2016

Probing Allosteric, Partial Inhibition of Thrombin Using Novel Anticoagulants

Stephen S. Verespy III
Virginia Commonwealth University

PROBING ALLOSTERIC, PARTIAL INHIBITION OF THROMBIN USING NOVEL ANTICOAGULANTS

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

STEPHEN S. VERESPY III
Bachelor of Science, Keystone College, 2010

Director: UMESH R. DESAI
DOCTOR OF PHILOSOPHY, DEPARTMENT OF MEDICINAL CHEMISTRY

Virginia Commonwealth University
Richmond, Virginia
August 2016

Acknowledgement

There are a number of people I would like to thank in aiding me with this achievement. First my advisor, Dr. Desai, has been pivotal in not only empirical direction for the work, but also in shaping me as a young scientist. He has been able to instill into me the essential questions to ask in research and how to answer them properly to the fullest extent. He has also been very supportive in my personal goals and of my past and future endeavors; to which I am most grateful.

To my wife, Kristen, who has stuck with me through thick and thin; from the beginning of my graduate work 6 years ago, to our soon-to-be new life in San Diego. She has always given me the motivation to continue and follow through with what I am passionate about. Kristen is vital to my success because she can show reason and logic when I lose sight of the bigger picture, as well as keep me from overthinking some minor details in work and life. She continues to be a fierce companion and my greatest friend, thank you and I love you, from the bottom of my heart.

To my committee members, Drs. Cropp, Hartman, and Safo, I truly appreciate your always kind and understanding demeanors. From doing rotations in your labs during my first year, to allowing me to use equipment in the last months of time here. You all have been crucial to my success at VCU. I also owe my colleagues and collaborators,

Drs. Akul Mehta, Nehru Viji Sankaranarayanan, Bassem Mohammad, Adam Hawkridge, and Donald Brophy a great deal of thanks for all our talks, equipment use, and time in aiding in the work performed herein.

Last, but not least, the list of my family and friends who have supported me throughout my academic career is exhaustive and cannot be mentioned here. I just want to extend a thank you to all of you.

Table of Contents

List of Tables

Page

List of Figures

Page

Abstract

PROBING ALLOSTERIC, PARTIAL INHIBITION OF THROMBIN USING NOVEL ANTICOAGULANTS

By Stephen S. Verespy III, B. S.

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2016

Major Director: Umesh R. Desai
Doctor of Philosophy, Department of Medicinal Chemistry

Thrombin is the key protease that regulates hemostasis; the delicate balance between procoagulation and anticoagulation of blood. In clotting disorders, like deep vein thrombosis or pulmonary embolism, procoagulation is up-regulated, but propagation of clotting can be inhibited with drugs targeting the proteases involved, like thrombin. Such drugs however, have serious side effects (e.g., excessive bleeding) and some require monitoring during the course of treatment. The reason for these side effects is the mechanism by which the drugs' act. The two major mechanisms are direct orthosteric and indirect allosteric inhibition, which will completely abolish the protease's activity. Herein

we sought an alternative mechanism called allosteric, partial inhibition, that has shown promise to truly regulate coagulation. Partial inhibition through allosteric mechanisms are well described for membrane-bound and oligomeric proteins. However proteases, specifically monomeric proteases (i.e., thrombin), have not shown this phenomenon until now. A small library of coumarin-based sulfated allosteric modulators (CSAMs) was synthesized to target a surface region called exosite 2; mainly composed of highly positively charged residues surrounded by hydrophobic patches. Studies revealed a non-competitive mechanism of binding with a range of IC_{50}s between 0.2-58 μM combined with inhibitory efficacies (ΔY) between 22-73%; indicative of allosteric, partial inhibition. The K_D was determined for the most potent compound (**3g**; IC_{50} = 0.2 μM, ΔY = 47%) at 0.15 μM. **3g** was observed to bind at exosite 2 through unfractionated heparin competition and thrombin mutant studies. Additional computational studies were in agreement with the mutant and competition studies, showing the sulfate of **3g** binding within a pocket containing R126 and R233. Fluorescence quenching and antithrombin inactivation rate studies described a conformational change to thrombin's active site in the presence of **3g**, supporting reduction of thrombin's catalytic efficiency, without complete inhibition of thrombin's proteolytic activities. Coupled enzyme assays and gel electrophoresis showed that in the presence of **3g**, hydrolysis of fibrinogen (IC_{50} = 0.51 μM, ΔY = 94%) and protein C activation (IC_{50} = 1.7 μM, ΔY = 91%) is fully inhibited. Alternatively, FXIII activation was shown to be only partially inhibited by the presence of **3g**, and FXI activation did not show any significant activation or inhibition. **3g** was also shown to be active in human plasma and whole blood, but requiring much higher concentrations to

induce an anticoagulant effect. Mice studies looking at the effects of **3g** *in vivo* showed that even at high concentrations, showed no abnormal bleeding or any other irregularities. This work highlights a novel occurrence regarding thrombin's allosteric functionality against multiple endogenous substrates. This library of compounds may be useful in the future development of allosteric inhibitors and probes that pose little to no risk of bleeding events by inducing partial inhibition.

Key Words: Partial inhibition, thrombin, glycosaminoglycan mimetics, protease, exosite, allosteric modulator, anticoagulants, catalytic efficiency, conformational allostery.

CHAPTER 1 Introduction

1.1 Proteases

Biological systems rely heavily on enzymatic processes to function through a number of catalytic pathways[1]. Of those, hydrolysis of peptide bonds through use of proteases (also known as proteinases and peptidases) play major roles in homeostatic regulation (e.g., immune response, blood coagulation, apoptosis, etc.)[1,2], as well as in disease (e.g., cancer, viral/bacterial infection, cardiovascular disease, etc.)[3–6]. The translation of proteases must be well regulated due to their highly efficient nature of catalysis. If up or down regulation of specific proteases is not maintained, several degrees of dysfunction could result from degradation of non-ideal substrates or buildup of substrates that could be potentially harmful. In the human genome alone 2-4% of the DNA is made up of genes encoding for proteases; with an estimated number of 550 protease genes known to date[1]. Proteases can be further categorized by traits relating to their mechanism of action; and are done so in the following order of prevalence in humans. Of the > 550 proteases, 34% make up the metalloproteases, 31% serine proteases, 26% cysteine proteases, 5% threonine proteases, and 4% aspartic proteases[1,2]. These distinctions are crucial in understanding the mechanism behind each individual class of proteases, so

selective treatments may be developed to augment their effects if over produced in a system.

1.1.1 Types and Mechanisms

Metalloproteases are named so because of their one or more cationic transition metals located within their active (orthosteric) site. The most prevalent metal found in metalloproteases (MPs) is zinc (Zn^{2+})[7,8], but in some circumstances cobalt (Co^{2+}) and manganese (Mn^{2+}) have also been found[7,8]. The transition metal within the active site is typically held in place by coordination with two histidine residues and a glutamate residue (three histidines may also be present)[9,10]. With regards to MP's catalytic mechanism, there are currently two theories on how the active site facilitates hydrolysis. The first is that upon hydrolysis of a peptidyl bond, a water molecule coordinated with the Zn^{2+} is activated by a secondary glutamate. In turn the Zn^{2+} also coordinates with the carbonyl oxygen from the substrate and facilitates the hydrolysis of the substrate by the nucleophilic water molecule[9–11]. The other theory suggests that the secondary glutamate acts as the nucleophile instead of a water molecule, and forms an acyl intermediate with the substrate carbonyl and the zinc. The theory suggests that a water molecule then behaves as a nucleophile, but attacks the carboxylate of the secondary glutamate, causing deacylation and hydrolysis product formation[11].

The second most abundant class, serine proteases, catalyze numerous hydrolytic reactions with high rates and efficiency for a multitude of endogenous macromolecules. These proteases contain a serine residue in their active site and is coupled to a histidine residue and an aspartate residue. This is generally known as a "catalytic triad" in which a

nucleophile, an acid, and a base reside in close proximity to facilitate hydrolysis of a peptide[12,13]. When a substrate enters the catalytic triad, the aspartate stabilizes the histidine while it hydrogen bonds with the proton of the hydroxyl nucleophile of the serine[14]. The hydroxyl then attacks the carbonyl forming a tetrahedral intermediate and the subsequent electron movement hydrolyzes the amide bond producing an amine terminus and is exported from the active site. As a water molecule enters the active site, the nitrogen of the histidine again hydrogen bonds the one of the water's protons, making the water nucleophilic, where it attacks the *in situ* ester and forms an acyl intermediate. This then causes a collapse of electrons back onto other half of the substrate forming the carboxy terminus, regenerating the serine as the second product is released from the active site[12–14]. Serine proteases also contain a small pocket within their active site known as an oxyanion hole, which uses backbone hydrogens from the amide bonds (from the serine and an adjacent amino acid, usually glycine) to stabilize the negatively charged intermediates[15].

Cysteine proteases, as their name suggests, utilize a thiol containing cysteine residue at the center of their active sites. Although cysteine proteases are similar to the serine proteases, the pK_a of the thiol (normally ~8.3) is significantly lower; hovering near 3.32[16,17]. This has been shown to be due to the thiolate-imidazolium (or thiolate-imidazolium-asparagine) ion pair that exists in the reactive site[17]. Thus depending on the environment of the active site and the family of cysteine protease, a catalytic triad or a catalytic dyad may exist[18]. With a cysteine protease's catalytic triad, the same mechanism as the serine protease can be observed, except the intermediary ester that is formed is now a thioester, and the asparagine stabilizes the thiolate-imidazolium interaction between the

Cys-His[16–18]. In the catalytic dyad, typically the same type of mechanism exists, in the absence of the Asn, and instead is stabilized by electrostatic effects from ionizable groups within 20 Å of the Cys-His; resulting in a nucleophilic attack from the thiolate on the peptidyl substrate[17]. The mechanism proceeds with the thioester intermediate, releasing the *N*-terminus product, addition of nucleophilic water to restore the cysteine thiol and release the *C*-terminus product.

Threonine proteases are unique in that they utilize a secondary alcohol for their nucleophile, rather than a primary nucleophile like the serine's and cysteine's. Due to the secondary alcohol present, the γ-methyl group sterically blocks basic residues (i.e., histidine) responsible for deprotonating the alcohol. To remedy this, nature has constructed threonine proteases to having the threonine be the *N*-terminal amino acid[19]. With this, the γ'-hydroxyl faces the *N*-terminal amine, allowing the amine to behave as a base[20]. The presence of a water molecule allows the amine to abstract a proton, and in turn the newly formed hydroxyl deprotonates the γ-hydroxyl of threonine[21]. The threonine then attacks the carbonyl of the substrate present in the active site, forming a tetrahedral intermediate, which goes on to hydrolyzing the amide bond. The threonine is restored when the water molecule stabilized by the *N*-terminal amine hydrolyzes the intermediary ester, releasing the *C*-terminal product.

Lastly, the aspartic proteases utilize two aspartic acid residues existing with one as the acid and the other as the conjugate base, aspartate. This class represents a small number of the total protease presence in the human genome; only 15 of the 21 identified so far have been discovered to be therapeutically relevant[2]. This however does not include non-

mammalian proteases from viral, fungal, and parasitic sources (composed of >20 individual aspartic proteases)[1]. The mechanism is very simple in that the conjugate base activates a water molecule, which in turn attacks the peptide bond of the substrate[22–24]. The electrons transferred to the oxygen of the carbonyl in the tetrahedral intermediate, abstract the proton of the second aspartic acid[22–25]. The hydrolysis of the substrate is then completed by the second aspartate removing a proton of the tetrahedral intermediate hydroxyl, causing the electrons to resonate back through the amide bond, forming a carboxy terminus and an amine terminus. The resonance restores the aspartic dyad for another substrate to be hydrolyzed.

1.1.2 Serine Protease Recognition Sequences

Comprehending the different protease mechanisms is an important aspect in understanding how the individual proteases behave. However for this work, serine proteases are the focal point. While their hydrolysis mechanism is straight forward, the serine protease's recognition of substrates is seemingly of more interest. Serine proteases are unique in that they are highly plastic enzymes and able to adopt numerous structural conformations to accommodate their substrate(s)[13,26]. This means that a protease can hydrolyze multiple, structurally dissimilar proteins or peptide networks with similar kinetic rates[13,27–32]. Although the specificity of substrate is very broad with the different families of serine proteases, sequence selectivity is usually conserved within each family[13,33,34]. In order words, the exact peptidyl cleavage sequence is crucial and varies between each protease.

For example, trypsin, is considered to be fairly promiscuous, and is most widely known for its use in biochemical laboratories for digestion of other proteins. Trypsin's proteolytic mechanism involves hydrolyzing the C-terminus of a peptidyl linkage when the position, P1 residue, is an arginine (R) or lysine (K). However, the residues proceeding and preceding the P1 position directly affect the recognition and affinity. For most cases, proline blocks cleavage if it is at the P1' position, a caveat though is that if the positional sequence, P4-P3-P2-P1-P1'-P2', is XXWKPX or XXMRPX (where X is any amino acid), the proline has no blocking capabilities[35,36]. Likewise, if the positional sequence is XXC/DKDX, XXCKH/YX, XXCRKX, or XXRRH/RX, the hydrolysis is blocked[35]. The understanding of this specificity and selectivity is critical when investigating active site inhibitors of proteases for drug development. However, these competitive inhibitors simply bind within the orthosteric pocket with a greater binding affinity (> K_D), outcompeting endogenous substrates. This route of inactivating a protease's activity has been studied for decades.

Another key example is the chymotrypsin-like protease, thrombin. It is broadly used *in vitro* for biochemical assays and is the key protease in blood coagulation, which will be discussed later, as the main topic of this work. Like trypsin, thrombin's proteolytic activity is at the *C*-terminal of arginine residues at P1 within a peptidyl sequence[35,37]. Additional optimal conditions include when arginines are flanked by glycines at the P2 and P1' positions. Other favorable conditions are when the P4 and P3 positions are occupied by hydrophobic residues, proline occupies P2, and for P1' and P2' to lack acidic residues[35,37]. Further investigations into thrombin's specificity of the P2-P4 positions when arginine is at

P1 revealed several interesting observations. At P4, long chain, hydrophobic residues (e.g., isoleucine, leucine, norleucine, phenylalanine and to some extent tryptophan) were ideal substitutions[38]. Most likely owing to thrombin's S4 pocket being considerably hydrophobic; containing isoleucine, methionine, leucine, and tryptophan. The P3 position was shown to be non-discriminatory; with the exception of glutamate, which produced little catalytic activity from thrombin[39]. Lastly at the P2 position, proline dominated showing no other residues capable of producing comparable catalytic efficiency[35,37–39]. This information is important in understanding thrombin's natural substrate sequence specificity, and how changes in thrombin's substrates can have remarkably different catalytic rates.

1.2 Venous and Arterial Thromboembolism

The relationship between protease activity and cardiovascular disease (CVD) is strongly correlated. Both genetic and environmental factors have been shown to be key activators of CVD[40]. Notably, there are two common types of CVD: venous thromboembolism (VTE), which includes deep vein thrombosis (DVT) and pulmonary embolism (PE); and arterial thromboembolism (ATE), which includes ischemic stroke and myocardial ischemia, leading to myocardial infarction[40,41]. Two simple distinctions between the two diseases are the locations of a thrombus (blood clot) formed, which in turn causes an embolism (occlusion) in either a vein or an artery, and the clinical symptoms each disease articulates[40]. Although these diseases are classified as separate entities, they share many common traits, risk factors, and have been shown be more related than not[40,42–44]. The general mechanism of VTE is a result unabridged up-regulation of

proteases within the coagulation system, while ATE has been understood to be directly caused by activated platelets[45,46]. However as more investigations ensue, they suggest a relationship between the two diseases[40,42–44].

Distinctions and symptoms relating to venous thromboembolism as stated involves thrombus formation within a vein resulting in occluded blood flow, swelling and pain. DVT presents itself mostly in the "deep veins" including the femoral vein, pelvic vein, and in some instances the cerebral veins[43,47]. PE blood clots are actually fragments of DVT clots, which break off and travel to the lungs. This results in partial or full occlusion of the pulmonary arteries which can be detrimental to the lung tissue containing the clot. It is important to note that although PE blood clots occur in arteries, they are included in the VTE description because they are caused directly by DVT clots.

Arterial thromboembolism generally considered to be more dangerous, due to their location of formation. ATE's typically form around areas that have experienced an arterial rupture called an atheroma. This is usually caused by the buildup of plaques within the arterial wall called atherosclerosis within the heart or brain[45,46]. Once the plaques induce an atheroma, thrombosis ensues forming an occlusion that partially or fully blocks blood flow to the specific area affected[48]. The result in the case of a blockage to heart is a myocardial ischemia and to the brain is an ischemic stroke. The tissue suffers from loss of blood flow, thus loss of oxygen, causing an infarction.

1.2.1 Prevalence

VTE and ATE have significant epidemiological impacts on world populations. It is currently estimated that in the United States alone, venous thromboembolic events result in

>250,000 incidence cases annually[49]. Concurrently, ~30% of patients have at least one recurring episode within a 10-year period following initial diagnosis[49]. The mortality rate varies year-to-year, but typically ranges from 60,000-100,000 deaths/year[49]. According to a report by the American Heart Association, ischemic strokes brought on by ATE, result in > 675,000 incidence cases/year in the US alone, with about 170,000 of those being reoccurrences[49]. Nearly 25% of those cases result in death each year. Looking at the world population, > 39 million identified cases/year from ischemic stroke, owing to 11.13% of deaths worldwide[50]. The same report goes on to stating that 735,000 Americans experience myocardial infarctions each year, with 210,000 of those being repeated episodes[49,50]. This results in ~370,000 deaths/year in the US. For myocardial infarctions worldwide, > 7.2 million deaths occur each year; making it the leading cause of death[49]. These unsettling statistics highlight the fact that although cardiovascular diseases have been relentlessly studied, available treatments are not enough in providing long-term stasis of patients.

1.2.2 Serine Proteases of the Coagulation Cascade

The end goal for many scientists interested in preventing or combating CVDs is to develop new, more effective treatments, or to identify new mechanisms that are able to be targeted. However before understanding the various pharmacological treatments, one must understand the concept behind one of the main driving forces for VTE and ATE, the coagulation cascade. Under normal conditions, endothelial cells, which coat the interior of blood vessels, produce prostacyclin (prostaglandin I2 or PGI_2) and nitric oxide (NO) and excrete them into blood flow. Combined, these two molecules act as vasodilators; while PGI_2 also acts to inhibit platelet activation[51]. Upon injury to the endothelium, collagen is

exposed and tissue factor is released[52]. The glycoprotein von Willebrand factor (vWF), binds to the collagen and mediates the binding of platelets to exposed collagen, which induces initial activation of platelets[53]. The activated platelets undergo major structural changes, becoming dendritic, and release granules containing a multitude of coagulation factors, adhesion molecules, polyphosphates, adenosine diphosphate (ADP), calcium (Ca^{2+}), and thromboxane A_2 (TXA_2). ADP and TXA_2 induce platelet recruitment, activation, and aggregation; while the calcium is essential for several of the coagulation factors to bind platelets[53,54]. The activated platelets expose phosphatidylserine on their extracellular surface, as well as release polyphosphates, giving way to the two coagulation pathways[53,55].

The first coagulation pathway is called the intrinsic or "contact-activation" pathway. The pathway is given that term because the protease cascade that ensues, is a result of polyanionic interactions with the surface of the protease, factor XII[55]. Upon the activation of platelets, small amounts of FXII come in contact with the phosphatidylserine or the polyphosphates released by the platelets, inducing a conformational change, activating FXII. The little FXIIa produced then activates kallikrein from prekallikrein. At the same time, high molecular weight kininogen (HMWK) binds to the negatively charged surface of the injured endothelium and the activated platelets, acting as a cofactor for more kallikrein and FXIIa generation[55,56]. FXIIa in turn, activates factor XI in the presence of calcium, initiating the intrinsic pathway[55]. As seen in **Figure 1**, FXIa goes on to activating factor IX in the presence of calcium[57]. At this point, the cascade enters the common pathway, where now FIXa and calcium activates factor X. Then FXa in complex with

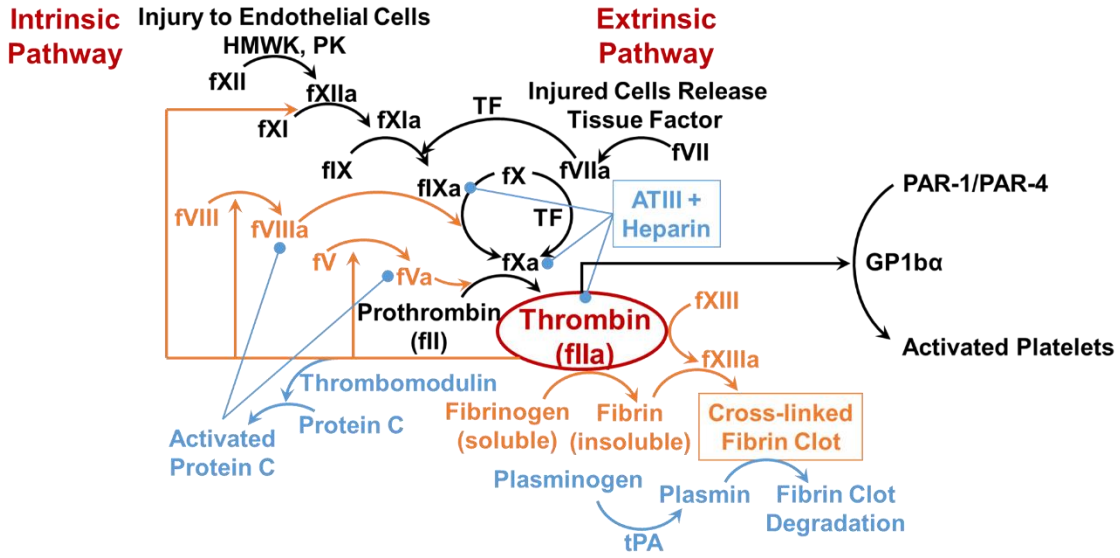


Figure 1: Diagram of the Intrinsic (contact-activation) pathway and the Extrinsic (Injury) pathway. Abbreviations of coagulation factors are denoted as "f", platelet activated receptor (PAR), glycoprotein 1bα (GP1bα), high molecular weight kininogen (HMWK), prekallikrein (PK), antithrombin III (ATIII), tissue factor (TF) and tissue plasminogen activator (tPA).

calcium, and phosphatidylserine activate small levels of thrombin from prothrombin, the key protease of the coagulation cascade[57–59].

To get to the same common pathway, the extrinsic pathway or "tissue factor" pathway, is initiated by the actual trauma of the damaged blood vessel. Damage to the endothelial cells exposes the subendothelium, which expresses high levels of thromboplastin (tissue factor; TF). The TF forms a complex with plasma factor VII, activating it, which subsequently activates both FIX and FX, completing the loop for the common pathway[59,60]. As stated above, the extrinsic and intrinsic pathway generate low levels of thrombin, which then goes on to performing multiple functions, including amplifying its own production. Thrombin behaves as a procoagulant by activating factor VIIIa, FIXa cofactor[61]; as well as activating factor V, a FXa cofactor[61]. As these cofactors are activated and complexes with their respected enzymes are formed, the production of

thrombin is greatly accelerated. When thrombin levels are sufficient, thrombin goes on to hydrolyzing its main substrate, fibrinogen, into insoluble fibrin monomers. Thrombin concurrently will activate factor XIII, a transglutaminase responsible for cross-linking the fibrin monomers, forming the initial fibrin mesh that holds the activated platelets together around the site of injury[62]. Thrombin can also recruit more activated platelets to the injury site by binding with glycoprotein 1bα (GP1bα) and cleaving the protease-activated receptors (PAR) 1 and 4 on the platelet surface[63,64]. Binding of thrombin to GP1bα (or polyphosphates) also induces a feedback loop of thrombin activation by independently activing more FXIa (see **Figure 1**)[65,66].

1.2.3 Hypercoagulation: Current Treatments, Mechanisms, and Side Effects

As insinuated above, the concentration of zymogen and active proteases released by platelets or that are free flowing in plasma play a critical role in contributing to blood clotting, as well as disorders like hypercoagulation. Hypercoagulable states (thrombophilia) are defined as when a patient has a higher tendency to form abnormal blood clots due to several factors including genetics, medical conditions, or medication being taken[67]. Predominantly, inherited factors are the most concerning, due to a patient being unaware that they are predisposed. The most common type of thrombophilia is factor V Leiden, a genetic mutation of the F5 gene, where factor V is translated[68]. The mutation causes a sequence alteration at the codon for Arg506 at the activated protein C cleavage site and substitutes with glutamine[69]. This mutation results in FVa:FXa complex remaining active for 10-fold longer, producing considerably more thrombin. There is also support that the lack of hydrolyzed FVa at R506, which is a cofactor for inactivation of FVIIIa, also

allows FXa generation to be accelerated for longer periods of time[69]. Combined, these would produce an abnormal acceleration of clot formation.

The second most common type of thrombophilia is the prothrombin gene mutation (G20210A mutation) on the F2 gene. This single nucleotide mutation is part of the noncoding region of the prothrombin gene, and thus does not cause any mutation to the prothrombin itself. However, the mutation does cause an up-regulation of prothrombin produced[69,70]. Statistical analysis of people with the G20210A mutation showed a 2-3-fold increase in having a clotting event[69,70], with those odds increasing if combined with other risk factors, especially cancer.

In order to combat these thrombotic events, whether the source is congenital or acquired, anticoagulant drugs have been used and continue to be in development for treatment of DVT, PE, and arterial thrombosis. These drugs act in one of two ways, direct or indirect. One of the first and most widely used anticoagulants is warfarin. Warfarin is a coumarin-based small molecule that acts as an indirect anticoagulant. The mechanism by which warfarin proceeds through targets the coagulation proteases requiring calcium, for binding platelet surfaces[71]. Better known as a vitamin K antagonist, warfarin inhibits vitamin K epoxide reductase, and effectively reduces the amount of vitamin K available as a cofactor for γ-glutamylcarboxylase[71–73]. The γ-glutamylcarboxylase adds a γ-carboxylic acid to glutamic acids on thrombin, FVII, FIX, FX, and protein C; called the Gla-domain, responsible for binding calcium[73]. The warfarin in turn, causes production of coagulation proteases that lack the Gla-domain, which renders them incapable to bind to platelets[71]. This greatly reduces the ability for coagulation to ensue. Other indirect inhibitors of

coagulation include glycosaminoglycans and their derivatives, which will be explained in greater detail in Chapter 1 Section 3.

Although warfarin and its derivatives function very well to inhibit coagulation, extensive monitoring of the drug per individual patient is required due to the complications associated with warfarin therapy[74]. The main issue relating to warfarin is the high risk of excessive bleeding events (hemorrhaging), even when taken as prescribed. Accidental overdose greatly increases the risk for hemorrhages and induction of hemophilia-like symptoms, with the only immediate reversal agent (antidote) being fresh frozen plasma[75]. Vitamin K can be introduced, but takes 4-6 hours before liver production of natural coagulation factors containing the Gla-domain will be synthesized[75].

Due to this potentially severe complication, researchers began investigating alternative anticoagulants. These direct anticoagulants, bind directly to a protease's active site, rendering them incapable of catalyzing hydrolytic reactions[76]. To date, there are two types of direct orthosteric anticoagulants available in the clinic, with more being investigated for alternative targets. The two types include thrombin-specific and FXa-specific inhibitors. For thrombin-specific there are also two subtypes, univalent and bivalent. The most notable univalent direct thrombin inhibitors (DTIs) are argatroban and dabigatran. Argatroban is an injectable small molecule that has been approved for the treatment of thrombosis and for patients experiencing heparin-induced thrombocytopenia (HIT); a immunological condition brought on by the anticoagulant heparin, causing a reduction in platelet concentrations[77]. The drug reversibly binds to thrombin's orthosteric site with an inhibitory constant (K_I) of 40 nM and half-maximal inhibitory concentrations

(IC_{50}) of ~10 nM against platelet aggregation for free thrombin and ~20 nM for clot-bound thrombin[77,78]. Dabigatran is an orally bioavailable small molecule that reversibly inhibits thrombin's orthosteric site with a K_I of 4.5 nM[79], as well as acts on thrombin in the clot-bound state with an IC_{50} of 10 nM[79]. It is dosed as a prodrug, dabigatran etexilate, which additionally contains an ethyl ester and hexyloxycarbonyl carbamide, to increase hydrophobicity and membrane absorption[80].

The bivalent DTIs indeed interact with the active site of thrombin, but also with a secondary site called exosite 1. Exosite 1 is an allosteric site on thrombin adjacent to the orthosteric site, and will be discussed in greater detail in Chapter 1 Section 4. The first bivalent DTI to be discovered was hirudin. Hirudin, a 65 amino acid peptide found the in the salivary glands of leeches, has an interesting mechanism of action[81]. It first recognizes thrombin by binding to exosite 1 through mainly electrostatic and hydrophobic interactions using its last 17 residues at the *C*-terminus[81]. Then the peptide chain, mainly through its hydrophobic residues of the *N*-terminus binds within the active site pocket of thrombin, occupying the S1-S4 pocket[81]. This discovery was a novel mechanism, both in the exosite binding and the way it binds to thrombin's active site pocket.

The direct FXa inhibitors (DXIs), like apixaban and rivaroxaban, act by targeting FXa's orthosteric site. The compounds are selective and do not inhibit thrombin or other coagulation proteases, which subsequently do not have any direct effects on platelet aggregation[82]. However, these two small molecules do inhibit both free and clot-bound FXa, which indirectly inhibit thrombin activation. Both compounds have been approved

for treatment of DVT and PE. Apixaban boasts a K_I of 0.08 nM[83] and rivaroxaban has a 0.4 nM K_I[84].

A large drawback of all of these anticoagulants is the risk of excessive and severe bleeding. Although the DTIs and DXIs effectively treat thrombotic complications better than warfarin, some side effects are more prevalent than compared to warfarin. Argatroban, dabigatran, and rivaroxaban have been shown to have almost a 2-fold increase in major gastrointestinal bleeding[85]. Argatroban also showed with >10% of patients in trial experiencing minor bleeding in urinary organs, coupled with increased events of low hemoglobin[85]. Hirudin clinical trials showed 2-10-fold increases in bleeding from injury sites, drops in hemoglobin, and non-threatening bleeding (i.e., nose bleeds)[85]. Until recently, none of these treatments had an approved, reliable emergency antidote. In 2015 however, idarucizumab was discovered to be a fast-acting antidote for dabigatran and its anticoagulant effects[86].

1.3 Glycosaminoglycans

An alternative treatment to the anticoagulants discussed above, along with the oldest anticoagulant known, heparin and its many derivatives form a group of compounds known as glycosaminoglycans (GAGs). GAGs as they exist in nature, are linear, complex, carbohydrate structures composed of repeating, disaccharide units[87]. The type of monomeric sugars that make up the disaccharides contain at least one amine in each disaccharide. GAGs are biosynthesized in virtually all mammalian cells and are usually attached to a protein or lipid, dubbed a proteoglycan or glycolipid[87]. Thus, GAGs are found intracellularly, on the cell surface, and extracellularly, by means of excretion by some

cells[87]. GAGs are one of the most unique biomolecules synthesized by cells because they are non-template-based and can have variable modifications through epimerization, sulfation, and acetylation.

GAGs can be broken down into five major classes: hyaluronan (HA), chondroitin sulfate (CS), dermatan sulfate (DS), keratin sulfate (KS), and heparin/heparin sulfate (H/HS). HA is the only non-sulfated GAG, composed of the disaccharide repeating units of *N*-acetylglucosamine and glucuronic acid, which contains glycosidic linkages that alternative between the anomeric carbons β1-4 and β1-3 (GlcNAcβ1-4GlcAβ1-3)[87]. CS is a variably *O*-sulfated GAG made up of *N*-acetylgalactosamine and glucuronic acid. The CS is linked through glycosidic bonds at the β anomeric carbons, 1-4 and 1-3 (GalNAcβ1-4GlcAβ1-3) with a sulfation pattern typically composed of the 4- and 6- position of the GalNAc and the 2-position of GlcA[87]; although analysis has shown that some disaccharides remain desulfated[88]. DS is very similar to CS and was once considered a derivative of chondroitin sulfate, but the GlcA is isomerized, giving way to iduronic acid (IdoA)[87]. Thus DS contains the GlcNAcβ1-4IdoAβ1-3 disaccharide linkage. KS is another variably *O*-sulfated GAG consisting of galactose and *N*-acetylglucosamine (Galβ1-4GlcNAcβ1-3), but interestingly can be part of a branched glycan system through mannose linking[89]. It can also be capped at the non-reducing end with fucoses or *N*-acetylneuraminic acid[89].

Lastly and more importantly for this work, heparin/heparan sulfate is a variably *O*-sulfated and *N*-sulfonated GAG composed of α-glucosamine and β-glucuronic or α-iduronic acid (GlcN[Ac/S]α1-4GlcAβ1-4/IdoAα1-4), depending on the individual

disaccharide unit[87]. Variability of acetylation or sulfonation at the C-2 amine, as well as isomerization within disaccharide units causing the anomeric carbon to change between α and β, makes H/HS the most complex GAG sequence[90]. Heparin and heparin sulfate are collective in grouping because their backbone disaccharides and sulfation/acetylation patterns are generally the same, however heparin has been shown to be more sulfated than HS; heparin contains ~2.7 sulfates per disaccharide *versus* HS having ≥ 1 sulfate per disaccharide[87]. In humans, free heparin only consists of ~5% of the total H/HS, with HS dominating in the proteoglycan form. Additionally, heparin is produced mainly in mast cells, only being excreted under affliction[87].

GAGs have a role in almost every biochemical mechanism in the body, with H/HS most notably being involved within blood coagulation[91]. H/HS are indirect anticoagulants that act as cofactors for endogenous serine protease inhibitors (serpins). These serpins, such as antithrombin and heparin cofactor II, act to covalently inactivate the serine protease by insertion of an inactivation loop into the active site of the protease. The serpin initially interacts with the protease non-covalently, through a Michaelis-like complex, similar to that of normal substrate interactions[92]. However, when the protease attempts to hydrolyze the inactivation loop, the hydrolysis reaction is halted when the serine covalently cleaves the *C*-terminal arginine within the inactivation loop[92]. Once the covalent linkage occurs, the protease-bound serpin undergoes a conformational change, irreversibly augmenting the active site pocket, disrupting the catalytic triad[92]. H/HS accelerates this reaction by binding tightly to a distal site on the serpin (K_D = low nM) and weakly binding to a distal site on the protease (K_D = low-to-mid μM)[93]. The protease therefore has a more

dynamic interaction with the H/HS, allowing it to "walk" along the H/HS chain until it can begin interacting with the serpin to form the initial Michaelis-like complex[92,94]. This acceleration by H/HS increases the rate of the inactivation by 2000-4000-fold[95,96].

In clinically relevant situations, exogenous heparin extracted from bovine or porcine sources is used for anticoagulant treatment. The heparin derived in this fashion must meet strict federal guidelines pertaining to the molecular weight, degree of sulfate, and animal source[97]. Because full length heparin is not yet possible to chemically synthesize, a number of enzymatic, chemoenzymatic, and synthetic steps have been taken to create heparin derivatives like low-molecular weight heparin (LMWH; enoxaparin)[98], ultra-low-molecular weight heparin (ULMWH)[99], and fondaparinux (pentasaccharide)[100].

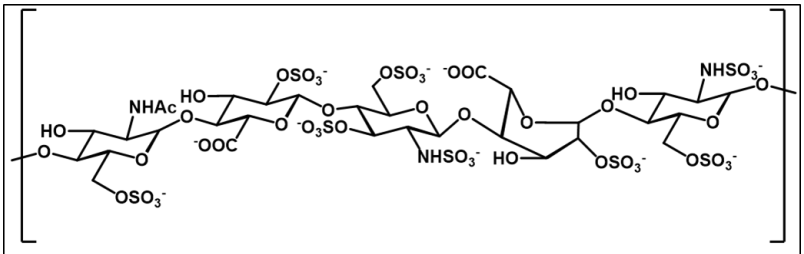


Figure 2: Heparin 'DEGFH' high affinity antithrombin binding sequence

These heparins show lower affinity towards plasma proteins and as they decrease in size, show less effects accelerating serpin-thrombin interactions, but focus mainly on facilitating antithrombin (AT)-FXa interactions; with fondaparinux exclusively targeting AT-FXa (see **Figure 2**). The reasoning behind making smaller heparin analogs is to combat heparin-induced thrombocytopenia (HIT), mentioned earlier. In some patients, heparin triggers the formation of antibodies, that act by activating platelets[101]. The activated platelets form microaggregates that lowers the free platelet count in blood, but also can cause occlusions to blood capillaries[101]. Lower molecular weight heparins reduce the amount of HIT observed as they get smaller.

Unfortunately the drawbacks besides HIT as seen with other anticoagulants, include major bleeding complications with heparin, LMWH, and fondaparinux. Although LMWH and fondaparinux have reduced the risk for HIT, both showed higher prevalence of major and minor bleeding after surgery, but overall equivalent or less bleeding for treatment of DVT and PE compared to heparin[102]. This information suggests that although there are downfalls for every type of anticoagulant, immense progress has been made regarding safety of the drugs. However the notion of an "ideal" anticoagulant, which would instill hemostasis to predisposed and afflicted patients, still looms.

1.4 Allostery and Conformational Change

Arguably one of the most important phenomenon that occurs with proteins and within the study of enzymology, is the concept of allosterism. Allostery can be defined as a distal interaction that induces a functional change to a protein (or enzyme), thereby altering the primary site's binding capability or catalytic efficiency[103]. Two major theories have been at the forefront when describing allostery. The first was introduced by the discovery of hemoglobin's quaternary structure[104]. The theory suggested that structural alterations of the protein's conformation, induced by allosteric effectors, changed a substrate's binding affinity to the protein. This brought about the common explanation of oxygen binding and the effect of 2,3-biphosphoglycerate (BPG) binding to hemoglobin and how it effects the tense (T) and relaxed (R) states. It stated that hemoglobin in the deoxy form (deoxyHb) has a low affinity for oxygen (T-state), but once an oxygen binds to one of the four subunits, cooperativity takes over. This cooperativity induces conformational change to the adjacent subunits, allowing those subunits to bind with

oxygen, transitioning to oxyhemoglobin (oxyHb) or R-state[105]. When BPG binds to the allosteric site of Hb, it stabilizes the T-state, in turn reducing the affinity for oxygen.

An alternative to this theory is the idea that an allosteric effector can broaden or narrow the free energy within a protein's conformational space[106,107]. This theory implies that entropic events within the conformational space of the protein actually affects the thermodynamics and the conformational dynamics, thus speeding up or slowing down the motions of the protein itself[106,107]. Simply put, a distinct structural change does not need to occur for a protein to respond to an allosteric effector[108]. Therefore when looking at allosteric effects as a whole, one must take into consideration that if the no structural change can be identified, then dynamic allostery must be taking precedence.

Allostery is a significant aspect of the coagulation cascade in that thrombin, the key regulatory protease of the system, is highly regulated by allosteric interactions[109,110]. Structurally, thrombin is a 37 kDa monomeric, serine protease with two distinct allosteric sites called exosite 1 (residues include: K36, H71, R73, R75, Y76, R77, K109, and K110)

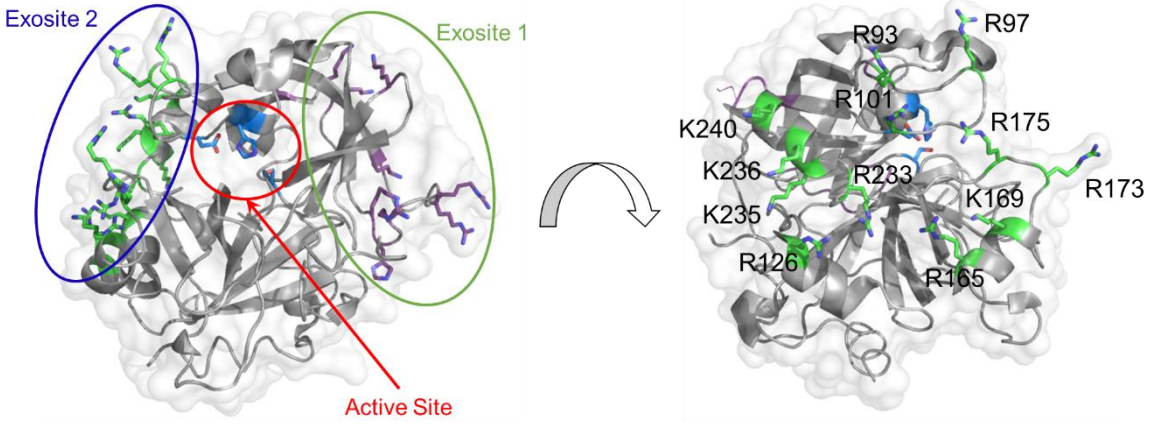


Figure 3: PDB ID: 3EQ0 showing the residues of the active site, exosite 1, and exosite 2. A 90° turn shows the face of exosite 2 and labels the basic residues.

and exosite 2 (residues include: R93, R97, R101, R126, R233, K235, K236, and K240); as seen in **Figure 3**, using PDB ID: 3EQ0. This is in addition to the typical serine protease catalytic triad composed of S195, H56, and D102. These exosites are highly cationic in nature, owing to their ability to bind with anionic species like the sulfates of GAGs, the tyrosine sulfates of many cofactors and peptides, and D/E residues found in repeat sequences of many of the same cofactors and peptides[109]. Additionally, the exosites contain many hydrophobic pockets that aid in ligand specificity and recognition[109,111].

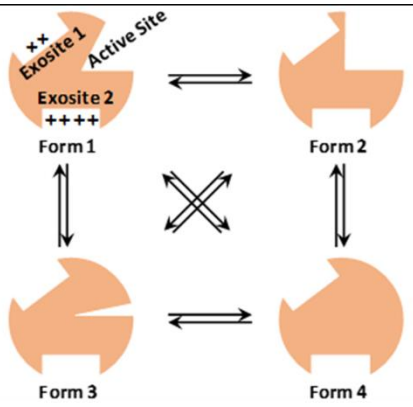


Figure 4: Diagram of thrombin depicting conformational plasticity; where each form has an equal chance of being present at any time.

In conjunction with the active site and exosite recognition, thrombin is known to have high degrees of conformational plasticity[112], allowing it to sample a number of conformational states at any time; shown in **Figure 4**. Thrombin is thus considered to have a large conformational space. Because of this flexibility, thrombin is able to recognize and hydrolyze multiple substrates with only three recognition sites (active site, exosite 1 and 2)[112]. Therefore, it is reasonable to assume that allosteric effectors can induce either conformational or dynamic allostery. For example, when heparin binds to exosite 2, no significant structural change is observed in crystallographic studies[113] or by using fluorescence techniques[114]. However, other endogenous ligands like γ'-fibrinogen[115] and GP1bα[116] induce significant alterations to thrombin's structure through exosite 2-mediated interactions; sharing similar binding modes and residues to heparin. It is also important to note that thrombin's S1 pocket is significantly affected by the

concentration of sodium in solution[117]. Na^+ has been shown to bind with the backbone carbonyl oxygens of thrombin's Y184, R221 and K224, as well as with four water molecules[117]. The binding of sodium coordinates the residues properly to give functional geometry to the S1 pocket, allowing it to recognize the arginine residue of substrates[117]. Thus, low sodium concentrations or blocking of this region would yield impaired recognition of the substrate and reduced catalytic activity.

1.5 Glycosaminoglycan Mimetics

Having recognized that the exosites of thrombin can be selective for specific sequences, to which the extent of these can be altered by the hydrophobicity of the ligand (i.e., increasing the non-ionic binding energy), it was apparent that the design of synthetic compounds could result in allostery-induced dysfunction of the catalytic site. Initial development and studies used chemoenzymatically-derived sulfated, low molecular weight lignins (LMWLs)[118–120], which were purposed to mimic heparin binding. The difference however was that backbone scaffold was significantly more hydrophobic by use of phenoxypropanoic acid and dihydrobenzofuran repeat units instead of GlcN and GluA/IdoA repeats[118]. Also the sulfation in the LMWLs was *O*-phenolic, replacing the GAGs alkyl *O*-/*N*-sulfonation[118].

Using chromogenic substrate hydrolysis assays, the LMWLs were shown to inhibit thrombin's activity by direct interaction of the heparin binding site (HBS), shown through heparin competition and Michaelis-Menten kinetic studies with low nanomolar IC_{50}s[118,120]. These were confirmed as non-saccharide glycosaminoglycan mimetics (NSGMs) by means of salt dependence studies and mutant studies. The LMWL's binding affinities were

reduced as the concentration of sodium was increased, suggesting that the multivalent interaction of the sulfates was taking place, as seen with heparin[119]. The mutant studies revealed that like heparin, R93 and R175 had significant impacts in binding with the LMWLs showing a 7-8-fold reduction in inhibition when mutated[121]. Additional electrostatic interactions included R101, K233, K235, and K240, which form the heparin binding site[121]. However the notable difference is that although the LMWLs and heparin bind at the same site, the LMWL induce allosteric inhibition of thrombin, while heparin does not. Thus, the inhibition aspect must arise from the hydrophobic interactions of the backbone scaffold.

Although the LMWLs displayed impressive inhibition concentrations, they were not selective. Along with inhibiting thrombin, they also inhibited plasmin, FIXa, FXa, FXIa, neutrophil elastase, and cathepsin G[120]. The LMWLs were also polydisperse, variably sulfated, and had non-uniform linkages[118]. To combat some of these issues, a synthetic polymer known as SbO4L was developed. Being SbO4L was fully synthetic and the chemistry restrictive to one linkage being formed by a phenolic S_N2 reaction, the polymer formed was linked uniformly[122]. The polydispersity and average sulfation presence were also uniform as compared to unfractionated heparin and LMWH used in the clinics. Significant advances over the LMWLs was SbO4L's $\log_{10}$-fold selectivity towards thrombin and its ability to inhibit both thrombin and thrombin's binding affinity to GP1bα on platelets, effectively acting as an antithrombotic and antiplatelet agent[122,123].

Stemming from the LMWL discoveries and concurrently with the SbO4L work was the group of molecules known as sulfated allosteric modulators (SAMs), which more

closely resembled small molecule-based inhibitors, rather than polymeric ligands. The initial SAMs included sulfated monomeric, dimeric, trimeric, and tetrameric benzofurans[124–126]. The synthesis of these molecules was to eliminate polydispersity, while increase the non-ionic contributions to LMWL scaffold. As mentioned previously, the LMWLs contained dihydrobenzofuran cores, which were easily attainable through the cyclization of 1,2-dihydroxybenzene and acetoacetate[124,125]. This gave the benzofuran heterocycle that was variously oligomerized in multiple reports and contained various sulfation patterns, as well as a variety of protected and deprotected scaffolds[124–126]. The general pattern of inhibition was that as the scaffolds elongated through oligomerization, the inhibition became more potent. Additionally, the BFs showed much better inhibition containing only one phenolic sulfate, and opening up the ester present in the structure to a carboxylate, did not increase the potency[124,125]. Suggesting that the BFs act through mainly hydrophobic interactions, while the sulfate is simply needed to "steer" the benzofuran to the exosite[127]. Using the same techniques as for the LMWL, the BF family of compounds showed that they bind to exosite 2 and specifically, the most potent BF dimer (IC_{50} = 5 μM) was shown to bind at R173[128], while the benzofuran trimer (IC_{50} = 0.67 μM) was shown to bind at the R233 residue of thrombin[125]. Modeling studies showed two distinctly different binding sites for the dimer and trimer, and the tetramer did not have high potency[125]. Overall this suggested that the inhibition of thrombin by the BFs was size dependent and relied on hydrophobic interactions as well as the thrombin residues donating hydrogen bonds to the H-bond acceptors of the BF molecules. It is important to specify that the BF trimer was not selective for thrombin, and inhibited FXIa with nearly

equipotent IC_{50}s (0.8 μM for FXIa *versus* 0.67 μM for thrombin), by binding to FXIa's HBS[129].

Other SAMs produced included quinazolinone, flavone, and tetrahydroisoquinoline scaffolds. These compounds were structurally different from the BFs and so were their specificity towards coagulation proteins. The quinazolinones/flavones (QAOs) were synthesized as monomers or as dimers (homodimers or heterodimers with a flavone moiety), which were linked by alkyl chains connected by 1,2,3-triazole(s), or by just alkyl chains of varying lengths. The compounds were shown to be selective toward FXIa, over thrombin, FXa, trypsin, and chymotrypsin[127]. The most potent QAO was a bisulfated homodimer with a 1,2,3-triazole containing a 5 carbon alkyl linker, possessing an IC_{50} of 59 μM[127]. Alteration of the linkage from the 1- and 4-position to the 1,2,3-triazole to 1- and 5-position led to ~2-fold drop in potency, suggesting that the triazole geometry plays a significant role in binding to FXIa[127]. Alternatively, the tetrahydroisoquinolines (THIQs) were found to be antithrombin accelerators for FXa inactivation. The THIQs accelerated the inactivation of FXa by ~79-fold, but doing so with low potency[130]. However, an interesting point to take away is that the relatively similar scaffolds have remarkably different specificities and induce very different effects. Some act as direct, allosteric inhibitors, that bind to the HBSs of proteases, while others bind to the HBSs and act as heparin does in acceleration of AT inactivation.

1.6 Partial Inhibition

One aspect of allosteric inhibition that has been hinted at by previous data, was the observation that some compounds, specifically SbO4L and BF trimer, exhibit submaximal

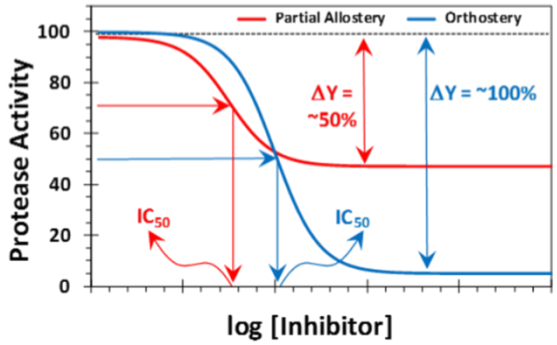


Figure 5: Example of dose-response profiles of a protease activity with an orthosteric inhibitor and an allosteric, partial inhibitor. Abbreviations: IC_{50} = half-maximal inhibitory concentration, ΔY = efficacy of inhibition.

inhibition of thrombin. In lay terms, even at saturation of the ligand, the protease remains active and able to hydrolyze substrate. In graphical terms (dose-response profiles), partial inhibition is identified by the ΔY (efficacy of inhibition) where the ΔY > 20% and < 80% (**Figure 5**). What is of interest is the principle that while the ligand is inhibiting the protease, it remains partially active towards the *small* substrate. This phenomenon could translate into multiple downstream effects. For instance, if there is an alteration to thrombin's pocket, endogenous substrate may not be as recognizable or there may be activation when there is typically inhibition[131]. Another possibility is that the catalytic efficiency is reduced, thereby slowing the enzyme's hydrolysis rate due to dysfunction of the S1 pocket or catalytic triad[132]. Furthermore, partial inhibition is a rare event in enzymes, especially with thrombin being a monomeric protein. The typical proteins or enzymes that experience partial inhibition are multimeric proteins such as receptors/membrane-bound proteins or multimeric proteins.

There are a multitude of examples for allosteric inhibitors inducing partial inhibition of receptor/membrane-bound proteins. Likewise, the examples include not only the typical structural conformational changes, but also conformational dynamic changes. Recently, two examples of partial agonism explained by conformational dynamics have been reported describing ionotropic and metabotropic glutamate receptors, respectively.

The first report used crystallography and computational simulations which compared agonist, partial agonist and antagonist ligand volumes bound to the binding site with calculated free energy surfaces. The study showed that increasing the ligand size, lowered the agonist efficacy; meaning that the free energy's minima narrowed, reduced, or a combination of the two[133]. This was indicative of small ligands being agonists, and as their size increased, they became partial agonists or antagonists. Thus, the free energy diagrams for each ligand may be used to predict if a ligand would act as a partial agonist or not. Furthermore, the curvature of the free energy graphs may be correlated with the energies relating to the glutamate channel's opening or closing in response to a ligand, but with a caveat that the simulations were performed using a monomer of the receptor and not the full tetrameric structure[133].

The second report takes this notion a step further where they investigated the glutamate receptors using a partial agonist and monitoring the FRET changes at the dimer interface of the ligand binding domain. They successfully demonstrated that the partial agonist, DCG-IV, did not induce full occupancy of the low FRET state, even at saturating levels; unlike the full agonist, which occupied the high FRET or active state[134]. Additionally, saturating concentrations of DCG-IV was able to keep dynamics near the basal state of the receptor. This was also compared to full agonist, which showed that its induced dynamics were dramatically slower, in looking at FRET traces[134]. The last major aspect described that mutations made at conserved sequences of transmembrane (TM) domains, positive allosteric modulators (PAMs) that bind at those TM domains drastically increase the efficacy and affinity of the DCG-IV[134]. This supported the notion that the TM

domains are directly linked to the ligand binding domain, an important feature the first report hypothesized.

A caveat regarding this partial agonism is that these reports indeed showed partial inhibition through conformational dynamics, but did not use allostery. This is an important aspect to take into consideration because again, receptors are multimeric proteins, but not enzymes. Therefore, their mechanisms of action are vastly different and are difficult to make comparisons to. Although many allosteric agonists exist, they do not all directly cause partial inhibition in receptors[135]. They can however enhance or modulate the partial agonist, as seen in the example of DCG-IV and the PAM of the TM domains[134]. An example however of an allosteric partial antagonist examines the structure activity relationship of the 5-(phenylethynyl)pyrimidine scaffold. They showed that by varying small structural changes to the scaffold, it changed the allosteric activity from partial antagonist (71% response) to full antagonist (99% response)[136]. They concluded from this study that the allosteric partial antagonism is not conserved to the general scaffold because of the large variations in the activity of the inhibitor on the receptor[136]. This demonstrates the high sensitivity of an allosteric site and that minor changes can have dramatic effects on the outcome, most likely due to dynamics. To look at more classical examples of allosteric partial inhibition, one must turn to a recently discovered inhibitor of a phosphodiesterase, which is most relevant to this work.

The study on phosphodiesterase 4 (PDE4) examines the notion of an allosteric partial inhibitor in order to aid in the prevention of side effects from patient using orthosteric inhibitors of PDE4, which can cause abnormal cAMP concentrations in cells[137].

Their evidence suggests that the PDE4 acts a dimer with negative cooperativity and can be within multiple conformations; so as in saying only one of the two binding sites may be open or closed at any time. Their lead designed compound was shown to be 1000-fold more selective for the PDE4 isoform[137]. The study had observed that their lead compound acted by shifting the binding of two key hydrophobic residues at the active site and predicted interactions with the residues involved in binding the Zn^{2+} metal[137]; which is linked to conformational allostery. The overall partial inhibition was > 50% with an IC_{50} of 2 nM, while also decreasing the k_{cat}/K_M of one of the active sites[137]. Additionally, the cAMP levels in the cell cultures were not shown to be elevated, insinuating that the allosteric partial inhibitor was indeed perform its enzymatic duties[137].

The work described herein provides several details of partial inhibition by allosteric and non-allosteric structural conformational changes and conformational dynamics. Convincingly, there is a plethora of information regarding partial inhibition, but only with regard to higher multimeric proteins, with the last study involving a dimer. One important factor that separates oligomeric proteins with enzymes is that the for enzymes, only allosteric effectors can induce partial inhibition[135]. With the receptors, the partial inhibition can be induced by an allosteric modulator or by a ligand within the main ligand binding domain, because the receptors act through channel opening and closing. If a ligand binds to the enzyme active site, there is an all-or-none phenomenon where partial inhibition fundamentally cannot occur. However, allostery offers up that accessibility if the ligand induces a specific change to the structure or to the dynamics of the enzyme. The following work describes the first known compound to induce allosteric partial inhibition of a

monomeric enzyme, thrombin, its mechanism of inhibition, as well as a number of interesting actions on endogenous substrates.

CHAPTER 2 Coumarin-Based Sulfated Allosteric Modulators

Over the past decade, our lab has presented multiple works on the development of non-saccharide glycosaminoglycan mimetics (NSGMs). These NSGMs began as chemoenzymatically and fully synthetically produced , lignin-based polymers presenting diverse arrangements of *O*-sulfates and carboxylates[114]. The first generation of NSGMs were plagued with difficult characterization and separation; resulting in heterogeneous mixtures and diverse conformational sequences[122]. Stemming from these polymers came small molecules composed of heterocyclic cores, some mimicking the lignins, decorated with *O*-sulfates synthesized from phenolic moieties. These heterocyclic cores included benzofurans, quinazolinones, and flavones that were analyzed from monomeric cores to tetrameric cores, in some instances. The addition of heterocyclic cores to produce monomers through tetramers was done in two manners, direct conjugation or with varying linkers. To this, the development of more than 150 diverse sulfated allosteric modulators (SAMs) has led to a number significant advances within the fields of regulating

coagulation[124,125,127,129,130,138] and inhibition of cancer stem cell proliferation and differentiation[139].

We chose to expand the diversity and functionality of the SAM libraries, which led us to study coumarins as the heterocyclic scaffold of choice for several reasons. A number of coumarins are commercially available with diverse, synthetically accessible handles, which greatly reduces the number of steps involved with synthesis. Coumarins are pharmacologically relevant; the best and most relevant example of this is the vitamin K antagonist, warfarin, which utilizes a coumarin backbone[71]. The coumarin scaffold offers a structurally diverse core compared to the other heterocycles in the library, like the flavones and quinazolinones.

We chose to focus on the coagulation cascade to examine whether these coumarin-based sulfated allosteric modulators (CSAMs) were efficacious against any of the pharmacologically relevant proteases involved. Our earlier work on designing/discovering allosteric modulators of thrombin showed that the sulfated scaffolds exhibited allosteric, full inhibition (ΔY = 80–100%) at saturation[124–127]. However, an interesting observation made with the lignin SbO4L and one of the benzofuran trimers was that at saturating concentrations of the inhibitors, thrombin remained partially active (ΔY = 20-79%)[122,124]. This idea of allosteric, partial inhibition was intriguing and sought after with the CSAMs.

2.1. Synthesis of First and Second Generation CSAMs

Sulfated coumarins were synthesized using either a protection–deprotection strategy or a selective reaction strategy, which enabled the generation of the desired number of free *O*-phenolic groups that could then be sulfated in a final step using sulfur trioxide–trimethylamine complex[140,141]. Initially, *O*-methylated analogs are deprotected using boron tribromide (BBr_3) producing the free hydroxycoumarin scaffolds. The hydroxycoumarins were then sulfated using microwave assisted sulfation (MWAS), requiring sulfur trioxide trimethylamine complex ($SO_3{\cdot}N(CH_3)_3$), triethylamine (TEA) as a base, and acetonitrile (ACN) as the solvent. First and second generation dimeric species were also considered due to previous studies indicating that spanning a greater radius with larger compounds increased potency. The second generation dimers were chosen to represent the most promising compounds, with increased sulfation. For the first generation dimeric species, the hydroxycoumarins were alkylated selectively with either propargyl bromide, 1-bromo-3-chloropropane, or 1-bromo-4-chlorobutane. Subsequently, the chloropropane or chlorobutane was reacted with sodium azide, to produce the alkylazide. To form the dimeric compounds, the coumarin alkylazide and propyne groups were reacted through copper (I)-assisted azide-alkyne cycloaddition (CuAAC). The dimeric precursors were then sulfated using MWAS, resulting in a *O*-sulfate trimethylamine complex. The second generation CSAMs first had their catechol moieties selectively protected using the diphenylmethane (DPM) group, proceeded through the first alkylation steps, and was deprotected. This deprotection prior to azidation and CuAAC was selected due to the lack of dimeric product formation when the DPM was present. The free hydroxyls were then

sulfated using the MWAS. For all sulfated compounds, an additional step using cation exchange chromatography ensured the alkylammonium sulfates were exchanged to sodium sulfates and lyophilized to produce the final compound.

Collectively, the overall synthesis resulted in a 36 compound library of first generation CSAMs (**Figure 9**). Then followed by two second generation compounds, which contained higher degrees of sulfation of the two most promising molecules. Thus in total, a 38 compound library was successfully synthesized exploiting various sulfation patterns, substituent placement, and monomeric/dimeric species.

2.1.1 Experimental Details

For *O*-methylated coumarins, 50-100 mg of the starting material were suspended in 5 mL of dry DCM and subjected to 1.5 eq. per oxygen atom of 1.0 M BBr_3 solution at -78 ˚C under N_2 atmosphere. The solutions were stirred for 5 h, allowed to reach room temperature (rt), following which the solutions were again cooled to 0 ˚C and 6 mL od water (or 1:1 water/methanol) were slowly added with stirring to quench the reaction over the course of 10 minutes. The biphasic solution was then condensed *in vacuo* and was taken up in 20 mL of ethyl acetate to which 15 mL of saturated ammonium chloride solution was sequentially added. The mixture was transferred to a separatory funnel where the organic phase was extracted and the aqueous phase was followed by two washes of 10-20 mL of ethyl acetate. The organic layers were pooled, washed with brine, dried over anhydrous Na_2SO_4, and then condensed *in vacuo*. The residues were taken up in minimal DCM (~2 mL), dry-loaded onto silica gel and purified via normal phase flash chromatography (0-70% ethyl acetate/hexanes). Fractions that contained the *O*-

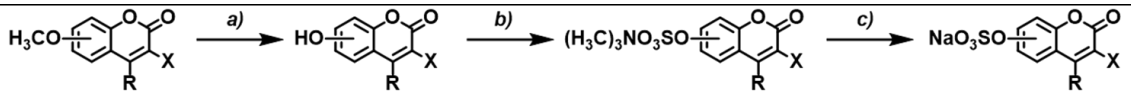


Figure 6: Synthetic scheme for demethylation of *O*-methoxy protected coumarins. Subsequently the sulfation of hydroxy-coumarins is shown:
a) 1.0M BBr_3, dry DCM, -78 °C to rt, 5h, 87-94%
b) SO_3•$N(CH_3)_3$, TEA, ACN, MW, 100 °C, 1-3h, 91-94%; ≥95% purity
c) Cation exchange chromatography.

demethylated products were pooled and condensed *in vacuo* to yield white or yellow solids in 87 – 98%; yielding precursors for products of **2w**-**2y** (**Figure 6**).

To prepare dimeric CSAMs **3a-3f** (**Figure 7**), we began with 150 mg of *O*-methoxy coumarins, which were suspended in 5 mL of DCM in a round-bottom flask, brought down to -78 °C and 1.5 equivalence per oxygen atom of 1.0M BBr_3 in DCM was added dropwise to the stirred solutions. The reactions proceeded for 5h as the temperature rose from -78 °C to room temperature. To quench the reactions, the round-bottom flasks were placed in ice baths to bring the reaction temperature to 0 °C. To the flask, 10 mL of 1:1 water:methanol was added to reactions, effectively demethylating the coumarins. The volatile components were condensed *in vacuo* and were taken up in 20 mL of ethyl acetate and 20 mL of saturated ammonium chloride solution. The solution was transferred to a separatory funnel, the organic layer was extracted and the aqueous phase was washed twice with 10 mL of ethyl acetate. The organic layers were pooled, washed with brine, and dried over anhydrous Na_2SO_4 and then concentrated *in vacuo*. The residues were taken up in minimal DCM (2-3 mL), dry-loaded onto silica and purified by normal phase flash chromatography (0-70% ethyl acetate:hexanes). The fractions containing the deprotected analogs were pooled and concentrated *in vacuo*, yielding 75-96%.

This yielded analogs containing two phenolic hydroxyl groups at the 7-position and the 3-(4'-phenyl)-position. These two phenols were predicted to have different pK_a's and thus by controlling the equivalence of base and electrophilic alkyl group, could selectively alkylate one hydroxyl over the other. These analogs were taken up in 5-10 mL of dry DMF and treated with 0.5-1.0 molar equivalence of Cs_2CO_3. The solution was stirred for 10 minutes and 0.9-1.0 molar equivalence of 1-bromo-3-chloropropane or 1-bromo-4-chlorobutane was added slowly to the solution over a period of 20 minutes. The reaction was allowed to continue stirring overnight (24 hours) at room temperature. Upon complete consumption of the starting material (monitored by TLC), the reaction was quenched with 5 mL of 1.0 M HCl, 5 mL of water, and extracted with 20 mL of ethyl acetate. The aqueous phase was washed twice with 10 mL of ethyl acetate and discarded. The organic layers were pooled, washed with 20 mL brine, dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. The residue was taken up in DCM, dry-loaded onto silica, and purified by normal phase flash chromatography. The separation yielded three major peaks that corresponded to three products; the first peak as the dialkyated species, the second and most prominent peak as the 7-alkylated species, and the third peak was the 3-(4'-hydroxyphenyl) alkylation. The major peak, which contained 66-83% yield of products was chosen to move forward as the one half of the dimer and characterized by ^{1}H-NMR to confirm structure and position of the alkyl group.

Subsequently, the alkylchloride moiety was subjected to an azide substitution reaction to afford the alkylazide handle. To proceed with this, 1.5 equivalence of sodium azide was added in one portion to a stirred solution of the coumarin-alkylchloride in 5 mL

of DMF within a 10 mL microwave tube. The tube was then moved a CEM Discover microwave reactor and reacted for 5 hours at 70 °C. Upon completion, the contents were

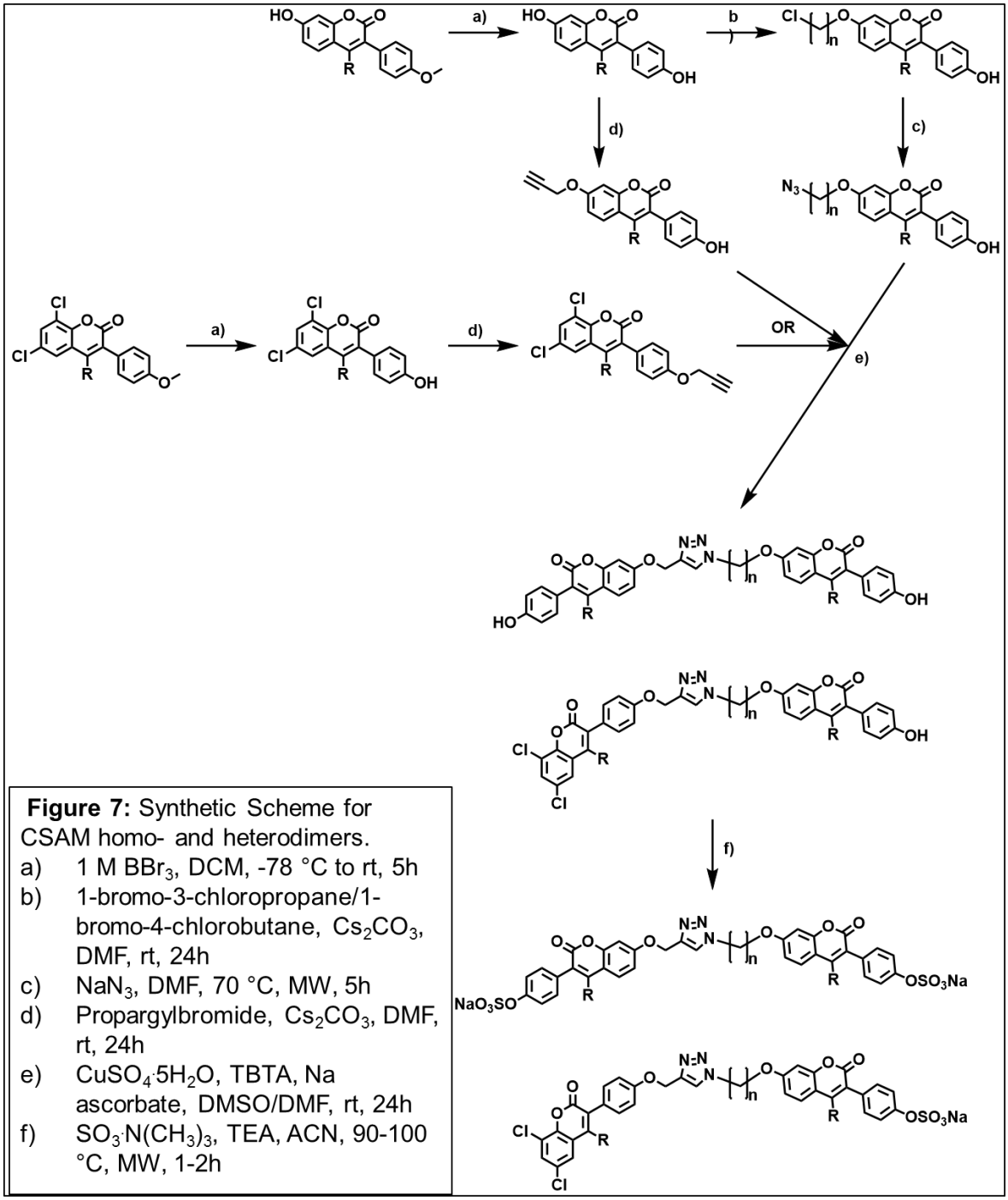


Figure 7: Synthetic Scheme for CSAM homo- and heterodimers.

a) 1 M BBr_3, DCM, -78 °C to rt, 5h
b) 1-bromo-3-chloropropane/1-bromo-4-chlorobutane, Cs_2CO_3, DMF, rt, 24h
c) NaN_3, DMF, 70 °C, MW, 5h
d) Propargylbromide, Cs_2CO_3, DMF, rt, 24h
e) $CuSO_4 \cdot 5H_2O$, TBTA, Na ascorbate, DMSO/DMF, rt, 24h
f) $SO_3 \cdot N(CH_3)_3$, TEA, ACN, 90-100 °C, MW, 1-2h

diluted with 3 washes of 10 mL ethyl acetate and separated with 30 mL of water. The organic layer was washed multiple times with 10 mL of water to remove excess DMF. The organic layer was then washed with brine, dried over anhydrous Na_2SO_4, and concentrated *in vacuo*. The residue was saved without further purification. The confirm the identity of the compound, ^{1}H-NMR showed a distinct up-field shift of the adjacent CH_2 and FT-IR was used to confirm the presence of the azide moiety ~2100 cm^{-1}.

Once the alkylazide was prepared, the alkyne handle needed for the cycloaddition was synthesized. The starting dihydroxy coumarin was transferred into a round bottom flask where 5 mL of dry DMF was added and stirred. To this, 0.5-1.0 molar equivalence of Cs_2CO_3 was added and stirred for 10 minutes. Following, 0.9-1.0 molar equivalence of 80% propargyl bromide solution was slowly added to the stirred solution over the course of 10 minutes. The reaction was allowed to proceed overnight and monitored by TLC until all the starting material was consumed. The reaction was then quenched with 10 mL of 1.0M HCl and diluted with water. 20 mL of ethyl acetate was added and the mixture was transferred to a separatory funnel. The organic layer was separated and the aqueous layer was washed twice with 10 mL of ethyl acetate to remove residual compound. The organic layers were pooled, washed with brine, dried over Na_2SO_4, and concentrated *in vacuo*. The residue was taken up in minimal DCM (2-3 mL) and dry-loaded onto silica gel for normal phase flash chromatography. The products, similarly to the alkylchloride, displayed three major peaks. The major second peak was identified by ^{1}H-NMR to confirm positional conformation of the propyne moiety in 58-89% yields.

Once both the azide and alkyne handle were produced, the coupling reaction could ensue. The coupling employed typical copper(I)-assisted azide alkyne cycloaddition (CuAAC) conditions where the azide and alkyne substituents were each dissolved in 1 mL of DMF mixed into a solution of 5 mL of 4:1 DMSO:DMF. After stirring for 10 minutes, 7 mol% of sodium ascorbate solubilized in water was added to the solution and stirred for an additional 10 minutes. To this, a solution of 3 mol% of 10 mM 1:1 $CuSO_4 \cdot 5H_2O$:tris-[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine in 55% DMSO was added to the stirred solution and allowed to react overnight at room temperature. After the allotted time, the stirred solution was quenched by adding a few drops of 10% NH_4OH (aq) and then flushed with 20 mL of water. A yellow to white precipitate formed and was filtered off. The precipitate was then taken up in 20 mL of ethyl acetate and washed with 5 mL of 1.0M HCl and 15 mL of water. The aqueous phase was discarded and the organic phase was washed with brine and dried over anhydrous Na_2SO_4. The organic phase was concentrated *in vacuo* and dry-loaded onto silica for normal phase flash chromatography. The desired product was isolated using 0-100 % ethyl acetate:hexanes, then 0-10 % methanol:DCM giving 72-94% yields and confirmed by ^{1}H-NMR.

The last synthetic step involved microwave-assisted sulfation. The free hydroxycoumarins (demethylated monomers (**2w-2y**), free hydroxycoumarins (**1a-1d**; **2a-2v**) commercially obtained and synthetic dimers (**3a-3g**)) were taken up in 5 mL of acetonitrile (ACN), placed into a 10 mL microwave tube, and to it 6 molar equivalence of sulfur trioxide trimethylamine complex ($SO_3 \cdot N(CH_3)_3$) per OH was added and stirred. Subsequently, 10 molar equivalence of triethylamine (TEA) per OH was added and the

solution was allowed to stir for 10 minutes. At this point, there was a color change, indicating the hydroxyl groups were deprotonated. The reaction tubes were then placed inside the CEM microwave reactor and was allowed to react at 90-100 °C for 1 hour per OH being sulfated. Upon completion, the tubes and their contents were washed with DCM into a round bottom flask, spotted on TLC to confirm appearance of a sulfated species (10 % methanol in DCM mobile phase; R_f between 15 and 30), and concentrated *in vacuo*. The residue was again taken up in minimal DCM and dry-loaded onto silica gel. Isolation of the sulfated compound utilized normal phase flash chromatography (0-30% methanol:DCM). Fractions containing the sulfated compound were concentrated *in vacuo*. The residue was taken up in minimal water >5 mL (or 3:1 water:methanol) and applied to a sephadex CM 25 sodium exchange column to produce a sodium sulfate salt from the alkylammonium sulfate salt. The fractions containing the compound were pooled, frozen at -80 °C, and lyophilized overnight. This produced a fine, off-white/yellow powder, which was analyzed by UPLC-MS, ^{1}H-NMR, and ^{13}C-NMR.

The synthesis for the second generation CSAMs (**3i** and **3j**) began by charging a 50 mL round-bottom flask with 250 mg of 3-(3',4'-dimethoxyphenyl)-7-methoxy-4-phenylcoumarin in 5 mL of dry DCM. The flask was then put in -78 °C acetone bath. While stirring, 7.5 molar equivalence of 1.0M BBr_3 in DCM was added dropwise to the stirred solution, flushed with nitrogen, and allowed to react to room temperature; over the course of 5 hours. Then, an ice bath replaced the acetone bath to bring the reaction to 0 °C and 10 mL of 1:1 methanol:water was slowly added to quench the reaction and allowed to stir for an additional 10 minutes. Upon completion, the solution was condensed *in vacuo* to

remove the majority of the organics and the slurry that remained was taken up in 20 mL of

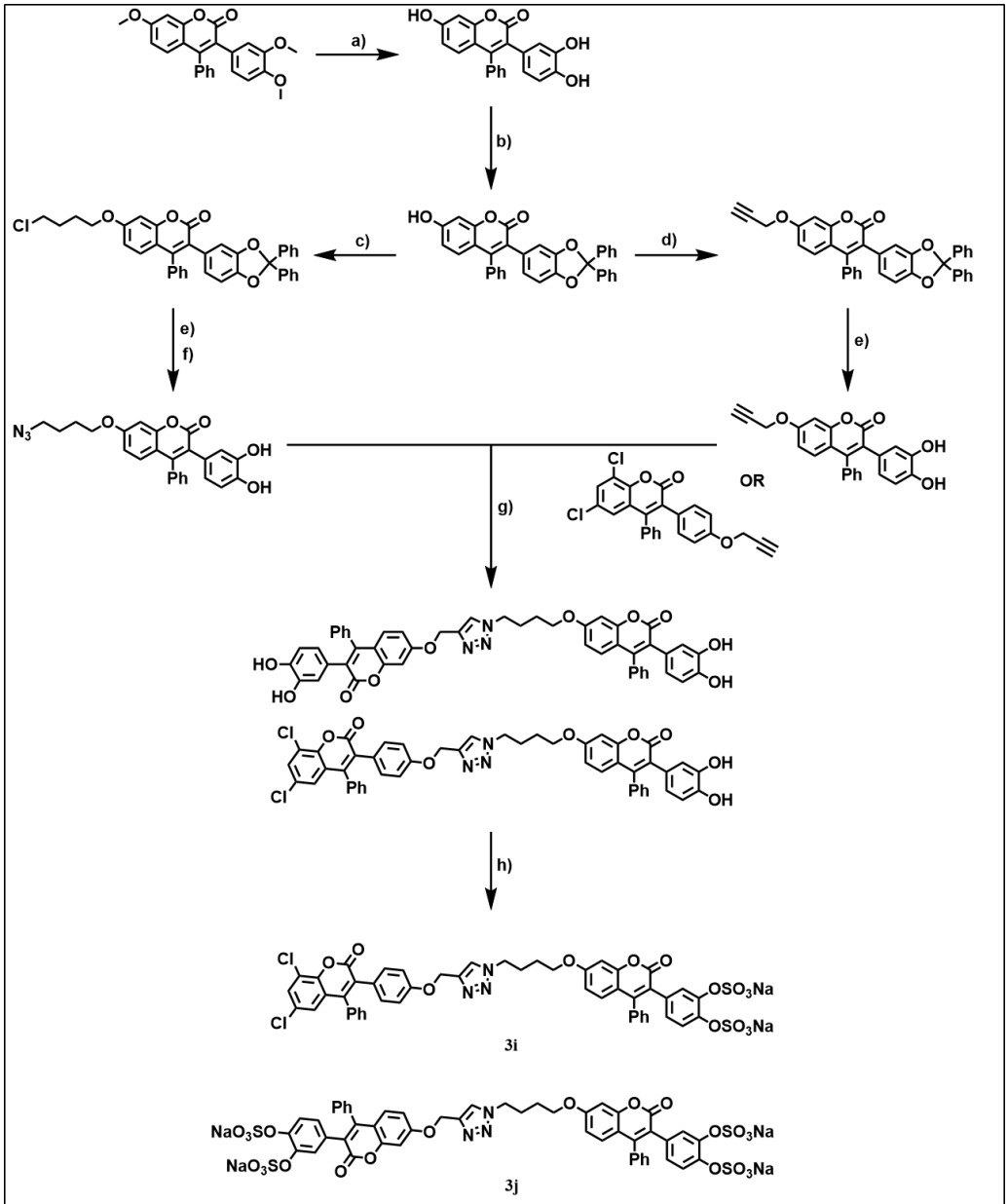


Figure 8: Synthetic Scheme for second generation CSAM homo- and heterodimers.

a) 1 M BBr_3, DCM, -78 °C to rt, 5h
b) Dichlorodiphenylmethane, Ph_2O, reflux, 5h
c) 1-bromo-4-chlorobutane, Cs_2CO_3, DMF, rt, 24h
d) Propargylbromide, Cs_2CO_3, DMF, rt, 24h
e) 80% acetic acid, reflux, 3h
f) NaN_3, DMF, 70 °C, MW, 5h
g) $CuSO_4{\cdot}5H_2O$, TBTA, Na ascorbate, DMSO/DMF, rt, 24h
h) $SO_3{\cdot}N(CH_3)_3$, TEA, ACN, 90-100 °C, MW, 2-4h

ethyl acetate and washed with 15 mL of saturated ammonium chloride solution (aq.). The organic layer was separated and the aqueous phase was washed twice with 10 mL of ethyl acetate. The organic phases were pooled and washed with brine and dried over Na_2SO_4, filtered, and placed in a 250 mL round bottom flask to be condensed *in vacuo*. TLC (35% ethyl acetate:hexanes, R_f = 0.10) showed no starting material and a single spot, so the pale yellow residue obtained after concentration was used without any further purification; yield of 93% (206 mg, **Figure 8**, Step A).

The now 3-(3',4'-dihydroxyphenyl)-7-hydroxy-4-phenyl coumarin (200 mg) was placed in a two-neck round bottom flask and 14 mL of diphenylether was added. To this, 1.5 molar equivalence of 1,1-dichloro-1,1-diphenylmethane was added and the flask was sealed and a reflux condenser was attached. The solution was littered with precipitant until refluxing conditions (175 °C) were reached in an oil bath for 5 hours, after which the homogenous solution cooled down to room temperature and crystals began forming. The crystals were filtered and washed with ice cold hexanes. The remaining crystals were transferred to a screw cap vial and dried *in vacuo* overnight, yielding 72% (219 mg, **Figure 8**, Step B) of a diphenylmethyl (DPM) protected 3-(3',4'-dihydroxyphenyl) moiety.

The next step was to obtain the alkyne moiety for later usage in the cycloaddition. For **3i**, the alkyne handle was the same as **3g**. However for **3j**, the alkyne handle contained the DPM protected 3-(3',4'-dihydroxyphenyl) moiety, which was synthesized by adding 1.2 molar equivalence of 80% propargyl bromide to a solution of 3 mL dry DMF, 103 mg of protected coumarin containing free 7-hydroxy, and 1.6 molar equivalence of Cs_2CO_3. The reaction took place overnight, and upon completion was quenched with 1.0 M HCl

and extracted and washed with ethyl acetate. The organic layer was washed with brine, dried over Na_2SO_4, concentrated, and dry loaded onto silica for flash chromatography (0-60% ethyl acetate:hexanes). The fractions containing the desired product were pooled and concentrated *in vacuo* for further use (76 % yield, 97 mg, **Figure 8**, Step D).

Subsequently, the free hydroxyl group on the 7-position of the DPM 3-(3',4'-dihydroxyphenyl) was also used with the electrophilic alkane chain. Thus, 1.2 molar equivalence of 1-bromo-4-chlorobutane was added to a stirred solution of the protected coumarin (103 mg) and 1.6 equivalence of Cs_2CO_3 in 3 mL dry DMF. The solution was allowed to react overnight and then quenched with 1.0 M HCl. 20 mL of ethyl acetate was added and the biphasic solution was put into a separatory funnel, where the organic layer was extracted. The aqueous layer was washed twice with 10 mL of ethyl acetate, after which the organic layers were pooled, washed with brine, dried over Na_2SO_4, filtered, and concentrated *in vacuo*. The remaining residue was taken up in ~5 mL of DCM and dry loaded onto silica gel. Flash chromatography was performed (0-60% ethyl acetate:hexanes) to isolate the protected, alkylated product. The fractions were pooled, concentrated, and dried under high vacuum overnight (89 % yield, 108 mg, **Figure 8**, Step C).

The DPM protected butyl chloride coumarin and the DPM protected propargyl coumarin were both subjected to catechol deprotection by dissolving 108 mg and 97 mg, respectively in 10 mL of 80% acetic acid and refluxing for 3 hours produced the deprotected products, identified by TLC (70:30 hexanes:ethyl acetate, R_f 0.35 and 0.33 respectively). The reaction mixture was cooled to room temperature and extracted with 25 mL x 3 of ethyl acetate. The pooled organic layer was washed with brine, dried over

Na_2SO_4, filtered, and concentrated *in vacuo*. The residue was taken up in minimal DCM to be dry loaded onto silica gel and was purified by flash chromatography (0-100% ethyl acetate:hexanes). The fractions containing the product were pooled and concentrated *in vacuo*, and placed on high vacuum overnight yielding 88 % (69 mg) and 83 % (56 mg), respectively (**Figure 8**, Step E).

The deprotected butyl chloride coumarin (69 mg) was dissolved in 3 mL of DMF and was reacted with 1.5 molar equivalence of NaN_3. The reaction was heated at 70 °C for 5 hours in a microwave tube. Once completed the reaction was diluted with 20 mL of ethyl acetate and washed several times with 10 mL of water to remove excess DMF. The organic layer was washed with brine, dried over Na_2SO_4, filtered, and concentrated *in vacuo*. No further purification was performed for this compound. Confirmation of the alkylazide (97 % yield, 68 mg) was done by FT-IR showing the N_3 stretch (2100 cm^{-1}) and the up-field shift of the adjacent CH_2 triplet on ^{1}H-NMR (**Figure 8**, Step F).

Like previously used, the copper-assisted azide-alkyne cycloaddition was employed for dimerization of **3i** precursor by dissolving 68 mg of deprotected butylazide in 1 mL of DMF. Then 97 mg of dichloro propargyl moiety of **3g** was dissolved in 1 mL of DMF. The two were mixed together in 3 mL of DMSO and 7 mol% of sodium ascorbate dissolved in water was added to the stirred solution. After 10 minutes, 3 mol% of 1:1 $CuSO_4 \cdot 5H_2O$:TBTA in a 55% DMSO solution was added to the reaction mixture and allowed to stir overnight. Upon completion, a few drops of 10% NH_4OH was added, the solution turned blue and 20 mL of cold water was poured into the reaction tube, where a yellow precipitate formed. After allowed to stir for 10 minutes, the precipitate was filtered

and kept for further use. The precipitate was taken up in 20 mL of ethyl acetate and 15 mL of 1.0M HCl was added. The biphasic solution was separated, the aqueous layer was washed twice with 10 mL of ethyl acetate, and the organic layers were pooled. The organics were then concentrated *in vacuo*, dry loaded onto silica gel, and the desired product was isolated by flash chromatography (0-100% ethyl acetate:hexanes to 0-10% methanol:DCM). The fractions containing the product were pooled and concentrated, yielding a yellow residue (65 % yield, 86 mg, **Figure 8**, Step G).

Likewise for the **3j** precursor, 77 mg (from a secondary batch) of the deprotected butylazide and 56 mg of the deprotected propargyl coumarin were each dissolved in 1 mL of DMF and mixed together in 3 mL of DMSO. Then, 7 mol% of sodium ascorbate was added to the stirred solution, where after 10 minutes, 3 mol % of 1:1 $CuSO_4 \cdot 5H_2O$:TBTA in a 55% DMSO solution was added to the reaction mixture and allowed to stir overnight. Upon completion, a few drops of 10% NH_4OH was added, the solution turned blue and 20 mL of cold water was poured into the reaction tube, where a yellow precipitate formed. After allowed to stir for 10 minutes, the precipitate was filtered and kept for further use. The precipitate was taken up in 20 mL of ethyl acetate and 15 mL of 1.0M HCl was added. The biphasic solution was separated, the aqueous layer was washed twice with 10 mL of ethyl acetate, and the organic layers were pooled. The organics were then concentrated *in vacuo*, dry loaded onto silica gel, and the desired product was isolated by flash chromatography (0-100% ethyl acetate:hexanes to 0-10% methanol:DCM). The fractions containing the product were pooled and concentrated, yielding an off-white residue (98 % yield, 125 mg, **Figure 8**, Step G).

The deprotected, coupled 3-(3',4'-dihydroxylphenyl) coumarins were then sulfated using the same method mentioned above. The coumarins 86 mg (**3i** precursor) and 125 mg (**3j** precursor), respectively, were taken up in 3 mL of dry acetonitrile in a microwave tube with a stir bar. To that was added 12-24 molar equivalence of $SO_3 \cdot N(CH_3)_3$ and allowed to stir for several minutes, after which 20-40 molar equivalence of triethylamine was added. The solution stirred for another 10 minutes and the reaction tube was placed in a microwave and heated to 100 °C for 2 and 4 hours, respectively. Upon completion, the mixture was diluted into a round bottom flask with DCM and concentrated *in vacuo*. The

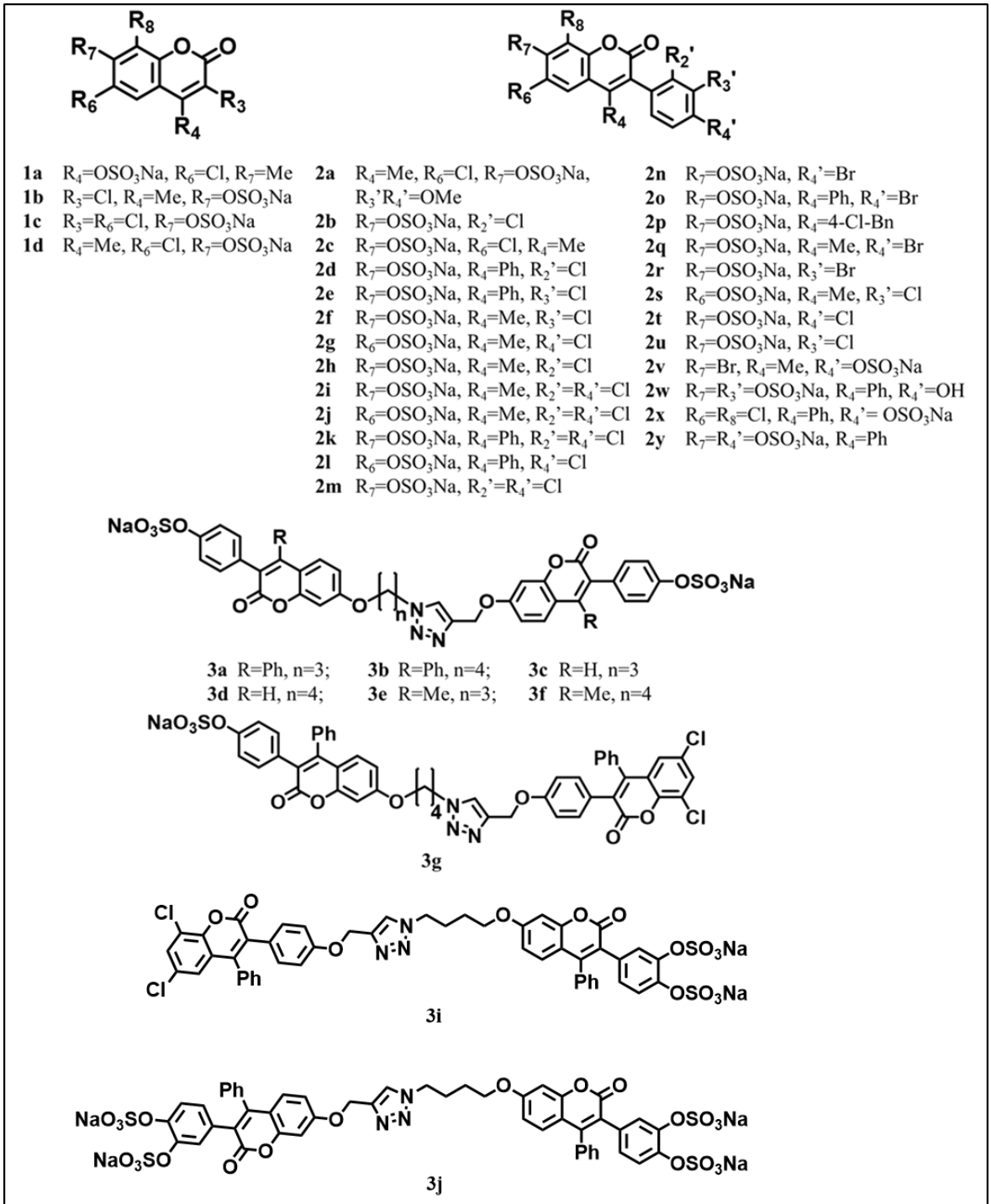


Figure 9: First and second generation CSAMs synthesized inclusive of monomer, homo- and heterodimers

residue was then taken up in 5 mL of DCM and dry loaded onto silica gel for separation on

normal phase flash chromatography (0-30% methanol:DCM). The fractions containing the sulfated compound were pooled, concentrated *in vacuo*, and taken up in minimal water. The solution was then applied to a sephadex CM 25 column charged with sodium acetate, for sodium exchange of the amine salt. The fractions collected containing the sodium salt were pooled, frozen in a -80°C refrigerator, and lyophilized. This resulted in a light yellow powder (97 %, 103 mg, **3i**) and an off-white powder (62%, 116 mg, **3j**), respectively, which was stored at -20 °C for future use.

2.2 CSAM Characterization

Characterization of the coumarin-based sulfated allosteric modulators utilized multiple techniques. ^{1}H-NMR and ^{13}C NMR characterizations was performed with a 400 MHz Bruker NMR, IR measurements were done using a Varian FTIR, and MS data was obtained from a Waters Triple Quad Mass Spectrometer.

3a1/3b1/3g1 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.36 (m, 3H), 7.17 (dd, J = 8 Hz, 1.8 Hz, 2H), 6.91 (d, J = 8 Hz, 2H), 6.88 (d, J = 8 Hz, 1H), 6.82 (d, J = 4 Hz, 1H), 6.74 (dd, J = 8 Hz, 2.36 Hz, 1H), 6.54 (d, J = 8 Hz, 2H).

3c1/3d1 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.81 (s, 1H), 7.50 (d, J = 8 Hz, 2H), 7.44 (d, J = 8 Hz, 1H), 6.77 (d, J = 8 Hz, 2H), 6.74 (dd, J = 8 Hz, 2.32 Hz, 1H), 6.65 (d, J = 8 Hz, 1H).

3e1/3f1 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.71 (d, J = 8 Hz, 1H), 7.21 (d, J = 8 Hz, 2H), 6.95 (d, J = 8 Hz, 2H), 6.91 (d, J = 8 Hz, 1H), 6.81 (d, J = 8 Hz, 1H), 2.33 (s, 3H).

3a2 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.23 (m, 3H), 7.10 (dd, *J* = 8 Hz, 2.48 Hz, 2H), 6.93 (m, 2H), 6.89 (d, *J* = 8 Hz, 2H), 6.80 (dd, *J* = 8 Hz, 2.54 Hz, 1H), 6.52 (d, *J* = 8 Hz, 2H), 4.28 (t, *J* = 8 Hz, 2H), 4.02 (t, *J* = 8 Hz, 2H), 2.21 (quint, *J* = 8 Hz, 2H).

3b2/3g2 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.24 (m, 3H), 7.09 (dd, *J* = 8 Hz, 1.84 Hz, 2H), 6.91 (m, 3H), 6.73 (dd, *J* = 8 Hz, 2.48 Hz, 1H), 6.50 (d, *J* = 8 Hz, 2H) 4.18 (t, *J* = 8 Hz, 2H), 3.96 (t, *J* = 8 Hz, 2H), 1.81 (quint, *J* = 8 Hz, 2H), 1.72 (quint, *J* = 8 Hz, 2H).

3c2 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.98 (s, 1H), 7.69 (d, *J* = 8 Hz, 1H), 7.61 (d, *J* = 8 Hz, 2H), 7.05 (s, 1H), 7.01 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.88 (d, *J* = 8 Hz, 2H), 4.32 (t, *J* = 8 Hz, 2H), 4.08 (t, *J* = 8 Hz, 2H), 2.34 (quint, *J* = 8 Hz, 2H).

3d2 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.96 (s, 1H), 7.67 (d, *J* = 8 Hz, 1H), 7.58 (d, *J* = 8 Hz, 2H), 7.02 (s, 1H), 6.99 (dd, *J* = 8 Hz, 2.36 Hz, 1H), 6.84 (d, *J* = 8 Hz, 2H), 4.15 (t, *J* = 8 Hz, 2H), 3.92 (t, *J* = 8 Hz, 2H), 1.86 (quint, *J* = 8 Hz, 2H), 1.75 (quint, *J* = 8 Hz, 2H).

3e2 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.61 (d, *J* = 8 Hz, 1H), 7.05 (d, *J* = 8 Hz, 2H), 6.87 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.80 (d, *J* = 4 Hz, 1H), 6.79 (d, *J* = 8 Hz, 2H), 4.19 (t, *J* = 8 Hz, 2H), 3.98 (t, *J* = 8 Hz, 2H), 2.31 (s, 3H), 2.21 (quint, *J* = 8 Hz, 2H), 2.18 (s, 3H).

3f2 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.73 (d, *J* = 8 Hz, 1H), 7.20 (d, *J* = 8 Hz, 2H), 6.98 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.93 (d, *J* = 8 Hz, 2H), 6.91 (d, *J* = 8 Hz, 1H), 4.22 (t, *J* = 8 Hz, 2H), 4.01 (t, *J* = 8 Hz, 2H), 2.31 (s, 3H), 1.98 (quint, *J* = 8 Hz, 2H), 1.87 (quint, *J* = 8 Hz, 2H).

3a3 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.23 (m, 3H), 7.10 (dd, *J* = 8 Hz, 2.48 Hz, 2H), 6.93 (m, 2H), 6.89 (d, *J* = 8 Hz, 2H), 6.80 (dd, *J* = 8 Hz, 2.54 Hz, 1H), 6.52 (d, *J* = 8 Hz,

2H), 4.28 (t, *J* = 8 Hz, 2H), 3.78 (t, *J* = 8 Hz, 2H), 2.21 (quint, *J* = 8 Hz, 2H). IR (neat): 2102 cm^{-1}.

3b3/3g3 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.24 (m, 3H), 7.09 (dd, *J* = 8 Hz, 1.84 Hz, 2H), 6.91 (m, 3H), 6.73 (dd, *J* = 8 Hz, 2.48 Hz, 1H), 6.50 (d, *J* = 8 Hz, 2H) 4.08 (t, *J* = 8 Hz, 2H), 3.34 (t, *J* = 8 Hz, 2H), 1.81 (quint, *J* = 8 Hz, 2H), 1.72 (quint, *J* = 8 Hz, 2H). IR (neat): 2098 cm^{-1}.

3c3 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.98 (s, 1H), 7.69 (d, *J* = 8 Hz, 1H), 7.61 (d, *J* = 8 Hz, 2H), 7.05 (s, 1H), 7.01 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.88 (d, *J* = 8 Hz, 2H), 4.32 (t, *J* = 8 Hz, 2H), 3.81 (t, *J* = 8 Hz, 2H), 2.34 (quint, *J* = 8 Hz, 2H). IR (neat): 2119 cm^{-1}.

3d3 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.96 (s, 1H), 7.67 (d, *J* = 8 Hz, 1H), 7.58 (d, *J* = 8 Hz, 2H), 7.02 (s, 1H), 6.99 (dd, *J* = 8 Hz, 2.36 Hz, 1H), 6.84 (d, *J* = 8 Hz, 2H), 4.15 (t, *J* = 8 Hz, 2H), 3.45 (t, *J* = 8 Hz, 2H), 1.86 (quint, *J* = 8 Hz, 2H), 1.75 (quint, *J* = 8 Hz, 2H). IR (neat): 2107 cm^{-1}.

3e3 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.61 (d, *J* = 8 Hz, 1H), 7.05 (d, *J* = 8 Hz, 2H), 6.87 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.80 (d, *J* = 4 Hz, 1H), 6.79 (d, *J* = 8 Hz, 2H), 4.19 (t, *J* = 8 Hz, 2H), 3.73 (t, *J* = 8 Hz, 2H), 2.31 (s, 3H), 2.21 (quint, *J* = 8 Hz, 2H), 2.18 (s, 3H). IR (neat): 2112 cm^{-1}.

3f3 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.73 (d, *J* = 8 Hz, 1H), 7.20 (d, *J* = 8 Hz, 2H), 6.98 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.93 (d, *J* = 8 Hz, 2H), 6.91 (d, *J* = 8 Hz, 1H), 4.22 (t, *J* = 8 Hz, 2H), 3.50 (t, *J* = 8 Hz, 2H), 2.31 (s, 3H), 1.98 (quint, *J* = 8 Hz, 2H), 1.87 (quint, *J* = 8 Hz, 2H). IR (neat): 2130 cm^{-1}.

3a4/3b4 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.23 (m, 3H), 7.10 (m, 2H), 6.95 (m, 2H), 6.88 (d, *J* = 8 Hz, 2H), 6.78 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.51 (d, *J* = 8 Hz, 2H), 4.81 (s, 2H), 3.03 (s, 1H).

3c4/3d4 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.86 (s, 1H), 7.54 (m, 3H), 6.88 (d, *J* = 8 Hz, 2H), 6.78 (d, *J* = 8 Hz, 2H), 4.82 (s, 2H), 3.04 (s, 1H).

3e4/3f4 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.61 (d, *J* = 8 Hz, 1H), 7.05 (d, *J* = 8 Hz, 2H), 6.89 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.80 (d *J* = 8 Hz, 1H), 6.78 (d, *J* = 8 Hz, 2H), 4.82 (s, 2H), 3.04 (s, 1H), 2.30 (s, 3H).

3g4 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.64 (d, *J* = 4 Hz, 1H), 7.29 (m, 3H), 7.17 (dd, *J* = 8 Hz, 1.92 Hz, 2H), 7.02 (d, *J* = 8 Hz, 2H), 6.90 (d, *J* = 2.44 Hz, 1H), 6.65 (d, *J* = 8 Hz, 2H), 4.79 (s, 2H), 3.09 (s, 1H).

3a5 – ^{1}H NMR (Acetone-d_6 400 MHz): 8.37 (s, 1H), 7.39 (m, 6H), 7.29 (s, 1H), 7.22 (m, 4H), 7.11 (s, 1H), 7.04 (m, 4H), 6.97 (m, 7H), 6.88 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 5.29 (s, 2H), 4.61 (t, 2H), 4.16 (t, 2H), 2.35 (quint, *J* = 8 Hz, 2H).

3b5 – ^{1}H NMR (Acetone-d_6 400 MHz): 8.35 (s, 1H), 7.37 (m, 6H), 7.26 (s, 1H), 7.15 (m, 4H), 7.08 (s, 1H), 7.04 (m, 4H), 6.91 (m, 4H), 6.87 (s, 1H), 6.54 (dd, *J* = 8 Hz, 2.36 Hz, 1H), 5.29 (s, 2H), 4.50 (t, *J* = 8 Hz, 2H), 4.15 (t, *J* = 8 Hz, 2H), 2.05 (quint, *J* = 8 Hz, 2H), 1.78 (quint, *J* = 8 Hz, 2H).

3c5 – ^{1}H NMR (Acetone-d_6 400 MHz): 8.35 (s, 1H), 8.08 (s, 2H), 7.68 (m, 2H), 7.58 (d, *J* = 8 Hz, 4H), 7.18 (s, 1H), 7.05 (m, 3H), 6.84 (d, *J* = 8 Hz, 2H), 5.29 (s, 2H), 4.61 (t, *J* = 8 Hz, 2H), 4.14 (t, *J* = 8 Hz, 2H) 2.38 (quint, *J* = 8 Hz, 2H).

3d5 – ^{1}H NMR (Acetone-d_6 400 MHz): 8.33 (s, 1H), 8.07 (s, 2H), 7.68 (m, 2H), 7.58 (d, *J* = 8 Hz, 4H), 7.18 (s, 1H), 7.04 (m, 2H), 6.85 (d, *J* = 8 Hz, 4H), 5.29 (s, 2H), 4.50 (t, *J* = 8 Hz, 2H), 4.13 (t, *J* = 8 Hz, 2H), 2.06 (quint, *J* = 8 Hz, 2H), 1.77 (quint, *J* = 8 Hz, 2H).

3e5 – ^{1}H NMR (Acetone-d_6 400 MHz): 8.37 (s, 1H), 8.00 (s, 2H), 7.57 (t, *J* = 8 Hz, 2H), 7.04 (s, 1H), 7.02 (d, *J* = 8 Hz, 4H), 6.92 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.89 (m, 2H), 6.74 (d, *J* = 8 Hz, 4H), 5.21 (s, 2H), 4.58 (t, *J* = 8 Hz, 2H), 4.08 (t, *J* = 8 Hz, 2H), 2.35 (quint, *J* = 8 Hz, 2H), 2.16 (s, 6H).

3f5 – ^{1}H NMR (Acetone-d_6 400 MHz): 8.33 (s, 1H), 7.96 (s, 2H), 7.75 (t, *J* = 8 Hz, 2H), 7.17 (s, 1H), 7.11 (d, *J* = 8 Hz, 4H), 7.07 (dd, *J* = 8 Hz, 2.52 Hz, 1H), 6.99 (m, 2H), 6.84 (d, *J* = 8 Hz, 4H), 5.30 (s, 2H), 4.51 (t, *J* = 8 Hz, 2H), 4.14 (t, *J* = 8 Hz, 2H), 2.25 (s, 6H), 2.04 (quint, *J* = 8 Hz, 2H), 1.77 (quint, *J* = 8 Hz, 2H).

3g5 – ^{1}H NMR (Acetone-d_6 400 MHz): 8.24 (s, 1H), 8.03 (s, 1H), 7.42 (m, 6H), 7.23 (m, 4H), 7.09 (m, 3H), 7.01 (m, 4H), 6.90 (m, 5H), 5.27 (s, 2H), 4.47 (t, *J* = 8 Hz, 2H), 4.13 (t, *J* = 8 Hz, 2H), 1.99 (quint, *J* = 8 Hz, 2H), 1.70 (quint, *J* = 8 Hz, 2H).

3i1/3j1 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.24 (m, 3H), 7.09 (dd, *J* = 8 Hz, 1.8 Hz, 2H), 6.84 (d, *J* = 8 Hz, 1H), 6.69 (s, 1H), 6.63 (dd, *J* = 8 Hz, 1.8 Hz, 1H), 6.58 (s, 1H), 6.45 (d, *J* = 8 Hz, 1H), 6.32 (dd, *J* = 8 Hz, 2 Hz, 1H).

3i2/3j2 – ^{1}H NMR (DMSO-d_6 400 MHz): 7.49 (m, 11H), 7.33 (m, 3H), 7.20 (m, 2H), 6.90 (d, *J* = 8 Hz, 1H), 6.83 (m, 3H), 6.75 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 6.61 (dd, *J* = 8H, 2.4 Hz, 1H).

3i3/3j3 – ^{1}H NMR ($CDCl_3$-d_1 400 MHz): 7.42 (m, 4H), 7.31 (m, 6H), 7.22 (m, 3H), 7.03 (3H), 6.82 (s, 1H), 6.66 (m, 4H), 3.99 (t, *J* = 8Hz, 2H), 3.55 (t, *J* = 8 Hz, 2H), 1.97 (m, 4H).

3j4 – ^{1}H NMR (DMSO-d_6 400 MHz): 7.49 (m, 11H), 7.33 (m, 3H), 7.20 (m, 2H), 6.90 (d, *J* = 8 Hz, 1H), 6.83 (m, 3H), 6.75 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 6.61 (dd, *J* = 8H, 2.4 Hz, 1H), 5.04 (s, 2H), 3.59 (s, 1H).

3i5/3j5 – ^{1}H NMR ($CDCl_3$-d_1 400 MHz): 7.33 (m, 3H), 7.14 (m, 3H), 6.93 (s, 1H), 6.79 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 6.73 (s 1H), 6.57 (d, *J* = 8 Hz, 1H), 6.39 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 4.16 (t, *J* = 8 Hz, 2H), 3.67 (t, *J* = 8 Hz, 2H), 2.03 (m, 4H).

3i6/3j6 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.39 (m, 3H), 7.24 (m, 2H), 7.04 (d, *J* = 8Hz, 1H), 6.98 (s, 1H), 6.86 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 6.74 (s, 1H), 6.61 (d, *J* = 8 Hz, 1H), 6.47 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 4.22 (t, *J* = 8 Hz, 2H), 3.49 (t, *J* = 8 Hz, 2H), 1.98 (quint, *J* = 8 Hz, 2H), 1.86 (quint, *J* = 8 Hz, 2H). IR (neat): 2109 cm^{-1} and 2115 cm^{-1}.

3i7 – ^{1}H NMR (DMSO-d_6 400 MHz): 8.16 (s, 1H), 7.93 (s, 1 H), 7.34 (m, 8H), 7.16 (m, 2H), 7.10 (m, 2H), 7.02 (m, 2H), 6.84 (m, 5H,), 6.54 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 5.00 (s, 2H), 4.40 (t, *J* = 8 Hz, 2H), 4.06 (t, *J* = 8 Hz, 2H), 1.95 (quint, *J* = 8 Hz, 2H), 1.67 (quint, *J* = 8 Hz, 2H).

3j7 – ^{1}H NMR (Acetone-d_6 400 MHz): 7.91 (s, 1H), 7.64 (s, 1H), 7.30 (m, 8H), 7.08 (m, 4H), 6.89 (s, 1H), 6.86 (s, 1H), 6.81 (s, 1H), 6.75 (d, *J* = 8 Hz, 2H), 6.70 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 6.59 (s, 1H), 6.46 (d, *J* = 8 Hz, 1H), 6.32 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 4.99 (s, 2H), 4.42 (t, *J* = 8 Hz, 2H), 4.06 (t, *J* = 8 Hz, 2H), 2.04 (quint, *J* = 8 Hz, 2H), 1.76 (quint, *J* = 8 Hz, 2H).

1a – Sodium 6-chloro-7-methyl-2-oxo-2*H*-chromen-4-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.72 (s, 1H), 7.63 (s, 1H), 6.28 (s, 1H), 2.76 (s, 3H). ^{13}C NMR (DMSO-d_6, 100 MHz): 161.35, 159.75, 151.46, 128.79, 122.15, 118.89, 114.85, 96.21, 52.79, 19.86. MS (ESI) calculated for $C_{10}H_6ClNaO_6S$ $[(M-Na)]^-$, *m/z* 289.6519, found $[(M-Na)]^-$, *m/z* 289.43.

1b – Sodium 3-chloro-4-methyl-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.81-7.79 (d, *J* = 8 Hz, 1H), 7.25-7.23 (d, *J* = 8 Hz, 2H), 2.57 (s, 3H). ^{13}C NMR (DMSO-d_6, 100 MHz): 156.74, 156.37, 151.67, 148.71, 126.38, 117.11, 116.78, 114.36, 106.58, 16.08. MS (ESI) calculated for $C_{10}H_6ClNaO_6S$ $[(M-Na)]^-$, *m/z* 289.6518, found $[(M-Na)]^-$, *m/z* 289.39.

1c – Sodium 6-chloro-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.99-7.97 (d, *J* = 8 Hz, 1H), 7.87 (s, 1H), 7.59 (s, 1H), 6.43-6.41 (d, *J* = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 159.71, 152.74, 151.98, 143.13, 128.31, 119.64, 114.68, 114.41, 107.76. MS (ESI) calculated for $C_9H_4ClNaO_6S$ $[(M-Na)]^-$, *m/z* 275.6248, found $[(M-Na)]^-$, *m/z* 275.37.

1d – **Sodium 6-chloro-4-methyl-2-oxo-2*H*-chromen-7-yl sulfate–** ^{1}H NMR (DMSO-d_6 400 MHz): 7.84 (s, 1H), 7.59 (s, 1H), 6.32 (s, 1H), 2.43 (s, 3H). ^{13}C NMR (DMSO-d_6, 100 MHz): 159.62,152.52, 152.11, 151.97, 125.60, 119.66, 115.31, 112.92, 107.73, 18.04. MS (ESI) calculated for $C_{10}H_6ClO_6S$ $[(M-Na)]^-$, *m/z* 289.6620, found $[(M-Na)]^-$, *m/z* 289.39.

2a – Sodium 6-chloro-3-(3,4-dimethoxyphenyl)-4-methyl-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.81 (s, 1H), 7.53 (s, 1H), 6.97-6.94 (d, *J* = 12 Hz 1H), 6.85 (s, 1H), 6.78-6.75 (d, J = 12 Hz, 1H), 3.73 (s, 3H), 3.67 (s, 3H), 2.20 (s, 3H). ^{13}C NMR (DMSO-d_6, 100M Hz): 159.80, 151.09, 148.32, 145.09, 126.76, 125.96, 124.84,

122.68, 119.67, 115.87, 114.05, 111.49, 107.46, 55.63, 55.53, 16.47. MS (ESI) calculated for $C_{18}H_{14}ClNaO_8S$ $[(M-Na)]^-$, *m/z* 425.8018, found $[(M-Na)]^-$, *m/z* 425.52.

2b – Sodium 3-(2-chlorophenyl)-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 8.00 (s, 1H), 7.63-7.61 (d, *J* = 8 Hz, 1H), 7.51-7.37 (m, 4H), 7.22 (s, 1H), 7.14-7.12 (d, *J* = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100M Hz): 159.00, 157.05, 154.21, 142.96, 134.26, 132.86, 131.82, 130.14, 129.26, 129.11, 127.19, 123.65, 116.82, 113.92. MS (ESI) calculated for $C_{15}H_8ClNaO_6S$ $[(M-Na)]^-$, *m/z* 351.7228, found $[(M-Na)]^-$, *m/z* 351.42.

2c – Sodium 6-chloro-4-methyl-2-oxo-3-phenyl-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.83 (s, 1H), 7.55 (s, 1H), 7.41-7.32 (m, 3H), 7.26-7.23 (d, *J* = 8 Hz, 2H), 2.17 (s, 3H). ^{13}C NMR (DMSO-d_6, 100M Hz): 159.71, 151.75, 151.18, 147.13,134.46, 130.19,130.12, 128.09, 127.87, 126.06, 124.92, 119.76, 115.77, 107.53, 16.39. MS (ESI) calculated for $C_{16}H_{10}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 365.7498, found $[(M-Na)]^-$, *m/z* 365.48.

2d – Sodium 3-(2-chlorophenyl)-2-oxo-4-phenyl-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.38-7.11 (m, 12H), 7.00-6.98 (d, *J* = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100M Hz): 161.49, 159.42, 156.92, 154.68, 153.58, 133.80, 133.54, 132.65, 132.51, 129.58, 128.82, 128.64, 128.49, 128.30, 127.99, 127.20, 126.62, 116.87, 113.42, 106.96, 102.33. MS (ESI) calculated for $C_{21}H_{12}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 427.8208, found $[(M-Na)]^-$, *m/z* 427.53.

2e – Sodium 3-(3-chlorophenyl)-2-oxo-4-phenyl-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.32-7.25 (m, 4H), 7.18-7.11 (m, 5H), 7.03-7.01 (dd, *J* = 8 Hz, 2H), 6.91-6.89 (d, *J* = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100M Hz): 177.67, 160.22,

156.83, 153.39, 151.66, 151.23, 136.60, 134.13, 132.00, 130.39, 129.37, 129.22, 128.96, 128.38, 128.23, 127.89, 127.19, 122.85, 116.62, 115.08, 106.80. MS (ESI) calculated for $C_{21}H_{12}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 427.8208, found $[(M-Na)]^-$, *m/z* 427.54.

2f – Sodium 3-(3-chlorophenyl)-4-methyl-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.80-7.78 (d, *J* = 8 Hz, 1H), 7.52-7.43 (m, 3H), 7.31-7.29 (m, 1H), 7.26-7.21 (m, 2H), 2.26 (s, 3H). ^{13}C NMR (DMSO-d_6, 100M Hz): ^{13}C NMR (DMSO-d_6, 100M Hz): 159.89, 156.59, 152.83, 148.59, 139.05, 136.87, 132.71, 130.00, 129.06, 127.79, 126.50, 122.71, 116.40, 114.96, 106.48, 16.42. MS (ESI) calculated for $C_{16}H_{10}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 365.7498, found $[(M-Na)]^-$, *m/z* 365.48.

2g – Sodium 3-(4-chlorophenyl)-4-methyl-2-oxo-2*H*-chromen-6-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.52-7.51 (d, *J* = 4 Hz, 1H), 7.47-7.45 (d, *J* = 8 Hz, 2H), 7.42-7.39 (dd, *J* = 8 Hz, 1H), 7.32-7.29 (d, *J* = 8 Hz, 3H), 2.16 (s, 3H). ^{13}C NMR (DMSO-d_6, 100M Hz): 159.79, 149.83, 148.02, 147.97, 133.49, 132.69, 132.31, 132.09, 128.16, 125.19, 124.77, 120.06, 116.74, 116.63, 116.57, 16.44. MS (ESI) calculated for $C_{16}H_{10}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 365.7498, found $[(M-Na)]^-$, *m/z* 365.49.

2h – Sodium 3-(2-chlorophenyl)-4-methyl-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.74-7.72 (d, *J* = 8 Hz, 1H), 7.54-7.51 (m, 1H), 7.41-7.32 (m, 3H), 7.20-7.15 (m, 2H), 2.09 (s, 3H). ^{13}C NMR (DMSO-d_6, 100M Hz): 159.10, 156.75, 153.02, 149.62, 133.63, 133.29, 132.04, 130.03, 129.23, 127.34, 126.51, 121.85, 116.55, 114.62, 106.64, 16.05. MS (ESI) calculated for $C_{16}H_{10}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 365.7498, found $[(M-Na)]^-$, *m/z* 365.49.

2i – Sodium 3-(2,4-dichlorophenyl)-4-methyl-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.72 (s, 1H), 7.54-7.33 (m, 5H), 2.09 (s, 3H). ^{13}C NMR (DMSO-d_6, 100 MHz): 158.97, 156.96, 153.08, 150.10, 134.48, 133.76, 133.43, 132.75, 128.83, 127.60, 126.57, 120.77, 116.58, 114.49, 106.63, 16.07. MS (ESI) calculated for $C_{16}H_9Cl_2NaO_6S$ $[(M-Na)]^-$, *m/z* 400.1918, found $[(M-Na)]^-$, *m/z* 399.46.

2j – Sodium 3-(2,4-dichlorophenyl)-4-methyl-2-oxo-2*H*-chromen-6-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.72 (s, 1H), 7.54-7.33 (m, 5H), 2.08 (s, 3H). ^{13}C NMR (DMSO-d_6, 100 MHz): 158.79, 154.01, 150.00, 149.72, 148.25, 134.28, 133.88, 133.24, 132.65, 128.85, 127.64, 125.32, 123.17, 119.53, 116.81, 16.10. MS (ESI) calculated for $C_{16}H_9Cl_2NaO_6S$ $[(M-Na)]^-$, *m/z* 400.1918, found $[(M-Na)]^-$, *m/z* 399.47.

2k – Sodium 3-(2,4-dichlorophenyl)-2-oxo-4-phenyl-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.32–7.25 (m, 4H), 7.18–7.12 (m, 5H), 7.03–7.01 (m, 2H), 6.91-6.89 (d, J = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 177.67, 170.10, 160.22, 156.83, 153.39, 151.66, 151.23, 136.60, 134.13, 132.00, 130.39, 129.37, 129.22, 128.96, 128.23, 127.89, 127.19, 122.85, 116.62, 115.08, 106.80. MS (ESI) calculated for $C_{21}H_{11}Cl_2NaO_6S$ $[(M-Na)]^-$, *m/z* 462.2628, found $[(M-Na)]^-$, *m/z* 461.07.

2l – Sodium 3-(4-chlorophenyl)-2-oxo-4-phenyl-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 7.31-7.26 (m, 4H), 7.17-7.08 (m, 7H), 7.03-7.00 (d, J = 12 Hz, 1H), 6.91-6.88 (d, J = 12 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 161.23, 160.47, 159.94, 154.41, 151.94, 134.40, 133.53, 132.55, 132.52, 132.47, 131.78, 128.97, 128.76, 128.28, 128.22, 127.48, 127.43, 121.15, 113.27, 112.31, 102.20. MS (ESI) calculated for $C_{21}H_{12}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 427.8208, found $[(M-Na)]^-$, *m/z* 427.56.

2m – Sodium 3-(2,4-dichlorophenyl)-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 8.02 (s, 1H), 7.68 (s, 1H), 7.63-7.61 (d, J = 8 Hz, 1H), 7.48 (s, 1H), 7.47 (s, 1H), 7.22 (s, 1H), 7.15-7.12 (d, J = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 158.88, 157.14, 154.25, 143.36, 133.96, 133.91, 133.25, 133.14, 129.26, 128.81, 127.42, 122.53, 116.86, 113.84, 106.55. MS (ESI) calculated for $C_{15}H_7Cl_2NaO_6S$ [(M-Na)]$^-$, *m/z* 386.1648, found [(M-Na)]$^-$, *m/z* 385.44.

2n – Sodium 3-(4-bromophenyl)-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6, 400 MHz): 8.20 (s, 1H), 7.65-7.57 (m, 5H), 7.19 (s, 1H), 7.13-7.10 (d, J = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 159.66, 156.87, 153.86, 140.92, 134.08, 131.12, 130.43, 129.15, 122.95, 121.57, 116.76, 114.50, 106.28, 104.80, 99.49. MS (ESI) calculated for $C_{15}H_8BrNaO_6S$ [(M-Na)]$^-$, *m/z* 396.1768, found [(M-Na)]$^-$, *m/z* 397.38.

2o – Sodium 3-(4-bromophenyl)-2-oxo-4-phenyl-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.40-7.36 (m, 6H), 7.23-7.21 (d, J = 8 Hz, 2H), 7.12-7.08 (m, 3H), 6.98-6.96 (d, J = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 179.59, 160.26, 156.73, 153.33, 151.40, 144.49, 134.22, 133.79, 133.00, 132.79, 130.40, 129.00, 128.37, 128.27, 127.84, 123.08, 120.61, 116.60, 115.13, 106.79, 90.08. MS (ESI) calculated for $C_{21}H_{12}BrNaO_6S$ [(M-Na)]$^-$, *m/z* 472.2748, found [(M-Na)]$^-$, *m/z* 471.56.

2p – Sodium 4-(4-chlorobenzyl)-2-oxo-3-phenyl-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.42-7.39 (d, J = 12 Hz, 1H), 7.36-7.28 (m, 3H), 7.26-7.18 (m, 5H), 7.11-7.09 (d, J = 8 Hz, 2H), 6.99-6.96 (dd, J = 10, 4 Hz, 1H), 3.97 (s, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz): 173.47, 160.75, 160.43, 154.47, 148.44, 137.10, 134.73, 134.62, 130.94, 129.78, 129.71, 128.49, 128.20, 127.85, 127.80, 124.36, 113.05, 111.10, 102.16,

99.38, 44.22, 33.95. MS (ESI) calculated for $C_{22}H_{14}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 441.8478, found $[(M-Na)]^-$, *m/z* 441.63.

2q – Sodium 3-(4-bromophenyl)-4-methyl-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.80-7.78 (d, *J* = 8 Hz, 1H), 7.67-7.65 (d, *J* = 8 Hz, 2H), 7.32-7.30 (d, *J* = 8 Hz, 2H), 7.25-7.22 (m, 2H), 2.26 (s, 3H). ^{13}C NMR (DMSO-d_6, 100 MHz): 159.89, 156.58, 155.99, 152.81, 148.33, 134.00, 132.52, 131.06, 126.44, 122.89, 121.16, 116.39, 115.01, 106.48, 101.41, 16.40. MS (ESI) calculated for $C_{16}H_{10}BrNaO_6S$ $[(M-Na)]^-$, *m/z* 410.2038, found $[(M-Na)]^-$, *m/z* 409.45.

2r – Sodium 3-(3-bromophenyl)-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.31 (s, 1H), 7.96 (s, 1H), 7.77-7.75 (d, *J* = 8 Hz, 1H), 7.72-7.70 (d, *J* = 8 Hz, 1H), 7.62-7.60 (d, *J* = 8 Hz, 1H), 7.46-7.42 (t, *J* = 8 Hz, 1H), 7.27 (s, 1H), 7.21-7.19 (d, *J* = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 159.65, 157.05, 153.95, 141.49, 137.70, 137.21, 130.93, 130.88, 130.30, 129.25, 127.40, 122.53, 121.41, 116.79, 114.43, 106.28. MS (ESI) calculated for $C_{15}H_8BrNaO_6S$ $[(M-Na)]^-$, *m/z* 396.1768, found $[(M-Na)]^-$, *m/z* 397.42.

2s – Sodium 3-(3-chlorophenyl)-4-methyl-2-oxo-2*H*-chromen-6-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.61-7.59 (m, 1H), 7.50-7.48 (m, 3H), 7.45 (s, 1H), 7.40-7.37 (m, 1H), 7.33-7.31 (m, 1H). 2.24 (s, 3H). ^{13}C NMR (DMSO-d_6, 100 MHz): 159.71, 149.90, 148.17, 148.05, 136.80, 132.75, 130.02, 129.87, 128.93, 127.92, 125.04, 124.84, 120.01, 116.73, 116.64, 44.20, 16.44. MS (ESI) calculated for $C_{16}H_{10}ClNaO_6S$ $[(M-Na)]^-$, *m/z* 365.7498, found $[(M-Na)]^-$, *m/z* 365.49.

2t – Sodium 3-(4-chlorophenyl)-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.27 (s, 1H), 7.79-7.77 (d, *J* = 8 Hz, 2H), 7.71-7.69 (d, *J* = 8 Hz, 1H), 7.54-7.52 (d, *J* = 8 Hz, 2H), 7.27 (s, 1H), 7.21-7.19 (d, *J* = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 167.40, 159.72, 156.85, 153.86, 140.95, 133.70, 132.94, 130.14, 129.15, 128.19, 122.90, 116.76, 114.50, 106.28, 44.21. MS (ESI) calculated for $C_{15}H_8ClNaO_6S$ $[(M-Na)]^-$, *m/z* 351.7228, found $[(M-Na)]^-$, *m/z* 351.46.

2u – Sodium 3-(3-chlorophenyl)-2-oxo-2*H*-chromen-7-yl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.33 (s, 1H), 7.83 (s, 1H), 7.73-7.69 (m, 2H), 7.53-7.49 (m, 2H), 7.28 (s, 1H), 7.22-7.20 (d, *J* = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 159.65, 157.00, 153.93, 141.50, 136.94, 132.87, 131.07, 130.04, 129.26, 128.06, 127.01, 122.59, 116.77, 114.44, 106.26. MS (ESI) calculated for $C_{15}H_8ClNaO_6S$ $[(M-Na)]^-$, *m/z* 351.7228, found $[(M-Na)]^-$, *m/z* 351.46.

2v – Sodium 4-(6-bromo-4-methyl-2-oxo-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.94 (s, 1H), 7.73 (dd, *J* = 8 Hz, 1H), 7.36 (d, *J* = 8 Hz, 1H), 7.17 (m, 4H), 2.22 (s, 3H). ^{13}C NMR (DMSO-d_6, 100 MHz): 159.46, 153.39, 151.06, 146.67, 133.93, 130.64, 128.48, 128.14, 126.97, 122.18, 119.88, 119.83, 118.45, 116.31, 26.17, 16.49. MS (ESI) calculated for $C_{16}H_{10}BrNaO_6S$ $[(M-Na)]^-$, *m/z* 410.2038, found $[(M-Na)]^-$, *m/z* 409.49.

2w – Sodium 2-hydroxy-4-(2-oxo-4-phenyl-7-(sulfonatooxy)-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.38-7.31 (m, 5H), 7.20 (d, *J* = 8 Hz, 2H), 7.07 (dd, *J* = 8 Hz, 2.32 Hz, 1H), 6.93 (d, *J* = 8 Hz, 1H), 6.64 (dd, *J* = 8 Hz, 2.16 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 160.46, 156.33, 153.16, 150.81, 143.27, 142.42, 134.41,

129.02, 128.97, 128.20, 128.13, 127.76, 127.47, 124.48, 123.98, 122.72, 118.27, 116.40, 115.30, 106.67, 99.49. MS (ESI) calculated for $C_{21}H_{12}Na_2O_{11}S_2$ $[(M-2Na)]^-$, *m/z* 504.4155, found $[(M-2Na)]^{2-}$, *m/z* 252.51.

2x – Sodium 4-(6,8-dichloro-2-oxo-4-phenyl-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.02 (d, *J* = 2.4 Hz, 1H), 7.44-7.37 (m, 3H), 7.25-7.23 (m, 2H), 7.07 (d, *J* = 8 Hz, 2H), 7.00 (d, *J* = 8 Hz, 2H), 6.86 (d, *J* = 8 Hz, 1H). ^{13}C NMR (DMSO-d_6, 100 MHz): 160.44, 154.36, 150.33, 148.27, 134.89, 132.31, 132.21, 131.93, 130.36, 130.01, 129.88, 129.83, 129.45, 129.24, 128.95, 126.46, 124.46, 123.26, 122.66, 120.26, 120.20. MS (ESI) calculated for $C_{21}H_{11}Cl_2NaO_6S$ $[(M-Na)]^-$, *m/z* 462.2628, found $[(M-Na)]^-$, *m/z* 461.58.

2y – Sodium 4-(2-oxo-4-phenyl-7-(sulfonatooxy)-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 7.40-7.31 (m, 4H), 7.23-7.21 (m, 2H), 7.08-6.93 (m, 6H). ^{13}C NMR (DMSO-d_6, 100 MHz): 160.62, 156.40, 153.15, 152.58, 150.83, 147.52, 134.60, 131.12, 129.04, 128.55, 128.23, 128.18, 127.71, 123.88, 118.91, 118.89, 118.82, 116.49, 115.36, 114.40, 106.74. MS (ESI) calculated for $C_{21}H_{12}Na_2O_{10}S_2$ $[(M-2Na)]^-$, *m/z* 488.4165, found $[(M-2Na)]^{2-}$, *m/z* 244.54.

3a – Sodium 4-(2-oxo-7-((1-(3-((2-oxo-4-phenyl-3-(4-(sulfonatooxy)phenyl)-2*H*-chromen-7-yl)oxy)propyl)-1*H*-1,2,3-triazol-4-yl)methoxy)-4-phenyl-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.37 (s, 1H), 7.40-7.32 (m, 6H), 7.29 (s, 1H), 7.22-7.19 (m, 4H), 7.11 (s, 1H), 7.04-7.01 (m, 4H), 6.97-6.93 (m, 7H), 6.88 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 5.29 (s, 2H), 4.61 (t, 2H), 4.16 (t, 2H), 2.35 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz): 193.62, 161.07, 160.73, 160.64, 154.10, 154.06, 152.58,

150.96, 150.94, 147.96, 143.19, 141.93, 134.61, 131.14, 129.02, 128.50, 128.38, 128.36, 128.24, 128.19, 127.46, 125.00, 123.60, 123.13, 123.02, 122.63, 118.84, 114.19, 113.92, 113.81, 112.85, 112.65, 102.77, 101.47, 101.18, 99.50, 67.38, 65.51, 65.49, 59.27, 52.77, 50.97, 48.67, 46.63, 46.59, 42.40, 29.22, 9.92. MS (ESI) calculated for $C_{48}H_{33}N_3Na_2O_{14}S_2$ $[(M-2Na)]^-$, *m/z* 939.8985, found $[(M-2Na)]^{2-}$, *m/z* 470.45.

3b – Sodium 4-(2-oxo-7-((1-(4-((2-oxo-4-phenyl-3-(4-(sulfonatooxy)phenyl)-2*H*-chromen-7-yl)oxy)butyl)-1*H*-1,2,3-triazol-4-yl)methoxy)-4-phenyl-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.35 (s, 1H), 7.40-7.31 (m, 6H), 7.29 (s, 1H), 7.21-7.19 (m, 4H), 7.11 (s, 1H), 7.04-7.01 (m, 4H), 6.97-6.94 (m, 4H), 6.92 (s, 1H), 6.90 (dd, *J* = 8 Hz, 2.36 Hz, 1H), 5.29 (s, 2H), 4.50 (t, *J* = 8 Hz, 2H), 4.15 (t, *J* = 8 Hz, 2H), 2.05 (quint, *J* = 8 Hz, 2H), 1.78 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz): 203.94, 196.89, 195.80, 175.54, 166.72, 165.45, 161.34, 161.31, 160.73, 157.49, 155.67, 154.06, 152.59, 141.90, 140.87, 139.56, 134.65, 134.61, 131.14, 129.52, 129.22, 129.15, 129.12, 129.03, 128.61, 128.52, 128.50, 128.45, 128.36, 128.31, 128.24, 124.82, 123.13, 120.66, 118.93, 118.83, 113.93, 112.68, 101.52, 101.12, 94.16, 88.00, 67.65, 61.78, 52.77, 51.47, 49.08, 26.41, 7.31. MS (ESI) calculated for $C_{49}H_{35}N_3Na_2O_{14}S_2$ $[(M-2Na)]^-$, *m/z* 953.9255, found $[(M-2Na)]^{2-}$, *m/z* 477.49.

3c – Sodium 4-(2-oxo-7-((1-(3-((2-oxo-3-(4-(sulfonatooxy)phenyl)-2*H*-chromen-7-yl)oxy)propyl)-1*H*-1,2,3-triazol-4-yl)methoxy)-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.38 (s, 1H), 8.18 (s, 2H), 7.72-7.62 (m, 6H), 7.25-7.22 (m, 5H), 7.05-7.04 (m, 2H), 6.99 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 5.30 (s, 2H), 4.62 (t, *J* = 8 Hz, 2H), 4.17 (t, *J* = 8 Hz, 2H), 2.37 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz):

175.55, 159.13, 144.57, 139.90, 139.49, 136.17, 135.28, 132.43, 131.38, 128.92, 128.85, 126.18, 123.16, 122.63, 119.91, 117.12, 116.36, 115.79, 110.72, 109.93, 108.24, 103.12, 102.74, 88.62, 79.35, 78.57, 72.02, 71.32, 66.46, 56.09, 47.12, 42.14, 40.72, 37.56, 13.22, 10.00. MS (ESI) calculated for $C_{36}H_{25}N_3Na_2O_{14}S_2$ $[(M-2Na)]^-$, *m/z* 787.7025, found $[(M-2Na)]^{2-}$, *m/z* 394.24.

3d – Sodium 4-(2-oxo-7-((1-(4-((2-oxo-3-(4-(sulfonatooxy)phenyl)-2*H*-chromen-7-yl)oxy)butyl)-1*H*-1,2,3-triazol-4-yl)methoxy)-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.35 (s, 1H), 8.18 (s, 2H), 7.71-7.62 (m, 6H), 7.25-7.22 (m, 5H), 7.04-6.97 (m, 3H), 5.29 (s, 2H), 4.51 (t, *J* = 8 Hz, 2H), 4.16 (t, *J* = 8 Hz, 2H), 2.06 (quint, *J* = 8 Hz, 2H), 1.79 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz): 181.55, 175.12, 174.05, 168.82, 160.00, 154.61, 154.52, 145.95, 141.93, 141.60, 139.90, 129.56, 129.39, 128.85, 124.83, 123.15, 122.93, 119.91, 113.39, 113.10, 112.90, 105.82, 101.07, 100.68, 94.85, 67.61, 61.76, 57.57, 49.07, 45.69, 44.08, 26.43, 25.47, 18.59, 13.88, 7.98, 3.25. MS (ESI) calculated for $C_{37}H_{27}N_3Na_2O_{14}S_2$ $[(M-2Na)]^-$, *m/z* 801.7295, found $[(M-2Na)]^{2-}$, *m/z* 401.27.

3e – Sodium 4-(4-methyl-7-((1-(3-((4-methyl-2-oxo-3-(4-(sulfonatooxy)phenyl)-2*H*-chromen-7-yl)oxy)propyl)-1*H*-1,2,3-triazol-4-yl)methoxy)-2-oxo-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.38 (s, 1H), 7.79 (d, *J* = 8 Hz, 1H) 7.76 (d, *J* = 8 Hz, 1H), 7.26-7.20 (m, 8H), 7.09-6.97 (m, 4H), 5.31 (s, 2H), 4.63 (t, *J* = 8 Hz, 2H), 4.17 (t, *J* = 8 Hz, 2H), 2.27 (s, 6H), 2.40 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz): 161.19, 160.63, 160.22, 153.66, 153.58, 153.06, 147.91, 147.87, 141.98, 130.79, 129.05, 127.00, 126.96, 124.80, 123.09, 122.90, 119.84, 113.84, 113.57,

112.68, 112.48, 101.33, 100.93, 84.69, 67.56, 61.73, 56.57, 52.78, 49.09, 44.34, 36.61, 26.59, 26.44, 25.48, 16.44, 13.45, 10.45, 3.42. MS (ESI) calculated for $C_{38}H_{29}N_3Na_2O_{14}S_2$ $[(M-2Na)]^-$, *m/z* 815.7565, found $[(M-2Na)]^{2-}$, *m/z* 408.71.

3f – Sodium 4-(4-methyl-7-((1-(4-((4-methyl-2-oxo-3-(4-(sulfonatooxy)phenyl)-2*H*-chromen-7-yl)oxy)butyl)-1*H*-1,2,3-triazol-4-yl)methoxy)-2-oxo-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.36 (s, 1H), 7.79 (m, 2H), 7.22-7.19 (m, 9H), 7.09-6.98 (m, 3H), 5.30 (s, 2H), 4.52 (t, *J* = 8 Hz, 2H), 4.16 (t, *J* = 8 Hz, 2H), 2.27 (s, 6H), 2.07 (quint, *J* = 8 Hz, 2H), 1.79 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz): 168.27, 166.21, 161.19, 160.63, 160.22, 153.66, 153.58, 153.06, 147.91, 147.87, 141.98, 130.79, 129.05, 127.00, 126.96, 124.80, 123.09, 122.90, 119.84, 113.84, 113.57, 112.68, 112.48, 101.33, 100.93, 84.69, 82.62, 67.56, 61.73, 56.57, 52.78, 49.09, 44.34, 26.59, 26.44, 25.48, 16.44, 13.45, 3.42. MS (ESI) calculated for $C_{39}H_{31}N_3Na_2O_{14}S_2$ $[(M-2Na)]^-$, *m/z* 829.7835, found $[(M-2Na)]^{2-}$, *m/z* 415.29.

3g – Sodium 4-(7-((1-(4-((3-(2,4-dichlorophenyl)-2-oxo-4-phenyl-2*H*-chromen-7-yl)oxy)butyl)-1*H*-1,2,3-triazol-4-yl)methoxy)-2-oxo-4-phenyl-2*H*-chromen-3-yl)phenyl sulfate – ^{1}H NMR (DMSO-d_6 400 MHz): 8.24 (s, 1H), 8.03 (s, 1H), 7.41-7.34 (m, 6H), 7.24-7.18 (m, 4H), 7.10-7.08 (m, 3H), 7.03-6.94 (m, 4H), 6.91-6.86 (m, 5H), 5.07 (s, 2H), 4.46 (t, *J* = 8 Hz, 2H), 4.11 (t, *J* = 8 Hz, 2H), 1.99 (quint, *J* = 8 Hz, 2H), 1.71 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz): 185.99, 161.31, 160.64, 159.15, 154.45, 154.14, 150.97, 148.97, 146.95, 142.47, 140.96, 134.63, 133.65, 131.67, 131.45, 131.12, 130.60, 130.11, 129.14, 129.07, 129.02, 128.63, 128.49, 128.33, 128.23, 127.96, 125.11, 124.46, 123,12, 122.95, 121.36, 119.36, 118.82, 114.16, 113.77, 112.64, 104.29, 101.20,

68.03, 67.65, 61.02, 56.52, 49.03, 47.23, 45.67, 26.37, 25.45, 16.06, 5.86. MS (ESI) calculated for $C_{49}H_{34}Cl_2N_3NaO_{10}S$ $[(M-Na)]^-$, *m/z* 927.7718, found $[(M-Na)]^-$, *m/z* 928.83.

3i – Sodium 4-(7-(4-(4-((4-(6,8-dichloro-2-oxo-4-phenyl-2H-chromen-3-yl)phenoxy)methyl)-1H-1,2,3-triazol-1-yl)butoxy)-2-oxo-4-phenyl-2H-chromen-3-yl)-1,2-phenylene bis(sulfate) – ^{1}H NMR (DMSO-d_6 400 MHz): 8.23 (s, 1H), 8.01 (s, 1H), 7.42 (m, 8H), 7.23 (d, *J* = 8 Hz, 2H), 7.17 (d, *J* = 8 Hz, 2H), 7.09 (d, *J* = 8 Hz, 2H), 6.91 (m, 5H), 6.59 (dd, *J* = 8 Hz, 2.4 Hz, 1H), 5.07 (s, 2H), 4.47 (t, *J* = 8 Hz, 2H), 4.13 (t, *J* = 8 Hz, 2H), 2.07 (quint, *J* = 8 Hz, 2H), 1.73 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6, 100 MHz): 161.26, 160.48, 159.14, 157.45, 154.15, 150.91, 148.95, 146.94, 143.24, 142.52, 142.45, 134.47, 133.65, 131.66, 130.59, 129.07, 128.94, 128.63, 128.49, 128.36, 128.20, 128.05, 127.96, 125.88, 125.11, 124.48, 124.39, 123.07, 122.70, 121.35, 118.29, 113.77, 113.59, 112.58, 101.09, 67.63, 61.04, 49.03, 26.40, 25.46. MS (ESI) calculated for $C_{49}H_{33}Cl_2N_3Na_2O_{14}S_2$ $[(M-Na)]^-$, *m/z* 1045.82, found $[(M-SO_3Na)]^-$, 942.13 *m/z*.

3j – **Sodium 4-(7-((1-(4-((3-(3,4-bis(sulfonatooxy)phenyl)-2-oxo-4-phenyl-2H-chromen-7-yl)oxy)butyl)-1H-1,2,3-triazol-4-yl)methoxy)-2-oxo-4-phenyl-2H-chromen-3-yl)-1,2-phenylene bis(sulfate)** – ^{1}H NMR (DMSO-d_6 400 MHz): 8.33 (s, 1H), 7.38 (m, 13H), 6.96 (m, 5H), 6.74 (m, 3H), 5.33 (s, 2H), 4.49 (t, *J* = 8 Hz, 2H), 4.13 (t, *J* = 8 Hz, 2H), 2.03 (quint, *J* = 8 Hz, 2H), 1.76 (quint, *J* = 8 Hz, 2H). ^{13}C NMR (DMSO-d_6 100 MHz): 203.94, 196.89, 195.80, 175.54, 166.72, 165.45, 161.34, 161.31, 160.73, 157.49, 155.67, 154.06, 152.59, 141.90, 140.87, 139.56, 134.65, 134.61, 131.14, 129.52, 129.22, 129.15, 129.12, 129.03, 128.61, 128.52, 128.50, 128.45, 128.36, 128.31, 128.24, 124.82, 123.13, 120.66, 118.93, 118.83, 113.93, 112.68, 101.52, 101.12, 94.16, 88.00, 67.65,

61.78, 52.77, 51.47, 49.08, 26.41, 25.43. MS (ESI) calculated for $C_{49}H_{33}N_3Na_4O_{22}S_4$ [(M-Na)]$^-$, *m/z* 1213.00, found [(M-SO_3Na)]$^{2-}$, 492.72 *m/z*.

CHAPTER 3 Evaluation of Inhibition and Thrombin-CSAM Binding Mechanism

Many glycosaminoglycans exist as highly negatively charged ligands, which typically interact with the surface of proteins through electrostatic forces by pairing with positively charged residues on the surface of said protein (lysine, arginine, and histidine residues). These interactions commonly occur at allosteric (distal) sites on the protein inducing some type of conformational change of the protein. The general theory behind how these interactions take place is the polyelectrolyte theory, which explains that the binding is actually an entropic event[143]. In this theory, we assume that sodium cations are bound to the sulfate groups in a manner to stabilize the large, negative charge density; which is thermodynamically unfavorable[143]. In the event the GAG's sulfate comes into contact with a protein's positively charged surface, for example an arginine, the sodium ion is displaced and the arginine's positive charge stabilizes the sulfate's negative charge; providing a more favorable interaction. Although these types of binding are generally homologous interactions (i.e., Na^+ v. $R\text{-}NH_4^+$), the sulfates present on GAGs have been shown to prefer the guanidinium group of the arginine residues over the ammonium group of lysine (Arg > Lys > His)[143]. This lays towards the specificity GAGs possess for their respected protein interactions.

GAG recognition and specificity relies on many other factors. First, spacing of Arg and Lys residues is essential for proper binding and decreases the possibility of non-specific interactions[143]. This is also done by the protein surface displaying other residues. For example, thrombin's exosite 2 contains hydrophobic patches around key residues, as well as strategically placed serines, glutamines, and asparagines for enhancement of binding. Another considerable aspect is the topography of the protein's surface, by which the GAG interacts with. For instance, heparin interacts with both thrombin and antithrombin; this facilitates the inactivation of thrombin by antithrombin through a heparin-induced, pseudo-intramolecular reaction. This is attained by thrombin having a lower binding affinity to heparin than antithrombin. Thrombin's binding affinity (K_D) for heparin is 6-10 μM[144], while antithrombin's binding affinity is ~10 nM[95]. This 1000-fold difference in affinities, allows tight binding of antithrombin to heparin, while thrombin is moves toward the antithrombin reactive loop. This information shows why heparin acts in an indirect manner. Thus, the binding of heparin to thrombin does not cause a dysfunctional conformational change of thrombin's orthosteric (active) site. Many studies have shown that thrombin's allosteric (or exo-) sites and the active site are interconnected; thus binding at one site induces a change to the subsequent sites[145], but not seen with heparin binding to thrombin[113].

Taking into consideration the combination of thermodynamic effects of polyelectrolytes, additional hydrophilic or hydrophobic interactions, and allosteric effects, compounds can be derived to specifically bind to thrombin's exosites. The coumarin-based sulfated allosteric modulators (CSAMs) synthesized herein represent such a group that

selectively target thrombin, at a specific exosite. Weighing into the conformational changes they induce, could alter the ability for thrombin to hydrolyze thrombin's substrates. Additionally, using minimal sulfates while increasing the hydrophobicity of the CSAMs eliminates the polyelectrolyte phenomenon, effectively using the sulfate as a steering group and increases the necessity for other binding examples.

3.1. Dose-Dependent Inhibition

Thrombin is the major regulatory serine protease within the coagulation cascade. Its mechanism within the active site utilizes a catalytic triad composed of histidine 57, aspartate 102, and serine 195, in collaboration with a water molecule to hydrolyze the carboxy-side of the arginine or lysine residues[13,109]. This mechanism allows for a much simpler approach to probing the active site's function of an activated thrombin. By obtaining synthetic peptides containing *p*-nitroanilide groups on the carboxy-side of the arginine's, when thrombin hydrolyzes the substrate, *p*-nitroaniline is released and can be monitored at 405 nm. This technique has been used extensively for thrombin, as well as other serine proteases like factor Xa and factor XIa[118,120,122,124–127,129,146].

The first generation library of CSAMs was initially screened against thrombin, factor Xa and factor XIa, which are considered major targets for anticoagulation. The experiments tested CSAMs at a high concentration of 250 μM in duplicates. Depending on the enzyme present in the well being screened, appropriate chromogenic substrates were used; *H*-D-cyclohexylalanyl-Ala-Arg-*p*-nitroanilide for thrombin, methoxycarbonyl-D-cyclohexylglycyl-Gly-Arg-*p*-nitroanilide for factor Xa, and L-pyroGlu-Pro-Arg-*p*-nitroaniline for factor XIa, as performed previously[120,124,125,127,129,146].

For thrombin, the experiments were carried out in 180 µL of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH = 7.4) at 25 °C, 5 µL of 240 nM thrombin and 10 µL of CSAM (or vehicle, DMSO) were incubated for 10 minutes. Then, with the addition of 5 µL of substrate (2 mM), the initial rates obtained from *p*-nitroaniline production were monitored over a period of 60 seconds on the Flexstation III plate reader. For factor Xa, 180 µL of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, 0.02% Tween 80, pH 7.4) was added to each well followed by 5 µL of 43.5 nM enzyme and 10 µL of CSAM or vehicle. After a 10 min incubation, 5 µL of 5 mM substrate was added to each well simultaneously and allowed to react for 300 seconds at 37 ºC. Lastly for factor XIa screens, 85 µL of buffer (50 mM Tris-HCl, 150 mM NaCl, 0.1% PEG 8000, 0.02% Tween 80, pH 7.4) was added to each well, followed by 7 µL of 15.3 nM enzyme and 3 µL of CSAM or vehicle and incubated for 10 min at 37 ºC. After the incubation period, 5 µL of 6.9 mM S-2366 was and allowed to react for 300 seconds at 37 ºC. Any sample that showed ≥20% reduction in initial rate compared to vehicle was considered a "hit". A total of eleven compounds produced hits, ten were identified for thrombin, six for factor Xa, and nine for factor XIa.

These eleven CSAMs were then tested in a dose-response manner to assess the half-maximal inhibitory concentration (IC_{50}), efficacy of inhibition (ΔY), and Hill slope (HS) against their respected enzymes, calculated through non-linear regression analysis $Y = Y_0 + \frac{Y_M - Y_0}{1+10^{(\log[Inhibitor]_0 - log\ IC_{50})\times HS}}$ (Equation 1). Similar to the protocol above, serial dilutions of stock CSAM in DMSO were prepared (0.04 - 5 mM) giving way to concentrations ranging from 0.002 - 250 μM in the wells. The experimental data produced dose-response curves calculating the fractional activity (Y) at each concentration of the

	Thrombin			Factor Xa			Factor XIa		
	***IC*50 (μM)**	**Δ*Y* (%)**	***HS***	***IC*50 (μM)**	**Δ*Y* (%)**	***HS***	***IC*50 (μM)**	**Δ*Y* (%)**	***HS***
2e	58 ± 7	66 ± 6	2.7	62 ± 20	60 ± 20	2.9	220 ± 12	86 ± 7	4.1
2k	7.2 ± 2	22 ± 4	2.9	65 ± 23	54 ± 14	1.9	137 ± 7	83 ± 6	2.8
2o	NI	-	-	53 ± 28	91 ± 26	1.2	NI	-	-
2p	35 ± 1	35 ± 1	5.3	NI	-	-	NI	-	-
3a	0.5 ± 0.1	73 ± 6	1.3	56 ± 29	92 ± 23	1.1	51 ± 2	88 ± 2	3.5
3b	1.0 ± 0.1	58 ± 2	2.6	13 ± 3	67 ± 14	6.7	59 ± 3	95 ± 3	4.3
3c	7.8 ± 1	60 ± 6	1.2	NI	-	-	78 ± 3	100 ± 3	4.3
3d	11 ± 2	34 ± 3	9.2	NI	-	-	91 ± 6	97 ± 5	3.4
3e	20 ± 3	52 ± 7	3.9	NI	-	-	97 ± 4	98 ± 4	4.0
3f	8 ± 3	36 ± 9	9.8	NI	-	-	73 ± 2	97 ± 2	3.9
3g	0.2 ± 0.1	47 ± 3	1.2	163 ± 2	83 ± 3	1.8	31 ± 0	99 ± 3	3.5

Table 1: Results of dose-response of CSAMs to three proteases: thrombin, factor Xa, and factor XIa. Abbreviations: half maximal inhibitory concentration (IC_{50}), efficacy of inhibition (ΔY), Hill slop (HS), No inhibition (NI)..

inhibitors tested against each protease; see **Table 1**.

The analysis provided extremely interesting data. The compounds against factor Xa were very mild in potency, having all IC_{50}s ≥ 13 μM, and were not very attractive to pursue further (**Table 1**). For factor XIa, the CSAMs behaved similarly to SAMs previously synthesized and tested[127,129]. The IC_{50}s for monomers were relatively high, 220 and 137 μM, respectively, but this was the first instance of monomeric and monosulfated

compounds to induce inhibition of factor XIa. The dimeric CSAMs ranged in IC_{50}s from 97 – 31 µM, of which **3g** was the most potent SAM identified to date (**Table 1**). Similarly to previous SAMs tested[127,129,146], the CSAMs showed high Hill slopes owing to the cooperativity of the dimeric factor XIa, but all possessed ΔY > 80%.

In the case of thrombin, the active CSAMs ranged in IC_{50}s from 0.2 – 58 µM, describing the most potent SAMs to date (**Figure 10**). More strikingly however, was the efficacies of inhibition. The ΔYs ranged from 22 – 73%, suggesting that the CSAM scaffold inherently induces partial inhibition. The exception to this was the compound **3h**, which is **3g** without the sulfate, or known as **3g5** as depicted by the NMR data. This could mean that the binding of CSAM to thrombin, induces a slight conformational change to the orthosteric site, allowing it to remain partially active. This also suggests that the sulfate is indeed required for proper binding of the CSAM to thrombin.

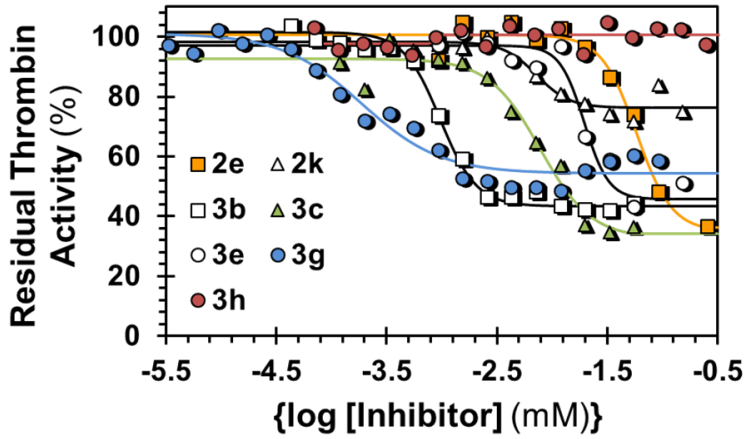


Figure 10: Sampling of the dose-response profiles of selected CSAMs against thrombin. Profiles show variable efficacies and potencies. CSAM **3h** is the non-sulfated derivative of **3g**, which shows no inhibition, indicating the sulfate is required for binding.

3.1.1 Structure-Activity Relationships

In order to fully understand what was causing the active CSAMs to inhibit thrombin, structure-activity relationships (SARs) were assessed for the compounds. SARs

identify the compound's structural groups necessary for activity by comparing against various structural differences at that position on the molecule, which is then compared to the activity it induces on the target. First, looking at the monomeric sulfated coumarins, only three were active toward thrombin (**2e**, **2k** and **2p**). These three possessed bulky phenyl (or 4-chlorobenzyl) groups at the 4-position. Also, the monomers all contained chloro groups, either on the 3-(phenyl) or 4-(benzyl) moiety, suggesting that significant hydrophobic interactions are at play inhibiting thrombin. It is worth noting that not all compounds with bulky moieties at the 4-position inhibited thrombin. Additionally, compounds containing the sulfate at the 7-position were favored over the 6-position, or at the 4-(4'-hydroxyphenyl) position. Increasing the sulfate count as in **2w** or **2y** was not favorable for inhibition of thrombin.

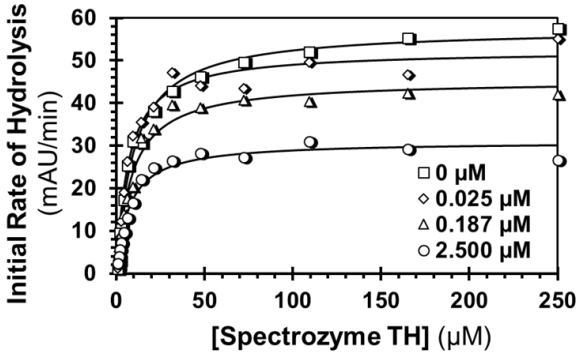
In contrast, all dimeric CSAMs displayed good inhibition potencies (IC_{50} < 20 μM). Multiple variations were involved with the dimers, including chain length, linker position, number of sulfates, and 4-position moieties. First linker length varied between 3 and 4 carbons, which did not alter the inhibition significantly. Linker position gave way to either alkylation at the 7-position yielding head-to-head or 7-7' dimers, and one head-to-tail or 7-4' dimer. Dimers **3a** – **3f** were 7-7' dimers consisting of a total of two sulfates, one at each 4-(4'-hydroxyphenyl) position, while **3g**, the 7-4' dimer contained one sulfate at the 4-(4'-hydroxylphenyl) position. Lastly, the 4-position varied by phenyl, methyl, and hydrogen. The original observation that the 4-phenyl group was more potent held true, and the methyl or hydrogen did not significantly increase the inhibition concentration. Overall,

3g contained all the structural elements of the most active dimers and monomers including a sulfated 4-(4'-hydroxyphenyl), 4-position phenyl, and aryl chloride substituents.

3.2 Michaelis-Menten Kinetics

The inhibition profiles of CSAMs show interesting results when compared to other compounds against the same proteases. For thrombin, the molecule **3g** presents a nanomolar IC_{50} while allowing the enzyme to remain active. Likewise, **3g** also is shown to be the most potent sulfated allosteric modulator of FXIa to date[127,129]. Thus, it is of high priority to understand the binding mechanism of **3g** and its represented library. Typically to achieve this, Michaelis-Menten kinetic studies are performed to observe the enzyme's catalytic rate to hydrolyze various concentrations of a substrate and applying the parameters to the equation: $v_0 = \frac{V_{MAX}[S]}{[S]+K_M}$ (Equation 2). This can be performed in the absence or presence of an inhibitor. As the concentration of the inhibitor increases from the control, there can be decreases in the V_{MAX} (maximum velocity or = $k_{cat}[E]_0$; k_{cat} is the catalytic rate and $[E]_0$ is the initial enzyme concentration), the K_M (Michaelis constant or the substrate concentration at ½ V_{MAX}), or both. Decreases in the K_M represent competitive inhibition, decreases in the V_{MAX} represent non-competitive inhibition, and decreases in both represent un-competitive inhibition.

Considering **3g** was the most potent compound against thrombin and factor XIa, we chose it to represent the CSAM library and test its kinetics, respectively. In a similar manner to the inhibition studies, 180 µL of thrombin buffer (20 mM Tris-HCl, 100 mM

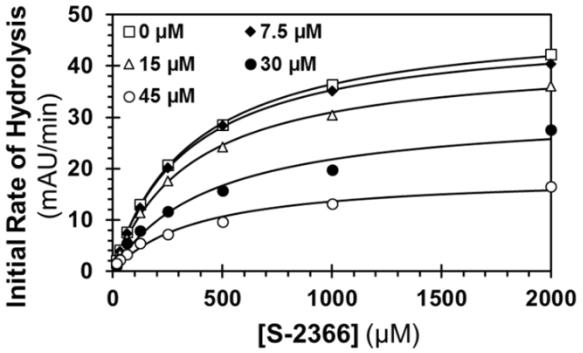


[3g] (μM)	K_M (μM)	V_{MAX} (mAU/min)
0	9.8 ± 1	57 ± 1.3
0.025	6.9 ± 1	52 ± 1.4
0.187	8.5 ± 1	45 ± 1.1
2.500	7.2 ± 1	31 ± 0.8

Figure 11: Profile of Michaelis-Menten kinetic studies using increasing concentrations of **3g** in the presence of thrombin, using Spectrozyme TH as the active site substrate. The general trend in the graph and table shows a decrease the V_{MAX}, but no significant changes to the K_M suggesting **3g**'s mechanism of binding is non-competitive.

NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH = 7.4) was incubated with 5 µL of 240 nM thrombin and 10 µL of vehicle, 0.5, 3.74, or 50 µM of **3g** (0.025, 0.187, 2.5 µM in the well) for 10 minutes at 25 °C. After the incubation period, 5 µL of Spectrozyme TH dilutions ranging from 0.019 – 10 mM were added and the initial rate of hydrolysis was monitored at 405 nm for *p*-nitroaniline production at a constant 25 °C. The slopes generated from the Flexstation III were inputted into SigmaPlot and applied to equation 2. The data revealed that as the concentration of **3g** increased from 0 to 2.5 µM, the K_M remained within the measured range for thrombin activity uninhibited as shown in **Figure 11**. However, the increase in **3g** did cause the V_{MAX} to decrease in a concentration dependent manner, shown in **Figure 11**. This supports a non-competitive mechanism of binding, suggesting allostery.

Likewise, with factor XIa, 85 µL of factor XIa buffer (50 mM Tris-HCl, 150 mM NaCl, 0.1% PEG 8000, 0.02% Tween 80, pH 7.4) was incubated with 7 µL of 15.3 nM enzyme and 3 µL of vehicle, 7.5, 15, 30, or 45 µM for 10 minutes at 37 ºC. After the incubation period, 5 µL of S-2366 ranging from 1.2 – 40 mM was added and monitored at



[3g] (µM)	K_M (µM)	V_{MAX} (mAU/min)
0	355 ± 7	49 ± 0.4
7.5	343 ± 5	47 ± 0.3
15	352 ± 14	42 ± 0.6
30	443 ± 92	31 ± 2.4
45	380 ± 72	19 ± 1.3

Figure 12: Profile of Michaelis-Menten kinetic studies using increasing concentrations of **3g** against factor XIa using S-2366 as substrate for the active sites. The general trend in the graph and table shows the same trend as seen in **Figure 11** against thrombin. There is a concentration-dependent decrease of the V_{MAX}, but no significant changes to the K_M suggesting **3g**'s mechanism of binding is again non-competitive.

405 nm for 300 seconds at a constant temperature of 37 °C. The slopes generated by the Flexstation III were applied to equation 2 in SigmaPlot. The data from this showed a non-competitive mechanism of binding as well. The K_M remained within standard deviation of uninhibited factor XIa. Although, the V_{MAX} decreased in a concentration dependent manner. Comparing the kinetics of **3g** with thrombin and factor XIa, both indeed maintain a non-competitive mechanism, however, at saturation of **3g** against thrombin, the rate of hydrolysis does not become diminished beyond ~50%. The same cannot be said for factor XIa, because as the concentration of **3g** reaches saturation, the rate of hydrolysis continues to become abolished. This further supports the phenomenon of allosteric, partial inhibition.

3.3 Exosite Competition

As mentioned earlier, thrombin possesses two distinct allosteric sites called exosite 1 and exosite 2. These exosites are selective to a number of different ligands. Exosite 1 binds with fibrinogen Aα, protease-activated receptor 1, the A1 domain of factor XI,

thrombomodulin, hirudin, and heparin cofactor II[109]. Conversely, exosite 2 binds exclusively with heparin/heparin sulfate, chondroitin sulfate of thrombomodulin, glycoprotein 1bα, and γ'-fibrinogen[109]. It is safe to assume that because of these sudden, yet significant differences in structural features, synthetic compounds can selectively target one exosite over the other.

To assess whether **3g** binds to exosite 1 or 2, competition studies are performed to investigate such a phenomenon. For exosite 1, hirudin peptide (HirP, [5F]-Hir-(54-65)-(SO_3^-)) is a truncated, fluoresceinylated version of the anticoagulant hirudin, which is found in the salivary glands of leeches[147]. The HirP does not possess the portion of the peptide that interacts and inactivates thrombin's active site, thus can be used to probe exosite 1. Likewise, unfractionated heparin (UFH) was used for probing exosite 2 interactions. As mentioned previously, UFH does not interfere with thrombin's ability to hydrolyze substrate. The experiments utilized the calculated K_D of HirP or UFH for thrombin and appropriate concentrations are used to show absence of competitor, the K_D concentration, and two-fold the K_D concentration. If there is a change in the dose-response curves, showing an increased IC_{50}, the competitor is competing with **3g** at that exosite.

For exosite 1 competition studies, appropriate volumes (175 or 170 µL) of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH = 7.4) were incubated with 5 µL of 240 nM thrombin, 10 µL of DMSO or varying concentrations of **3g** (50 – 0.2 µM) giving 2.5 – 0.001 µM, and 5 µL of vehicle, 5 µL of 500 nM HirP, or 10 µL

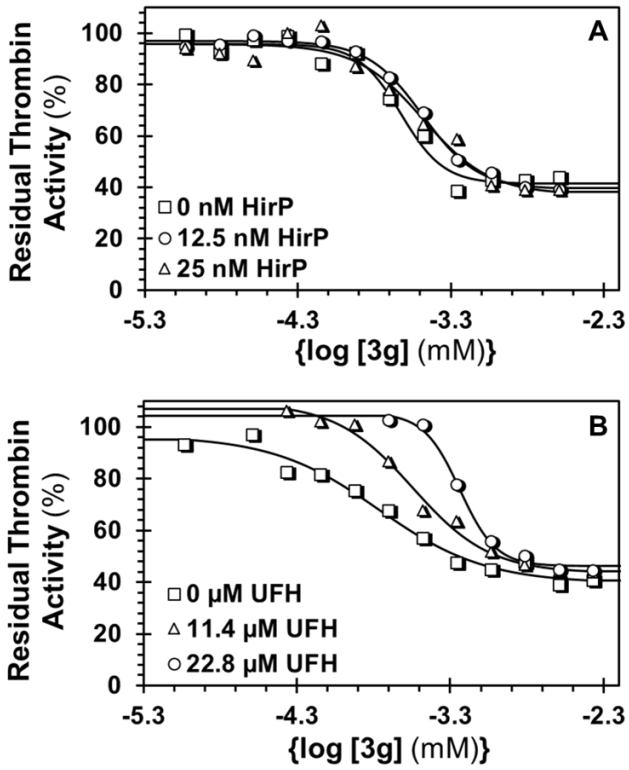


Figure 13: Dose-response profiles of varying concentrations of **3g** in the presence of a constant concentration of thrombin. **A)** Increasing concentrations of HirP (0, 12.5, 25 nM) show no considerable changes in the profiles. **B)** The increasing concentrations of UFH (0, 11.4, 22.8 uM) show right shifts of the curves, as well as increasing Hill Slopes suggesting that binding of **3g** and UFH is cooperative.

of 500 mM HirP were added. After ten minutes of incubation at 25 °C, 5 µL of Spectrozyme TH (2 mM) was added to the reaction and monitored for 60 seconds at 405 nm. Slopes generated from the Flexstation III were input into SigmaPlot using Equation 1.

Similarly for exosite 2, adjusted volumes (177 or 174 µL) of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH = 7.4) were incubated with 5 µL of 240 nM of thrombin, 10 µL of DMSO or varying concentrations of **3g** (50 – 0.2 µM) giving 2.5 – 0.001 µM, and 3 µL of vehicle, 3 µL of 761 µM, or 6 µL of 761 µM UFH were added. After the ten-minute incubation, 5 µL of Spectrozyme TH was added and monitored for initial rate of hydrolysis (60 seconds) at 405 nm. The slopes generated from the Flexstation III were input into SigmaPlot using Equation 1.

Hirudin peptide's binding affinity for thrombin has been determined to be 12.5 nM, so the concentration in each set of experiments varied at 0, 12.5, or 25 nM. The dose-

	$IC_{50,app}$ (μM)	ΔY (%)	HS	$IC_{50,predicted}$ (μM)
[**HirP**] (μM)				
0	0.22 ± 0.02	52 ± 2	3.3	0.24
0.0125	0.33 ± 0.02	58 ± 1	2.4	0.45
0.025	0.33 ± 0.06	59 ± 6	2.0	0.67
[**UFH**] (μM)				
0	0.17 ± 0.02	56 ± 3	1.3	0.17
11.4	0.29 ± 0.03	66 ± 5	1.7	0.29
22.8	0.58 ± 0.03	59 ± 3	3.7	0.45

Table 2: Dixon-Webb analysis displayed the dose-response data in the presence of the exosite 1 competitor, hirudin peptide (HirP), and exosite 2 competitor, unfractionated heparin (UFH). The $IC_{50,\ predicted}$ values compared against the $IC_{50,\ apparent}$ showed **3g** competed ideally with UFH for exosite 2. Abbreviations: half maximal inhibitory concentration (IC_{50}), efficacy of inhibition (ΔY), Hill slope (HS).

response profiles generated did not seem to differ significantly, nor did the IC_{50} generated (**Table 2** and **Figure 13A**). However in the presence of HirP, the IC_{50} of **3g** increased by 50% (from 236 nM to 327 and 330 nM). This phenomenon led us to believe that the interactions at exosite 1 caused a slight conformational change to exosite 2, resulting in the decreased potency of **3g** to exosite 2. For this to hold true, there must be competition at exosite 2 with UFH. Analysis of the dose-response profiles of UFH (at 0, 11.4, and 22.8 μM) showed concentration-dependent increases of the IC_{50} from 170 nM at 0 μM UFH, to 285 nM and 577 nM for 11.4 and 22.8 μM, respectively (**Table 2** and **Figure 13B**). An interesting observation of the dose-response profiles of exosite 2 competition was that as the concentration of UFH increased, the Hill slope also increased. As increased Hill slopes generally mean enhancing or inducing cooperativity of ligand binding to protein[148], this same phenomenon may be occurring to the system considering **3g** has a 100-fold higher binding affinity than UFH (as shown in Chapter 3, Section 6). Further Dixon-Webb analysis using $IC_{50,predicted} = \left(IC_{50,apparent}\right)\left(1 + \frac{[comp]}{K_{comp}}\right)$ (Equation 3; **Table 2**) showed that the **3g** did not compete ideally with exosite 1, thus ruling it out as the binding site, but did in fact compete with

UFH for exosite 2. This supports exosite 2 as the primary binding site of **3g** and that it indeed binds in an allosteric region of thrombin.

3.4 Mutant Studies

Considering the general binding region of **3g** was identified as exosite 2, further investigation into the specific residues involved with binding was essential. Typically, mutation studies such as alanine scans, give pertinent information regarding what single residues are responsible for binding by replacing them with an alanine residue. Our previous work involving exosite 2 binders have led us to using mutants that have arginines and lysines of exosite 2 mutated to alanines[124,125,128]. The information obtained from dose-response profiles generally show a 2 to 10-fold decrease in potency when the arginine or lysine involved in binding the sulfate(s) of SAMs is mutated. These residues can be mapped out to assist in the identification of the specific binding pocket. A few difficulties are observed with this technique however, especially with **3g** considering it only contains one sulfate. First, mutating surface arginines and lysines may affect the folding of thrombin or changes in structure and conformation. This must be taken into account because thrombin's activity is so highly influenced by small changes on the surface, it may cause alterations in activity and binding. Second, usually compounds tested using this method contain multiple sulfates at different positions. This gives way to sulfate-R/K interaction mapping. With **3g** possessing one sulfate, we may be only able to identify one residue responsible for binding; unless the R/K mutation influences a favorable hydrogen bonding of ligand to another residue, thus increasing the potency of **3g**. Lastly, mutants of thrombin are recombinant and are expressed in *E. coli*, which lack the ability to glycosylate

thrombin. The lack of glycosylation has been shown to cause improper quaternary (4°) folding of proteins and change thermodynamic properties necessary for ideal function and propensity[149]. This may affect thrombin in the same manner.

To obtain useful data from the thrombin mutants, dose-response profiles generated from incubation with dilutions of **3g** in the presence of recombinant wild type thrombin were compared against the mutants K236A, K235A, R233A, R175A, R173A, K169A, R165A, R126A, R101A, R97A, and R93A (K240A was found to be completely inactive and removed from the experimental data). The mutants were kindly provided by the Rezaie laboratory (St. Louis University), where their individual optical densities (ODs) were reported. Prior to performing the hydrolysis assays, the concentrations of the obtained mutants were determined from their ODs and used accordingly. Similar to the previous dose-response experiments, 180 µL of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH 7.4) was incubated at 25 °C with 5 µL of 240 nM recombinant thrombins (determined from OD) and 10µL of DMSO or **3g** (0.04 - 5 mM). After ten minutes, 5 µL of 2 mM Spectrozyme TH was added and monitored at 405 nm for 60 seconds at a constant 25 °C.

As shown in **Figure 14**, the wild type thrombin results in an IC_{50} ~400 nM; about two-fold higher than the IC_{50} with thrombin obtained from human plasma. Understanding

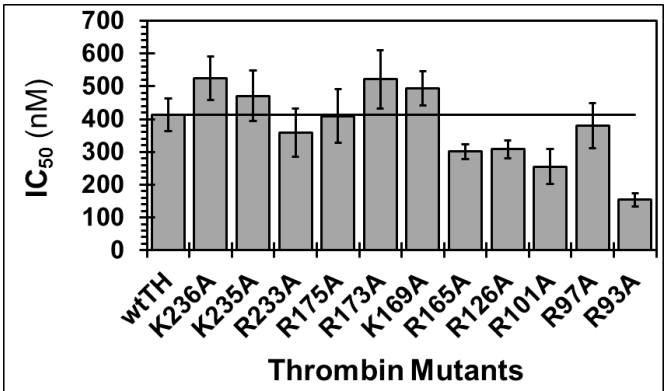


Figure 14: A bar graph representing the **3g** IC_{50}s generated against recombinant wild-type thrombin (wtTH) and the 11 single exosite 2 alanine mutants. A horizontal black line represents the wtTH IC_{50} threshold.

this is most likely due to decreased activity and incomplete quaternary conformation from lacking glycosylation, we used this as our control IC_{50}. The subsequent thrombin mutants showed varying effects resulting from **3g** interactions (**Table 3**). Most notably was the lack of statistically significant increases of IC_{50} with a single residue; going against what was originally predicted to be observed. What we instead observed for R101A and R93A, was a 2.3-fold increase in potency, yielding an IC_{50} of 170 nM. We could only make the rationale that the mutation and loss of the arginine(s) at those positions resulted in increasing the binding affinity of **3g** to thrombin, by most likely inducing a more favorable interaction or conformation.

	$IC_{50,app}$ (μM)	ΔY (%)	HS
wtTH	0.413 ± 0.05	39 ± 3	4.0
K236A	0.525 ± 0.07	64 ± 4	2.6
K235A	0.471 ± 0.08	33 ± 3	1.9
R233A	0.359 ± 0.07	33 ± 4	2.0
R175A	0.409 ± 0.08	42 ± 5	2.2
R173A	0.522 ± 0.09	36 ± 3	2.0
K169A	0.494 ± 0.05	38 ± 3	7.6
R165A	0.301 ± 0.02	28 ± 2	4.2
R126A	0.308 ± 0.03	41 ± 2	3.2
R101A	0.255 ± 0.05	35 ± 4	2.5
R97A	0.380 ± 0.07	41 ± 4	2.4
R93A	0.154 ± 0.02	51 ± 2	1.3

Table 3: Dose-response experimental data showing the apparent IC_{50} for each recombinant thrombin. The inhibition efficacy (Δ Y) was recorded, as well as the Hill slope (HS).

3.5 Computational Docking

In lieu of the mutant studies not identifying a specific residue responsible for binding the sulfate of **3g**, we took to using computational docking and scoring methods to enrich the data we obtained

from the mutants' data. Our lab has previously used docking functions to aid with identification of allosteric binding sites for glycosaminoglycans, their derivatives, and GAG mimetics[150]. Typically two major scoring functions are used, GoldScore and ChemPLP, which cluster and rank the best binding pose over a number of iterations. ChemPLP was used first to rank and score the docking poses and GoldScore was subsequently used to rescore the docking poses. Using the algorithms in this manner inserts a checks and balance system. ChemPLP is much more efficient at analyzing hydrophobic and salt interactions, while GoldScore is more efficient at hydrophilic interactions. Thus because **3g** contains mainly hydrophobic with the highly likelihood of an electrostatic interaction, ChemPLP was chosen as the first filter.

To assess **3g** as a ligand for thrombin, it was first drawn in SYBYL-X, a molecular modeling and docking software. To the structure energy minimization was executed to give the optimal conformation and torsion angles of the rotatable bonds in the structure. Then, the x-ray crystal structure of thrombin (3EQ0) bound to a hirudin variant at exosite 1 and the active site was uploaded into SYBYL-X from the Protein Data Bank. The crystal structure's ligands such as water, sodium, and the hirudin variant were removed from the structure and side chains were minimized to their lowest energies. The remaining outlier rotamers not visible in the original crystal structure, were added minimized and fixed. Once completed, the file containing the minimized thrombin and **3g** were transferred to the GOLD software program, which contains the executable algorithms for GoldScore and ChemPLP.

To perform the initial wild type thrombin docking, a centroid was set in the middle of exosite 2 with a radius of 20 Å. Upon **3g** docking, 300 generic algorithm (GA) iterations were performed in triplicate and binding poses were clustered with an RMSD ≤ 2.5 Å to categorize into tight clusters. The initial ranking and scoring was performed using ChemPLP and was rescored with GoldScore. Subsequently, each alanine mutation was installed into the thrombin structure individually and the methyl group of the alanine was used as the centroid of the 20 Å radius. Likewise, **3g** was docked into these individual mutated structures using the parameters as described above. **Figure 15** shows the top scored poses for each set of runs for wild type and the individual mutations.

Analysis of the docking poses show several interesting aspects. To examine the relevance of exosite 2 binding, we aligned the crystal structure 1XMN, which shows thrombin bound to heparin with the minimized 3EQ0 crystal structure to define the heparin binding pocket **Figure 15A**. Applying the poses of **3g** predicted to bind to thrombin showed that **3g** binds in the same binding pocket as heparin. This was regarded as a positive observation, because experimental evidence shows **3g** binding at the heparin binding site. Upon further investigation, it seems that phenyl *O*-sulfate of **3g** may be binding in a pocket made up of R126 and R233. This influence of two arginines binding one sulfate, may explain why mutations at these residues do not alter the IC_{50}s observed in the mutant studies. We suggest that the sulfate, in the absence of one binding partner

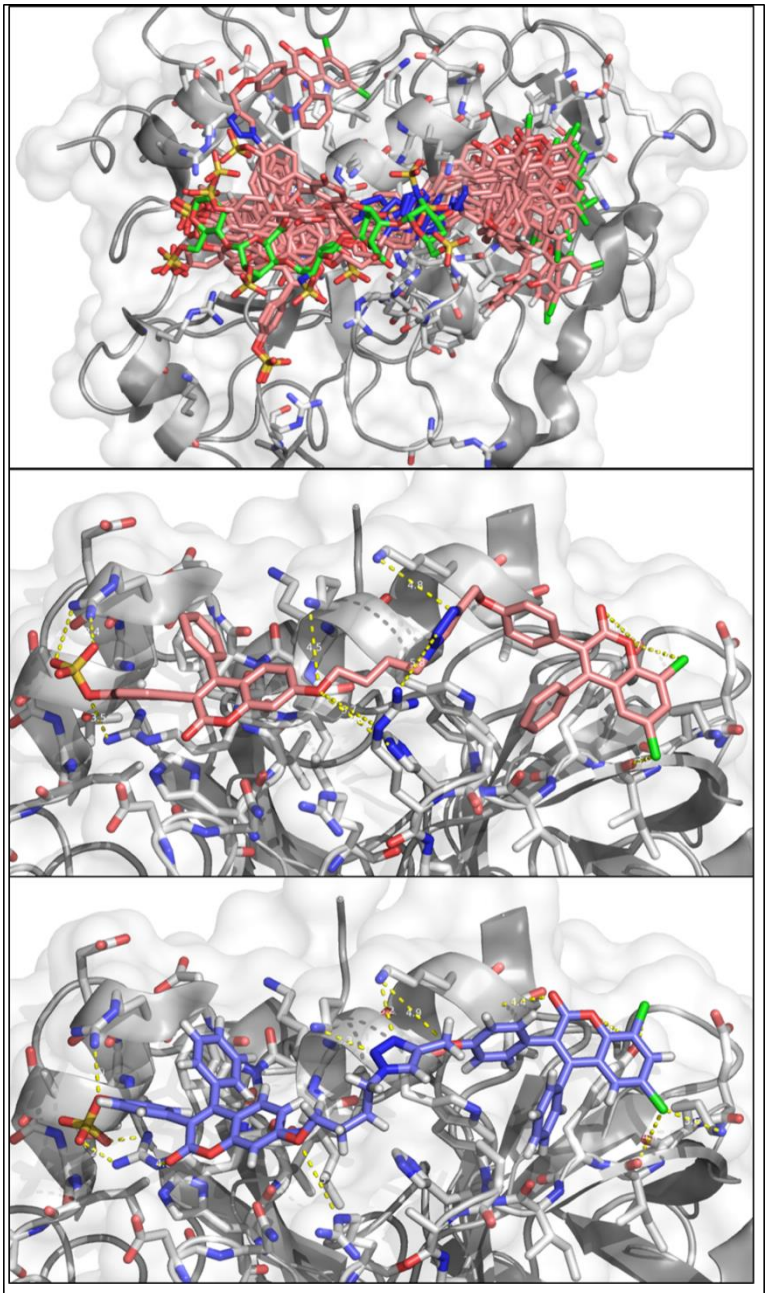


Figure 15: Pymol images of **3g** docking to thrombin (PDB ID: 3EQ0). Individual residues are kept rigid after initial minimization. **A)** Overlay of multiple grouped clusters representing predicted poses of **3g** (pink). Heparin (green) is also overlaid to show the heparin binding site. **B)** Predicted **3g** pose binding to wild type thrombin (IC_{50} = 410 nM) showing likely interactions. **C)** Predicted **3g** binding R93A mutant thrombin (IC_{50} = 154 nM) showing a shift in the triazole ring favoring a hydrogen bond between the ring nitrogens and K240 and K236 of thrombin.

(R126 or R233), simply forms a stronger electrostatic interaction with the remaining arginine available; instead of sharing a partial charge by the two arginines. To confirm this experimentally however, a double mutation (R126A and R233A) would have to be obtained and the dose-response profile would need to be evaluated. It is important to note that when **3g** lacks a sulfate and is the hydroxyl, no inhibition is observed. This suggests that the sulfate is indeed essential for binding and inducing inhibition of thrombin.

Further insight into the binding of **3g** to thrombin

was obtained by investigating the individual mutations of exosite 2 itself; most notably, the R93A and R101A mutations. Structurally, these two residues are adjacent to one another and tend to show similar effects to how **3g** binds when they are individually mutated. For both mutations versus the wild type, the docking showed the triazole ring shifting to a more favorable position, where the amines can act as hydrogen bond acceptors for K240 and K236. These interactions are posed as too far away (> 4 Å) for the wild type docking, but still may be involved depending on the rotamers of K240 and K236 (**Figure 15B and C**).

Additionally, the unique feature of **3g** is that it contains a very hydrophobic tail with two aryl chlorides in addition to the necessity of the 4-phenyl on the coumarin cores; suggested to be important for high affinity binding. In the docked poses it shows that both 4-phenyls are set in hydrophobic pockets where π-π and cation-π interactions can have much greater influence over the compounds that contained the methyl or hydrogen groups at the 4-position. In reference to the dichloro-tail, it appears to show halogen bonding with the ammonium group of K87 at an acceptable angle (observed θ = 125-155°; reported θ = 165-180° for R-Cl---NR_3) and distance (observed distance ~ 3.2 Å; reported distance 3.27 Å) (**Figure 15B and C**). All of the tail end docked poses, with the exception of R175A, are relatively conserved to the K87 ammonium group. It is important to note that with regard to the halogen bonding angle to the ammonium group, the K87 side chain rotamer is fixed and under no constriction would most likely have a more acceptable angle (closer to θ = 165-180°). This collective information strongly suggests the necessity of the dichloro moiety needed to enhance binding. Overall, the docking supports the experimental exosite

competition studies and clarifies the results of the mutant studies performed. The data obtained suggest viable poses and binding partners of **3g** to thrombin's exosite 2.

3.6 Thermodynamic Binding Affinities

According to our IC_{50} data, **3g** binds quite tightly to have such high potency for an allosteric interaction with an enzyme. To obtain more applicable information of its binding characteristics, we obtained the binding affinities through fluorescence changes as a ligand is titrated into the system, taking into account dilution factors. The characteristic loss or gain in fluorescence intensity as the ligand concentration increases, results directly from the ligand binding to thrombin. This loss or gain of fluorescence can be then used to calculate the binding affinity (K_D) using the quadratic equation

$$\frac{\Delta F}{F_o} = \frac{\Delta F_{MAX}}{F_o}\left\{\frac{[Th]_o+[3g]_o+K_D-\sqrt{([Th]_o+[3g]_o+K_D)-4[Th]_o[3g]_o}}{(2[Th]_o)}\right\}$$ (Equation 4) because the free enzyme (initial enzyme) concentration is known.

The binding affinity of **3g** was therefore determined by first, addition of 180 µL of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM CaCl2, 0.1% PEG 8000, pH 7.4) to a quartz cuvette and placed inside the fluorometer to measure background fluorescence. Fluorescence intensities were monitored by excitation at 490 nm and emission at 525 nm using slit widths of 0.5 mm throughout the experiments. Then 20 µL of 2 µM fluorescein-labeled (*f*-FPRck-) human α-thrombin (*f*-TH) was added to the cuvette, mixed, and placed into the fluorometer to measure the initial fluorescence intensity. Following this was a

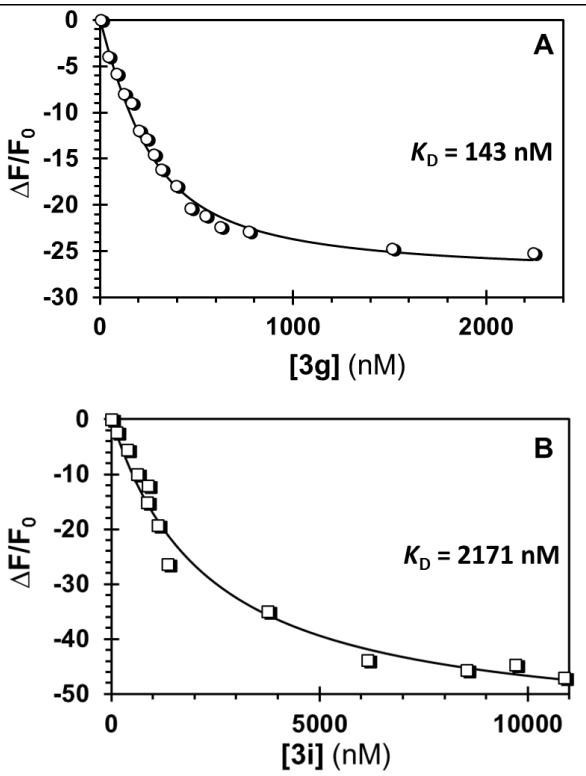


Figure 16: Thermodynamic binding affinity profiles of **3g** and **3i** for fluorescein-labeled thrombin. **A)** Depicts **3g** binding resulting in a K_D of 143 nM with a maximum change in fluorescence intensity (ΔF_{MAX}) of 28%. **B)** Profile showing the bisulfated **3g**, **3i**, as much weaker binder to thrombin, displaying a K_D of 2171 nM and a ΔF_{MAX} of ~57%

titration of 1 µL additions of **3g**, which was monitored after each addition, until the intensity began to remain unchanged with excess **3g**. The K_D of the **3g**–thrombin complex was calculated using Equation 4, in which ΔF is the change in fluorescence from the formation of the **3g**–thrombin complex with each addition of **3g**. F_0 is the initial fluorescence of *f*-TH with buffer, ΔF_{MAX} represents the maximum observed change in fluorescence due to **3g** saturation of *f*-TH, and $[Th]_0$ is the initial concentration of *f*-TH unbound. The binding was assumed to possess a stoichiometry of 1:1. The data obtained from the fluorescence studies was input into SigmaPlot, using Equation 4. The studies revealed that **3g** has a binding affinity for thrombin of 143 ± 12 nM, with a ΔF_{MAX} of ~28%, as seen in **Figure 16A**. This thermodynamic affinity is comparable and is within the deviation of the IC_{50} (187 ± 40 nM).

With this information in hand along with the docking results, we wanted to investigate whether addition of a sulfate at the 3-(3'-hydroxylphenyl) position would incur more beneficial binding, considering the proposed binding pocket of the lone sulfate of **3g**

contained two arginines. As synthesized and described in Chapter 2, Section 1, the compound **3i** contained the 3-(3',4'-di-*O*-phenylsulfate). To assess the binding affinity of **3i**, the same experiment as stated above was implemented, but instead of titrating in **3g**, **3i** was used instead. The results we observed were to our surprise that, **3i** showed a > 15-fold loss in affinity (K_D = 2171 nM; ΔF_{MAX} = ~57%) (**Figure 16B**). This detrimental loss in affinity suggested that the *meta*-sulfate adjacent to the *para*-sulfate was most likely causing steric clash or hindrance with the binding pocket. This further supported the likelihood of the phenyl sulfate of **3g** binding within the pocket made up of R126 and R233, because the bisulfate is too bulky to make a proper fit.

3.7 Allostery-Induced Dysfunction of Thrombin's Orthosteric Site

Understanding the mechanism behind binding and the location of CSAMs binding to thrombin is an essential aspect of drug discovery. One elusive detail missing from these GAG mimetics was how the CSAMs were inducing allosteric, partial inhibition of thrombin. We knew that proteins, like receptors, had the ability to adopt ligand-induced conformations that allowed the receptor to not produce a signal at maximal efficiency. The ligand can bind to an allosteric site or to the agonist's binding site and induce the change from there[134,136]. Likewise, ligands can also induce a type of partial inhibition on soluble, smaller, dimeric enzymes like phosphodiesterase 4[137], or hemoglobin, a tetrameric protein when CO competes with O_2 for the heme[151].

A significant difference in those systems compared to thrombin is that it is not only a conformationally flexible protease, but is also monomeric. This dramatically complicates the understanding of such a mechanism. Typically, partial response or partial ability to

bind natural substrate is viewed as a cooperative effect. For example with hemoglobin, one molecule of oxygen binds to one of the four subunits of the tetramer at slow relative rate. This induces a conformational change, and subsequent oxygen molecules can then bind at a faster rate; causing a conformational change every time[152]. For thrombin however, it does not contain multiple domains or even a dimeric structure. Therefore, comprehending on a structural level how CSAMs induce submaximal enzymatic activity, without crystallographic information, casts doubt on the phenomenon. Thus strong experimental information to support such a claim is desired. The other option is that CSAMs induce dynamic allostery, which alters the free energy of thrombin's conformational space through enthalpic and entropic changes. Distinguishing between the two type of allostery would greatly assist in understanding the mechanism of the observed partial inhibition.

3.7.1 Collisional Fluorescence Quenching

There are limited techniques available to probe conformational change of a protein without the use of NMR or x-ray crystallography, that do not require expensive technologies (i.e. surface plasmon resonance or mass spectroscopy)[153–155]. To overcome this hurdle, we implemented a system where an active site fluorophore-bound thrombin (*f*TH) could be utilized as an internal probe. The system would work as such, because when the chloromethylketone of *f*-FPRck covalently binds to thrombin's active site serine (Ser195), the arginine lays within the S1 pocket, proline within the S2 pocket, phenylalanine within the S3 pocket, and the fluorophore (in this case fluorescein) lays within the S4 pocket of thrombin's active site. Under normal and uninhibited dynamic conditions, the S4 pocket is most accessible to solvent or the environment. Thus, if a

fluorescein molecule resides in the pocket, it would be available to collisional quenching agents[156]. In this case, two collisional quenching agents were sampled, acrylamide and NaI, however the iodide quencher gave more significant quenching and was used throughout the experiments[157]. The basis of the experiments was that if **3g** or another SAM were causing full or partial inhibition, then there would be noticeable differences in the two SAMs versus the uninhibited thrombin. This is due to the conformational change in and around the active site, which would cause changes to the iodide's ability to access the fluorophore. This could be measured using the Stern-Volmer equation $\frac{F_o}{F} = 1 + K_{SV}[Q]$ (Equation 5), where the change in fluorescence intensity versus the concentration of collisional quencher is linearized; then the slopes of each experiment could be compared.

To perform the initial uninhibited experiments, 180 µL of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH 7.4) was added to a clean quartz cuvette and the background fluorescence was measured (constant λ_{EX} = 490 nm, λ_{EM} = 525 nm, 0.5 mm slit widths). To this, 20 µL of 2 µM *f*TH was added and its basal fluorescence

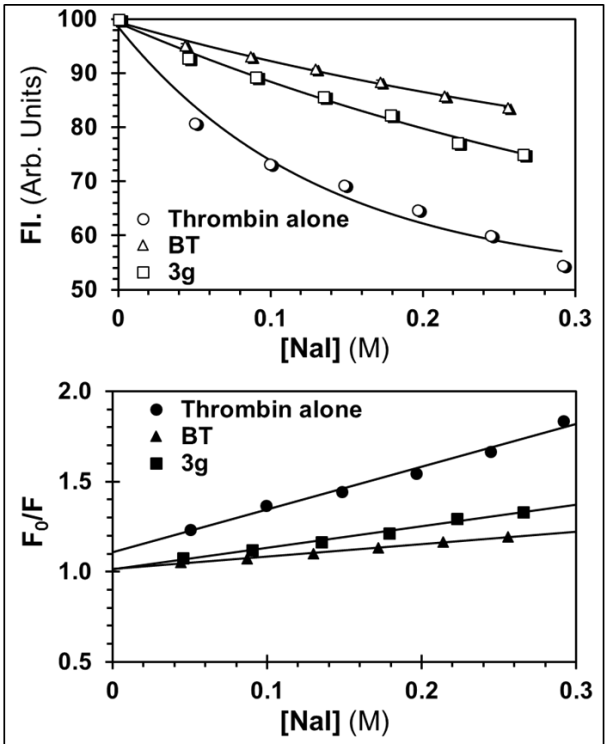


Figure 17: Fluorescence collisional quenching assays using fluorescein-labeled thrombin, sodium iodide (NaI) as a collisional quenchers in the presence of vehicle (DMSO), **3g** or **BT**. **A)** Profiles showing the differences in fluorescence quenching with respect to each compound; where FI is the normalized fluorescence intensity. **B)** The data collected from the quenching experiments were linearized using the Stern-Volmer relationship, describing the change in slope (K_{SV}) if there is a change in quenching.

intensity was measured. Then, 1 µL of 10 M NaI was added, mixed, and measured for intensity. This we repeated five times until equilibrium was reached and no more loss in fluorescence was observed. For quenching in the presence of inhibitors (**3g** and benzofuran trimer (**BT**)), we chose to keep the concentration of the inhibitors at 90% saturation (**3g** = 1 µM for 90% and **BT** = 2.5 µM for 90%), due to the compounds precipitating at high concentrations. The **3g** experiments were carried out using 177.5 µL of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH 7.4) in a quartz cuvette, monitored the same as above. Then 20 µL of 2 µM *f*TH was added and a reading was taken for the initial fluorescence. Following 2.5 µL of 80 µM **3g**, the solution was mixed, incubated at room temperature for 5 minutes, and measured for fluorescence intensity. Lastly, the 1 µL increments of 10 M NaI were added and measured individually six times. Similarly for **BT** (with the exception of volumes; 173.8 µL buffer and 6.2 µL of 80 µM **BT**), the experiments were carried out to

observe the changes in fluorescence intensity (**Figure 17A**). The intensities measured were input into an Excel file and dilution factors were taken into account to receive accurate quenching agent concentrations. The data was then applied to Equation 5, where the term K_{SV} represents the Stern-Volmer constant, which portrays the slope of the linear relationship, [Q] is the quencher concentration, F_0 is the initial fluorescence and F is the fluorescence intensity at a particular [Q].

Analysis of the data revealed that *f*TH in the absence of inhibitor produced a K_{SV} of $2.37 \pm 0.14\ M^{-1}$, corresponding to saturated quenching of the fluorescein. However in the presence of **3g**, the K_{SV} reduced to $1.19 \pm 0.06\ M^{-1}$, see **Figure 17B**. Such a significant decrease suggests that **3g** induces a conformational change of the protein's region surrounding the fluorescein in the S4 pocket. Likewise, *f*TH in the presence of **BT** showed a more significant reduction in the K_{SV} ($0.69 \pm 0.02\ M^{-1}$). This result suggests that **BT** induces a larger conformational change to the surround residues around S4. In comparison to the control (absence of inhibitor) to **3g** and **BT**, the **3g**'s K_{SV} is 50% of the uninhibited, while the **BT**'s K_{SV} is 71% of the uninhibited. If this corresponds to the ΔY values generated through dose-response profiles, the ΔY of **3g** (47%) compares very well with the percentage of K_{SV} (%K_{SV}) observed in the quenching studies. The comparison of **BT**'s ΔY (79%) also matches well with the %K_{SV} of 71%. These results signify that differential conformational changes can be achieved with small molecule, allosteric modulators of thrombin, and the extent of the changes depends on the ligand used for inhibition.

3.7.2 Antithrombin Inactivation Rate of Thrombin

The results discussed above show definitive examples that there is indeed a modulation of conformational changes to thrombin's structure, using different exosite 2 targeted ligands; resulting in reduced catalytic efficiency (k_{cat}/K_M). However to be thorough and subject this theory to scrutiny, we chose to develop a label-free assay to investigate the change to thrombin's active site induced by **3g** using an endogenous biomolecule known to exclusively interact with thrombin's active site. This assay utilized antithrombin III (AT), an endogenous serpin, known to contain an inactivation loop that covalently inactivates thrombin's active site serine[92,96]. The methodology includes using excess AT in the presence of thrombin over a period of 6 half-lives, to obtain the kinetic rate of inactivation. If the presence of increased concentrations of **3g** (from 0 to near saturation) show a decrease in the rate compared to an uninhibited system, then **3g** is truly causing allosteric dysfunction of the catalytic triad residues, and embellishes on the biophysical mechanism of how **3g** is partially inhibiting thrombin. It is of importance to note that the antithrombin reaction is uncatalyzed, meaning unfractionated heparin was not used within the experiment. We eliminated the use of UFH because in the presence of heparin, antithrombin inactivation of thrombin is accelerated by ~4000-fold. This would render us unable to perform this type of experiment because the reaction kinetics would be too fast to observe.

Initial experiments using the uninhibited system contained a fixed concentration of 6 nM of α-thrombin incubated with 0 mM **3g** (vehicle = DMSO) in buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH 7.2) for 10 minutes at 25 °C. The reaction was initiated by addition of a fixed concentration of 100 nM of AT. For the first

time point (t = 0 seconds), an aliquot of sample was mixed with 2 mM Spectrozyme TH and monitored on the Flexstation for 60 seconds at 405 nm. The Spectrozyme was used to quantify the amount of active thrombin available, and the t = 0-time point would be used to normalize the data. Subsequent time points were taken at 240, 480, 960, 1920, 3840, and 7680 seconds by removing aliquots of the reaction and mixing with Spectrozyme TH. Similarly, the inhibited system containing either 75, 225, or 675 nM of **3g** were investigated using the same time points and procedure for monitoring.

Analysis of the raw data involved first inputting the slopes obtained from the Flexstation into an Excel spread sheet, taking averages of the individual experiments for each concentration of **3g** used. To normalize the data, each averaged slope at its respected time point was divided into the zero-time point, where $t_{0s}/t_{0s} = \%_{\text{normalized 0s}}$, $t_{240s}/t_{0s} = \%_{\text{normalized 240s}}$, etc. The measured activity from the beginning of the experiment to the last time point was input into the standard exponential decay equation $[A] = [A]_0 e^{-k_{obs}t}$ (Equation 6) assuming pseudo-first order rate constant, the observed rate (k_{obs}), at each concentration of **3g**. The intrinsic second-order rate constant of antithrombin inhibition of thrombin (k_{int}) was calculated by inputting the k_{obs} from Equation 6 into $k_{int} = k_{obs}[AT]_0$, Equation 7. The intrinsic rate constant (k_{int}) was then plotted against the % concentration of thrombin–**3g** complex, obtained from quadratic equation 8,

$$[TH-3g] = \left\{\frac{[Th]_o+[3g]_o+K_D-\sqrt{([Th]_o+[3g]_o+K_D)-4[Th]_o[3g]_o}}{(2)}\right\}$$, to derive the linear relationship of the different k_{int}s of antithrombin inhibition of thrombin–**3g** complex.

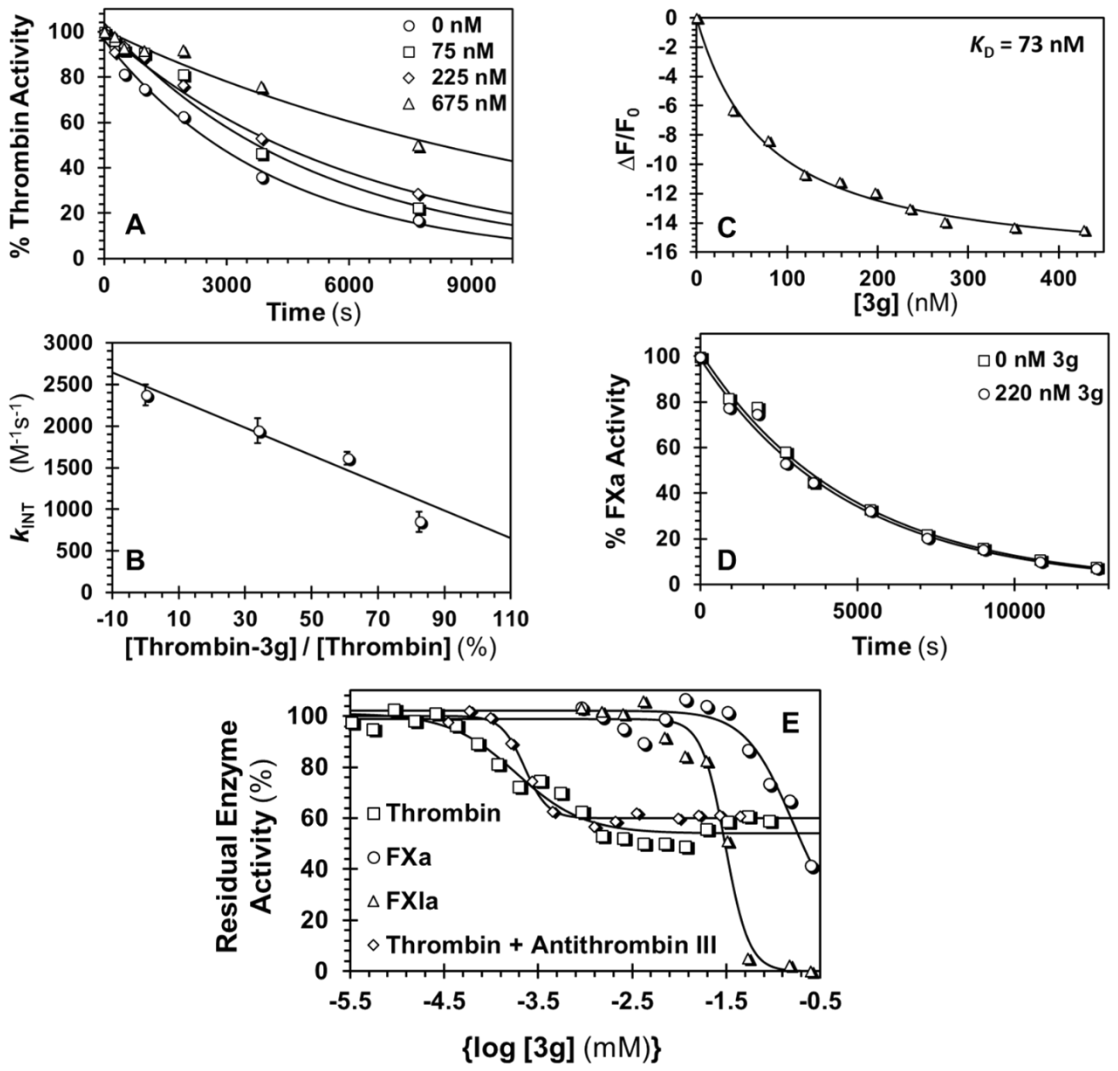


Figure 18: Antithrombin-induced thrombin inactivation studies describing **3g** inducing dysfunction of thrombin's active site, catalytic triad through an allosteric mechanism. **A)** Psuedo-first order kinetic profiles that show **3g** reducing the inactivation rate of antithrombin for thrombin in a concentration-dependent manner. **B)** The linear profile relates the calculated second-order rate constant (k_{INT}) for each concentration of **3g** *versus* the % concentration of thrombin-**3g** complex present calculated from the quadratic equation 8. **C)** The thermodynamic binding affinity of **3g** for antithrombin, calculated by 1,8-ANS displacement as described in the text. **D)** Pseudo-first order kinetic profile of antithrombin inactivation of Fxa in the presence and absence of **3g**, showing no difference in the rate of inactivation. This confirms the effects seen by **3g** for thrombin, were due to **3g** interaction with thrombin and not antithrombin. **E)** A comparative dose-response profile for **3g** against thrombin, factor Xa, factor XIa, and thrombin + antithrombin III. This aids in describing that thrombin in the presence or absence of antithrombin does not alter **3g**'s IC_{50} for thrombin.

Figure 18A displays the time profiles of thrombin inactivation by AT at the

varying **3g** levels discussed above. The comparison of the decay profiles using the control versus the **3g** experiments, show a concentration dependent reduction in the rate of antithrombin inactivation. This suggested that in the presence of **3g**, thrombin's active site is protected from antithrombin's inactivation loop. However, to understand if thrombin remains active, even at 100% **3g** saturation, the derived intrinsic rate constant must be used to compare in an expected linear manner against the calculated thrombin-**3g** concentration; which we derived with Equation 8. **Figure 18B** showed that k_{int} versus the concentration of thrombin-**3g** complex as a percent, would be expected to remain 33 ± 12% active at saturation. The ability of thrombin to remain active at **3g** saturation compares relatively well with the ΔY observed from dose-response profiles (ΔY = 47%), as well as the reduction of quenching (%K_{SV} = 50%). This observation fully supports the hypothesis that an intermediate conformational change of thrombin's active site induced by the allosteric effector **3g** can cause alteration in the catalytic efficiency and by the same principle, to thrombin's ability to cleave substrate. We propose that the inability of antithrombin to interact and inhibit thrombin in the presence of **3g** is directly caused by this conformational change of the active site (described in **Figure 19**).

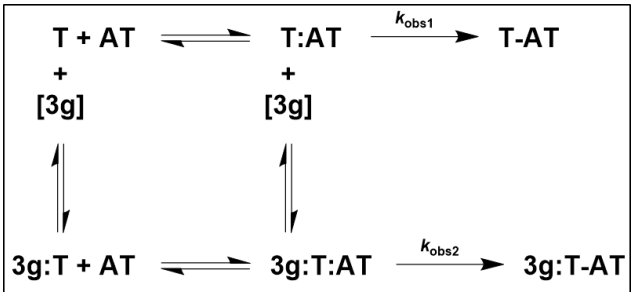


Figure 19: Proposed kinetic equation for antithrombin inactivation of thrombin in the absence and presence of **3g**. The addition of **3g** slows the rate of thrombin's inactivation, therefore k_{obs2} shown above is variable, depicted by the dotted arrow.

To further validate this, several experiments were performed to assess **3g**'s effect on thrombin inhibition in the presence of antithrombin and **3g**'s effects on antithrombin

itself. This ensured that effects observed in the inactivation assay are due solely to **3g**'s interaction with thrombin. The primary experiment was a dose-response inhibition, similar to the ones described in Section 1 of this chapter. In brief, 6 nM of thrombin, various **3g** concentrations, and 100 nM of antithrombin were incubated together in thrombin buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH = 7.4) over the course of five minutes at 25 °C. Then Spectrozyme TH was added to the solution and monitored for *p*-nitroaniline release at 405 nm. The linear slopes of the residual thrombin activity for each concentration of **3g** were input into SigmaPlot and applied to Equation 1. As seen in comparative **Figure 18E**, the IC_{50}s between the (-) AT and (+) AT show no significant difference, suggesting that the presence of AT does not affect **3g**'s ability to inhibit thrombin.

We then wanted to establish whether **3g** was binding to antithrombin. For this we used a fluorometric thermodynamic binding assay similar to the one used in Section 6. There was however a significant difference in that the antithrombin did not possess a labeled fluorophore. We turned to use of 1-anilinonapthalene-8-sulfonic acid (ANS) which non-covalently binds to protein surface, causing an increase in the basal fluorescence emission of the molecule (λ_{EX} = 405 nm and λ_{EM} = 532 nm). Then once the saturation of ANS to antithrombin is established, a subsequent titration of ligand is added stepwise and if the molecule is binding, the intensity will decrease. Once ligand binding saturation is reached, the decreases in fluorescence will cease. Thus, 1 µM of AT was incubated in thrombin buffer in a quartz cuvette and monitored at the wavelengths described above. To this, 5 x 1 µL increments of 200, 1000, and 5000 µM was added to the cuvette, mixed, and

individually measured. To remove background fluorescence, the same procedure was done in the absence of AT. Once the background was subtracted from the AT sample, it was plotted as $\Delta F/\Delta F_0$ *versus* concentration of ANS (corrected for dilution) (figure not shown). This displayed what seemed like a biphasic curve, suggesting a high and low affinity binding site. Focusing on the high affinity site, the calculated concentration of ANS needed for the displacement assay was 35 µM.

With this information, the displacement assay was constructed using in the quartz cuvette, thrombin buffer, 1 µM of AT, and 35 µM of ANS. The fluorescence intensities were taken at each addition of volume and the titration of **3g** began using 7 x 1 µL increments of 8 µM **3g**, followed by 3 x 1 µL of 16 µM **3g**. Again dilution factors were taken into consideration and the readings of $\Delta F/F_0$ versus the concentration of **3g** was input into Equation 4. The calculated K_D from the experiment showed that **3g**'s binding affinity to antithrombin was 74 nM and the ΔF_{MAX} was 17 % (**Figure 18C**). That is 2-fold higher affinity than **3g** for thrombin. This was worrisome, so one last experiment was developed to resolve the potential issue.

Through our previous studies on **3g** with factor Xa, we knew that **3g**'s IC_{50} for FXa was ~160 µM, well beyond the K_D's of thrombin and antithrombin. Additionally, antithrombin inactivates FXa using the same mechanism as it does for thrombin. Thus, we developed a FXa inactivation assay to probe whether the binding of **3g** to antithrombin affected the rate of inactivation of FXa. The assay was set up the same as above, where instead of using FXa was incubated in the presence of excess AT and vehicle (DMSO) or 220 nM **3g**. We used this concentration of **3g** because a significant change in antithrombin

inactivation of thrombin was seen at the same concentration and we know **3g** will not inhibit FXa at this concentration. The reaction was initiated by incubating 1.1 nM of FXa with 100 nM of AT with vehicle or **3g**. A total of ten time points between 0 and 12600 seconds were taken by aliquoting a volume of reaction and measuring activity with Spectrozyme Xa. The two sets of raw data were input into SigmaPlot and assessed using Equation 6. The results shown in **Figure 18D**, reveal no observable change in the k_{obs} for 0 or 220 nM of **3g**.

The results taken herein overwhelming support the theory that **3g** was in fact inducing direct, allosteric, partial inhibition of thrombin. **3g** acts by causing dysfunction of thrombin's active site where it remains active against small chromogenic substrate, even at saturated inhibitor concentration. Although experimental results suggest **3g** also binds tightly with antithrombin, there is no evidence that the binding caused a disruption of antithrombin's inactivation loop. Most likely, **3g** binds to a distal location on antithrombin that does not cause any significant structural changes that would render the inactivation loop compromised.

CHAPTER 4 Endogenous Substrate Hydrolysis: Examination of Inhibition *in vitro*, *ex vivo*, and *in vivo*

The notion of using thrombin as a key target for anticoagulant therapy arises from experimental evidence that the alteration of its activity leads to a range of different effects downstream in the system. It has been shown that thrombin's reactive specificity is contingent on its binding partner[109,158,159]. In the example of thrombomodulin (TM), thrombin bound to thrombomodulin loses the ability to hydrolyze fibrinogen, but increases the k_{cat}/K_M by ~1000-fold for hydrolysis of protein C[160]. This phenomenon is theorized to occur for several reasons. First, thrombomodulin binding to exosite 1 and to some extent exosite 2 (through interaction with chondroitin sulfate present on 20-35% of TM) may act as a supplementary exosite for protein C to bind and alter conformation to favorably bind within thrombin's active site[161,162]. Another theory is that TM binding alters thrombin's active site to favor protein C hydrolysis and not fibrinogen[163,164]. Additionally, it has been suggested that TM may conformationally alter the sodium binding site on thrombin, which in turn reduces the cationic effects seen by the activation peptide of protein and thus influences a favorable conformation for protein C to be hydrolyzed[160]. These effects may act individually or collaboratively.

This notable, induced change to the k_{cat}/K_M, which results in the lower activation energy for protein C in the presence of TM was partially the driving force behind the development of NSGMs and the CSAM library itself. We have shown that **3g** induced a reduction in the catalytic efficiency of thrombin for small substrate, but up until this point, we have not investigated whether the same type of partial inhibition is carried through for the endogenous substrates of thrombin. These substrates of particular interest include fibrinogen, factor XI, protein C, and factor XIII. Therefore, we began studies into the understanding of the mechanism of **3g**-induced partial inhibition of thrombin as it pertains to endogenous substrates.

4.1 In vitro Fibrinogenolysis

The major substrate for thrombin and the protein responsible for fibrin clots, fibrinogen exists as a soluble, free-flowing 340 kDa protein in plasma at concentrations of 2-4 g/L[165]. It is a dimeric protein consisting of three pairs of nonidentical subunits (Aα, Bβ, and γ). Upon hydrolysis by thrombin, insoluble fibrin monomers are produced and are subsequently crosslinked through a transglutamination reaction by factor XIIIa. The crosslinked fibrin forms a mesh with platelets, trapping red and white blood cells, forming a thrombus or occlusion within a blood vessel. As previously mentioned, current therapies that target thrombin result in the absence of fibrin production, which can have detrimental effects on a patient[74,85,102]. With **3g** showing partial inhibition of thrombin against small substrate, it created curiosity to whether the same phenomenon would occur against fibrinogen. The possibility that **3g** could induce partial fibrinogenolysis, could allow for a

more sensitive regulation of thrombosis, without abolishing thrombin's innate activity towards the substrate.

To investigate this property of sensitizing thrombin to fibrinogen by **3g**, a transmittance assay was performed in triplicates to examine the amount of fibrin clot formed in the presence of inhibitor. Human fibrinogen was purchased from Haematologic Technologies Inc. (Essex Junction, CT) and used in place of Spectrozyme TH as thrombin's substrate. In brief, 85 µL of buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH = 7.4), 5 µL of 120 nM thrombin, and 5 µL of varying concentrations of **3g** (0.002-250 µM) were mixed in wells of a 96-well plate. Once thoroughly mixed, 5 µL of 5.3 mg/mL of fibrinogen was added to each well simultaneously and placed in the Flexstation. To assess transmittance, the plate reader was set to scan the wells at 600 nm for the first 100 seconds after the addition of fibrinogen. The slopes generated are directly related to the amount of fibrin clot formed in each well, thus the amount of light transmitted through the well during clot formation, assessing thrombin's rate of activity toward fibrinogen. The slopes and corresponding concentrations (log_{10}) were inputted into SigmaPlot and Equation 1 was applied to calculate IC_{50}, ΔY, and HS.

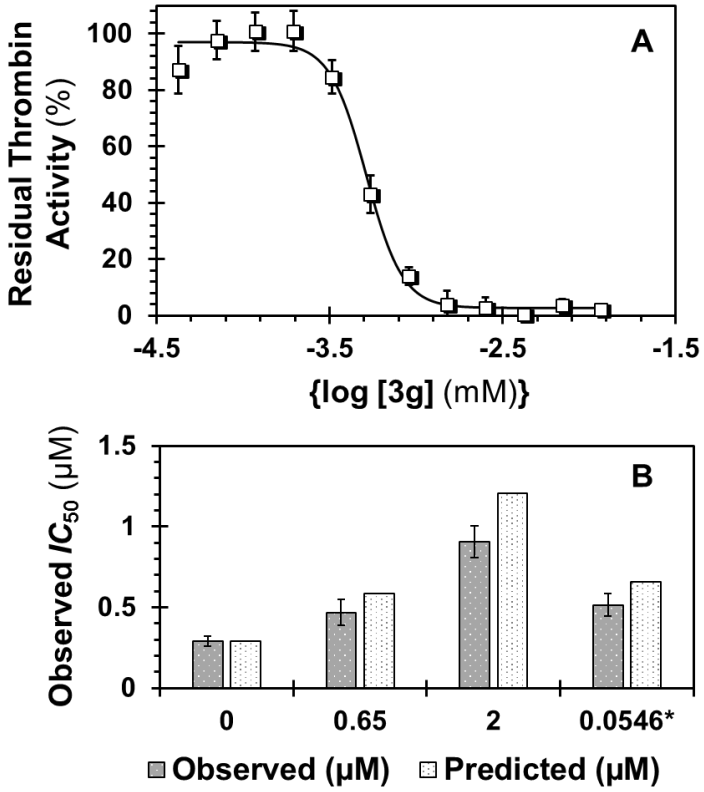


Figure 20: A) Dose-response profile of **3g** and thrombin in the presence of fibrinogen as the substrate; using a transmittance assay to show the amount of fibrin produced. **B)** Stern-Volmer relationship observing the apparent competitive trend at exosite 2 between **3g** and γ'-fibrinogen. The last entry "0.0546*" is the calculated amount of γ'-fibrinogen in the sample in **A**.

Much to our surprise, the dose-response profiles (**Figure 20A)** generated from the experiments showed that **3g** indeed inhibited thrombin from hydrolyzing fibrinogen, but full, not partial inhibition was observed (ΔY = 94%). Additionally, the IC_{50} increased 2.3-fold to 514 nM, which was a peculiar event considering the thrombin concentration was kept at 6 nM throughout all experiments performed, as used previously. After investigation of the literature, a few explanations could be drawn about why there may be loss of partial inhibition and loss of potency. About 7% of human fibrinogen contain an additional 20 amino acid sequence called γ'-fibrinogen, caused by an alternative mRNA processing event in the γ-chain[167], as well as significant posttranslational modifications[168]. This results in a sequence which contains a tyrosine sulfate and binds with thrombin's exosite 2 with a K_D of ~0.18 μM and ~0.63 μM[168–170], respectively. This γ'-chain may be interacting with thrombin within the experiment and compete with **3g**, causing the increase in IC_{50}.

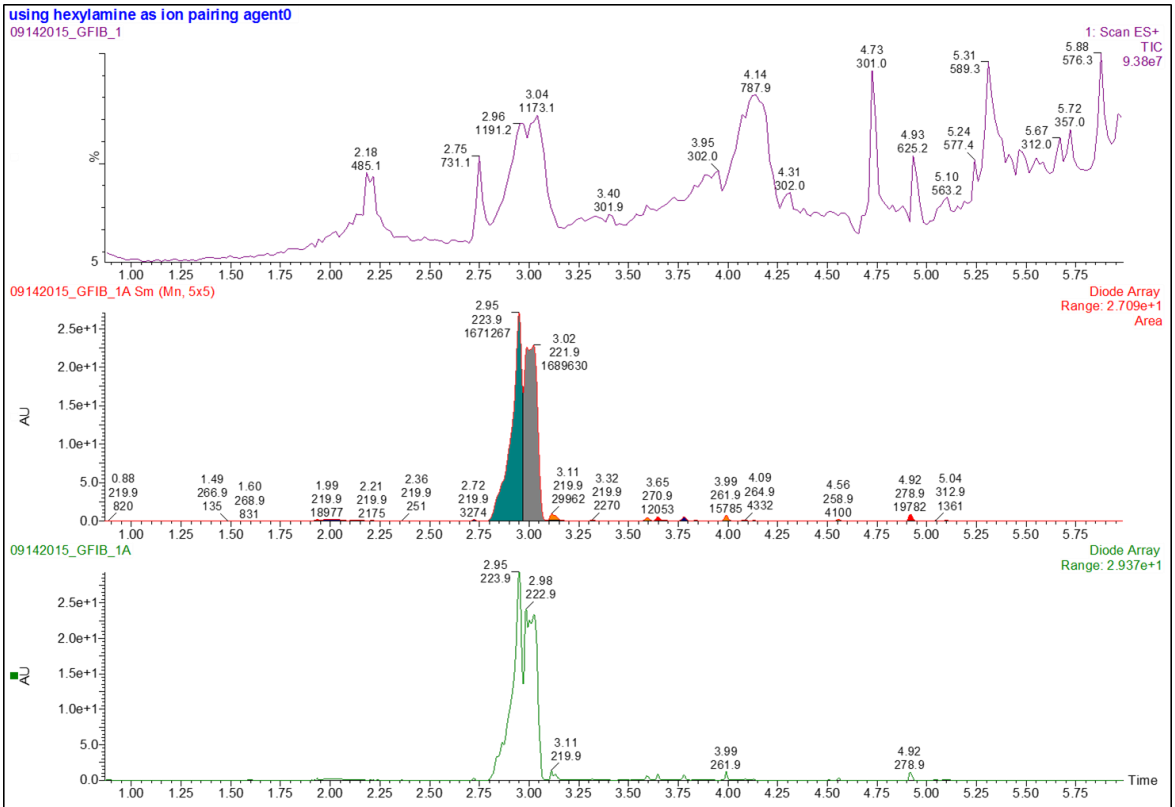


Figure 21: Mass spectra coupled with chromatography with a UPLC showing the retention times and masses of the recombinant γ'-fibrinogen sample. Top graph is MS scan, middle graph is the integrated chromatograph, bottom is the original chromatography from 0-6 minutes.

A recombinant γ'-peptide with phosphorylated tyrosines (PEHPAETE-Y(PO_3^-)-DSL-Y(PO_3^-)- PEDDL, replacing the tyrosine sulfates, was obtained from Dr. Farrell at Oregon Health and Science University, as a free flowing powder. We performed competition studies using 0, 0.65, and 2 μM of γ'-chain (γ'fibP) using the same protocol as in Chapter 3 Section 3, however the results were inconclusive and showed no competition. We then looked at the γ'fibP itself and calculated the molar extinction coefficient at a wavelength of 280 nm (ε_{280}) for tyrosine and tryptophans present on the peptide. Assuming the sample was pure, the γ'fibP powder was weighed, and calculated to make an 80 μM solution using its known average molecular weight (2280.12 Da). The solution was

transferred to a quartz cuvette and placed inside a UV spectrophotometer to observe the absorption at 280 nm. The absorption revealed that compared to a calculated 80 µM, the concentration in the sample was 70% lower than that according to the Beer's Law (A = εLc; where A is absorbance, ε is the molar extinction coefficient, L is the pathlength of the cell, and c is concentration of solution).

Subsequently, the same solution was injected into the UPLC-MS to analyze the contents and to verify the peptide was in the sample. The analysis showed that at a retention time of 3 minutes, two major peaks corresponding to 1191.2 m/*z* and 1173.1 m/*z*, respectively (as seen in **Figure 21**) were present. After examining the MS data in conjunction with the chromatograph, we determined that the sample actually contained ~1:1 mixture of *N*-acetylated (2322.15 Da) and *N*-free (2280.12 Da) γ'fibP. This was confirmed by integrating the chromatograph peaks and taking the percentage of the *N*-acetylated peak corresponding to 1191.2 m/*z*, which related to the doubly charged species with hexylamine, $[M+C_6H_{13}N]^{2+}$ and comparing it against the second peak corresponding to the *N*-free γ'fibP, 1173.1 m/*z*, relating to the doubly charged species with sodium $[M+Na]^{2+}$. Both peptides are considered to be in their protonated forms, and the MS was in positive mode, so only one counter ion was observed for each.

After confirming the sample did contain variants of γ'fibP, we adjusted the stock volumes to use in the competition assay to again use 0, 0.65, and 2 µM and repeated the assay; assuming the K_D = 0.63 µM as described in literature for the peptide. In this instance, the comparison of IC_{50}s generated from the dose-response profiles with γ'fibP (**Figure 20B**), showed a general trend of increasing IC_{50}s as the concentration of γ'fibP

increased using Dixon-Webb analysis. However the system does not seem to behave as a truly ideal competitor, as expected for a competitor of exosite 2, because there remains a significant difference in the observed *versus* predicted IC_{50}s. Considering the trend seen in **Figure 20B**, it can be suggested that the 1:1 mixture of *N*-acetylated and free-NH_2 γ'fibP causes a slight difference in binding and thus *are* competitors of **3g** against thrombin.

With this information, we then used the average of 7% to calculate the theoretical amount of γ'-fibrinogen present in the original 5.3 mg/mL (15.59 μM) fibrinogen sample for the fibrinogenolysis assay. The concentration of γ'-fibrinogen expected to be present was 0.0546 μM in each well. The IC_{50} generated from the fibrinogenolysis (514 nM) was input into the Dixon-Webb analysis (Equation 3) using a K_D of 0.2 μM, corresponding to the literature[169]. This result showed a predicted IC_{50} of 655 nM compared to the 514 nM reported by the assay (**Figure 20B**). Although not absolutely convincing, the analysis suggests that the natural γ'-fibrinogen may be competing with **3g** at exosite 2. We do know the exact concentration of γ'-fibrinogen present in the sample considering the concentration varies by source, but this is a relative estimate and the results support competition.

These data give support for the reasoning of the 2.3-fold loss in potency of **3g** for thrombin using fibrinogen as the substrate, but does not account for the loss of partial inhibition. This explanation requires a more theoretical approach. Considering the exosites and active site are interconnected[145], the major binding of fibrinogen to exosite 1, coupled with **3g** binding, may cause an overall structural change. so that the subsequent conformational change completely blocks the fibrinogen from gaining access to the active

site; but at lower potency because the conformation of exosite 2 is not favorable for **3g**. A similar event was seen in the HirP competition studies, where the IC_{50} increased by ~50%, but did not go beyond that; even as the concentration of HirP increased. The conformational changes resulting from the HirP at exosite 1, while **3g** was binding to exosite 2, supports an altered binding site for **3g**. Although we did not see a loss of partial inhibition, the sheer size difference between a small peptide and a 340 kDa protein binding to thrombin at both exosites, may cause enough conformational changes to completely lose partial inhibition.

4.2 In vitro Protein C Activation

Another protease within the web of coagulation, protein C in its activated form is a natural anticoagulant found endogenously in the plasma of mammals. The purpose of activated protein C (aPC), is to inactivate the cofactors for factor IXa and factor Xa, factors VIIIa and Va respectively; through a proteolytic event. The inactivation of these result in a vast reduction in the rate of IXa and Xa's activity; therefore causing a significant loss of thrombin production *in vivo*. As mentioned, thrombin is the protease responsible for activation of protein C through a hydrolysis reaction with its cofactor thrombomodulin. Upon thrombin binding to thrombomodulin, it is believed that the complex facilitates first binding with protein C through a supplementary exosite on thrombomodulin, then inducing a conformational change of protein C. There are also arguments as to whether thrombomodulin induces allosteric changes to the active site of thrombin upon binding. According to work done by Fruentes-Prior, P. and Bode, W. *et al.*, the binding of thrombomodulin to thrombin does not show marketed structural changes to thrombin's

active site, as shown by high resolution crystal structures; thereby suggesting the thrombin-TM complex's activity is solely due to conformational changes of protein C caused by TM[171]. Opposition to that theory from Koeppe, J. R. and Komive, E. A *et al*., suggests that the binding of TM to thrombin causes a long-range tightening of the loop on thrombin containing residues 126-132, which in turn alters substrate specificity[164]. This theory was supported by $^{1}H/^{2}H$ exchange studies, fluorescent probing of the active site, and crystallography.

Once in the activated form, aPC utilizes a cofactor, protein S, to facilitate and accelerate the inactivation of factor Va[172]. The aPC-protein S complex also overcomes a biochemical hurdle by increasing the rate of hydrolyzing Arg^{506} of factor Va by ~20-fold[173], as well as thwarting the innate protection of FVa's Arg^{506} by FXa when in complex[173]. Due to aPC's important role as an *in vivo* anticoagulant, there is an absolute necessity to understand how **3g** affects not only activated protein C itself, but to discover whether **3g** has any effect on the thrombin-TM complex as it activates protein C.

To probe this system, activated protein C was obtained from Haematologic Technologies Inc. (Essex Junction, CT) and used in a chromogenic substrate hydrolysis assay. In 85 µL of buffer (50 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG

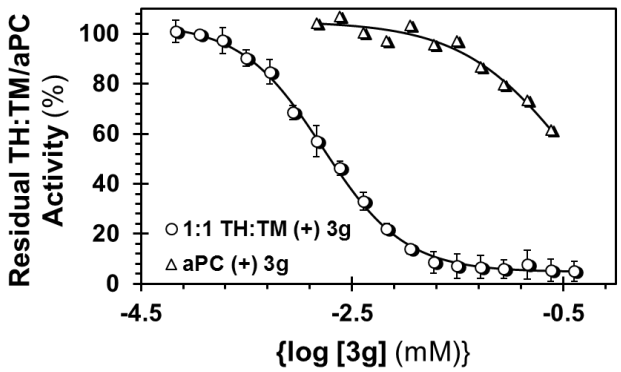


Figure 22: Herein represent two different dose-response profiles. First of **3g** toward activated protein C (aPC) (triangle). The other profiles are coupled assays where thrombin and **3g** are incubated in the presence of cofactor thrombomodulin (TM), then protein C is added as the initial substrate. Thrombin's activity is then abolished with argatroban and S-2366 is added to asses the amount of active protein C at each **3g** concentration.

8000, pH 7.2), 5 µL of 500 nM of activated protein C was incubated with 5 µL of varying concentrate ions of **3g** or vehicle (DMSO) for 10 minutes. Then, 5 µL of 6.9 mM S-2366 was added and monitored at 405 nm for 300 seconds. The slopes generated from the Flexstation were input into SigmaPlot and applied to Equation 1 for IC_{50}, ΔY, and HS.

The dose-response profile generated showed that **3g** does not inhibit aPC significantly, with an IC_{50} of 346 µM. This aids in confirming **3g**'s selectivity towards thrombin, but also that the activated protein C produced in the following coupled experiment, would not interfere with the readings obtained. From here we devised a coupled thrombin-protein assay, where a fixed concentration of thrombin (6 nM) would be incubated with trabbit-derived thrombomodulin (10 nM) in the presence of a range of **3g** concentrations (0.002-250 µM) or vehicle. To the incubated solution, excess protein C (400 nM) would be added and the solution would react for 10 minutes. To quench the reaction, argatroban (800 nM) in a buffer containing 0.02% tween 80 would be added to inhibit thrombin fully and break up any aggregates. After the quenching, chromogenic substrate (S-2336) for protein C would be added and monitored at 405 nm for production of *p*-nitroaniline. The slopes generated would be fitted to Equation 1 using SigmaPlot.

Completion of these experiments yielded some interesting results. First the experiments containing 0 nM thrombomodulin (TM) did not show any response with the addition of S-2366, signifying that **3g** in the presence of thrombin, without TM does not activate protein C. We anticipated this result, as the activation of protein C requires TM, but we did not know if **3g** would also induce activation of protein C. However, as the concentration of TM increased to 10 nM, we saw a dose-dependent inhibition of thrombin-TM activation of protein C having IC_{50} of 1.65 μM. Under normal conditions, the thrombin-TM complex would activate protein C without issue, as seen at low concentrations of **3g** (**Figure 22**). However in a dose-dependent manner with **3g**, the thrombin-TM complex was inhibited, not allowing for protein C to become activated. This posed an interesting problem in terms of this scaffold becoming a viable drug candidate. If **3g** was inducing inhibition of protein C activation in the presence of thrombin and TM, then it would be inhibiting a pathway crucial for the homeostasis of coagulation.

On the other hand, if one looks at these results from a biophysical standpoint, the dose-dependent inhibition of thrombin-TM complex demonstrates several things. It first validates the potential for a small molecule to induce significant alteration to a protease complex through allostery. Although the partial inhibition as seen with small substrate was not carried through to this substrate, the **3g** still has a significant impact on the complex. The **3g**-thrombin-TM complex formed in the assays reversed the innate mechanism of thrombin-TM significantly. This suggests that **3g** significantly altered the active site of thrombin even when TM was bound. The phenomenon of a small molecule, exosite 2

inhibitor reversing protein C activation is the first of its kind and has not been reported in literature to our knowledge.

4.3 In vitro Factor XIIIa Activation

The general trend seen with thrombin is that it has many endogenous substrates and considerably more when bound cofactors. Under normal and thrombotic states, after thrombin cleaves fibrinogen into fibrin monomers, it also activates factor XIII (existing in either plasma or on platelet surfaces). The activated FXIII (FXIIIa) is a transglutaminase, which exists *in vivo* as a zymogen (FXIII). The plasma zymogen (320 kDa) is a tetramer in the absence of calcium and consists of two non-identical pairs of α and β subunits ($\alpha 2\beta 2$)[62]. Upon the activation of FXIII, thrombin hydrolyzes the bond between the Arg36 and Gly37 of the α subunits and releases a two signal peptides per tetramer. This forms the activated intermediate FXIIIa (α_2'β2), which after binding to calcium causes dissociation of the β subunits from the α. The FXIIIa now possesses the ability to cross-link the fibrin monomers through a transglutamination reaction[174]. This cross-linking accelerates the build-up of a thrombus by forming the platelet-fibrin mesh. In an up-regulated thrombotic event, this reaction continues, which eventually can lead to a blockage or a distal embolism. However, absence of FXIIIa activity causes severe bleeding complications[175]. Because of this, thrombin-**3g** interaction in the presence of FXIII zymogen was studied to assess whether **3g** would induce full, partial, or no inhibition of FXIII activation.

The method of validation utilized SDS-PAGE to observe the activation of FXIII by thrombin through monitoring the α and β subunit migration on gradient 4-15% gels. Under non-reducing conditions in the presence of calcium, the two subunits separate giving two

distinguishable bands on the gel (**Figure 23A**). When the α subunit is activated by thrombin, the α_2 band loses ~ 8 kDa from the activation peptide cleavage and migrates down towards the β subunit. This distinct band shift in migration can used to analyze activity of thrombin in the presence and absence of **3g**.

To first understand the kinetics of 1 NIH unit of thrombin (7.8 nM) against 200 nM of FXIII, the reaction was carried out in 200 µL containing the 7.8 nM thrombin and 200 nM FXIII from 0-30 minutes, and 0, 20-60 minutes taking 20 µL aliquots every 2.5 minutes and 5 minutes, respectively and mixing 1:1 with 2x Laemlli buffer (non-reducing), and denaturing the sample at 95 °C for 5 minutes. From each sample, 20 µL was transferred into wells of 7.5% Mini-Protean TGX precast 10-well gels (BioRad) and electrophoresed using 1x SDS running buffer at a constant 30 mA. After 50 minutes the gels were washed with high purity water and SDS was removed from the gels by microwaving the gels in water baths three times and each time allowing them to cool down for five minutes on an orbital shaker. The water from the baths was discarded and the gels were stained using Coomassie brilliant blue G-250. After ~1 hour, the gels were destained and analyzed via an Amersham Imager 600 (GE).

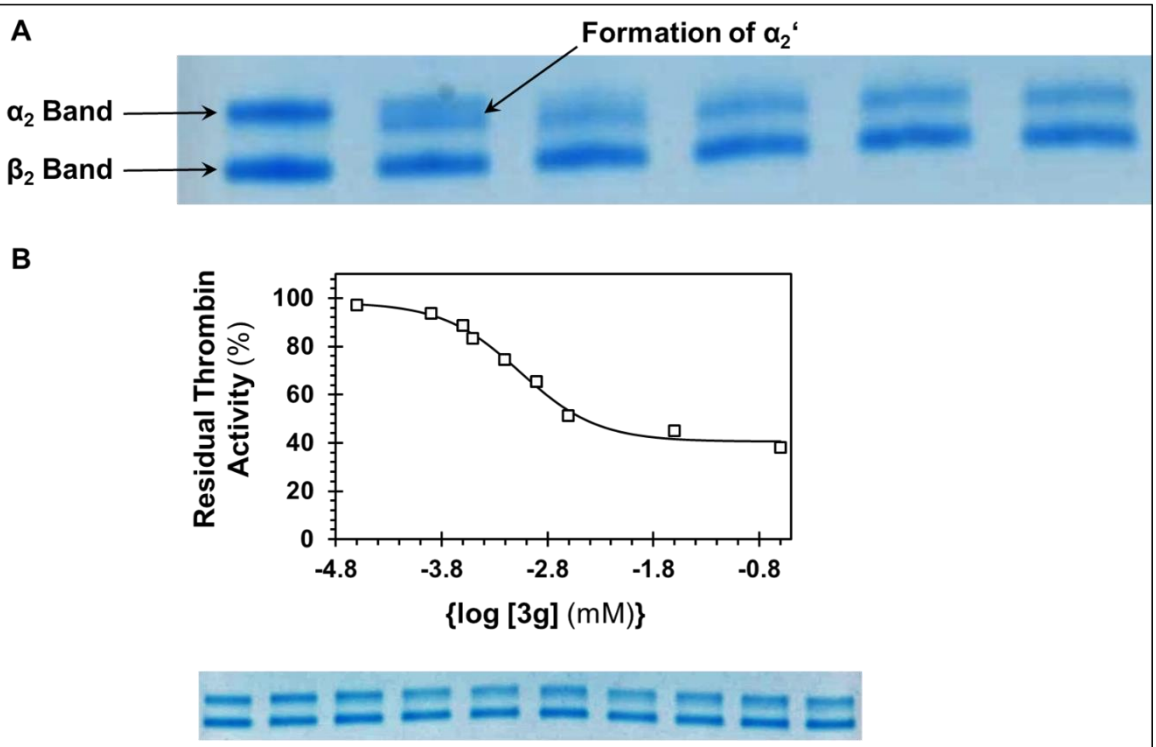


Figure 23: Monitoring FXIII activation by thrombin under various conditions by SDS-PAGE. **A)** The bands on the gel show the α_2 and β_2 bands in the first well prior to thrombin hydrolysis. Over time as thrombin hydrolyses FXIII, the α_2' band begins to form and migrates toward the β_2 band. **B)** Dose-response profile of FXIII activation by thrombin in the presence of 250-0 μM of **3g** generated from the gel below the profile.

The first gel showed that after 30 minutes, the appearance of a migratory band representing the α_2' subunit was beginning to appear. The second gel, which monitored a longer period of time (0, 20-60 minutes), showed that a time-dependent increase of α_2' was being produced (see **Figure 23A**) The reaction was not driven to completion at 60 minutes, but the substantial visual amount of α_2' present on the stained gel would allow for a dose-dependent experiment to be performed using an incubation time of 60 minutes. The experiment would show whether thrombin in the presence of **3g** would inhibit, reduce or accelerate the production of FXIIIa *in vitro*.

In brief, the reaction took place using 600 µL microcentrifuge tubes, solutions of 200 µL buffer (20 mM Tris-HCl, 100 mM NaCl, 2.5 mM $CaCl_2$, 0.1% PEG 8000, pH 7.3) contained 7.8 nM of thrombin and varying concentrations of **3g** (0-250 µM) or vehicle. After a five-minute incubation period, FXIII was added to each tube to bring the concentration to 200 nM and allowed to react for 60 minutes. Aliquots of the reaction were quenched by mixing 20 µL of the sample with 20 µL of 2x Laemlli sample buffer (non-reducing). Each mixed sample was denatured by incubating at 90 °C for 5 minutes and then immediately placed on ice until ready for analysis. For gel electrophoresis, a 4-15% Mini-Protean TGX precast 10-well gel (BioRad) was used to which 10 µL of PageRuler 10-250 kDa (Thermo Scientific) was added to well number 1, while 20 µL of denatured, mixed sample was added to the subsequent wells (2-10) in the order of increasing concentration (0-250 µM). The chamber and gel cassette were filled with 1x SDS running buffer and the gel was run at a constant 30 mA for ~55 minutes. Upon completion, the gel was removed from its casing, washed with high purity water, heated by microwave briefly, and was placed on an orbital shaker for 5 minutes. The water was discarded, and the wash/heating step was repeated twice more. After the last water discard, a solution containing Coomassie brilliant blue G-250 dye was added to cover the gel, heated briefly, and placed on an orbital shaker for 1 hour. After the hour incubation, the stain was removed and high purity was added to destain the gel. The gel was then photographed by an Amersham Imager 600 (GE) for further analysis.

Images of the gel seemed to reveal a dose-dependent inhibition of FXIII activation, by analysis of the presence of the α_2' band discussed above. Using Quantity One gel

analysis software (BioRad), intensities of the α_2' band in each well were compared against the mean background intensity of the lanes and the α_2' band in the well containing 0 nM **3g**. The values were normalized (X_{norm}) using the equation $\frac{X-\mu}{X_{max}-\mu} = X_{norm}$, where X is the α_2' band intensity of wells containing **3g**, X_{max} is the α_2' band intensity of the well containing 0 nM **3g**, and μ is the mean background intensity; also known as the minimum X value. The data received was then input into SigmaPlot to calculate the IC_{50}, ΔY, and HS, outputting an excited dose-response profile.

As seen in **Figure 23B**, there is a concentration dependent inhibition of FXIIIa generation in the presence of **3g**. The α_2' band shows partial generation at saturating levels of **3g** indicative of partial activation of FXIIIa by thrombin. The slope generated by SigmaPlot from the gel shows an IC_{50} of 0.87 nM and ΔY of 55%. This is the first example of the thrombin-**3g** complex showing partial inhibition of hydrolysis for an endogenous substrate. As exciting as this is, it is important to mention that this is an uncatalyzed reaction. *In vivo,* fibrin acts as cofactor, binding to exosite 1 of thrombin and accelerates the hydrolysis of FXIII by 80-fold[109]. The lack of fibrin at exosite 1 may distort the results seen here. The observation that exosite 1 cofactors remove the potential for partial inhibition is an important aspect of this work and will need further investigation to understand if the partial inhibition is carried through in the presence of fibrin.

4.4 In vitro Factor XIa Activation

Thrombin has the ability to up-regulate itself through auto-activation of the intrinsic pathway. In this mechanism, thrombin activates factor XIa (FXIa), causing an

uninhibited cascade of reactions ultimately activating thrombin from prothrombin and continuing through that feedback loop[176]. A means to circumvent this phenomenon is to inhibit FXIa entirely, because it has been shown to have little bleeding consequences, and hemostasis remains largely intact[177]. Although CSAMs are shown to inhibit FXIa, the most potent inhibitor (**3g**) does so with a 150-fold less potency of FXIa compared to thrombin. However, it may stand to reason that **3g**'s partial inhibition of thrombin may affect activation of FXIa by thrombin. Therefore an assay can be used to probe the activation of FXIa by thrombin in the presence of **3g**.

Although this seemed relatively simple, a few caveats were found in the literature. The relative rate of thrombin activating FXI is very slow; so slow in fact that rate has been shown to be only 2.5 x 10^2 $M^{-1}s^{-1}$ when the concentrations of thrombin and FXI are 1:1[178,179]. Though, if a polyanionic cofactor is added (i.e., 70-mer polyphosphate or 500 kDa dextran sulfate) to a reaction where thrombin is 40-fold lower concentration than FXI, the rate drastically increases (7.0 x 10^5 $M^{-1}s^{-1}$)[178,179]. An issue with this however, is that FXI intrinsically auto activates after 5-10 minutes of incubation with a polyanion[179]. Therefore, if use of a polyanion is necessary, a control to monitor the autoactivation must be taken into account. Lastly, considering FXI is a dimer with one active site on each of its monomeric subunits, traditional monitoring via chromogenic substrate would not allow for us to distinguish between ½ FXIa of full FXIa.

To circumvent the hurdles with FXI activation, the initial steps taken involved first studying the plasma-obtained FXI in the presence of 1 NIH unit of thrombin alone. This required an 8-hour study to observe how much ½ FXIa and full FXIa was generated, by

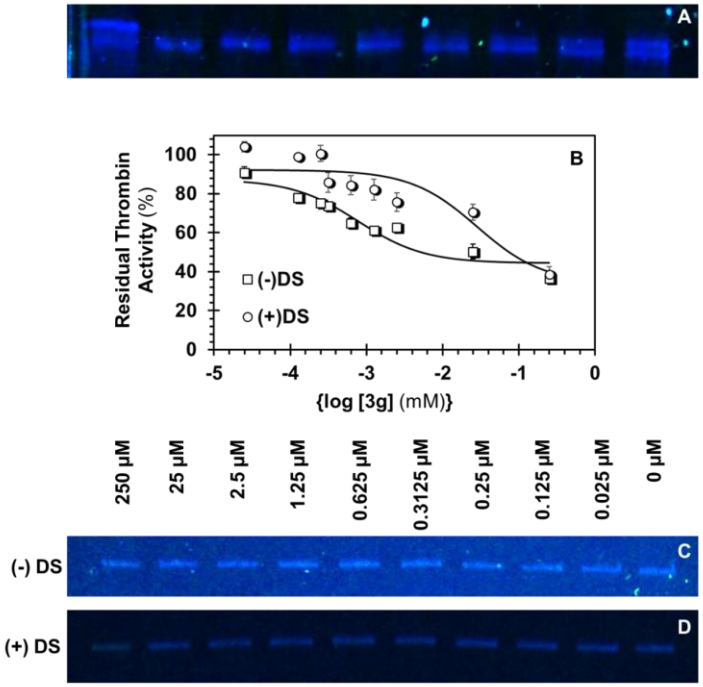


Figure 24: Monitoring FXI activation by thrombin under various conditions by SDS-PAGE. **A**) Time-based uninhibited activation of 1/2FXIa and full FXIa. The band separation indicates conformational change from 1/2FXIa to full FXIa. **B**) The dose-response profiles depict the individual concentrations of 3g and their effect on the uncatalyzed ((-)DS) (**C**) and catalyzed ((+)DS) (**D**) thrombin-FXI system. Visualization of 1/2FXIa was performed by saturating the system with fluorescein-FPRck to act as a covalent quencher of the active sites of both thrombin and FXIa, as well as a fluorogenic label.

observing activation of FXI on SDS-PAGE. As FXI is activated to first ½ FXIa and then FXIa, there is a conformational change that allows differential migration on low percentage gels, as seen in **Figure 24A**. Additionally, being ½ FXIa and FXI migrate very similarly on the gels, active site fluorophore implementation was used. Fluorescein-labeled FPRck or dansyl-EGRck was used to enhance sensitivity for visualization of the bands as seen in **Figure 24A**.

Two experiments were undertaken to assess the thrombin-**3g** complex effects on FXI. The first was to observe dose-dependence of the uncatalyzed complex on FXI over a 5-hour period using 1 NIH unit of thrombin and 200 nM of FXI. The second experiment looked at the same dose-dependence effects, but using 1 µg/mL of dextran sulfate (DS; mol. Wt. 500 kDa) as a the polyanionic cofactor over a 5 and 10-minute period. The experiments proceeded by taking a fixed concentrations of thrombin (7.8 nM) in buffer (50 mM Tris-HCl, 100 mM NaCl, 0.1% PEG 8000, pH 7.4) and incubating with increasing amounts of **3g** (0-250 µM), and a fixed volume of FXI zymogen (owing to 200 nM) at 37

°C. For the catalyzed experiment, 5 µL of 20 µg/mL of DS was added to each reaction to give 1 µg/mL in each tube. To quench all of the reactions, 3-fold molar equivalence of fluorescein-FPRck would be added and allowed to react over 30 minutes at room temperature. The tubes were mixed 1:1 with 2x Laemlli buffer (non-reducing) and 20 µL of mixed sample were loaded on 7.5% Mini-Protean TGX precast 10-well gel (BioRad) and run at 30 mA for 60 minutes using 1x SDS running buffer. The gels were removed of SDS by washing and heating with distilled water 3x. The gels were scanned using the fluorescent setting of the Amersham 600 gel imager (GE). The gels were then stained with coomassie G-250, destained, and imaged with the Amersham 600.

Upon imaging the uncatalyzed and catalyzed gels, the bands for each individual concentration of **3g** appeared to not be significantly different from one another. To obtain clarification of this, we employed the use of Quantity One gel analysis software (BioRad) to analyzed the individual bands, by comparing against the 0 nM 3g and the background. First, the uncatalyzed gels showed a dose-dependent decrease in the amount of ½ FXIa produced by the reaction. Normalizing the intensities and comparing them *versus* the concentration of **3g** showed partial inhibition with an IC_{50} of 0.782 µM (**Figure 24 B** and **C**). Alternatively, for the catalyzed gels, partial inhibition was still observed, but with an IC_{50} of 28 µM (**Figure 24 B** and **D**). This suggested that the dextran sulfate played a major role in competing with **3g** for exosite 2 of thrombin and that **3g** induced catalytic dysfunction to thrombin's active site resulting partial inhibition. This competition and change in IC_{50} is similar to what is seen in the competition studies with UFH.

4.5 Ex vivo Human Plasma Studies

One of the most important aspects involving anticoagulant drug discovery is not necessarily understanding the precise target protein or complex, but finding out if the novel agent retains its ability to inhibit clot formation in plasma and blood. It is crucial in the development process to gather as much information as possible regarding the drug's potency and if there are other proteins within the blood or plasma that it may be binding to. In this case, the CSAMs are known to bind thrombin tightly, but it also behaves relatively promiscuous with other coagulation proteins. To investigate this further, we used an *ex vivo* assay and human citrated plasma to probe whether the CSAMs act towards the intrinsic or extrinsic pathways. This is performed by measuring the activated partial thromboplastin time (aPTT) for the intrinsic pathway, and prothrombin time (PT) for the extrinsic pathway. In brief, citrated human plasma is used to chelate any calcium ions within the plasma, to ensure there is no premature activation of coagulation enzymes. The two assays differ in how they are initiated. The aPTT assay incubates with ellagic acid, to simulate "contact"-activation. Once the plasma and reagents are incubated and mixed properly, a calcium chloride solution is added, which triggers a procoagulant response. Alternatively, the PT assay incubates the citrated plasma with thromboplastin-D (tissue factor) and is then initiated by addition of calcium.

To begin, one-stage recalcification assays were performed, using a BBL fibrometer (Beton-Dickinson, Sparles, MD). To measure the aPTT, a number of various **3g** concentrations or vehicle (DMSO) were added to citrated human plasma and aPTT reagent (0.2% ellagic acid) which is incubated at 37 ºC for 4 min. Prewarmed 25 mM $CaCl_2$ is added to initiate clot formation. The time to clot formation was then recorded in seconds.

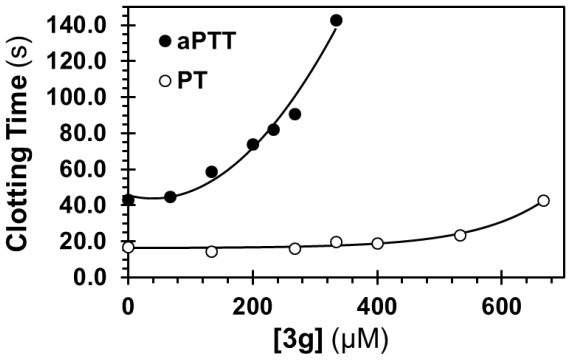


Figure 25: Plasma studies examining the intrinsic (aPTT) and extrinsic (PT) pathways of coagulation.

In the PT assay, **3g** or vehicle is added to citrated human plasma and incubated at 37 ºC for 30 sec. After the short incubation period, the pre-warmed thromboplastin-D solution was added and the clotting time is established. The concentration required to double the clotting time (2xaPTT or 2xPT) was considered the effective concentration in plasma[180].

Results from the plasma studies revealed that **3g** targeted the intrinsic pathway, having a 2xaPTT of 220 μM compared to the PT having a 2xPT of 660 μM; see **Figure 25**. This ensured that **3g** indeed behaved as an anticoagulant, but displayed a 1000-fold loss in potency in plasma. In order to evaluate this loss, we looked at the binding of **3g** to serum albumin, considering that ~55% of plasma consists of albumin and is the likely candidate for binding exogenous compounds like drugs[181]. An intrinsic fluorescence experiment used albumin at plasma concentration (45 mg/mL) with titrations of **3g**, up to 250 μM. The study showed that K_D was 184 μM, well above the K_D for **3g** to thrombin (147 nM). This led us to believe that the **3g** was not binding to serum albumin. Therefore, the binding must be occurring with immunoglobulins in the plasma or with some other serum protein. More investigations into the binding partner would need to be established to identify the mechanism, but falls out of the scope of these basic *ex vivo* experiments and more into a pharmacokinetic aspect.

4.6 Ex vivo Whole Blood (Thromboelastography) Studies

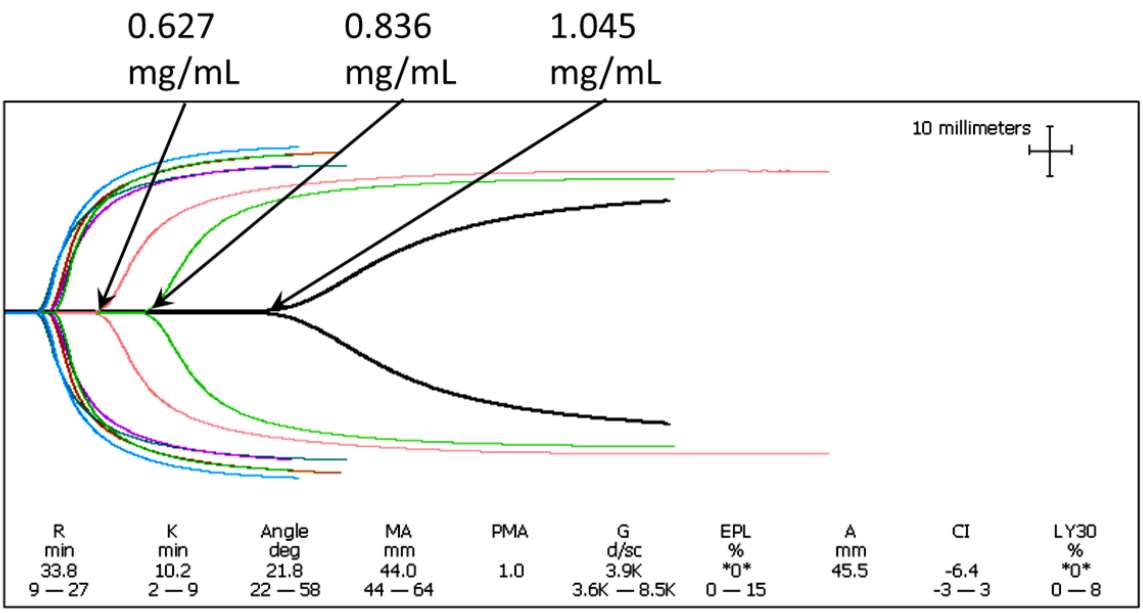


Figure 26: Whole blood studies using thromboelastography to examine **3g**-induced features of human blood. The concentrations of **3g** that show anticoagulant effects are labeled above

To establish the efficiency of coagulation, fibrin build-up, and clot strength (i.e., thrombodynamic properties) in the presence of **3g**, thromboelastography (TEG) was used. These studies were performed in the laboratory of Dr. Don Brophy (School of Pharmacy, VCU). Briefly, the thrombodynamic properties were monitored as a blood sample is induced to clot through a recalcification assay. The study was kept at a low-shear environment, representing a sluggish venous flow. The properties are expected to drastically change when comparing a blood sample with and without an inhibitor. Clot formation was monitored via a pin connected to a torsion wire located at the center of a cup, where the blood is held. The strength and elasticity of the clot applies a force to the pin and changes its rotation, which is subsequently monitored and recorded. The data is then applied to a set of parameters; maximum amplitude (MA) of clot formation (i.e.,

stability), amplification of clot strength (G), initial reaction time to the start of clotting (R), and the angle (α), which is representative of fibrin build-up and cross-linking[182].

The experiments utilized the 2-fold below the aPTT concentration to begin the experiment, and increased to the 2xaPTT concentration and increased each subsequent concentration by the same factor. The analysis showed an anticoagulant TEG profile beginning to take effect at ~3x the 2xaPTT concentration, and equivalent to the 2xPT concentration at 627 mg/mL (**Figure 26**). Clot initiation time was doubled for 0.627 mg/mL tripled for 0.836 mg/mL and at 1.045 mg/mL was having a 6-fold increase compared to the control. Additionally at 1.045 mg/mL of **3g**, the maximum amplitude and angle both drastically decrease, which may be indicative of having both anticoagulant and

	Clot initiation time	Time until fixed clot strength	Rate of clot formation	Clot strength
[3g] (mg/mL)	**R (min)**	**K (min)**	**Angle (degree)**	**MA (mm)**
1.045	33.8	10.2	21.8	44.0
0.836	17.6	3.7	47.8	51.6
0.627	11.8	3.2	49.3	53.6
0.418	5.7	1.7	67.6	64.2
0.209	4.4	1.8	64.7	58.2
0.105	5.9	1.9	62.7	59.3
Vehicle	6.4	1.6	68.5	63.5
Blank	4.8	1.5	67.7	66.1

Table 4: Whole data from each individual TEG study of varying **3g**. Blank was whole blood, vehicle was 14% ethanol solution, and **3g** concentrations increased from that point. The red box shows the concentrations that begin to display anticoagulant properties.

antiplatelet effects (**Table 4**). However, that is only based off of this empirical evidence and was not further investigated due to the extremely high concentration. It is however important to note that **3g** does behave as an anticoagulant in whole blood, albeit being at a substantial dose.

4.7 Mouse Tail-Bleed Studies

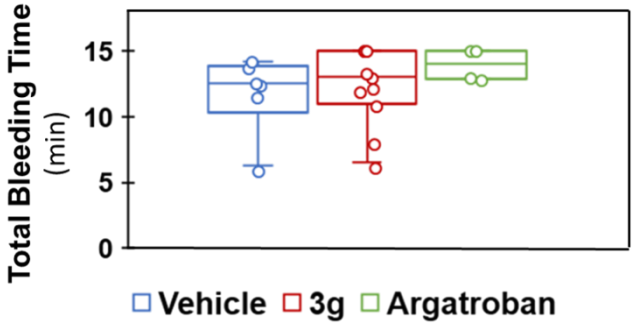


Figure 27: Statistical analysis of bleeding times from mouse tail-bleed assays are shown. Experiments used 14% ethanol as the vehicle for a negative control. Mice that were dosed (2 mg **3g**/mouse or 0.66 ug argatroban/mouse) are compared to study if any excess bleeding occurs with **3g**.

A common judgement of the efficacy of an anticoagulant is the utilization of a mouse tail-bleed assay, which is relatively non-invasive, very simple, and contain a considerable amount of information[183]. The experimental procedure is rudimentary and straight forward. A number of mice are dosed with either controls (positive: argatroban or negative: vehicle) or with drug by injection to the tail vein. A short period of time passes, imitating an incubation period, and the tip of their tail is cut. Once the initial cut is made, the timer starts and continues until a clot has formed. These times are then compared to observe if the drug causes any type of excessive bleeding at the concentration dosed.

For experiments involving **3g**, the CSAM was dissolved in 30% ethanol and diluted to 14% to be given as a 2 mg dose/mouse. An injection of the same volume containing either 14% ethanol as a negative control or 0.66 µg of argatroban as a positive control were used. The single doses were given as bolus injections to the tail vein, and incubated for 5 minutes prior to the tail cut. The times were recorded for each bleed. Analysis of the control bleeds *versus* the **3g** doses showed no significant bleeding for any of the experiments. There were no statistically significant changes in **3g** from the negative control or from the positive control as seen in **Figure 27**. This indicates that at the dosage in mice, **3g** does not cause excessive bleeding, but neither did the pharmacologically relevant dose

of argatroban; which is not surprising being < 5% of patients see adverse bleeding with argatroban and only 4 data points were taken for the positive control. To gain more confidence in these experiments, higher doses of **3g**, as well as argatroban would need be done to obtain clinical relevance. An indirect observation was that even at a 2 mg dose, the mice seemed unaffected and did not display any acute symptoms even after having an average of 1.4 mM of **3g** in their systems.

CHAPTER 5 Conclusions and Future Work

Exploitation of thrombin's exosites and inherent plasticity may be a key factor in regulation of its endogenous activity and maintaining hemostasis *in vivo*. Previous studies have used similar techniques, which utilize the notion of partial inhibition for alleviation of side effects; while still asserting response from the target protein[137]. The results reported herein tie together a number of theories with regard to a new, plausible mechanism for how to inhibit thrombin while reducing or eliminating the risk of abnormal bleeding. The potential to remove the risk of bleeding comes from the idea that partial inhibition allows for either reduced catalytic efficiency for a specific substrate, or the active site conformation may become unfavorable for interaction with certain substrates (i.e., inducing previously unobserved selectivity).

From a mechanistic point-of-view, a monomeric protease displaying partial inhibition suggests that the protease is behaving as an oligomeric protein. This is most likely due to the large conformational dynamics inherent to the protease. However, as seen in the fluorescence quenching experiments and the antithrombin inactivation, the conformation of and around the active site is changing dramatically. Leaving only the observation that structural conformational change is being induced, and inhibition is not due to conformational dynamics. This is fully supported by the collisional quencher not

being able to gain access to the fluorescein at the S4 pocket. Additional support for this is the significant change observed to thrombin-**3g** complex's recognition of protein C when it is bound to thrombomodulin. The **3g** induced a complete change to thrombin's recognition for protein C from activation under normal conditions, to inhibition of protein C activation in the presence of **3g**. Similarly, the partial inhibition of FXI and FXIII activation *versus* the full inhibition of fibrinogen cleavage *versus* recognition of protein C suggests that the alteration of thrombin's structure induced by **3g** may cause a meta-stable state of thrombin where its activity is suppressed or altered. Therefore, the extent of inhibition may have a direct correlation to the cofactors bound to exosite 1 and how that alters the thrombin-**3g** complex. Crystallographic data may be able to clarify the exact mechanism of **3g**-induced partial inhibition and work on the co-crystal structure is currently underway using thrombin-**3g** and thrombin-TM-**3g** complexes.

From a drug development aspect, **3g** is the most potent exosite 2 inhibitor of thrombin to date, but does not seem to present a viable route to the clinic; considering its loss of potency in plasma and blood, partial inhibition of FXIII activation, inhibition of protein C activation, and partial inhibition of FXI activation. The identification of the endogenous mechanism by which **3g** binds significantly is needed, so further alterations of the structure can be performed to regain the nanomolar K_D's seen *in vitro*. The use of **3g** does present itself as a valuable tool for probing thrombin and the newly discovered possibility to induce allosteric partial inhibition. If a compound could be derived from **3g** that serves as a partial inhibitor towards fibrinogen or a selective inhibitor to reduce the feedback loop of thrombin up-regulation through FXI, FV, and FVIII activation, there will

be great clinical potential in that aspect. Maintaining thrombin's activity towards fibrinogen without the side effects of the bleeding in addition to reduction of thrombin propagation would make for an ideal anticoagulant. Future work on these compounds of the partial inhibition of thrombin is needed to exploit the features mentioned here. Additional work into how **3g** affects FV and FVIII activation is crucial in this. Ultimately obtaining structural information through NMR or x-ray crystallography is of key importance to aid in the design of more potent and selective partial inhibitors.

Literature Cited

(1) López-Otín, C.; Matrisian, L. M. *Nat. Rev. Cancer* **2007**, *7* (10), 800–808.

(2) Lopez-Otin, C.; Bond, J. S. *J. Biol. Chem.* **2008**, *283* (45), 30433–30437.

(3) Koblinski, J. E.; Ahram, M.; Sloane, B. F. *Clinica Chimica Acta*. **2000**, 113–135.

(4) van Kempen, L. C. L.; de Visser, K. E.; Coussens, L. M. *Eur. J. Cancer* **2006**, *42* (6), 728–734.

(5) Quirós, P. M.; Langer, T.; López-Otín, C. *Nat. Rev. Mol. Cell Biol.* **2015**, *16* (6), 345–359.

(6) Desk, R.; Williams, L. *Biochemistry* **2006**, *114* (10), 1070–1077.

(7) Rawlings, N. D.; Barrett, A. J. In *Handbook of Proteolytic Enzymes: Second Edition*; **2004,** *1*, 231–268.

(8) Fukasawa, K. M.; Hata, T.; Ono, Y.; Hirose, J.; Fukasawa, K. M.; Hata, T.; Ono, Y.; Hirose, J. *J. Amino Acids* **2011**, 1–7.

(9) Cerdii-Costa, N.; Gomis-Rooth, F. X. *Protein Science*. **2014**, 123–144.

(10) Hooper, N. M. *FEBS Letters*. **1994**, 1–6.

(11) Wu, S.; Zhang, C.; Xu, D.; Guo, H. *J. Phys. Chem. B* **2010**, *114* (28), 9259–9267.

(12) Blow, D. M. *Trends in Biochemical Sciences*. **1997**, 405–408.

(13) Hedstrom, L. *Chem. Rev.* **2002**, *102* (12), 4501–4524.

(14) Derewenda, Z. S.; Derewenda, U.; Kobos, P. M. *J. Mol. Biol.* **1994**, *241* (1), 83–93.

(15) Henderson, R. *J. Mol. Biol.* **1970**, *54* (2), 341–354.

(16) Turk, B.; Turk, D.; Turk, V. *Biochim. Biophys. Acta*. **2000**, 98–111.

(17) Pinitglang, S.; Watts, A. B.; Patel, M.; Reid, J. D.; Noble, M. A.; Gul, S.; Bokth, A.;

Naeem, A.; Patel, H.; Thomas, E. W.; Sreedharan, S. K.; Verma, C.; Brocklehurst, K. *Biochemistry* **1997**, *36* (33), 9968–9982.

(18) Barrett, A. J.; Rawlings, N. D. *Biol. Chem.* **2001**, 727–733.

(19) Peräkylä, M.; Rouvinen, J. *Chem. - A Eur. J.* **1996**, *2* (12), 1548–1551.

(20) Seemüller, E.; Lupas, A.; Stock, D.; Löwe, J.; Huber, R.; Baumeister, W. *Science* **1995**, *268* (5210), 579–582.

(21) Brannigan, J.; Dodson, G.; Duggleby, H. J.; Moody, P. C.; Smith, J. L.; Tomchick, D. R.; Murzin, a G. *Nature* **1995**, *378* (6555), 416–419.

(22) Coates, L.; Erskine, P. T.; Mall, S.; Gill, R.; Wood, S. P.; Myles, D. A. A.; Cooper, J. B. *Eur. Biophys. J.* **2006**, 559–566.

(23) Coates, L.; Tuan, H.-F.; Tomanicek, S.; Kovalevsky, A.; Mustyakimov, M.; Erskine, P.; Cooper, J. *J. Am. Chem. Soc.* **2008**, *130* (23), 7235–7237.

(24) Coates, L.; Erskine, P. T.; Wood, S. P.; Myles, D. A. A.; Cooper, J. B. *Biochemistry* **2001**, *40* (44), 13149–13157.

(25) Cooper, J. B. *Curr. Drug Targets* **2002**, *3* (2), 155–173.

(26) Blankenship, E.; Vukoti, K.; Miyagi, M.; Lodowski, D. T. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2014**, *70* (3), 833–840.

(27) Bone, R.; Silen, J. L.; Agard, D. A. *Nature* **1989**, *339* (6221), 191–195.

(28) Davis, J. H.; Agard, D. A. *Biochemistry* **1998**, *37* (21), 7696–7707.

(29) Bone, R.; Fujishige, A.; Kettner, C. A.; Agard, D. A. *Biochemistry* **1991**, *30* (43), 10388–10398.

(30) Pasternak, A.; Ringe, D.; Hedstrom, L. *Protein Sci.* **1999**, *8* (1), 253–258.

(31) Liu, S. Q., *et al. Pesticides in the modern world*. **2011**, *16*, 333-376.

(32) Hansen, G.; Gielen-Haertwig, H.; Reinemer, P.; Schomburg, D.; Harrenga, A.; Niefind, K. *J. Mol. Biol.* **2011**, *409* (5), 681–691.

(33) Ain, Q. U.; Méndez-Lucio, O.; Ciriano, I. C.; Malliavin, T.; van Westen, G. J. P. *Integr. Biol.* **2014**, *6* (11), 1023–1033.

(34) Dorovska, V. N.; Varfolomeyev, S. D.; Kazanskaya, N. F.; Klyosov, A. A.; Martinek, K. *FEBS Lett.* **1972**, *23* (1), 122–124.

(35) Cottrell, G. S.; Turner, A. J. *Handbook of Proteolytic Enzymes.* **2013.**

(36) Olsen, J. V.; Ong, S.-E.; Mann, M. *Mol. Cell. Proteomics* **2004**, *3* (6), 608–614.

(37) Gallwitz, M.; Enoksson, M.; Thorpe, M.; Hellman, L. *PLoS One* **2012**, *7* (2), e31756.

(38) Backes, B. J.; Harris, J. L.; Leonetti, F.; Craik, C. S.; Ellman, J. A. *Nat. Biotechnol.* **2000**, *18* (2), 187–193.

(39) Newell-Caito, J. L.; Griffiths, A. E.; Fay, P. J. *Biochemistry* **2012**, *51* (16), 3451–3459.

(40) Prandoni, P. *European Journal of Internal Medicine*. **2009**, 660–661.

(41) Goldhaber, S. Z.; Bounameaux, H. In *The Lancet* **2012**, *379*, 1835–1846.

(42) Prandoni, P.; Bilora, F.; Marchiori, A.; Bernardi, E.; Petrobelli, F.; Lensing, A. W.; Prins, M. H.; Girolami, A. *N. Engl. J. Med.* **2003**, *348* (15), 1435–1441.

(43) Eliasson, Å.; Bergqvist, D.; Björck, M.; Acosta, S.; Sternby, N. H.; Ögren, M. *J. Thromb. Haemost.* **2006**, *4* (9), 1897–1902.

(44) Hong, C.; Zhu, F.; Du, D.; Pilgram, T. K.; Sicard, G. A.; Bae, K. T. *Atherosclerosis* **2005**, *183* (1), 169–174.

(45) Voetsch, B.; Loscalzo, J. *Arterioscl. Thromb. Vascul. Biol.* **2004**, 216–229.

(46) Previtali, E.; Bucciarelli, P.; Passamonti, S. M.; Martinelli, I. *Blood Transfus.* **2011**, *9* (2), 120–138.

(47) Wolberg, A. S.; Mackman, N. *Arterioscler. Thromb. Vasc. Biol.* **2009**, *29* (3), 296–297.

(48) Sen, S.; Oppenheimer, S. M.; Lima, J.; Cohen, B. *Stroke.* **2002**, *33* (4), 930–935.

(49) Mozaffarian, D.; Benjamin, E. J.; Go, A. S.; Arnett, D. K. *Circulation* **2015**, *131* (4), e29–e39.

(50) Finegold, J. A.; Asaria, P.; Francis, D. P. *Int. J. Cardiol.* **2013**, *168* (2), 934–945.

(51) Gryglewski, R. J.; Chlopicki, S.; Uracz, W.; Marcinkiewicz, E. *Med Sci Monit* **2001**, *7* (1), 1–16.

(52) Monroe, D. M.; Hoffman, M. *Arteriosclerosis, Thrombosis, and Vascular Biology*. **2006**, 41–48.

(53) Heemskerk, J. W. M.; Bevers, E. M.; Lindhout, T. *Thromb. Haemost.* **2002**, *88* (2), 186–193.

(54) Palta, S.; Saroa, R.; Palta, A. *Indian Journal of Anaesthesia*. **2014**, 515–523.

(55) Renn, T.; Schmaier, A. H.; Nickel, K. F.; Blombock, M.; Maas, C. *Blood.* **2012**, 4296–4303.

(56) Cochrane, C. G.; Revak, S. D.; Wuepper, K. D. *J. Exp. Med.* **1973**, *138* (6), 1564–1583.

(57) Gailani, D.; Renné, T. *Arterioscl. Thromb. Vascul. Biol*. **2007**, 2507–2513.

(58) Mann, K. G. In *Thromb. Haemost.* **1999**, *382*, 165–174.

(59) Furie, B.; Furie, B. C. *N. Engl. J. Med.* **1992**, *326* (12), 800–806.

(60) Mackman, N.; Tilley, R. E.; Key, N. S. *Arterioscl. Thromb. Vascul. Biol*. **2007**, 1687–1693.

(61) Di Cera, E. *J. Thromb. Haemost.* **2007**, 196–202.

(62) Ariëns, R. A. S.; Lai, T.-S.; Weisel, J. W.; Greenberg, C. S.;. *Blood.* **2002**, *100* (3), 743–754.

(63) De Candia, E.; Hall, S. W.; Rutella, S.; Landolfi, R.; Andrews, R. K.; De Cristofaro, R. *J. Biol. Chem.* **2001**, *276* (7), 4692–4698.

(64) Coughlin, S. R. *Nature* **2000**, *407* (6801), 258–264.

(65) Choi, S. H.; Smith, S. A.; Morrissey, J. H. *Blood* **2011**, *118* (26), 6963–6970.

(66) Yun, T. H.; Baglia, F. A.; Myles, T.; Navaneetham, D.; Lopez, J. A.; Walsh, P. N.; Leung, L. L. K. *J. Biol. Chem.* **2003**, *278* (48), 48112–48119.

(67) Lipets, E. N.; Ataullakhanov, F. I.; Sinauridze, E.; Panteleev, M.; *et al. Thromb. J.* **2015**, *13* (1), 4.

(68) Major, D. A.; Sane, D. C.; Herrington, D. M. *Am Hear. J* **2000**, *140* (2), 189–195.

(69) Rosendaal, F. R.; Reitsma, P. H. *Journal of Thrombosis and Haemostasis*. **2009**, 301–304.

(70) Poort, S. R.; Rosendaal, F. R.; Reitsma, P. H.; Bertina, R. M. *Blood.* **1996**, *88* (10), 3698–3703.

(71) Bell, R. G.; Sadowski, J. A.; Matschiner, J. T. *Biochemistry* **1972**, *11* (10), 1959–1961.

(72) Stenfio, J.; Suttie, J. W. *Ann. Res. Biochem.* **1977**, *46*, 157–172.

(73) Helgeland, L. *Biochem. Educ.* **1980**, *8* (3), 66–69.

(74) Wysowski, D. K.; Nourjah, P.; Swartz, L. *Arch Intern Med* **2007**, *167* (13), 1414–1419.

(75) Warkentin, T. E.; Crowther, M. A. *Can. J. Anesth.* **2002**, *49* (6), S11–S25.

(76) Lee, C. J.; Ansell, J. E. *British Journal of Clinical Pharmacology*. **2011**, 581–592.

(77) Bush, L. R. *Cardiovasc. Drug Rev.* **1991**, *9* (3), 247–263.

(78) Lunven, C.; Gauffeny, C.; Lecoffre, C.; O'Brien, D. P.; Roome, N. O.; Berry, C. N. *Thromb. Haemost.* **1996**, *75* (1), 154–160.

(79) Wienen, W.; Stassen, J. M.; Priepke, H.; Ries, U. J.; Hauel, N. *Thromb. Haemost.* **2007**, *98* (2), 333–338.

(80) Huttunen, K. M.; Raunio, H.; Rautio, J. *Pharmacol. Rev.* **2011**, *63* (3), 750–771.

(81) Grütter, M. G.; Priestle, J. P.; Rahuel, J.; Grossenbacher, H.; Bode, W.; Hofsteenge, J.; Stone, S. R. *EMBO J.* **1990**, *9* (8), 2361–2365.

(82) Ringwala, S. M.; DiBattiste, P. M.; Schneider, D. J. *J. Thromb. Thrombolysis* **2012**, *34* (3), 291–296.

(83) Agrawal, R.; Jain, P.; Dikshit, S. N. *Curr Drug Targets* **2012**, *13* (6), 863–875.

(84) Harder, S. *Thromb. J.* **2014**, *12* (1), 22.

(85) Crowther, M. A.; Warkentin, T. E. *Blood* **2008**, *111* (10), 4871–4879.

(86) Schiele, F.; Van Ryn, J.; Canada, K.; Newsome, C.; Sepulveda, E.; Park, J.; Nar, H.; Litzenburger, T. *Blood* **2013**, *121* (18), 3554–3562.

(87) Imamura, A.; Ando, H.; Korogi, S.; Tanabe, G.; Muraoka, O.; Ishida, H.; Kiso, M. *Essentials Glycobiol.* **2009**, 784.

(88) Sakko, A. J.; Butler, M. S.; Byers, S.; Reinboth, B. J.; Stahl, J.; Kench, J. G.; Horvath, L. G.; Sutherland, R. L.; Stricker, P. D.; Henshall, S. M.; Marshall, V. R.; Tilley, W. D.; Horsfall, D. J.; Ricciardelli, C. *Cancer Epidemiol. Biomarkers Prev.* **2008**, *17* (9), 2488–2497.

(89) Funderburgh, J. L. *IUBMB Life* **2002**, *54* (4), 187–194.

(90) Shriver, Z.; Capila, I.; Venkataraman, G.; Sasisekharan, R. *Handb. Exp. Pharmacol.* **2012**,

207 (207), 159–176.

(91) Liu, J.; Pedersen, L. C. *Appl. Microbiol. Biotechnol*. **2007**, 263–272.

(92) Huntington, J. A. *J. Thromb.Haemost.* **2013**, 254–264.

(93) Griffith, M. J. *J. Biol. Chem.* **1979**, *254* (23), 12044–12049.

(94) Baglin, T. P.; Carrell, R. W.; Church, F. C.; Esmon, C. T.; Huntington, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (17), 11079–11084.

(95) Olson, S. T.; Björk, I.; Sheffer, R.; Craig, P. A.; Shore, J. D.; Choay, J. *J. Biol. Chem.* **1992**, *267* (18), 12528–12538.

(96) Desai, U. R.; Petitou, M.; Björk, I.; Olson, S. T. *J. Biol. Chem.* **1998**, *273* (13), 7478–7487.

(97) Liu, H.; Zhang, Z.; Linhardt, R. J. *Nat. Prod. Rep.* **2009**, *26* (3), 313–321.

(98) Hirsh, J. *Circulation* **1998**, *13* (15), 1575–1582.

(99) Xu, Y.; Masuko, S.; Takieddin, M.; Xu, H.; Liu, R.; Jing, J.; Mousa, S. A.; Linhardt, R. J.; Liu, J. *Science (80-.).* **2011**, *334* (6055), 498–501.

(100) Bauer, K. A.; Hawkins, D. W.; Peters, P. C.; Petitou, M.; Herbert, J.-M.; van Boeckel, C. a a; Meuleman, D. G. *Cardiovasc. Drug Rev.* **2002**, *20* (1), 37–52.

(101) Franchini, M.; Chong, B.; Jang, I.-K.; Hursting, H.; Warkentin, T. *et al*. *Thromb. J.* **2005**, *3* (1), 14.

(102) Levi, M.; Eerenberg, E.; Kamphuisen, P. W. *J. Thromb. Haemost.* **2011**, 1705–1712.

(103) Guo, J.; Zhou, H.-X. *Chem. Rev.* **2016**, *116* (11), 6503-6515.

(104) Monod, J.; Wyman, J.; Changeux, J. P. *J. Mol. Biol.* **1965**, *12* (1), 88–118.

(105) Mihailescu, M. R.; Russu, I. M. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98* (7), 3773–3777.

(106) Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J. *Nature* **2014**, *508* (7496), 331–339.

(107) Cooper, A.; Dryden, D. T. F. *Eur. Biophys. J.* **1984**, *11* (2), 103–109.

(108) Nussinov, R.; Tsai, C. J. *Curr. Opin. Struct. Biol.* **2015**, *30*, 17–24.

(109) Huntington, J. A. *J. Thromb. Haemost.* **2005**, *3*, 1861–1872.

(110) Lechtenberg, B. C.; Freund, S. M. V; Huntington, J. A. *Biol. Chem.* **2012**, *393*, 889–898.

(111) Pechik, I.; Madrazo, J.; Mosesson, M. W.; Hernandez, I.; Gilliland, G. L.; Medved, L. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (9), 2718–2723.

(112) Huntington, J. A. *Biochimica et Biophysica Acta*. **2012**, 246–252.

(113) Carter, W. J.; Cama, E.; Huntington, J. A. *J. Biol. Chem.* **2005**, *280* (4), 2745–2749.

(114) Henry, B. L.; Desai, U. R. *Thromb. Res.* **2014**, *134* (5), 1123–1129.

(115) Pineda, A. O.; Chen, Z. W.; Marino, F.; Mathews, F. S.; Mosesson, M. W.; Di Cera, E. *Biophys. Chem.* **2007**, *125* (2-3), 556–559.

(116) Dumas, J. J.; Kumar, R.; Seehra, J.; Somers, W. S.; Mosyak, L. *Science* **2003**, *301* (5630), 222–226.

(117) Di Cera, E.; Guinto, E. R.; Vindigni, A.; Dang, Q. D.; Ayala, Y. M.; Wuyi, M.; Tulinsky, A. *J. Biol. Chem.* **1995**, *270* (38), 22089–22092.

(118) Henry, B. L.; Aziz, M. A.; Zhou, Q.; Desai, U. R. *Thromb. Haemost.* **2010**, *103* (3), 507–515.

(119) Henry, B. L.; Connell, J.; Liang, A.; Krishnasamy, C.; Desai, U. R. *J. Biol. Chem.* **2009**, *284* (31), 20897–20908.

(120) Henry, B. L.; Thakkar, J. N.; Liang, A.; Desai, U. R. *Biochem. Biophys. Res. Commun.* **2012**, *417* (1), 382–386.

(121) Abdel Aziz, M. H.; Mosier, P. D.; Desai, U. R. *Biochem. Biophys. Res. Commun.* **2011**,

413 (2), 348–352.

(122) Mehta, A. Y.; Thakkar, J. N.; Mohammed, B. M.; Martin, E. J.; Brophy, D. F.; Kishimoto, T.; Desai, U. R. *J. Med. Chem.* **2014**, *57* (7), 3030–3039.

(123) Mehta, A. Y.; Mohammed, B. M.; Martin, E. J.; Brophy, D. F.; Gailani, D.; Desai, U. R. *J. Thromb. Haemost.* **2016**, *14* (4), 828–838.

(124) Sidhu, P. S.; Liang, A.; Mehta, A. Y.; Abdel Aziz, M. H.; Zhou, Q.; Desai, U. R. *J. Med. Chem.* **2011**, *54* (15), 5522–5531.

(125) Sidhu, P. S.; Abdel Aziz, M. H.; Sarkar, A.; Mehta, A. Y.; Zhou, Q.; Desai, U. R. *J. Med. Chem.* **2013**, *56* (12), 5059–5070.

(126) Verghese, J.; Liang, A.; Sidhu, P. P. S.; Hindle, M.; Zhou, Q.; Desai, U. R. *Bioorganic Med. Chem. Lett.* **2009**, *19* (15), 4126–4129.

(127) Karuturi, R.; Al-Horani, R. A.; Mehta, S. C.; Gailani, D.; Desai, U. R. *J. Med. Chem.* **2013**, *56* (6), 2415–2428.

(128) Abdel Aziz, M. H.; Sidhu, P. S.; Liang, A.; Kim, J. Y.; Mosier, P. D.; Zhou, Q.; Farrell, D. H.; Desai, U. R. *J. Med. Chem.* **2012**, *55* (15), 6888–6897.

(129) Argade, M. D.; Mehta, A. Y.; Sarkar, A.; Desai, U. R. *J. Med. Chem.* **2014**, *57* (8), 3559–3569.

(130) Al-Horani, R. A.; Liang, A.; Desai, U. R. *J. Med. Chem.* **2011**, *54* (17), 6125–6138.

(131) Lane, D. A.; Philippou, H.; Huntington, J. A. *Blood*. **2005**, 2605–2612.

(132) Adams, T. E.; Huntington, J. A. *Arterioscl. Thromb. Vascul. Biol*. **2006**, 1738–1745.

(133) Dai, J.; Zhou, H. X. *Structure* **2015**, *23* (1), 228–236.

(134) Vafabakhsh, R.; Levitz, J.; Isacoff, E. Y. *Nature* **2015**, *524* (7566), 497–501.

(135) Christopoulos, A. *Nat Rev Drug Discov* **2002**, *1* (3), 198–210.

(136) Sharma, S.; Rodriguez, A. L.; Conn, P. J.; Lindsley, C. W. *Bioorganic Med. Chem. Lett.* **2008**, *18* (14), 4098–4101.

(137) Burgin, A. B.; Magnusson, O. T.; Singh, J.; Witte, P.; Staker, B. L.; Bjornsson, J. M.; Thorsteinsdottir, M.; Hrafnsdottir, S.; Hagen, T.; Kiselyov, A. S.; Stewart, L. J.; Gurney, M. E. *Nat. Biotechnol.* **2010**, *28* (1), 63–70.

(138) Al-Horani, R. A.; Karuturi, R.; White, D. T.; Desai, U. R. *Molecules* **2015**, *20* (1), 608–624.

(139) Patel, N. J.; Karuturi, R.; Al-Horani, R. A.; Baranwal, S.; Patel, J.; Desai, U. R.; Patel, B. B. *ACS Chem. Biol.* **2014**, *9* (8), 1826–1833.

(140) Al-Horani, R. A.; Desai, U. R. *Tetrahedron.* **2010**, 2907–2918.

(141) Raghuraman, A.; Riaz, M.; Hindle, M.; Desai, U. R. *Tetrahedron Lett.* **2007**, *48* (38), 6754–6758.

(142) Rostovtsev, V. V; Green, L. G.; Fokin, V. V; Sharpless, K. B. *Angew. Chemie - Int. Ed.* **2002**, *41* (14), 2596–2599.

(143) Hileman, R. E.; Fromm, J. R.; Weiler, J. M.; Linhardt, R. J. *BioEssays*. **1998**, 156–167.

(144) Olson, S. T.; Halvorson, H. R.; Bjork, I. *J. Biol. Chem.* **1991**, *266* (10), 6342–6352.

(145) Petrera, N. S.; Stafford, A. R.; Leslie, B. A.; Kretz, C. A.; Fredenburgh, J. C.; Weitz, J. I. *J. Biol. Chem.* **2009**, *284* (38), 25620–25629.

(146) Al-Horani, R. A.; Ponnusamy, P.; Mehta, A. Y.; Gailani, D.; Desai, U. R. *J. Med. Chem.* **2013**, *56* (3), 867–878.

(147) Gomez-Outes, A.; Luisa Suarez-Gea, M.; Calvo-Rojas, G.; Lecumberri, R.; Rocha, E.;

Pozo-Hernandez, C.; Isabel Terleira-Fernandez, A.; Vargas-Castrillon, E. *Curr. Drug Discov. Technol.* **2012**, *9* (2), 22.

(148) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem. Int. Ed.* **1998**, *37* (20), 2754–2794.

(149) Shental-Bechor, D.; Levy, Y. *Proc. Natl. Acad. Sci.* **2008**, *105* (24), 8256–8261.

(150) Sankaranarayanan, N. V.; Sarkar, A.; Desai, U. R.; Mosier, P. D. *Methods Mol. Biol.* **2015**, *1229*, 289–314.

(151) Ruiz-Larrea, M. B. *Biochem. Mol. Biol. Educ.* **2002**, *30* (4), 235–238.

(152) Lee, A. W.; Karplus, M. *Proc. Natl. Acad. Sci. U. S. A.* **1983**, *80* (December), 7055–7059.

(153) Palmer, A. G. *Curr. Op. Struct. Biol.* **1997**, 732–737.

(154) S. Boussaad; J. Pean, A.; Tao*, N. J. *Anal. Chem.* **2000**, *72* (1), 222–226.

(155) Chowdhury, S. K.; Katta, V.; Chait, B. T. *J. Am. Chem. Soc.* **1990**, No. 112, 9013–9015.

(156) Keizer, J. *J. Am. Chem. Soc.* **1983**, *105* (6), 1494–1498.

(157) Phillips, S. R.; Wilson, L. J.; Borkman, R. F. *Curr Eye Res* **1986**, *5* (0271-3683), 611–619.

(158) Adams, T. E.; Huntington, J. A. *Arterioscler. Thromb. Vasc. Biol.* **2006**, *26* (8), 1738–1745.

(159) Lane, D. a; Philippou, H.; Huntington, J. a. *Bloodjournal* **2008**, *106* (8), 2605–2612.

(160) Xu, H.; Bush, L. A.; Pineda, A. O.; Caccia, S.; Di Cera, E. *J. Biol. Chem.* **2005**, *280* (9), 7956–7961.

(161) Ye, J.; Esmon, C. T.; Johnson, A. E. *J. Biol. Chem.* **1993**, *268* (4), 2373–2379.

(162) Fuentes-Prior, P.; Iwanaga, Y.; Huber, R.; Pagila, R.; Rumennik, G.; Seto, M.; Morser, J.; Light, D. R.; Bode, W. *Nature* **2000**, *404* (March), 518–525.

(163) Yang, L.; Manithody, C.; Walston, T. D.; Cooper, S. T.; Rezaie, A. R. *J. Biol. Chem.* **2003**, *278* (39), 37465–37470.

(164) Koeppe, J. R.; Seitova, A.; Mather, T.; Komives, E. A. *Biochemistry* **2005**, *44* (45), 14784–14791.

(165) Tennent, G. A.; Brennan, S. O.; Stangou, A. J.; O'Grady, J.; Hawkins, P. N.; Pepys, M. B. *Blood* **2007**, *109* (5), 1971–1974.

(166) Ariëns, R. A. S.; Lai, T.-S.; Weisel, J. W.; Greenberg, C. S. *et al*. *Blood* **2002**, *100* (3), 743–754.

(167) Farrell, D. H. *Blood* **2014**, *124* (9), 1389–1390.

(168) Farrell, D. H. *Clinical Chemistry and Laboratory Medicine*. **2012**, 1903–1909.

(169) Meh, D. A.; Siebenlist, K. R.; Mosesson, M. W. *J. Biol. Chem.* **1996**, *271* (38), 23121–23125.

(170) Lovely, R. S.; Moaddel, M.; Farrell, D. H. *J. Thromb. Haemost.* **2003**, *1* (1), 124–131.

(171) Fuentes-Prior, P.; Iwanaga, Y.; Huber, R.; Pagila, R.; Rumennik, G.; Seto, M.; Morser, J.; Light, D. R.; Bode, W. *Nature* **2000**, *404* (March), 518–525.

(172) Kalafatis, M.; Rand, M. D.; Mann, K. G. *J. Biol. Chem.* **1994**, *269* (50), 31869–31880.

(173) Rosing, J.; Hoekema, L.; Nicolaes, G. A. F.; Thomassen, M. C. L. G. D.; Hemker, H. C.; Varadi, K.; Schwarz, H. P.; Tans, G. *J. Biol. Chem.* **1995**, *270* (46), 27852–27858.

(174) Hornyak, T. J.; Bishop, P. D.; Shafer, J. A. *Biochemistry* **1989**, *28*, 7326–7332.

(175) Hsieh, L.; Nugent, D. *Haemophilia* **2008**, *14* (6), 1190–1200.

(176) Kravtsov, D. V; Matafonov, A.; Tucker, E. I.; Sun, M. F.; Walsh, P. N.; Gruber, A.; Gailani, D. *Blood* **2009**, *114* (2), 452–458.

(177) Guoguen, P.; Galinat, H.; Blouch, M. T.; Bridey, F.; Duchemin, J.; Le Gal, G.; Abgrall, J. F.; Pan-Petesch, B. *Br. J. Haematol.* **2012**, *156* (2), 245–251.

(178) Matafonov, A.; Sarilla, S.; Sun, M. F.; Sheehan, J. P.; Serebrov, V.; Verhamme, I. M.; Gailani, D. *Blood* **2011**, *118* (2), 437–445.

(179) Geng, Y.; Verhamme, I. M.; Smith, S. B.; Sun, M. F.; Matafonov, A.; Cheng, Q.; Smith, S. A.; Morrissey, J. H.; Gailani, D. *Blood.* **2013**, *121* (19), 3962–3969.

(180) Kershaw, G.; Orellana, D. *Semin. Thromb. Hemost.* **2013**, *39* (3), 283–290.

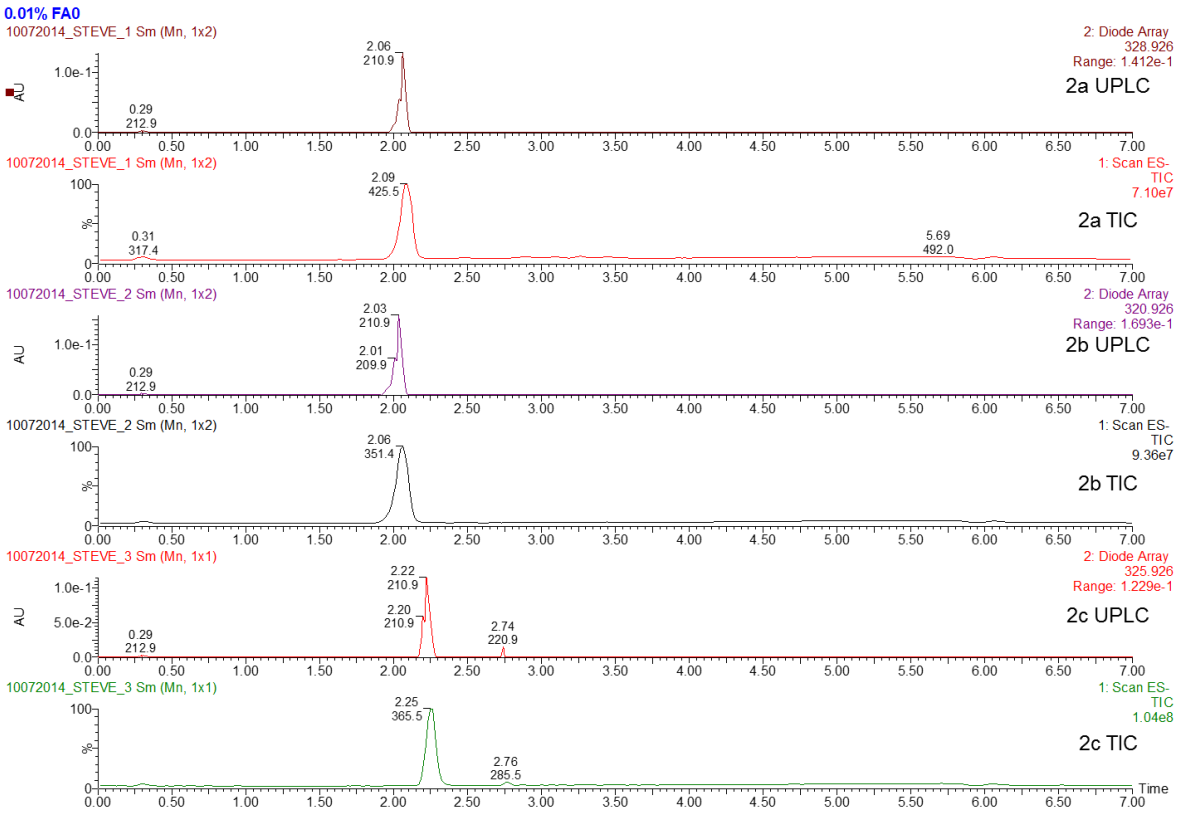
(181) Arroyo, V. *Hepatology*. Wiley Subscription Services, Inc., A Wiley Company August **2009**, 355–357.

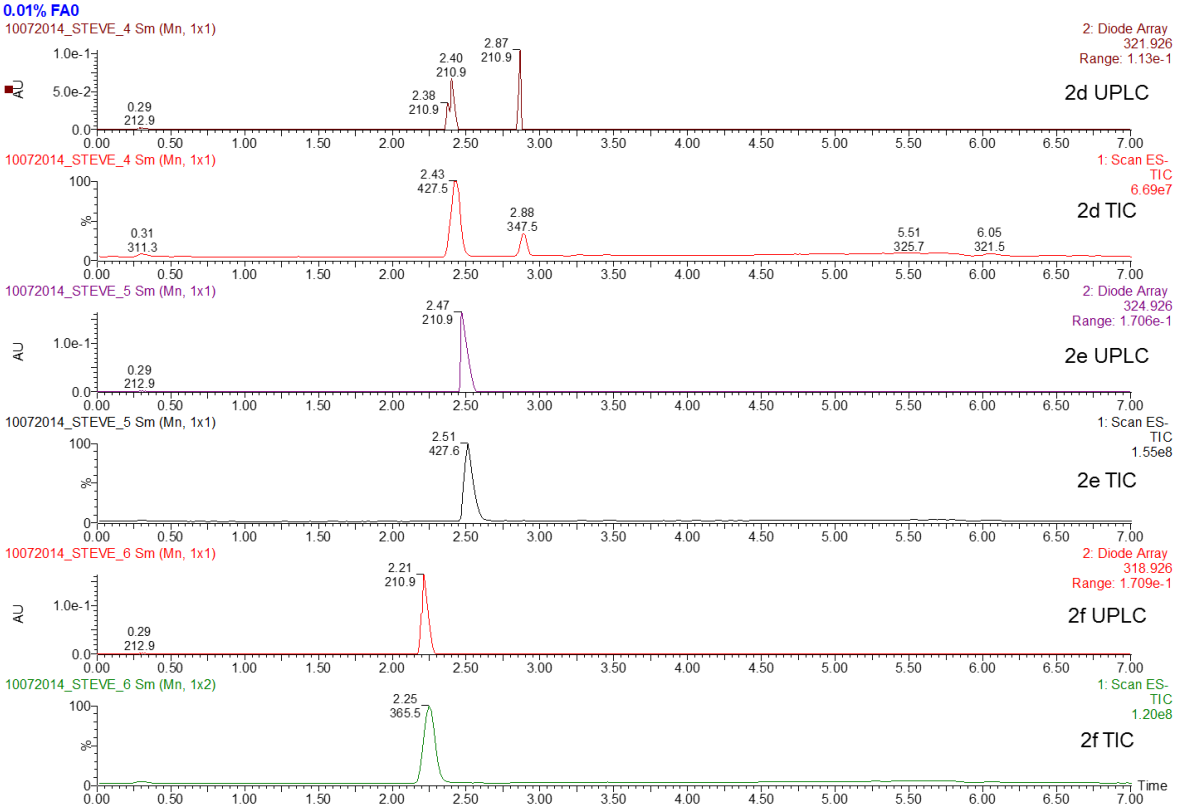
(182) Bolliger, D.; Seeberger, M. D.; Tanaka, K. A. *Transfusion Medicine Reviews*. **2012**, 1–13.

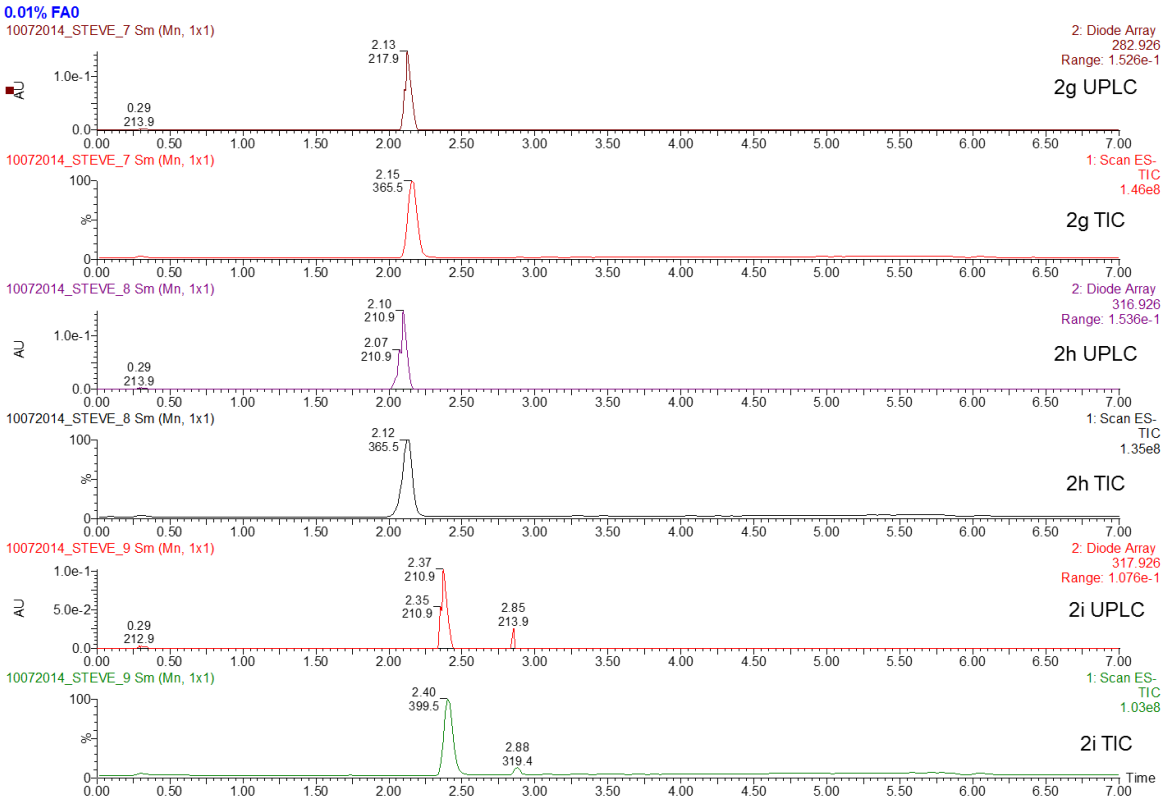
(183) Greene, T. K.; Schiviz, A.; Hoellriegl, W.; Poncz, M.; Muchitsch, E. *J. Thromb. Haemost.* **2010**, *8* (12), 2820–2822.

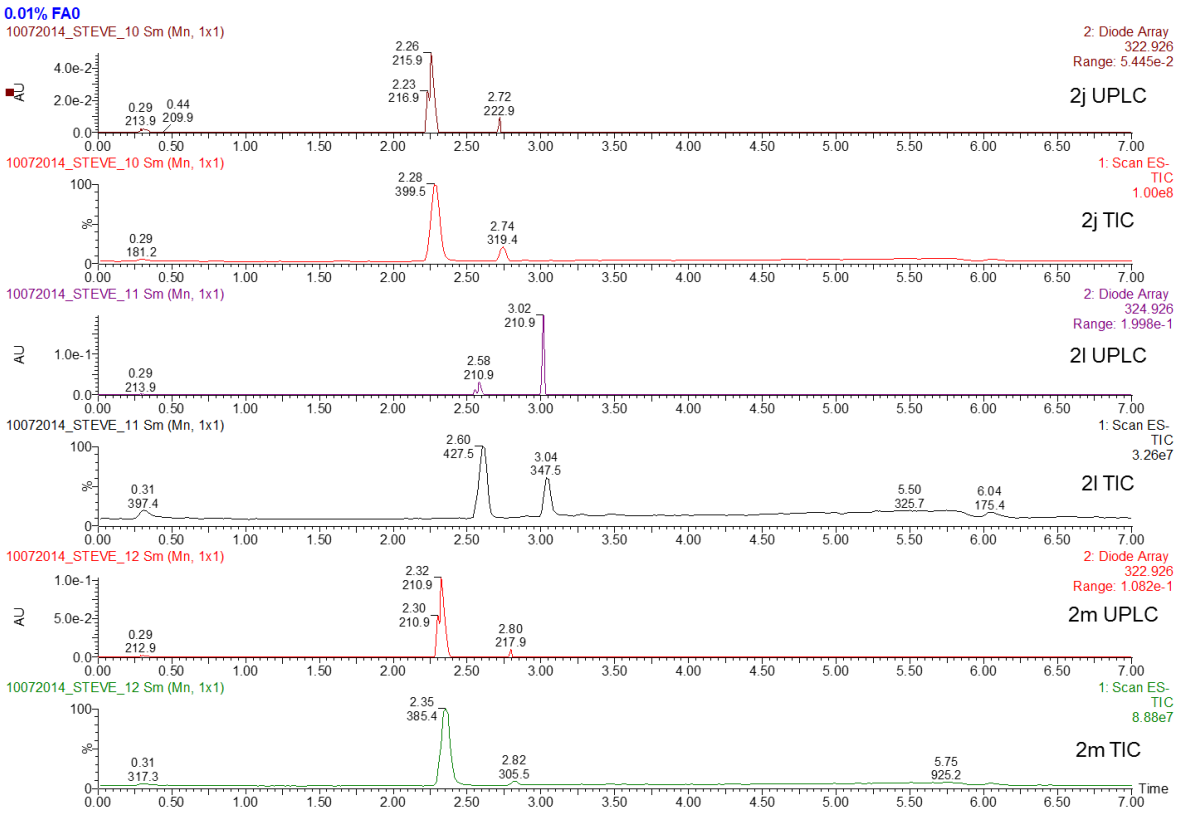
APPENDIX A

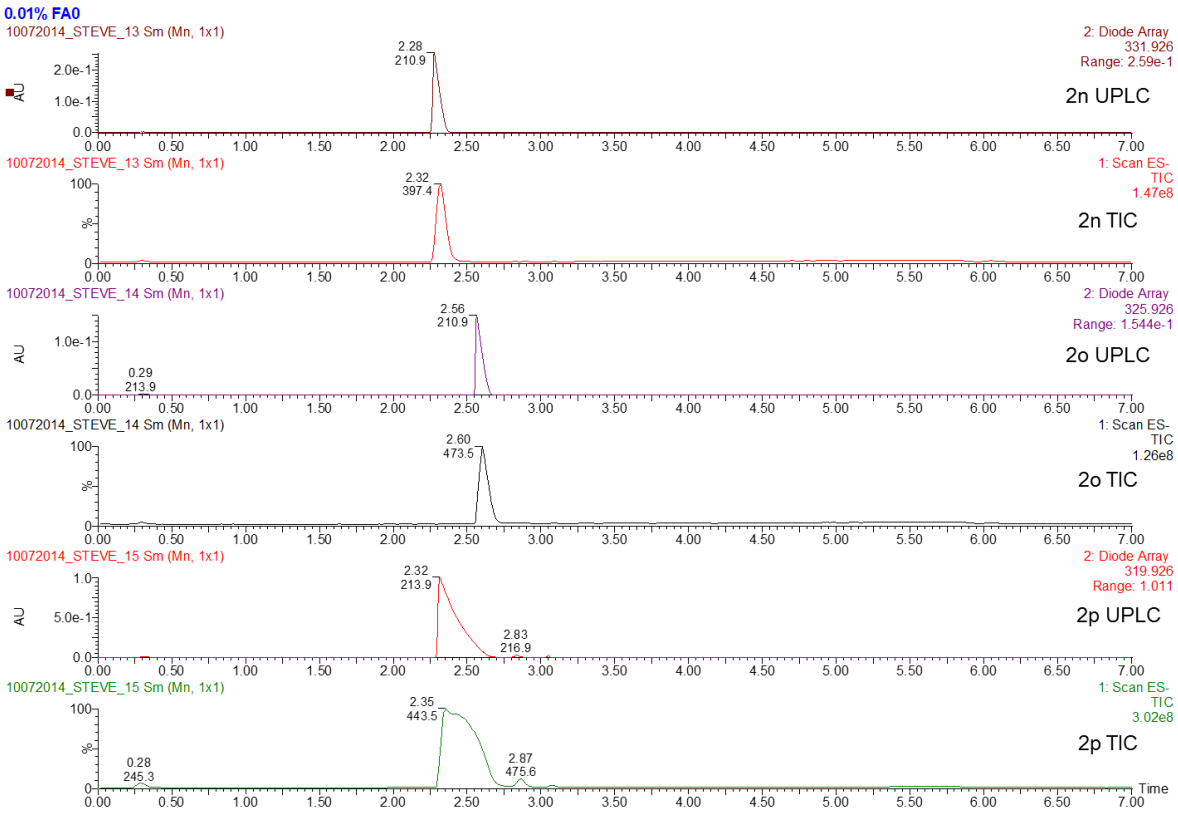
Ultra-Pressure liquid chromatography (UPLC) coupled to a low-resolution mass spectrometer (MS) produced the following spectra. The spectra are labeled on the right for the name of the compound with UPLC or TIC, representing the chromatograph and total-ion count, respectively. The solvent system contained 0.01% formic acid with a gradient of 10:90 acetonitrile:water, to 90:10 acetonitrile:water.

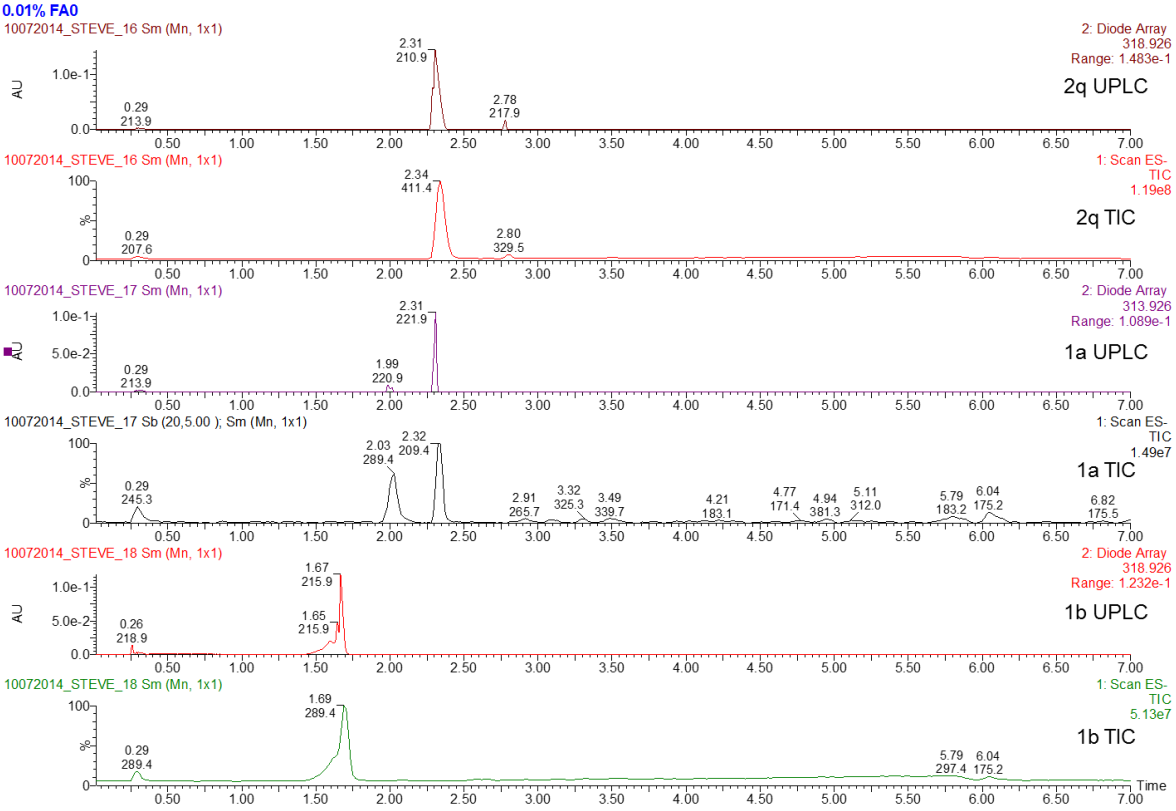


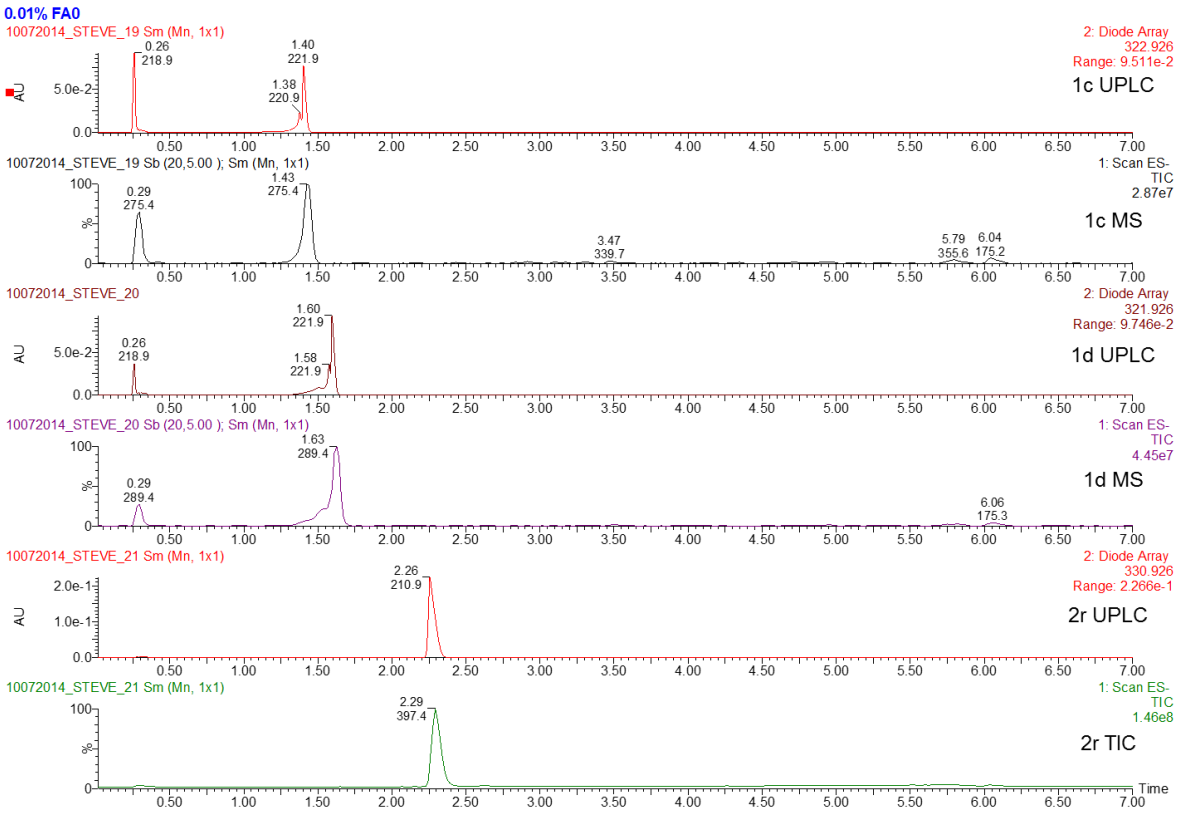


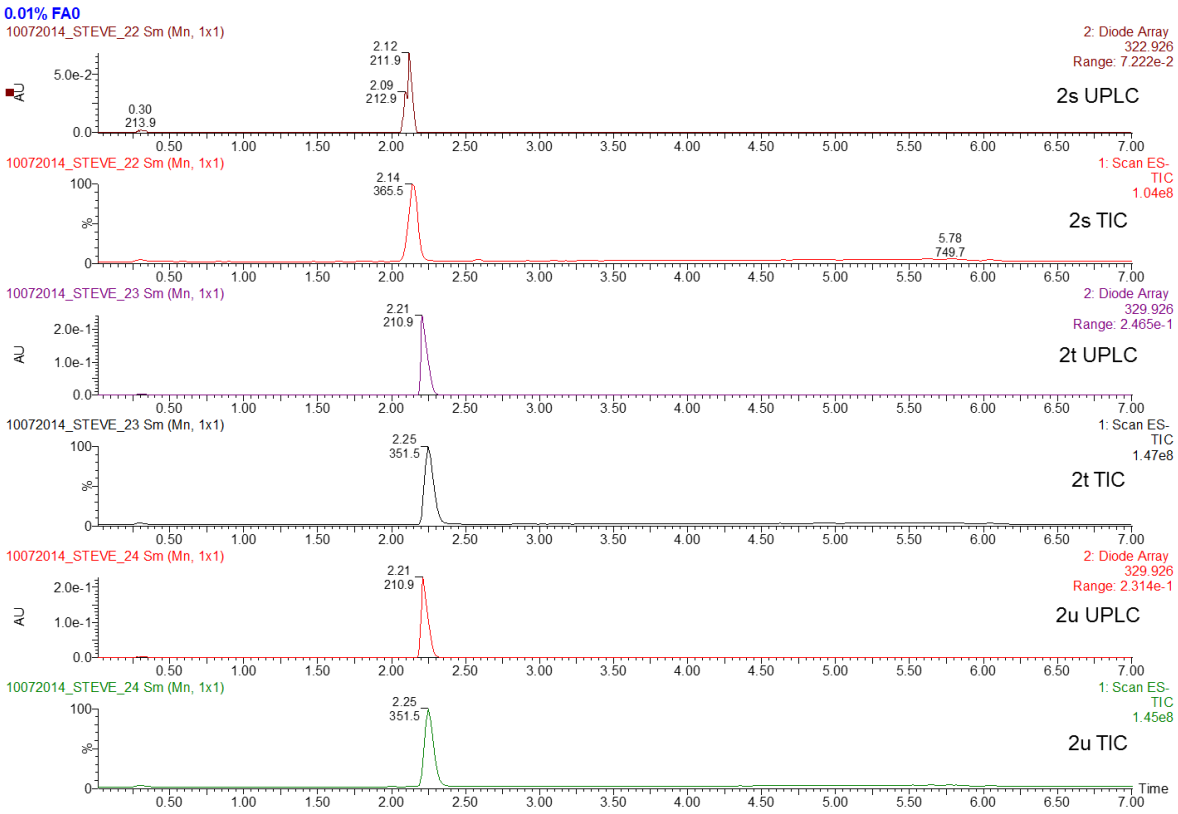


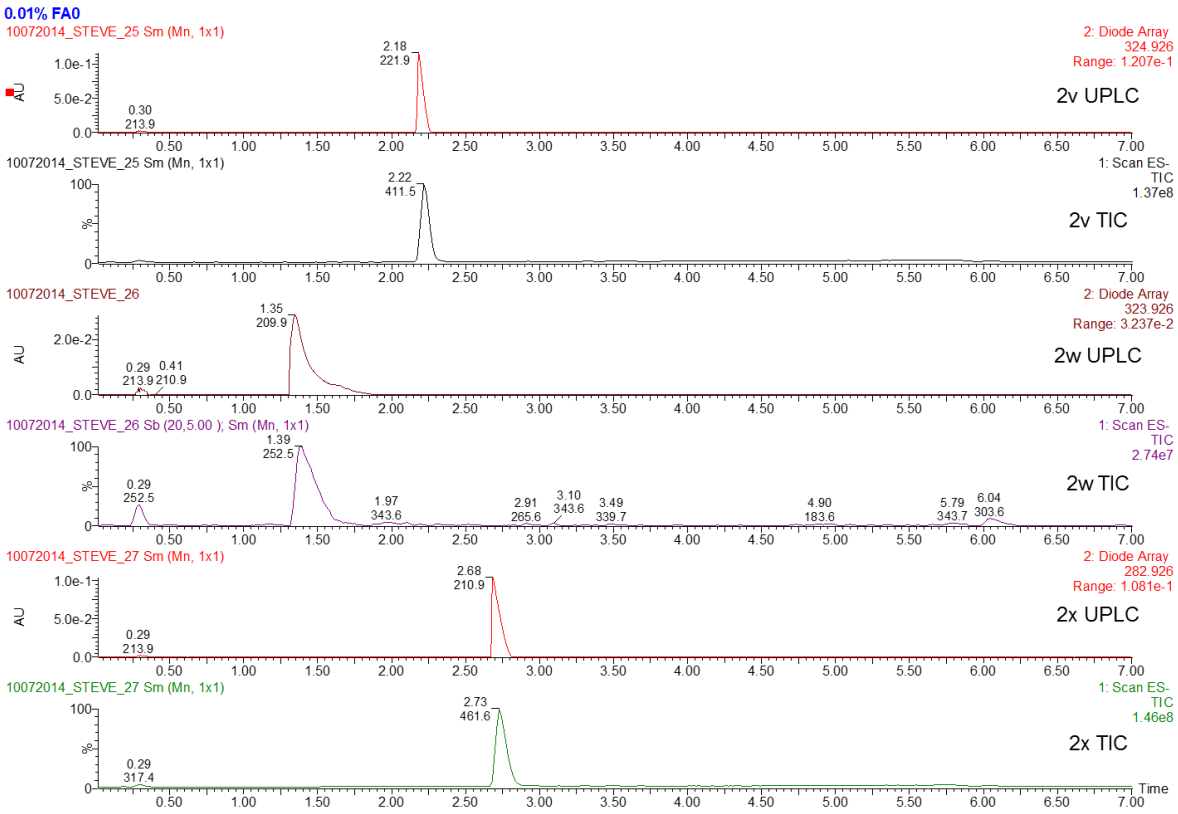


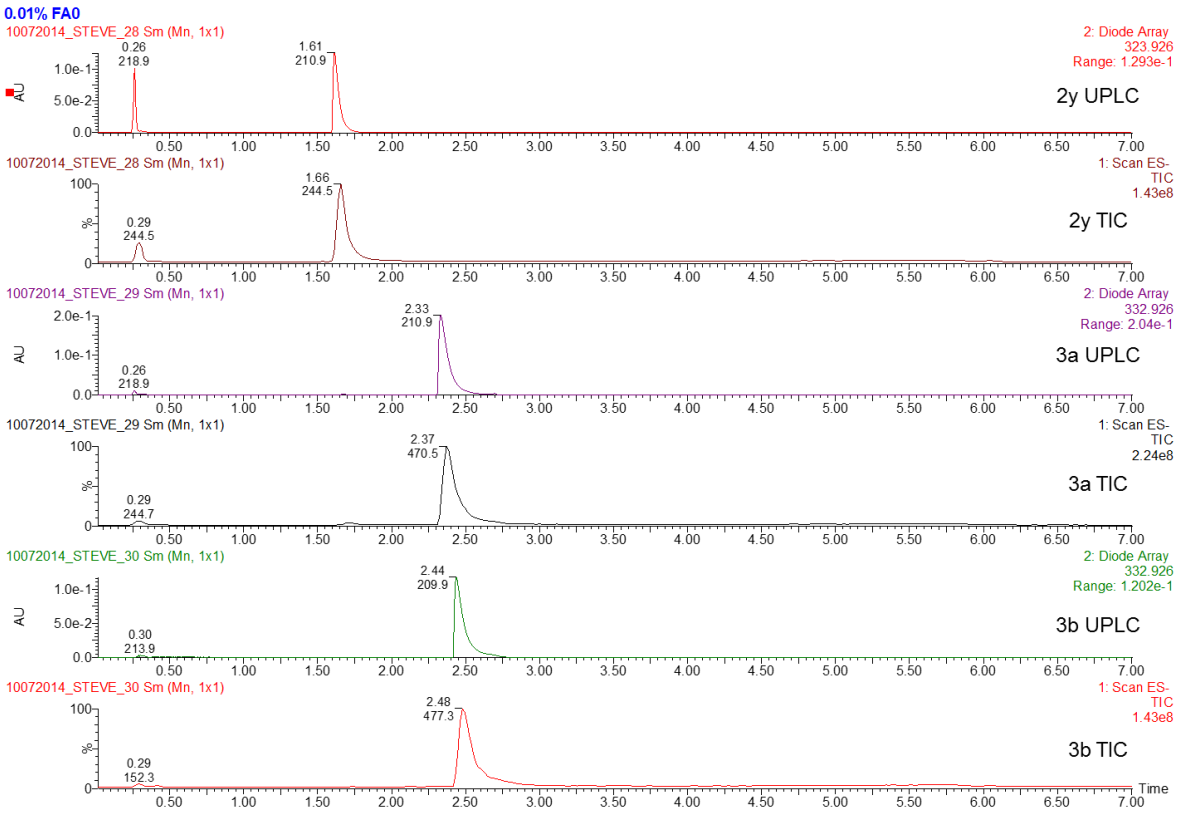


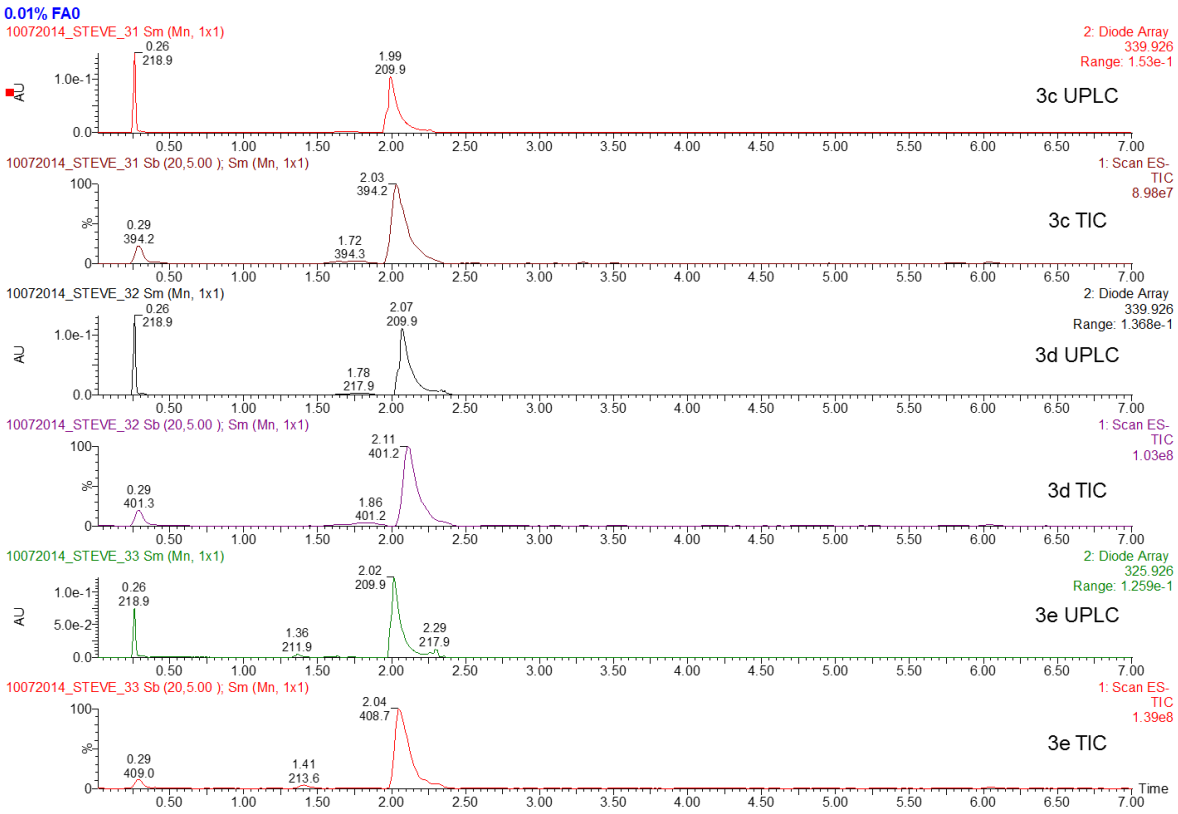


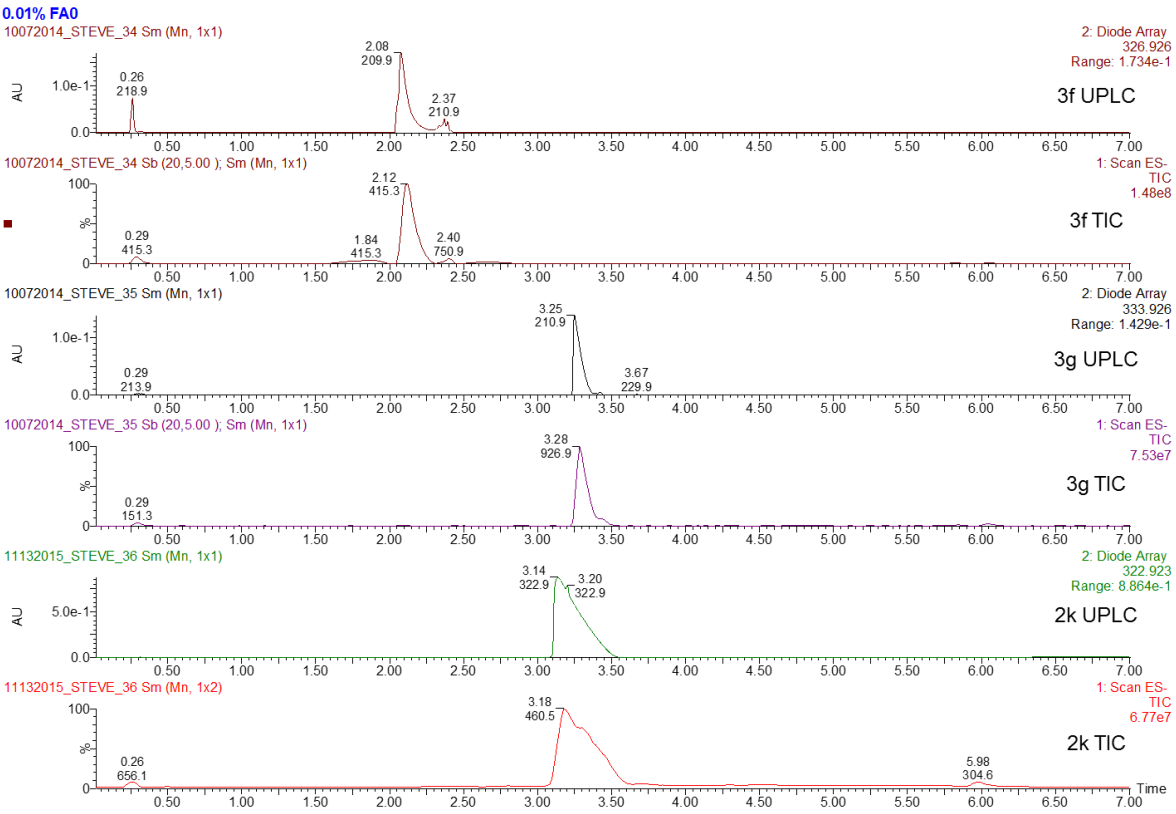






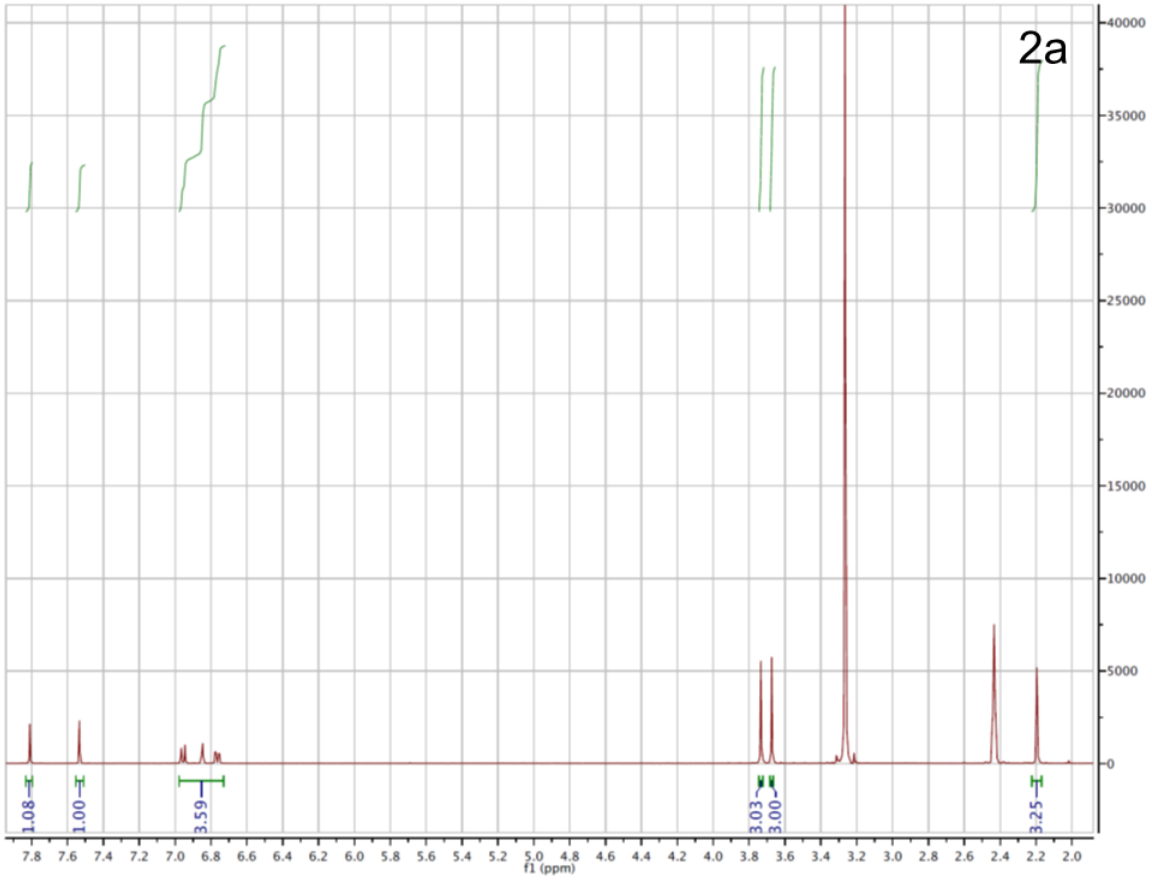


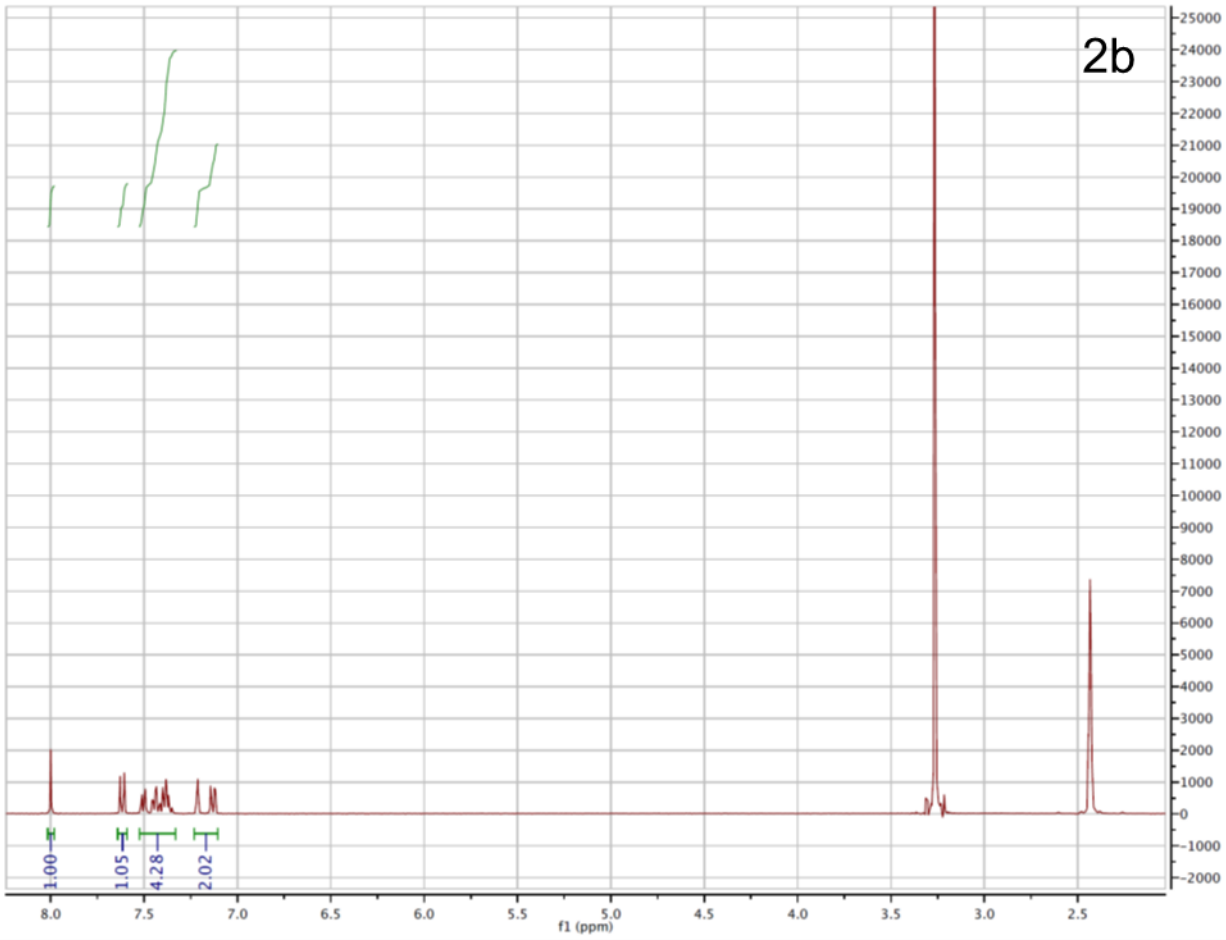


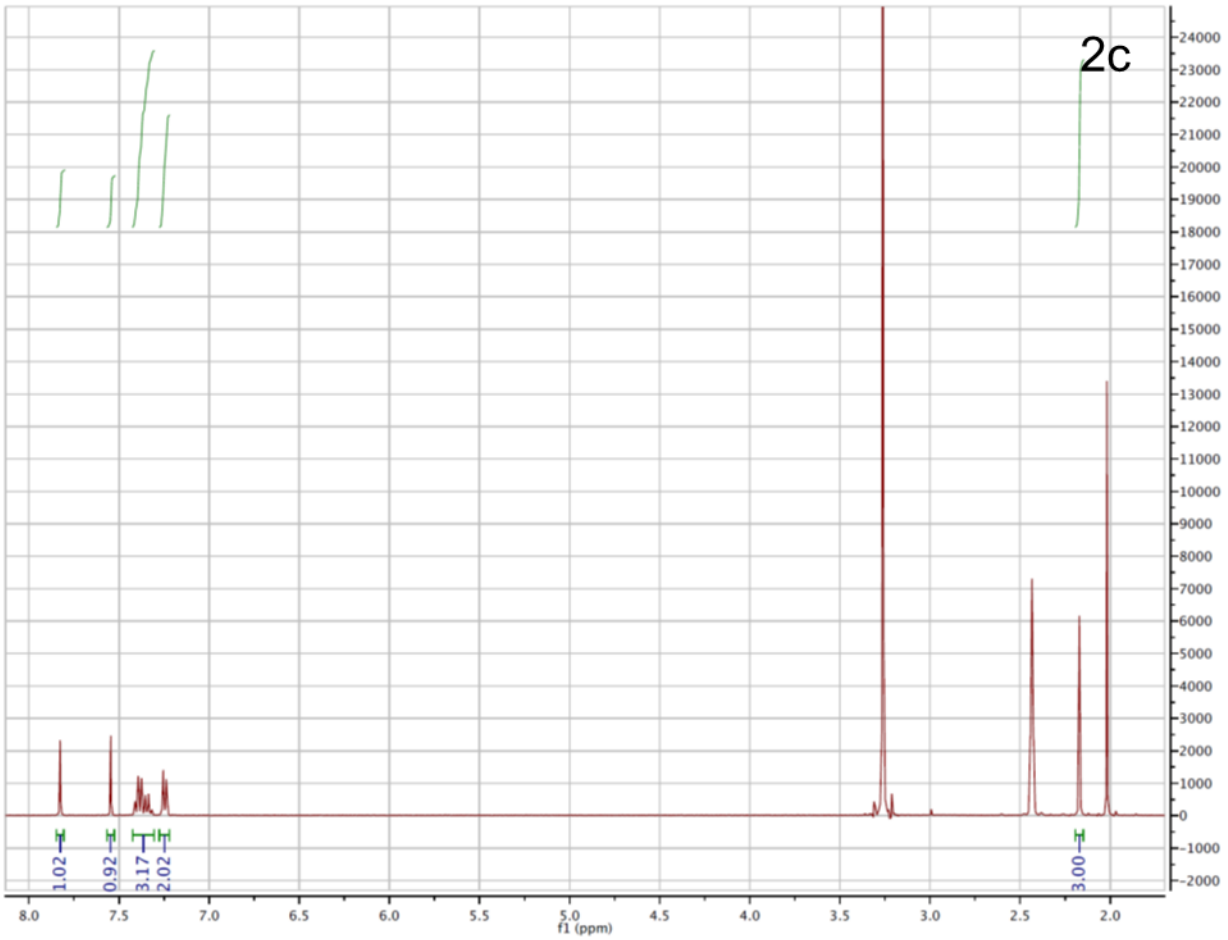


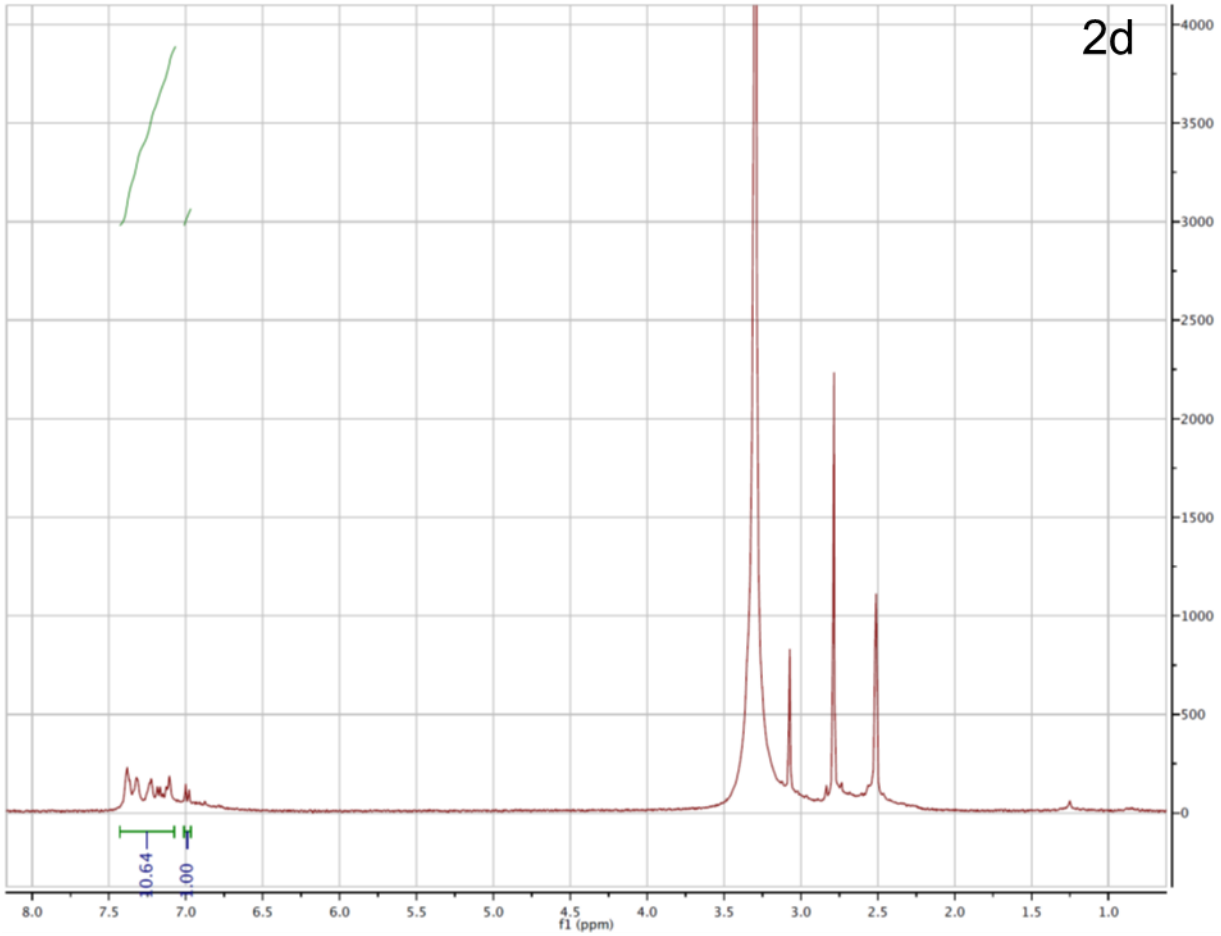
APPENDIX B

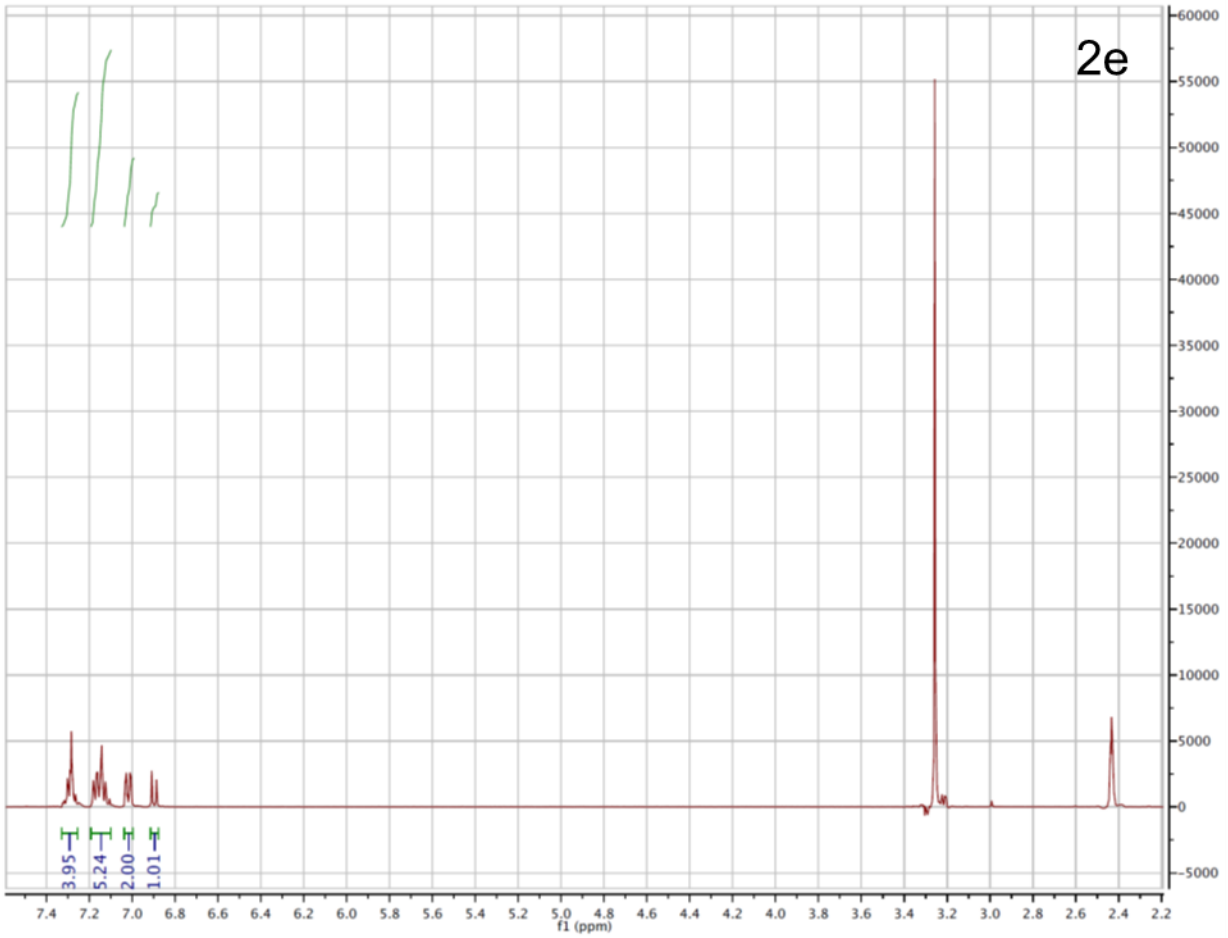
The following ^{1}H-NMR and ^{13}C-NMR graphs are taken from a Bruker 400 MHz NMR, processed to show the integrations of each peak relevant the compound named on the upper-right. Full characterization is found in Chapter 2, Section 2. Unlabeled peaks are solvent or impurities within the solvent (> 99 % d_6-DMSO).

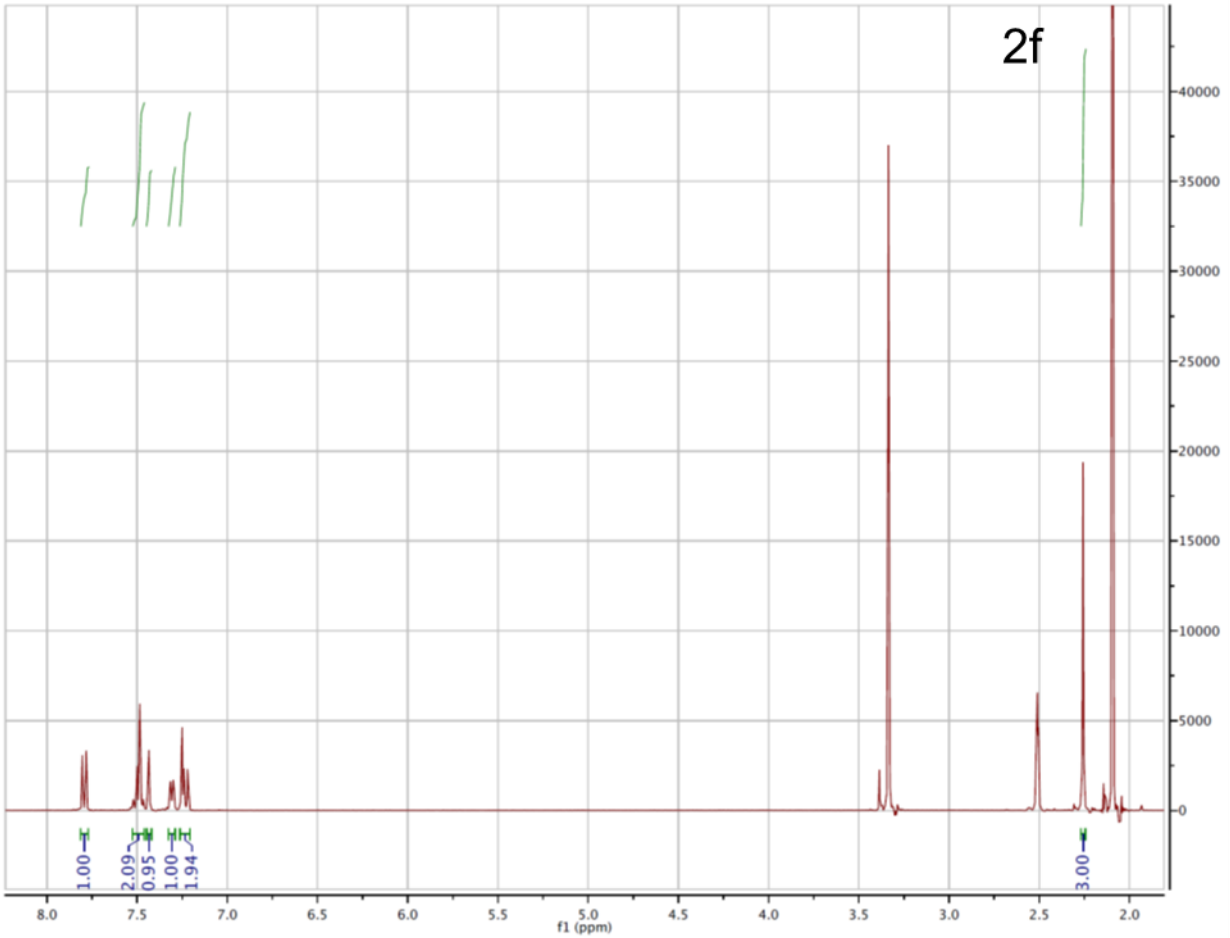


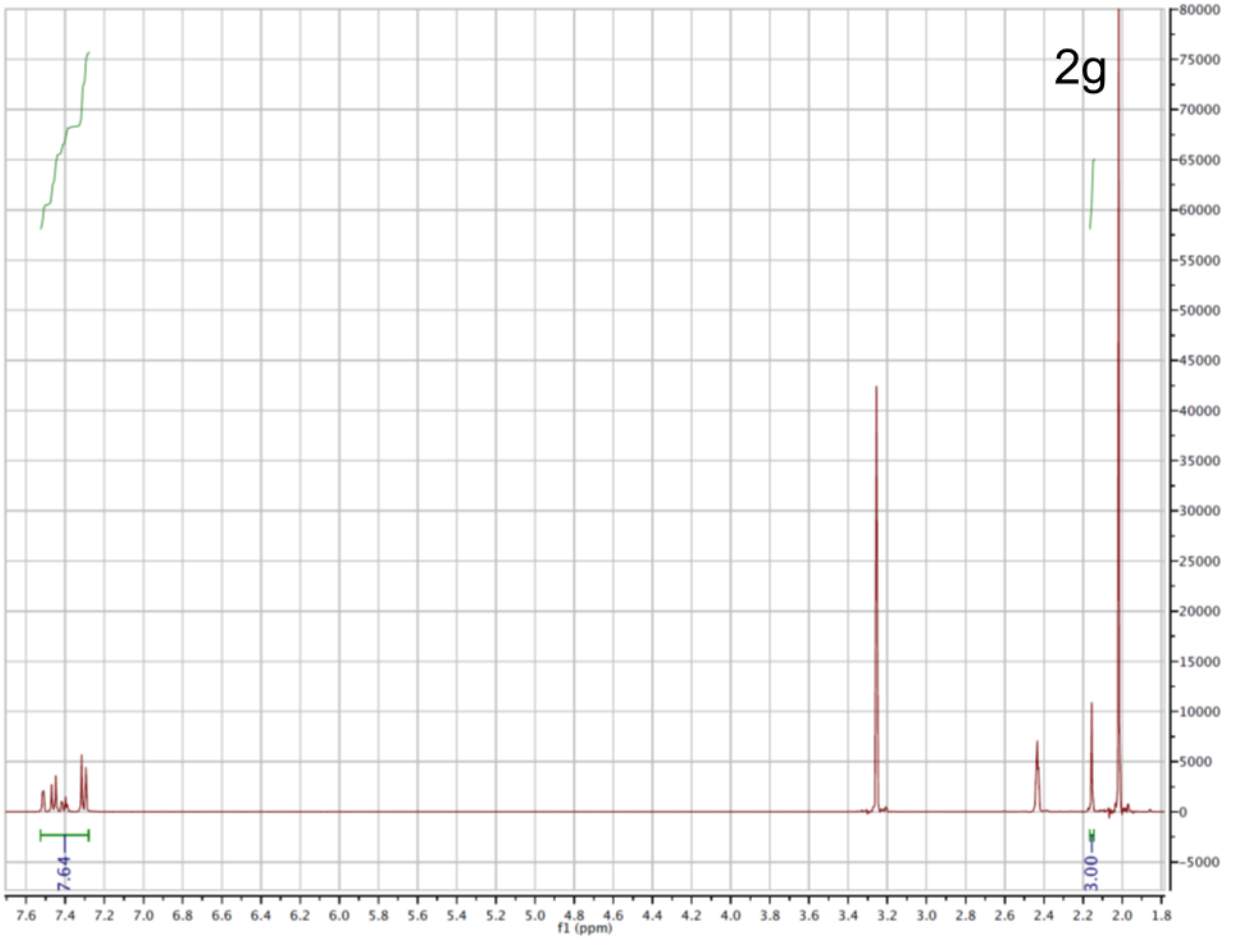


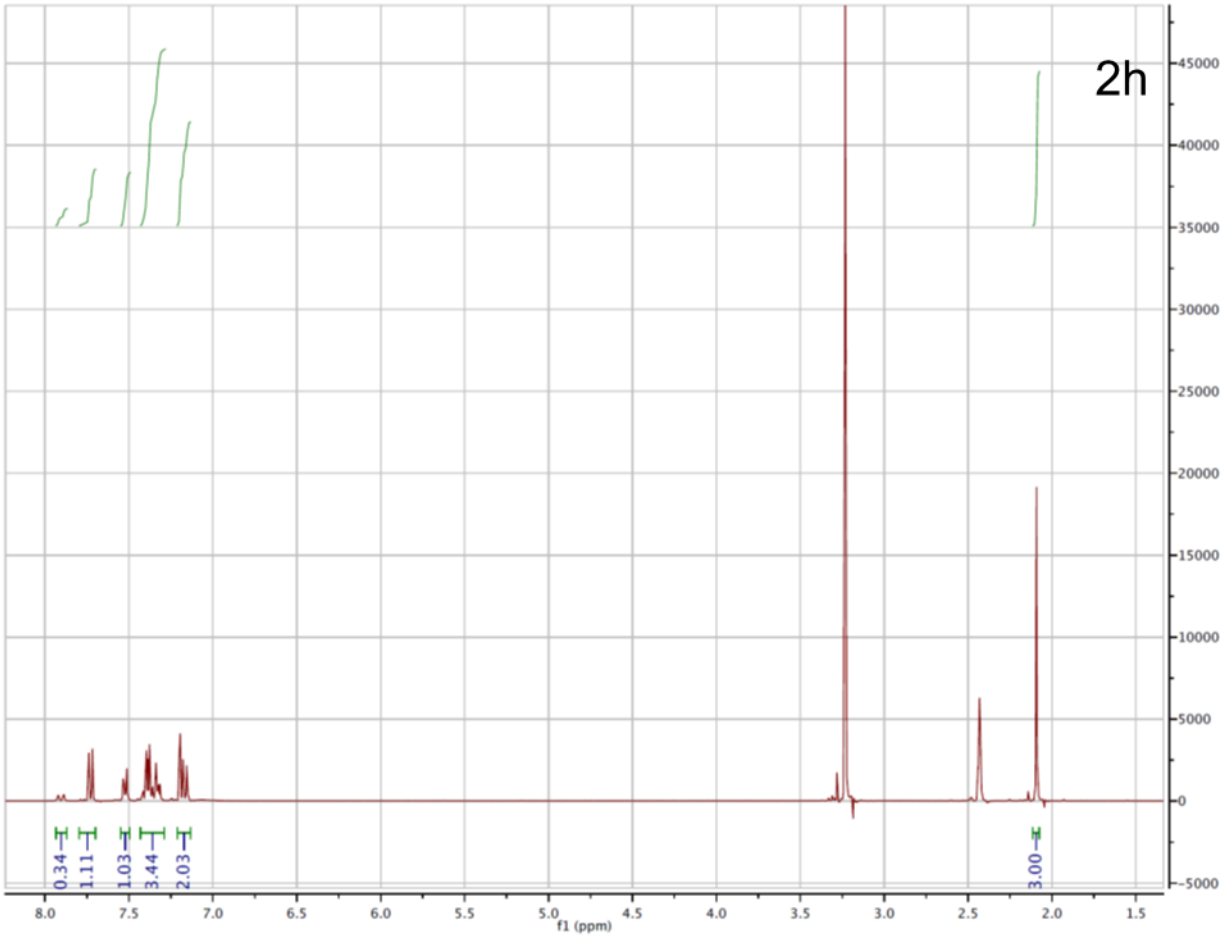


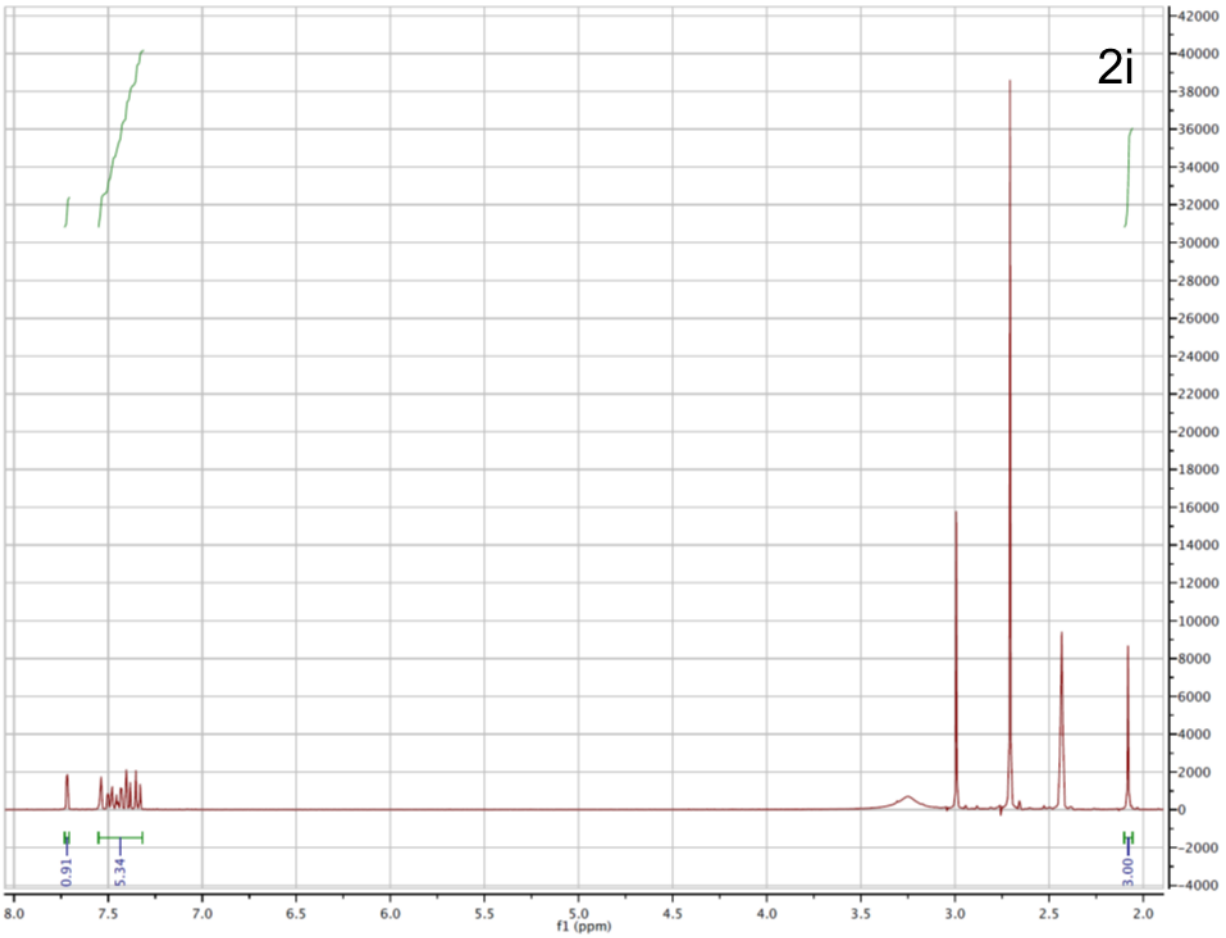


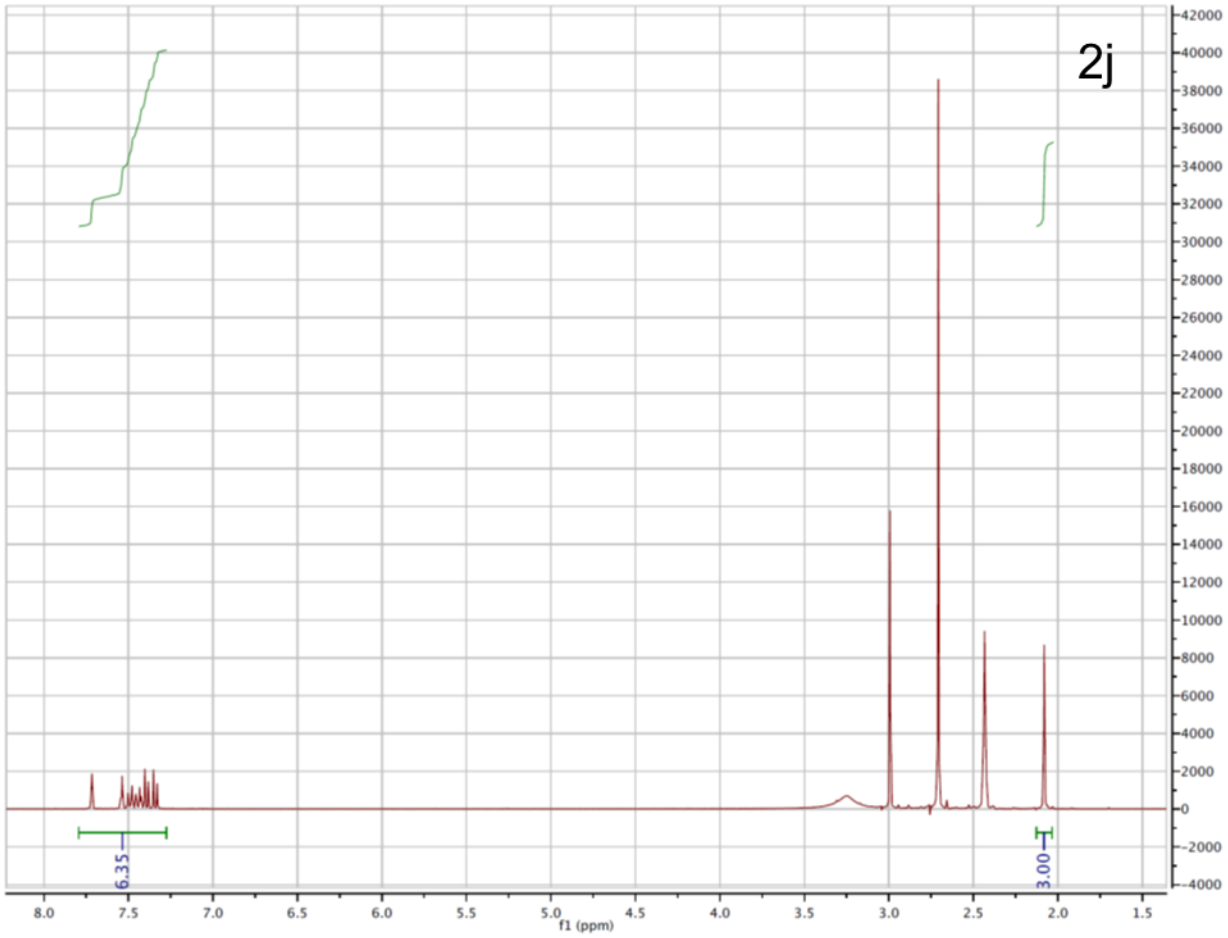


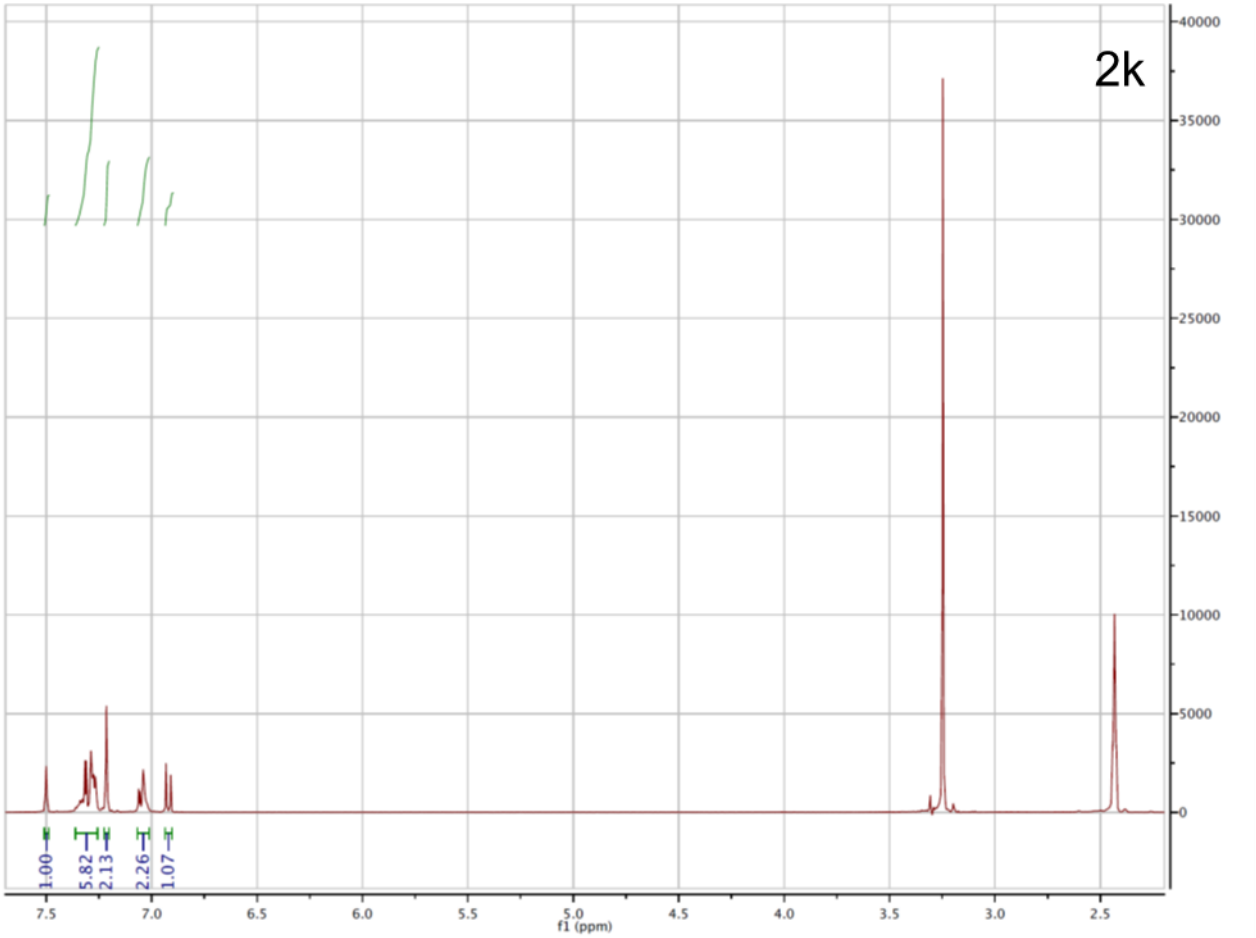


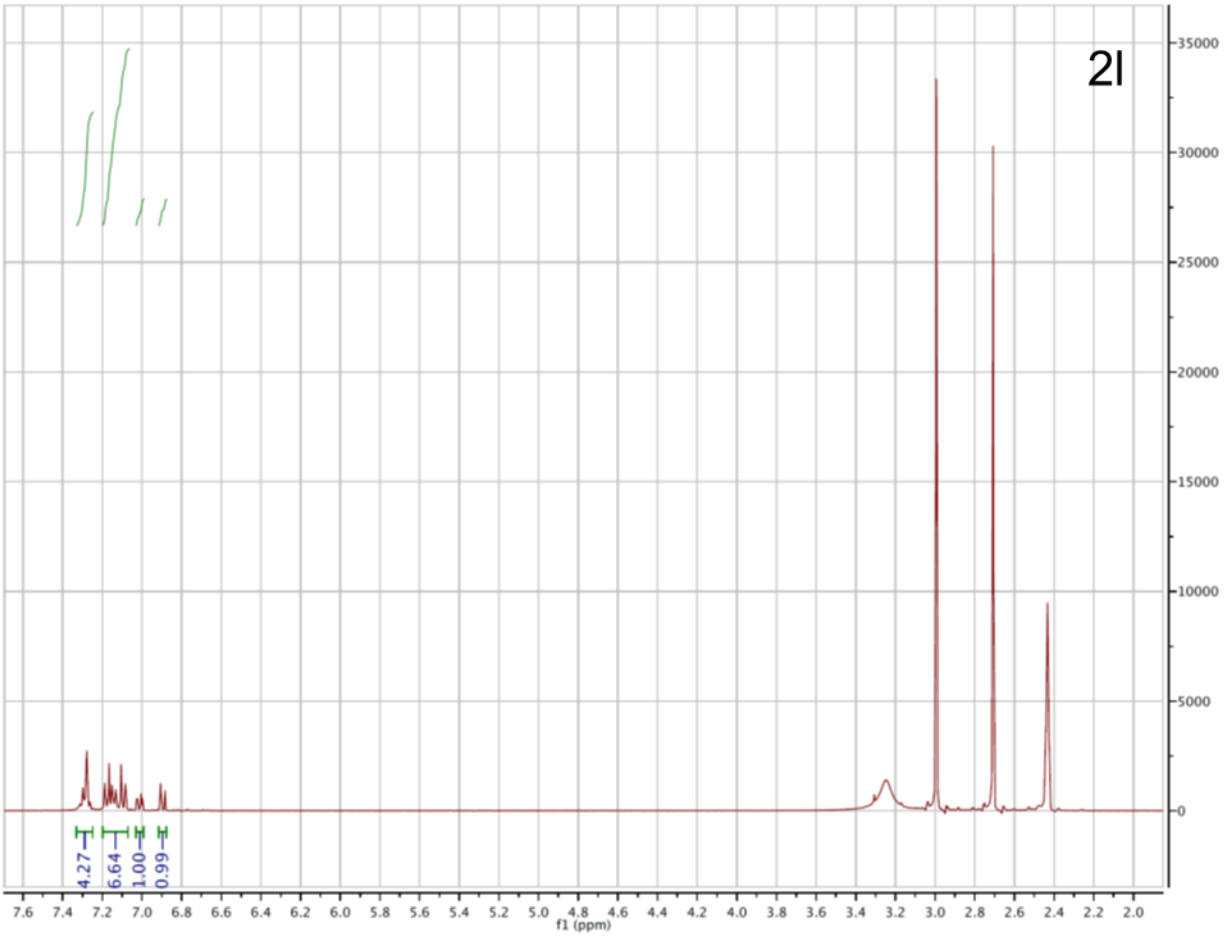


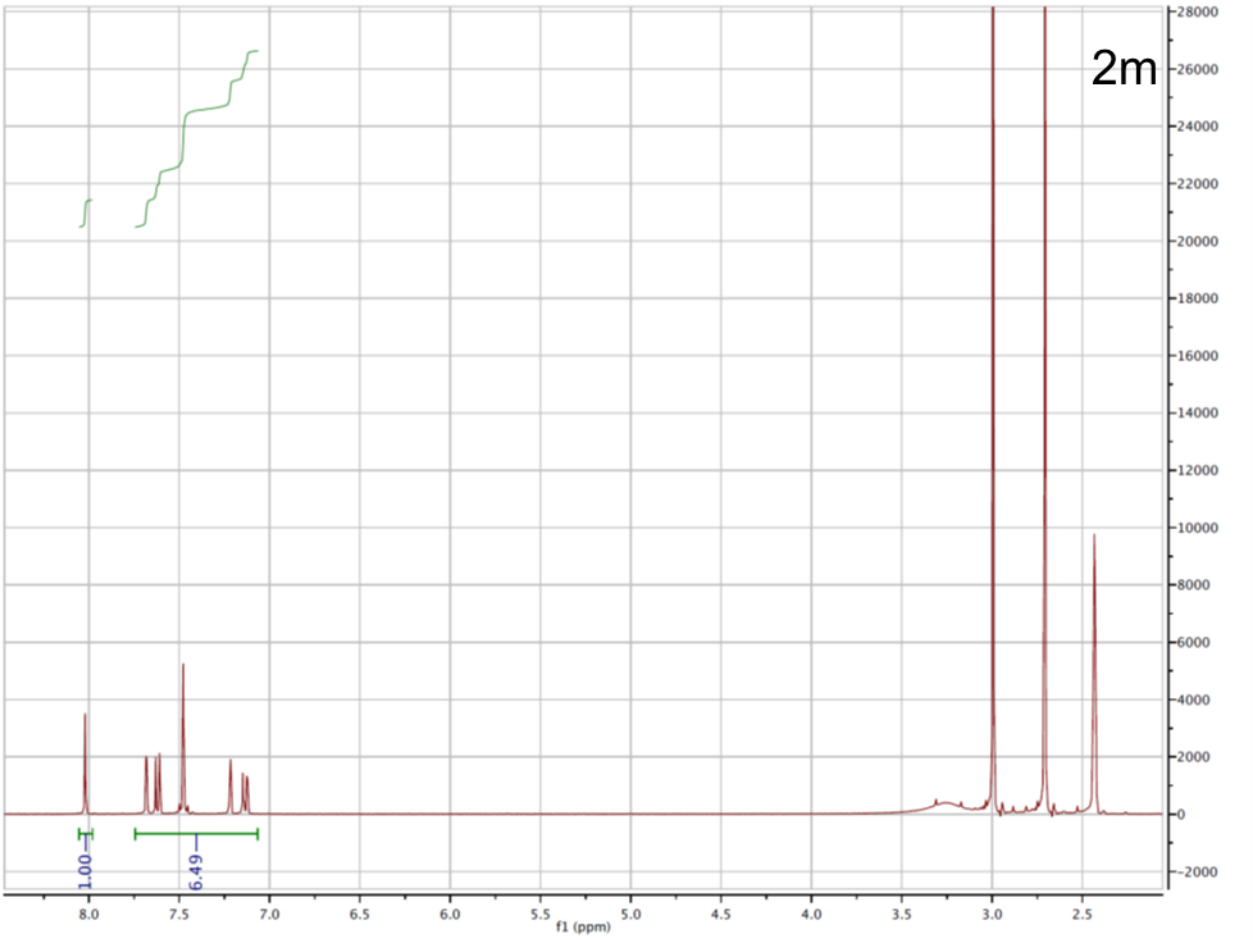


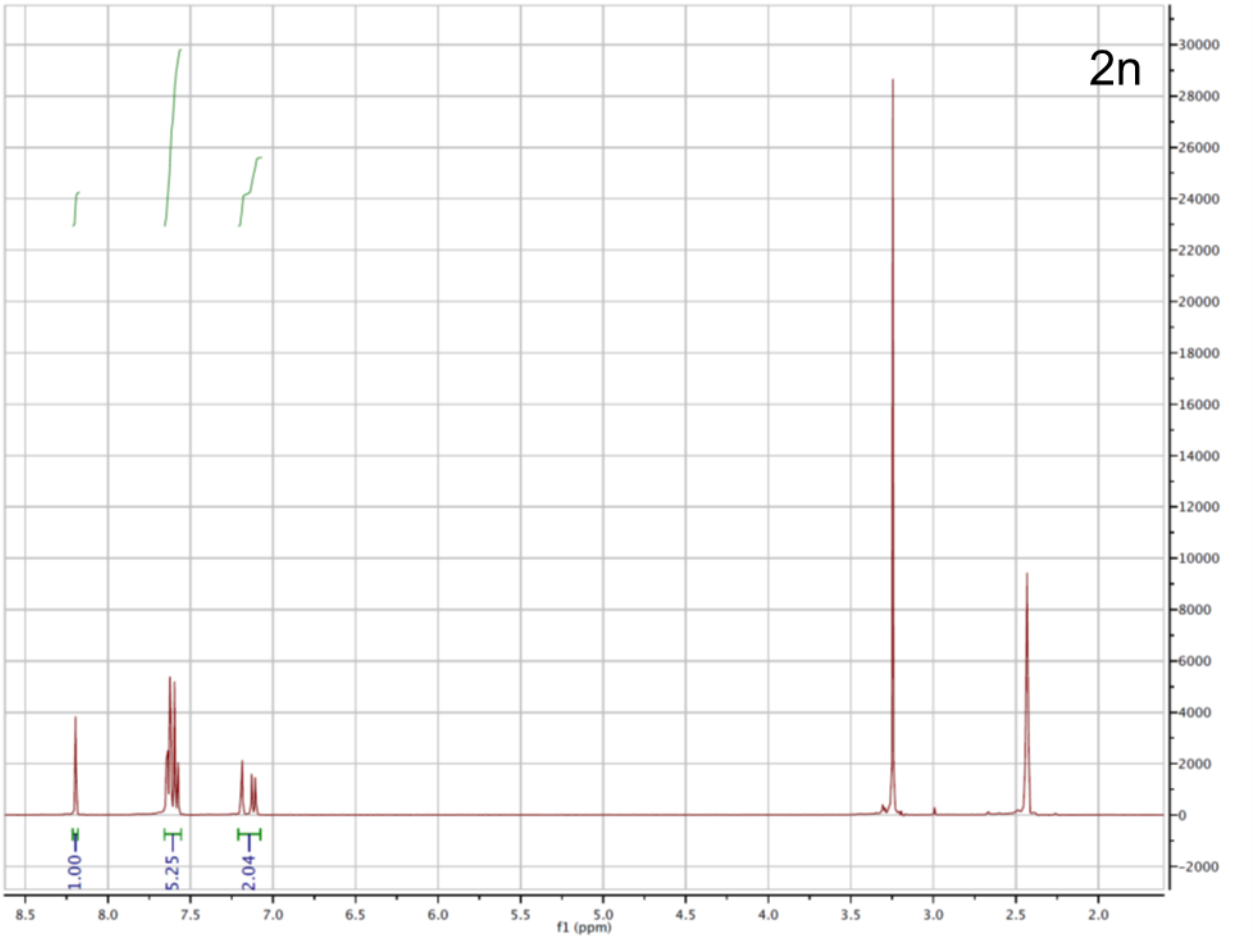


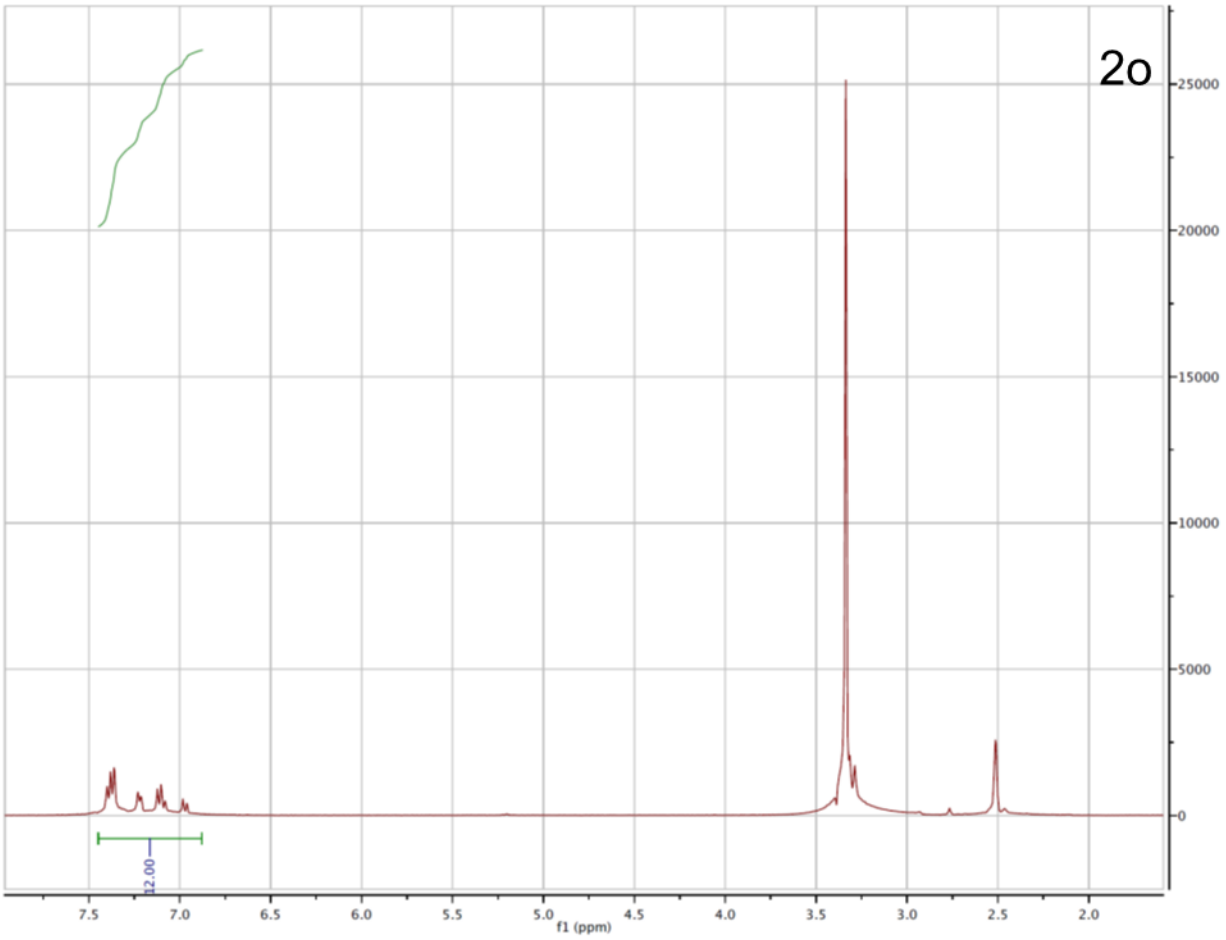


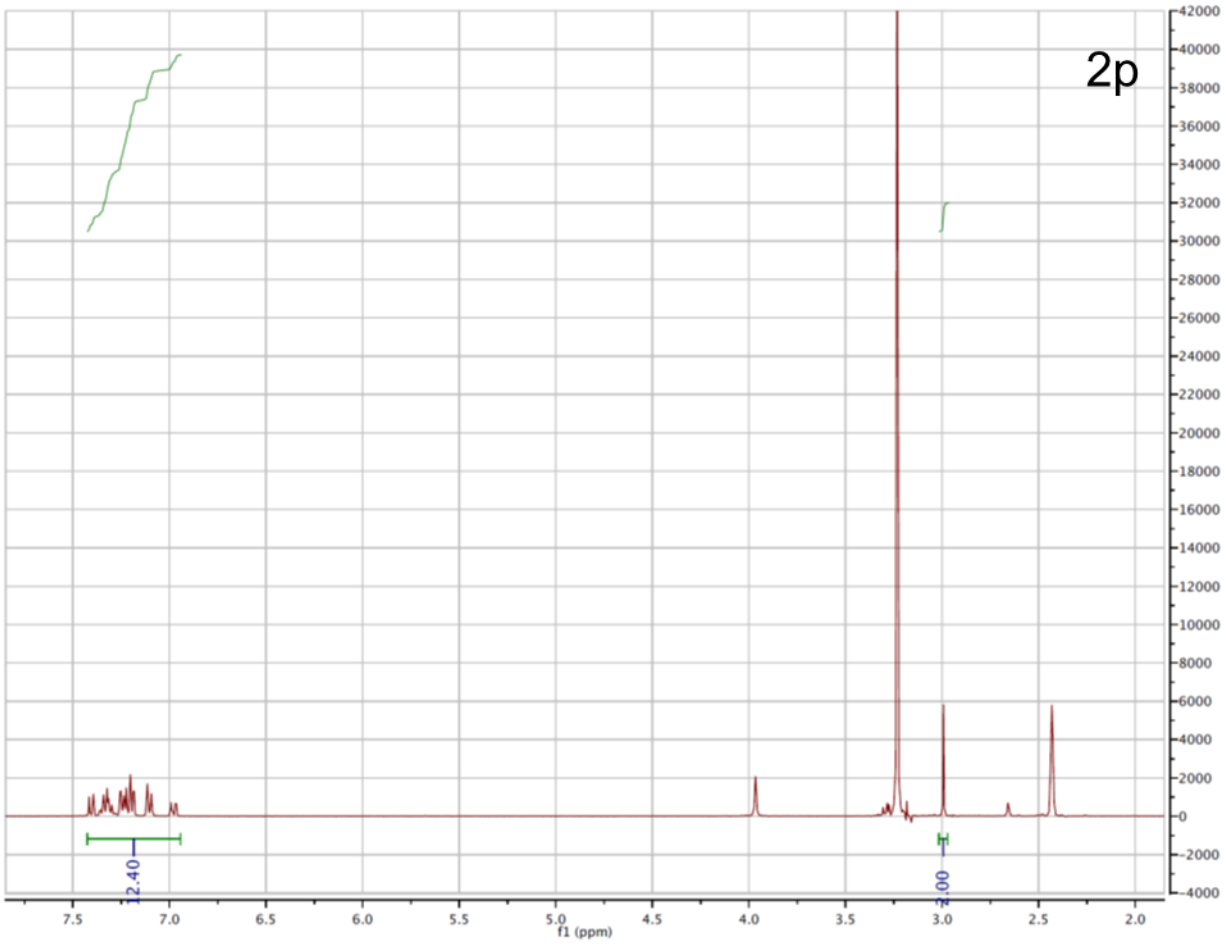


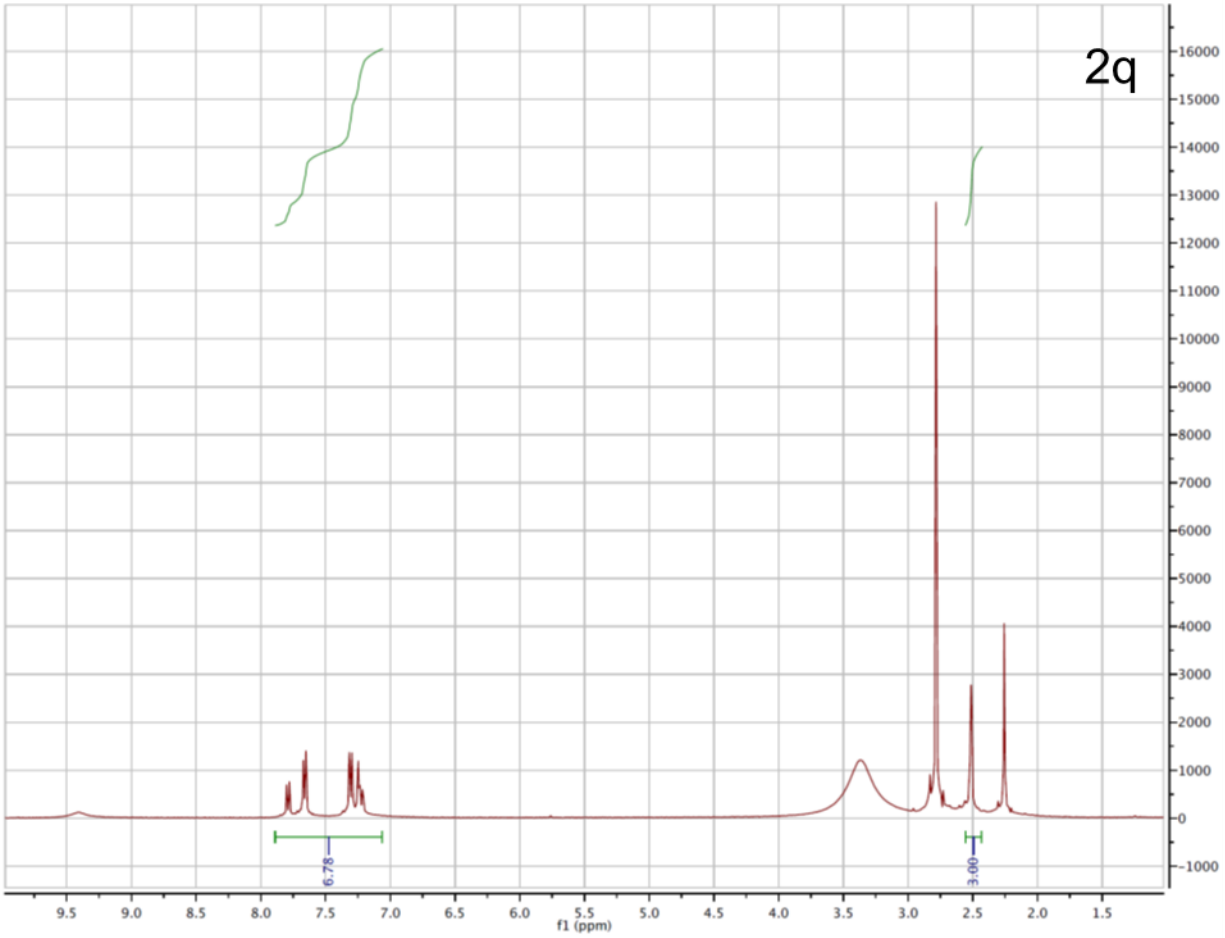


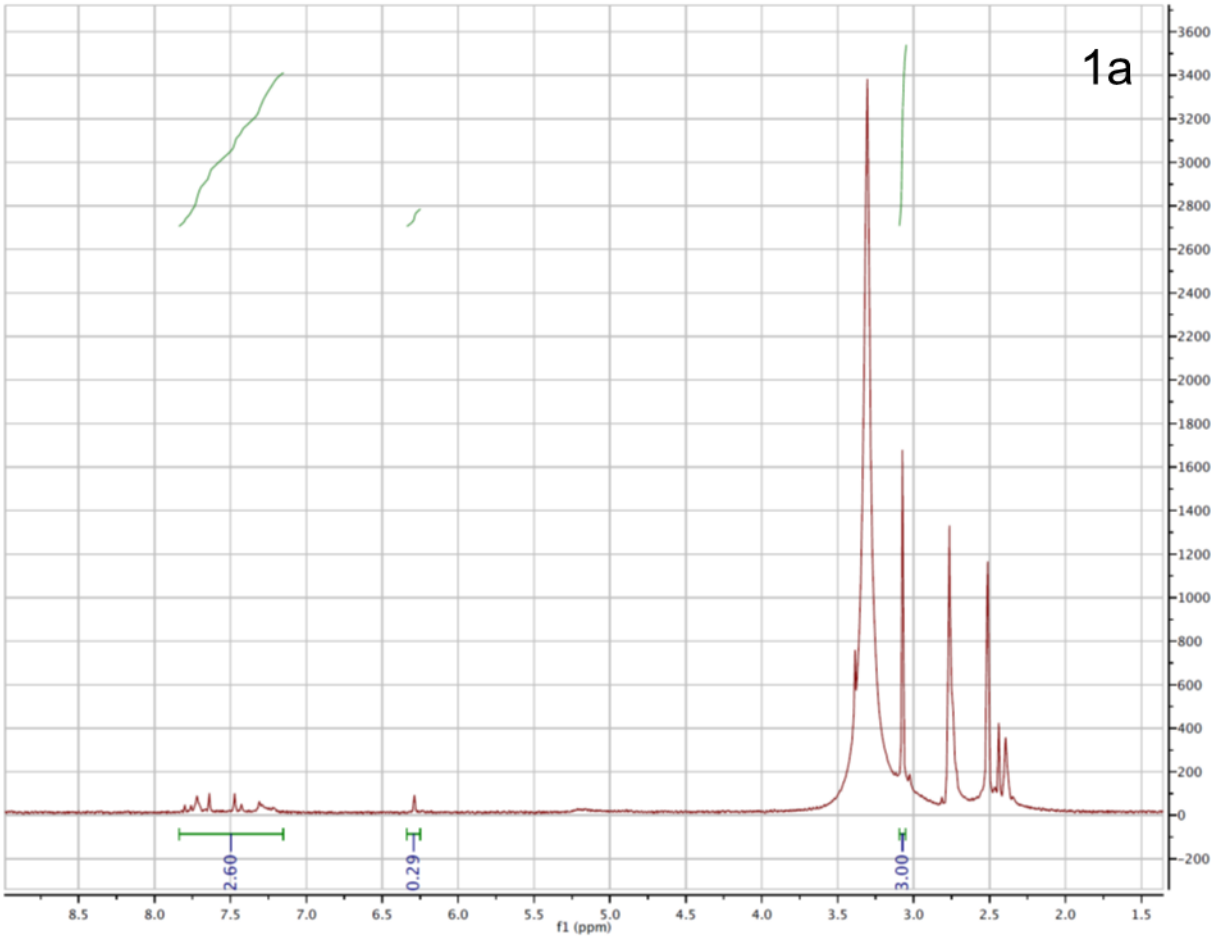


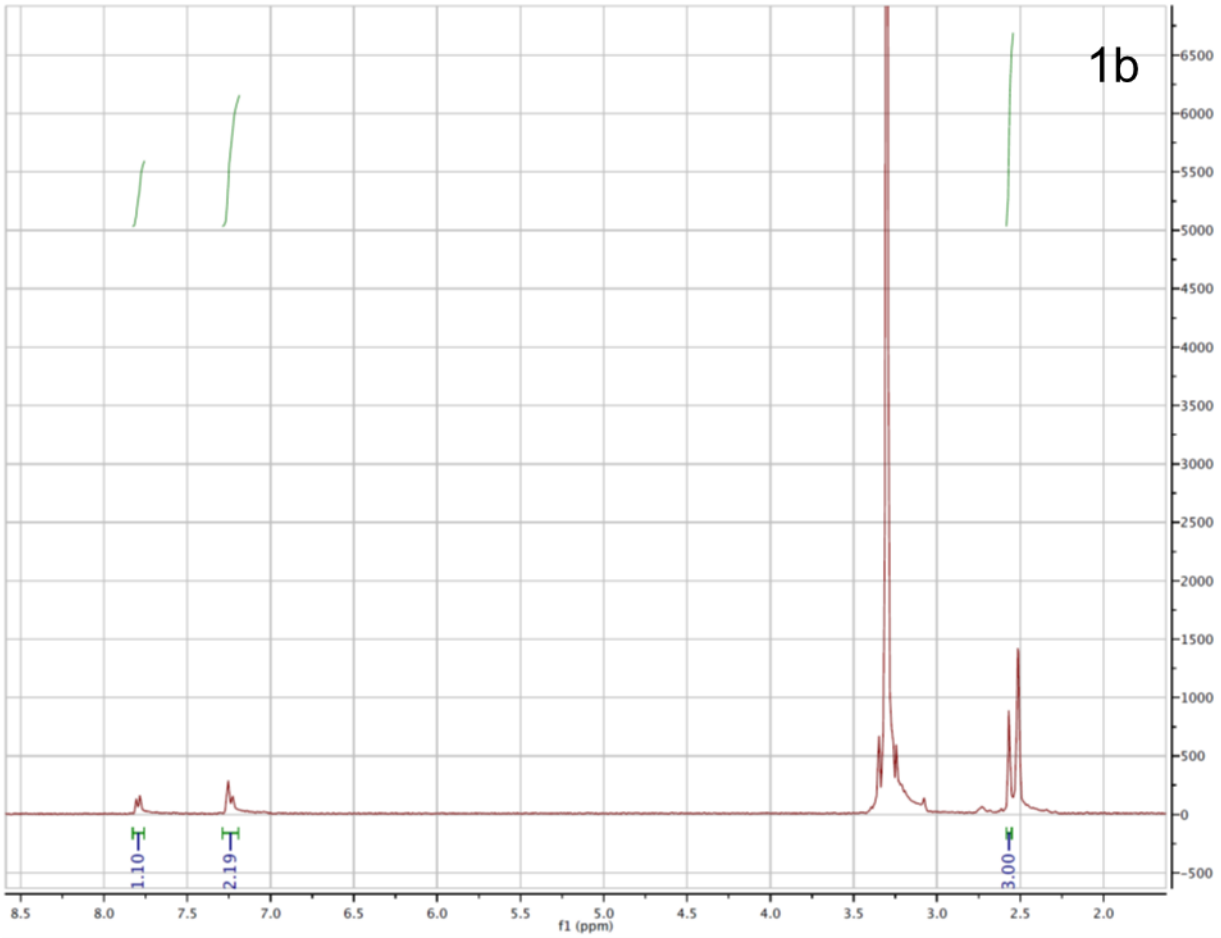


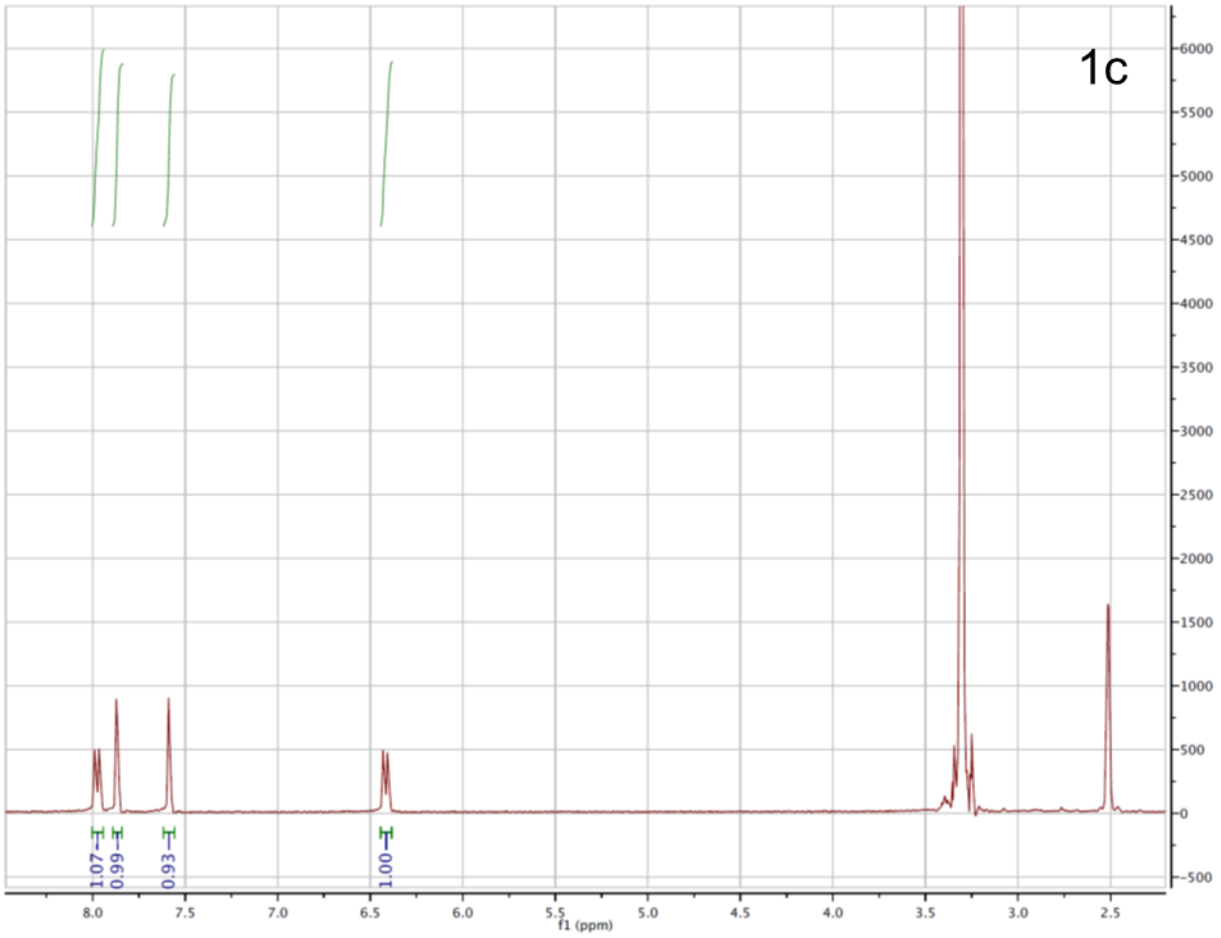


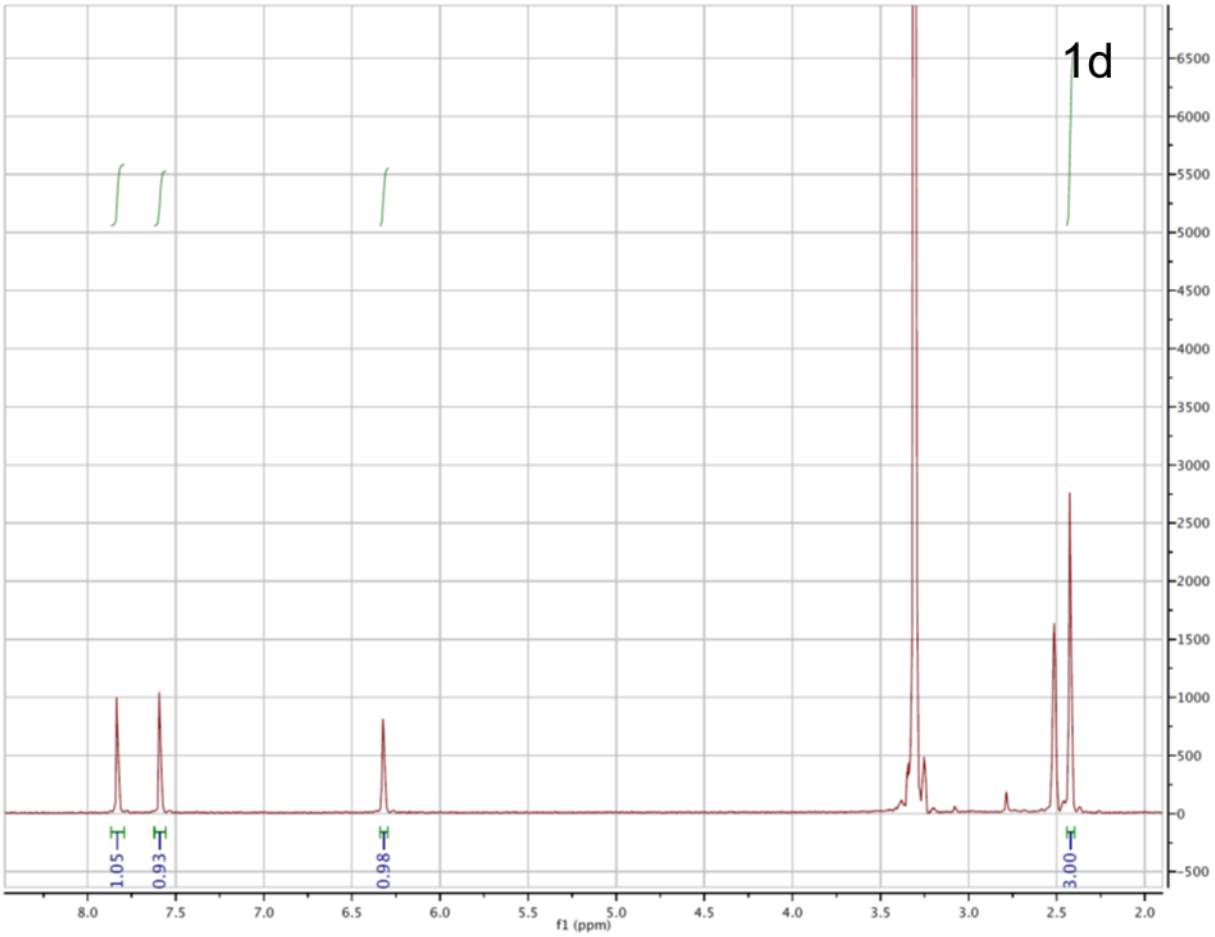


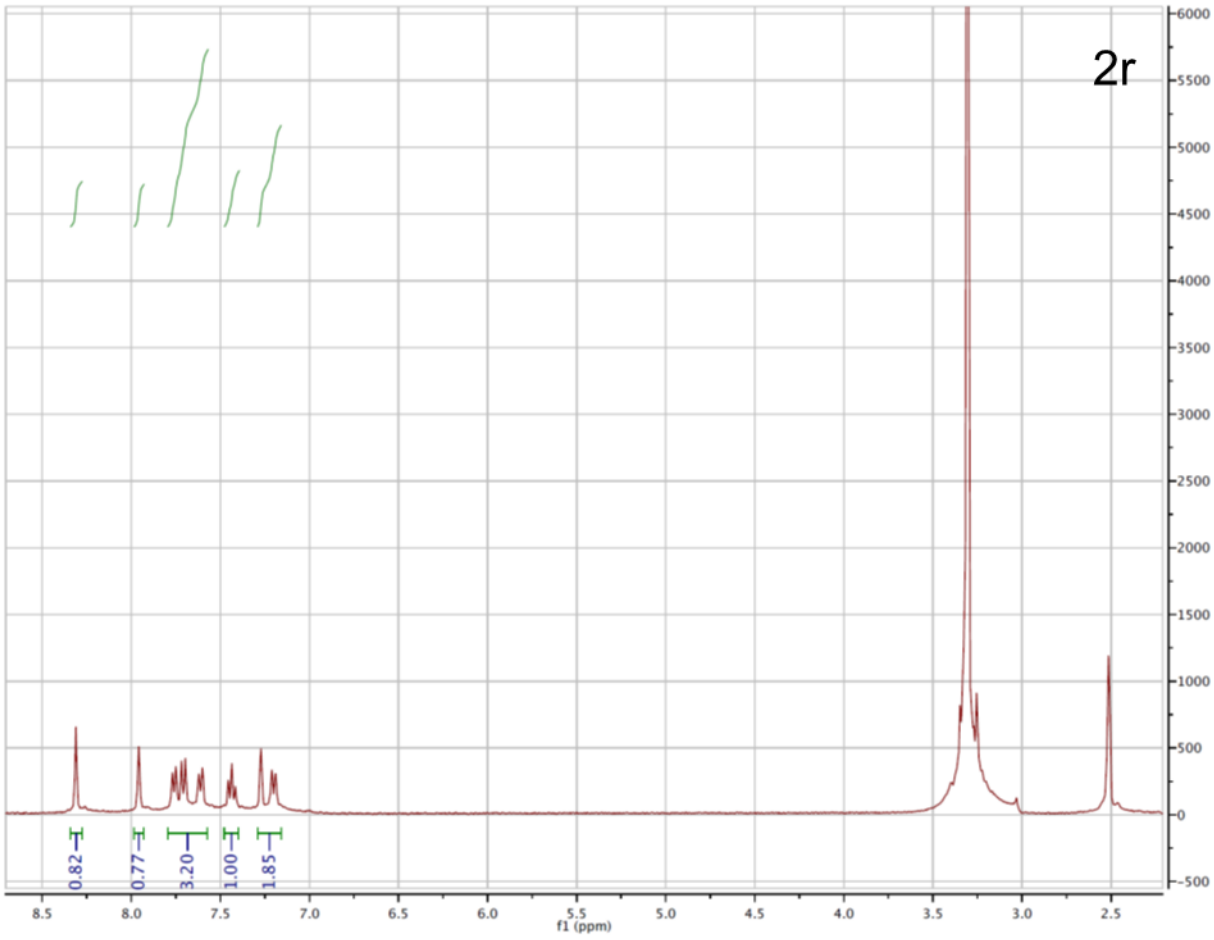


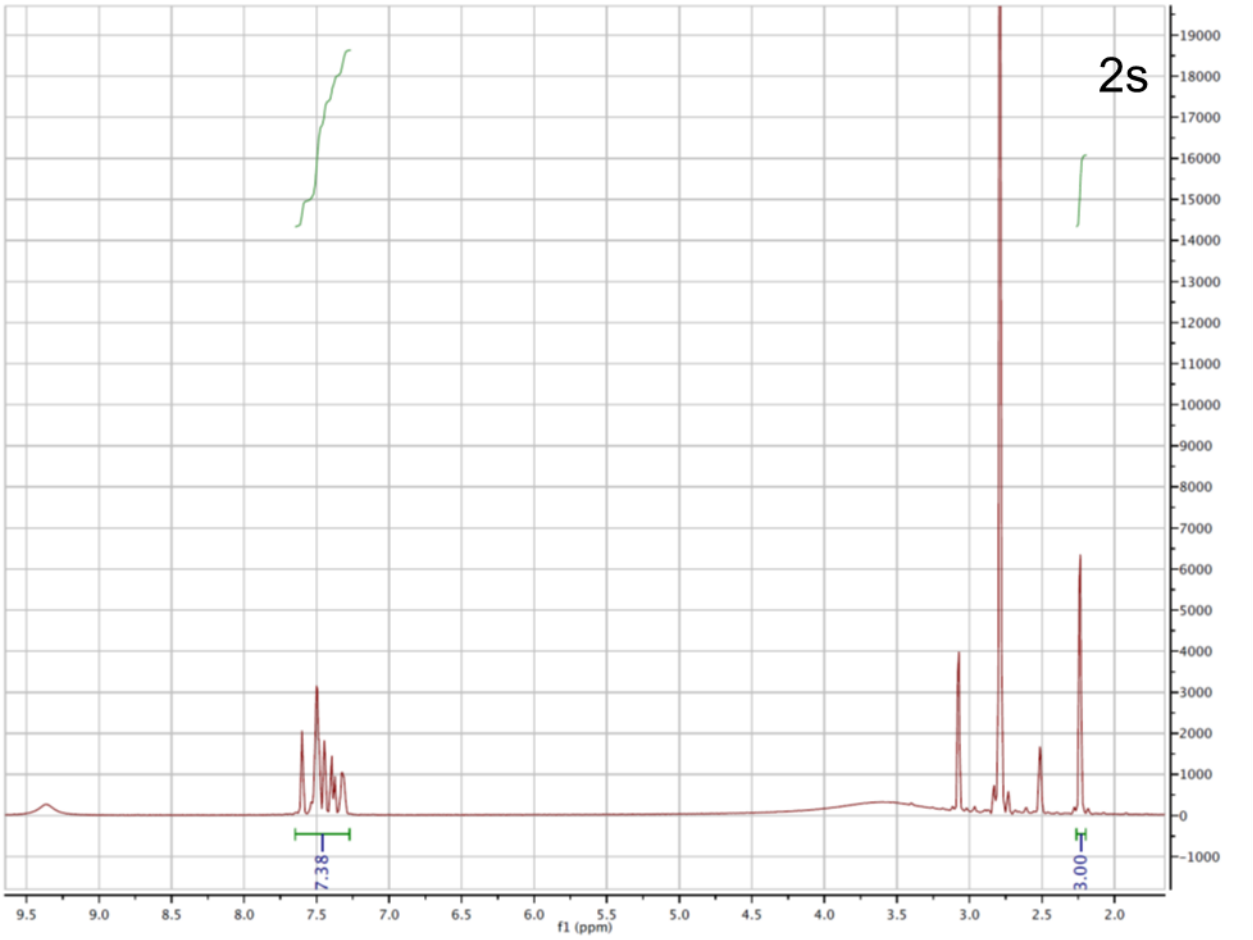


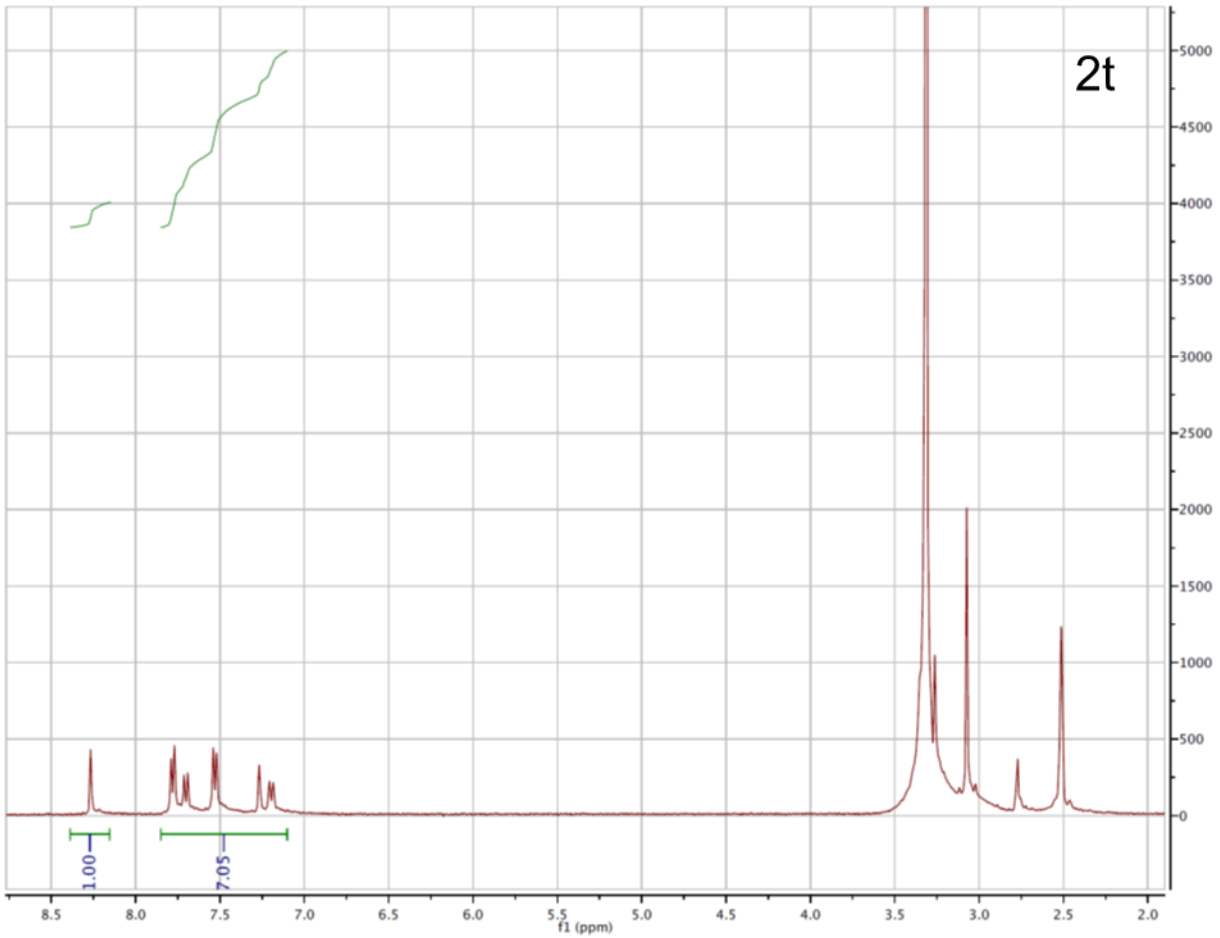


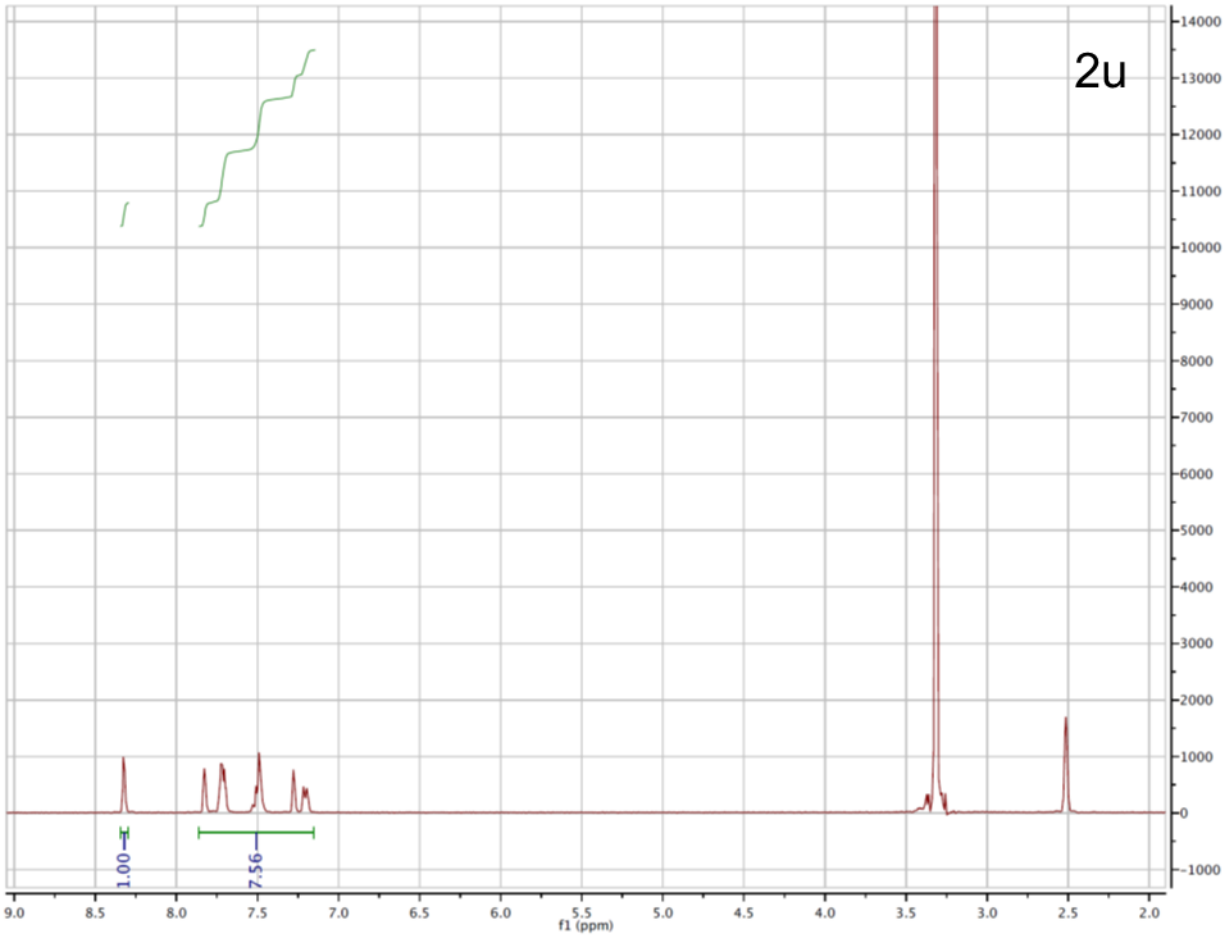


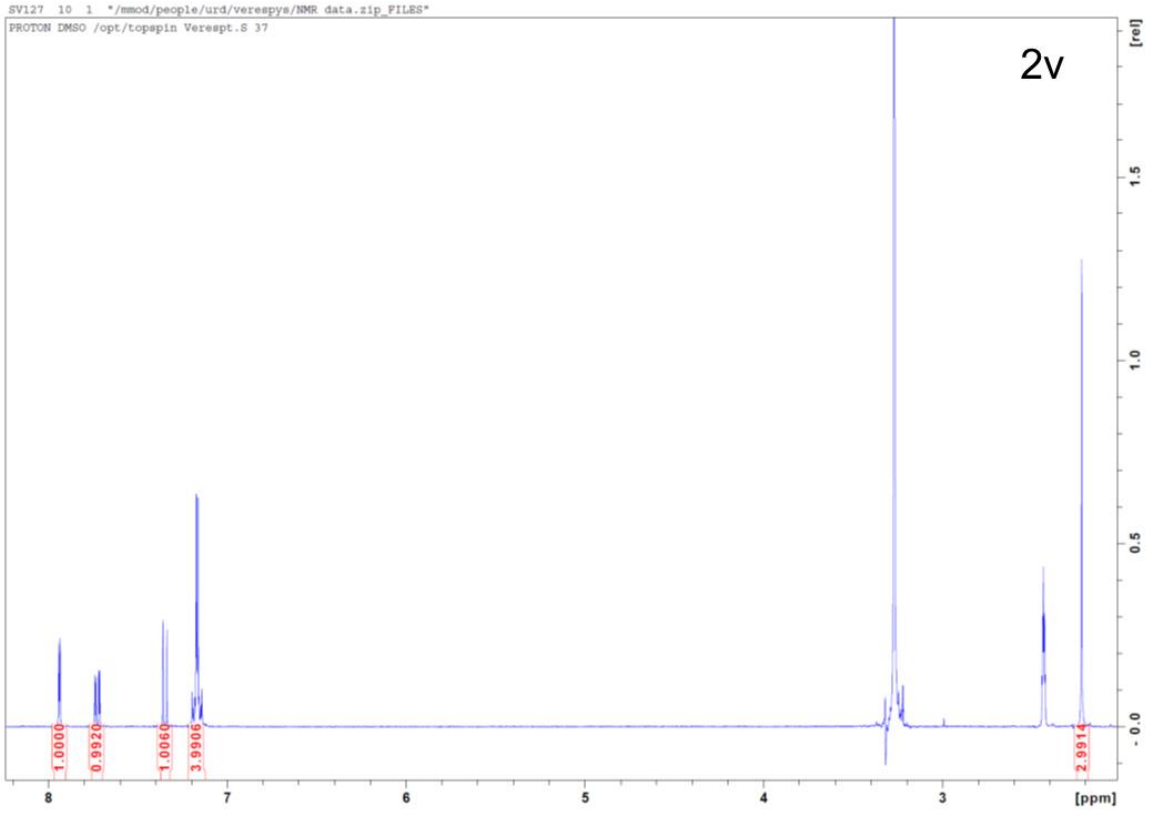


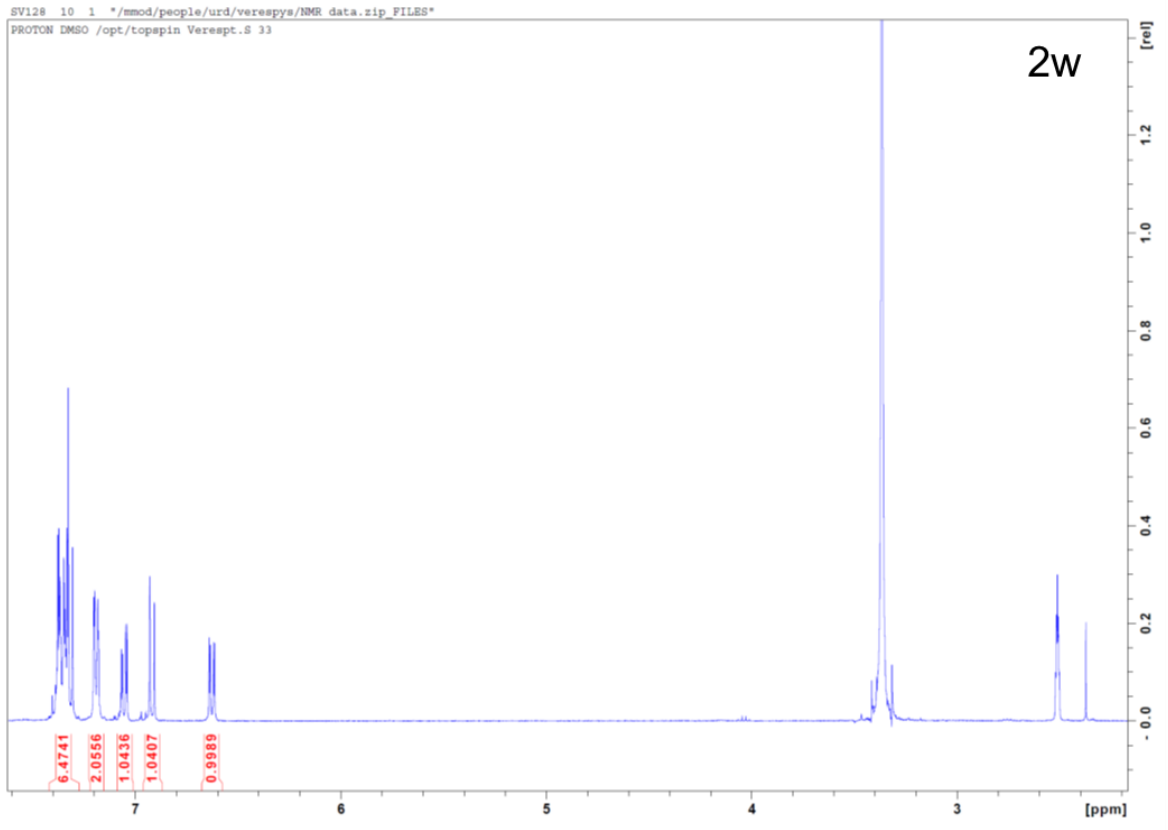


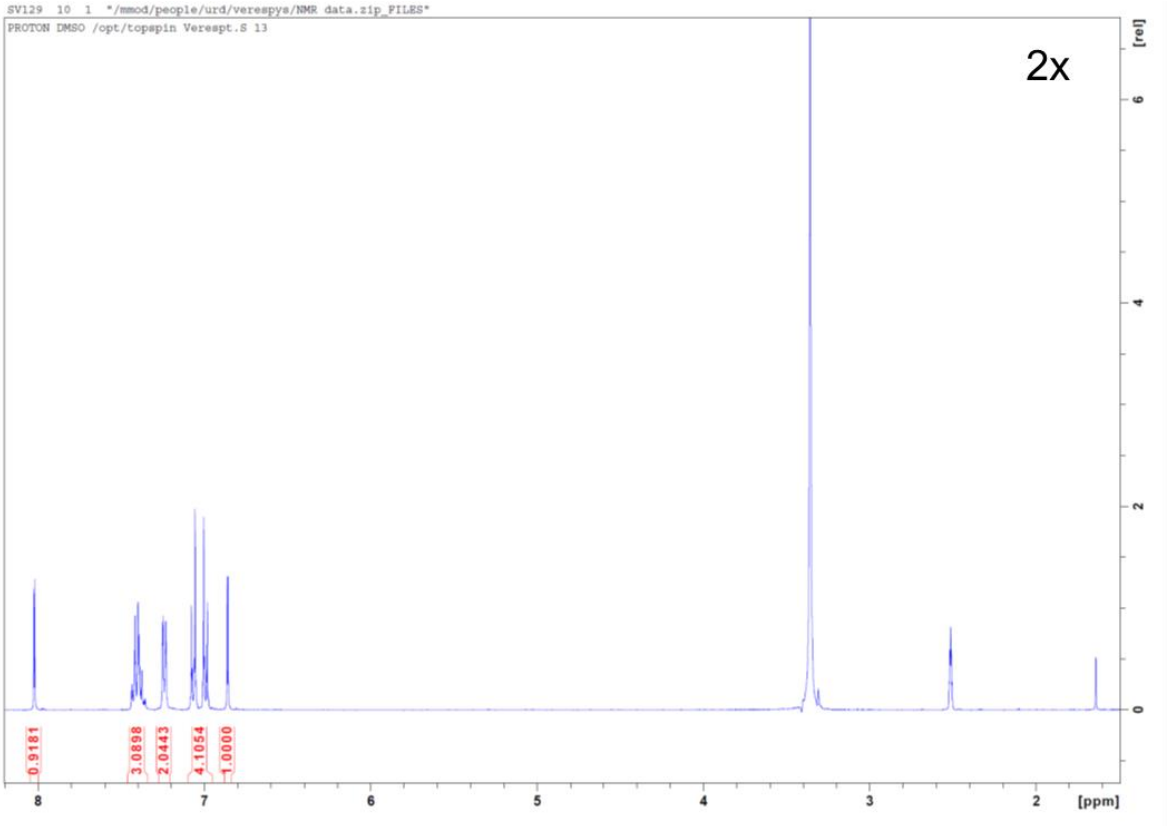


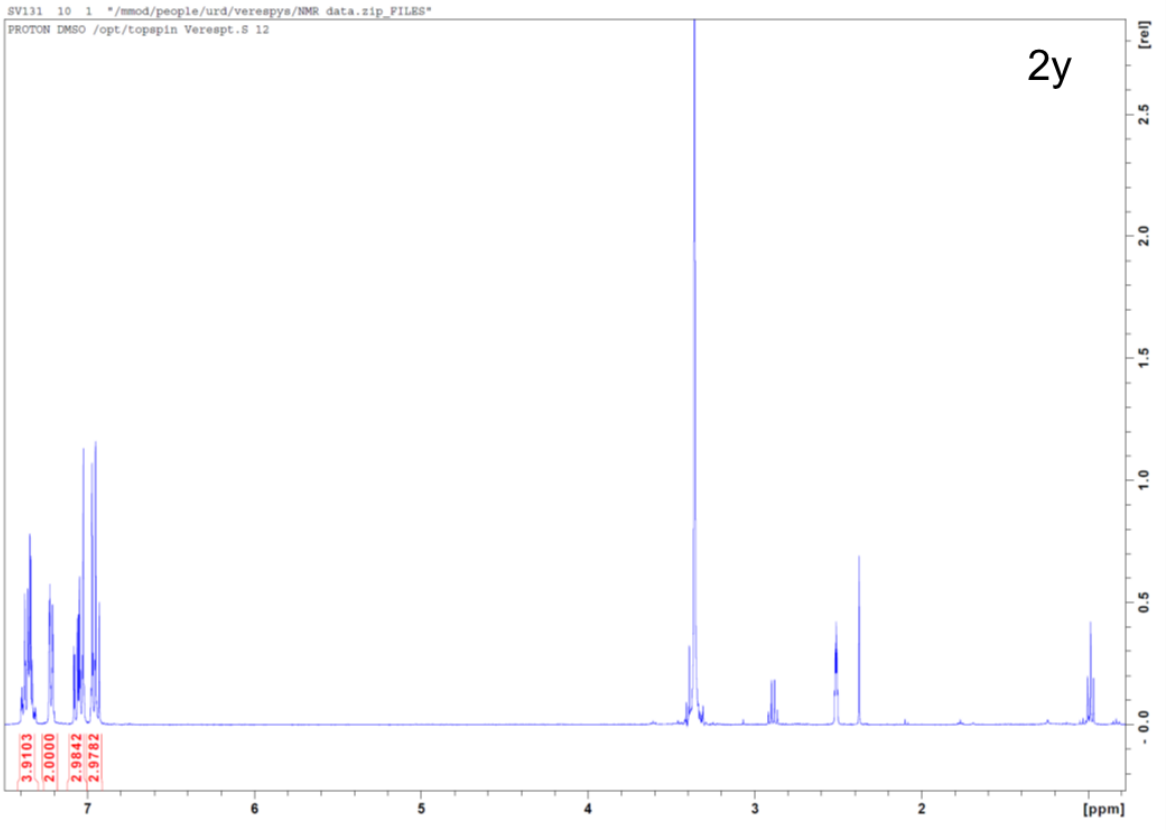


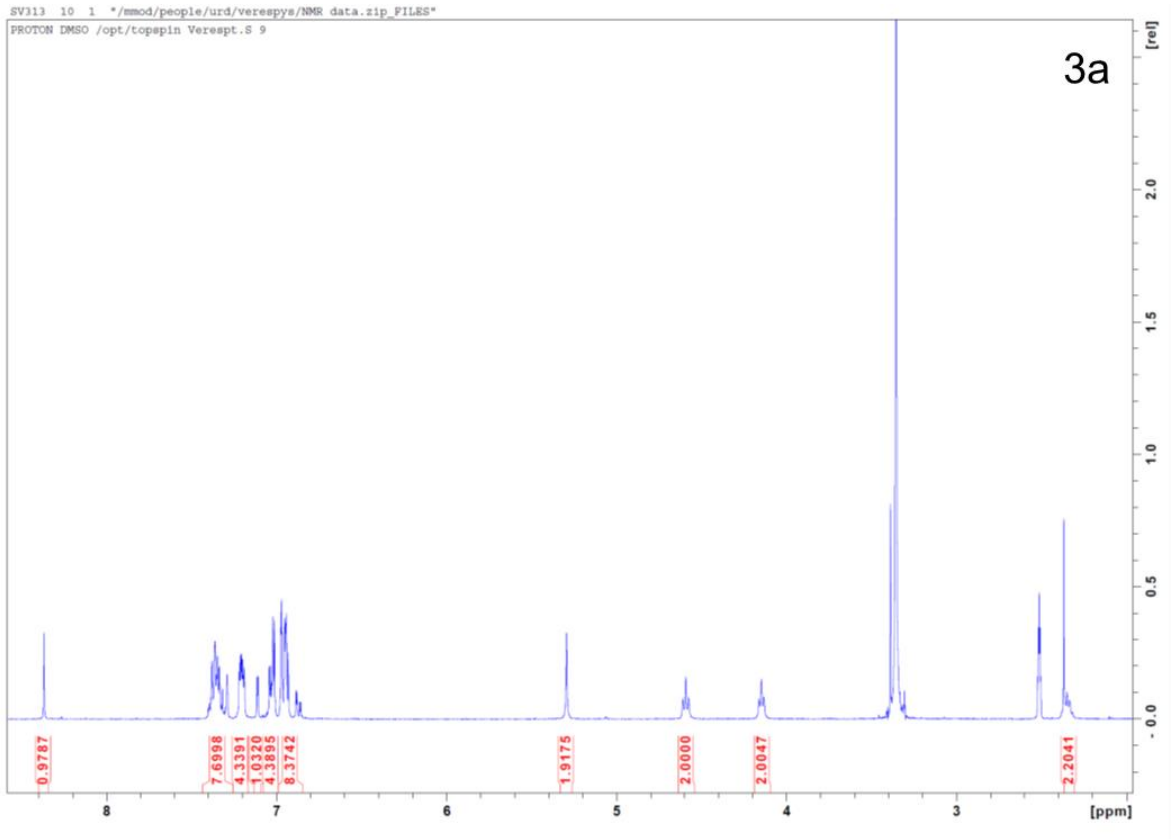


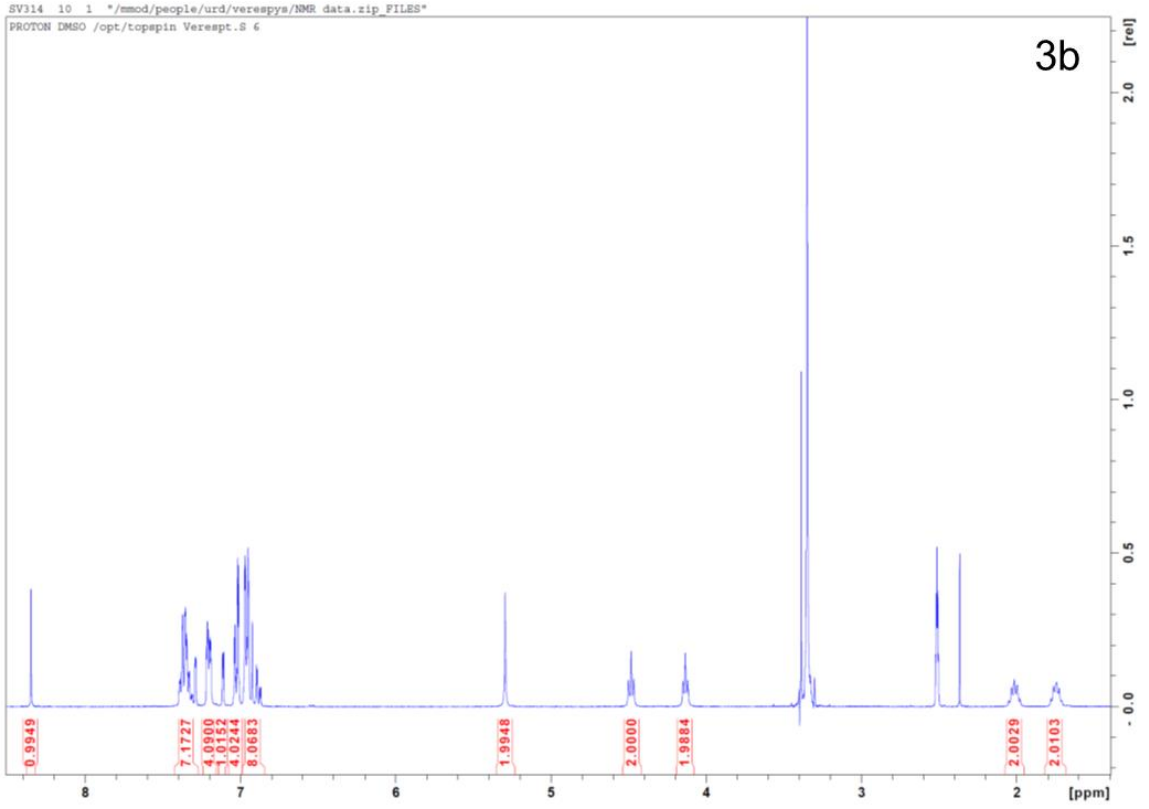


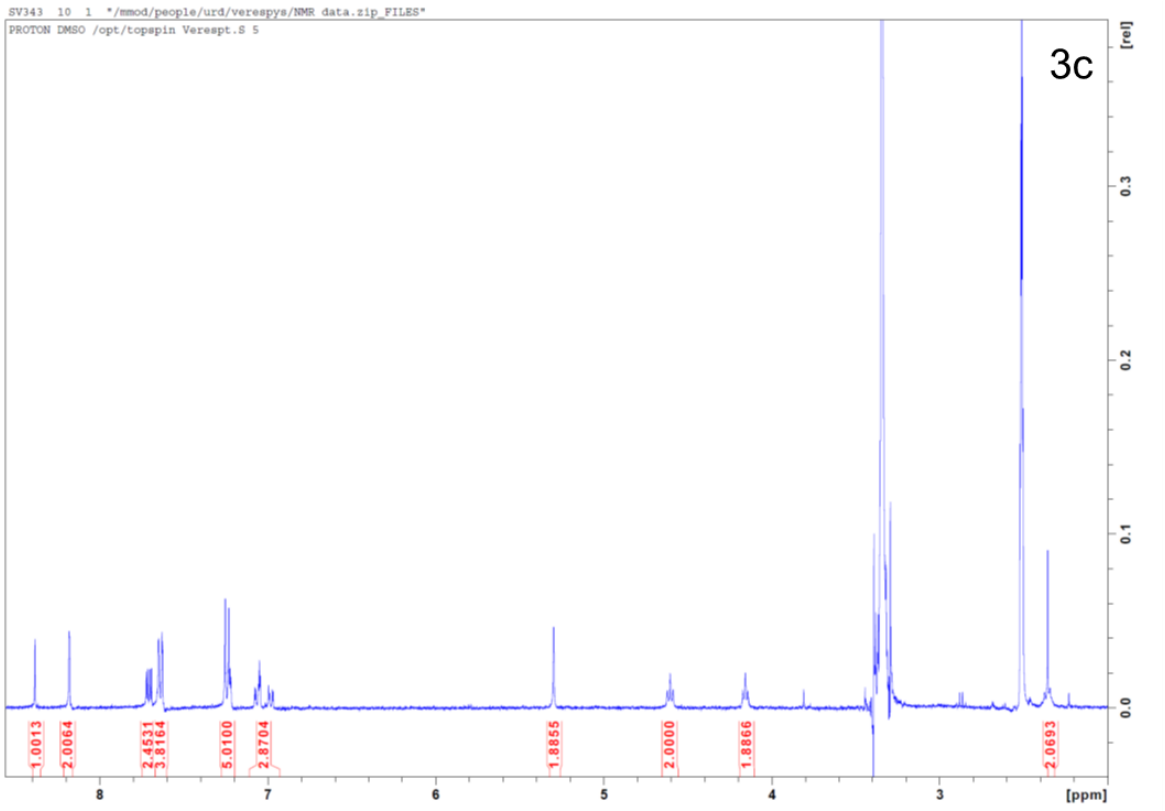


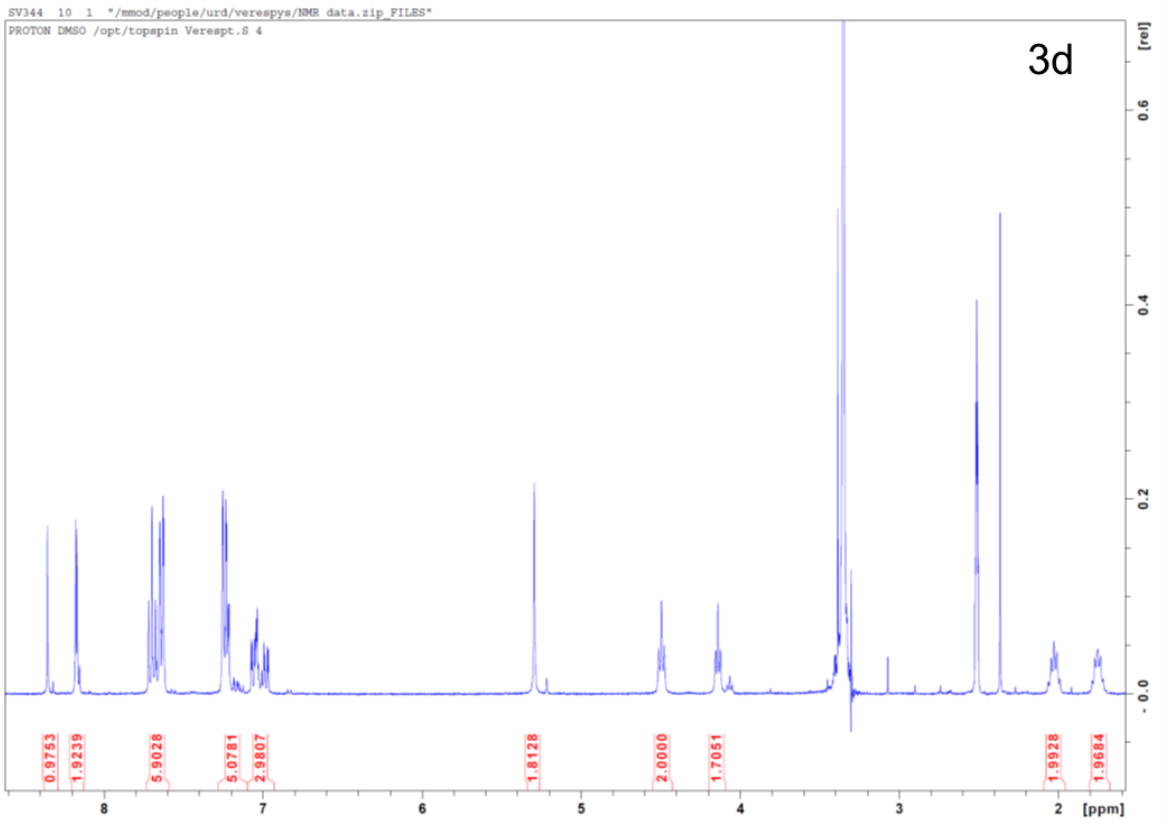


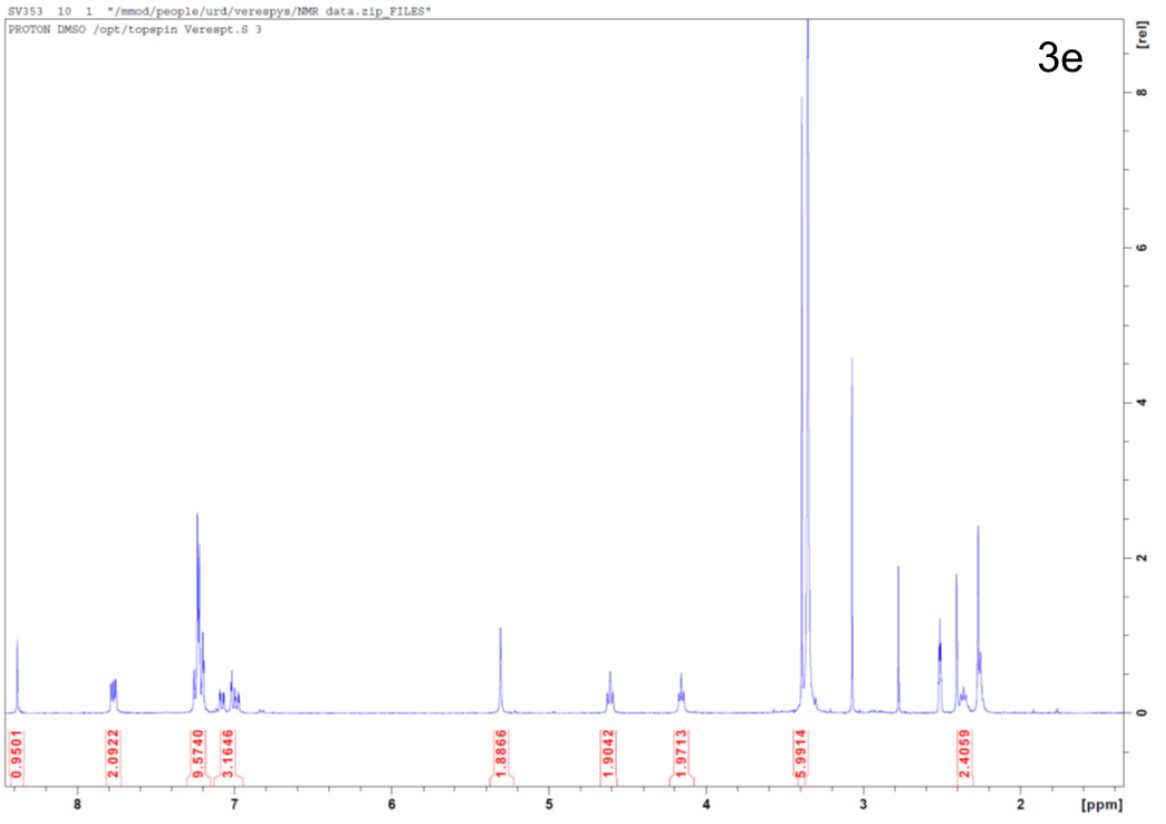


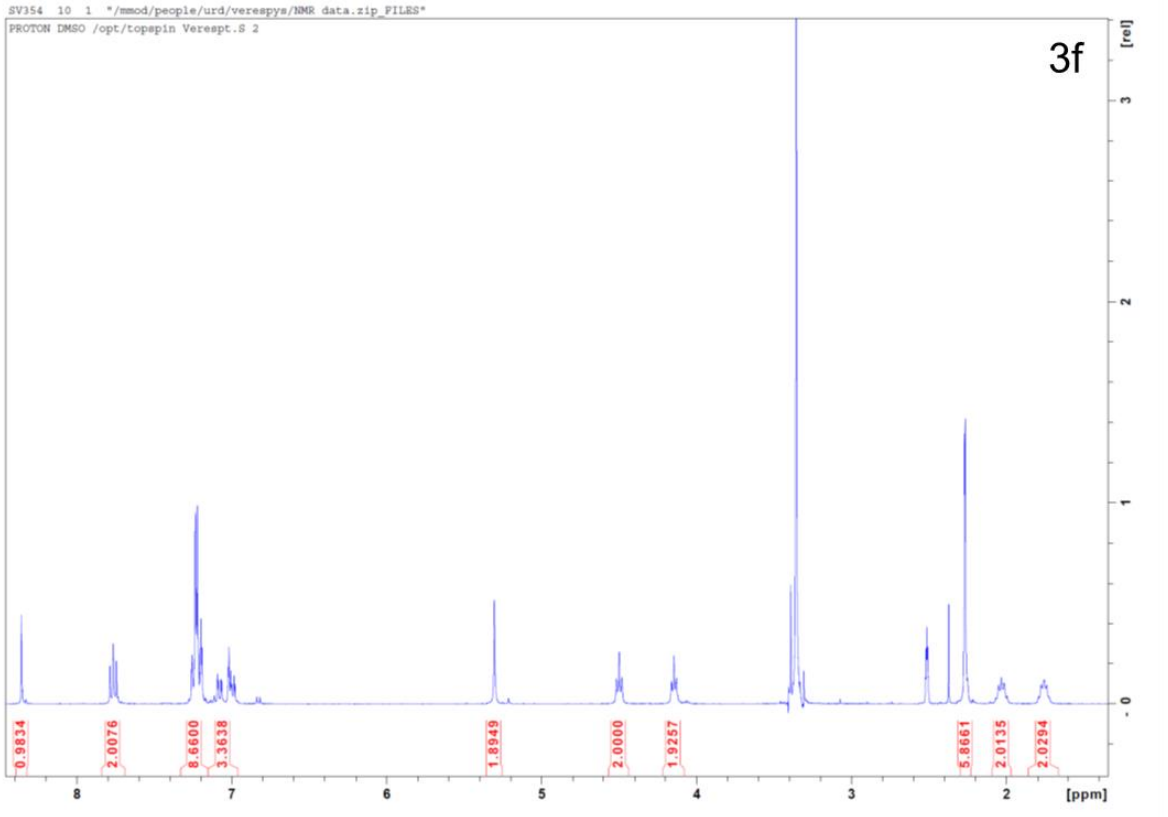


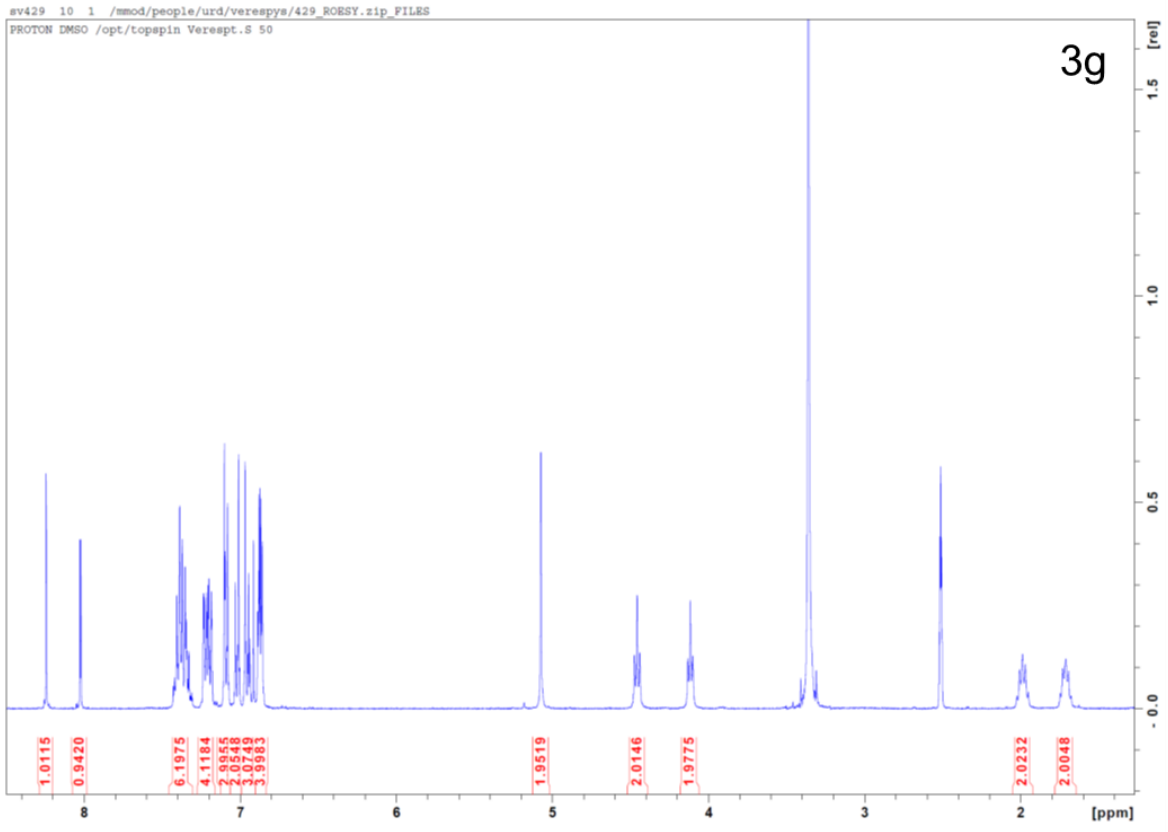


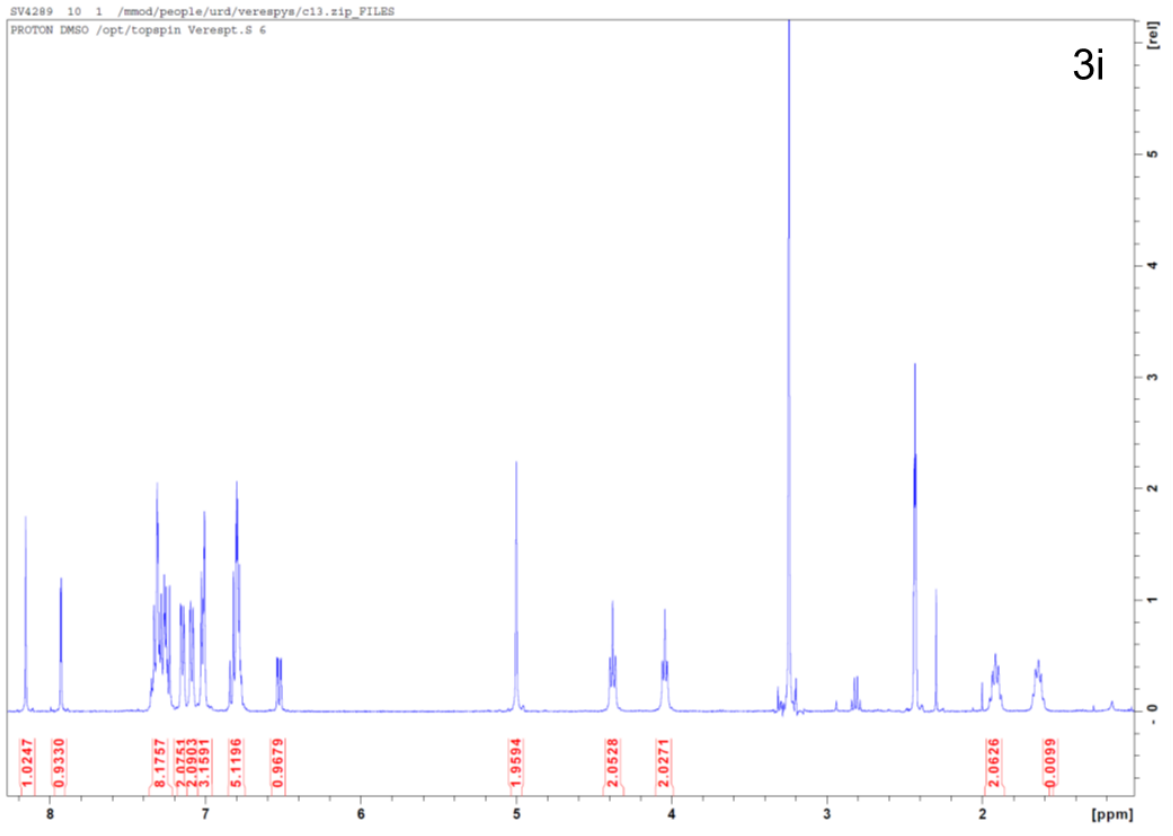


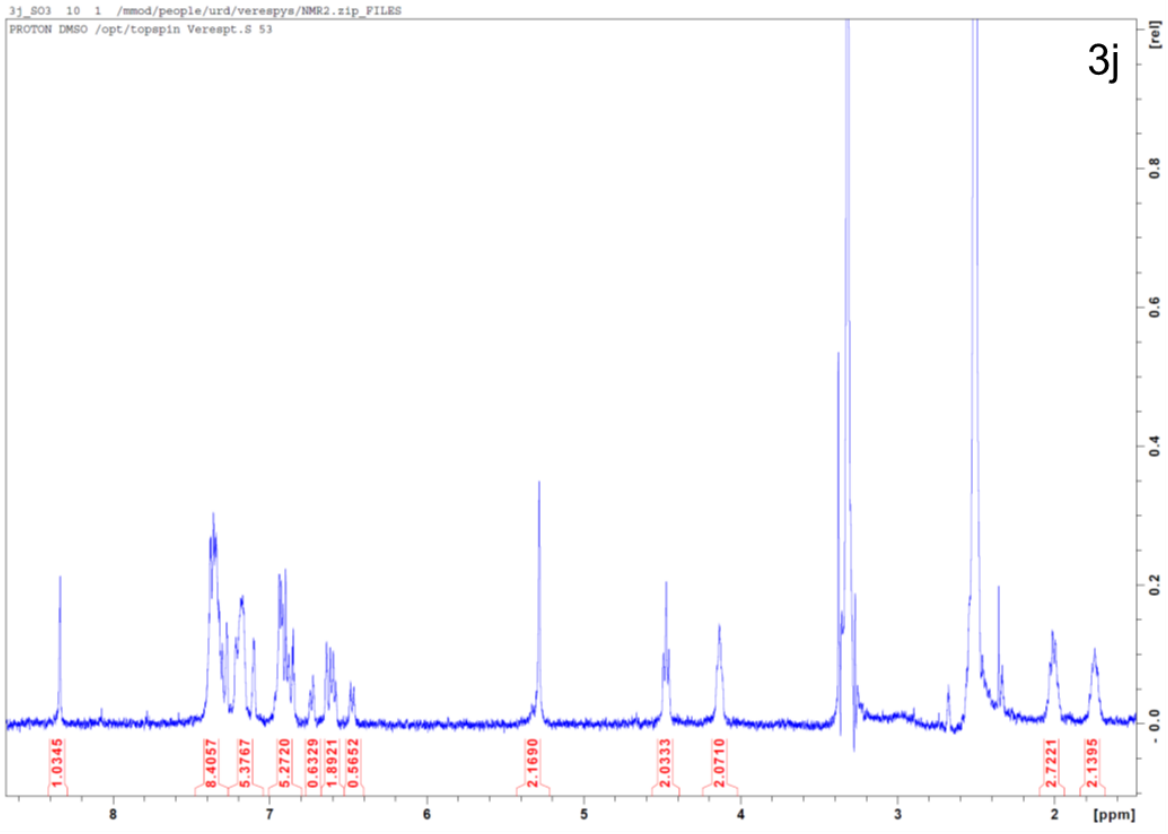


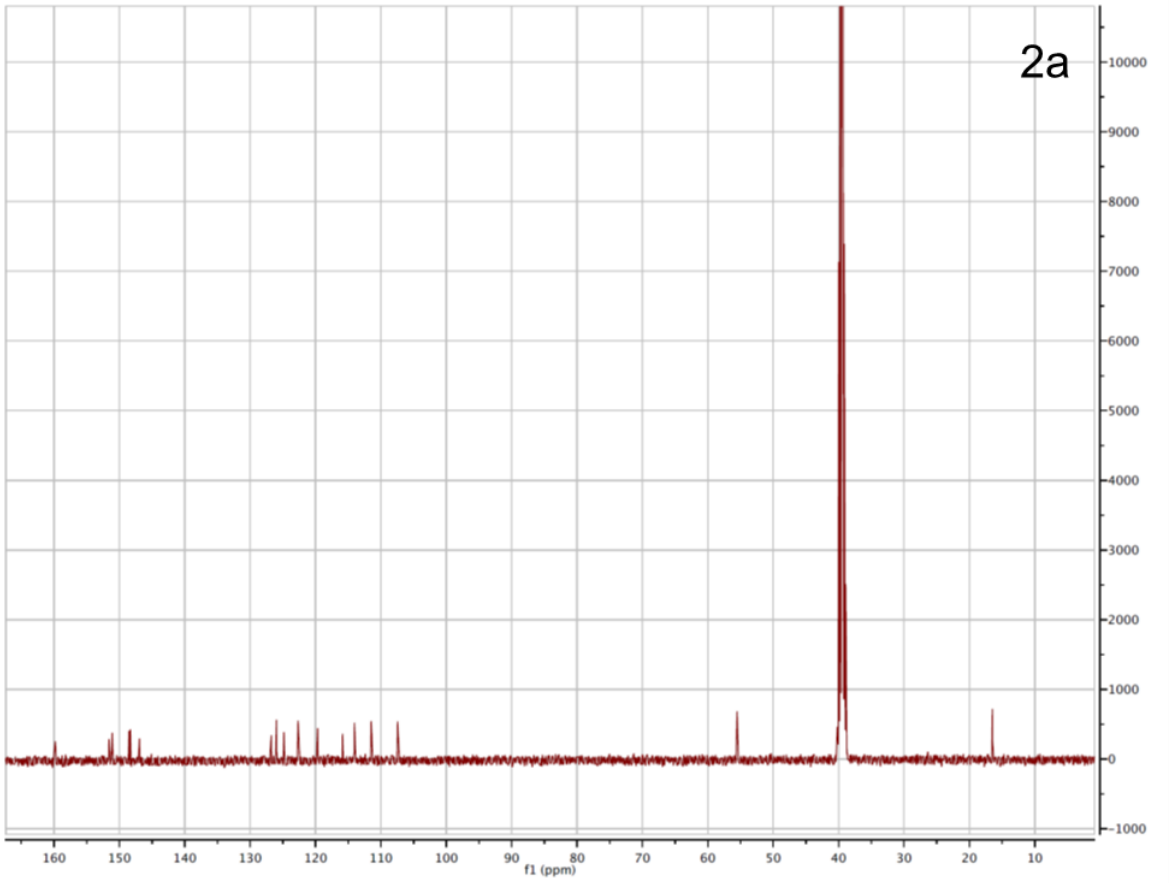


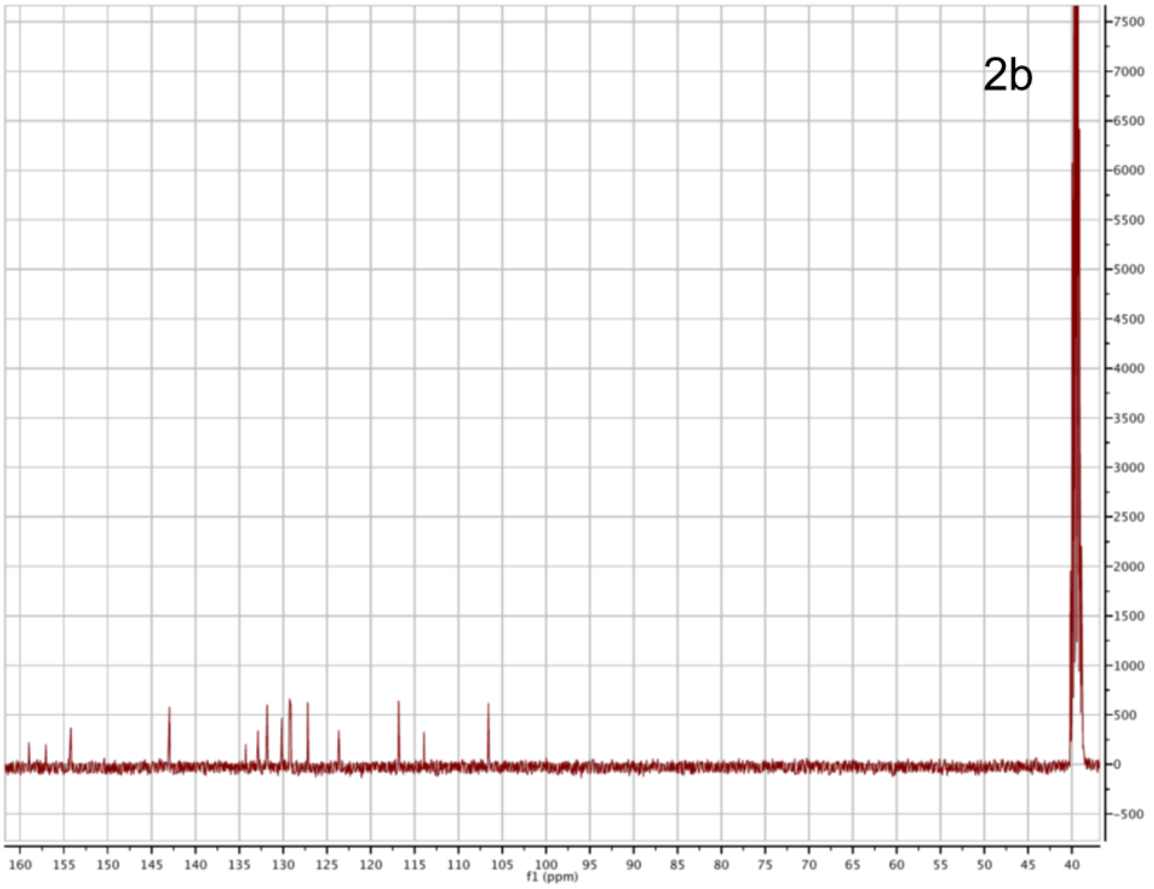


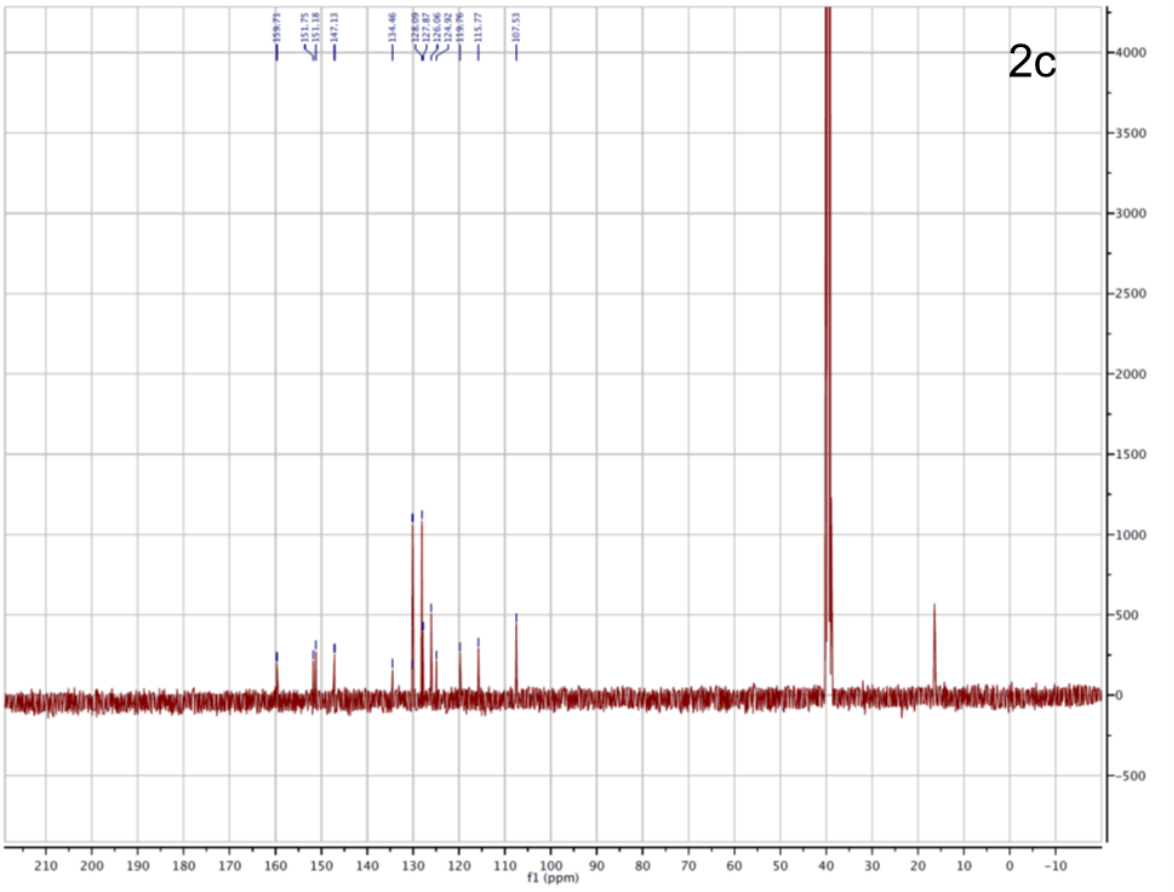


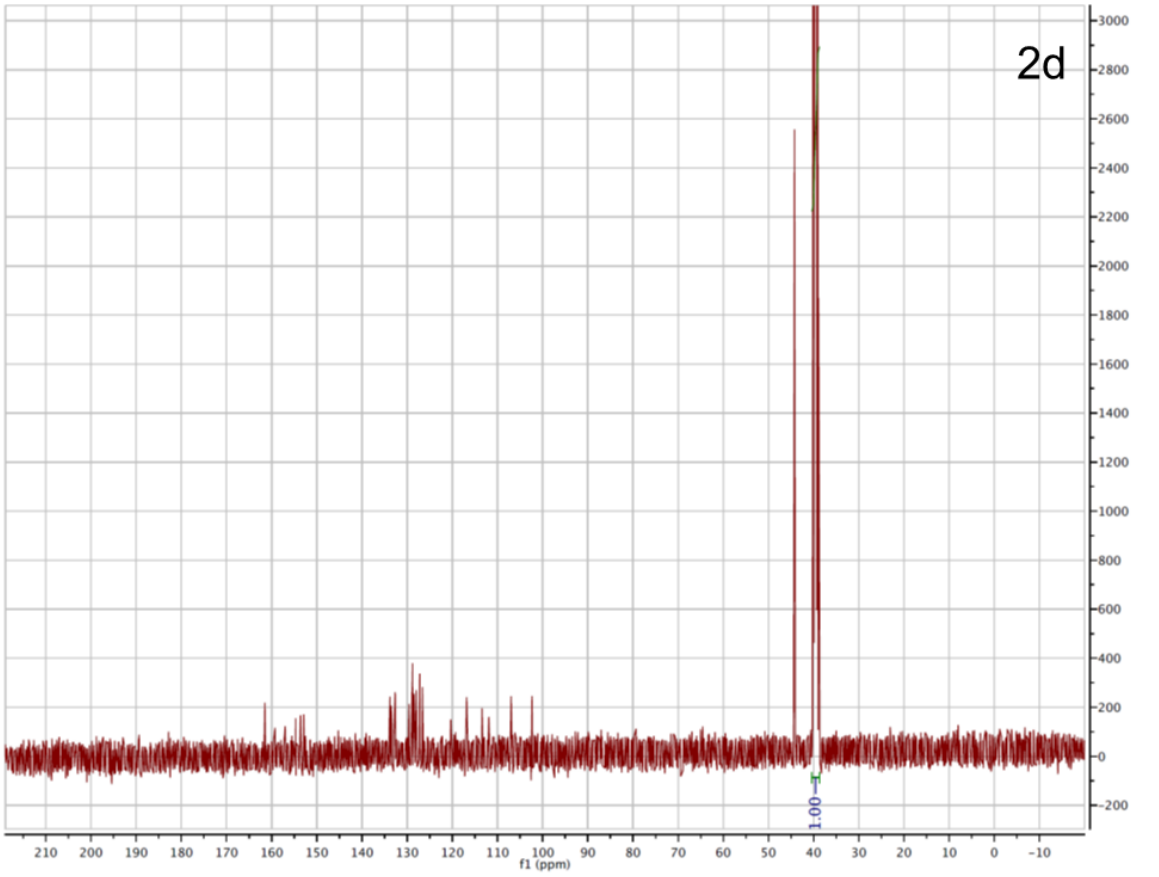


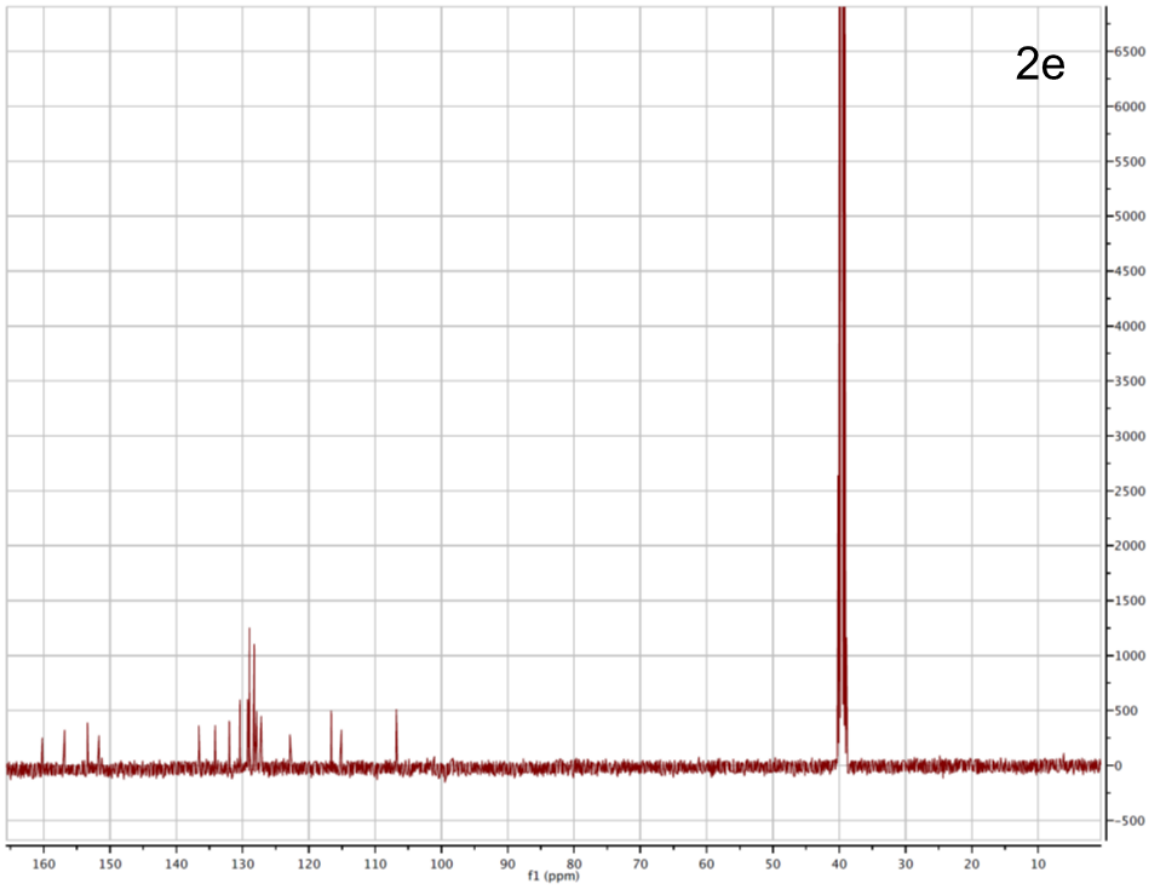


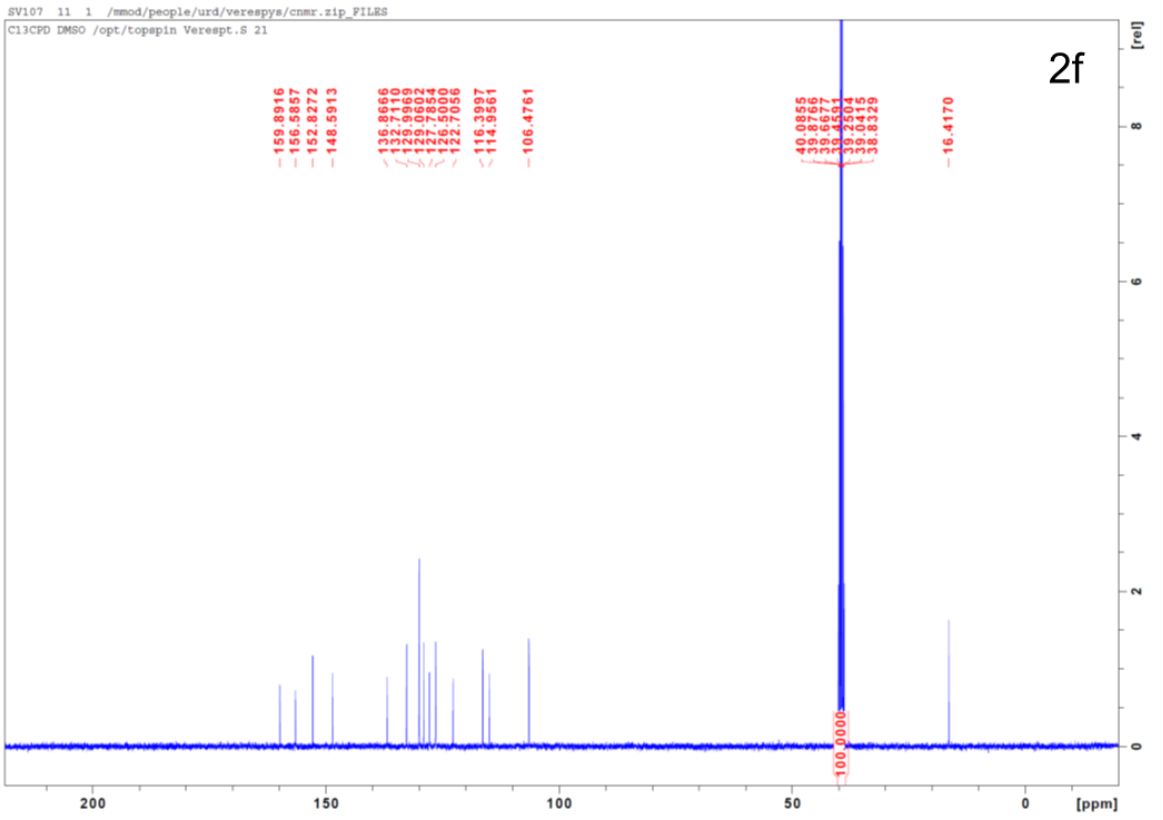


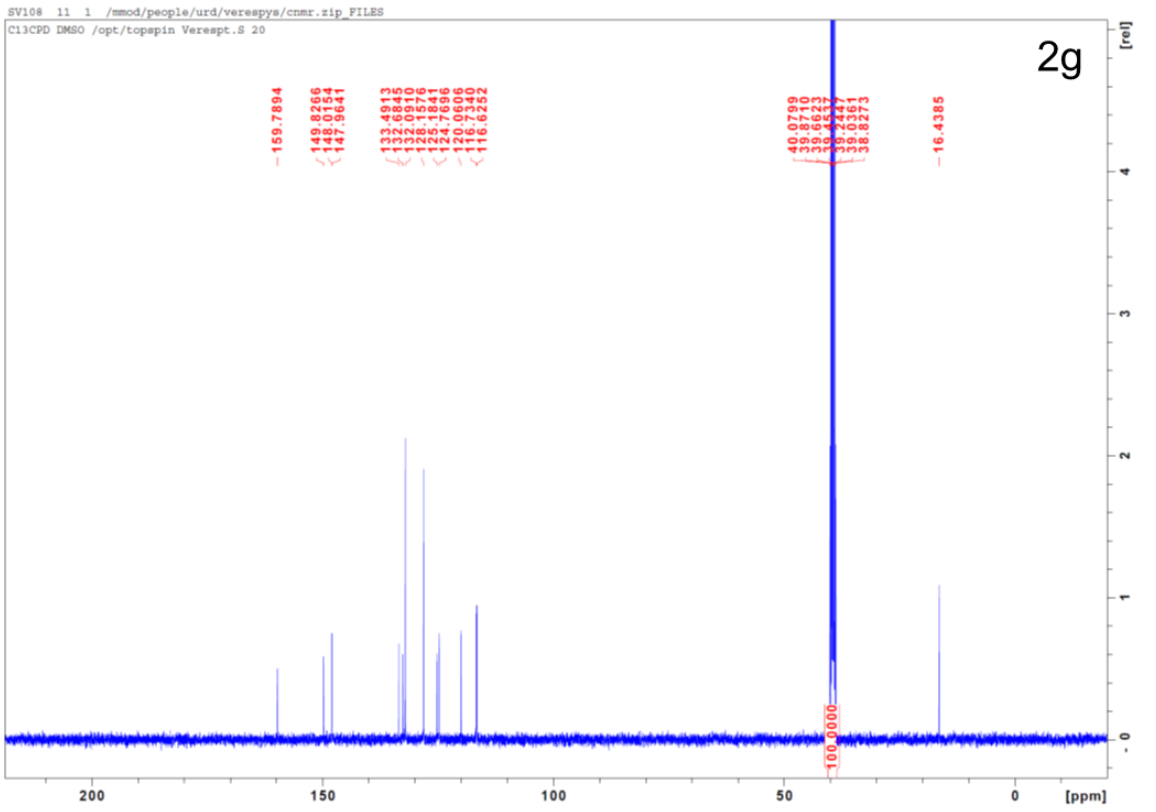


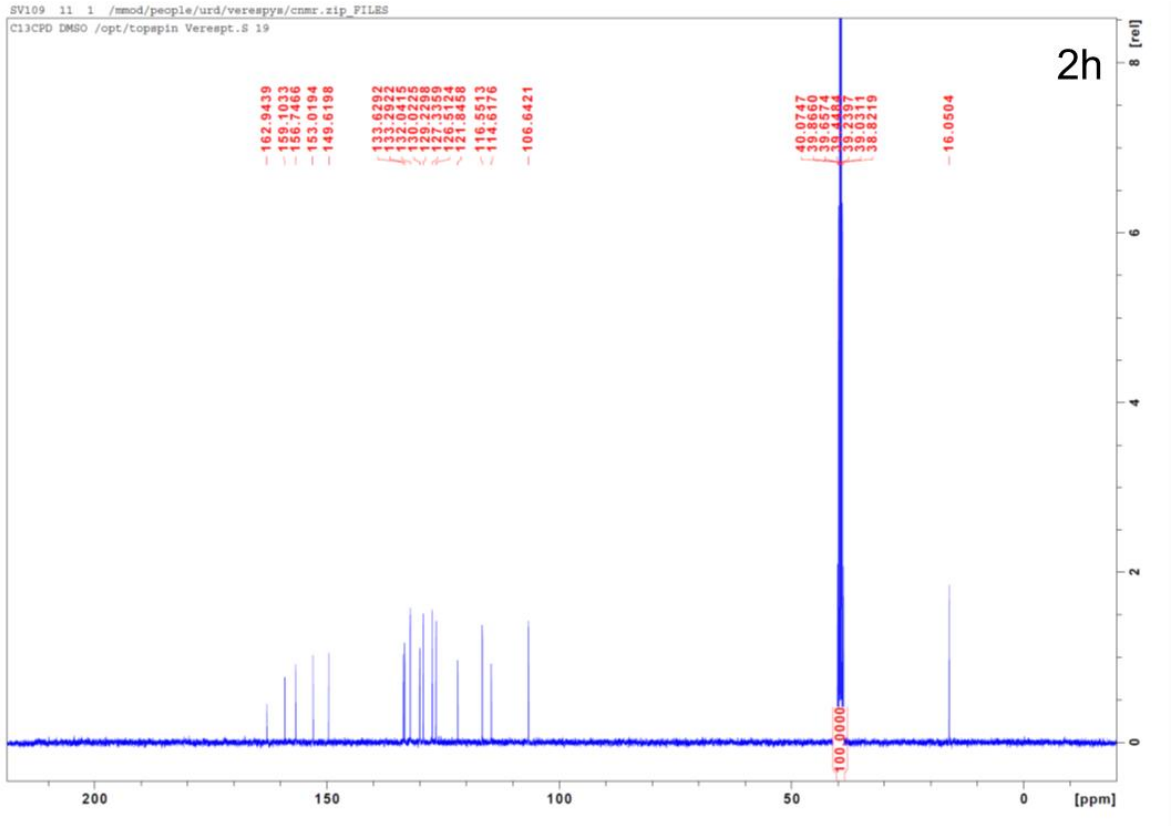


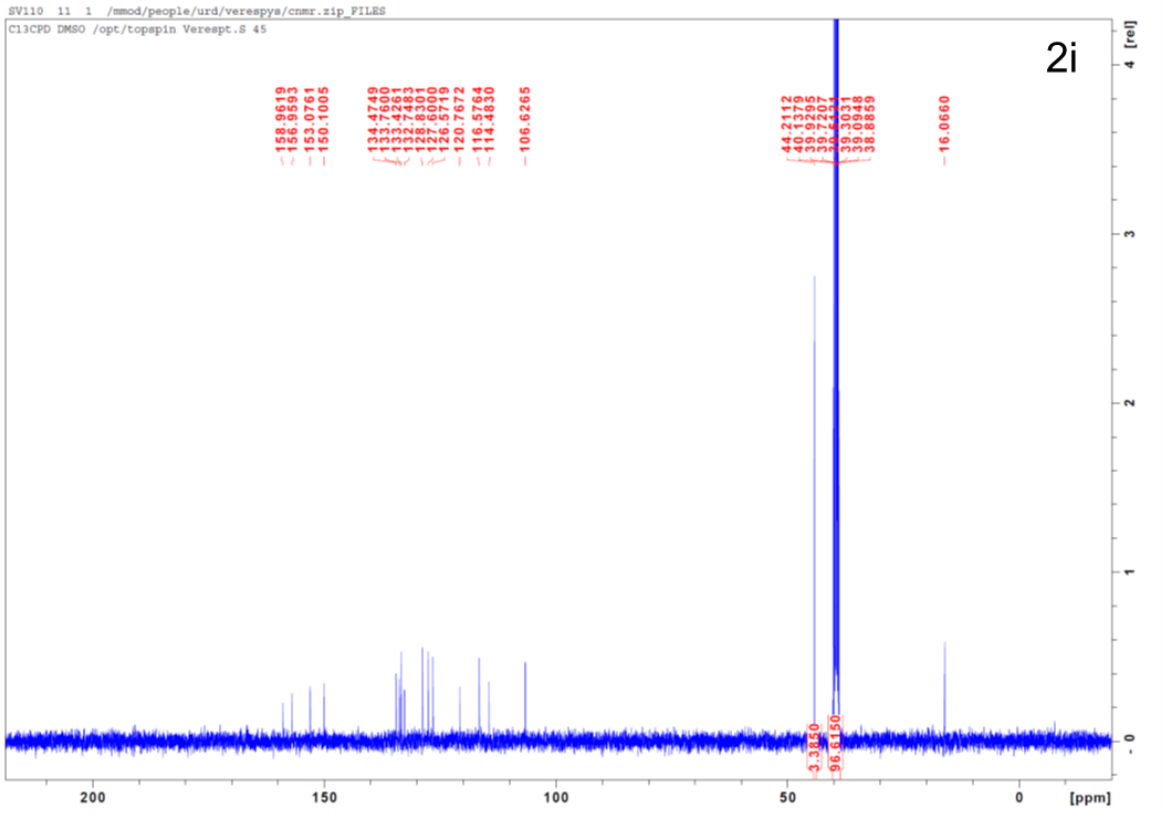


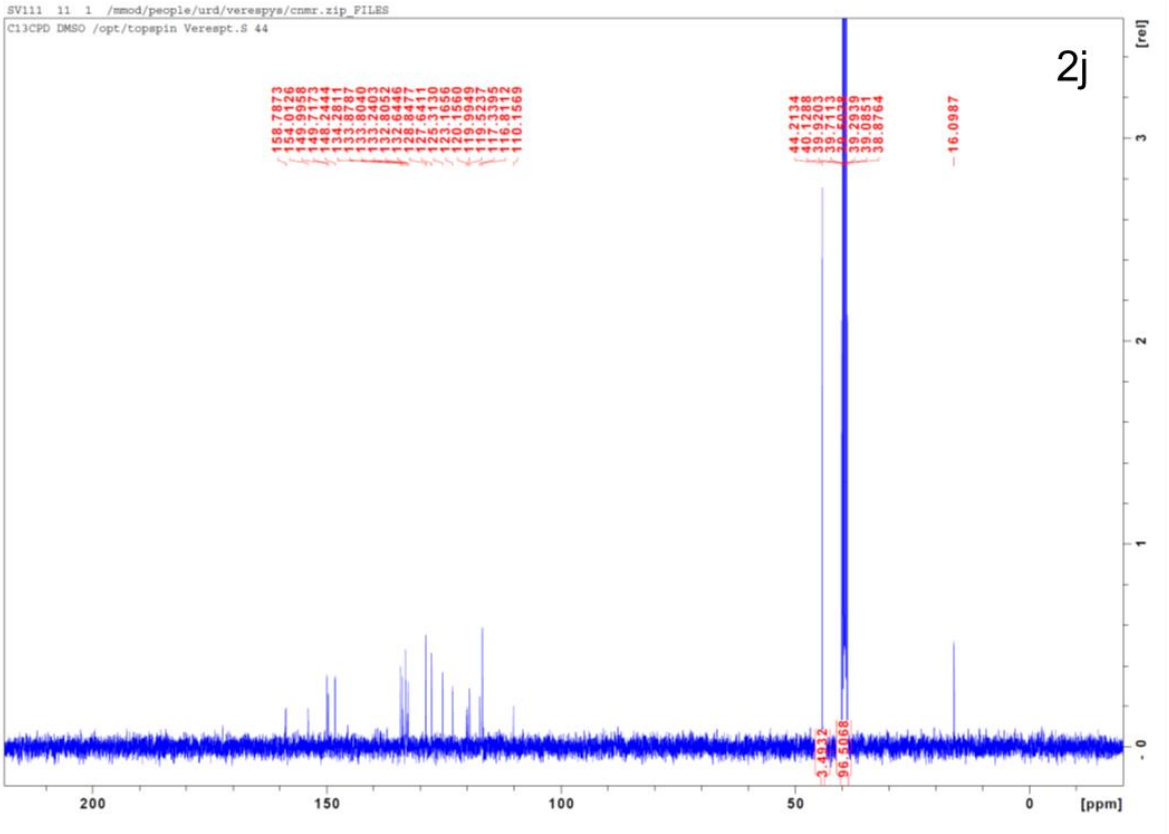


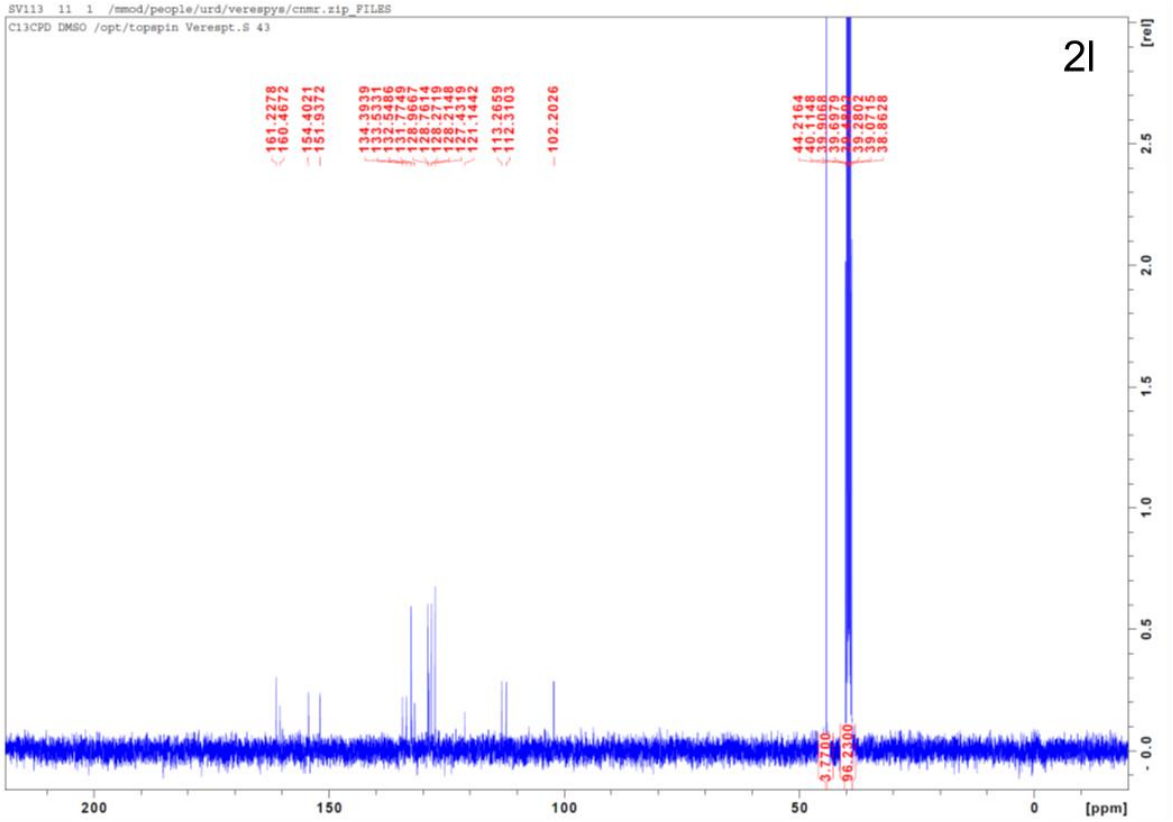


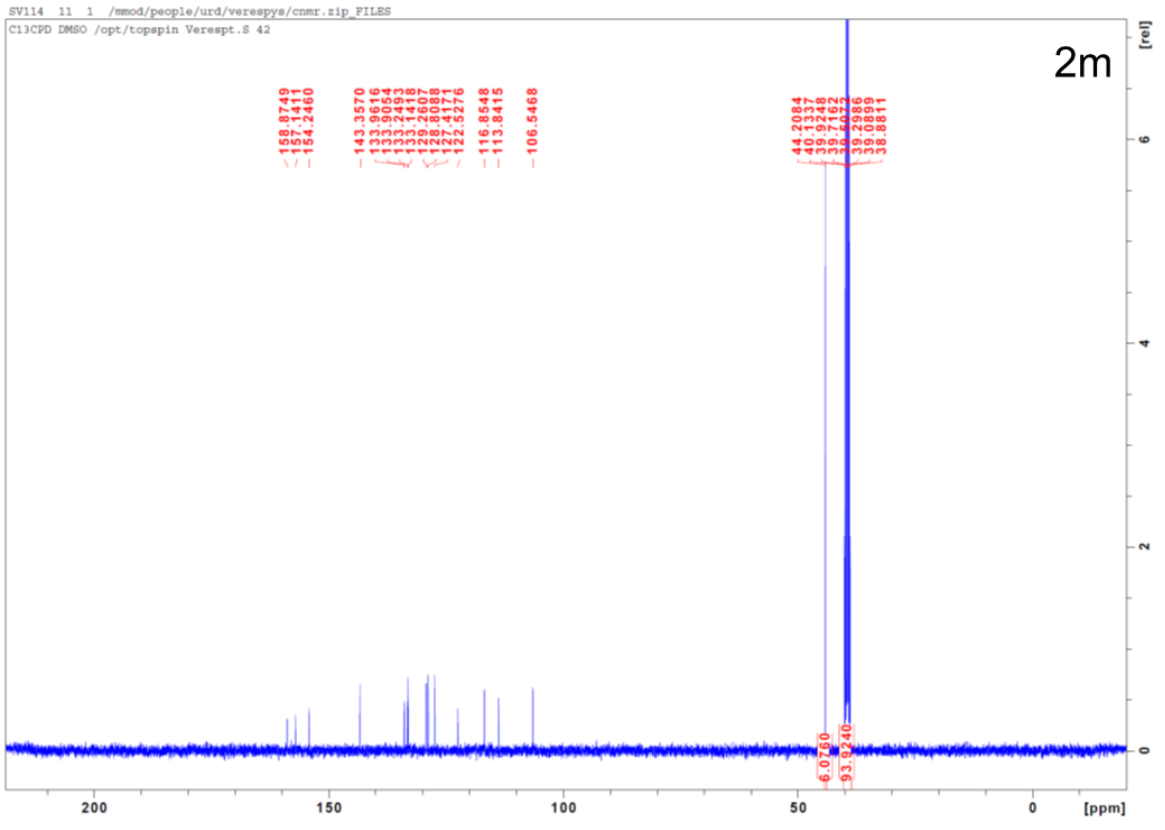


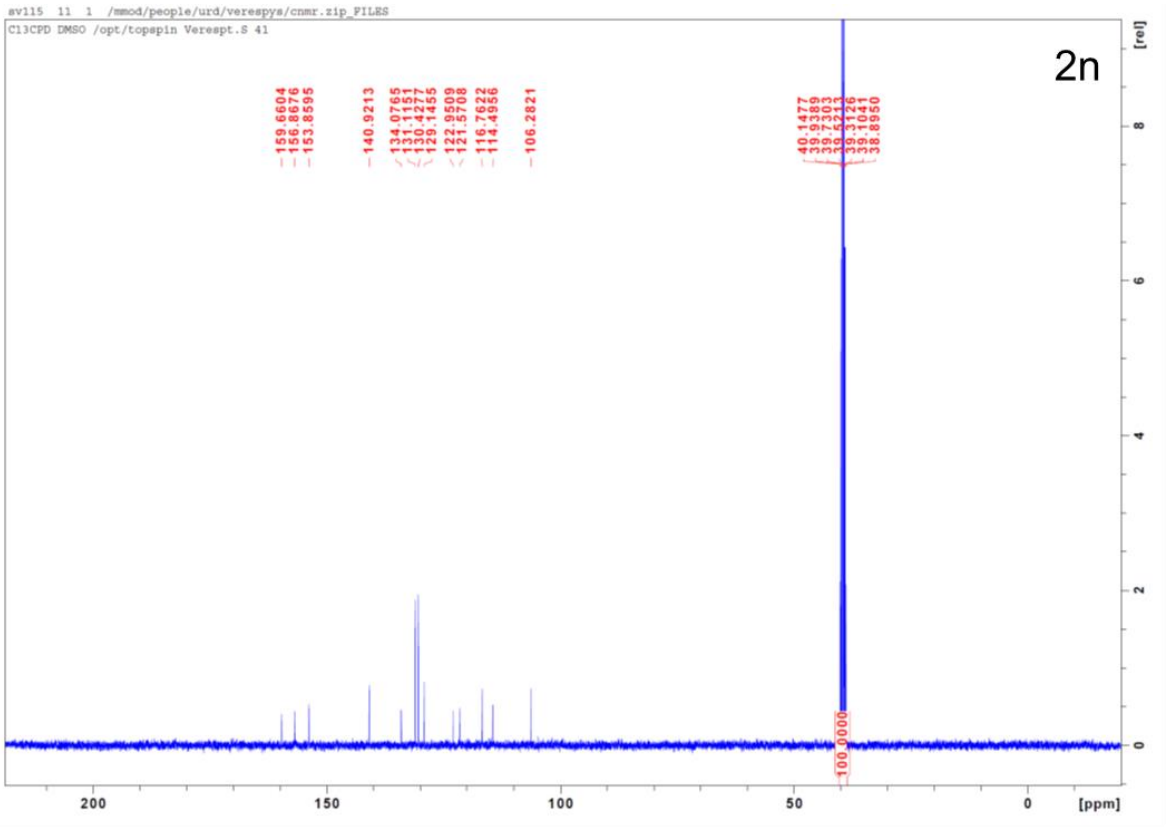


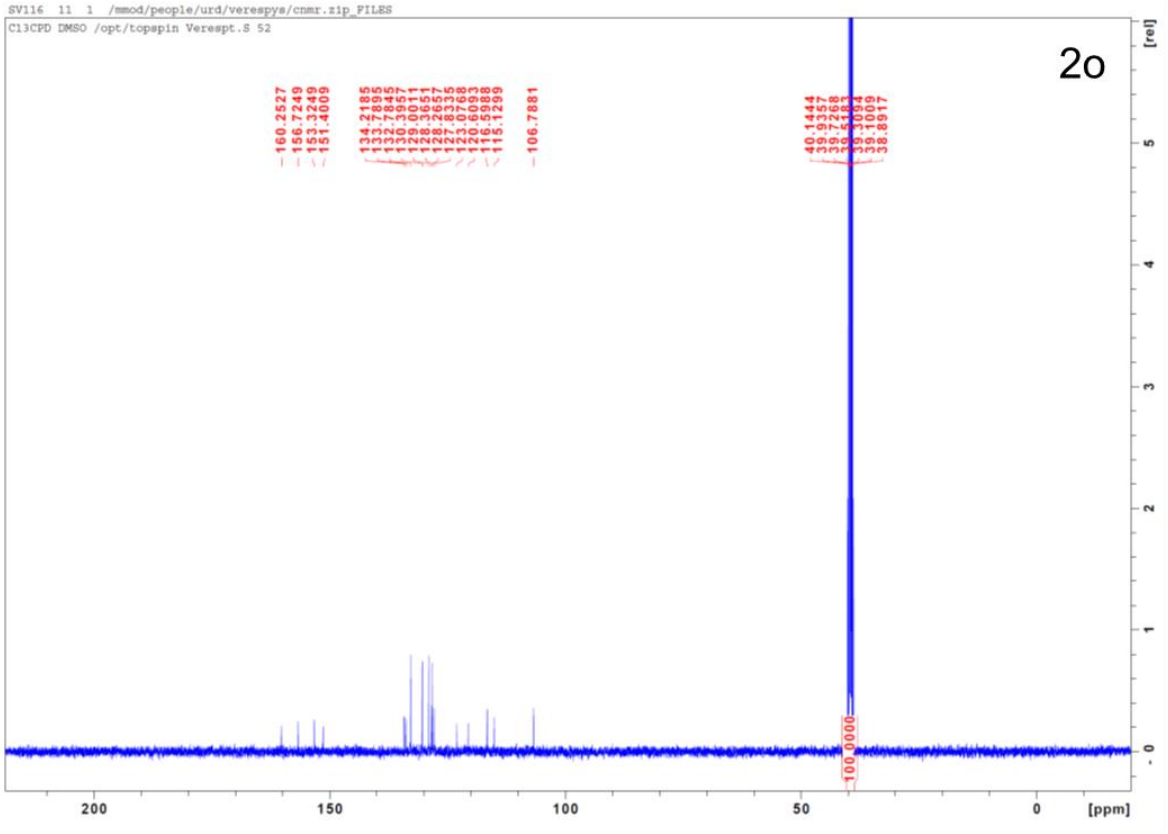


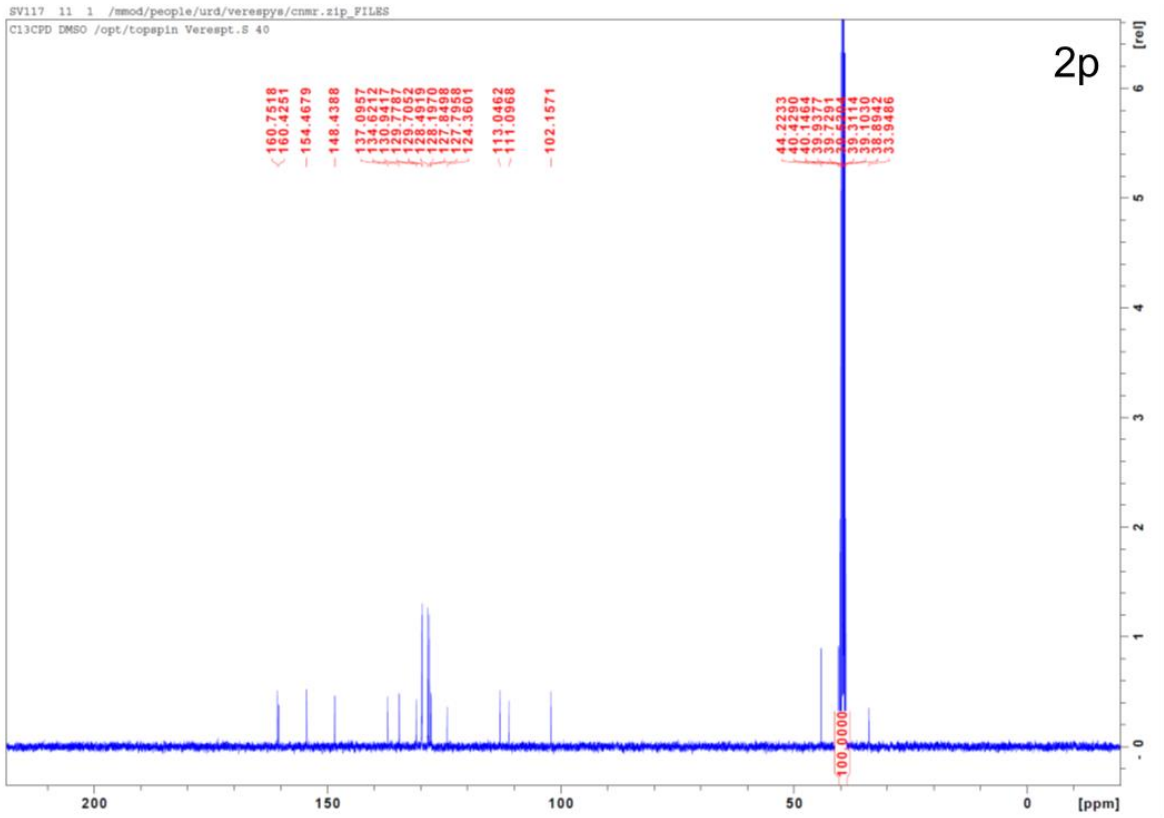


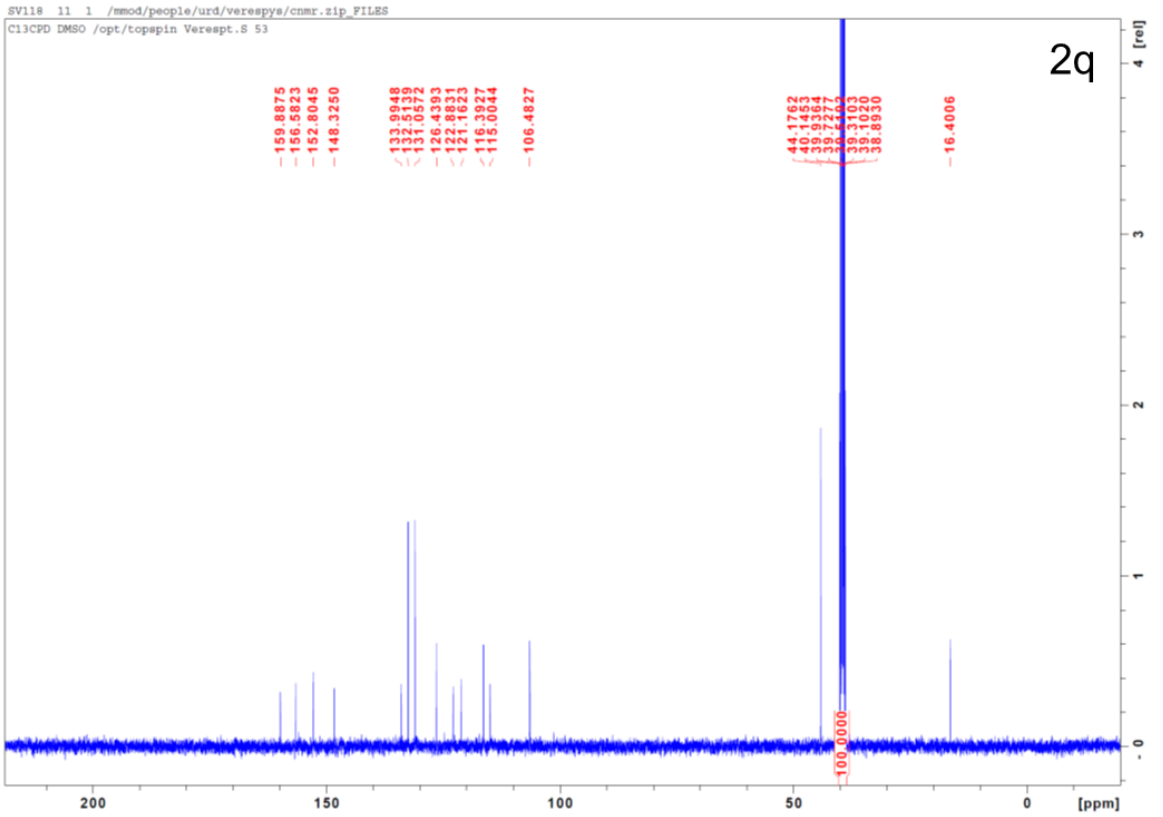


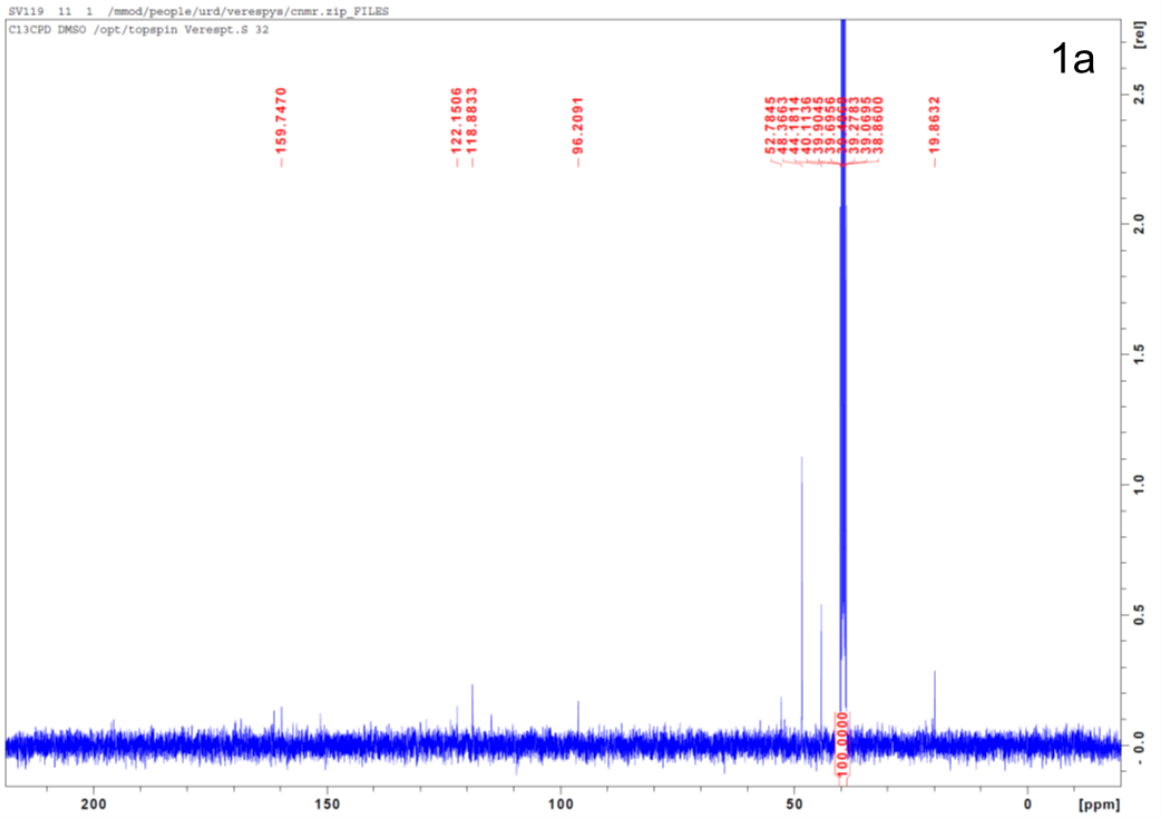


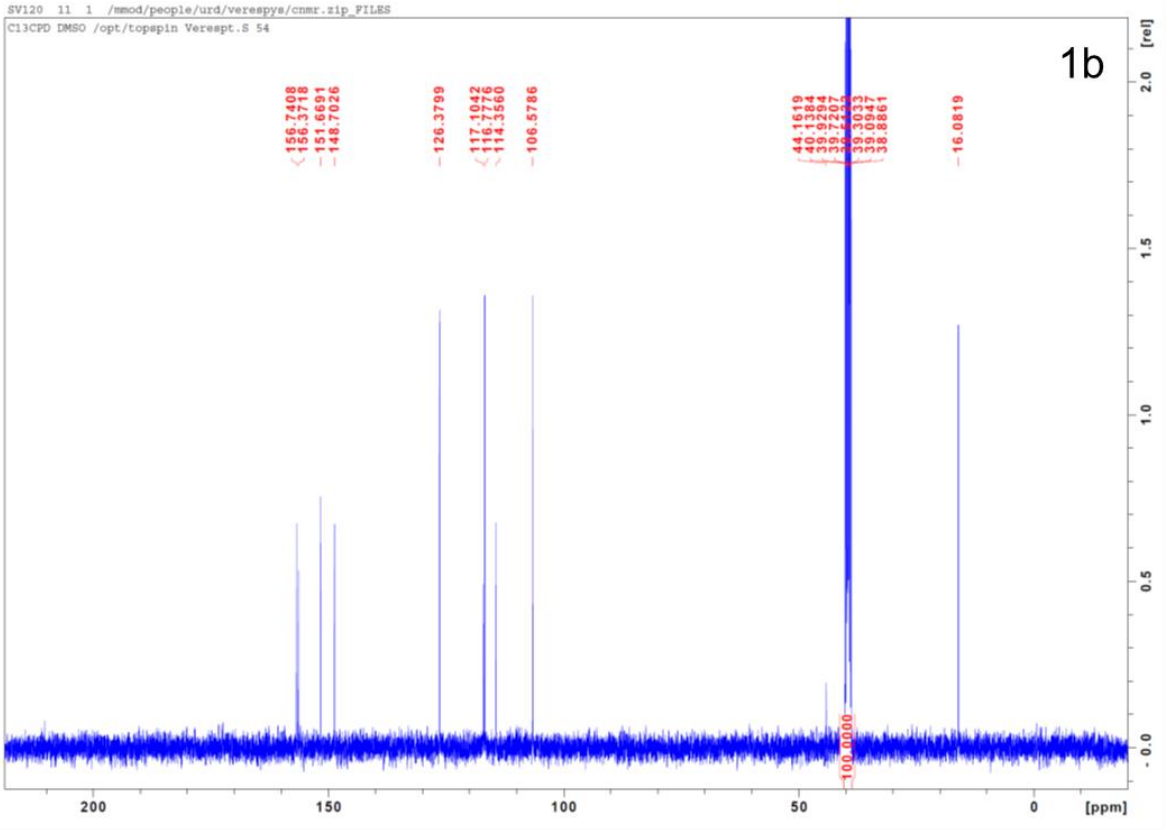


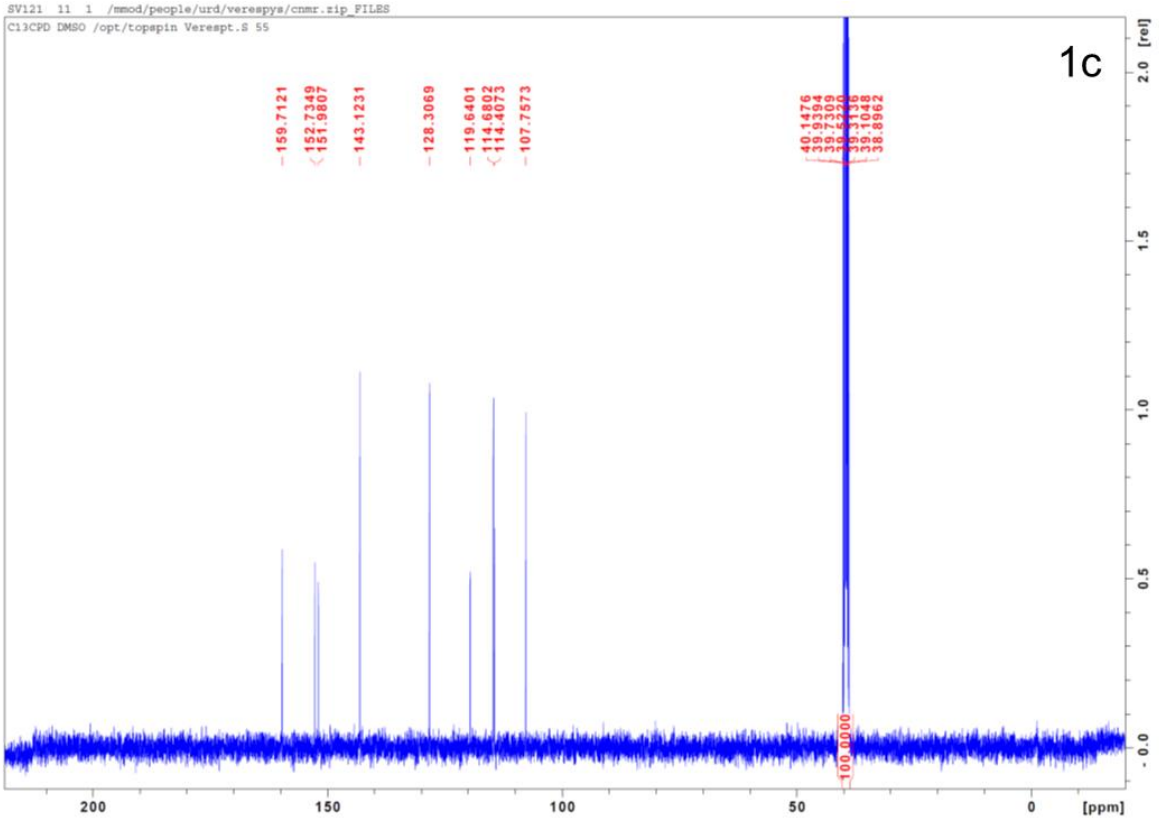


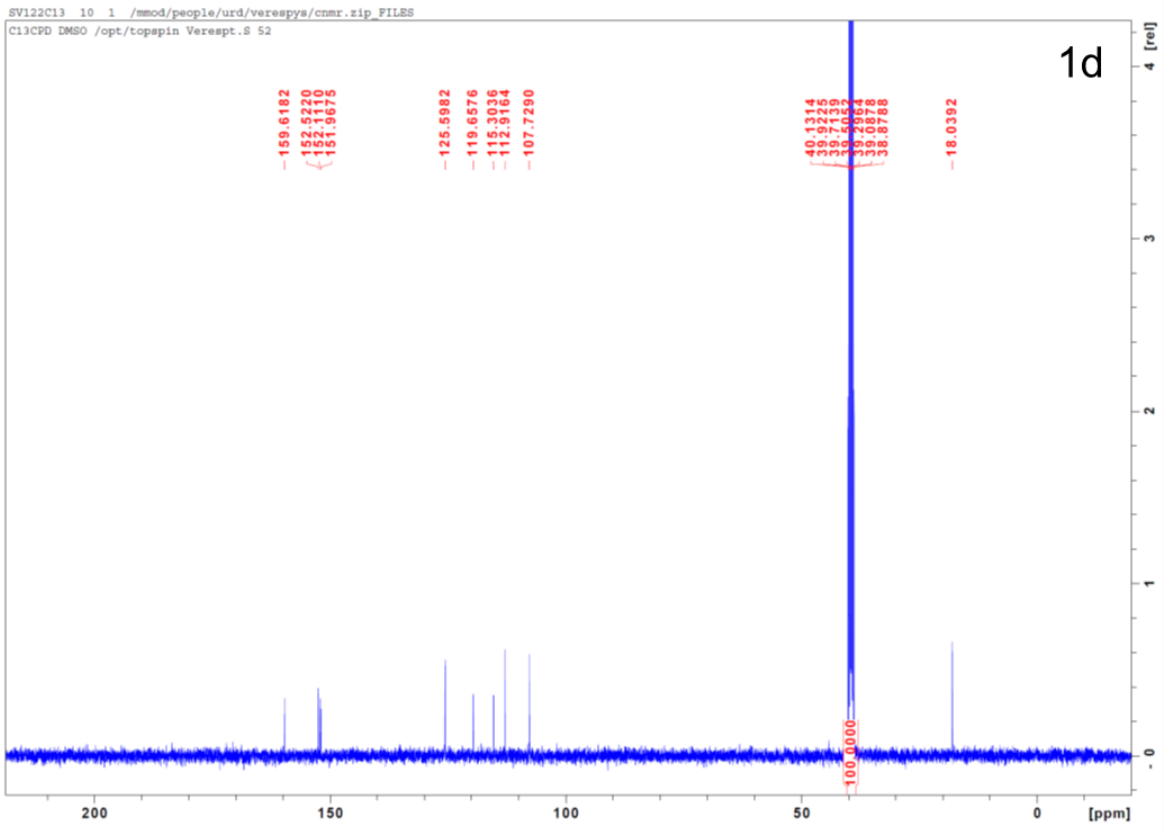


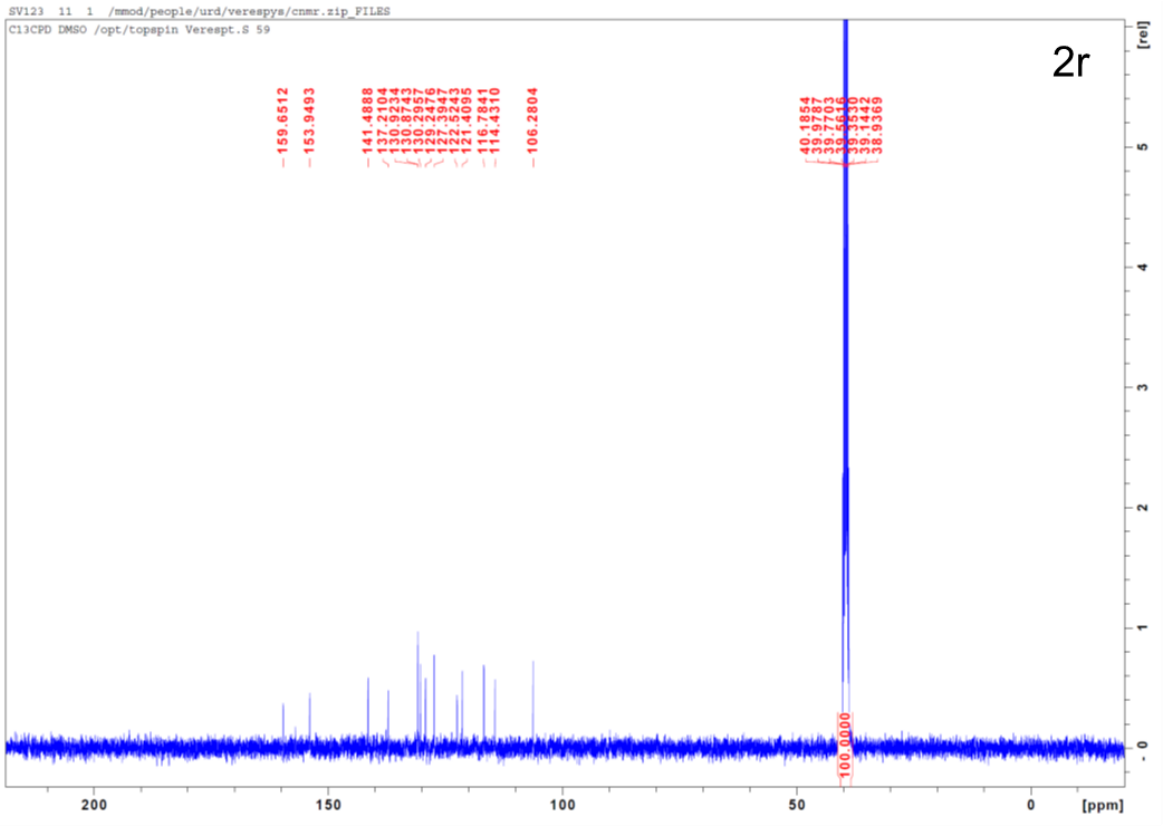


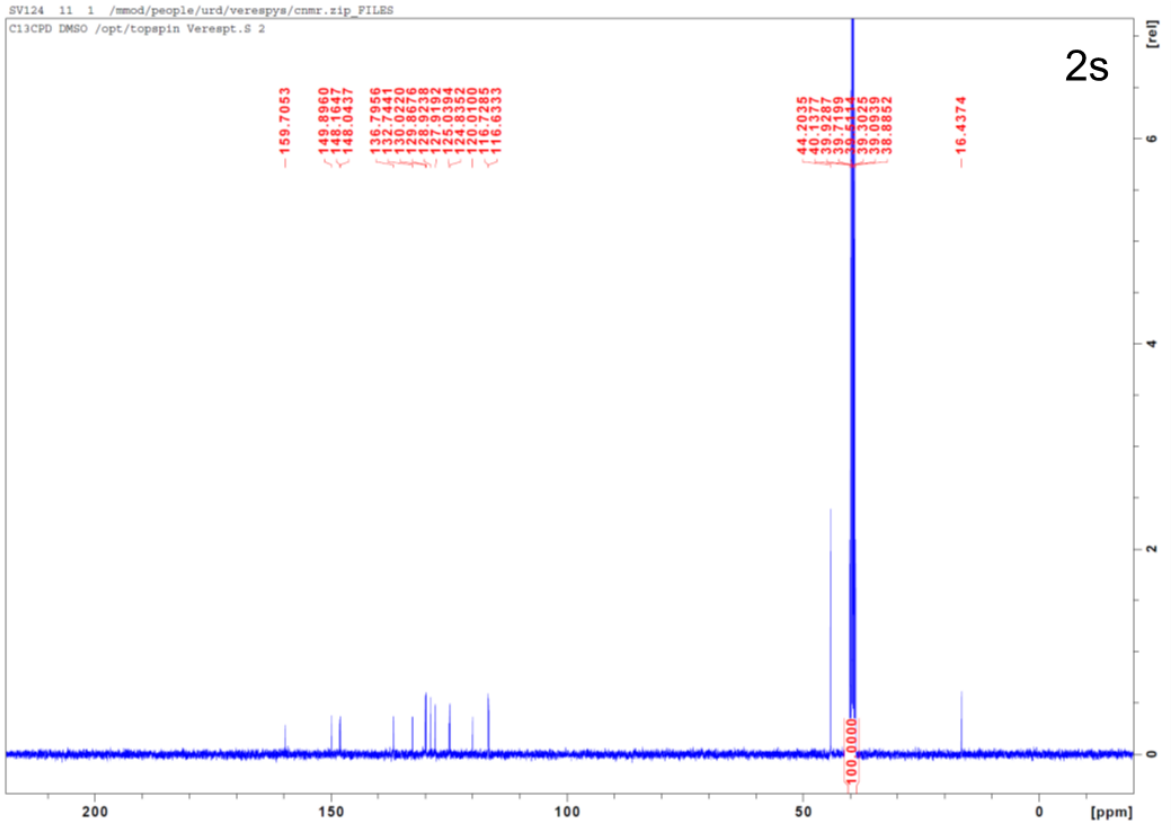


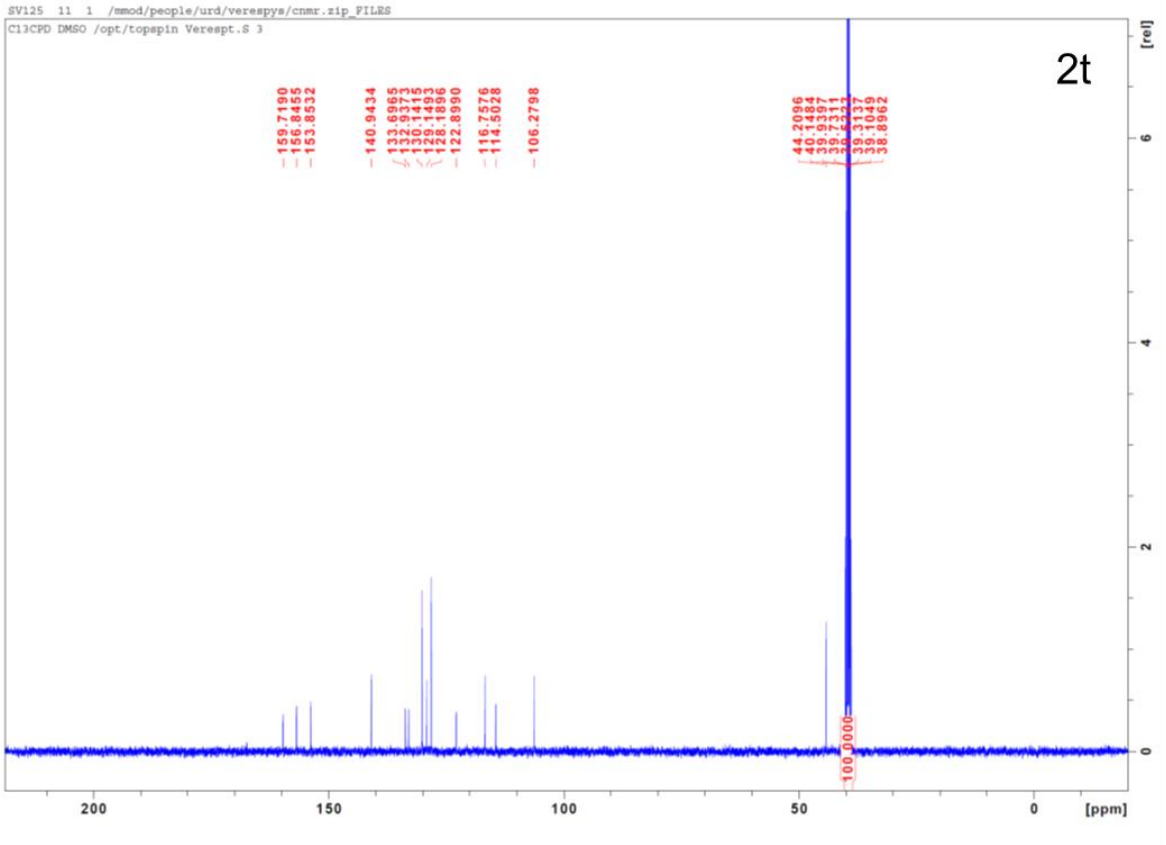


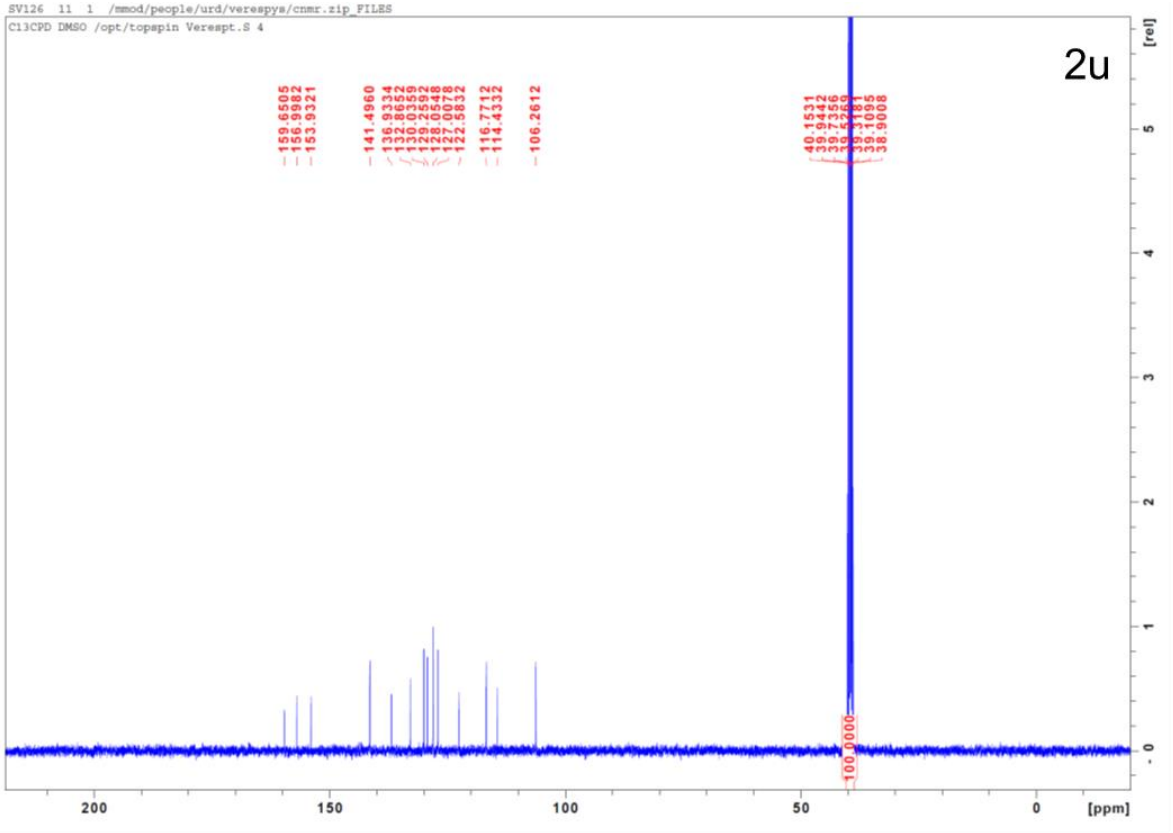


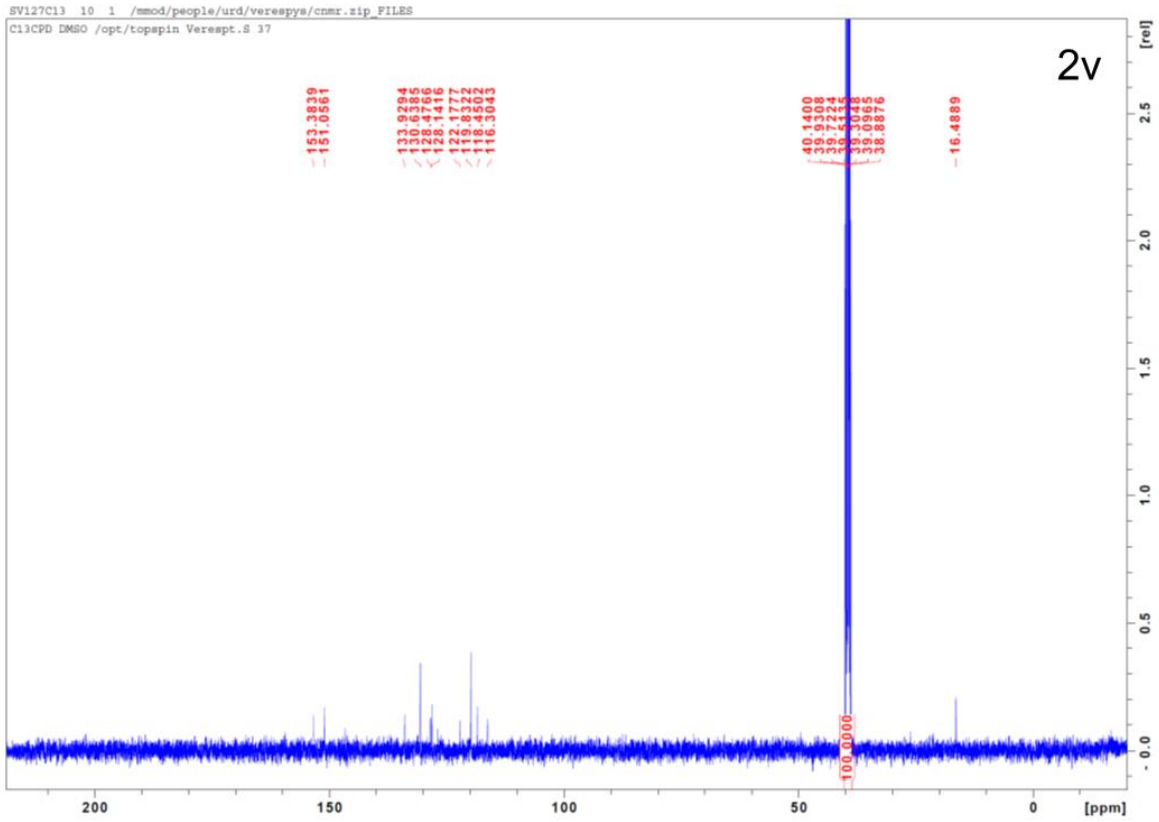


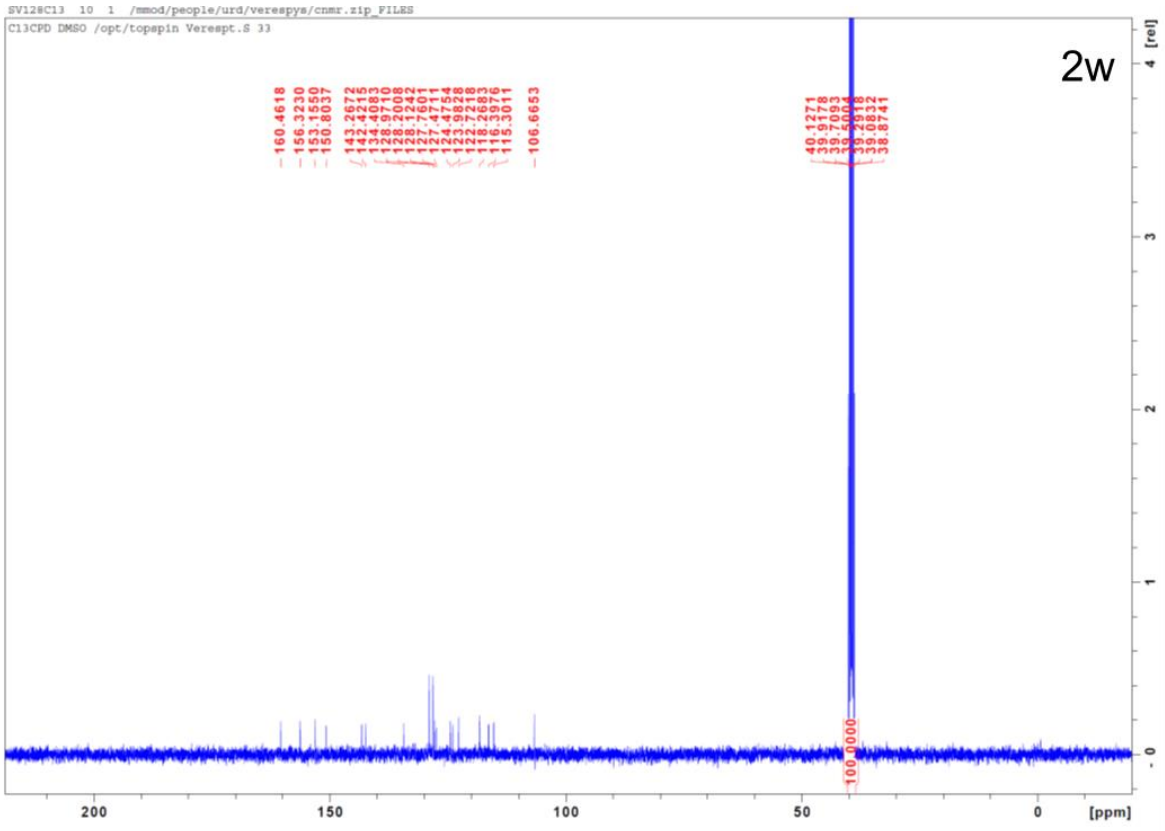


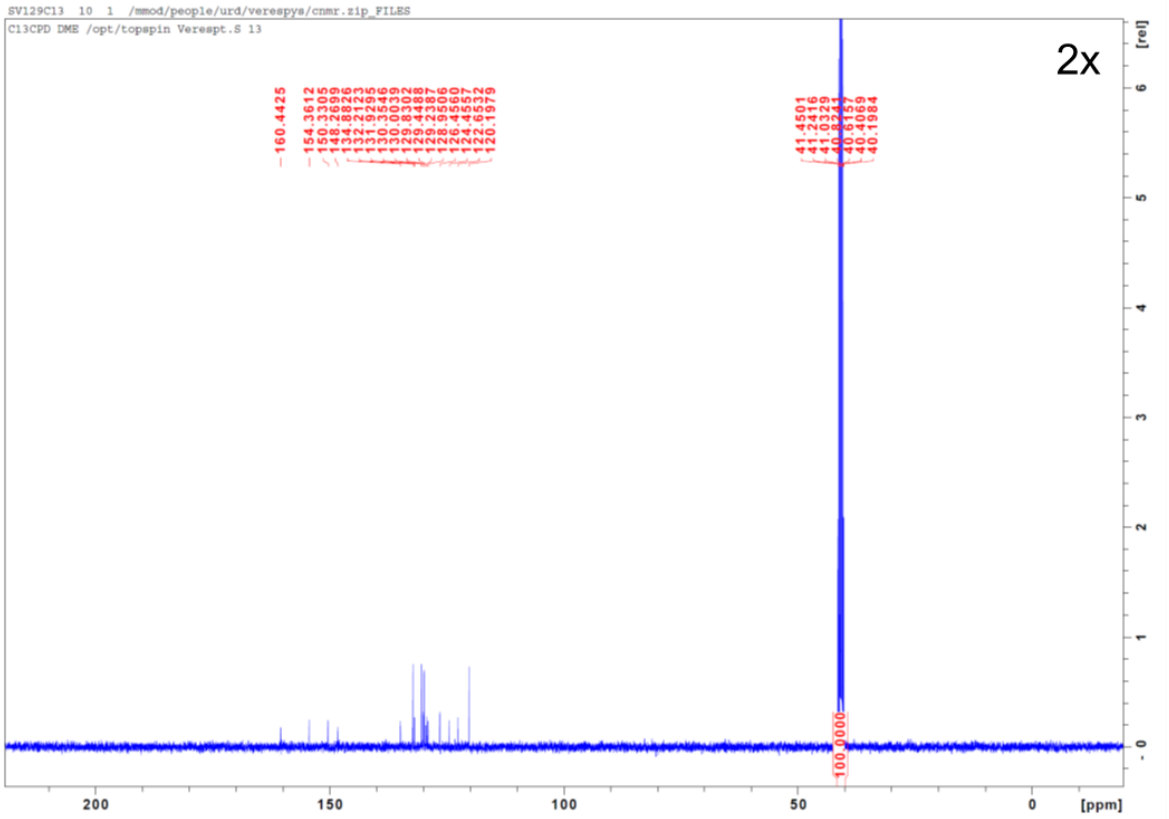


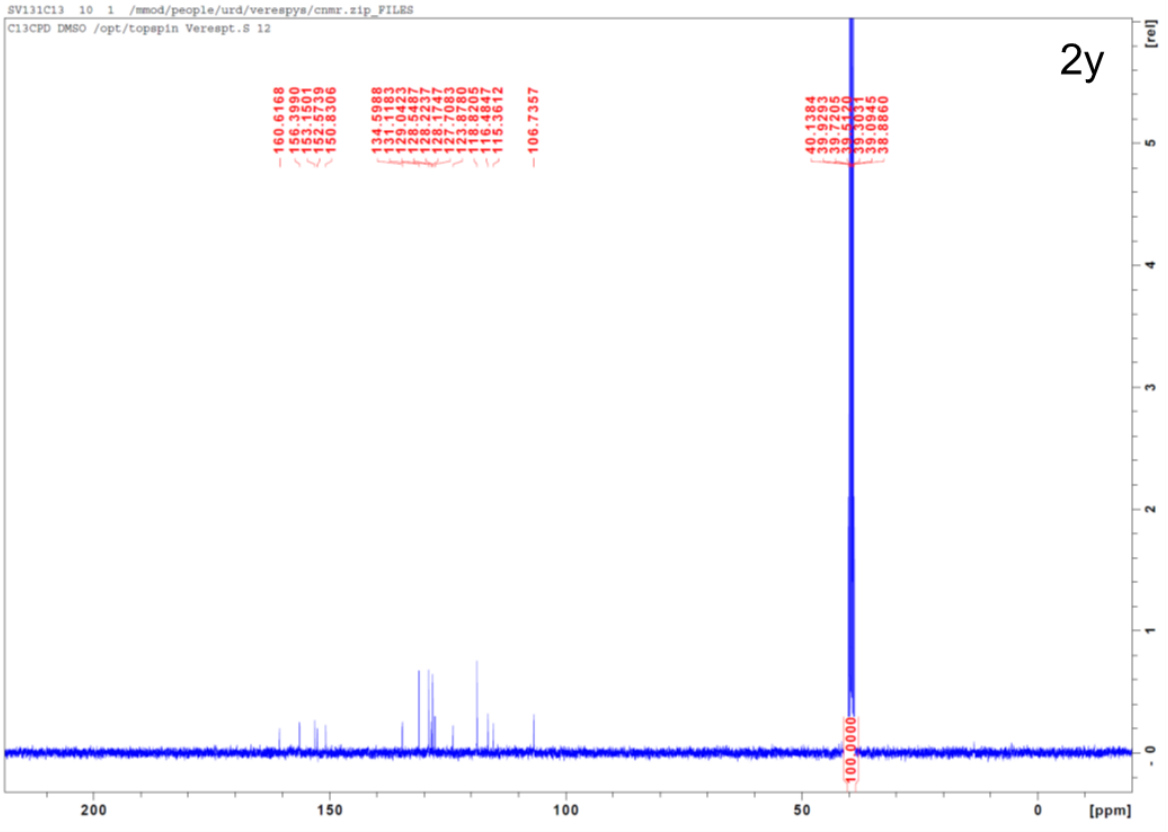


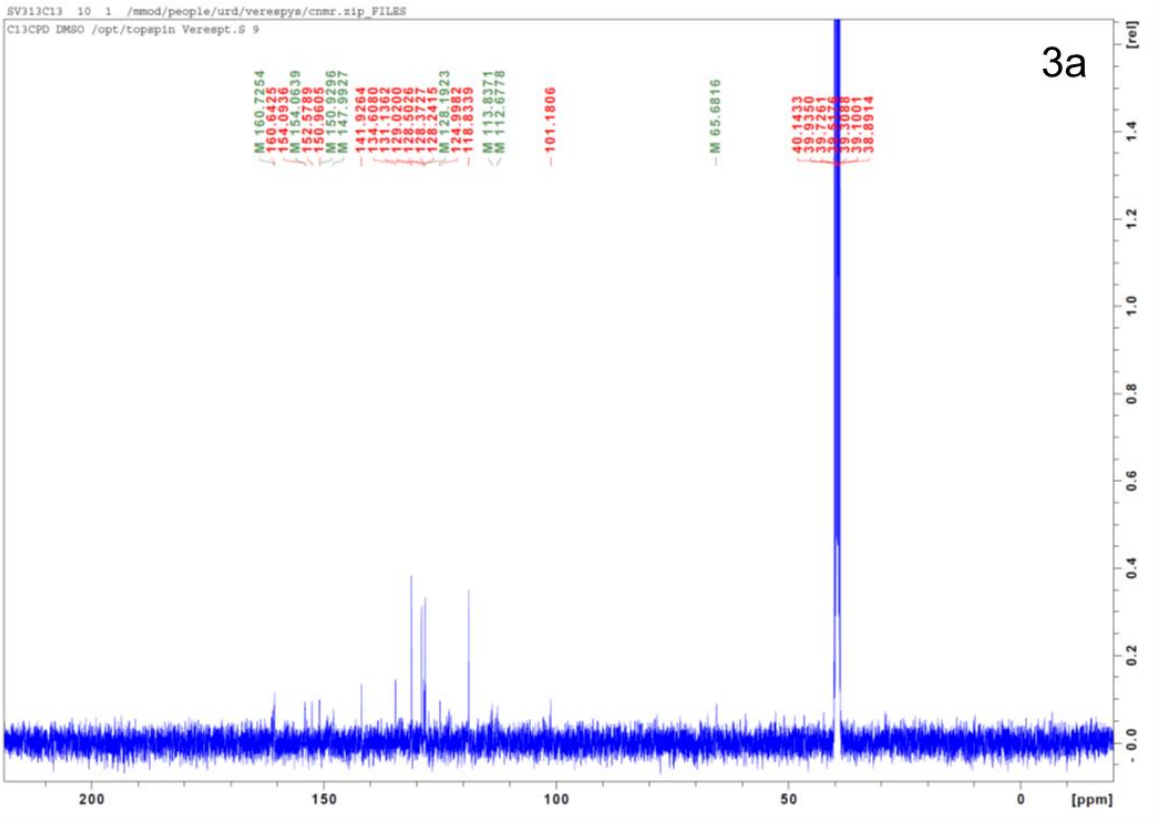


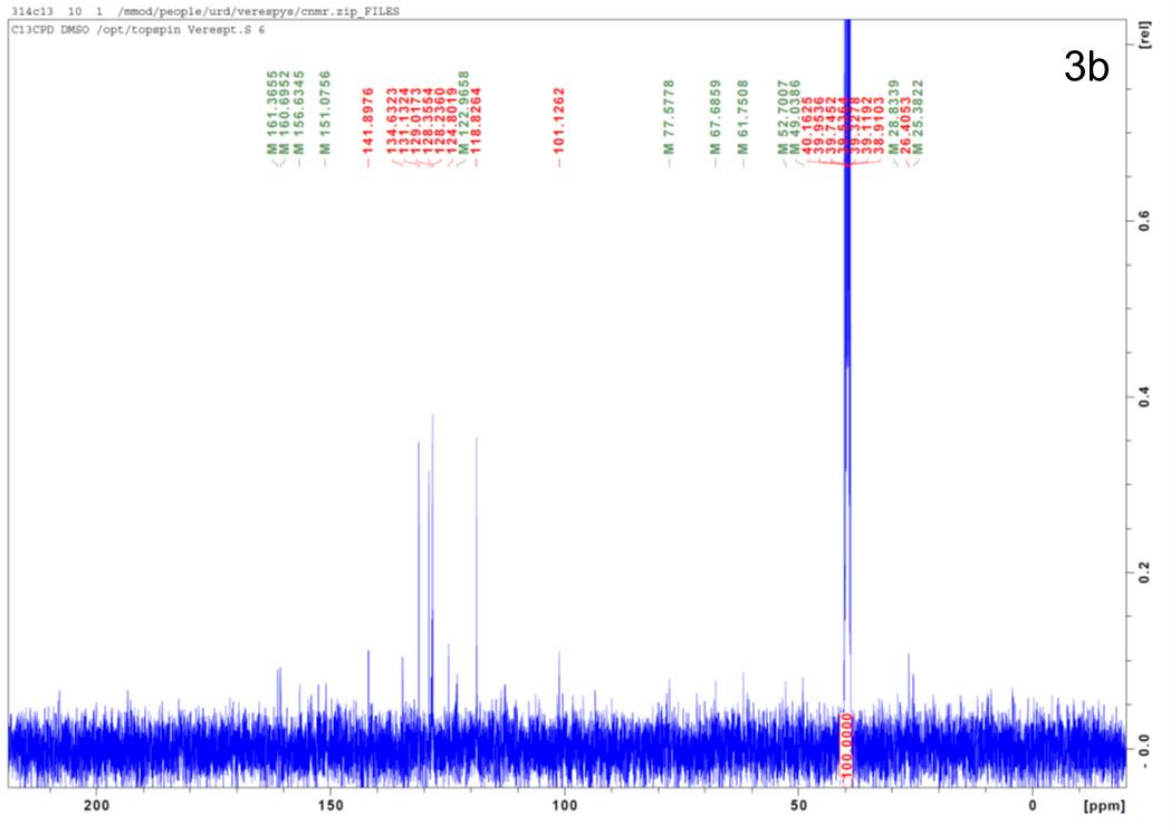


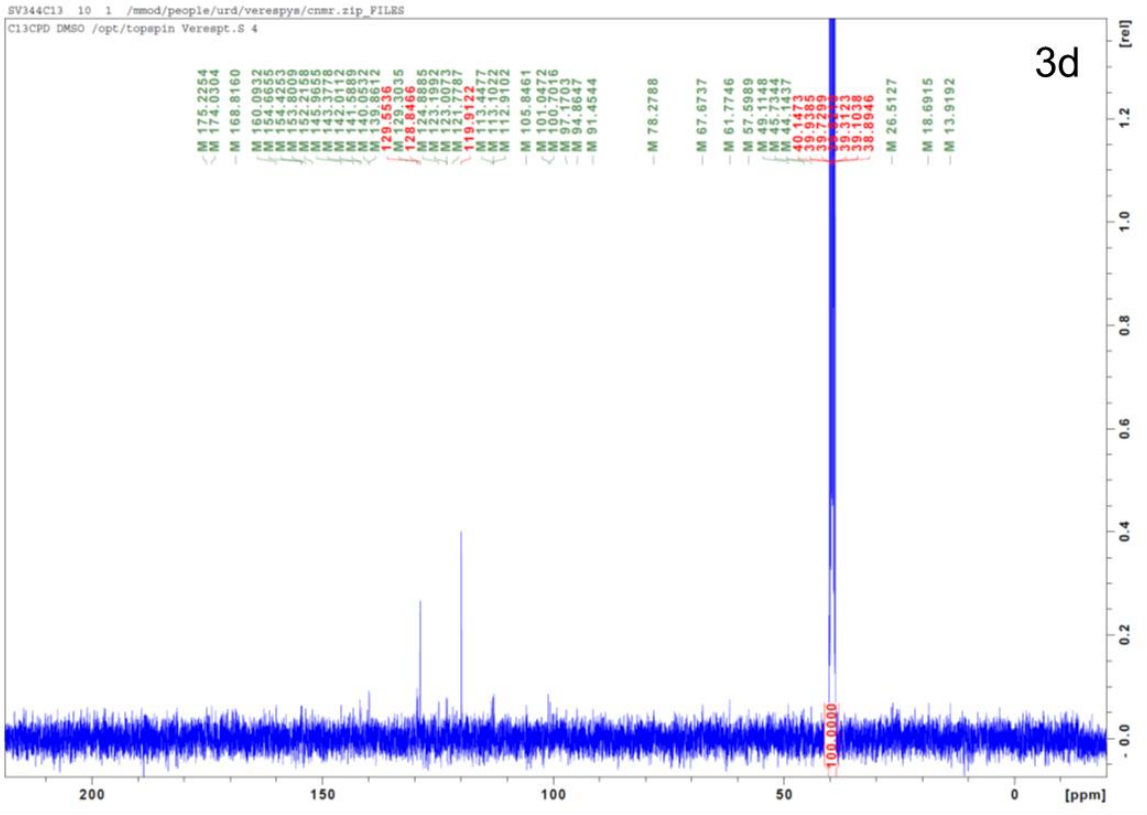


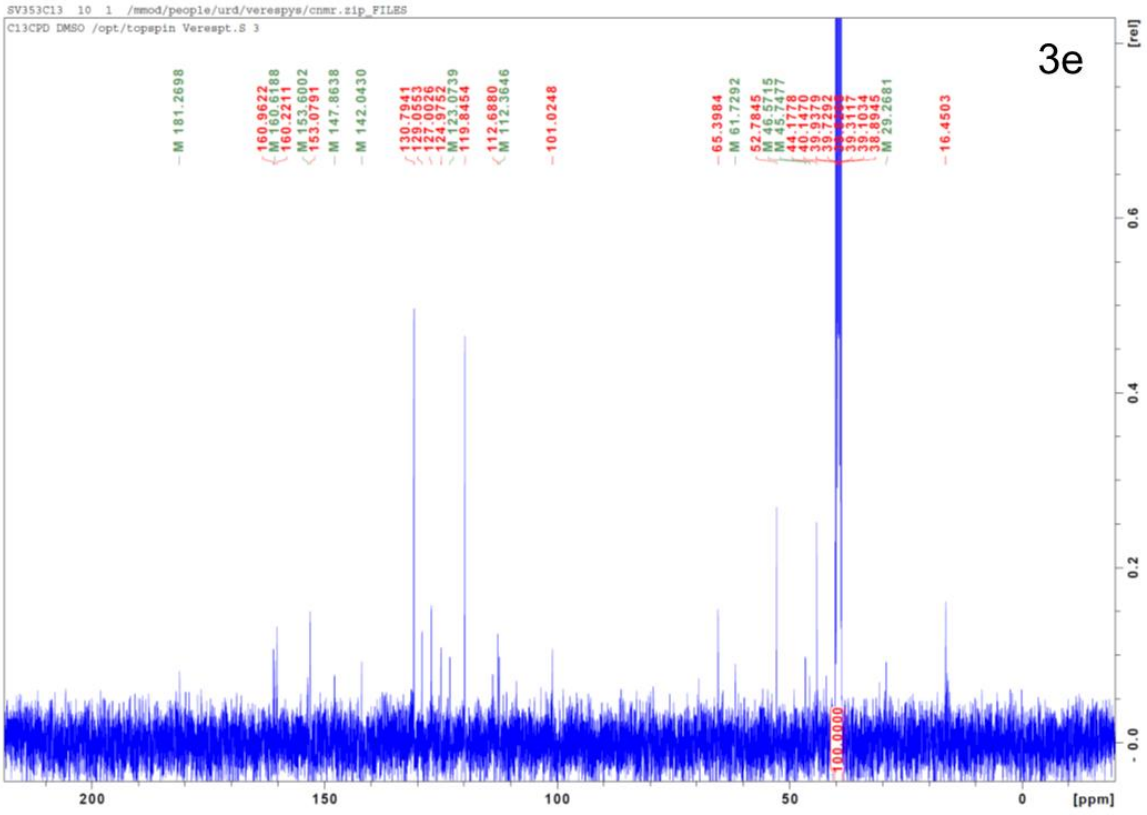


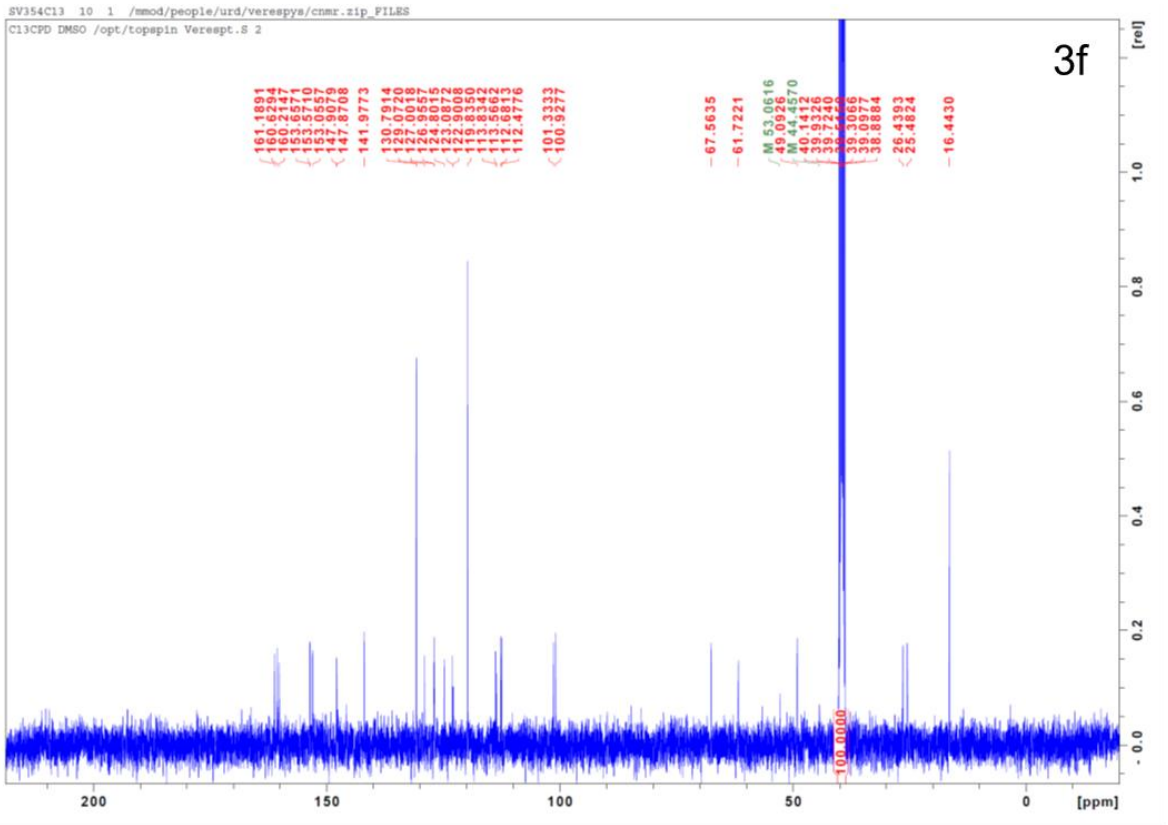


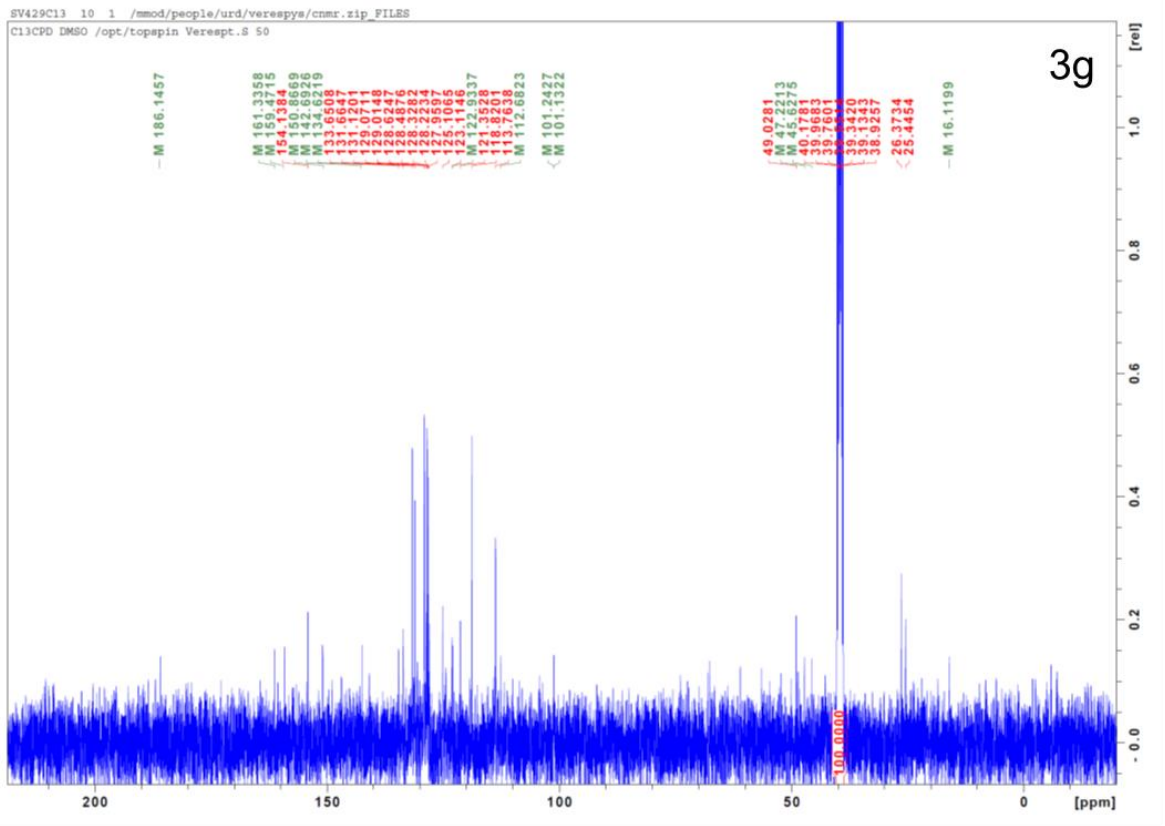


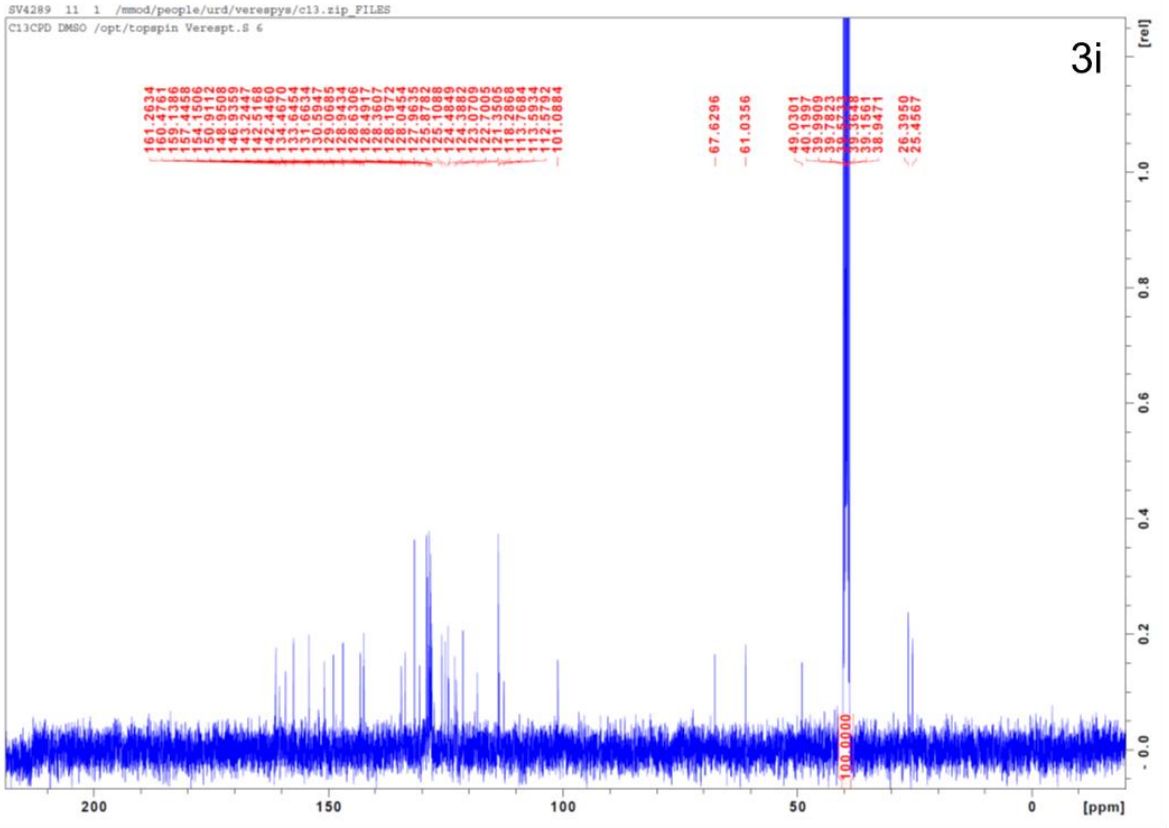












VITA

Stephen S. Verespy III was born in Olyphant, Pennsylvania on April 12, 1988; the first son to Stephen S Verespy Jr. and Judy Verespy. Stephen completed his Bachelor of Science degree in Biology, with minors in Chemistry and Mathematics at Keystone College in Spring 2010. The following Fall, he enrolled in the Master of Arts program in Biochemistry at the University of Scranton. Upon completion of his coursework, in 2012 he began his Doctoral work at Virginia Commonwealth University; completing his Ph. D. requirements in the Chemical Biology program in August 2016. During his tenure at VCU, Stephen accumulated authorship of a book chapter, a journal article, and an invention disclosure. Stephen has presented work at multiple conferences including the Drug Discovery and Therapy World Congress and the American Society for Biochemistry and Molecular Biology. He also was awarded First Place at the Institute for Structural Biology, Drug Discovery and Development's Graduate Student and Postdoctoral Fellow Research Poster Competition.

Publications:

Verespy III, S., Mehta, A. Y., Afosah, D., Al-Horani, R. A., Desai, U. R. (2016) Allosteric partial inhibition of monomeric proteases. Sulfated coumarins induce regulation, not just inhibition of thrombin. *Sci. Rep.* 6, 24043.

Al-Horani, R. A., Karuturi, R., Verespy III, S., Desai, U. R. (2015) Synthesis of glycosaminoglycan mimetics through sulfation of polyphenols. *Methods Mol. Biol.* 1229, 49-67.