DEVELOPMENT OF CLICK NUCLEIC ACIDS FOR BIOSENSING

by

Derek J. Bischoff

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Bachelor of Chemical Engineering

Spring 2019

DEVELOPMENT OF CLICK NUCLEIC ACIDS FOR BIOSENSING

by

Derek J. Bischoff

Approved: _____
Christopher J. Kloxin, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____
Millicent O. Sullivan, Ph.D.
Committee member from the Department of Chemical & Biomolecular Engineering

Approved: _____
Carlton R. Cooper, Ph.D.
Committee member from the Board of Senior Thesis Readers

Approved: _____
Michael Chajes, Ph.D.
Chair of the University Committee on Student and Faculty Honors

ACKNOWLEDGMENTS

First, I would like to thank Dr. Christopher Kloxin for giving me the opportunity to join his research group at the University of Delaware prior to starting my first year as an undergraduate. With my limited background in laboratory research, Dr. Kloxin paired me with a graduate student just beginning his own doctoral research. As I gained experience and took core courses like organic chemistry, I received more trust from both Dr. Kloxin and the research group which gave me the freedom to begin my own work.

I would like to thank Bryan Sutherland, the Ph.D. student that has been with me working on this project since its conception many years ago. The vast majority of laboratory skills I now possess are a direct result of Bryan's excellent teaching. Both Dr. Kloxin and Bryan have always been very supportive and willing to give me advice, whether related to my thesis, my courses, or my decision to attend graduate school.

I would also like to recognize the other members of the Kloxin research group that have helped me learn laboratory skills. I would like to specifically thank Bassil El-Zaatari who mentored me on a different project during the summer between my sophomore and junior years. This experience working with polymer networks broadened my experience to include new chemistries, new laboratory techniques, and polymer characterization.

In addition, I would like to thank the Undergraduate Research Program at UD for providing me with the opportunity to participate in academic research. Specific

opportunities that allowed me to continue my academic pursuit include the Summer Fellows, Summer Scholars, and the senior thesis programs. These programs gave me the chance to network with other students with similar interests and provided financial support to make it all possible.

Finally, I would like to thank Dr. Sullivan and Dr. Cooper for their roles as the second and third readers of my thesis, respectively. I also want to thank my family and friends for their support during my time as an undergraduate.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Herein, a new type of nucleic acid analog referred to as a ‘click nucleic acid’ (CNA) is described. Nucleic acids, such as deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), as well as current analogs, such as peptide nucleic acid (PNA) and locked nucleic acid (LNA), have been used in many lab scale applications that leverage the complementarianism between two sequence specific strands of the polymer. Limitations of the current methods not only prevent widespread use *in vivo* due to enzymatic degradation and hydrolysis, but also are quite costly to scale up and utilize inefficient reactions to make these periodic polymers. The proposed CNA material that can overcome these issues is built using the same four nucleobases that are found in naturally occurring DNA with the naturally occurring phosphate-sugar backbone replaced with a neutral, primarily alkane backbone that could lead to tighter binding of complementary strands. The ‘click’ portion of its name refers to click chemistry, a subset of chemical reactions that are highly efficient and proceed under ambient conditions. This CNA material has the benefits of reusing excess monomer, efficient click chemistry, scale up, and cleaner purification. The synthesis of the four nucleotide mimicking monomers follows a four or five step synthetic pathway involving commonly employed organic reactions such as an alkylation, hydrolysis, reductive amination, and substitution/elimination. Subsequently, a solid phase synthesis protocol is outlined to provide sequence specific polymers using a protecting group strategy similar to typical Fmoc peptide synthesis. Initial results show that solubility is an issue that is preventing polymers larger than three units long from being both synthesized and characterized. In overcoming this challenge, a variation on

the four nucleobase mimicking monomers has begun development that adds an extra carboxylic acid functionality to the backbone that should help with solubility.

Chapter 1

INTRODUCTION

1.1 Properties of Deoxyribonucleic Acid (DNA)

DNA is the molecule used by living organisms to direct life and is appropriately positioned as the first step of the central dogma of biology. Its name was chosen for its discovery in the nucleus and the presence of phosphate groups as it relates to phosphoric acid. It is a long biomacromolecule, for example in humans reaching three billion repeat units, that consists of three main components to form a polymer: a sugar (deoxyribose), a phosphate, and a nucleobase. A nucleobase can be one of four molecules covalently attached to the sugar of the backbone. The four nucleobases are adenine, cytosine, guanine, and thymine which are further broken down into two categories. Adenine and guanine are purines, consisting of a six-membered and a five-membered nitrogen containing ring, whereas cytosine and thymine are pyrimidines, a six-membered nitrogen containing ring. An important feature of the nucleobases is that they can form hydrogen bonds with other specific nucleobases: adenine-guanine and cytosine-thymine. These hydrogen bonding pairs, often referred to as Watson-Crick base pairs, are highly specific when binding with their partner and are shown in Figure 1. Due to the negative charge on the phosphate group, DNA is polyanionic and is water soluble.

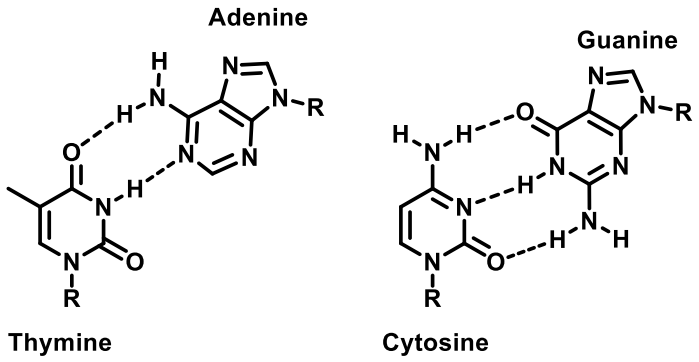


Figure 1: Hydrogen bonding (dashed lines) among preferred nucleobase pairs

DNA is found in cells in a condensed form creating supramolecular structures like the double stranded helix, the G-quadruplex, or the intercalated motif (i-motif). The structure of the DNA molecules in its various forms is shown in Figure 2. In biological systems, the DNA can exist in both single stranded (ss) forms and double stranded (ds) forms depending on which stage of the cell cycle is currently taking place. Another feature of DNA and RNA is their directionality, denoted by 3' and 5', where the numbers refer to the carbon of the sugar and the prime (') notation refers to the sugar and not the nucleobase. The carbon that connects to the nucleobase is number one, the third carbon is connected to a hydroxyl group pointing in the 3' direction, and the fifth carbon is connected to the phosphate pointing in the 5' direction. This directionality is important in determining which direction the molecule is replicated and the sequence of nucleobases. These naturally occurring DNA properties have allowed engineers to construct various useful structures.

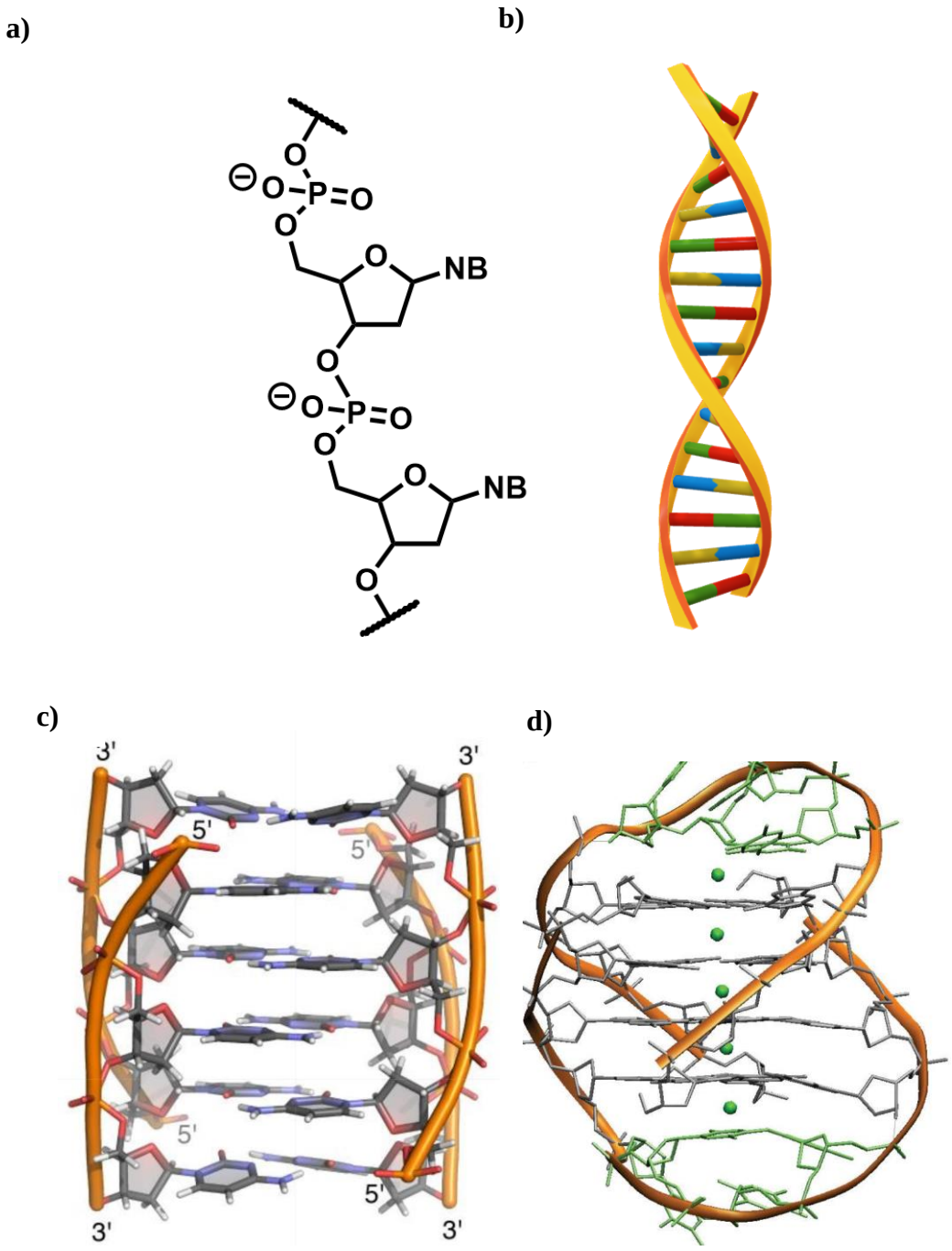


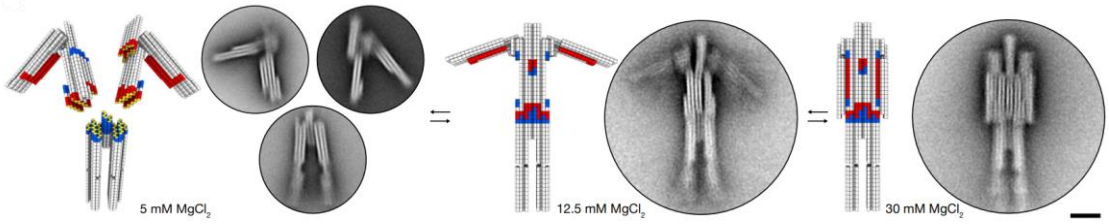
Figure 2: **a)** chemical structure of ss DNA; NB = nucleobase; **b)** a cartoon view of dsDNA in the double helix conformation; **c)** a cartoon view of the i-motif adopted from ref. 17; **d)** a cartoon view of the G-quadruplex adopted from ref. 18

1.2 Leveraging the Natural Properties of DNA to Make Materials

1.2.1 DNA Origami

DNA origami is a technique used to construct nanoscale architectures by combining hundreds of small DNA molecules called staple strands with longer single stranded DNA molecules called scaffold strands to form a target shape. DNA was chosen to be the building blocks due to its predictable and specific base pairing, its well understood dimensions including the diameter, twist, and rise per turn, and finally its ease of synthesis allowing for many strands with unique sequences to be made quickly.[1] The ability to engineer these higher order structures relies on the ability of DNA to form branch junctions that can be further joined together in such a way as to limit sequence symmetry, as found in naturally occurring dsDNA. This work was largely pioneered by Nadrian Seeman and resulted in junctions of up to 12 arms, showing that DNA junction molecules can be highly branched. These multi-arm junctions are quite flexible and have conformations that are difficult to predict, making them unsuitable as building blocks from which higher order structures are generated. To overcome this problem, DNA double crossover molecules were created by combining two 4-way junctions into a single motif resulting in a rigid DNA molecule held together by two crossovers. This motif led to the formation of the first higher order DNA lattice structure in two dimensions in 1998 by Nadrian Seeman and coworkers.[2] In 2006, Paul Rothumend was the first to coin the DNA origami technique and construct many single layered structures such as a five pointed star and a disk with three holes, shown in Figure 3.[3] As DNA origami technology has progressed, more advanced structures dependent upon other factors such as

temperature and divalent ion concentrations to induce motion in the structures. As seen in Figure 4



, Dietz and coworkers have shown that changes in Mg^{2+} concentration can cause a heterotrimeric nanorobot to switch between three different states: disassembled, assembled with arms up, and assembled with arms down.

Figure 3: Atomic force microscopy images of a five pointed star (left) and a disk with three holes (right) constructed from DNA origami; both images are 165nm by 165nm; adapted from ref. 3

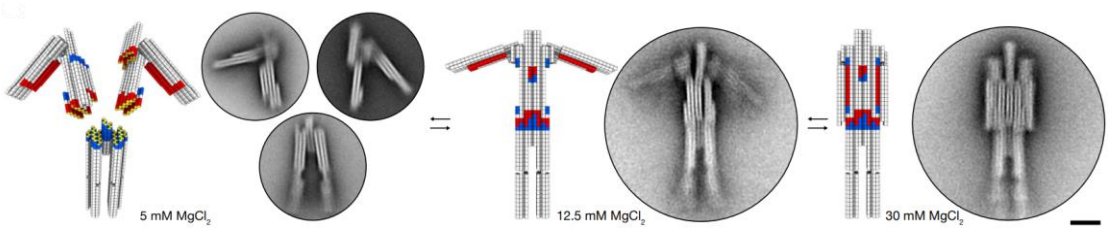


Figure 4: Average negative-stain TEM micrographs of disassembled, assembled with arms up, and assembled with arms down conformations; scale bar is 25nm; adapted from ref. 4

1.2.2 Spherical Nucleic Acids (SNAs)

In 1996, Mirkin et al. described a method in which small gold nanoparticles functionalized with DNA sequences can self-assemble.[16] Two non-complementary DNA sequences that are thiol functionalized attach to gold nanoparticles. When the two types of nanoparticles are introduced into the same solution with a DNA duplex with sticky ends that are complementary to those sequences on the nanoparticles, they self-assemble into aggregates (Figure 5). Since this aggregation is caused by secondary interactions, it can be easily reversed with heat to force denaturation. These aggregates can have their optical, electronic, and structural properties adjusted by using the DNA nucleobase interactions via specific sequences and the length of the oligomer. The self-assembly is further controlled by the solvent, temperature, and electrolyte concentration. As compared to other DNA assemblies (eg. origami), these aggregates are easily characterized by techniques such as atomic force microscopy (AFM), transmission electron microscopy (TEM), and other spectroscopic techniques conventionally used with DNA.

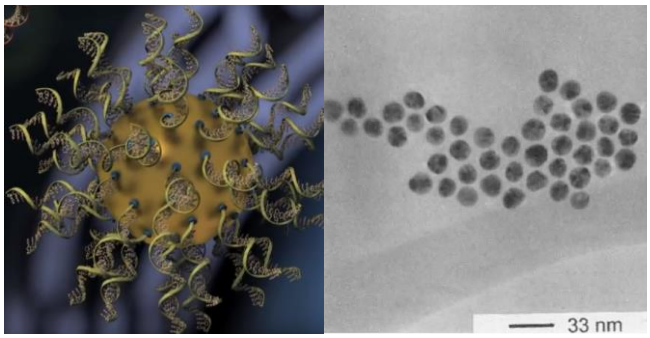


Figure 5: Cartoon view of DNA attached to gold nanoparticle forming a SNA (left) and a 2D colloidal aggregate of SNAs; adapted from ref. 16

1.2.3 DNA Hydrogels

A hydrogel is a type of material made from polymers that are crosslinked in order to absorb large amounts water in a 3D network. Typically, once a hydrogel is swelled with water, there is a larger weight fraction of water than the crosslinked polymer network. By adjusting the polymer chemistry and the density of crosslinks, the hydrophilicity of the hydrogel can be adjusted. These crosslinking points can either be physical or covalent, making it possible for DNA nucleobase interactions to hold the polymeric network together. Advantages of using DNA in hydrogels include being able to control the physical properties with temperature as well as controlling the strength of the crosslinks by adjusting the sequence of DNA, either changing its length or the number of C:G pairs. One use of hydrogels that is of particular interest is their ability to hold and release cargo, such as a therapeutic (Figure 6). In order for these materials to be used in vivo, they must be biocompatible, have sensical mechanical properties, release their cargo slowly, and degrade within a reasonable amount of time.

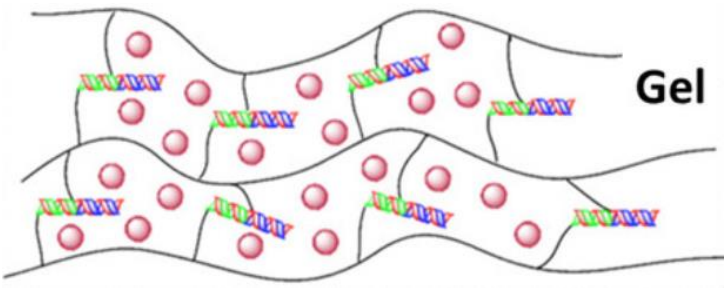


Figure 6: Cartoon of DNA hydrogel with payload (red circles) adapted from ref. 19

1.2.4 DNA-Polymer Hybrid Micelles

Micelles are a supramolecular structure that is obtained when amphiphilic molecules, a molecule with a hydrophilic and hydrophobic portion, aggregate together when placed in water. In order to minimize interactions between hydrophobic regions and water, many hydrophobic regions will pack together to form a micelle core which leaves the hydrophilic portions exposed to the water which is a more favorable interaction. The micelles can only form once the critical micelle concentration and temperature are surpassed. For example, DNA conjugated to a lipid has the potential to be a therapeutic, however, the stability of these micelles in complex organisms is not guaranteed since different locations in large organisms have different environments. Efforts have been made to develop controllable micelles in order to maintain stability regardless of spatiotemporal variability.

In one report[20], photosensitive dissociation of G-quadruplexes has been used to alter the stability of lipid conjugated DNA micelles to avoid disruption by serum albumin (lipid binding) while in blood (Figure 7). As the micelles bind to the blood protein, they are routed to organs such as the liver where functionality in other target organs is severely limited. Upon exposure to light, cytosine residues are released which bind to the residues responsible for G-quadruplex and the G-quadruplex

formation is blocked. In the presence of serum albumin, the micelles are no longer stable and cannot undergo cellular uptake. Only after successful cellular uptake can micelles deliver a therapeutic.

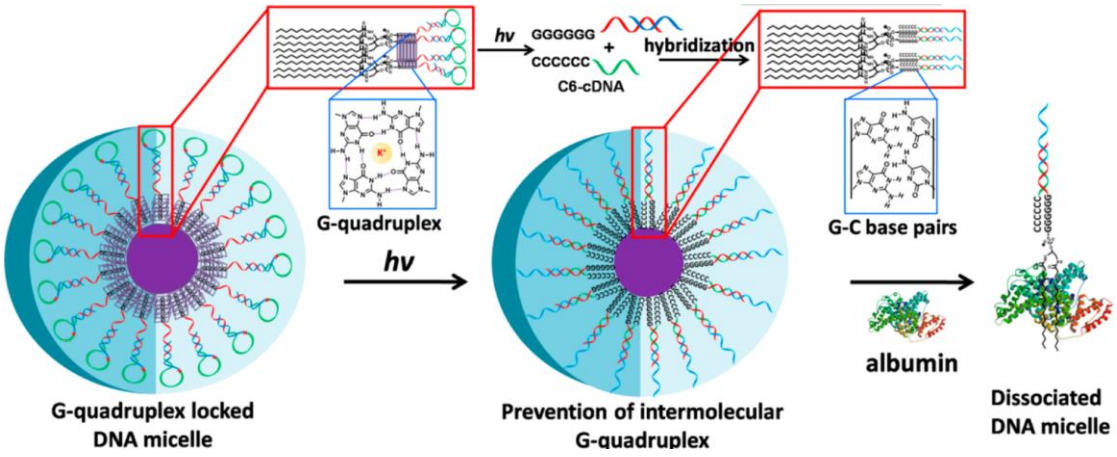


Figure 7: Scheme depicting the stability regulation by light; adapted from ref. 20

1.3 Phosphoramidite Nucleic Acid Synthesis

The need for synthetic, sequence-controlled DNA molecules is clear as they have uses in medicine and antisense applications. The synthesis of DNA is carried out using a solid phase strategy (discussed further later) in which the resin is supported on a filter. Advantages of the solid phase organic synthesis technique over the typical solution phase synthesis are the ability to use large excess of reagents, fewer purification steps, and the ability to be easily automated with a robot. The current phosphoramidite method pioneered by Marvin Caruthers is used to produce large quantities of DNA relatively quickly. The synthesis proceeds in cycles of multiple steps in which DNA is elongated in the 3' to 5' direction, opposite to natural DNA replication *in vivo*. After attaching the first nucleoside to the solid phase support, the

chain elongation process takes place in which cycles of four steps (Figure 8) are repeated to obtain the desired sequence. The sequence is then cleaved from the resin and its protecting groups are removed. The couplings are typically over 99% efficient[28], however, the need for protecting groups and multiple steps per nucleoside addition leave room for improvement.

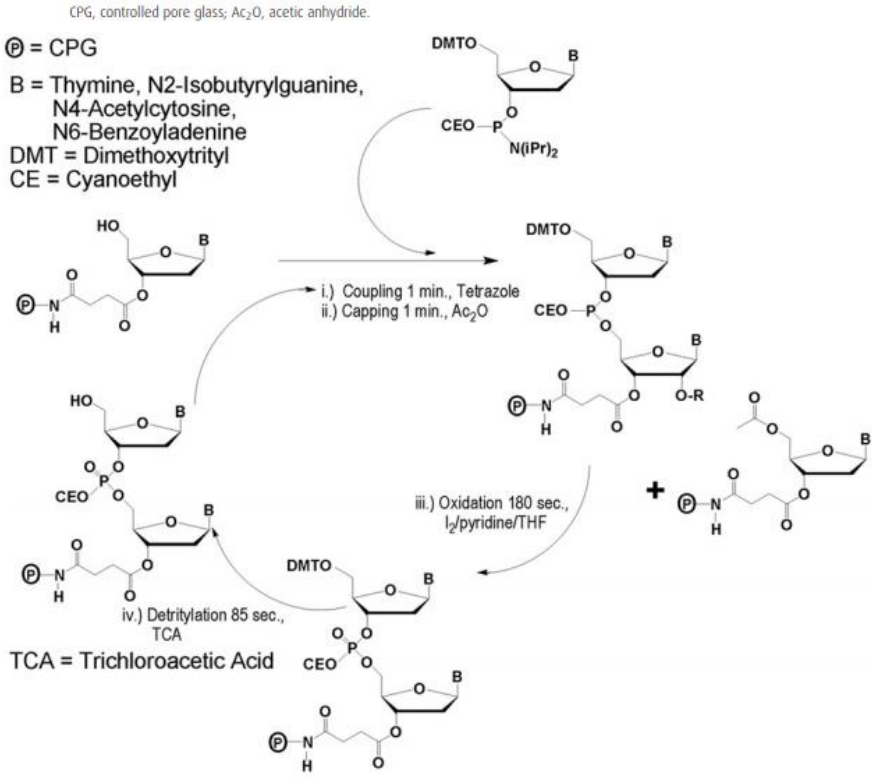


Figure 8: Synthetic cycle of DNA constructed using solid phase synthesis; adapted from ref. 27

1.4 Synthetic Nucleic Acid Analogs

Using naturally occurring nucleic acids for inspiration, researchers have developed synthetic nucleic acid analogs that take advantage of natural base pairing

capabilities but change the structure to overcome weaknesses. Often times, the backbone is altered to increase stability *in vivo* to enzymes or to temperature. The strength of the base-pairing can also be altered. Many applications of these materials exist such as biosensing, gene regulation, and sequencing.

1.4.1 Peptide Nucleic Acids (PNAs)

The term peptide nucleic acid was chosen to emphasize its synthesis as it relates to solid phase peptide synthesis and its relationship to nucleic acids even though the material is neither a peptide nor a nucleic acid. First reported in 1991 by Nielsen et al.[21], PNAs utilize the same four nucleobases as DNA with an achiral polyamide backbone that is neutral. The neutral backbone results in weaker electrostatic repulsion among the polyanionic DNA backbone allowing the PNA to bind more tightly and with more specificity to a complementary DNA strand than two complementary DNA strands do with each other. The synthesis of PNAs is straightforward and is carried out using well-established solid phase peptide synthesis techniques. An issue with PNAs is that their synthesis relies on inefficient amide bond formation chemistries requiring the use of at least 2.5 equivalents of PNA monomer, sometimes requiring as high as 10 equivalents.[22] This synthetic approach causes accumulation of waste, difficulty in scaling, and high costs. PNAs have several applications, however, these synthetic drawbacks and other poor material properties such as poor cellular uptake limit their large-scale production and leaves room for improvement.

1.4.2 Locked Nucleic Acids (LNAs)

In this nucleic acid analog, a methylene bridge formed between the 2'-O and the 4'-C positions locks the ribose in the 3'-endo conformation (natural B-DNA is in 2'-endo conformation). Here the term 'endo' refers to the pucker of the sugar moiety. LNAs take on the A form of duplexes. LNAs have the ability to bind with DNA and RNA following Watson-Crick base pairing rules. The introduction of the methylene bridge forces the monomer into the ideal conformation for base pairing which results in more rapid and stable base pairing, as observed by an increase in melting temperature of double stranded oligomers with the introduction of LNA monomers. Other advantages of LNAs include resistance to nucleases, higher base pairing specificity, and better base mismatch discrimination in antisense applications. For example, in the design of quantitative polymerase chain reaction (qPCR) systems, it is necessary for probe sequences to not only be highly specific but also have specific melting temperatures in order to quantify DNA yields.[24] With an increase in duplex melting temperature comes the ability for primer sequences to be shortened while retaining their specificity and allowing polymerases to begin replication. A comparison of the chemical structure of PNA, LNA, and DNA is shown in Figure 9.

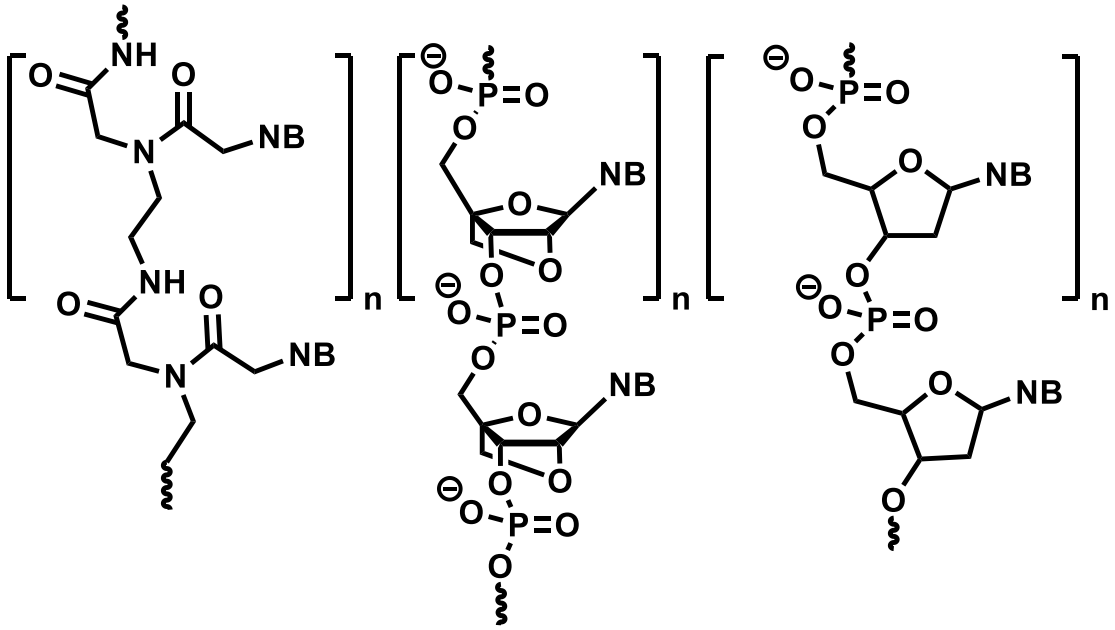


Figure 9: Comparison of PNA (left), LNA (middle), and DNA (right); NB = nucleobase

1.4.3 Previous Generation of CNA

The first generation of CNAs reported by Christopher Bowman and Christopher Kloxin[25] developed a framework from which the current CNA generation is built. Two types of CNA monomers were previously described that would allow for both sequence specific oligomer construction through the thiol-Michael click reaction and homopolymerization of an oligomer sequence through the thiol-ene click reaction as shown in Figure 10. Drawbacks from this specific monomer design include a nonoptimal electron withdrawing group (the acrylamide) and the inability to synthesize all four nucleobase monomers (only adenine and thymine shown). As shown in Appendix A, the vinyl sulfonamide goes to a higher conversion in less time. In their work, it was shown that an organogel was formed when the poly(T-*alt*-A) CNA polymer was covalently attached to a tetra-PEG-thiol. Since nearly half of the

human genome consists of repeating short sequences and many diseases are caused by trinucleotide repeat disorders, this scheme of developing short oligomers and then polymerizing the short oligomers can be incredibly powerful in biosensing with these naturally occurring DNA sequences. The CNAs aim to combine the ease of synthesis found in PNAs and the coupling efficiencies of the phosphoramidite couplings in DNA synthesis.

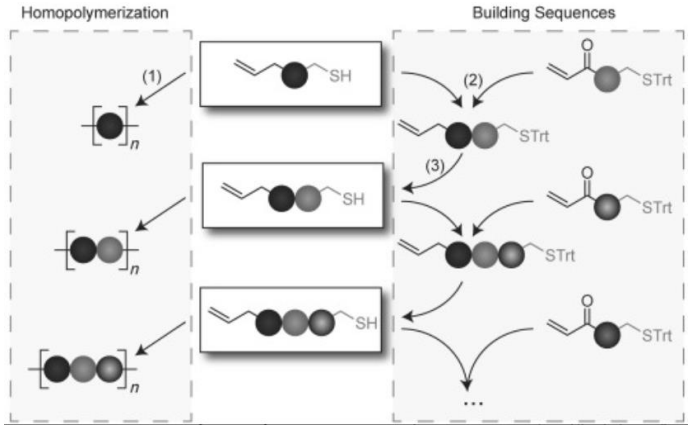


Figure 10: Reaction scheme depicting sequence specific oligomer synthesis through the thiol-Michael reaction and homopolymerization through the thiol-ene reaction; adapted from ref. 25

1.5 Click Chemistry

Click chemistry utilizes a subset of chemical reactions available to organic chemists that are high yielding, proceed under mild conditions, and possess a large thermodynamic driving force. This term for these reactions were first coined by Sharpless et al. in 2001[1] where an approach was outlined to create carbon-heteroatom bonds from readily available starting materials. Leveraging these desirable

characteristics, these reactions have found many uses in the engineering fields thanks to their ease of use and installation of functional handles.

One example of the click reactions discovered thus far is the thiol-Michael addition between a thiol functionality and an electron poor alkene to form a thioether bond. This reaction is either base or nucleophile catalyzed and proceeds through the proposed mechanism shown in Figure 11. In the first step of the base-catalyzed reaction, the base catalyst deprotonates the thiol, forming a thiolate. This nucleophilic thiolate attacks the electron deficient alkene, resulting in the thioether bond. The carbanion is resonance stabilized by the adjacent electron withdrawing group (EWG), making the reaction possible. This pair of electrons finally abstracts a proton from another free thiol, thus regenerating the thiolate.

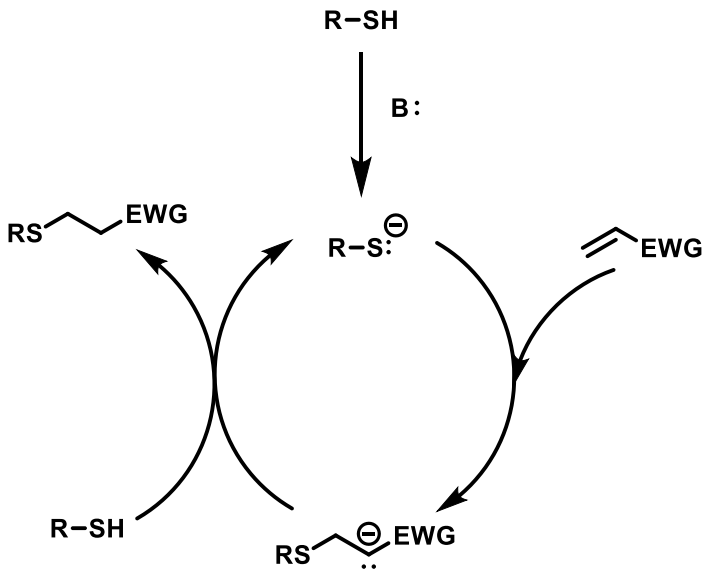


Figure 11: Mechanism for the Thiol-Michael addition click reaction; EWG = electron withdrawing group; adapted from ref. 6

1.6 Photochemistry

One important class of photochemical reactions that will now be briefly discussed is ortho nitrobenzene derivatives. Of particular interest is the 3-amino-3-(2-nitrophenyl)propanoic acid (ANP) linker that will be used to covalently connect a solid phase resin bead to a growing oligomer. The ANP linker will cleave to afford the oligomer product as a primary carboxamide. The mechanism for this ANP linker is thought to be an intramolecular hydrogen abstraction in the ortho position relative to the excited nitro group.[9] The nitrobenzyl based linker tends to follow a Norrish Type II mechanism. By introducing light energy at 365 nm, the N=O π bond is broken yielding two radicals. Subsequently, the linker becomes excited and enters the singlet state where the ground state electron is spin paired with the excited electron. Following an intersystem crossing event, where the spin paired electrons transition to an unpaired state, the lower energy triplet state is reached. The oxygen radical then abstracts a hydrogen from the benzylic carbon and the resonance of the π electrons yields an aci-nitro compound. Finally, a five membered ring is formed leading to the cleavage forming the product of the photocleavage reaction as shown in Figure 12.

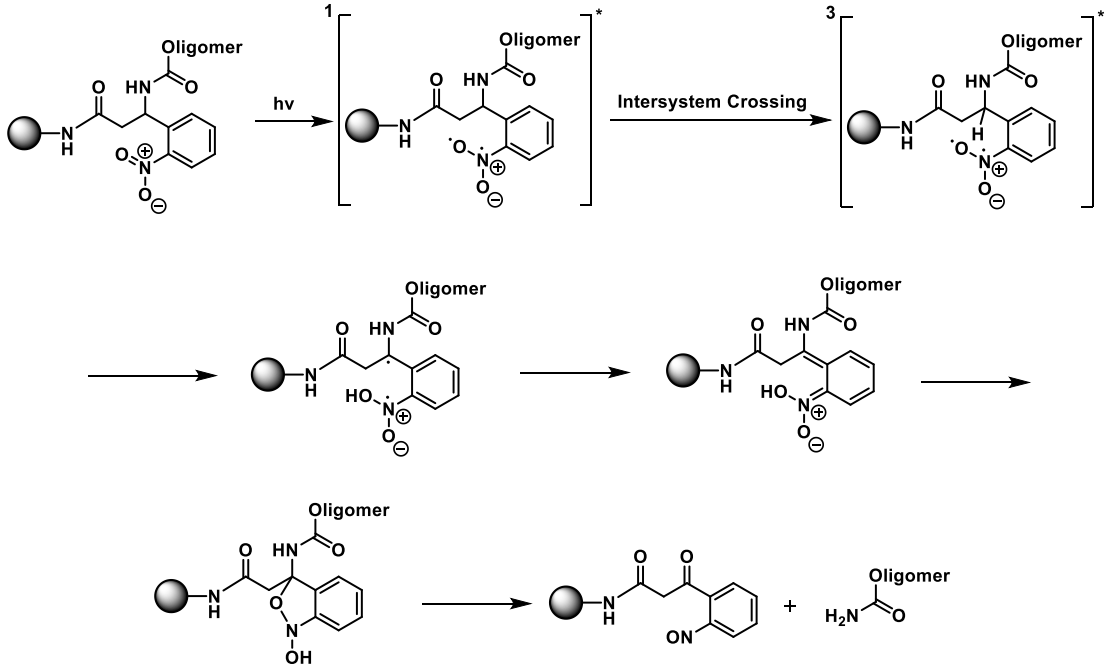


Figure 12: Proposed mechanism photocleavage of ANP linker

Chapter 2

CLICK NUCLEIC ACID MONOMER DESIGN AND SYNTHESIS

2.1 CNA Structure

The design of the CNA monomer was inspired by DNA in conjunction with protected amino acids used in solid phase synthesis. Both DNA and CNA molecules share the same four nucleobases: adenine, cytosine, guanine, and thymine, however, the backbone structure is different. In DNA, the backbone consists of alternating phosphate and deoxyribose sugar molecules whereas in CNA the backbone is composed of mainly alkanes linked together by thioether bonds. Since a similar structure is used, the CNA molecules should have many of the same features as DNA, primarily the ability for complementary strands to interact via nucleobase pairing and to form a secondary structure. At this time, it is unclear as to the ability of this generation CNA to bind with DNA. By removing the phosphodiester bonds in DNA, the CNA is no longer susceptible to hydrolysis via nucleases, a class of biological enzymes that is responsible for DNA repair, replication, and degradation. A similar spacing of nucleobases is also used; in DNA there are six atoms separating each monomer whereas in CNA there are seven atoms. This difference is not ideal but is necessary due to the current synthetic pathway used to create the CNA monomers.

The goal of this project is to create sequence-controlled CNA polymers and to enable this, a solid phase synthesis approach will be used. In solid phase synthesis, monomers are added one at a time in a step by step, controlled manner. Each monomer that is coupled utilizes a protecting group on its terminus to prevent multiple monomer additions leading to the loss of sequence control. The need for this terminal protecting group has largely driven the monomer synthesis and is the reason for multiple versions

of each of the four basic monomers. In this generation of CNA monomers, the trityl group is used to protect the thiol functionality. Modifications to this basic structure will be discussed later with minor modifications to the synthetic pathways.

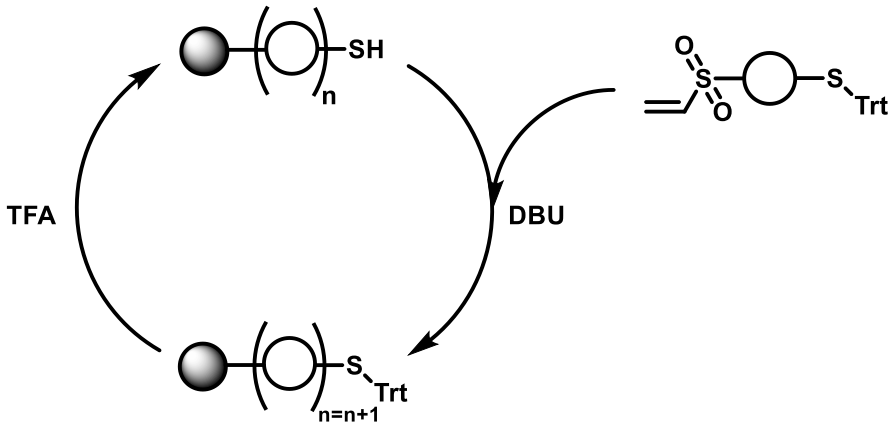


Figure 13: Schematic representing subsequent monomer coupling strategy; the empty circle represents any of the four nucleobase monomers and the filled circle represents the resin with the linker

2.2 CNA Thiol-Michael Monomer Synthesis

The synthesis of each of the four CNA monomers follows a similar path with slight adjustments due to the inherent chemistry of each nucleobase. The general scheme is to start with the nucleobase, then add an acetaldehyde diethyl acetal functionality, convert this to an aldehyde/diol, add the backbone, and finally add the vinyl sulfone. This general scheme is outlined in Figure 14. This final monomer will be referred to as the thiol-Michael monomer because this monomer structure will be used to participate in the thiol-Michael click reaction discussed earlier. This synthetic

route affords the opportunity for all four nucleobases to be made unlike the previous generation of CNAs.

In the first step, the secondary amine of the nucleobase heterocycle acts as a nucleophile to undergo a S_N2 reaction with the alkyl halide bromoacetaldehyde diethyl acetal. The tertiary amine product can theoretically react once more to form a quaternary amine that has been alkylated twice, however, this is not observed experimentally. Slight variations in the starting material are required in order to direct the proper reaction because of the multiple potential heterocycle amines.

In the second step, the acetal is cleaved giving an aldehyde/diol product, often called a geminal diol. An equilibrium exists between the two forms and the major species can be changed by altering the solvent. The difference between the two forms is a molecule of water so it is often understood that the diol is the hydrate of the aldehyde. This is an important distinction because the subsequent reaction requires an aldehyde, but the NMR spectrum (in deuterated water) obtained for the geminal diol shows only the diol.

The next step is a reductive amination reaction between the aldehyde intermediate and the backbone. Whereas direct alkylation of an amine can happen multiple times, the addition of an aldehyde to form an imine will only result in a single addition. In this reaction, first an imine intermediate is formed when the primary amine of the backbone and the aldehyde of the nucleobase combine, releasing a water molecule. The double bond of the imine is then reduced using a reducing agent, either soidum cyanoborohydride or sodium triacetoxyborohydride.

Finally, the electron deficient alkene functionality, the vinyl sulfone is attached. In this reaction, triethylamine deprotonates the secondary amine on the

backbone allowing it to perform a substitution and a subsequent elimination to yield a vinyl sulfonamide.

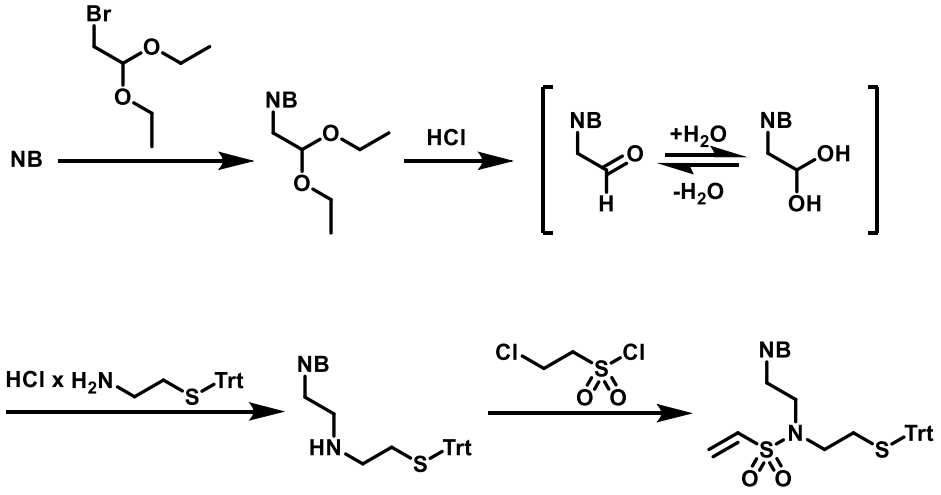


Figure 14: General scheme for the synthesis of thiol-Michael CNA monomers

2.2.1 Thiol-Michael Adenine Monomer Synthesis

Adenine and thymine are the two nucleobases that follow the straightforward synthetic route outlined above without exception. During the synthesis, it was noted that the diol/aldehyde intermediate is sensitive to light, as determined by the white solid turning yellow/brown upon too much exposure to ambient laboratory and fume hood light. Thus, precautions such as working without lights on and using aluminum foil to cover the mixture were taken to avoid degradation of the material. The overall yield for this multistep synthetic route is 9.5% (see Figure 15).

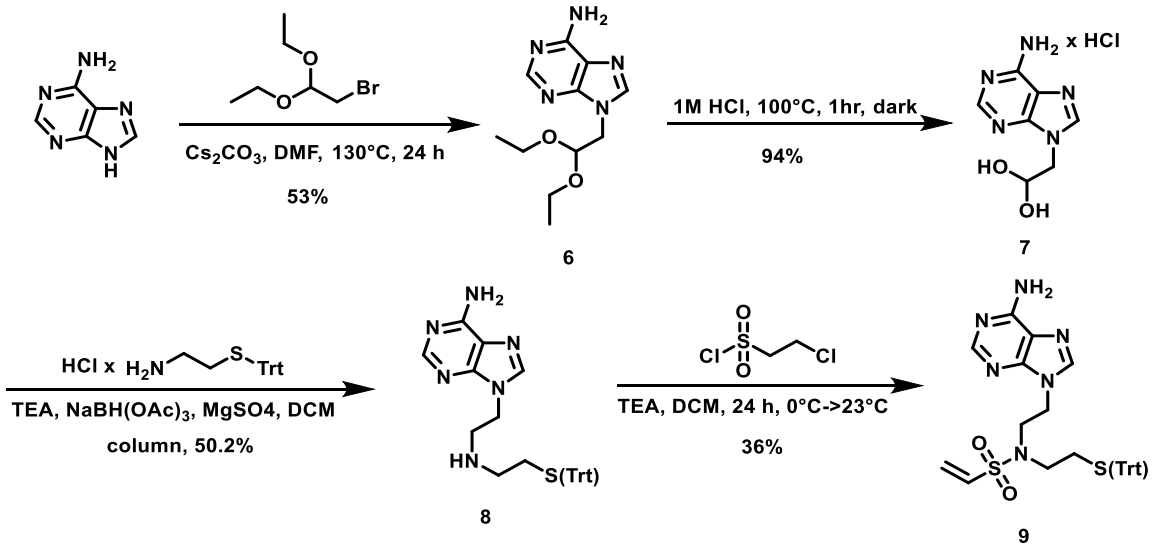


Figure 15: Synthetic pathway for the thiol-Michael adenine monomer

2.2.2 Thiol-Michael Cytosine Monomer Synthesis

Cytosine is the first exception to the general reaction scheme shown in Figure 14. The issue with using cytosine as the starting material is that there are three possible locations at which the reaction could take place. This behavior of having multiple side products was observed when using cytosine as the starting material as opposed to using N4-acetylcytosine. The desired addition of the acetaldehyde diethyl acetal to the secondary amine does occur, but a mixture of other products makes separation difficult. By using the N4-acetylcytosine as the starting material, the desired reaction is maximized and the side product is easily separated out. The removal of the acetyl protecting group is easily accomplished with 2M ammonia in methanol. The overall yield for this multistep synthetic route is 12.7% (see Figure 16).

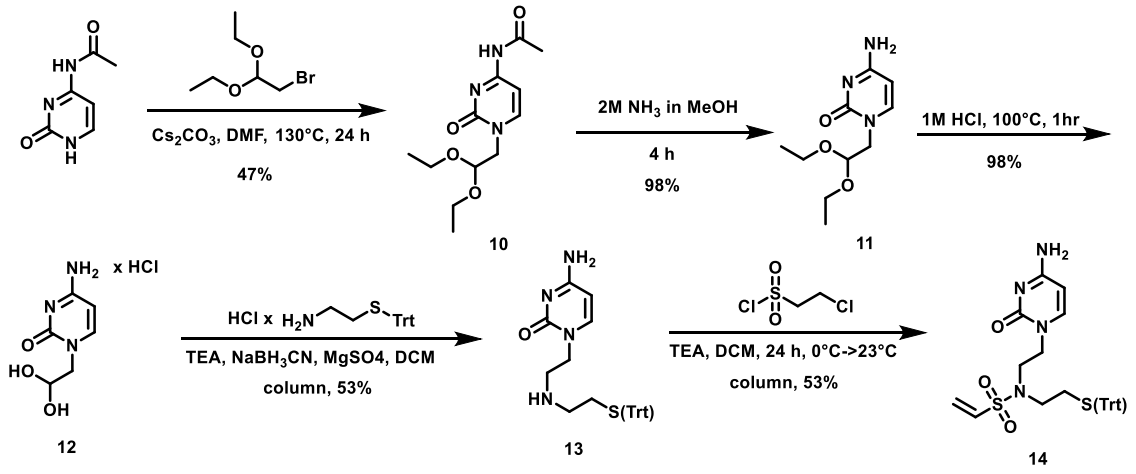


Figure 16: Synthetic pathway for the thiol-Michael cytosine monomer

2.2.3 Thiol-Michael Guanine Monomer Synthesis

The guanine monomer is the other exception to the general scheme in Figure 14. The starting material used is 2-amino-6-chloropurine as opposed to guanine. In this case the desired reaction on the secondary amine (N9 position) can take place. By using the 2-amino-6-chloropurine, the pKa of the desired reactive site on the N9 position is increased to approximately 10.12 which makes it the most likely to react. The guanine starting material has a more basic pKa of approximately 11.4 at the N1 position, which would yield the wrong intermediate. After the N-alkylated product is obtained, the 2-amino-6-chloropurine is converted back to guanine by oxidizing the 6-chloro moiety to a 6-oxo functionality. The overall yield for this multistep synthetic route is 6.6% (see Figure 17).

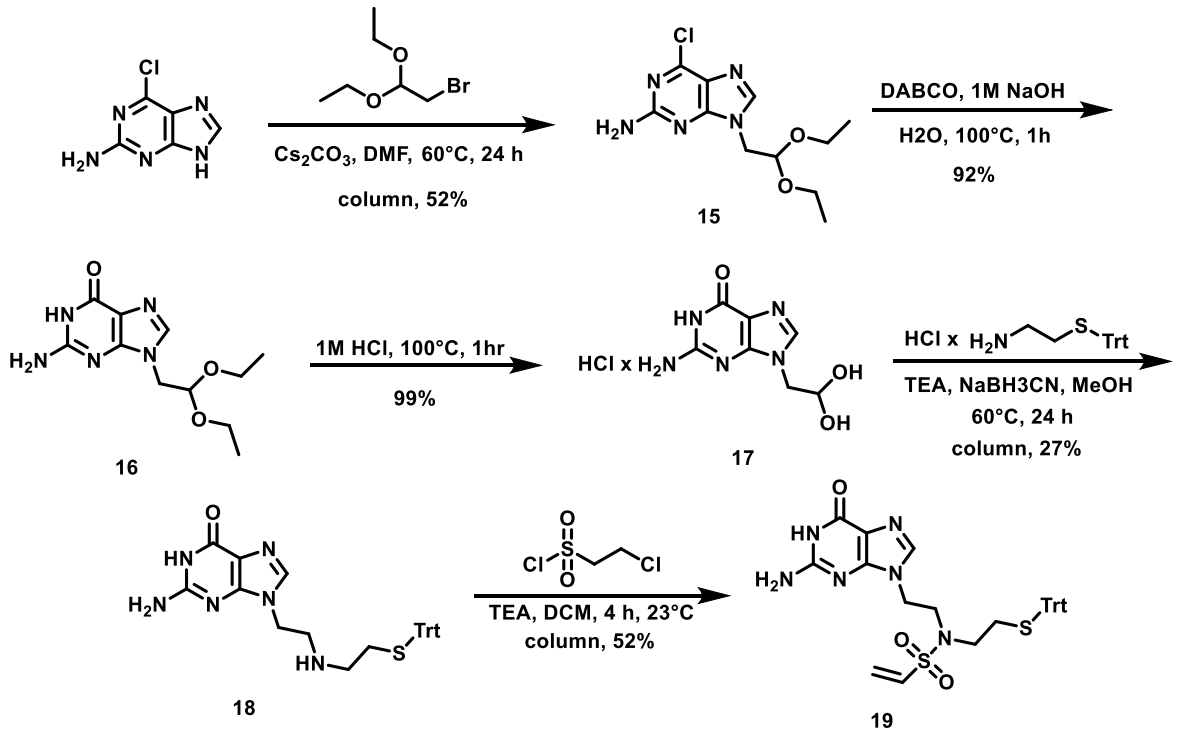


Figure 17: Synthetic pathway for thiol-Michael guanine monomer

2.2.4 Thiol-Michael Thymine Monomer Synthesis

The thymine monomer synthesis is straight forward and has the best yields of the four monomers with an overall yield for this multistep synthetic route is 17.3% (see Figure 18). For this reason, it is used in further experiments thus saving time synthesizing starting material.

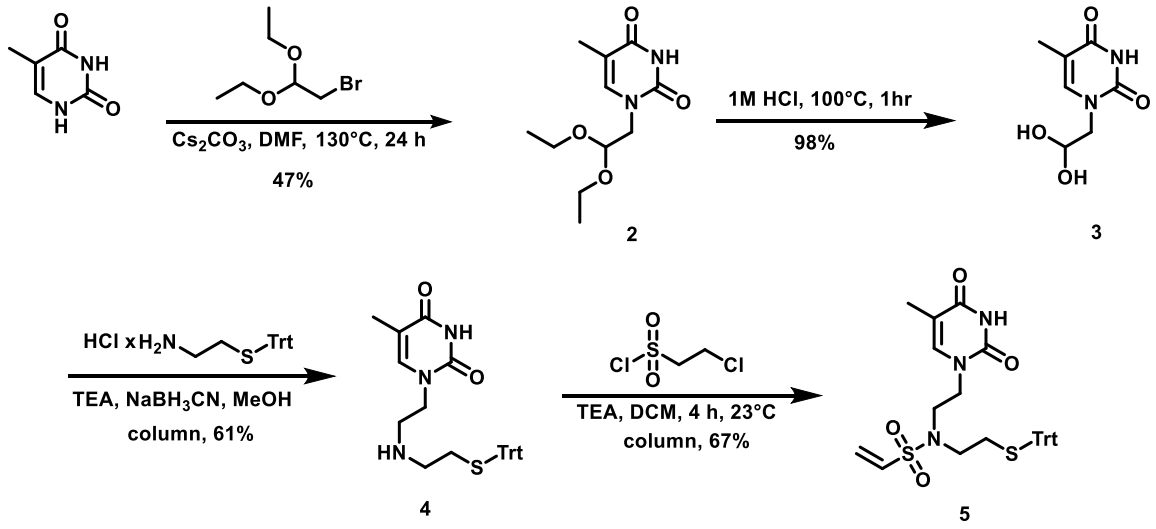


Figure 18: Synthetic pathway for the thiol-Michael thymine monomer

Chapter 3

SOLID PHASE, SEQUENCE-SPECIFIC CNA SYNTHESIS

3.1 Solid Phase Oligomer Synthesis

A procedure pioneered by Merrifield[15] to develop peptides was used as inspiration for the design and synthesis of the click nucleic acids. The scheme developed by Merrifield involved the use of a solid phase support polymer covalently bound to a growing peptide chain. The peptide chain is grown by coupling fluorenylmethyloxycarbonyl (Fmoc) protected amino acids to the peptide chain. The synthesis of a peptide is accomplished by repeating cycles for the addition of amino acids: couple Fmoc protected amino acid, wash away unreacted amino acid and coupling reagents, deprotect Fmoc group, wash away Fmoc and deprotection reagents. This work was of such importance that Merrifield was awarded the Nobel Prize in 1984.

In the case for CNA synthesis, the amino protecting group Fmoc has been replaced with a thiol protecting group called triphenylmethyl (trityl or Trt) or one of its derivatives such as monomethoxytrityl (Mmt). Along the same line at the other terminus, the carboxylic acid has been replaced with an electron deficient alkene. A comparison of the monomer units used in traditional peptide synthesis and CNA synthesis are shown in Figure 19. In traditional peptide synthesis, a carboxylic acid reacts with an amine to form amide bonds whereas in CNA synthesis, a thiol reacts with an alkene to form thioether bonds. A major advantage of the CNA coupling is that a click reaction is used so that conversions are quantitative whereas in traditional peptide synthesis, couplings are not as efficient under ambient conditions.

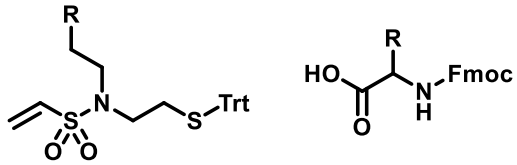


Figure 19: Comparison between current CNA monomers (left) and traditional Fmoc protected amino acids (right)

3.2 Solid Phase Scheme Development

The design of the solid phase synthesis scheme relies on a few key parameters that will now be discussed. Based on these criteria, two separate solid phase schemes have been developed. The first critical parameter for solid phase synthesis is the solid phase support, otherwise known as the resin. This resin is a polymer that is not soluble and has a functionality with variable loading capacities that the oligomer of interest can be grown from. In both of the systems developed, the resin uses a TentaGel S structure with slight differences at each functional handle. TentaGel S is a copolymer consisting of a low crosslinked polystyrene matrix on which poly(ethylene glycol) (PEG) is grafted. The majority of the weight of the resin (50-70%) is PEG and thus its properties dominate the resin. Attached to the PEG is a functional group of your choice that allows for a linker to be installed. This is the next important parameter and is the reason for two separate schemes. In the first scheme, the TentaGel S NH_2 scheme, a linker is not already attached upon receipt from a commercial source. In the second scheme, the TentaGel S RAM scheme, a linker is attached upon receipt from the manufacturer. It is crucial that the conditions used to eventually cleave the oligomer from resin are orthogonal to the conditions used for monomer coupling and deprotection. This need for different chemistries gives rise to the two schemes that are now developed.

3.2.1 TentaGel S NH_2 Resin

Using this type of solid phase support requires that a custom linker be installed prior to oligomer synthesis. A photolabile linker that can cleave when irradiated by particular wavelengths of light is desirable due to the orthogonality of light to the linker and not any part of the oligomer. The nucleobases and the trityl group have UV absorbances around 260 nm and 420 nm, respectively. Due to the complex nature of the nucleobase monomer structure, a linker was needed that would not interfere with these other chemistries. The ANP photolabile linker was chosen for its compatibility with both acid and base sensitive protecting groups as well as its reported high recovery after photolysis.[11] The photocleavage product is a primary carboxamide shown in Figure 20 that should not have a detrimental effect on base pairing.

The first modification made to this resin is the addition of the ANP linker using standard amide coupling conditions (HBTU, NMM, DCM) in Appendix B.5. The ANP linker itself has an amino functionality that is Fmoc protected. The Fmoc group is then removed using a solution of 20% piperidine in DCM. Verification that the ANP linker has been added is carried out using the Kaiser test for free amines as described in Appendix B.2.

At this point, the construction of the oligomer begins. Based on initial runs not shown in this thesis, it is known that solubility would be an issue. To overcome this issue, two Alloc protected lysine amino acid residues were added first. The coupling of these residues follows the procedure in Appendix B.3. After removing the Fmoc, a short molecule, 2-(tritylthio)acetic acid, is added to provide a trityl protected thiol functionality. The trityl group is removed using 20% trifluoroacetic acid (TFA) with 5% triethylsilane (TES) as a scavenger. These are relatively harsh conditions that would otherwise cleave a linker if not chosen carefully. This finally yields a free thiol

is ready to build the CNA oligomer. The Ellman's test is carried out to verify the presence of free thiols after the deprotection of the trityl group.

The monomer used in this strategy utilizes a trityl protecting group on the thiol functionality as developed in Chapter 2. The monomer coupling is carried out using one equivalent of the base catalyst 1,8-diazabicyclo(5.4.0)undec-7-ene) (DBU) with four equivalents of monomer in two separate coupling events in an attempt to get full conversion. Currently, there has been no attempt in optimizing this system in terms of monomer equivalents, amount of catalyst, temperature, or duration of coupling. The trityl functional group is removed by acid, as described earlier allowing for future couplings. Each coupling and deprotection are verified using the Ellman's test for free thiols.

The terminal monomer after the final coupling retains its trityl protecting group in order to prevent the formation of disulfides between two oligomers after resin cleavage. A microcleavage is then carried out in a solution of 4:1 (v/v) of water:methanol under 365 nm light at an intensity of 20 mW/cm^2 for one hour as described in Appendix B.6.1. The resulting brown solution is then injected into a mass spectrometer for initial characterization following Appendix B.7.

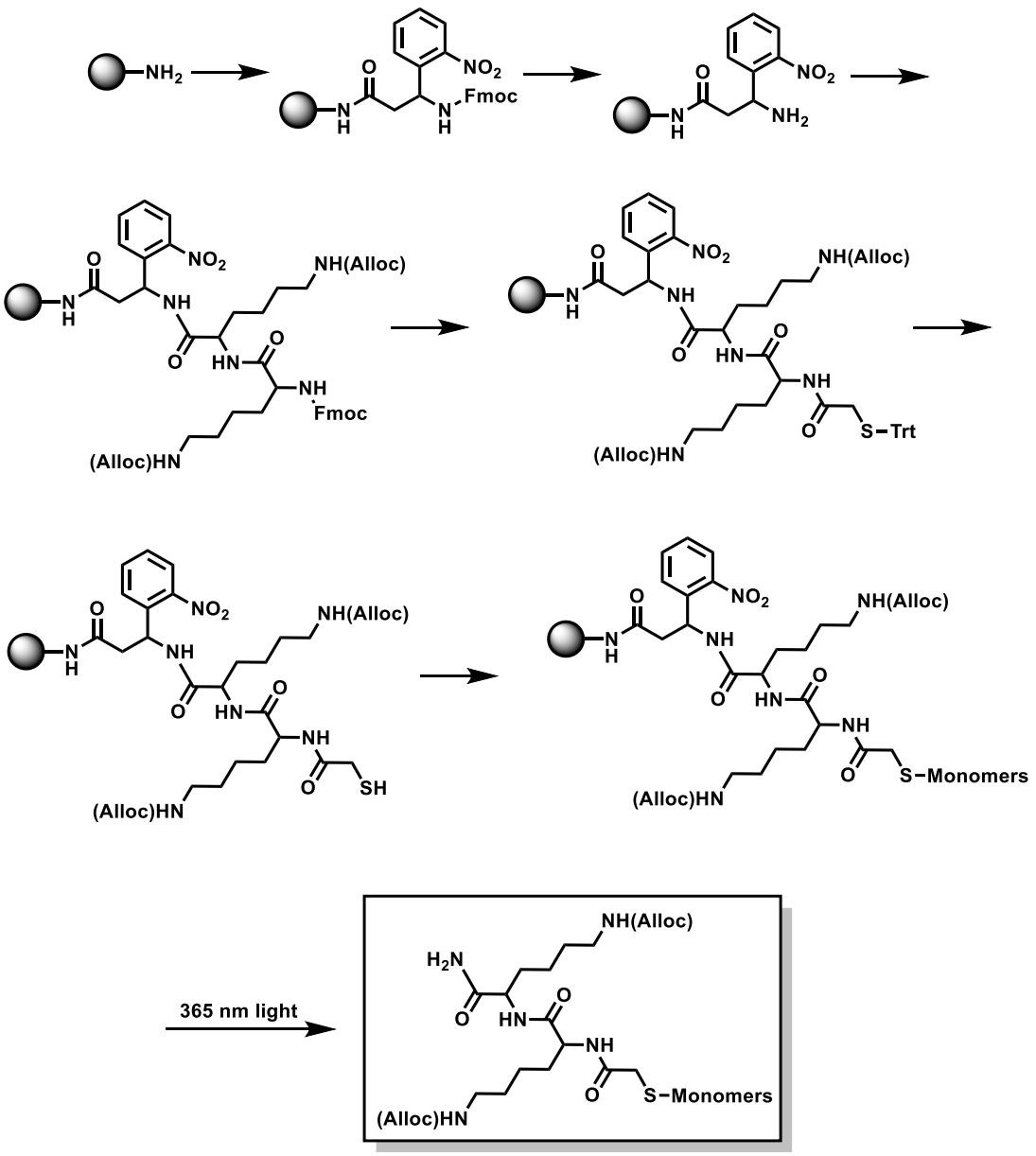


Figure 20: Installation of ANP linker and primary carboxamide product; dark circle represents the TentaGel S resin

3.2.2 TentaGel S Rink Amide (RAM) Resin

This type of commercially available solid phase support comes preinstalled with a Fmoc protected amine functionality. Since this is a common type of linker, this

strategy is more likely to be utilized in future scale up. This rink amide linker has the structure shown in Figure 21. This structure has greater acid sensitivity due to the two additional methoxy groups' electron donating properties compared to previous generation methylbenzhydrylamine (MBHA) linkers. The sensitivity of this linker prevents the use of the trityl group on our monomers. Thus, a protecting group with more sensitivity to acid is required. To meet this requirement, the Mmt group was substituted for the Trt group. Luckily, the synthetic route to make these monomers follows the previously outlines path with the exception of a Mmt protected backbone. The additional methoxy group makes the protecting more sensitive to acid. In this case, the Mmt group is more sensitive to acid compared to the linker allowing for the selective removal of the monomer's Mmt group without cleaving from the resin. The rationale described for the TentaGel S NH_2 linker scheme in terms of lysine residues is replicated here instead using Boc protection as opposed to Alloc protection. An acetyl capped cysteine residue with the thiol Mmt protected is also used instead of the 2-(Tritylthio)acetic acid linker. The cysteine was used here since it is a more common reagent that would have a lower cost for scale up.

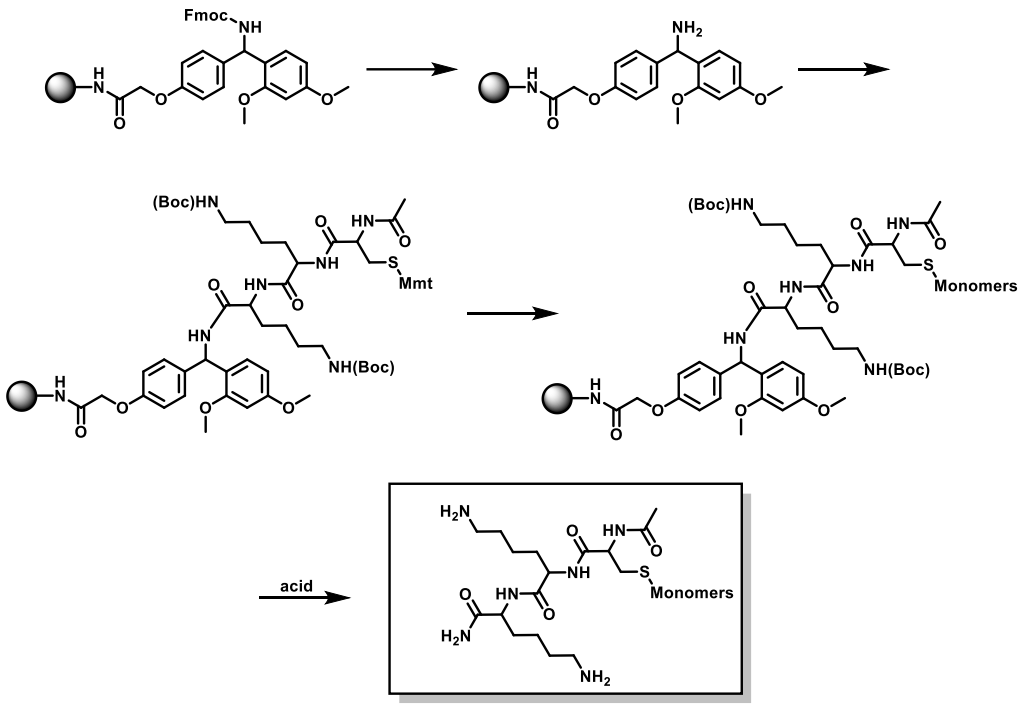


Figure 21: Installation of rink amide (RAM) linker and primary carboxamide product; dark circle represents the TentaGel S resin

3.3 Solid Phase Results for Both Schemes

3.3.1 Results Using the TentaGel S NH_2 Resin Strategy

The first attempt at using the ANP photocleavable linker with the thiol-Michael thymine monomer yielded mixed results. In this experiment, four equivalents of thiol-Michael monomer compared to the linker thiol functionality were used with a 0.4 equivalent of DBU base catalyst. In the following plot, each sample was not taken at the same concentration since each was the result of a qualitative microcleavage that was not purified. Therefore, only the relative intensities for each run individually can be compared against one another. In these ultra-performance liquid chromatography (UPLC) traces, the largest absorbance comes from the trityl groups that are associated

with successful monomer additions. During the ionization of the sample (electron spray ionization in the mass spectrometer used), the trityl group often fragments and the product is found where there is a large absorbance. These results are positive since they show that the sequences can in fact be made. The downside to these results is that the conversions are not quantitative as determined by remnants of the previous oligomer in subsequent runs.

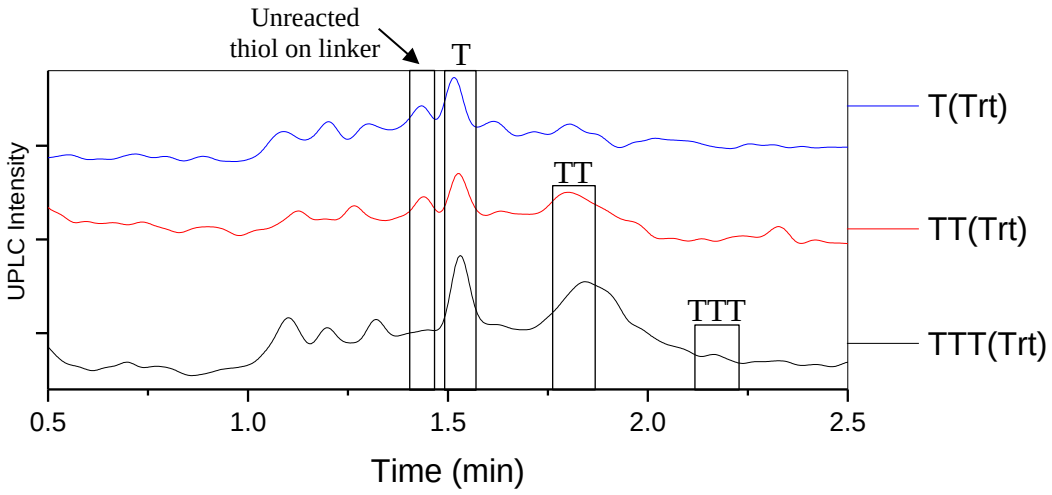


Figure 22: UPLC traces of ANP photodegradable linker oligomers

3.3.2 Results Using the TentaGel S RAM Resin Strategy and Mmt Protected Thiol-Michael Monomers

From previous experience with the synthesis of the monomers as well as previous solid phase attempts, it was known that solubility of the oligomers would be a major concern.[26] It was also known that due to the harsh deprotection conditions of the trityl group, the thiol-Michael monomers developed previously could not be used. As a result, the RAM resin was first loaded with two Boc (tert-Butyloxycarbonyl) protected lysine amino acid residues using the standard amino acid coupling procedure

in Appendix B.3. Once deprotected, the free amines on the lysine residues would assist solubility by providing a charge and hydrogen bonding. A cysteine residue was then installed following the same protocol in order to have a thiol functionality that the oligomer can be grown from. The Fmoc protecting group on the cysteine is subsequently removed using the piperidine as before and it was capped using a solution of 5% acetic anhydride and 5% DIPEA in DMF. This capping procedure was repeated to ensure full conversion and was verified using a Kaiser test. At this point, the RAM resin has two Boc protected lysine residues and a Mmt protected cysteine residue that had been capped with an acetyl group. As a first attempt without any attempt at optimization, the Mmt group on the cysteine was deprotected using 15 washes of 2% TFA in DCM for 30 sec each since no scavenger was present to remove impurities. During the first few washes, a bright yellow/orange color was observed indicating that Mmt was being removed as shown in Figure 23 a). By the time 15 washes were completed, the orange color had disappeared and the solution was clear. An Ellman's test for free thiols was carried out and the resulting solution was yellow, indicating the presence of thiols as shown in Figure 23 b).

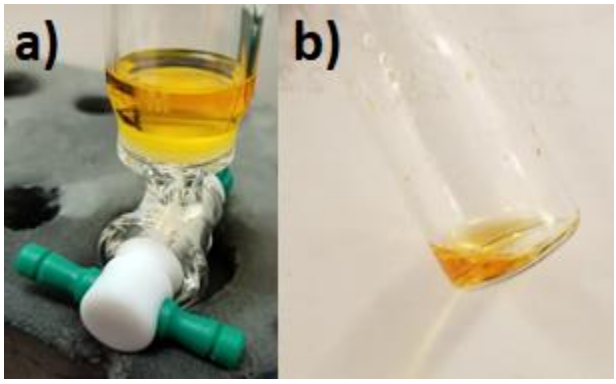


Figure 23: a) Peptide synthesis vessel containing resin during first wash of Mmt deprotection; b) Ellman's test for free thiols following 15 washes of Mmt deprotection

With the free thiol, the addition of the first monomer can occur. Relative to the number of functionalities present on the resin, five equivalents of the thiol-Michael thymine monomer, 1 equivalent of DBU, and DMF were mixed in a separate vial. The vial's contents were transferred to the peptide synthesis vessel and the reaction was vortexed for one hour. After the hour, DMF was used to wash the resin to remove excess monomer and base. A microcleavage following the protocol in Appendix B.6.2 and subsequent analysis in a mass spectrometer were carried out. This resulted in a mixture of both Lys-Lys-Cys-Thymine-SH and Lys-Lys-Cys-SH. During the resin cleavage, the strong acidic conditions remove the Boc and Mmt protecting groups. There was a need for a second coupling of the first monomer in order to reduce the amount of unreacted thiol on the cysteine. Thus, a second coupling with conditions identical to those outlined above was carried out. As seen in the UPLC trace in Figure 24, this helped by adjusting the relative intensity of the reacted to unreacted thiols, but there is still some unreacted thiol.

With these results, it is clear that some optimization in terms of the amount of monomer, the amount of base catalyst, and the length of the coupling must be completed. To this end, a brand new batch of RAM resin loaded with Lys-Lys-Cys with the acetyl capped N-terminus was prepared. The same general sequence of deprotections and washing occurred. For the coupling, new conditions of 10 equivalents of monomer, 2 equivalents of DBU, and one hour reaction time were used. The results from this run are not shown, however, the ratio of the intensities of the successful coupling to the free thiol is much better than the first run. At this point, the addition of the second monomer was carried out using the same conditions. After the microcleavage, the UPLC trace was not as good as the first monomer in terms of the amount of reacted thiol. Thus, a second coupling of the second monomer was carried out. The UPLC trace showed only a slight increase in the ratio of reacted to unreacted thiol.

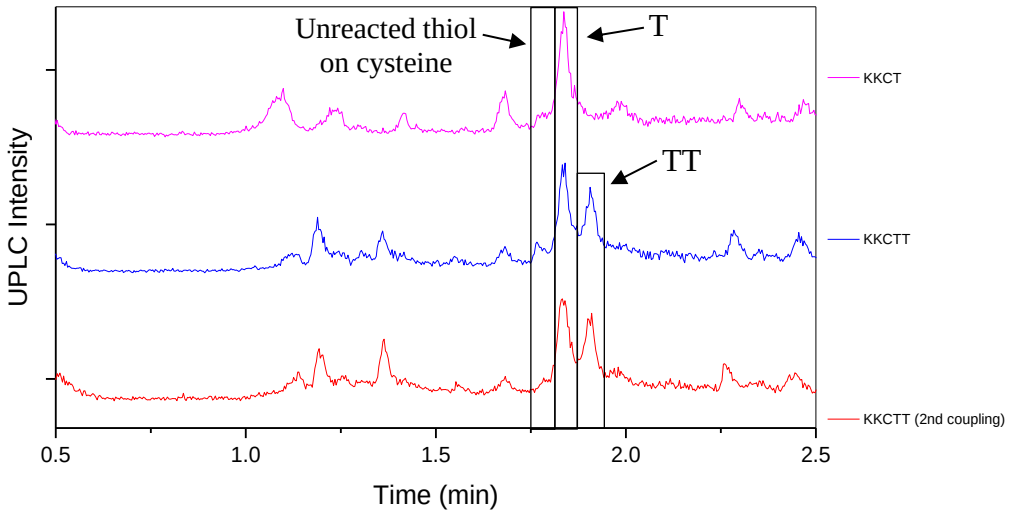


Figure 24: UPLC trace prior to running sample through mass spectrometer; growth of oligomer from a single T monomer to a dimer of T

In response to the low conversions observed in prior attempts, a new trial was conducted in which 50 equivalents of thymine monomer and 20 equivalents of DBU catalyst were used with a one hour reaction time. This trial showed that all of the free thiol attached to the cysteine was converted with the addition of one thymine monomer. A second thymine coupling was attempted following the same conditions as the first. This resulted in a mixture of single T and double TT, indicating that the conversion was not quantitative (see Figure 25). These results indicate that the thiol on the monomer is not as efficient as the thiol on the cysteine.

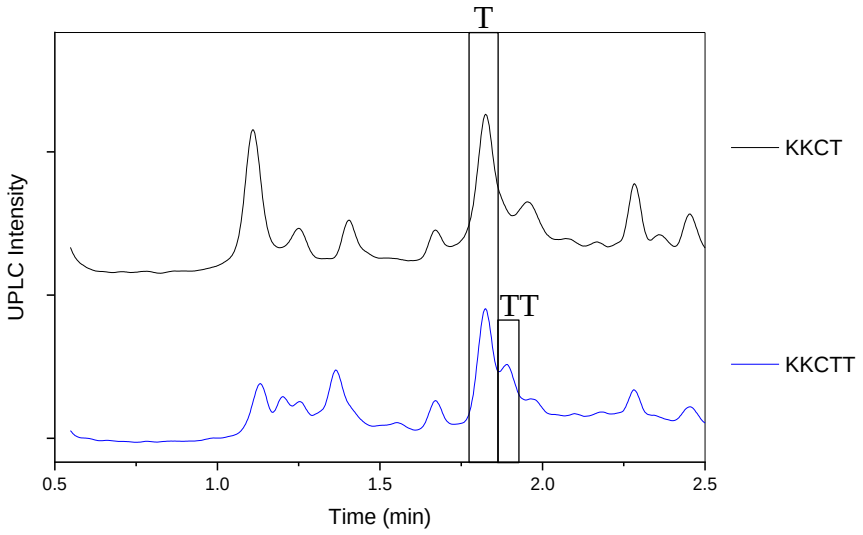


Figure 25: UPLC trace of trial using 50 eq. thymine monomer and 20 eq. DBU catalyst for one hour for each coupling

3.3.3 Discussion

The trials conducted have shown that the thiol-Michael monomers synthesized can be added to a solid phase support. These results, however, show that the coupling

is not always quantitative as one would expect from a click reaction such as the thiol-Michael. A clear difference in the degree of reactivity was observed between the thiol of the cysteine and the thiol of the monomer unit. One explanation for this is because the cysteine has a pKa of about 8.3 and thioglycolic acid linker has a pKa of around 9.3 whereas the pKa on the thiol-Michael monomers is in the range of 10.06-10.3 according to calculation from Chemicalize. The lower pKa of the thiol increases the amount of thiolate formed to proceed with the mechanism. Another potential issue is disulfide formation amongst free thiols on the resin. This can be addressed by using a resin with less densely packed functional groups or by using an organic buffer solution to prevent their formation. Further complicating the reaction is that the thiols are on an insoluble bead whereas the alkene in the desired reaction is in the liquid phase. Another approach that can be tried is using heat to push the reaction forward similarly to peptide synthesizers that utilize a microwave. The last issue could be formation of some secondary structure while still on resin that could be hiding the thiol functionality from reacting with its intended partner.

Chapter 4

CONCLUSIONS AND FUTURE DIRECTION

4.1 Conclusions

In this thesis, many objectives have been accomplished that have culminated in the development of click nucleic acids (CNAs). This new class of DNA analog has the potential to displace other DNA analogs due to its utilization of the highly efficient click reaction, the thiol-Michael reaction, and its *in vivo* resistance to enzymatic degradation. Additionally, the sequence specific nature of the synthesized oligomers has the ability be used in traditional DNA analog applications such as qPCR probes or in biosensing.

A synthetic route to synthesize all four nucleobases including adenine, cytosine, guanine, and thymine, has been identified and executed. Variations on these four monomers have subsequently been used in solid phase oligomer synthesis to yield sequence specific nucleic acid analogs. A photocleavage protocol has successfully been developed to separate the growing oligomer from the solid phase support. The cleaved oligomers have been characterized using mass spectrometry.

4.2 Future Direction

The oligomers made thus far are short and have largely been limited by their solubility. In order to address this, work has begun to find a synthetic approach to make a more soluble monomer. As shown in Appendix D, an initial route to make such a monomer is described. The addition of the carboxylic acid (once the methyl ester is cleaved with base) should increase the solubility of the monomers making coupling and characterization easier.

Many potential applications using the CNA sequences exist now that the framework for their synthesis has been established. One potential application is in the formation of hydrogels. With the sequences synthesized via solid phase, they can be grafted to a polymer of PEG, for example. The hydrogel would then be held together by physical crosslinking mediated by Watson-Crick base pairing. Another application of the CNA sequences could be in the construction of micelles. The CNA sequences would be conjugated to a polymer with hydrophilic properties to create an amphiphilic molecule. Using potential CNA-protein or CNA-drug interactions, a cargo can be physically attached to the amphiphilic molecules. These molecules could then be placed in an aqueous solution to self-assemble into cargo-carrying micelles.

With these CNA oligomers, characterization in terms of their duplexing capability with other CNA oligomers needs to be completed. More importantly, the duplexing ability of CNA with DNA as well as RNA is extremely important as this would determine what *in vivo* applications this material would have. Further investigation into the limits in length of the oligomers and the secondary structure formation, if any, must also be investigated. The secondary structure of the duplex with a complementary oligomer can be determined using circular dichroism spectroscopy and the binding strength between strands can be determined using UV-Vis spectroscopy to generate melting curves.

Future research into the optimization of the solid phase coupling conditions for both developed schemes needs to be done. Work into the optimization of this coupling has begun. As shown in Appendix E, compound **27** has been synthesized to act as a model system. With the added acetyl group, the free thiol can react with any of the four standard thiol-Michael monomers. This could help determine the kinetics of this

reaction so that the timing can be tuned and the final conversion can be determined (hopefully quantitative). It is key that conditions for complete or near complete conversion are found. For example, the overall yield for a 20-mer with 95% coupling efficiencies is only 36% whereas with 99% coupling efficiencies is 82%.

REFERENCES

1 F. Hong, F. Zhang, Y. Liu and H. Yan, Chemical Reviews, 2017, 117, 12584–12640.

2 E. Winfree, F. Liu, L. A. Wenzler, N. C. Seeman, Nature, 1998, 394, 539-544.

3 P. W. K. Rothemund, Nature, 2006, 440, 297–302.

4 T. Gerling, K. F. Wagenbauer, A. M. Neuner and H. Dietz, Science, 2015, 347, 1446–1452.

5 H. C. Kolb, M. G. Finn and K. B. Sharpless, Angewandte Chemie International Edition, 2001, 40, 2004–2021.

6 D. P. Nair, M. Podgórski, S. Chatani, T. Gong, W. Xi, C. R. Fenoli and C. N. Bowman, Chemistry of Materials, 2013, 26, 724–744.

7 C. E. Hoyle and C. N. Bowman, Angewandte Chemie International Edition, 2010, 49, 1540–1573.

8 A. Dondoni and A. Marra, Chem. Soc. Rev., 2012, 41, 573–586.

9 B. B. Brown, D. S. Wagner and H. M. Geysen, Molecular Diversity, 1995, 1, 4–12.

10 H. Li, J.-M. Hah and D. S. Lawrence, Journal of the American Chemical Society, 2008, 130, 10474–10475.

11 S. M. Sternson and S. L. Schreiber, Tetrahedron Letters, 1998, 39, 7451–7454.

12 E. Kaiser, R. L. Colescott, C. D. Bossinger and P. I. Cook, Analytical Biochemistry, 1970, 34, 595–598.

13 C. B. Bottom; S. S. Hanna; D. D, Siehr. Mechanism of the ninhydrin reaction. Biochem. Educ. 1979, 6, 4-5.

14 G. L. Ellman, Archives of Biochemistry and Biophysics, 1959, 82, 70–77.

15 R. B. Merrifield, Journal of the American Chemical Society, 1963, 85, 2149–2154.

16 C. A. Mirkin, R. L. Letsinger, R. C. Mucic, J. J. Storhoff. Nature 1996, 382, 607.

17 H. Abou Assi, M. Garavís, C. González and M. J. Damha, Nucleic Acids Research, 2018, 46, 8038–8056.

18 S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, Nucleic Acids Research, 2006, 34, 5402–5415.

19 M.-A. Shahbazi, T. Bauleth-Ramos and H. A. Santos, Advanced Therapeutics, 2018, 1, 1800042.

20 C. Jin, X. Liu, H. Bai, R. Wang, J. Tan, X. Peng and W. Tan, ACS Nano, 2017, 11, 12087–12093.

21 P. Nielsen, M. Egholm, R. Berg and O. Buchardt, Science, 1991, 254, 1497–1500.

22 C. Avitabile, L. Moggio, L. D. D'Andrea, C. Pedone and A. Romanelli, Tetrahedron Letters, 2010, 51, 3716–3718.

23 N. Umezawa, Y. Noro, K. Ukai, N. Kato and T. Higuchi, ChemBioChem, 2011, 12, 1694–1698.

24 K. N. Ballantyne, R. A. H. van Oorschot and R. J. Mitchell, Genomics, 2008, 91, 301–305.

25 W. Xi, S. Pattanayak, C. Wang, B. Fairbanks, T. Gong, J. Wagner, C. J. Kloxin and C. N. Bowman, Angewandte Chemie International Edition, 2015, 54, 14462–14467.

26 R. Behrendt, P. White and J. Offer, Journal of Peptide Science, 2016, 22, 4–27.

27 M. H. Caruthers, Biochemical Society Transactions, 2011, 39, 575–580.

28 R. A. Hughes and A. D. Ellington, Cold Spring Harbor Perspectives in Biology, 2017, 9, a023812.

Appendix A

SUPPORTING DATA

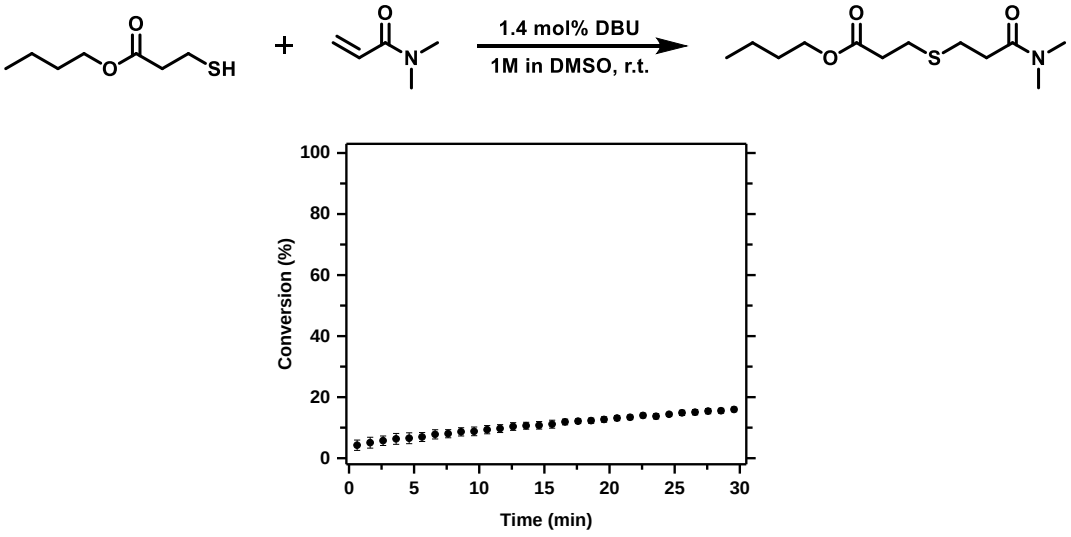


Figure A1: Conversion profile of a model thiol and a model acrylamide

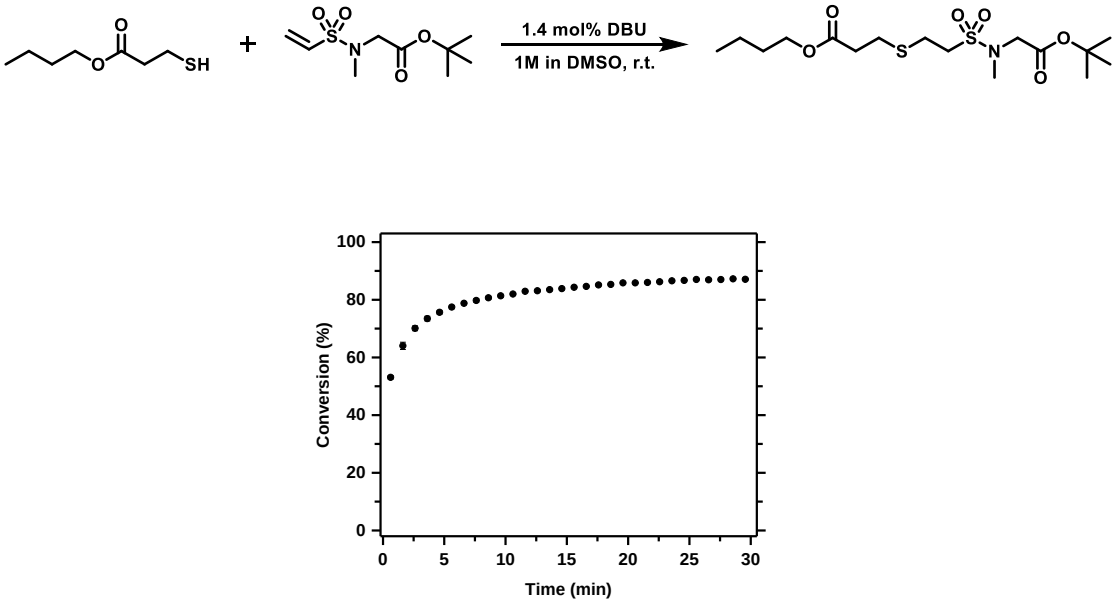


Figure A2: Conversion profile of a model thiol and a model vinyl sulfonamide

Appendix B

GENERAL PROCEDURES

B.1 Ellman's Test for Free Thiols

Ellman's reagent[14], 5,5'-dithiobis-(2-nitrobenzoic acid), is used to qualitatively or quantitatively determine if free thiol functional groups are present during the oligomer synthesis. Ellman's reagent reacts with a thiol to yield 2-nitro-5-thiobenzoate or TNB^{-2} which is a bright yellow color. In order to carry out the test, a handful of resin beads are transferred to a small vial. 100 μL of Ellman's reagent, 100 μL of DCM, and one drop of triethylamine are then added. A yellow color is a positive test and indicates the presence of free thiols whereas a colorless solution is a negative test and indicates that no thiols are present.

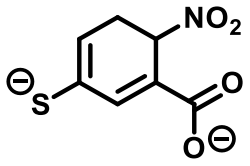


Figure B1: 2-nitro-5-thiobenzoate or TNB^{-2}

B.2 Kaiser Test for Free Amines

The Kaiser test[12] is used qualitatively to determine quickly if successive coupling to amines has been successful. During the cycle of adding monomers and removing the protecting group, free amines are absent and then present, respectively. The Kaiser test is used to verify that either a coupling reaction or a deprotection reaction has gone to completion. The test is carried out by transferring a handful of solid phase resin beads to a small vial followed by the addition of three drops each of

three solutions: 500 mg ninhydrin in 10 mL ethanol, 40 g phenol in 10 mL ethanol, and pyridine. The mixture was subsequently heated to 100 °C using a heat gun for 20 seconds. A dark blue color is a positive test for free amines and a yellow color is a negative test. The dark blue color results from the compound referred to as Ruhemann's purple which is formed when ninhydrin reacts with free amines shown below in Figure B2[13].

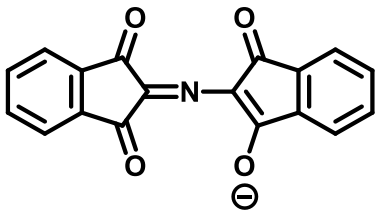


Figure B2: Ruhemann's purple

B.3 Amino Acid Coupling Protocol

The Fmoc protected resin was first swelled in DCM upon removal from the refrigerator. The DCM was removed and the resin was washed with DMF. The resin was then vortexed in a solution of 20% piperidine in DMF for 15 min twice in order to remove the Fmoc. After washing away the Fmoc, a Kaiser test for free amines was carried out to ensure complete Fmoc deprotection. Next, the amino acid (either Fmoc-Lys(Boc)-OH, Fmoc-Lys(Alloc)-OH, or Fmoc-Cys(Mmt)-OH) was added to a separate vial followed by 4 equivalents of HBTU, DMF and 8 equivalents of NMM (N-Methylmorpholine). This solution was transferred to the resin and it reacted on the vortexer for 30 min using two couplings. The unreacted residues and by-products were then washed away using DMF.

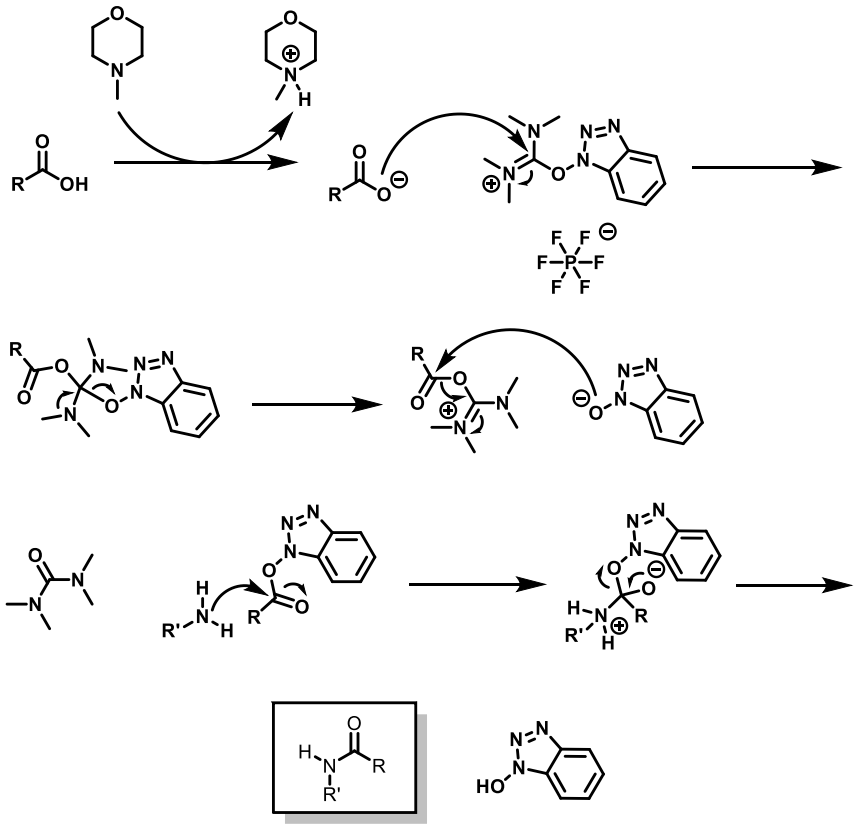


Figure B3: Mechanism of HBTU amide coupling

B.4 Monomer Coupling Protocol

The thiol-Michael monomer is added to a separate vial followed by DMF and DBU. The solution is well mixed and is transferred to the resin. The coupling takes one hour to complete. The exact equivalences of monomer and DBU in addition to the number of coupling events are variable are detailed in the main text for each experiment.

B.5 Coupling ANP Linker to TentaGel S NH_2 Resin

The resin (100 mg, 0.026 mmol functionality) was first swelled in DCM for 15 min and a Kaiser test was performed to ensure a free amine was present. The resin was

then swelled in DMF for 5 min. In a separate vial, the ANP linker (4 eq) was dissolved in 3 mL DMF. HBTU (39.4 mg, 104 μmol) was added to the separate vial followed by NMM (22.9 μL, 208 μmol) and it was vortexed. The separate vial's contents were transferred to the peptide synthesis vessel and the reaction proceeded for 30 min while vortexing. This coupling procedure was then repeated two more times to ensure near quantitative conversion of the free amine.

B.6 Resin Microcleavage Protocols

B.6.1 TentaGel S NH_2 Resin

Dry beads that had previously been swelled in DCM were transferred to a 1.5 mL centrifuge tube. 1 mL of a solution of 4:1 water:methanol was then added to the tube and it was vortexed.[9] The tube was subsequently irradiated at 20 mW/cm^2 using 365 nm light in an amber light filtered room. After one hour, the now dark yellow solution was removed.

B.6.2 TentaGel S RAM Resin

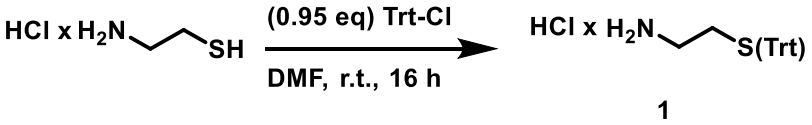
Dry beads that had previously been swelled in DCM were transferred to a 1 dram glass vial. 0.5 mL trifluoroacetic acid was added and the vial vortexed for one hour. The TFA was then evaporated by running nitrogen over the vial and the residue was resuspended in 0.5 mL DCM. The DCM was subsequently evaporated to remove any remnants of TFA. Finally, 0.5 mL of methanol were added to the vial to prepare for characterization.

B.7 Characterization of Oligomers via Mass Spectrometry

Following microcleavage procedures, 500 µL the solution containing the resin beads and oligomer product were transferred to a Thomson filter vial with a 0.45µm pore size in order to prevent the beads from entering the machine. The vial was then placed in an Acquity UPLC H-Class/SQD2 machine. The machine then injects the sample and runs a UltraPerformance Liquid Chromatography (UPLC) separation on the injection running a gradient of acetonitrile (ACN) in water from 5% ACN to 100%. A mass spectrum is obtained for each interval along the UPLC trace.

Appendix C

SYNTHETIC PROCEDURES FOR THIOL-MICHAEL MONOMERS



1

2-(Tritylthio)ethylamine hydrochloride (Backbone): Cysteamine hydrochloride (87 mmol, 10 g) was added to an oven dried and nitrogen purged 500 mL round-bottom flask followed by 75 mL of anhydrous DMF. Trityl chloride (82.7 mmol, 23 g) was subsequently added and the mixture was left stirring for 24 hours. Ethyl acetate (100 mL) was added to the round bottom flask and approximately 300 mL of distilled water was added to the flask. The solution was then transferred to a 1 L beaker and water was added to fill the beaker. A white solid precipitated and was filtered using a medium fret. The cake was then washed with 100 mL of diethyl ether to afford the product as a white powder, hydrochloric acid salt. The material was further dried in a desiccator overnight to give compound **1** with over 99% yield (30.9 g). The product (backbone) was used in later synthesis without further purification.

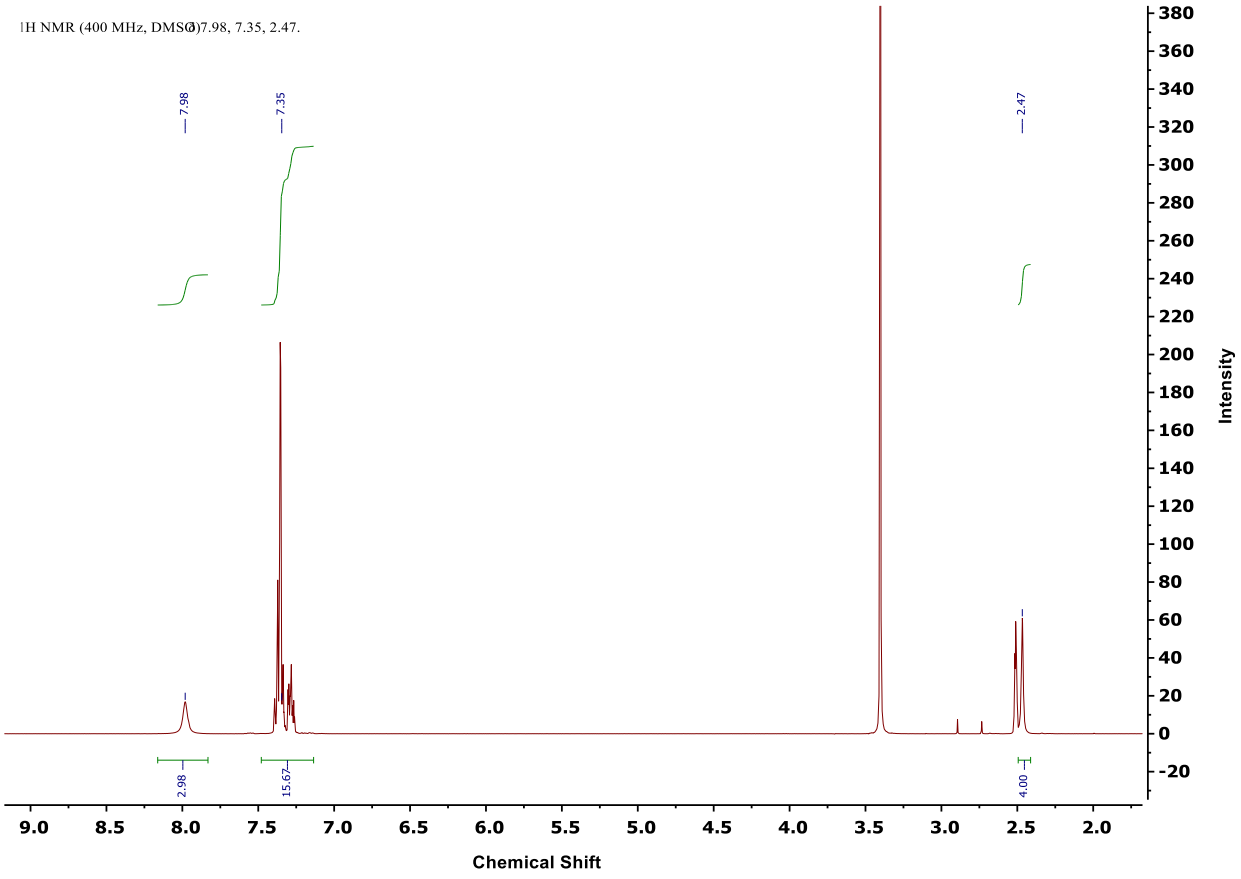


Figure C1: HNMR spectrum of compound **1**

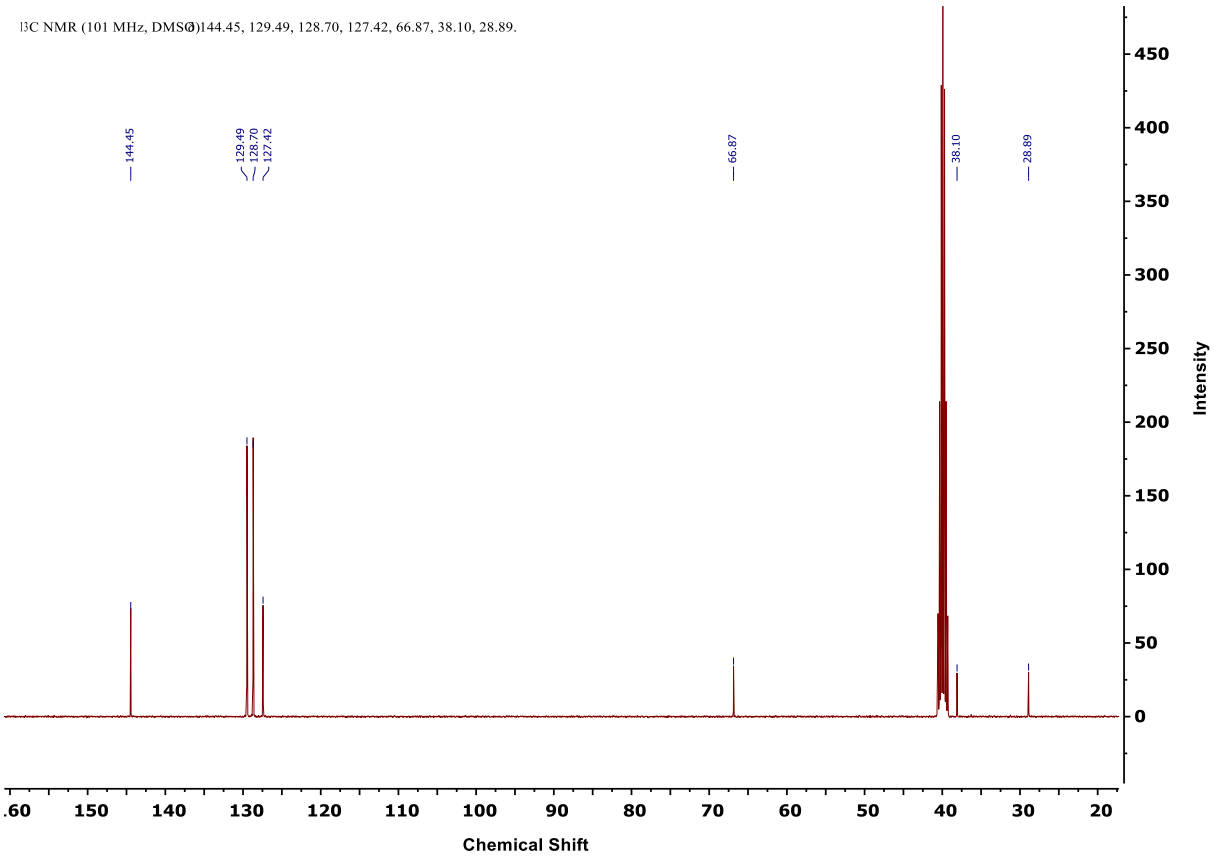
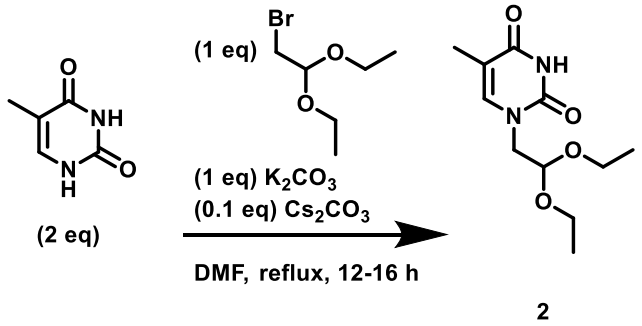


Figure C2: CNMR spectrum of compound **1**



2

1-(2,2-diethoxyethyl)-5-methyl-1,2,3,4-tetrahydropyrimidine-2,4-dione: To an oven-dried, nitrogen purged round bottom flask was added thymine (159 mmol, 20 g, 2 eq), potassium carbonate (79.4 mmol, 10.96 g, 1 eq), and cesium carbonate (7.94 mmol, 2.58 g, 0.1 eq). Anhydrous DMF (150 mL) was added and allowed to mix for 10 min. Next, bromoacetaldehyde diethyl acetal (79.4 mmol, 11.94 mL, 1 eq) was added in one portion. The flask was then refluxed overnight at 130 ℃ while stirring. The following day, the solution was filtered while hot to remove unreacted thymine and salts. The cake was washed with minimal DMF before being concentrated via rotary evaporation. Once most of the DMF was removed, the residual oil was dissolved in ethyl acetate (300 mL) and extracted with water (5x 150 mL) to remove residual DMF. The organic layer was collected, dried over Na_2SO_4, filtered, and evaporated to dryness to produce a viscous oil that crystallized upon standing in a fridge. After 8 h, the solid was triturated with 2:1 petroleum ether to ether and filtered. The cake was washed with minimal 2:1 petroleum ether to ether, collected, and dried *in vacuo* to produce compound **2** (9.02 g, 47%).

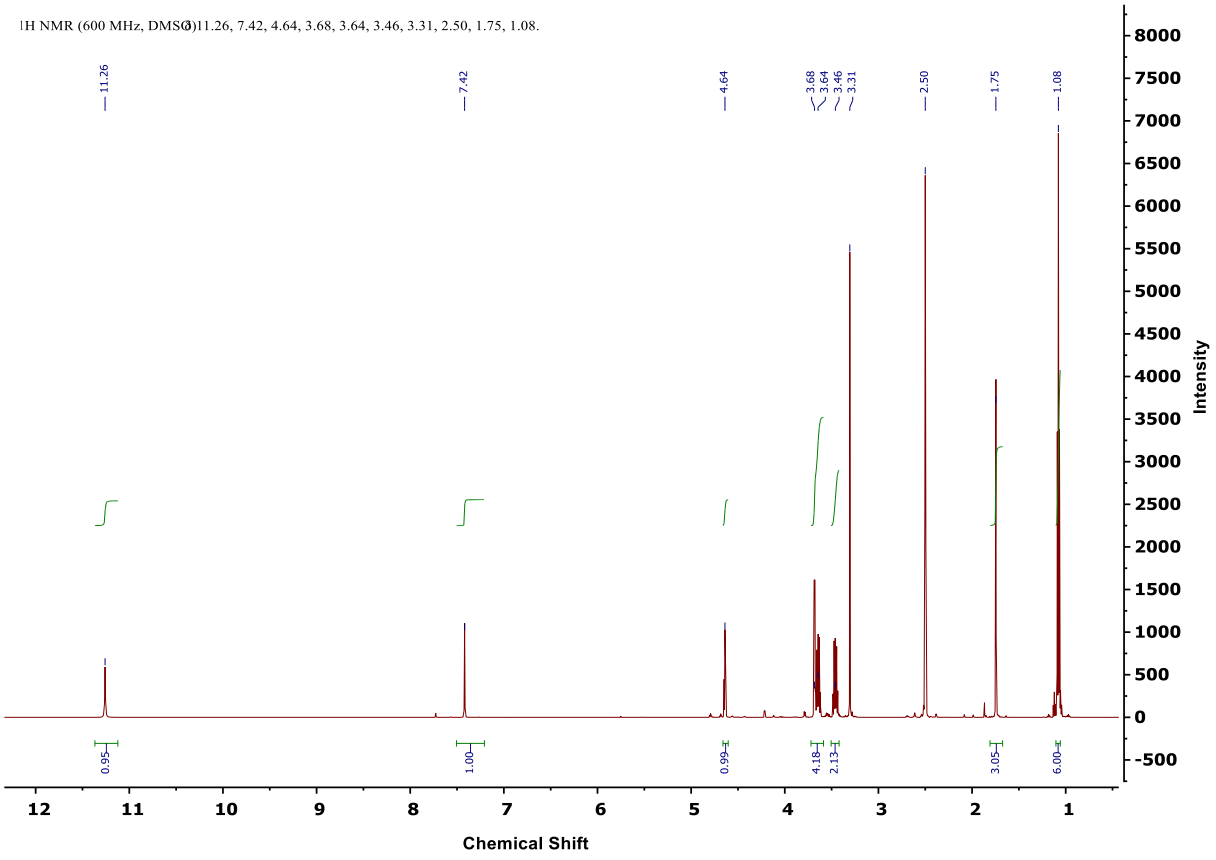


Figure C3: HNMR spectrum of compound **2**

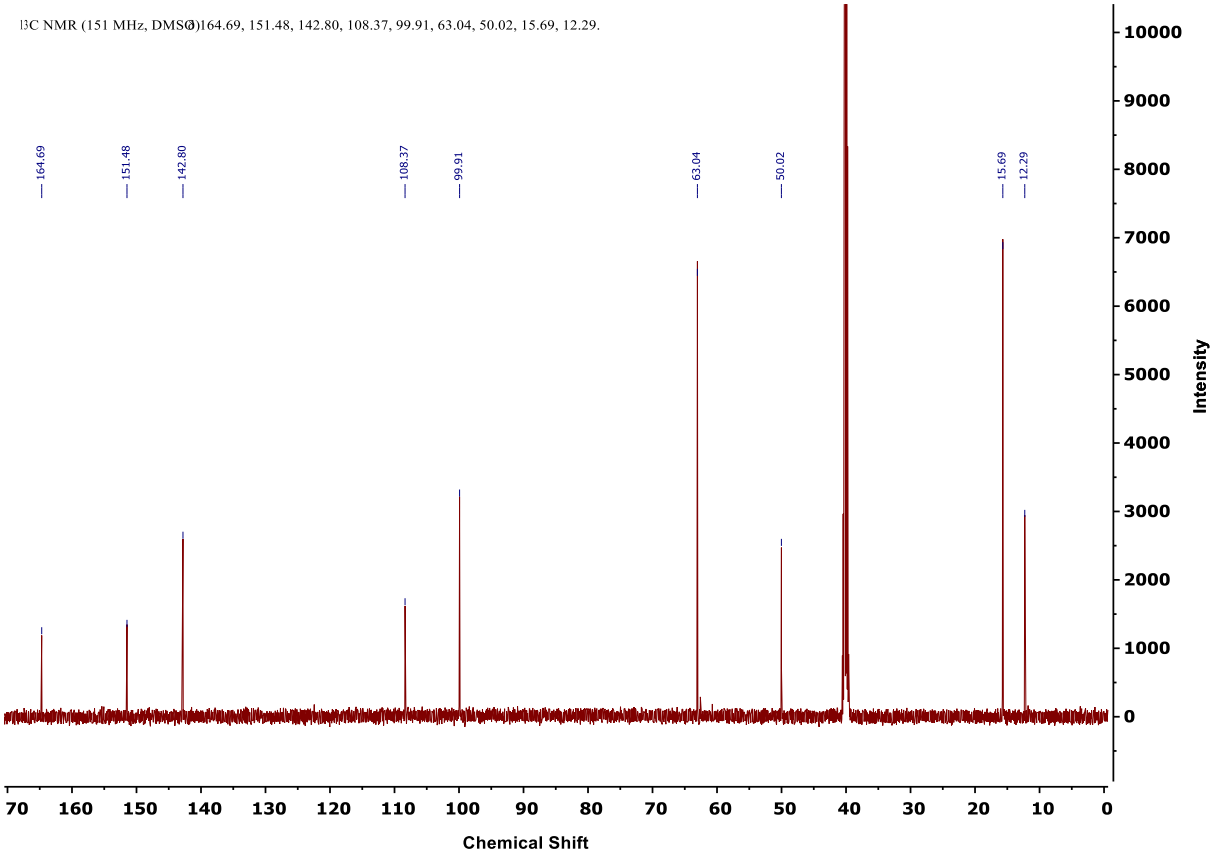
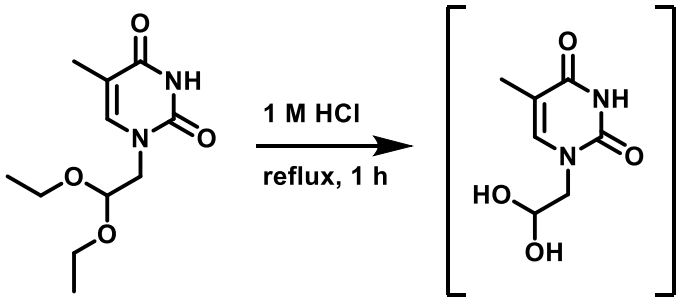


Figure C4: CNMR spectrum of compound **2**



2 3

1-(2,2-dihydroxyethyl)-5-methyl-1,2,3,4-tetrahydropyrimidine-2,4-dione: Compound **2** (10 g, 41.3 mmol, 1 eq) was added to a 500 mL round bottom flask followed by 1 M HCl (100 mL). The solution was then refluxed for 1 h. Following the completion of the reaction, the solution was evaporated to dryness via rotary evaporation producing a white solid. To the white solid was added THF (50 mL) and cooled in an ice bath for 1 h before being filtered and subsequently washed with diethyl ether. Drying in vacuo produced 7.6 g (98 %) of the hydrated aldehyde **3** which was used in the next step without further purification.

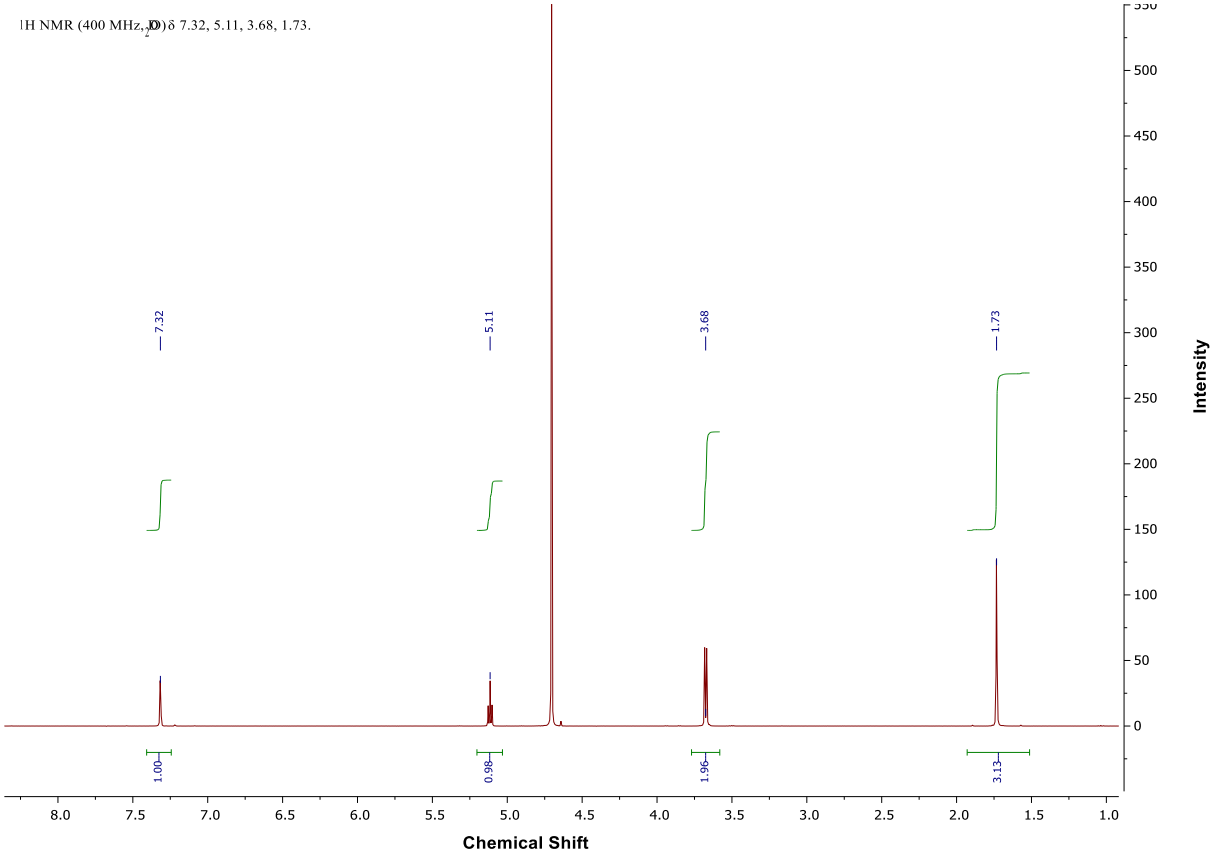


Figure C5: HNMR spectrum of compound **3**

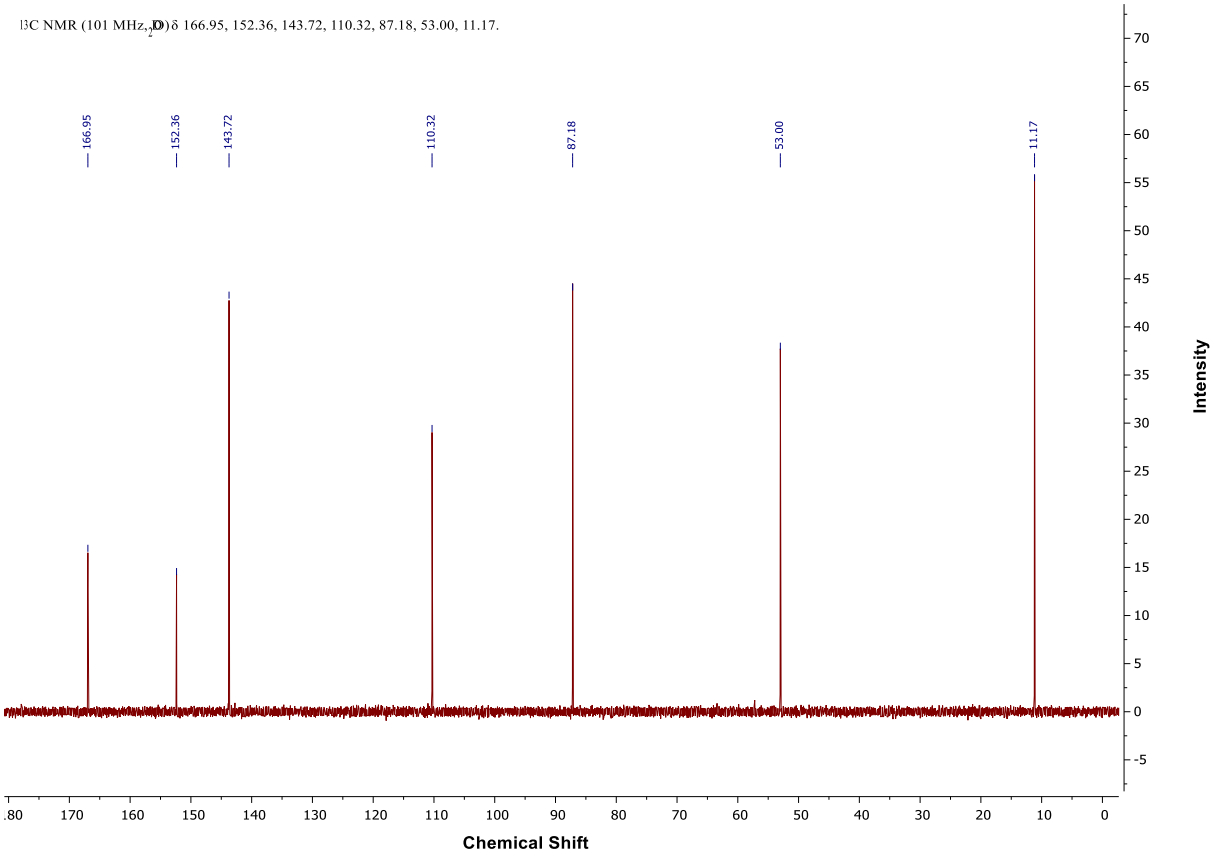
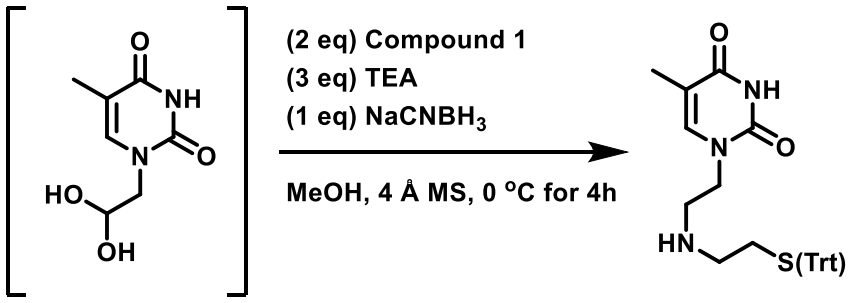


Figure C6: CNMR spectrum of compound **3**



3 4

5-methyl-1-[2-({2-[(triphenylmethyl)sulfanyl]ethyl}amino)ethyl]-1,2,3,4-tetrahydropyrimidine-2,4-dione: To an oven dried nitrogen purged 500 mL round bottom flask was added 4 Å molecular sieves. Compound **1** (11.4 g, 32.2 mmol, 2 eq**)** was added to the round bottom flask followed by 200 mL of anhydrous methanol. Anhydrous triethylamine (6.74 mL, 48.4 mmol, 3 eq) and compound **3** (3 g, 16.1 mmol, 1 eq) were added sequentially. The reaction was cooled in an ice bath before sodium cyanoborohydride (1g, 16.1 mmol, 1 eq) was added to the solution. The reaction was monitored by TLC for 4 h prior to filtering. Once filtered, the cake was washed with 50 mL of methanol and 100 mL of DCM. The eluent was then collected and transferred to a separatory funnel. An additional 300 mL of DCM and 400 mL of water were added and the organic layer was collected. The aqueous layer was washed with DCM (2x, 100 mL) and the organic layers were combined, dried over magnesium sulfate, and filtered. The solution was then concentrated and purified by flash chromatography with a gradient of 0 to 5% MeOH in DCM to afford the compound **4** as a white solid (4.62 g, 61%).

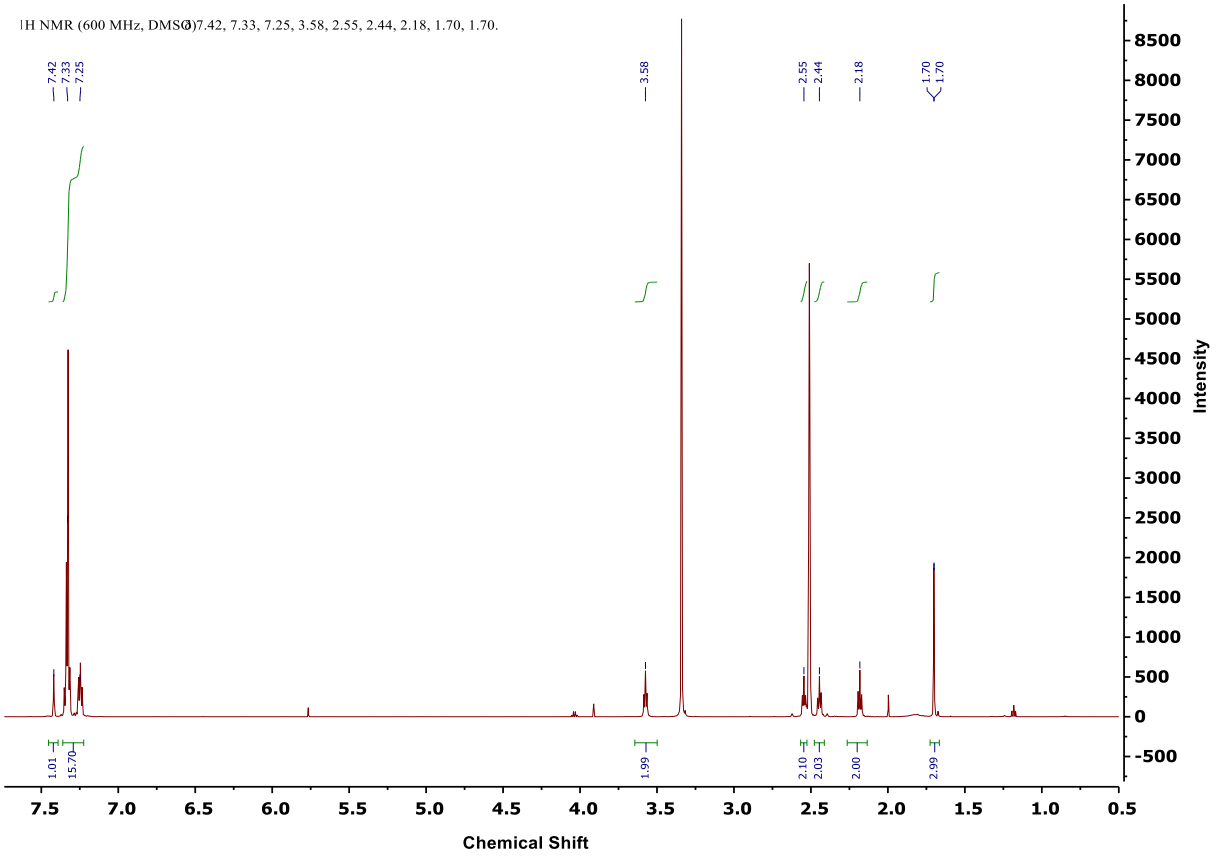


Figure C7: HNMR spectrum of compound **4**

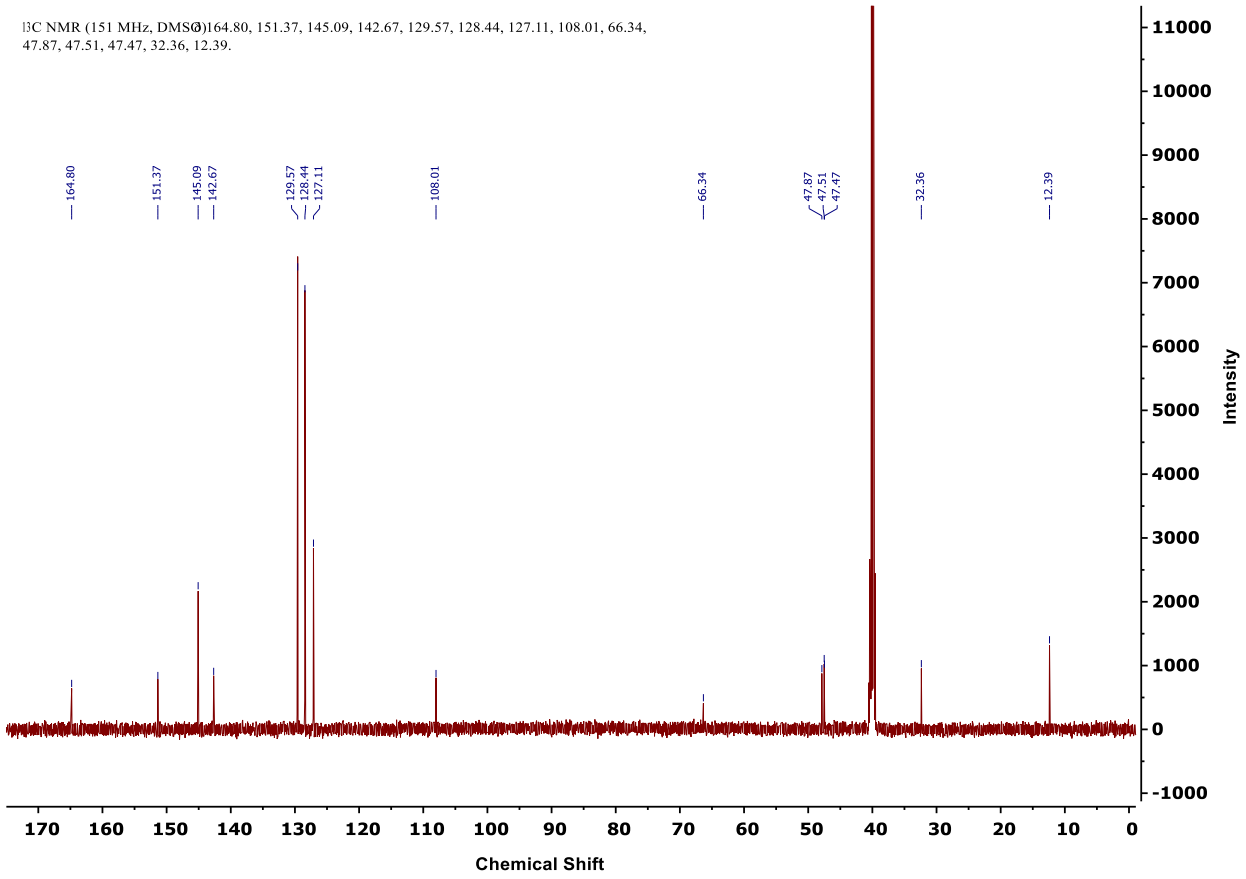
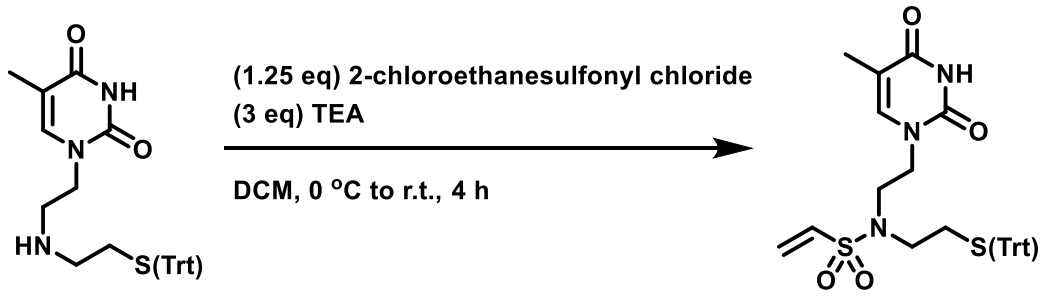


Figure C8: CNMR spectrum of compound **4**



4 5

N-[2-(5-methyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-1-yl)ethyl]-N-{2-[(triphenylmethyl)sulfanyl]ethyl}ethene-1-Sulfonamide: To an oven dried, nitrogen purged round bottom flask was added compound **4** (6.5 g, 13.8 mmol, 1 eq) followed by 100 mL of DCM. Once compound **4** dissolved, the round bottom flask was placed in an ice bath and allowed to cool for 15 min. Next, anhydrous triethylamine (5.76 mL, 41.4 mmol, 3 eq) was added followed by dropwise addition of 2-chloroethylsulfonyl chloride (1.802 mL, 17.25 mmol. 1.25 eq) over the course an hour. Once the addition was finished, the ice bath was removed and the reaction was allowed to stir at room temperature for 4 h. 100 mL of water was added to quench the reaction and the mixture was transferred to an extraction vessel. An additional 200 mL of DCM was added and the organic layer was washed with water (2x 100 mL) followed by drying over Na_2SO_4, the Na_2SO_4 was filtered and the organic solution concentrated before being purified by flash chromatography with a gradient of 0% to 5% methanol in DCM to afford product **5** (5.19 g, 67%) as a tan solid.

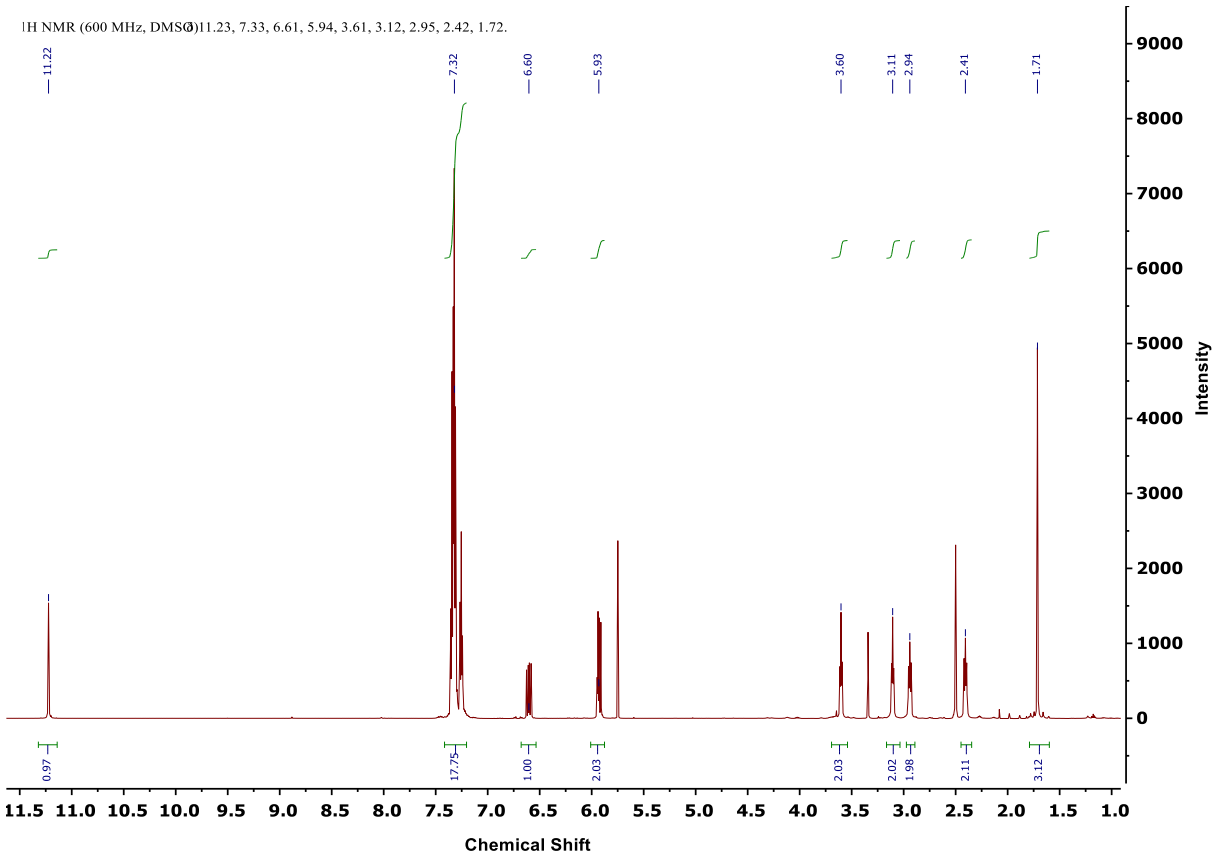


Figure C9: HNMR spectrum of compound **5**

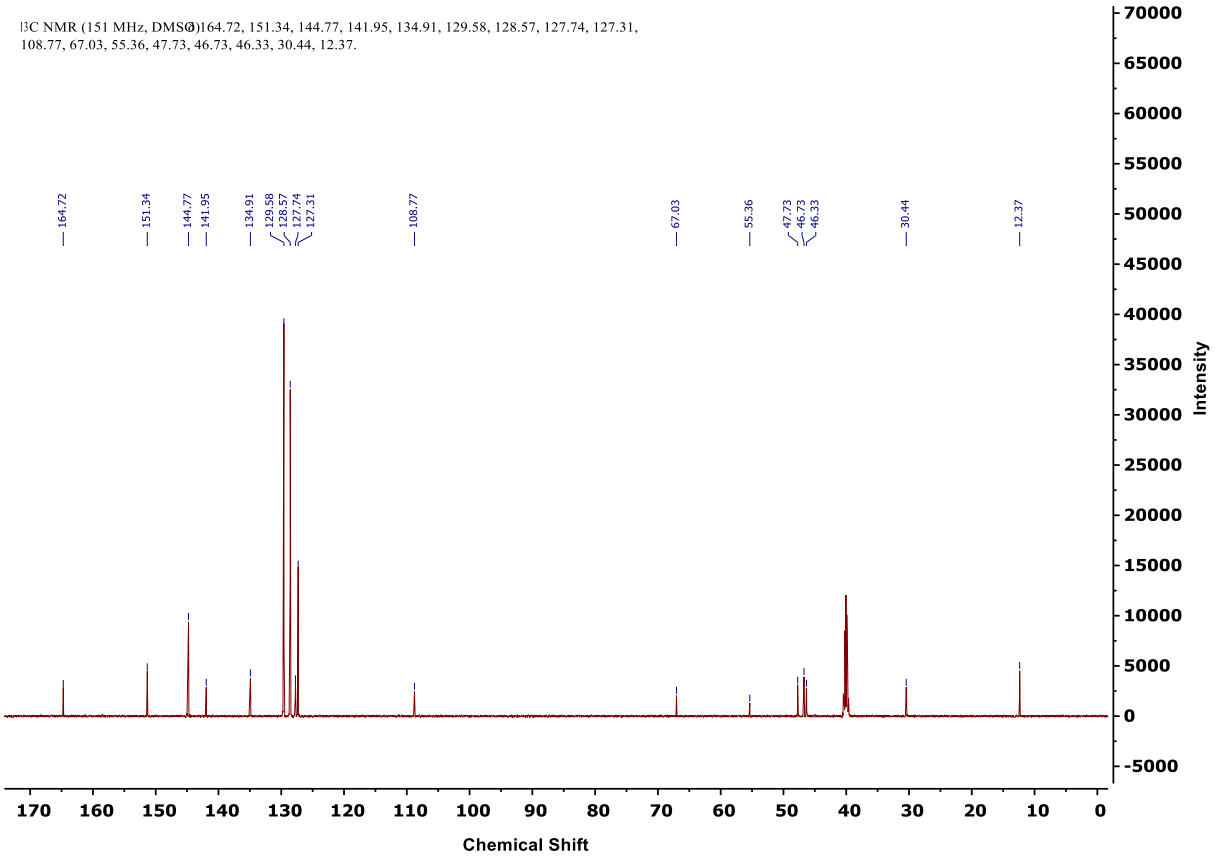
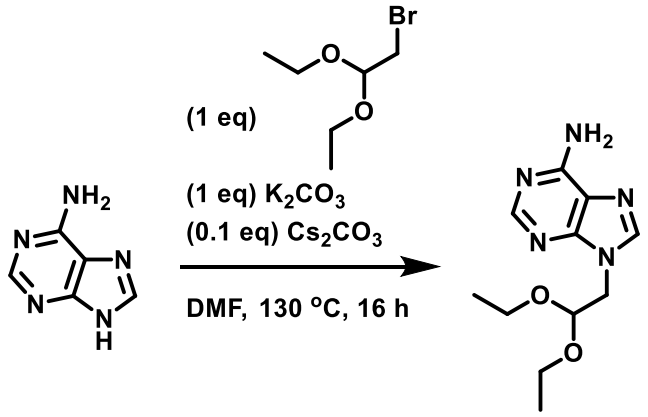


Figure C10: CNMR spectrum of compound **5**



6

9-(2,2-diethoxyethyl)-9H-purin-6-amine: Adenine (74.1 mmol, 10g, 1 eq) was added to an oven-dried round bottom flask that had been purged with nitrogen gas. Anhydrous potassium carbonate (74.1 mmol, 10.2 g, 1 eq) and cesium carbonate (7.41 mmol, 2.41 g, 0.1 eq) were added followed by 150 mL anhydrous DMF. After stirring for 15 minutes, bromoacetaldehyde diethyl acetal (74.1 mmol, 11.14 mL, 1 eq) was added in one portion. The reaction was then refluxed at 130 ℃ for 16 h. Upon completion of the reaction, the flask contents were filtered while hot. The cake was washed with an additional 50 mL of DMF and the solvent was subsequently evaporated using a rotary evaporator. Once dry, 200 mL of ethanol was added to the flask. The flask was stirred at 85 ℃ until all product was dissolved, leaving only insoluble salts as the precipitate. The solution was then filtered and the eluent was refrigerated overnight to recrystallize the product. The solid was filtered once more and the cake was washed with 100 mL of ethanol and 250 mL of diethyl ether. The product dried under *in vacuo* to produce compound **6** as a white solid (9.9 g, 53%).

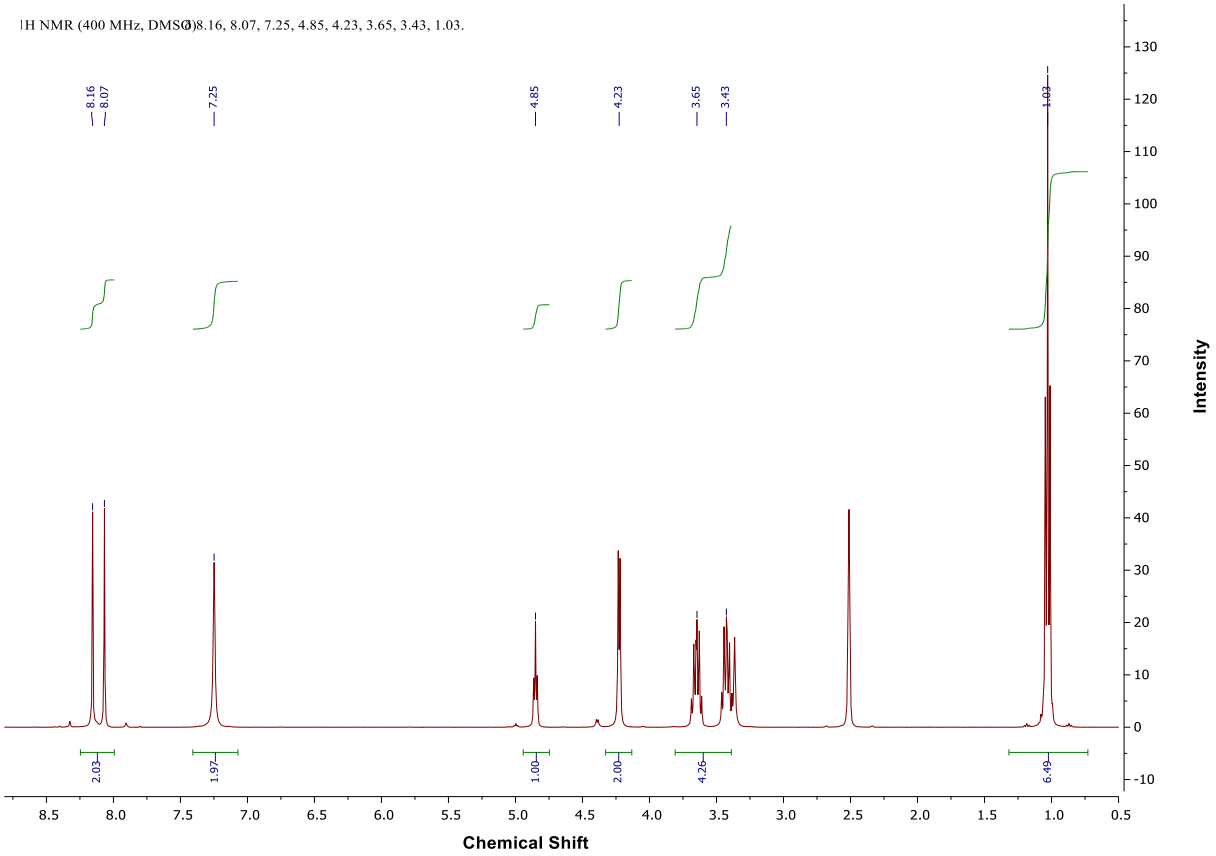


Figure C11: HNMR spectrum of compound **6**

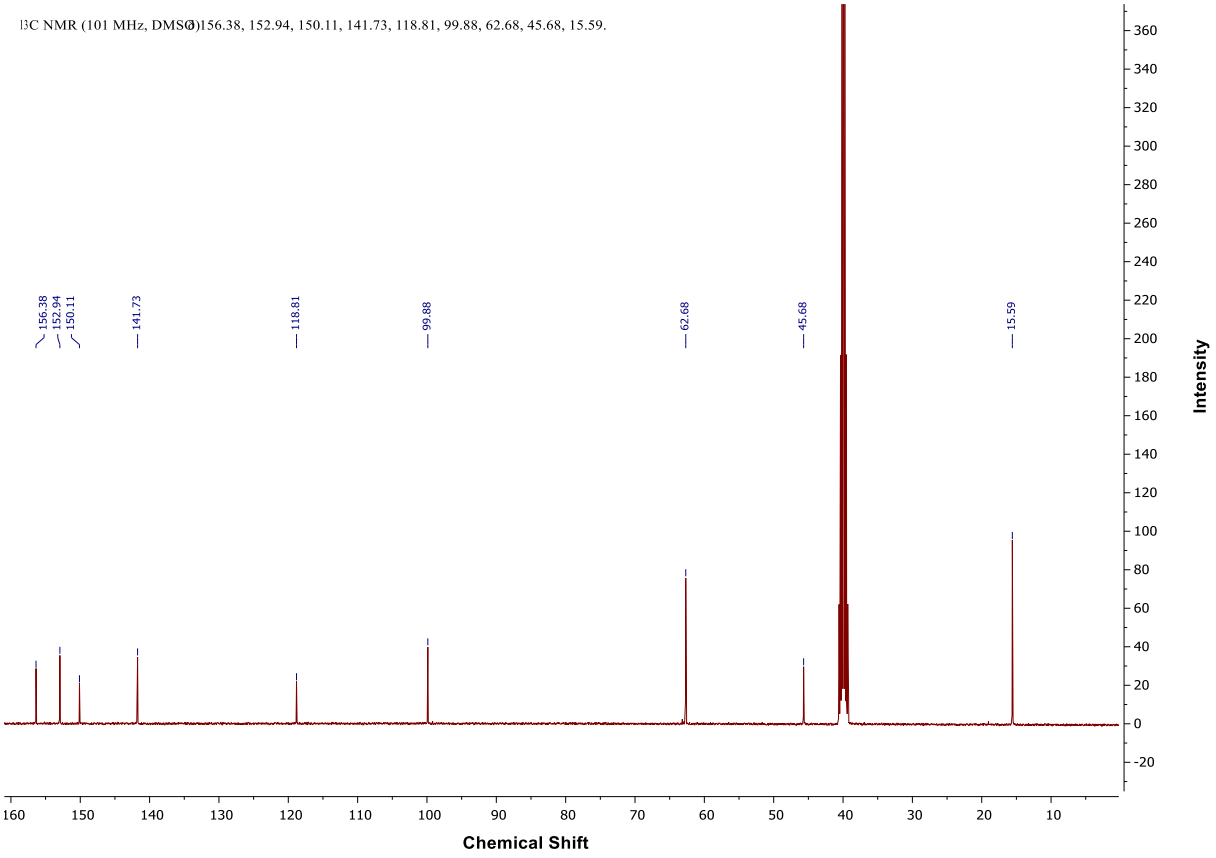
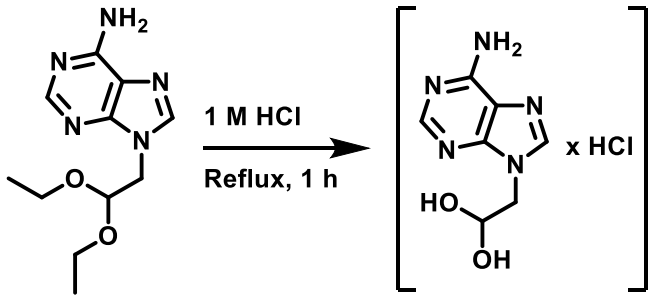


Figure C12: CNMR spectrum of compound **6**



6 7

2-(6-amino-9H-purin-9-yl)ethane-1,1-diol hydrochloride: Compound **6** (8 g, 31.8 mmol) was added to an oven-dried round bottom flask that had been purged with nitrogen gas. 100 mL of 1 M HCl was then added and the round bottom flask was refluxed for 1 h at 100 ℃ in the dark. While still in the dark, the water was then removed via rotary evaporation to give an off-white solid. To the solid was added 125 mL of tetrahydrofuran and the mixture was cooled in an ice bath for 1 h. The solid was filtered and washed with diethyl ether. The solid was dried *in vacuo* in the dark to afford compound **7** as an off-white solid (6.97 g, 94%). The product was found to be isolated as the hydrate hydrochloride salt and was used without any further purification.

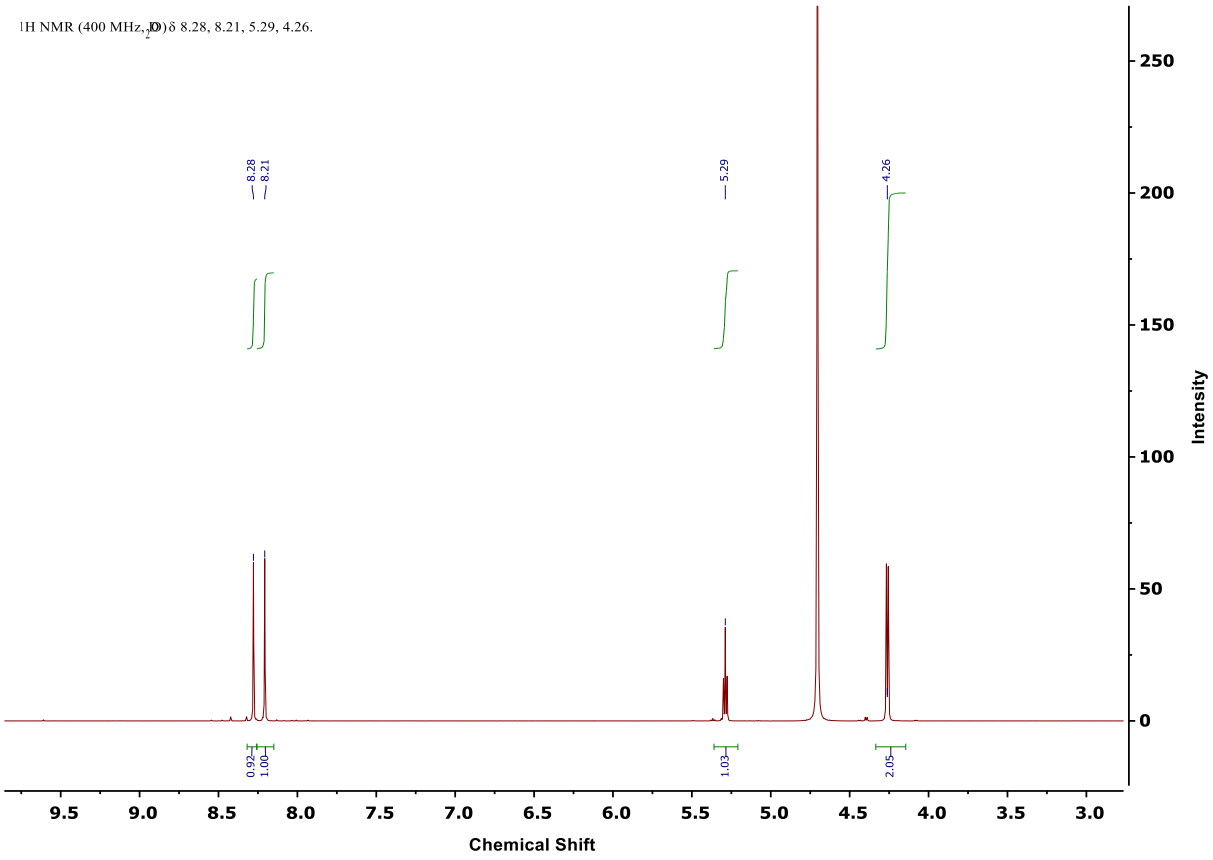


Figure C13: HNMR spectrum of compound **7**

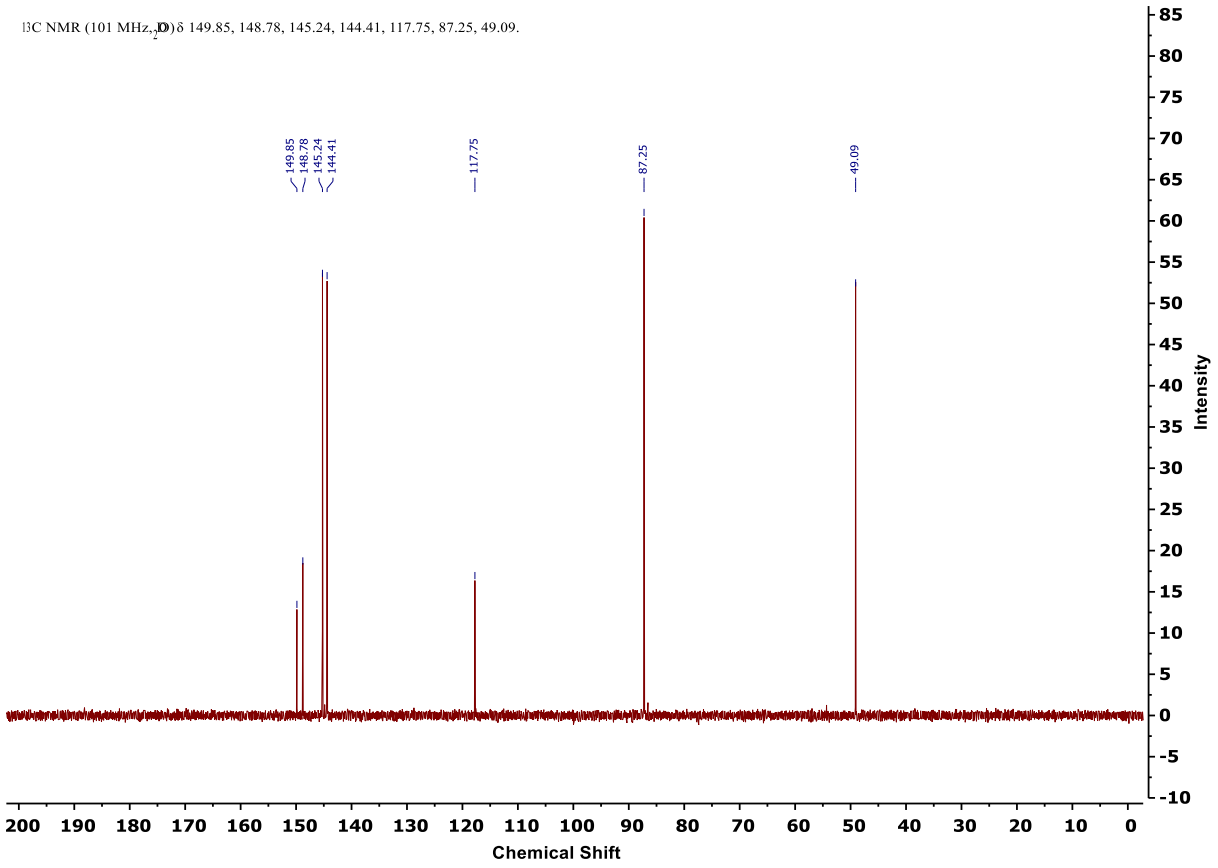
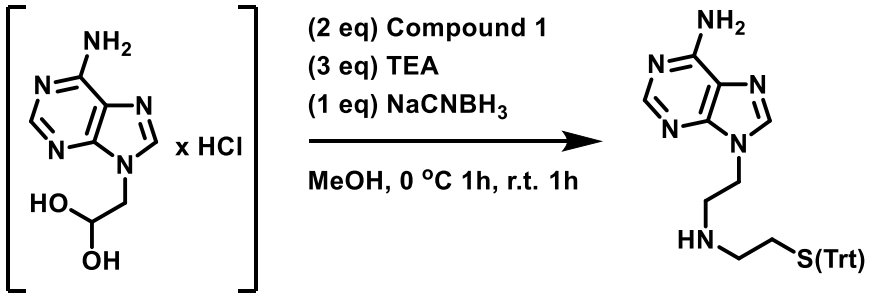


Figure C14: CNMR spectrum of compound **7**



7 8

9-[2-({2-[(triphenylmethyl)sulfanyl]ethyl}amino)ethyl]-9H-purin-6-amine: A 250 mL round bottom flask was oven dried and nitrogen purged. Molecular sieves, grade 514, were added to cover the bottom of the flask. Compound **7** (5.53 g, 15.6 mmol) was added followed by anhydrous methanol (100 mL). Anhydrous triethylamine (6.52 mL, 46.8 mmol) was added and allowed to stir for 30 min. The flask was then placed in an ice bath and sodium cyanoborohydride (0.97 g, 15.6 mmol) was added. The flask warmed to room temperature and reacted overnight. The next day, 150 mL of DCM was added to the flask and it was filtered. The mixture was then extracted with 300 mL of water and an additional 150 mL of DCM. The aqueous layer was washed twice with 100 mL of DCM and the DCM layers were collected and dried over anhydrous sodium sulfate. The DCM was then filtered and dry loaded for flash chromatography. A gradient of 0% to 10% methanol in DCM was used to purify the product giving a yield of 3.30 g (53%).

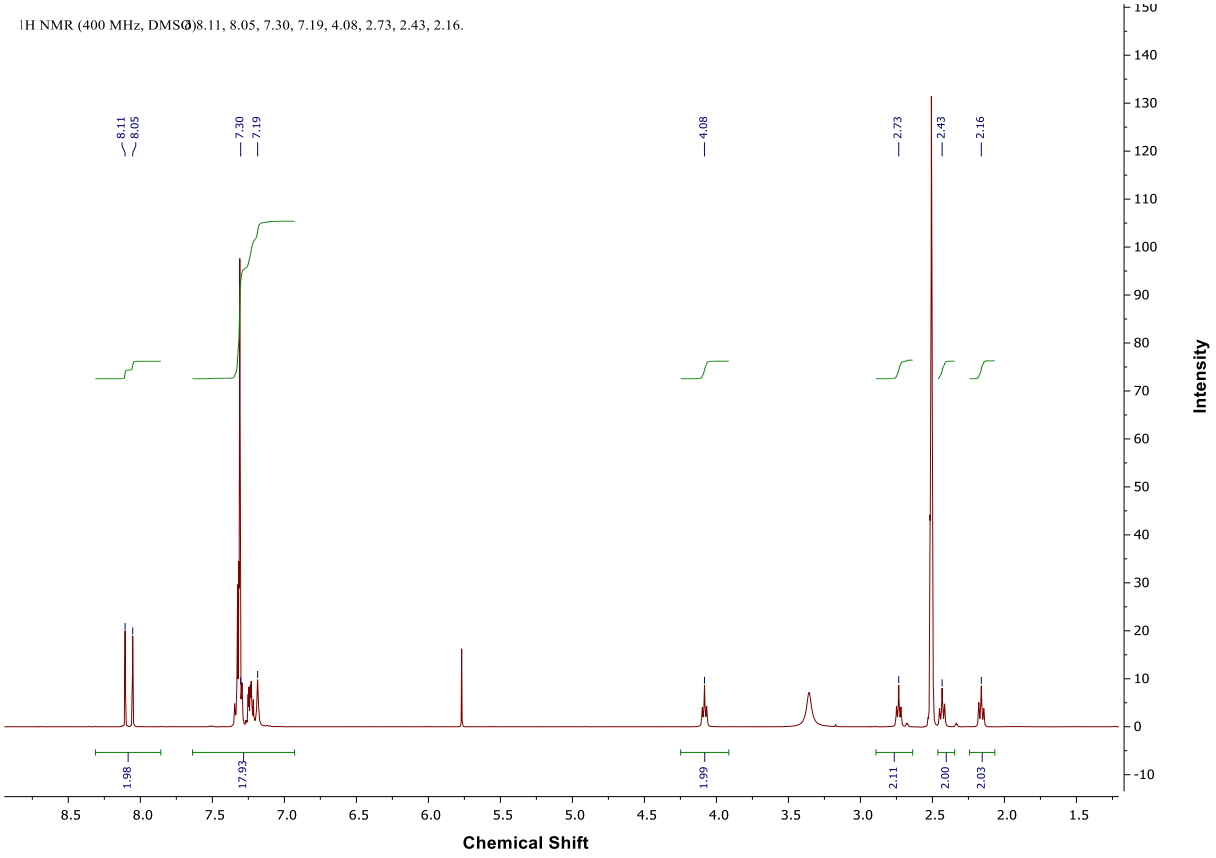


Figure C15: HNMR spectrum of compound **8**

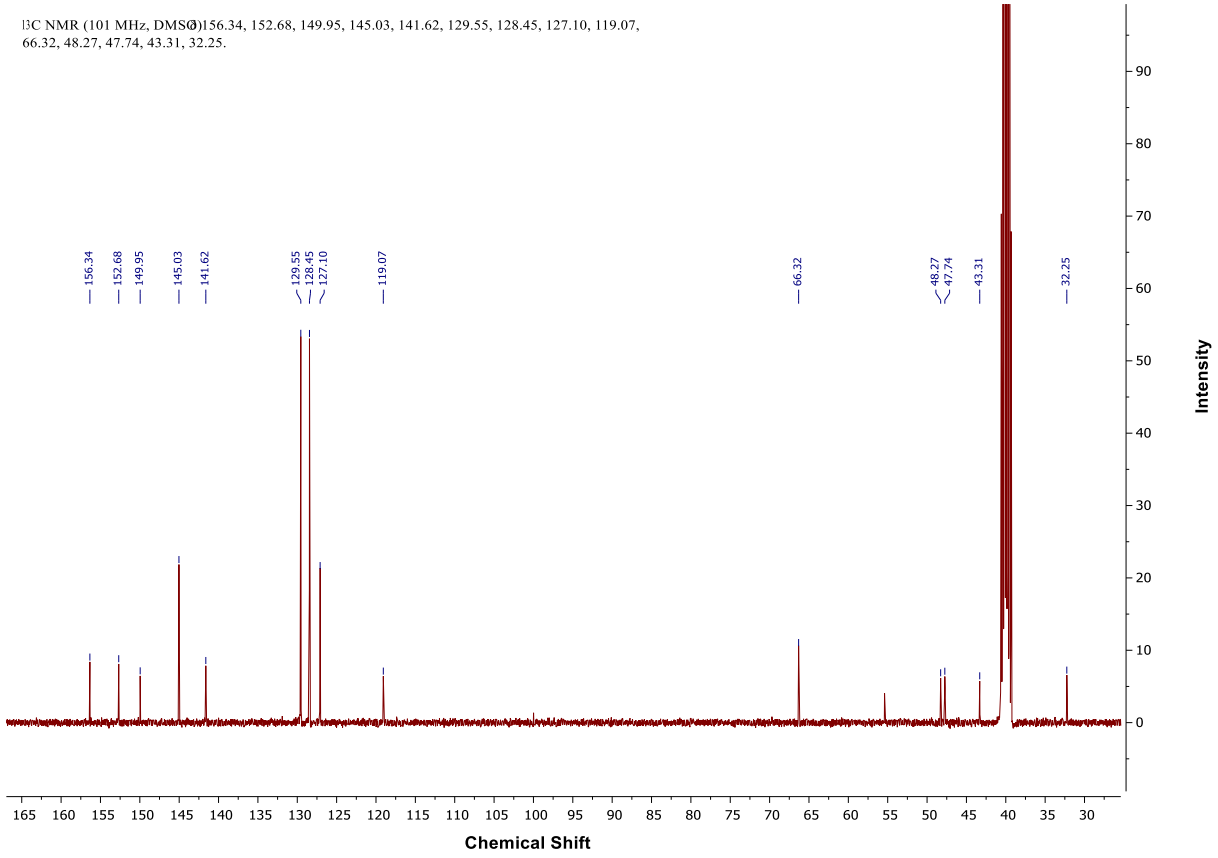
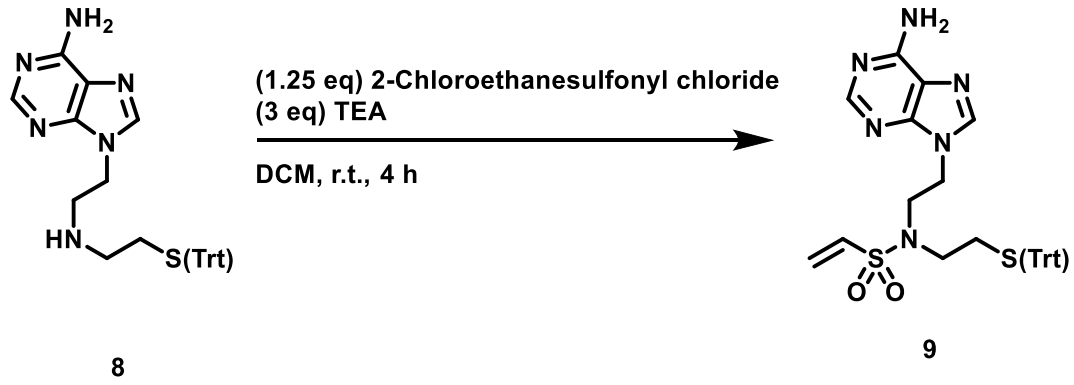


Figure C16: CNMR spectrum of compound **8**



8 9

N-[2-(6-amino-9H-purin-9-yl)ethyl]-N-{2-[(triphenylmethyl)sulfanyl] ethyl}ethene-1-sulfonamide: A 500 mL round bottom flask was oven dried and nitrogen purged. Compound **8** (6.36 g, 13.3 mmol) was added followed by 150 mL of anhydrous DCM. Anhydrous triethylamine (5.54 mL, 39.8 mmol) was added and the flask was placed in an ice bath for 15 minutes. 2-Chloroethanesulfonyl chloride (1.66 mL, 15.9 mmol) was added dropwise and the mixture warmed to room temperature over four hours. The next day, the mixture was transferred to an extraction vessel. 150 mL of DCM, a small amount of saturated sodium chloride solution, and 300 mL of water were added to the vessel. The organic layer was washed with 300 mL of a saturated sodium chloride solution. An additional 100 mL of DCM and a small amount of ethanol were added to assist in emulsion separation. The DCM was collected and dried over anhydrous sodium sulfate. The sample was then purified using flash chromatography with a gradient from 0% to 4% methanol in DCM. 2.75 g of compound **9** was collected with a 36% yield.

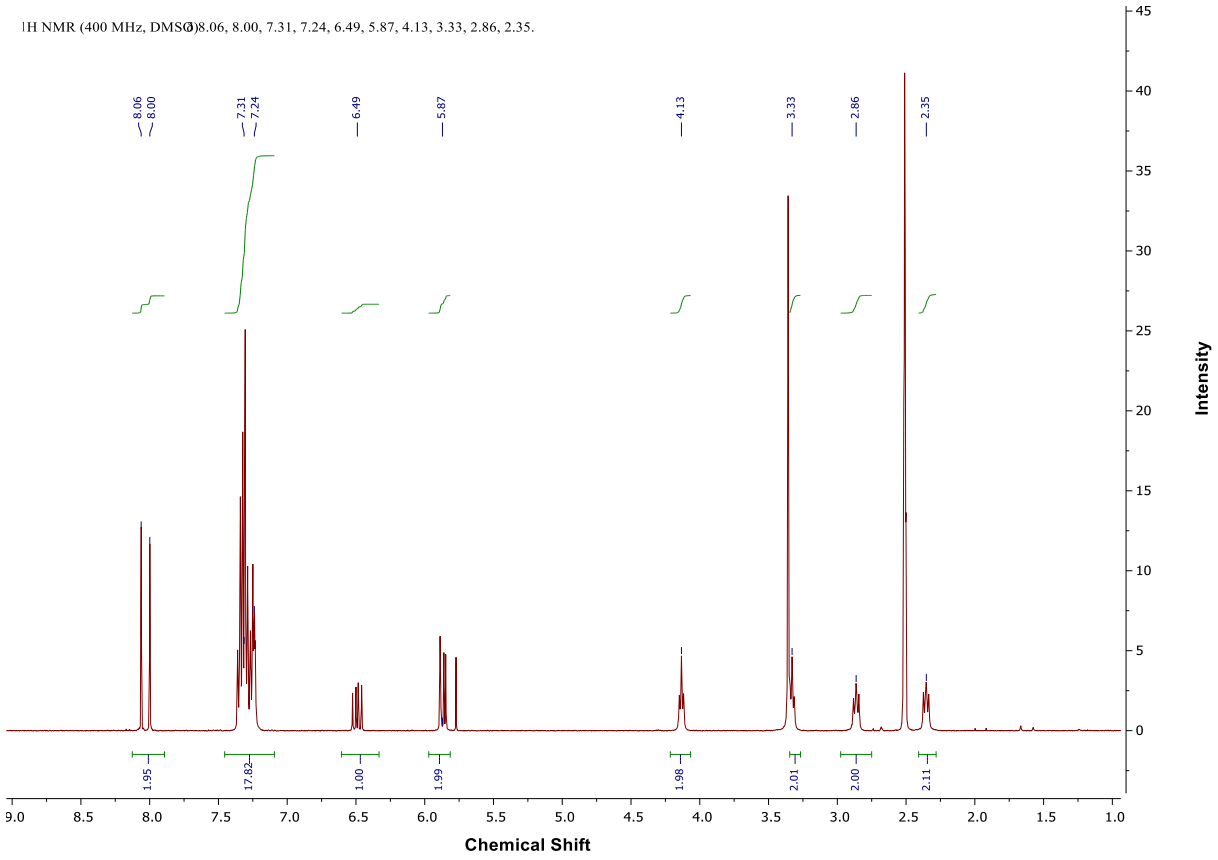


Figure C17: HNMR spectrum of compound **9**

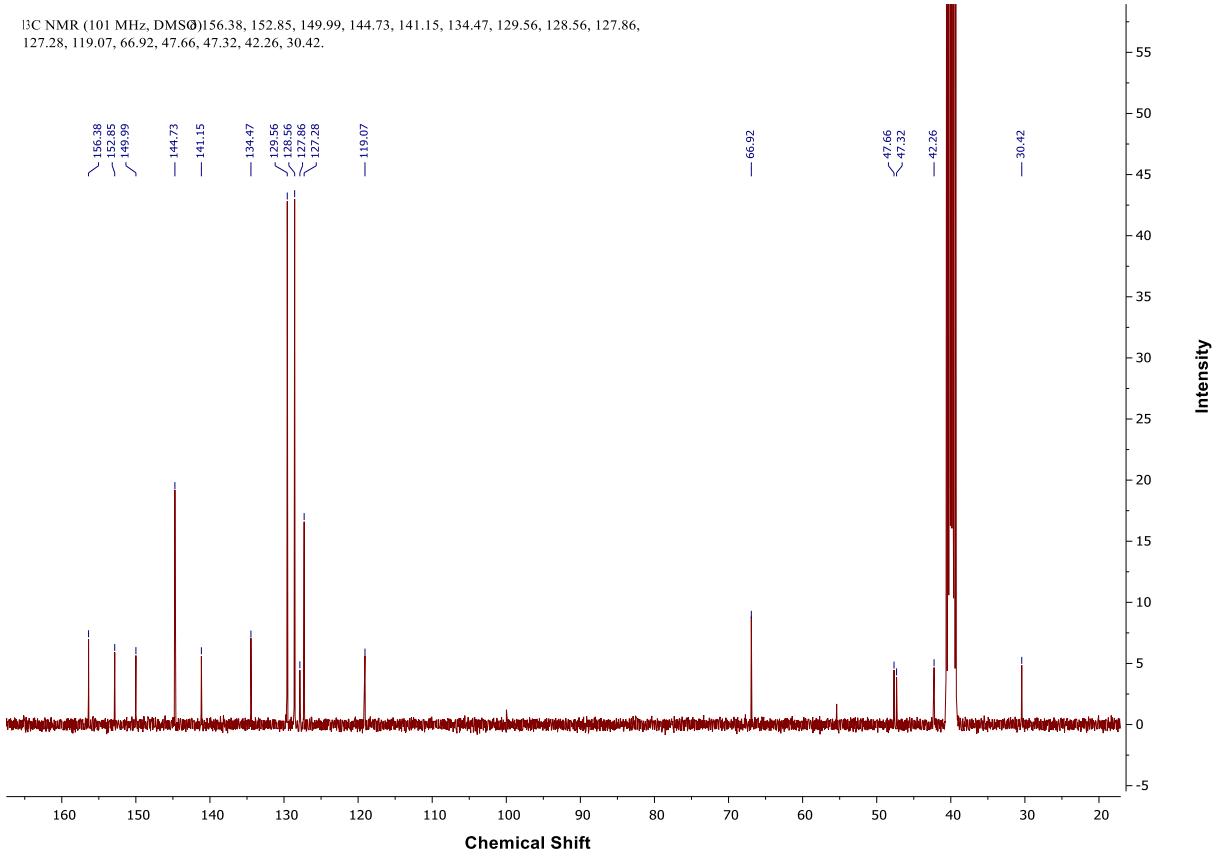
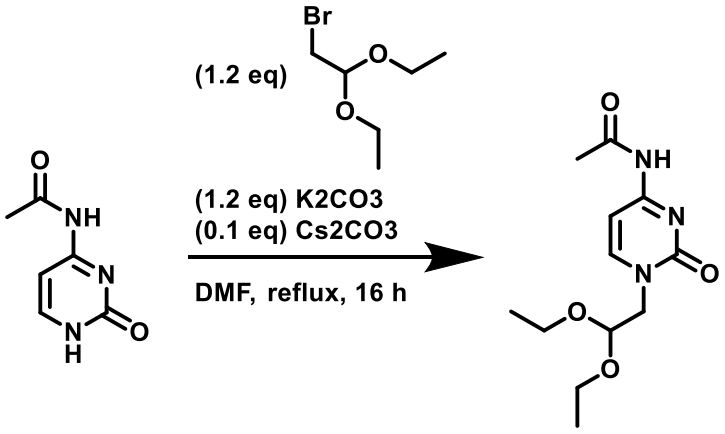


Figure C18: CNMR spectrum of compound **9**



10

N-[1-(2,2-diethoxyethyl)-2-oxo-1,2-dihydropyrimidin-4-yl]acetamide: N4-acetylcytosine (65 mmol, 10 g, 1 eq) was added to an oven-dried N_2 purged round bottom flask. Anhydrous potassium carbonate (78 mmol, 10.8 g, 1.2 eq) and cesium carbonate (6.5 mmol, 2.11 g, 0.1 eq) were then added. Anhydrous DMF (150 mL) was then added followed by one portion of bromoacetaldehyde diethyl acetal (78 mmol, 11.73 mL, 1.2 eq). The flask was then refluxed at 130 ℃ for 16 h. Once the reaction was finished, the flask was filtered while hot. The filtered cake was washed with minimal DMF before the eluent was concentrated via rotary evaporation resulting in a brown solid. The crude product was then dissolved in chloroform (300 mL) and was extract with water (3x, 200 mL). The organic layer was then collected and dried over anhydrous sodium sulfate. The chloroform was then filtered and evaporated at reduced pressure resulting in a brown solid. The flask was then triturated with 60 mL of diethyl ether producing a white solid. The mixture was cooled in a fridge over night before being filtered. The cake was then washed with 40 mL diethyl ether and was collected and dried *in vacuo* to afford compound **10** as a white solid (8.1 g, 46%).

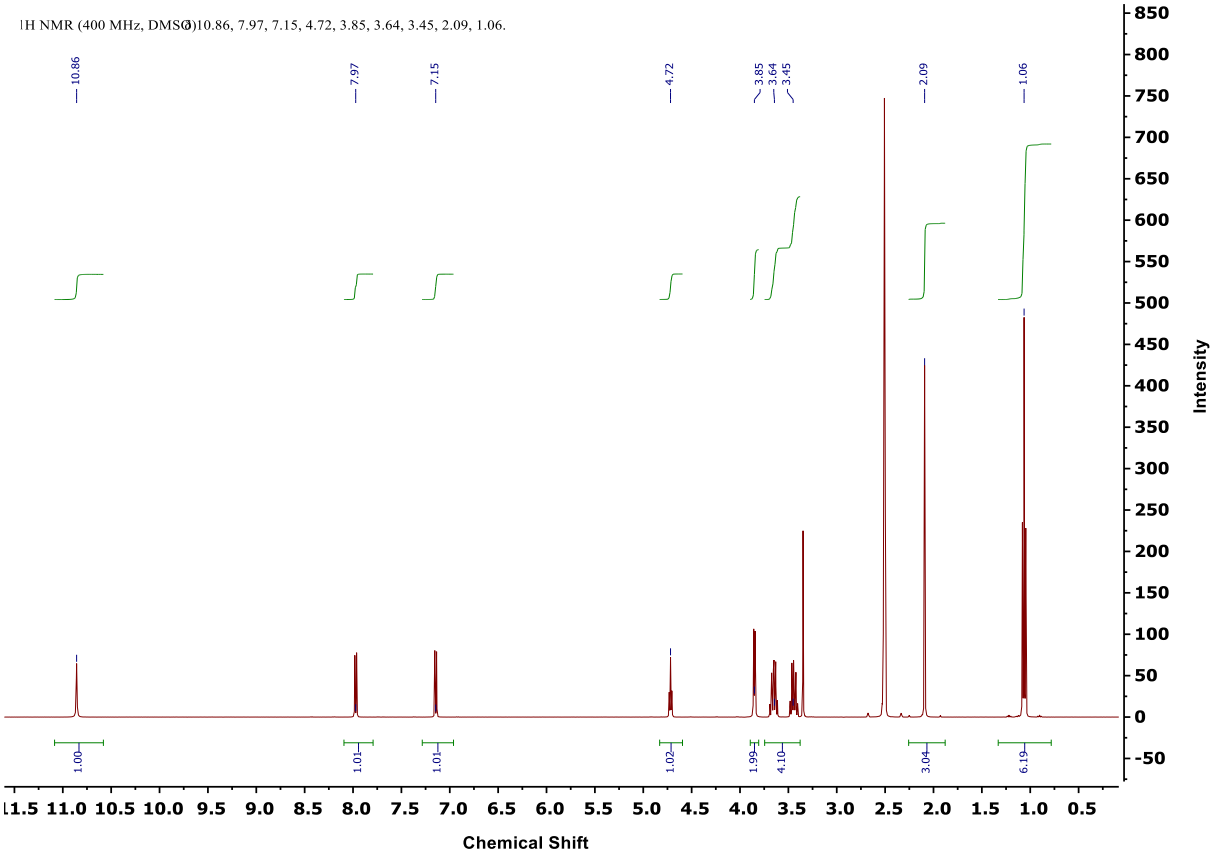


Figure C19: HNMR spectrum of compound **10**

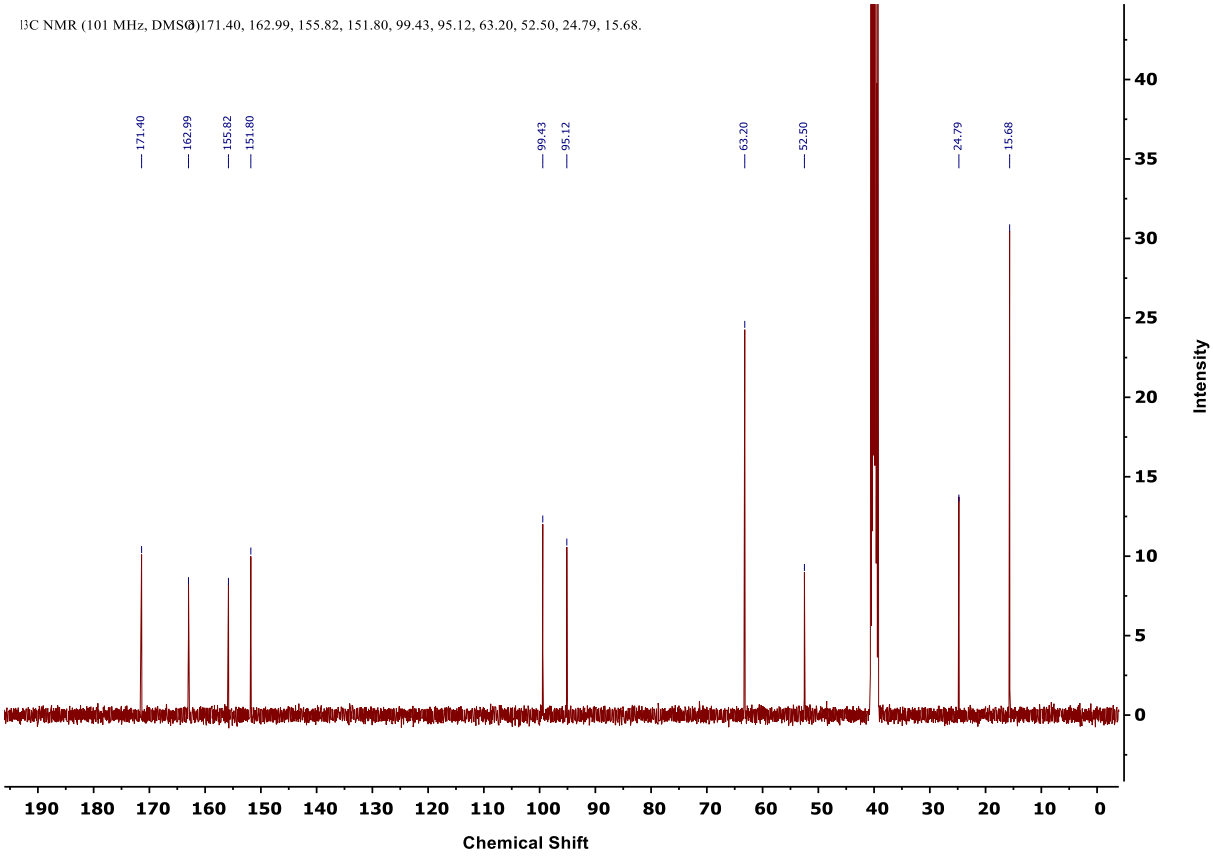
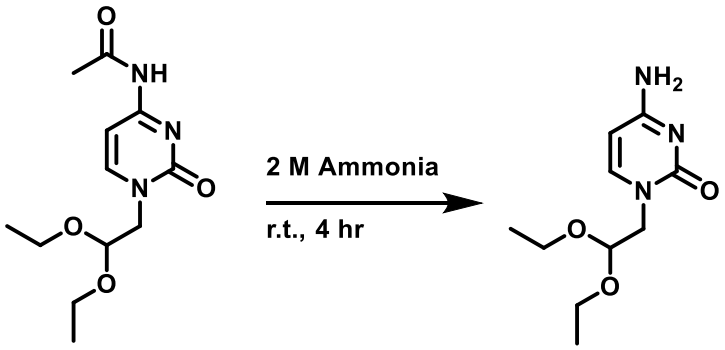


Figure C20: CNMR spectrum of compound **10**



10 11

4-amino-1-(2,2-diethoxyethyl)-1,2-dihydropyrimidin-2-one: Compound **10** (8.2 g, 30 mmol) was added to a 500 mL round bottom flask that had been oven dried and nitrogen purged. An anhydrous solution of 2 M ammonia in methanol (160 mL) was added and the reaction stirred for 4 h at room temperature. The flask was then evaporated at a reduced pressure to dryness and diethyl ether (100 mL) was added. The contents were then filtered and the cake was washed with 50 mL diethyl ether. The solid was collected and dried *in vacuo* to produce compound **11** as a white solid (6.8 g, 98%).

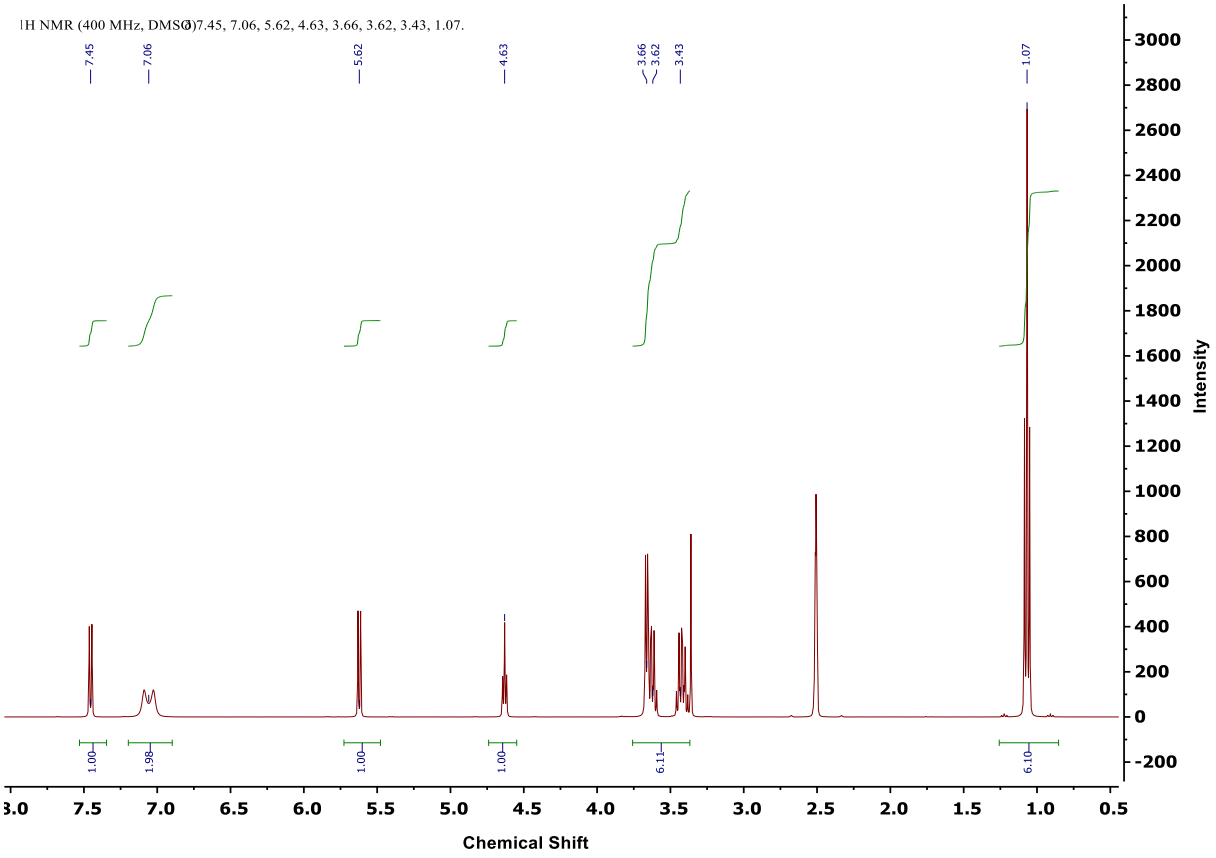


Figure C21: HNMR spectrum of compound **11**

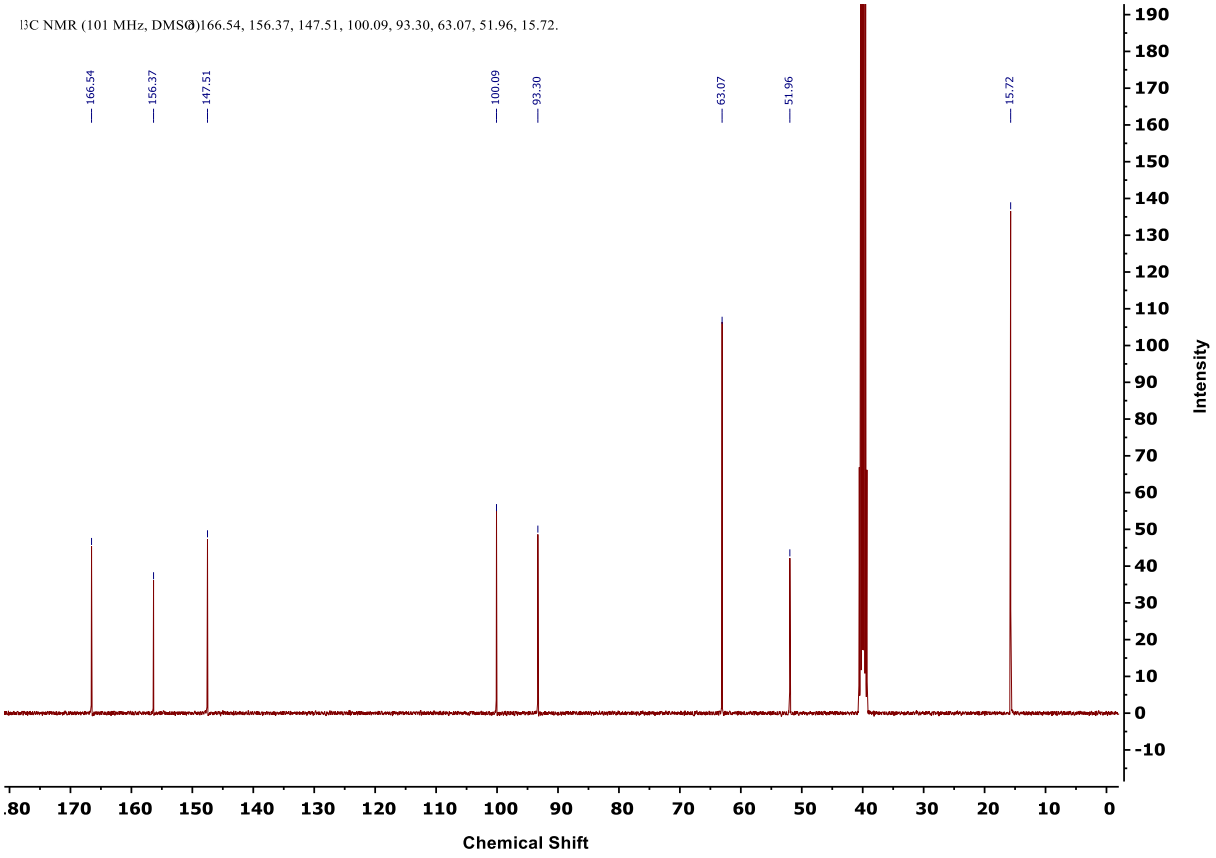
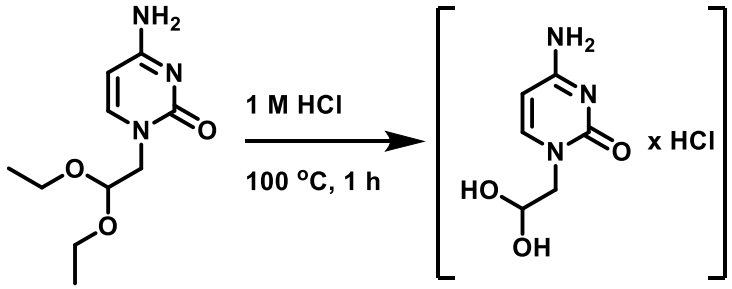


Figure C22: CNMR spectrum of compound **11**



11 12

4-amino-1-(2,2-dihydroxyethyl)-1,2-dihydropyrimidin-2-one hydrochloride: Compound **11** (6.6 g, 29.1 mmol) was added to an oven-dried round bottom flask that had been purged with nitrogen gas. 100 mL of 1 M HCl was then added and the round bottom flask was refluxed for 1 h at 100 ℃. Upon completion of the reaction, the water was removed via rotary evaporation. Upon drying, a white solid formed. To the solid was added 100 mL of tetrahydrofuran and the mixture was cooled in an ice bath for 1 h. The solid was filtered and washed with diethyl ether. The solid was dried under high vac to afford compound **12** as a white solid (4.87 g, 98%). The product was found to be isolated as the hydrate hydrochloride salt and was used without any further purification.

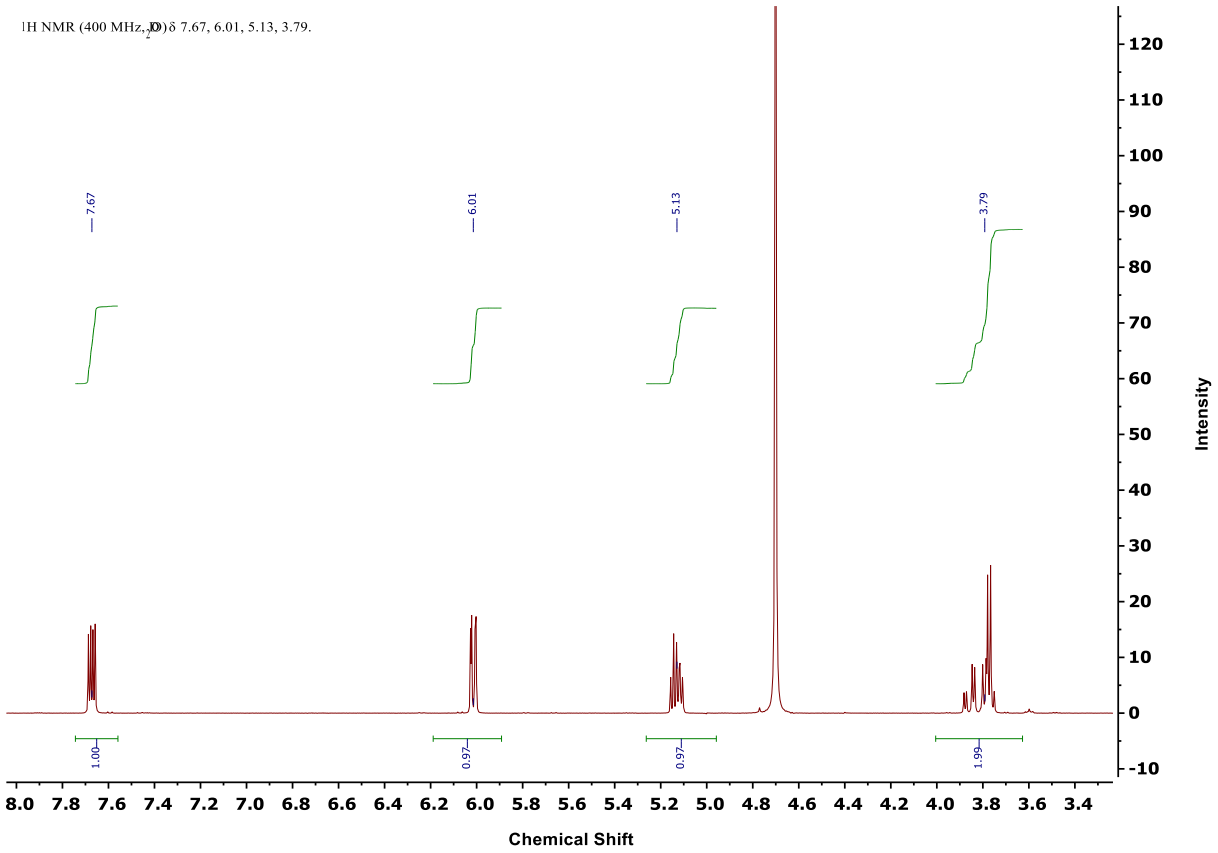


Figure C23: HNMR spectrum of compound **12**

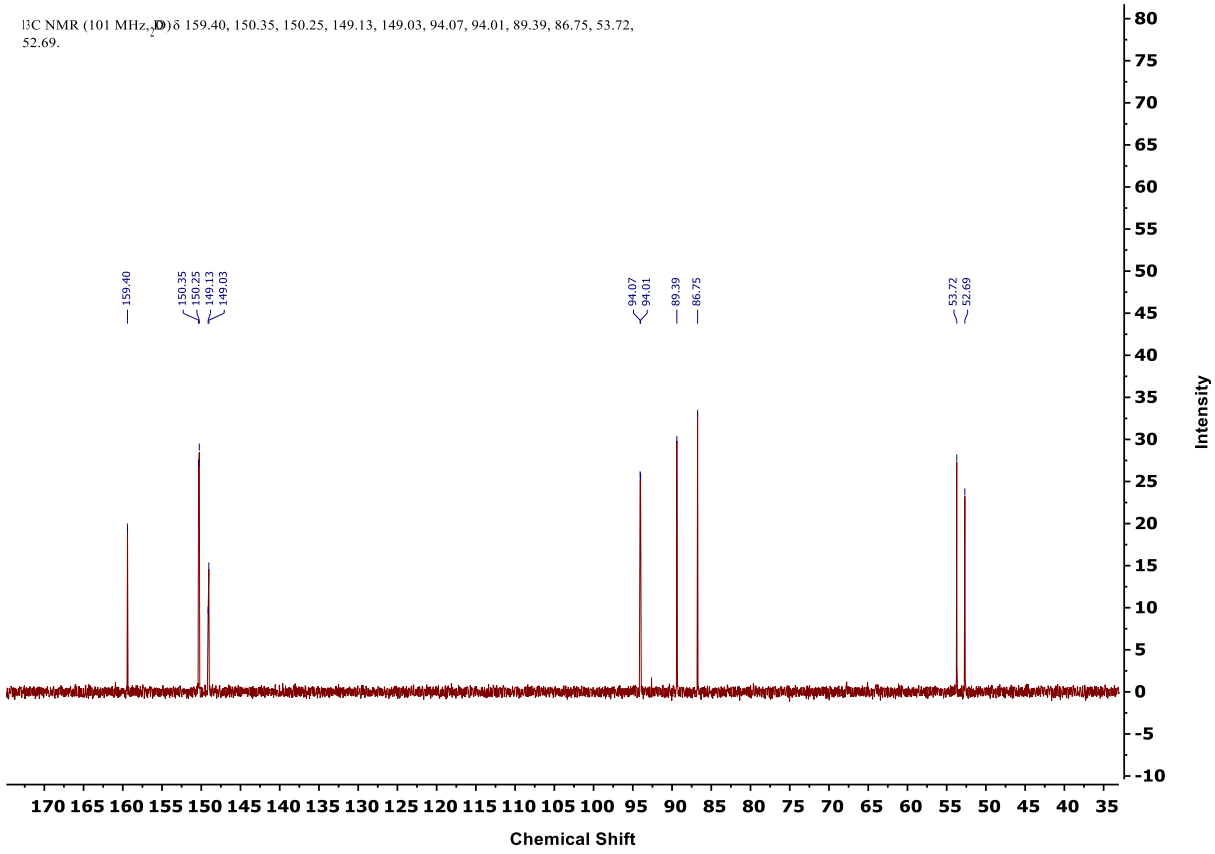
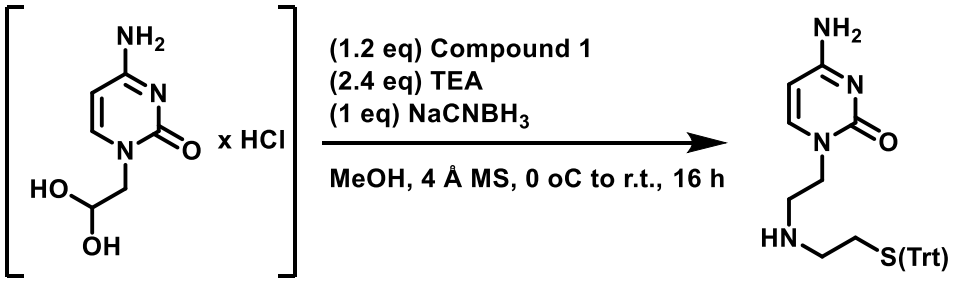


Figure C24: CNMR spectrum of compound **12**



12 13

4-amino-1-[2-({2-[(triphenylmethyl)sulfanyl]ethyl}amino)ethyl]-1,2-dihydropyrimidin-2-one: Molecular sieves, grade 514, were added to a 500 mL round bottom flask that had been oven dried and nitrogen purged. Compound **1** (10.2 g, 28.9 mmol), anhydrous methanol (200 mL), anhydrous triethylamine (6.1 mL, 43.5 mmol), and compound **12** (3 g, 14.5 mmol) were added. The flask was then submerged in an ice bath for 15 minutes. Sodium cyanoborohydride (0.91 g, 14.5 mmol) was then added in one portion as the flask remained on ice for 15 additional minutes. After reacting at room temperature for two hours, the mixture was filtered over a medium fret. The solution was extracted with 400 mL DCM and 400 mL water. The solid product **13** was collected to give a yield of 53% (3.4 g).

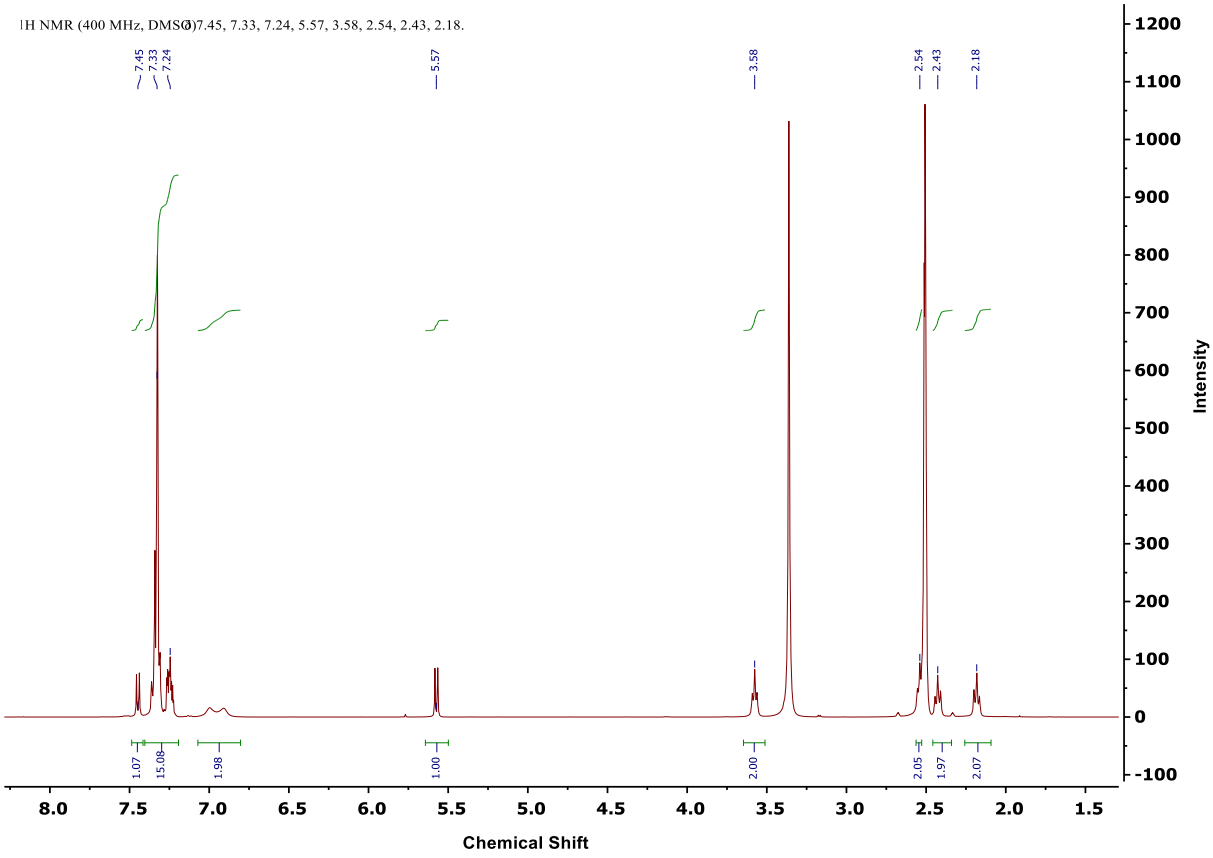


Figure C25: HNMR spectrum of compound **13**

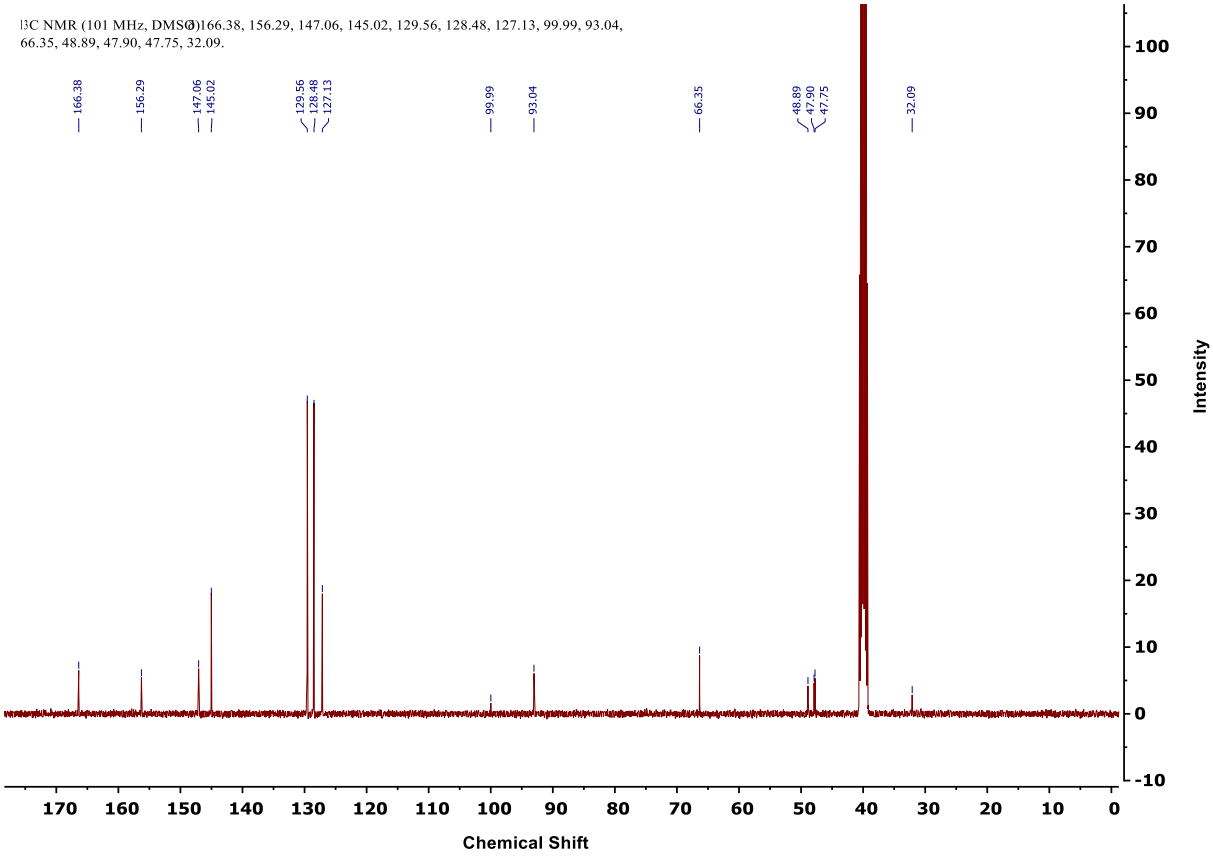
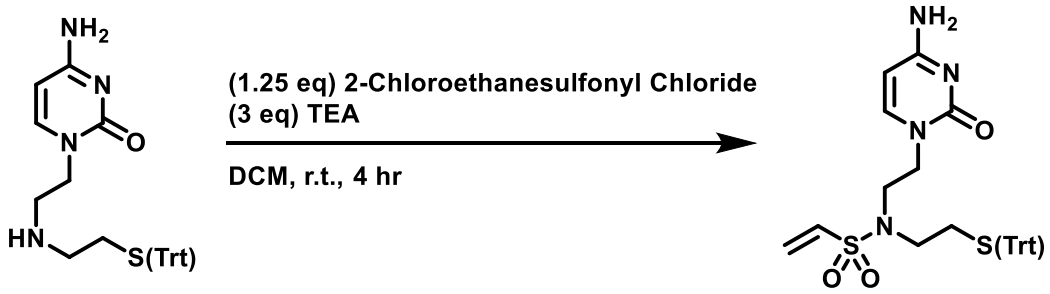


Figure C26: CNMR spectrum of compound **13**



13 14

N-[2-(4-amino-2-oxo-1,2-dihydropyrimidin-1-yl)ethyl]-N-{2-[(triphenylmethyl)sulfanyl]ethyl} ethene-1-sulfonamide: To an oven dried, nitrogen purged round bottom flask was added compound **13** (4.93 g, 10.8 mmol, 1 eq) followed by 100 mL of DCM. Next, anhydrous triethylamine (4.51 mL, 32.4 mmol, 3 eq) was added followed by dropwise addition of 2-chloroethylsulfonyl chloride (1.36 mL, 13 mmol. 1.2 eq) over the course an hour. Once the addition was finished, the reaction was allowed to stir at room temperature for 4 hr. Once finished, 100 mL of water was added to quench the reaction and was transferred to an extraction vessel. An additional 200 mL of DCM was added and the organic layer was washed with water (2x 100 mL) followed by drying over Na_2SO_4. The Na_2SO_4 was filtered and the organic solution concentrated before being purified by flash chromatography with a gradient of 0% to 10% methanol in DCM to afford product **14** (3.1 g, 52%) as a tan solid.

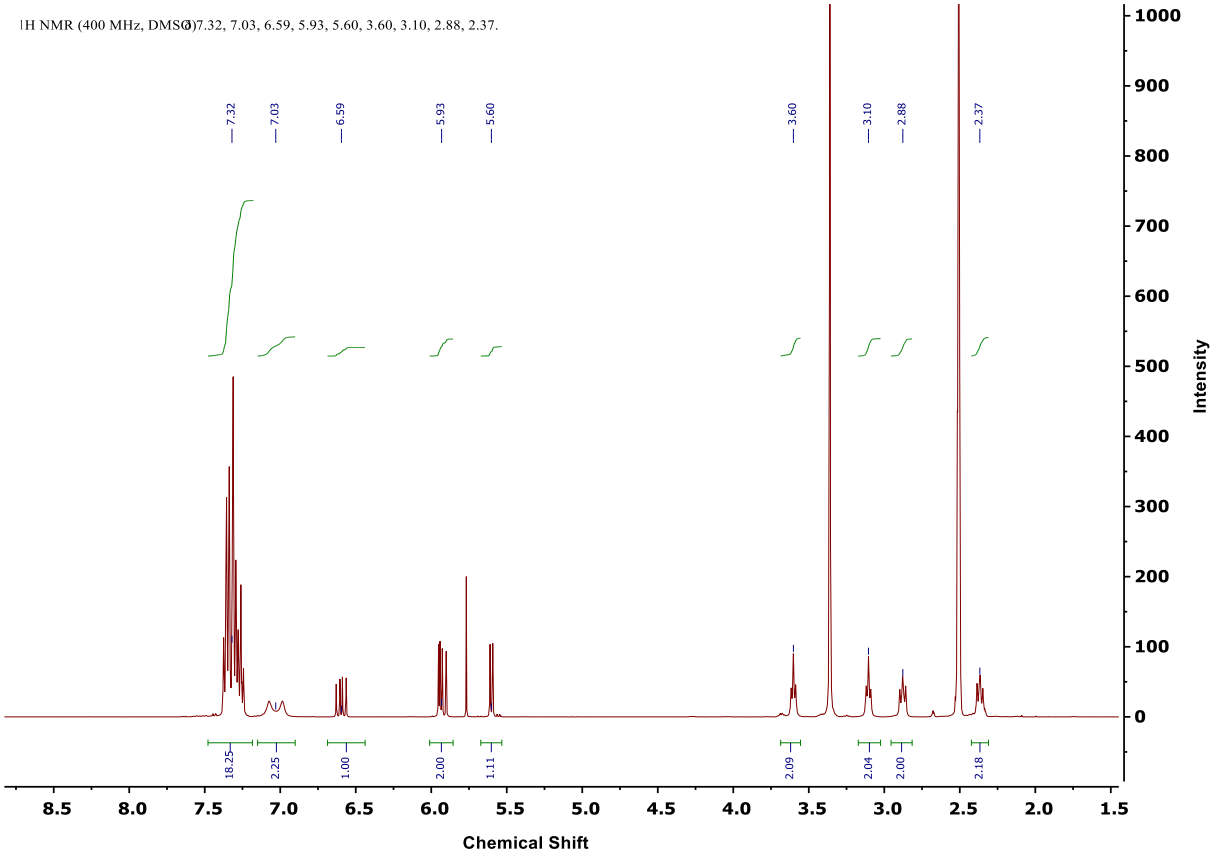


Figure C27: HNMR spectrum of compound **14**

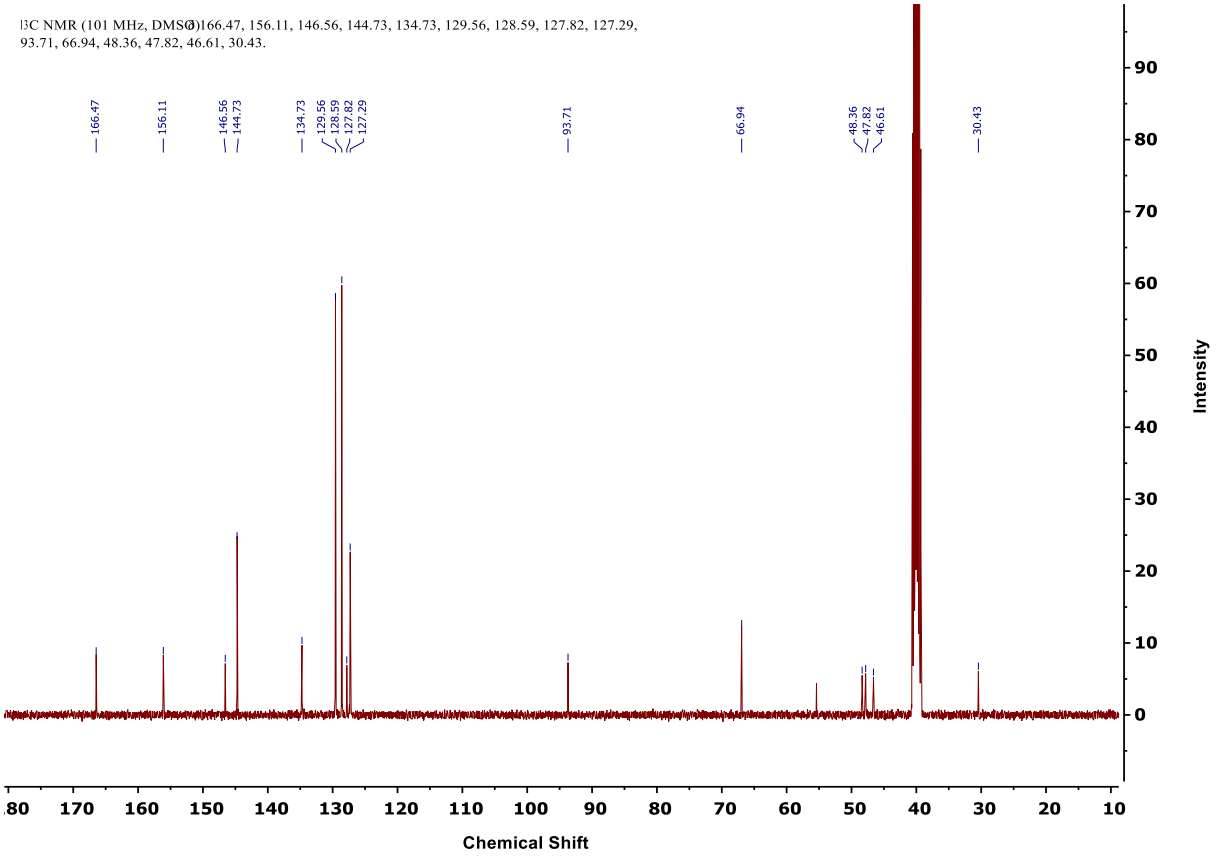
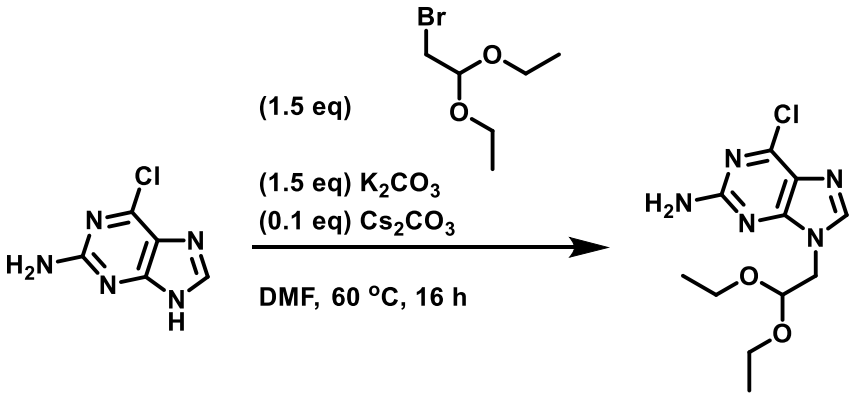


Figure C28: CNMR spectrum of compound **14**



6-chloro-9-(2,2-diethoxyethyl)-9H-purin-2-amine: 2-amino-6-chloropurine (59 mmol, 10 g, 1 eq) was added to an oven-dried round bottom flask that has been purged with nitrogen gas. Anhydrous potassium carbonate (88.5 mmol, 12.2 g, 1.5 eq) and cesium carbonate (5.9 mmol, 1.9 g, 0.1 eq) were added to the flask. Anhydrous DMF (150 mL) was added followed by bromoacetaldehyde diethyl acetal (88.5 mmol, 13.3 mL, 1.2 eq) in one portion. The flask was then heated at 60 ℃ overnight. Upon completion of the reaction, the flask was filtered while hot. The cake was washed with minimal DMF and the eluent was evaporated under reduced pressure to produce a maroon oil. The resulting oil was dissolved in chloroform (300 mL) and washed with water (5x, 200 mL). The organic layer was collected and dried over anhydrous magnesium sulfate. After filtering the magnesium sulfate, the chloroform was evaporated resulting in an orange oil. The resulting oil was then purified using flash column chromatography with a gradient of 0% to 75% ethyl acetate in hexanes. The product was collected and dried *in vacuo* producing compound **15** (8.8 g, 52%).

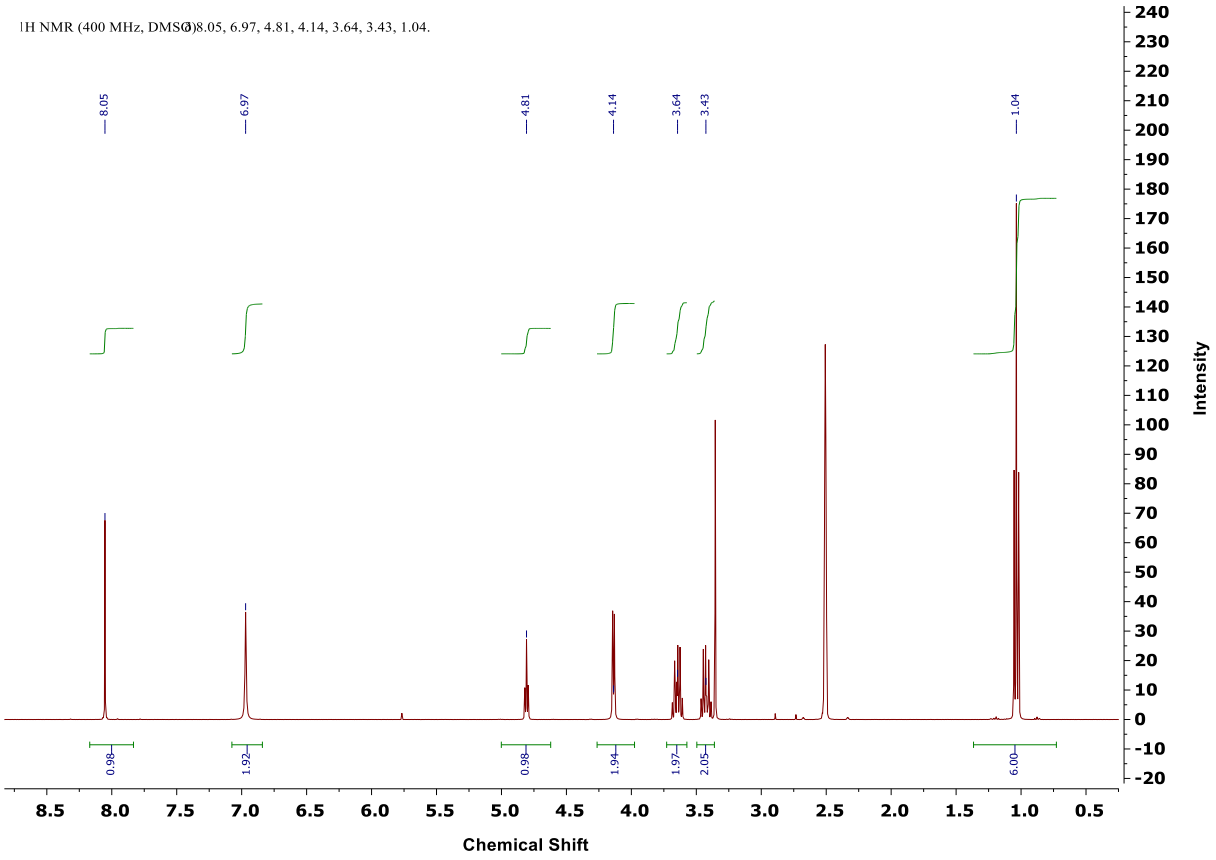


Figure C29: HNMR spectrum of compound **15**

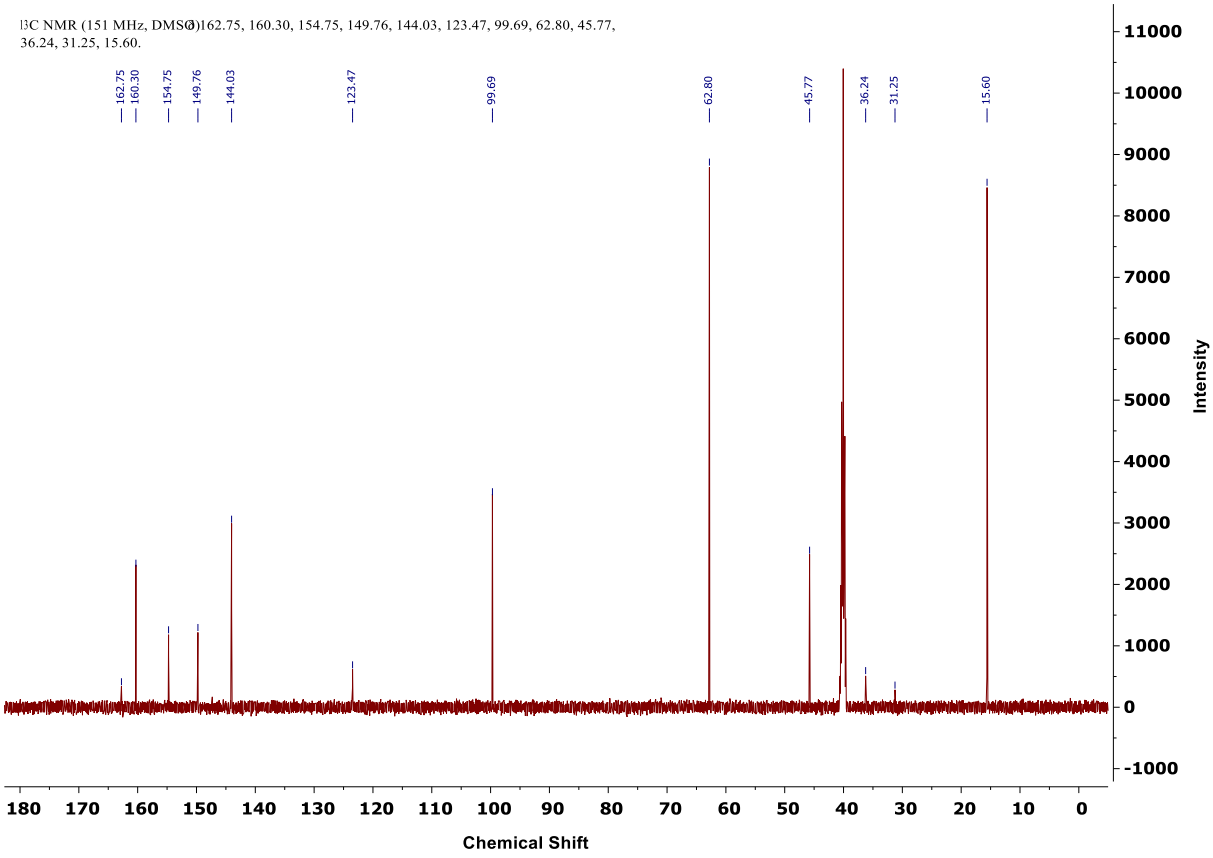
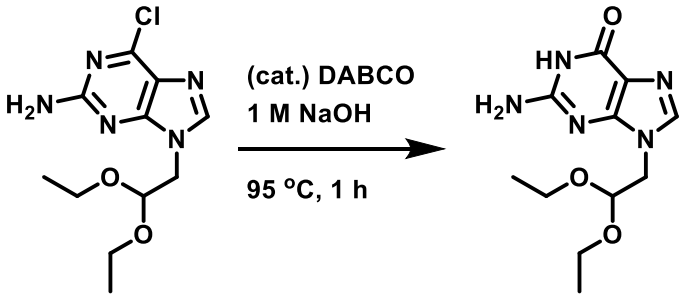


Figure C30: CNMR spectrum of compound **15**



15 16

2-amino-9-(2,2-diethoxyethyl)-6,9-dihydro-1H-purin-6-one: Compound **15** (15.6 mmol, 4.45 g, 1 eq) was added to an oven-dried and nitrogen purged round bottom flask followed by 1,4-diazabicyclo[2.2.2]octane (0.156 mmol, 17 mg, 0.01 eq) and 50 mL of 1M NaOH. The flask was then heated to 95 ℃ for one hour. Upon completion of the reaction, the flask was cooled in an ice bath for 1 h and the pH was adjusted to seven while stirring by dropwise addition of concentrated HCl. Once pH 7 was reached, a white precipitate was formed. The solution was filtered and the cake was washed with a minimal amount of water and 50 mL diethyl ether. The solid product was then dried *in vacuo* to produce compound **16** as a white solid (3.8 g, 92%).

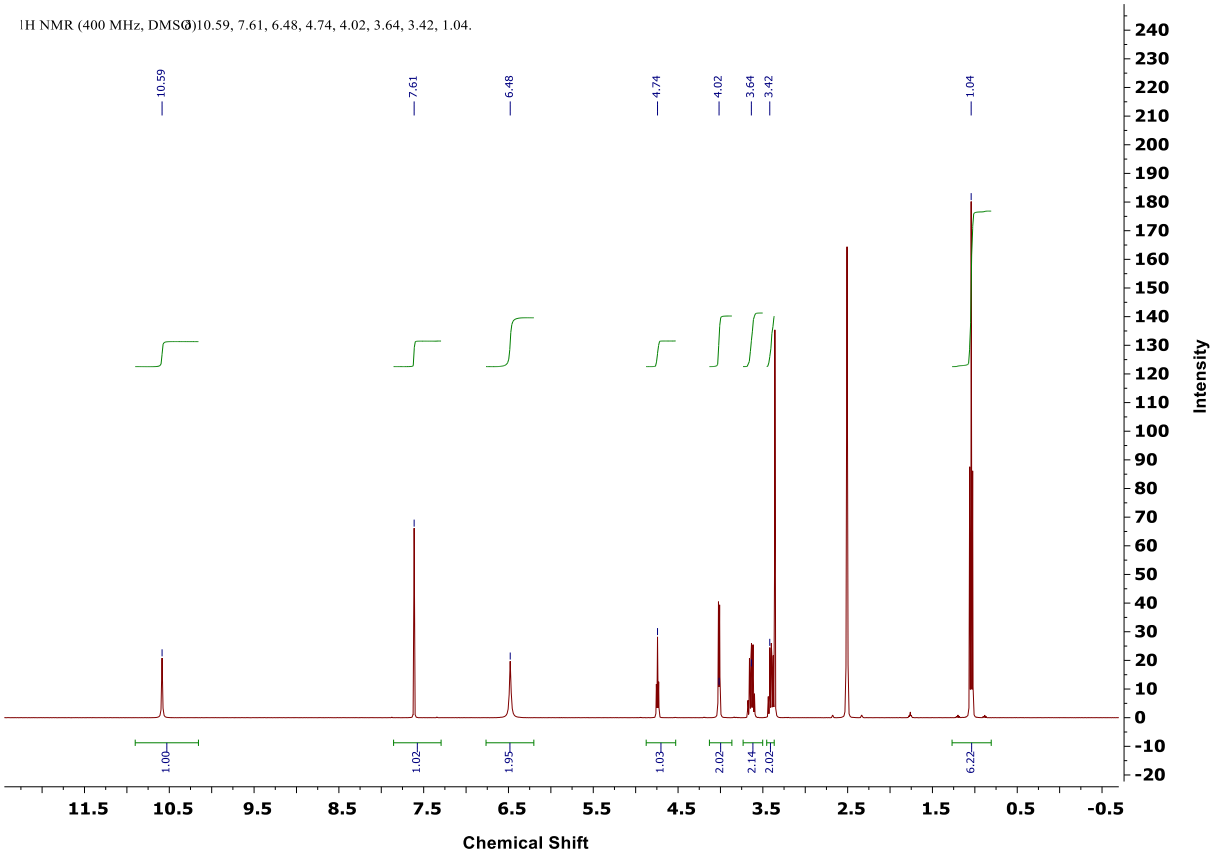


Figure C31: HNMR spectrum of compound **16**

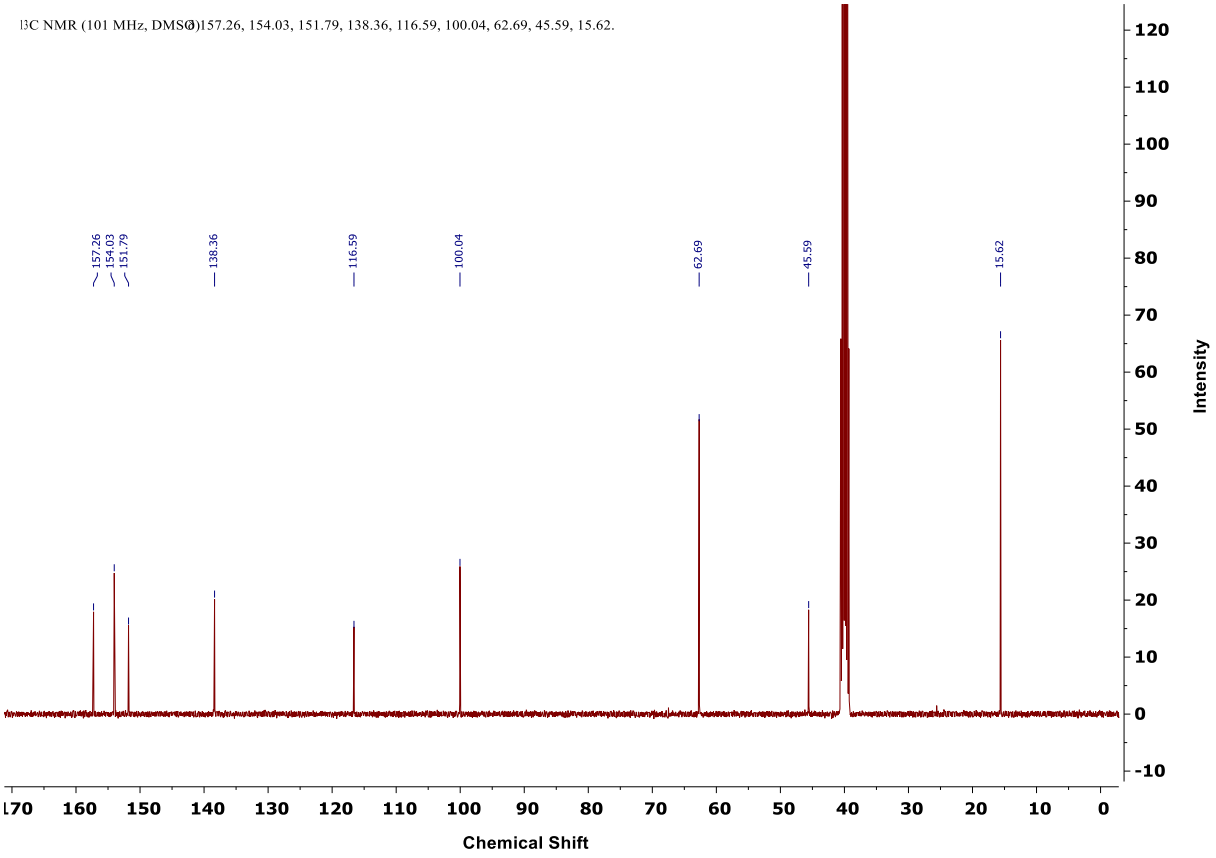
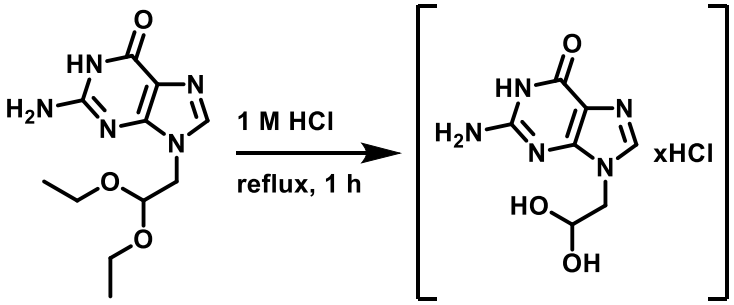


Figure C32: CNMR spectrum of compound **16**



16 17

2-amino-9-(2,2-dihydroxyethyl)-6,9-dihydro-1H-purin-6-one: Compound **16** (51.5 mmol, 13.8 g) was added to a 250 mL oven-dried round bottom flask that had been purged with nitrogen gas followed by 125 mL of 1 M HCl. The reaction was refluxed at 100 ℃ in the dark for one hour. Once the reaction was complete, the water was then evaporated via rotary evaporation in the dark producing a white solid. Subsequently, 50 mL of tetrahydrofuran was added to the round bottom flask and the mixture was cooled in an ice bath for one hour. The solid was then filtered and the cake was washed with diethyl ether before being dried *in vacuo* to afford compound **17** (12.6 g, 99%). The product was isolated as the dihydrate hydrochloride salt.

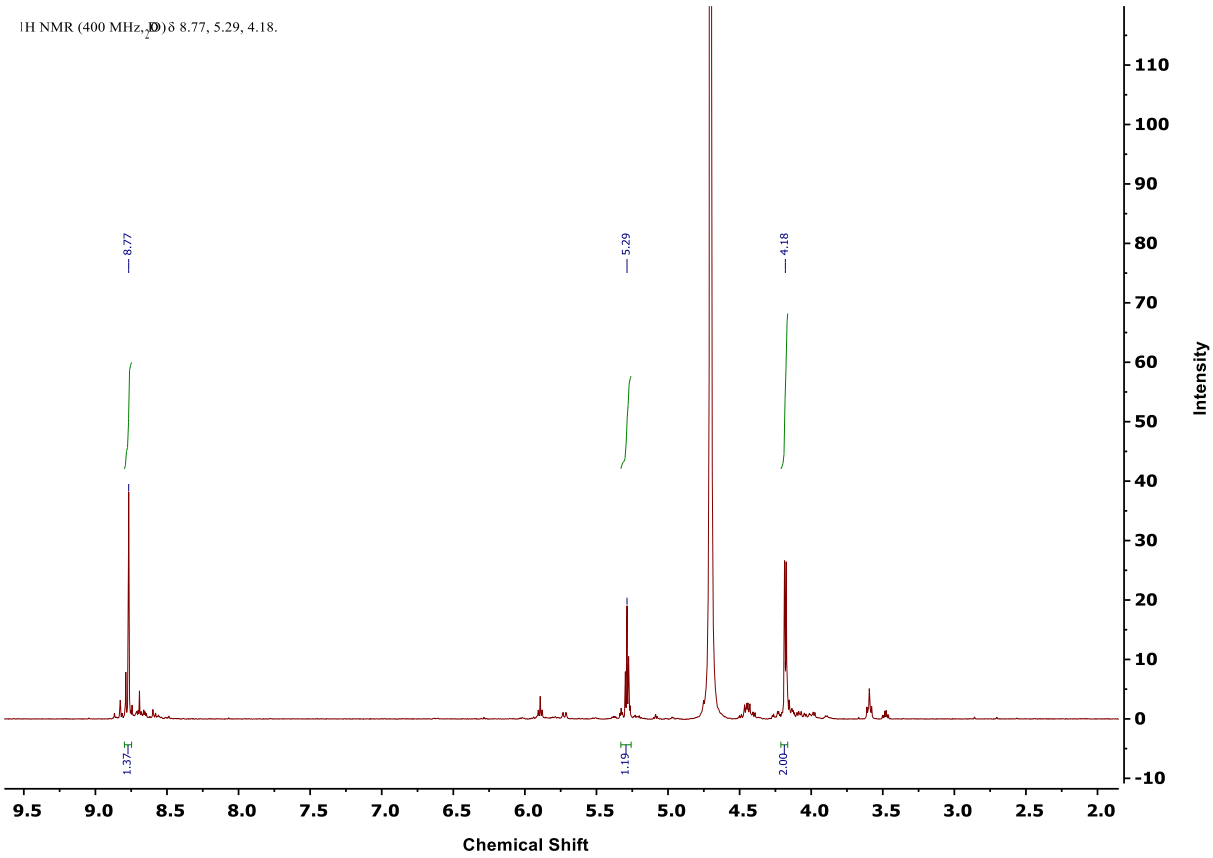


Figure C33: HNMR spectrum of compound **17**

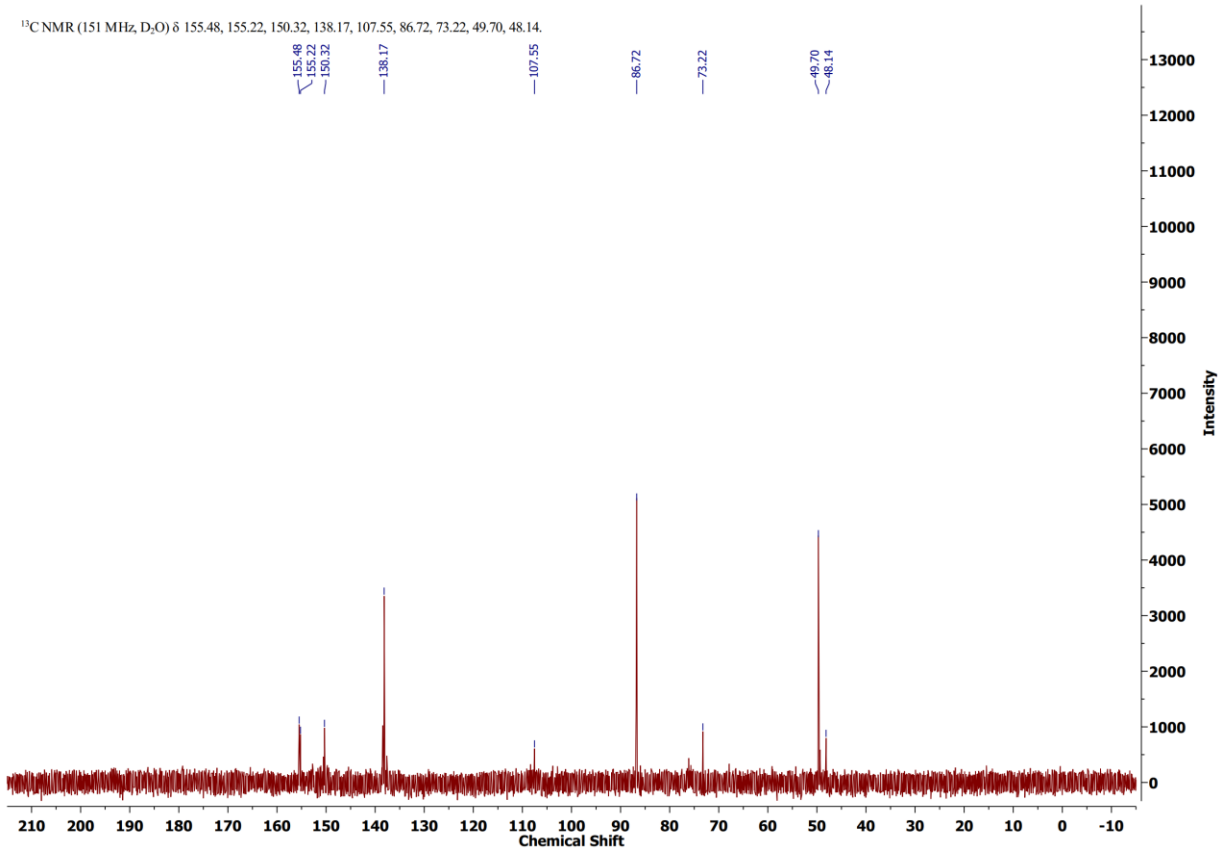
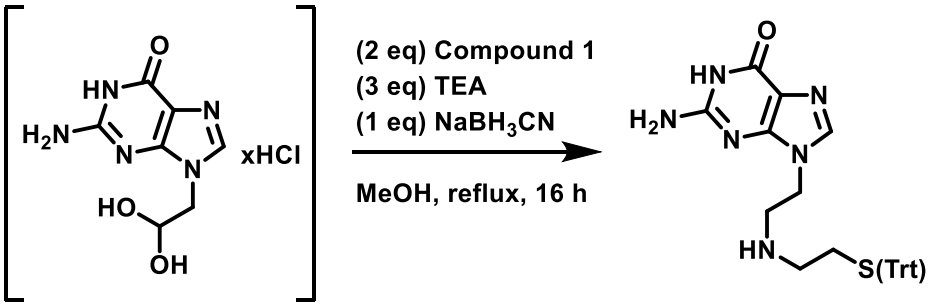


Figure C34: CNMR spectrum of compound **17**



17 18

2-amino-9-[2-({2-[(triphenylmethyl)sulfanyl]ethyl}amino)ethyl]-6,9-dihydro-1H-purin-6-one: To a 500 mL round bottom flask that had been oven dried and nitrogen purged was added molecular sieves. Compound **1** (8.6 g, 24.3 mmol, 2 eq), 300 mL anhydrous methanol, and anhydrous triethylamine (5.1 mL, 36.4 mmol, 3 eq) were subsequently added followed by mixing for 15 min. Compound **17** (3 g, 12.1 mmol, 1 eq) was then added followed by sodium cyanoborohydride (0.76 g, 12.1 mmol, 1 eq). After the addition, the round bottom flask was heated to reflux at 75 °C overnight. Upon completion of the reaction, the mixture was filtered and the cake was washed with 100 mL DCM. The eluent was collected and concentrated to roughly 200 mL and transferred to an extraction vessel in which an additional 300 mL of DCM was added. The organic solution was extracted with water (400 mL). The water layer was then washed twice with 200 mL DCM. The DCM was subsequently dried over sodium sulfate and filtered. The solution was then concentrated via rotary evaporation. Flash chromatography was used to purify the product using a gradient from 0% MeOH in DCM to 15% MeOH in DCM. The solvent was collected and evaporated to produce compound **18** as a white solid (1.62 g, 27%).

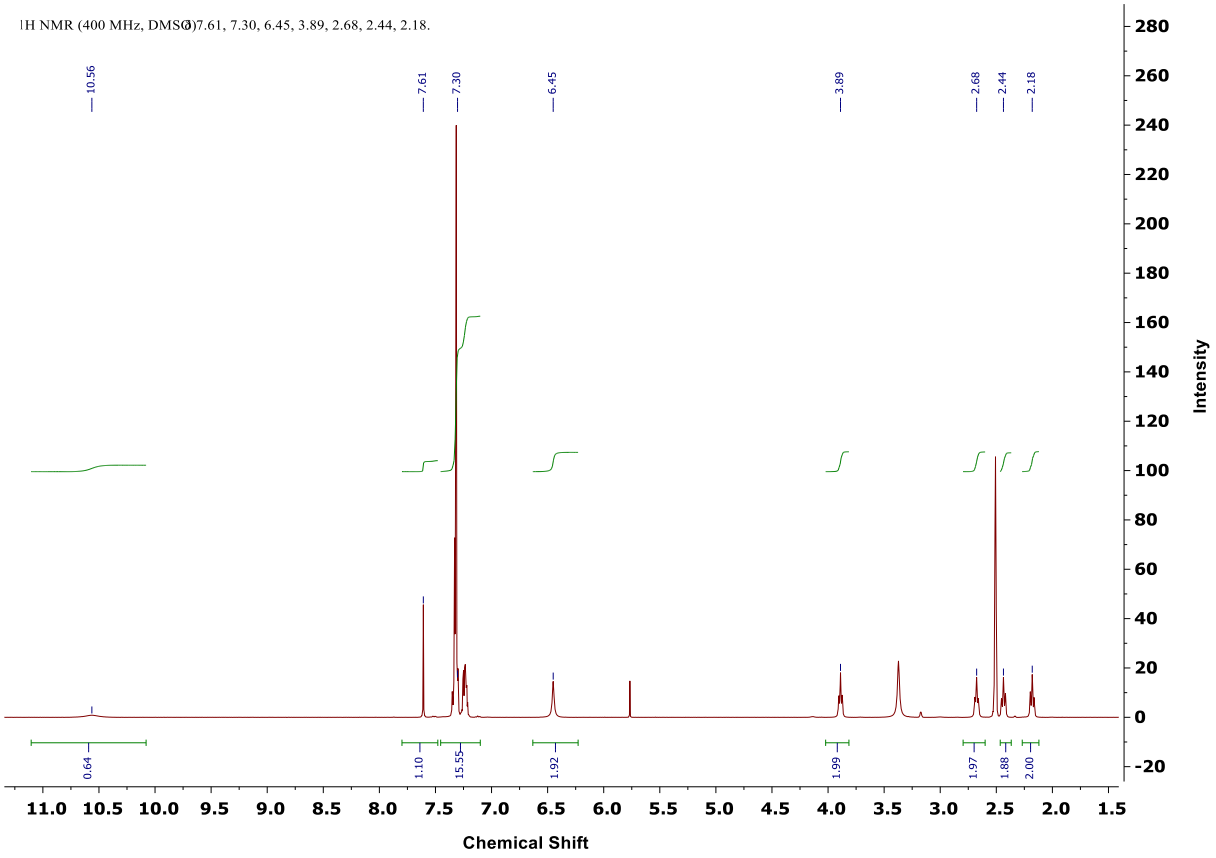


Figure C35: HNMR spectrum of compound **18**

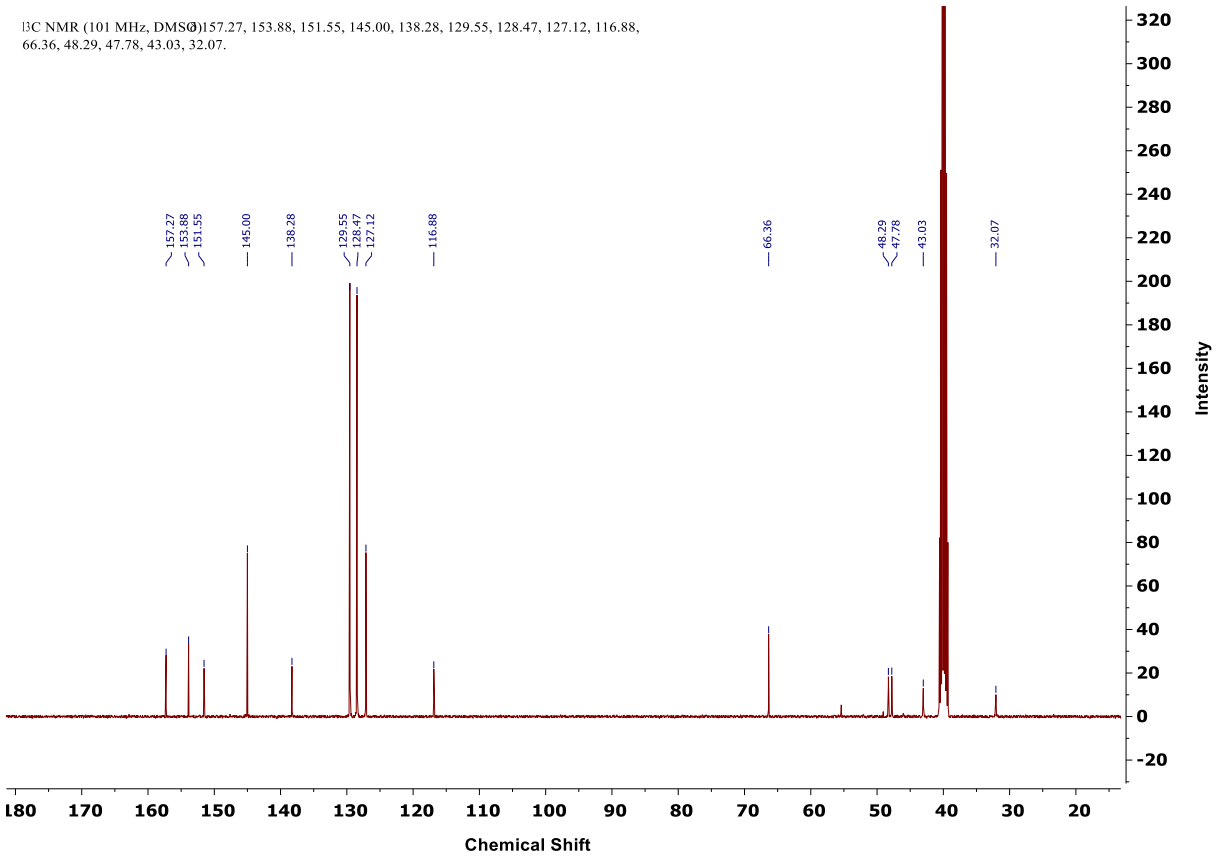
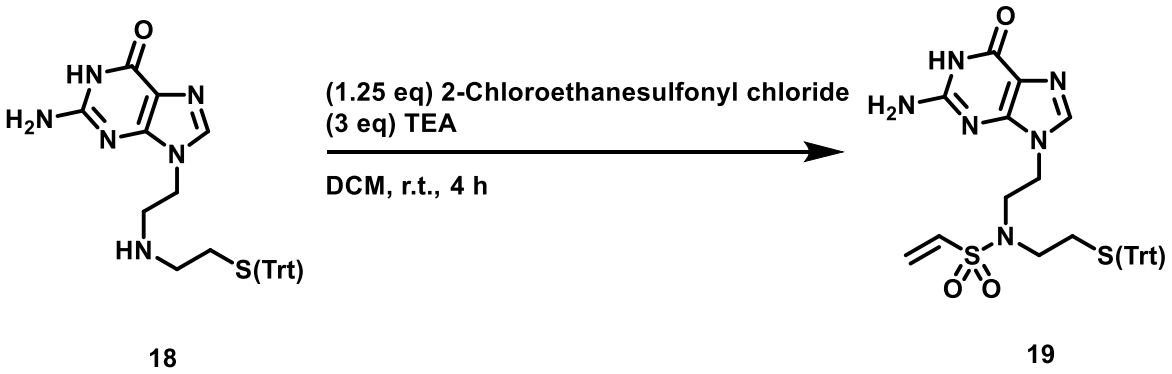


Figure C36: CNMR spectrum of compound **18**



N-[2-(2-amino-6-oxo-6,9-dihydro-1H-purin-9-yl)ethyl]-N-{2-[(triphenylmethyl)sulfanyl]ethyl}ethene-1-sulfonamide: Compound **18** (3 g, 6.05 mmol, 1 eq) was added to a 500 mL round bottom flask that had been oven dried and N_2 purged. Anhydrous DCM (300 mL) and anhydrous triethylamine (2.53 mL, 18.15 mmol, 3 eq) were added. After mixing for 15 min, 2-Chloroethanesulfonyl chloride (790 µL, 7.56 mmol) was added dropwise over the course of one hour. Once the addition was finished, the mixture was allowed to react for 4 h before water (100 mL) was added to quench the reaction. The organic phase was extracted, dried over Na_2SO_4, filtered, and concentrated via rotary evaporation. The resulting oil was column purified with a gradient of 0% to 10% methanol in DCM to produce compound **19** as a white solid (1.83 g, 52%).

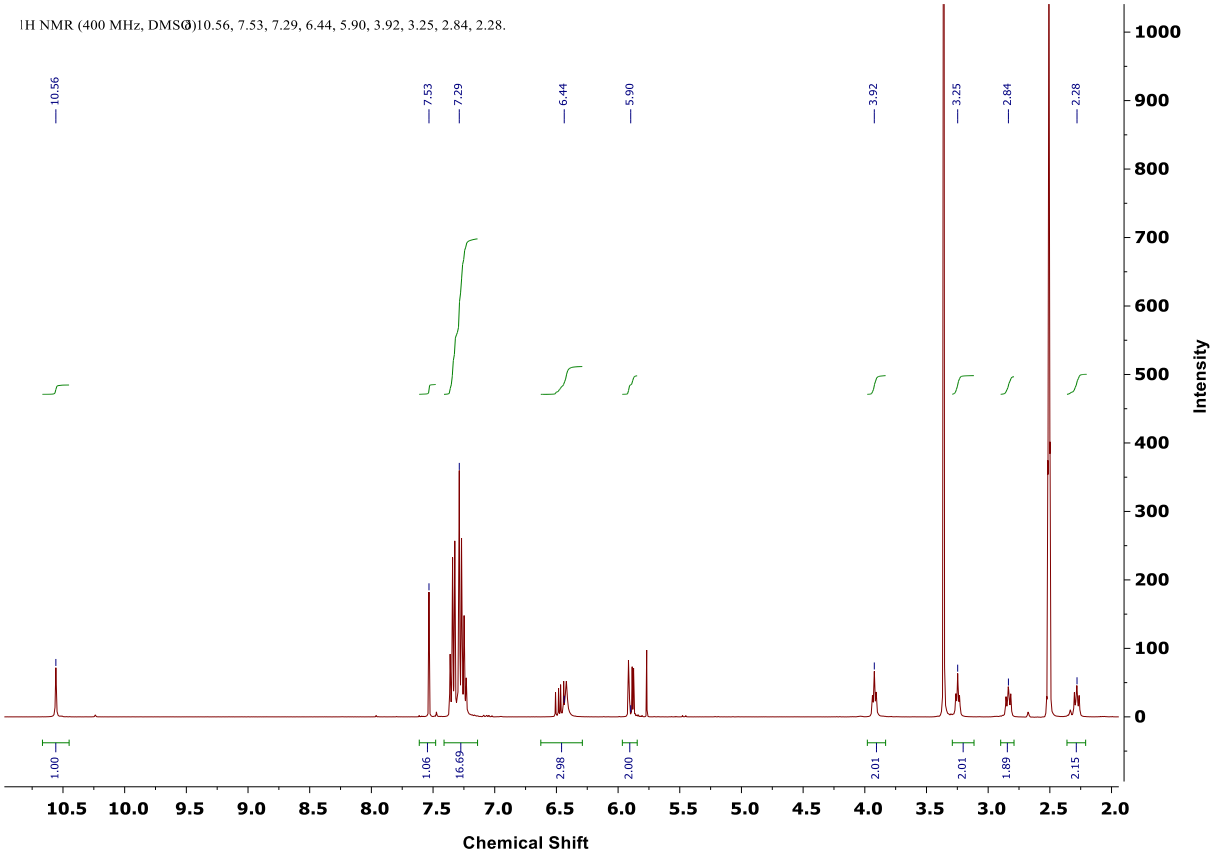


Figure C37: HNMR spectrum of compound **19**

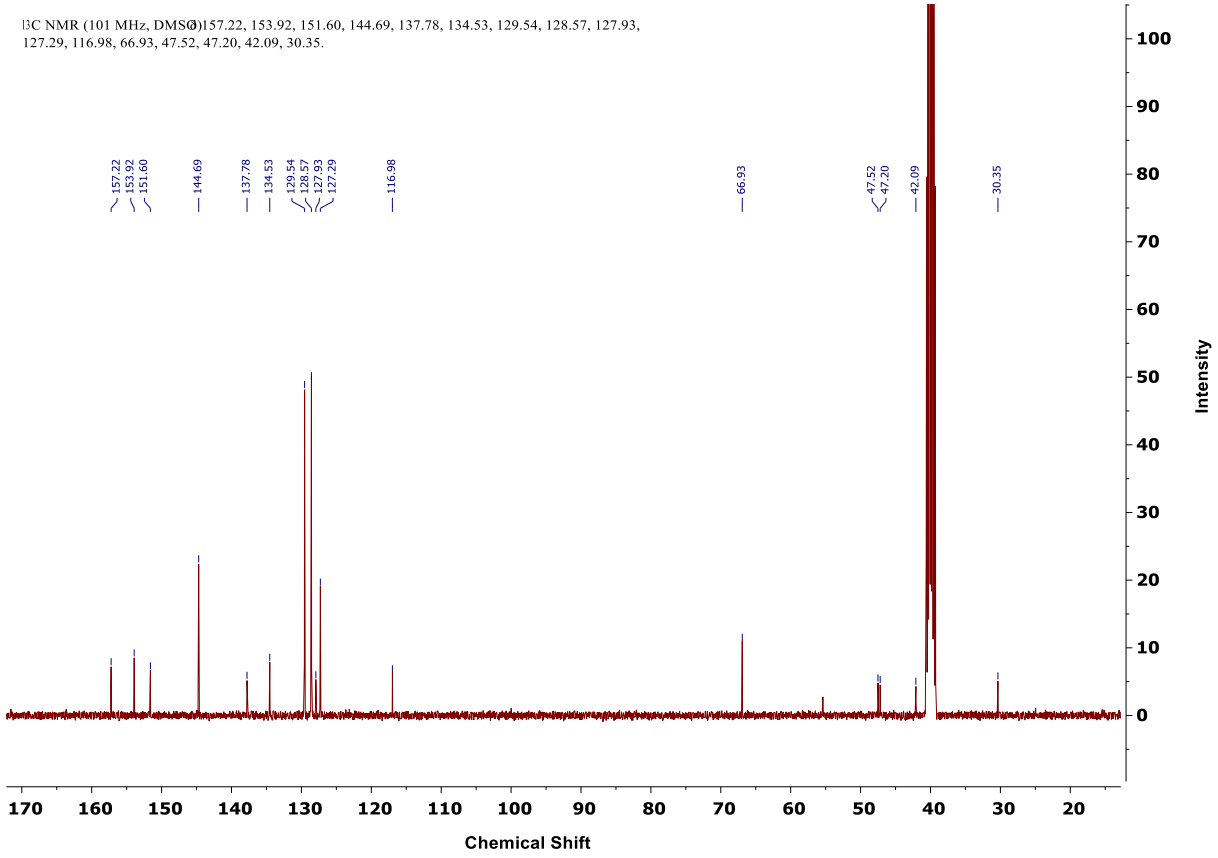
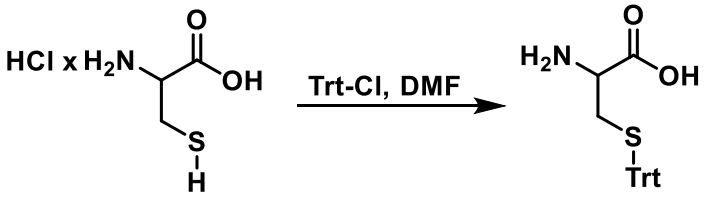


Figure C38: CNMR spectrum of compound **19**

Appendix D

SYNTHETIC PROCEDURES FOR ACID-FUNCTIONALIZED THIOL-MICHAEL MONOMERS



20

S-trityl-L-cysteinate: L-cysteine hydrochloride (10 g, 63.4 mmol) was added to a round bottom flask that had been oven dried and purged with nitrogen gas. Anhydrous DMF (40 mL) was added followed by trityl chloride (17.7 g, 63.4 mmol). After stirring for 48 hours at room temperature, the mixture was added to a beaker containing 350 mL of a 10% (w/v) sodium acetate solution. The beaker's contents were filtered and the cake was washed with water. The solid was redissolved in acetone at 50 °C for 30 minutes. The mixture was subsequently cooled in an ice bath for one hour. The mixture was then filtered and the cake was washed with both acetone and diethyl ether affording compound **20** as a white powder in 57.6% yield (13.3 g).

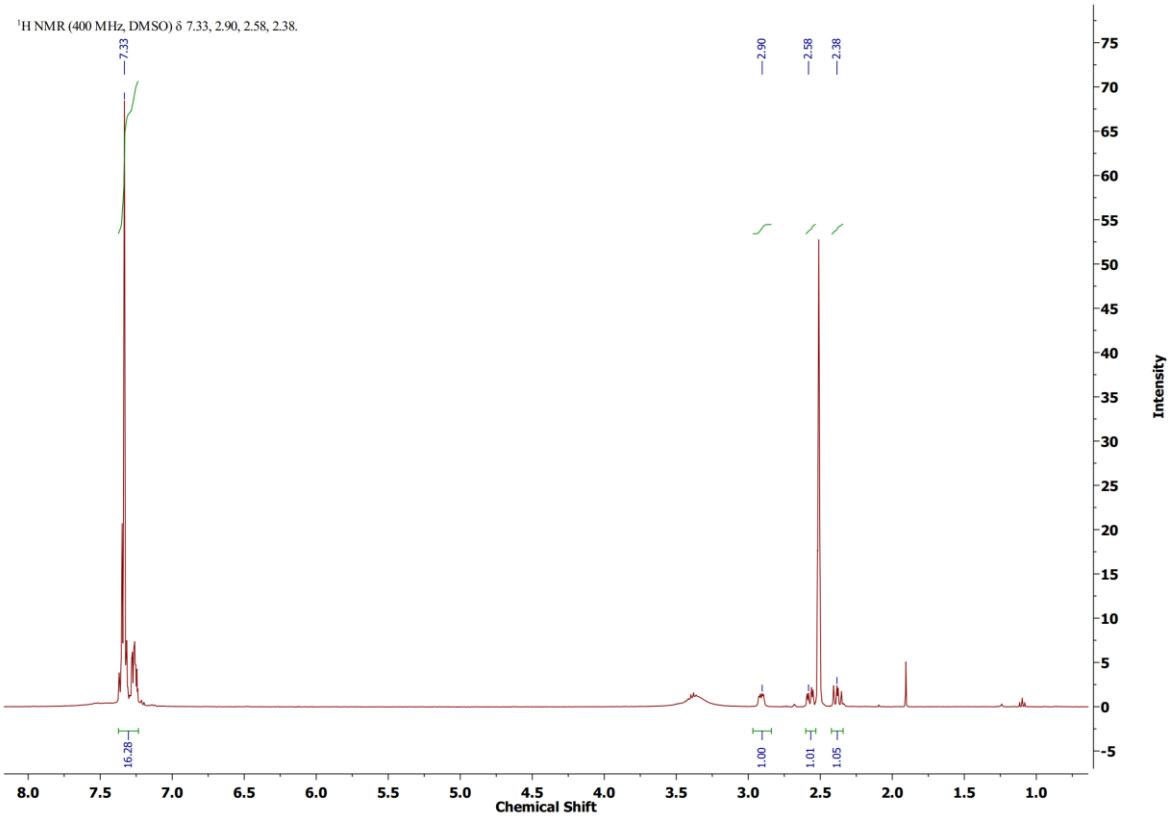


Figure D1: HNMR spectrum of compound **20**

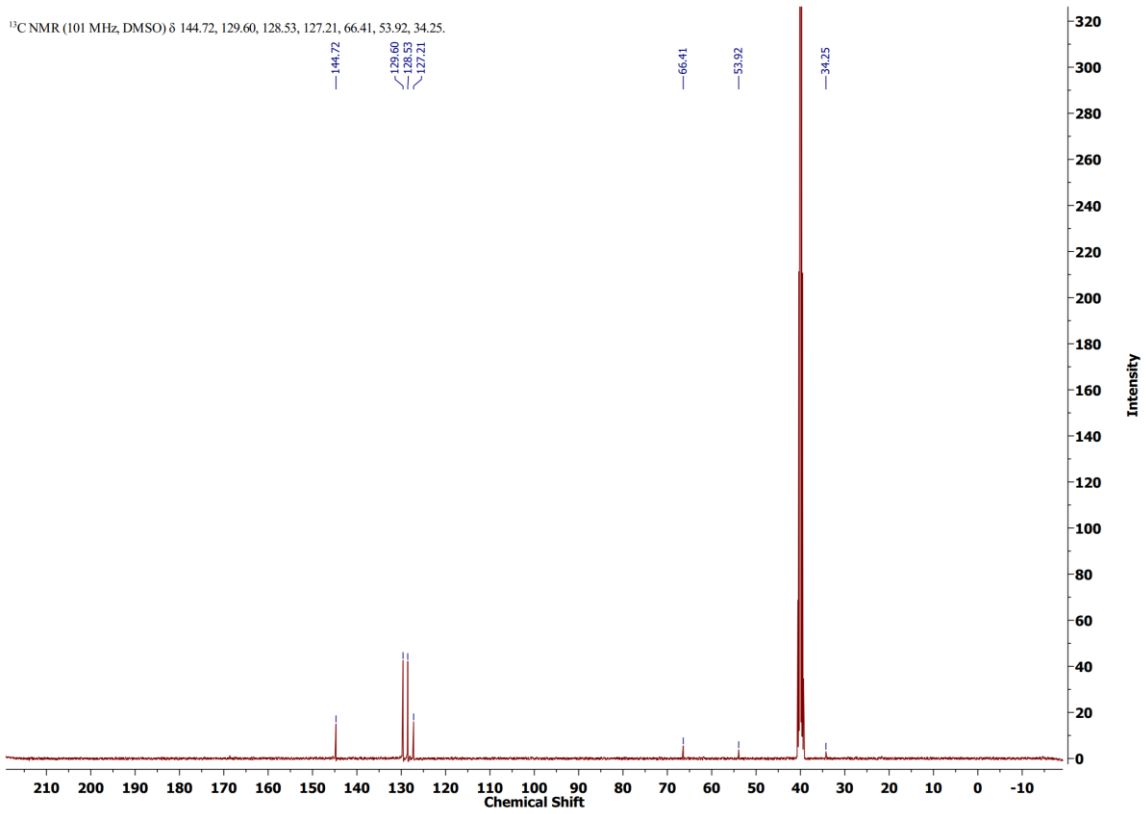
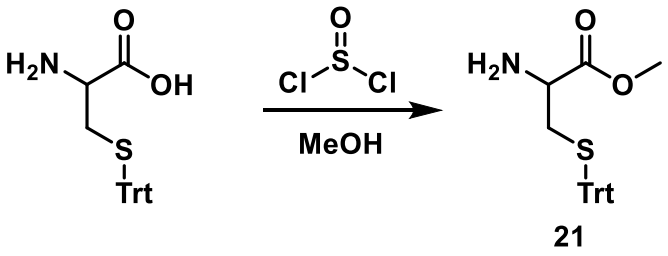


Figure D2: CNMR spectrum of compound **20**



21

L-(S)-tritylcysteine methyl ester: S-trityl-L-cysteine (1 g, 2.75 mmol) was added to a round bottom flask that had been oven dried and purged with nitrogen gas. Anhydrous methanol (50 mL) was added and the flask was cooled in an ice bath for 30 minutes. Thionyl chloride (1.59 mL, 22.0 mmol) was added in a drop wise manner

after which the flask warmed to room temperature. The mixture was then refluxed at 50 °C for 5 hours. The mixture was evaporated at a reduced pressure giving and the crude product was extracted using 100 mL of ethyl acetate. The organic layer was washed five times with 50 mL of a saturated sodium bicarbonate solution. The ethyl acetate was collected and dried over magnesium sulfate. The solution was filtered and evaporated under reduced pressure to yield compound **21** a viscous yellow oil in 96.8% yield (1 g).

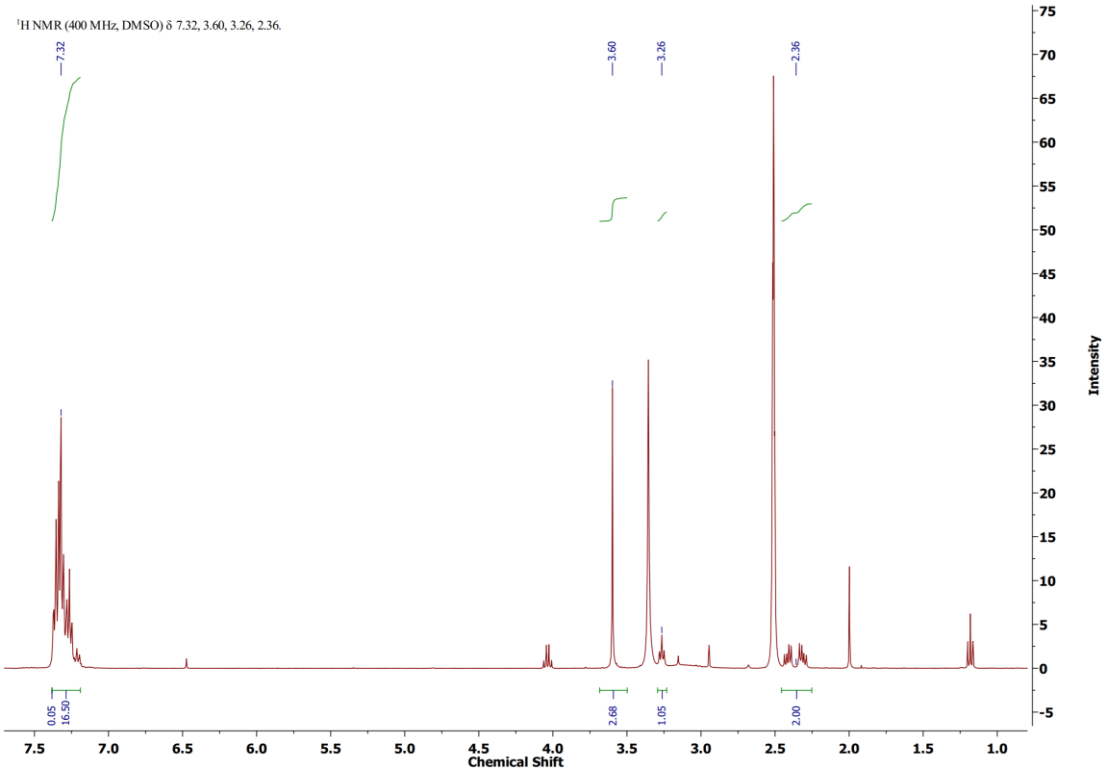


Figure D3: HNMR spectrum of compound **21**

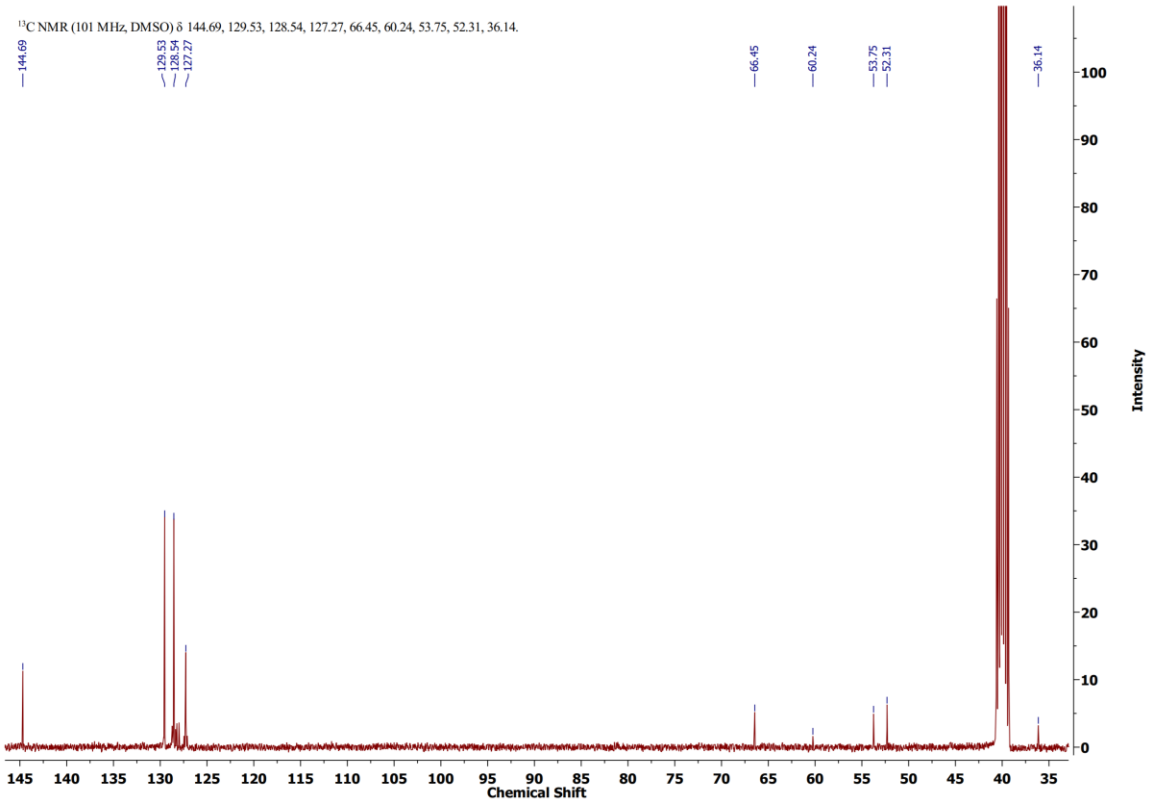
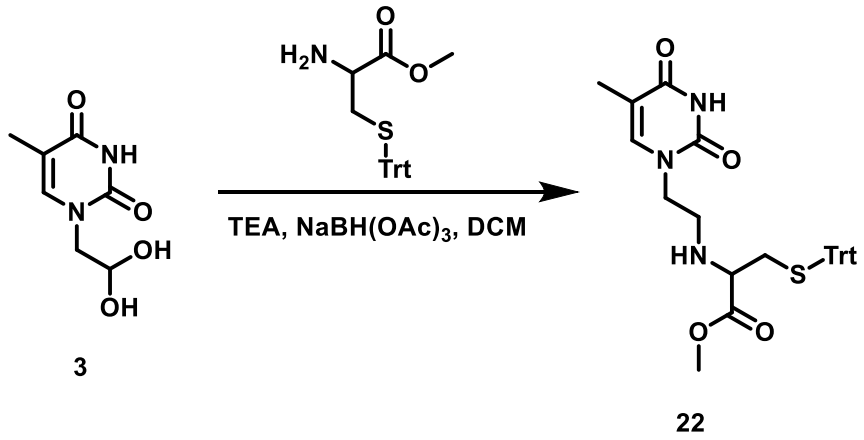


Figure D4: CNMR spectrum of compound **21**



3

22

Methyl 2-{[2-(5-methyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-1-yl)ethyl]amino}-3-[(triphenylmethyl)sulfanyl]propanoate: Tritylcysteine methyl ester (2.67 g, 7.08 mmol) was added to a 250 mL round bottom flask that had been oven dried and purged with nitrogen. Anhydrous DCM (100 mL) was added followed by compound **3** (1.09 g, 5.89 mmol) and triethylamine (1.64 mL, 11.79 mmol). After stirring for one hour at room temperature, the flask was cooled in an ice bath for 15 min. Sodium triacetoxyborohydride (1.75 g, 8.26 mmol) was added and the flask warmed to room temperature as it stirred overnight. The flask's contents were transferred to an extraction vessel with 200 mL water and an additional 100 mL DCM. The aqueous layer was washed twice with 100 mL water, the organic layers were combined, and then dried over magnesium sulfate. After filtering, the DCM was concentrated and the product was purified using flash chromatography with a gradient to 5% MeOH in DCM. Collecting the desired fractions yielded 0.97 g of compound **22** as an oil (26%).

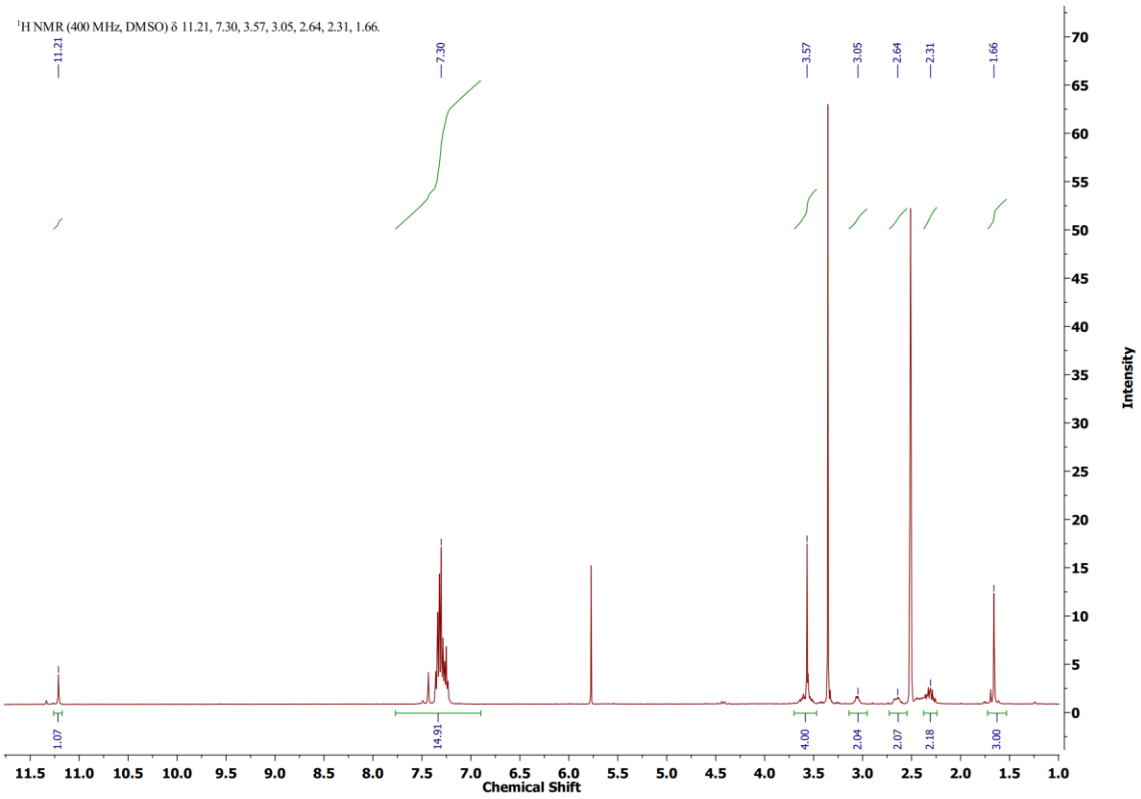


Figure D5: HNMR spectrum of compound **22**

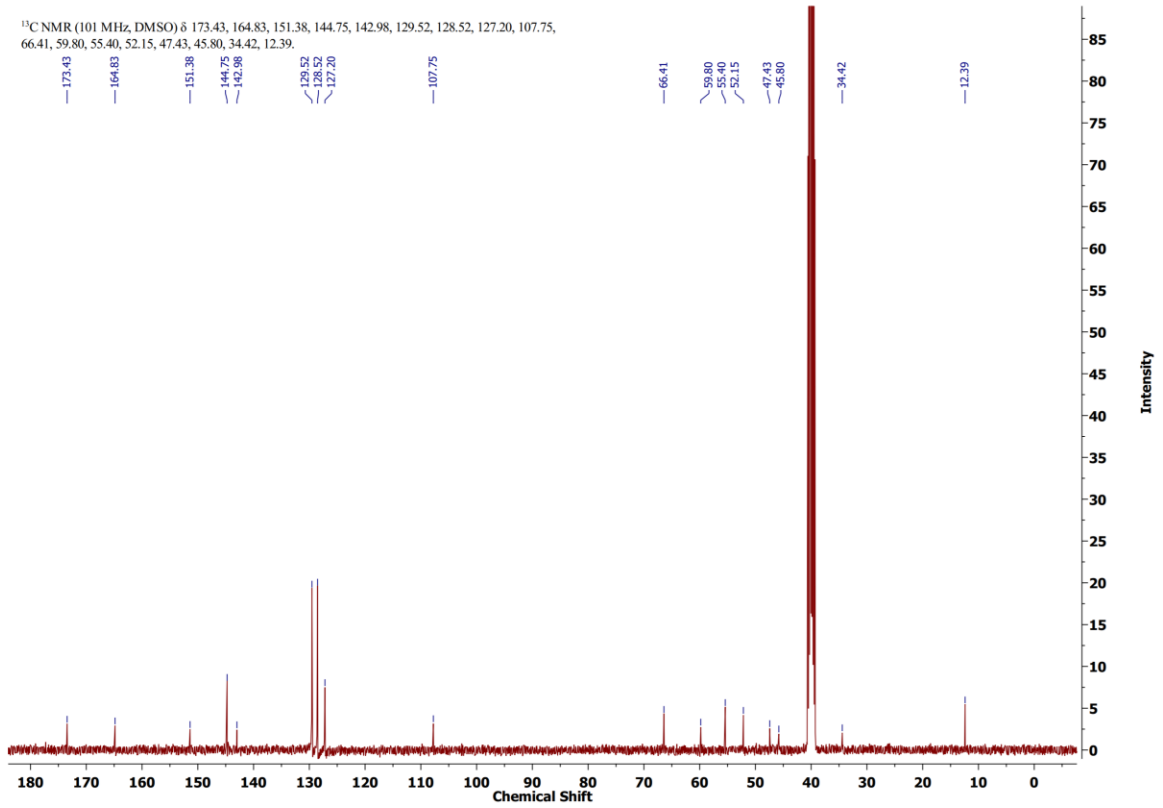
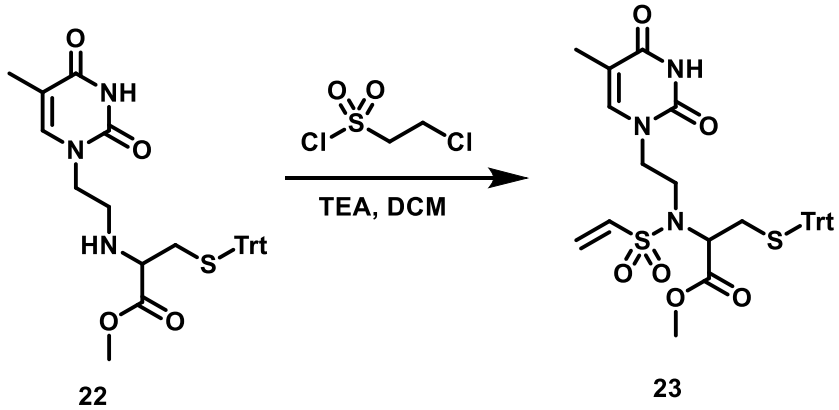


Figure D6: CNMR spectrum of compound **22**



22 **23**

Methyl 2-{N-[2-(5-methyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-1-yl)ethyl]ethenesulfonamido}-3-[(triphenylmethyl)sulfanyl]propanoate: Compound **22** (1.99 g, 3.77 mmol) was added to a 100 mL round bottom flask that had been oven

dried and nitrogen purged. 50 mL anhydrous DCM was added followed by anhydrous triethylamine (1.58 mL, 11.30 mmol). The flask was then cooled in an ice bath for 15 min after which 2-chloroethanesulfonyl chloride (492 µL, 4.71 mmol) was added slowly. The mixture stirred at room temperature for four hours. The flask's contents were transferred to an extraction vessel with 200 mL water, 100 mL saturated sodium chloride solution, and an additional 100 mL DCM. The aqueous layer was washed twice with 100 mL water, the organic layers were combined, and then dried over magnesium sulfate. After filtering, the DCM was concentrated and the product was purified using flash chromatography with a gradient to 2.5% MeOH in DCM. Collecting the desired fractions and evaporating the solvent under reduced pressure yielded 779 mg of compound **23** as an orange solid (33.4%).

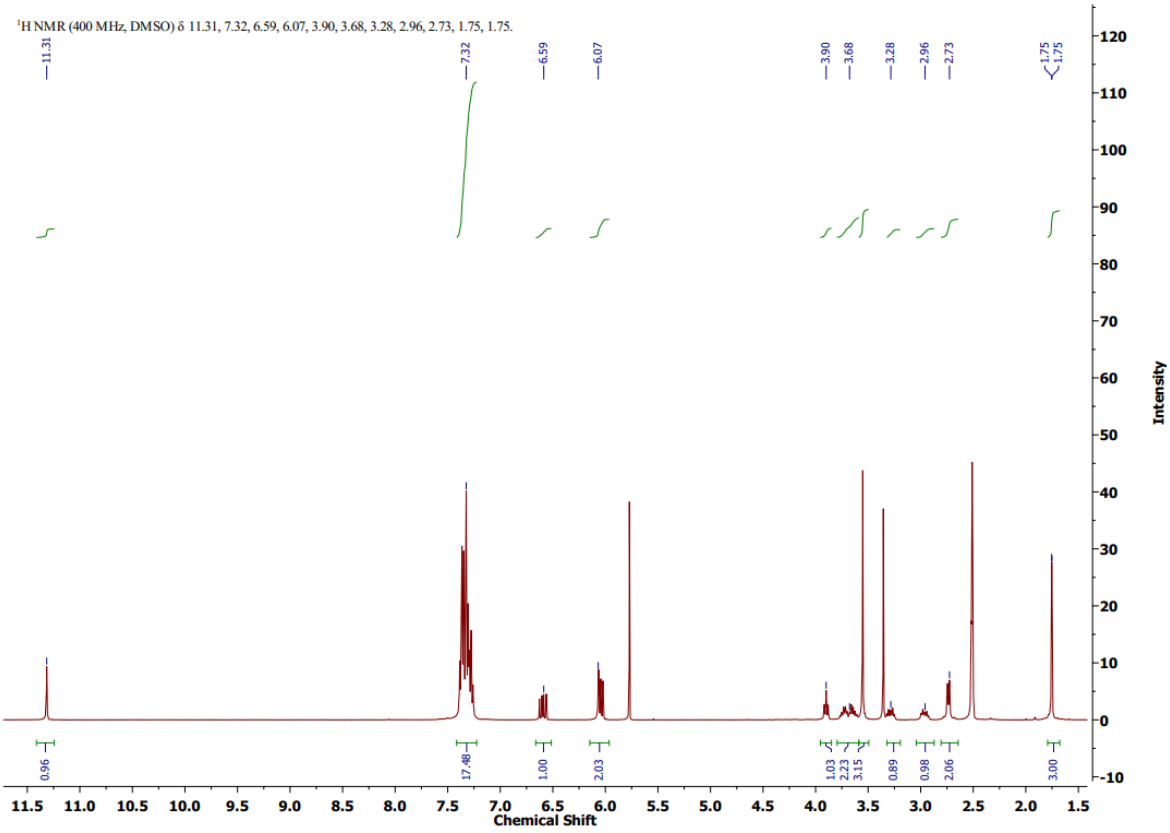


Figure D7: HNMR spectrum of compound **23**

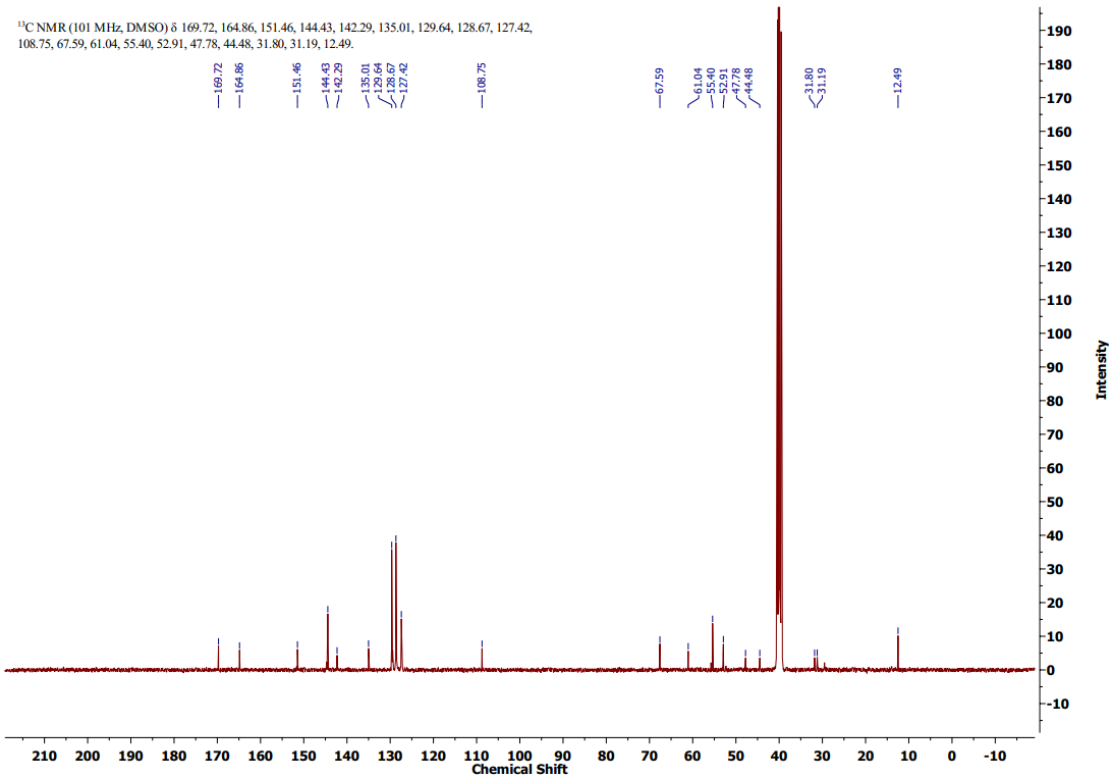
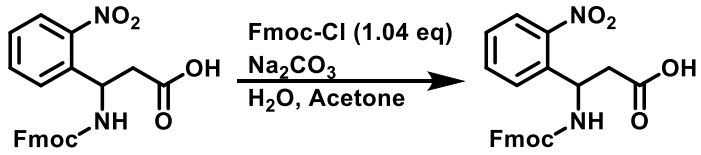


Figure D8: CNMR spectrum of compound **23**

Appendix E

OTHER SYNTHETIC PROCEDURES



24

3-(9-fluorenylmethyloxycarbonyl)amino-3-(2-nitrophenyl)propionic acid (ANP) (adapted from ref. 10): Sodium carbonate (2.5 g) was added to a 250 mL round bottom flask that had been oven dried and nitrogen purged. Water (25 mL) and acetone (12 mL) were then added to the flask followed by the addition of 2-nitrophenylglycine hydrochloride (2.79 g, 12 mmol). In a separate 100 mL round bottom flask, fluorenylmethyloxycarbonyl chloride (3.24 g, 12.5 mmol) was dissolved in acetone (40 mL). The Fmoc solution was added slowly over 20 min on ice and the mixture remained on ice for an hour. The mixture warmed to room temperature overnight. The reaction mixture was transferred to a 1 L extraction vessel and 400 mL of cooled water were added with diethyl ether (300 mL). The aqueous layer was collected in a beaker and the pH was adjusted to one. The solid precipitate was filtered over a medium fret and dried. Compound **24** was collected to give a yield of 71% (3.7 g).

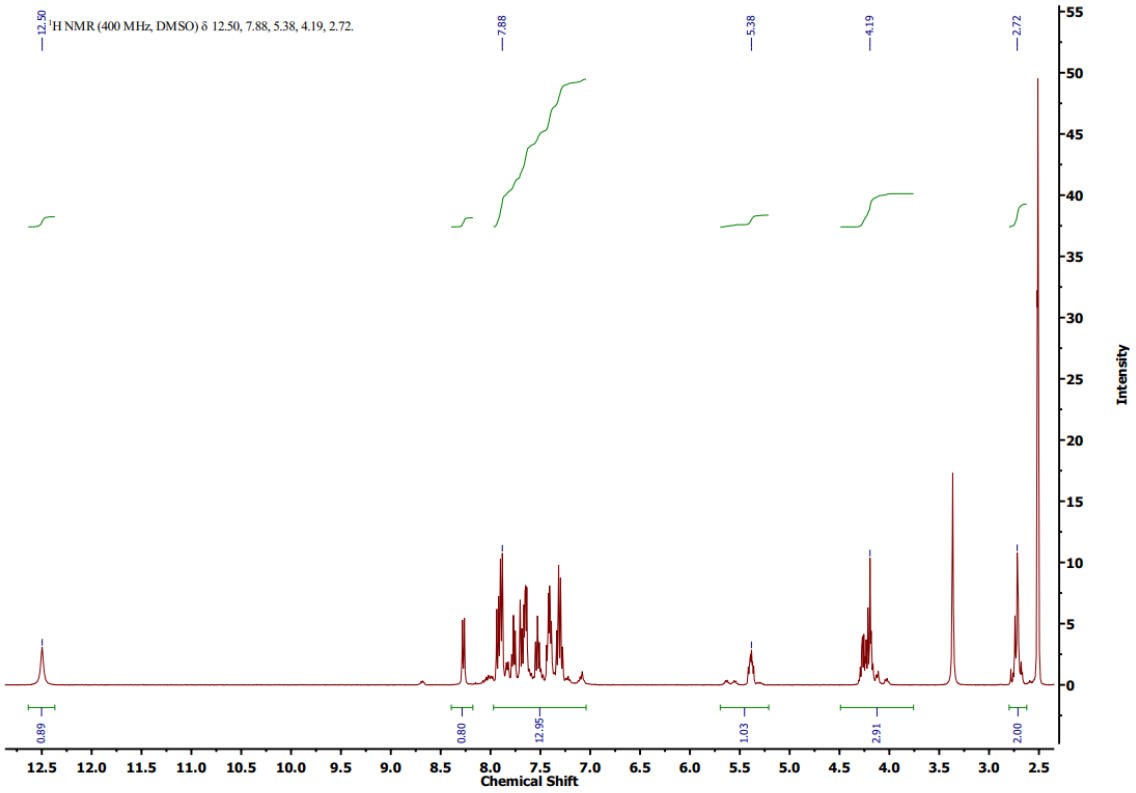


Figure E1: HNMR spectrum of compound **24**

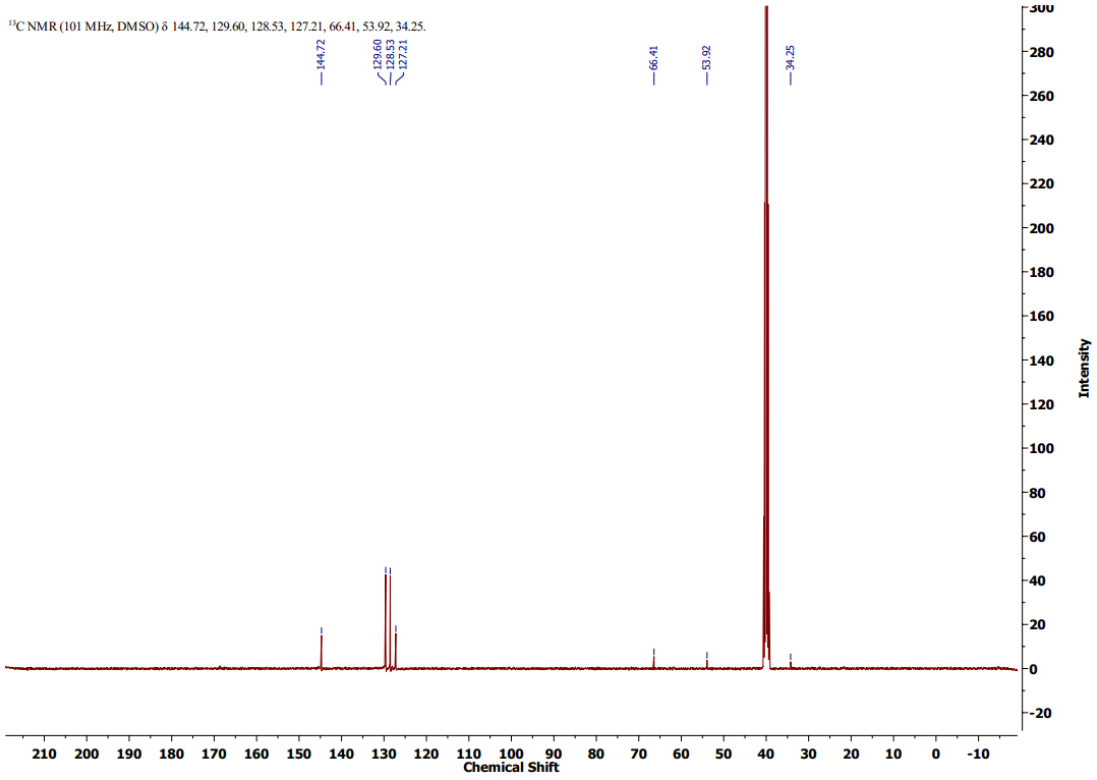
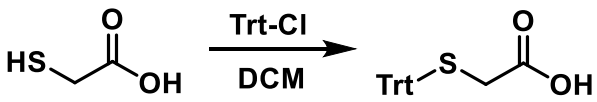


Figure E2: CNMR spectrum of compound **24**



25

2-(tritylthio)acetic acid: 3-mercaptopropanoic acid (2.40 g; 22.61 mmol) was added to a round bottom flask that had been oven dried and purged with nitrogen gas. Anhydrous DCM (150 mL) was added followed by trityl chloride (16.8 g, 60 mmol). The mixture stirred for 6 hours and the solvent was evaporated at a reduced pressure. The crude residue was dissolved in 50 mL diethyl ether and it was evaporated at a reduced pressure. The crude product was cooled to room temperature and it was then

suspended in water, filtered, and triturated with water (2x 50 mL) and diethyl ether (50 mL) to afford compound **25** as a white powder with 80% yield (18.7 g).

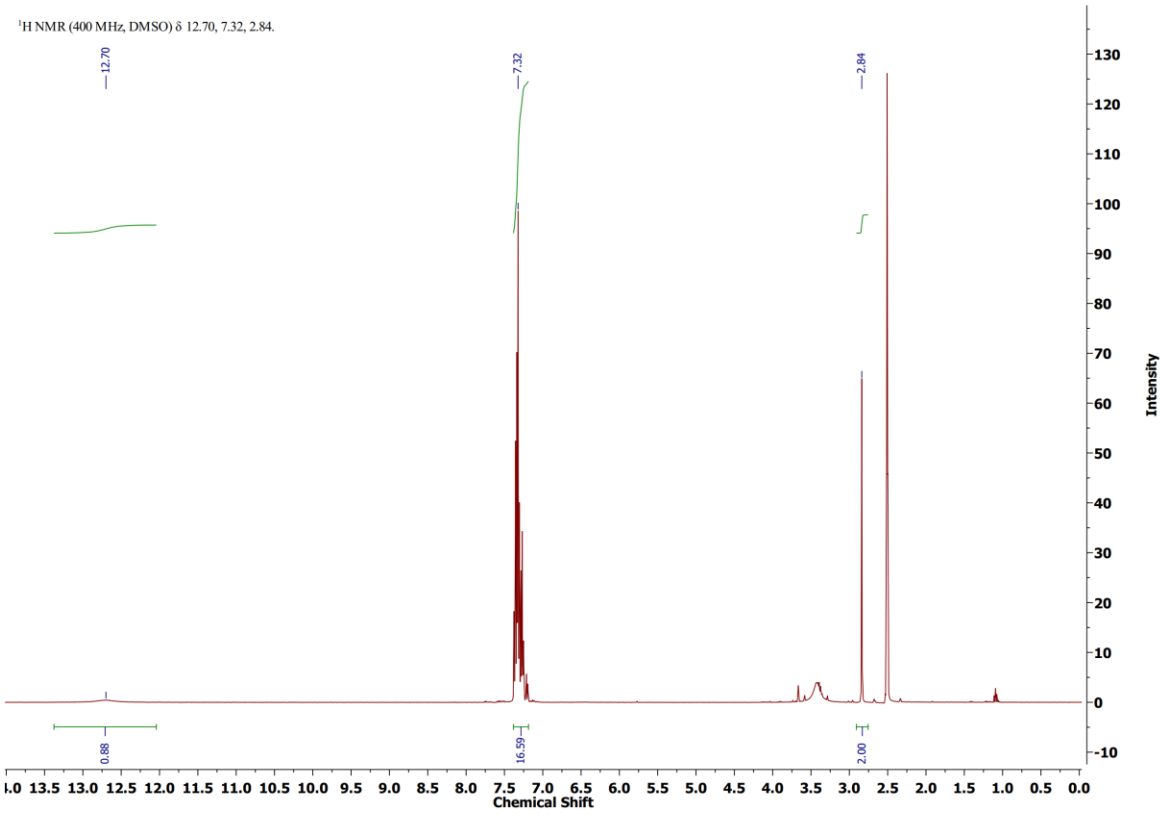
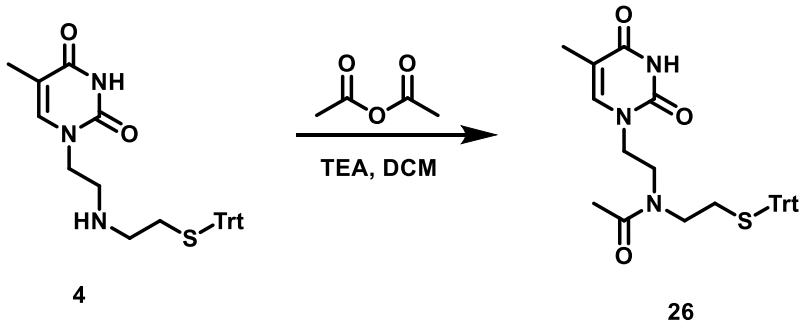


Figure E3: HNMR spectrum of compound **25**



4 26

N-[2-(5-methyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-1-yl)ethyl]-N-{2-[(triphenylmethyl)sulfanyl]ethyl}acetamide: Compound 4 (2 g, 4.24 mmol) was added to a 100 mL round bottom flask that had been oven dried and nitrogen gas purged. Anhydrous DCM (40 mL) was added followed by anhydrous triethylamine (1.18 mL, 8.49 mmol). Acetic anhydride (0.80 mL, 8.49 mmol) was added dropwise. After 2 hours of monitoring via TLC, another equivalence of TEA was added. The reaction stirred overnight at room temperature. The next day, 50 mL of water was added and the organic layer was extracted with two washed of 50 mL of water. The organic layer was collected and dried over anhydrous sodium sulfate. The sample was then dry loaded for purification with flash chromatography using a gradient of 0% to 5% methanol in DCM. The product was collected to yield 2.16 g of compound 26 (99%).

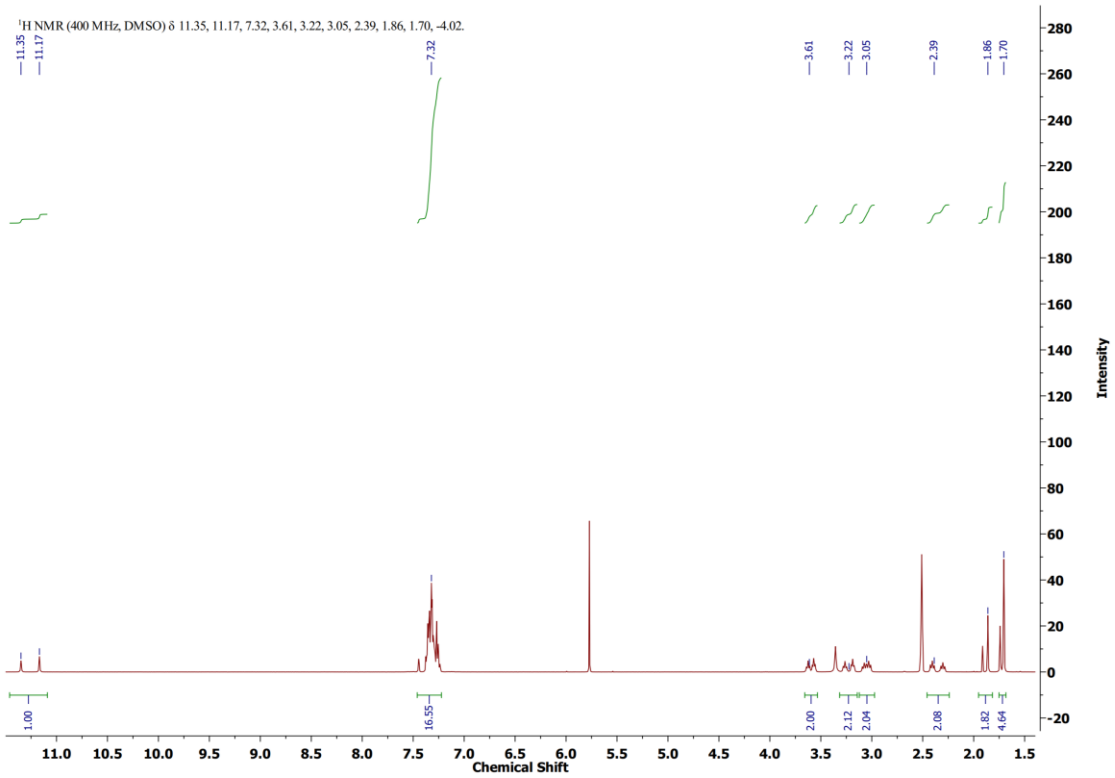


Figure E4: HNMR spectrum of compound **26**

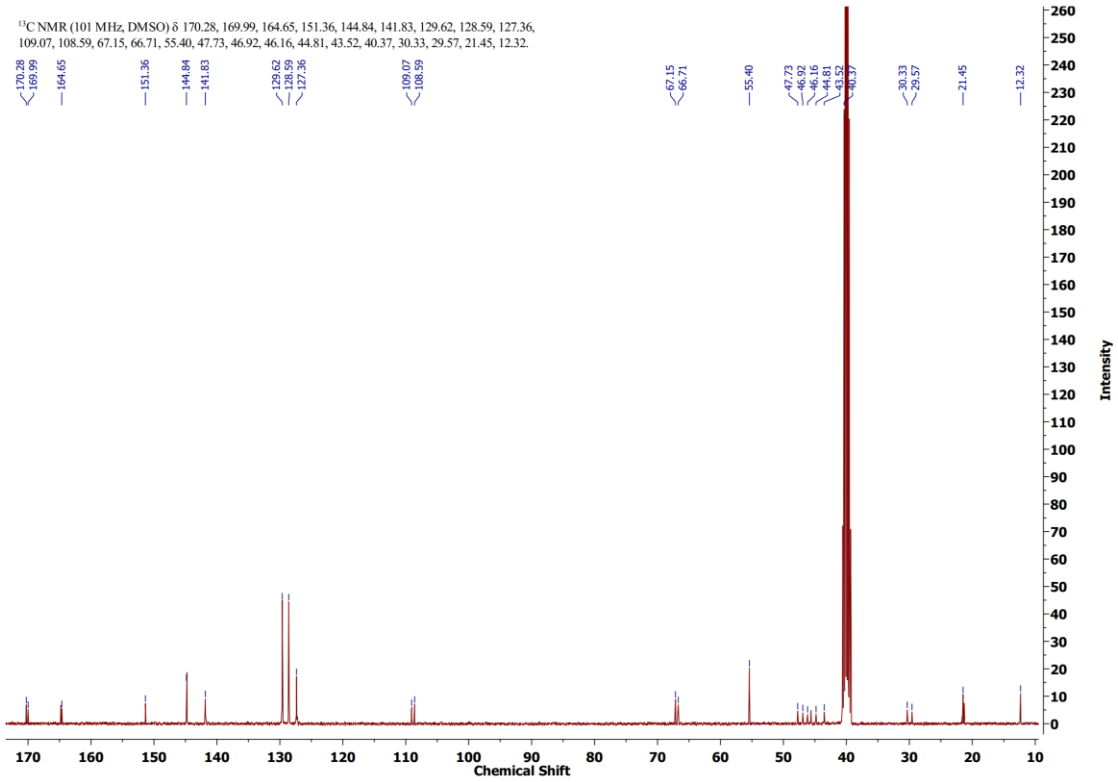
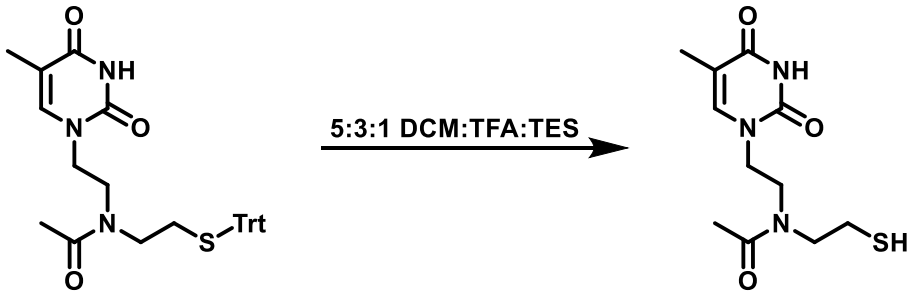


Figure E5: CNMR spectrum of compound **26**



26 **27**

N-[2-(5-methyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidin-1-yl)ethyl]-N-(2-sulfanylethyl)acetamide: Compound **26** (2.16 g, 4.21 mmol) was added to a 100 mL round bottom flask that had been oven dried and nitrogen gas purged. 45 mL of a solution containing a volumetric ratio of 5:3:1 of DCC:TFA:TES was added. The

reaction stirred for 15 min after which it was concentrated using the rotary evaporator (not to dryness). The concentrate was transferred to an extraction vessel and a saturated sodium bicarbonate solution was added slowly to avoid excessive bubbling. The DCM was collected and dried over anhydrous sodium bicarbonate. The DCM was then evaporated via rotary evaporation and 50 mL of diethyl ether was added. The product crystalized over 48 hours in a refrigerator. The resulting solid was filtered and collected as 71 mg compound **27** (6.2%).

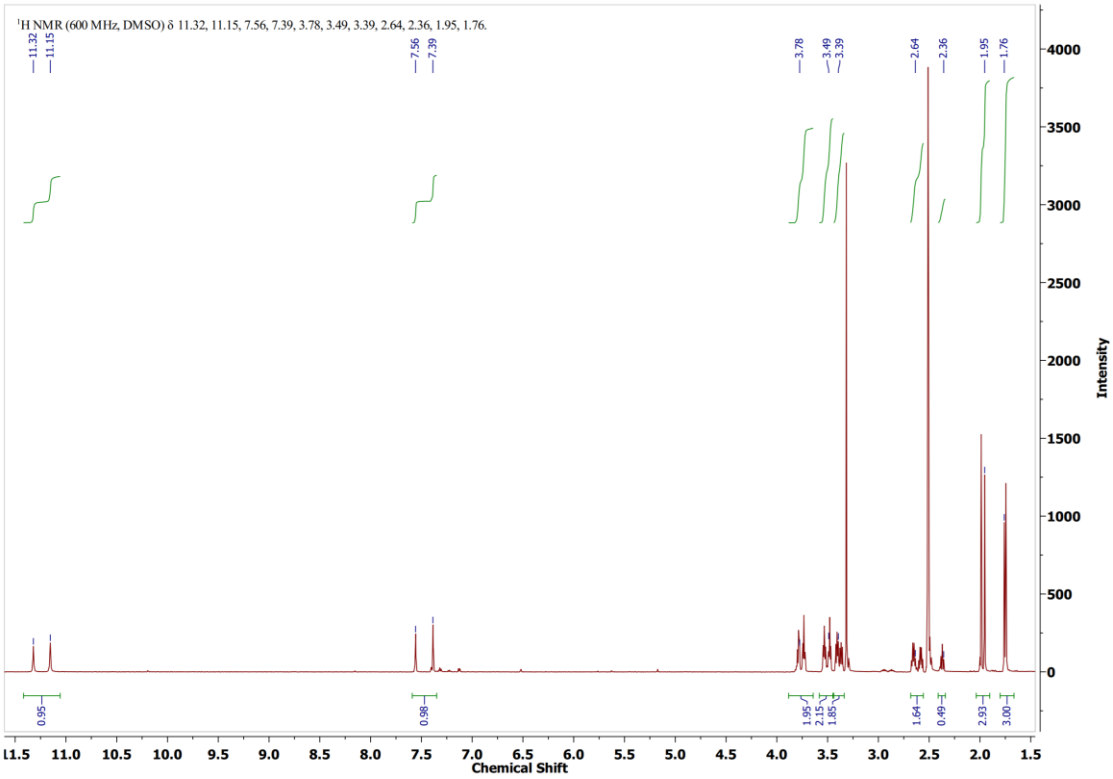


Figure E6: HNMR spectrum of compound **27**

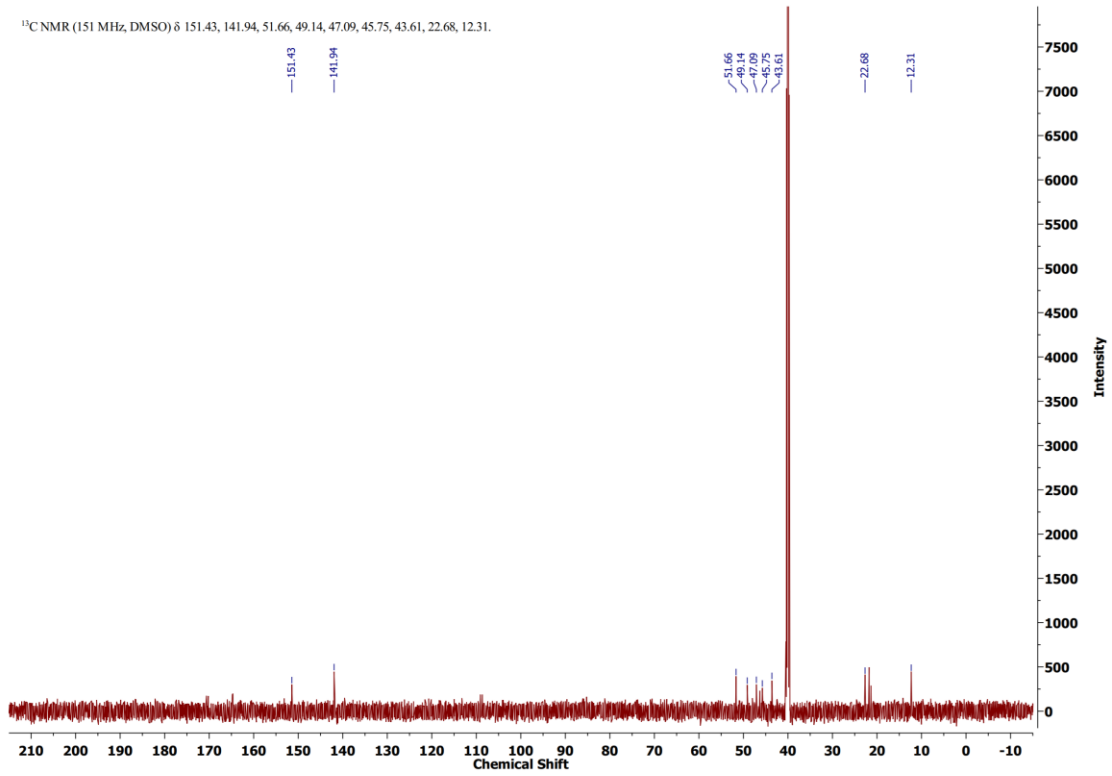


Figure E7: CNMR spectrum of compound **27**