inflammatory assay data (continued)



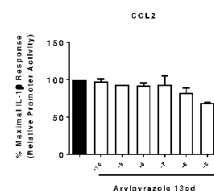
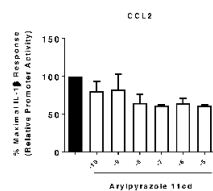
General Structure of arylpyrazole-based GC scaffold



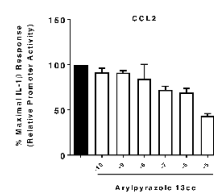
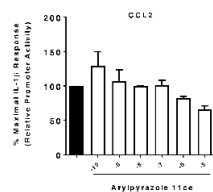
GC Analog



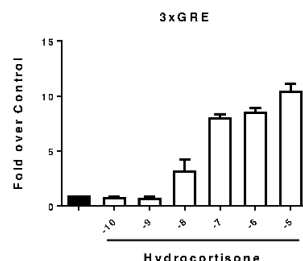
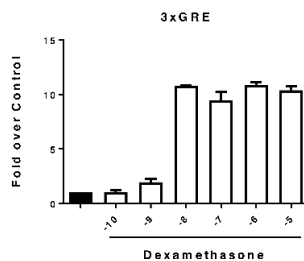
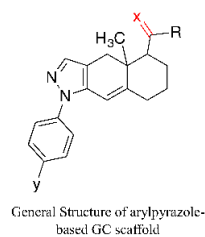
y = methoxy
R = 2,3-difluoroanisoyl



y = methoxy
R = 3-thiophenyl



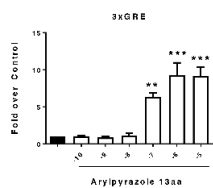
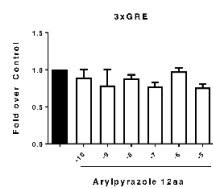
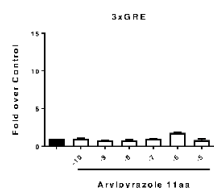
1.8.2 – 3xGRE Transcription assay data



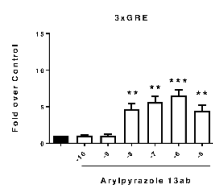
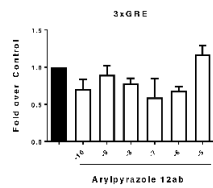
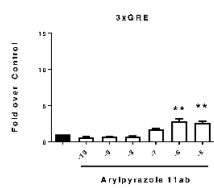
GC Analog



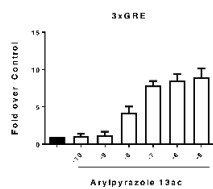
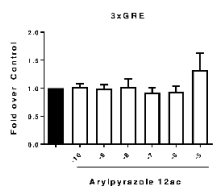
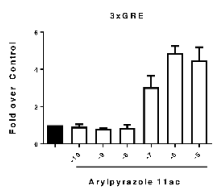
y = fluoro
R = methyl



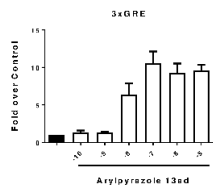
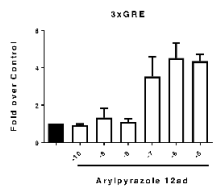
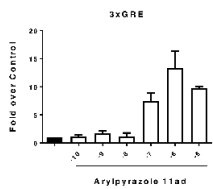
y = fluoro
R = phenyl



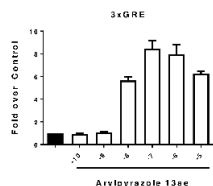
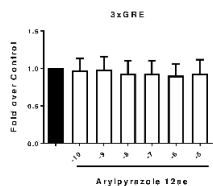
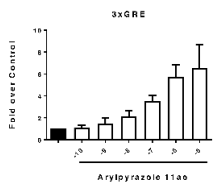
y = fluoro
R = 4-fluorophenyl



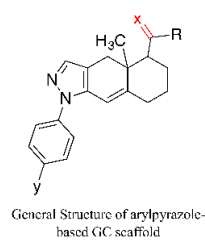
y = fluoro
R = 2,3-difluoroanisyl



y = fluoro
R = 3-thiophenyl



1.8.2 – 3xGRE transcription assay data (continued)

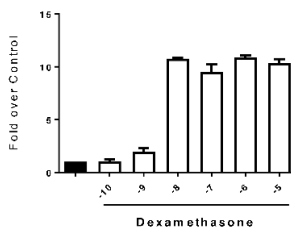


GC Analog

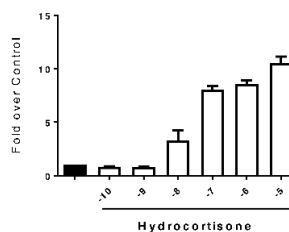
X =



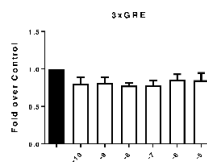
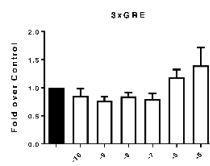
3xGRE



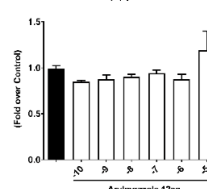
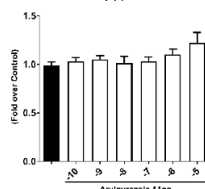
3xGRE



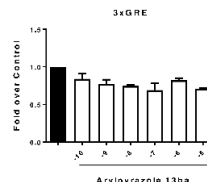
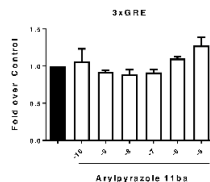
y = fluoro
R = cyanomethyl



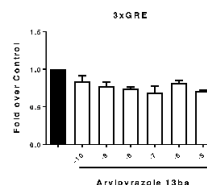
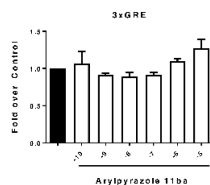
y = fluoro
R = t-butyl



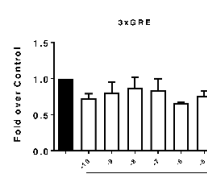
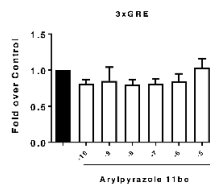
y = methyl
R = methyl



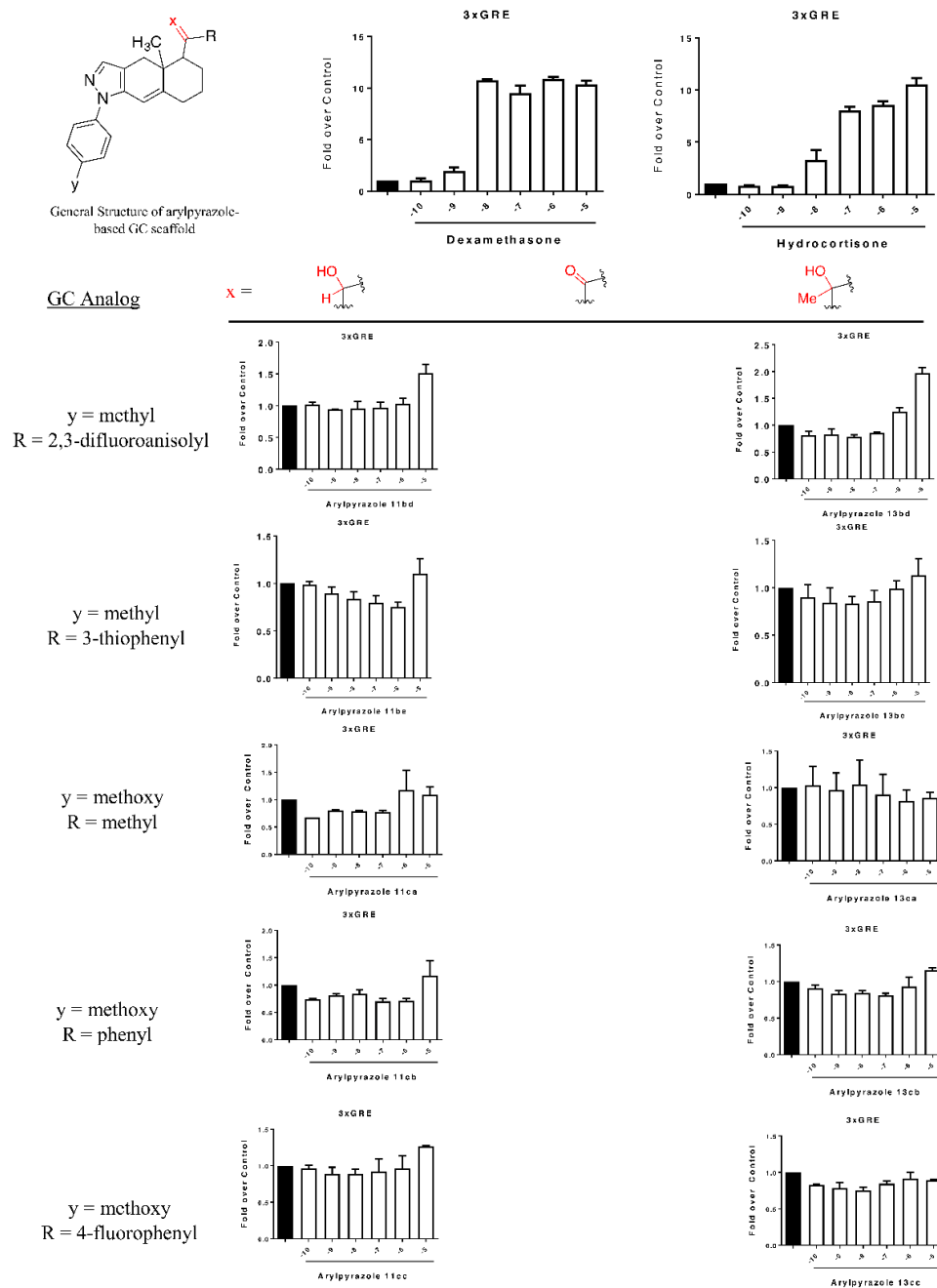
y = methyl
R = phenyl



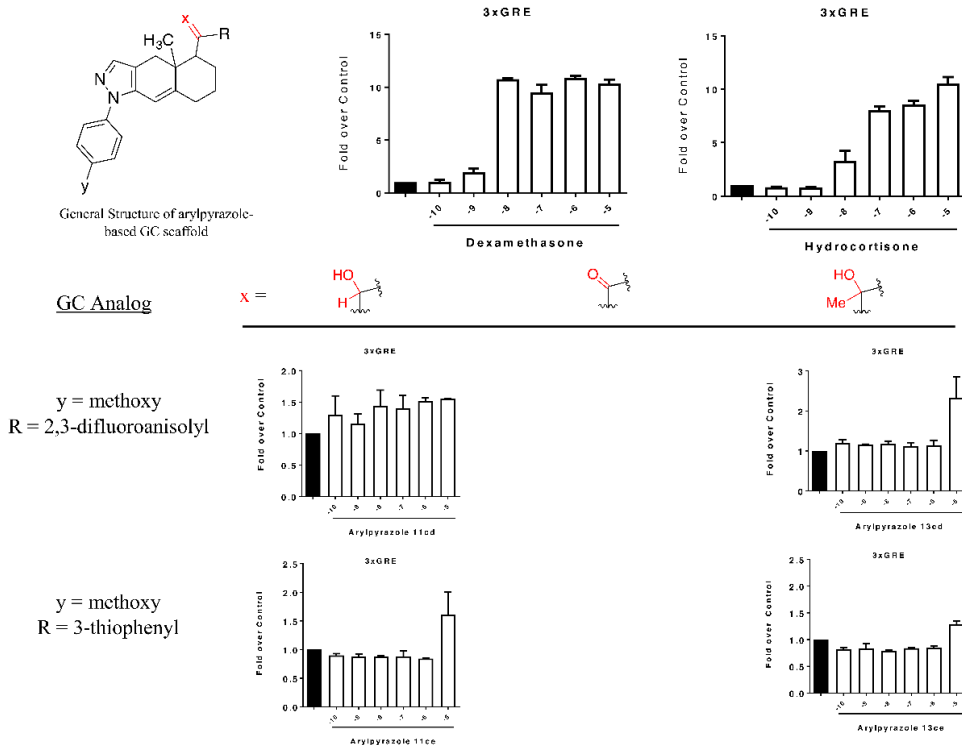
y = methyl
R = 4-fluorophenyl



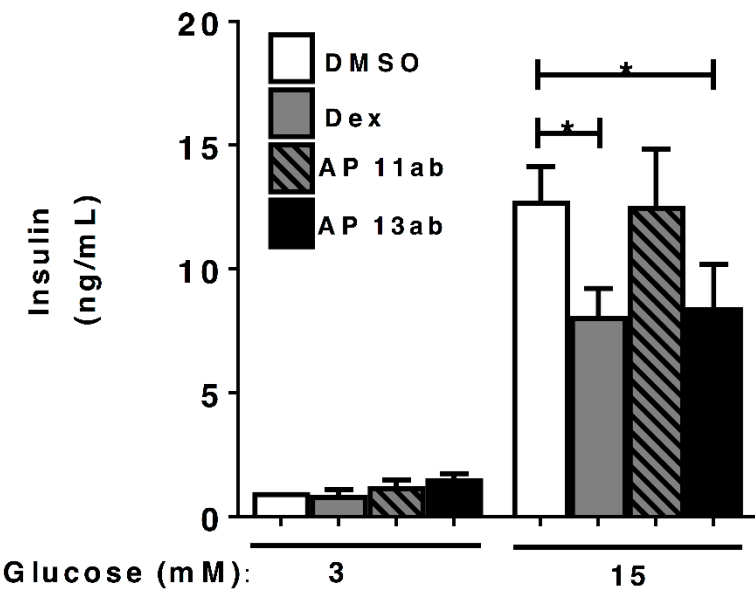
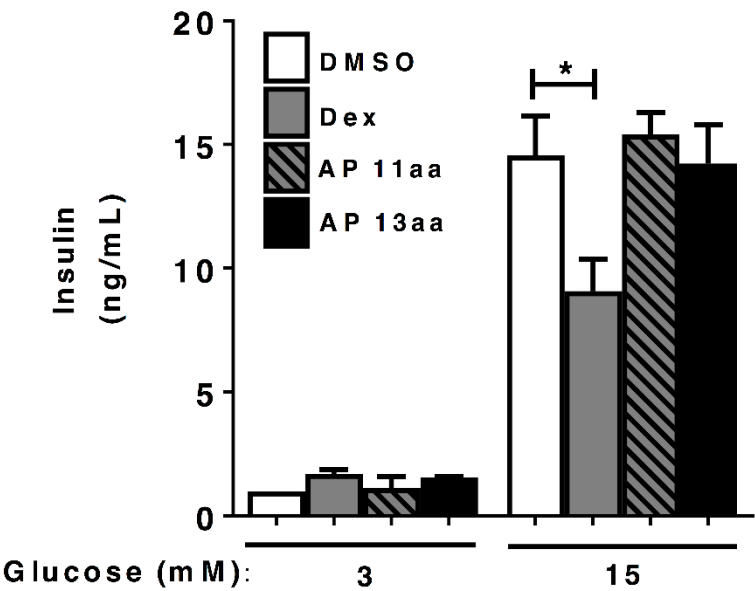
1.8.2 – 3xGRE transcription assay data (continued)



1.8.2 – 3xGRE transcription assay data (continued)



1.8.3 – Glucose-stimulated insulin secretion (GSIS) assay data



1.8.4 – Full characterization data of novel and previously reported compounds

4-Fluorophenyl analogues

1-(4-fluorophenyl)-4a-methyl-1,4,4a,6,7,8-hexahydrospiro[benzo[f]indazole-5,2'-[1,3]dioxolane] (7a): ¹H-NMR (CDCl₃, 500 MHz) δ 1.21 (s, 3H), 1.65-1.82 (m, 4H), 2.27-2.30 (d, 1H), 2.42-2.52 (m, 2H), 3.14-3.17 (d, 1H), 3.97-4.10 (m, 4H), 6.20 (s, 1H), 7.14 (m, 2H), 7.41-7.44 (m, 3H).

1-(4-fluorophenyl)-4a-methyl-1,4,4a,6,7,8-hexahydro-5H-benzo[f]indazol-5-one (8a): ¹H-NMR (CDCl₃, 500 MHz) δ 1.25 (s, 3H), 1.63-1.72 (m, 1H), 2.06-2.12 (m, 1H), 2.51-2.57 (m, 2H), 2.60-2.69 (m, 2H), 2.87-2.91 (s, 2H), 6.29 (s, 1H), 7.15-7.18 (t, 2H), 7.44-7.47 (m, 2H), 7.48 (s, 1H).

1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazole-5-carbaldehyde (10a): ¹H-NMR (CDCl₃, 500 MHz) δ 1.25 (s, 3H), 1.35-1.45 (m, 2H), 1.68-1.77 (m, 1H), 1.89-1.95 (m, 2H), 2.32-2.42 (m, 3H), 2.87-2.90 (d, 1H), 3.07-3.10 (d, 1H), 6.18 (s, 1H), 7.12-7.16 (t, 2H), 7.41-7.44 (m, 3H), 9.93 (s, 1H).

1-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (11aa): ¹H-NMR (CDCl₃, 500 MHz) δ 1.14 (s, 3H), 1.26-1.28 (d, 3H), 1.41-.147 (dd, 2H), 1.63-1.68 (dd, 1H), 1.73-1.80 (m, 1H), 1.90-1.97 (m, 1H), 2.31-2.43 (m, 2H), 2.46-2.52 (d, 1H), 2.98-3.03 (d, 1H), 4.25 (s, 1H), 6.10 (s, 1H), 7.11-7.19 (m, 2H), 7.41-7.47 (m, 3H).

2-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)propan-2-ol (13aa): ¹H-NMR (CDCl₃, 300 MHz) δ 1.14 (s, 3H), 1.26-1.28 (d, 3H), 1.35-1.47 (m, 2H), 1.58 (s, 3H), 1.63-1.68 (dd, 1H), 1.73-1.80 (m, 1H), 1.89-1.98 (m, 1H), 2.26-

2.42 (m, 2H), 2.46-2.52 (d, 1H), 2.98-3.03 (d, 1H), 6.10 (s, 1H), 7.12-7.18 (m, 2H), 7.41-7.47 (m, 3H).

(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(phenyl)methanol (**11ab**): ¹H-NMR (CDCl₃, 500 MHz) δ 1.18 (s, 3H), 1.52-1.58 (m, 2H), 1.60-1.63 (dd, 1H), 1.70-1.74 (m, 2H), 2.17-2.21 (m, 1H), 2.28-2.36 (m, 1H), 2.66-2.69 (d, 1H), 3.08-3.11 (d, 1H), 5.10 (d, 1H), 6.02 (s, 1H), 7.04-7.08 (m, 2H), 7.16-7.18 (m, 1H), 7.27-7.29 (m, 4H), 7.35-7.38 (m, 3H).

1-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-phenylethan-1-ol (**13ab**): ¹H-NMR (CDCl₃, 500 MHz) δ 1.27 (s, 3H), 1.56-1.65 (m, 2H), 1.67 (s, 3H), 1.74-1.80 (m, 1H), 2.02-2.08 (dd, 1H), 2.23-2.44 (m, 2H), 2.58-2.64 (d, 1H), 3.12-3.17 (d, 1H), 6.09 (s, 1H), 7.10-7.16 (t, 2H), 7.27-7.52 (m, 8H).

(4-fluorophenyl)(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)methanol (**11ac**): ¹H-NMR (CDCl₃, 300 MHz) δ 1.26 (s, 3H), 1.65-1.69 (m, 3H), 1.77-1.86 (m, 1H), 2.26-2.41 (m, 2H), 2.72-2.77 (d, 1H), 3.17-3.22 (d, 1H), 5.19-5.22 (s, 1H), 6.12 (s, 1H), 7.02-7.08 (t, 2H), 7.13-7.19 (t, 2H), 7.30-7.34 (dd, 2H), 7.44-7.48 (m, 3H).

1-(4-fluorophenyl)-1-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (**13ac**): ¹H-NMR (CDCl₃, 500 MHz) δ 1.25 (s, 3H), 1.51-1.62 (m, 2H), 1.66 (s, 3H), 1.72-1.81 (m, 2H), 1.95-2.05 (m, 1H), 2.21-2.40 (m, 2H), 2.58-2.64 (d, 1H), 3.14-3.20 (d, 1H), 6.09 (s, 1H), 7.00-7.06 (t, 2H), 7.11-7.17 (t, 2H), 7.25-7.28 (m, 1H), 7.41-7.49 (m, 4H).

(3,4-difluoro-5-methoxyphenyl)(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)methanol (11ad): ¹H-NMR (CDCl₃, 300 MHz) δ 1.26 (s, 3H), 1.54-1.57 (m, 1H), 1.64-1.72 (m, 1H), 1.82-1.86 (m, 2H), 2.28-2.46 (m, 3H), 2.72-2.77 (d, 1H), 3.17-3.22 (d, 1H), 3.92 (s, 3H), 5.12 (s, 1H), 6.12 (s, 1H), 6.72-6.80 (m, 2H), 7.15-7.21 (m, 2H), 7.45-7.54 (m, 3H).

1-(3,4-difluoro-5-methoxyphenyl)-1-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (13ad): ¹H-NMR (CDCl₃, 500 MHz) δ 1.26 (s, 3H), 1.54-1.60 (m, 2H), 1.62 (s, 3H), 1.64-1.66 (m, 1H), 1.80-1.85 (m, 1H), 1.96-1.99 (dd, 1H), 2.24-2.30 (m, 1H), 2.34-2.42 (m, 1H), 2.57-2.60 (d, 1H), 3.07-3.11 (d, 1H), 3.91 (s, 3H), 6.10 (s, 1H), 6.87-6.91 (m, 2H), 7.12-7.16 (t, 2H), 7.28 (s, 1H), 7.42-7.45 (m, 2H).

(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(thiophen-3-yl)methanol (11ae): ¹H-NMR (CDCl₃, 500 MHz) δ 1.24 (s, 3H), 1.65-1.93 (m, 5H), 2.27-2.48 (m, 2H), 2.67-2.72 (d, 1H), 3.08-3.15 (d, 1H), 5.24 (s, 1H), 6.12 (s, 1H), 6.94-6.99 (m, 1H), 7.12-7.24 (m, 3H), 7.30-7.33 (m, 1H), 7.43-7.48 (m, 3H).

1-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-(thiophen-3-yl)ethan-1-ol (13ae): ¹H-NMR (CDCl₃, 500 MHz) δ 1.27 (s, 3H), 1.52-1.62 (m, 3H), 1.68 (s, 3H), 2.01-2.06 (m, 1H), 2.24-2.43 (m, 3H), 2.64-2.69 (d, 1H), 3.24-3.30 (d, 1H), 6.10 (s, 1H), 7.14-7.17 (m, 3H), 7.27-7.30 (m, 1H), 7.33 (s, 1H), 7.42-7.47 (m, 3H).

3-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-3-hydroxypropanenitrile (11af): white solid obtained in 46% yield; mp 121-123.8 °C; TLC (Silica G w/UV254) 25% EtOAc in hexanes, R_f = 0.14; [α]_D²⁰ = +60 (c 1.0, MeOH); ¹H-NMR (CDCl₃, 500 MHz) δ 1.04-1.13 (s, 3H), 1.25-1.46 (m, 2H), 1.61-1.71 (m, 2H), 1.78-1.97 (m, 2H), 2.30-2.42 (m, 2H), 2.47-2.72 (m, 3H), 2.93-3.21 (d, 1H), 4.16-4.40 (m, 1H), 6.12 (s,

1H), 7.14-7.44 (m, 5H); ¹³C-NMR (CDCl₃, 500 MHz) δ 19.32, 20.91, 25.37, 26.20, 32.98, 33.19, 34.64, 52.82, 66.62, 109.76, 115.96, 116.14, 125.32, 137.86; IR (NaCl, thin film) ν_{max} (cm⁻¹): 3335 (OH), 3030 (w), 2934, 2867, 2249 (CN), 1619, 1518, 1225, 840; HRMS (DART-RTOF) *m/z*: [M+H]⁺ Calcd for C₂₁H₂₃FN₃O, 352.18197; found 352.18293.

3-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-3-hydroxybutanenitrile (**12af**): white solid obtained in 16% yield; mp 192-194 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.11; [α]_D²⁰ = +60 (c 1.0, CHCl₃); ¹H-NMR (CDCl₃, 500 MHz) δ 1.20 (s, 3H), 1.45-1.49 (m, 2H), 1.82-1.87 (t, 2H), 1.93-1.97 (d, 1H), 2.32-2.44 (m, 3H), 2.74-2.77 (d, 1H), 2.82-2.88 (m, 2H), 3.53-3.61 (s, 2H), 6.15 (s, 1H), 7.14-7.44 (m, 5H); ¹³C-NMR (CDCl₃, 500 MHz) δ 18.53, 25.21, 25.61, 31.72, 33.85, 34.75, 41.16, 59.60, 110.18, 112.91, 113.49, 116.00, 135.61, 136.29, 137.72, 146.60, 160.56, 162.53, 199.38; IR (NaCl, thin film) ν_{max} (cm⁻¹) 2955, 2920, 2850, 1723 (m), 1514 (m), 1462, 1377, 1222; HRMS (DART-RTOF) *m/z*: [M+H]⁺ Calcd for C₂₁H₂₁FN₃O, 350.16632; found 350.16529.

1-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-2,2-dimethylpropan-1-ol (**11ag**): ¹H-NMR (CDCl₃, 500 MHz) δ 0.97 (s, 9H), 1.08 (s, 3H), 1.38-1.45 (m, 1H), 1.59-1.67 (m, 1H), 1.72-1.79 (m, 2H), 1.83-1.87 (m, 1H), 2.30-2.35 (m, 1H), 2.36-2.44 (m, 1H), 2.45-2.48 (d, 1H), 2.82-2.85 (d, 1H), 3.50 (s, 1H), 6.13 (s, 1H), 7.13-7.17 (t, 2H), 7.41 (s, 1H), 7.43-7.47 (m, 2H).

2-(1-(4-fluorophenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-3,3-dimethylbutan-2-ol (**12ag**): beige solid obtained in 14% yield; mp 192-194 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.54; [α]_D²⁰ = +8.7 (c 1.0, MeOH); ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.19 (s, 9H), 1.33 (s, 3H), 1.40-1.50 (m, 1H), 1.55 (s, 1H), 1.70-1.74 (dd, 1H), 2.27-2.33 (m, 1H), 2.42-2.53 (m, 1H), 2.56-2.68 (m, 2H), 3.08-3.13 (m, 1H), 6.11 (s, 1H), 7.12-7.18 (t, 2H), 7.36 (s, 1H), 7.42-7.46 (dd, 2H); ¹³C-NMR (CDCl₃, 125.66 MHz) δ 19.22,

26.03, 27.02, 27.50, 31.99, 34.63, 41.82, 45.42, 53.04, 109.06, 113.60, 115.92, 119.70, 125.32, 135.81, 136.61, 137.77, 149.32, 160.46, 162.43, 218.05; IR (NaCl, thin film) ν_{max} (cm^{-1}) 2927, 2855, 1699, 1516, 1479, 1464, 1375, 1366, 1292, 1224, 1176, 1152, 1134, 1000; HRMS (DART-RTOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{23}\text{H}_{28}\text{FN}_2\text{O}$, 367.2180; found 367.2179.

4-methylphenyl analogues

4a-methyl-1-(p-tolyl)-1,4,4a,6,7,8-hexahydrospiro[benzo[f]indazole-5,2'-[1,3]dioxolane]

(7b): $^1\text{H-NMR}$ (CDCl_3 , 300 MHz) δ 1.14 (s, 3H), 1.56-1.79 (m, 5H), 2.19-2.22 (m, 1H), 2.32 (s, 3H), 2.42-2.45 (d, 1H), 3.08-3.11 (d, 1H), 3.92-4.01 (m, 4H), 6.18 (s, 1H), 7.17-7.19 (m, 2H), 7.27-7.29 (m, 2H), 7.33 (s, 1H).

4a-methyl-1-(p-tolyl)-1,4,4a,6,7,8-hexahydro-5H-benzo[f]indazol-5-one **(8b)**: $^1\text{H-NMR}$ (CDCl_3 , 300 MHz) δ 1.24 (s, 3H), 1.59-1.79 (m, 2H), 2.03-2.13 (m, 1H), 2.40 (s, 3H), 2.49-2.70 (m, 3H), 2.90 (s, 2H), 6.32 (s, 1H), 7.26-7.28 (m, 2H), 7.35-7.38 (m, 2H), 7.47 (s, 1H).

4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazole-5-carbaldehyde **(10b)**: $^1\text{H-NMR}$ (CDCl_3 , 300 MHz) δ 1.12 (s, 3H), 1.33-1.50 (m, 1H), 1.67-1.76 (m, 1H), 1.81-1.98 (m, 2H), 2.34-2.46 (m, 6H), 2.88-2.94 (d, 1H), 3.08-3.13 (d, 1H), 6.21 (s, 1H), 7.25-7.28 (m, 2H), 7.34-7.37 (m, 2H), 7.43 (s, 1H), 9.91 (s, 1H).

1-(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol

(11ba): white solid obtained in 50% yield; mp 178.1-181.0 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.16; $[\alpha]_{\text{D}}^{20}$ = -16 (c 1.0, MeOH); $^1\text{H-NMR}$ (CDCl_3 , 500 MHz) δ 1.13 (s, 3H), 1.23-1.29 (m, 3H), 1.34-1.45 (m, 2H), 1.56-1.62 (m, 2H), 1.72-1.78 (m, 1H), 1.90-1.94 (m, 1H), 2.27-2.30 (m, 1H), 2.32-2.37 (m, 4H), 2.47-2.50 (d, 1H), 2.97-3.00 (d, 1H), 6.14 (s, 1H), 7.23-7.26 (d, 2H), 7.34-7.36 (d, 2H), 7.40 (s, 1H); $^{13}\text{C-NMR}$ (CDCl_3 , 500 MHz) δ 19.21, 20.81, 24.34, 25.85, 33.32, 34.67, 41.14, 54.89, 56.00, 66.11, 109.50, 113.47,

123.47, 129.65, 136.81, 137.28, 137.60, 149.57; IR (NaCl, thin film) $\nu_{\max}$ (cm⁻¹): 3315 (OH), 3043 (w), 2962, 2936, 2862, 1620, 1518 (w), 1425, 1132, 820; HRMS (DART-TOF) m/z : [M+H]⁺ Calcd for C₂₁H₂₇N₂O, 323.21179; found 323.21001.

2-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)propan-2-ol (**13ba**): white solid obtained in 55% yield; mp 173-176 °C; TLC (Silica G w/UV254) 25% EtOAc in hexanes, R_f = 0.17; $[\alpha]_D^{20}$ = -75 (c 1.0, MeOH); ¹H-NMR (CDCl₃, 500 MHz) δ 1.19 (s, 3H), 1.33-1.37 (s, 8H), 1.38-1.39 (d, 1H), 1.66-1.69 (d, 1H), 1.77-1.80 (dt, 1H), 1.84-1.87 (dt, 1H), 2.31 (m, 1H), 2.39 (s, 3H), 2.73-2.76 (d, 1H), 3.53-3.57 (d, 1H), 6.16 (s, 1H) 7.24-7.25 (d, 2H), 7.35-7.38 (d, 2H), 7.40 (s, 1H); ¹³C-NMR (CDCl₃, 500 MHz) δ 19.65, 26.38, 27.48, 29.70, 33.91, 34.52, 36.19, 43.57, 57.61, 75.52, 109.76, 115.22, 123.26, 129.63, 136.59, 137.38, 150.95; IR (NaCl, thin film) $\nu_{\max}$ (cm⁻¹): 3395 (OH), 2928, 2855, 1611 (w), 1519 (m); HRMS (ESI-TOF) m/z : [M+H]⁺ Calcd for C₂₂H₂₉N₂O, 337.2274; found 337.2200.

(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(phenyl)methanol (**11bb**): white solid obtained in 28% yield; mp 184-187 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.20; $[\alpha]_D^{20}$ = -12 (c 1.0, MeOH); ¹H-NMR (CDCl₃, 500 MHz) δ 1.26 (s, 3H), 1.60-1.81 (m, 5H), 2.23-2.28 (d, 1H), 2.36-2.43 (m, 4H), 2.74-2.77 (d, 1H) 3.15-3.18 (d, 1H), 5.17 (s, 1H), 6.15 (s, 1H), 7.23-7.27 (d, 3H), 7.34-7.38 (m, 6H), 7.41 (s, 1H); ¹³C-NMR (CDCl₃, 500 MHz) δ 19.90, 21.05, 25.79, 41.28, 56.66, 71.95, 109.51, 113.67, 123.52, 125.38, 126.78, 128.16, 129.68, 137.60, 145.70, 149.77; IR (NaCl, thin film) $\nu_{\max}$ (cm⁻¹): 3319 (OH), 3039 (w), 2929, 2868, 1610 (w), 1519 (m), 821; HRMS (DART-TOF) m/z : [M+H]⁺ Calcd for C₂₆H₂₉N₂O, 385.22744; found 385.22612.

1-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-phenylethan-1-ol (**13bb**): white solid obtained in 82% yield; mp 172-174.5 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.27; $[\alpha]_D^{20}$ = -26 (c 1.0, MeOH); ¹H-NMR (CDCl₃,

500 MHz) δ 1.27 (s, 3H), 1.62-1.66 (m, 2H), 1.68 (s, 3H), 1.74-1.79 (m, 2H), 2.04-2.07 (d, 1H), 2.24-2.28 (d, 1H), 2.33-2.40 (m, 4H), 2.60-2.63 (d, 1H), 3.10-3.13 (d, 1H), 6.14 (s, 1H), 7.23-7.29 (m, 4H), 7.33-7.36 (m, 4H), 7.49-7.52 (m, 2H); ^{13}C -NMR (CDCl_3 , 500 MHz) δ 20.41, 21.02, 25.82, 26.11, 28.40, 29.68, 34.20, 35.90, 44.00, 56.78, 78.79, 109.52, 115.37, 123.26, 125.56, 126.74, 128.02, 129.62, 136.05, 136.53, 137.41, 137.72, 151.03; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3365 (OH), 3054 (w), 2928, 2858, 1610 (w), 1519 (m), 821; HRMS (ESI-RTOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{27}\text{H}_{31}\text{N}_2\text{O}$, 399.2430; found 399.2368.

(4-fluorophenyl)(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)methanol (**11bc**): white solid obtained in 11.5% yield; mp 188-191 °C; TLC (Silica G w/UV254) 25% EtOAc in hexanes, R_f = 0.24; $[\alpha]_{\text{D}}^{20}$ = -24 (c 1.0, MeOH); ^1H -NMR (CDCl_3 , 500 MHz) δ 1.26 (s, 3H), 1.50-1.61 (m, 3H), 1.67-1.71 (m, 2H), 1.79-1.82 (d, 1H), 2.25-2.28 (d, 1H), 2.39 (s, 3H), 2.74-2.77 (d, 1H), 3.16-3.20 (d, 1H), 5.19 (s, 1H), 6.16 (s, 1H), 7.02-7.46 (m, 9); ^{13}C -NMR (CDCl_3 , 500 MHz) δ 19.19, 21.04, 25.74, 33.25, 35.24, 41.24, 56.75, 71.47, 109.62, 113.54, 114.58, 123.50, 126.92, 129.69, 136.68, 136.90, 137.23, 137.58, 141.26, 149.48, 160.71, 162.66; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3312 (OH), 3041 (w), 2931, 2867, 2830, 1898, 1603, 1518, 1219, 819; HRMS (DART-RTOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{26}\text{H}_{28}\text{FN}_2\text{O}$, 403.21802; found 403.21649.

1-(4-fluorophenyl)-1-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (**13bc**): white solid obtained in 80% yield; mp 188.5-191 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.25; $[\alpha]_{\text{D}}^{20}$ = -9 (c 1.0, MeOH); ^1H -NMR (CDCl_3 , 500 MHz) δ 1.26 (s, 3H), 1.60-1.63 (m, 3H), 1.66 (s, 3H), 1.75-1.79 (d, 1H), 1.99-2.02 (d, 1H), 2.24-2.27 (d, 1H), 2.39 (s, 3H), 2.60-2.63 (d, 1H), 3.12-3.15 (d, 1H), 5.30 (s, 1H), 6.14 (s, 1H), 7.01-7.48 (m, 9H); ^{13}C -NMR (CDCl_3 , 500 MHz) δ 20.40, 21.02, 25.83, 26.16, 28.41, 34.15, 36.03, 43.98, 56.89, 78.55, 109.60, 114.59, 115.25, 123.26, 127.24, 129.63, 136.58, 137.69, 150.94; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3315 (OH), 3020 (w), 2928,

2860, 1600, 1519, 1222; HRMS (DART-RTOF) m/z : $[M+H]^+$ Calcd for $C_{27}H_{30}FN_2O$, 417.23367; found 417.23232.

(3,4-difluoro-5-methoxyphenyl)(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)methanol (**11bd**): white solid obtained in 76% yield; mp 97-101 °C; TLC (Silica G w/UV254) 25% EtOAc in hexanes, R_f = 0.22; $[\alpha]_D^{20}$ = +31 (c 1.0, MeOH); 1H -NMR ($CDCl_3$, 500 MHz) δ 1.25 (m, 3H), 1.37-1.50 (m, 2H), 1.54-1.65 (m, 2H), 1.76-1.79 (d, 1H), 2.04 (s, 3H), 2.26-2.29 (d, 1H), 2.72-2.75 (d, 1H), 3.14-3.17 (d, 1H), 3.88 (s, 3H), 5.10 (s, 1H), 6.16 (s, 1H), 6.72-6.78 (m, 2H), 7.24-7.46 (m, 6H); ^{13}C -NMR ($CDCl_3$, 500 MHz) δ 14.18, 19.94, 21.04, 25.70, 33.17, 35.31, 41.20, 56.71, 60.38, 71.28, 105.52, 106.04, 109.70, 113.35, 123.36, 129.66, 136.67, 136.99, 137.17, 137.55, 149.25; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3387 (OH), 3050 (w), 2927, 2850, 1623, 1522, 1454, 1426, 1339, 1226, 1097; HRMS (DART-RTOF) m/z : $[M+H]^+$ Calcd for $C_{27}H_{29}F_2N_2O_2$, 451.21916; found 451.21790.

1-(3,4-difluoro-5-methoxyphenyl)-1-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (**13bd**): white solid obtained in 94% yield, mp 181-183 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.19; $[\alpha]_D^{20}$ = +13 (c 3.0, MeOH); 1H -NMR ($CDCl_3$, 500 MHz) δ 1.22-1.35 (m, 6H), 1.58-1.62 (m, 4H), 1.79-1.83 (d, 1H), 1.96-1.99 (d, 1H), 2.25-2.28 (d, 1H), 2.34-2.44 (s, 3H), 2.57-2.61 (d, 1H), 3.05-3.08 (d, 1H), 3.91-3.96 (s, 3H), 6.15 (s, 1H), 6.87-6.92 (m, 2H), 7.23-7.30 (m, 4H), 7.36 (s, 1H); ^{13}C -NMR ($CDCl_3$, 500 MHz) δ 20.53, 25.64, 26.24, 28.90, 33.96, 35.77, 44.02, 56.34, 123.27, 129.65, 136.66; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3348 (OH), 2927, 2853, 1622, 1520, 1451, 1420, 1334, 1228, 1101; HRMS (ESI-RTOF) m/z : $[M+H]^+$ Calcd for $C_{28}H_{31}F_2N_2O_2$, 465.2348; found 465.2283.

(4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(thiophen-2-yl)methanol (**11be**): white solid obtained in 36% yield; mp 109-114 °C; TLC (Silica G w/UV254) 25% EtOAc in hexanes, R_f = 0.26; $[\alpha]_D^{20}$ = -17 (c 1.0, MeOH); $^1\text{H-NMR}$ (CDCl_3 , 500 MHz) δ 1.23 (s, 3H), 1.62-1.91 (m, 6H), 2.26-2.31 (m, 1H), 2.40 (s, 3H), 2.67-2.72 (m, 1H), 3.06-3.14 (m, 1H), 5.39 (s, 1H), 6.16 (s, 1H), 6.95-7.43 (m, 8H); $^{13}\text{C-NMR}$ (CDCl_3 , 500 MHz) δ 19.07, 21.03, 21.64, 23.60, 24.35, 25.83, 33.32, 33.51, 34.67, 35.08, 40.64, 41.13, 54.89, 56.02, 66.12, 67.89, 109.57, 114.13, 123.42, 129.65, 136.46, 137.59, 149.61; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3319 (OH), 3106 (w), 3045 (w), 2928, 2856, 2836, 1611, 1540 (m); HRMS (DART-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{24}\text{H}_{27}\text{N}_2\text{OS}$, 391.18386; found 391.13976.

1-((4aR)-4a-methyl-1-(p-tolyl)-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-(thiophen-2-yl)ethan-1-ol (**13be**): white solid obtained in 76% yield; mp 90-95 °C; TLC (Silica G w/UV254) 25% EtOAc in hexanes, R_f = 0.28; $[\alpha]_D^{20}$ = -12 (c 3.0, MeOH); $^1\text{H-NMR}$ (CDCl_3 , 500 MHz) δ 1.25 (s, 3H), 1.40-1.41 (d, 1H), 1.57-1.59 (d, 1H), 1.65 (s, 3H), 1.69-1.70 (d, 2H), 1.76-1.78 (d, 1H), 2.03-2.06 (d, 1H), 2.26-2.28 (d, 1H), 2.39 (s, 3H), 2.66-2.69 (d, 1H), 3.22-3.25 (d, 1H), 6.15 (s, 1H), 7.12-7.37 (m, 8H); $^{13}\text{C-NMR}$ (CDCl_3 , 500 MHz) δ 20.13, 21.02, 26.16, 28.19, 29.68, 34.31, 35.80, 43.88, 56.70, 109.66, 115.39, 119.69, 123.26, 125.64, 126.49, 129.63, 136.56, 137.73, 150.91; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3381 (OH), 3105 (w), 3042 (w), 2928, 2859, 1611, 1518 (m), 821; HRMS (DART-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{25}\text{H}_{29}\text{N}_2\text{OS}$, 405.1995; found 405.1997.

4-methoxyphenyl analogues

1-(4-methoxyphenyl)-4a-methyl-1,4,4a,6,7,8-hexahydrospiro[benzo[f]indazole-5,2'-[1,3]dioxolane] (**7c**): $^1\text{H-NMR}$ (CDCl_3 , 300 MHz) δ 1.21 (s, 3H), 1.64 (m, 2H), 1.84 (m, 2H), 2.24-2.30 (m, 1H), 2.38-2.45 (m, 1H), 2.48-2.53 (d, 1H), 3.13-3.19 (d, 1H), 3.86 (s, 3H), 3.99-4.08 (m, 4H), 6.20 (s, 1H), 6.95-6.98 (m, 2H), 7.36-7.39 (m, 3H).

1-(4-methoxyphenyl)-4a-methyl-1,4,4a,6,7,8-hexahydro-5H-benzo[f]indazol-5-one (8c):

¹H-NMR (CDCl₃, 300 MHz) δ 1.24 (s, 3H), 1.59-1.79 (m, 2H), 2.03-2.13 (m, 1H), 2.40 (s, 3H), 2.49-2.70 (m, 3H), 2.90 (s, 2H), 6.32 (s, 1H), 7.26-7.28 (m, 2H), 7.35-7.38 (m, 2H), 7.47 (s, 1H).

1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazole-5-

carbaldehyde (10c): ¹H-NMR (CDCl₃, 300 MHz) δ 1.11 (s, 3H), 1.33-1.48 (m, 1H), 1.65-1.80 (m, 1H), 1.88-1.96 (m, 2H), 2.32-2.45 (m, 3H), 2.85-2.92 (d, 1H), 3.06-3.11 (d, 1H), 3.84 (s, 3H), 6.16 (s, 1H), 6.96-6.99 (m, 2H), 7.36-7.40 (m, 3H), 9.89 (s, 1H).

1-(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-

yl)ethan-1-ol (11ca): white solid obtained in 58% yield; mp 156-159 °C; TLC (Silica G w/UV254) 25% EtOAc in hexanes, R_f = 0.11; [α]_D²⁰ = +25 (c 1.0, MeOH); ¹H-NMR (CDCl₃, 500 MHz) δ 1.13 (s, 3H), 1.23-1.29 (m, 4H), 1.34-1.45 (m, 3H), 1.74-1.78 (d, 1H), 1.88-1.95 (d, 1H), 2.28-2.31 (d, 1H), 2.35-2.39 (m, 1H), 2.47-2.51 (d, 1H), 2.98-3.01 (d, 1H), 3.84-3.87 (s, 3H), 6.10 (s, 1H), 6.96-7.40 (m, 5H); ¹³C-NMR (CDCl₃, 500 MHz) δ 19.21, 20.80, 24.36, 25.85, 33.31, 34.68, 41.16, 54.89, 55.49, 66.10, 109.38, 113.42, 114.25, 133.01, 136.73, 137.32, 149.51, 158.54; IR (NaCl, thin film) ν_{max} (cm⁻¹): 3376 (OH), 3020 (w), 2962, 2931, 2862, 1612, 1517, 1250, 831; HRMS (DART-TOF) m/z: [M+H]⁺ Calcd for C₂₁H₂₇N₂O₂, 339.20670; found 339.20344.

2-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-

yl)propan-2-ol (13ca): white solid obtained in 58% yield; mp 194-197 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.08; [α]_D²⁰ = -17 (c 1.0, MeOH); ¹H-NMR (CDCl₃, 500 MHz) δ 1.19 (s, 3H), 1.30-1.38 (d, 6H), 1.48-1.55 (m, 1H), 1.66-1.69 (d, 1H), 1.76-1.87 (dd, 2H), 2.27-2.37 (m, 3H), 2.73-2.76 (d, 1H), 3.53-3.57 (d, 1H), 3.85 (s, 3H), 6.11 (s, 1H), 6.95-7.41 (m, 5H); ¹³C-NMR (CDCl₃, 500 MHz) δ 19.66, 26.38, 27.48, 28.39, 29.68, 33.91,

34.51, 36.20, 43.60, 55.52, 57.61, 75.49, 109.57, 114.24, 114.96, 124.86, 133.10, 136.29, 137.41, 150.90, 158.42; IR (NaCl, thin film) $\nu_{\max}$ (cm⁻¹): 3349 (OH), 3018, 2928, 2857, 1612, 1517, 1248, 831; HRMS (ESI-RTOF) m/z : [M+H]⁺ Calcd for C₂₂H₂₉N₂O₂, 353.2223; found 353.2184.

1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(phenyl)methanol (**11cb**): beige solid obtained in 73% yield; mp 109.3-111.5 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.11; $[\alpha]_D^{20}$ = -10 (*c* 1.0, MeOH); ¹H-NMR (CDCl₃, 500 MHz) δ 1.27 (s, 3H), 1.56-1.81 (m, 6H), 2.24-2.27 (d, 1H), 2.36-2.44 (m, 1H), 2.75-2.78 (d, 1H), 3.17-3.20 (d, 1H), 3.85 (s, 3H), 5.20 (s, 1H), 6.11 (s, 1H), 6.96-7.44 (m, 9H); ¹³C-NMR (CDCl₃, 500 MHz) δ 19.90, 25.78, 33.28, 35.23, 41.30, 56.65, 72.02, 114.27, 125.09, 125.36, 126.84, 128.20, 132.99, 145.61, 149.62, 158.58; IR (NaCl, thin film) $\nu_{\max}$ (cm⁻¹): 3307 (OH), 3056 (w), 2932, 2868, 2835, 1611, 1518, 1250, 833; HRMS (DART-RTOF) m/z : [M+H]⁺ Calcd for C₂₆H₂₉N₂O₂, 401.22235; found 401.21952.

1-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-phenylethan-1-ol (**13cb**): beige solid obtained in 100% yield, mp 105.1-108 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.20; $[\alpha]_D^{20}$ = -30 (*c* 1.0, CHCl₃); ¹H-NMR (CDCl₃, 500 MHz) δ 1.27 (s, 3H), 1.55-1.63 (m, 3H), 1.68 (s, 3H), 1.75-1.77 (m, 1H), 2.04-2.07 (m, 1H), 2.23-2.26 (m, 1H), 2.33-2.38 (m, 1H), 2.60-2.63 (d, 1H), 3.10-3.13 (d, 1H), 3.84 (s, 3H), 6.08 (s, 1H), 6.95-7.51 (m, 10H); ¹³C-NMR (CDCl₃, 500 MHz) δ 20.43, 25.82, 26.20, 28.38, 29.68, 34.19, 35.91, 44.03, 55.48, 56.79, 78.80, 109.33, 114.22, 115.10, 125.55, 128.01, 133.11, 136.12, 137.46, 151.03, 158.39; IR (NaCl, thin film) $\nu_{\max}$ (cm⁻¹): 3338 (OH), 3020 (w), 2925, 2854, 1611, 1517, 1250, 833; HRMS (DART-RTOF) m/z : [M+H]⁺ Calcd for C₂₇H₃₁N₂O₂, 415.23800; found 415.23680.

(4-fluorophenyl)(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzof[*f*]indazol-5-yl)methanol (**11cc**): white solid obtained in 69% yield; mp 178-182 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, $R_f = 0.26$; $[\alpha]_D^{20} = +4$ (c 4.0, MeOH); $^1\text{H-NMR}$ (CDCl_3 , 500 MHz) δ 1.26 (s, 3H), 1.56-1.58 (m, 2H), 1.66-1.70 (m, 2H), 1.79-1.82 (m, 1H), 2.25-2.28 (m, 1H), 2.37-2.40 (m, 1H), 2.73-2.76 (d, 1H), 3.17-3.20 (d, 1H), 3.85 (s, 3H), 5.20 (d, 1H), 6.11 (s, 1H), 6.96-7.06 (m, 4H), 7.30-7.44 (m, 5H); $^{13}\text{C-NMR}$ (CDCl_3 , 500 MHz) δ 19.84, 20.59, 21.04, 25.80, 29.68, 33.28, 35.05, 41.22, 55.69, 69.85, 109.63, 113.56, 119.81, 123.50, 125.51, 126.00, 129.68, 137.60, 147.31, 149.4; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3399 (OH), 2930, 2850, 1602, 1512, 1223, 824; HRMS (DART-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{26}\text{H}_{28}\text{FN}_2\text{O}_2$, 419.21293; found 419.20660.

1-(4-fluorophenyl)-1-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzof[*f*]indazol-5-yl)ethan-1-ol (**13cc**): white solid obtained in 49% yield; mp 182.1-185 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, $R_f = 0.28$; $[\alpha]_D^{20} = -51$ (c 1.0, MeOH); $^1\text{H-NMR}$ (CDCl_3 , 300 MHz) δ 1.25 (s, 3H), 1.53-1.62 (m, 2H), 1.67 (s, 3H), 1.73-1.78 (m, 1H), 1.99-2.02 (d, 1H), 2.24-2.36 (m, 3H), 2.60-2.63 (d, 1H), 3.12-3.15 (d, 1H), 3.85 (s, 3H), 6.09 (s, 1H), 6.95-7.49 (m, 9H); $^{13}\text{C-NMR}$ (CDCl_3 , 500 MHz) δ 20.42, 25.82, 26.18, 28.46, 29.69, 34.14, 36.05, 44.02, 55.53, 56.90, 78.55, 109.48, 114.23, 114.59, 114.76, 124.84, 127.24, 136.09, 137.48, 151.60; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3363 (OH), 2929, 2840, 1600, 1517, 1250, 834; HRMS (DART-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{27}\text{H}_{30}\text{FN}_2\text{O}_2$, 433.22858; found 433.22897.

(3,4-difluoro-5-methoxyphenyl)(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzof[*f*]indazol-5-yl)methanol (**11cd**): white solid obtained in 71% yield, mp 174.1-177.4 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, $R_f = 0.16$; $[\alpha]_D^{20} = -50$ (c 1.0, MeOH); $^1\text{H-NMR}$ (CDCl_3 , 500 MHz) δ 1.26 (s, 3H), 1.52-1.56 (m, 2H), 1.66-1.69 (m, 2H), 1.80-1.84 (m, 2H), 2.26-2.29 (m, 1H), 2.73-2.76 (d, 1H), 3.15-3.18 (d, 1H), 3.85 (s, 3H),

3.91 (s, 3H), 5.12 (s, 1H), 6.11 (s, 1H), 6.72-6.79 (m, 2H), 6.96-6.99 (m, 2H), 7.37-7.44 (m, 3H); ^{13}C -NMR (CDCl_3 , 500 MHz) δ 20.00, 25.70, 33.15, 35.35, 41.23, 55.54, 56.72, 71.34, 106.04, 109.62, 113.09, 114.29, 125.09, 136.72, 137.36, 149.14, 158.65; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3400 (OH), 2930, 2868, 1611, 1517, 1250, 833; HRMS (DART-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{27}\text{H}_{29}\text{F}_2\text{N}_2\text{O}_3$, 467.21408; found 467.21488.

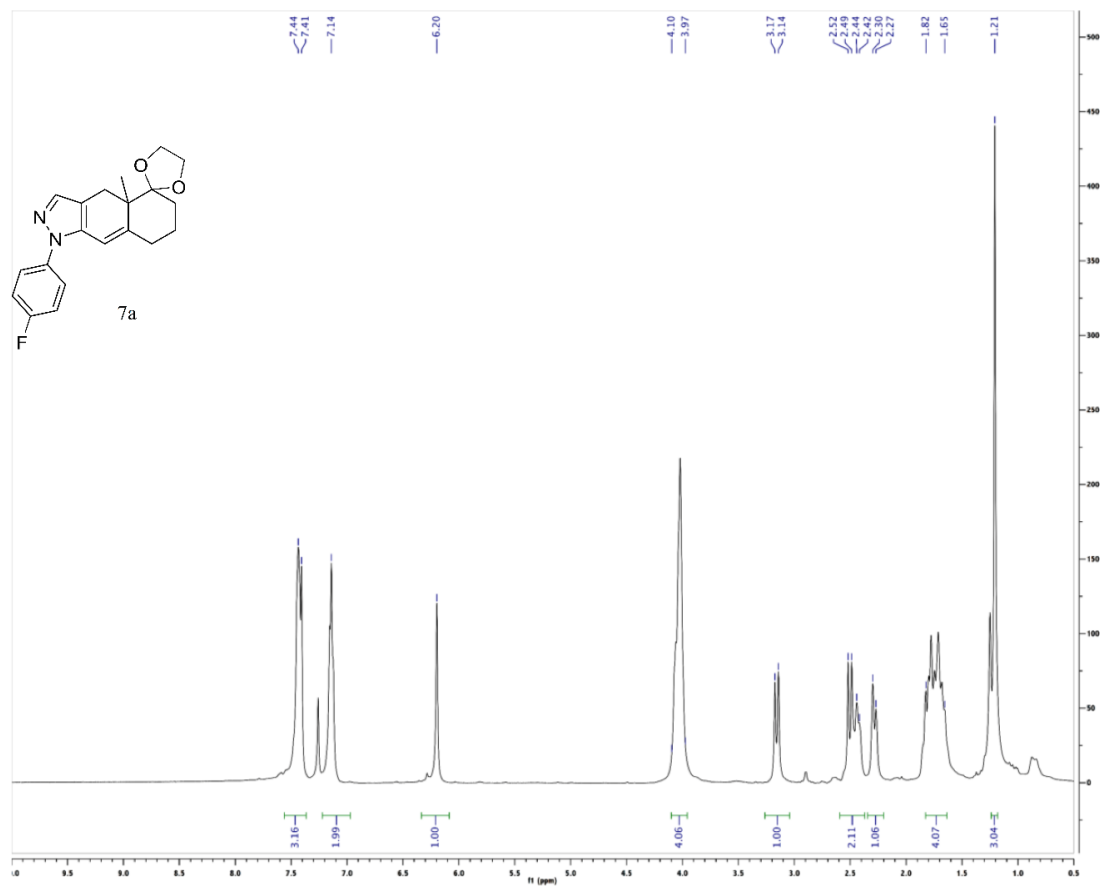
1-(3,4-difluoro-5-methoxyphenyl)-1-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)ethan-1-ol (**13cd**): white solid obtained in 65% yield, mp 184-187 °C, TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.13; $[\alpha]_{\text{D}}^{20}$ = +55 (c 1.0, MeOH); ^1H -NMR (CDCl_3 , 500 MHz) δ 1.26 (s, 3H), 1.60-1.67 (m, 6H) 1.79-1.82(d, 1H), 1.96-1.99 (d, 1H), 2.24-2.39 (m, 2H), 2.57-2.60 (d, 1H), 3.05-3.08 (d, 1H), 3.85 (s, 3H), 3.91 (s, 3H), 6.10 (s, 1H), 6.87-7.01 (m, 4H), 7.32-7.45 (m, 3H); ^{13}C -NMR (CDCl_3 , 500 MHz) δ 20.54, 25.65, 26.24, 28.89, 33.95, 35.79, 44.05, 55.53, 56.35, 56.78, 106.27, 106.78, 109.50, 114.25, 114.75, 124.85, 133.03, 137.42, 150.81; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3329 (OH), 2932, 2850, 1622, 1518, 1251, 1100, 834; HRMS (DART-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{27}\text{H}_{30}\text{FN}_2\text{O}_2$, 433.22858; found 433.22897.

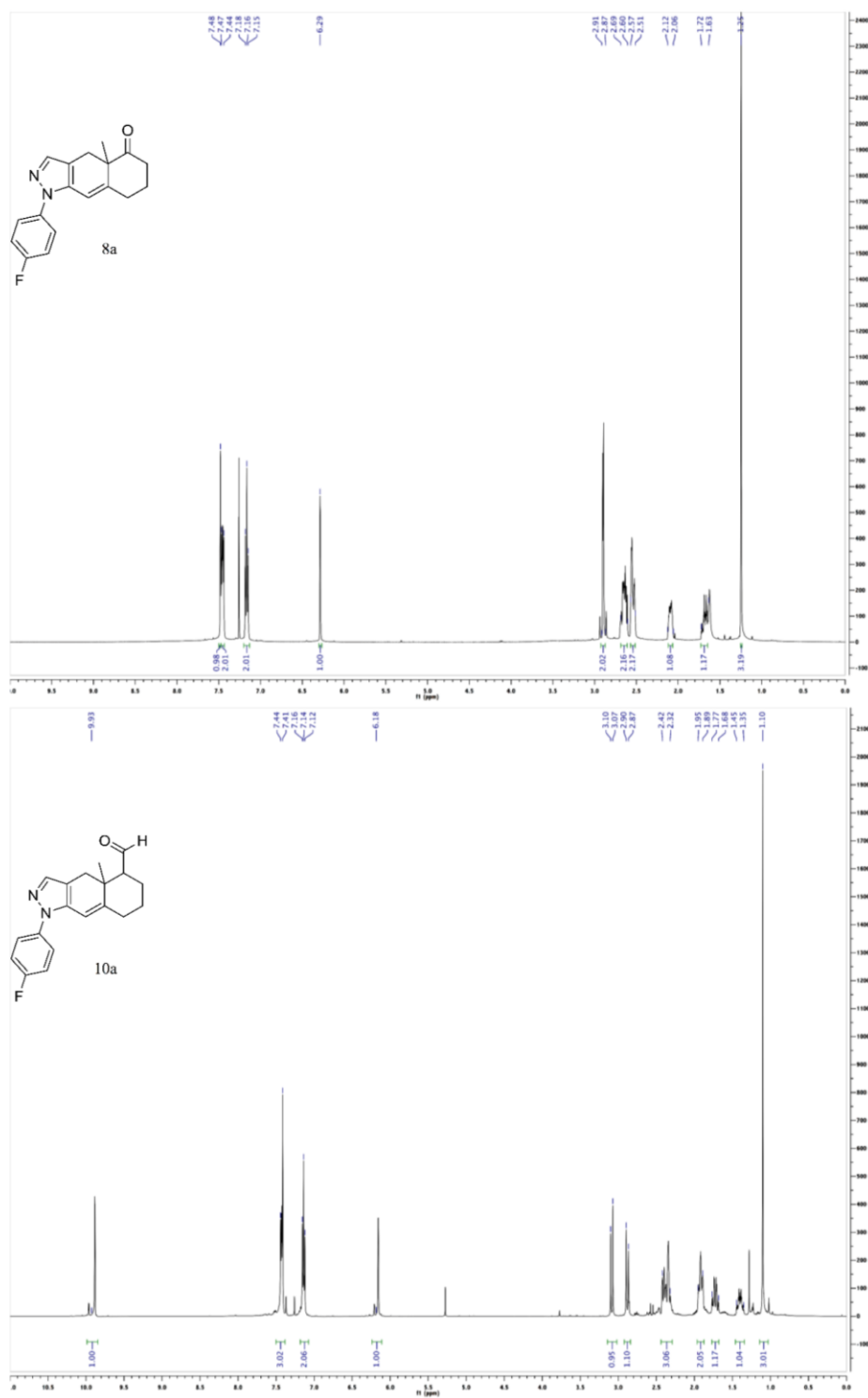
(1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)(thiophen-2-yl)methanol (**11ce**): white solid obtained in 79% yield; mp 171-173.6 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.17; $[\alpha]_{\text{D}}^{20}$ = -6.2 (c 1.0, MeOH); ^1H -NMR (CDCl_3 , 500 MHz) δ 1.26 (m, 4H), 1.62-1.72 (m, 2H), 1.77-1.85 (m, 2H), 2.27-2.41 (m, 2H), 2.69-2.72 (d, 1H), 3.12-3.15 (d, 1H), 3.85 (s, 3H), 5.25 (s, 1H), 6.12 (s, 1H), 6.96-7.40 (m, 8H); ^{13}C -NMR (CDCl_3 , 500 MHz) δ 19.87, 20.59, 25.80, 33.28, 35.04, 41.26, 55.54, 55.69, 69.83, 109.44, 113.31, 114.29, 119.82, 125.11, 125.51, 126.00, 137.26, 147.31, 158.65; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3310 (OH), 3090 (w), 2920, 2860, 1610 (w), 1519, 1230, 830; HRMS (ESI-TOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{24}\text{H}_{27}\text{N}_2\text{O}_2\text{S}$, 407.1787; found 407.1720.

1-((4aR)-1-(4-methoxyphenyl)-4a-methyl-4,4a,5,6,7,8-hexahydro-1H-benzo[f]indazol-5-yl)-1-(thiophen-2-yl)ethan-1-ol (**13ce**): white solid obtained in 48% yield; mp 146-148.5 °C; TLC (Silica G w/UV254) 30% EtOAc in hexanes, R_f = 0.16; $[\alpha]_D^{20}$ = +12 (c 1.0, MeOH); $^1\text{H-NMR}$ (CDCl_3 , 500 MHz) δ 1.25 (s, 3H), 1.59-1.65 (m, 2H), 1.68 (s, 3H), 1.73-1.78 (m, 2H), 2.03-2.06 (d, 1H), 2.25-2.36 (m, 2H), 2.65-2.69 (d, 1H), 3.22-3.25 (d, 1H), 3.84 (s, 3H), 6.10 (s, 1H), 6.94-7.40 (m, 8H); $^{13}\text{C-NMR}$ (CDCl_3 , 500 MHz) δ 20.16, 26.16, 28.19, 34.31, 35.82, 43.91, 55.48, 56.71, 77.52, 109.47, 114.23, 115.13, 119.72, 124.88, 126.51, 133.11, 136.13, 137.47, 150.93, 152.99, 158.40; IR (NaCl, thin film) ν_{max} (cm^{-1}): 3365 (OH), 2930, 2850, 1610, 1517, 1250, 835; HRMS (ESI-RTOF) m/z : $[\text{M}+\text{H}]^+$ Calcd for $\text{C}_{25}\text{H}_{29}\text{N}_2\text{O}_2\text{S}$, 421.1944; found 421.1907.

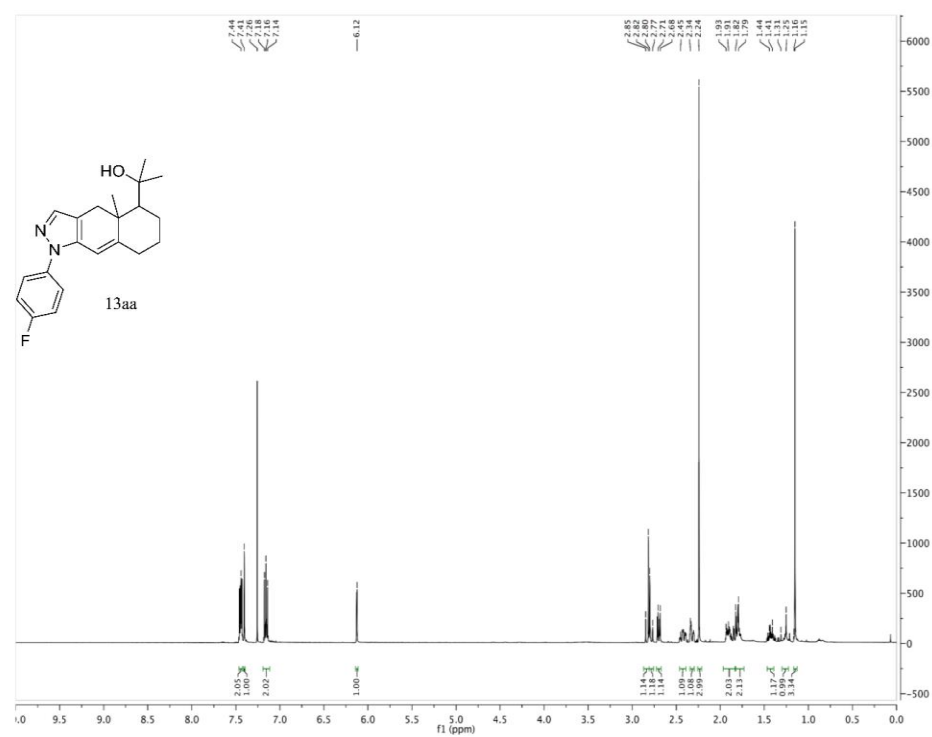
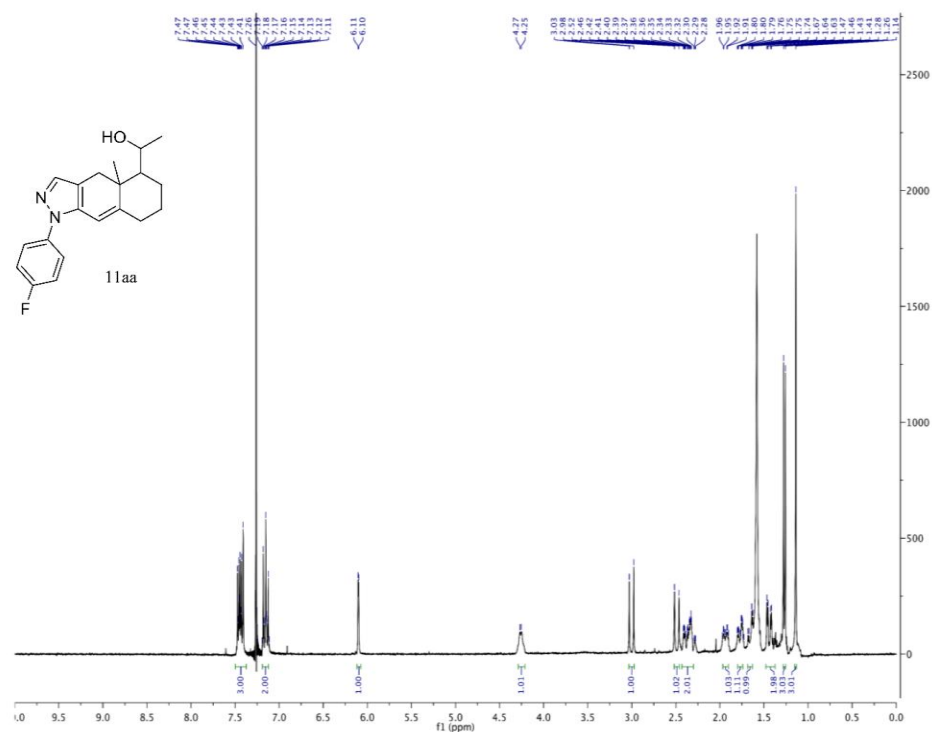
1.8.5 – ^1H -NMR and ^{13}C -NMR spectra

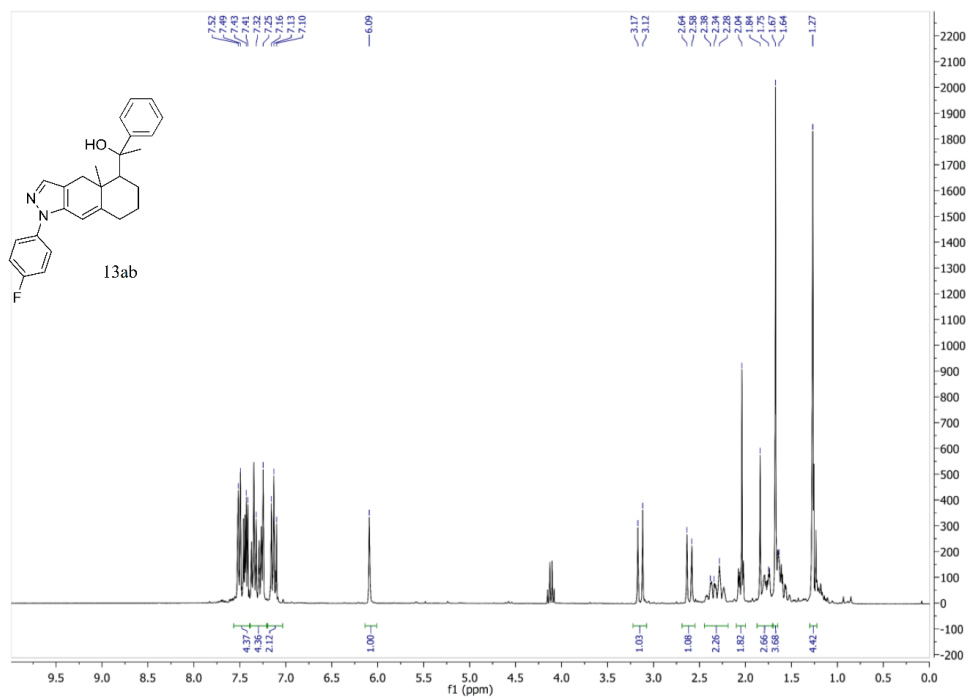
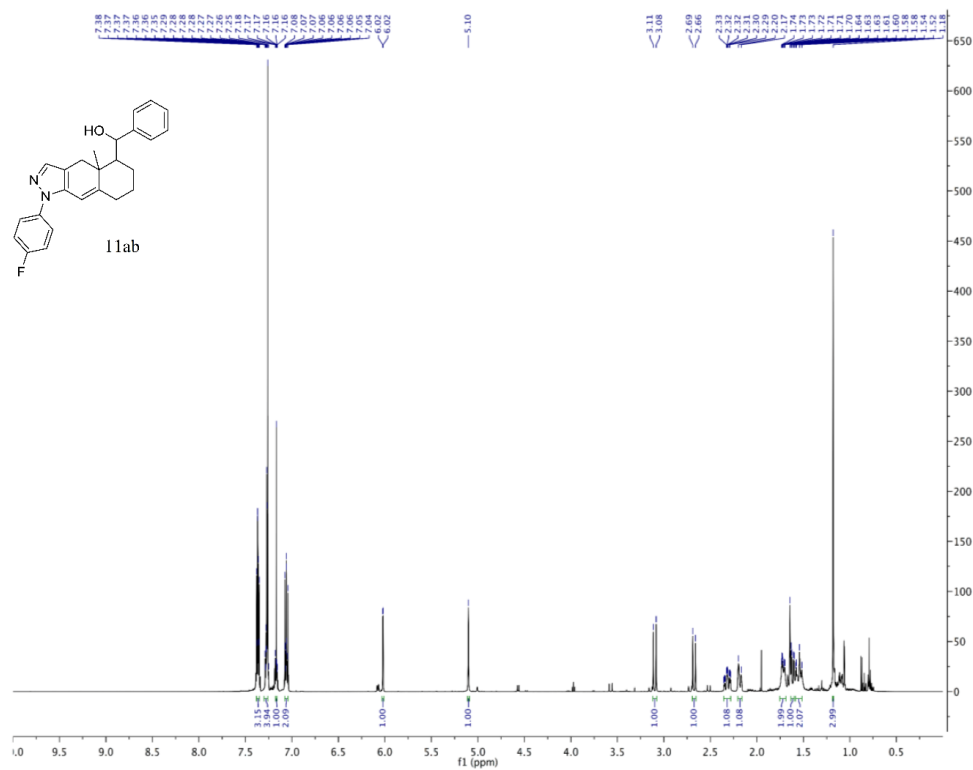
^1H -NMR for compound **7a**



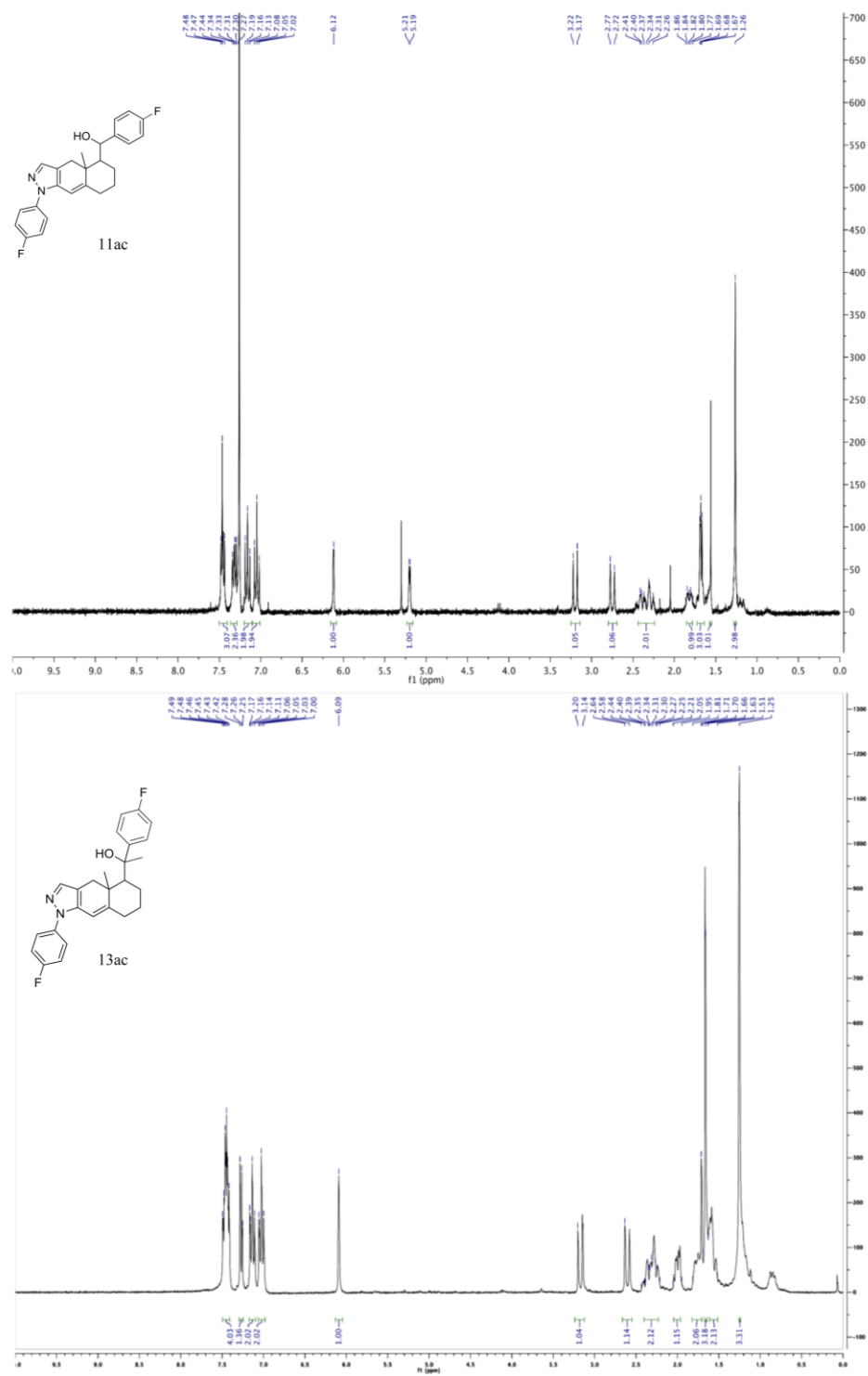
¹H-NMR for compounds **8a** and **10a**

¹H-NMR for compounds **11aa** and **13aa**

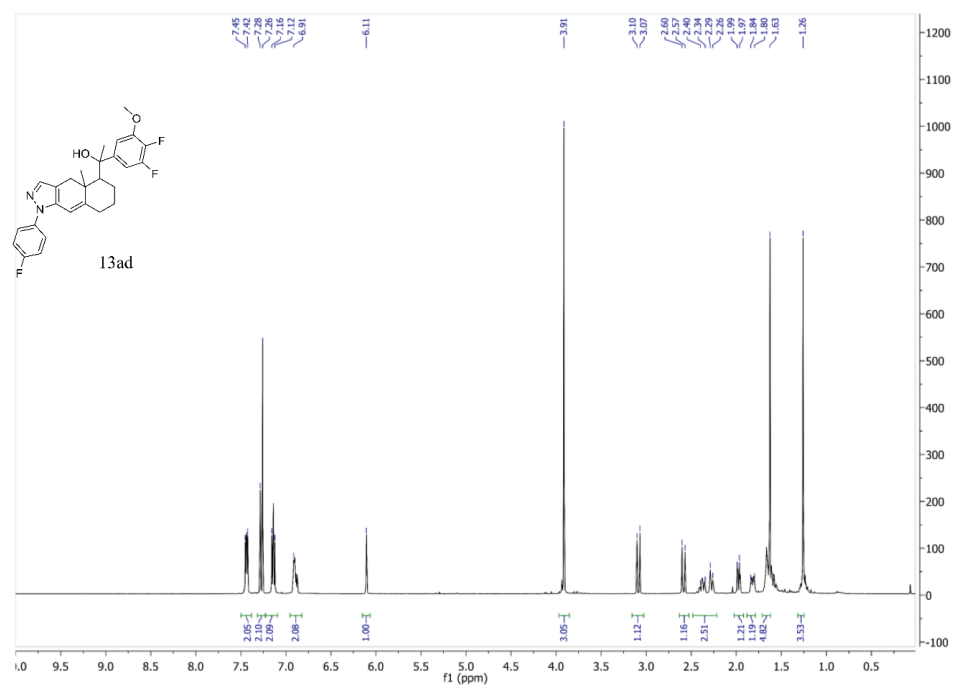
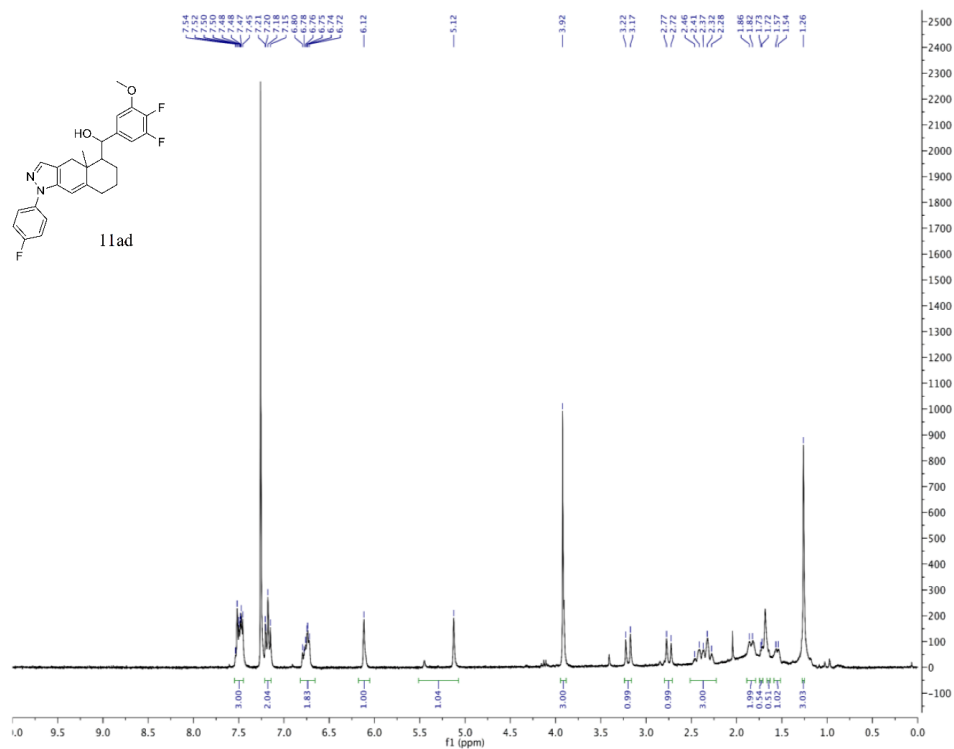


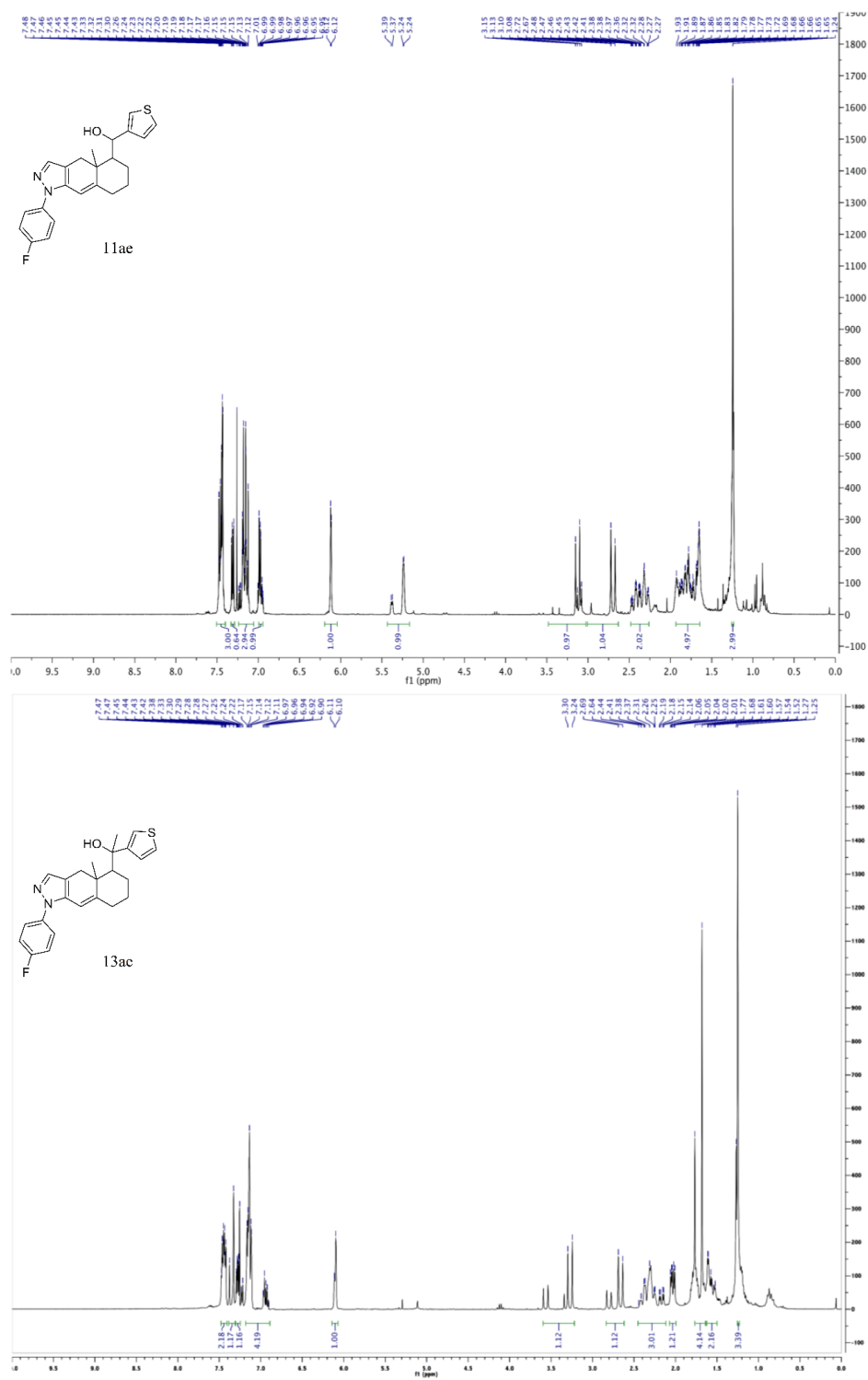
¹H-NMR for compounds **11ab** and **13ab**

^1H -NMR for compounds **11ac** and **13ac**

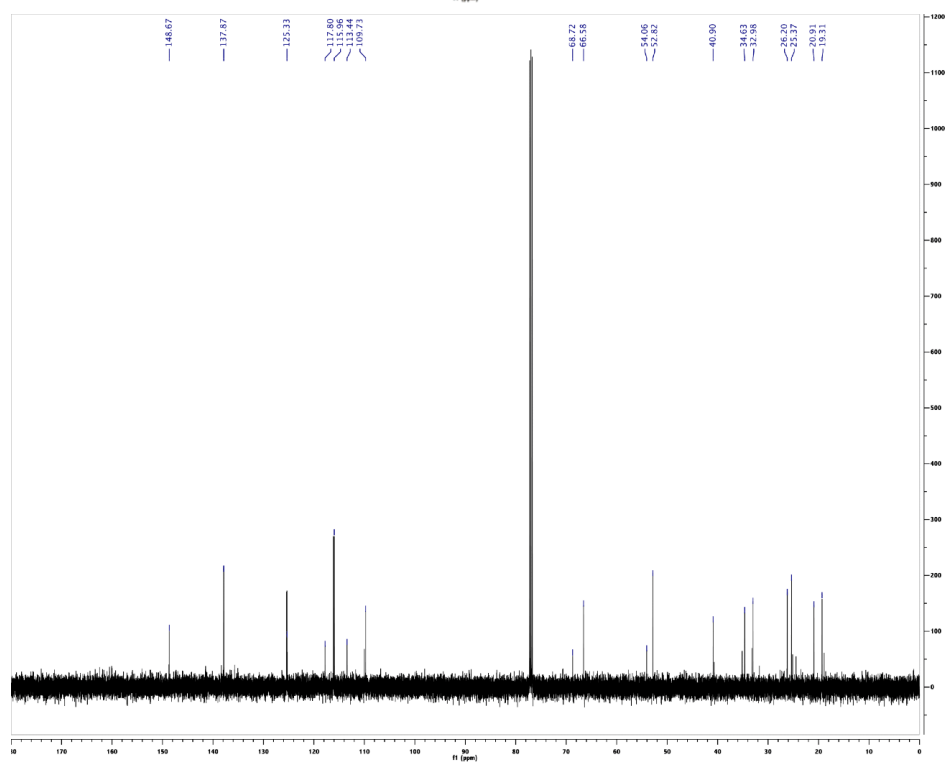
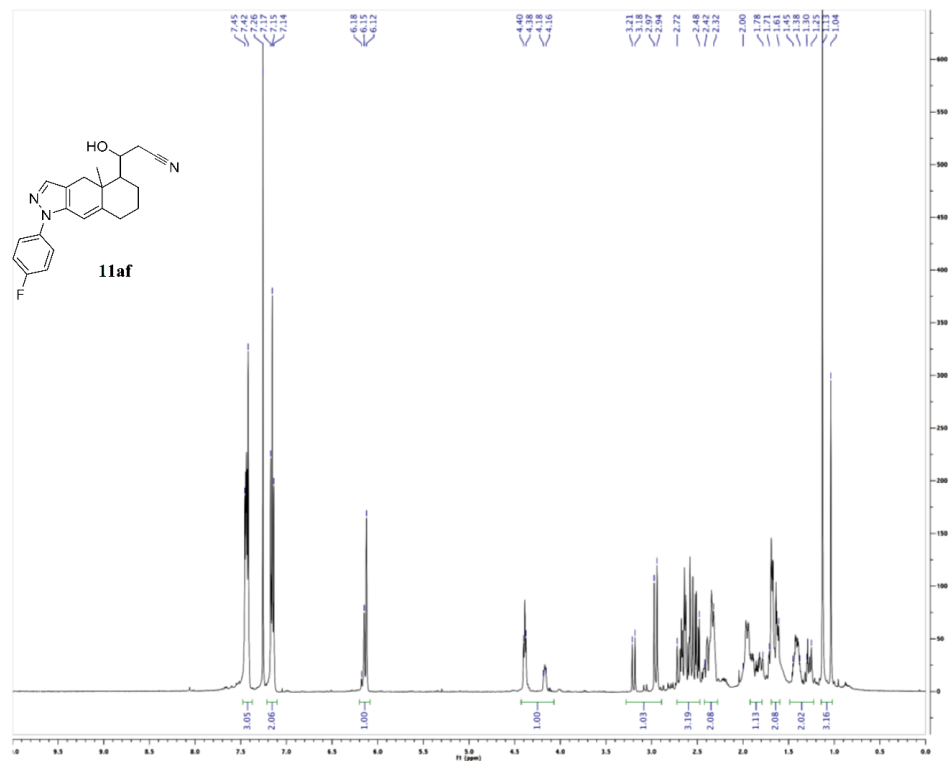


¹H-NMR for compounds **11ad** and **13ad**

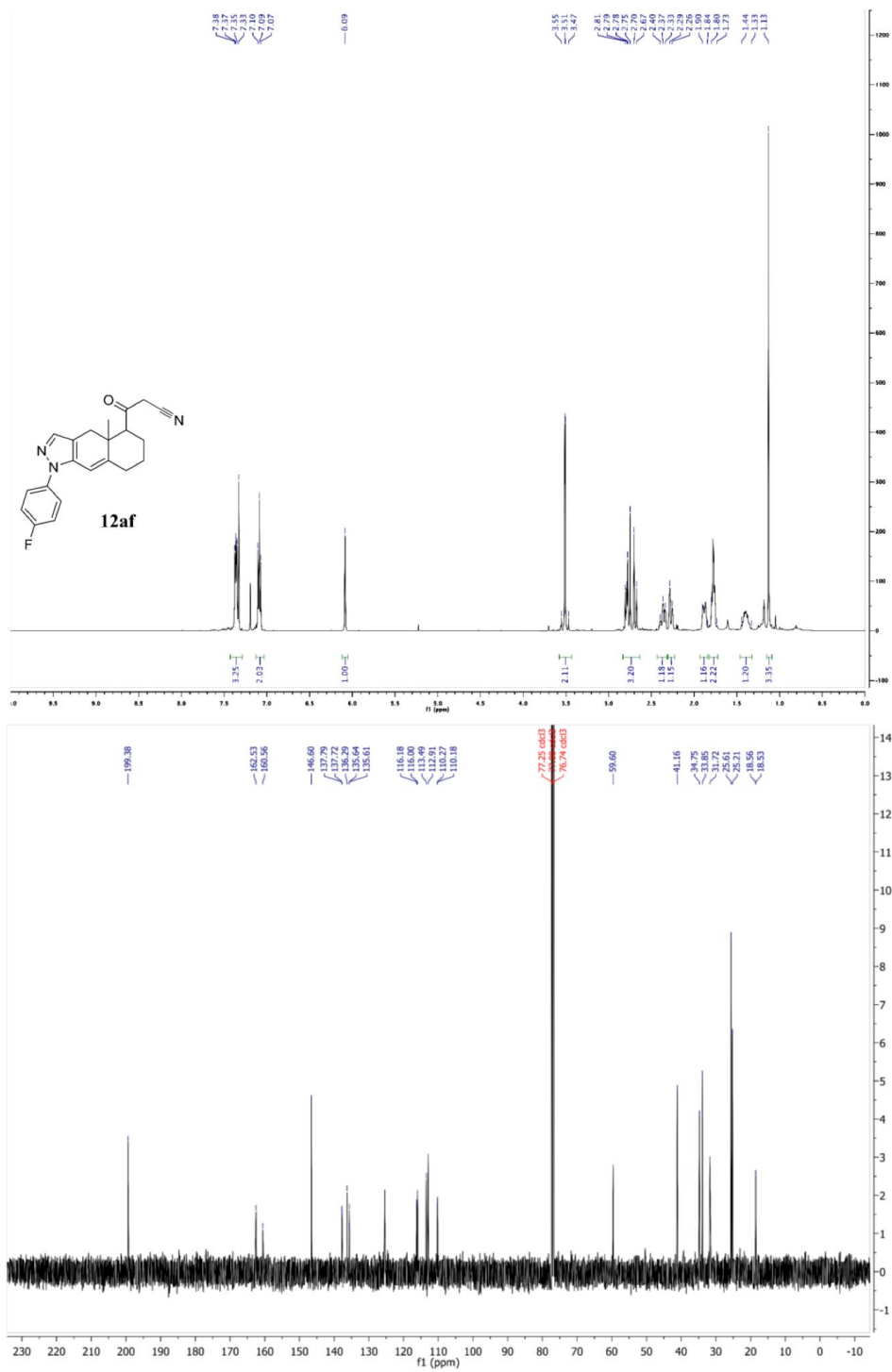


¹H-NMR for compounds **11ae** and **13 ae**

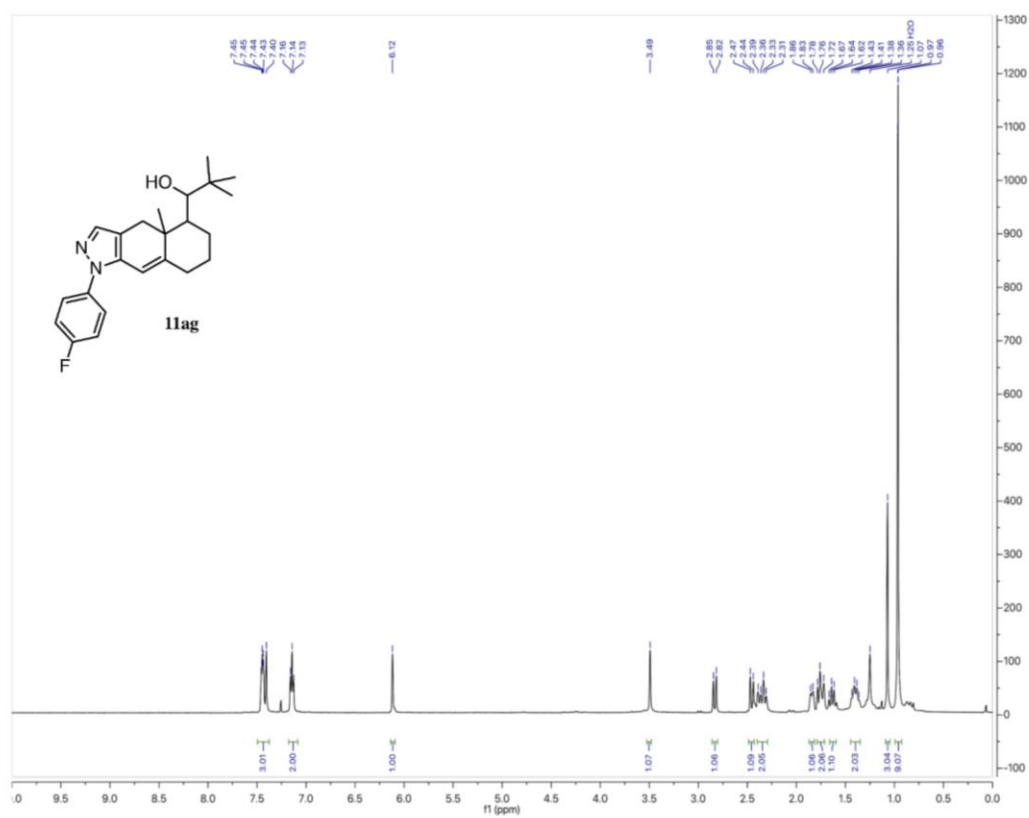
^1H -NMR and ^{13}C -NMR for compound **11af**



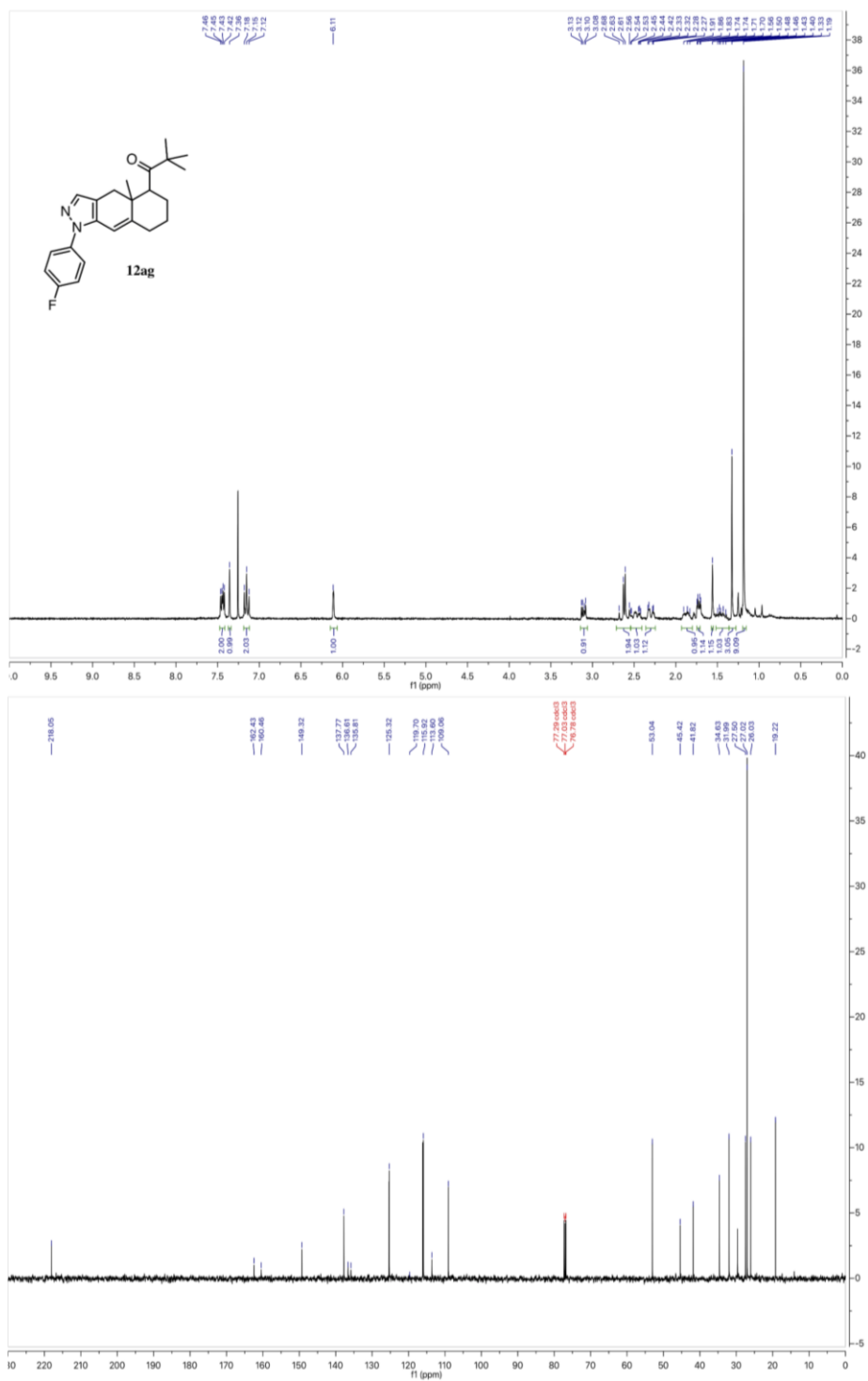
^1H -NMR and ^{13}C -NMR for compound **12af**



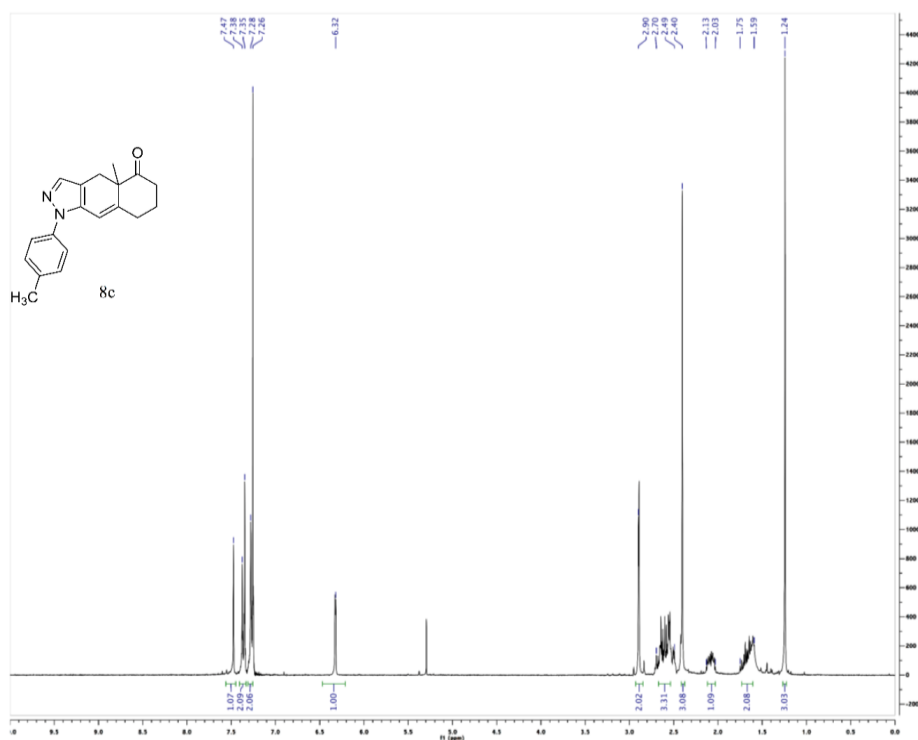
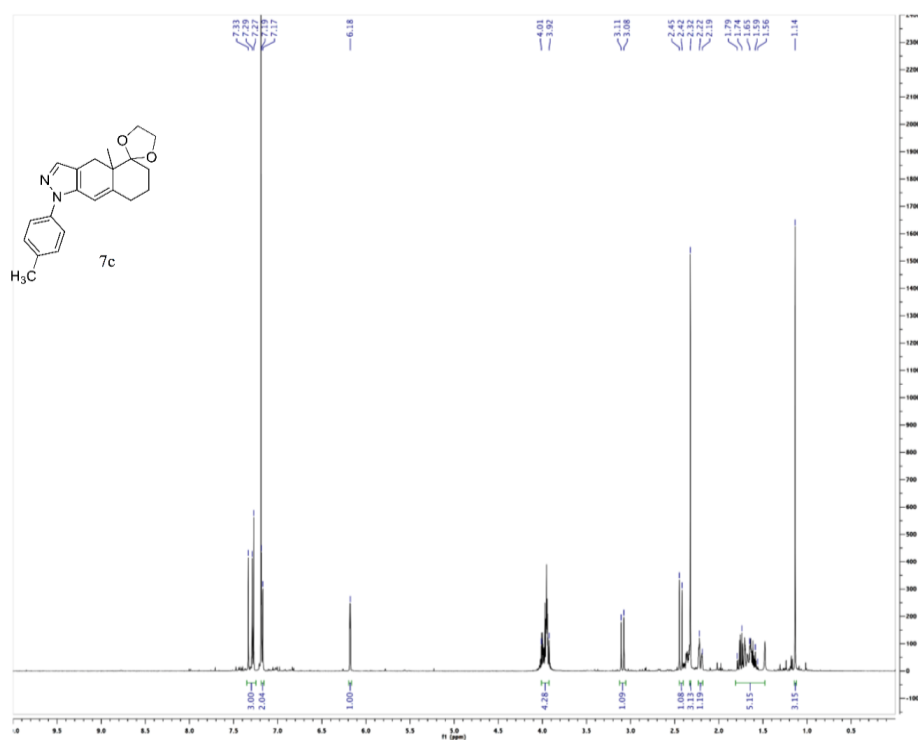
¹H-NMR for compound **11ag**

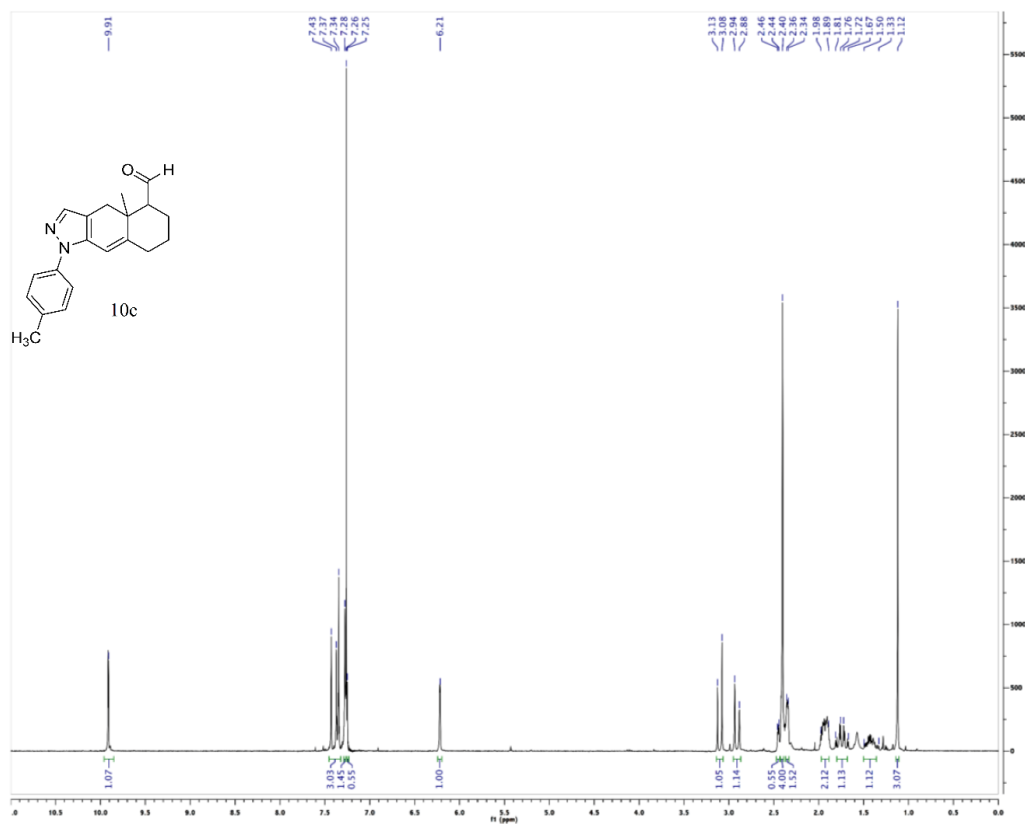


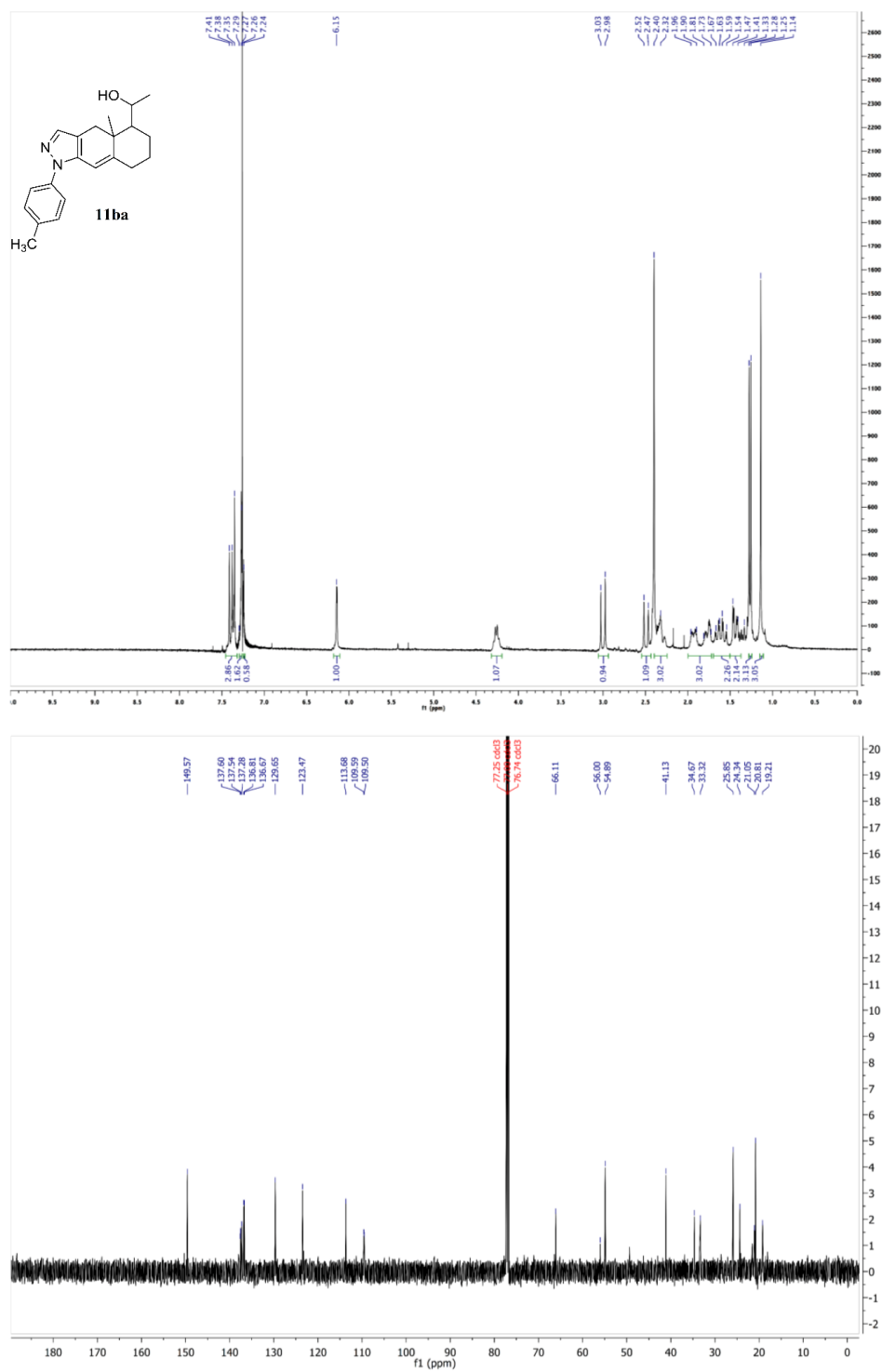
^1H -NMR and ^{13}C -NMR for compound **12ag**



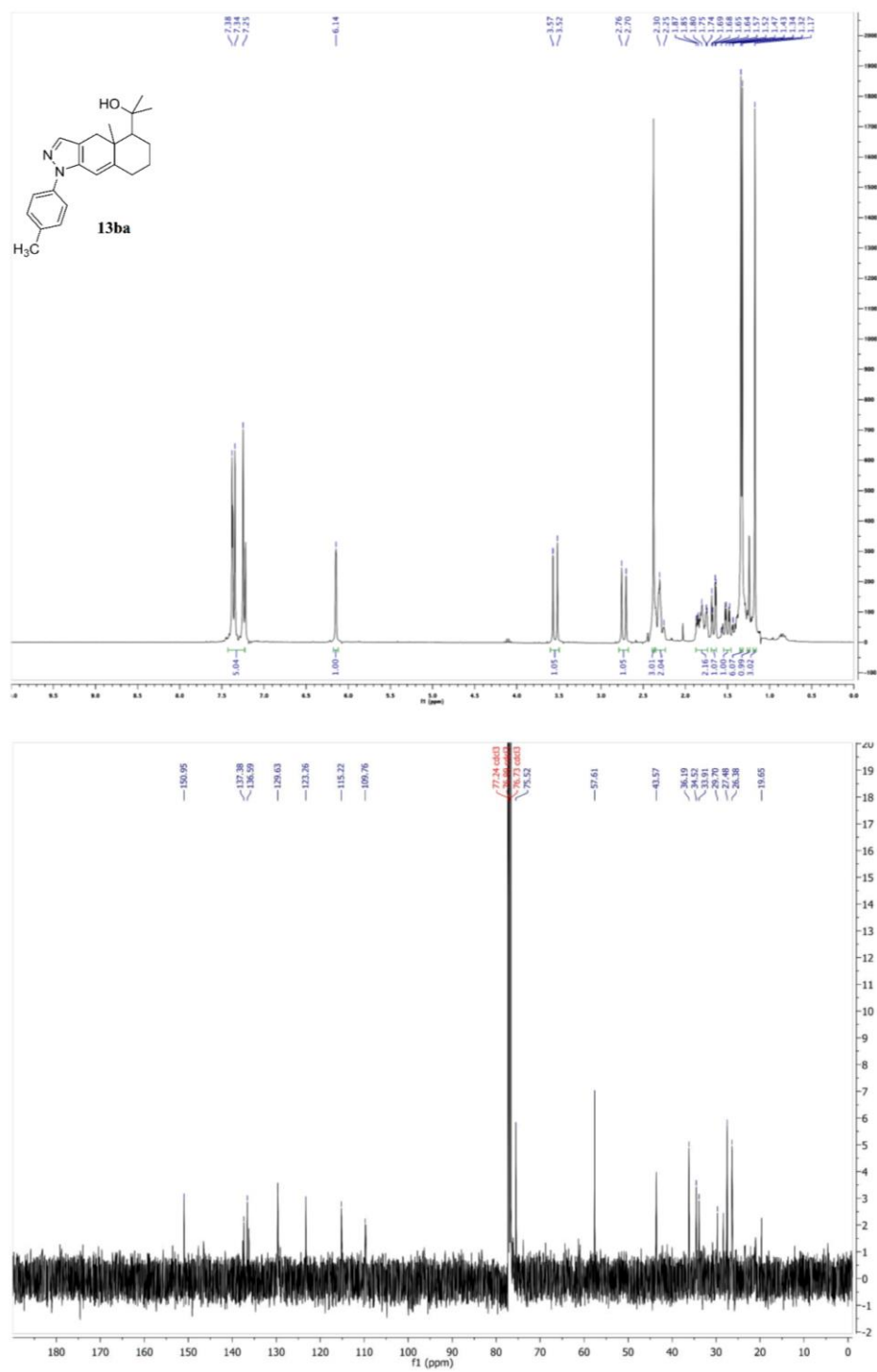
¹H-NMR for compounds **7c** and **8c**



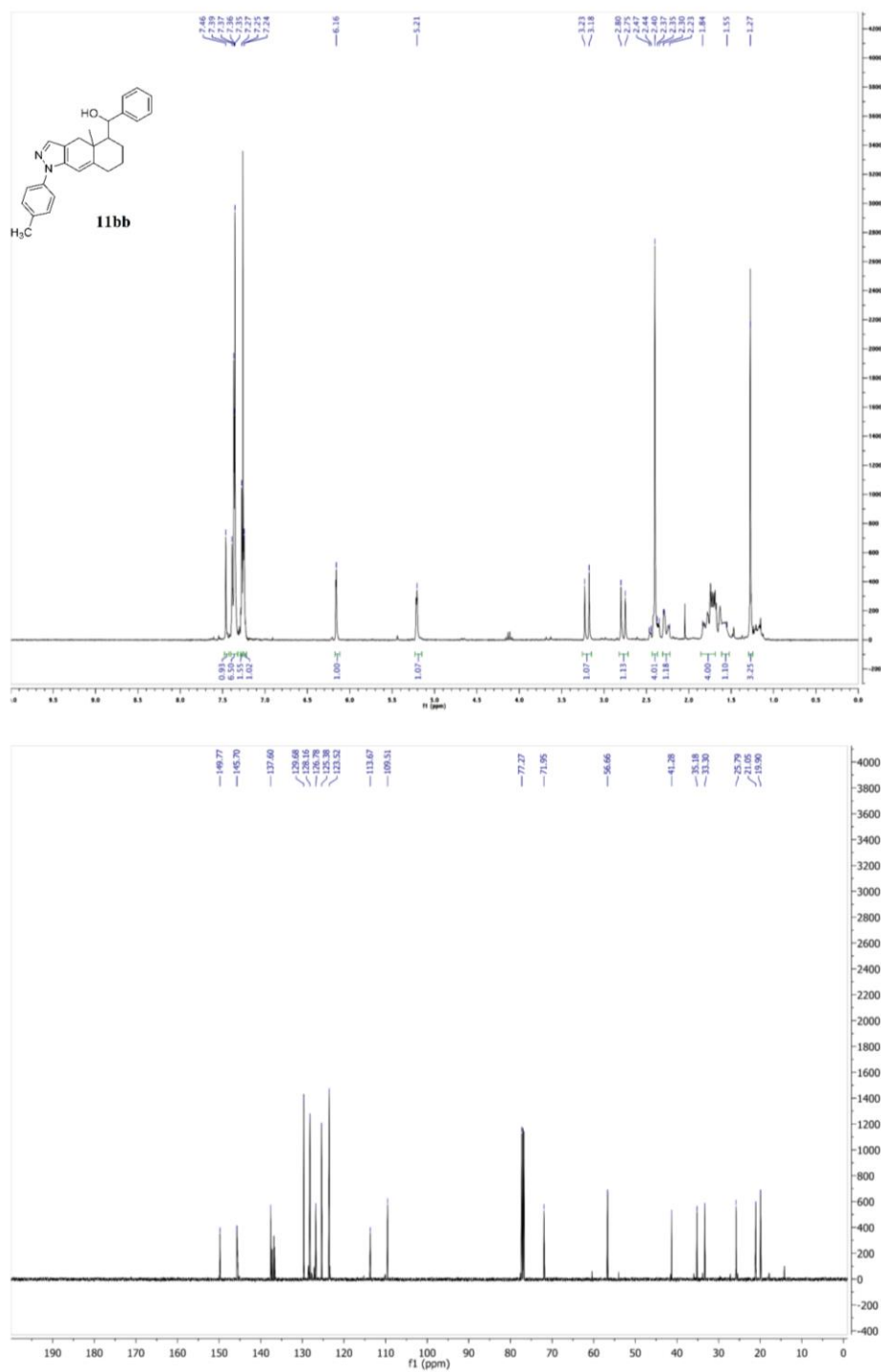
¹H-NMR for compound **10c**

^1H -NMR and ^{13}C -NMR for compound **11ba**

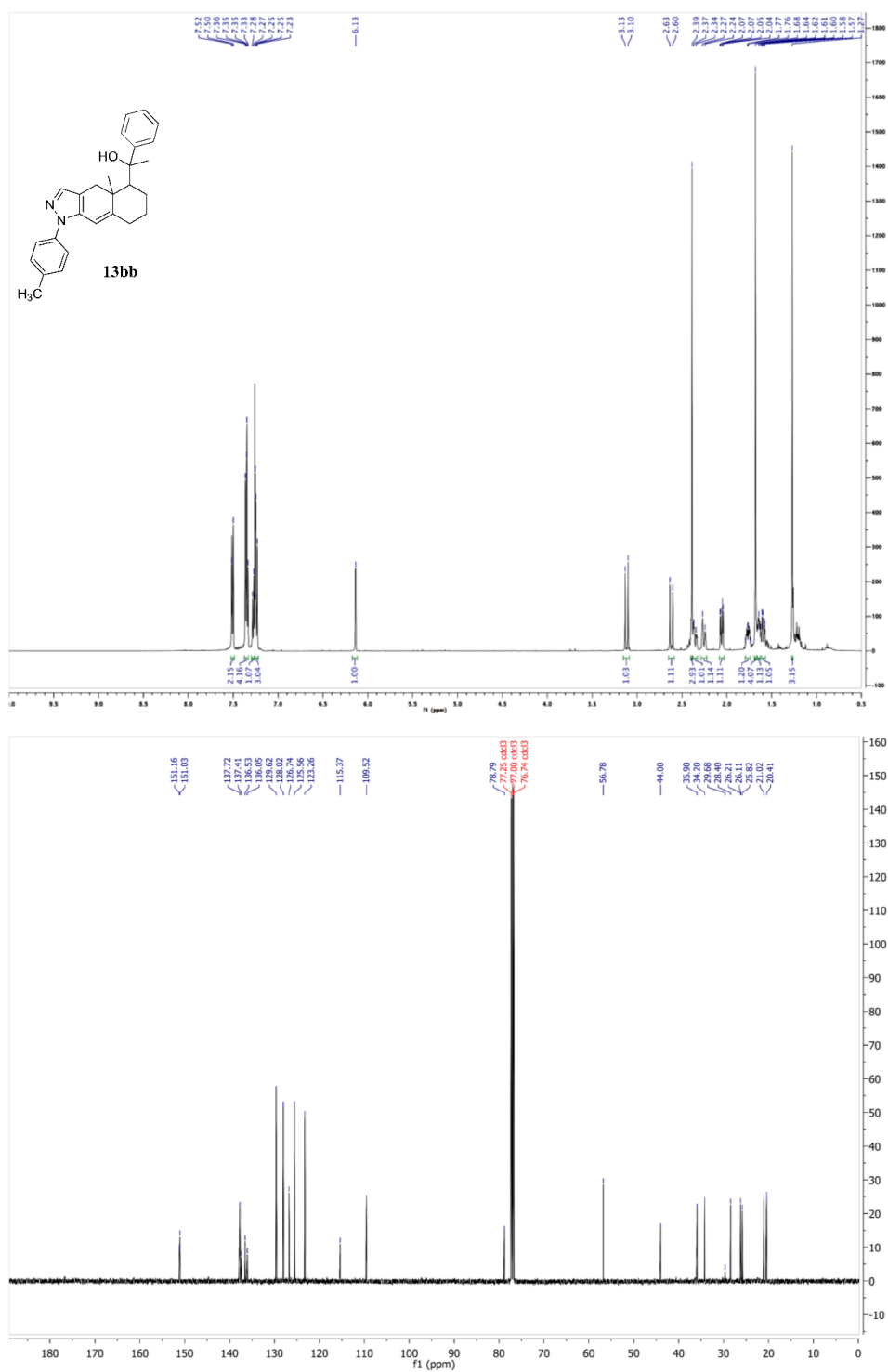
^1H -NMR and ^{13}C -NMR for compound **13ba**

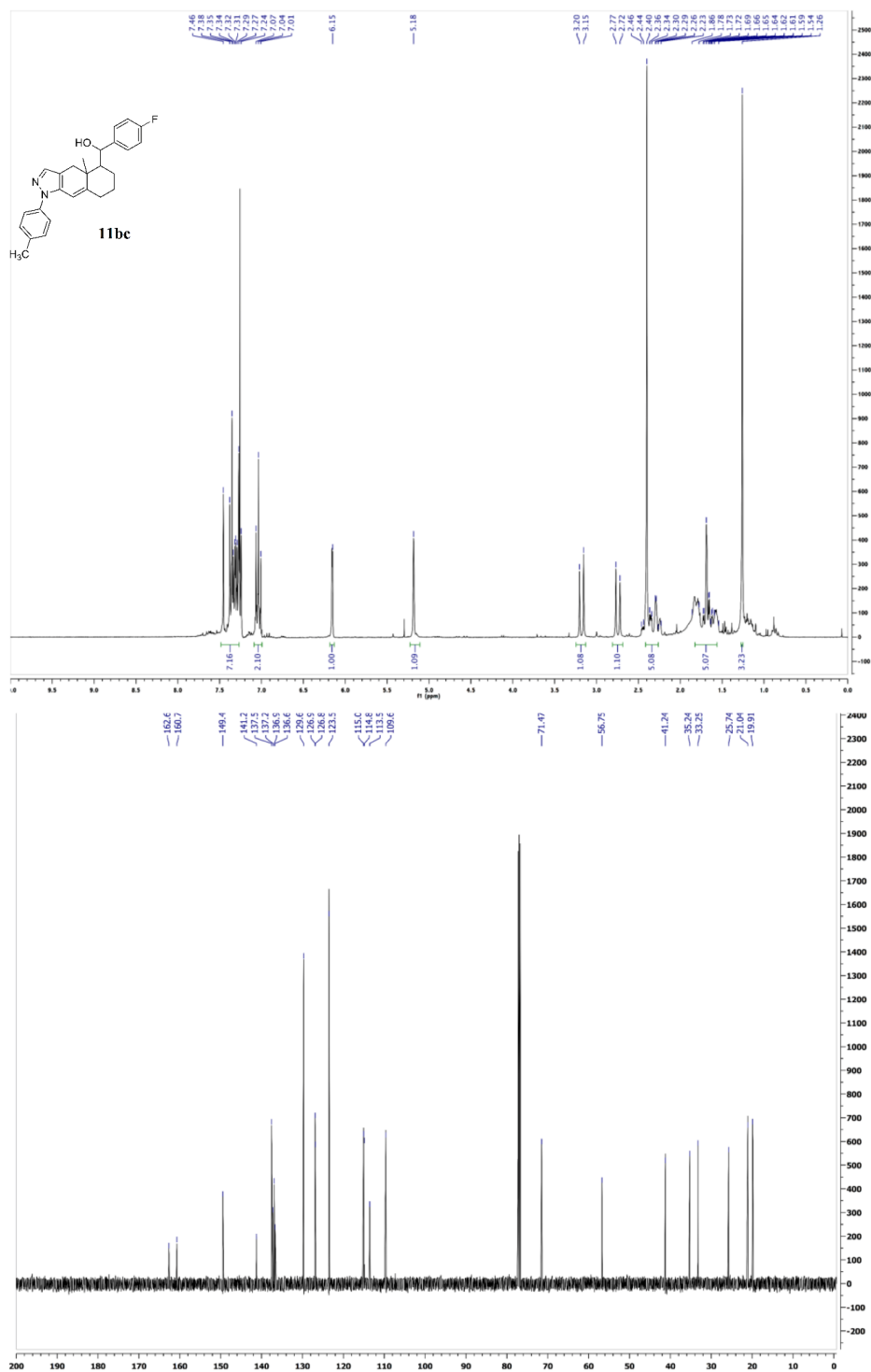


^1H -NMR and ^{13}C -NMR for compound **11bb**

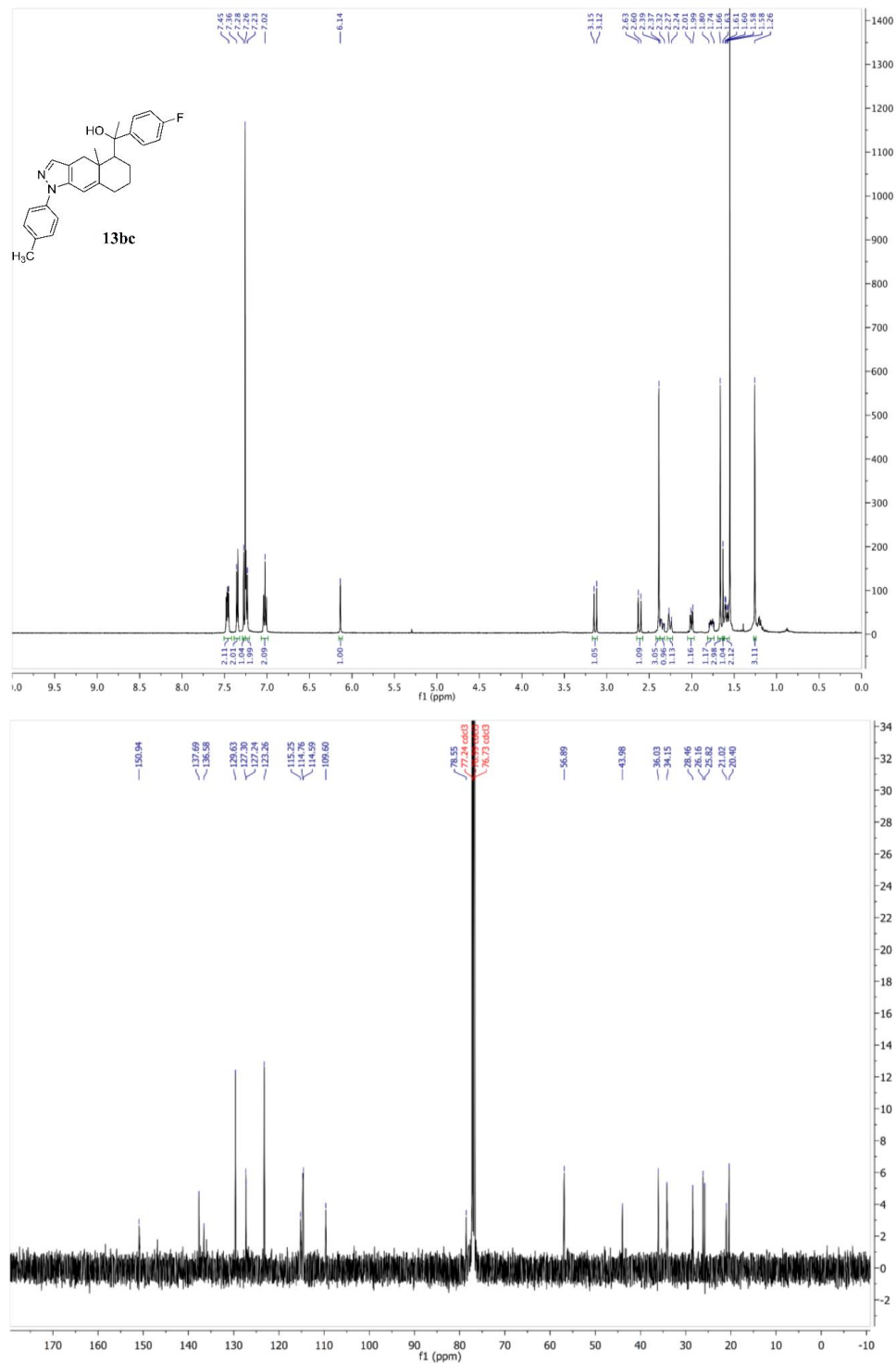


^1H -NMR and ^{13}C -NMR for compound **13bb**

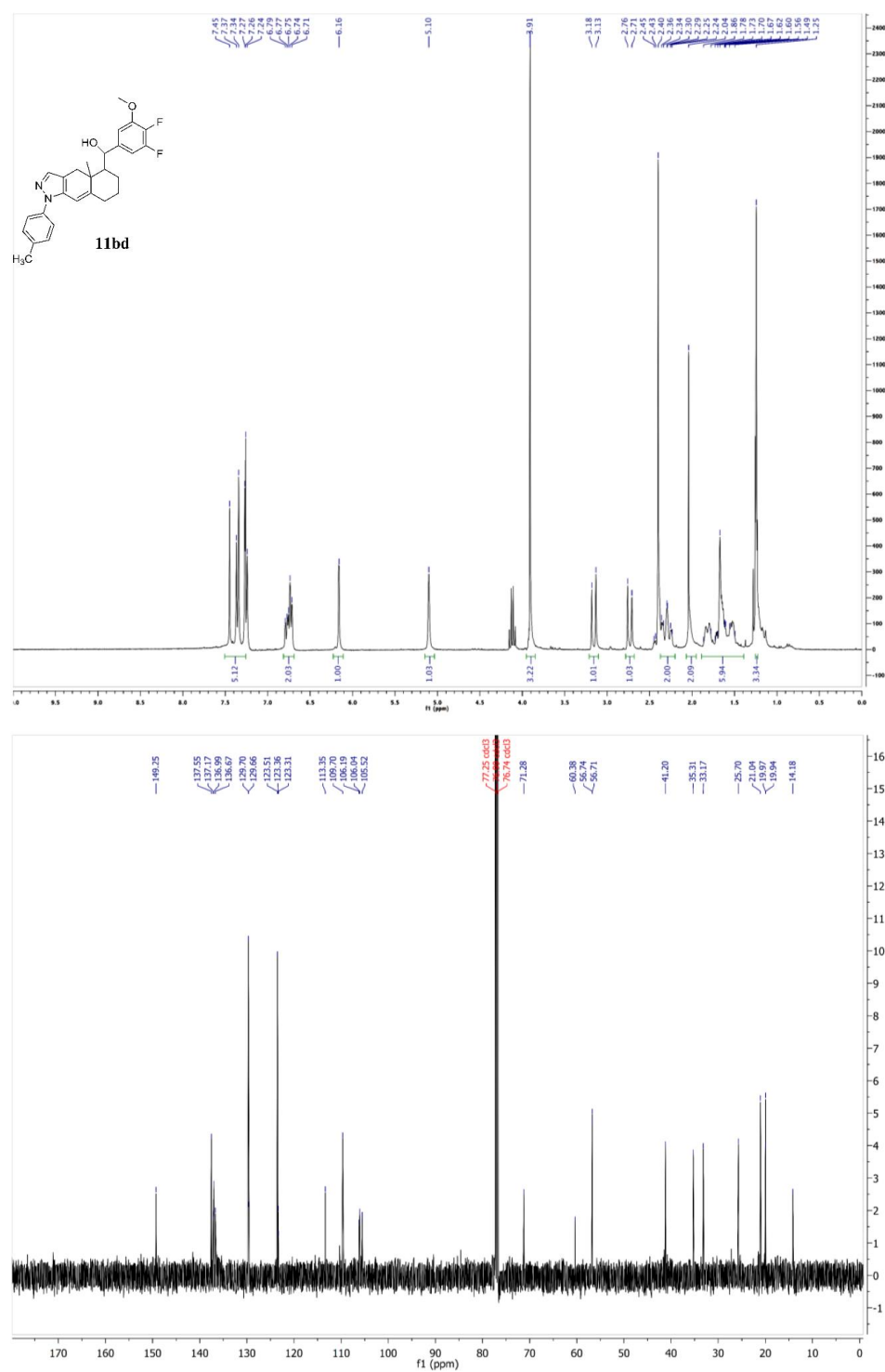


^1H -NMR and ^{13}C -NMR for compound **11bc**

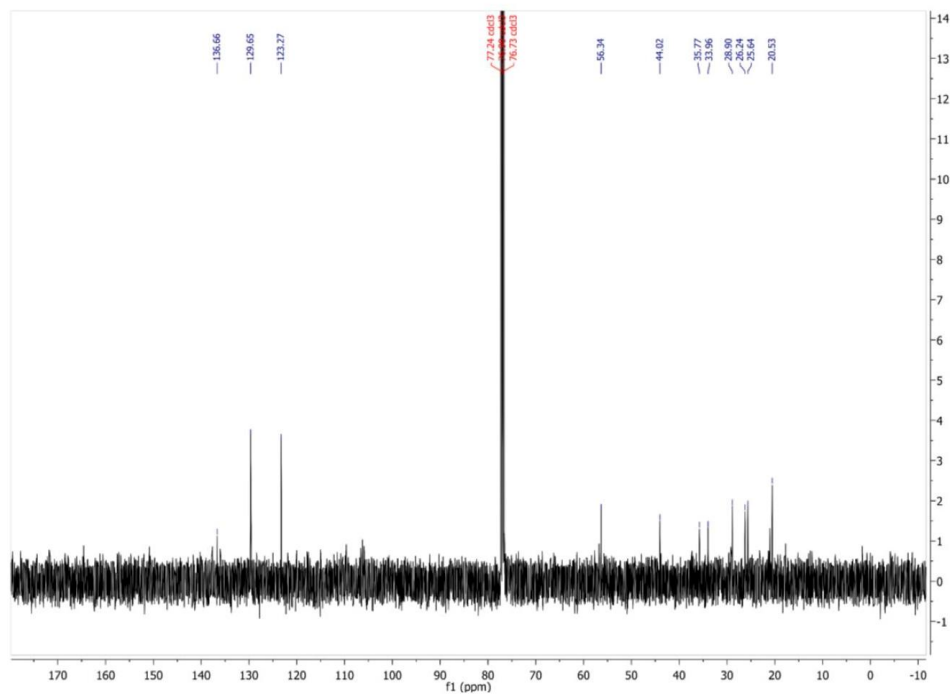
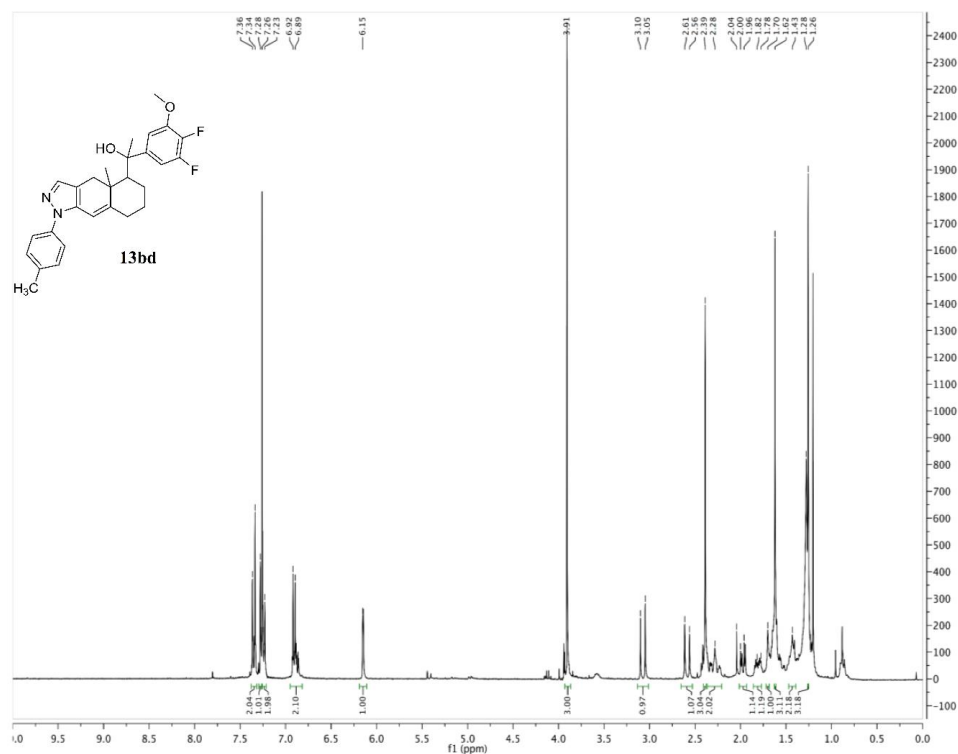
^1H -NMR and ^{13}C -NMR for compound **13bc**



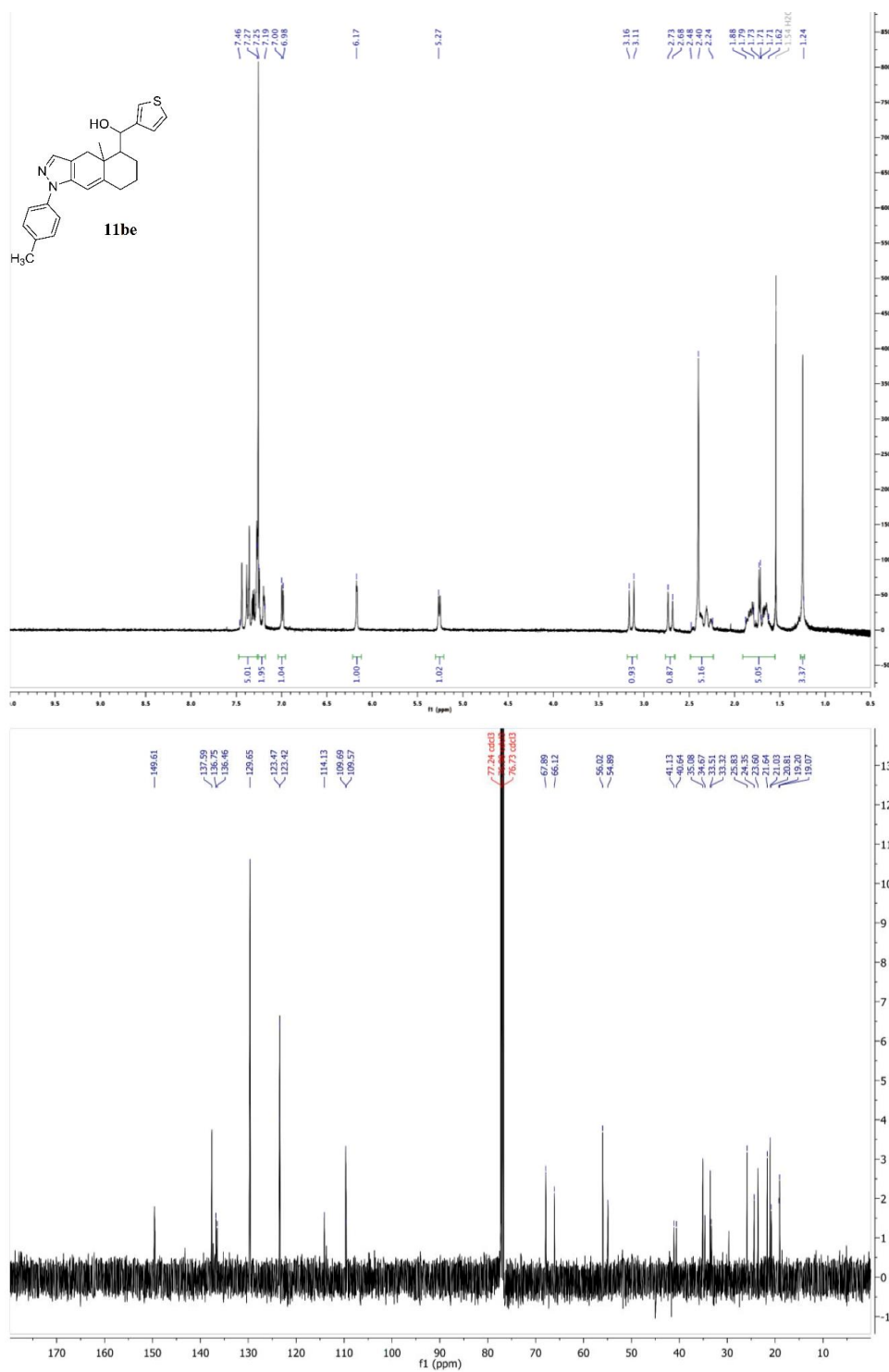
^1H -NMR and ^{13}C -NMR for compound **11bd**



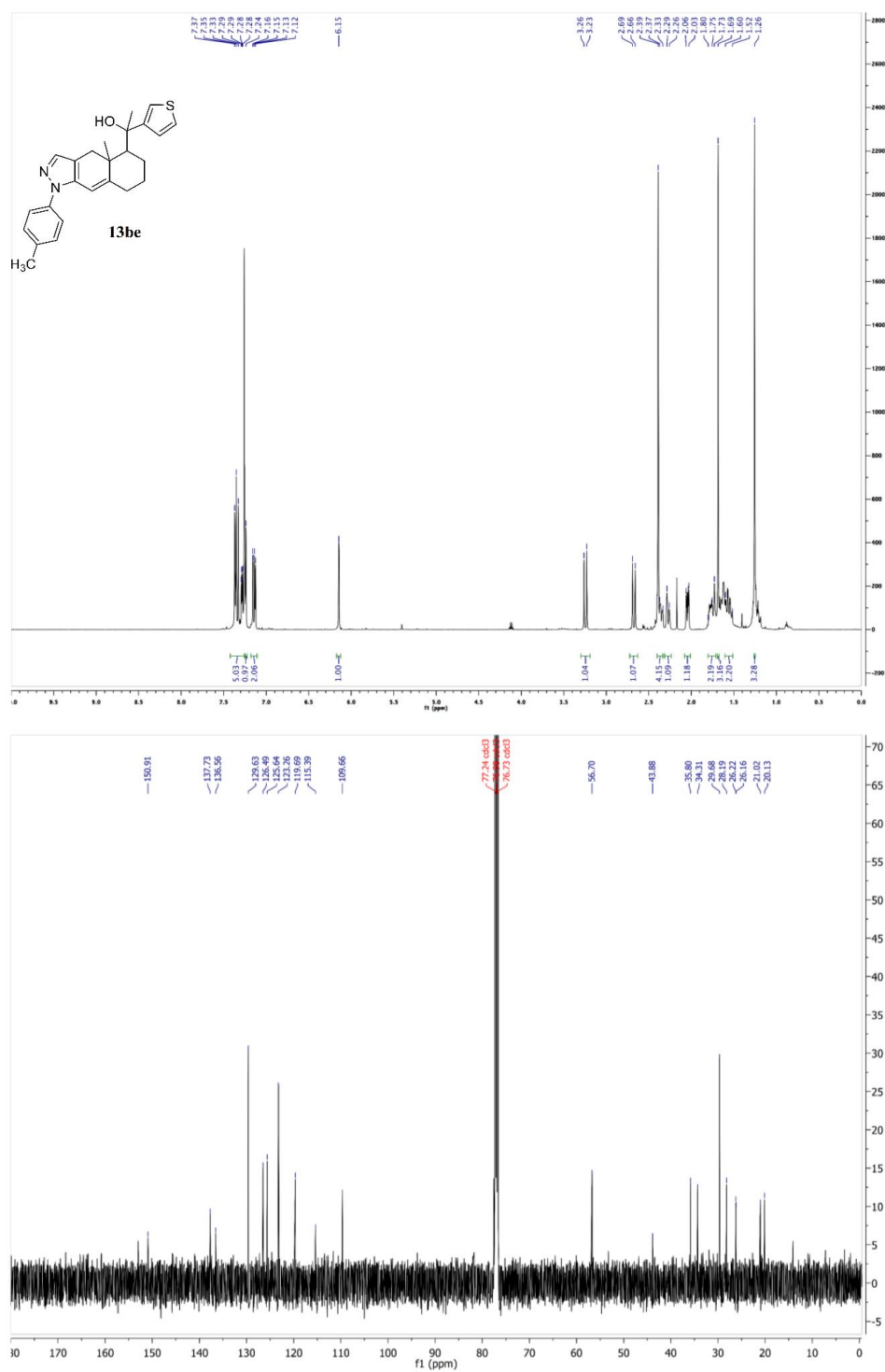
^1H -NMR and ^{13}C -NMR for compound **13bd**



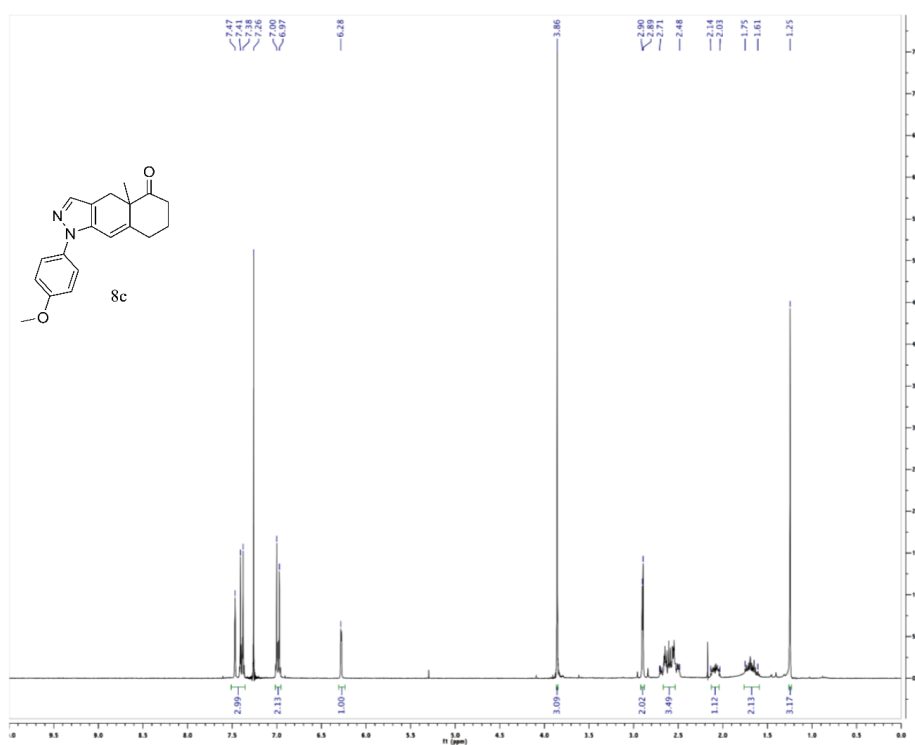
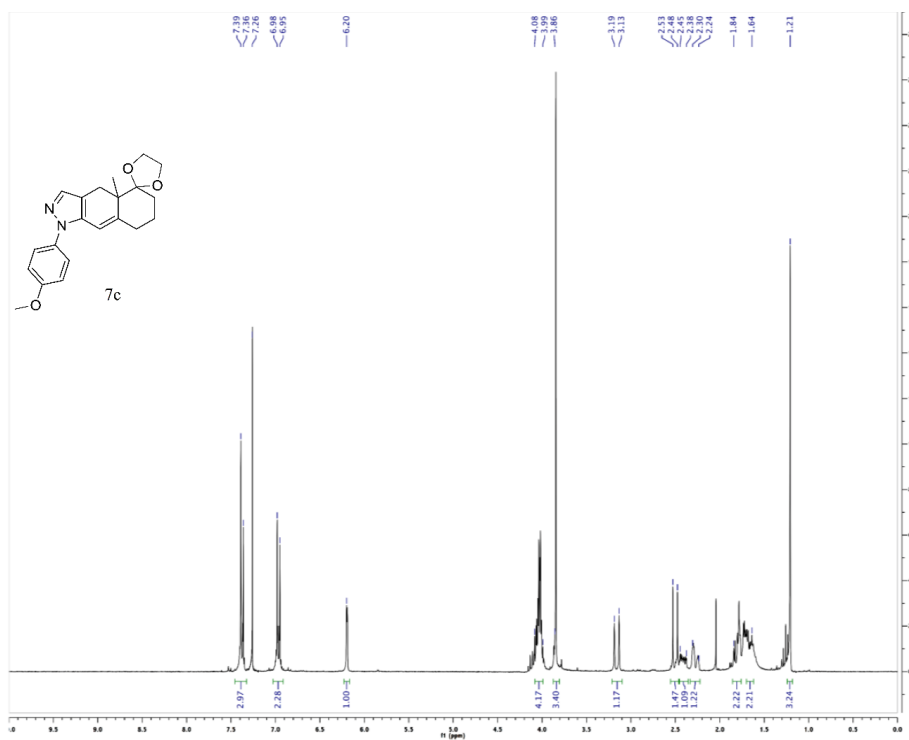
^1H -NMR and ^{13}C -NMR for compound **11be**



^1H -NMR and ^{13}C -NMR for compound **13be**



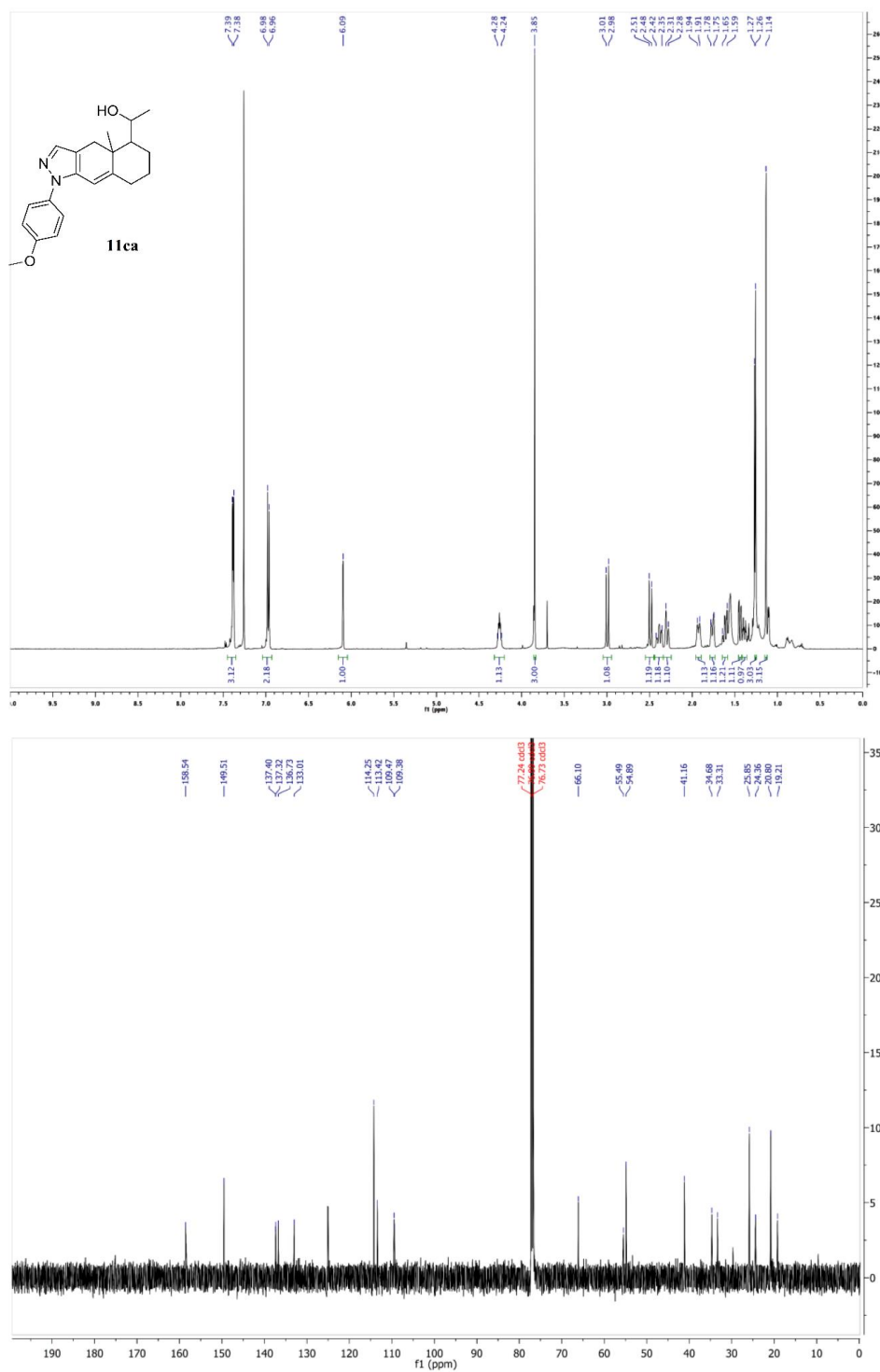
¹H-NMR for compounds **7c** and **8c**



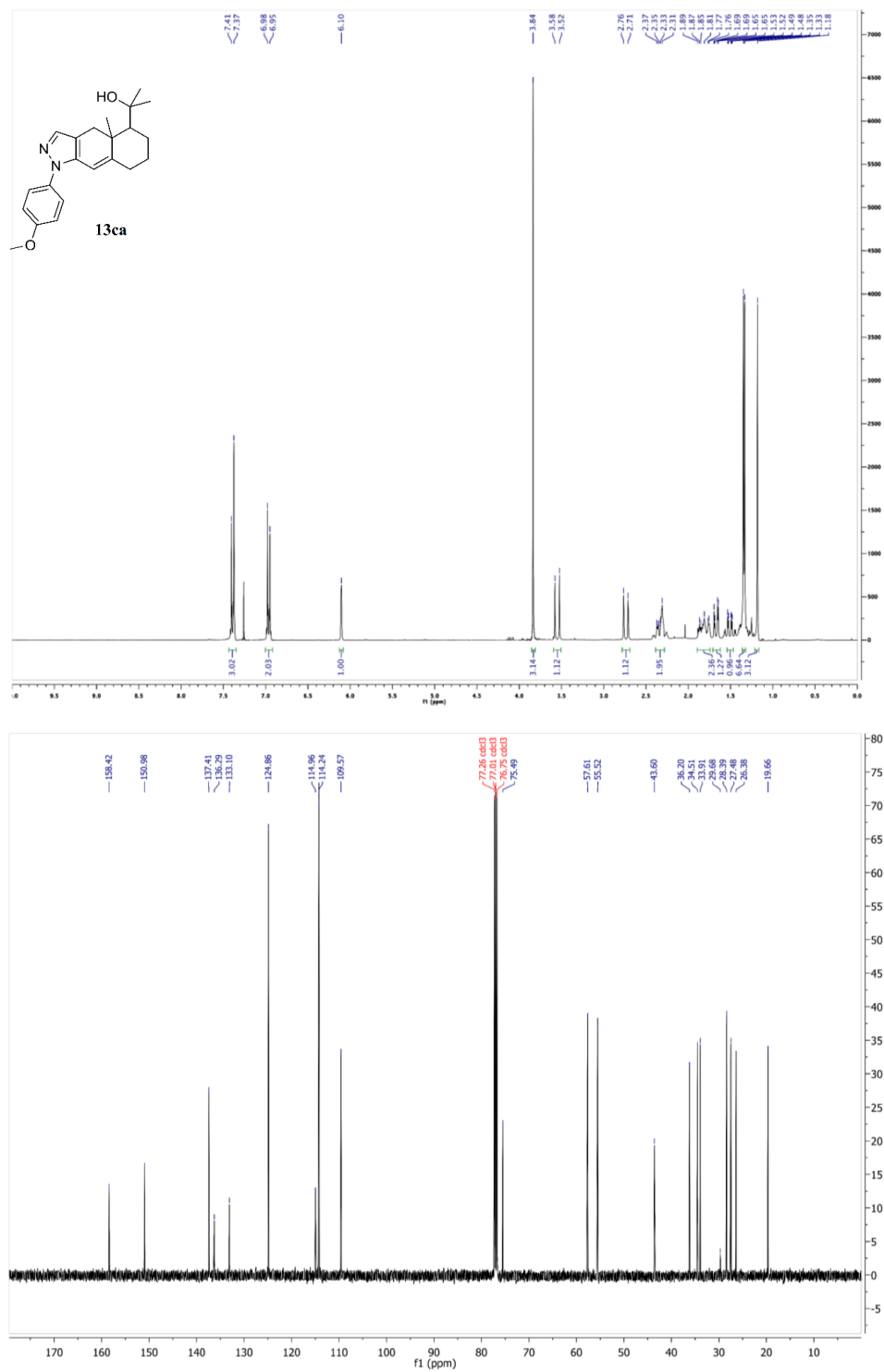
¹H-NMR compound **10c**



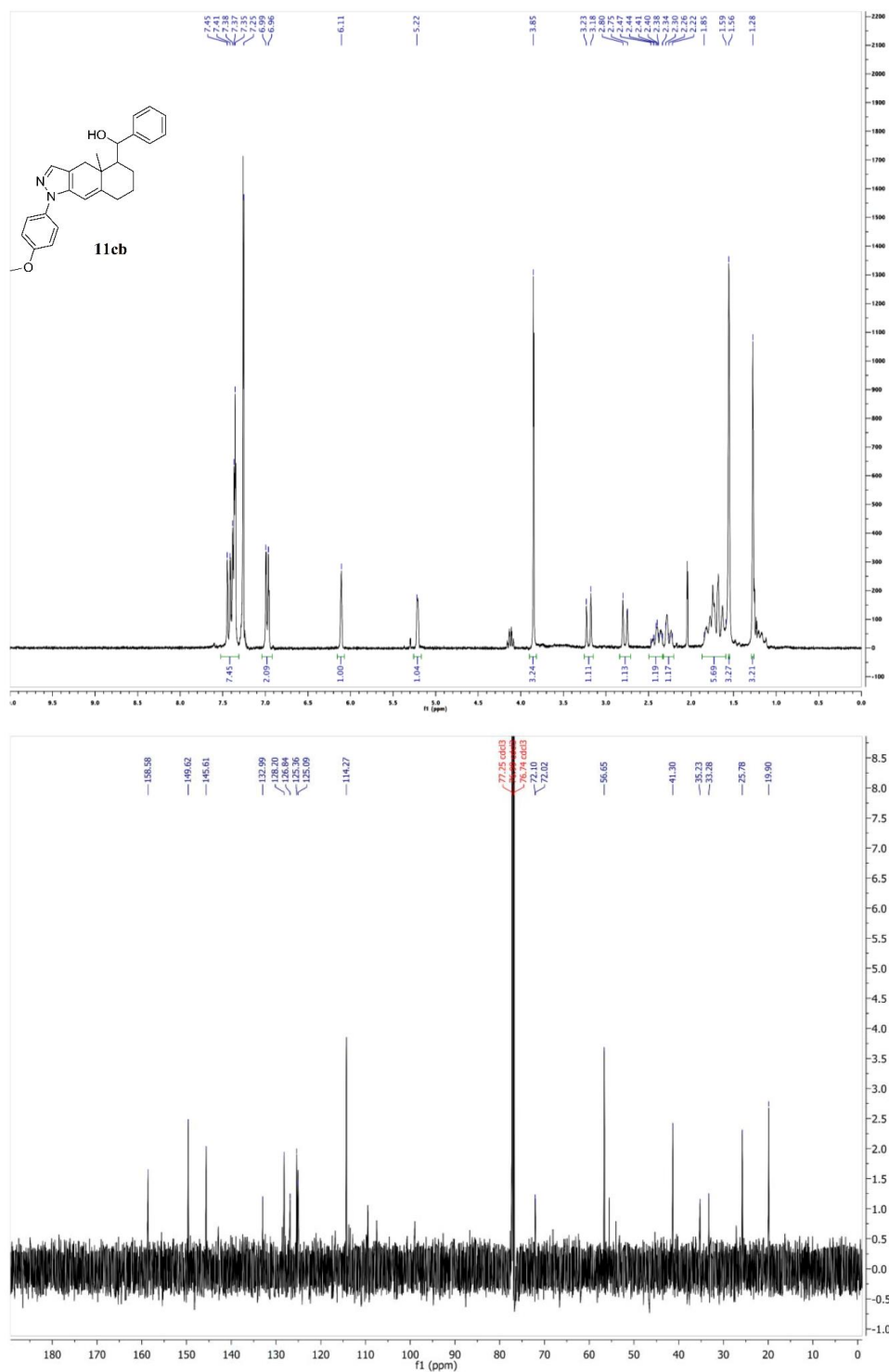
^1H -NMR and ^{13}C -NMR for compound **11ca**



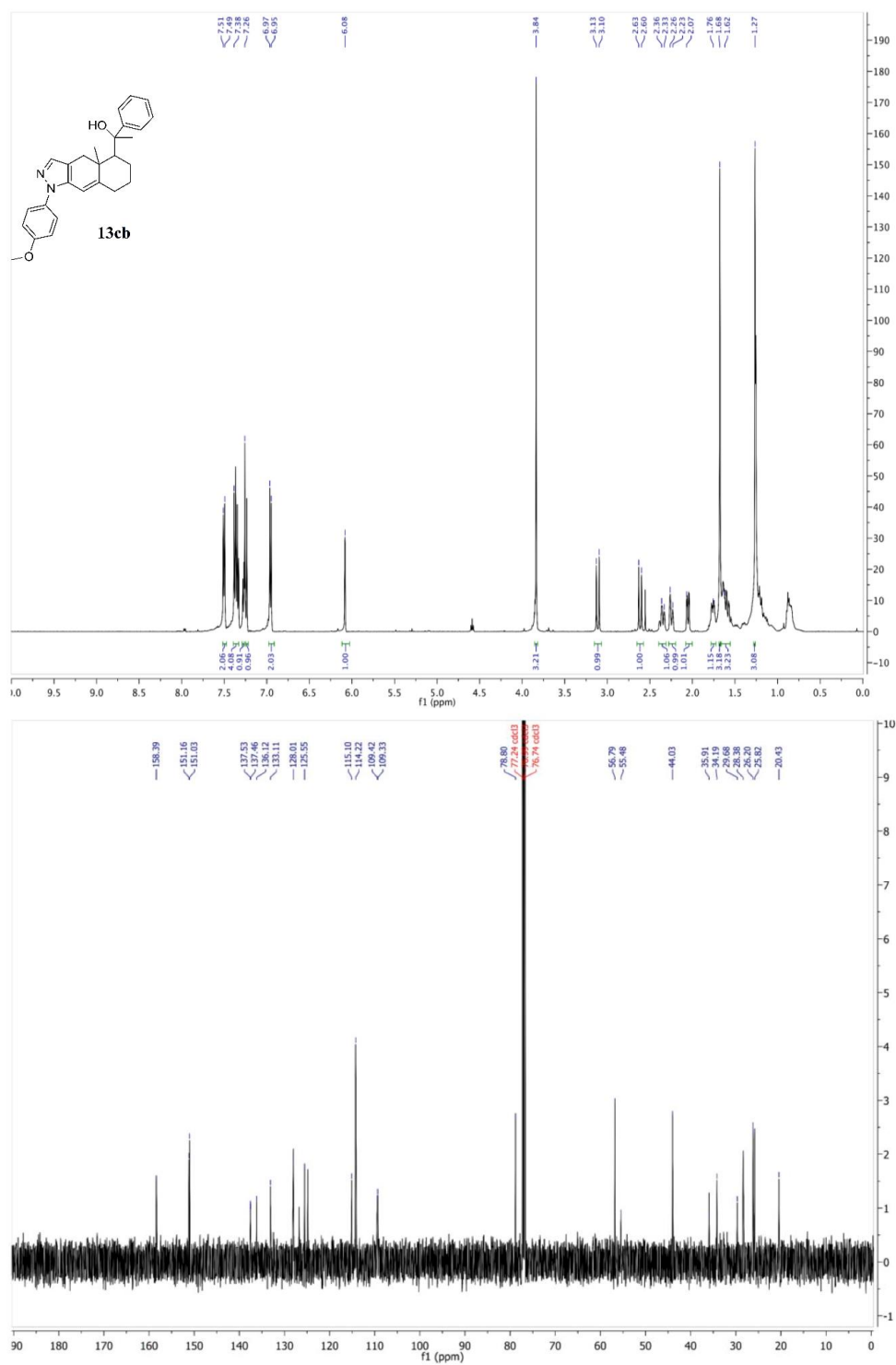
^1H -NMR and ^{13}C -NMR for compound **13ca**



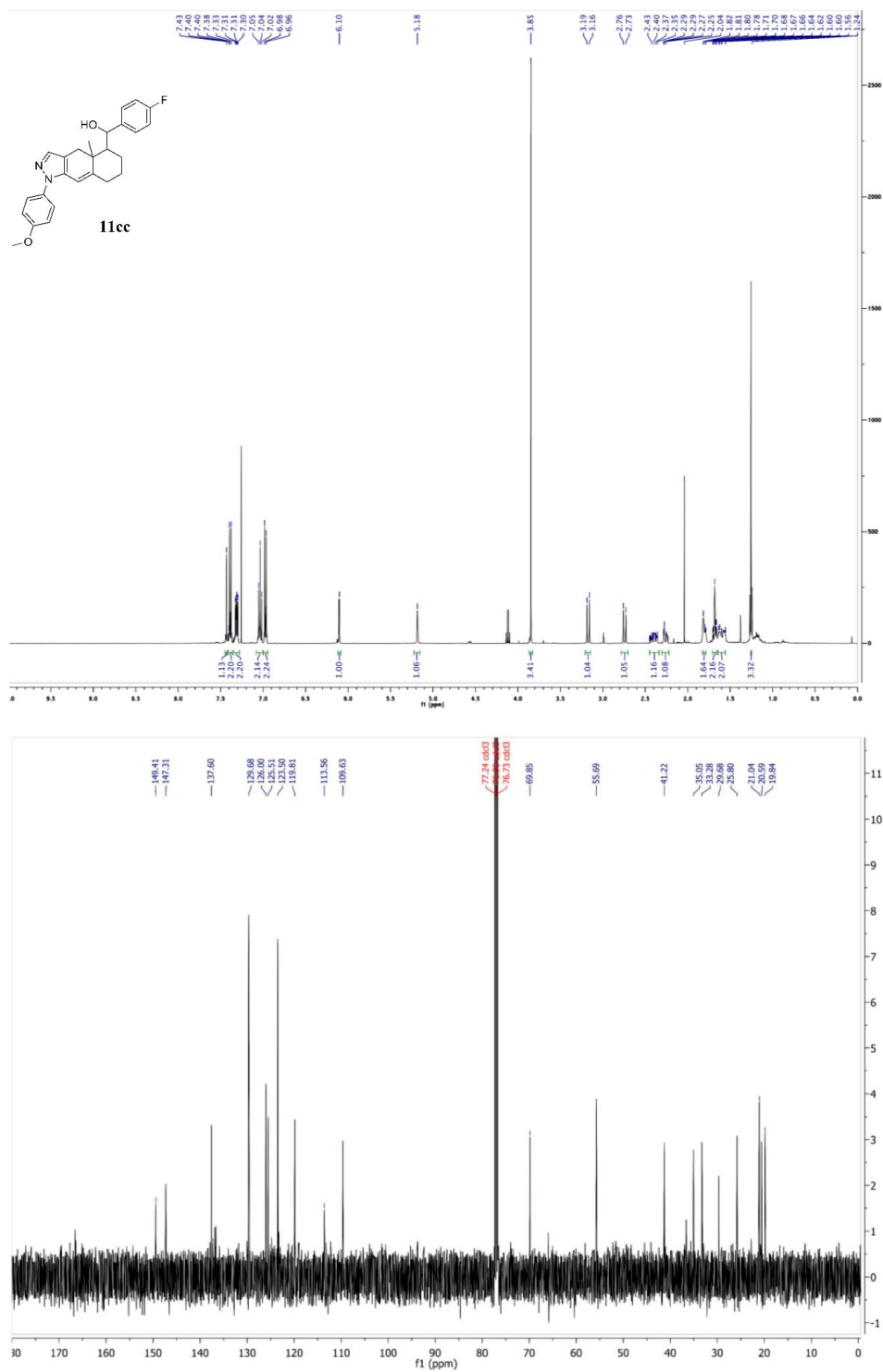
^1H -NMR and ^{13}C -NMR for compound **11cb**



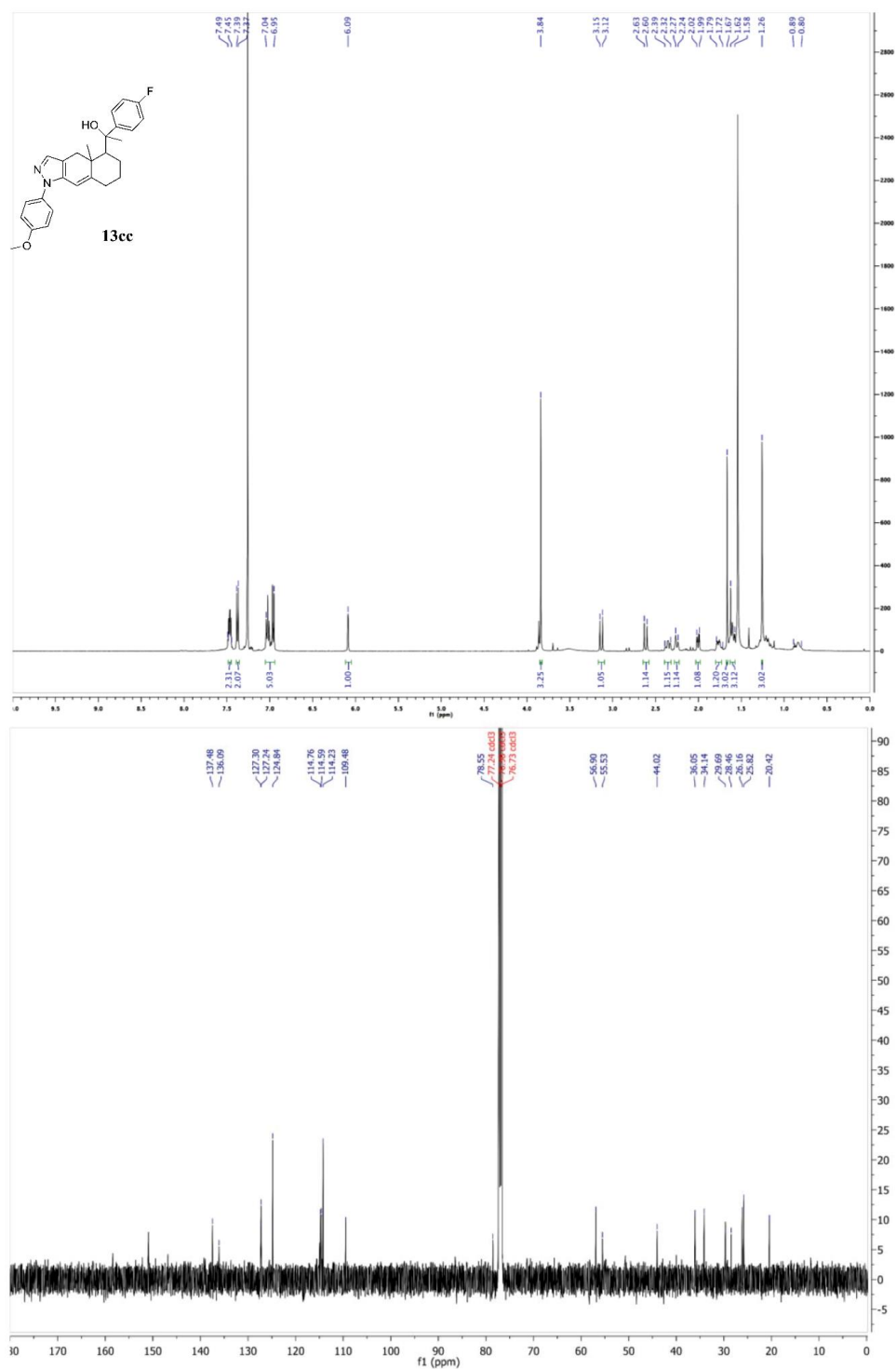
^1H -NMR and ^{13}C -NMR for compound **13cb**

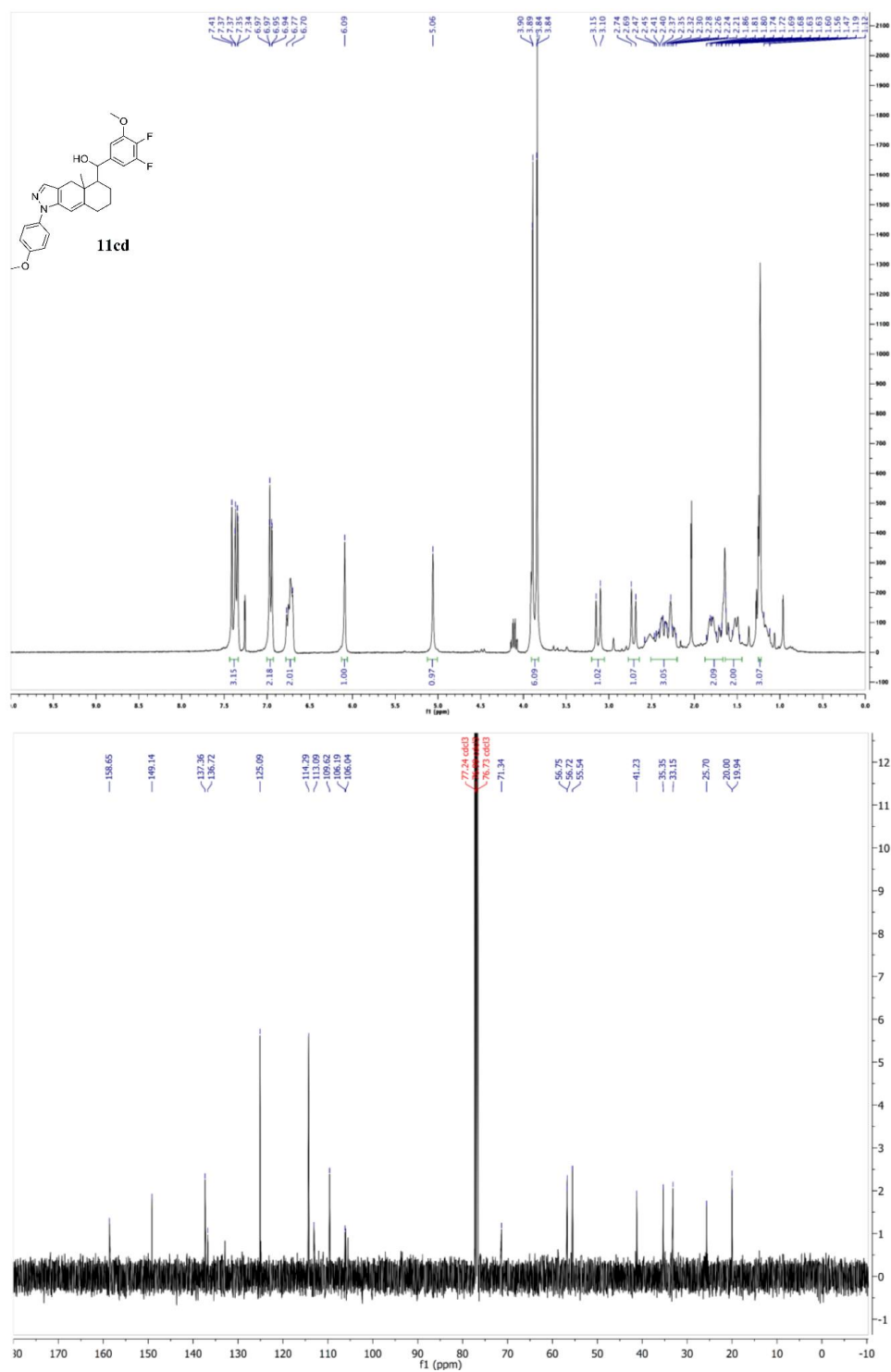


^1H -NMR and ^{13}C -NMR for compound **11cc**

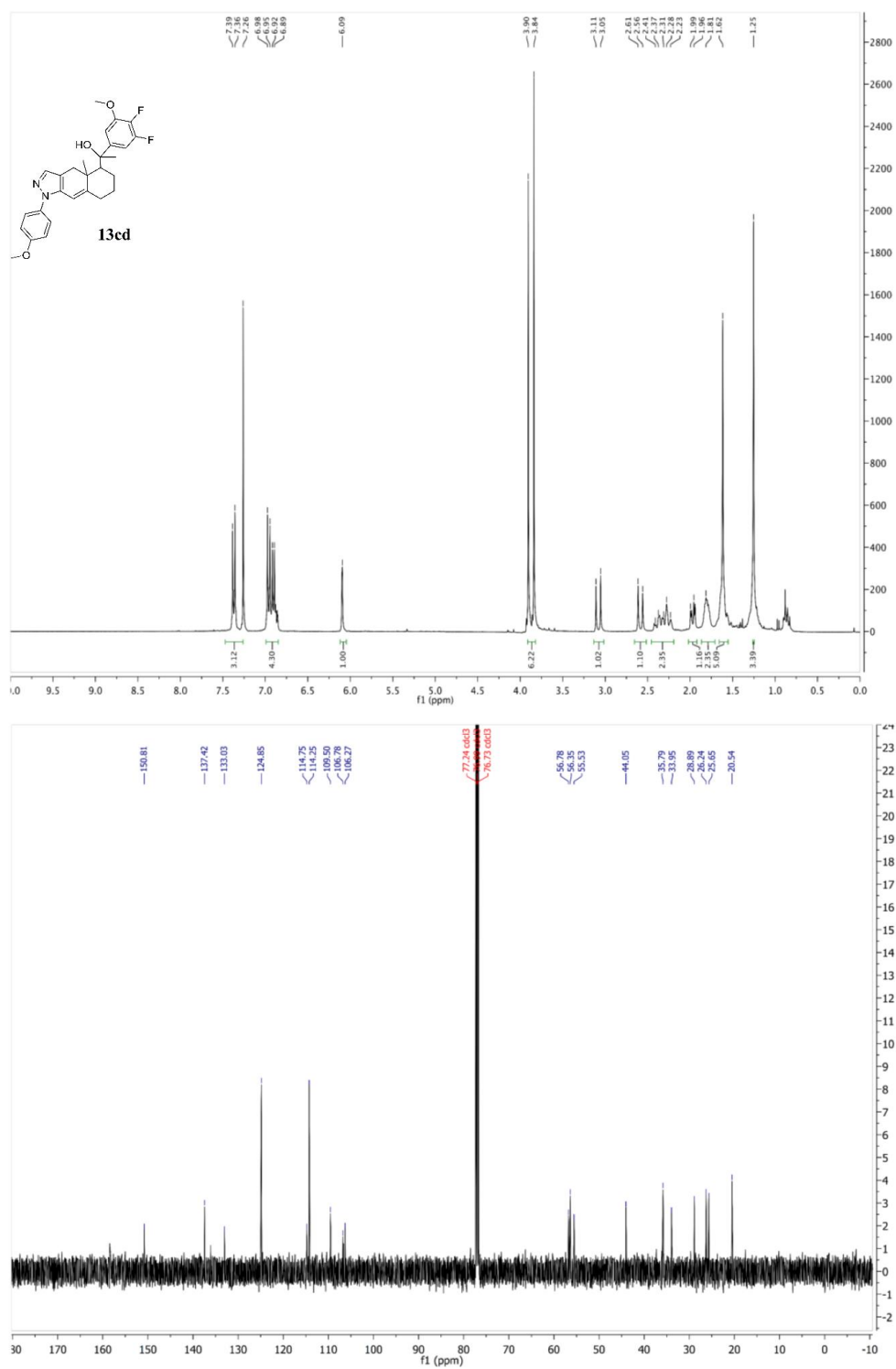


^1H -NMR and ^{13}C -NMR for compound **13cc**

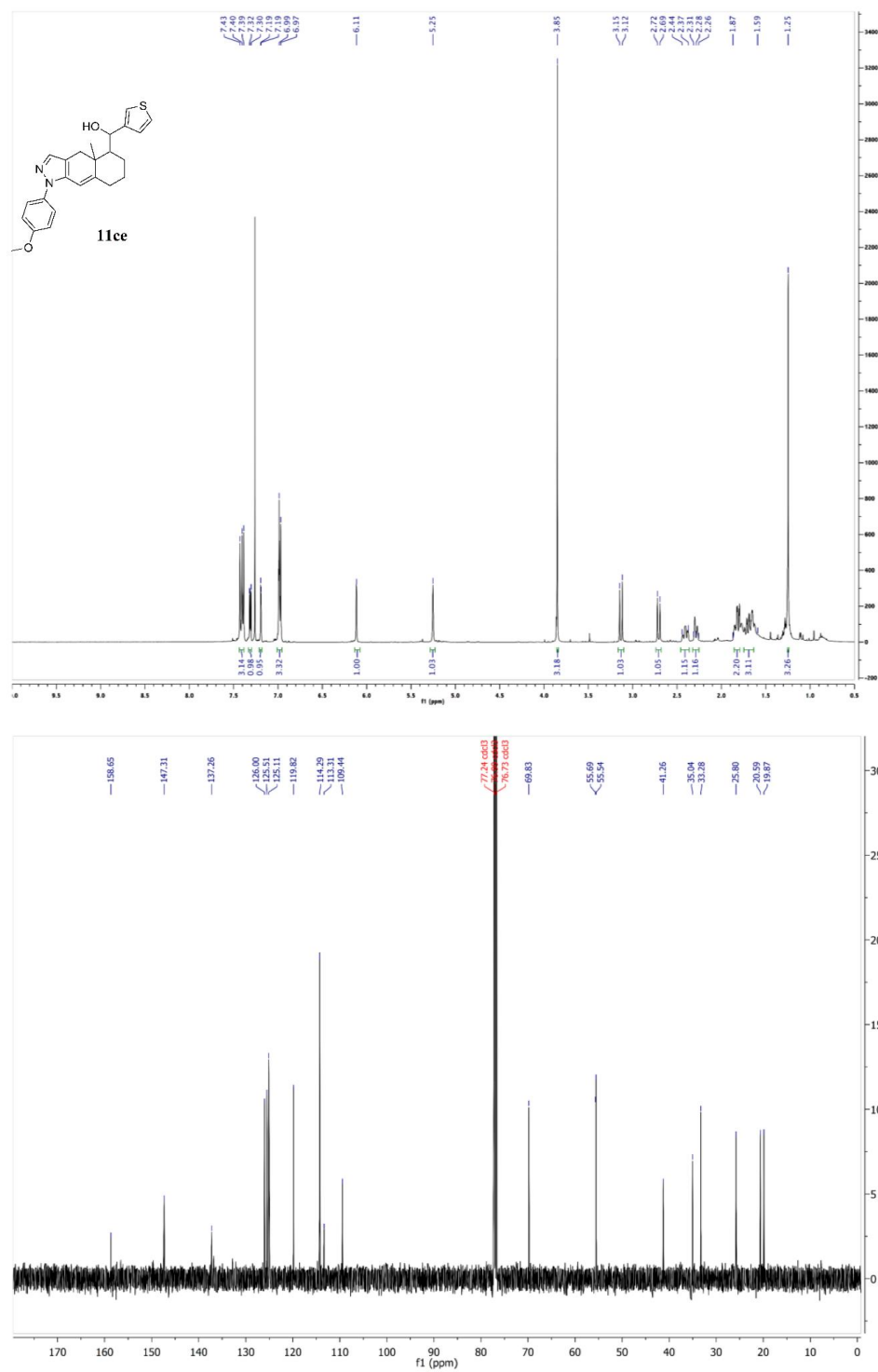


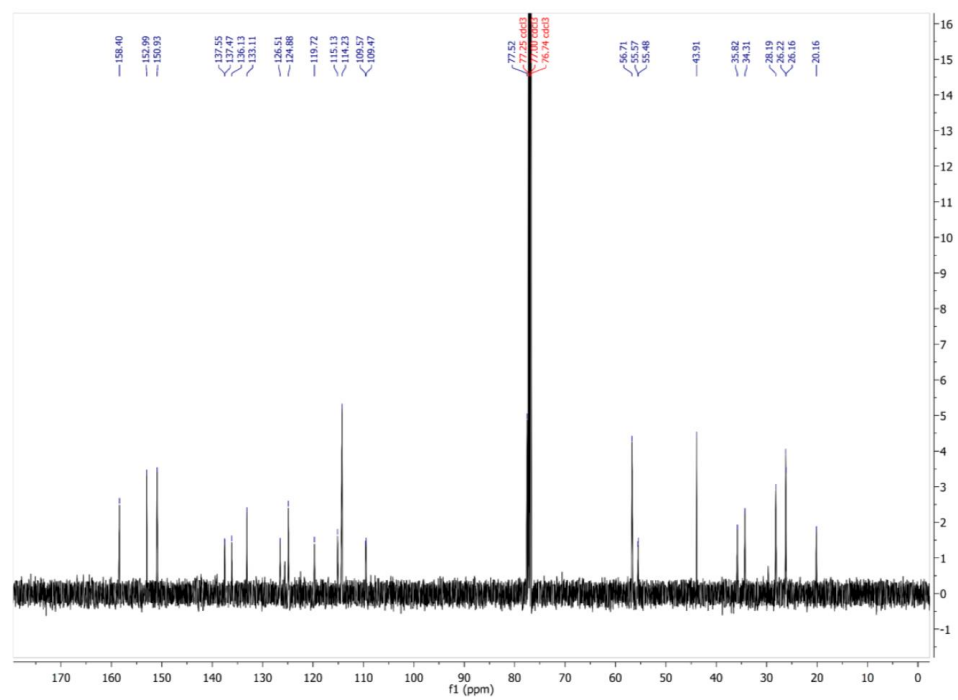
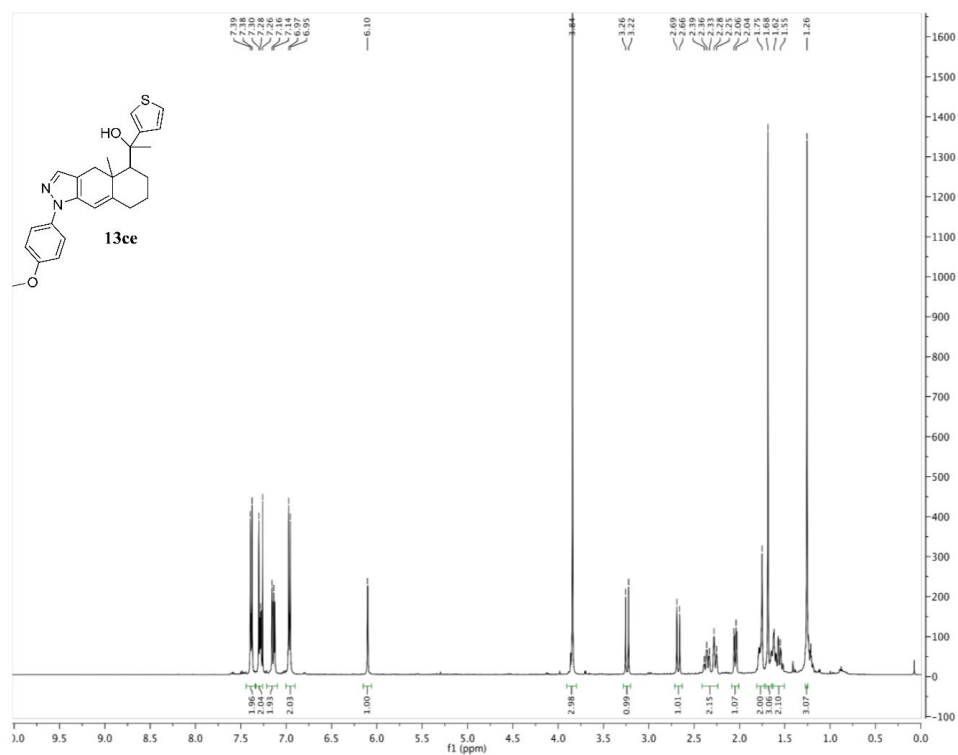
^1H -NMR and ^{13}C -NMR for compound **11cd**

^1H -NMR and ^{13}C -NMR for compound **13cd**



^1H -NMR and ^{13}C -NMR for compound **11ce**

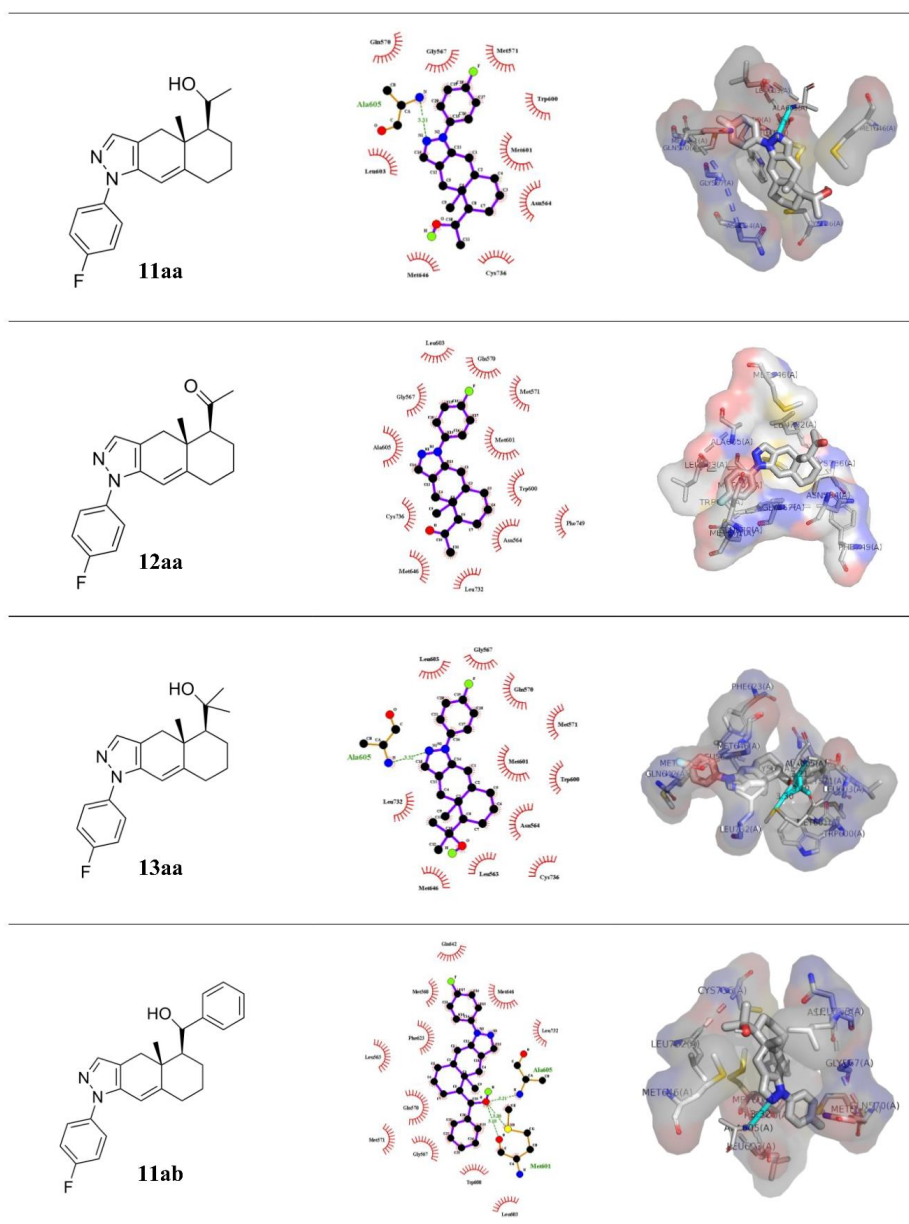


^1H -NMR and ^{13}C -NMR for compound **13ce**

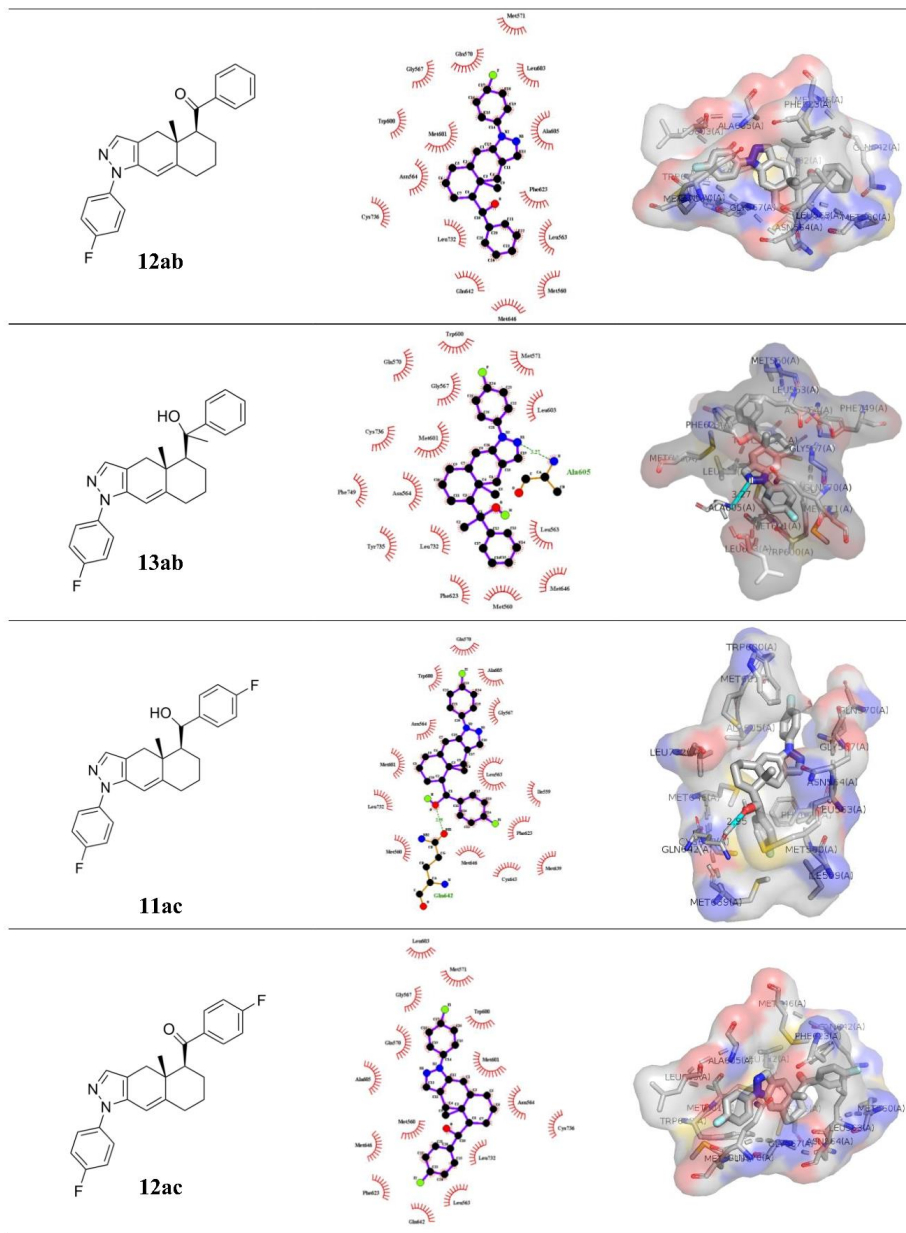
1.8.6 – 2D/3D representations of ligand-receptor interactions

Compounds are shown in the orientation of highest calculated binding affinity, which are shown in **section 1.8.7**.

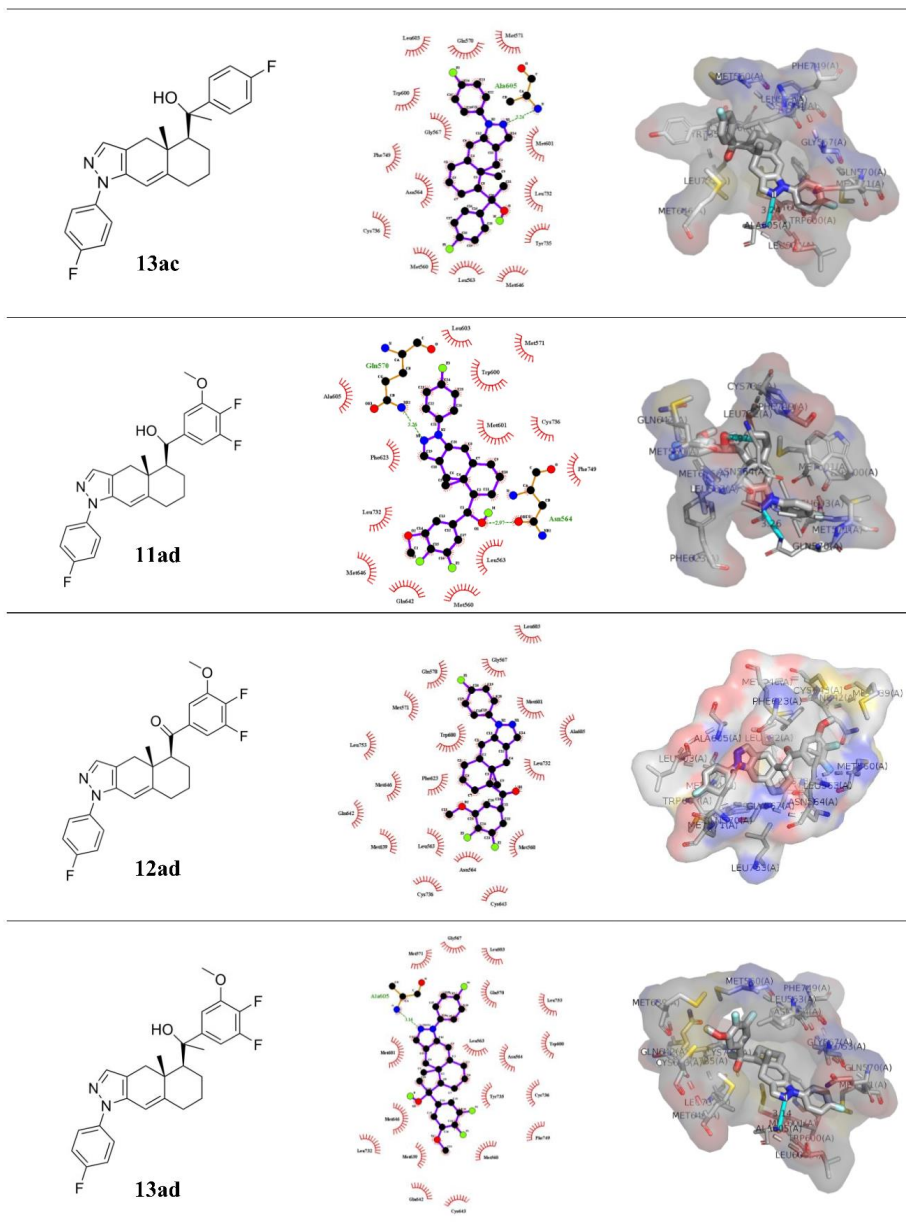
2D/3D ligand-residue interactions for compounds **11aa**, **12aa**, **13aa**, and **11ab**



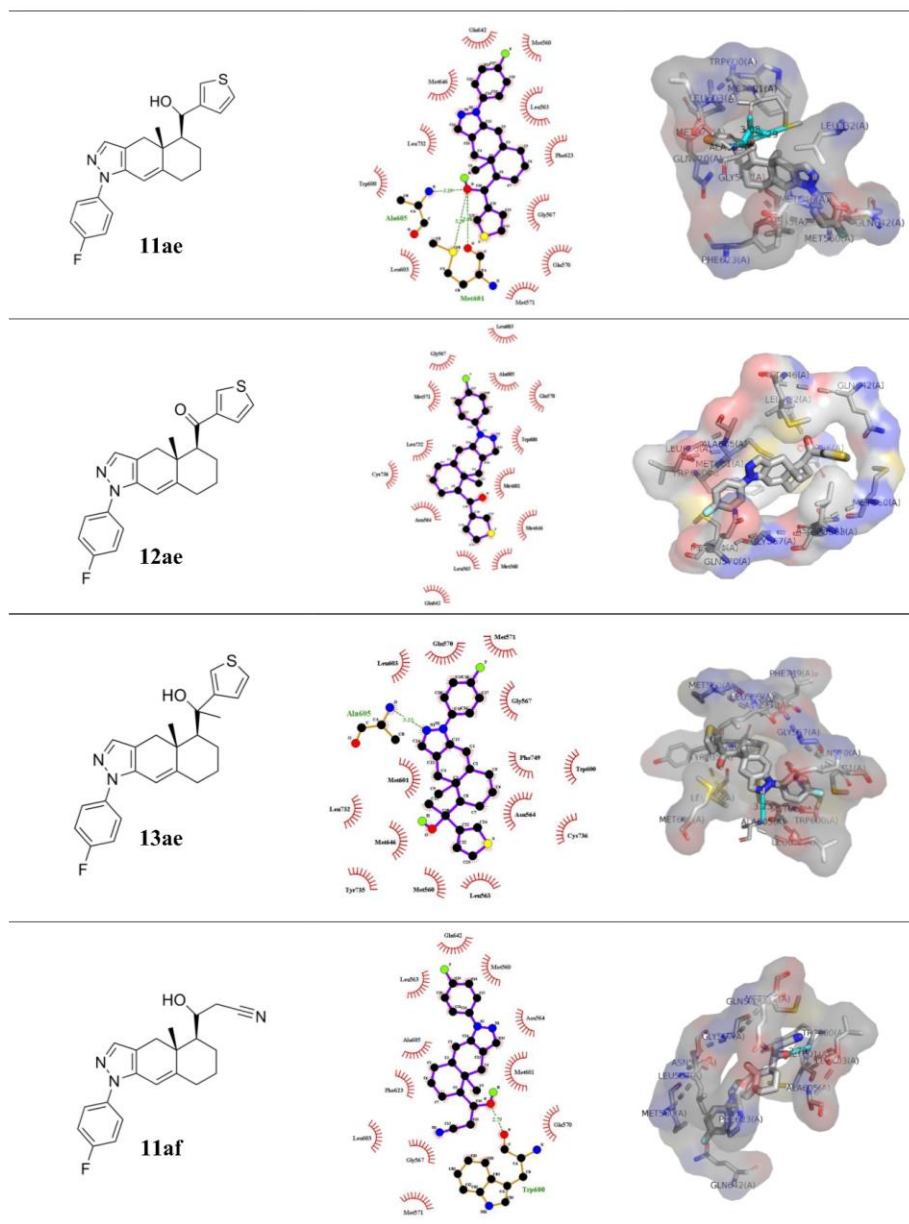
2D/3D ligand-residue interactions for compounds **12ab**, **13ab**, **11ac**, and **12ac**



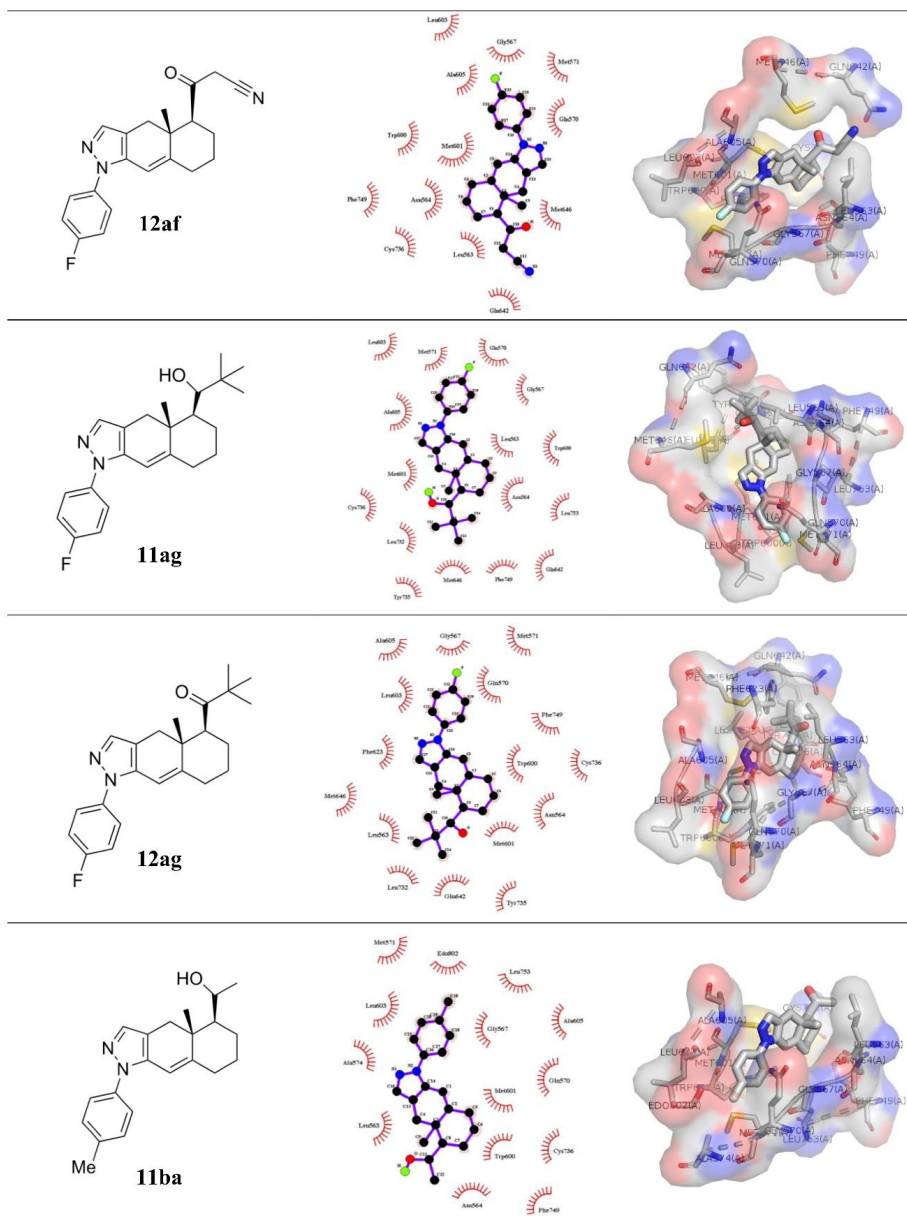
2D/3D ligand-residue interactions for compounds **13ac**, **11ad**, **12ad**, and **13ad**



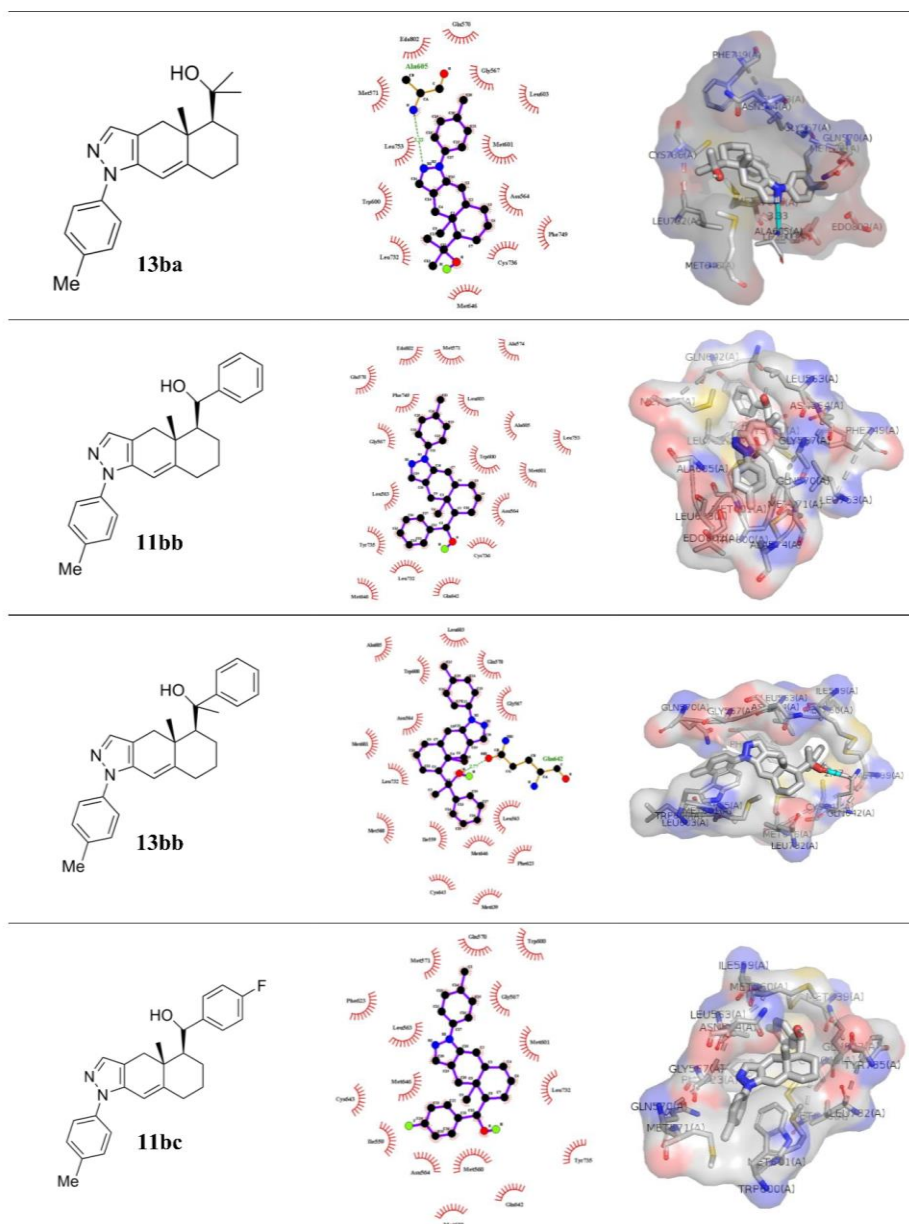
2D/3D ligand-residue interactions for compounds **11ae**, **12ae**, **13ae**, and **11af**

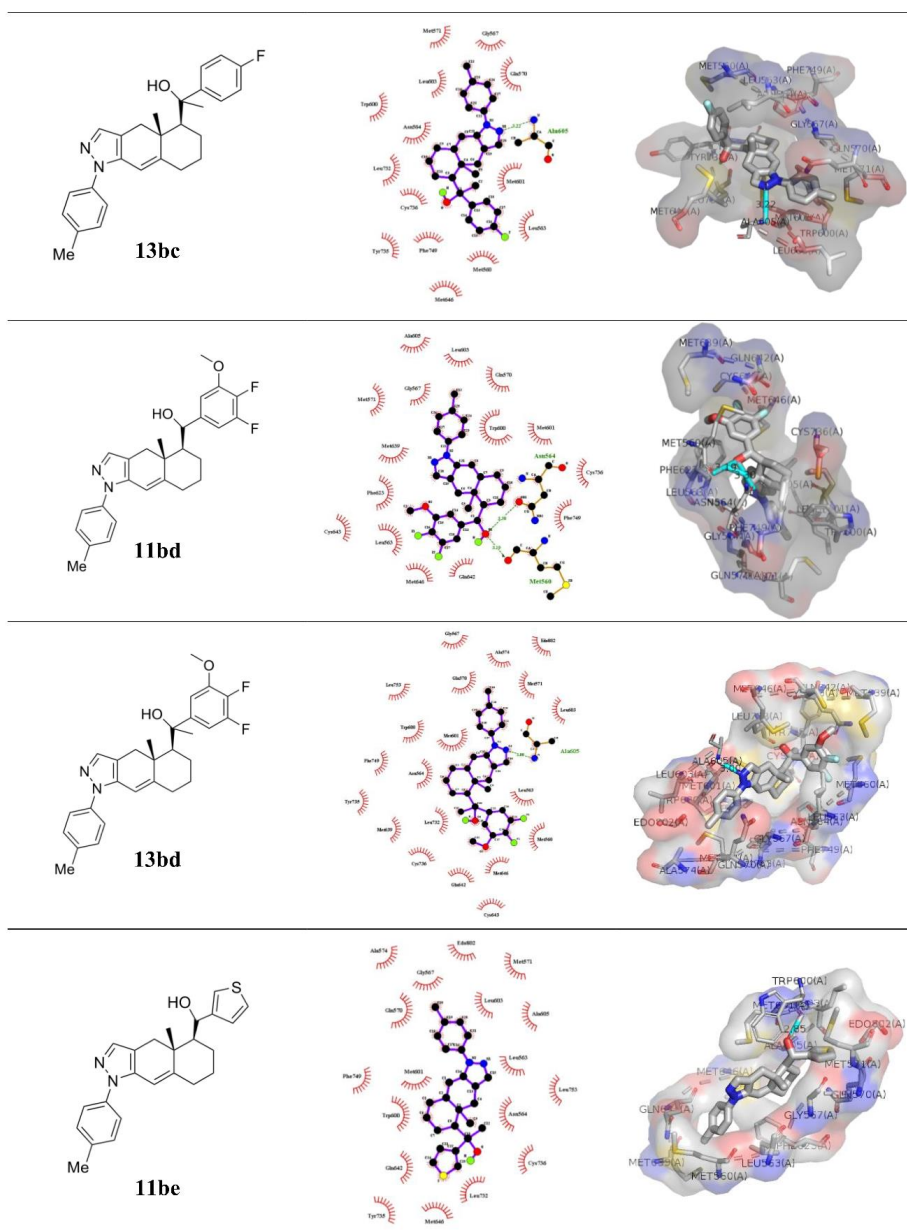


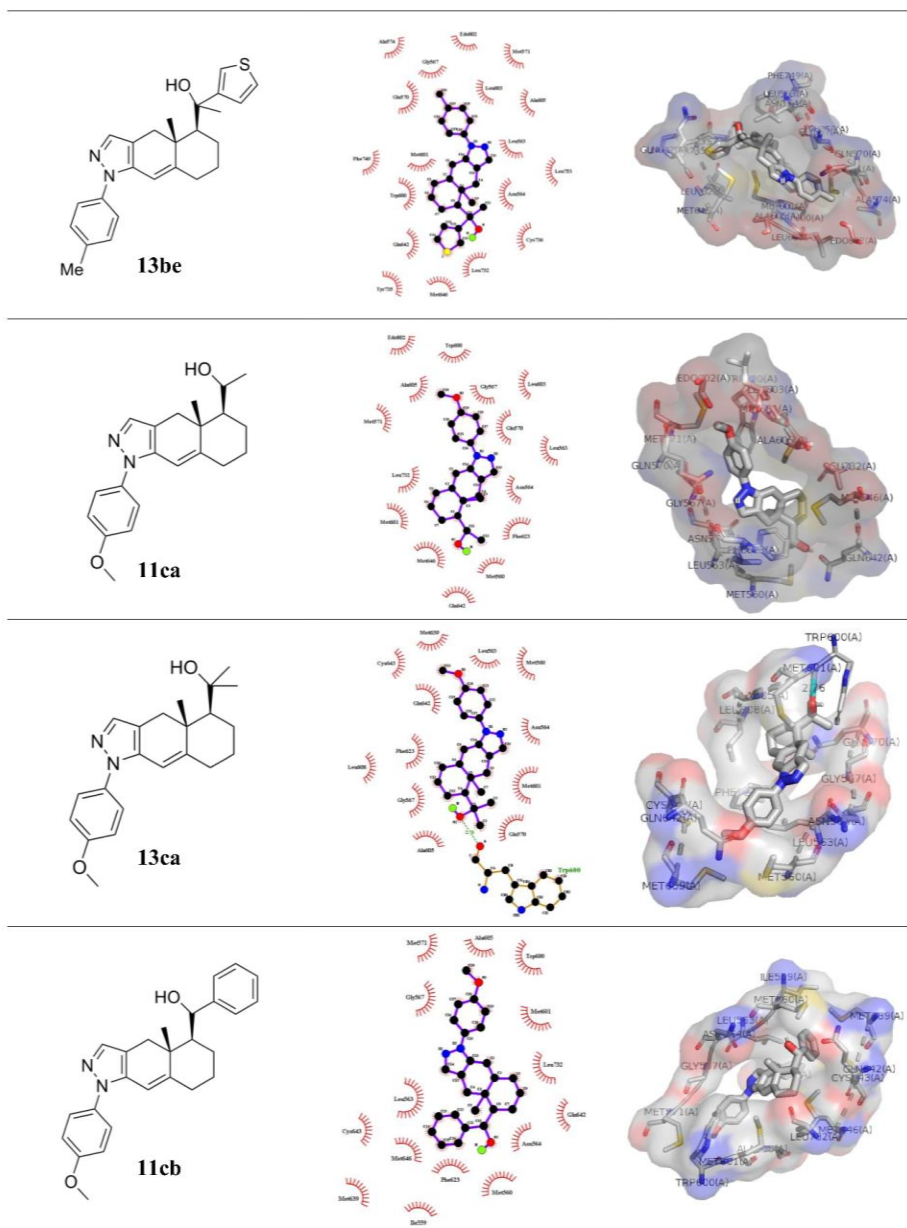
2D/3D ligand-residue interactions for compounds **12af**, **11ag**, **12ag**, and **11ba**

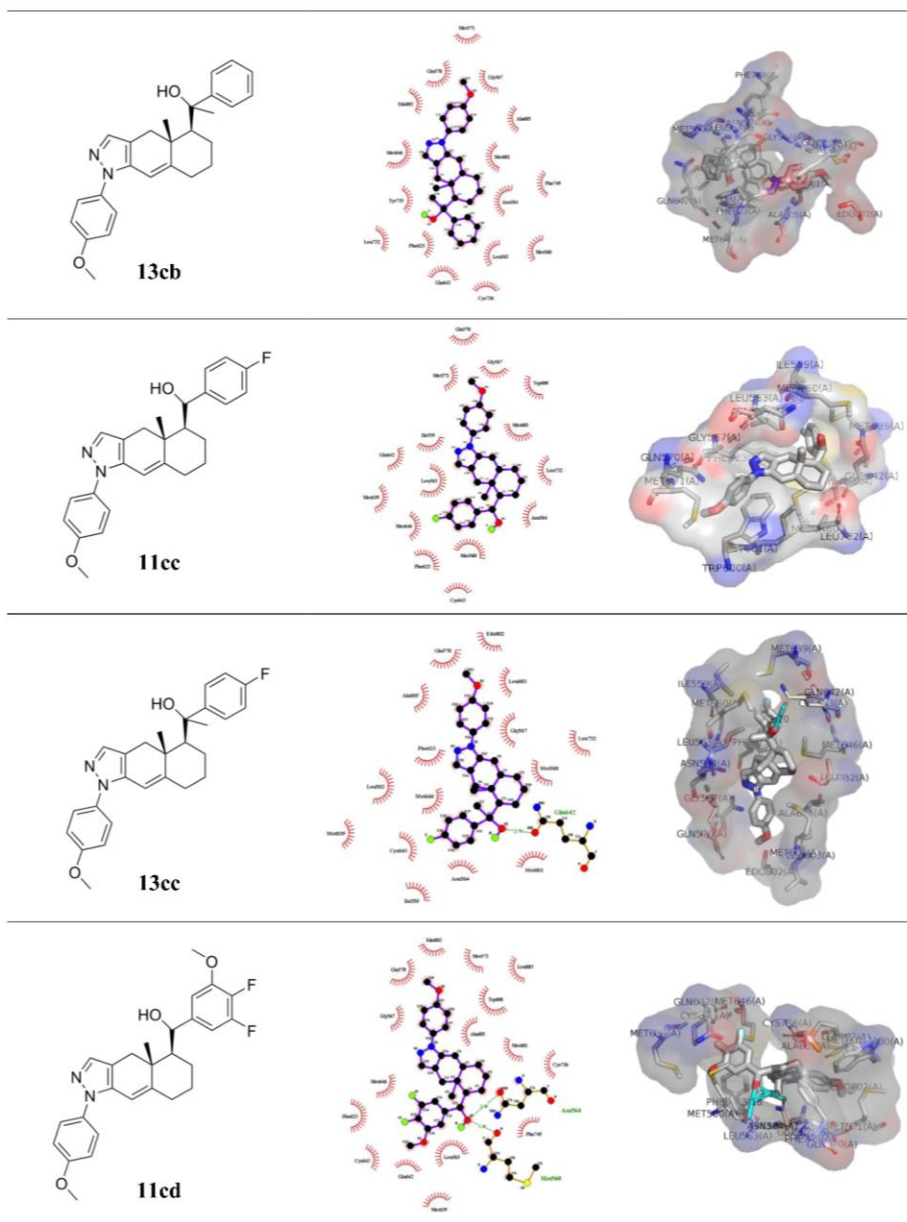


2D/3D ligand-residue interactions for compounds **13ba**, **11bb**, **13bb**, and **11bc**

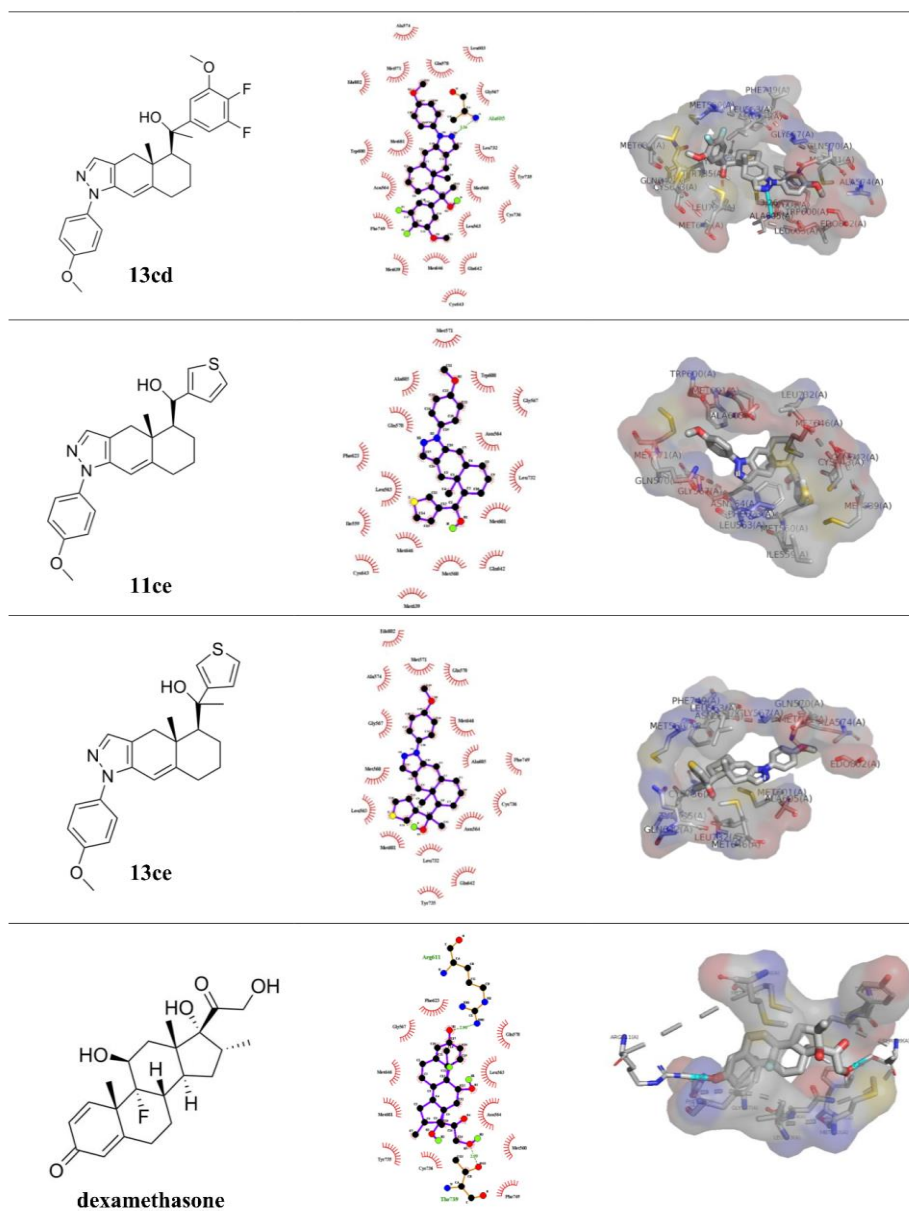


2D/3D ligand-residue interactions for compounds **13bc**, **11bd**, **13bd**, and **11be**

2D/3D ligand-residue interactions for compounds **13be**, **11ca**, **13ca**, and **11cb**

2D/3D ligand-residue interactions for compounds **13cb**, **11cc**, **13cc**, and **11cd**

2D/3D ligand-residue interactions for compounds **13cd**, **11ce**, **13ce**, and **Dex**



1.8.7 – Tables of ligand orientations in 5NFP pocket

Glucocorticoid	Top binding conformations	binding affinity (kcal/mol)	orientation in 5NFP			
			1	2	3	4
dexamethasone	1	-10.6		X		
	2	-9.0	X			
	3	-8.7			X	
	4	-8.7	X			
11aa	1	-9.5	X			
	2	-9.4	X			
	3	-9.3				X
	4	-9.0				X
13aa	1	-10.0	X			
	2	-9.9				X
	3	-9.7	X			
	4	-9.6	X			
11ab	1	-10.4				X
	2	-10.0	X			
	3	-9.7	X			
	4	-9.7	X			
13ab	1	-10.9	X			
	2	-10.4	X			
	3	-10.1				X
	4	-10.1	X			
11ac	1	-10.6	X			
	2	-10.1	X			
	3	-9.5				X
	4	-9.5				X
13ac	1	-11.3	X			
	2	-10.5	X			
	3	-10.4	X			
	4	-10.2				X
11ad	1	-9.6	X			
	2	-9.5	X			
	3	-7.4				X
	4	-7.2	X			
13ad	1	-11.2	X			
	2	-9.7	X			
	3	-9.7	X			
	4	-8.5				X

Tables of ligand orientations in 5NFP pocket (continued)

Glucocorticoid	Top binding conformations	binding affinity (kcal/mol)	orientation in 5NFP			
			1	2	3	4
11ae	1	-10.0				X
	2	-9.8	X			
	3	-9.2	X			
	4	-9.1	X			
13ae	1	-10.3	X			
	2	-10.3	X			
	3	-9.8				X
	4	-9.8	X			
11af	1	-9.6				X
	2	-9.5	X			
	3	-9.2				X
	4	-9.1	X			
11ag	1	-10.4	X			
	2	-9.7	X			
	3	-9.5	X			
	4	-9.4	X			
11ba	1	-9.3	X			
	2	-9.3				X
	3	-9.1	X			
	4	-8.9	X			
13ba	1	-9.8	X			
	2	-9.7				X
	3	-9.5	X			
	4	-9.4	X			
11bb	1	-10.5	X			
	2	-10.2				X
	3	-10.2	X			
	4	-10.1				X
13bb	1	-10.6	X			
	2	-10.2	X			
	3	-9.5	X			
	4	-8.6	X			
11bc	1	-10.3	X			
	2	-10.2				X
	3	-10.2	X			
	4	-9.6	X			

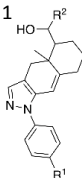
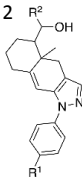
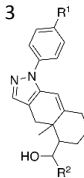
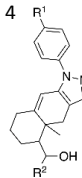
Tables of ligand orientations in 5NFP pocket (continued)

Glucocorticoid	Top binding conformations	binding affinity (kcal/mol)	orientation in 5NFP			

Tables of ligand orientations in 5NFP pocket (continued)

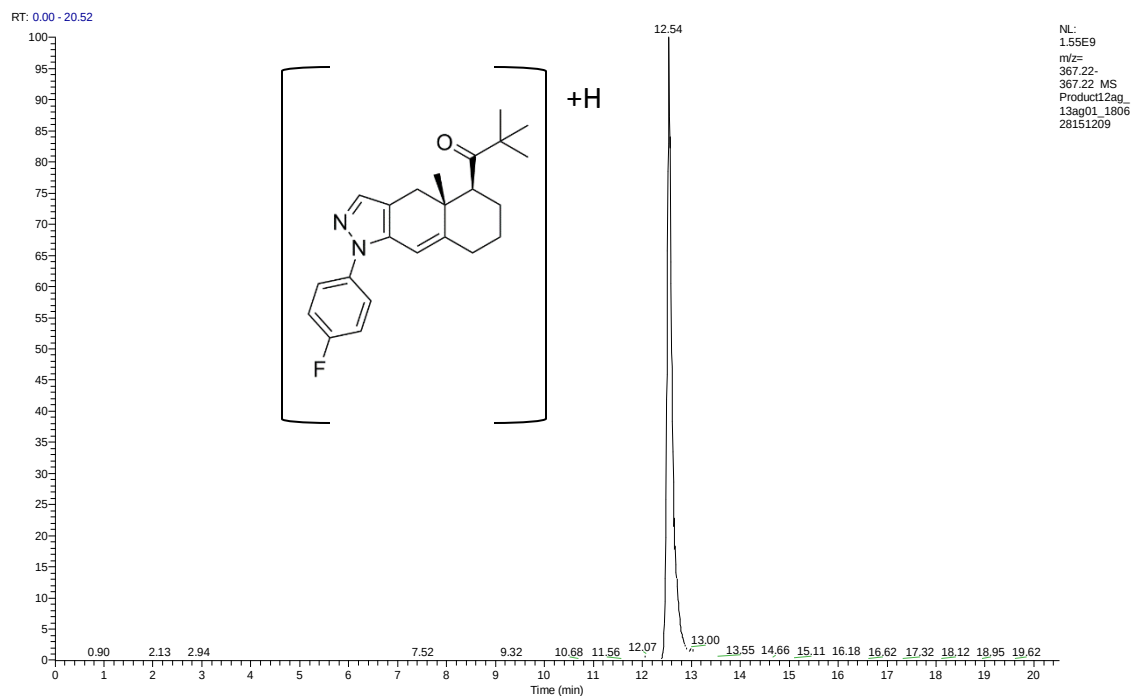
Glucocorticoid	Top binding conformations	binding affinity (kcal/mol)	orientation in 5NFP			
			1	2	3	4
11cc	1	-10.2	X			
	2	-9.6	X			
	3	-9.5	X			
	4	-9.2	X			
13cc	1	-11	X			
	2	-10.4	X			
	3	-9.2	X			
	4	n/a				
11cd	1	-9.4	X			
	2	-8.6	X			
	3	-6.9		X		
	4	n/a				
13cd	1	-10.1	X			
	2	-9.6	X			
	3	-9.3	X			
	4	n/a				
11ce	1	-9.3	X			
	2	-9.3	X			
	3	-9.1				X
	4	-8.8	X			
13ce	1	-9.8	X			
	2	-9.7	X			
	3	-9.5	X			
	4	-9.2	X			
12aa	1	-9.9	X			
	2	-9.7	X			
	3	-9.6	X			
	4	-9.3				X
12ab	1	-11.1	X			
	2	-10.5	X			
	3	-10.2	X			
	4	-10.0				X
12ac	1	-11.5	X			
	2	-10.6	X			
	3	-10.5				X
	4	-10.3	X			

Tables of ligand orientations in 5NFP pocket (continued)

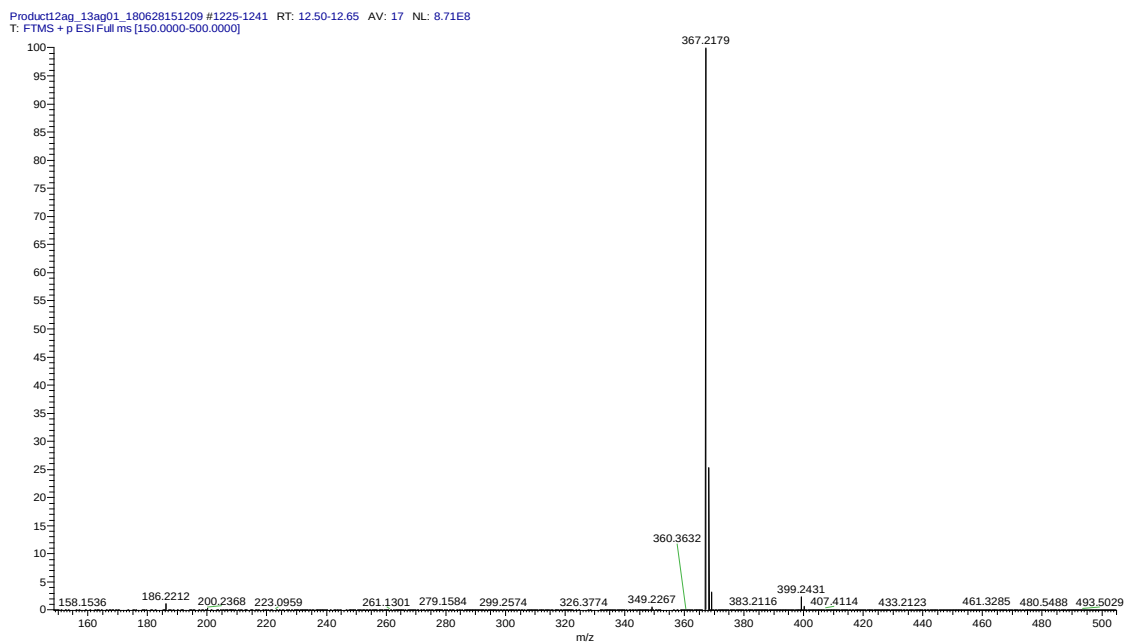
Glucocorticoid	Top binding conformations	binding affinity (kcal/mol)	orientation in 5NFP			
			1 	2 	3 	4 
12ad	1	-11.0	X			
	2	-9.9	X			
	3	-8.6				X
	4	-8.5				X
12ae	1	-10.6	X			
	2	-10.3	X			
	3	-9.8	X			
	4	-9.8				X
12af	1	-10.1	X			
	2	-9.8	X			
	3	-9.7				X
	4	-9.5				X
12ag	1	-10.3	X			
	2	-9.8	X			
	3	-9.8				X
	4	-9.5	X			

1.8.8 – EICs and LCMS method for analysis of crude product 13ag

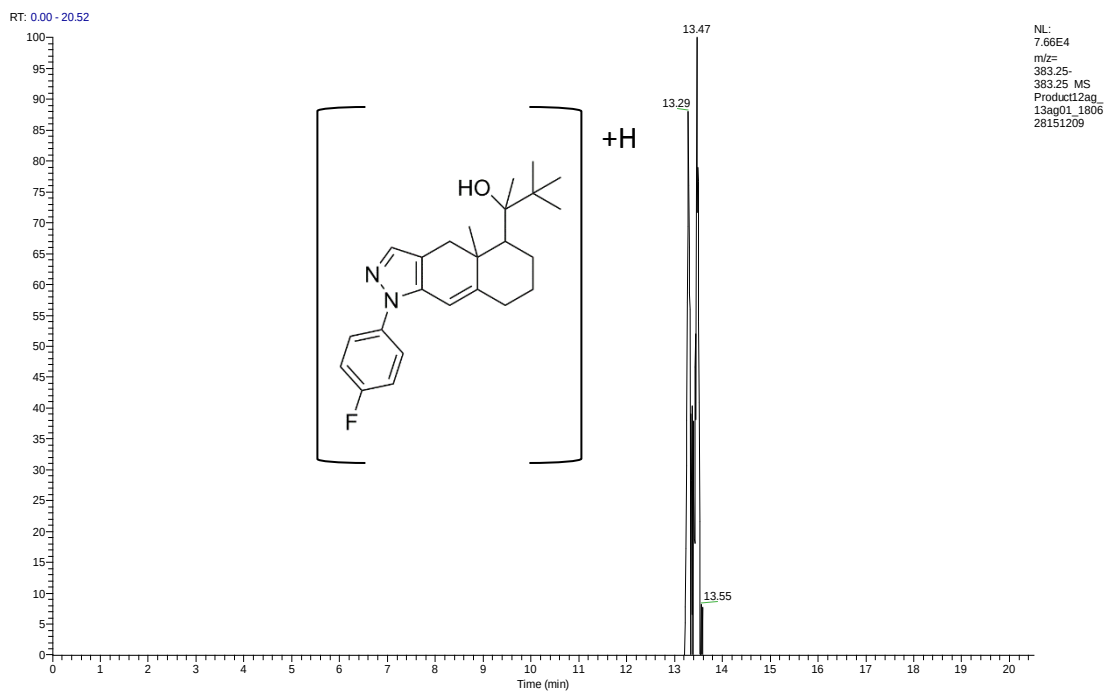
EIC – [M+H]⁺ 12ag



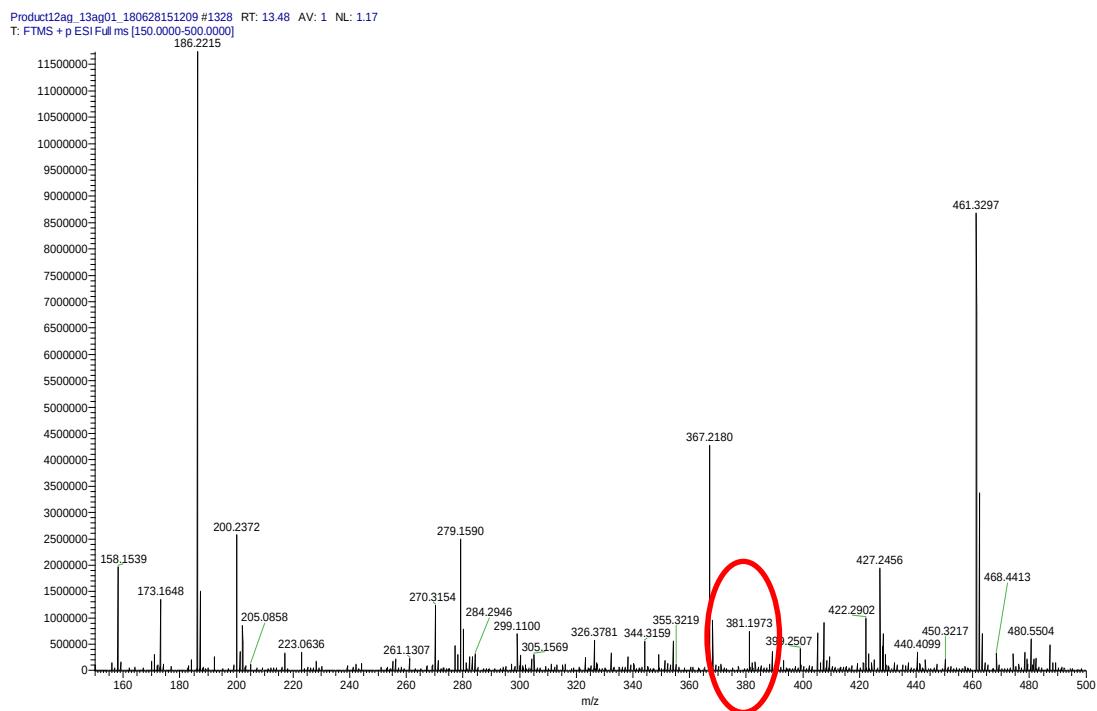
[M+H]⁺ calc'd for C₂₃H₂₈FN₂O⁺ = 367.2180 m/z; found – 367.2179 m/z; Δppm = -0.27



EIC – [M+H]⁺ 13ag

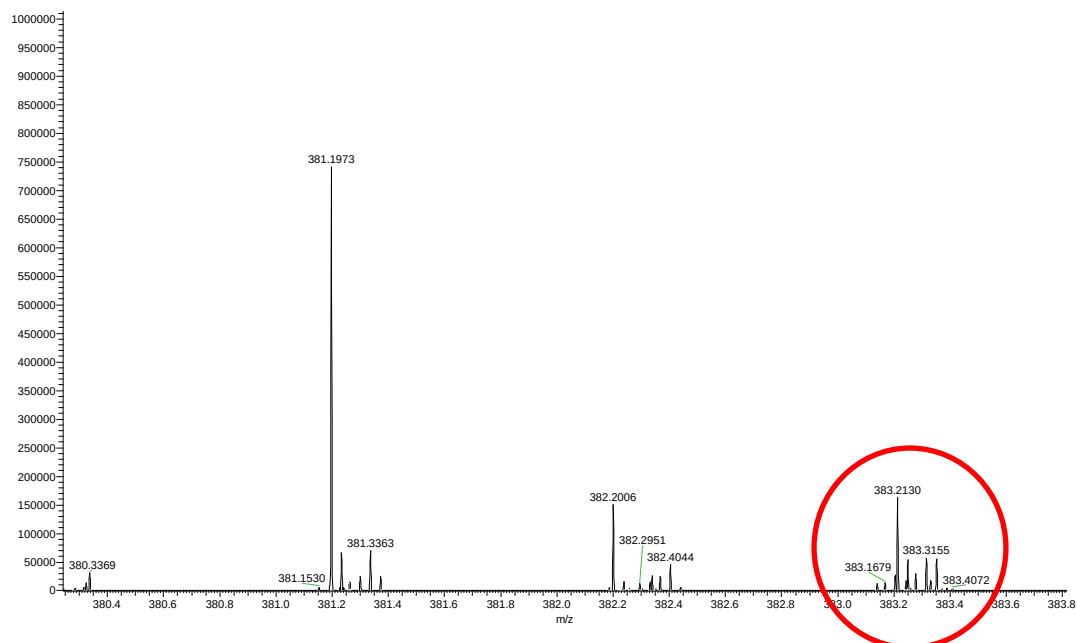


Mass spectrum under EIC of 13ag



Zoomed-in mass spectrum under **13ag** peak

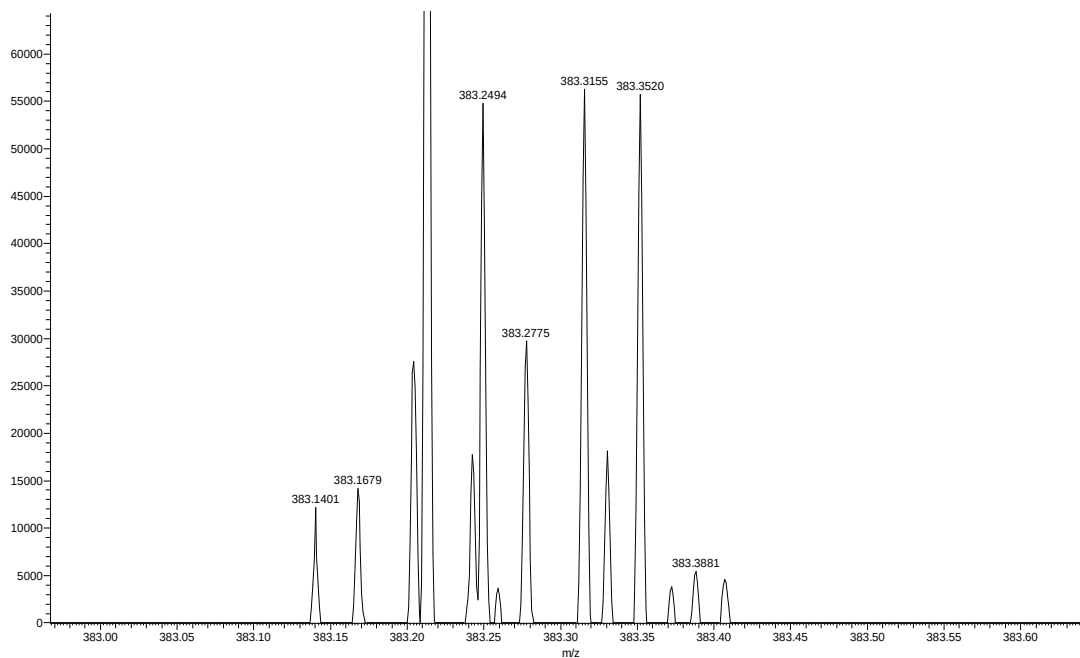
Product12ag_13ag01_180628151209 #1328 RT: 13.48 AV: 1 NL: 7.41"
T: FTMS + p ESI Full ms [150.0000-500.0000]



$[M+H]^+$ calc'd for $C_{24}H_{32}FN_2O^+$ = 383.2493 m/z; found – 383.2494 m/z; Δ ppm = 0.21

Relative abundance of **13ag** to **12ag** = 1.55×10^9 ions **12ag** : 5.5×10^4 ions **13ag**

Product12ag_13ag01_180628151209 #1328 RT: 13.48 AV: 1 NL: 1.63F6
T: FTMS + p ESI Full ms [150.0000-500.0000]



CHAPTER II

SYNTHETIC LIPIDS AS ISOTOPE TRACERS TO DIFFERENTIATE BRANCHED-CHAIN LIPIDS FROM THEIR STRAIGHT-CHAIN ISOBARS IN LCMS ANALYSES.

Portions of this chapter will be included in a future publication with a working title of:
Utility of (D₃)-isovaleric acid as a stable isotope tracer for differentiating isolipids from straight-chain isobars in LC-ESI-MS lipidomic analyses.

Brandon J. Kennedy, Liz Fozo, Shawn R. Campagna

2.1 – Abstract

Absolute identification of individual lipid species is a challenge in lipidomic analyses that is often dealt with by clustering groups of lipids by class, chain length, and degree of unsaturation (**Figure 2.1.1**). This is an issue when considering that isobaric lipids of the same class may have differing activities *in vivo* due to the structural differences between them. An example of these co-identified lipid species in lipidomic analyses employing liquid chromatography electrospray ionization mass spectrometry (LC-ESI-MS)

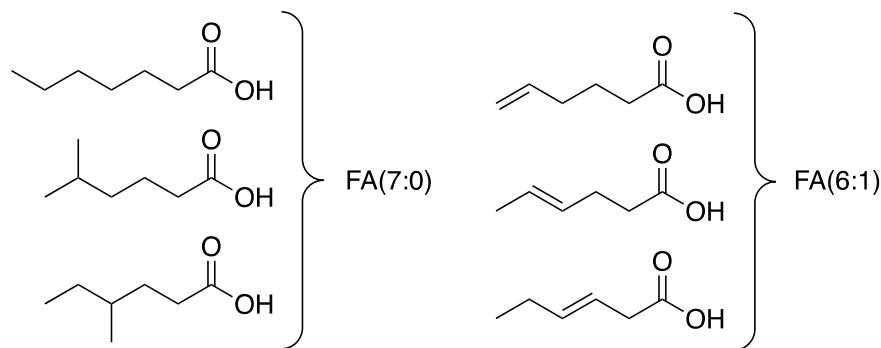


Figure 2.1.1 – Lipid Isobars with the same compound abbreviation

Lipids are typically annotated with a nomenclature that does not specify tail branching or unsaturation by location, but only indicates the number of carbons and degree of unsaturation

are straight-chain lipids with their respective isobaric counterpart containing a terminal isopropyl moiety. These isolipids are impossible to differentiate from straight-chain lipids since they have the same exact mass and very similar retention profiles in LC separations. This type of lipid isobar is easy to identify by routine gas chromatography electron impact mass spectrometric (GC-EI-MS) analysis of these compounds for a couple of reasons. Due to the higher theoretical plate count in typical gas chromatography versus liquid chromatography, there is greater physical separation of these species. Also, by using a hard ionization source like EI, the analytes are detected as characteristic fragments that illuminate their structural differences.⁵²

In order to make many larger lipids and lipid species with bulky headgroups fly in a gas chromatograph, they must be derivatized first, either into silyl ethers or into methyl esters as with fatty acid methyl ester (FAME) analysis. There are other derivitization methods specifically used for isolipid differentiation such as creating dimethyloxazolinyl or picolinyl esters (**Figure 2.1.2**); although these procedures are useful for differentiating branched chain lipids from straight-chain lipids, they require chemistries more complicated than methylation or silylation and are typically not pursued.⁵³⁻⁵⁴ The biggest downside to these procedures is the loss of connectivity with head groups, which makes it impossible to correlate what higher-order lipids these fatty acid tails were component to. Biologically relevant lipid species can be analyzed by mass spectrometry by softer ionization sources such as ESI, which keeps the biogenic lipid intact throughout liquid chromatographic (LC) separation and subsequent ionization and mass analysis. But this lessens the ability of the analyst to differentiate structural isomers of the same exact mass.

Although instruments that couple LC systems to mass spectrometers with an EI source exist, these are not a common instrument to most research facilities due to issues with the incompatibility of LC and EI. To mitigate these issues, an isotope tracing

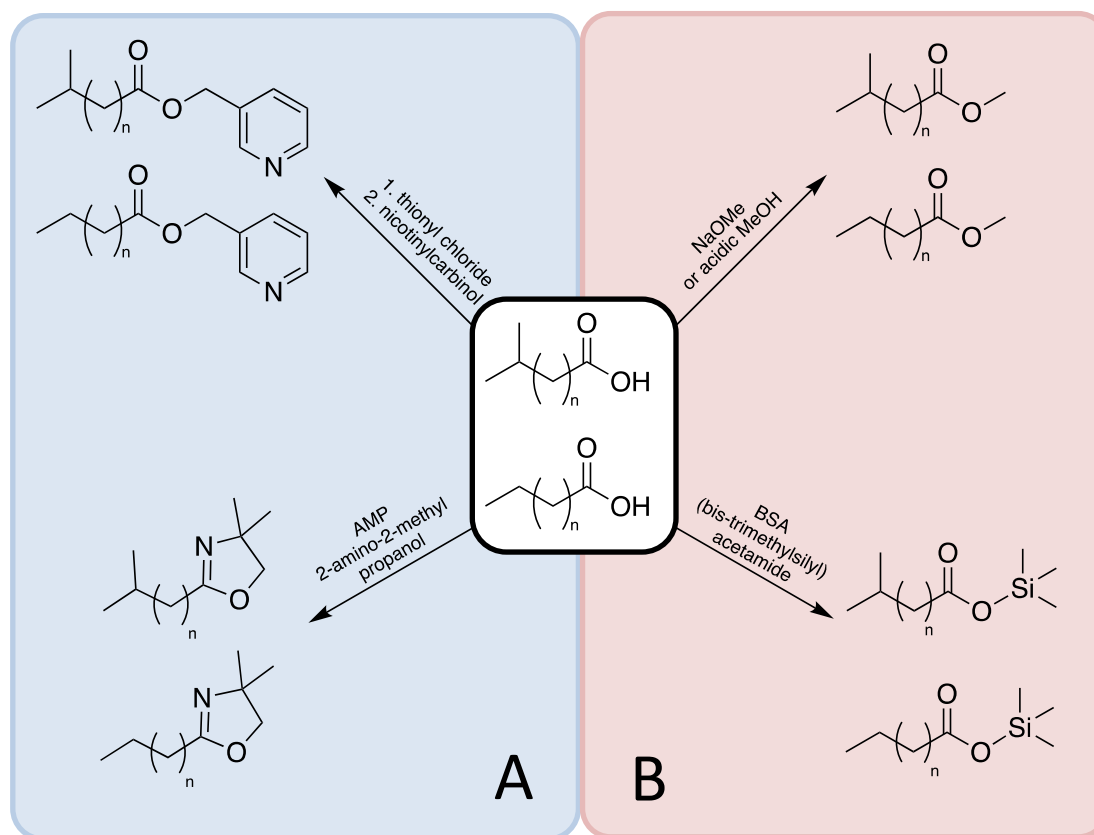


Figure 2.1.2 – Fatty acid derivitization methods for GC-EI-MS analysis

GCMS analysis of lipids usually require sample derivitization to either generate EI spectra that give specific structural information such as with nicotinylic esters and oxazolones (**A**) or to increase volatility such as with methyl esters and trimethylsilyl esters (**B**).

experiment was devised to tag lipids with a stable isotope for LC-ESI-MS analysis (**Figure 2.1.3**). The method would permit the biological generation of isotopically labeled isolipids from a feedstock of deuterated isovaleric acid (d_3 -(iso5:0):“D₃-iC₅”). Through this strategy, microbes supplemented with D₃-iC₅ use the heavy lipid analogue to incorporate into many biological molecules that feature a terminal isopropyl unit. In this way, cultures of bacteria can be grown on agar that has labeled and unlabeled isovaleric acid, lipids are extracted, and analyzed by LCMS instrumentation. The spectral features found in the D₃-iC₅ supplemented cultures that contain a population of isotopically labeled analogues not seen in the control cultures would indicate the presence of an isolipid. This strategy has many implications for LC-ESI-MS analyses of a multitude of biogenic compounds, especially in the structural elucidation of an unknown spectral feature with bioactivity.

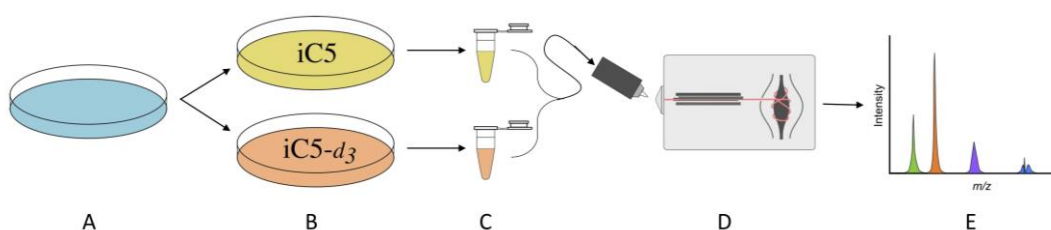


Figure 2.1.3 – iC5-d₃ isotope tracer experimental workflow

Shown above is the experimental workflow for the isotope tracer study. Cells are grown normally (**A**), then cultures are then transferred to new plates doped with either labeled or non-labeled isovaleric acid (**B**). Once enough time has passed for nutrient uptake, cells are harvested and extracted for LCMS analysis (**C**). Cell extracts are separated and analyzed by LC-ESI-MS (**D**), and the data is analyzed for the presence of isotopically labeled molecules that correlate to non-labeled molecules in the unlabeled sample (**E**).

2.2 – Introduction

2.2.1 – *Lipidomics and lipid semiochemicals*

Chemical signals, or semiochemicals, govern microbial life, from guided movement of individual cells, to initiating cascades of developmental change within a community of cells.⁵⁵ Understanding these signals may provide insight into how unicellular organisms communicate, how individual cells within multicellular organisms communicate with each other, and ultimately a greater fundamental understanding of cell-cell signaling. Identification and quantification of small molecules in biological systems has been of paramount importance to biochemical research, due to the influence the small molecule space has on the physiology and function of organisms. Lipids are a class of molecules that have important functional role in cell biology, normal and dysfunctional physiology, microbial virulence, and disease pathology.⁵⁶⁻⁵⁷ Lipids as signaling molecules has been an ongoing paradigm conceived of in the 1950's concerning the function of lipids in cell-cell communication and their role as biomarkers for disease.⁵⁸⁻⁶¹ These lipid species can be analyzed by a number of well-established instrumental methods such as LCMS, GCMS, MALDI-MS, NMR, and usually fall under the umbrella term of metabolomic analysis.

Generally, metabolomic studies compare some aspect of the metabolome for some control specimen to that of a mutant with some observable phenotypic differences to that of the control. By determining relative or absolute concentrations of metabolites representative of the system being studied, hypotheses on the functional role of metabolites that differ in concentration between the experimental organisms may be generated or substantiated. Lipidomics is a sub-class of metabolomics describing the analysis of all the detectable lipid species present within a sample. This can include non-polar compounds such as terpenes, steroids, long-chain fatty acids, and isoprenoids as well as polar compounds such as phospholipids, saccharolipids, short-chain fatty acids,

and polyketides.⁶² Several biologically relevant lipid species resemble a similar structure where a polar head-group is connected to one, or several, lipid tails of varying length, degrees of unsaturation, and functionality. Determining the exact structure of the lipid tails in these compounds is often not a trivial task. Characterizing these features require isolating the species of interest and performing either 1D/2D NMR experiments or GCMS analyses involving some chemical derivatization on the isolated compound(s). These derivatized species have much lower vapor pressures (volatilities) making them more amenable to analysis by gas chromatography. This gives the analyst the ability to separate lipids by chain-length and degree of unsaturation as well as provides fragmentation data from the hard ionization source to gather information on molecular structure.

Most lipidomic analyses are performed on biological systems with the intent to better understand a function or loss of function due to some lipid(s) either being present/absent or in higher/lower concentration to that in a biological control specimen. Lipids species that are typically targeted and focused on in a lipidomic analysis are those that are part of cellular membranes, which vary in composition and structure. Changes to membrane lipid content can have a variety of effects on the physical and chemical attributes of the cell, such as membrane rigidity/fluidity, temperature tolerance, export and import of external and internal compounds, and can affect the binding of proteins and small molecules to cell membrane-localized structures.⁶³ These targeted studies utilize a library of analytical standards to annotate the data obtained from chromatography coupled mass spectrometry to somewhat ambiguously identify the lipids present in the system being studied. These libraries consist of several lipid classes that are known to exist in cells and cellular compartments; each with a combination of previously characterized lipid tails to assist in identifying lipids

Often, the goal of microbial lipidomics is to gather information on unknown molecule(s) that may have significance in the initiation of a behavioral response in microbes, such as the stringent response in the soil bacterium *Myxococcus xanthus*.

Recently, previously uncharacterized lipids with cytotoxic activity that are specifically formed during the stringent response of *M. xanthus* have been reported as containing terminal isoprenoids, which was determined through a combination of MS and NMR experiments. Hoffmann and co-workers (*ACS Chemical Biology*, 2018) isolated the cytotoxic lipids (**Figure 2.2.1**) by semipreparative chromatography of cell extracts of nearly 500 plated cultures.⁶⁴ The inclusion of the NMR experiment was necessary due to the limitations of current LCMS techniques in differentiating between straight-chain and branched-chain lipids. The amount of work needed to arrive at the same conclusion could be minimized by the use of a targeted isotope tracing experiment only requiring the use of a single LC-ESI-MS analysis.

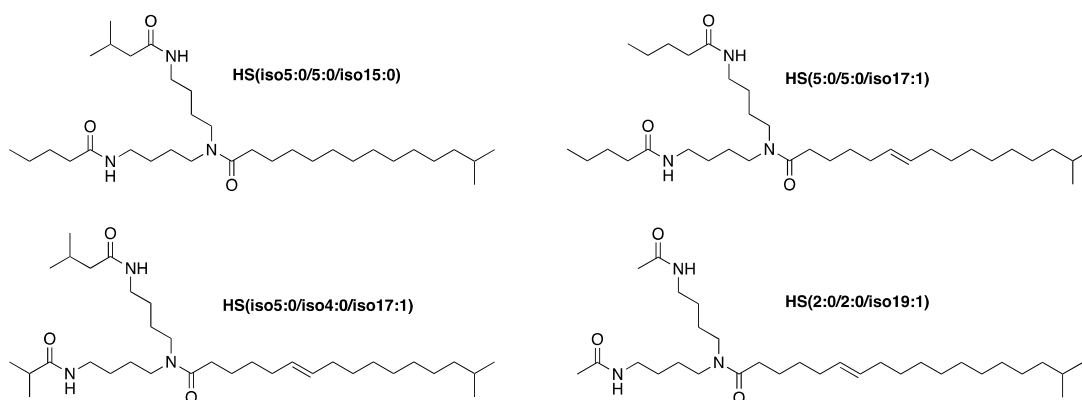


Figure 2.2.1 – Lipids expressed during *M. Xanthus* stringent-response

Homospermidine (HS) lipids discovered by Hoffman *et. al.* to be important for initiating fruiting body formation in *M. xanthus*.⁶⁴

2.2.2 – Differentiating isobars in LCMS analyses

Structural isomers are a confounding aspect of mass spectrometric analyses. Isobaric compounds are a problem often encountered in lipidomic space, where structural isomers of lipids obscure the exact knowledge that can be obtained from such an analysis. As an example: the position of a methyl group along the length of a lipid tail cannot be determined without fragmentation data (**Figure 2.2.2**). This issue can sometimes be resolved under optimal chromatographic conditions to separate the isobars, if their retention factors are different enough to allow for such separation; However, if the isobars of interest cannot be separated through chromatographic techniques prior to introduction into a mass spectrometer, determining relative concentrations between these species is practically impossible. The parent mass of these kinds of structural isomers are the same, therefore they cannot be filtered from one-

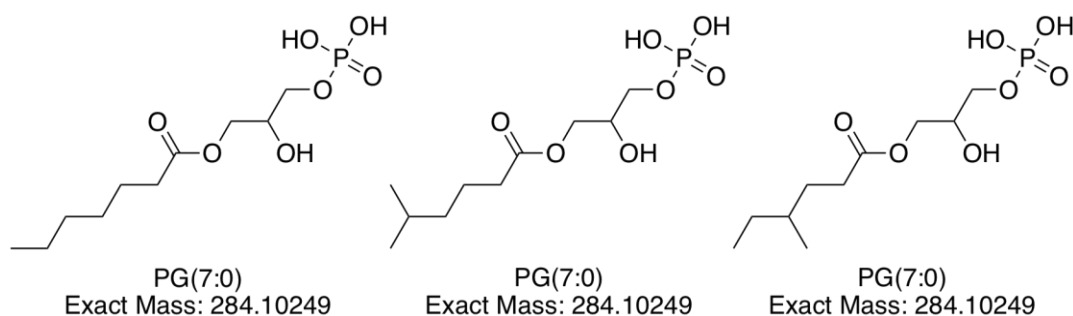


Figure 2.2.2 – Isobaric compounds indistinguishable by LC-ESI-MS analysis

Three examples of isobaric compounds that would be indistinguishable by traditional lipidomic LC-ESI-MS analysis.

another during mass filtering in instrumental ion optics – such as with a triple-quadrupole mass spectrometer. So, even when generating informative fragmentation data on mixtures of isobars, it may be very difficult to determine relative concentrations between these compounds; Although, there have been successes in differentiating between isobaric compounds by LCMS analysis with pre-column derivatization. Marfey's reagent is one such success that is used to modify mixtures of amino acid enantiomers and enforce chiral separation on non-chiral stationary phases (**Figure 2.2.3**).⁶⁵⁻⁶⁶ This allows for enantioselective separations to be performed on cheaper and more routinely used HPLC columns. Another example of the utility of pre-column derivatization is in the analysis of sugars, which have either been phosphorylated or sulfonated that are isobaric in lower-resolution (<118,000) mass spectrometers. These species can be distinguished from one another by using an ion-pairing tripeptide (Lys-Lys-Lys; aka "K3") to engender characteristic fragments that illuminate these functional group differences.⁶⁷

The specific distinction of sulfonation versus phosphorylation or between enantiomers are of great importance to biochemical research. These species are physiologically and chemically distinguishable by their reactivities and bioactivities, but they are indistinguishable in MS-space without the use of a high-resolution instrument such as a Fourier transform ion cyclotron resonance (FT-ICR) or Orbitrap mass spectrometer. Isobaric compounds that do not contain the exploitable physical and chemical properties as just described are simple carbonaceous structural isomers that have similar chromatographic properties, which are challenging to distinguish in LCMS analyses. These compounds are simple to identify and determine relative concentrations of these species in a sample with NMR techniques but require a large amount of fairly pure compound in order to make NMR analysis feasible. This is where small molecule isotope tracing experiments have an advantage by not having to rely on pre-analysis sample derivatization, or unique chemistries induced by instrumental fragmentation. By introducing an isotopically labeled molecular congener to a biological system it's possible

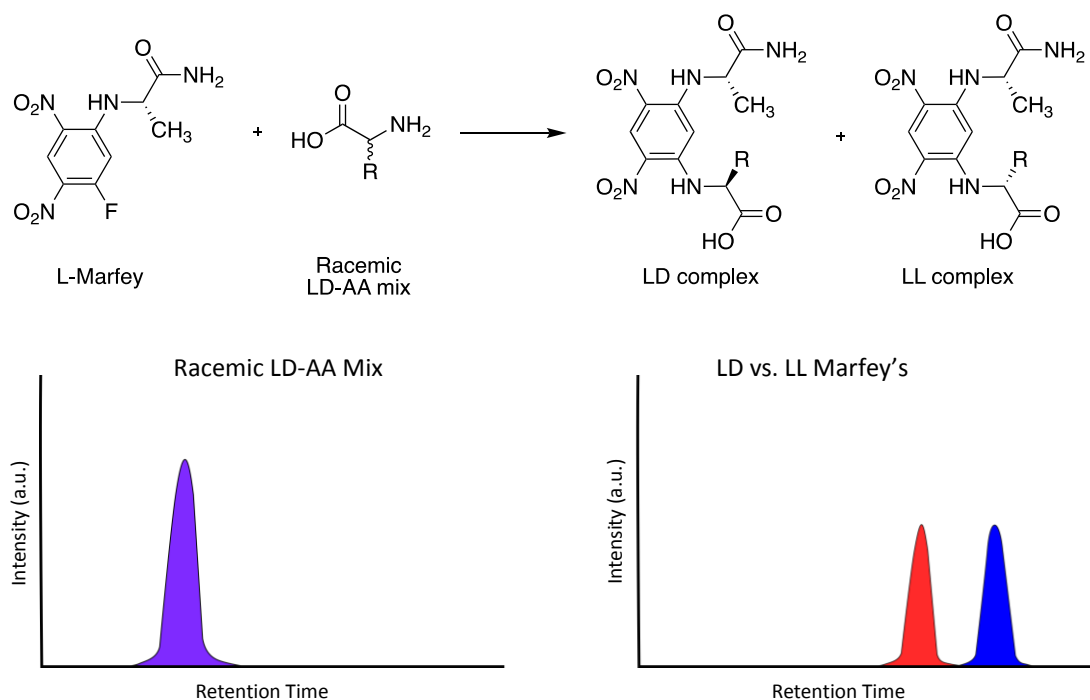


Figure 2.2.3 – Marfey's reagent for the separation of amino acid enantiomers

Following derivatization with 1-fluoro-2,4-dinitrophenyl-5-L-alanine amide (Marfey's reagent) of a racemic mixture of amino acids, separation of these enantiomers can be achieved using non-chiral stationary phases making these types of analyses more cost-effective than by conventional chiral-column LCMS.

to generate metabolites built of the introduced tracer, which would provide a distinguishable signal of a multitude of structural isomers.

2.2.3 – Lipid biosynthesis and isotope tracing

Lipids are synthesized biologically by finding and absorbing necessary lipids from the environment, or by utilizing small ubiquitous metabolites such as acetate or branched-chain amino acids. Straight-chain fatty acids are derived from acetyl-CoA or other small straight-chain fatty acids, which undergo a few priming conversions before entering the lipid elongation pathway. Branched-chain fatty acids arise from branched-chain amino acids leucine, isoleucine, and valine or from free branched-chain fatty acids such as isovaleric acid acquired from the environment or as byproducts from their own metabolism such as the metabolic degradation of leucine (**Figure 2.2.4**).⁶⁸⁻⁶⁹ It is conceivable that these biosynthetic pathways could be taken advantage of by introducing an isotopically labeled synthon such as those just described, and allowing the cells to make a multitude of isotopically enriched primary and secondary metabolites.

Stable isotopes have been used for decades in MS analyses for a number of applications, especially in the realm of studying metabolism, or the chemical transformations occurring in biological systems.⁷⁰ Beginning in the 1920's, the work of Francis W. Aston using a primitive mass spectrometer was instrumental in solidifying John Dalton's theory that the existence of many stable isotopes per element was credible. The fact that stable isotopes exist and would be detectable from molecules that are not isotopically enriched gave scientists an avenue to explore metabolism *in vivo*. Stable isotope-labeled molecules have a number of applications in mass spectrometry such as with metabolic flux studies, isotope dilution, internal standardization, and isotope tracing experiments. Isotope tracing is the focus of this chapter, more specifically the approach of providing an organism an isotopically labeled feed-stock and observing what metabolic products are formed. Using this method of introducing a stable isotope into metabolic products of bacterial systems is not a novel concept, but it has not yet been used to probe

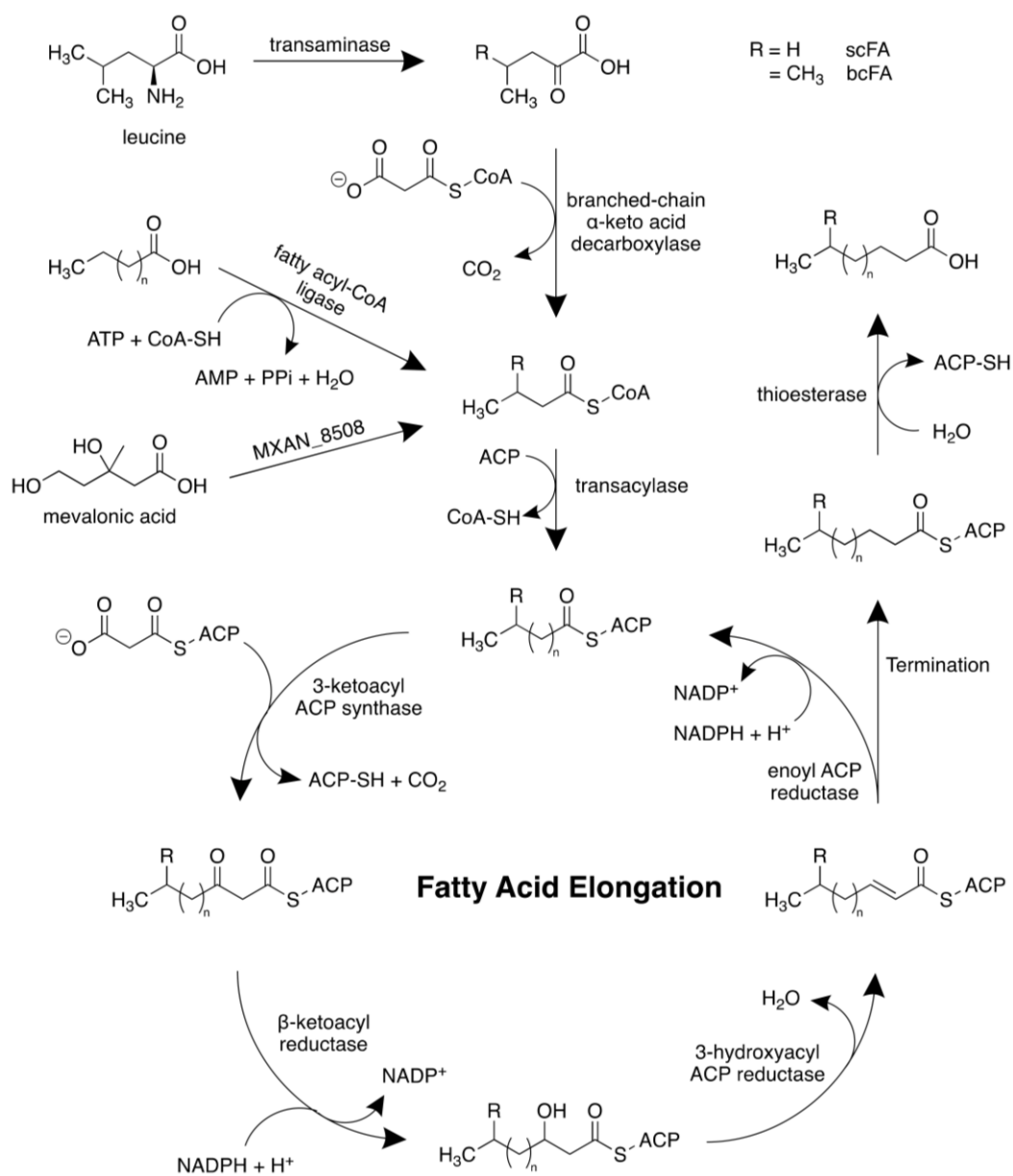


Figure 2.2.4 – Branched-chain and straight-chain lipid biosynthesis

Branched-chain and straight-chain lipids of varying lengths may be biologically synthesized by the elongation of primed branched-chain fatty acids such as isovaleric acid and butyric acid (shown above; $n=1$).

the metabolome for structural information about the molecules that adopt an isotopically enriched moiety.

2.2.4 – Isolipid model organism: *Myxococcus xanthus*

The canonical prokaryotic model organism for studying social behavior and cell-cell signaling is the soil bacterium *Myxococcus xanthus*; a Gram-negative proteobacteria with a large genome that exists as a predatory saprophytic single-species biofilm.⁵⁵ This organism is well suited for studying the effects of chemical signals, due to the nearly comprehensive understanding what chemical signals are produced and how colonies of these bacteria are affected developmentally by these signals. When a colony of *M. xanthus* faces starvation, vegetative growth ceases and the bacterial colony instead shifts all effort to the formation of macroscopic mounds of dormant myxospores, called fruiting bodies. This shift in behavior is known as the stringent response and is initiated by amino acid limitation. The response promotes the down-regulation of the gene responsible for cell growth, *SocE*, and the simultaneous transcriptional up-regulation of *CsgA*, which produces the inner-membrane protein CsgA - essential for cell aggregation, lipid body production, and fruiting body formation.⁷¹⁻⁷² Although the action of CsgA has been determined to be a cardiolipin phospholipase that populates the cell with diacylglycerides (DAGs) and diacyl phospholipids that are believed to be essential in fruiting body formation.⁷³⁻⁷⁴

Another important feature of the stringent response in *M. xanthus* is the production of homospermidine lipids that are cytotoxic to competing bacteria, a defense mechanism that uses lipids available to the organism to produce these compounds. These lipids were recently discovered and characterized by Müller and co-workers by a combination of LCMS and NMR techniques.⁶⁴ An alternative strategy to the work that went into this discovery could have been the use of a logical isotope tracer to determine structural motifs of these lipids. These compounds are easily differentiable since they can be identified by their exact mass in a typical LCMS-mediated lipidomic analysis. Using *M.*

xanthus as a model organism for tagging the production of isolipid species via supplementation with an isotopically labeled isolipid carbon-source will provide definitive proof-of-concept for this isotope tracing method that can be applied to numerous microbial systems.

2.3 – Results and discussion

2.3.1 – Analysis of *M. xanthus* lipids

In a collaboration with the microbiology lab of Dr. Larry M. Shimkets at the University of Georgia, Athens; samples of whole-cell lipid extracts were obtained to determine the lipidomic profile of a set of *M. xanthus* strains that behaved differently under nutrient-limitation. Wild-Type sp. DK1622 was used to be the control specimen while sp. LS1991 was a strain deficient in the production of branched-chain lipids, which did not undergo sporulation under these limiting conditions. It is because this mutant (LS1191) lacks a majority of pathways available to *M. xanthus* to synthesize isovaleryl-CoA from leucine. The pathways in question are functions of branched-chain ketoacid decarboxylase and a secondary pathway mediated partially by MXAN_4265, a species-specific glutaconyl-CoA transferase that is part of the mevalonate-dependant isoprenoid biosynthetic pathway.⁷⁵ Specifically, this mutant is restored to normal function by supplementation with either exogenous iso15:0 or iso5:0 fatty acids.⁵⁵ It was shown through previous publications that by supplying *M. xanthus* mutants unable to undergo sporulation with lipids extracted from sp. DK1622, the mutant strains were able to undergo successful sporulation into viable fruiting bodies.⁷⁶⁻⁷⁷ In order to profile the lipids present in *M. xanthus* whole-cell lipid extracts were subjected to lipidomic analysis by an in-house lipidomics method in the Biological Small Molecule Mass Spectrometry Core facility (BSMMSC) at the University of Tennessee, Knoxville. Profiling of lipids present during vegetative growth would allow for the distinction of the lipids present during this growth-phase to be compared to the lipid profile of sp. DK1662 during sporulation. Taking

these data into account and comparing them to the profiles of mutant strains, which cannot undergo sporulation, could provide evidence of some mass feature not present in the mutant strain but readily detectable in the wild-type organism. These detectable biomarkers would be the top candidates for either initiating this phenotypic behavior in sp. DK1622 and would need to be isolated and structurally characterized by liquid chromatography tandem mass spectrometry (LC-MS/MS) with additional structural information gathered by NMR analyses of the isolated compound(s).

Upon lipidomic analysis of sp. DK1622 and sp. LS1191 it was apparent that the lipid profiles between these organisms were similar in terms of detectable lipid species but differed in detected concentrations of these lipids between organisms (**Figure 2.3.1**). It could be that there is a necessary quorum of certain lipid species that could be contributing to the reason that mutant cells that produced less lipids could not undergo fruiting body formation. An alternative and more likely reason for this observation is that the structural identity of the lipids present between these organisms may have a greater effect on the phenotypic differences between them. Since the *esg* (E-signal) mutant sp. LS1191 is deficient in branched-chain lipid production, the low concentration or absence of certain isolipids could be the determining factor in the lack of developmental changes when colonies of this mutant face starvation.⁷⁸⁻⁷⁹ If this structural difference in semiochemicals is the causal factor for loss of function in sp. LS1191, any of the lipids detected between these strains could contain this structural feature as straight-chain lipids are indistinguishable from lipids branched-chain lipids of the same chemical formula in most LCMS analyses.

The BSMMSC list of lipid species cataloged in the in-house lipidomics method covers a range of phospholipid species of differing carbon count and degree of unsaturation, and the lack of drastic differences between these experimental organisms suggests that structural components of these lipids may be contributing to the phenotypic differences; However, it is quite possible that there are unknown mass features

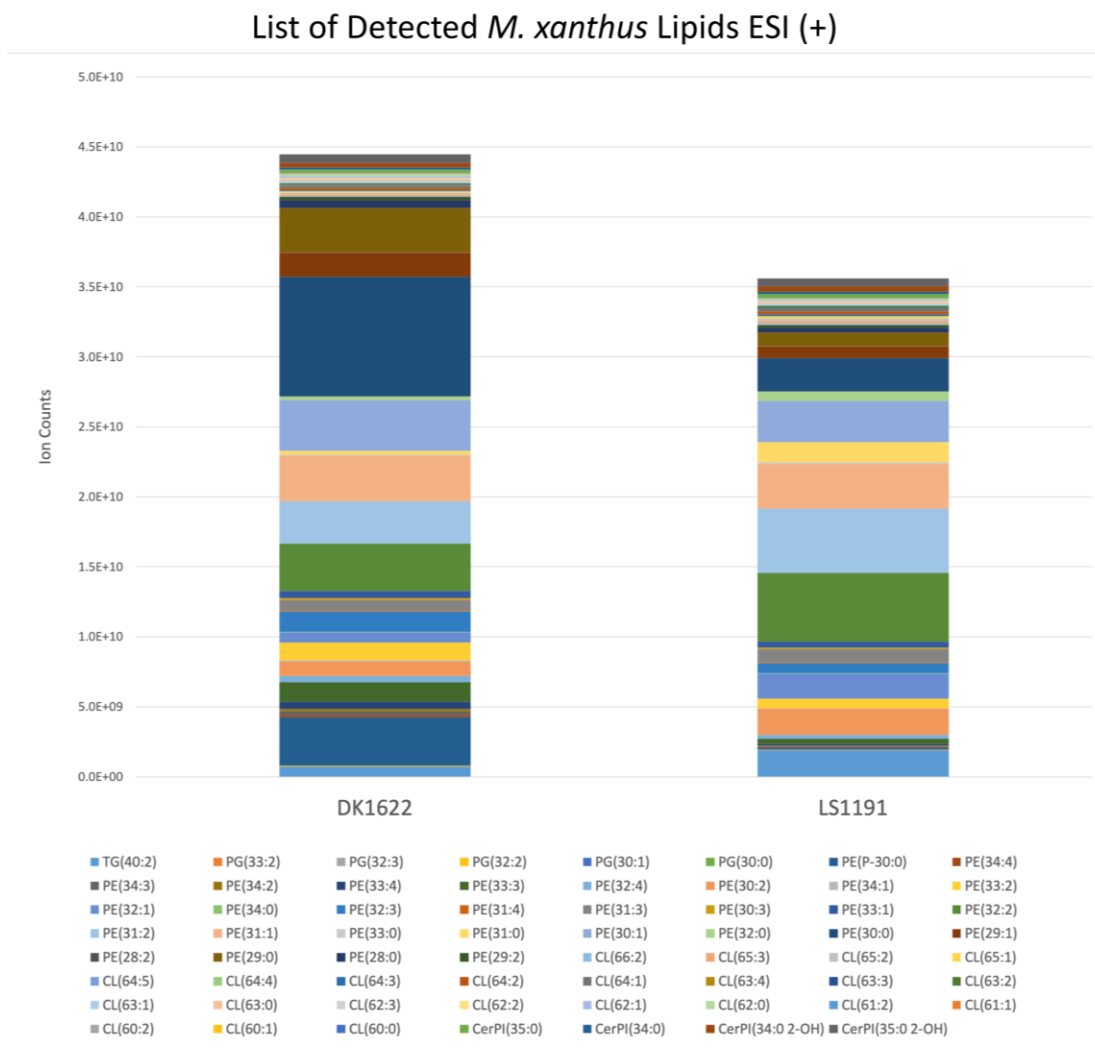


Figure 2.3.1 – Lipid composition between *M. xanthus* strains DK1622 and LS1191

Lipid species detected in wild-type *M. xanthus* DK1622 and *esg* mutant LS1191. This list of detected lipids in *M. xanthus* was previously established by Dr. Helge Bode.⁸⁰

not cataloged in the BSMMSC's lipidomics library that are contributing to the observed differences in these two organisms. The unknown space of LCMS analyses is very complex and contain thousands of mass features. In an attempt to curb the complexity of the whole-cell lipid extracts, preparative TLC was performed on *M. xanthus* extracts and a set of five distinct bands along the TLC plate revealed restorative activity to sp.LS1191. These "bioactive bands" were subjected to lipidomic analysis to determine what unknown spectral features in these samples may have contributed to the observed activity. Even after performing statistical analyses such as PLSDA and extracting the variable projection of importance (VIP) scores of these mass features to determine the features that significantly drive statistical separation between samples, the list of statistically impactful mass features can still be between hundreds to thousands of significant features (**Figure 2.3.2**). With this in mind, being able to spectral features with a specific structural component when considering all the detectable masses in a routine lipidomics experiment would be notably helpful to the analyst. Including another means of determining the compounds in the unknowns to be more relevant to the study of isobaric compounds that are structural isomers of each other would reduce the list of important features significantly.

2.3.2 – Isotope tracing in *M. xanthus*

M. xanthus is a well-studied organism and much is known about the behavioral, structural, and chemical characteristics that make this organism unique. The lipid composition of the cell membrane in *M. xanthus* has been established to contain an unusually high level of 13-methyltetradecanoic acid (iso15:0), which makes up nearly 41% of the fatty acid composition of sp. DK1622 cells.⁸¹ Taking into consideration the natural abundance of isolipids in this organism, it seemed practical to synthesize and supplement a deuterated version of the naturally occurring 15-carbon lipid, D₆-(iso15:0) – also known as iC₁₅-d₆ – to see what higher-order compounds the organism used that lipid for,

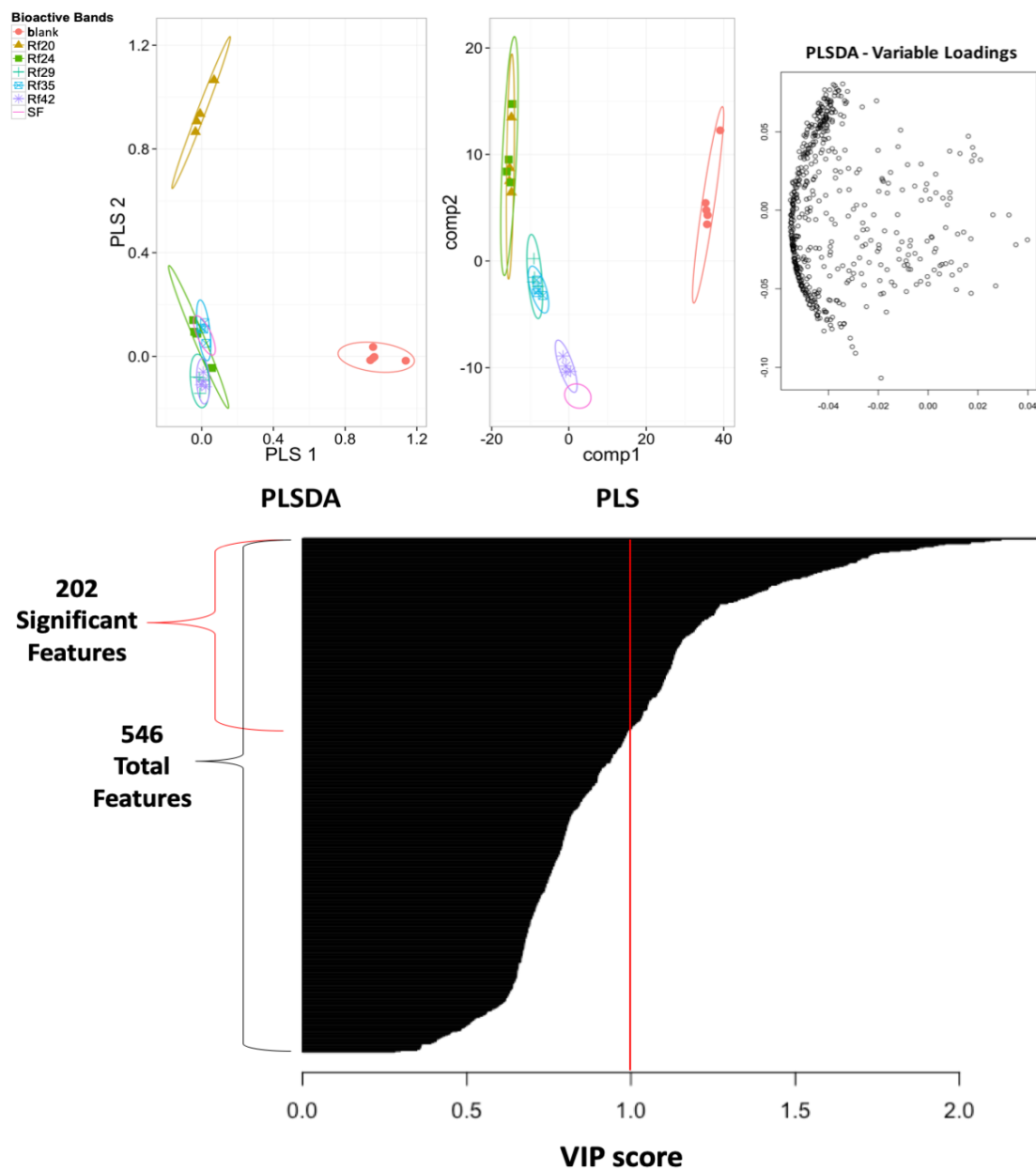


Figure 2.3.2 – PLSDA and VIP score analysis of unknown spectral features between *M. xanthus* strains DK1622 and LS1191

Shown above is an example of the complexity of the unknown mass features in an untargeted LCMS analysis of *M. xanthus* strains DK1622 and LS1191.

especially during the stringent response. The *M. xanthus* mutant LS1191 is deficient in branched-chain fatty acid synthesis, making it the ideal organism to probe the uptake and incorporation of iso-fatty acids into the cell. Developmentally significant lipid species containing (iso15:0) can be investigated by presenting sp. LS1191 with synthetic $iC_{15}-d_6$ as a component of the media used to grow the cells on. Analysis of the lipids extracted from the ($iC_{15}-d_6$)-doped LS1191 colonies will demonstrate the uptake and incorporation of the synthetic lipid, generating a detectable population of isolipid species that can be further targeted for fragmentation studies as putative signaling molecules with a partially known structure.

The data obtained from this experiment could be used to optimize detection methodology in preparation for an untargeted isotope tracing experiment. A critical aspect of this project is whether or not *M. xanthus* will pick up exogenous $iC_{15}-d_6$ and incorporate it into higher-order compounds such as phospholipids, cardiolipin, or other lipid classes. By adding both synthetic lipids iso15:0 and $iC_{15}-d_6$ to the media of vegetative cells and comparing the lipid composition against the lipid extract of wild-type *M. xanthus* cell grown on undoped media, it can be determined if the synthetic lipid was taken up by the cells, as well as if the addition of exogenous lipids change the lipid profile of the cells themselves. This approach with d_6 -(iso15:0) was attempted but met with suboptimal results. To combat this apparent lack of isotope tracer uptake, a smaller more permeable and metabolically useful tracer was considered. Since the primary biologic synthon for the biosynthesis of most isolipids in *M. xanthus* is isovaleryl-CoA, it was obvious that the ideal isotope tracer for this study was not $iC_{15}-d_6$, but rather the precursor to isovaleryl-CoA, isovaleric acid. All synthetic efforts shifted to the generation of the small isotopically labeled compound iC_5-d_3 for the tracing study. After synthesis of this isotope tracer iC_5-d_3 and supplementation of the tracer into cultures of DK1622 and the isolipid-deficient mutant LS1191, there was a great degree of labeling observed in lipidomic analyses of these samples (**Appendix 2.6.2**). This was not unexpected, as *M. xanthus* synthesizes most

of its isolipid components from isovaleryl-CoA, except for the only other smaller isobutyrate species. Since these microbes were also supplemented with iC_5-d_3 during amino acid limitation, these microbes should have been producing homospermidine lipids as anti-microbial compounds during the stringent response. These molecules were also detected in this analysis, and labeled lipids were detected in the samples supplemented with iC_5-d_3 , where the degree of labeling corresponds to the number of isolipid tails present in the previously identified compounds by Müller and co-workers (**section 2.6.3**).

2.3.3 – Synthesis of isotope tracers

The synthesis of the preliminary isotope tracer target molecule $iC_{15}-d_6$ was adapted from an article published by the Williams group of the University of Melbourne, Australia (**Figure 2.3.3**).⁸² The synthesis of these long-chain fatty acids starts with the commercially available macrocyclic musk 13-methyltetradecanolide and then relies on recursive dehomologation to achieve the isolipid of the desired tail-length. Briefly, double addition of D_3 -methyl units, via freshly generated Grignard, to the 15-carbon lactone **1** provides diol **2**. Selective reduction of the tertiary alcohol with a silane reducing agent in the presence of $BF_3 \cdot Et_2O$ provides **3** with traces of alkene by-product. After purification of the desired product, the residual alkene by-product was reduced by Pd-catalyzed hydrogenation, purified, and collected with the rest of the fatty alcohol **3**. Recursive dehomologation was achieved through a Chugaev reaction with **3**, proceeding via intramolecular elimination of intermediate xanthate **4** to alkene **5**. Oxidation of the 17-carbon alkene **5** with potassium permanganate giving the 16-carbon iso-fatty acid **6**. After reduction of **6** by borane-dimethylsulfide, the resultant alcohol **7** was taken through the same transformations as molecule **3**, yielding the final product $iC_{15}-d_6$ (**10**). This synthesis could have been shortened by a few steps by using cyclotridecanolide as the starting material, though this would have been extremely cost-prohibitive due to the fact that cyclotridecanolide is ~50,000 times more expensive than 13-methyltetradecanolide. For the synthesis of the 5-carbon isotope tracer iC_5-d_3 , the strategy of attaining this molecule

1 $\xrightarrow[\text{THF, } \sim 98\%]{\text{D}_3\text{-MeMgI}}$ **2** $\xrightarrow[\text{DCM, } \sim 96\%]{\text{BF}_3 \cdot \text{Et}_2\text{O}, \text{Et}_3\text{SiH}}$ **3** $\xrightarrow[\text{THF, quant.}]{\text{1. NaH, CS}_2, \text{2. MeI}}$ **4**

4 $\xrightarrow[\sim 89\%]{\text{reflux}}$ **5** $\xrightarrow[\text{DCM, } \sim 93\%]{\text{KMnO}_4, \text{Bu}_4\text{NBr}, \text{AcOH}, \text{H}_2\text{O}}$ **6** $\xrightarrow[\text{THF, } \sim 78\%]{\text{BH}_3 \cdot \text{SMe}_2}$ **7**

7 $\xrightarrow[\text{THF, quant.}]{\text{1. NaH, CS}_2, \text{2. MeI}}$ **8** $\xrightarrow[\sim 89\%]{\text{reflux}}$ **9** $\xrightarrow[\text{DCM, } \sim 86\%]{\text{KMnO}_4, \text{Bu}_4\text{NBr}, \text{AcOH}, \text{H}_2\text{O}}$ **10**

Synthetic procedure of iso-fatty acids adapted from the Williams work to include the incorporation of fully deuterated terminal methyl groups.⁸²

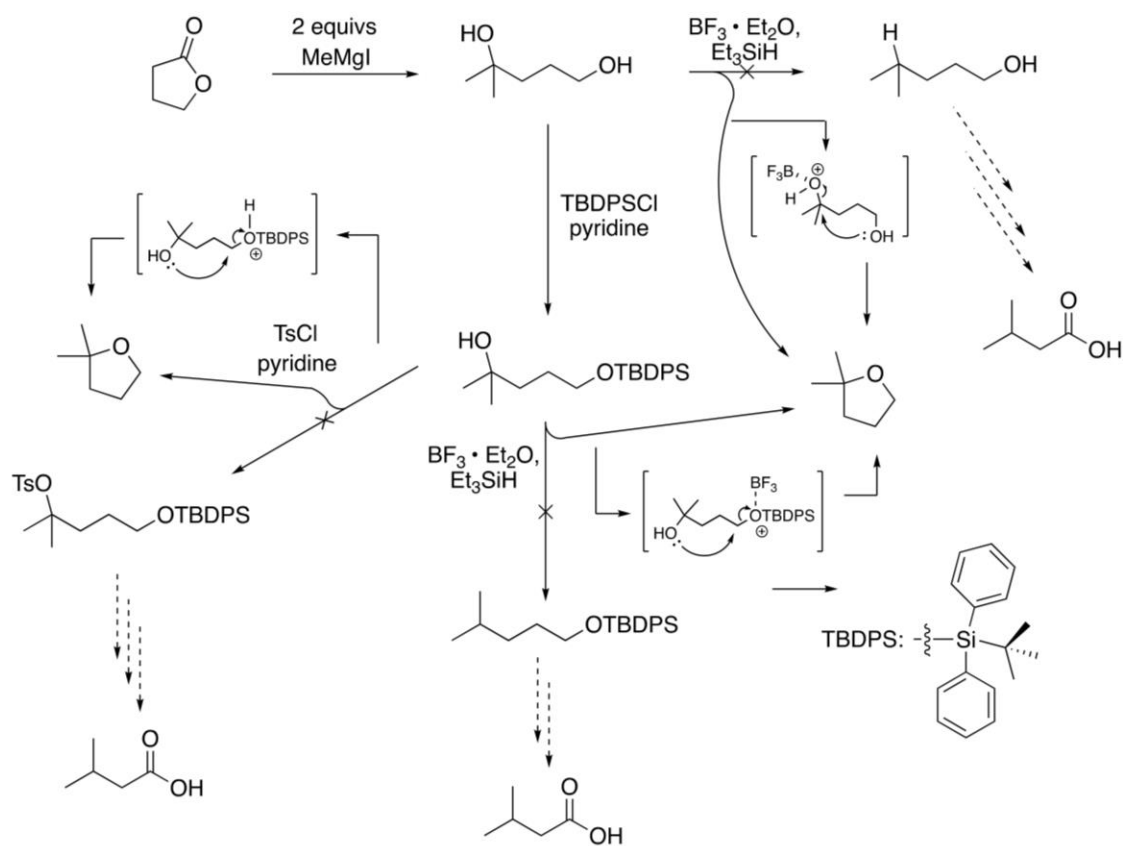


Figure 2.3.4 – Attempted synthetic routes toward isovaleric acid

Attempted syntheses of isovaleric acid through the Williams strategy, which was met with abject failure due to self-cyclization of intermediates.

Even though the reaction conditions of this protection did not result in cyclization, and even survived purification via flash chromatography using ammonia-doped silica as the stationary phase, efforts to remove the tertiary alcohol all met in failure. Either by using similar chemistry to the TBDPS protection to tosylate the remaining alcohol, or under reducing conditions with $\text{BF}_3 \cdot \text{Et}_2\text{O}$, the compound cyclized presumably by some nucleophilic mechanism that was not further explored.

Alternatively, a new synthetic route was devised after searching the literature for reactions allowing the addition of a carbonaceous nucleophile to yield a carboxylic acid with a terminal isopropyl group. This approach has been shown to be achievable with a number of small strained lactones that gave iso- and anteiso-fatty acids in good yields.⁸³⁻⁸⁴ This synthetic route (**Figure 2.3.5**) was much simpler, giving small iso-fatty acids directly from starting material, and more economical due to the low cost and availability of the appropriate synthon for the target molecule, β -butyrolactone (**11**). The desired compound was achieved in a single step by the addition of a D_3 -methyl unit via a Gilman-type diorganocuprate, which proceeded through a direct β -carbon fission between the isotopically labeled reagent and the lactone **11**, giving the product $\text{iC}_5\text{-d}_3$ (**12**) in a reasonable yield.

2.4 – Methods

2.4.1 – Synthesis of isotopically labeled tracing lipids

The synthetic procedures used to obtain the isotope tracers $\text{iC}_{15}\text{-d}_6$ and iC_5d_3 were obtained from previously published procedures by Müller and Ito, respectively.⁸²⁻⁸³ Deuterium was incorporated into the target compounds by the action of a Grignard reagent generated from deuterated methyl iodide (CD_3I). Chemical reagents were purchased commercially, and most were used without further purification. Carbon disulfide, methyl iodide and (D_3)-methyl iodide were purified by distillation immediately prior to use. Dry tetrahydrofuran (THF) was obtained by distillation from a potassium-

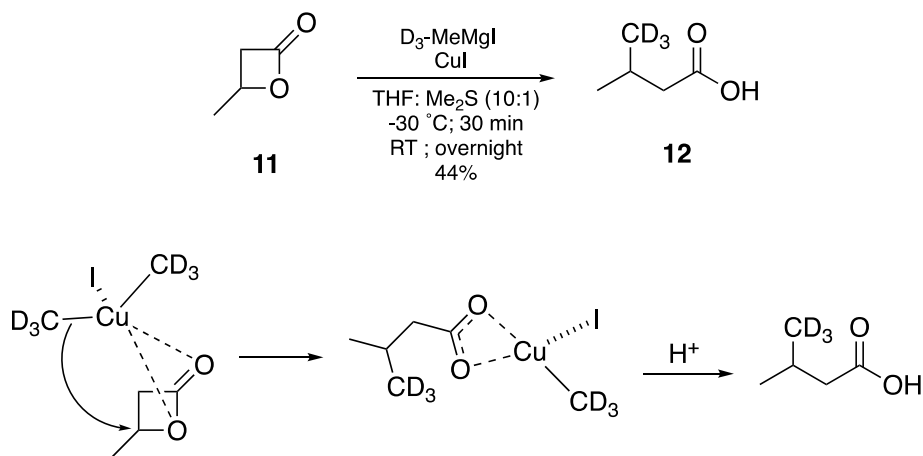
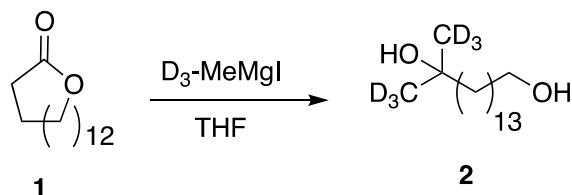


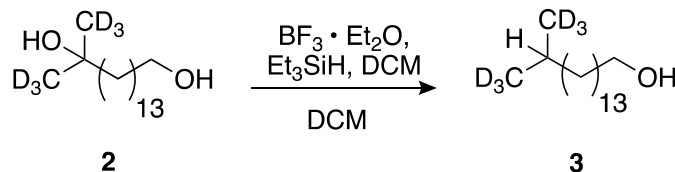
Figure 2.3.5 – Synthetic route toward $i\text{C}_5\text{-}d_3$

Synthetic procedure for the preparation of deuterated isovaleric acid with proposed mechanism, adapted from a previously published work on nucleophilic opening of β -propiolactones.⁸⁴

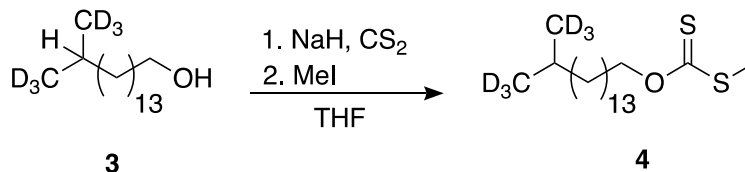
(Drierite®). All glassware that was used in reactions requiring dry solvents were flame-dried. Thin-layer chromatography (TLC) was carried out using Sorbent Technologies silica G w/UV254 TLC plates. All purified compounds of interest were visualized as single spots using short-wave UV light; more permanent staining of TLC spots was achieved using vanillin, p-anisaldehyde, bromothymol blue, or KMnO_4 TLC staining solutions. $^1\text{H-NMR}$ spectra were obtained on an Oxford 300 Mhz or Varian InNova 500 MHz instruments in solutions of CDCl_3 . Spectra of isolated compounds are cataloged in the appendix for chapter two. High resolution mass spectrometry (HRMS) data was obtained using a Thermo Scientific Exactive Plus Orbitrap mass spectrometer equipped with an ESI source.



15-(methyl- d_3)hexadecane-16,16,16- d_3 -1,15-diol (D_6 -iC17-diol; **2):** To a stirring and chilled solution ($-78\text{ }^\circ\text{C}$) of petadecanolid (**1**, 1.44 g, 6 mmol) in dry THF (40 mL) was added freshly generated D_3 -methlmagnesium iodide in Et_2O (1 M, 20 mmol) dropwise via cannula. The resulting thick and chunky solution was allowed to warm to room temperature and stir for an additional 16 hours. At this time the reaction mixture was cooled to $0\text{ }^\circ\text{C}$ and a saturated solution of NH_4Cl was added to the slurry until all magnesium salts were dissolved and the pH of the biphasic mixture had reached a slight acidity (pH ~ 6.0). The organic layer was washed with brine (100 mL) and collected while the remaining aqueous solution was back-extracted with Et_2O (3 x 50 mL). The pooled organic extracts were dried over anhydrous $MgSO_4$, filtered, and concentrated in vacuo. Crude product was purified recrystallization by dissolution of the crude product in a solution of 9:1 [hexanes: $EtOAc$], and chilled in a $-20\text{ }^\circ\text{C}$ freezer overnight. Yield = 98%. TLC R_f = 0.17 (25% $EtOAc$ /hexanes). HRMS (ESI-Orbitrap) m/z : $[M+H]^+$ Calcd for $C_{17}H_{31}D_6O_2$, 279.3165; found 279.3161. 1H -NMR ($CDCl_3$, 300.08 MHz) δ 1.25 (br s, 22H), 1.41 (br s, 2H), 1.53-1.58 (m, 2H), 3.60-3.65 (t, 2H)



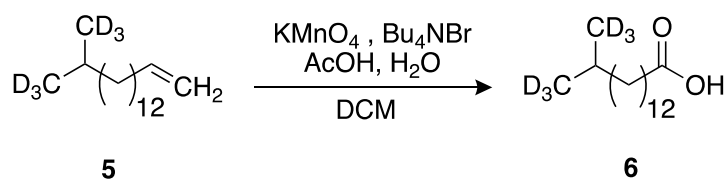
15-(methyl-d₃)hexadecan-16,16,16-d₃-1-ol (D₆-iC₁₇-ol; **3):** To a cool (0 °C) solution of **2** (5.5 g, 20.19 mmol) and Et₃SiH (6.45 mL, 40.4 mmol) in dry DCM (40 mL) was added BF₃•Et₂O (24.93 mL, 202 mmol) in DCM (10 mL) dropwise via a dropping addition funnel. Once added, the reaction was allowed to stir for an additional 45 min at 0 °C. At this time the reaction mixture was diluted with ice-cold water (2 x 40 mL) and extracted with DCM (2 x 50 mL). The pooled organic solvent was washed with a saturated solution of NaHCO₃. The organic layer was dried over anhydrous MgSO₄, filtered, and concentrated *in vacuo*. Crude product was purified via normal-phase flash chromatography. Yield = 96%. TLC R_f = 0.95 (10% Et₂O/hexanes). HRMS (ESI-Orbitrap) *m/z*: [M+H]⁺ Calcd for C₁₇H₃₁D₆O, 263.3215; found 263.3125. ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.12-1.16 (m, 2H), 1.26 (br s, 22H), 1.46-1.49 (t, 1H), 1.54-1.59 (m, 2H), 3.63-3.65 (t, 2H)



S-methyl-O-(15-(methyl-d₃)hexadecyl-16,16,16-d₃)carbonodithioate (D₆-iC17-xanthate;

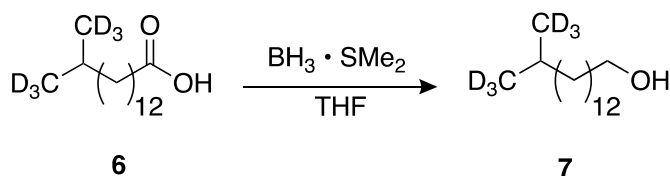
4): The first step of this one-pot reaction sequence starts with NaH (60% dispersion in mineral oil, 1.73 g, 43.33 mmol) being added in one portion to a chilled (0 °C) solution of **3** (4.44 g, 17.33 mmol) in dry THF (60 mL) and allowed to stir for 1 hour at this temperature. The next step forms the sodium dithiolate intermediate by slow addition of CS₂ (1.57 mL, 26 mmol) via syringe to the cool (0 °C) reaction solution and is again allowed to proceed at this temperature for 1 hour. Formation of the methyldithiolate ester is accomplished by the addition of iodomethane (1.62 mL, 26 mmol) to the stirred reaction at 0 °C, which is then allowed to warm to room temperature naturally over the course of 5 hours. At this time the reaction is quenched by the addition of a saturated solution of NH₄Cl (70 mL) and extracted with EtOAc (3 x 100 mL). The organic layer was dried over anhydrous MgSO₄, filtered, and concentrated *in vacuo*. The rusty-brown product was taken up in enough pentane to drop the viscosity of the oil and passed through a silica plug (95:5 [pentane:EtOAc]). The orange solution of product was again concentrated *in vacuo* and residual solvent removed on a strong vacuum manifold to give a bright orange oil in quantitative yield of sufficient purity to move forward. HRMS (ESI-Orbitrap) *m/z*: [M+H]⁺ Calcd for C₁₉H₃₃D₆OS₂, 353.2813; found 353.2844. ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.12-1.16 (m, 2H), 1.26 (br s, 20H), 1.38-1.42 (m, 2H), 1.46-1.50 (m, 1H), 1.77-1.83 (p, 2H), 2.56, 2.77 (sp s, 3H), 4.20-4.23, 4.58-4.60 (sp t, 2H)

flame-dried short-path distillation head equipped with a round-bottom receiving flask of appropriate size to accommodate purified product. The closed system is evacuated and charged with argon several times, then the product is distilled by gently heating the reaction mixture again with a Meker-Fischer burner to push the alkene to the receiving flask. Heating should continue till the remaining liquid in the distilling flask is a thick black oil that no longer distills alkene **5**. At this time the pungent crude and faint yellow product is further purified via normal-phase flash chromatography with 100% pentane to give a clear and colorless oil **5** after solvent is removed. Yield = 89%. TLC R_f = 1.0 (100% pentane). HRMS (ESI-Orbitrap) m/z : $[M-H]^-$ Calcd for $C_{17}H_{27}D_6$, 243.2964; found 243.2986. 1H -NMR ($CDCl_3$, 300.08 MHz) δ 1.12-1.16 (m, 2H), 1.26 (br s, 20H), 1.35-1.39 (m, 2H), 1.46-1.49 (t, 1H), 2.02-2.06 (qt, 2H), 4.91-4.94 (dp, 1H), 4.97-5.02 (dq, 1H), 5.78-5.86 (m, 1H)



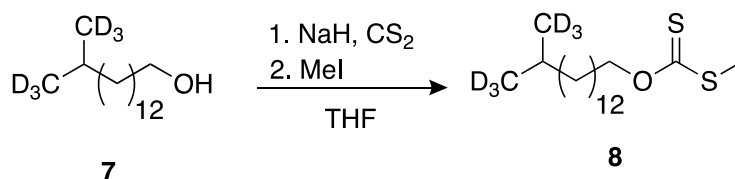
14-(methyl-d₃)pentadecanoic-15,15,15-d₃ acid (D₆-iC16-FA; **6):** To a 1 L flask equipped with a hefty stir bar is added a solution of **5** (2.0 g, 8.1 mmol) and Bu₄NBr (1.92 g, 4.1 mmol) in DCM and chilled to 0 °C with an ice/water bath. In a separate Erlenmeyer flask KMnO₄ (6.4 g, 40.1 mmol) is fully dissolved in DI H₂O (110 mL) and the homogenous solution was poured directly into the reaction flask. Finally, glacial acetic acid (23 mL) was added to the well-stirred biphasic system and was allowed to warm to room temperature and stir for 24 hours. The resultant dark brown sludge was chilled to 0 °C and acidified by the addition of a cold solution of HCl (5M, 110 mL). A saturated solution of Na₂SO₃ (typically >500 mL) was added in small portions to this mixture with vigorous gas evolution

until the aqueous portion of the mixture appears slightly cloudy and colorless, with the organic layer being a clear pale-yellow solution. The organic layer was separated from the mixture, and the aqueous phase extracted with DCM (5 x 50 mL). The pooled organic solution was dried over anhydrous MgSO₄, filtered, and concentrated *in vacuo*. Crude product was purified via normal-phase flash chromatography. Yield = 93%. TLC R_f = 0.76 (78:20:2 [pentane:EtOAc:AcOH]). HRMS (ESI-Orbitrap) *m/z*: [M-H]⁻ Calcd for C₁₆H₂₅D₆O₂, 261.2706; found 261.2708. ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.12-1.16 (m, 2H), 1.25 (br s, 18H), 1.46-1.48 (m, 1H), 1.60-1.63 (m, 2H), 2.34-2.37 (t, 2H)

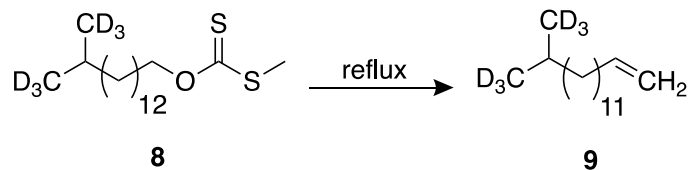


14-(methyl-d3)pentadecan-15,15,15-d3-1-ol (D₆-iC16-ol; 7): A flame-dried round bottom flask equipped with a dropping addition funnel was charged with compound **6** (2.71 g, 10.6 mmol) and 20 mL of dry THF was added to the starting material in the round bottom flask and chilled to 0 °C with an ice-water bath. 10 mL of dry THF was loaded into the dropping addition funnel along with BH₃·SMe₂ (1.21 mL, 12.8 mmol), and subsequently added dropwise to the chilled solution of **6** under an inert atmosphere. This reaction mixture was allowed to stir at this temperature for about 45 minutes, or until the cessation of gas evolution was observed. At this time, the reaction was warmed to room temperature and was stirred for an additional 12 hours. Once the reaction was complete, MeOH was added to the reaction mixture slowly to quench the remaining borane reagent. The crude solution was dried over anhydrous MgSO₄, filtered, and concentrated *in vacuo*. Crude product was purified via normal-phase flash chromatography. Yield = 78%. TLC R_f =

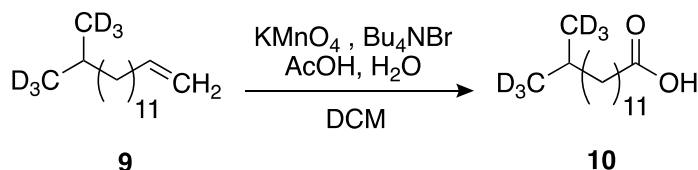
0.95 (10% Et₂O/hexanes). HRMS (ESI-Orbitrap) m/z : [M+H]⁺ Calcd for C₁₆H₂₉D₆O, 249.3059; found 249.3068. ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.12-1.16 (m, 2H), 1.26 (br s, 20H), 1.46-1.49 (t, 1H), 1.54-1.59 (m, 2H), 3.63-3.65 (t, 2H)



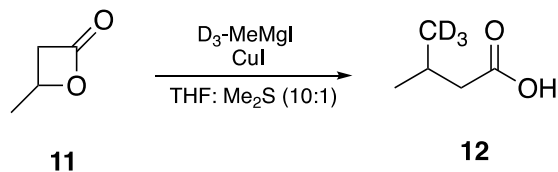
S-methyl O-(14-(methyl-d3)pentadecyl-15,15,15-d3) carbonodithioate (D₆-iC16-xanthate; 8): Refer to the procedure used to generate compound **4** using the same molar equivalents of reagents and relative solvent ratios to amount of starting material. The rusty-brown product **8** was taken up in enough pentane to drop the viscosity of the oil and passed through a silica plug (95:5 [pentane:EtOAc]). The orange solution of product was again concentrated *in vacuo* and residual solvent removed on a strong vacuum manifold to give the bright orange oil **8** in quantitative yield of sufficient purity to move forward. HRMS (ESI-Orbitrap) m/z : [M+H]⁺ Calcd for C₁₈H₃₁D₆OS₂, 339.2657; found 339.2695.



14-(methyl-d3)pentadec-1-ene-15,15,15-d3 (D₆-iC16-alkene; 9): Refer to the procedure used to generate compound **5** using the same molar equivalents of reagents and relative solvent ratios to amount of starting material. The pungent crude and faint yellow product **9** is further purified via normal-phase flash chromatography with 100% pentane to give a clear and colorless oil **9** after solvent is removed. Yield = 89%. TLC R_f = 1.0 (100% pentane). HRMS (ESI-Orbitrap) m/z : $[\text{M}-\text{H}]^-$ Calcd for C₁₆H₂₅D₆, 229.2808; found 229.2810. ¹H-NMR (CDCl₃, 300.08 MHz) δ 1.11-1.15 (m, 2H), 1.25 (br s, 16H), 1.34-1.39 (m, 2H) 1.45-1.49 (t, 1H), 2.01-2.05 (qt, 2H), 4.90-4.93 (dp, 1H), 4.96-5.00 (dq, 1H), 5.76-5.85 (m, 1H)



13-(methyl-d3)tetradecanoic-14,14,14-d3 acid (D₆-iC15-FA; 10): Refer to the procedure used to generate compound **6** using the same molar equivalents and relative solvent ratios. Crude product **10** was purified via normal-phase flash chromatography. Yield = 96%. TLC R_f = 0.68 (78:20:2 [pentane:EtOAc:AcOH]). HRMS (DART) m/z : $[\text{M}-\text{H}]^-$ Calcd for C₁₅H₂₃D₆O₂, 247.2550; found 247.2553. ¹H-NMR (CDCl₃, 499.73 MHz) δ 1.12-1.17 (m, 2H), 1.26 (br s, 16H), 1.46-1.50 (m, 1H), 1.59-1.67 (p, 2H), 2.33-2.38 (t, 2H)



3-methylbutanoic-4,4,4-d3 acid (D₃-iC5-FA; **12):** To a 3 L flame-dried 3-neck flask equipped with a mechanical stirring rod was added anhydrous copper(I) iodide (32.6 g, 171 mmol) followed by freshly distilled CS₂ (80 mL, 1.32 mol), which was allowed to dissolve fully before addition of dry THF (800 mL) all under an inert atmosphere. This clear yellow solution was chilled to -78 °C and a freshly generated solution of D₃MeMgI in Et₂O (3.4 M, 100 mL) was added dropwise via cannula, which produced a fairly viscous and cloudy yellow slurry of the diorganocuprate intermediate. This solution was warmed to -40 °C and allowed to stir for an additional 30 minutes prior to addition of electrophile. At this time the reaction was kept at -40 °C during addition of lactone **11** (12.7 mL, 156 mmol), which was added dropwise via syringe. During addition of the lactone, the yellow color subsided very minimally, and this mixture was allowed to naturally warm to room temperature overnight with constant vigorous stirring.

Overnight the cloudy solution had turned a gnarly deep brown color with a rancid odor of isovaleric acid. The reaction was quenched by the addition of copious amounts of a saturated solution of NH₄Cl until slightly acidic (pH ~ 6.0), which gave a biphasic system consisting of a rusty orange organic phase and a deep blue aqueous phase. The layers were separated, the organic phase was filtered first through a paper filter covered with sand to remove most of the insoluble Mg/Cu salts, and then filtered again through a plug of Celite™ 545 filter aid. The clear organic solution was again washed with NH₄Cl to remove any residual copper species indicated by the blue Cu-NH₄ complex, washed with

brine, dried over anhydrous MgSO_4 , filtered, and concentrated *in vacuo*. Crude product was purified via fractional distillation through a glass bead-packed fractionation column. Inseparable by-product of crotonic acid was reduced to isovaleric acid with Pd/C under 1 atm of H_2 for 3-hours. Yield = 44%. HRMS (ESI-Orbitrap) m/z : $[\text{M}-\text{H}]^-$ Calcd for $\text{C}_5\text{H}_6\text{D}_3\text{O}_2$, 104.0796; found 104.0782. ^1H -NMR (CDCl_3 , 499.73 MHz) δ 0.96-0.98 (d, 3H), 2.05-2.12 (q, 1H), 2.21-2.23 (d, 2H).

2.4.2 – Cell culturing and biological incorporation of tracing lipids

The procedure for culturing of *M. xanthus* cells was provided by Dr. Tye O'hara Boynton of the Shimkets group and began with growing 20 mL cultures overnight in CYE broth (1% casitone, 0.5% yeast extract, 0.1% MgSO_4 , and 10 mM MOPS) at 32 °C. Each culture was poured into its own respective 1 L of CYE broth and incubated further until the culture reached an optical density (OD) of $\sim 3 \times 10^8$ cells/mL. These inoculated broths were centrifuged at 10,000 RPM for 10 minutes, then the large cell pellet was resuspended in fresh CYE broth until an OD of 5×10^9 cells/mL was reached. At this point 8 mL of culture was spread onto 500 mL of TPM agar (10mM Tris-HCl, 8mM MgSO_4 , 1mM potassium phosphate buffer @ pH = 7.6, 1.5% agar) and allowed to absorb to dryness, then covered and incubated at 32 °C for 24 hours. This was the procedure for growing the control samples for lipid analysis. For supplemented cultures of *M. xanthus* synthetic lipids or synthetic isotope tracing lipids were introduced during the last step by including the lipids in the TPM agar at a fairly high concentration (100 μM) to ensure suitable cellular uptake.

2.4.3 – Lipid extraction, preparative TLC, and sample preparation

Lipids were extracted following a method developed and provided by Shimkets and co-workers. From the previous culturing procedure (**section 2.4.2**), cell cultures from the TPM plates that had been incubated for 24 hours were harvested into a 50 mL conical tube, two 500 mL plates per tube. At this time harvested cells were resuspended in 6 mL

of Milli-Q H₂O and then 22.5 mL of a [1:2] mixture of CHCl₃:MeOH was added to the tube, vortexed for three minutes and chilled on ice for 1 hour. After chilling, an additional 6 mL of CHCl₃ was added and vortexed for 1 minute, then a 6 mL portion of Milli-Q H₂O was added to the tube and vortexed for another minute. The extraction mixture was then centrifuged at 10,000 RPM for 10 minutes to separate the biphasic mixture and the organic phase was collected into a separate glass container and the solvent was removed under a stream of nitrogen. Dried lipids were re-dissolved in a [1:2] mixture of CHCl₃:MeOH and quantitatively transferred to a clean and pre-weighed glass vial. Solvent was again removed under a stream of nitrogen and the dried lipid samples were weighed prior to storage in a -80 °C freezer.

These whole-cell extracts were further fractionated to isolate the lipid species present in the sample from other chloroform-soluble metabolites and proteinaceous cell components following a procedure previously published by Hamilton and Comai.⁸⁵ It should be noted that this procedure separated lipids into different lipid classes, but the lipid fractions collected at the end of this procedure were pooled and sent to the BSMMS for total lipid analysis. This procedure was also used to isolate partial glycerides for preparative TLC, so this procedure served a dual-purpose. Briefly, the lipid fraction of whole-cell extracts was achieved using prepacked silica Sep-pak columns (Supelco; Supelclean LC-Si SPE; 12 mL; 2 g) by the following procedure. Preparation of a Sep-pac column for cell extract fractionation began with washing the column with 16 mL of a [96:4] mixture of *n*-hexane:MTBE followed by 48 mL of *n*-hexane. A 10 mg portion of whole-cell lipid extract was dissolved in 16 mL of a [200:3] *n*-hexane:MTBE mixture and loaded onto the prepared Sep-pak column. Ether lipids were eluted as the first fraction by passing 40 mL of [200:3] *n*-hexane:MTBE through the lipid-loaded column. Triglycerides were eluted in the second fraction by passing 48 mL of [96:4] *n*-hexane:MTBE through the column. Free fatty acids were eluted in the third fraction by first acidifying the loaded column with 40 mL of [100:0.2] *n*-hexane:AcOH then by passing

48 mL of [100:2:0.2] *n*-hexane:MTBE:AcOH through the column. Monoacyl- and diacyl-glycerides were eluted in the fourth fraction by passing 48 mL of [100:0.2] MTBE:AcOH through the column. The fifth and final fraction eluted from the column contained a mixture of phospholipids and xanthene compounds by passing 48 mL of [100:4:1] MTBE:methanol:(1 mM NH₄OAc). For total lipid analysis these separated lipid fractions were pooled and dried under a stream of nitrogen prior to storage at -20 °C.

Preparative TLC was performed on partial glycerides isolated as fraction four by the previous procedure. Normal-phase silica TLC plates were spotted with partial glycerides dissolved in [2:1] CHCl₃:MeOH and spotted on the TLC plates. Once dry, the plates were run with freshly prepared [85:15:5] toluene:CHCl₃:MeOH and visualized with transient iodine vapor staining. Five lipid fractions (R_f = 0.20, 0.24, 0.29, 0.35, and 0.42) from this separation were found to restore fruiting body formation in sp. LS1191 by extracellular complementation, were extracted from preparative TLC plate and submitted to the BSMMSC for lipidomic analysis. All lipid samples were prepared in the same manner prior to lipidomic analysis. For LCMS analysis 10-15 mg of extracted lipid were weighed into 1-dram glass vials, resuspended in 300 μ L of [9:1] CHCl₃:MeOH and sonicated briefly to ensure total dissolution. These dissolved lipid samples were centrifuged at 13,000 RPM for 5 minutes, quantitatively transferred to autosampler vials and subjected to lipidomic analysis.

2.4.4 – LC-HRMS method

The data was collected on an Exactive Plus Orbitrap mass spectrometer with an electrospray ionization (ESI) source coupled with an Ultimate3000 ultra high-performance liquid chromatography (UHPLC) system (Thermo Fischer Scientific, Pittsburgh, PA). Separations were performed on a Kinetex HILIC HPLC column (150×2.1 mm, 2.6 mm) (Phenomenex, Torrance, CA, USA) maintained at 25 °C. Mobile phase A was 10 mM ammonium formate adjusted with formic acid to pH = 3 and Mobile phase B was 93/7 (v/v) ACN/10 mM ammonium formate with the aqueous portion of the mixture adjusted

with formic acid to pH = 3. The flow rate was 0.2 mL/min and the mobile phase gradient used was 0-1 min 100% B, 1-15 min 100-81% B, 15.1-25 min 48% B, 25.1-35 min 100% B. The injection volume was 10 μ L for each sample, which was kept at 4 °C for the entire analysis. The ESI parameters were as follows; scan range 100-1500 m/z, resolution 140,000, AGC target 3.0×10^6 , MaxIT of 200 ms, sheath gas flow rate 25 units, aux gas flow rate 10 units, spray voltage ± 4.0 kV, and capillary temperature 350 °C. The mass spectrometry method was operated in both positive and negative mode using full scan analysis for each sample, once per ionization mode, for a total run-time of 1 hour and 10 minutes per sample.

2.4.5 – Data analysis

Untargeted lipidomics data was analyzed using a MAVEN, which is an open-source program used to generate extracted ion chromatograms (EIC) and align samples according to retention time. MAVEN is capable of compiling a list of chromatographic peaks based on a set of parameters which determine how selective the program is in the peaks it extracts from the total ion chromatogram (TIC). This feature is used to gather potentially meaningful data within a sample set in an untargeted fashion. A pseudo-targeted approach to analyzing this data uses a pre-generated list of chemical formulas, which provide the program with exact masses to extract chromatographic peaks from the data generated by the mass spectrometer. The list used for the targeted method was adopted from a previously published study by Dr. Helge B. Bode, which detailed a comprehensive list of lipids found in wild-type *M. xanthus*.⁸⁰ This exhaustive and comprehensive list of lipids and their relative abundance in *M. xanthus* was used to mine the data obtained from this lipidomics study to determine what lipid species were significantly different between the experimental groups.

2.5 – Conclusion

The initial goal of investigating lipids produced during starvation-induced sporulation in *M. xanthus* was to determine what molecule(s) were responsible for the initiation of fruiting body formation in this microbial system. Differences in the lipid profile between wild-type and mutant cell lines did not provide strong evidence that pointed to a single compound or class of compounds that could explain this phenotypic difference; Although, it was apparent that the wild-type strain DK1622 was better-able to produce larger total quantities of lipids than their mutant counterpart, strain LS1191. The relative ratios of detected lipid species were similar between these experimental groups. This finding was not sufficient to explain why LS1191 did not produce lipid-bodies under nutrient limitation. Being that LS1191 is a strain deficient in the production of branched-chain lipids, it could be that the structural attributes of the lipids produced in this mutant were contributing to the lack of a successful sporulation event. Due to the large abundance of isolipid species in *M. xanthus*, especially isoC15:0, this led to the idea that the presence of certain lipid(s) that contained an isopropyl moiety may be the determining factor in initiating fruiting body formation; However, these isolipids have many isobaric counterparts with differing degrees of tail branching or lack-there-of. The isotope tracing experiments devised to mitigate this issue were quite successful in terms of allowing for the biological generation of isotopically labeled isolipids in LS1191 to be compared against lipids identified in DK1622. In a parallel complementation study headed by the lab of Dr. Shimkets, preparative TLC fractions of whole-cell lipid extracts were assayed for their ability to restore fruiting body formation in LS1191, but their findings showed that several separate TLC fractions were able to restore sporulation behavior in the mutant strain. Analysis of these fractions by LCMS again showed that the lipid profiles between TLC fractions that did not result in fruiting bodies and the fractions that did encourage fruiting body production were very similar in relative ratios of lipids but showed different concentrations of lipids being contained within those fractions. It was

then thought that the unknown lipid(s) that could be initiators of fruiting body formation maybe some previously unobserved lipid species that contained the isolipid structural feature, and could be mined from the unannotated LCMS data, or what is referred to as the “unknowns”. From here, a lipid tracer study was implemented to make identification of these possible unknown compounds much more achievable.

The first labeling study attempted to incorporate an isotopically labeled synthetic lipid, isoC₁₅-d₆, which was intended to be taken up directly by starving *M. xanthus* cells and directly used to make the signaling lipid(s) thought to contain at least one isoC₁₅ tail. The uptake of this lipid was not ideal, and lead to a dearth of isotope signal in the wild-type and mutant cultures. This was likely caused by the innate hydrophobicity of the lipid tag, which lead to poor cellular uptake, and ultimately a lack of isotope incorporation into biogenic lipids. To get around this issue of poor cellular uptake of the isotope tracer, it was postulated that incorporation of a much smaller, more permeable, and biosynthetically more useful lipid synthon to the organism could be used. While keeping the chemistries of synthesizing the isotopically labeled isopropyl moiety similar to that in the strategy of that for isoC₁₅-d₆ by using a D₃-methyl Grignard reagent, isovaleric acid became the synthetic target for the new isotope tracer. LCMS analysis of the lipid fraction of whole-cell extracts from *M. xanthus* strains DK1622 and LS1191 following supplementation with iC₅-d₃ showed a high-level of incorporation of the D₃ isotope signal. This allowed for the distinction of what mass features contained a terminal isopropyl unit, but also made the task of determining the lipids that may initiate sporulation to be a daunting task, especially when considering the unknown mass features present in the sample. This finding though that the organisms incorporated the isolipids tracer into so many things was not unlikely since this organism has been previously shown to survive quite well without the ability to make straight-chain lipids. Dr. Helge B. Bode had shown that a mutant of *M. xanthus* that could not build its own straight-chain lipids shifted practically all lipid production to that of branched-chain lipids with no observable changes

in morphology or behavioral phenotype.⁸¹ Because of this, the LS1191 mutant that could not make its own branched chain lipids from amino acids proved to be quite effective in a proof-of-concept microbe for testing the incorporation of deuterated isovaleric acid into the organisms lipidome.

A single lipid, or class of lipids, responsible for fruiting body formation was not identified by either isotope tracing, or complementation assays of preparative TLC fractions. Due to the complexity of lipid fractions that prompted this phenotypic response by complementation studies, and the promiscuity of the iC_5-d_3 isotope tracer, the focus of the project quickly shifted toward exploiting iC_5-d_3 as a means to differentiate straight-chain lipids from isobars containing a terminal isopropyl unit. This was verified by identifying specific labeling of homospermidine lipids produced during fruiting body formation, corroborating the findings of Müller and co-workers.

The work presented in this chapter describes the utility of isotope tracing in lipidomic analyses of complex biological samples using LCMS instrumentation. Through this work it has been found that there are clear advantages in using small polar lipids as an isotope tracer even when the target lipids species are larger molecules. The author has presented an isotope-tracing experiment for detecting compounds in a typical lipidomic analysis to differentiate straight-chain lipid species from isobaric congeners containing a terminal isopropyl moiety. These isobaric species are difficult to separate chromatographically and/or differentiate in the mass-domain without the use of hard-ionization sources to induce characteristic fragmentation. This method of feeding bacterial cultures iC_5-d_3 has an important application for those undertaking LCMS analyses of biological systems, since a large and diverse pool of biogenic lipids can be derived from isovaleric acid. The results from an experiment such as this would illuminate mass features that contain terminal isopropyl units, while the task of determining the exact structure of the parent molecule would still need considerable work in determining the precise chemical formula and exact structure of the detected compound.

2.6 – Chapter II appendix

Table 2.6.1 – Table of Bode lipidomics data (DK1622 vs LS1191)

Shown below is a table of detected masses representative of lipid species that are known to be present in wild-type *M. xanthus* DK1622.⁸⁰

Potential Compound	RT (min)	Molecular Species	Detected Mass (<i>m/z</i>)	Ion Counts		p-value
				DK1622	LS1191	
Cer(d19:0/16:0)	1.6	[M+H] ⁺	554.5506	4.06E+06	2.35E+07	4.93E-04
Cer(d19:0/i17:0 2-OH)	1.6	[M+H] ⁺	584.5613	8.27E+06	7.49E+07	3.36E-04
Cer(d19:0/i17:0)	1.7	[M+H] ⁺	568.5665	9.30E+06	1.32E+08	1.04E-04
CerPI(d19:0/16:0 2-OH)	12.0	[M+H] ⁺	812.5635	4.14E+06	3.17E+06	4.76E-02
CerPI(d19:0/16:0)	10.9	[M+H] ⁺	796.5689	3.05E+06	2.79E+06	6.27E-01
Cer-PI(D19:0/i17:0 2-OH)	11.9	[M+H] ⁺	826.5796	1.23E+07	5.66E+06	1.96E-03
CerPI(D19:0/i17:0)	11.9	[M+H] ⁺	810.5851	8.41E+06	4.23E+06	4.19E-03
CL(60:0)	9.7	[M+H] ⁺	1297.924	4.16E+06	2.45E+06	5.72E-02
CL(60:1)	9.9	[M+H] ⁺	1295.906	9.31E+05	9.75E+05	8.12E-01
CL(60:2)	9.9	[M+H] ⁺	1293.892	1.22E+05	1.55E+05	4.15E-01
CL(60:3)	-	-	-	-	-	-
CL(61:1)	9.5	[M+H] ⁺	1309.922	1.82E+06	2.82E+06	3.39E-02
CL(61:2)	9.8	[M+H] ⁺	1307.909	5.90E+05	7.25E+05	4.48E-02
CL(61:3)	9.7	[M+H] ⁺	1305.893	9.12E+04	1.48E+05	1.88E-02
CL(61:4)	-	-	-	-	-	-
CL(62:0)	9.8	[M+H] ⁺	1325.952	1.08E+07	9.36E+06	1.86E-01
CL(62:1)	9.4	[M+H] ⁺	1323.938	3.97E+06	6.20E+06	7.98E-04
CL(62:2)	9.6	[M+H] ⁺	1321.923	1.62E+06	2.65E+06	6.27E-03
CL(62:3)	9.3	[M+H] ⁺	1319.908	3.61E+05	7.60E+05	3.81E-03
CL(62:4)	9.2	[M+H] ⁺	1317.893	5.26E+04	9.65E+04	1.33E-01
CL(62:5)	-	-	-	-	-	-
CL(63:0)	9.4	[M+H] ⁺	1339.968	1.06E+07	8.58E+06	1.16E-01
CL(63:1)	9.6	[M+H] ⁺	1337.954	6.72E+06	8.58E+06	4.99E-02
CL(63:2)	9.2	[M+H] ⁺	1335.939	3.24E+06	5.80E+06	1.53E-03
CL(63:3)	9.6	[M+H] ⁺	1333.925	1.21E+06	2.56E+06	7.99E-04
CL(63:4)	9.5	[M+H] ⁺	1331.91	2.74E+05	4.69E+05	2.38E-02
CL(63:5)	-	-	-	-	-	-
CL(64:0)	9.6	[M+H] ⁺	1353.98	3.54E+06	5.74E+06	4.27E-03
CL(64:1)	9.7	[M+H] ⁺	1351.97	7.28E+06	8.64E+06	8.06E-02
CL(64:2)	9.4	[M+H] ⁺	1349.953	4.76E+06	7.90E+06	8.55E-05
CL(64:3)	9.5	[M+H] ⁺	1347.94	2.57E+06	5.05E+06	5.59E-04
CL(64:4)	9.5	[M+H] ⁺	1345.922	8.10E+05	1.58E+06	2.87E-03
CL(64:5)	9.5	[M+H] ⁺	1343.909	1.89E+05	2.63E+05	1.32E-01
CL(64:6)	-	-	-	-	-	-
CL(65:1)	9.5	[M+H] ⁺	1365.982	2.45E+06	5.98E+06	3.94E-04
CL(65:2)	9.4	[M+H] ⁺	1363.97	3.65E+06	6.74E+06	9.91E-03

Table 2.6.1 (continued)

Potential Compound	RT (min)	Molecular Species	Detected Mass (<i>m/z</i>)	Ion Counts		p-value
				DK1622	LS1191	
CL(65:3)	9.6	[M+H] ⁺	1361.958	2.98E+06	5.67E+06	6.11E-03
CL(66:1)	9.8	[M+H] ⁺	1380.003	9.77E+05	2.87E+06	8.48E-04
CL(66:2)	9.7	[M+H] ⁺	1377.988	1.72E+06	4.06E+06	2.12E-04
DG(30:0)	1.6	[M+NH ₄] ⁺	558.5092	3.49E+06	1.90E+06	2.48E-02
DG(31:1)	1.6	[M+NH ₄] ⁺	570.5092	5.31E+05	1.18E+06	1.45E-02
DG(32:0)	1.7	[M+NH ₄] ⁺	586.5405	3.45E+05	1.20E+06	5.77E-05
DG(32:2)	1.6	[M+NH ₄] ⁺	582.5092	4.45E+05	2.67E+06	1.19E-05
DG(33:0)	1.7	[M+NH ₄] ⁺	600.5561	6.45E+05	4.32E+06	2.85E-07
DG(33:1)	1.6	[M+NH ₄] ⁺	598.5405	2.58E+05	4.28E+05	1.40E-02
DG(33:2)	1.6	[M+NH ₄] ⁺	596.5248	2.37E+05	8.81E+05	1.03E-04
PE(28:0)	9.6	[M+H] ⁺	636.459	3.89E+07	2.58E+07	5.67E-02
PE(28:1)	10.0	[M+H] ⁺	634.4438	8.19E+06	9.56E+06	6.36E-01
PE(28:2)	9.3	[M+H] ⁺	632.4258	1.87E+06	1.37E+06	2.09E-01
PE(29:0)	9.6	[M+H] ⁺	650.475	1.97E+08	9.36E+07	1.77E-02
PE(29:1)	9.9	[M+H] ⁺	648.4591	5.34E+07	5.11E+07	5.33E-01
PE(29:2)	9.5	[M+H] ⁺	646.4433	6.62E+06	7.05E+06	6.12E-01
PE(30:0)	10.0	[M+H] ⁺	664.4906	3.12E+08	1.51E+08	2.25E-03
PE(30:1)	8.1	[M+H] ⁺	662.4734	8.31E+07	1.37E+08	8.19E-04
PE(30:2)	9.4	[M+H] ⁺	660.4596	2.49E+07	5.03E+07	7.26E-04
PE(30:3)	9.6	[M+H] ⁺	658.4434	3.30E+06	3.65E+06	2.03E-01
PE(31:0)	9.9	[M+H] ⁺	678.5056	7.36E+06	7.05E+07	2.11E-03
PE(31:1)	9.5	[M+H] ⁺	676.4907	1.20E+08	1.41E+08	1.05E-01
PE(31:2)	9.7	[M+H] ⁺	674.4752	7.04E+07	1.09E+08	1.11E-02
PE(31:3)	9.5	[M+H] ⁺	672.4574	1.75E+07	2.50E+07	2.56E-02
PE(31:4)	9.5	[M+H] ⁺	670.4428	1.14E+06	1.03E+06	2.32E-01
PE(32:0)	10.0	[M+H] ⁺	692.5217	2.61E+07	3.15E+07	2.58E-01
PE(32:1)	9.6	[M+H] ⁺	690.5069	1.95E+07	7.77E+07	1.16E-03
PE(32:2)	9.5	[M+H] ⁺	688.4913	9.11E+07	1.37E+08	3.06E-02
PE(32:3)	9.6	[M+H] ⁺	686.4743	5.43E+07	4.92E+07	3.76E-01
PE(32:4)	9.6	[M+H] ⁺	684.4586	1.31E+07	7.98E+06	2.92E-02
PE(33:0)	9.4	[M+H] ⁺	706.5372	7.85E+05	6.89E+06	2.47E-02
PE(33:1)	9.6	[M+H] ⁺	704.5215	1.47E+07	1.54E+07	7.37E-01
PE(33:2)	8.4	[M+H] ⁺	702.5044	2.58E+07	1.64E+07	2.22E-03
PE(33:3)	9.2	[M+H] ⁺	700.4893	3.40E+07	9.88E+06	2.20E-05
PE(33:4)	9.2	[M+H] ⁺	698.4742	1.69E+07	2.76E+06	3.25E-03
PE(34:0)	9.4	[M+H] ⁺	720.5526	7.60E+05	1.29E+06	1.20E-01
PE(34:1)	9.5	[M+H] ⁺	718.5378	1.92E+06	9.78E+05	3.88E-03
PE(34:2)	8.2	[M+H] ⁺	716.5219	4.31E+06	1.66E+06	2.02E-03
PE(34:3)	9.2	[M+H] ⁺	714.5054	7.21E+06	1.64E+06	3.77E-03
PE(34:4)	9.3	[M+H] ⁺	712.4892	5.08E+06	1.93E+06	4.56E-04
PE(P-30:0)	9.5	[M+H] ⁺	648.4957	8.46E+07	2.96E+06	2.54E-02
PE(P-32:0)	8.3	[M+H] ⁺	676.5249	5.96E+05	-	3.72E-03

Table 2.6.1 (continued)

Potential Compound	RT (min)	Molecular Species	Detected Mass (<i>m/z</i>)	Ion Counts		p-value
				DK1622	LS1191	
PG(29:0)	5.1	[M+H] ⁺	681.4681	1.73E+06	5.41E+05	2.72E-04
PG(30:0)	5.6	[M+H] ⁺	695.4847	2.88E+06	1.50E+06	1.59E-05
PG(30:1)	5.1	[M+H] ⁺	693.469	1.49E+06	1.66E+06	5.84E-02
PG(31:1)	4.2	[M+H] ⁺	707.4837	2.34E+06	3.64E+06	3.96E-03
PG(31:2)	4.3	[M+H] ⁺	705.4707	8.95E+05	9.23E+05	6.87E-01
PG(32:1)	3.8	[M+H] ⁺	721.5027	1.46E+06	3.34E+06	7.66E-06
PG(32:2)	4.4	[M+H] ⁺	719.4869	2.80E+06	1.70E+06	4.49E-04
PG(32:3)	5.4	[M+H] ⁺	717.4675	5.74E+05	2.83E+05	1.02E-02
PG(33:0)	3.5	[M+H] ⁺	737.5277	-	3.73E+04	2.01E-01
PG(33:1)	4.5	[M+H] ⁺	735.5156	1.48E+06	8.20E+05	7.06E-06
PG(33:2)	5.0	[M+H] ⁺	733.4995	6.47E+05	1.70E+05	5.12E-05
PG(34:0)	5.7	[M+H] ⁺	751.5456	0.00E+00	3.21E+03	9.15E-02
PG(34:1)	5.2	[M+H] ⁺	749.5319	5.72E+04	1.48E+04	3.02E-02
TG(38:1)	1.6	[M+NH ₄] ⁺	782.7232	9.28E+06	2.98E+06	5.10E-02
TG(39:0)	1.6	[M+NH ₄] ⁺	684.6136	7.50E+05	2.68E+05	5.17E-03
TG(39:1)	1.5	[M+NH ₄] ⁺	698.6293	7.11E+05	2.60E+05	2.36E-03
TG(40:1)	1.5	[M+NH ₄] ⁺	696.6136	2.48E+05	2.16E+05	2.49E-01
TG(40:2)	1.5	[M+NH ₄] ⁺	710.6293	4.89E+05	2.89E+05	5.02E-02
TG(41:1)	1.5	[M+NH ₄] ⁺	708.6136	2.13E+05	1.26E+05	7.66E-02
TG(41:2)	1.5	[M+NH ₄] ⁺	724.645	6.32E+05	5.07E+05	3.97E-04
TG(41:3)	1.5	[M+NH ₄] ⁺	722.6293	2.56E+05	2.21E+05	3.59E-01
TG(42:1)	1.5	[M+NH ₄] ⁺	720.6136	5.39E+04	3.66E+04	2.68E-01
TG(42:2)	1.5	[M+NH ₄] ⁺	738.6606	1.47E+06	8.97E+05	3.50E-02
TG(43:0)	1.5	[M+NH ₄] ⁺	736.6449	5.05E+05	4.42E+05	1.95E-01
TG(43:1)	1.5	[M+NH ₄] ⁺	754.6919	5.64E+06	1.54E+06	4.52E-03
TG(43:2)	1.7	[M+NH ₄] ⁺	752.6762	2.35E+06	1.74E+06	6.17E-02
TG(43:3)	1.5	[M+NH ₄] ⁺	750.6606	9.26E+05	8.28E+05	5.46E-02
TG(44:0)	1.5	[M+NH ₄] ⁺	748.6449	2.08E+05	1.97E+05	4.04E-01
TG(44:1)	1.6	[M+NH ₄] ⁺	768.7075	8.04E+06	2.08E+06	2.82E-02
TG(44:2)	1.5	[M+NH ₄] ⁺	766.6919	4.65E+06	2.50E+06	2.82E-03
TG(44:3)	1.6	[M+NH ₄] ⁺	764.6762	2.00E+06	1.53E+06	9.29E-03
TG(44:4)	1.5	[M+NH ₄] ⁺	762.6606	4.45E+05	4.40E+05	8.74E-01
TG(45:0)	1.5	[M+NH ₄] ⁺	760.6449	7.71E+04	7.86E+04	8.88E-01
TG(45:1)	1.5	[M+NH ₄] ⁺	780.7075	5.46E+06	3.82E+06	1.99E-01
TG(45:2)	1.5	[M+NH ₄] ⁺	778.6919	2.77E+06	2.19E+06	2.44E-02
TG(45:3)	1.6	[M+NH ₄] ⁺	776.6762	8.28E+05	6.81E+05	2.38E-02
TG(45:4)	1.5	[M+NH ₄] ⁺	774.6606	1.62E+05	1.38E+05	1.55E-01
TG(46:0)	1.7	[M+NH ₄] ⁺	796.7389	3.41E+06	2.87E+06	3.58E-01
TG(46:1)	1.6	[M+NH ₄] ⁺	794.7232	3.79E+06	3.66E+06	8.64E-01
TG(46:2)	1.6	[M+NH ₄] ⁺	792.7075	3.64E+06	2.33E+06	5.81E-03
TG(46:3)	1.5	[M+NH ₄] ⁺	790.6919	1.17E+06	9.02E+05	1.03E-02
TG(47:0)	1.7	[M+NH ₄] ⁺	810.7545	2.85E+06	2.35E+06	5.06E-01

Table 2.6.1 (continued)

Potential Compound	RT (min)	Molecular Species	Detected Mass (<i>m/z</i>)	Ion Counts		p-value
				DK1622	LS1191	
TG(47:1)	1.5	[M+NH ₄] ⁺	808.7388	1.58E+06	2.33E+06	1.69E-01
TG(47:2)	1.6	[M+NH ₄] ⁺	806.7232	2.16E+06	1.65E+06	3.99E-02
TG(47:3)	1.5	[M+NH ₄] ⁺	804.7075	9.57E+05	7.89E+05	5.90E-02
TG(47:4)	1.5	[M+NH ₄] ⁺	802.6919	3.14E+05	2.31E+05	2.35E-02
TG(48:0)	1.9	[M+NH ₄] ⁺	824.7701	9.59E+05	1.46E+06	4.49E-01
TG(48:1)	1.8	[M+NH ₄] ⁺	822.7545	7.00E+05	1.07E+06	7.76E-02
TG(48:2)	1.6	[M+NH ₄] ⁺	820.7388	9.18E+05	8.16E+05	5.95E-01
TG(48:3)	1.6	[M+NH ₄] ⁺	818.7232	6.33E+05	4.37E+05	2.56E-02
TG(48:4)	1.5	[M+NH ₄] ⁺	816.7075	2.49E+05	1.43E+05	4.73E-03
TG(49:0)	2.2	[M+NH ₄] ⁺	838.7858	6.39E+05	6.07E+05	8.82E-01
TG(49:1)	1.6	[M+NH ₄] ⁺	836.7701	2.99E+05	3.63E+05	3.64E-01
TG(49:2)	1.5	[M+NH ₄] ⁺	834.7545	3.47E+05	2.32E+05	1.71E-01
TG(50:0)	2.2	[M+NH ₄] ⁺	852.8014	1.99E+05	1.47E+05	3.63E-01
TG(50:1)	1.8	[M+NH ₄] ⁺	850.7858	1.79E+05	8.81E+04	5.29E-02
TG(51:0)	2.0	[M+NH ₄] ⁺	866.8171	5.91E+04	2.10E+04	1.99E-01
TG(51:1)	2.0	[M+NH ₄] ⁺	864.8014	6.13E+04	9.00E+03	9.61E-02
TG(52:0)	2.2	[M+NH ₄] ⁺	880.8327	5.63E+04	2.53E+03	1.73E-01

Table 2.6.2 – Table of lipidomics data showing incorporation of *iC5-d₃*

Shown below are the signal intensities for the potential lipid species known to be present in *M. xanthus* DK1622.⁸⁰ Lipid species that were tagged with the *iC5-d₃* isotope tracer are annotated as detected in the *iC5-d₃*-doped LS1191 sample by a checkmark. Intensities for the tagged lipid species are not detailed.

Potential Compound	Control Ion Counts		IVA Suppl. Ion Counts		HIVA Suppl. Ion Counts		² H in LS1191-HIVA			
	DK1622	LS1191	DK1622	LS1191	DK1622	LS1191	D ₃	D ₆	D ₉	D ₁₂
Cer(d19:0/16:0)	2.36E+07	1.76E+08	6.50E+06	4.10E+07	1.55E+07	8.17E+06	✓	✓	-	-
Cer(d19:0/i17:0 2-OH)	6.08E+07	7.95E+08	1.36E+07	1.14E+08	4.89E+07	9.47E+06	✓	✓	-	-
Cer(d19:0/i17:0)	5.44E+07	1.09E+09	1.22E+07	1.33E+08	5.65E+07	9.18E+06	✓	✓	-	-
CerPI(d19:0/16:0 2-OH)	3.40E+08	4.02E+08	3.01E+08	1.97E+08	1.88E+08	-	✓	✓	-	-
CerPI(d19:0/16:0)	1.78E+08	2.67E+08	1.58E+08	1.79E+08	1.68E+08	2.97E+07	✓	✓	-	-
Cer-PI(D19:0/i17:0 2-OH)	8.89E+08	8.84E+08	7.45E+08	6.21E+08	6.52E+08	2.59E+07	✓	✓	-	-
CerPI(D19:0/i17:0)	3.09E+08	3.88E+08	2.54E+08	4.06E+08	4.25E+08	1.11E+07	✓	-	-	-
CL(60:0)	4.02E+07	2.13E+07	6.39E+07	6.29E+07	3.75E+07	2.52E+06	✓	✓	✓	✓
CL(60:1)	1.35E+07	1.08E+07	2.61E+07	3.82E+07	1.40E+07	4.32E+05	✓	✓	✓	✓
CL(60:2)	2.27E+06	2.45E+06	5.12E+06	9.13E+06	2.67E+06	1.73E+05	✓	✓	✓	✓
CL(60:3)	-	-	-	-	-	-	✓	✓	-	-
CL(61:1)	4.85E+07	4.73E+07	6.76E+07	7.50E+07	5.18E+07	3.33E+06	✓	✓	✓	✓
CL(61:2)	1.32E+07	1.95E+07	2.22E+07	2.58E+07	1.50E+07	1.01E+06	✓	✓	✓	✓
CL(61:3)	1.39E+06	2.97E+06	3.82E+06	6.73E+06	1.89E+06	1.71E+05	✓	✓	✓	-
CL(61:4)	-	-	-	-	-	-	✓	✓	-	-
CL(62:0)	1.87E+08	1.38E+08	1.84E+08	1.73E+08	2.03E+08	2.07E+07	✓	✓	✓	✓
CL(62:1)	1.16E+08	1.34E+08	1.33E+08	1.15E+08	1.23E+08	8.32E+06	✓	✓	✓	✓
CL(62:2)	4.90E+07	8.01E+07	6.45E+07	6.41E+07	5.15E+07	2.85E+06	✓	✓	✓	✓
CL(62:3)	1.09E+07	2.39E+07	1.82E+07	2.14E+07	1.21E+07	1.38E+06	✓	✓	✓	-
CL(62:4)	8.42E+05	1.93E+06	2.43E+06	3.81E+06	7.71E+05	1.76E+05	✓	✓	-	-
CL(62:5)	-	-	-	-	-	-	-	✓	-	-
CL(63:0)	1.71E+08	1.74E+08	1.52E+08	1.44E+08	1.74E+08	3.30E+07	✓	✓	✓	✓
CL(63:1)	1.92E+08	2.35E+08	1.88E+08	1.59E+08	1.97E+08	2.01E+07	✓	✓	✓	✓
CL(63:2)	1.16E+08	1.95E+08	1.29E+08	1.06E+08	1.16E+08	8.35E+06	✓	✓	✓	✓
CL(63:3)	4.11E+07	8.74E+07	5.61E+07	5.17E+07	3.99E+07	3.97E+06	✓	✓	✓	✓
CL(63:4)	7.76E+06	1.65E+07	1.36E+07	1.74E+07	6.93E+06	1.52E+06	✓	✓	✓	-
CL(63:5)	1.79E+05	5.88E+05	1.07E+06	2.52E+06	2.22E+05	1.08E+05	✓	✓	-	-
CL(64:0)	8.17E+07	1.53E+08	6.92E+07	5.77E+07	8.61E+07	4.82E+07	✓	✓	✓	✓

Table 2.6.2 (continued)

Potential Compound	Control Ion Counts		IVA Suppl. Ion Counts		HIVA Suppl. Ion Counts		² H in LS1191-HIVA			
	DK1622	LS1191	DK1622	LS1191	DK1622	LS1191	D ₃	D ₆	D ₉	D ₁₂
CL(64:1)	1.85E+08	2.74E+08	1.70E+08	1.32E+08	1.81E+08	2.42E+07	✓	✓	✓	✓
CL(64:2)	1.79E+08	3.02E+08	1.77E+08	1.45E+08	1.60E+08	1.13E+07	✓	✓	✓	✓
CL(64:3)	9.40E+07	1.89E+08	1.05E+08	9.24E+07	8.12E+07	6.39E+06	✓	✓	✓	✓
CL(64:4)	2.83E+07	5.52E+07	3.73E+07	3.42E+07	2.47E+07	3.50E+06	✓	✓	✓	-
CL(64:5)	4.24E+06	6.30E+06	7.08E+06	8.49E+06	3.15E+06	1.41E+06	✓	✓	-	-
CL(64:6)	-	-	-	-	-	-	✓	✓	-	-
CL(65:1)	9.12E+07	1.94E+08	7.57E+07	5.84E+07	8.34E+07	2.17E+07	✓	✓	✓	✓
CL(65:2)	1.39E+08	2.71E+08	1.21E+08	9.34E+07	1.17E+08	1.63E+07	✓	✓	✓	✓
CL(65:3)	1.17E+08	2.30E+08	1.16E+08	8.38E+07	9.52E+07	9.00E+06	✓	✓	✓	✓
CL(66:1)	3.59E+07	9.32E+07	2.49E+07	1.61E+07	2.92E+07	1.72E+07	✓	✓	✓	✓
CL(66:2)	6.91E+07	1.50E+08	5.44E+07	3.28E+07	5.52E+07	1.29E+07	✓	✓	✓	✓
DG(30:0)	2.59E+07	1.32E+07	1.27E+07	2.26E+07	5.85E+06	-	-	✓	-	-
DG(31:1)	2.66E+06	7.45E+06	2.41E+06	2.29E+06	1.16E+06	-	✓	-	-	-
DG(32:0)	5.93E+06	1.40E+07	2.56E+06	4.79E+06	1.84E+06	6.64E+04	-	-	-	-
DG(32:2)	2.33E+06	1.55E+07	2.84E+06	1.89E+06	8.09E+05	-	✓	✓	-	-
DG(33:0)	5.18E+06	2.64E+07	1.64E+06	2.80E+06	3.83E+06	6.70E+05	-	-	-	-
DG(33:1)	1.30E+06	4.92E+06	1.08E+06	9.25E+05	9.60E+05	4.68E+05	✓	-	-	-
DG(33:2)	1.53E+06	7.00E+06	1.88E+06	8.67E+05	1.03E+06	3.15E+05	✓	✓	-	-
PE(28:0)	2.90E+08	2.75E+08	4.35E+08	1.91E+08	3.25E+08	2.12E+07	✓	✓	-	-
PE(28:1)	4.15E+07	5.09E+07	8.42E+07	9.92E+07	5.19E+07	1.98E+07	✓	✓	-	-
PE(28:2)	2.75E+07	2.12E+07	5.50E+07	1.19E+08	2.21E+07	4.94E+07	✓	✓	-	-
PE(29:0)	3.72E+09	2.03E+09	4.58E+09	2.84E+09	3.60E+09	9.90E+07	✓	✓	-	-
PE(29:1)	1.90E+09	1.27E+09	2.67E+09	3.51E+09	1.83E+09	2.11E+08	✓	✓	-	-
PE(29:2)	2.78E+08	3.05E+08	3.94E+08	5.79E+08	3.00E+08	2.08E+08	✓	✓	-	-
PE(30:0)	9.81E+09	5.01E+09	1.02E+10	1.40E+10	1.18E+10	1.88E+07	✓	✓	-	-
PE(30:1)	4.05E+09	5.00E+09	3.56E+09	2.85E+09	4.05E+09	2.70E+08	✓	✓	-	-
PE(30:2)	1.20E+09	2.63E+09	1.34E+09	1.07E+09	1.29E+09	6.06E+08	✓	✓	-	-
PE(30:3)	1.66E+08	1.82E+08	2.58E+08	3.70E+08	1.41E+08	2.15E+08	✓	✓	-	-
PE(31:0)	2.93E+08	3.40E+09	2.28E+08	1.35E+08	2.89E+08	3.82E+06	✓	✓	-	-
PE(31:1)	3.78E+09	5.55E+09	3.08E+09	2.13E+09	3.38E+09	1.03E+08	✓	✓	-	-
PE(31:2)	3.53E+09	6.42E+09	3.39E+09	2.51E+09	3.49E+09	6.97E+08	✓	✓	-	-
PE(31:3)	8.86E+08	1.26E+09	9.55E+08	8.48E+08	9.06E+08	4.50E+08	✓	✓	-	-
PE(31:4)	5.20E+07	4.89E+07	6.73E+07	1.03E+08	4.29E+07	3.00E+07	✓	✓	-	-
PE(32:0)	7.41E+08	1.66E+09	4.97E+08	2.90E+08	2.90E+08	-	✓	✓	-	-

Table 2.6.2 (continued)

Potential Compound	Control Ion Counts		IVA Suppl. Ion Counts		HIVA Suppl. Ion Counts		² H in LS1191-HIVA			
	DK1622	LS1191	DK1622	LS1191	DK1622	LS1191	D ₃	D ₆	D ₉	D ₁₂
PE(32:1)	8.64E+08	3.15E+09	6.38E+08	5.44E+08	8.09E+08	2.51E+06	✓	✓	-	-
PE(32:2)	3.94E+09	6.98E+09	3.57E+09	2.05E+09	3.63E+09	8.53E+08	✓	✓	-	-
PE(32:3)	2.66E+09	2.79E+09	2.90E+09	2.19E+09	2.04E+09	1.08E+09	✓	✓	-	-
PE(32:4)	5.08E+08	3.38E+08	6.78E+08	8.92E+08	2.86E+08	3.71E+08	✓	✓	-	-
PE(33:0)	2.61E+07	3.38E+08	1.72E+07	8.01E+06	7.55E+06	-	-	-	-	-
PE(33:1)	5.87E+08	7.22E+08	4.16E+08	2.05E+08	4.20E+08	1.31E+06	✓	-	-	-
PE(33:2)	1.46E+09	9.59E+08	1.30E+09	7.43E+08	1.28E+09	4.34E+06	✓	✓	-	-
PE(33:3)	1.74E+09	5.11E+08	1.77E+09	1.43E+09	1.24E+09	7.24E+07	✓	✓	-	-
PE(33:4)	6.41E+08	1.28E+08	7.23E+08	8.68E+08	3.40E+08	4.11E+07	✓	✓	-	-
PE(34:0)	3.50E+07	6.47E+07	2.39E+07	1.22E+07	1.06E+07	6.89E+05	-	-	-	-
PE(34:1)	1.18E+08	6.21E+07	8.81E+07	5.80E+07	7.58E+07	-	✓	✓	-	-
PE(34:2)	2.33E+08	7.89E+07	1.93E+08	1.45E+08	1.61E+08	-	✓	✓	-	-
PE(34:3)	3.35E+08	7.51E+07	3.05E+08	2.89E+08	2.31E+08	1.52E+07	✓	✓	-	-
PE(34:4)	1.96E+08	9.36E+07	1.98E+08	2.10E+08	1.51E+08	5.10E+07	✓	✓	-	-
PE(P-30:0)	3.92E+09	1.53E+08	2.60E+09	9.82E+08	2.38E+09	3.32E+07	✓	✓	-	-
PE(P-32:0)	7.50E+06	-	3.61E+06	2.06E+05	2.95E+04	-	-	✓	-	-
PG(29:0)	8.42E+07	3.36E+07	1.31E+08	1.01E+08	5.52E+07	1.29E+06	✓	✓	-	-
PG(30:0)	1.55E+08	1.21E+08	2.07E+08	2.58E+08	1.95E+08	2.91E+06	✓	✓	-	-
PG(30:1)	7.56E+07	8.95E+07	8.56E+07	6.02E+07	6.87E+07	2.04E+07	-	✓	-	-
PG(31:1)	1.32E+08	2.22E+08	1.22E+08	8.34E+07	1.98E+08	2.40E+07	✓	✓	-	-
PG(31:2)	5.24E+07	5.46E+07	5.26E+07	3.65E+07	6.05E+07	2.28E+07	-	✓	-	-
PG(32:1)	9.97E+07	2.36E+08	7.65E+07	3.48E+07	6.49E+07	1.77E+07	-	✓	-	-
PG(32:2)	1.60E+08	1.12E+08	1.49E+08	1.04E+08	1.35E+08	7.38E+07	-	-	-	-
PG(32:3)	4.76E+07	2.87E+07	5.53E+07	8.02E+07	5.54E+07	1.82E+07	✓	-	-	-
PG(33:0)	-	-	-	-	-	-	-	-	-	-
PG(33:1)	8.12E+07	4.84E+07	6.89E+07	4.65E+07	5.97E+07	1.92E+06	-	✓	-	-
PG(33:2)	3.64E+07	8.38E+06	3.70E+07	3.53E+07	3.51E+07	3.01E+05	-	✓	-	-
PG(34:0)	-	-	-	-	-	-	-	-	-	-
PG(34:1)	2.02E+06	5.20E+05	2.11E+06	1.73E+06	2.00E+06	2.67E+06	-	✓	-	-
TG(38:1)	1.58E+08	5.03E+07	8.44E+07	2.22E+07	4.01E+07	2.07E+06	-	✓	✓	-
TG(39:0)	4.35E+06	1.34E+06	2.13E+06	3.85E+05	4.51E+05	4.29E+05	-	-	-	-
TG(39:1)	5.20E+06	1.53E+06	5.28E+06	6.15E+05	5.79E+05	5.52E+04	✓	-	-	-
TG(40:1)	1.95E+06	1.39E+06	1.64E+06	3.79E+05	3.13E+05	-	-	-	-	-
TG(40:2)	3.19E+06	1.93E+06	3.67E+06	5.23E+05	4.09E+05	-	-	-	-	-

Table 2.6.2 (continued)

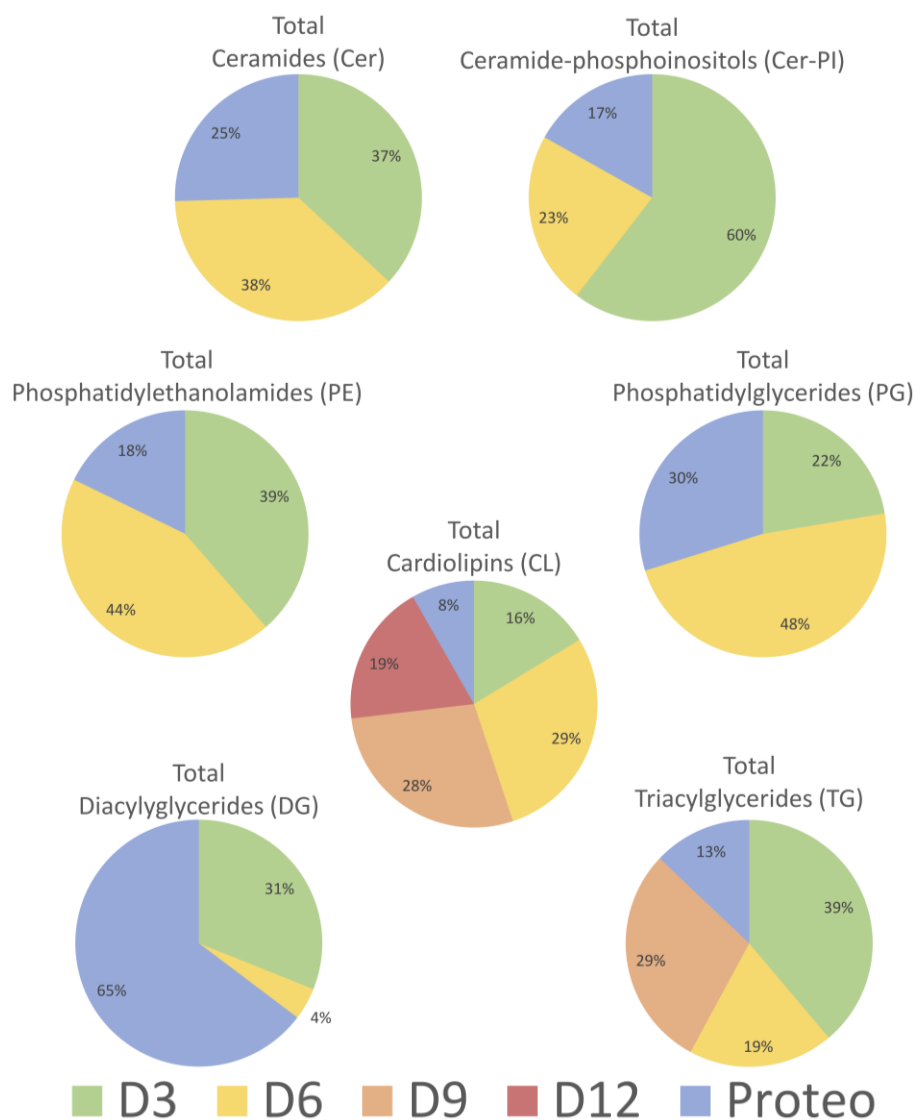
Potential Compound	Control Ion Counts		IVA Suppl. Ion Counts		HIVA Suppl. Ion Counts		² H in LS1191-HIVA			
	DK1622	LS1191	DK1622	LS1191	DK1622	LS1191	D ₃	D ₆	D ₉	D ₁₂
TG(41:1)	1.23E+06	8.08E+05	1.62E+06	2.05E+05	1.72E+05	2.97E+04	✓	-	-	-
TG(41:2)	5.21E+06	4.06E+06	6.27E+06	8.64E+05	1.85E+06	5.79E+05	-	-	-	-
TG(41:3)	1.54E+06	1.39E+06	2.32E+06	3.55E+05	3.03E+04	-	-	✓	-	-
TG(42:1)	2.40E+05	1.87E+05	2.34E+05	9.80E+03	5.38E+04	-	✓	-	-	-
TG(42:2)	1.26E+07	7.79E+06	1.31E+07	1.71E+06	2.02E+06	-	✓	-	-	-
TG(43:0)	3.49E+06	3.04E+06	4.02E+06	6.20E+05	4.15E+05	5.83E+03	✓	✓	✓	-
TG(43:1)	6.60E+07	1.87E+07	5.40E+07	6.29E+06	1.14E+07	3.14E+05	✓	✓	✓	-
TG(43:2)	2.45E+07	1.69E+07	2.33E+07	2.52E+06	5.19E+06	1.79E+05	✓	-	✓	-
TG(43:3)	7.24E+06	6.44E+06	8.17E+06	1.09E+06	1.11E+06	3.02E+05	✓	✓	✓	-
TG(44:0)	1.16E+06	1.02E+06	1.28E+06	8.71E+04	1.78E+05	1.13E+05	✓	✓	-	-
TG(44:1)	1.11E+08	3.15E+07	7.82E+07	1.09E+07	1.70E+07	1.53E+05	✓	✓	✓	-
TG(44:2)	5.35E+07	2.97E+07	4.31E+07	7.39E+06	1.29E+07	-	✓	✓	✓	-
TG(44:3)	1.79E+07	1.33E+07	1.61E+07	3.19E+06	3.62E+06	2.74E+05	✓	✓	✓	-
TG(44:4)	2.92E+06	2.88E+06	3.30E+06	4.06E+05	4.19E+05	2.88E+05	✓	✓	✓	-
TG(45:0)	4.00E+05	3.38E+05	4.91E+05	-	5.35E+04	9.05E+04	✓	✓	✓	-
TG(45:1)	6.49E+07	4.43E+07	4.07E+07	5.43E+06	1.87E+07	4.42E+05	✓	✓	✓	-
TG(45:2)	2.73E+07	2.13E+07	2.16E+07	1.92E+06	6.27E+06	1.14E+05	✓	✓	-	-
TG(45:3)	6.04E+06	4.88E+06	5.70E+06	4.37E+05	7.55E+05	-	✓	✓	✓	-
TG(45:4)	8.40E+05	7.03E+05	9.92E+05	6.22E+04	9.06E+04	-	✓	✓	✓	-
TG(46:0)	6.11E+07	5.40E+07	3.26E+07	3.71E+06	9.28E+06	1.17E+06	✓	-	-	-
TG(46:1)	5.10E+07	4.74E+07	3.11E+07	4.84E+06	2.01E+07	9.40E+05	✓	✓	-	-
TG(46:2)	3.79E+07	2.54E+07	2.52E+07	2.59E+06	8.79E+06	2.50E+05	✓	✓	-	-
TG(46:3)	9.15E+06	7.47E+06	6.42E+06	4.40E+05	1.40E+06	4.69E+04	✓	✓	✓	-
TG(47:0)	4.81E+07	4.45E+07	2.40E+07	3.58E+06	7.89E+06	1.71E+06	-	-	-	-
TG(47:1)	2.77E+07	3.59E+07	1.56E+07	2.22E+06	8.56E+06	1.07E+06	✓	-	-	-
TG(47:2)	2.58E+07	1.96E+07	1.53E+07	1.28E+06	6.92E+06	4.18E+05	✓	-	-	-
TG(47:3)	8.05E+06	6.37E+06	5.01E+06	1.89E+05	1.47E+06	3.28E+05	✓	✓	-	-
TG(47:4)	1.83E+06	1.26E+06	1.55E+06	-	1.65E+05	6.38E+05	✓	✓	-	-
TG(48:0)	1.40E+07	2.65E+07	5.14E+06	1.66E+05	2.64E+06	1.14E+06	-	-	✓	-
TG(48:1)	1.40E+07	1.98E+07	7.18E+06	9.57E+05	5.33E+06	1.51E+06	-	-	✓	-
TG(48:2)	1.17E+07	1.05E+07	5.96E+06	1.76E+04	2.83E+06	4.54E+05	✓	-	-	-
TG(48:3)	5.01E+06	3.52E+06	2.88E+06	1.37E+06	7.97E+05	-	✓	✓	-	-

Table 2.6.2 (continued)

Potential Compound	Control Ion Counts		IVA Suppl. Ion Counts		HIVA Suppl. Ion Counts		² H in LS1191-HIVA			
	DK1622	LS1191	DK1622	LS1191	DK1622	LS1191	D ₃	D ₆	D ₉	D ₁₂
TG(48:4)	1.15E+06	6.25E+05	9.09E+05	4.89E+03	1.71E+05	1.46E+04	✓	✓	-	-
TG(49:0)	8.37E+06	1.10E+07	2.71E+06	1.02E+05	1.43E+06	4.52E+05	-	-	-	-
TG(49:1)	5.33E+06	6.15E+06	1.39E+06	-	1.96E+06	7.42E+05	-	-	-	-
TG(49:2)	4.67E+06	2.72E+06	1.77E+06	1.44E+04	9.34E+05	1.34E+05	-	-	-	-
TG(50:0)	2.05E+06	2.07E+06	3.78E+05	4.77E+05	6.37E+05	4.10E+05	-	-	-	-
TG(50:1)	2.25E+06	1.07E+06	4.38E+05	1.87E+04	1.31E+06	9.19E+05	-	-	-	-
TG(51:0)	6.57E+05	1.15E+05	1.26E+05	1.43E+04	5.79E+02	7.15E+04	-	-	-	-
TG(51:1)	4.31E+05	2.10E+04	3.83E+03	-	1.04E+05	3.84E+04	-	-	-	-
TG(52:0)	-	-	-	-	-	-	-	-	-	-

2.6.1 – Percent-labeling for detected lipids in *M. xanthus* DK1622

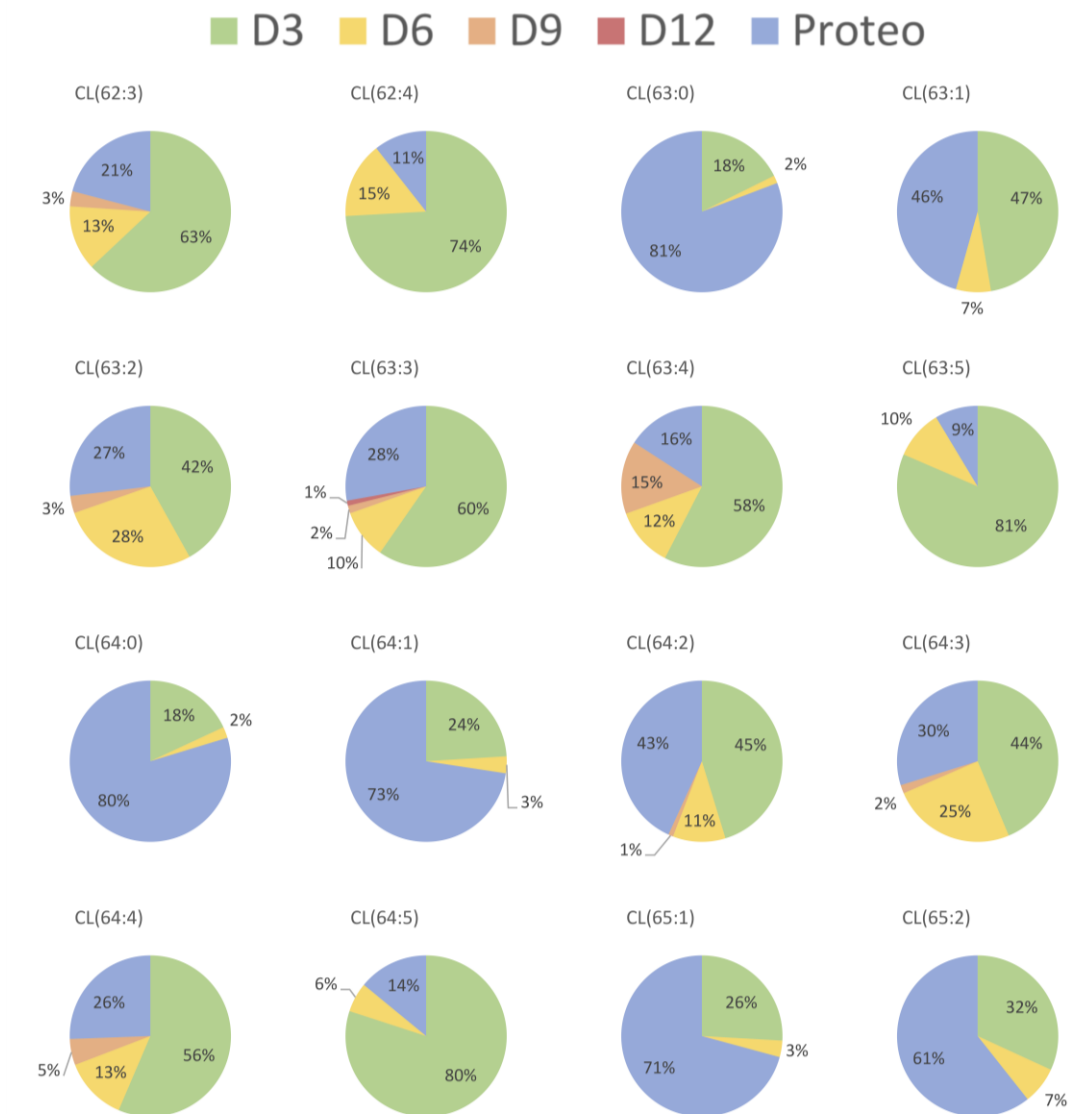
Below are a compilation of pie charts describing the populations of labeled lipids observed in each of the detected lipids by class as well as by individual species. Since a D₃ isolipid tracer was used in this experiment, the labels D3, D6, D9, and D12 describe the detected lipid bearing one, two, three, and four deuterated isolipid tails respectively.



2.6.1 – Percent labeling for detected lipids in *M. xanthus* DK1622 (continued)



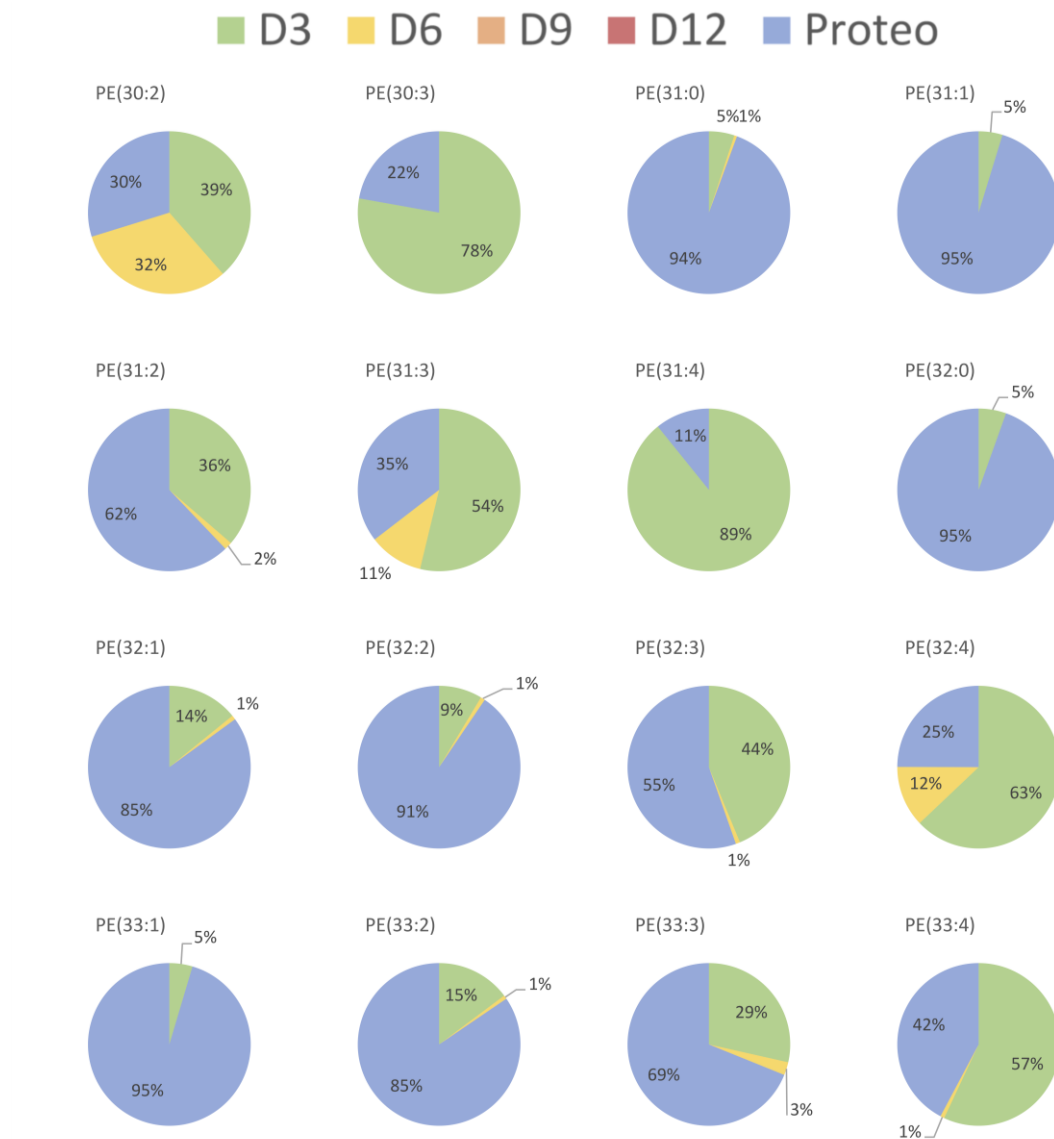
2.6.1 – Percent labeling for detected lipids in *M. xanthus* DK1622 (continued)



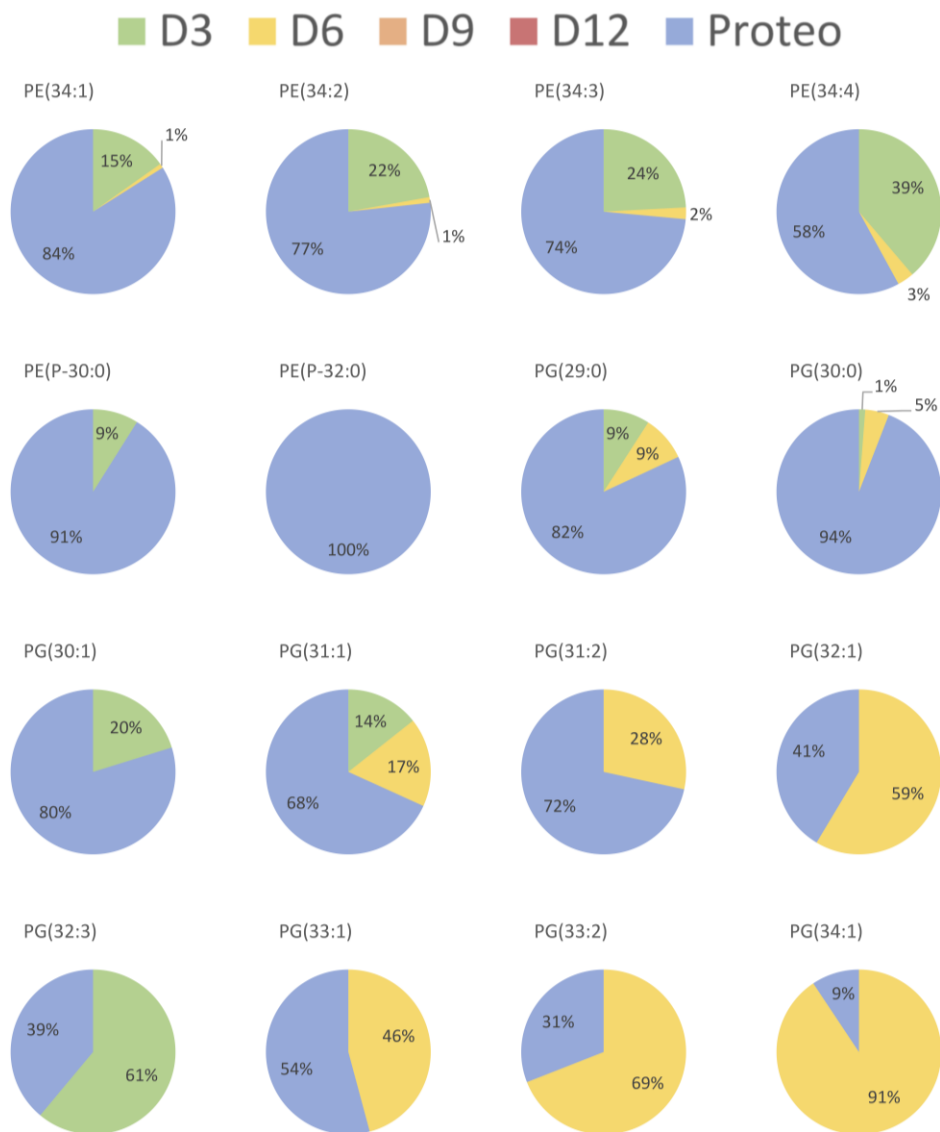
2.6.1 – Percent labeling for detected lipids in *M. xanthus* DK1622 (continued)



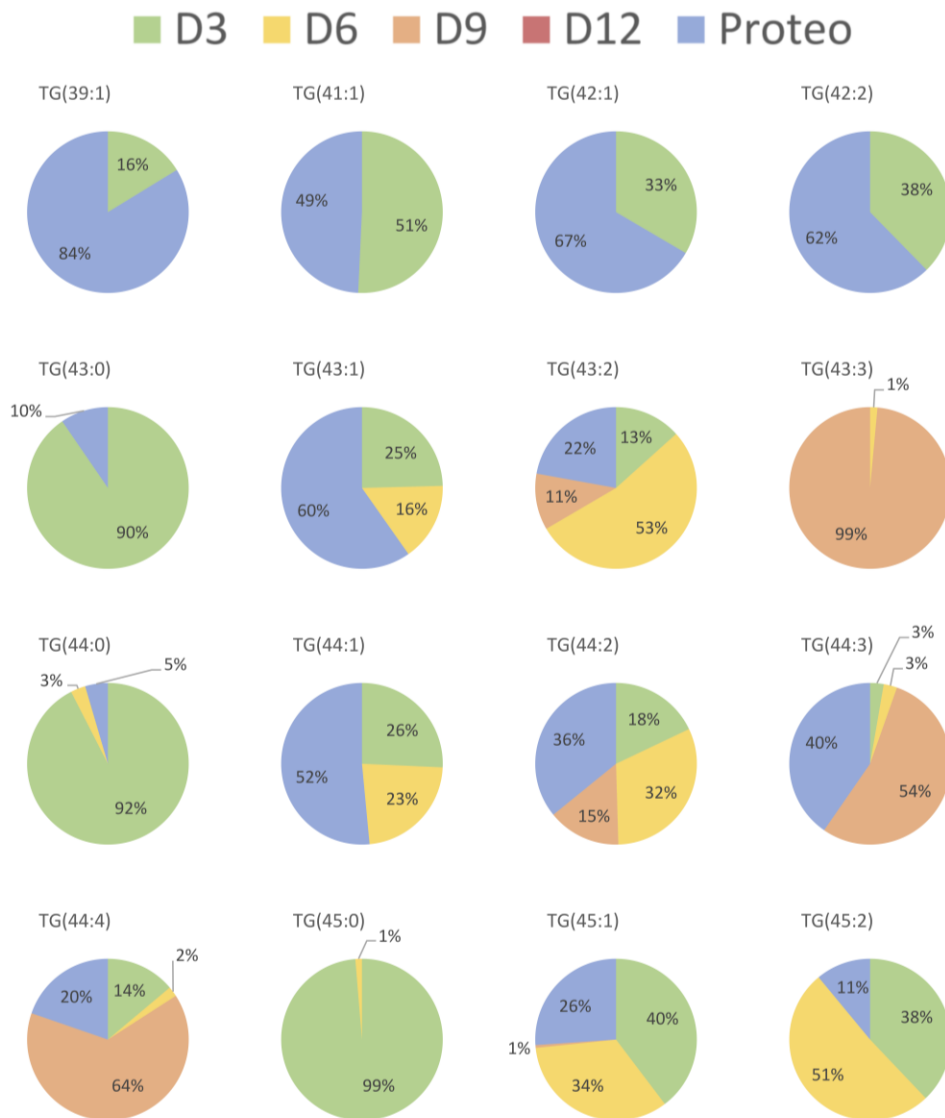
2.6.1 – Percent labeling for detected lipids in *M. xanthus* DK1622 (continued)



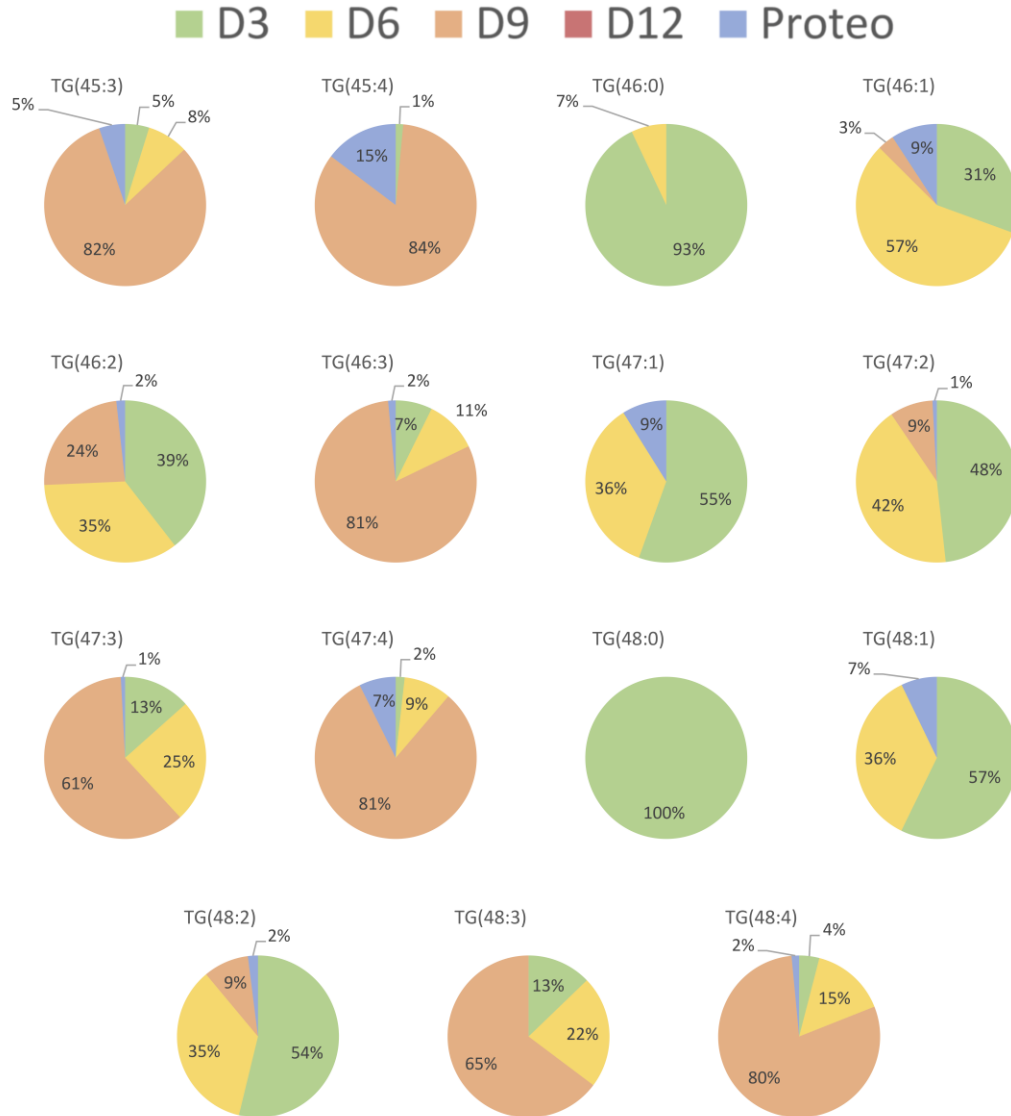
2.6.1 – Percent labeling for detected lipids in *M. xanthus* DK1622 (continued)



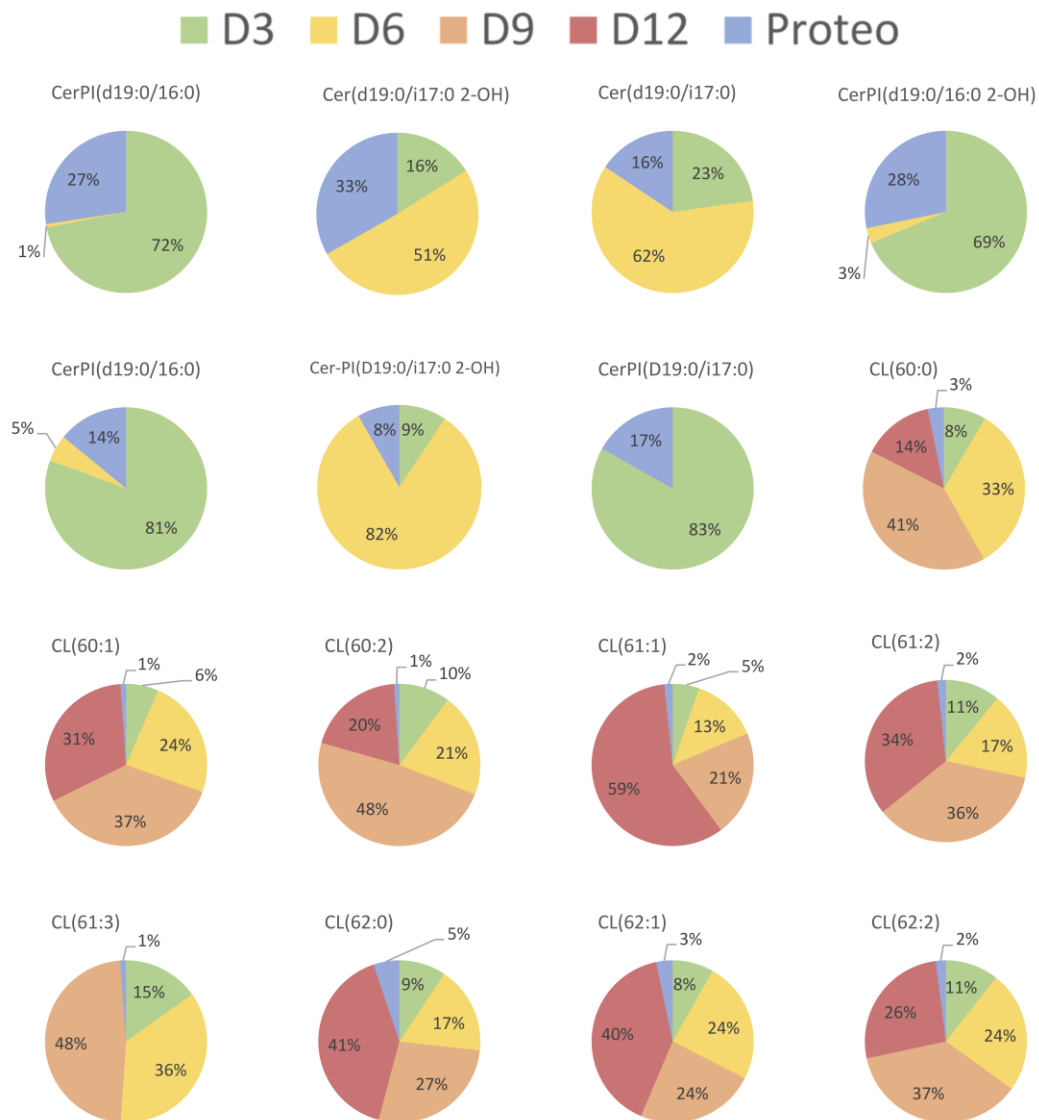
2.6.1 – Percent labeling for detected lipids in *M. xanthus* DK1622 (continued)



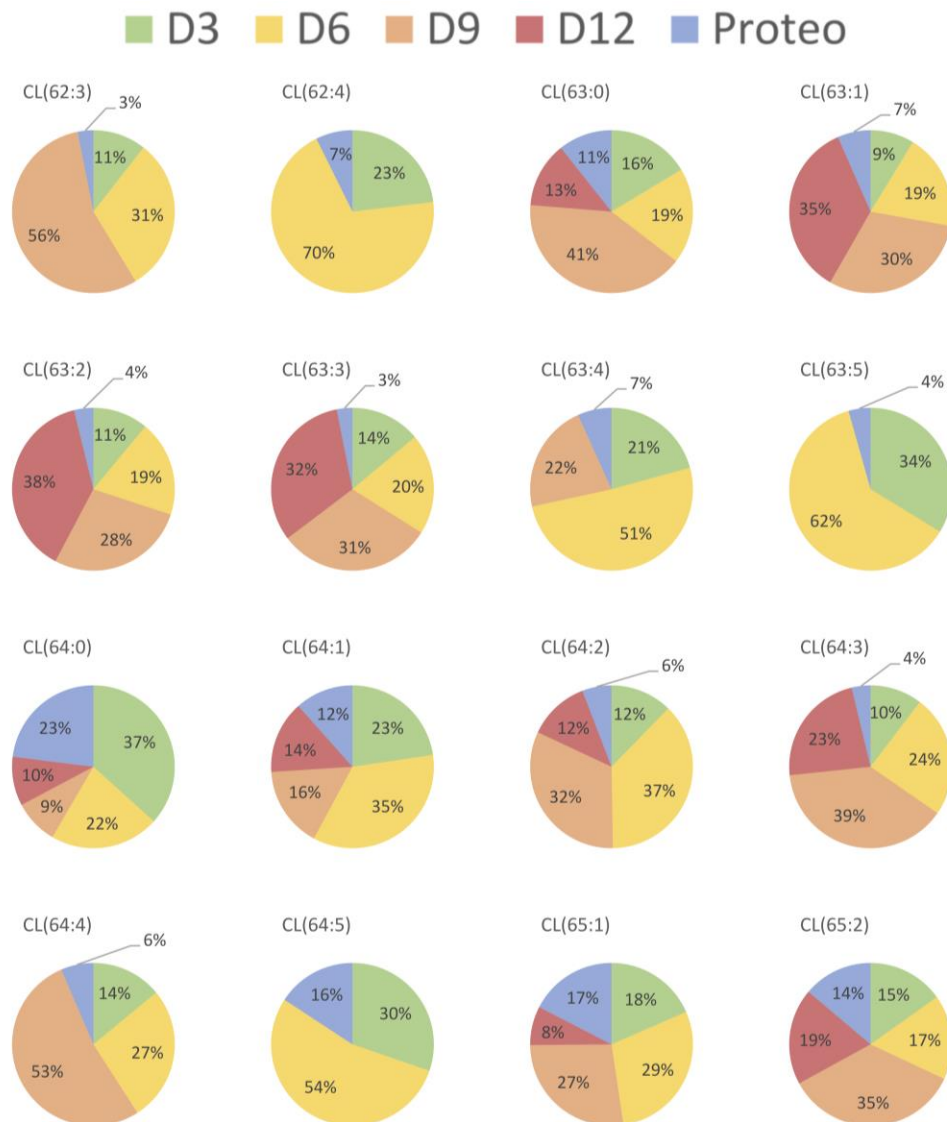
2.6.1 – Percent labeling for detected lipids in *M. xanthus* DK1622 (continued)



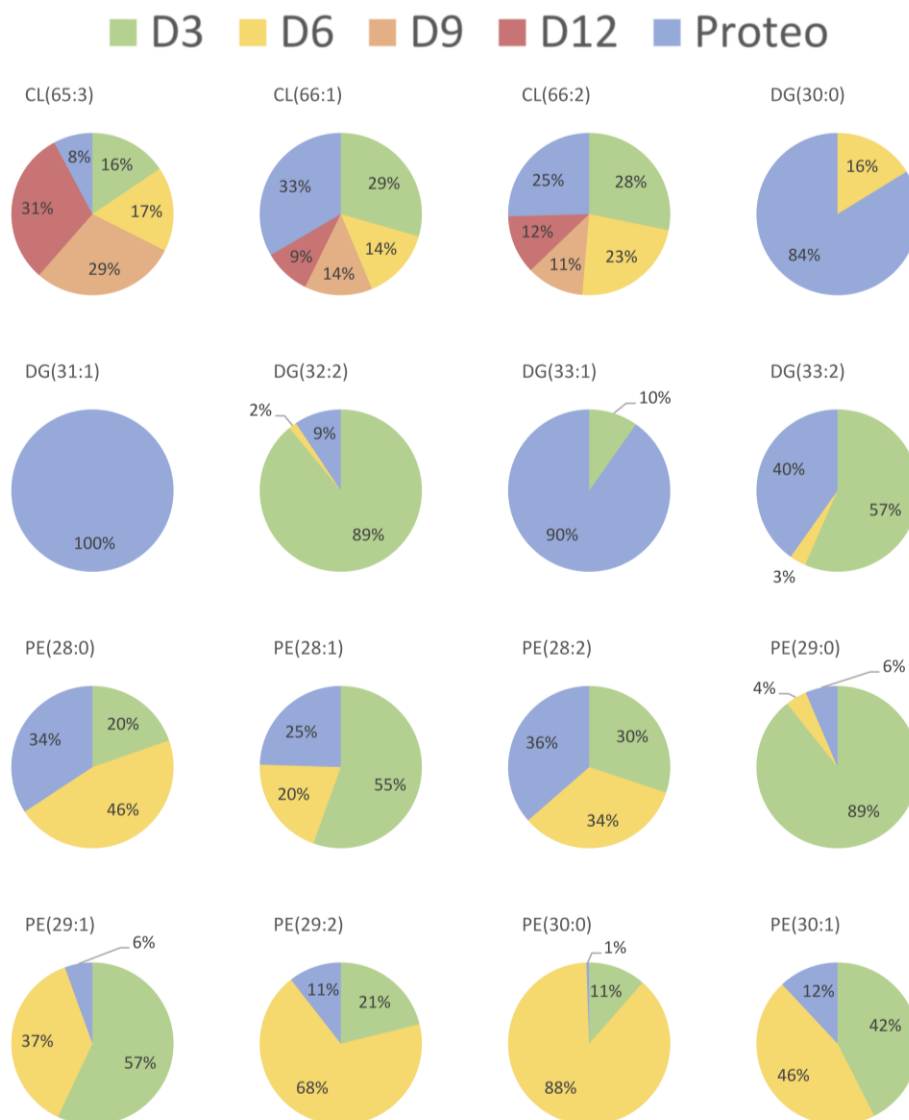
2.6.2 – Percent-labeling for detected lipids in *M. xanthus* LS1191



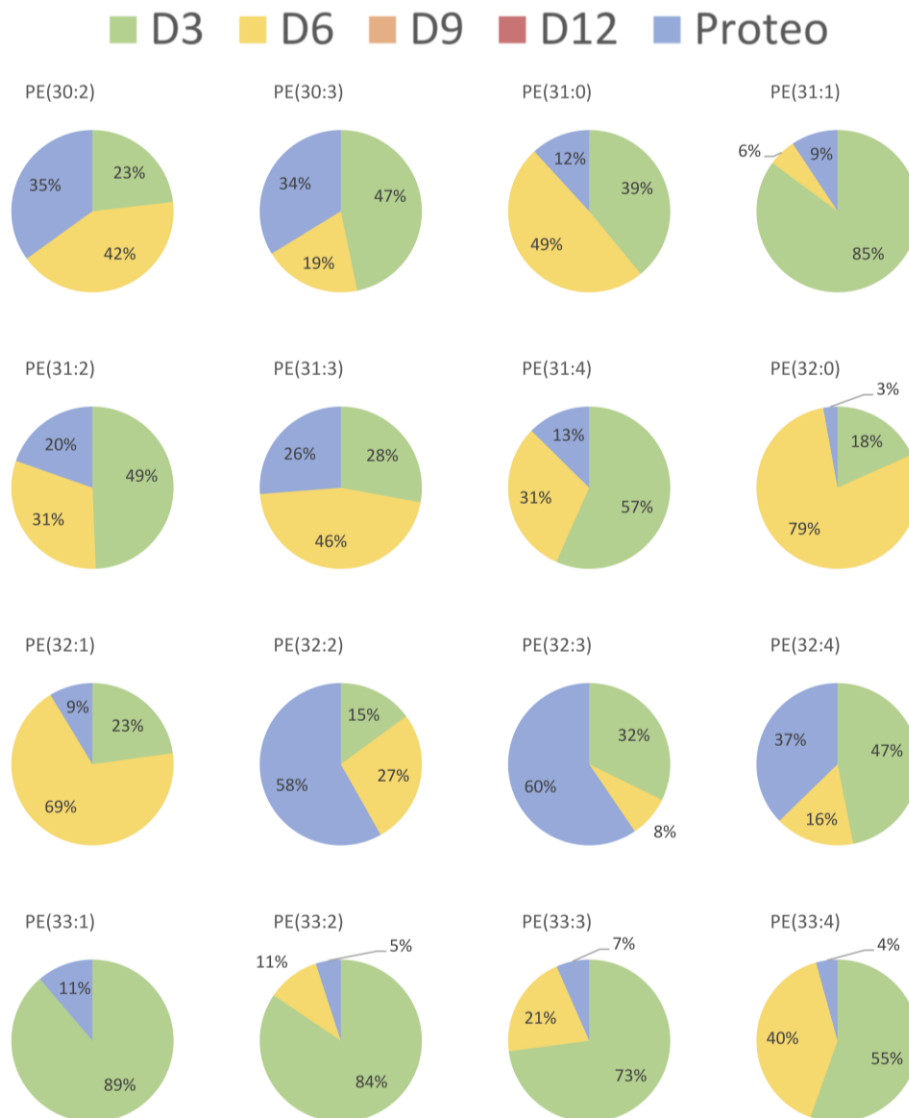
2.6.2 – Percent-labeling for detected lipids in *M. xanthus* LS1191 (continued)



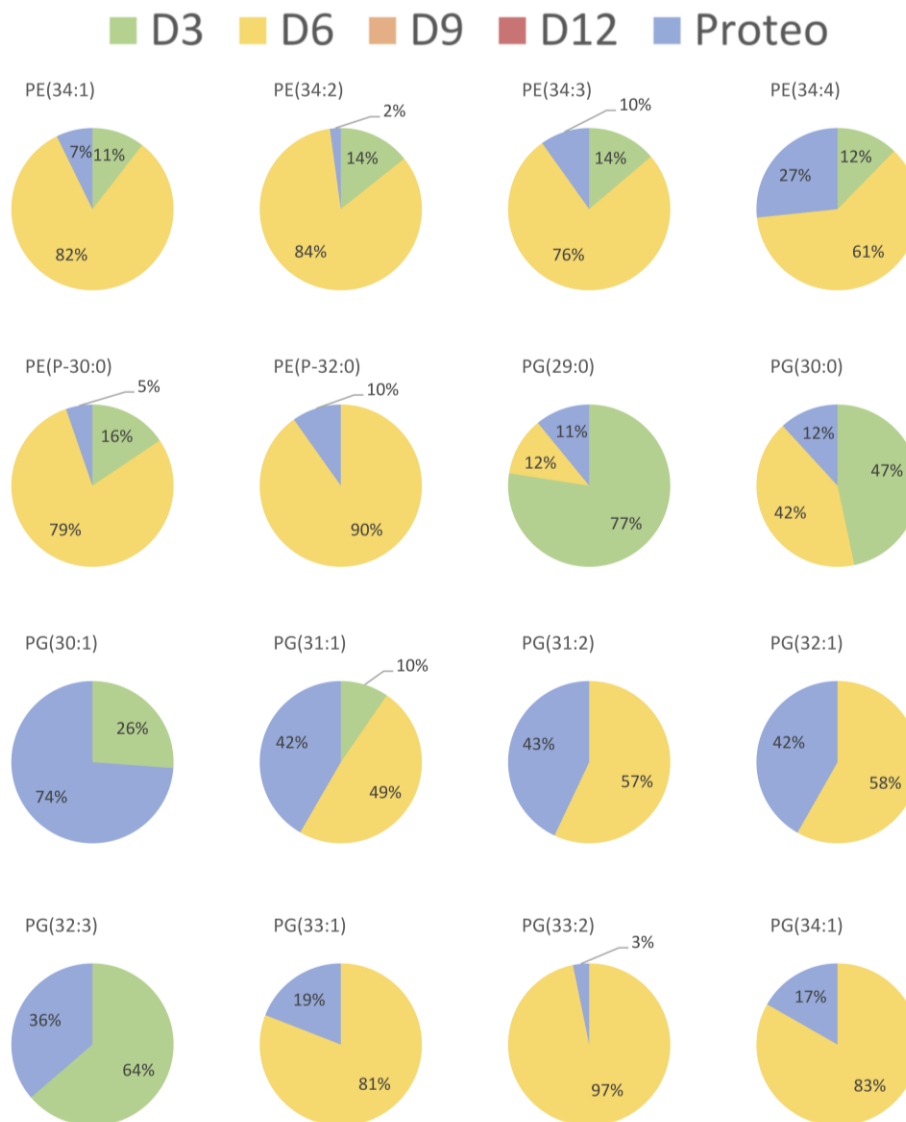
2.6.2 – Percent-labeling for detected lipids in *M. xanthus* LS1191 (continued)



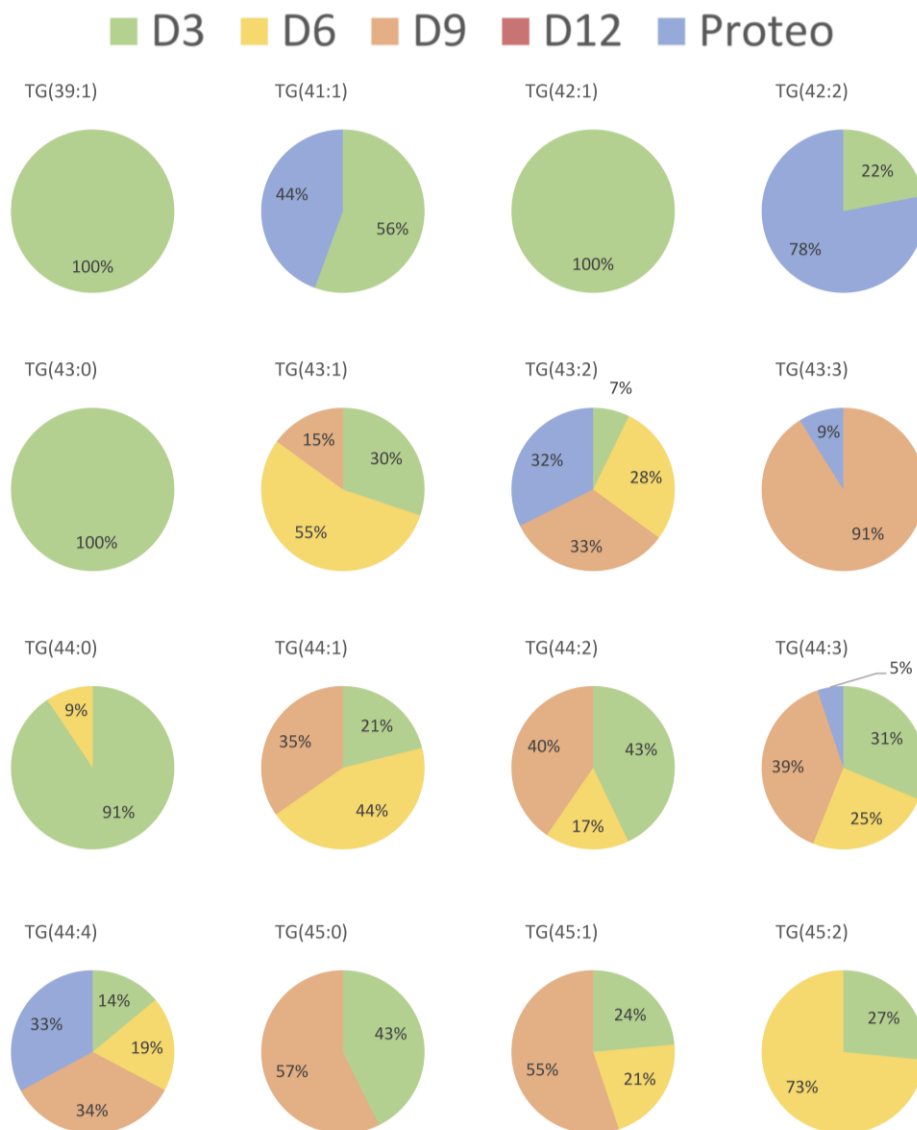
2.6.2 – Percent-labeling for detected lipids in *M. xanthus* LS1191 (continued)



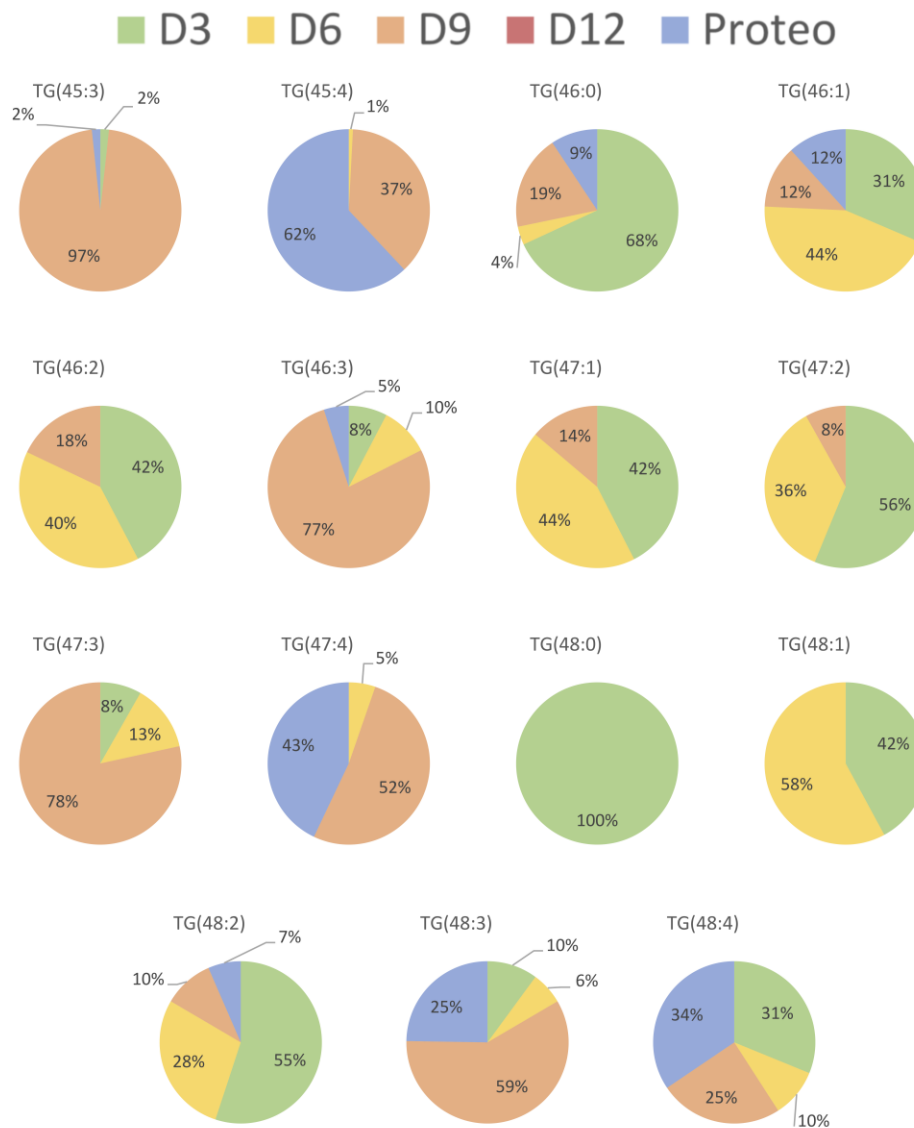
2.6.2 – Percent-labeling for detected lipids in *M. xanthus* LS1191 (continued)



2.6.2 – Percent-labeling for detected lipids in *M. xanthus* LS1191 (continued)



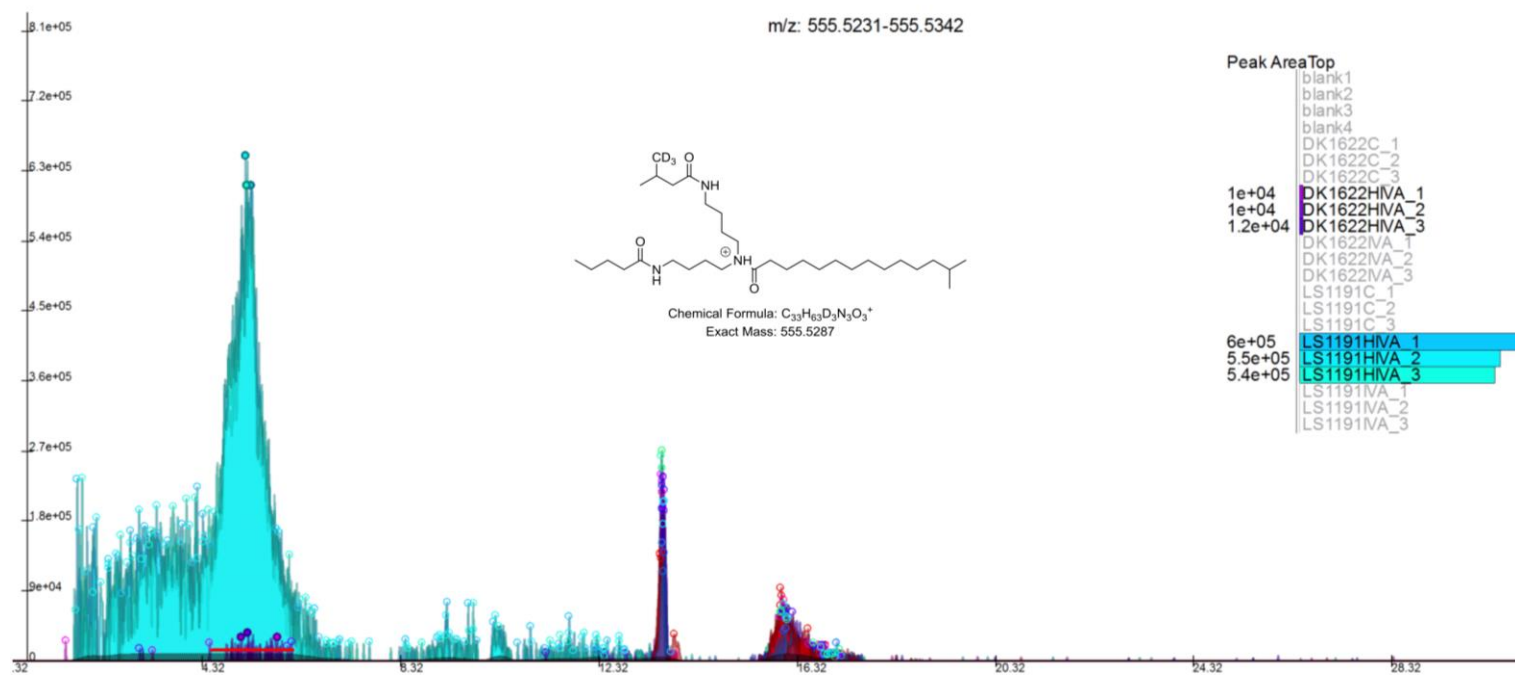
2.6.2 – Percent-labeling for detected lipids in *M. xanthus* LS1191 (continued)



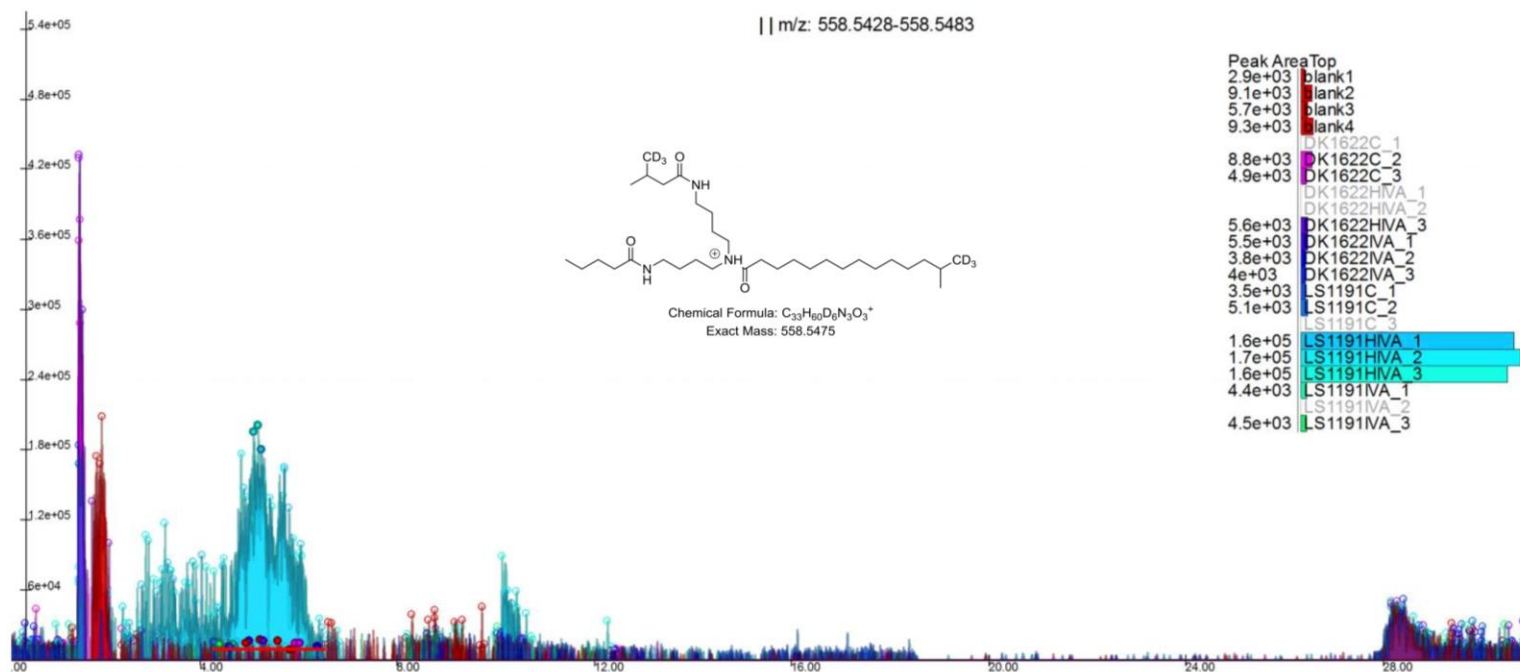
Extracted ion chromatogram for naturally occurring HS(5:0/iso5:0/iso15:0)



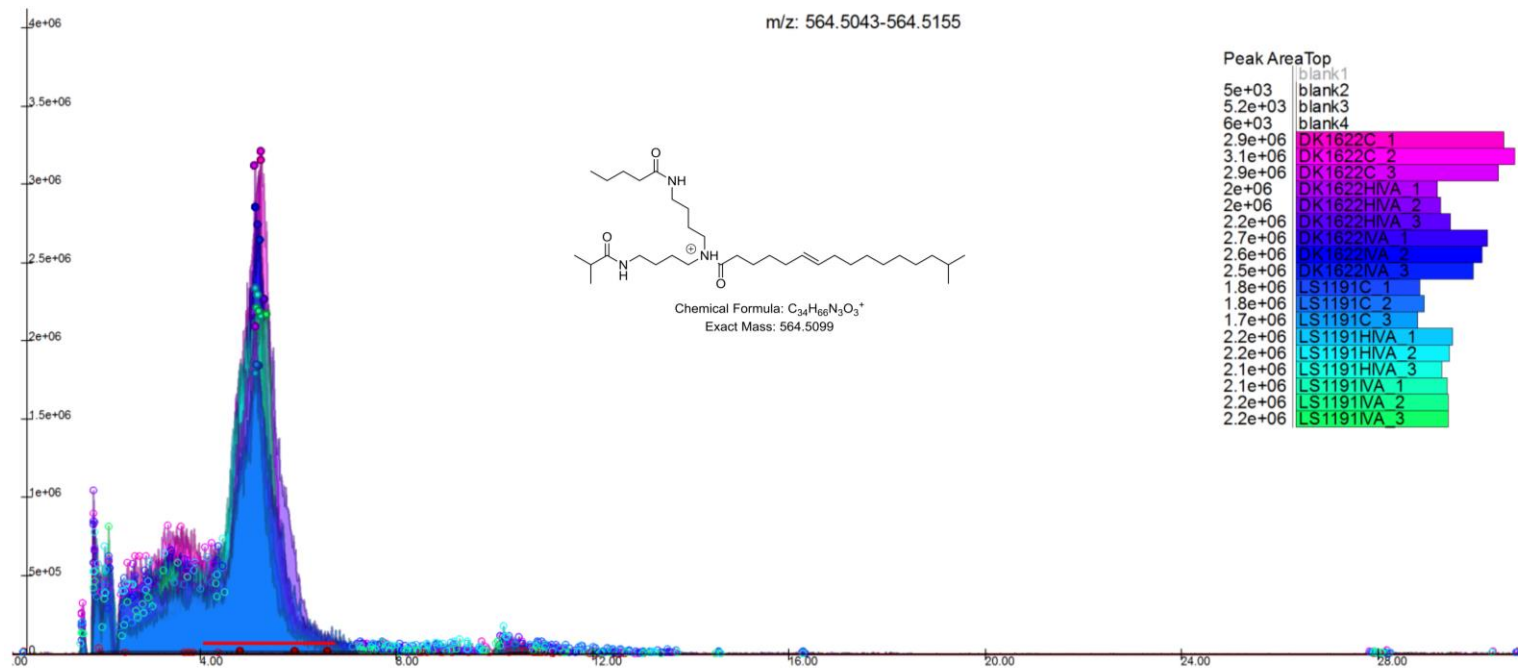
Extracted ion chromatogram for singly labeled (D3)-HS(5:0/iso5:0/iso15:0)



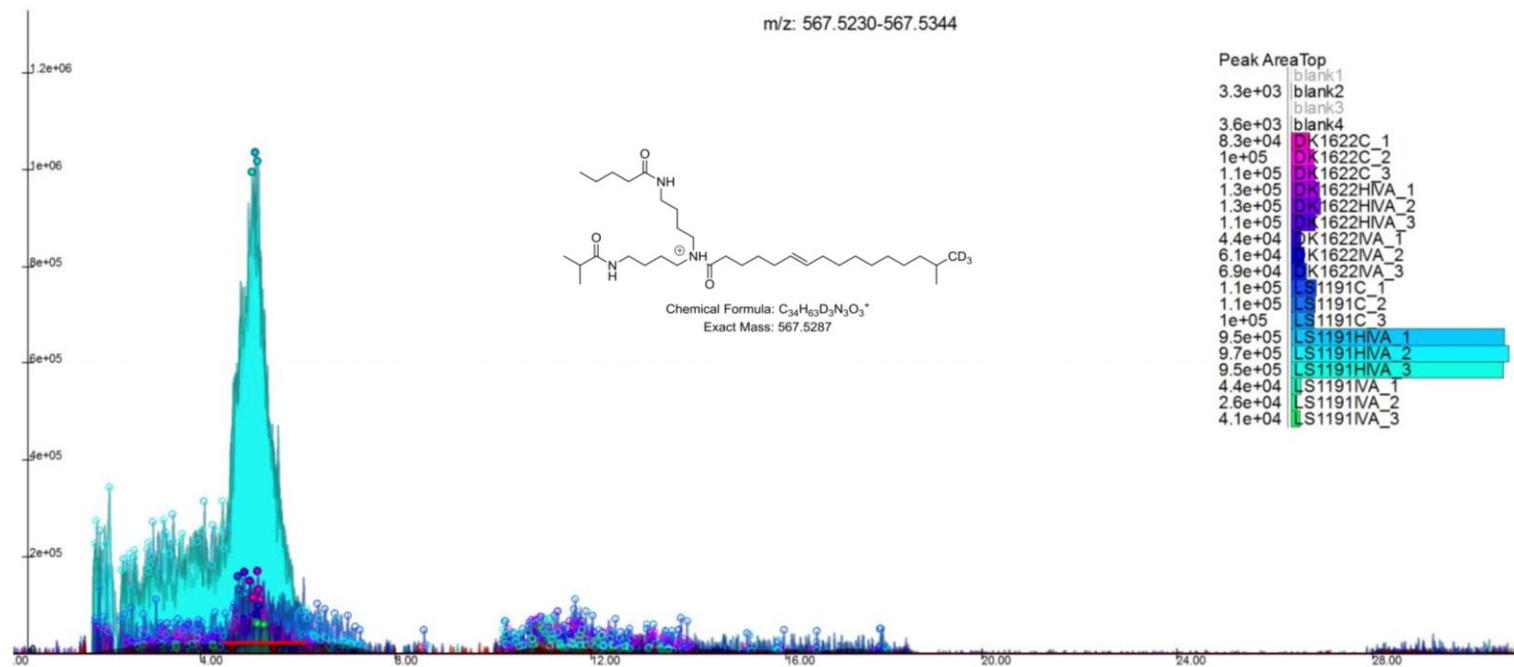
Extracted ion chromatogram for doubly labeled (D_6)-HS(5:0/iso5:0/iso15:0)



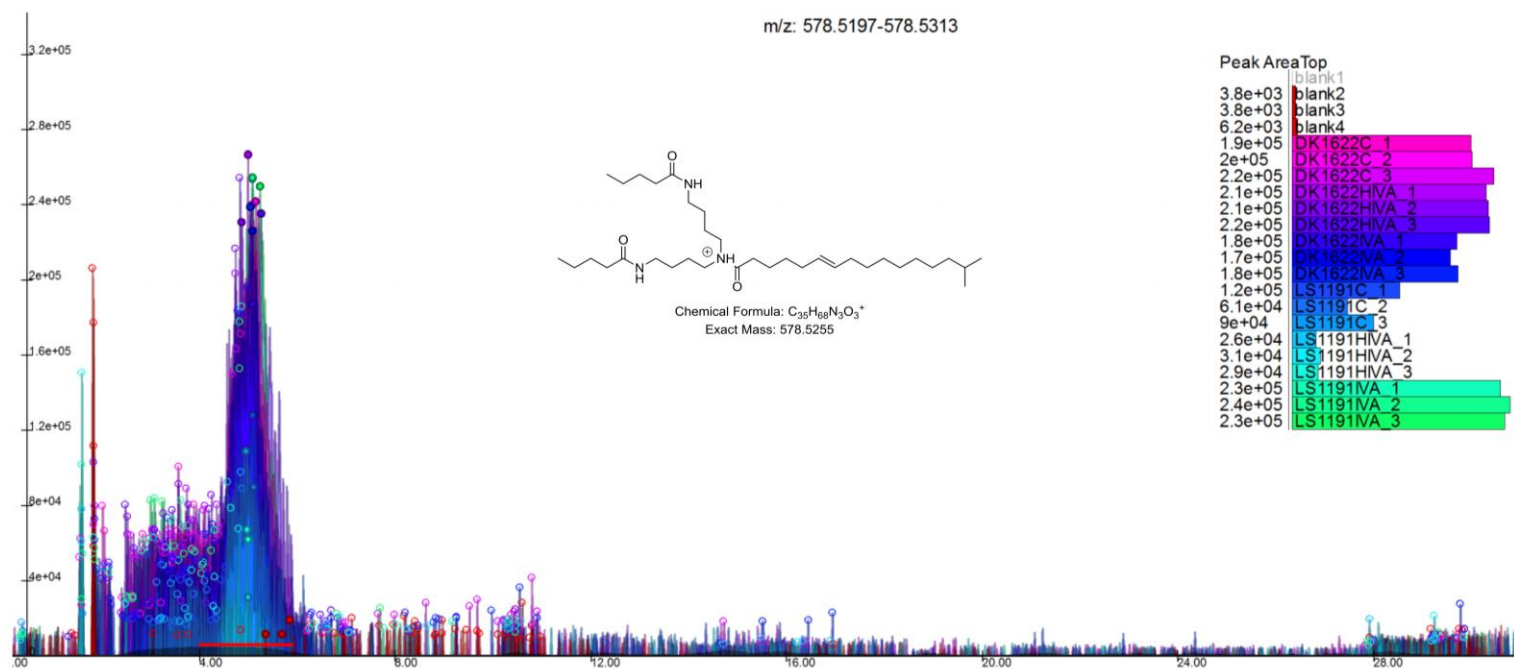
Extracted ion chromatogram for naturally occurring HS(5:0/iso4:0/iso17:1)



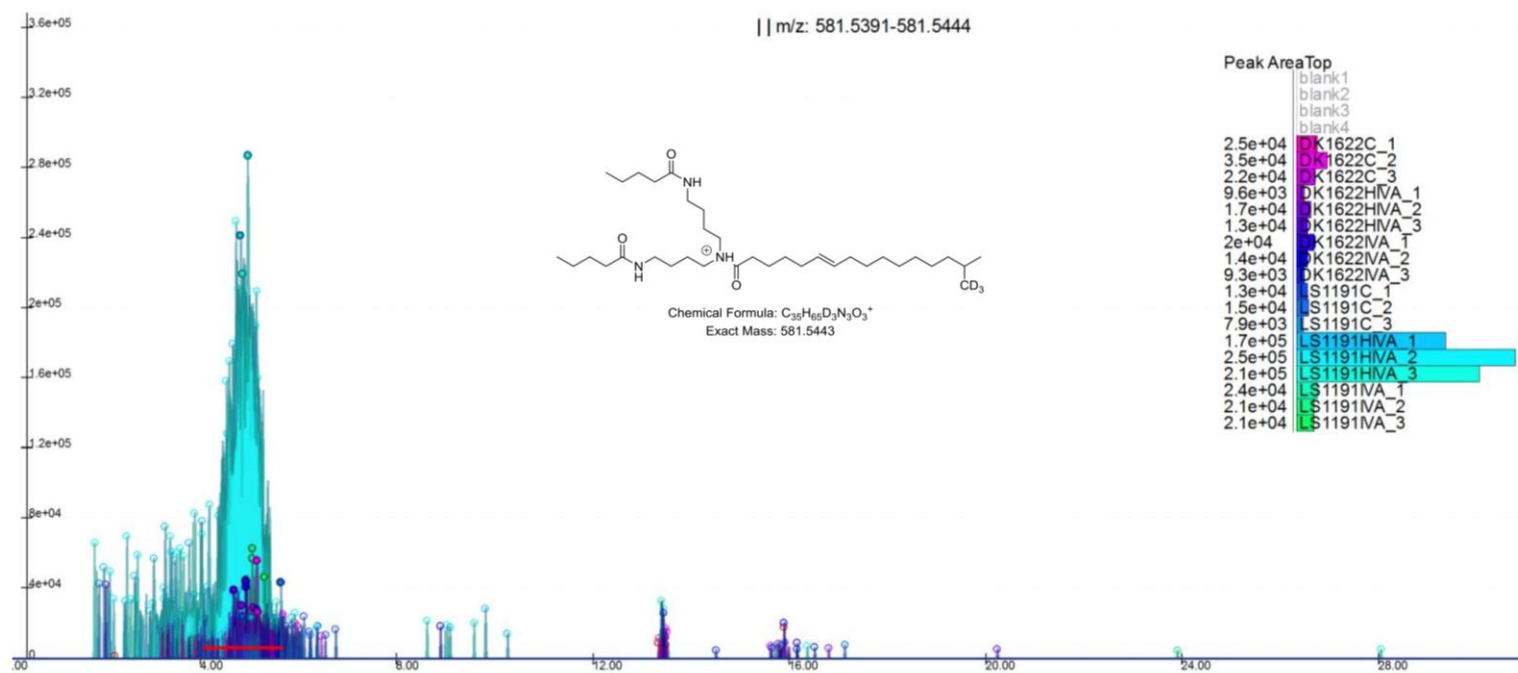
Extracted ion chromatogram for singly labeled (D_3)-HS(5:0/iso4:0/iso17:1)



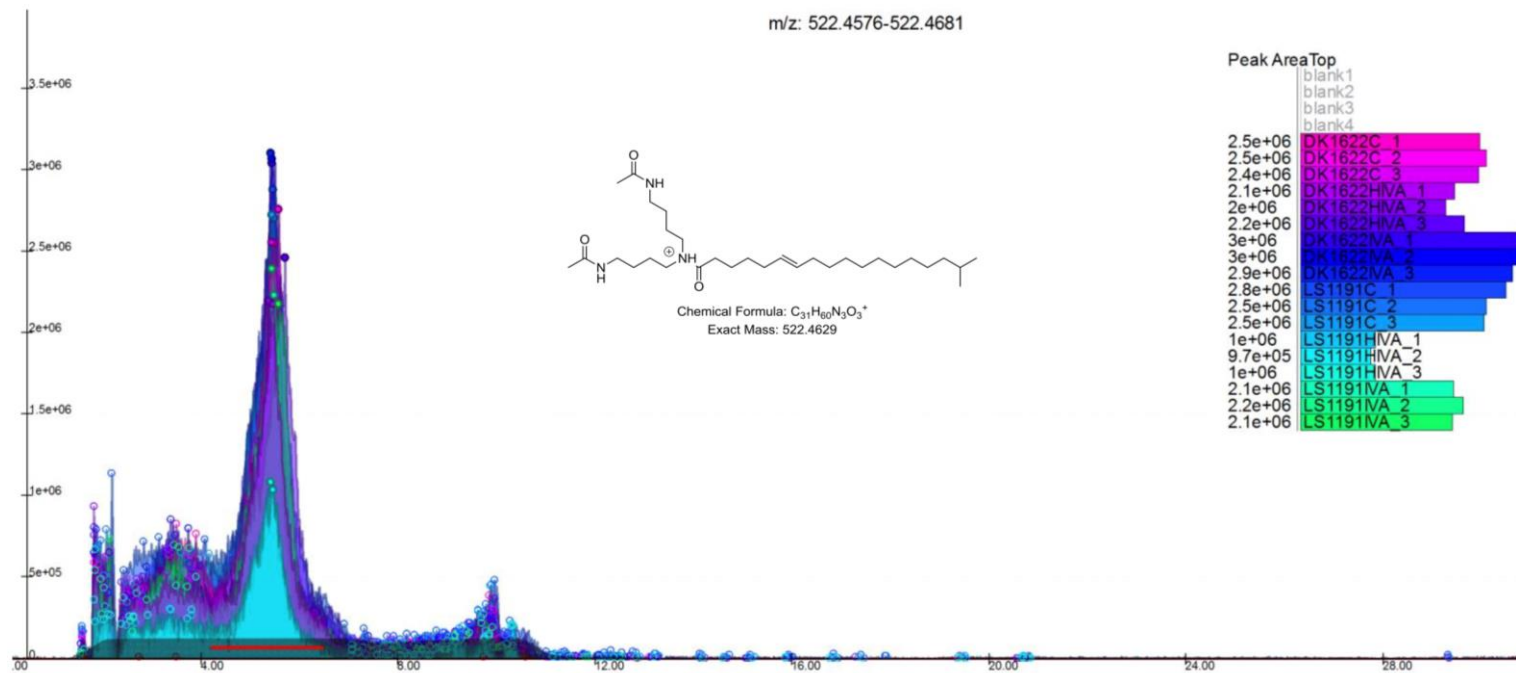
Extracted ion chromatogram for naturally occurring HS(5:0/5:0/iso15:0)



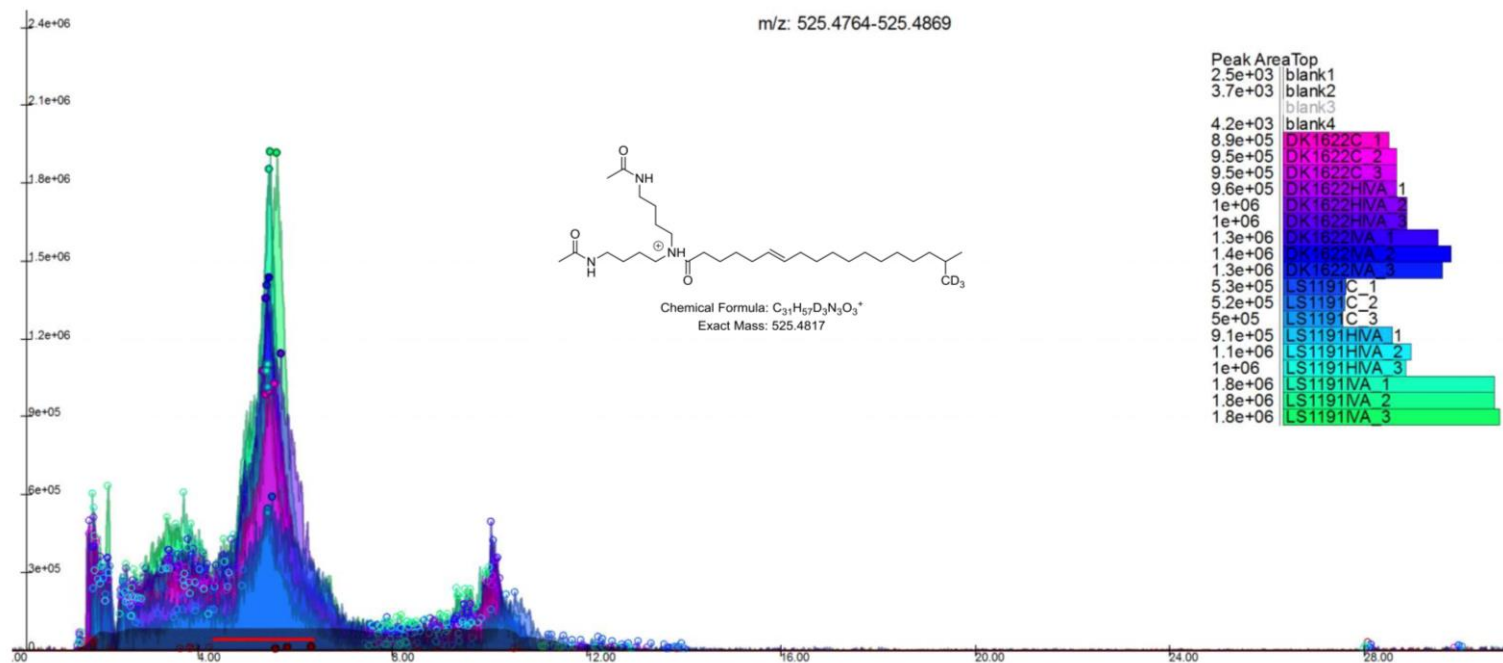
Extracted ion chromatogram for singly labeled (D₃)-HS(5:0/5:0/iso17:1)



Extracted ion chromatogram for naturally occurring HS(2:0/2:0/iso19:1)

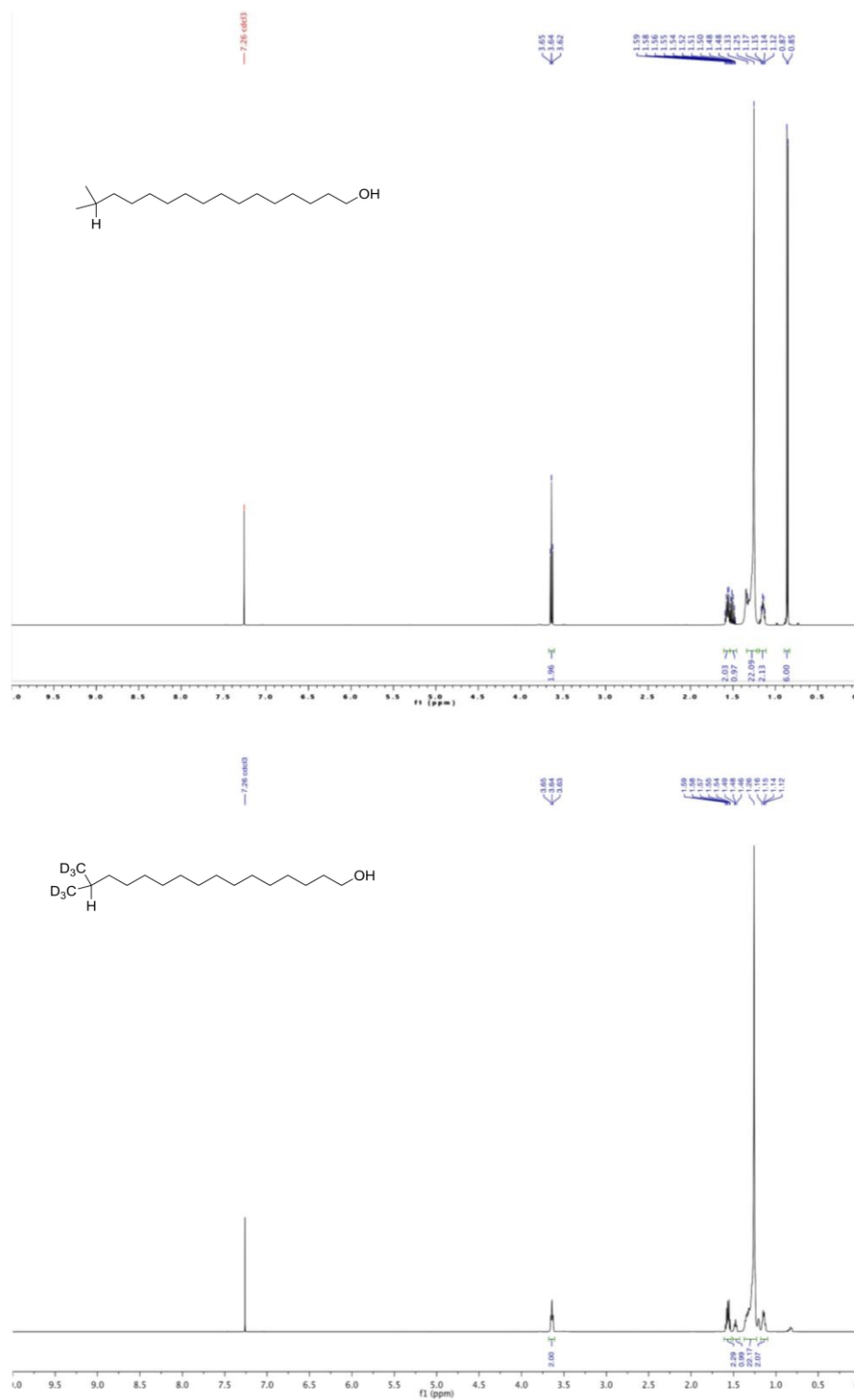


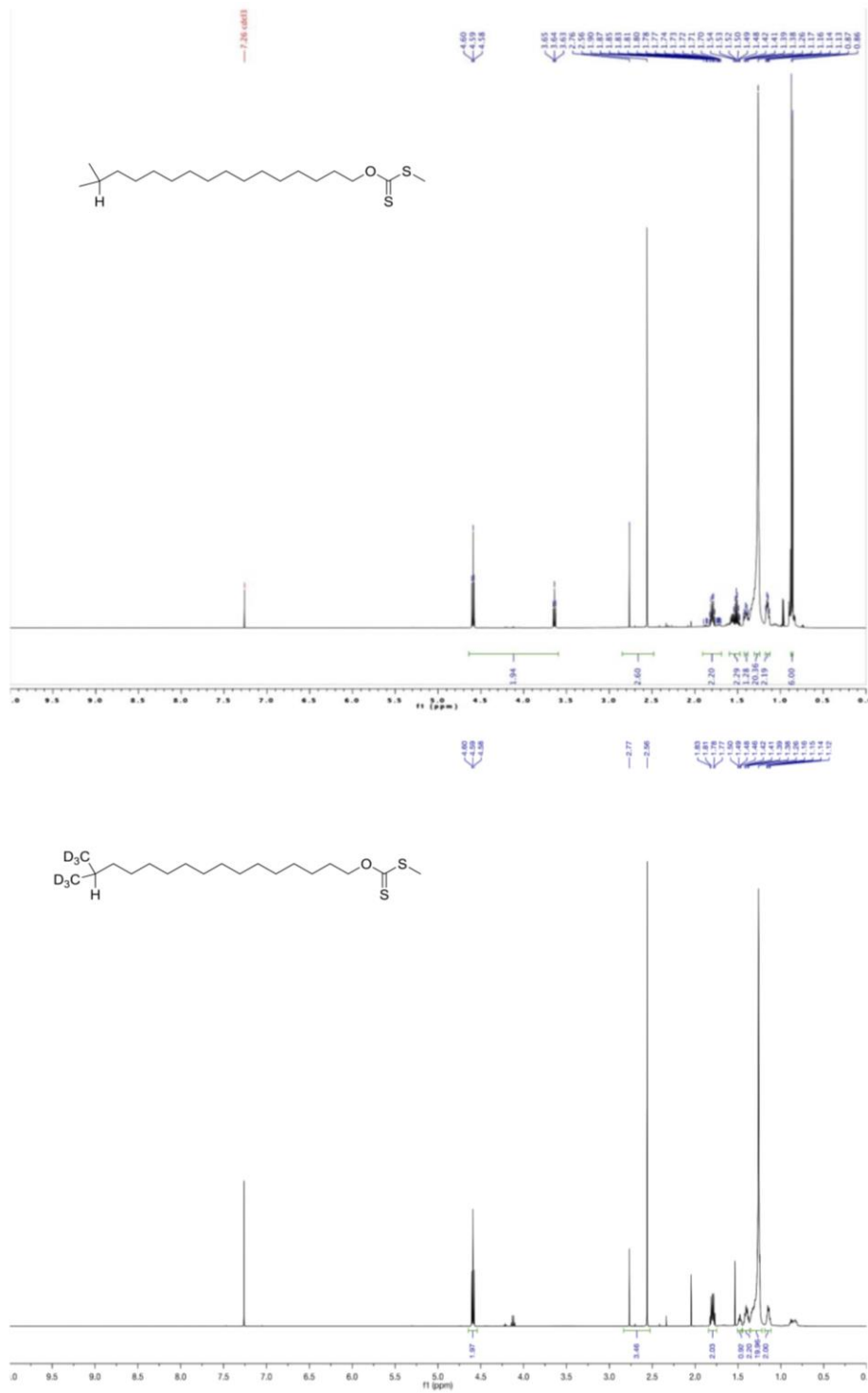
Extracted ion chromatogram for singly labeled (D_3)-HS(2:0/2:0/iso19:1)



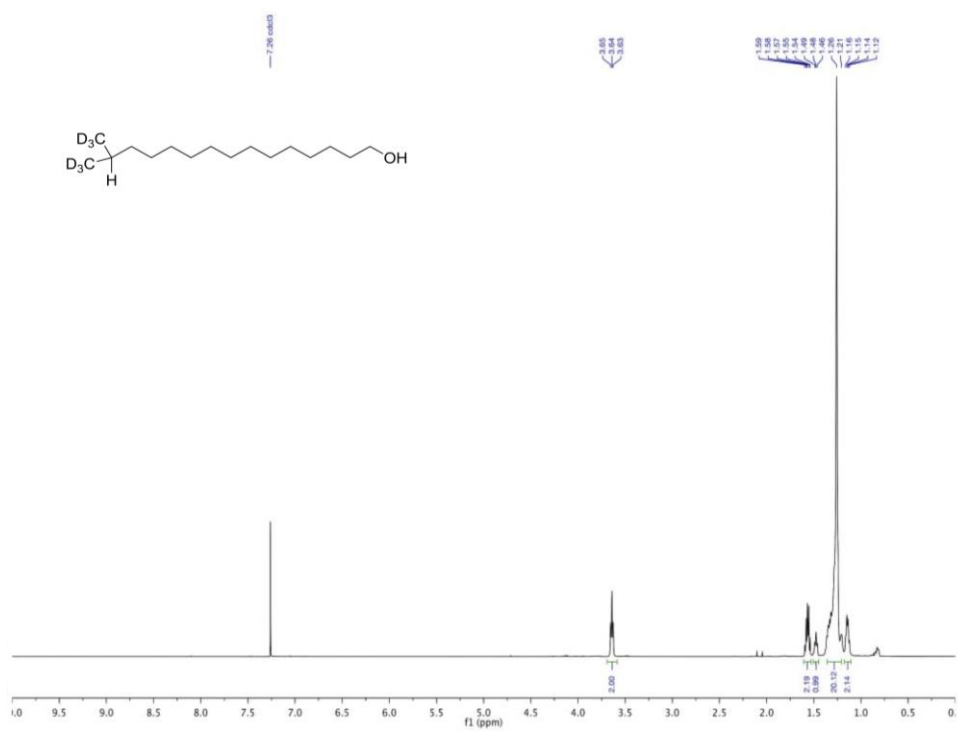
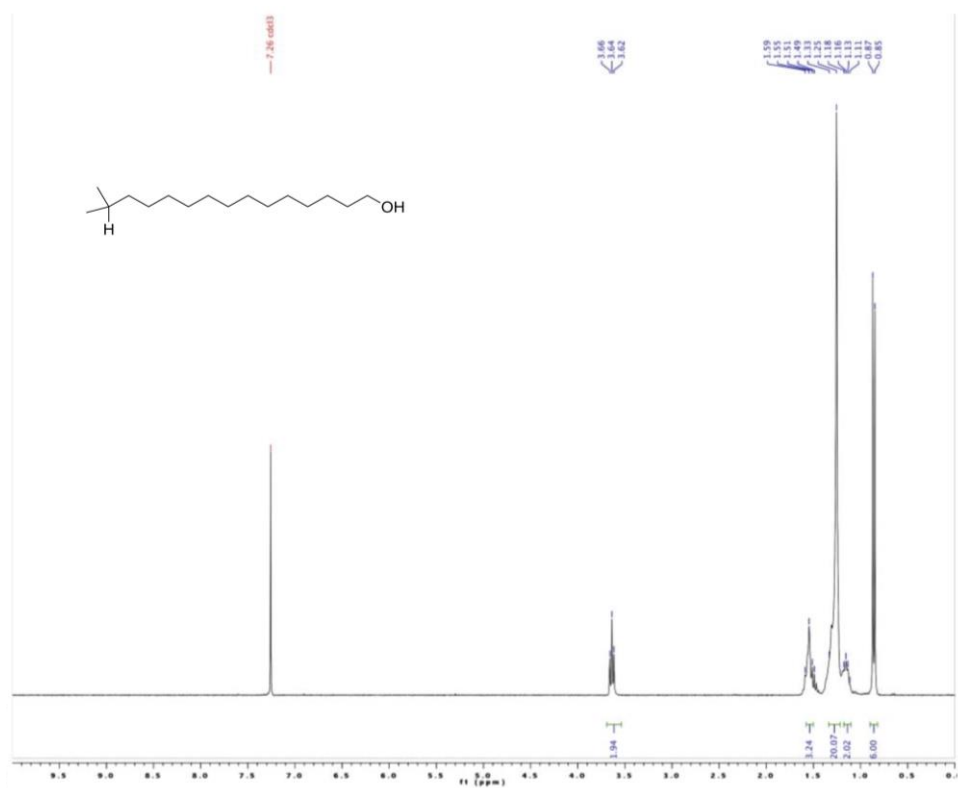
2.6.4 – ^1H -NMR of synthetic intermediates and isolipid tracers $\text{iC}_{15}\text{-d}_6$ and $\text{iC}_5\text{-d}_3$

^1H -NMR of compounds **iC17-ol** and **D₆-iC17-ol**

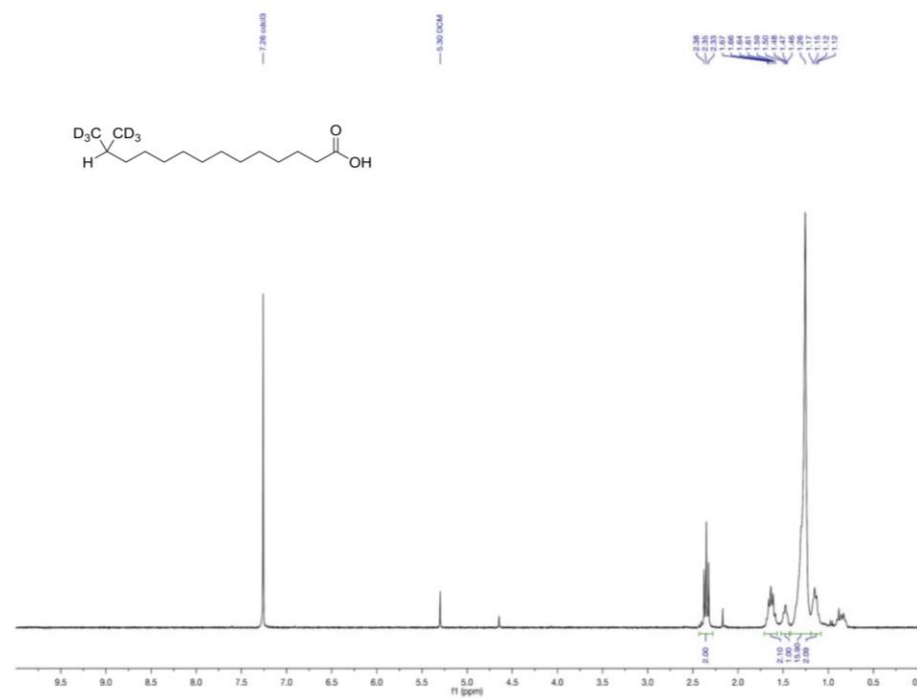
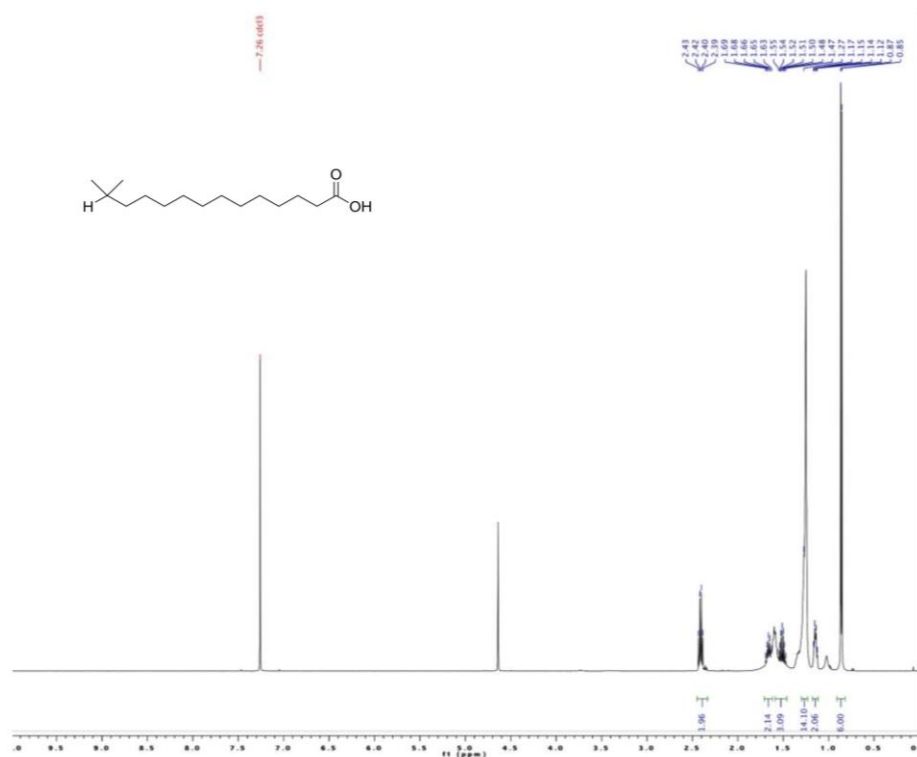


¹H-NMR of compounds **iC17-xanthate** and **D₆-iC17-xanthate**

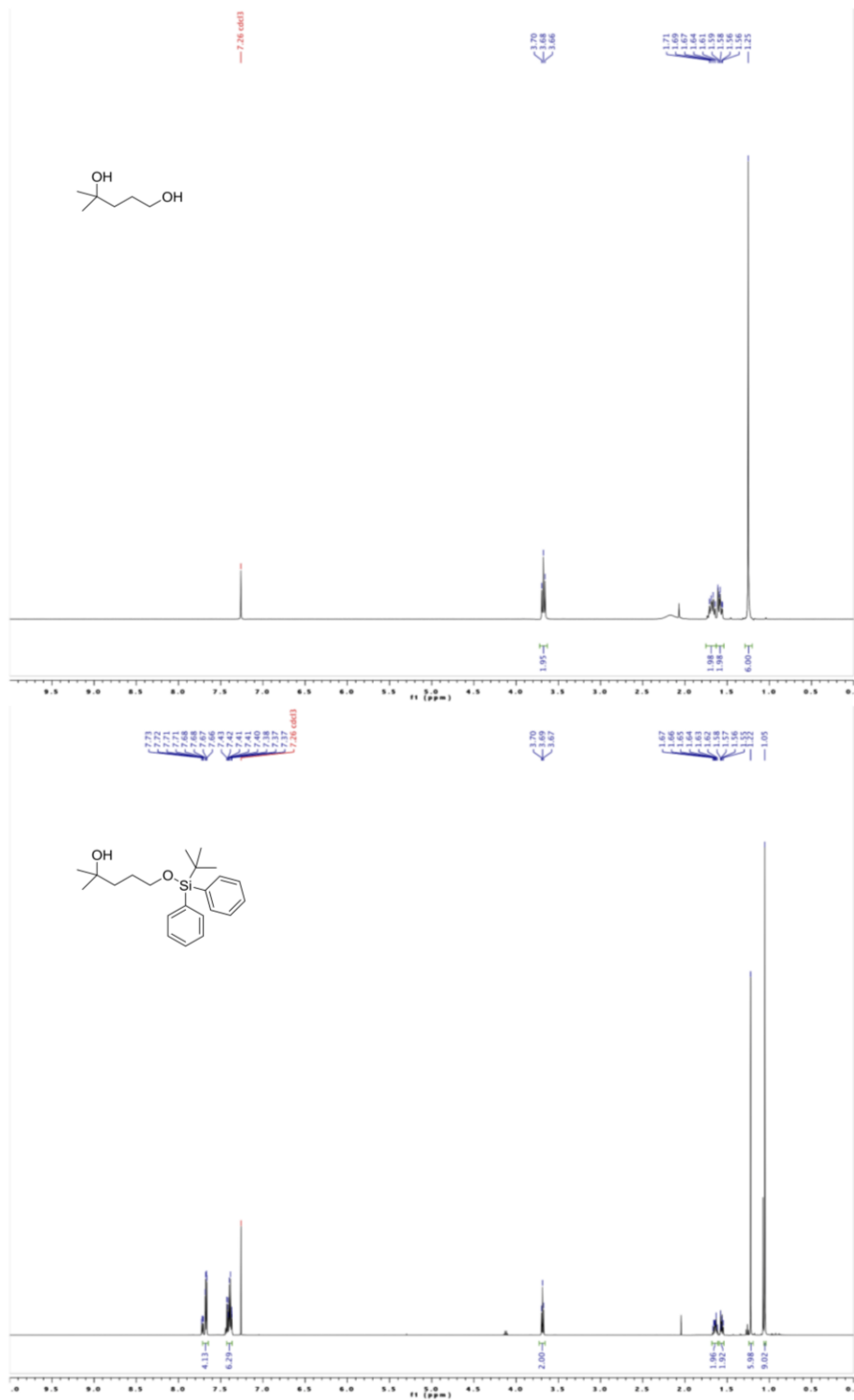
^1H -NMR of compounds **iC16-ol** and **D₆-iC16-ol**



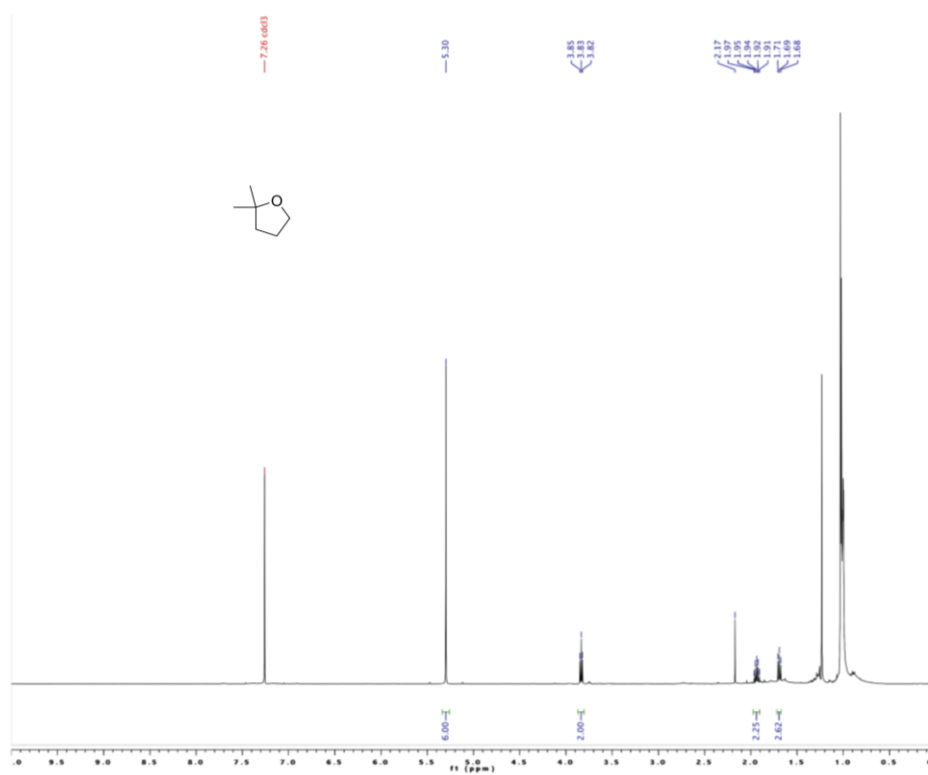
^1H -NMR of compounds **iC16-FA** and **D₆-iC16-FA**



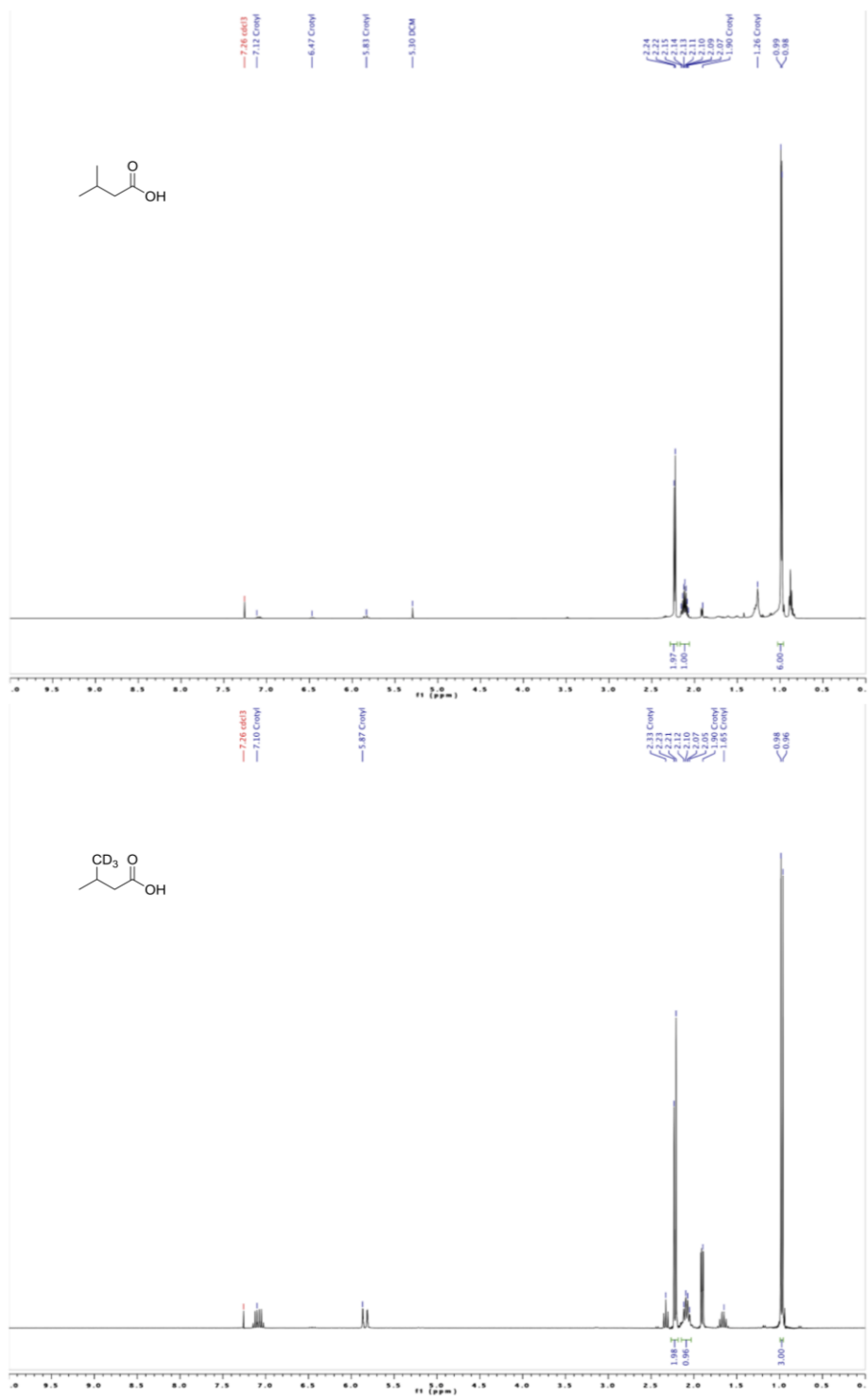
¹H-NMR of compounds **iC6-diol** and **TBDPS-iC6-diol**



¹H-NMR of compounds **1,1-dimethyl-THF**



^1H -NMR of compounds **iC5-FA** and **D₃-iC5-FA**



CHAPTER III

ISOLATION AND PRELIMINARY STRUCTURAL ELUCIDATION OF UNKNOWN SIDEROPHORE FROM THE MARINE BACTERIUM *SULFITOBACTER* *LOKTANELLA* SP. SE62.

The contents of this chapter are preliminary findings that will be part of a future publication titled:

Isolation, purification, and characterization of a potent siderophore abundantly produced by the Roseobacterium *Sulfitobacter Loktanella* sp. SE62.

Brandon J. Kennedy, Ashley M. Lato, Gary Lecleir, Shawn R. Campagna, and Stephen Wilhelm

3.1 – Abstract

The marine organism *Sulfitobacter loktanella* sp. SE62 is an α -proteobacterium belonging to the large and diverse Rhodobacteraceae genus, *Roseobacter*. In an unpublished study lead by Drs. Steven Wilhelm and Alison Buchan they attempted to determine the relative production of iron-binding compounds (siderophores), using *Roseobacters* among other marine microbes, sp. SE62 was found to be very active in the production of siderophore-like compounds with unknown structure. The conclusion that the apparent siderophore being uncharacterized was determined from thin-layer chromatographic (TLC) analyses of extracts from large cultures of microbes known to produce these compounds. Strain SE62 produced an iron-binding compound(s) with TLC features of different retention factors (R_f), different spot color after staining, and a larger spot-area compared to control strains (**Figure 3.1.1**). This observation of larger spot size was presumably due to a higher concentration of siderophore per unit volume of extract. This finding parallels the results from plating these microbes on plates of Chrome azurol sulfate (CAS)-doped agar (**Figure 3.1.2**) where the presence of a significantly larger halo around the cell culture indicates the microbes are producing an iron-binding compound in high amounts. These results lead to the hypothesis that this organism is producing a siderophore compound of unknown structure that may have large implications for microbial iron-cycling in marine ecosystems.

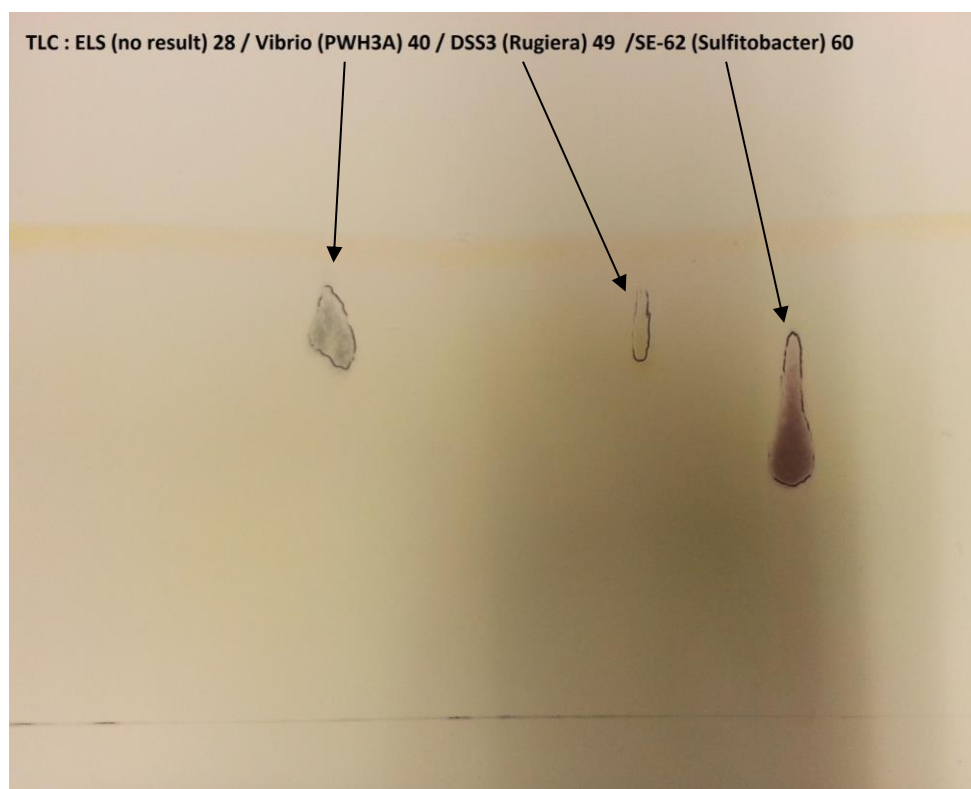


Figure 3.1.1 – TLC analysis of crude siderophore extract from Vib1, DSS3 and SE62

Shown above is a TLC plate that was developed with iodine-saturated MeOH, and subsequently visualized by exposing the plate to a 1% solution of FeCl_3 in EtOH. This TLC analysis was performed by Sebastien Guilmot whom was a student in the research lab of Dr. Steven Wilhelm.

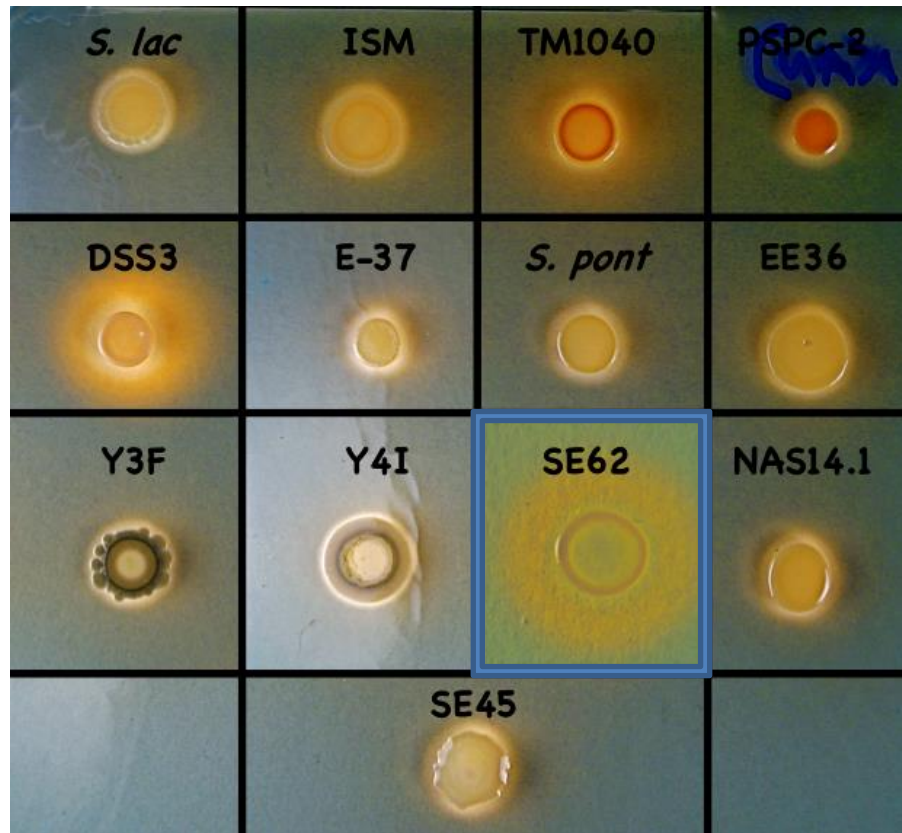


Figure 3.1.2 – Chrome azurol sulfate assay of iron-binding activity

Image used with permission from Dr. Gary LeCleir, unpublished data (created 2015). The figure shown above was received from Dr. LeCleir, a research scientist in the lab of Dr. Steven Wilhelm. The yellow halo around the cell culture indicates the ability of the microbe to leach iron from its surroundings.

Efforts toward the absolute isolation of the compound(s) responsible for the high-affinity for iron in these microbial extracts were undertaken and brought to light some important information in identifying the structure of this unknown iron-binding compound. Preparative TLC with more selective solvent systems for siderophore chromatography were employed, revealing that there are several iron-binding compounds present in the crude culture extract. A single highly concentrated fraction from sp. SE62 was isolated by preparative TLC and subjected to untargeted UHPLC-MS analysis. Two intense masses co-eluted from the LCMS experiment, showing a mass pattern indicative of these spectral features being annotated as the iron-bound and iron-unbound siderophore of interest. A UHPLC-MS/MS experiment to obtain more structural information from these putative siderophore masses revealed some unusual fragmentation patterns that did not align with typical peptide neutral mass losses, but also provided a few fragments indicative of known peptide neutral mass losses.

After optimization of TLC conditions for isolating the putative siderophore, both normal phase and reverse phase flash chromatography were performed on ~100 mg of crude extract in order to isolate a large enough sample of partially purified siderophore for NMR studies. Although the amount of compound isolated from flash chromatography (~20 mg) was enough to begin seeing single-compound definition in the ^1H -NMR spectrum and ^1H - ^1H coupling in a 2D- ^1H -NMR correlation spectrum, but the purified sample was still too impure to generate truly meaningful data. The complexity of the sample was emphasized by the absence of peaks in the ^{13}C -NMR spectrum of the entire isolated sample in a reduced-volume NMR tube. The spectra that were generated revealed that the chromatographic methods employed were able to isolate the desired compound(s) from the crude mixture, but more material was needed and other purification techniques such as semi-preparative HPLC needed to be performed in order to isolate the compound(s) absolutely for confirmatory structural elucidation by several NMR experiments.

3.2 – Introduction

3.2.1 – Siderophores

Iron is an essential element to organisms across all animal kingdoms. Its low concentration and bioavailability in certain environments provide an evolutionary pressure to develop mechanisms to mitigate this issue.⁸⁶ Iron (III) (Fe^{3+}) species are insoluble at biological pH and must therefore must be taken up by mechanisms other than diffusion.^{8, 87} Siderophores are specialized small molecules that allow microbes to leach Fe^{3+} from their surrounding environment through chelation and cellular transport mechanisms of the complexed iron into the cell to use in various enzymatic processes.⁸⁸ Free iron in the environment is difficult to come by ($\sim 1 \times 10^{-9} - 1 \times 10^{-18}$ M Fe^{3+}) as it is typically bound in minerals or horded by higher-order organisms where iron is bound to functional proteins like hemoglobin or sequestered in specialized storage proteins such as ferritins.⁸⁹⁻⁹¹

Bacterial siderophores can have a variety of functions and a variety of structural attributes that make this class of compounds diverse and challenging to characterize (**Figure 3.2.1**).⁹² Microbes in both solid surface and aquatic environments produce these compounds to not only aid in the extraction of iron from surrounding sources, but can also serve as a defense mechanism. Some bacteria produce anti-microbial siderophores that upon being absorbed by rival organisms become toxic and kill off competitors for sources of nutrition.⁹³ Many microbes have been shown to produce multiple types of siderophore compounds in iron-limited environments.⁹⁴ This is presumably to mitigate the uptake of certain siderophore compounds by competing microbes, and allow for species-selective uptake of iron-siderophore complex using compounds that other microbes do not recognize.⁹⁵

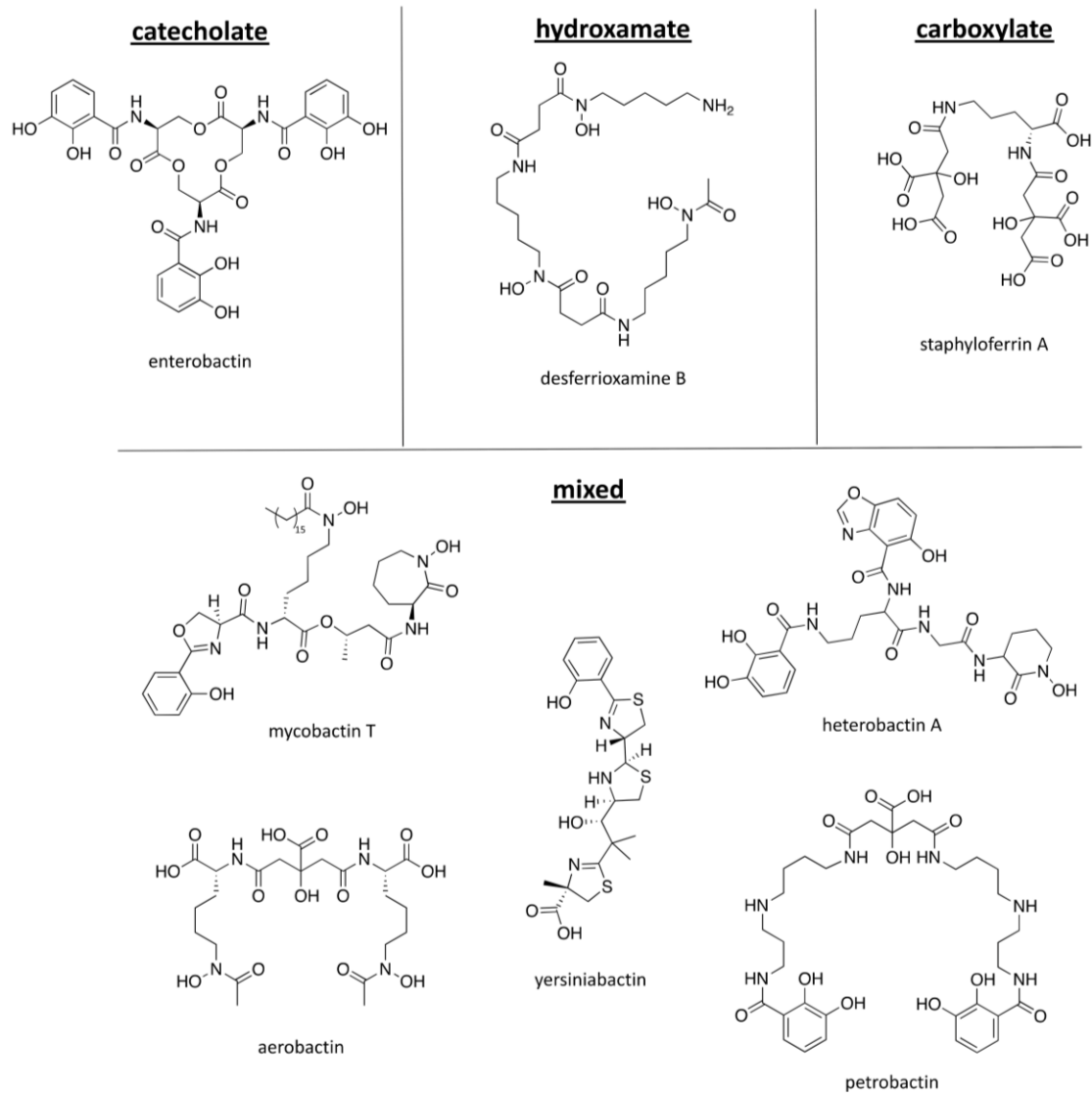


Figure 3.2.1 – Diversity of structure and functionality in siderophores

There are several types of iron-complexing compounds known as siderophores categorized by the functional groups present. What makes these compounds so challenging to characterize is that many siderophores belong to more than one class known as a mixed-type siderophore.

Iron is essential to the function of many enzymes as co-factors for metabolic processes and are universally important in the maintenance and synthesis of DNA in all organisms. The transport of extracellular iron from the environment into the cell begins with the biosynthesis of these compounds within the cytosol of bacterial cells. Once these compounds are produced, they are then excreted into the environment by membrane-bound transport proteins to diffuse into the surrounding environment and scavenge for available ferric iron species. to perform redox chemistries and complex with iron bound within minerals or to simply complex soluble iron sources. Once these compounds have sequestered iron, they may be transported into the cell by specialized membrane-localized transport proteins and subsequently reduced to release the bound iron and is mediated by siderophore complexation.

3.2.2 – *Sulfitobacter loktanella* sp. SE62

Nearly twenty years ago Dr. Alison Buchan isolated the microbe, *Sulfitobacter loktanella* sp. SE-62 (SE62), from the coastal waters of Sapelo Island, Georgia, which has spiked interest in the secondary metabolite(s) it produces that strongly bind iron.⁹⁶ The *Roseobacters* are a large clade of α -proteobacteria that make up ~20% of the microbial consortium of coastal marine environments and ~5% of the bacterial population in ocean surface waters.⁹⁷ They are a group of organisms known for their prevalence in marine environments, as well as their abundance genes coding for non-ribosomal peptide synthetases and polyketide synthases, which are an abundant source of secondary metabolites.⁹⁸ This organism, SE62, when compared to other *Roseobacters* had shown to be a significant producer of iron-binding compounds by an experiment using CAS-doped agar plates (**Figure 3.1.2**). This assay uses an agar that is treated with an equimolar combination of CAS and a ferric iron compound (typically FeCl_3) that when mixed produce a CAS-Iron complex that is blue in color. When there is a compound with a higher affinity for iron than that of CAS introduced to the blue agar, the iron is stripped from the CAS complex and the blue color is quenched.⁹⁹ It is apparent by the large halo of uncomplexed

CAS appearing around a culture of SE62 that this microbe is either producing a large concentration of iron-binding species, or that the compound(s) it produces have a significantly higher affinity for iron than the compounds other *Roseobacters* are producing, which also may result in a larger sequestration of iron from the CAS-doped agar plate. Identifying this compound, or compounds that this microbe is producing could have large implications for marine iron cycling not only in the coastal waters of Sapelo Island in Georgia, but if this microbe has a presence in other aquatic environments this compound could be used by several organisms to support their own survival.

3.2.3 – Natural product isolation

The isolation of natural products produced by living organisms is a field of chemistry that seeks to purify molecules with potent bioactivities for medical purposes such as in the case of the chemotherapeutic agent, taxol. In a review by Otto Sticher, it is mentioned that natural products are going to always be a source of potent new drugs due to the vast diversity of compounds produced as secondary metabolites in living systems that have equally diverse activities *in vivo*.¹⁰⁰ Sources of these medically relevant natural products come from innumerable biological sources from all the animal kingdoms. This work involves first identifying an organism that produces a compound or set of compounds that have bioactivity, gathering the biological matrix that contains the compound(s) of interest, extraction of said compounds, absolute isolation of said compounds, and finally performing a slew of instrumental and chemical analyses that generate data sufficient to determining the structure of the isolated natural product. This type of work often provides very small amounts of the target material from a much larger sample of raw biological material, which is where chemical synthesis is of great benefit. Once the structure of the compound is determined, it is a matter of developing a synthetic method that can achieve the desired molecule. Once a potent natural product is synthesized, many research groups look to increase the atom economy of the synthetic

strategy employed so that sufficient quantities can be made in a cost-effective and timely manner.

3.2.4 – Technologies for small molecule structure elucidation

There are several typical means of determining the structure of an unknown organic compound from chemical syntheses, where there are quantities of by-product that can be isolated from easily scalable reactions. These technologies are also used to determine the structure of an unknown compound isolated from nature with interesting properties. Tools such as nuclear magnetic resonance (NMR), full-scan and targeted tandem mass spectrometry (MS / MS-MS), or spectrophotometric instrumentation (IR, UV-Vis, polarimetry) are employed and typically used in conjunction to give enough detail to definitively identify a structure. X-ray diffractometry is also another powerful, albeit less commonly employed, technique for structural elucidation of small molecules since this technique requires very pure crystalline samples. Obtaining samples of pristine purity from biological matrices in quantities to grow crystals requires a fairly large amount (≥ 1 g) of biological material to isolate enough pure compound for appreciable crystal growth. Current NMR technologies can acquire reasonably informative spectra (1- and 2-dimensional homo/heteronuclear experiments) with as little as 100 μ g of pure samples.¹⁰¹ Spectrophotometric analytical methods such as UV-Vis and IR may be routinely used with sample quantities in this region, and mass spectrometry is so sensitive that this amount of sample would be more than adequate even when coupled to chromatographic instrumentation. Although these techniques require a small amount of pure sample, the absolute isolation of natural products still may require a considerably larger amount of biological material to isolate the compound of interest. As previously mentioned, the Pacific Yew tree is a natural source of the powerful anti-cancer drug Paclitaxel (Taxol), but the pure compound only makes up on average 7.5×10^{-5} weight-% of the dry bark harvested from the tree.¹⁰²⁻¹⁰³ This should emphasize the need for large quantities of raw material to isolate single compounds for structural elucidation.

3.3 – Results and discussion

3.3.1 – First-pass screening of culture extracts by UHPLC-UV-MS

Samples that were received from the Wilhelm group were siderophore-selective extracts of presumably simple composition, which seems to be a common misconception that research groups in the biological sciences tend to make. This is not because they do not understand the complexity of biological samples, but rather the opposite – they are aware that extracts of cells or higher-order tissues are extremely complicated in composition, therefore purification methodologies that are well-established are typically employed to isolate target compounds. These methodologies are successful in isolating target compound classes but are typically laden with numerous molecules of specific functional chemistries that behave chromatographically similar or have other comparable physical properties that may sometimes be overlooked, which results in numerous compounds being isolated alongside target molecules or classes of molecules. In the instance of isolating the compound(s) responsible for binding iron toward characterization of an unknown siderophore produced by the *Sulfitobacter loktanella* SE-62, the use of poorly selective TLC conditions (normal-phase silica and “iodine-saturated” MeOH as mobile phase) assumed that there was primarily a single compound isolated by the methods employed, which under more scrutinous investigation revealed that this single “compound” was in fact a very complicated and diverse sample in and of itself.

Before any rigorous cleaning of the samples was employed in the analysis of the siderophore selective extracts from Vib1, DSS3, and SE62 obtained from the lab of Dr. Steven Wilhelm, it seemed practical to first get an idea of the complexity of the samples by means other than TLC. Some type of analysis that could possibly illuminate a set of mass features unique to SE62 that may be contributing to the strong ability to leach iron from its environment. This was accomplished using ultra high-performance liquid chromatography coupled mass spectrometry (UHPLC-MS) to gather preliminary data and

compare spectral features between samples. Since these samples were extracted from cell cultures, and the overall polarity of the iron-binding compound was not known, but it should be quite polar in nature in order to act as a ligand to complex with ionic iron species. From this imagining of physical characteristics of the target molecule, it was decided to use a HILIC column for analytical separations. There is importance with characterizing an unknown molecule to gather as much information you can on a sample, this includes spectrophotometric means of obtaining vague insight into the functionality of the compounds present in the sample. Information on how these samples respond to spectroscopic detection such as with a flow-through photodiode array (PDA) detector gave another level of qualitative data to view similarities and differences between the samples. As expected, the total ion chromatogram (TIC) (**Figure 3.3.1**) for the crude samples was extremely complex, and had spectral features that were similar, but also dissimilar, which is not unexpected. The PDA chromatograms (**Figure 3.3.2**) from this experiment also gave somewhat similar results between samples. It was obvious that the samples would have to be purified prior to analysis to give a better idea of which mass features were responsible for the iron-binding activity of these extracts. Having this data from the crude extracts allows for searching for the extracted ion chromatogram of the mass features in the purified samples and lining them up with their respective PDA responses at different excitation wavelengths giving a vague idea of functional groups that may be present in the detected compounds.

3.3.2 – Colorimetric analysis and preparative thin-layer chromatography

The compound class a siderophore belongs to can be qualitatively determined with colorimetric tests such as Arnow's assay and Atkin's assay. Arnow's assay is specifically suited for catechol-type siderophore, which upon treatment of the siderophore with a nitrous acid, and subsequent treatment with sodium hydroxide gives a pink color. Atkin's assay tests for the presence of a hydroxamate-type siderophore by treatment of the sample with iron(III) perchlorate, which gives an orange color. These

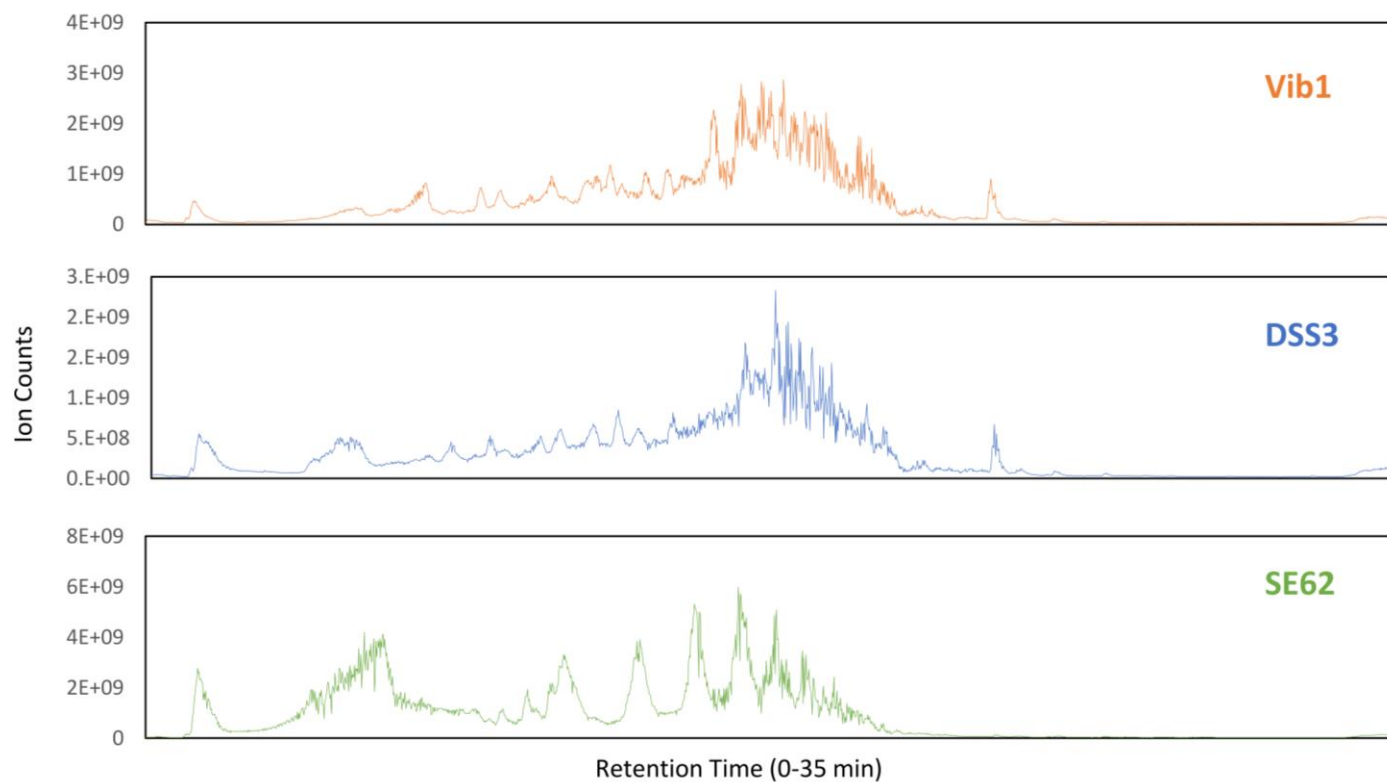


Figure 3.3.1 – Total ion chromatograms from the UHPLC-MS analysis of crude siderophore extracts

Shown above are the total ion chromatograms for the crude siderophore extracts of Vib1, DSS3, and SE62 samples.

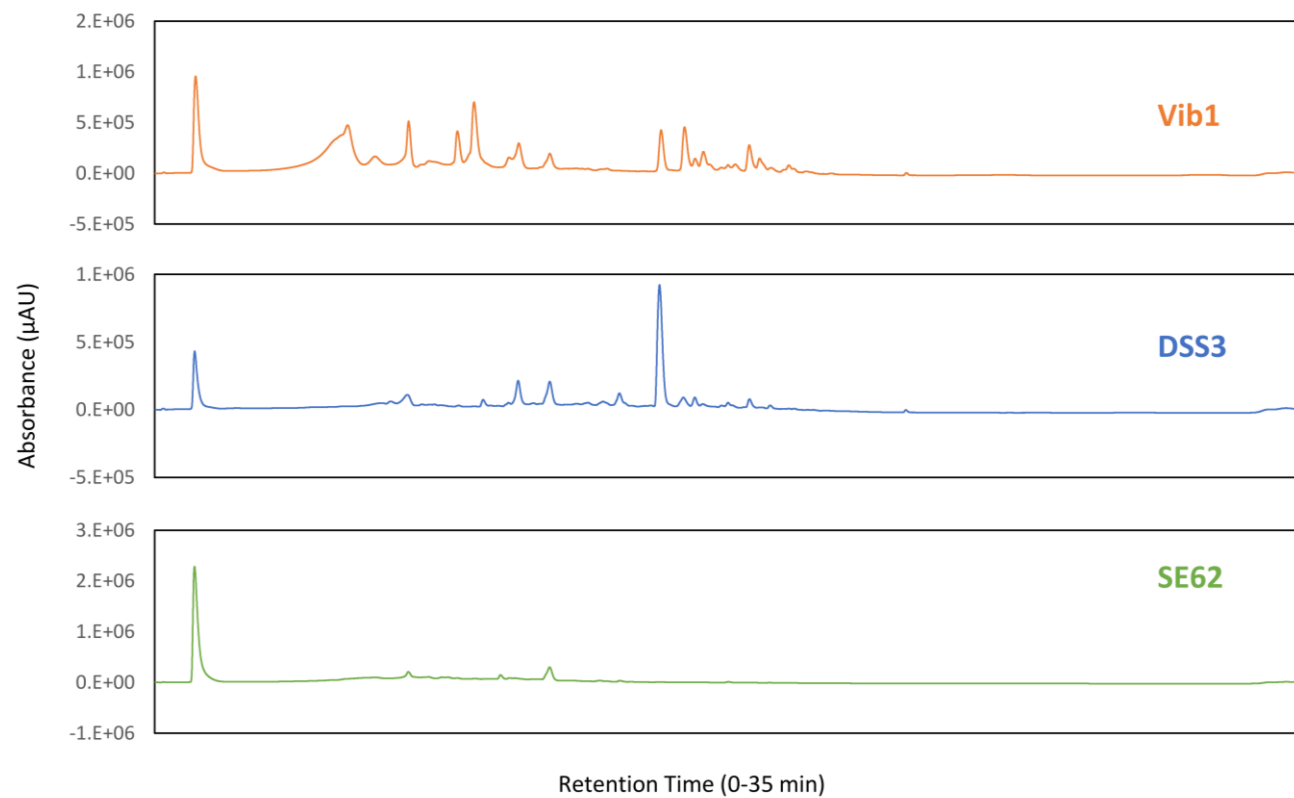


Figure 3.3.2 – Select PDA chromatogram from UHPLC-MS analysis of crude siderophore extract

Shown above is the PDA chromatogram being collected at 254 nm (bw = 16 nm) for the crude siderophore extracts of Vib1, DSS3, and SE62 samples

methods are good for non-mixed classes of siderophores that only respond to one or the other assay. Developed TLC plates with the crude SE62 sample upon colorimetric analysis via assays of Arnow and Atkin did not yield any coloration indicating these isolated compounds were neither of a hydroxamate or catecholate type. The approach of this experiment was only to identify TLC spots as having an affinity to iron, which was tested by selectively staining of the developed TLC plate with a solution of FeCl_3 . By this method, it would be apparent if spots on the TLC plate bound to iron, they would have a red/orange color, compared to the rest of the yellow plate. This was also the method of choice since it best replicated the TLC analysis performed by the Wilhelm group – using normal-phase silica TLC plates as the stationary phase, and iodine-saturated methanol as the mobile phase. By replicating their method, using iodine-saturated methanol then staining with 5% FeCl_3 in MeOH exposed spots that were different in color and retention factor (R_f). Vib1 gave spots that were dark grey/blue at $R_f = 0.8-0.9$, DSS3 gave a faint orange spot also around $R_f = 0.8$, and SE62 gave a large dark merlot spot at $R_f = 0.5-0.7$ (**Figure 3.1.1**). Modifications to the collaborators TLC mobile phase to be less polar and chemically reactive (exclusion of iodine in mobile phase) resulted in separations that resulted in more selective chromatography than just the iodine saturated MeOH (**Figure 3.3.3A**), but only revealed orange spots after staining. Reverse-phase separations were also attempted with the crude extract using C_{18} -silica TLC plates and mobile phases of varying $\text{H}_2\text{O}:\text{ACN}$ ratios, but all separations proved to be poorly selective giving a high R_f (≥ 0.95) value for all iron-binding compounds in a single TLC spot. This confirmed that normal-phase chromatography would be the best method of isolating these compounds, which should be apparent due to the polar nature of these metal chelators.

Staining the optimized TLC plate separations with 5% FeCl_3 in MeOH revealed that there were several spots in the SE62 sample responding to stain containing iron, meaning there were several compounds present on the plate that technically could be identified

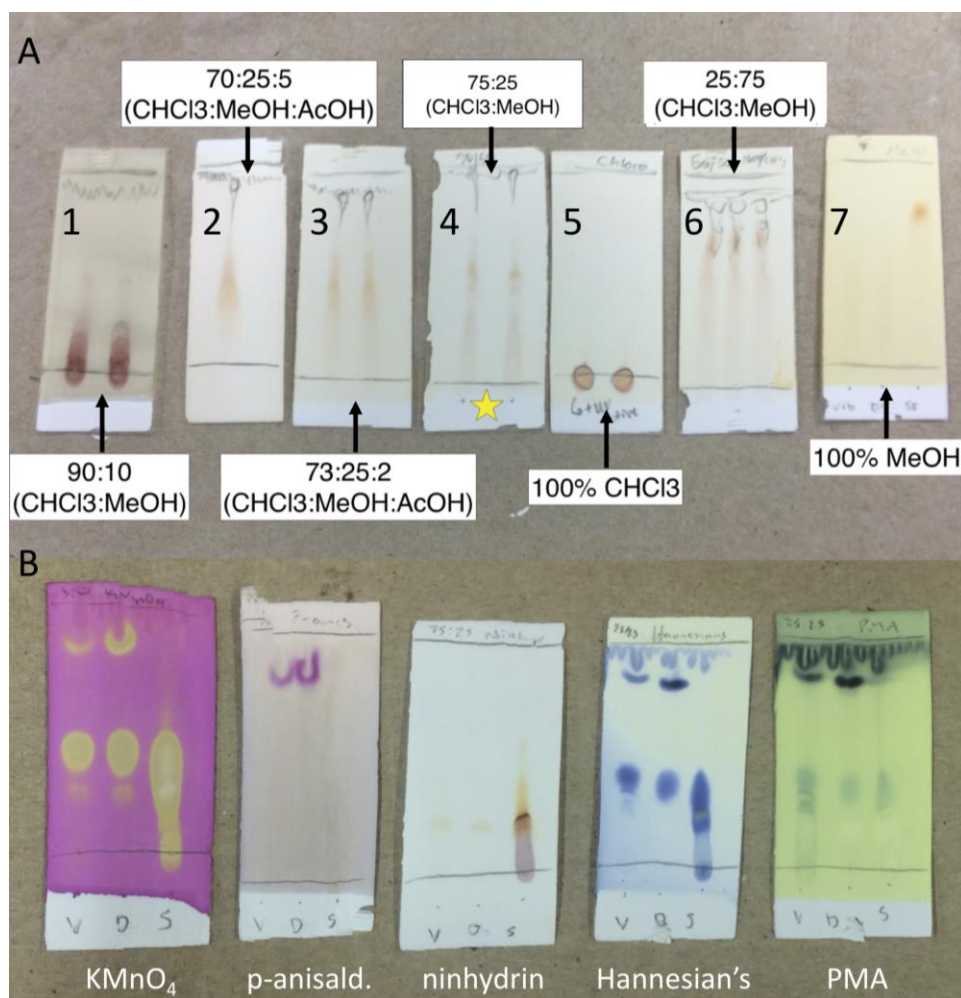


Figure 3.3.3 – Thin layer chromatographic analysis of crude siderophore extracts

A.) Thin layer chromatography (TLC) optimization for fractionation of crude siderophore extracts of vib1(v), DSS3 (d), and SE62 (s), where plate 4 mobile phase conditions were found to give the best separation. Plates were stained with an ethanolic solution of FeCl₃.

B.) Differential staining procedures using potassium permanganate, p-anisaldehyde, ninhydrin, Hannesian's, and phosphomolybdic acid to get an idea of functional groups present in the iron binding compounds.

as siderophores, or at least respond/react with ferric iron. To obtain more information from this preparative TLC method, several replicate plates were developed and visualized with a variety of TLC stains that would reveal more about the chemical identity of whatever was being separated on the plate (**Figure 3.3.3B**). Narrowing down what functional groups may be present in the iron-binding fraction versus the other non-iron reactive regions of the TLC plate will be useful for determining possible functional groups constituent of the unknown compound(s) of interest and how the crude extracts compare to each other. Preliminary TLC analyses with an array of TLC stains show an assortment of structural features for the iron-binding regions in the SE62 sample due to differential staining properties of a variety of possible functional groups.

Staining a plate with an alcoholic solution of KMnO_4 indicates the presence of oxidizable functional groups such as alcohols, amines, alkenes, and alkynes - the resulting compounds from the action of KMnO_4 stain yellow against a purple plate. Amines and amino acids stain brown against a fairly colorless background with a solution of ninhydrin. Several general stains were also employed to determine the true complexity of the crude samples. From these staining solutions alone, it was determined that the compound of interest was likely a peptide due to the brown staining via ninhydrin, and likely contained either oxidizable functional groups or itself contained carboxylic acid moiety by the yellow staining behavior with KMnO_4 . Although some TLC staining methods are “universal” they do not always stain the same things or with the same intensity. Therefore a few plates were treated with Hannesian’s stain (acidic solution of ceric ammonium molybdate) and alcoholic solutions of p-anisaldehyde and phosphomolybdic acid which stain compounds unselectively giving spots of blue, pink, and green, respectively. These staining procedures did not all behave the same with this sample attesting to a diversity of molecule classes or functional groups present in the compounds separated by the optimized TLC method.

One benefit of preparative TLC is that separations may be done on very small amounts of sample and that once the R_f for the compounds of interest have been

established, the experimenter may then excise the desired fraction from the TLC plate by scraping off the silica. This silica sample has the compound(s) of interest adsorbed to the surface of the silica, and can be extracted with small volumes of solvent, and directly injected onto an analytical column for LCMS analysis to determine the masses of the compounds present in that fraction of the TLC analysis. Using the excised TLC fractions between all strains analyzed can assist in pinpointing the mass of the molecules unique to the SE62 strain from any other more common cellular compounds that may also be present in the other strain extracts.

3.3.3 – UHPLC-HRMS analysis of Fe-binding TLC fractions

Mass spectrometric analysis of the TLC fraction with the strongest staining behavior is a starting point towards isolating and identifying the compound(s) responsible for binding to iron in this sample. The TLC spots from the preparative TLC experiment with SE62 were excised from the plate and extracted with MeOH, filtered, and transferred directly to an autosampler vial for injection onto an analytical column for UHPLC-MS analysis. Controls to account for metabolites and lipids that may be present in the iron-binding TLC fraction were used to assist in pinpointing spectral features that were unique to the SE62 strain. These controls were prepared by preparative TLC of the extracts from Vib1 and DSS3 cell cultures, excising the same R_f region as with the SE62 sample, extracting with MeOH, filtering, and transferring the supernatant into autosampler vials for analysis.

Samples were subjected to UHPLC separation on a Kinetex reverse phase C18 (100 x 2.1 mm, 2.6 μ , 100 Å) analytical column prior to mass spectrometric analysis. The choice of using reverse phase chromatography in this situation was due to the fact that normal phase chromatography was already performed on these samples (prep-TLC). This would perhaps enforce some separation in lipophilic space if amphiphilic molecules retained similarly to the target compounds with this method of sample preparation. Siderophores are highly polar compounds, meaning that the compounds of interest in these samples

would likely be pushed off a reverse phase column first, leaving behind compounds with a degree of hydrophobicity that were likely components of cell membranes. Using an electrospray ionization (ESI) source operating in positive mode, the column effluent was introduced into the mass spectrometer as positively charged gas-phase ions to be detected by an orbitrap mass analyzer.

This analysis proved to be quite informative by the identification of co-eluting mass features in the total ion chromatogram of the SE62 sample that indicated a siderophore-like compound. These detected ions shared similar ^{13}C -isotope peaks, a mass pattern indicative of being bound and unbound to iron, and the correct $^{54/56}\text{Fe}$ isotopic ratio for the iron-bound ion (**Figure 3.3.4**). Although the detected ionic species took some thought to realize that these were related ions bound and unbound to a ferric ion, and not just unrelated masses that co-eluted, the $[\text{M}+\text{H}]^+$ and $[\text{M}-2\text{H}+\text{Fe}^{3+}]^+$ ions were annotated as such, and the detected mass of 671.3230 m/z mass was determined to be a strong candidate for the iron-binding behavior of the siderophore sample.

3.3.4 – HRMS data to calculate potential molecular formula

Data from the UHPLC-MS experiment was obtained using the mass spectrometers highest resolution setting (resolution = 140,000), which can be used to generate a list of reasonable chemical formulas based on atomic constraints such as the number and identity of atoms the program can use to generate formulas close in mass to the masses observed in the spectrum. There are a number of programs that can do this task, but Xcalibur (Thermo Scientific) was used in this instance to generate a plausible molecular formula based off the masses that were representative of the iron-bound and -unbound siderophore species. Since this molecule class (siderophores) is not traditionally observed in metabolomic mass analyses, it seemed unnecessary to use the Seven Golden Rules method of chemical formula calculation. The Seven Golden Rules is an Excel program developed by Dr. Tobias Kind and Dr. Oliver Fiehn that takes certain atomic ratios that are typically found in biogenic compounds as well as considers isotope patterns and the

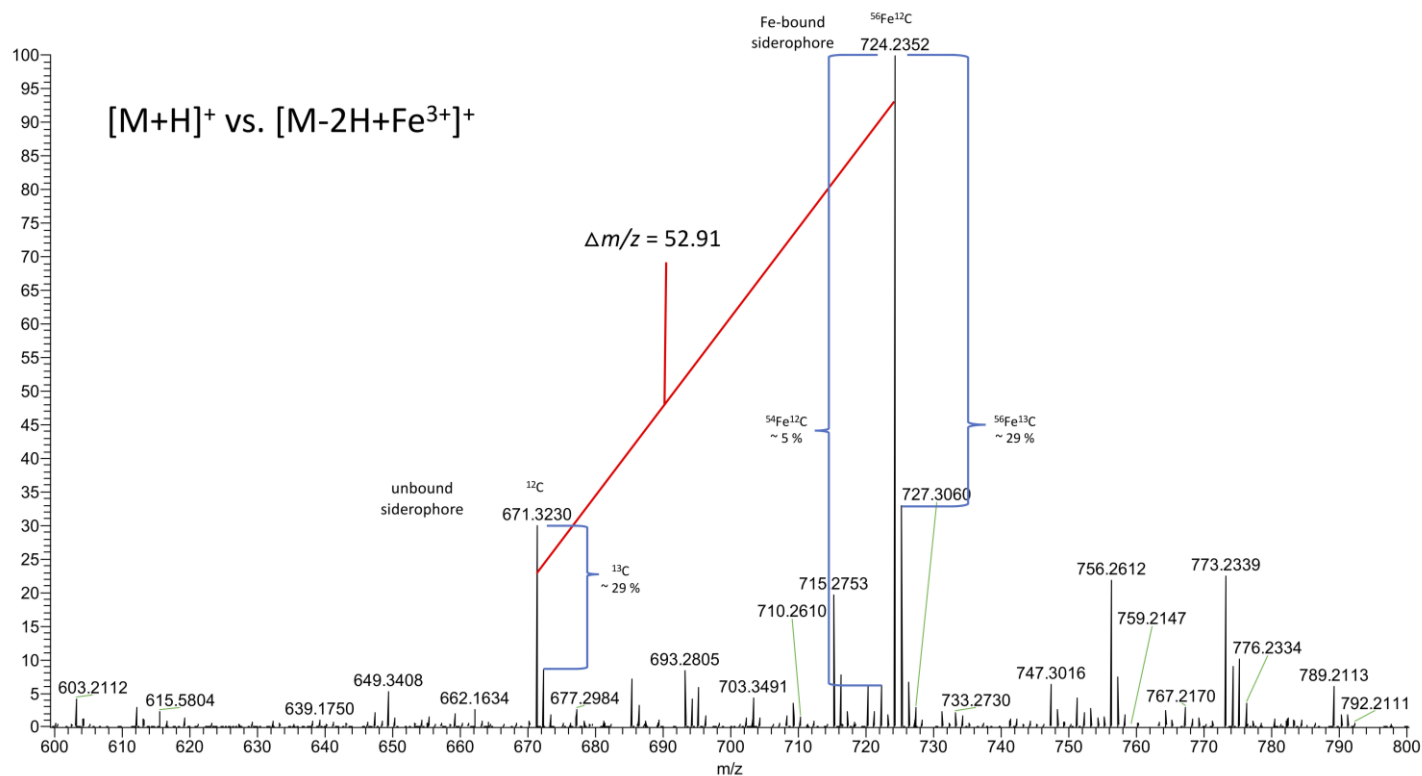


Figure 3.3.4 – Mass spectrum of putative siderophore in *Sulfitobacter laktanella* sp. SE62

With a similar ^{13}C isotope ratio for these two masses, and the correct mass difference for the ions being bound and unbound to iron, this pair of masses have been preliminarily annotated as the putative siderophore of interest.

presence of the penultimate formulae in online mass databanks before finalizing the list of possible chemical formulas.¹⁰⁴ Typically, siderophores contain atoms of carbon, hydrogen, oxygen, nitrogen, and sulfur to a lesser extent – this lead to constraining the Xcalibur formula generator to these atoms and keeping the mass error below 5 ppm, as dictated by the instrumentation (orbitrap mass analyzer) used during mass analysis. By this method, the formula generator came up with 13 plausible formulas. Of the top hits from this list, there was a formula that aligned well with the isotopic pattern observed (~29-30 carbon atoms) from the HRMS analysis of this sample as well as a rough 2:1 ratio of oxygen to nitrogen that would be present in a peptidic carboxylate-type siderophore. It was this formula ($C_{29}H_{46}N_6O_{12}$) that was used to further extrapolate a potential molecular structure.

3.3.5 – LC-MS-MS analysis of putative siderophore

Previous preparative TLC and UHPLC-MS experiments have shown that although the unknown iron-binding compound of interest can be partially isolated by these methods, it was not isolated absolutely. Data acquired from the UHPLC-MS study of the preparative TLC sample shows there is still much sample complexity in the total ion chromatogram where the characteristic iron-binding compound elutes, even after liquid chromatographic separation on an analytical column. In order to obtain data that could be used to construct a plausible structure for the unknown iron-binding compound of interest, tandem mass spectrometry (MS-MS) was performed on a Thermo Scientific Q-Exactive mass spectrometer via parallel reaction monitoring (PRM). A PRM method uses what is known as a mass inclusion list, which is a list of masses that the ion optics isolate, fragment, and inject into the mass analyzer in succession in-time with the duty cycle of the instrument. Unlike a traditional triple-quadrupole instrument that uses the tandem quadrupoles to isolate, fragment, and analyze a user-selected ionic species, the Q-Exactive uses a segmented quadrupole to select a parent ion to enter a higher-energy collisional dissociation cell (HCD cell) to fragment the ions prior to being introduced into

an orbitrap mass analyzer, which provides the high-resolution mass analysis. As determined by the UHPLC-MS analysis of the strongly iron-binding fraction from preparative TLC of a crude siderophore sample, the iron-unbound species of interest that was detected was 671.3230 m/z , this was the ion selected for fragmentation analysis to determine structural features of this unknown compound.

3.3.6 – Scale-up purification and NMR analysis of unknown siderophore

Once preparative TLC was shown to be effective in isolating the mass detected in previous experiments showing a mass pattern indicative of a siderophore bound and unbound to ferric iron, it was time to apply the optimized chromatography to the rest of the crude siderophore extract. This was accomplished using normal-phase silica gel packed into a flash column, and the polarity of the most effective TLC mobile phase was reduced by about 60% (~10% MeOH in CHCl_3 versus 25% MeOH in CHCl_3) to ensure the chromatography was not too rushed. When adopting TLC separations to flash chromatography, the polarity of the mobile phase needs to be reduced because of the change in how the mobile phase runs through the stationary phase. In a TLC-based separation, mobile phase relies on capillary action to pull the liquid up and through the stationary phase of the plate, whereas with a flash column, the mobile phase either is allowed to flow through the stationary phase via gravity or is pushed through with air pressure. In both instances with a flash column, the retention of compounds on the stationary phase changes drastically compared to being “pulled up” a TLC plate, and therefore the mobile phase used must compensate for this change in fluid dynamics. Throughout the flash column, very small (5 mL) fractions were collected in clean test tubes and tested for their ability to bind iron via spotting each fraction onto a TLC plate, running the plate with the optimal TLC mobile phase, and finally staining the plate with a 3% solution of FeCl_3 in EtOH. Using this method, the fractions that gave a TLC spot with an equivalent R_f to that of the spot of interest in the crude mixture were pooled and solvent was removed prior to NMR analysis. Proton NMR (^1H -NMR) revealed fairly

resolved peaks in the spectrum, but contained a number of peaks that looked to be unrelated to the compound of interest due to odd integration patterns and messy splitting of the proton signals (**Figure 3.3.5**). The sample complexity was further confirmed by the lack of signal intensity in the carbon NMR (^{13}C -NMR) spectrum (**Figure 3.3.6**), which was probably due in part to the small amount of sample obtained. However, 20 mg of a pure sample in a reduced-volume NMR tube should give appreciable signal, especially after 1000 scans. Therefore the assumption was made that this sample was still complex – due to a potentially large number of carbon atoms experiencing different chemical environments and relaxing at a variety of frequencies, which upon subjecting to a large number of scans would not sum to a noticeable peak in the spectrum. Homonuclear correlation spectroscopy (COSY-NMR) was performed in a final attempt to deconvolute the data obtained from the ^1H -NMR experiment on the isolated iron-binding species of interest (**Figure 3.3.7**). Solution-phase ^1H - ^1H - COSY NMR spectrum gives the analyst the ability to see which neighboring proton signals are spin-coupled, which would give some clarity to complex splitting patterns and what proton signals are likely coming from a single molecule.

3.3.7 – Data analysis and construction of a potential siderophore structure

Considering the exact masses of the iron-unbound parent compound and the masses of the fragments generated from tandem-mass analysis of the intact siderophore, a plausible molecular structure could be constructed from these data. Information gathered from the fragmentation experiment revealed there were some fragments that were characteristic of peptide neutral losses, so the parent compound may have been a peptide of some sort (**Figure 3.3.8**). An important neutral mass loss of $\sim 130.05\ m/z$ was observed between three of the fragments generated, indicating that there may be three glutamate residues present in the putative siderophore, which would make sense as these residues under the correct pH conditions would be ideal for chelation to ferric iron in solution. There were also several fragments indicative of loss of ammonia from the

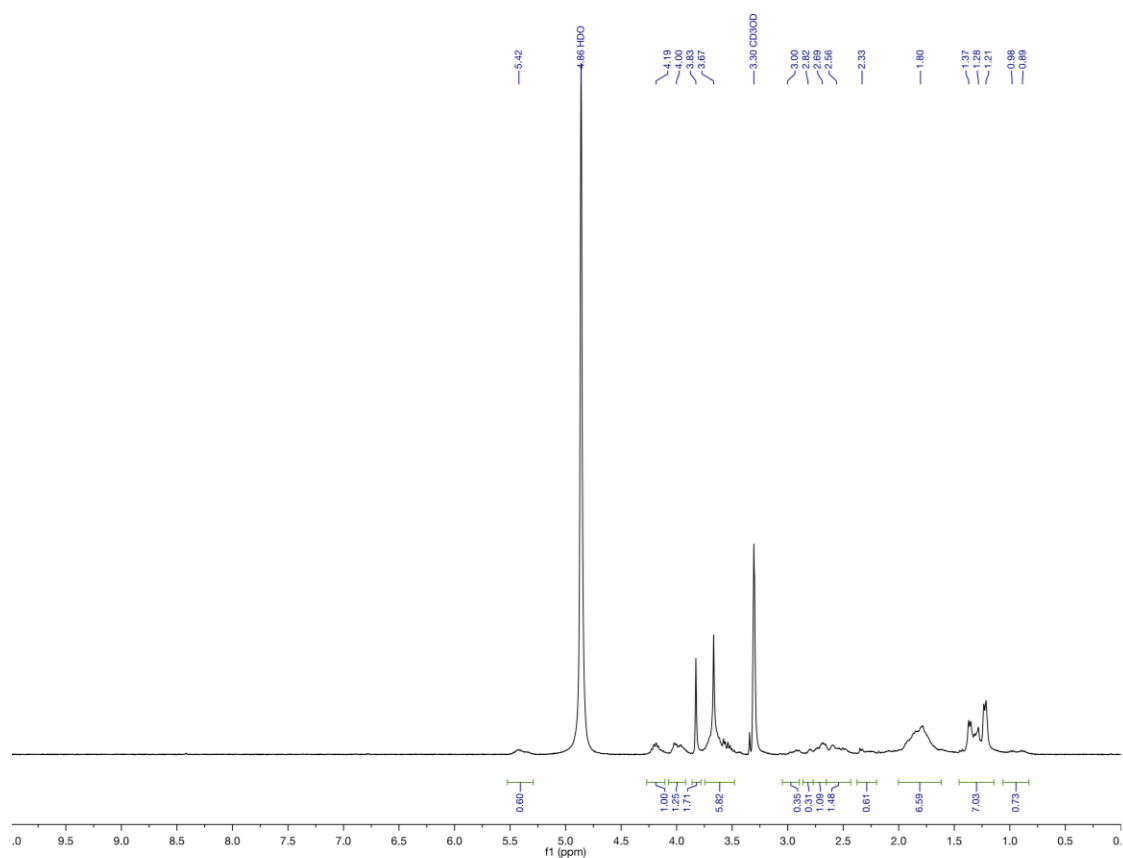


Figure 3.3.5 – ^1H -NMR of partially purified SE62 siderophore

Shown above is a proton NMR spectrum of the “purified” siderophore sample with strong iron-binding behavior by TLC staining. Peak integrals were normalized to a single proton for the multiplet at 4.19 ppm, as this may be a methine unit typical of a peptide. The spectrum was taken in deuterated methanol, and there is a fair degree of water present in the sample.

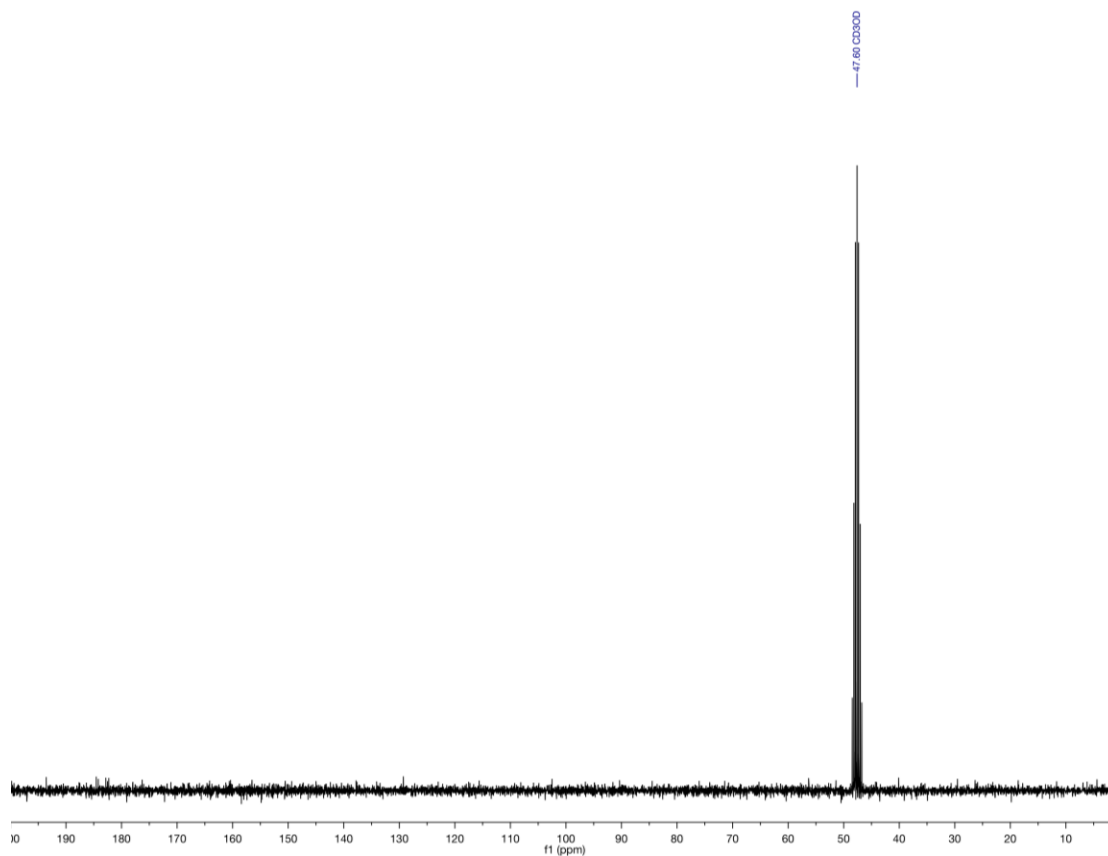


Figure 3.3.6 – ¹³C-NMR of partially purified SE62 siderophore

Shown above is a carbon-13 NMR spectrum that practically shows zero signal intensity from the sample, and only a solvent residual peak. This may be due to the complexity of the isolated sample, since the concentration and number of scans taken (1000) for this spectrum should give appreciable signal if the sample was pure.

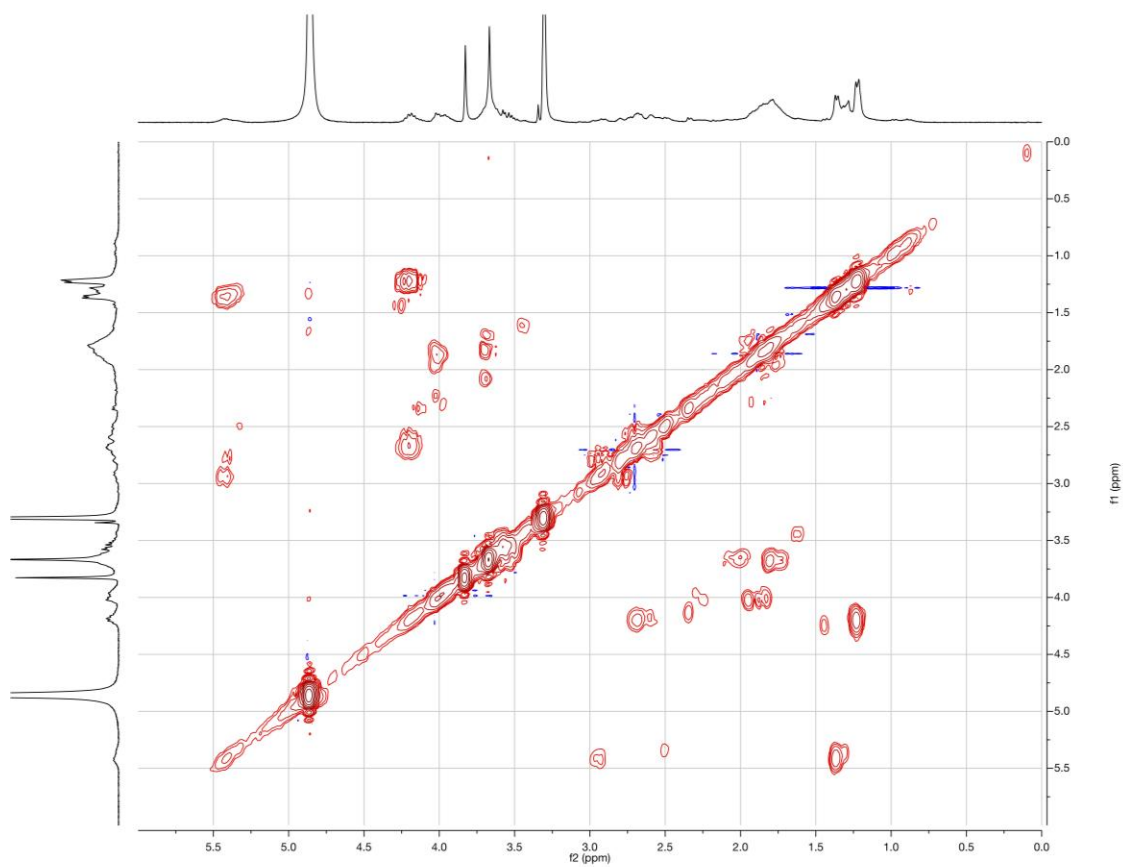


Figure 3.3.7 – ^1H - ^1H homonuclear correlation spectrum of partially purified SE62 siderophore

Shown above is a COSY NMR spectrum that indicates what proton signals are coupled to each other.

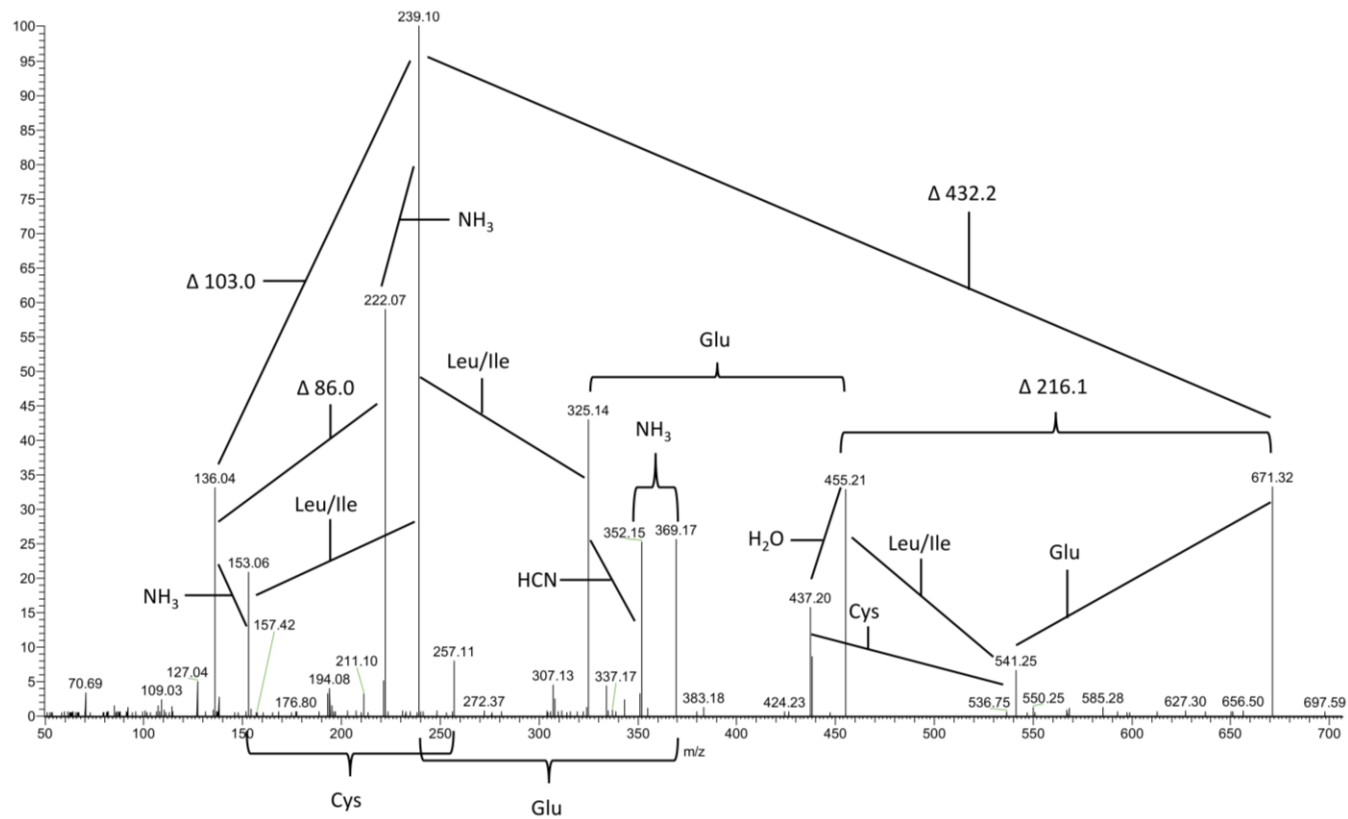


Figure 3.3.8 – Tandem mass spectrometric analysis of putative SE62 siderophore

MS/MS of iron-unbound siderophore of interest showing possible peptide neutral mass losses

structure, also indicating that this compound may be a peptide. By piecing together the data from the original UHPLC-MS analysis which detected masses of the iron-bound and unbound siderophore, fragments from the MS-MS experiment, keeping amino acid structures in mind, and calculating the remaining atoms available from the generated molecular formula, the cyclic hexapeptide ESEKEA (Glu-Ser-Glu-Leu-Glu-Ala) was constructed and has a calculated mass error of 1.63 ppm from the observed mass (**Figure 3.3.9**) indicating that this combination of atoms was a likely candidate for serious consideration as the putative siderophore of interest, but the configuration of those atoms was not necessarily correct, which NMR experiments would be essential for determining.

This mass analysis alone yielded a promising conclusion for the potential identification of the unknown iron-binding compound present in the SE62 sample but needed confirmation with more analytical data better suited to give structural information on the isolated compound. Therefore, flash chromatography was performed on the SE62 sample, in order to isolate a large enough amount of this compound for NMR studies such as ^1H -NMR, ^{13}C -NMR and a few 2D NMR studies. Normal-phase chromatography was found to be better suited in the separation of iron-binding species than reverse-phase chromatography (C_{18} silica plates) by TLC analysis, so flash chromatography was performed on the crude SE62 siderophore extract yielding ~20 mg of product for NMR analysis. Unfortunately, the spectra generated from the NMR analysis did not give much meaningful information due to the lack of signal clarity in the ^1H -NMR spectrum, the lack of signal intensity in the ^{13}C -NMR spectrum, or the cross-peaks, although quite well-resolved, observed in the ^1H - ^1H COSY spectrum. There was a fair degree of signals observed in the ^1H -NMR spectrum that aligned well with a predicted NMR spectrum generated by ChemDraw Professional 16.0 software (**Figure 3.3.10**). The solvent setting used for predicting the ^1H -NMR spectrum was CDCl_3 , so the predicted peaks are likely going to be shifted from the experimental NMR, which was taken in

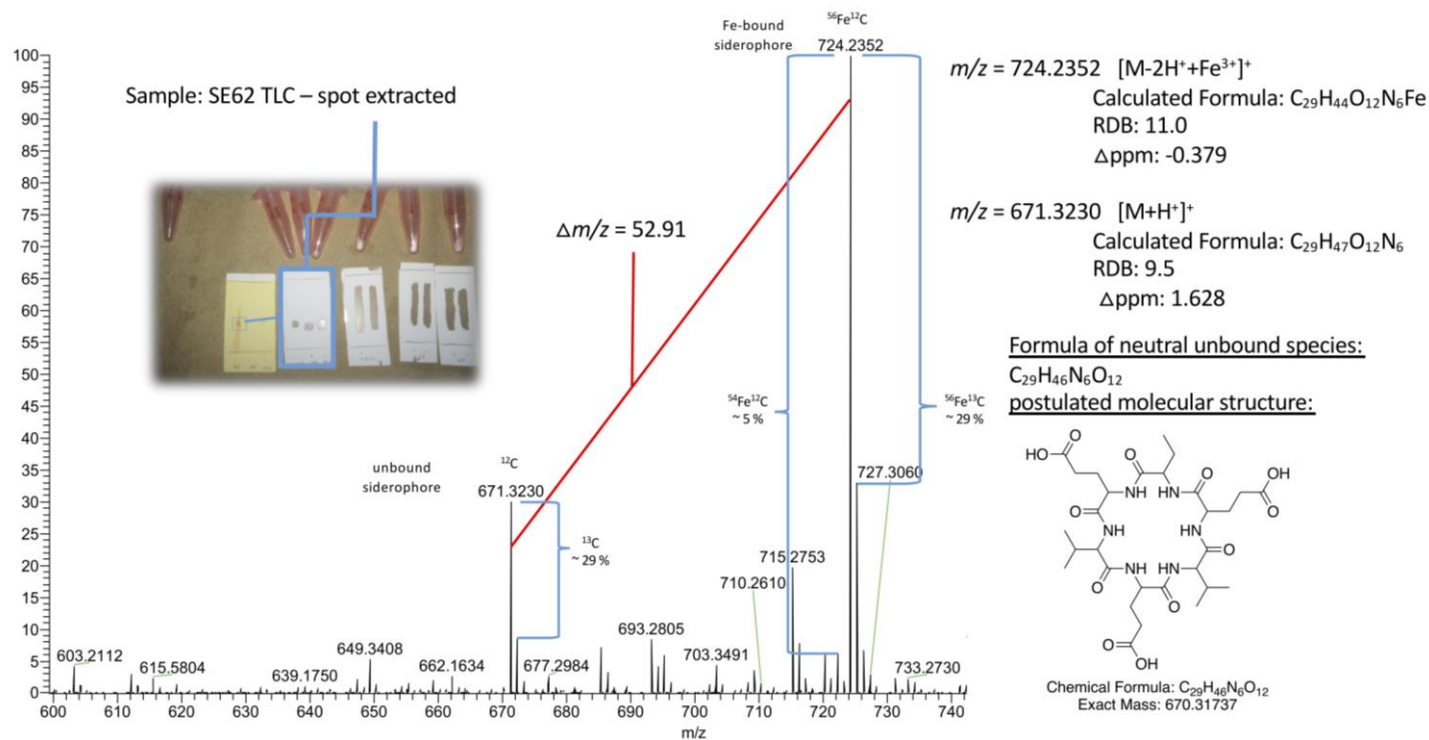


Figure 3.3.9 – Possible SE62 siderophore structure from MS data

Using exact masses, plausible chemical formulae were calculated, and a structure was put together using MS/MS data. The proposed structure has a very small mass-error to that of the detected mass and is a good candidate for preliminary identification.

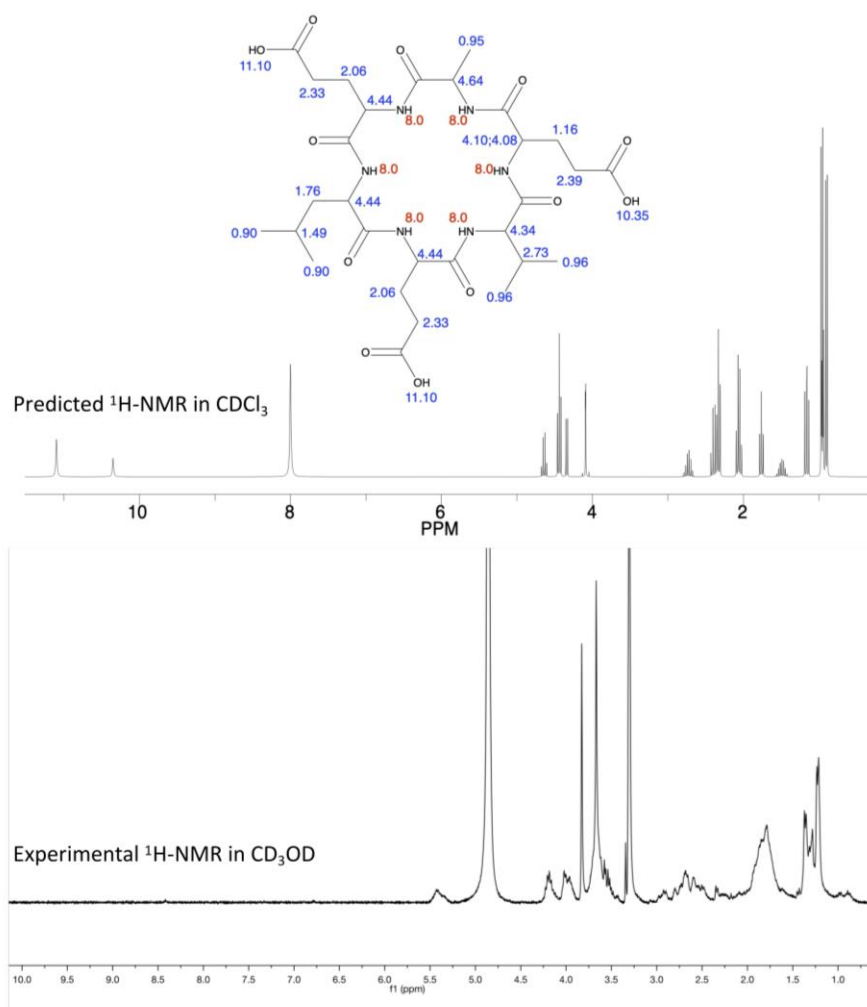


Figure 3.3.10 – Predicted ^1H -NMR spectrum of the putative SE62 siderophore structure

A structure that was created from MS data was used to generate a predicted ^1H -NMR spectrum to compare against the experimental spectrum. Although the experimental sample was taken in a different solvent than the program calculated the predicted spectrum, the two spectra line up fairly well, but further work needs to be done to obtain a spectrum with more definitive peaks and splitting patterns.

CD₃OD. The choice of using a deuterated solvent with exchangeable deuteriums unfortunately would have quenched the signals arising from any exchangeable protons in the compound of interest, such as amines and carboxylic acids, but this solvent seemed to be the one that dissolved the sample the best and was therefore used.

3.4 – Methods

3.4.1 – Iron-limited cell culturing

All procedures for culturing were obtained from Dr. Gary LeClerc, a research scientist working under Dr. Steven Wilhelm at the University of Tennessee, where the culturing took place. In order to encourage the production of siderophores in sp. SE62, extra care was taken to limit the amount of iron present in their environment. All containers were acid-washed to ensure the removal of trace metals and liquid media was prepared with Chelex-100-treated water in order to achieve an iron-deficient culturing environment. A special room equipped with large high-efficiency particulate air (HEPA) filters was prepared in Dr. Wilhelm's lab to simulate a "clean room" following trace metal-free atmosphere protocol. This was done in order to minimize the incidence of contamination during liquid transfers; since trace amounts of iron could be introduced into the liquid media from sources as simple as a fleck of dust or a stray droplet of iron-contaminated water.¹⁰⁵⁻¹⁰⁶

Cell cultures of *Vibrio* phage PWH3a-P1 (Vib1), along with *Roseobacter* strains *Silicibacter pomeroyi* DSS-3 (DSS3) and *Sulfitobacter loktanella* SE-62 (SE62), were prepared by personnel in the Wilhelm group adhering to the following procedure: Resurrection of previously collected specimen began with plating frozen bacterial isolates on YTSS agar (yeast, tryptone, sea salts, ~1.5% agar) to achieve viable cultures. Single colonies of revived organisms were transferred to plates with a defined marine media, "C-bud" (artificial seawater with varying levels of yeast extract and tryptone), to produce a batch of viable specimen to transfer into an iron-limited liquid C-bud medium.¹⁰⁷ C-bud

liquid media was passed through a column of Chelex X-100 ion exchange resin to remove iron and were subsequently sterilized prior to inoculation. Small 50 mL batches of this liquid media were inoculated with cultures grown previously on C-bud agar to encourage iron-limited growth in a liquid medium before inoculation of a large batch (~50 L) of iron-depleted liquid media. All cultured media (agar and liquid) were allowed to grow at room temperature for the duration of growth time (non-specified).

3.4.2 – Isolation of siderophores

The extraction procedure that was used to preliminarily isolate the iron-binding compounds of interest was adopted from a previously published work by Drs. Wilhelm and Trick (*Limnol. Oceanogr.*, 1994)⁹⁴. Large cultures of microbes grown in large 50 L carboys were centrifuged at 3700 X *g* for about 30 minutes to pelletize a majority of the cells present in the liquid media, which was decanted from the cell pellet and vacuum-filtered through a fine glass microfiber filter to remove the cells remaining in the media. This process allowed for the collection of extracellular siderophores solubilized in the liquid media, which requires further purification to extract the iron-binding compounds with a fair degree of selectivity. All reagents and solvents used to isolate the organic iron-binding compounds from the growth media were treated with Chelex X-100 resin to remove trace amounts of iron that could be present. Isolation of the siderophores began with the acidification of the filtered growth media (pH ~ 2.5) in order to reduce the water-solubility of solubilized organic compounds as well as dissociate any iron-bound siderophores present. The acidified medium was passed through a column of Amberlite XAD-16 resin, which retained the organic compounds and removed all the salts and organic-unbound ions present after flushing the column with several column volumes of Milli-Q H₂O, which was previously adjusted to pH = 3.0 with HCl_(aq). The organic components of the media along with the desired iron-binding compounds were eluted from the Amberlite resin by passing a few column volumes of MeOH through the resin,

collected, and concentrated to yield a concentrated supernatant extract, which was stored at -20 °C until it was processed further.

Siderophores were isolated further from the concentrated supernatant extract by solid-phase extraction utilizing prepacked C18 cartridges (Agilent Bond Elut C18). This was accomplished by taking the previously isolated concentrated methanolic supernatant extract and removing the remaining solvent to near-dryness and taking the sample up in acetonitrile to load the C18 cartridge. The loaded cartridge was eluted with aqueous solutions in a step-wise gradient of increasing methanol concentrations of 10%, 20%, 30%, and 70%. These fractions were dried in a SAVANT Speed Vac Plus SC210A. Dried fractions were taken up in Milli-Q H₂O and subjected to CAS-agar assay to determine the presence of siderophores in the fractions, which were then pooled together upon a positive result for iron-binding species.

This final pooled sample was the crude siderophore extract that was received from the Wilhelm lab and further purified by preparative TLC. Several mobile phases of varying ratios of chloroform, acetic acid, and methanol were assessed to determine a chromatographic system of optimized selectivity (**Figure 3.3.3A**). Other aprotic polar solvents such as THF and EtOAc were also considered, but these trials proved fruitless as alcohols were the only solvents that would move the iron-binding compounds up the TLC plate. The optimal mobile phase to enforce the most separation was determined to be 25 % MeOH in CHCl₃, which was the solvent system employed during preparative TLC of the siderophore samples acquired from the Wilhelm lab. Once the plates were developed in this solvent system, the plates were reversibly stained with iodine vapor, and the desired spots were outlined in pencil. Once the iodine was removed from the plates with gentle heating, the TLC spots of interest were scraped from the plate into their respective 1-mL Eppendorf tubes per sample, and the iron-binding compounds were extracted from the excised spots with 300 µL of MeOH, sonicated briefly, and centrifuged before

transferring the extracted samples into autosampler vials for injection onto a UHPLC-MS instrument.

3.4.3 – UHPLC-MS and UHPLC-MS-MS analysis of preparative TLC samples

Prepared samples were subjected to the following analysis for both the crude siderophore extract and the prep-TLC samples. The instrumentation used was an Exactive Plus Orbitrap mass spectrometer with an electrospray ionization (ESI) source coupled with an Ultimate3000 ultra high-performance liquid chromatography (UHPLC) system (Thermo Scientific, Pittsburgh, PA). For each sample analysis 20 μ L of sample, which were kept at 4 °C for the entire analysis, was injected onto a Kinetex reverse phase C18 (100 x 2.1 mm, 2.6 μ , 100 Å) analytical column maintained at 25 °C and were separated by the following LC method: Mobile phase A was 0.1 % formic acid in H₂O and Mobile phase B was 0.1 % formic acid in MeOH. The flow rate was 0.2 mL/min and the mobile phase gradient used was 0-6 min 2-30 % B, 6.01-12 min 30-70 % B, 12.01-15 min 70-100 % B, 15.01-20 min 100% B, 20.01-25 min 100-50 % B, 25.01-30 min 50-2 % B. The ESI parameters were as follows: scan range 120-1800 m/z, resolution 140,000, AGC target 3.0×10^6 , MaxIT of 200 ms, sheath gas flow rate 25 units, aux gas flow rate 10 units, spray voltage ± 4.0 kV, and capillary temperature 350 °C. The mass spectrometry method was operated in both positive and negative mode using full scan analysis for each sample, once per ionization mode, for a total run-time of 30 minutes per sample, per ionization mode. A post-column in-line PDA detector (Ultimate3000 PDA detector; Thermo Scientific, Pittsburgh, PA) was used prior to mass analysis to scan several wavelengths obtaining some crude structural information. Light absorption was observed at 254 nm, 280 nm, 310 nm, and 380 nm with a bandwidth of 16 nm, a narrow slit-width, and a scan rate of 5 Hz.

For the fragmentation studies, a separate LC method was used to increase the number of scanning events the instrument could use during the method. Since reverse-phase chromatography barely retained the desired compound, the instrument would

have a smaller time-window to isolate, fragment, and analyze the desired mass. HILIC chromatography was better able to retain the compound but did not give a well-resolved EIC for the target compound. Instead, the compound of interest smeared through the column during much of the chromatographic method, which provided plenty of time for the instrument to isolate the desired mass and perform fragmentation analysis. For this analysis the samples were analyzed on a Q-Exactive Orbitrap mass spectrometer coupled to an Ultimate3000 UHPLC system (Thermo Scientific, Pittsburgh, PA). The mass spectrometer was operated in parallel reaction monitoring (PRM) mode for this analysis scanning for masses corresponding to both the iron-bound and unbound siderophore of interest in both a singly charged and doubly charged state. Mobile phase A was 0.1 % formic acid in H₂O and Mobile phase B was 0.1 % formic acid in MeOH. The flow rate was 0.2 mL/min and the mobile phase gradient used was 0-3 min 1-10 % B, 3.01-6 min 10-50 % B, 6.01-9 min 50-100 % B, 9.01-9.5 min 100 % B, 9.51-15 min 100-1 % B. The ESI parameters were as follows for each isolated mass in the PRM method; an isolation window of 0.4 *m/z*, resolution of 140,000, AGC target of 1.0×10^5 , MaxIT of 100 ms, sheath gas flow rate of 25 units, aux gas flow rate of 10 units, spray voltage of ± 4.0 kV, and a capillary temperature of 350 °C. The mass spectrometry method was operated in positive mode using full scan analysis for each isolated mass for each sample, for a total run time of 15 minutes.

3.4.4 – Flash chromatography and subsequent NMR analysis

Normal phase flash chromatography was performed on the crude siderophore extract to isolate the fraction that had the strongest staining behavior when applied to a TLC plate and stained with an ethanolic solution of FeCl₃. To achieve this, the optimal mobile phase from TLC separations (25 % MeOH in CHCl₃) was reduced in polarity to 10 % MeOH in CHCl₃. The column (1" diameter) was packed with a slurry of silica (~50 g) and CHCl₃ and the crude sample (~100 mg) was taken up in a minimal amount of MeOH and loaded onto the packed column, and mobile phase loaded into the reservoir above the

column. The column was not pressurized during separations and was allowed to flow naturally at about 5 mL per minute. Small 5 mL fractions were collected and tested for their ability to bind iron by spotting a TLC plate and subsequently staining the plate with an ethanolic solution of FeCl_3 . Once staining behavior was observed, TLCs were developed using the previously described method and the desired fraction was collected, solvent removed, and dried under high vacuum to give about 20 mg of “purified” compound. This entire sample was taken up in a small volume of deuterated methanol (~ 1 mL CD_3OD) and placed in a reduced-volume NMR tube for NMR analysis. ^1H -NMR, ^{13}C -NMR, and COSY spectra were taken on an Oxford 300 MHz instrument operating at 300.08 MHz for the proton experiments and at 75.46 MHz for the carbon experiment.

3.4.5 – Data analysis

All mass spectrometry data in this chapter was analyzed using the Thermo Scientific Xcalibur software package. This program was used to generate all ion chromatograms, photo-diode array chromatograms, and mass spectra obtained from the siderophore sample analyses. Xcalibur was used to also analyze tandem mass spectrometry data. Proton, carbon, and homonuclear correlation NMR data was analyzed with MestReNova computer software to annotate proton and carbon-13 signal chemical shifts and integrate peak areas.

3.5 – Conclusion

TLC analysis of extracts from three marine organisms – *Vibrio* phage PWH3a-P1, *Rugeria pomeroyi* DSS-3, and *Roseobacter loltanella* sp. SE62 – revealed there were differences in concentration and molecular composition for the iron-binding compounds present in the crude extracts. Optimization of a preparative TLC separation system allowed for the isolation of compounds significantly different in the sp. SE62 sample than the others, with this strain having an abundance of iron-binding activity compared to the other strains. Even within this sample’s multitude of iron-responsive compounds, there

was one single spot that dominated the others in strength of color, presumably by the high concentration of the compound and/or by being highly reactive to the ferric chloride staining solution. LCMS analysis of these crude extracts had shown that between the three marine microbes there were similar mass patterns and mass features, and to a lesser extent PDA-mediated photo-responsiveness, but in the sp. SE62 sample there was a high-intensity mass-feature that had an interesting isotope pattern characteristic of the natural isotopic distribution of iron, and a mass difference exactly matching that of a singly-charged iron-bound species. This finding solidified that the putative siderophore was isolated in an iron-bound and unbound state in the LCMS analysis. Their co-elution may have been a coincidence, or in-source fragmentation had been energetic enough to release the bound iron from the siderophore, giving both masses as detectable mass features. Upon LCMS analysis of the strongly-staining TLC spot by preparative TLC, it was found that the unique mass found in the crude extract analysis was in the strongly staining TLC spot pointing to this mass feature being the compound of interest for this study. It was then apparent that more work needed to be done in isolating this spot from the crude extract.

Both normal-phase and reverse-phase flash chromatography was attempted on the crude sp. SE62 siderophore extract to isolate the compound of interest. Reverse-phase separation provided no retention of the iron-binding species, and TLC analysis post-column revealed that the sample flushed through the column without removing any of the other iron-binding compounds. Normal-phase chromatography proved to be more effective as the TLC separation that worked well was a normal-phase separation. Even though that chromatography worked well, the fractions collected that were related only to the single strongly-staining TLC spot proved to be very complex by LMCS analysis. Regardless, the ^1H -NMR that was obtained to get an idea of the structural features present in the sample was fairly uninformative, which was due to a combination of the complexity of the sample as well as the limited quantity isolated for analysis. It's not to

say that the NMR data was not helpful, but the messy splitting patterns and interconnectivities 2D-NMR experiments gave with these samples are not conclusive for structural elucidation and must be processed further to obtain cleaner spectra. Finally, the last attempt at gathering structural information from this partially purified siderophore sample was to perform a targeted LCMS-MS experiment to gather fragmentation data on the iron-unbound compound. This resulted in a complex spectrum with fragments that do not correspond to typical neutral mass losses, further increasing the complexity of this dataset, and thusly definitive structure determination.

Though these methods of analysis resulted in the realization that more sample would need to be obtained to further purify and gather larger quantities for NMR studies, not all the work was futile. LCMS analysis gave an exact mass of the compound in an iron-bound and unbound state, which could be then used to generate the plausible molecular formulas $C_{29}H_{43}O_{12}N_6Fe$ and $C_{29}H_{46}O_{12}N_6$ respectively. Piecing together the molecular formula into a putative structure gave a cyclic hexapeptide ESEKEA, that tandem-MS fragmentation data did not fully corroborate, other than the few apparent fragments of glutamic acid. Many of the fragments were not of the typical peptide neutral mass losses but piecing together the fragmentation data was unsuccessful and challenging, perhaps due to unusual chemistries this molecule may exhibit as an ion undergoing high-energy collisional dissociation.

Absolute isolation and structural elucidation of the unknown siderophore extracted from the marine bacterium *S. loktanella* sp. SE62 was not achieved, there was enough data gathered to generate a potential molecular formula to determine a putative structure for this unknown compound. Further work in isolating a large (~1 g) batch of crude siderophore extract will be necessary in order to purify large enough quantities of this compound for NMR studies. This task has been undertaken by one of the junior graduate students in Dr. Campagna's research lab, Ashley M. Lato.

CONCLUSION

Synthetic organic chemistry and analytical chemistry, although different in subject matter and philosophy, are very much intertwined and reliant on each other. This is especially true in the study of chemical biology, where in order to observe the effects that specific chemicals have on a biological system with analytical instrumentation, an organic chemist is called upon to synthesize a molecular probe to facilitate this work. This paradigm has certainly not exhausted its usefulness and will be a source of informative research for an unforeseeable amount of time. The work presented in these few chapters exemplifies the informative nature of these disciplines. Pioneering work in synthetic organic chemistry is not only limited to novel molecular transformations, but also includes investigating innovative scaffolds for drug therapies and to learn more about the structure-activity relationship (SAR) between chemicals and their intended biological targets. The knowledge attained from SAR studies leads to better-tailored molecules that stimulate a biological response with high specificity. These SAR studies can be extended to the examination of functional group involvement in spatial, kinetic, and pharmacodynamic conclusions, which ultimately result in medicine unparalleled in efficacy. Whereas within the purview of analytical chemistry, a scientist is able to utilize modern instrumentation to study the molecular makeup of living systems. This is not limited to the analysis of known compounds in metabolomic analyses, but more importantly the unknown space of biogenic chemical compounds. Understanding the chemical transformations available to a specific organism can lead to new ways to manage invasive species or treatment-resistant pathogens.

The first chapter of this document showcases how incredibly useful and versatile synthetic organic chemistry is in the development of optimized drug therapies. Through the synthesis and *in vitro* evaluation of previously reported *N*-arylpyrazolyl-based glucocorticoid analogues, it has been demonstrated that these molecules have great potential to replace some modern-day GC-based therapies. These compounds have been

shown to be quite effective in managing inflammation beforehand, but with the addition of a few novel compounds and modifications not attempted in earlier work, the literature will reflect a larger body of work that validates these findings. As already stated, these compounds have potent anti-inflammatory action, but more importantly they are also limiting the production of pro-inflammatory cytokines unlike currently available GC-based therapies. This permits a medical practitioner to administer these compounds without the unintended stimulation of the immune system to attack the cells that respond to these molecules. The application of these GCs in pancreatic organ transplants could revolutionize their success rate due to their lack of impact on β -cell function. Though these molecules have been assayed in an earlier publication, there is a lack of molecular modeling to describe the SARs, which engender the desired biological response. Computational evaluation of these compounds revealed that small aliphatic substitutions in conjunction with a 4-fluorophenyl group appended on the pyrazole portion of the Merck scaffold reduced the incidence of cytokine transcription. It was also shown by DFT modeling calculations that this likely had been mediated by the binding orientation of the GC in the GR pocket, resulting in a GC-GR complex with an appropriate shape to bind specifically to nuclear transcription factors that elicited the desired anti-inflammatory response. Higher-order computations will be needed to understand these specific GC-GR complex shapes and how they bind with nuclear GREs and other transcription factors to not only determine the mechanism of provoking the desired therapeutic response, but to also understand what transcription factors, previously thought to be undruggable targets respond best to these therapies.

In a separate study of how chemicals influence biology, an investigation of the presence of some unknown semiochemical in the soil bacterium *M. xanthus* led to the development of a stable isotope tracer for the identification of a specific structural motif in routine LC-ESI-MS analyses. The application of synthetic chemistry in biology for the purpose of understanding how chemicals encourage a phenotype or behavioral response

largely involves the incorporation of a stable isotope-labeled congener to the system being investigated. Analytical chemistry is fraught with challenges in applying certain instrumentation to the analysis of analytes poorly compatible with accessible instruments. The analysis of isobaric compounds without the availability of a hard ionization source in mass spectrometry is one such case where isotope-labeled compounds mitigate this issue. Through the work presented in the second chapter, a synthetic isotope tracer (iC5- d_3) was used to highlight compounds that contain a specific structural feature. Absorption of a labeled feedstock that is a biosynthetic precursor to lipid species bearing an isopropyl moiety allows for the generation of a diverse pool of isotopically labeled compounds. Using this strategy, it is possible to use LC-ESI-MS instrumentation traditionally used in lipidomic investigations to qualitatively determine the presence of this moiety in an untargeted fashion. The value of this methodology was demonstrated by confirming the presence of previously reported compounds in *M. xanthus* documented to bear a terminal isopropyl unit in homospermidine lipids. The degree of labeling detected established that the isotope tracer would only be incorporated into compounds with this specific moiety; However, it also seems that the tracer was incorporated into everything that could possibly bear this structural pattern. Though this can illuminate the presence of this single structural feature within an unknown mass feature, determining a structure from a mass is a daunting task. This method could potentially make the structural elucidation of unknown mass features in these experiments less intimidating.

The final chapter of this dissertation is an account of the processes attempted to isolate and elucidate the structure for an unknown biological molecule with interesting metal-transporting and cytotoxic properties. This compound was extracted from a marine bacterium belonging to the large *Roseobacter* clade and may have important implications for iron cycling in marine environments, not only for bacterial use, but ultimately for all marine organisms. Through TLC analysis of a crude siderophore-selective extract from *S.*

loktanella sp. SE62 (SE62), it was apparent that there were several iron-binding compounds present in the extract. Comparing the TLC plates from other siderophore-producing microbes revealed that SE62 was producing a compound that stained strongly upon exposure to a solution of ferric chloride with a unique retention factor to the compounds produced by the other marine microbes. Extraction of this compound by preparative TLC allowed for mass analysis of the compound(s) present in the excised spot, which exhibited a mass pattern indicative of the excised compound in a bound and unbound state with ferric iron. Column chromatography of the crude siderophore extract was successful in isolating the single TLC spot with strong staining behavior but proved to be of insufficient quantity to obtain a meaningful NMR spectrum. LCMS analysis of the isolated compound(s) had shown that the flash column did not isolate a single compound but that several compounds were present in the purified sample. It was not feasible to try and purify the sample any further for NMR studies, as the final isolated sample was less than 20 mg. Work is currently being done to collect a larger sample size to be subjected to several columns of differing selectivity to absolutely isolate the compounds present in the sample with a high affinity for ferric iron.

Chemistry is a scientific discipline that will always be changing the world we live in and challenging the status-quo, from the medicines we take to the understanding of how living systems function. Even when solely focusing on chemistry, there is no conceivable end to the improvements that can be made to the specificity of chemical reactions or to the sensitivity of analytical instrumentation. It is a large responsibility of the scientific community to make logical advancements with current technology and techniques in order to influence future generations of scientists and to continue expanding human knowledge *ad infinitum*. The findings presented in this dissertation will hopefully inspire new ideas to further investigate the understanding of the few biological systems explored in this work or be able to be applied in other areas of research.

REFERENCES

1. Mellman, I.; Coukos, G.; Dranoff, G., Cancer immunotherapy comes of age. *Nature* **2011**, *480*, 480.
2. Lesterhuis, W. J.; Haanen, J. B. A. G.; Punt, C. J. A., Cancer immunotherapy – revisited. *Nature Reviews Drug Discovery* **2011**, *10*, 591.
3. Uzu, T.; Harada, T.; Sakaguchi, M.; Kanasaki, M.; Isshiki, K.; Araki, S.; Sugimoto, T.; Koya, D.; Haneda, M.; Kashiwagi, A.; Yamauchi, A., Glucocorticoid-Induced Diabetes Mellitus: Prevalence and Risk Factors in Primary Renal Diseases. *Nephron Clinical Practice* **2007**, *105* (2), c54-c57.
4. André, I.; Gonzalez, A.; Wang, B.; Katz, J.; Benoist, C.; Mathis, D., Checkpoints in the progression of autoimmune disease: lessons from diabetes models. *Proceedings of the National Academy of Sciences* **1996**, *93* (6), 2260.
5. Kadmiel, M.; Cidlowski, J. A., Glucocorticoid receptor signaling in health and disease. *Trends in Pharmacological Sciences* **2013**, *34* (9), 518-530.
6. Clark, M.; Kroger, C. J.; Tisch, R. M., Type 1 Diabetes: A Chronic Anti-Self-Inflammatory Response. *Frontiers in Immunology* **2017**, *8*, 1898.
7. Collier, J. J.; Burke, S. J.; Eisenhauer, M. E.; Lu, D.; Sapp, R. C.; Frydman, C. J.; Campagna, S. R., Pancreatic β -Cell Death in Response to Pro-Inflammatory Cytokines Is Distinct from Genuine Apoptosis. *PLOS ONE* **2011**, *6* (7), e22485.
8. Butler, A., Marine Siderophores and Microbial Iron Mobilization. *Biometals* **2005**, *18* (4), 369-374.
9. Ali, A.; Thompson, C. F.; Balkovec, J. M.; Graham, D. W.; Hammond, M. L.; Quraishi, N.; Tata, J. R.; Einstein, M.; Ge, L.; Harris, G., Novel N-arylpyrazolo [3, 2-c]-based ligands for the glucocorticoid receptor: receptor binding and in vivo activity. *Journal of medicinal chemistry* **2004**, *47* (10), 2441-2452.
10. Barnes, P. J., Anti-inflammatory Actions of Glucocorticoids: Molecular Mechanisms. *Clinical Science* **1998**, *94* (6), 557.

11. Wood, J. N., From plant extract to molecular panacea: a commentary on Stone (1763) 'An account of the success of the bark of the willow in the cure of the agues'. *Philosophical Transactions of the Royal Society B: Biological Sciences* **2015**, 370 (1666), 20140317.
12. Sneader, W., The discovery of aspirin: a reappraisal. *BMJ : British Medical Journal* **2000**, 321 (7276), 1591-1594.
13. Benedek, T. G., History of the development of corticosteroid therapy. *Clinical and Experimental Rheumatology* **2011**, 29 (5 (Suppl. 68)), 7.
14. Hellal-Levy, C.; Couette, B.; Fagart, J.; Souque, A.; Gomez-Sanchez, C.; Rafestin-Oblin, M.-E., Specific hydroxylations determine selective corticosteroid recognition by human glucocorticoid and mineralocorticoid receptors. *FEBS Letters* **1999**, 464 (1-2), 9-13.
15. Ferreira, S. H.; Moncada, S.; Vane, J. R., Indomethacin and Aspirin abolish Prostaglandin Release from the Spleen. *Nature New Biology* **1971**, 231 (25), 237-239.
16. Ballinger, A.; Smith, G., COX-2 inhibitors vs. NSAIDs in gastrointestinal damage and prevention. *Expert Opinion on Pharmacotherapy* **2001**, 2 (1), 31-40.
17. De Bosscher, K.; Beck, I. M.; Haegeman, G., Classic glucocorticoids versus non-steroidal glucocorticoid receptor modulators: Survival of the fittest regulator of the immune system? *Brain, Behavior, and Immunity* **2010**, 24 (7), 1035-1042.
18. Beck, I. M. E.; Vanden Berghe, W.; Vermeulen, L.; Yamamoto, K. R.; Haegeman, G.; De Bosscher, K., Crosstalk in Inflammation: The Interplay of Glucocorticoid Receptor-Based Mechanisms and Kinases and Phosphatases. *Endocrine Reviews* **2009**, 30 (7), 830-882.
19. Atlanta, G. C. f. D. C. a. P., US Department of Health and Human Services, National Diabetes Statistics Report, 2017. *Centers for Disease Control and Prevention* **2017**.

20. Bougnères, P.; Valleron, A.-J., Causes of early-onset type 1 diabetes: toward data-driven environmental approaches. *The Journal of Experimental Medicine* **2008**, 205 (13), 2953.
21. Rother, K. I., Diabetes Treatment — Bridging the Divide. *New England Journal of Medicine* **2007**, 356 (15), 1499-1501.
22. Frigerio, S.; Junt, T.; Lu, B.; Gerard, C.; Zumsteg, U.; Holländer, G. A.; Piali, L., β cells are responsible for CXCR3-mediated T-cell infiltration in insulitis. *Nature Medicine* **2002**, 8, 1414.
23. Mathis, D.; Vence, L.; Benoist, C., β -Cell death during progression to diabetes. *Nature* **2001**, 414, 792.
24. Collier, J. J.; Fueger, P. T.; Hohmeier, H. E.; Newgard, C. B., Pro- and Antiapoptotic Proteins Regulate Apoptosis but Do Not Protect Against Cytokine-Mediated Cytotoxicity in Rat Islets and β -Cell Lines. *Diabetes* **2006**, 55 (5), 1398.
25. Leonard, B.; Ralph, P. R., The Glucocorticoid Receptor: Molecular Mechanism and New Therapeutic Opportunities. *Current Drug Targets - Inflammation & Allergy* **2002**, 1 (2), 127-136.
26. Schulz, M.; Eggert, M., Novel Ligands: Fine Tuning the Transcriptional Activity of the Glucocorticoid Receptor. *Current Pharmaceutical Design* **2004**, 10 (23), 2817-2826.
27. Michael, J. C.; Steven, W. E.; Philip, R. K.; Michael, E. K., The Pursuit of Differentiated Ligands for the Glucocorticoid Receptor. *Current Topics in Medicinal Chemistry* **2003**, 3 (14), 1617-1635.
28. Rogatsky, I.; Luecke, H. F.; Leitman, D. C.; Yamamoto, K. R., Alternate surfaces of transcriptional coregulator GRIP1 function in different glucocorticoid receptor activation and repression contexts. *Proceedings of the National Academy of Sciences* **2002**, 99 (26), 16701.

29. Meijsing, S. H.; Pufall, M. A.; So, A. Y.; Bates, D. L.; Chen, L.; Yamamoto, K. R., DNA Binding Site Sequence Directs Glucocorticoid Receptor Structure and Activity. *Science* **2009**, 324 (5925), 407.
30. Rosenfeld, M. G.; Lunyak, V. V.; Glass, C. K., Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes & Development* **2006**, 20 (11), 1405-1428.
31. Burke, S. J.; May, A. L.; Noland, R. C.; Lu, D.; Brissova, M.; Powers, A. C.; Sherrill, E. M.; Karlstad, M. D.; Campagna, S. R.; Stephens, J. M.; Collier, J. J., Thiobenzothiazole-modified Hydrocortisones Display Anti-inflammatory Activity with Reduced Impact on Islet β -cell Function. *Journal of Biological Chemistry* **2015**.
32. Corbett, J. A.; Wang, J. L.; Misko, T. P.; Zhao, W.; Hickey, W. F.; McDaniel, M. L., Nitric Oxide Mediates IL-1 β -Induced Islet Dysfunction and Destruction: Prevention by Dexamethasone. *Autoimmunity* **1993**, 15 (2), 145-153.
33. Xuan-Yu, M.; Hong-Xing, Z.; Mihaly, M.; Meng, C., Molecular Docking: A Powerful Approach for Structure-Based Drug Discovery. *Current Computer-Aided Drug Design* **2011**, 7 (2), 146-157.
34. Freddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K., Ten-Microsecond Molecular Dynamics Simulation of a Fast-Folding WW Domain. *Biophysical Journal* **2008**, 94 (10), L75-L77.
35. Hemmerling, M., Nilsson, S., Edman, K., Eirefelt, S., Russell, W., Hendrickx, R., Johnsson, E., Karrman Mardh, C., Berger, M., Rehwinkel, H., Abrahamsson, A., Dahmen, J., Eriksson, A.R., Gabos, B., Henriksson, K., Hossain, N., Ivanova, S., Jansson, A.H., Jensen, T.J., Jerre, A., Johansson, H., Klingstedt, T., Lepisto, M., Lindsjo, M., Mile, I., Nikitidis, G., Steele, J., Tehler, U., Wissler, L., Hansson, T., Selective Nonsteroidal Glucocorticoid Receptor Modulators for the Inhaled Treatment of Pulmolnary Diseases. *J. Med. Chem.*, 2017; Vol. 60, pp 8591-8605.

36. O. Trott, A. J. O., 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.
37. Schrödinger, L., The PyMOL Molecular Graphics System, Version ~1.8. **2015**.
38. Laskowski R A, S. M. B., LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling* **2011**, *51* (10), 2778-2786.
39. TURBOMOLE V7.2 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007.
40. Becke, A. D., Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A* **1988**, *38* (6), 3098-3100.
41. Perdew, J. P., Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Physical Review B* **1986**, *33* (12), 8822-8824.
42. Weigend, F., Accurate Coulomb-fitting basis sets for H to Rn. *Physical Chemistry Chemical Physics* **2006**, *8* (9), 1057-1065.
43. Weigend, F.; Ahlrichs, R., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Physical Chemistry Chemical Physics* **2005**, *7* (18), 3297-3305.
44. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics* **2010**, *132* (15), 154104.
45. Sierka, M.; Hogekamp, A.; Ahlrichs, R., Fast evaluation of the Coulomb potential for electron densities using multipole accelerated resolution of identity approximation. *The Journal of Chemical Physics* **2003**, *118* (20), 9136-9148.
46. Vosko, S. H.; Wilk, L.; Nusair, M., Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics* **1980**, *58* (8), 1200-1211.

47. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B* **1988**, 37 (2), 785-789.
48. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993**, 98 (7), 5648-5652.
49. Hohmeier, H. E.; Mulder, H.; Chen, G.; Henkel-Rieger, R.; Prentki, M.; Newgard, C. B., Isolation of INS-1-derived cell lines with robust ATP-sensitive K⁺ channel-dependent and -independent glucose-stimulated insulin secretion. *Diabetes* **2000**, 49 (3), 424.
50. Shah, P.; Westwell, A. D., The role of fluorine in medicinal chemistry. *Journal of Enzyme Inhibition and Medicinal Chemistry* **2007**, 22 (5), 527-540.
51. Smart, B. E., Fluorine substituent effects (on bioactivity). *Journal of Fluorine Chemistry* **2001**, 109 (1), 3-11.
52. Ran-Ressler, R. R.; Lawrence, P.; Brenna, J. T., Structural characterization of saturated branched chain fatty acid methyl esters by collisional dissociation of molecular ions generated by electron ionization. *Journal of Lipid Research* **2012**, 53 (1), 195-203.
53. Harvey, D. J., Picolinyl esters as derivatives for the structural determination of long chain branched and unsaturated fatty acids. *Biomedical Mass Spectrometry* **1982**, 9 (1), 33-38.
54. Yu, Q. T.; Liu, B. N.; Zhang, J. Y.; Huang, Z. H., Location of methyl branchings in fatty acids: Fatty acids in uropygial secretion of Shanghai Duck by GC-MS of 4,4-dimethyloxazoline derivatives. *Lipids* **1988**, 23 (8), 804-810.
55. Bhat, S.; Ahrendt, T.; Dauth, C.; Bode, H. B.; Shimkets, L. J., Two Lipid Signals Guide Fruiting Body Development of *Myxococcus xanthus*. *mBio* **2014**, 5 (1).
56. Hu, C.; van der Heijden, R.; Wang, M.; van der Greef, J.; Hankemeier, T.; Xu, G., Analytical strategies in lipidomics and applications in disease biomarker discovery. *Journal of Chromatography B* **2009**, 877 (26), 2836-2846.

57. Chan, J.; Fujiwara, T.; Brennan, P.; McNeil, M.; Turco, S. J.; Sibille, J. C.; Snapper, M.; Aisen, P.; Bloom, B. R., Microbial glycolipids: possible virulence factors that scavenge oxygen radicals. *Proceedings of the National Academy of Sciences* **1989**, *86* (7), 2453.
58. Afshinnia, F.; Rajendiran, T. M.; Wernisch, S.; Soni, T.; Jadoon, A.; Karnovsky, A.; Michailidis, G.; Pennathur, S., Lipidomics and Biomarker Discovery in Kidney Disease. *Seminars in Nephrology* **2018**, *38* (2), 127-141.
59. Fernandis, A. Z.; Wenk, M. R., Membrane lipids as signaling molecules. *Current Opinion in Lipidology* **2007**, *18* (2).
60. Divecha, N.; Irvine, R. F., Phospholipid signaling. *Cell* **1995**, *80* (2), 269-278.
61. González-Domínguez, R.; García-Barrera, T.; Gómez-Ariza, J. L., Metabolomic study of lipids in serum for biomarker discovery in Alzheimer's disease using direct infusion mass spectrometry. *Journal of Pharmaceutical and Biomedical Analysis* **2014**, *98*, 321-326.
62. Fahy, E.; Cotter, D.; Sud, M.; Subramaniam, S., Lipid classification, structures and tools. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **2011**, *1811* (11), 637-647.
63. Spector, A. A.; Yorek, M. A., Membrane lipid composition and cellular function. *Journal of Lipid Research* **1985**, *26* (9), 1015-1035.
64. Hoffmann, M.; Auerbach, D.; Panter, F.; Hoffmann, T.; Dorrestein, P. C.; Müller, R., Homospermidine Lipids: A Compound Class Specifically Formed during Fruiting Body Formation of *Myxococcus xanthus* DK1622. *ACS Chemical Biology* **2018**, *13* (1), 273-280.
65. Berna, M. J.; Ackermann, B. L., Quantification of serine enantiomers in rat brain microdialysate using Marfey's reagent and LC/MS/MS. *Journal of Chromatography B* **2007**, *846* (1), 359-363.
66. Bhushan, R.; Brückner, H., Marfey's reagent for chiral amino acid analysis: A review. *Amino Acids* **2004**, *27* (3), 231-247.

67. Zhang, Y.; Go, E. P.; Jiang, H.; Desaire, H., A Novel Mass Spectrometric Method to Distinguish Isobaric Monosaccharides that are Phosphorylated or Sulfated Using Ion-Pairing Reagents. *Journal of the American Society for Mass Spectrometry* **2005**, *16* (11), 1827-1839.
68. Ara, K.; Hama, M.; Akiba, S.; Koike, K.; Okisaka, K.; Hagura, T.; Kamiya, T.; Tomita, F., Foot odor due to microbial metabolism and its control. *Canadian Journal of Microbiology* **2006**, *52* (4), 357-364.
69. Kaneda, T., Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. *Microbiological reviews* **1991**, *55* (2), 288-302.
70. Lehmann, W. D., A timeline of stable isotopes and mass spectrometry in the life sciences. *Mass Spectrometry Reviews* **2016**, *36* (1), 58-85.
71. Crawford, E. W.; Shimkets, L. J., The stringent response in *Myxococcus xanthus* is regulated by SocE and the CsgA C-signaling protein. *Genes & Development* **2000**, *14* (4), 483-492.
72. Crawford Jr, E. W.; Shimkets, L. J., The *Myxococcus xanthus* socE and csgA genes are regulated by the stringent response. *Molecular Microbiology* **2000**, *37* (4), 788-799.
73. Boynton, T. O. H.; Shimkets, L. J., *Myxococcus* CsgA, *Drosophila* Sniffer, and human HSD17B10 are cardiolipin phospholipases. *Genes & Development* **2015**.
74. Shimkets, L. J.; Rafiee, H., CsgA, an extracellular protein essential for *Myxococcus xanthus* development. *Journal of Bacteriology* **1990**, *172* (9), 5299.
75. Bode, H. B.; Ring, M. W.; Schwär, G.; Kroppenstedt, R. M.; Kaiser, D.; Müller, R., 3-Hydroxy-3-Methylglutaryl-Coenzyme A (CoA) Synthase Is Involved in Biosynthesis of Isovaleryl-CoA in the Myxobacterium *Myxococcus xanthus*; during Fruiting Body Formation. *Journal of Bacteriology* **2006**, *188* (18), 6524.
76. Hagen, D. C.; Bretscher, A. P.; Kaiser, D., Synergism between morphogenetic mutants of *Myxococcus xanthus*. *Developmental Biology* **1978**, *64* (2), 284-296.

77. McVittie, A.; Messik, F.; Zahler, S. A., Developmental Biology of *Myxococcus*. *Journal of Bacteriology* **1962**, *84* (3), 546.
78. Toal, D. R.; Clifton, S. W.; Roe, B. A.; Downard, J., The *esg* locus of *Myxococcus xanthus* encodes the E1 α and E1 β subunits of a branched-chain keto acid dehydrogenase. *Molecular Microbiology* **1995**, *16* (2), 177-189.
79. Shimkets, L. J., Intercellular Signaling During Fruiting-Body Development of *Myxococcus xanthus*. *Annual Review of Microbiology* **1999**, *53* (1), 525-549.
80. Lorenzen, W.; Bozhüyük, K. A. J.; Cortina, N. S.; Bode, H. B., A comprehensive insight into the lipid composition of *Myxococcus xanthus* by UPLC-ESI-MS. *Journal of lipid research* **2014**, *55* (12), 2620-2633.
81. Bode, H. B.; Ring, M. W.; Kaiser, D.; David, A. C.; Kroppenstedt, R. M.; Schwär, G., Straight-Chain Fatty Acids Are Dispensable in the Myxobacterium *Myxococcus xanthus* for Vegetative Growth and Fruiting Body Formation. *Journal of Bacteriology* **2006**, *188* (15), 5632.
82. M.B. Richardson, S. J. W., A practical synthesis of long-chain iso-fatty acids (iso-C12–C19) and related natural products. *Belstein Journal of Organic Chemsitry* **2013**, *9*, 5.
83. Fujisawa, T.; Sato, T.; Kawara, T.; Kawashima, M.; Shimizu, H.; Ito, Y., A three carbon homologation by the ring-opening of β -propiolactones with diorganocuprates. *Tetrahedron Letters* **1980**, *21* (22), 2181-2184.
84. Kawashima, M.; Sato, T.; Fujisawa, T., A facile method for synthesis of three carbon-homologated carboxylic acid by regioselective ring-opening of β -propiolactones with organocopper reagents. *Tetrahedron* **1989**, *45* (2), 403-412.
85. Hamilton, J. G.; Comai, K., Rapid separation of neutral lipids, free fatty acids and polar lipids using prepacked silica sep-Pak columns. *Lipids* **1988**, *23* (12), 1146-1149.
86. Vraspir, J. M.; Butler, A., Chemistry of Marine Ligands and Siderophores. *Annual Review of Marine Science* **2009**, *1* (1), 43-63.

87. Butler, A.; Theisen, R. M., Iron(III)–siderophore coordination chemistry: Reactivity of marine siderophores. *Coordination Chemistry Reviews* **2010**, *254* (3), 288-296.
88. Expert, D., Iron, an Element Essential to Life. In *Molecular Aspects of Iron Metabolism in Pathogenic and Symbiotic Plant-Microbe Associations*, Expert, D.; O'Brian, M. R., Eds. Springer Netherlands: Dordrecht, 2012; pp 1-6.
89. Harrison, P. M.; Arosio, P., The ferritins: molecular properties, iron storage function and cellular regulation. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1996**, *1275* (3), 161-203.
90. Robinson, A. K.; Rodríguez-Quinones, F.; Andrews, S. C., Bacterial iron homeostasis. *FEMS Microbiology Reviews* **2003**, *27* (2-3), 215-237.
91. Zhang, G.; Amin, S. A.; Küpper, F. C.; Holt, P. D.; Carrano, C. J.; Butler, A., Ferric Stability Constants of Representative Marine Siderophores: Marinobactins, Aquachelins, and Petrobactin. *Inorganic Chemistry* **2009**, *48* (23), 11466-11473.
92. Neilands, J. B., Siderophores: Structure and Function of Microbial Iron Transport Compounds. *Journal of Biological Chemistry* **1995**, *270* (45), 26723-26726.
93. Miethke, M.; Marahiel, M. A., Siderophore-Based Iron Acquisition and Pathogen Control. *Microbiology and Molecular Biology Reviews* **2007**, *71* (3), 413.
94. Wilhelm, S. W.; Trick, C. G., Iron-limited growth of cyanobacteria: Multiple siderophore production is a common response. *Limnology and Oceanography* **1994**, *39* (8), 1979-1984.
95. Challis, G. L.; Hopwood, D. A., Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proceedings of the National Academy of Sciences* **2003**, *100* (suppl 2), 14555.
96. Buchan, A.; Neidle, E. L.; Moran, M. A., Diversity of the Ring-Cleaving Dioxygenase Gene *pcaH* in a Salt Marsh Bacterial Community. *Applied and Environmental Microbiology* **2001**, *67* (12), 5801-5809.

97. Luo, H.; Moran, M. A., Evolutionary Ecology of the Marine *Roseobacter* Clade. *Microbiology and Molecular Biology Reviews* **2014**, 78 (4), 573.
98. Martens, T.; Gram, L.; Grossart, H.-P.; Kessler, D.; Müller, R.; Simon, M.; Wenzel, S. C.; Brinkhoff, T., Bacteria of the *Roseobacter* Clade Show Potential for Secondary Metabolite Production. *Microbial Ecology* **2007**, 54 (1), 31-42.
99. Loudon, B. C.; Haarmann, D.; Lynne, A. M., Use of Blue Agar CAS Assay for Siderophore Detection. *Journal of microbiology & biology education* **2011**, 12 (1), 51-53.
100. Sticher, O., Natural product isolation. *Natural Product Reports* **2008**, 25 (3), 517-554.
101. Bucar, F.; Wube, A.; Schmid, M., Natural product isolation – how to get from biological material to pure compounds. *Natural Product Reports* **2013**, 30 (4), 525-545.
102. Jennings, D. W.; Deutsch, H. M.; Zalkow, L. H.; Teja, A. S., Supercritical extraction of taxol from the bark of *Taxus brevifolia*. *The Journal of Supercritical Fluids* **1992**, 5 (1), 1-6.
103. Zurer, P.; Baum, R., Chemists Closing In on Synthesis of Scarce Anticancer Agent Taxol. *Chemical & Engineering News Archive* **1988**, 66 (41), 22-24.
104. Kind, T.; Fiehn, O., Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics* **2007**, 8 (1), 105.
105. Bruland, K. W.; Franks, R. P.; Knauer, G. A.; Martin, J. H., Sampling and analytical methods for the determination of copper, cadmium, zinc, and nickel at the nanogram per liter level in sea water. *Analytica Chimica Acta* **1979**, 105, 233-245.
106. Price, N. M.; Harrison, C. I.; Hering, J. C.; Hudson, R. J.; Nirel, P. V.; Palenik, B.; Morel, F. M. M., Preparation and chemistry of the artificial algal culture medium Aquil. *Biological Oceanography* **1989**, 443-461.
107. Charles, R. B.; Hollibaugh, J. T., Ecophysiology of a Mono Lake Picocyanobacterium. *Limnology and Oceanography* **2007**, 52 (6), 2484-2495.

VITA

Brandon J. Kennedy was born on December 29th, 1989 and grew up in the once-bustling town of Oil City, Pennsylvania where his friends, family, teachers, and townspeople instilled a curiosity for knowledge and experimentation. His love for science and the arts were cultured through the years by way of inspirational mentors, and a natural aptitude for absorbing information and testing the boundaries of his own talents.

Upon graduation from Oil City Senior High School in 2008, he attended Westminster College in New Wilmington, Pennsylvania - majoring in chemistry with a minor in music. During his time at Westminster, he explored his passion for fragrance by acquiring an internship as a fragrance chemist at the creative office of Givaudan in Manhattan, NY. From this experience came the desire to pursue a career in the fragrance industry developing novel odor chemicals to expand not only his experience of fragrance, but to also influence the creative potential for perfumers world-wide.

After graduating from Westminster College in 2012, he went on to study both analytical and synthetic organic chemistry in the research group of Dr. Shawn R. Campagna at the University of Tennessee Knoxville in Knoxville, Tennessee. He worked on several research projects utilizing his abilities as an organic and analytical chemist to investigate various biological systems.