7-26-2018 10:45 AM

Synthesis of 2-thiouracil and 2-thiothymine peptide nucleic acid monomers compatible with fmoc-based oligomerization

Timothy Martin-Chan, *The University of Western Ontario*

Supervisor: Hudson, Robert H.E., *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Chemistry

Abstract

Pseudo-complementary nucleobases are modified nucleobases which exhibit normal base pairing with unmodified complementary nucleobases; however, will not undergo stable base pairing with other pseudo-complementary nucleobases. The pseudo-complementary nature of these bases is of interest for the synthesis of pseudo-complementary peptide nucleic acid (pcPNA) oligomers, which are capable of binding targets in double-stranded DNA through a double duplex strand invasion. Until recently, pcPNAs were synthesized via Boc-based oligomerization chemistry which is currently less popular than the Fmoc-based strategy.

This research describes the development of a Fmoc protected thiouracil PNA monomer with an acid labile protecting group for use in standard SPPS. The synthesis uses an orthogonal protection strategy where the nucleobase is deprotected under standard acidic deprotection and resin cleavage conditions, while oligomerization occurs under basic conditions.

Keywords:

Peptide nucleic acid, PNA, pseudo-complementary nucleic acid, pcPNA, double duplex invasion, thiouracil, protecting groups

Acknowledgements

I would first like to thank Dr. Robert H. E. Hudson, my supervisor, for his guidance, support and patience throughout my thesis. His guidance and endless patience have made this difficult journey possible.

I am grateful to have worked with amazing labmates and colleagues whose friendship and advice has always been appreciated. Their ideas, discussions and assistance have been invaluable in this process. I am especially grateful for the assistance from Ali Heidari towards the completion of this project.

I would like to acknowledge the hard work of Mathew Willans and Aneta Borecki in maintaining the NMR facilities which made characterization of compounds possible. I would also like to thank Doug Hairsine for running mass spectrometry on my compounds.

Finally, words cannot express my gratitude for the support, encouragement and love I have received from my parents and from my girlfriend, Meagan Wiederman. I am grateful for their support and encouragement which has made this journey easier.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

A	Adenine
AEG	Aminoethyl glycine
amu	Atomic mass units
bhoc	Benzhydryloxycarbonyl
bn	Benzyl
boc	*Tert*-butyloxycarbonyl
br	broad
C	Cytosine
cbz	Carboxybenzyl
CD	Circular dichromism
^{13}C NMR	Carbon nuclear magnetic resonance
D	2,4-Diaminopurine
d	doublet
DCC	*N,N'*-dicyclohexylcarbodiimide
dd	Doublet of doublets
DIPEA	Diisopropylethylamine
DMF	*N,N*-dimethylformamide
DMSO-d_6	Deuterated dimethyl sulfoxide
DNA	Deoxyribose nucleic acid
EI	Electron ionization
ESI	Electrospray ionization

Et	Ethyl
Fmoc	9-fluorenylmethyloxybarbonyl
G	Guanine
HBTU	Hexafluorophosphate benzotriazole tetramethyl uranium
HMBC	Heteronuclear multiple-bond correlation
^{1}H NMR	Proton nuclear magnetic resonance
HOBt	Hydroxybenzotriazole
HPLC	High-pressure liquid chromatography
HRMS	High resolution mass spectrometry
K	Lysine
LRMS	Low resolution mass spectrometry
Me	Methyl
MMOB	Monomethoxybenzyl
mRNA	Messenger ribose nucleic acid
NCS	*N*-chlorosuccinimide
NMR	Nuclear magnetic resonance
NXS	*N*-halosuccinimide
ODN	Oligodeoxynucleotide
pcPNA	Pseudo-complementary peptide nucleic acid
PNA	Peptide nucleic acid
q	Quartet
RNA	Ribose nucleic acid

RNase H	Ribonuclease H
s	Singlet
SPPS	Solid phase peptide synthesis
^{S}T	2-Thiothymine
^{S}U	2-Thiouracil
t	Triplet
T	Thymine
TFA	Trifluoroacetic acid
TFMSA	Trifluoromethylsulfonic acid
THF	Tetrahydrofuran
T_m	Melting temperature
TMSCl	Trimethylsilane chloride
TPP	Triphenylphosphine
TPPO	Triphenylphosphine oxide

Chapter 1: General introduction

1.1. Discovery of DNA

Nuclein, first isolated by Swiss doctor Friedrich Miescher from leukocytes in the pus on surgical bandages as early as 1868, became one of the most important discoveries in the fields of biochemistry and chemical biology. While studying the chemical composition of leukocyte cells, Miescher discovered a precipitate which was resistant to protease digestion, indicating the novel compound was neither a protein nor a lipid and was instead a new class of biological molecules.[1] Miescher termed his discovery nuclein due to its extraction from the nucleus of the cells. Miesher's discovery was later renamed as 'nucleic acid' due to its acidic properties.

In 1909, Phoebus Levene discovered the sugar, phosphate, and nucleobase components of deoxyribose nucleic acids (DNA), and identified four different nucleobases: adenine, thymine, cytosine, and guanine.[2] However, it was not until the mid-twentieth century with foundational discoveries in DNA research that the importance of nucleic acids in biological chemistry became clear. In 1944, Oswald T. Avery, Colin MacLeod and Maclyn McCarty proposed that nucleic acids were the carrier of genetic information, rather than protein which had long been believed.[3] Research by Erwin Chargaff showed that purine and pyrimidine bases exist in a 1:1 ratio, demonstrating that the number of adenine bases equaled that of thymine and the number of guanine bases equaled that of cytosine.[4-7] Finally, in 1953, working from x-ray analysis by Rosalind Franklin and Maurice Wilkins, James Watson and Frances Crick proposed the double helical structure of DNA and the mechanism by which hereditary information is copied.[8]

1.2. Nucleic acid chemistry

Nucleic acids are an essential component of all known life, with DNA containing the hereditary information for all prokaryotes and eukaryotes. DNA contains a sugar-phosphodiester backbone, composed of β-2-deoxyribose sugars combined by phosphate at the 3-to-5 hydroxyl of the sugars. Each sugar contains one of four nucleobases adenine, cytosine, guanine, or thymine attached to the C1 position of the sugar. Nucleic acid strands can undergo hybridization between complementary base pairs adenine and thymine, and cytosine and guanine to form a double stranded duplex. The hydrogen bonding pattern between nucleobases provides specificity of base pairing, with adenine forming two hydrogen bonds with thymine while cytosine forms three hydrogen bonds with guanine, preventing mismatched hybridization (Figure 1.1). The duplex is stabilized due to hydrophobic interactions between the nucleobases and hydrophilic interactions between the charged phosphate backbone and an aqueous environment, hydrogen bonding between complementary nucleobases, and π-stacking from the aromatic purine and pyrimidine nucleobases. Mismatched nucleobases destabilize the duplex, disfavouring hybridization from occurring, resulting in sequence specificity for strand hybridization.

The specificity resulting from base complementarity provides a foundation for the design of nucleic acid oligomers capable of specific hybridization with DNA or RNA targets, known as molecular probes. This specificity allows the design of nucleic acid probes for research and genetic screening, as well as potential design of nucleic acid based therapeutic agents which rely on high specificity targeting of a genetic sequence.

Oligonucleotide sequences provide the required specificity through base pairing between complementary nucleobases. This targeting of a specific genetic sequence can prevent undesired off-target effects.

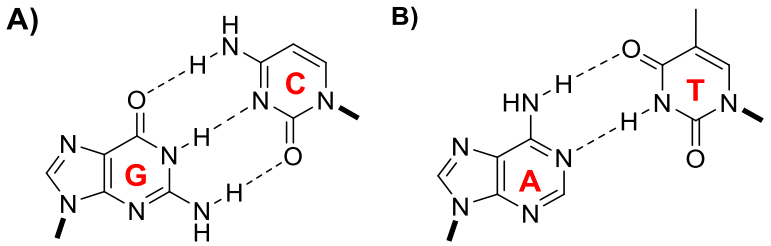


Figure 1.1. Base paring between nucleobases A) guanine and cytosine and B) adenine and thymine

1.3. Modified nucleic acids and DNA analogs

A major limitation to oligodeoxynucleotides (ODNs) used for *in vivo* applications results from the chemical instability of nucleic acid strands. ODN strands are susceptible to hydrolysis by nucleases, enzymes responsible for cleaving phosphodiester bonds in the breakdown of DNA and RNA, preventing ODNs from reaching the desired target. Many chemical modifications have been developed to increase the stability of the nucleic acid oligomers. In many cases, modifications to the sugar-phosphodiester backbone increase the stability *in vivo* as it prevents recognition and degradation by nucleases. Modification such as replacing the phosphate units with analogs such as phosphothioates or phosphoborates, or replacing the ribose sugar provide resistance to nucleases. However, these modifications often introduce chiral centers, and reduce the binding affinity between the oligomer and the target DNA or RNA strand.

1.4. Peptide nucleic acid chemistry

One such backbone modification replaces the sugar-phosphodiester backbone with an aminoethyl glycine backbone with the nucleobases attached via a methylene carbonyl linker to form a peptide nucleic acid (PNA) (Figure 1.2.).[9] The modified backbone is oligomerized through peptide bonds between a N-terminal amine, and a C-terminal carboxylic acid to form a PNA oligomer. Advantages of PNAs include, the aminoethyl glycine backbone maintains natural spacing observed with the sugar-phosphodiester backbone, allowing normal Watson-Crick base pairing with unmodified DNA, RNA and PNA strands. The unnatural backbone is resistant to degradation from either nucleases or proteases as the backbone is not recognized as either an oligonucleotide or a polypeptide for cleavage. The PNA oligomers undergo more stable hybridization with DNA and RNA than ODNs due to the neutral aminoethyl glycine backbone, ODNs experience anionic repulsion due to the negatively charged phosphate groups present in the backbone. Additionally, due to the neutral backbone, PNA hybridization is independent of ionic concentrations. PNAs have high sequence specificity for hybridization, with a single base mismatch significantly destabilizing a PNA-DNA duplex. The aminoethyl glycine backbone is achiral, allowing for easy synthesis of PNA monomers without difficult purification, and the amide bonds allow for easy oligomerization using standard solid phase peptide synthesis (SPPS) techniques.

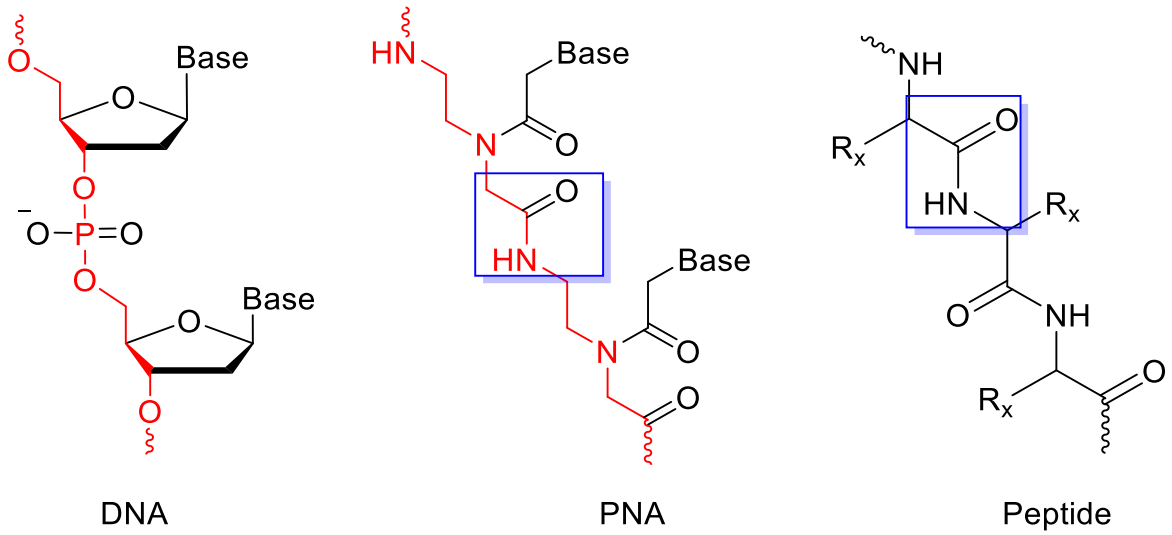


Figure 1.2. Structure of deoxyribose nucleic acid oligomer (DNA), peptide nucleic acid oligomer (PNA) and amino acid oligomer (peptide)

1.5. Peptide nucleic acids as therapeutic agents and molecular probes

PNA oligomers have potential applications in the design of gene targeting therapeutic agents. These therapeutic agents typically target a mRNA sequence; hybridization with the target mRNA inhibits translation preventing protein synthesis, known as the antisense strategy. Alternatively, gene therapeutic agents can be designed to specifically recognize and hybridize a target gene, thereby inhibiting transcription and preventing gene expression (antigene strategy). Studies of PNA oligomers have shown potential for applications as both antisense and antigene therapeutic agents, demonstrating the ability to inhibit both translation and transcription.

Antisense inhibition using PNA oligomers has been demonstrated to reduce translation in *in vitro* experiments. Inhibition occurs through steric blocking of translation which occurs either through hybridization overlapping the AUG start codon blocking

assembly of the ribosome complex at the start codon or through triplex formation with the coding region, arresting elongation resulting in the formation of truncated peptide.[10-12] The limitation of PNA based antisense therapies is that hybridization with target mRNA does not activate ribonuclease H (RNase H) -mediated cleavage of the RNA transcript, which is a more effective method of antisense inhibition. RNase H is an endonuclease enzyme which cleaves an RNA target using DNA (or a DNA analog) as a template causing inhibition of mRNA translation.[13, 14] RNase H catalyzes the hydrolysis of the target RNA which is more efficient inhibition compared to steric blocking of the mRNA transcript which requires a 1:1 ratio of PNA:mRNA to inhibit expression. Unlike some oligonucleotide analogs such as phosphorothioates,[15] PNA:RNA hybridization causes a structural change which is not recognized by the RNase H enzyme, limiting its effectiveness as an antisense gene therapy.

Antigene therapies involve the invasion of the DNA duplex using an oligonucleotide analog to hybridize to the target gene and sterically block the function of RNA polymerase. *In vitro* studies suggest that PNA oligomers are capable of inhibiting transcriptional expression of both prokaryotic and eukaryotic RNA polymerase by hybridizing to either the gene promoter region or downstream of the promoter blocking RNA polymerase from accessing the template DNA strand.[16, 17] Hybridization to the promoter region inhibits RNA expression by sterically blocking assembly of the RNA polymerase complex to the gene, while triplex invasion downstream of the promoter sterically blocks the progression of RNA polymerase, arresting transcription and resulting in the formation of truncated RNA transcripts. PNA oligomers have demonstrated

potential as antigene therapies, however PNA invasion of a DNA duplex requires the formation of a PNA_2:DNA triplex for invasion to occur limiting targets to homopurine sequences, which can undergo Watson-Crick and Hoogsteen base pairing. Furthermore, triplex formation is kinetically slow at physiological salt conditions, limiting PNA_2:DNA triplex hybridization.[18]

1.6. PNA hybridization

The stability of oligonucleotide duplexes can be determined as a measure of the thermal stability of the duplex. The thermal melting temperature (T_m) is defined as the temperature at which 50 % of the oligonucleotide duplex has dissociated. The T_m of oligonucleotides can be easily measured spectroscopically, due to the hypochromatic effect observed during hybridization. Free nucleobases absorb strongly around 260 nm, upon hybridization base pairing causes a reduction in the nucleobase absorbance.[19] This change in absorbance allows easy determination of T_m values, using absorption spectroscopy, which is a convenient method of analyzing relative oligonucleotide duplex stability. Typically, PNA-DNA duplexes have higher T_m values compared to a DNA-DNA duplex of identical sequence, while a PNA-PNA duplex has a higher T_m than the PNA-DNA duplex, resulting in a relative duplex stability: PNA-PNA>PNA-DNA>DNA-DNA.

The stability of the PNA-DNA duplex over the DNA-DNA duplex results from reduced electrostatic repulsion from the neutral aminoethyl glycine backbone, compared to the phosphate diester backbone. Thermodynamic studies of PNA-PNA and PNA-DNA hybridization show the lack of charged backbone typically results in more favourable

enthalpies of hybridization for the formation of PNA-PNA and PNA-DNA duplexes compared to the formation of a DNA-DNA duplex. In the case of PNA-DNA hybridization, a significant entropic contribution results from reduced ion association with the duplex backbone. The polyanionic phosphodiester backbone of DNA associates with Na^+ and K^+ counterions, while the neutral aminoethyl glycine backbone of PNA does not require counterions. Therefore, the reduced number of counterions associated with the PNA-DNA hybrid compared to the DNA-DNA duplex increases the disorder of the system, giving a more favourable entropic contribution. Circular dichromism (CD) spectroscopy shows helices formed by PNA-DNA hybridization closely resembles the B-helix formed by DNA-DNA hybridization suggesting hydrophobic interactions contribute to the stabilization of the PNA-DNA duplex. In contrast, CD spectroscopy of the PNA-PNA hybridization shows a rearrangement in the helical structure of the duplex, resulting in a more farourable enthalpic value, as a result of reduced electrostatic repulsion.[18, 20, 21]

1.7. DNA strand invasion

A major limitation of PNAs for many biological applications involving gene targeting is that PNA oligomers do not undergo duplex invasion of a double stranded DNA duplex. During duplex invasion, the displaced complementary strand remains bound upstream and downstream of invasion complex, thus the entropic benefit of replacing the anionic DNA strand with a neutral PNA strand is not obtained which decreases the free energy of hybridization making strand invasion unfavourable. One method to overcome this entropic restriction is to hybridize a complementary PNA oligomer to both strands of the complementary DNA duplex, thereby maximizing the enthalpic contribution from

strand hybridization. Lowering the energy of the invaded complex allows duplex invasion to occur in a process known as a double duplex invasion (Figure 1.3.). However, the double duplex invasion using PNA oligomers is hindered by the strong hybridization of PNAs, which results in highly stable PNA-PNA hybridization which is more favourable than PNA-DNA hybridization. As a result, the use of complementary PNA oligomers in a double duplex invasion, instead results in the formation of a stable PNA-PNA duplex leaving the target DNA duplex intact.

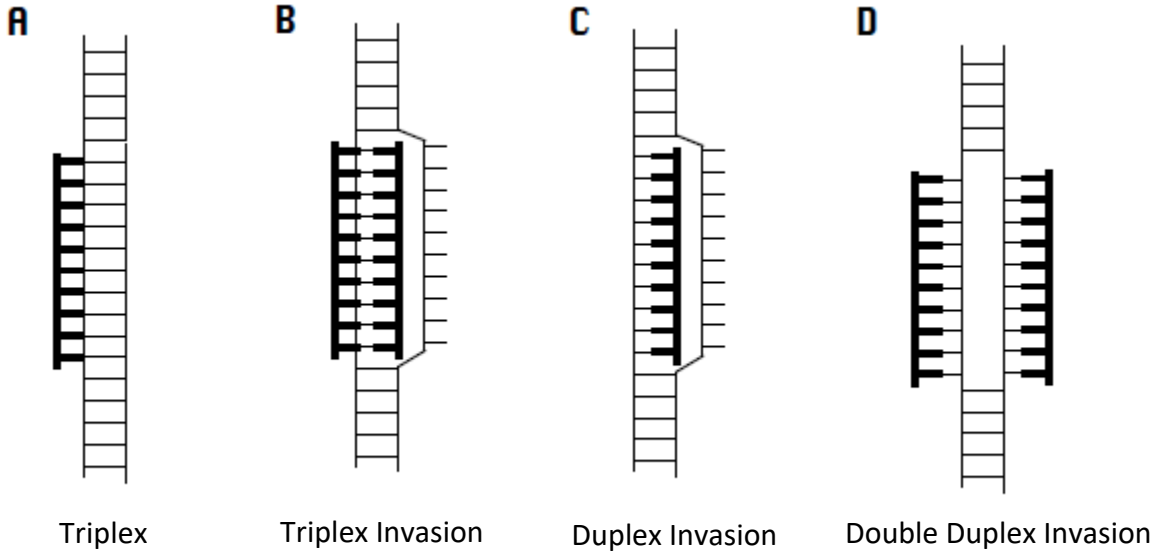


Figure 1.3. Modes of hybridization between PNA oligomer and nucleic acid duplex

A method of destabilizing the PNA-PNA duplex was developed by Nielsen *et al*[22] which utilized modified pseudo-complementary nucleobases to synthesis selectively binding oligomers, or pseudo-complementary PNAs (pcPNA). Pseudo-complementary nucleobases (Figure 1.4.), first reported by Gamper *et al*[23] in DNA oligomers, are modified complementary nucleobases which, when paired with an unmodified complementary

nucleobase undergo normal base pairing; however, when combined with another pseudo-complementary nucleobase steric hinderance prevents normal base pairing. The pseudo-complementary nucleobases are synthetic analogs of adenine and thymine, 2,4-diaminopurine (D) and 2-thiothymine (^{S}T) respectively, which in combination form pseudo-complementary base pairs. D, an analog of adenine containing an additional exocyclic amine at the 2 position, undergoes base pairing with thymine forming a more stable base pairing compared to A:T due to additional hydrogen bonding from the 2 position amine. ^{S}T, an analog of thymine which replaces the 2-position carbonyl with a thiol, undergoes normal base pairing with an unmodified adenine base. D:^{S}T base pairing is destabilized due to steric hinderance between the 2-position exocyclic amine on D and the larger sulfur atom which replaces oxygen in ^{S}T. Nielsen *et al*[22] reported the first synthesis of pcPNA oligomers using a boc-based SPPS approach and demonstrated their ability to undergo a double duplex invasion of a target DNA duplex.

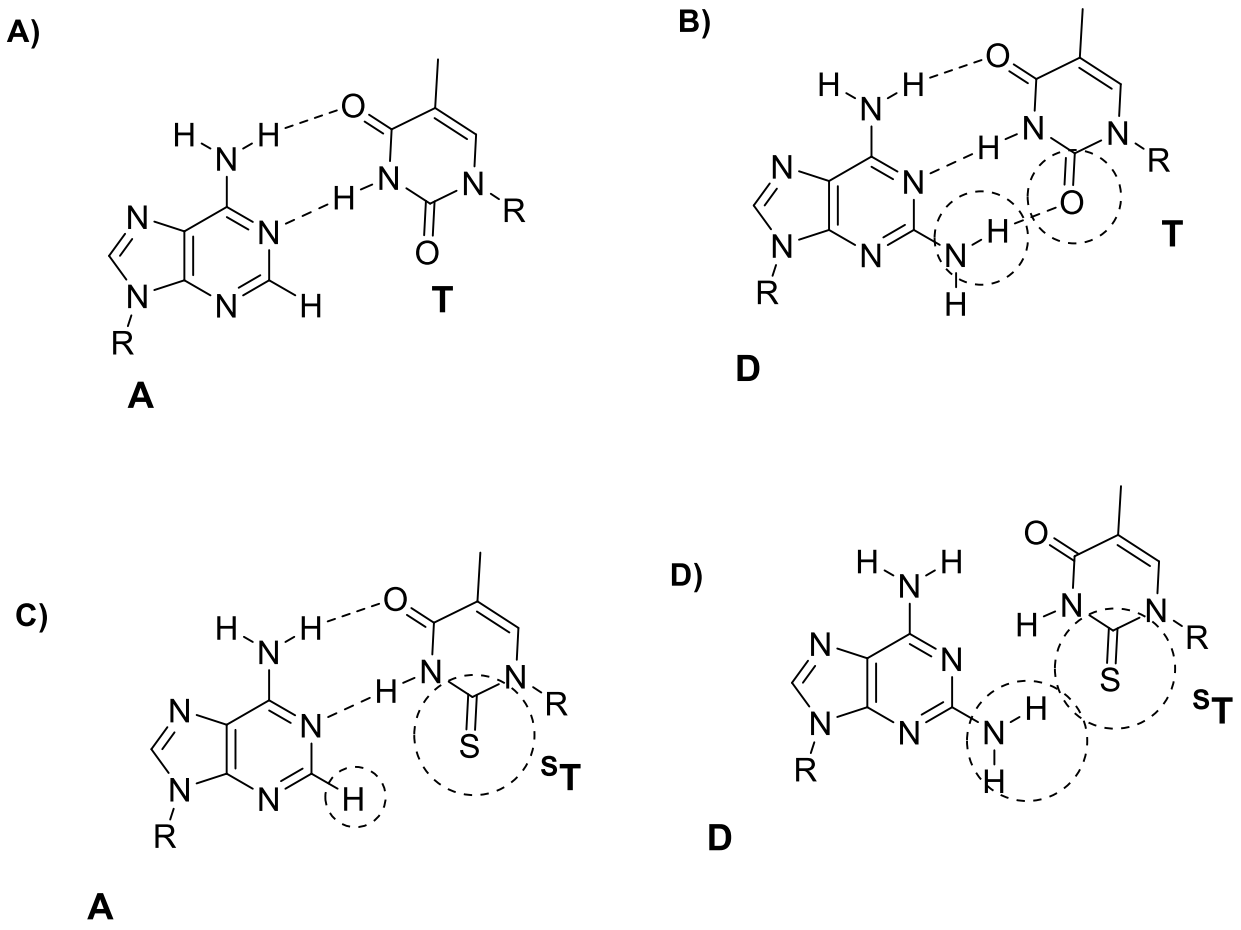


Figure 1.4. Base pairing of natural and pseudo-complementary nucleobases A) adenine and thymine, B) diaminopurine and thymine, C) adenine and thiothymine, D) diaminopurine and thiothymine

1.8. Solid phase peptide synthesis and protecting group strategy

Two methods of SPPS have primarily been used for chemical oligomerization in peptide synthesis, which primarily differ in the protecting group strategy. All amino acids contain both a carboxylic acid and an amine, which undergo a condensation reaction to form an amide linkage during oligomerization, which must be protected and require different deprotection conditions for control of sequence specific oligomerization. In addition, many amino acids contain reactive sidechains which require separate functional

group protection. Similarly, PNAs contain both a carboxylic acid and a primary amine which require separate protection and deprotection strategies, with most nucleobases requiring further protection during oligomerization. Historically, peptide synthesis was performed using a gradient acidolysis protection strategy where N-terminal boc-deprotection occurred under relatively weak acidic conditions, while C-terminal deprotection requires relatively stronger acidic conditions. This strategy for peptide synthesis was originally developed by Bruce Merrifield using a solid support resin for peptide synthesis,[24, 25] for which he was awarded the 1984 Nobel Prize in chemistry. Advances in chemical synthesis lead to the development of an orthogonal protection strategy with N-terminal fmoc group for which deprotection occurs under weakly basic conditions and C-terminal deprotection occurring under weakly acidic conditions. This strategy results in higher yielding oligomerization, and as a result can be used to synthesize longer peptide sequences. Both strategies have been modified for standard PNA oligomerization, with the standard monomers commercially available for synthesis

The original synthesis of PNA oligomers was reported by Egholm, Buchardt, Nielsen and Berg in 1992[26] which utilized a *tert*-butyloxycarbonyl group on the terminal amine, while the C-terminus of the oligomer was attached to a solid support, typically a cross-linked polystyrene resin. The nucleobases were protected using acid labile protecting groups so oligomer deprotection could occur concurrently with resin cleavage, under strongly acidic conditions. Thymine requires no protection as it is unreactive under oligomerization conditions, while the exocyclic amines on cytosine, adenine and guanine are protected using a carboxybenzyl (cbz) protecting group.[9, 26, 27] During monomer

synthesis, the C-terminal carbonyl is protected using a alkyl ester protecting group to prevent coupling between aminoethyl glycine molecules and to improve solubility in organic solvents.[28] The *N*α-boc protection is typically removed by treatment with trifluoroacetic acid (TFA), freeing the N-terminus for oligomerization from C-terminus to N-terminus. Oligomerization uses a coupling reagent such as hydroxybenzotriazole (HOBt) and *N,N'*-dicyclohexylcarbodiimide (DCC) which reacts with the carboxylic acid to form an active ester for amide bond formation. Post-synthesis treatment of the resin with hydrofluoric acid (HF) or trifluoromethylsulfonic acid (TFMSA) is used to cleave the synthesized oligomer from the solid support and to remove the cbz and bn protection of the nucleobases (Figure 1.5.).

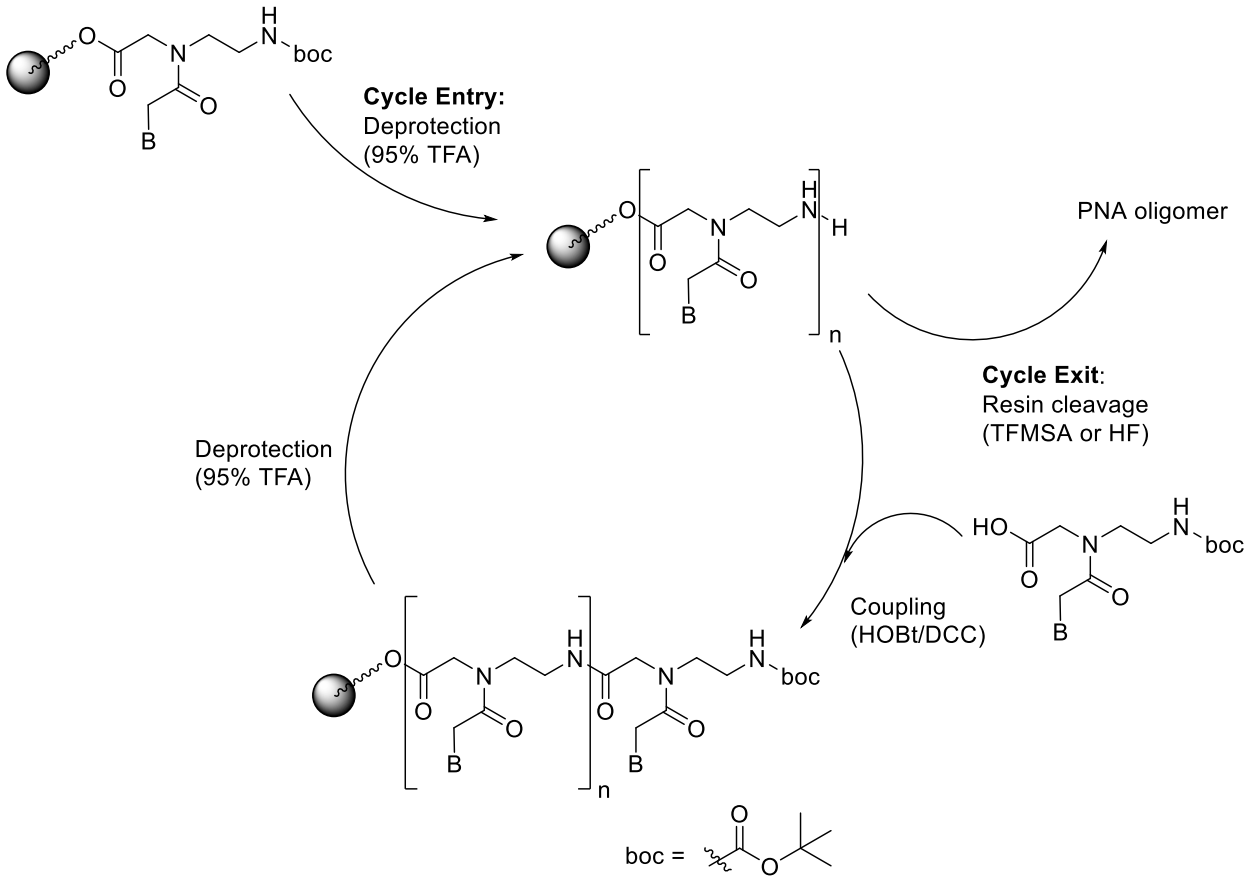


Figure 1.5. SPPS cycle for oligomerization with boc/cbz based protecting strategy

The limitation of boc-based peptide synthesis results from the gradient protection strategy. This can result in undesired cleavage of the PNA oligomer from the solid support resin, lowering the yield of oligomer synthesis. Low yielding oligomerization limits the maximum length of oligomer synthesis. Additionally, the final deprotection and resin cleavage requires strongly acidic treatment with hydrofluoric acid which can damage the oligomer and the harsh reagents can be dangerous to work with. An orthogonal protection strategy which uses a fluorenylmethyloxycarbonyl (fmoc) protecting group on the terminal amine, which requires weakly basic conditions for deprotection. Typical deprotection conditions use a solution of 20 % v/v piperidine solution in dimethylformamide to remove the fmoc protection, for oligomerization on the N-terminal (Figure 1.6.). The fmoc protecting group allows the use of more acid labile protecting groups for the C-terminal and for nucleobase protection. An acid labile linker is used to attach the polystyrene resin to the C-terminal of the oligomer. Nucleobase protection uses a more acid labile boc or benzhydryloxycarbonyl (bhoc) group to protect the exocyclic amines on cytosine, adenine, and guanine allowing resin cleavage and the final deprotection of the oligomer to be performed using a solution of 95 % v/v trifluoroacetic acid and 5 % triethylsilane as a cation scavenger. This orthogonal protection strategy uses milder deprotection conditions, and typically results in higher yielding oligomerization compared to the boc-based protection strategy. Due to these advantages, the fmoc-based protection strategy has become the standard for the chemical synthesis of peptides and PNA oligomers.

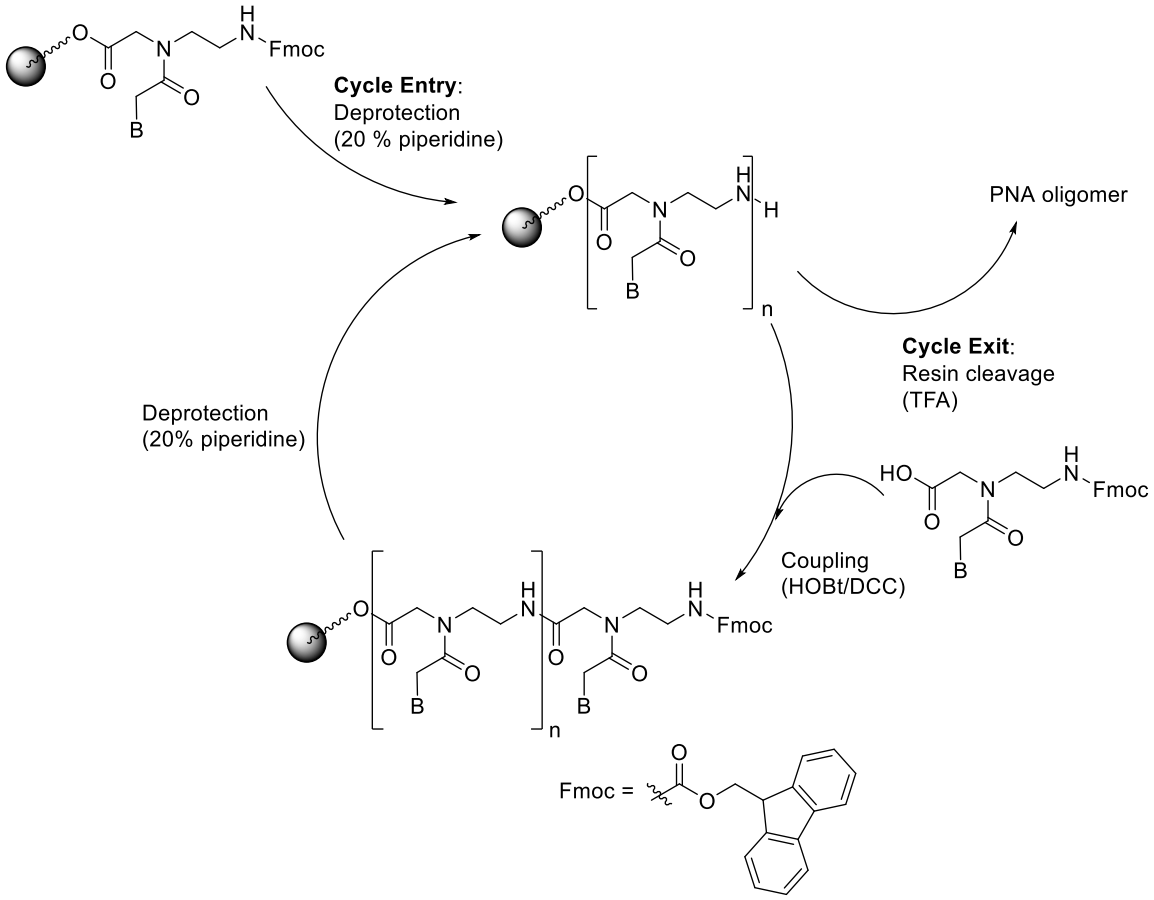


Figure 1.6. SPPS cycle for oligomerization with fmoc/boc based protection strategy

The pcPNA synthesis developed by Nielsen *et al* utilized the boc-based protection strategy, which at the time of development was the standard for peptide synthesis. Advances in peptide synthesis have limited the use of pcPNAs in nucleic acid research due to the undesirable method of oligomerization. However, several applications have been developed which demonstrates the potential for pcPNAs and the need for further research. Until recently,[29] only the boc-based ^{S}U monomer had been synthesized while both the boc protected and fmoc protected[30] D monomers, therefore the focus of this research is to synthesize a fmoc protected ^{S}U monomer compatible with standard fmoc-based SPPS.

1.9. Goals of thesis

Until recently, the synthesis of pcPNA oligomers was carried out using boc-based oligomerization strategy which has become unfavourable due to harsh deprotection conditions which result in lower yielding reactions thus limiting oligomer length. Therefore, the primary aim of this work is to develop a highly acid labile protecting group for thiouracil so that the PNA monomer is compatible with fmoc-based PNA oligomerization (Figure 1.7.). The 4-methoxybenzyl protecting group first used for boc-based synthesis was reportedly removed upon treatment with TFMSA or HF used for resin cleavage and final deprotection. To improve the lability of the protecting group, electron donating groups will be added to the benzyl based protecting group and the acidolysis of the protecting group tested under milder deprotection conditions.

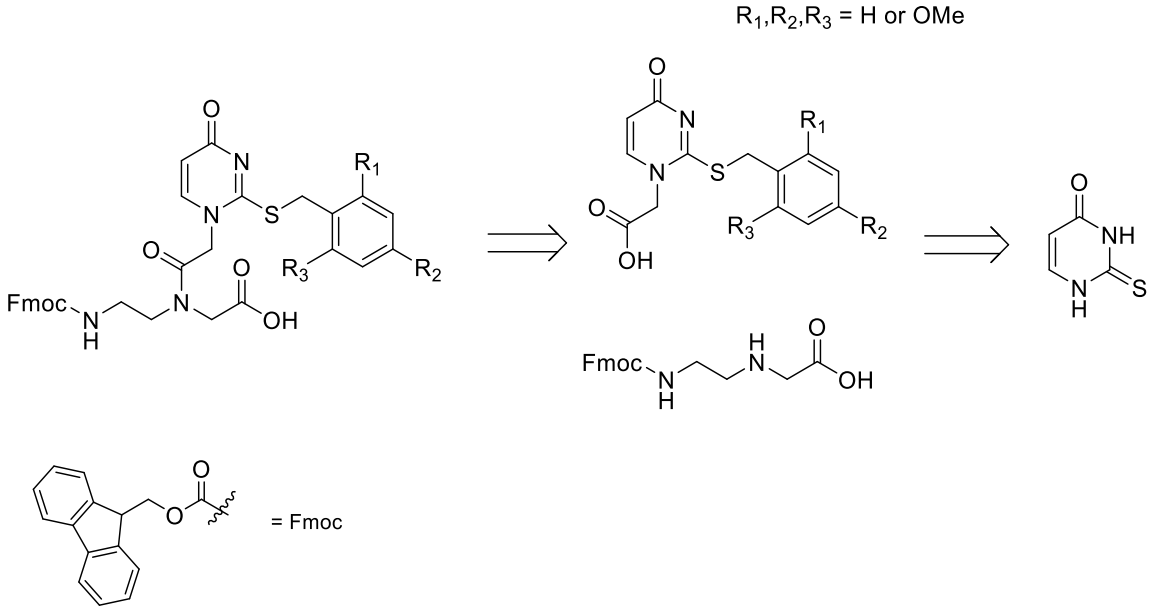


Figure 1.7. Proposed synthesis of highly acid labile protecting group for thiouracil PNA monomer

1.10. References

[1] Miescher, Friedrich, *Medicinisch-chemische Untersuchungen* **1871**, *4*, 441-460.

[2] Levene, Phoebus A., *Biochem. Z.* **1909**¸*17*, 120-131.

[3] Avery, Oswald; MacLeod, Colin M.; McCarty, Maclyn, *J. Exp. Med.* **1944**, *79*, 137-158.

[4] Chargaff, Erwin; Vischer, Ernst; Doniger, Ruth; Green, Charlotte; Misani, Fernanda, *J. Biol. Chem.* **1949**, *177*, 405-416

[5] Chargaff, Erwin; Zamenhof, Stephen; Green, Charlotte, *Nature* **1950**, *165*, 756-757.

[6] Chargaff, Erwin, *Experiencia* **1950***, 6*, 201-209.

[7] Zamenhof, Stephen; Brawermann, George; Chargaff, Erwin, *Biochim. Biophys. Acta.* **1951***, 9*, 402-405.

[8] Watson, James D.; Crick, Francis H.C., *Nature* **1953**, *171*, 737-738.

[9] Egholm, Michael; Buchardt, Ole; Nielsen, Peter E.; Berg, Rolf H., *J. Am. Chem. Soc.* **1992**, *114*, 1895-1897.

[10] Mologni, Luca; leCoutre, Philipp; Nielsen, Peter E.; Gambacorti-Passerini, Carlo, *Nucleic Acids Res.* **1998***, 26*, 1934-1938.

[11] Good, Liam; Nielsen, Peter E., *Proc. Natl. Acad. Sci.* **1998**, *95*, 2073-2076.

[12] Good, Liam; Nielsen, Peter E., *Nature Biotechnol.* **1998**, *16*, 355-358.

[13] Stein, Hans; Hausen, Peter, *Science* **1969**, *166*, 393-395.

[14] Hausen, Peter; Stein, Hans, *Euro. J. Biochem.* **1970**, *14*, 278-283.

[15] Furdon, Paul J.; Dominski, Zbigniew; Kole, Ryszard, *Nucleic Acids Res.* **1989**, *17*, 9193-1204.

[16] Nielsen, Peter E.; Egholm, Michael; Buchardt, Ole, *Gene* **1994**, *149*, 139-145.

[17] Praseuth, Danièle; Grigoriev, Mikhaïl; Guieysse, Anne-Laure; Pritchard, Linda L.; Harel-Bellan, Annick; Nielsen, Peter E.; Hélène, Claude, *Biochim. Biophys. Acta*. **1996**, *1309*, 226-238.

[18] Tomac, Sebastian; Sarkar, Munna; Ratilainen, Tommi; Wittung, Pernilla; Nielsen, Peter E.; Nordén, Bengt; Gräslund, Astrid, *J. Am. Chem. Soc.* **1996,** *118*, 5544-5552.

[19] Thomas, René, *Biochim. Biophys. Acta.* **1954**, *14*, 231-240.

[20] Ratilainen, Tommi; Holmén, Anders; Tuite, Eimer; Haaima, Gerald; Christensen, Leif; Nielsen, Peter E.; Nordén, Bengt, *Biochemistry* **1998**, *37*, 12331-12342.

[21] Ray, Arghya, and Nordén, Bengt, *FASEB J.* **2000**, *14*, 1041-1060.

[22] Lohse, Jesper; Dahl, Otto; Nielsen, Peter E., *Proc. Natl. Acad. Sci.* **1999***, 96,* 11804-11808.

[23] Kutyavin, Igor V.; Rhinehart, Rebecca L.; Lukhtanov, Eugeny A.; Gorn, Vladimir V.; Meyer, Rich B.; Gamper, Howard B., *Biochemistry* **1996**, *35*, 11170-11176.

[24] Merrified, Robert B., *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.

[25] Merrifield, Robert B., *J. Am. Chem. Soc.* **1964**, *86*, 304-305.

[26] Egholm, Michael; Buchardt, Ole; Nielsen, Peter E.; Berg, Rolf H., *J. Am. Chem. Soc.* **1992**, *114*, 9677-9678.

[27] Nielsen, Peter E.; Egholm, Michael; Buchardt, Ole, *Biochonjugate Chem*. **1994**, *5*, 3-7.

[28] Dueholm, Kim L.; Egholm, Michael; Behrens, Carsten; Christensen, Leif; Hansen, Henrick F.; Vulpius, Tore; Petersen, Kenneth H.; Berg, Rolf H.; Nielsen, Peter E.; Buchardt, Ole, *J. Org. Chem.* **1994**, *59*, 5767-5773.

[29] Sugiyama, Toru; Hasegawa, Genki; Niikura, Chie; Kuwata, Keiko; Imamura, Yasutada; Demizu, Yosuke; Kurihara, Masaaki; Kittaka, Atsushi, *Bioorganic Med. Chem. Lett.* **2017**, *27*, 3337-3341.

[30] St. Amant, André H.; and Hudson, Robert H.E., *Org. Biomol. Chem.* **2012**, *10*. 876-881

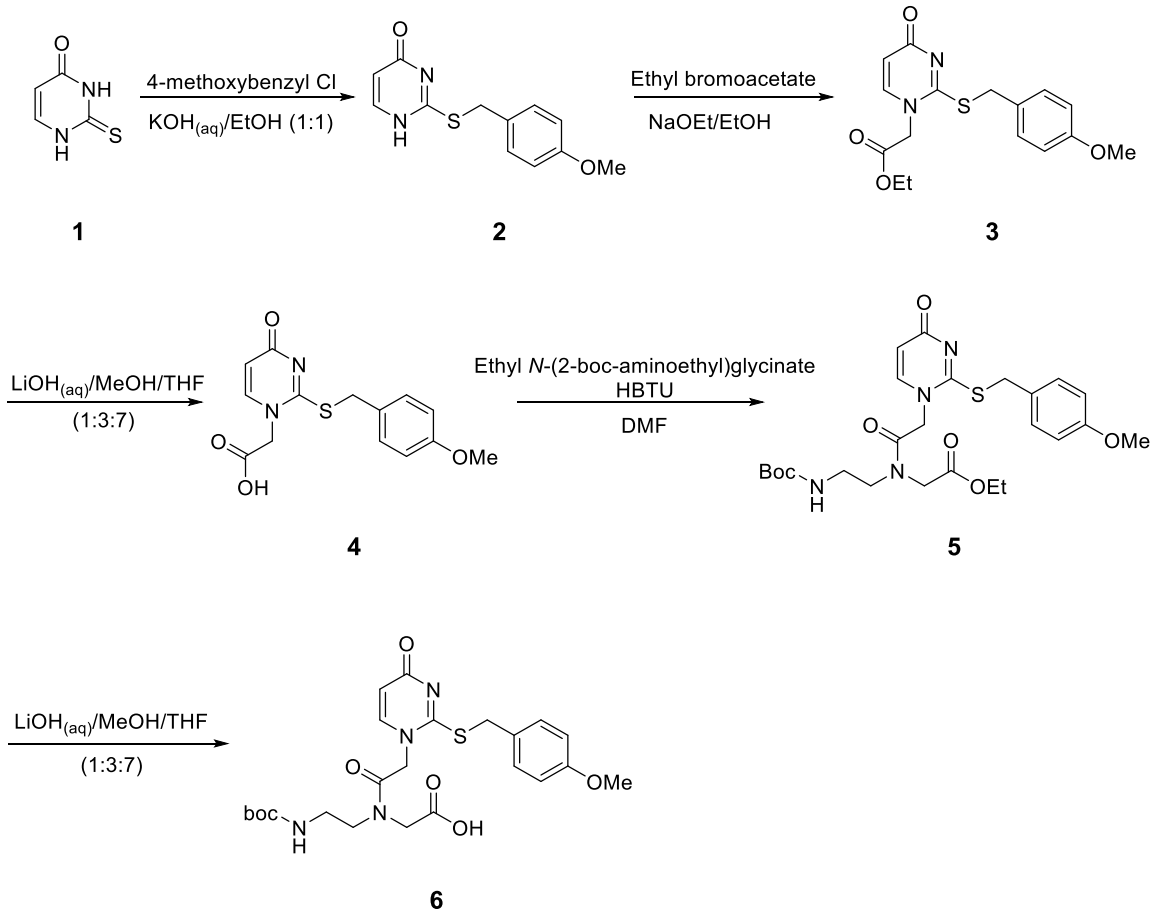
Chapter 2: Highly acid labile *S*-protection of thiouracil and thiothymine

2.1. Introduction

The oligomerization of peptide nucleic acids (PNAs) was designed based on the existing standard for peptide synthesis.[1] At the time of development, standard solid phase peptide synthesis (SPPS) utilized gradient acidolysis, with relatively weakly acidic conditions removing *N*-terminal amine protection and stronger acidic conditions utilized for sidechain deprotection and cleavage from the solid support resin. Typically, thymine and uracil nucleobases do not require protection as there they do not contain nucleophilic exocyclic amino groups. However, replacing the carbonyl group of T and U with a thionyl group at the 2 position of the nucleobase creates a nucleophilic site. Unlike with T and U nucleobases, the thiol at the 2 position of 2-thiouracil (^{S}U) is the most nucleophilic site due to the greater nucleophilicity of sulfur compared to nitrogen, as a result alkylation of the unprotected nucleobase with the methylene carbonyl linker results in *S*-alkylation rather than the desired *N* 1-alkylation. Therefore, protection of the ^{S}U nucleobases in necessary due to the nucleophilicity of the thiol group.

Nielsen *et al* developed the synthesis of a protected thiouracil (^{S}U) compatible with PNA oligomerization under standard SPPS conditions.[2] This synthesis reports a 4-methoxybenzyl protecting group on the 2-thiol position of the nucleobase, converting the thiol (**1**) to a thioether (**2**), which is removed under strongly acidic conditions used for global deprotection. The protected thiouracil nucleobase (**3**) is then attached to the protected aminoethyl glycine (AEG) backbone via a methylene carbonyl linker to yield the

^{S}U monomer for oligomerization (Scheme 2.1.). The ^{S}U monomer is then incorporated into PNA oligomers using standard boc-based SPPS procedures.



Scheme 2.1. Synthesis of thiouracil PNA monomer compatible with boc-based peptide synthesis[2]

Modern SPPS techniques utilize an orthogonal protection system which removes protecting groups and cleaves the peptide from the solid support resin under milder acidic conditions. Due to the mild conditions used for global deprotection, this work aims to develop a more acid labile protecting strategy for the ^{S}U monomer which is easily removed under these conditions. The electron donation from the 4-methoxy substituents on the benzyl protecting group helps stabilize the carbocation formed from acidolysis of

the protecting group. We propose that increasing the electron donation into the benzylic position will increase the stability of the cleaved product (Figure 2.1.). The increased stabilization of the cationic product results in a more acid labile protecting group which can be removed under milder acidolysis conditions. This chapter focuses on the protection of the thiouracil nucleobase using an increased number of methoxy groups on the aromatic ring to tune the removal of the benzyl protecting group for fmoc-based synthesis.

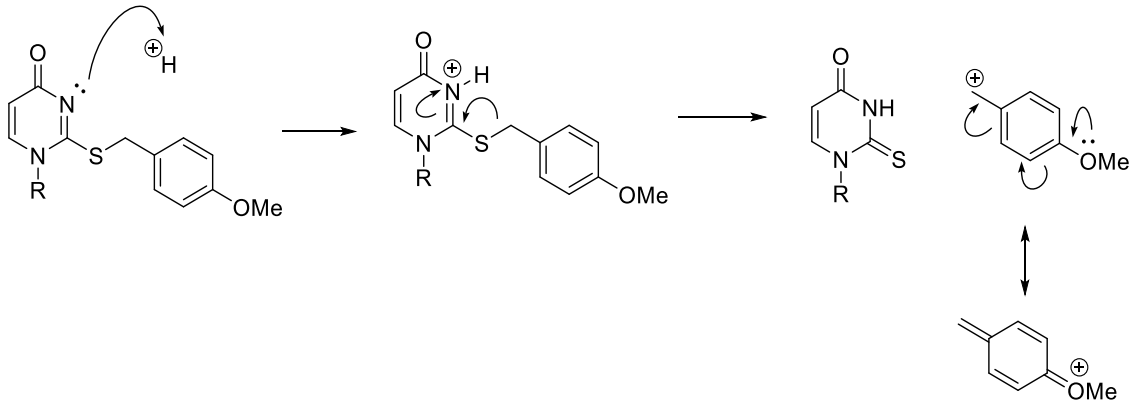


Figure 2.1. Mechanism of thiol protecting group acidolysis

2.2. Synthesis

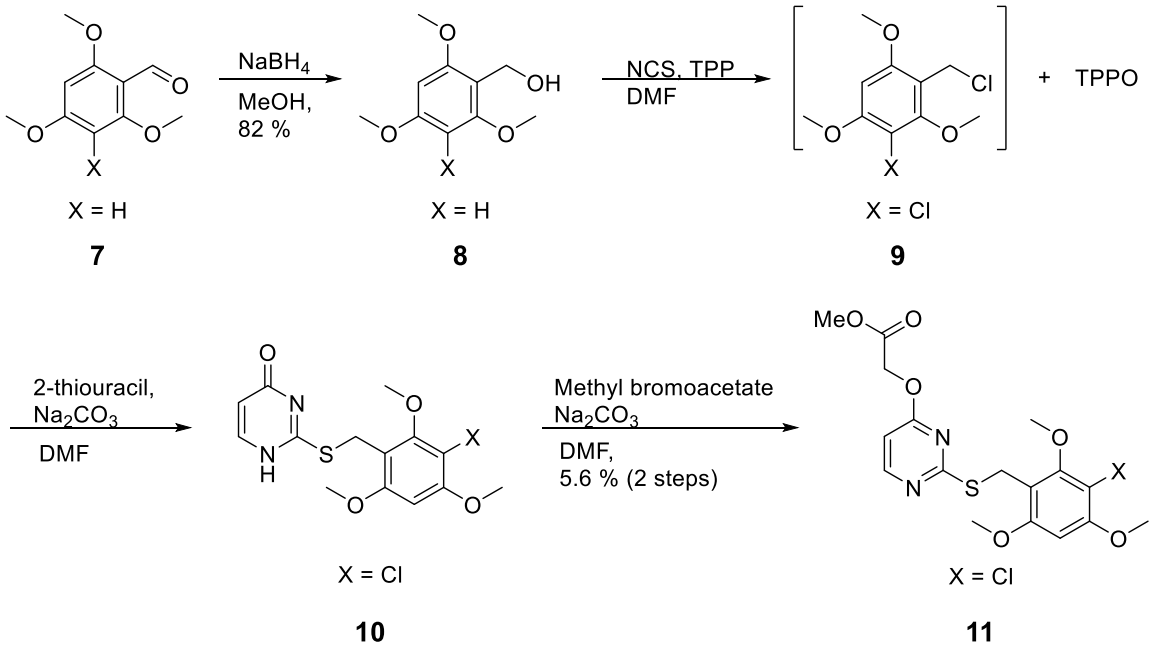
2.2.1. Thiouracil protection with 2,4,6-trimethoxybenzyl group

The 4-methoxybenzyl protecting group used by Nielsen *et al*., was initially replaced by a 2,4,6-trimethoxybenzyl group for protection of the thiol group. Synthesis (Scheme 2.2.) started with 2,4,6-trimethoxybenzaldehyde (**7**) which was reduced using sodium borohydride, to yield 2,4,6-trimethoxybenzyl (**8**) alcohol as the product. The benzyl alcohol was then halogenated to replace the alcohol group with a good leaving group for addition onto the thiouracil nucleobase. Initial attempts at halogenation used

thionyl chloride as a chlorinating agent, however instability of the 2,4,6-trimethoxybenzyl chloride resulted in rapid decomposition of the intermediate, preventing addition to the nucleobase. The electron donation from the three methoxy substituents on the benzyl group greatly increases the electron density at the benzylic position which stabilizes the carbocation formed from the elimination of the benzylic chloride. However, the electron donation to the benzylic carbocation allows the rapid decomposition of the 2,4,6-trimethoxybenzyl chloride upon workup of the reaction, preventing the isolation of the desired intermediate for the protection of the thiouracil nucleobase.

To overcome the issue with the workup and isolation of the 2,4,6-trimethoxybenzyl chloride from the solution of thionyl chloride, a different chlorination method was utilized which did not require workup prior to addition of thiouracil. The 2,4,6-trimethoxybenzyl alcohol was added to a solution of triphenylphosphine (Ph_3P) and *N*-chlorosuccinimide as the chlorinating agent for the 2,4,6-trimethoxybenzyl alcohol. The solution was then added to a basic solution of thiouracil without workup of the reactive 2,4,6-trimethoxybenzyl chloride intermediate (**9**). This enabled the synthesis of the 2,4,6-trimethoxybenzyl protected ^{S}U nucleobase (**10**), however the reaction resulted in low yields of the desired product and triphenylphosphine oxide (Ph_3PO) was a major byproduct of the reaction which could not be easily separated from the desired product. Due to the difficulty in purification, alkylation of the protected thiouracil was performed using the crude mixture of the nucleobase and Ph_3PO. We believed that the triphenylphosphine oxide would not affect the alkylation of the nucleobase, and that once alkylated the thiouracil could be easily separated from the triphenylphosphine oxide by

column chromatography. The crude mixture reacted with methyl bromoacetate as the alkylating agent in a solution of sodium carbonate in dimethylformamide, and the resulting product (**11**) was purified by column chromatography.



Scheme 2.2. Attempted synthesis of 2,4,6-trimethoxybenzyl protected thiouracil nucleobase

^{1}H NMR analysis of the purified product indicated addition of the 2,4,6-trimethoxybenzyl protection to the thiol group, however the ^{1}H NMR spectra also indicated a loss of symmetry in the protecting group, with the detection of 3 distinct signals from the methoxy substituents on the aromatic ring. Furthermore, the integration of the aromatic protons on the benzyl ring decreased from 2 H to 1 H, suggesting a substitution on the aromatic ring. The low-resolution mass spectra of identified a compound with a molecular mass of 414.09 amu, which is about 34 amu larger than then expected mass of the compound (380.10 amu). Based on the mass of the synthesized compound, and the loss of symmetry in the 2,4,6-trimethoxybenzyl protecting group

suggest that chlorination of the aromatic ring had occurred in addition to chlorination of the benzylic carbon. A search of the literature found metal-free electrophilic aromatic halogenation of electron rich aromatic rings using *N*-halosuccinimide (NXS) as a halogenating agent, and catalytic trimethylsilyl chloride (TMSCl).[3] The proposed mechanism for this reaction suggests that a small amount of chloride ions dissociate from TMSCl and reacts with NXS to form a diatomic interhalogen molecule which acts as the halogenating agent. This demonstrates the aromatic ring, when activated by electron donating groups, can undergo halogenation with *N*-halosuccinimide as a halogenating agent. Similarly, the highly electron rich aromatic ring of the 2,4,6-trimethoxybenzyl group undergoes a metal-free electrophilic aromatic halogenation from the *N*-chlorosuccinimide and triphenylphosphine chlorination conditions.

HMBC NMR analysis of the compound was performed to determine the structure of the synthesized compound. Three possible constitutional isomers exist from alkylation of the protected ^{S}U nucleobase, with alkylation of either the *N*-1 position or the *N*-3 position common with alkylation of uracil and *O*-alkylation of the carbonyl at the 4-position also possible. Analysis of the HMBC NMR spectra looks at interactions between the methylene carbonyl protons with the carbon atoms of the thiouracil ring. For the desired *N*-1 alkylation, the expected interaction would be observed with the C-2 and C-6 atoms while *N*-3 alkylation would result in interaction between the methylene carbonyl protons and the C-2 and C-4 atoms. Neither of these outcomes are observed in the HMBC NMR spectra, instead interactions between the methylene carbonyl protons and the C-4 and C-5 atoms indicating *O*-alkylation of the thiouracil. Due to the low yield of *S*-alkylation

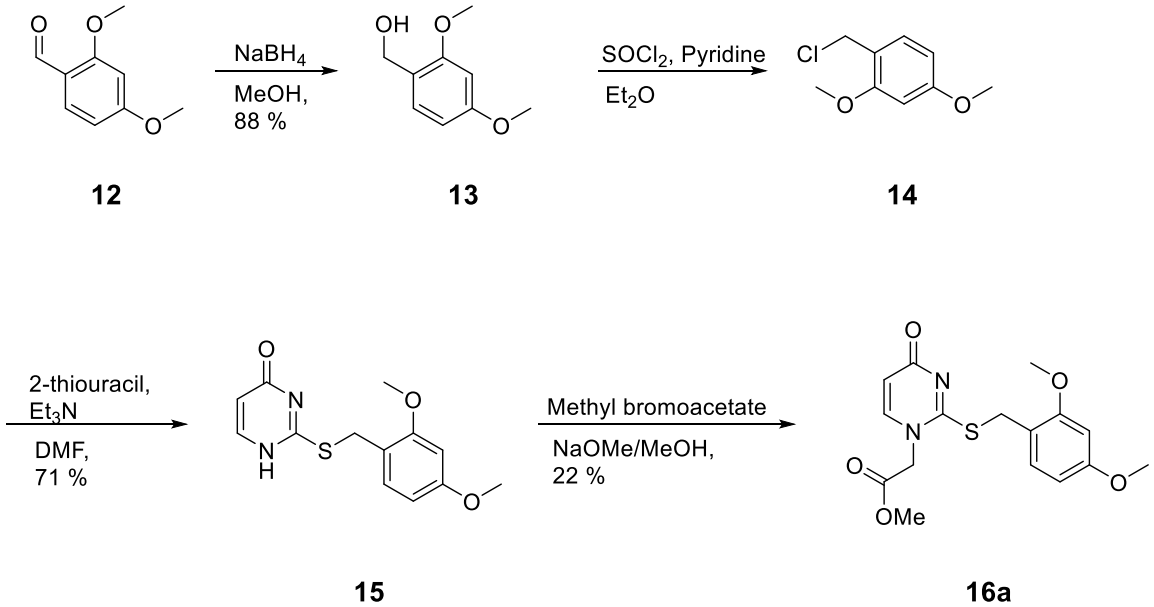
of the thiouracil with the 2,4,6-trimethoxybenzyl protecting group, the subsequent halogenation of the electron rich aromatic ring of the benzyl group, and the difficulty in alkylating the *N*-1 position, a less electron donating group was used for the thiol protection.

2.2.2. Thiouracil protection with 2,4-dimethoxybenzyl group

The destabilizing effect from the electron donation of the three methoxy groups into the benzylic carbon on the 2,4,6-trimethoxybenzyl chloride can be reduced by removing one of the methoxy substituents from the aromatic ring, instead using the 2,4-dimethoxybenzyl chloride. However, reducing the electron donation into the benzylic carbon will reduce the acid lability of the protecting group, as it reduces the stabilization of the carbocation formed from protecting group acidolysis. The 2,4-dimethoxybenzyl protecting group was used for protection of the nucleobase, as this would provide a more stable chloride intermediate than the 2,4,6-trimethoxybenzyl chloride, but would also provide more electron donation than the 4-methoxybenzyl group allowing easier acidolysis.

The protected thiouracil nucleobase was synthesized (Scheme 2.3.) starting with 2,4-dimethoxybenzaldehyde (**12**), which was first reduced using sodium borohydride in methanol to yield the 2,4-dimethoxybenzyl alcohol (**13**). The benzyl alcohol is then halogenated using a solution of thionyl chloride to form 2,4-dimethoxybenzyl chloride (**14**). The reaction was quenched using aqueous sodium bicarbonate, then the 2,4-dimethoxybenzyl chloride was extracted in ether to yield the compound as a clear

colourless oil. The 2,4-dimethoxybenzyl group was then added to the thiol group of the thiouracil using triethylamine as the base, which yielded the 2-(2,4-dimethoxybenzyl)thiouracil (**15**) as a white solid. The protected thiouracil was then alkylated using methyl bromoacetate in a solution of sodium methoxide, which yielded a mixture containing both *N*-1-alkylation (**16**) and *O*-4-alkylation of the nucleobase which was easily separated by column chromatography. The *N*-1-alkylated product was synthesized in just a 22 % yield, while the *O*-alkylated product resulted in a 13 % yield. Despite the low yield of the synthesis, the unreacted *S*-(2,4-dimethoxybenzyl)thiouracil was easily separated by column chromatography and reacted with methyl bromoacetate and sodium methoxide in a second alkylation reaction to increase the yield of the desired *N*-1-alkylated product.



Scheme 2.3. Synthesis of 2,4-dimethoxybenzyl protected thiouracil nucleobase

The *N*-1- and *O*-alkylated constitutional isomers were characterized with ^{1}H NMR spectroscopy. A comparison of the spectra showed similar chemical shifts with the

methylene carbonyl proton shifts as the major identifying feature of the constitutional isomers. The alkylation of the *O*-4 resulted in a downfield shift in the methylene carbonyl signals compared to alkylation at the *N*-1 position. Further HMBC NMR analysis of the compound suggested *N*-1 alkylation due to the interaction between the methylene carbonyl protons and the carbon atoms at the 2 and 5 positions of the pyrimidine ring, while the *O*-alkylation product showed coupling between the methylene carbonyl protons and the C-4 atom on the pyrimidine ring (Figure 2.2.).

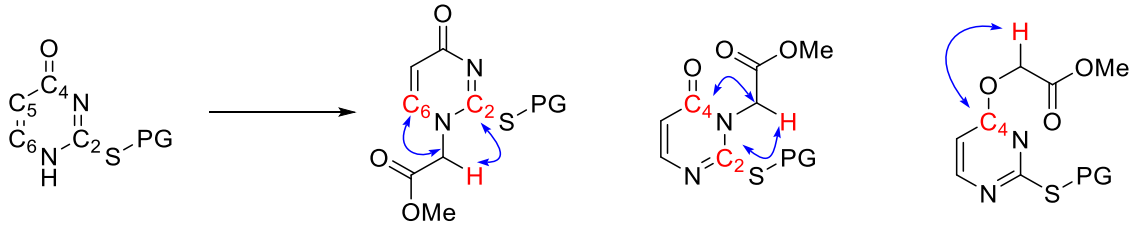


Figure 2.2. HMBC NMR interactions of methylene carbonyl protons with pyrimidine ring carbons when alkylated in the *N*-1, *N*-3 and *O*-4 positions

The acid cleavage of the thiol protecting group was confirmed under standard deprotection conditions used for global deprotection of PNA oligomers synthesized using fmoc-based oligomerization. The *N*-1-alkylated *S*-(2,4-dimethoxybenzyl)thiouracil was mixed with a solution of 5 % triethylsilane in trifluoroacetic acid, then the acid was evaporated under a nitrogen stream then the product was triturated with ether. The ^{1}H NMR spectra of the obtained product shows removal of the thiol protecting group under standard deprotecting conditions. The removal of the 2,4-dimethoxybenzyl group is seen as the absence of the methoxy protons, the aromatic benzyl protons and the benzylic protons in the NMR spectra. This test confirmed the acid lability of the thiol protecting group under standard deprotection conditions suggesting that once incorporated into a

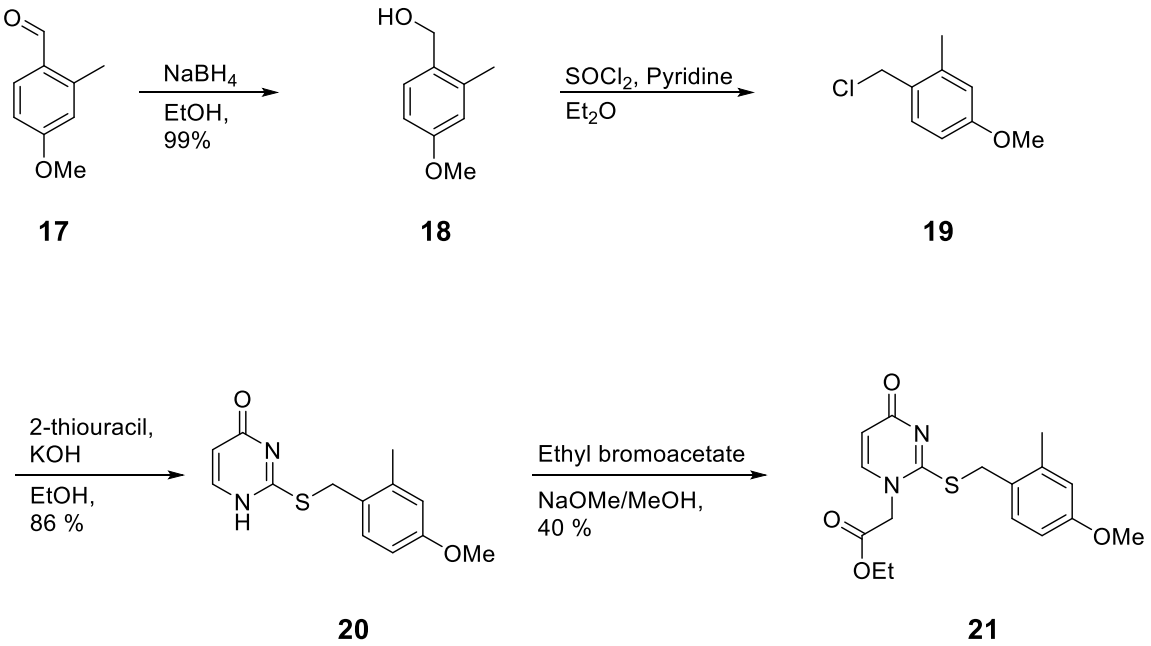
PNA oligomer, using fmoc-based chemistry, the thiouracil nucleobases could be easily deprotected during the global deprotection and resin cleavage of the PNA oligomer.

2.2.3. Thiouracil protection with 2-methyl-4-methoxybenzyl group

Recently published work reports the synthesis of a thiouracil monomer compatible with fmoc-based oligomerization conditions using a 2-methyl-4-methoxybenzyl group for protection of the nucleobase.[4] Electron donation from the methyl group into the aromatic ring, through hyperconjugation, increased electron density at the benzylic position which stabilizes the benzylic carbocation. The hyperconjugation of the methyl group provides weaker electron donation than the resonance from the methoxy groups; however, the electron donation from the methyl substituent combines with the electron donation of the *para* methoxy substituent to increase electron density at the benzylic position compared to the 4-methoxybenzyl group. Therefore the 2-methyl-4-methoxybenzyl protected ^{S}U nucleobase was synthesized (Scheme 2.4.) following the procedures described by Kittaka *et al*.[4] in order to test the rate of the protecting group acidolysis compared to the rate of acidolysis of the newly synthesized 2,4-dimethoxybenzyl protected nucleobase.

The 2-methyl-4-methoxybenzyl aldehyde (**17**) was reduced to the benzyl alcohol (**18**) using sodium borohydride. The alcohol was then substituted with a chloride using thionyl chloride to yield the 2-methyl-4-methoxybenzyl chloride (**19**) for the protection of thiouracil. The benzyl group was added to the thiol group under basic conditions to yield the *S*-(2-methyl-4-methoxybenzyl)thiouracil (**20**) which was *N*-1 alkylated (**21**) with ethyl

bromoacetate using a solution of sodium ethoxide in ethanol. Under these conditions, the *O*-alkylated product was formed as the minor product, and the products were easily separated by column chromatography. ^{1}H NMR analysis of the product matched the literature spectra reported by Kittaka *et al*., therefor no further characterization was performed.[4]



17 **18** **19**

20 **21**

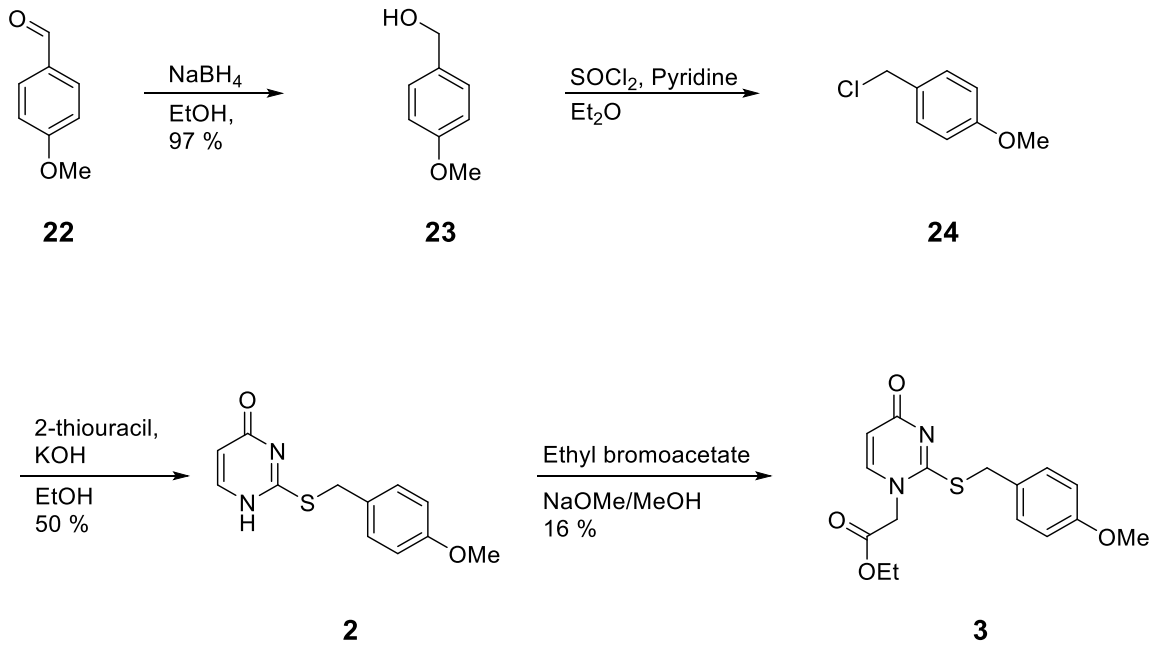
Scheme 2.4. Synthesis of 2-methyl-4-methoxybenzyl protected thiouracil nucleobase

2.2.4. Thiouracil protection with 4-methoxybenzyl group

To study the effects of the addition of methoxy substituents to the *ortho*- and *para*- positions of the benzene ring on the kinetics of the acidolysis of the benzyl protecting group from the thiol of the thiouracil nucleobase, the 4-methoxybenzyl protected thiouracil was synthesized. The 4-methoxybenzyl thiouracil being used in the

original synthesis of the thiouracil monomer for boc-based oligomerization, it was synthesized as a comparison for the more highly acid labile thiol protected monomers.

The *S*-(4-methoxybenzyl)thiouracil was synthesized (Scheme 2.5.) starting with the 4-methoxybenzyl alcohol (**22**), which was reduced to the benzyl alcohol (**23**) using sodium borohydride. The benzylic alcohol was substituted with chloride using thionyl chloride, and the resultant 4-methoxybenzyl chloride (**24**) was added to the thiouracil under basic conditions to obtain *S*-(4-methoxybenzyl)thiouracil (**2**). The protected nucleobase was then alkylated using ethyl bromoacetate and sodium ethoxide to yield *N*-1 alkylated 4-methoxybenzyl protected thiouracil nucleobase (**3**). The ^{1}H NMR spectra of the product matched the literature spectra reported by Nielsen *et al*.[2]



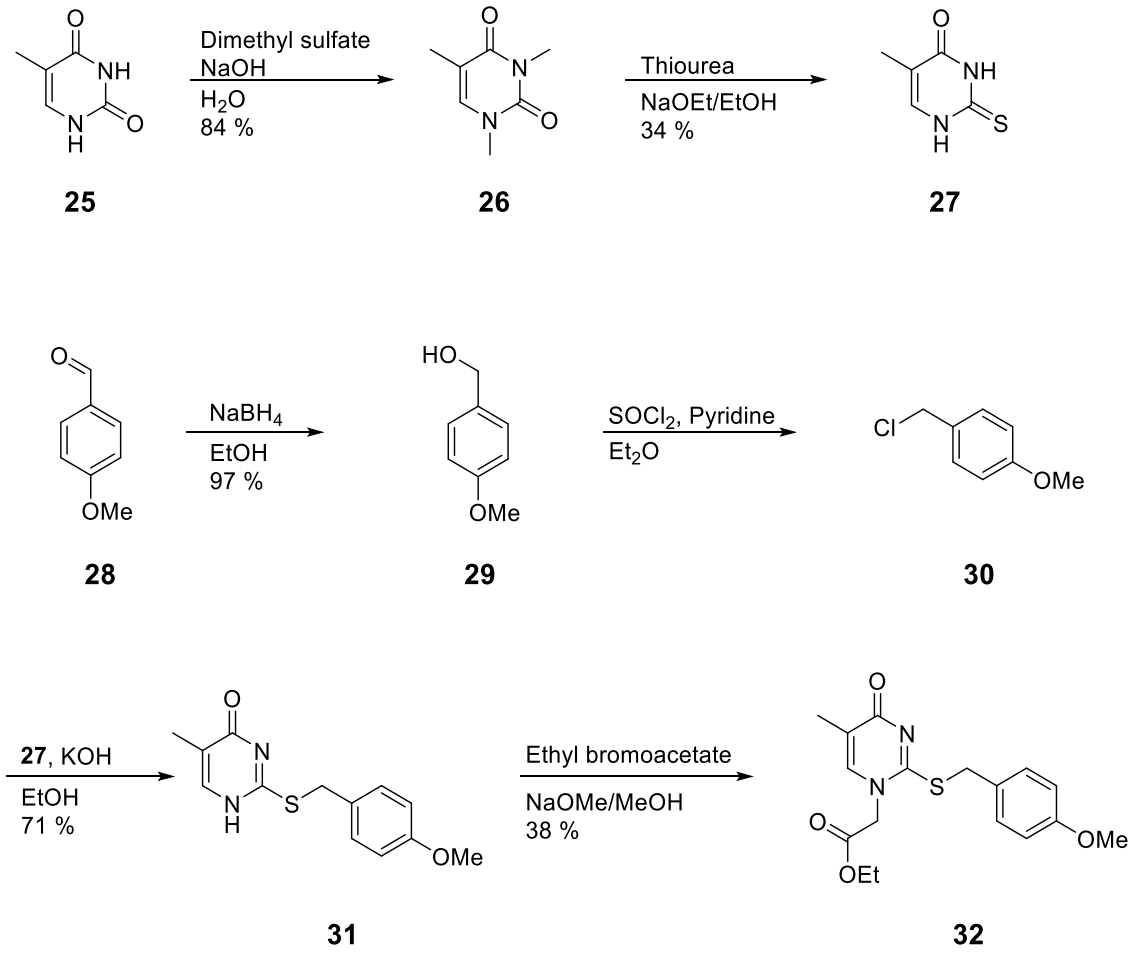
22 **23** **24**

2 **3**

Scheme 2.5. Synthesis of 4-methoxybenzyl protected thiouracil nucleobase

2.2.5. 4-Methoxybenzyl protected thiothymine

Thymine and uracil are nucleobases which differ only in the substituent at the 5-position of the pyrimidine ring, and both undergo complementary base pairing with adenine. To date, only the thiouracil nucleobase has been studied as a pseudo-complementary nucleobase in PNA oligomers; however, pseudo-complementary base pairing was first demonstrated using thiothymidine as a thymine analog in combination with 2,4-diaminopurine in oligodeoxynucleotides.[5] Studies comparing the melting temperatures of oligodeoxynucleotides containing either thymine or uracil as the complementary base with adenine show an increased hybridization stability of oligodeoxynucleotides containing thymine compared to the stability of those containing uracil.[6] To study to effect of thiothymine on the hybridization stability when incorporated into pcPNA oligomers, a protected thiothymine monomer was synthesized (Scheme 2.6.).



Scheme 2.6. Synthesis of 4-methoxybenzyl protected thiothymine nucleobase

Thiothymine was synthesized from the thymine nucleobase by first methylating the *N*-1 and *N*-3 positions (**26**) of the thymine ring using dimethyl sulfate. The dimethyl urea moiety was then substituted with thiourea to replace the carbonyl with a thiol (**27**). Due to the higher nucleophilicity of sulfur compared to nitrogen, the thiol undergoes nucleophilic attack of the C-6 position in a Michael addition followed by β-elimination of the methylated amine. The thiourea then undergoes nucleophilic attack at the C-4 position with elimination of *N,N'*-dimethylurea to form a 1,3-thiazine ring (Figure 2.3.).[7] The 1,3-thiazine undergoes rearrangement to the pyrimidine ring under alkali conditions

to form thiothymine.[8-10] The thiol group of the thiothymine was then protected with 4-methoxybenzyl chloride (**30**) under basic conditions. The nucleobase (**31**) was then alkylated with ethyl bromoacetate using sodium ethoxide to yield the N-1 alkylated product (**32**) as the major product and the *O*-alkylated product as the minor product.

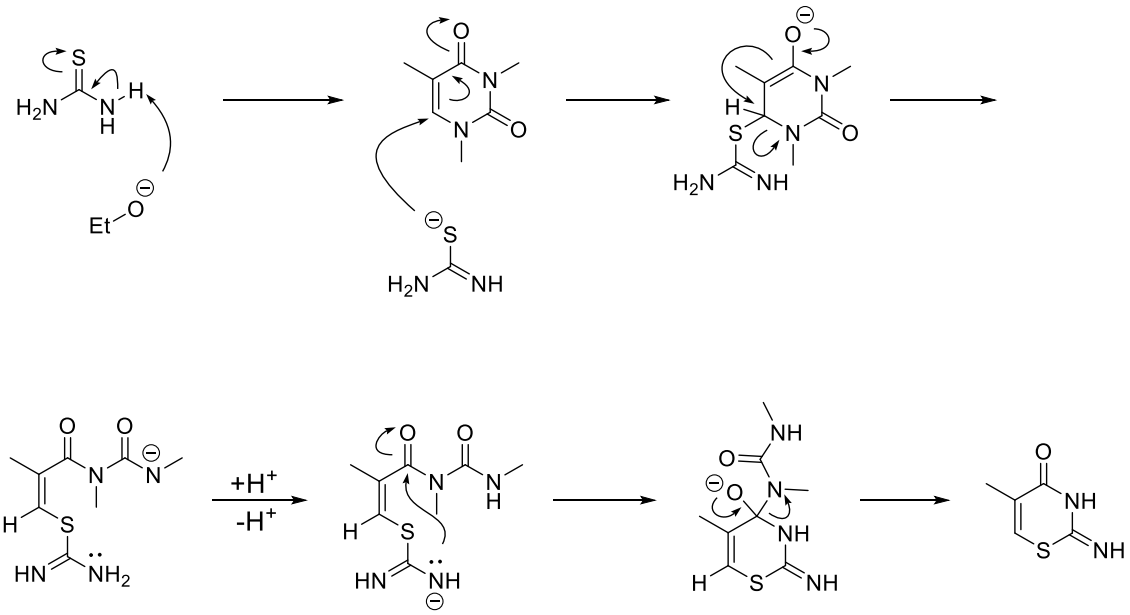


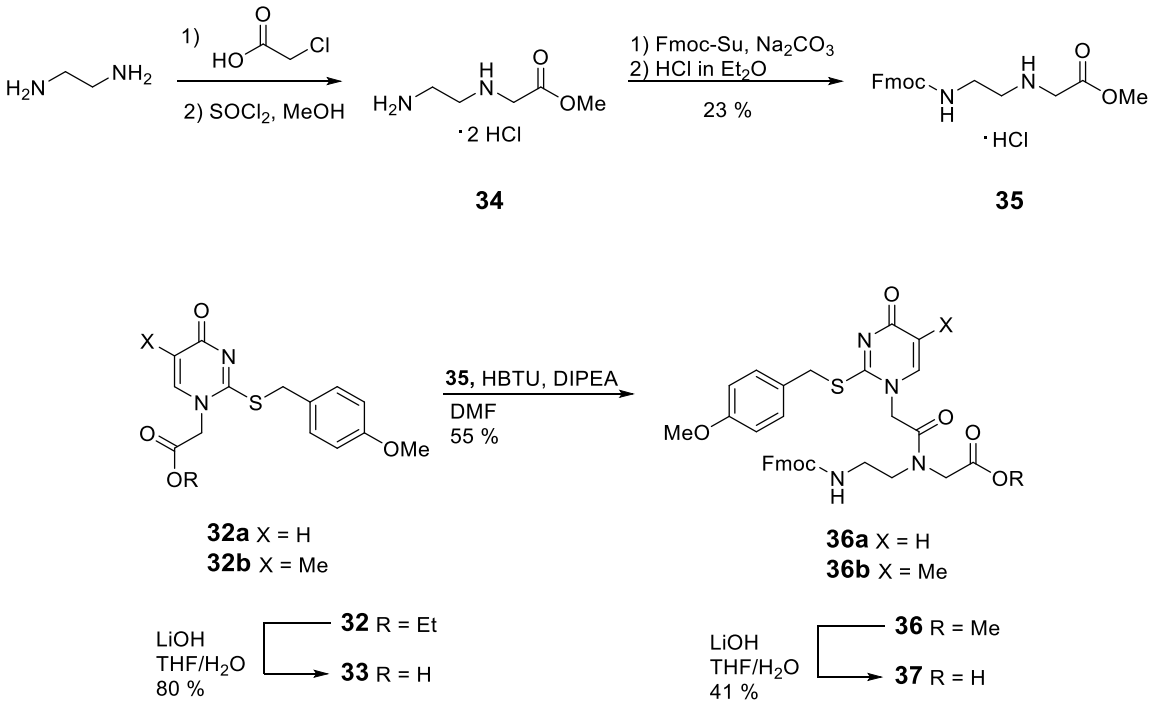
Figure 2.3. Mechanism of nucleophilic attack of 1,3-dimethyl thymine by thiourea to form 1,3-thiazine

2.2.6. Thiouracil and thiothymine PNA monomers

The aminoethylglycine backbone was synthesized (Scheme 2.7.) with a methyl ester protecting group on the C-terminal carboxylic acid and the fmoc protecting group on the N-terminal amine. The backbone was previously synthesized as the dihydrochloride salt following literature procedures, an excess of ethylenediamine was alkylated with 2-chloroacetic acid, then protected using thionyl chloride and methanol (**34**).[11, 12] The fmoc group was added to the N-terminus of the methyl aminoethyl glycinate

using fmoc-succinimide under basic conditions, then the protected backbone was isolated and stored as the hydrochloride salt (**35**).[13]

The alkylacetate of the protected nucleobase (**32**) was hydrolyzed using a solution of aqueous lithium hydroxide, to obtain the carboxylic acid functional group (**33**) for attachment to the aminoethylglycine backbone. The protected nucleobase was then reacted with a solution of hexafluorophosphate benzotriazole tetramethyl uroniom (HBTU) and diisopropylethyl amine (DIPEA), which acts as the coupling agent by forming an active ester on the methylene carbonyl linker of the nucleobase. The activated nucleobase is then reacted with the protected aminoethylglycine backbone to form the PNA monomer (**36**). The methyl ester protecting the C-terminus of the PNA monomer is then hydrolyzed using a solution an aqueous lithium hydroxide to form the C-terminus carboxylic acid (**37**). The kinetics of the reactivity of the C-terminus protecting group compared to the N-terminus protecting group under aqueous alkali conditions favours hydrolysis of the methyl ester rather than removal of the fmoc group, giving selective deprotection of the PNA backbone.



Scheme 2.7. Synthesis of thiouracil and thiothymine PNA monomer (**37**)

Basic conditions to remove the fmoc protecting group uses a solution of 20 % piperidine (pK_a = 11), a weaker base than lithium hydroxide (H_2O pK_a = 16), however selectivity for the hydrolysis of the methyl ester is due to faster kinetics of the ester hydrolysis using a hydroxide base. Amine bases are effective in the removal of the fmoc group since the acidolysis of the fmoc group results in the formation of highly reactive dibenzofulvene as a byproduct of the reaction, which is trapped by the amine base to form stabile adducts. Amine bases such as piperidine effectively trap the dibenzofulvene intermediate which drives the deprotection reaction to completion.[14-16] This driving force of the reaction is not present using a hydroxide base, which slows the rate of fmoc deprotection, and kinetically favours hydrolysis of the methyl ester giving selectivity for deprotection of the C-terminus.

Hydrolysis of the C-terminus methyl ester yields the desired thiouracil or thiothymine PNA monomer for oligomerization. The monomers where characterized using ^{1}H NMR and ^{13}C NMR spectroscopy and high-resolution mass spectrometry (HRMS). The ^{1}H NMR spectra shows several rotamer structures resulting from the amide bond formed between the methylene carbonyl linker and the aminoethyl glycine backbone and resulting from the carbamate resulting from fmoc protection of the N-terminus. The rotamers result from resonance of the nitrogen electrons into the π-bonding of the amide group. The greater electronegativity of the oxygen pulls greater electron density, causing partial π-bonding between the nitrogen and the carbon which prevents free rotation around the amide bonds. This results in resonance structures with E and Z conformational isomers from the amide bond and results in major and minor ^{1}H NMR signals for many of the alkyl protons (Figure 2.4.).

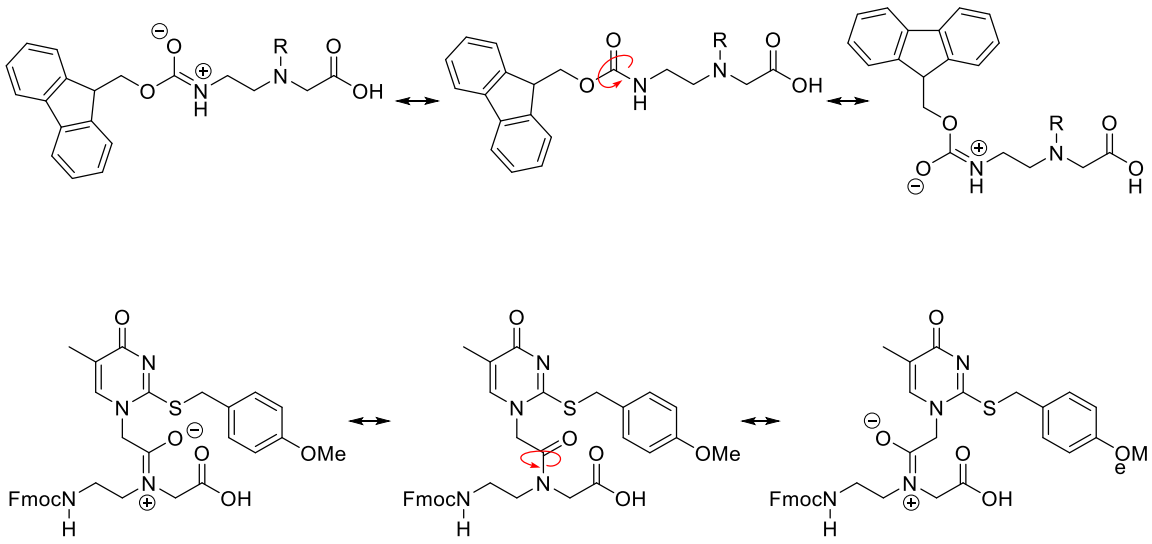
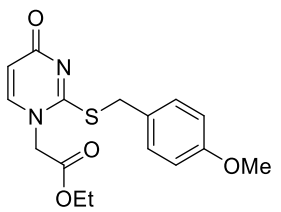


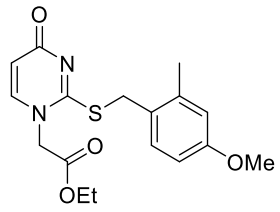
Figure 2.4. Resonance structures from amide and carbamate rotamers of PNA monomer **37**

2.3. Conclusions

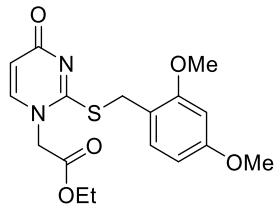
This chapter demonstrates the synthesis of thiouracil and thiothymine nucleobases with various benzyl based protecting groups (Figure 2.5.) on the 2-position thiol group. The addition of electron donating functional groups at the *ortho*- and *para*-positions of the benzene ring increases the electron density on the benzylic position. The addition of methoxy substituents onto the *ortho*- and *para*- positions donates electron density to the benzylic position through resonance through the aromatic ring. The next chapter will investigate the rate of acidolysis of the protecting group under standard fmoc-based SPPS deprotection conditions. This research aims to develop a highly acid labile protecting group for the thiol protecting group compared to the *para*-methoxybenzyl group used for boc-based pcPNA oligomerization. Furthermore, the deprotection rate was compared to the rate of acidolysis of the recently reported 2-methyl-4-methoxybenzyl protecting group developed for fmoc-based synthesis of pcPNA. Finally, using common coupling conditions the protected nucleobase was attached to the PNA backbone to synthesize the thiouracil and thiothymine PNA monomer for PNA oligomerization.



S-(4-methoxybenzyl)thiouracil (**3**)



S-(2-methyl-4-methoxybenzyl) thiouracil (**21**)



S-(2,4-dimethoxybenzyl)thiouracil (**16a**)

Figure 2.5. Benzyl based protecting groups on thiouracil

2.4 References

[1] Nielsen, Peter E.; Egholm, Michael; Berg, Rolf H.; Buchardt, O., *Science* **1991**, *254*(5037), 1497-1500.

[2] Lohse, Jasper; Dahl, Otto; Nielsen, Peter E., *Proc. Nat. Acad. Sci.* **1999**, *96*, 11804-11808.

[3] Maibunkaew, Tapanee; Thongsornkleeb, Charnsak; Tummatorn, Jumreang; Bunrit, Anon; Ruchirawat, Somsak, *Synlett*. **2014***, 25*, 1769-1775.

[4] Sugiyama, T.; Hasegawa, G.; Niikura, C.; Kuwata, K.; Imamura, Y.; Demizu, Y.; Kurihara, M.; Kittaka, A., *Bioorganic Med. Chem. Lett.* **2017**, *27*, 3337-3341.

[5] Kutyavin, Igot V.; Rhinehart, Rebecca L.; Lukhtanov, Eugeny A.; Gorn, Vladimir V.; Meyer, Rich B.; Gamper, Howard B., *Biochemistry* **1996**, *35*, 11170-11176.

[6] Sági, J.; Brahms, S. Brahms, J.; Ötvös, L., *Nucleic Acids Res.* **1979**, *6* (8), 2839-2848.

[7] Hirota, Kosaku; Watanabe, Kyoichi A.; Fox, Jack J., *J. Org. Chem.* **1978**, *43* (6), 1193-1197.

[8] Atkinson, M. R.; Shaw, G.; Schaffner, K.; Warrener, R. N., *J. Chem. Soc.* **1956**, *0*, 3847-3849.

[9] Cain, E. N.; Warrener, R. N., *Tetrahedron Lett*. **1967,** *49,* 4953-4957.

[10] Cain, E. N.; Warrener, R. N., *J. Org. Chem., Aust. J. Chem.* **1970**, *23*, 51-72.

[11] Farèse, Audrey; Patino, Nadia; Condom, Roger; Dalleu, Sandrine; Guedj, Roger, *Tetrahedron Lett.* **1996**, *37*(9), 1413-1416.

[12] Bialy, Laurent; Díaz-Mochón, Juan José; Specker Edgar, Keinicke, Lise; Bradley, Mark, *Tetrahedron* **2005**, *61*, 8295-8305.

[13] Hudson, Robert H.E.; Lui, Yuhong; Wojciechowski, Filip, *Can. J. Chem.* **2007**, *43*, 1193-1197.

[14] Luna, Omar F.; Gomez, Johana; Cárdenas, Constanza; Albericio, Fernando; Marshall, Sergio H.; Guzmán, Fanny, *Molecules* **2016**, *21*, 1542-1553.

[15] Hachmann, John; Lebl, Michal, *J. Comb. Chem.* **2006**, *8*, 149.

[16] Zinieris, Nikos; Leondiadis, Leondios; Ferderigos, Nikolas, *J. Comb. Chem.* **2005**, *7*, 4-6.

Chapter 3: Deprotection and oligomerization of thiouracil and thiothymine monomers

3.1. Introduction

The Hammett principle describes the linear effects of substituents in the *meta* and *para* positions of a benzene ring on the rate and on the equilibrium of a reaction on a side chain attached to the aromatic ring.[1, 2] The addition of electron donating groups and electron withdrawing groups to the benzene ring affect the electron density at the benzylic position. The Hammett principle describes the cumulative effect of electron donating and withdrawing groups at the *meta*- and *para*- position on the rate or equilibrium of reactions at the benzylic position caused by displacement of electron density on the aromatic ring. This linear relationship is not maintained with substituents at the *ortho*- position of the benzylic ring due to steric and electronic interactions between the *ortho*- and benzylic substituents.

The use of the 4-methoxybenzyl protecting group for the protection of the thiouracil nucleobase for boc-based peptide synthesis demonstrated the cleavage of the protecting group under treatment with strongly acidic trifluoromethylsulfonic acid. Thus, to improve the acid lability of the thiol protecting group, electron donating groups were added to the aromatic ring to push electron density to the benzylic position. The greater electron density on the benzylic position serves to stabilize the benzylic carbocation formed from the cleavage of the benzyl based protecting group. Several protected thiouracil nucleobases were synthesized with benzyl based protecting groups to test the cleavage of the protecting group under treatment with milder acidic conditions used for

fmoc-based oligomerization. The addition of methyl and methoxy groups to the aromatic ring were used to increase electron density on the benzylic position and drive the acidolysis of the protecting group.

3.2. Results of benzyl based protection acidolysis

The stability of the various methyl and methoxy benzyl protecting groups was tested under standard deprotection conditions which uses a solution of 95 % trifluoroacetic acid (TFA) and 5 % triethyl silane which quenches the resulting benzylic carbocation generated by acidolysis. The deprotection was monitored by ^{1}H NMR and the integration of the benzylic protons was used to monitor the progression of the reaction, while the integration of the methylene carbonyl protons was used as a reference for the integration. A drop of chloroform was added to the TFA solution as a solvent reference for the NMR spectra.

Initial studies of the deprotection of the 4-methoxybenzyl protected thiouracil found that rapid acidolysis of the thioether occurred under standard deprotection conditions. The initial recording of the ^{1}H NMR spectra occurred 4 min after the addition of the TFA solution to the 4-methoxybenzyl protected thiouracil. The initial spectra showed complete deprotection of the thiouracil had occurred prior to the initial measurement, suggesting rapid acidolysis of the thioether was occurring. This study suggested that the 4-methoxybenzyl protecting group originally developed for boc-based oligomerization was easily removed under TFA deprotection conditions. Due to the acid lability of the 4-methoxbyenzyl protecting group upon treatment with TFA, the thiouracil

protecting group was likely removed during boc-based synthesis of the pcPNA oligomer. Since this study showed the 4-methoxybenzyl protecting group is removed under the same deprotection conditions as the *N*-boc protecting group, the thiouracil protection was likely cleaved during boc-based oligomerization. This suggests that the thiol protecting group is not necessary for the oligomerization, and protection of the thiol group is only necessary for blocking *S*-alkylation and directing alkylation to the *N*-1 position.

Due to the high lability of the 4-methoxybenzyl group when treated with standard deprotection conditions, milder deprotection conditions were used to compare the relative rates of acidolysis of the benzyl based protecting groups. The acidolysis of the benzyl protecting groups was tested using solutions containing a lower concentration of TFA, and the deprotection of the 4-methoxybenzyl group was easily monitored using a solution of 2 % TFA and 1% triethylsilane in deuterated chloroform. The acidolysis of the protecting group was monitored by ^{1}H NMR; the integration of the benzylic protons and the methylene carbonyl protons was measured at 2 min intervals over 2 hours or until complete deprotection had occurred (Figure 3.1.). The integration of the methylene carbonyl protons was used as a reference for the integration and was recorded as 2 protons, then the relative integration of the benzylic protons was plotted against the reaction time. The half-life of the thioether under acidolysis conditions was used to compare the stability of the benzyl based protecting groups (Figure 3.2.).

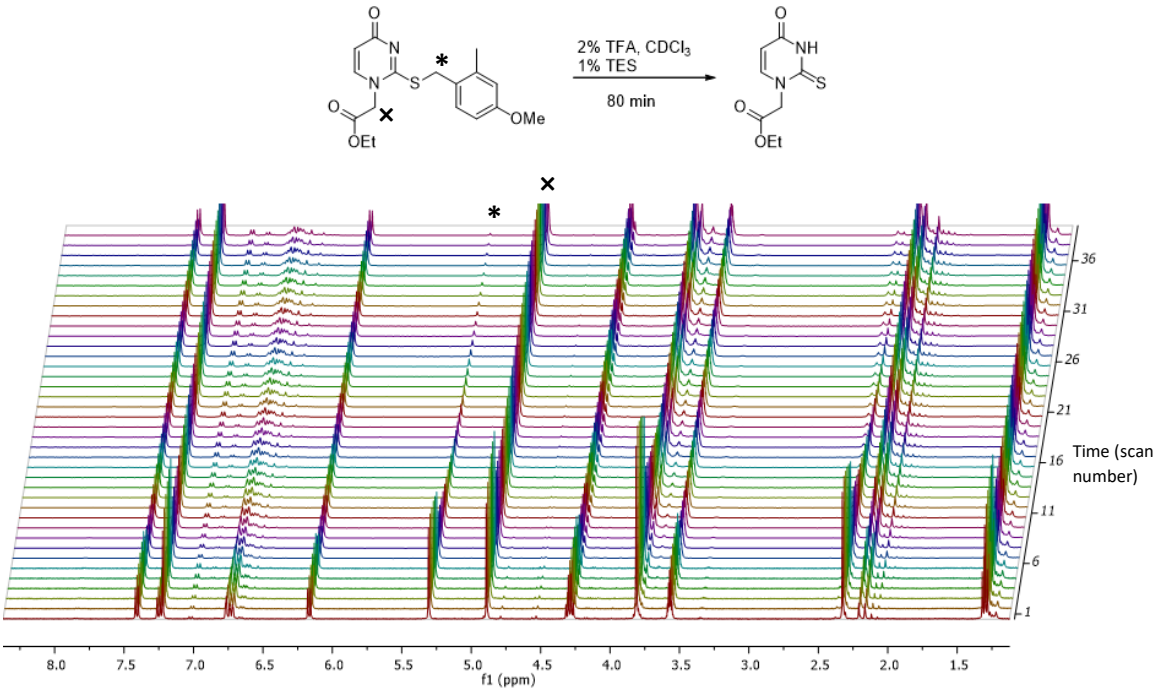


Figure 3.1. ^{1}H NMR spectra of *S*-(2-methyl-4-methoxybenzyl)thiouracil acidolysis in a solution of 2 % TFA and 1 % triethylsilane in $CDCl_3$.

Under the 2 % TFA deprotection conditions, the *S*-(4-methoxybenzyl)thiothymine and the *S*-(4-methoxybenzyl)thiouracil deprotected at similar rates, with the deprotection of thiothymine slightly slower than the deprotection of thiouracil. The slightly higher stability of the *S*-(4-methoxybenzyl)thiothymine is likely due to the formation of a more stable intermediate upon protonation of the nucleobase. The pK_a of thymine (9.86) is slightly higher than that of uracil (9.36) which is likely due to the C-5 methyl group, which can donate electron density into the pyrimidine ring making thymine a slightly more basic nucleobase than uracil.[3] Similarly, upon treatment with TFA, the electron donation from the methyl group helps to stabilize the positive charge resulting from protonation of the nucleobase. The higher stability of the protonated thiothymine compared to the thiouracil nucleobase results in a slightly slower acidolysis of the thioether. In the case of both the *S*-(4-methoxybenzyl)thiouracil and *S*-(4-methoxybenzyl)thiothymine, the

majority of the nucleobase was deprotected after 2 hours when treated with mild deprotection conditions. This contradicts previous assumptions that strongly acidic conditions were necessary for the acidolysis of the protecting group, and as a result a more acid labile thiol protecting group was necessary for adapting the thiouracil PNA monomer for a standard fmoc-based oligomerization strategy.[4, 5] However, this assumption was never tested and the conditions necessary for acidolysis were not previously examined. Acidolysis of the *S*-(2-methyl-4-methoxybenzyl)thiouracil was faster than the acidolysis of the *S*-(4-methoxybenzy)thiouracil. The reaction went to completion within 90 min, and the protected nucleobase had a half-life of around 13 min when treated with the solution of 2 % TFA (Figure 3.2.). The 2,4-dimethoxybenzyl was the most acid labile of the protecting groups studied, with complete deprotection of the *S*-(2,4-dimethoxybenzyl)thiouracil occurring within 4 min of treatment with a solution of 2 % TFA. Since the first ^{1}H NMR measurement was recorded after 4 min of treatment with TFA, no measurements of the acidolysis were recorded and the half-life of the reaction could not be determined. The short reaction time for the acidolysis of the 2,4-dimethoxybenzyl protecting group demonstrates the highly acid labile nature of the newly developed thiouracil protecting group, however kinetic studies of the existing 4-methoxybenzyl protecting group suggest that protection of the thiouracil nucleobase is not necessary during PNA oligomerization.

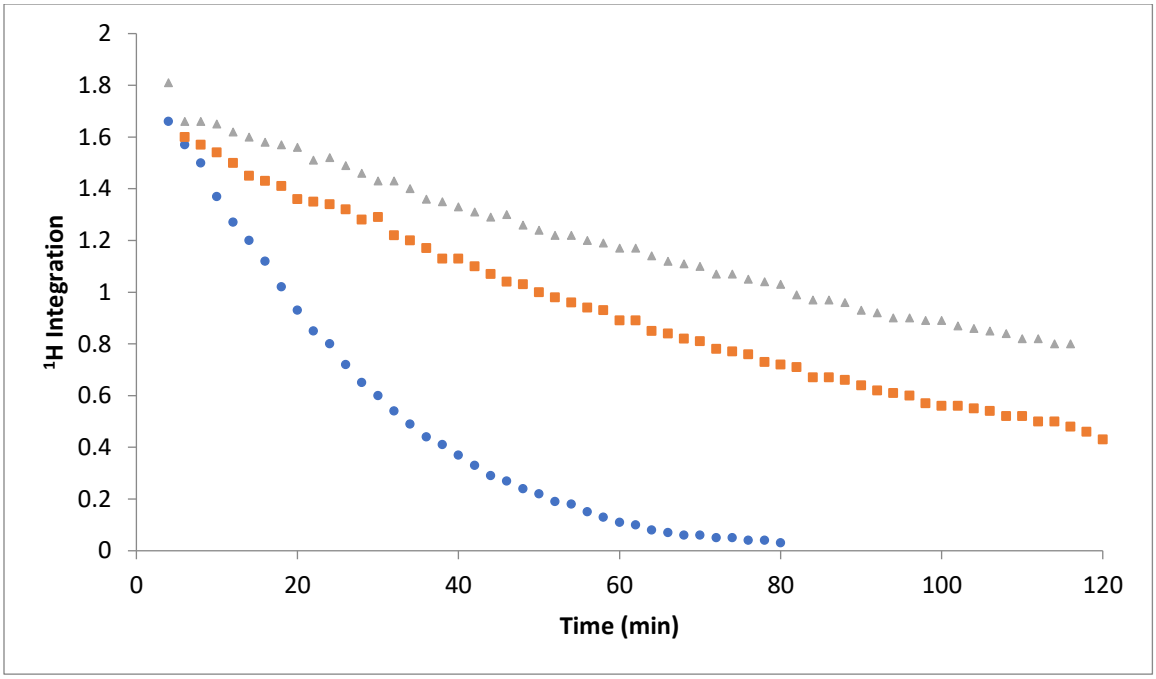


Figure 3.2. Relative ^{1}H integration of benzylic protons of • *S*-(2-methyl-4-methoxybenzyl)thiouracil, ▪ *S*-(4-methoxybenzyl)thiouracil, and ▲ *S*-(4-methoxybenzyl)thiothymine, over time.

The change in the benzylic integration over time was plotted to fit a first order reaction curve with the natural logarithm of the benzylic proton integration against the reaction time. The linear trend of the graph suggests the kinetics of the acidolysis reaction as a first order reaction, dependent on the rate of decomposition of the protonated nucleobase. Upon treatment of the nucleobase with the TFA solution, the acid-base reaction occurs quickly resulting in the protonation of the nucleobase. Protonation of the nucleobase occurs rabidly, with the rate limiting step being the decomposition of the protonated nucleobase resulting in the cleavage of the thioether. From the first order reaction plots, the half-life of the reaction is calculated to compare the relative rate of acidolysis of the benzyl based protecting groups.

A first order reaction proceeds dependent linearly on the concentration of a single reactant (**Eq. 1**), so the equation for the rate is expressed as:

$$rate = -\frac{d[A]}{dt} = k[A] \quad \textbf{(1)}$$

Where the rate of the reaction is the change in concentration of the reactants over time, which for a first order, unimolecular reaction is dependent on the concentration of a single reactant ($[A]$), and a rate constant (k). The integration of the first order rate equation (**Eq. 1**) yields a linear equation (**Eq. 4**) with respect to the natural logarithm of the concentration of the reactant.

$$\int_{[A]_o}^{[A]} \frac{d[A]}{[A]} = \int_{t_o}^{t} -k dt \quad \textbf{(2)}$$

$$\ln[A] - \ln[A]_o = -kt \quad \textbf{(3)}$$

$$\ln[A] = \ln[A]_o - kt \quad \textbf{(4)}$$

The integration of the benzylic protons is used as the concentration of the reactant and the natural logarithm of the integration plotted with time yields a linear first order rate, with the slope of the plot as the rate constant (Figure 3.3.). Using the rate constant, the half-life of the acidolysis of the benzyl based protecting groups is calculated to compare the relative stability of the protecting groups to deprotecting conditions.

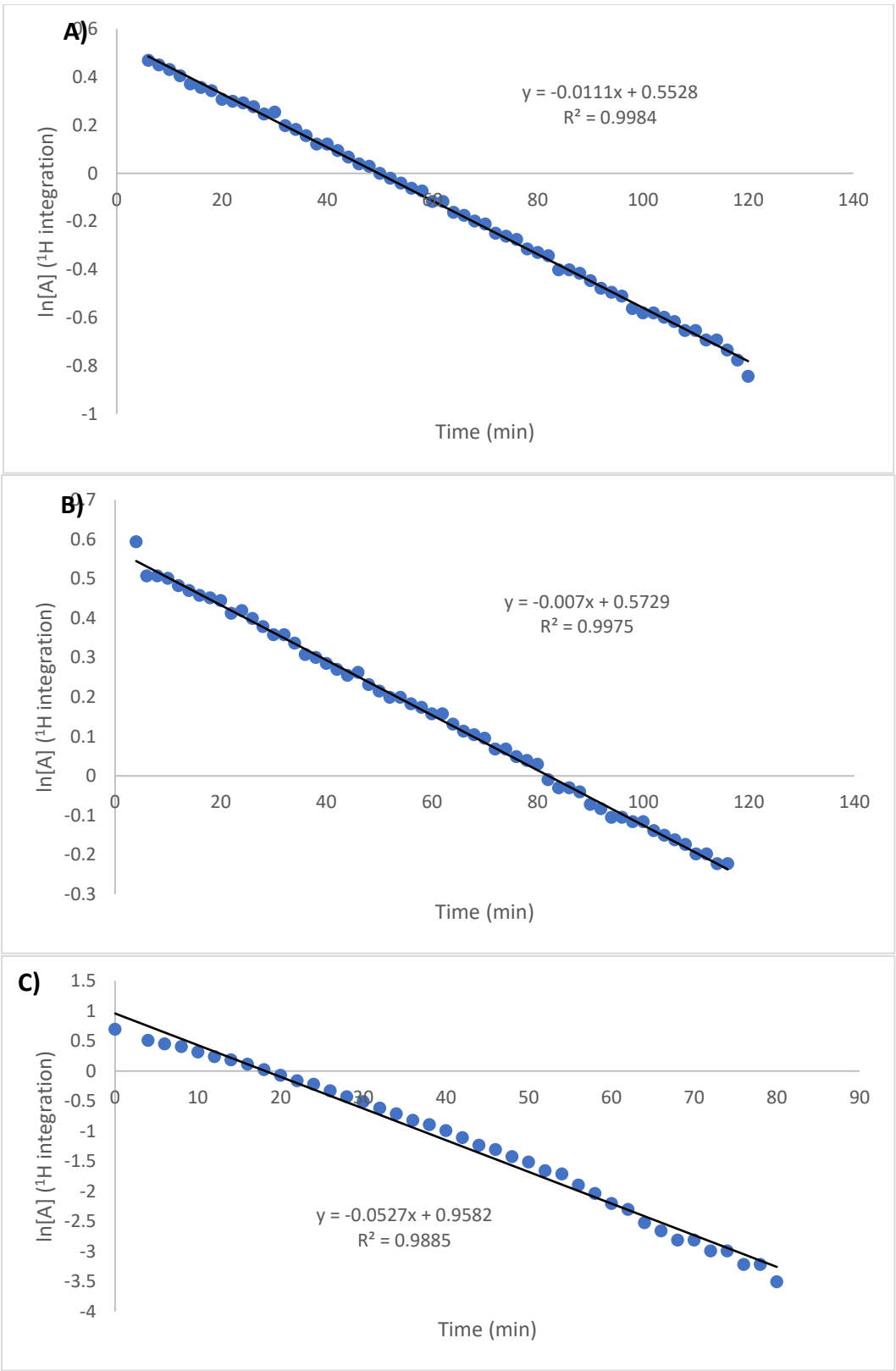


Figure 3.3. First order reaction curves for a) *S*-(4-methoxybenzyl)thiouracil, b) *S*-(4-methoxybenzyl)thiothymine, and c) *S*-(2-methyl-4-methoxybenzyl)thiouracil treated with TFA

Solving equation **4** for the half-life of the reaction, the time is calculated for half the initial concentration (**Eq. 6,** Table 3.1).

$$\frac{1}{2} = e^{-kt_{\frac{1}{2}}} \quad \textbf{(5)}$$

$$t_{\frac{1}{2}} = -\frac{\ln 2}{k} \quad \textbf{(6)}$$

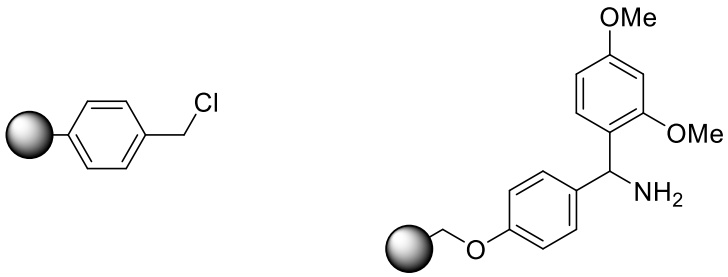
Table 3.1. Half-life of acidolysis of benzyl based protecting groups of thiouracil and thiothymine nucleobases

Protecting group	**Nucleobase**	**Rate constant (k)**	**Half-life ($t_{1/2}$) (min)**
4-methoxybenzyl	^{S}U	0.0111	62
4-methoxybenzyl	^{S}T	0.0070	99
2-methyl-4-methoxybenzyl	^{S}U	0.0527	13
2,4-dimethoxybenzyl	^{S}U	-	<1

The acidolysis of the benzyl based protecting groups of thiouracil and thiothymine demonstrated the highly acid-labile nature of the newly synthesized 2,4-dimethoxybenzyl protected thiouracil nucleobase. This protecting group was rapidly removed under treatment with mild TFA conditions, demonstrating a high sensitivity to deprotecting conditions. Furthermore, this study demonstrated the previously unstudied acid-lability of the 2-methyl-4-methoxybenzyl and 4-methoxybenzyl protecting previously developed for the synthesis of thiouracil PNA monomers. The unexpected acidolysis of the 4-methoxybenzyl protecting group suggests the protecting group is removed under standard boc-based oligomerization conditions, thus the thiol protection is likely unnecessary for oligomerization. To test this hypothesis, oligomerization was performed using an unprotected thiouracil PNA monomer and compared to previous oligomerization conditions.

3.3. Synthesis of pcPNA oligomers

The pcPNA oligomers were synthesized using standard fmoc-based solid phase peptide synthesis conditions. A series of 10-mer PNA oligomers were synthesized with the sequence NH_2-Lys-GXAGAXCACX-Lys-OH, where either the unmodified thymine PNA monomer, or the novel thiouracil or thiothymine PNA monomers were used in the synthesis of the 10-mer sequences. The PNA oligomers were synthesized on a polystyrene resin with a Rink amide linker (Figure 3.4). The polystyrene resin provides an insoluble support on which the PNA is oligomerizes which allows easy synthesis. The solid support allows the oligomer to be isolated by filtration and excess reagent is easily washed with dimethylformamide and dichloromethane. Washing the solid support with dichloromethane swells the resin, while washing with dimethylformamide contracts the resin. Alternating washes of dimethylformamide and dichloromethane results in a swelling and contracting of the resin which washes excess reagents through the polystyrene resin. The polystyrene resin is functionalized with Rink amide linkers to which the C-terminus of the peptide is attached. The Rink amide linker was developed as a more acid labile linker, compared to the Merrifield linker (Figure 3.4), which allows the oligomer to be removed from the polystyrene resin under treatment with milder acidic conditions. Linkers originally developed for boc-based peptide synthesis required stronger acidic conditions to remove the oligomer, as the linkage had to be resistant to treatment with TFA used for boc deprotection.



Merrifield linker Rink amide linker

Figure 3.4. Merrifield linker and Rink amide linker

Prior to automated oligomerization the polystyrene resin is preloaded with *Nα*-fmoc-*Nε*-boc-Lysine residue using HBTU and DIPEA as coupling agents to couple the C-terminus of the lysine to the Rink amide linker. Uncoupled linkers are capped using acetic anhydride to prevent additional coupling to the linker during oligomerization. Automated oligomerization is then carried out on the preloaded resin (K-PS) from the C-terminus to the N-terminus of the peptide using commercially available fmoc/bhoc protected PNA monomer and newly synthesized ^{S}U or ^{S}T PNA monomers. The *FastMoc* oligomerization procedure first uses a solution of 20 % 4-methylpiperidine to remove the N-terminal fmoc protecting group. The deprotection removal of the fmoc group results in the formation of dibenzofulvene piperidine adduct and piperidine-carbamate salt (Figure 3.5.), which are washed away with excess reagent. The deprotection of the terminal amine is monitored by conductivity monitoring, which detects conductivity resulting from the presence of the carbamate salt during deprotection. The prepared PNA monomers are mixed with a solution of HBTU and DIPEA to form an active ester on the C-terminus of the monomer. The activated PNA monomer is then mixed with the resin bound oligomer with the unprotected N-terminus to grow the oligomer strand. The N-terminus of the unreacted

oligomers are capped using acetic anhydride to prevent the synthesis of oligomers with a nucleotide deletion. Capping with the acetic anhydride truncates the oligomerization of the PNA strand. The deprotection, oligomerization and capping are repeated with each PNA monomer to synthesize the PNA strand. The PNA oligomer is cleaved from the polystyrene resin and removal of the bhoc and benzyl based protecting group from the nucleobases using a solution of 95 % TFA. The PNA oligomers were purified by high-pressure liquid chromatography (HPLC), and the purified oligomers were characterized by high-resolution mass spectrometry (Table 3.2).

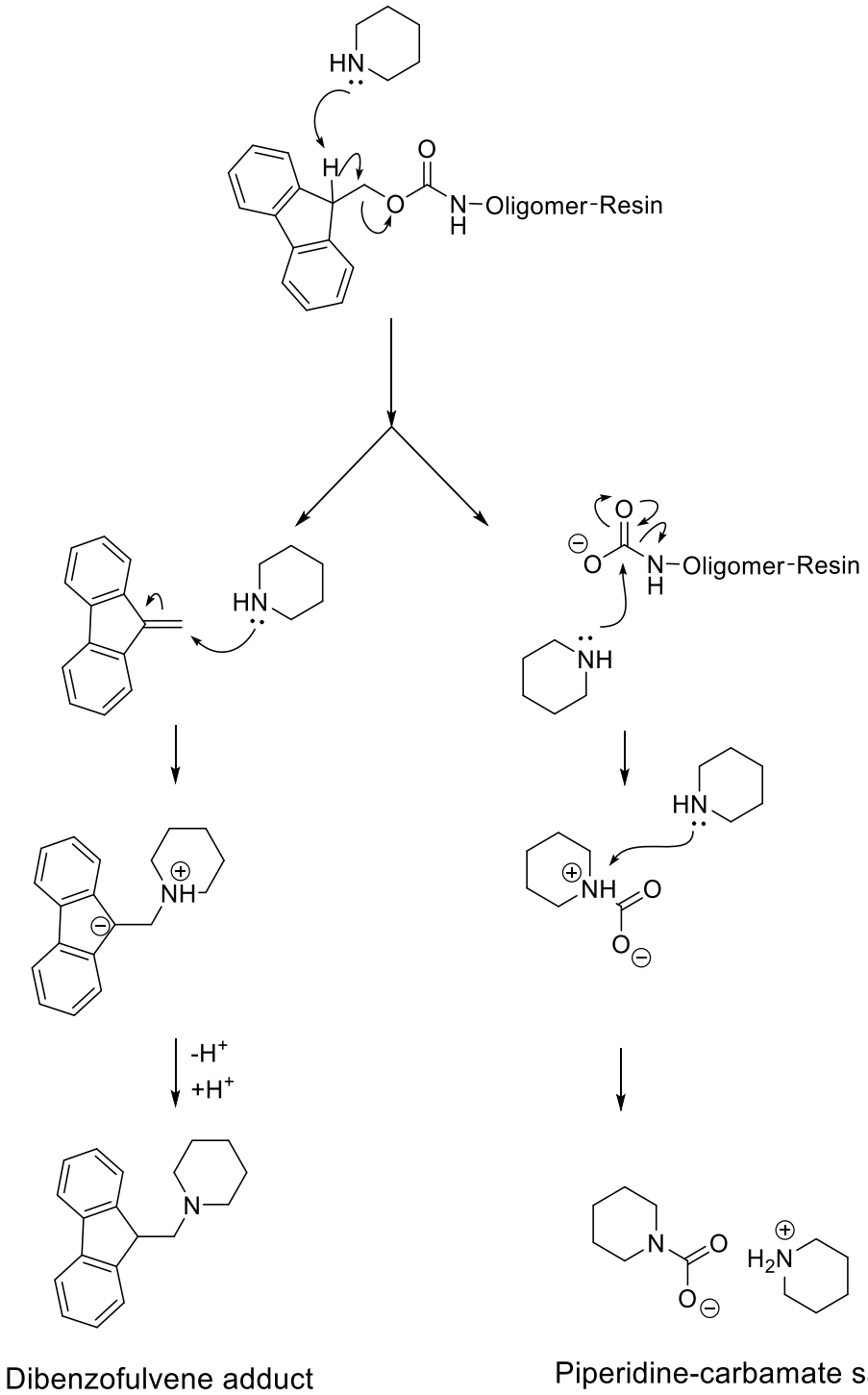


Figure 3.5. Byproducts of the cleavage of fmoc protecting group from *N*-terminus

To test the need for a protecting group on the thiouracil nucleobase during oligomerization, a sample of the prepared thiouracil PNA monomer (^{S}U(MMOB)) was

deprotected prior to oligomerization. The ^{S}U(MMOB) monomer was treated with the 95 % TFA solution for deprotection of the thiouracil nucleobase to yield the fmoc-^{S}U-AEG-OH PNA monomer for oligomerization. The thiouracil monomer without a protecting group on the nucleobase was then used for oligomerization under standard fmoc-based oligomerization conditions. The synthesized oligomer was purified by HPLC. The success of the oligomerization without a nucleobase protecting group was qualitatively analyzed by comparing the chromatogram of the unpurified oligomer to the chromatogram of the pcPNA oligomer synthesized using the ^{S}U(MMOB) monomer.

Table 3.2. Observed high-resolution mass of synthesized PNA oligomers

PNA	Protecting group	PNA sequence	Molecular formula	Calculated $[M+Na]^+$	Observed $[M+Na]^+$
PNA-T	None	KGTAGATCACTK	$C_{120}H_{159}N_{61}O_{33}$	3005.2536	3005.2551
PNA-^{S}U	4-methoxybenzyl	KG^{S}UAGA^{S}UCAC^{S}UK	$C_{117}H_{153}N_{61}O_{30}S_3$	3011.1382	3011.2430
PNA-^{S}U	None	KG^{S}UAGA^{S}UCAC^{S}UK	$C_{117}H_{153}N_{61}O_{30}S_3$	3011.1382	3011.2624
PNA-^{S}T	4-methoxybenzyl	KG^{S}TAGA^{S}TCAC^{S}TK	$C_{120}H_{159}N_{61}O_{30}S_3$	1538.0874[a]	1538.0963[a]

[a] Oligomer calculated and observed as the dicationic $[M+2Na]^{2+}$

The mass spectra of the synthesized PNA oligomers closely correlates with the calculated mass for the PNA oligomer. Additionally, mass signals were detected corresponding to the calculated $[M+2Na]^{2+}$ and $[M+3Na]^{3+}$ masses for all oligomers. The close correlation with the calculated masses suggests the complete synthesis of the PNA oligomers. This indicates the successful incorporation of the novel PNA monomers into the oligomer using standard fmoc-based oligomerization procedures. Additionally, using the thiouracil monomer with no protecting group on the thiol resulted in the successful

synthesis of the desired PNA oligomer. The mass of the oligomer synthesized without nucleobase protecting group matched that of the oligomer synthesized with the protected thiouracil nucleobase, while the HPLC chromatogram showed no significant impurities resulting from truncated oligomers during oligomerization. The oligomerization demonstrated that the protecting group on the thiol of the nucleobase was unnecessary for oligomerization.

3.4 Conclusions and future work

Studies of the deprotection rates of benzyl based protecting groups on the thiouracil nucleobase demonstrated the successful synthesis of a highly acid labile protecting group for fmoc-based synthesis of pcPNA oligomers. The addition of electron donating groups to the benzyl group stabilizes the benzylic carbocation resulting from the acidolysis of the thioether. Stabilizing the carbocation intermediate allows cleavage of the protecting group under milder acidic conditions. In addition to confirming the cleavage of the novel thiouracil protecting group under standard deprotection conditions, the study of the acidolysis discovered the cleavage of the previously synthesized 4-methoxybenzyl protecting group, used for boc-based oligomerization, under standard boc deprotection conditions. This suggested that a protecting group on the thiol was not necessary for PNA oligomerization.

The novel thiouracil and thiothymine PNA monomers were incorporated into PNA oligomers using standard fmoc-based oligomerization procedures. Characterization of the oligomers by mass spectrometry demonstrates the successful incorporation of the novel

thiouracil and thiothymine monomers into the PNA oligomer using standard fmoc-based oligomerization. Furthermore, the PNA oligomers were synthesized using thiouracil monomers without nucleobase protection, demonstrating the thiol protecting group is only necessary for directing alkylation to the *N*-1 position and is not necessary for oligomerization.

Future work will aim to test the stability of the synthesized PNA oligomer hybridization with complementary PNA and DNA oligomers. Hybridizing the PNA oligomers with a complementary target oligomer, then heating the duplex to determine the melting temperature of the duplex can be used to determine the stability of the PNA-PNA and PNA-DNA duplexes. Comparing the melting temperature of the thiouracil PNA monomer with the thiothymine PNA monomer can be used to determine the effect of replacing the thiouracil nucleobase with the thiothymine nucleobase on the stability of the PNA hybridization. Further future work will aim to develop a more effective synthesis of the fmoc-^{S}U-AEG-OH monomer. The synthesis of the PNA oligomer using the unprotected thiouracil nucleobase demonstrated the thiol group does not need protecting during oligomerization. Therefore, a more effective monomer synthesis could use a fmoc-AEG-O^{t}butyl backbone rather than the fmoc-AEG-Omethyl backbone. The ^{t}butyl ester protection of the C-terminus is removed under acidic conditions, rather than the basic conditions necessary for removal of the methyl ester protecting group. This is orthogonal to the fmoc deprotection conditions and would prevent unwanted deprotection of the N-terminus during ester hydrolysis. The acidic conditions would

remove the protecting group from both the nucleobase and the C-terminus, yielding the PNA monomer for oligomerization.

3.5. References

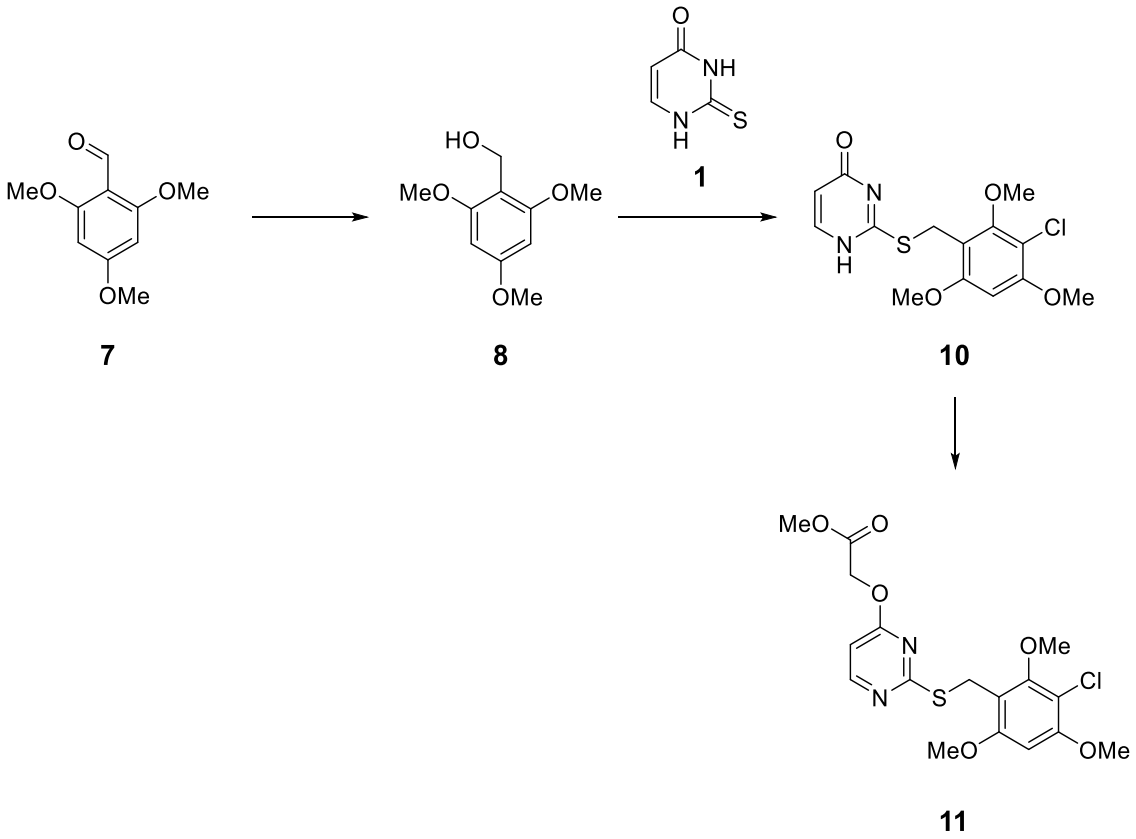
[1] Hammett, Louis P., *Chem. Rev.* **1935**, *17*, 125-136.

[2] Hammett, Louis P., *J. Am. Chem. Soc*. **1937**, *59*, 96-103.

[3] Ganguly, Sonali; Kundu, Kiron K., *Can. J. Chem.* **1994**, *72*, 1120-1126.

[4] Lohse, Jesper; Dahl, Otto; Nielsen, Peter E., *Proc. Natl. Acad. Sci.* **1999**, *96*, 11804-11808

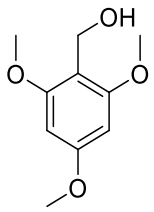
[5] Sugiyama, Toru; Hasegawa, Genki; Niikura, Chie; Kuwata, Keiko; Imamura, Yasutada; Demizu, Yosuke; Kurihara, Masaaki; Kittaka, Atsushi, *Bioorg. Med. Chem. Lett*. **2017**, *27*, 3337-3341



7 **8** **10** **11**

Scheme 4.1. Attempted synthesis of methyl 2-((2-((2,4,6-trimethoxybenzyl)thio)pyrimidin-4-yl)oxy)acetate

2,4,6-Trimethoxybenzyl alcohol (8)

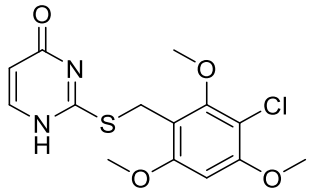


8

To a solution of 2,4,6-trimethoxybenzaldehyde (1.00 g, 5.10 mmol) and 0.5 M sodium hydroxide in methanol (30 mL) was added sodium borohydride (290 mg, 7.65 mmol) portionwise. The resultant solution was stirred for 2 h, then the reaction was

quenched with 30 mL water. The solution was concentrated under reduced pressure to approximately 30 mL, then the product was extracted into ether (3 x 20 mL). The organic layers were combined, washed with brine, dried over sodium sulfate, then the solvent evaporated to yield a clear, colourless oil which crystallized overnight (830 mg, 4.19 mmol, 82%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 6.12 (s, 2H), 4.71 (s, 1H), 4.69 (s, 1H), 3.82 (s, 6H), 3.81 (s, 3H). This conformed closely with the ^{1}H NMR spectrum previously reported in the literature.[1]

2-(2,4,6-Trimethoxy-3-chlorobenzyl)thiouracil (10)

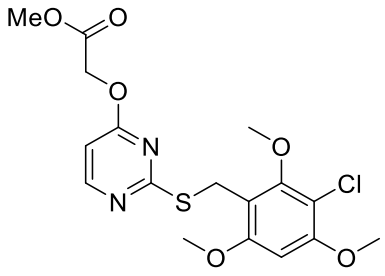


10

To a solution of **8** (2.00 g, 10.1 mmol) and triphenyl phosphine (1.94 g, 15.1 mmol) in 100 mL dichloromethane at 0 ^{o}C, was added *N*-chlorosuccinimide (1.62mg, 12.1 mmol). The solution was stirred for 1 h., then added to a suspension of 2-thiouracil (1.29 g, 10.1 mmol) and sodium carbonate (2.79 g, 20.2 mmol) in 100 mL dimethylformamide. The mixture was stirred for 90 min., then neutralized with 40 mL saturated citric acid. The solution was diluted with 150 mL water, then extracted in ethyl acetate (3 x 150 mL). The organic layers were combined, washed with water (4 x 150 mL), and brine (150 mL), dried over sodium sulfate, then the solvent was evaporated under reduced pressure to yielded

crude compound **10** combined with triphenylphosphine oxide as an off white solid (2.55 g). The crude product was used for further synthesis.

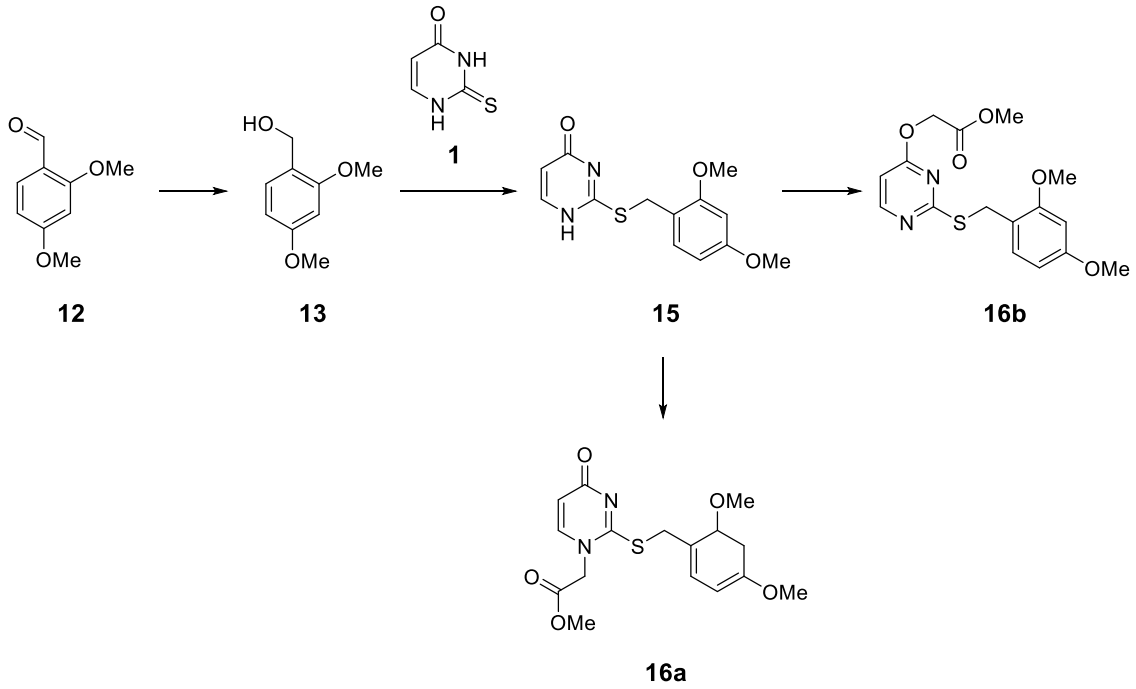
Methyl 2-((2-((2,4,6-trimethoxy-3-chlorobenzyl)thio)pyrimidin-4-yl)oxy)acetate (11)



11

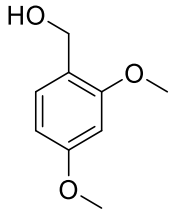
To a solution of crude **10** (2.55 g) and sodium carbonate (2.03 g, 11.1 mmol) in 120 mL dimethylformamide at 0 ^{o}C was added methyl bromoacetate (1.05 mL, 11.1 mmol), then the reaction was stirred at 0 ^{o}C for 15 min, then allowed to warmed to room temperature. The reaction was stirred for 4 h, then diluted with water (200 mL) and extracted with ethyl acetate (3 x 120 mL). The organic layers were combined, washed with water (4 x 100 mL) and brine (100 mL), dried with sodium sulfate, filtered, then the solvent was evaporated under reduced pressure, yielding a crude orange oil (3.84 g). The crude product was purified by column chromatography on silica gel (1:9 to 1:1 ethyl acetate: hexanes) to yield the pure product as light-yellow crystals (234 mg, 0.56 mmol, 5.6 % from (**8**)). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.24 (d, *J* = 5.7 Hz, 1H), 6.48 (d, *J* = 5.7 Hz, 1H), 6.27 (s, 1H), 4.89 (s, 2H), 4.37 (s, 2H), 3.85 (s, 3H), 3.82 (s, 3H), 3.79 (s, 3H), 3.71 (s, 3H). **^{13}C NMR** (400 MHz, $CDCl_3$) δ 172.2, 168.7, 167.5, 157.9, 157.7, 156.3, 156.0, 111.5, 108.6, 103.4,

92.6, 62.5, 61.8, 56.4, 56.1, 52.3, 24.7. **LRMS** (EI) calculated for $C_{17}H_{19}ClN_2O_6S$ $[M]^+$ 414.07, found 414.09.



Scheme 4.2. Synthesis of methyl 2-(2-((2,4-dimethoxybenzyl)thio)-4-oxopyrimidin-1(4H)-yl)acetate (**16a**)

2,4-Dimethoxybenzyl alcohol (13)

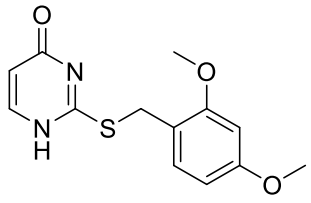


13

To a solution of 2,4-dimethoxybenzaldehyde (5.00 g, 30.09 mmol) and 0.5 M sodium hydroxide in 200 mL methanol was added sodium borohydride (1.37 g, 36.11 mmol) portionwise. The solution was stirred for 1 h, then the reaction was quenched with water. The solution was concentrated under reduced pressure, then the product was

extracted into ether. The organic layers were combined, washed with brine, dried over sodium sulfate, filtered, then the solvent was evaporated under reduced pressure to yield **13** as a clear, colourless oil which crystalized overnight (4.45 g, 26.46 mmol, 88 %). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.18 (d, *J* = 8.0 Hz, 1H), 6.52-6.42 (m, 2H), 4.63 (s, 2H), 3.86 (s, 3H), 3.82 (s, 3H), 2.27 (s, 1H). **^{13}C NMR** (400 MHz, $CDCl_3$) δ 160.7, 158.6, 129.7, 121.8, 103.9, 98.6, 61.7, 55.4, 55.3.

2-(2,4-Dimethoxybenzyl)thiouracil (15)

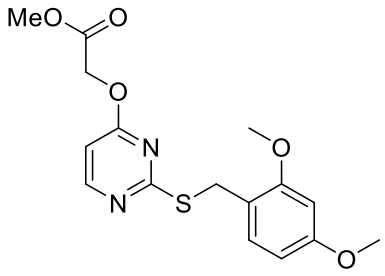


15

A solution of 2,4-dimethoxybenzyl alcohol (**13**) (4.12 g, 24.5 mmol) and pyridine (3.95 mL, 49.0 mmol) in 50 mL dry ether was cooled to 0 ^{o}C, then thionyl chloride (4.89 mL, 67.4 mmol) was added dropwise to solution over 15 min. The solution was stirred for 90 min at 0 ^{o}C, then quenched with 50 mL ice cold water. The ethereal layer was collected, and the aqueous layer was extracted with ether (3 x 25 mL). The organic layers were combined, washed with water (50 mL), and 5:1 saturated sodium chloride: saturated sodium bicarbonate (3 x 60 mL), and then dried over sodium sulfate. Dimethylformamide (10 mL) was added to solution, then the ether was evaporated under reduced pressure. The resulting dimethylformamide solution was added dropwise to a solution of 2-thiouracil (**1**) (1.57 g, 12.26 mmol) and triethylamine (2.05 mL, 14.71 mmol) in 50 mL

dimethylformamide at 0 °C. The resultant solution was stirred for 4 h at 0 °C, then diluted with water and extracted with dichloromethane (3 x 40 mL). The dichloromethane layers were combined, washed with water (4 x 50 mL), and brine (50 mL), dried over sodium sulfate then the solvent was evaporated under reduced pressure. The resulting white power was triturated with cold ether, yielding **14** as a pure white solid (2.41 g, 8.66, 71 %). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.89 (d, *J* = 6.6 Hz, 1H), 7.31 (d, *J* = 8.2 Hz, 1H), 6.49-6.40 (m, 2H), 6.20 (d, *J* = 6.6 Hz, 1H), 4.42 (s, 2H), 3.86 (s, 3H), 3.81 (s, 3H). **^{13}C NMR** (400 MHz, $CDCl_3$) δ 164.4, 162.6, 160.9, 158.4, 154.8, 131.4, 116.6, 110.8, 104.3, 98.5, 55.5, 55.4, 29.7. **HRMS** (EI) calculated for $C_{13}H_{14}N_2O_3S$ $[M]^+$ 278.0725, found 278.0720.

Methyl 2-((2-((2,4-dimethoxybenzyl)thio)pyrimidin-4-yl)oxy)acetate (16b)

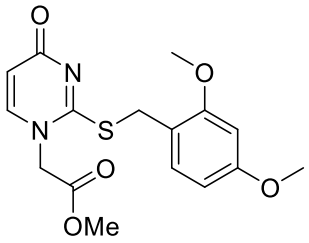


16b

To a solution of **14** (1.00 g, 3.59 mmol) and sodium carbonate (1.24 g, 8.98 mmol) in 60 mL dimethylformamide was added methyl bromoacetate (0.52 mL, 5.39 mmol) then the solution was heated to 80 °C. The reaction was stirred for 2 h at 80 °C, then diluted with water (70 mL) and extracted with dichloromethane (4 x 50 mL). The dichloromethane layers were combined, washed with water (4 x 60 mL) and brine (60 mL), dried over sodium sulfate, filtered, then the solvent was evaporated under reduced pressure,

yielding a yellow solid (1.23 g, 3.37 mmol, 94%). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.31 (d, *J* = 5.7 Hz, 1H), 7.31 (d, *J* = 8.2 Hz, 1H), 6.55 (d, *J* = 5.7 Hz, 1H), 6.50-6.40 (m, 2H), 4.93 (s, 2H), 4.34 (s, 2H), 3.86 (s, 3H), 3.81 (s, 3H), 3.78 (s, 3H). **^{13}C NMR** (400 MHz, $CDCl_3$) δ 172.0, 168.6, 167.5, 160.4, 158.6, 157.8, 130.8, 117.6, 104.1, 103.5, 98.6, 62.4, 55.5, 55.4, 52.2, 29.7.

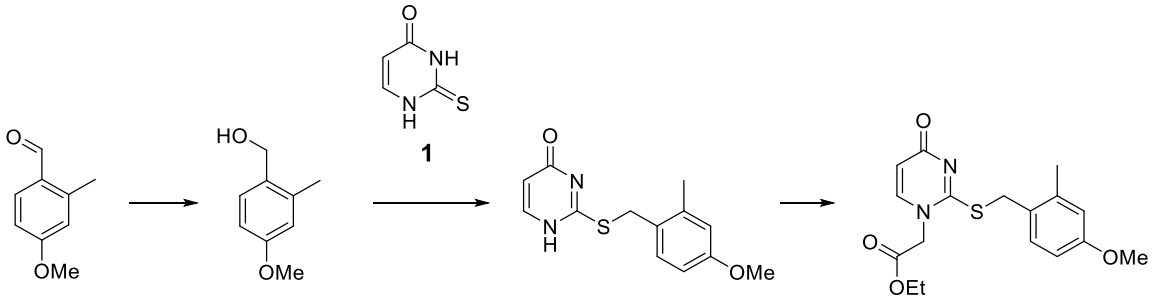
Methyl 2-(2-((2,4-dimethoxybenzyl)thio)-4-oxopyrimidin-1(4H)-yl)acetate (16a)



16a

To a solution of sodium (165 mg, 7.18 mmol) in 15 mL methanol was added **7** (1.00 g, 3.59 mmol), then the solution was stirred and heated to reflux. Methyl bromoacetate (0.79 mL, 7.18 mmol) was added to the reaction, the solution was stirred for 2 h at reflux, then cooled to room temperature and the solvent was evaporated under reduced pressure. The resulting residue was dissolved in 25 mL water then extracted with a solution of 3:1 dichloromethane: methanol (2 x 80 mL). The organic layers were combined, dried over sodium sulfate, filtered, then the solvent was evaporated under reduced pressure to yield a white residue. The residue was purified by column chromatography on silica gel (for **15**, 1% methanol in dichloromethane; for **16a**, 5% methanol in dichloromethane to 10% methanol in dichloromethane), yielding **16a** as a pure white solid (289 mg, 0.79 mmol, 22 %). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.34 (d, *J* = 8.9

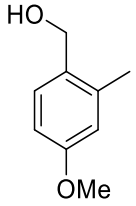
Hz, 1H), 7.07 (d, *J* = 7.6 Hz, 1H), 6.41-6.34 (m, 2H), 6.02 (d, *J* = 7.6 Hz, 1H), 4.47 (s, 2H), 4.44 (s, 2H), 3.77 (s, 3H), 3.74 (s, 3H), 3.73 (s, 3H). **^{13}C NMR** (400 MHz, $CDCl_3$) δ 167.8, 166.4, 163.8, 160.9, 158.6, 143.6, 131.8, 115.9, 110.0, 104.1, 98.5, 55.3, 53.1, 52.6, 41.9, 31.3. **HRMS** (EI) calculated for $C_{16}H_{18}N_2O_5S$ $[M]^+$ 350.0936, found 350.0949.



17 **18** **20** **21**

Scheme 4.3. Synthesis of ethyl 2-(2-((2-methyl-4-methoxybenzyl)thio)-4-oxopyrimidin-1(4H)-yl)acetate (**21**)

2-Methyl-4-methoxybenzyl alcohol (18)

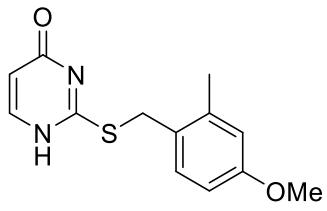


18

To a solution of 2-methyl-4-methoxybenzaldehyde (2.03 g, 12.8 mmol) in 32 mL ethanol was added sodium borohydride (519 mg, 13.7 mmol). The resultant solution was stirred for 15 min., then the reaction was quenched by the addition of 240 mL water. The solution was acidified with 4 M hydrochloric acid, then the product was extracted in diethyl ether (3 x 80 mL). The ethereal layers were combined, washed with water (60 mL), and brine (60 mL), then dried over sodium sulfate. The solvent was evaporated under reduced pressure to yield the pure product as a clear, colourless oil (2.00 g, 12.8 mmol).

^{1}H NMR (400 MHz, $CDCl_3$) δ 7.19 (d, 1H, *J* = 8.3 Hz), 6.71 (d, 1H, *J* = 2.7 Hz), 6.68 (dd, 1H, *J* = 2.7, 8.3 Hz), 4.59 (s, 2H), 3.76 (s, 3H), 2.33 (s, 3H), 1.59 (s, 1H). This closely corresponded to the ^{1}H NMR spectrum previously reported in the literature.[2]

2-(2-Methyl-4-methoxybenzyl)thiouracil (20)

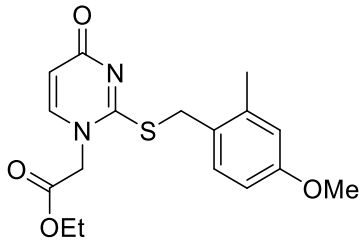


20

To a solution of 2-methyl-4-methoxybenzyl alcohol (2.00 g, 12.8 mmol) in 80 mL dry dichloromethane at 0 °C was added thionyl chloride (1.05 mL, 14.4 mmol) dropwise, then the solution was stirred for 2 h at 0 °C. The reaction was added to a solution of ice cold saturated sodium bicarbonate (250 mL), the dichloromethane layer was separated, then the aqueous layer was extracted with diethyl ether (2 x 60 mL). The organic layers were combined, washed with water (80 mL), and brine (80 mL), then dried over sodium sulfate. The solvent was evaporated under reduced pressure to yield a clear colourless oil. The resulting oil was dissolved 8 mL ethanol then added to a solution of 2-thouracil (833 mg, 6.58 mmol) in 8 mL ethanol and 8 mL aqueous potassium hydroxide (474 mg, 4.45 mmol KOH in H_2O). The resultant reaction was stirred overnight at room temperature, then the solvent was evaporated under reduced pressure. A solution of saturated sodium bicarbonate (20 mL) was added to the resultant residue then the precipitate was collected by filtration, washed with water, ethanol, ethyl acetate, and

diethyl ether to yield the product as pure white crystals (1.49 g, 5.60 mmol, 86 %). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 7.29 (d, 1H, *J* = 8.3 Hz), 6.79 (d, 1H, *J* = 2.1 Hz), 6.72 (dd, 1H, *J* = 2.1 Hz, 8.3 Hz), 4.35 (s, 2H), 3.72 (s, 3H), 2.32 (s, 3H). This closely corresponded to the ^{1}H NMR spectrum previously reported in the literature.[2]

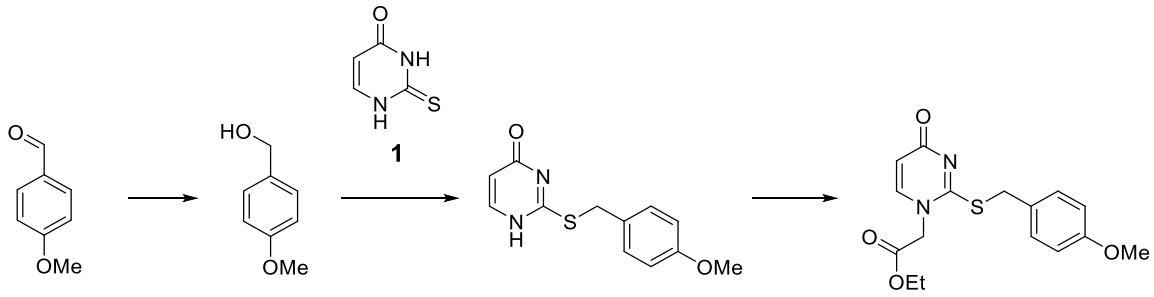
Ethyl 2-(2-((2-methyl-4-methoxybenzyl)thio)-4-oxopyrimidin-1(4H)-yl)acetate (21)



21

To a solution of sodium (273 mg, 11.4 mmol) dissolved in ethanol (13 mL) was added **2-**(2-methyl-4-methoxybenzyl)thiouracil (1.49 g, 5.60 mmol). The resultant suspension was stirred and heated to reflux until the solids dissolved, then ethyl bromoacetate (1.26 mL, 11.4 mmol) was added dropwise, then the solution was stirred at reflux for 2 h. The mixture was cooled to room temperature, then the ethanol was evaporated under reduced pressure. Water (15 mL) was added to the resulting residue, and the product was extracted in a solution of 3:1 dichloromethane: methanol (2 x 60 mL). The organic layers were combined, dried over sodium sulfate, then evaporated under reduced pressure. The resultant residue was purified by column chromatography on silica gel (ethyl acetate to 9:1 ethyl acetate: methanol) to yield the product as a pure white solid (781 mg, 2.2 mmol, 40 %). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.25 (d, 1H, *J* = 8.2 Hz), 7.08 (d, 1H, *J* = 7.6 Hz), 6.69 (d, 1H, *J* = 2.6 Hz), 6.66 (dd, 1H, *J* = 2.6, 8.2 Hz), 6.08 (d, 1H, *J*

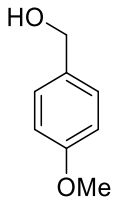
= 7.6 Hz), 4.45 (s, 2H), 4.44 (s, 2H), 4.20 (q, 2H, *J* = 7.1 Hz), 3.74 (s, 3H), 2.30 (s, 3H), 1.23 (t, 3H, *J* = 7.1 Hz). This corresponded closely to the ^{1}H NMR spectrum previously reported in the literature.[2]



22 **23** **2** **3**

Scheme 4.4. Synthesis of ethyl 2-(2-((4-methoxybenzyl)thio)-4-oxopyrimidin-1(4H)-yl)acetate (**3**)

4-Methoxybenzyl alcohol (23)

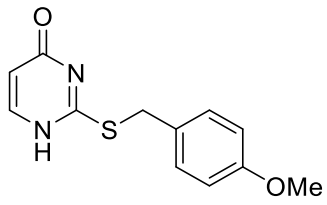


23

To a solution of 4-methoxybenzaldehyde (7.00 g, 51.4 mmol) in 60 mL ethanol was added sodium borohydride (2.92 g, 77.1 mmol). The resultant solution was stirred for 30 min, then the reaction was quenched by the addition of 250 mL water. The solution was acidified with 4 M hydrochloric acid, then the product was extracted into diethyl ether (3 x 100 mL). The ethereal layers were combined, washed with water (80 mL), and brine (80 mL), then dried over sodium sulfate. The solvent was evaporated under reduced pressure to yield the pure product as a clear, colourless oil (6.93 g, 50.1 mmol, 97 %). **^{1}H NMR** (400

MHz, $CDCl_3$) δ 7.27 (d, *J* = 8.4 Hz, 2H), 6.86 (d, *J* = 8.4 Hz, 2H), 4.58 (s, 2H), 3.77 (s, 3H), 1.61 (br s, 1H). This closely corresponded to the ^{1}H NMR spectrum previously reported in the literature.[3]

2-(4-Methoxybenzyl)thiouracil (2)

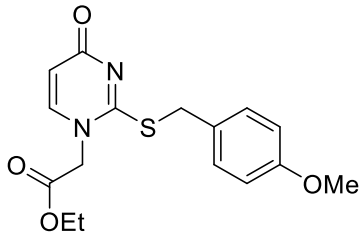


2

To a solution of 4-methoxybenzyl alcohol (6.93 g, 50.1 mmol) in 150 mL dry dichloromethane at 0 °C was added thionyl chloride (4.11 mL, 56.4 mmol) dropwise, then the solution stirred at 0 °C. After 2 h., the reaction was added to a solution of ice cold saturated sodium bicarbonate (300 mL), the dichloromethane layer was separated and the aqueous layer was extracted with diethyl ether (2 x 100 mL). The organic layers were combined, washed with water (120 mL), and brine (120 mL), then dried over sodium sulfate. The solvent was evaporated under reduced pressure to yield a clear colourless oil. The resultant oil was dissolved in 15 mL ethanol and then added to a solution of 2-thouracil (5.77 g, 45.8 mmol) in 15 mL ethanol and 15 mL aqueous potassium hydroxide (3.34 g, 56.5 mmol KOH in H_2O). The resultant reaction was stirred overnight at room temperature, then the solvent was evaporated under reduced pressure. A solution of saturated sodium bicarbonate (50 mL) was added to the resultant residue, then the precipitate was collected by filtration, washed with water, ethanol, ethyl acetate, and

diethyl ether to yield the product as pure white crystals (5.61 g, 22.6 mmol, 50 %). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 11.69 (br s, 1H), 7.88 (d, *J* = 6.6 Hz, 1H), 7.32 (d, *J* = 8.7 Hz, 2H), 6.86 (d, *J* = 8.7 Hz, 2H), 6.22 (d, *J* = 6.6 Hz, 1H), 4.42 (s, 2H), 3.81 (s, 3H). This corresponded closely with the ^{1}H NMR spectrum previously reported in the literature.[3]

Ethyl 2-(2-((4-methoxybenzyl)thio)-4-oxopyrimidin-1(4H)-yl)acetate (3)



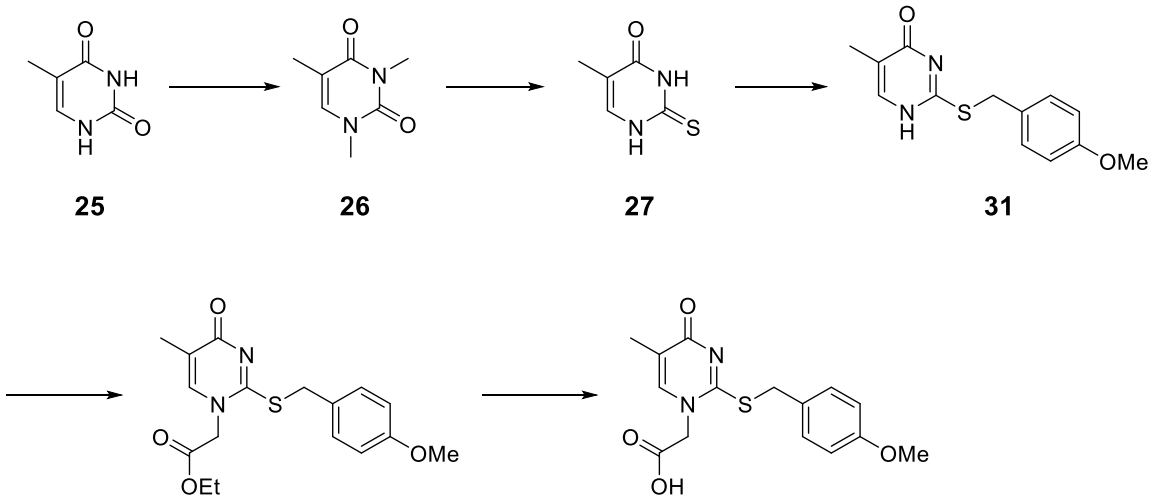
3

To a solution of refluxing ethanol 80 mL was dissolved sodium (370 mg, 16.1 mmol), then *S*-(4-methoxybenzyl)thiouracil (**2**) (3.34 g, 13.4 mmol) was added to the solution. Once the 2-(4-methoxybenzyl)thiouracil dissolved, ethyl bromoacetate (1.78 mL, 16.1 mmol) was added dropwise to the solution, and reaction was stirred at reflux for 90 min. The reaction was evaporated to dryness, then water was added to the residue (100 mL). The aqueous solution was extracted with a solution of 3:1 dichloromethane: methanol (2 x 200 mL). The organic layers were combined, dried over sodium sulfate, filtered, then evaporated to dryness under reduced pressure. The resultant residue was taken up into a solution of 1:1 ethyl acetate: hexanes, the solvent was evaporated under reduced pressure, and the resulting residue was triturated with ice cold ethyl acetate to yield the pure product as a white solid (712 mg, 2.12 mmol, 16 %). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.34 (d, *J* = 8.6 Hz, 2H), 7.11 (d, *J* = 7.6 Hz, 1H), 6.87 (d, *J* = 8.6 Hz, 2H), 6.13 (d, *J*

= 7.6 Hz, 1H), 4.504 (s, 2H), 4.499 (s, 2H), 4.28 (q, *J* = 7.2 Hz, 2H), 3.81 (s, 3H), 1.29 (t, *J* = 7.2 Hz, 3H). This closely corresponded with the ^{1}H NMR spectra previously reported in the literature.[3]

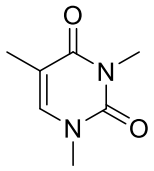
Thioester acidolysis

To a solution of 0.02 mL trifluoroacetic acid, 0.02 mL triethylsilane and 0.96 mL deuterated chloroform was added to 3 mg of protected thiouracil or thiothymine (**3, 16a, 21, 32**) and the reaction was monitored by ^{1}H NMR at 2 min. intervals until the reaction was completed. The reaction was carried out at 25 ^{O}C and monitored for 2 h, or until reaction had gone to completion. Integration of benzylic protons was monitored to determine the rate of deprotection.



Scheme 4.5. Synthesis of 2-(2-((4-methoxybenzyl)thio)-5-methyl-4-oxo-3,4-dihydropyrimidin-1(2H)-yl)acetic acid (**33**)

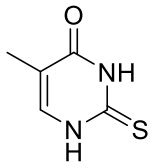
1,3-Dimethylthymine (26)



26

To a solution of sodium hydroxide (9.5 g, 237.9 mmol) and thymine (5.00 g, 39.6 mmol) in a solution of 20 mL water was added dimethyl sulfate (18.8 mL, 198.2 mmol) dropwise over 30 min. The solution was stirred for 2 h. until the starting material was consumed. The solution was extracted with dichloromethane (3 x 150 mL), the organic layers were combined, dried over sodium sulfate, filtered, then the solvent evaporated under reduced pressure to yield a white solid (5.1 g, 33.1 mmol, 84 %). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 7.58 (s, 1 H), 3.33 (s, 3 H), 3.27 (s, 3 H), 1.80 (s, 3 H). This closely corresponded to the ^{1}H NMR spectrum previously reported in the literature.[4]

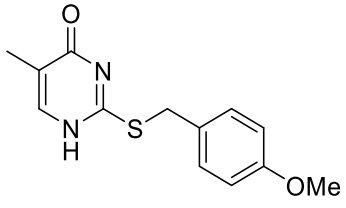
Thiothymine (27)



27

To a solution of sodium (2.7 g, 115.8 mmol) in 80 mL ethanol was added 1,3-dimethylthymine (5.1 g, 33.1 mmol) and thiourea (10.1 g, 132.3 mmol). The solution was stirred at reflux for 48 h., then the solvent was evaporated under reduced pressure. The resultant residue was dissolved in 30 mL water and acidified with concentrated hydrochloric acid. The precipitated was collected by filtration, yielding a pure white solid (1.6 g, 11.3 mmol, 34 %). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.39 (s, 1H), 12.13 (s, 1H), 7.32 (s, 1H), 1.78 (s, 3H). **^{13}C NMR** (400, DMSO-d_6) δ 175.1, 162.4, 138.5, 114.3, 12.4. This closely corresponded to the ^{1}H NMR spectrum previously reported in the literature.[5]

2-(4-Methoxybenzyl)thiothymine (31)

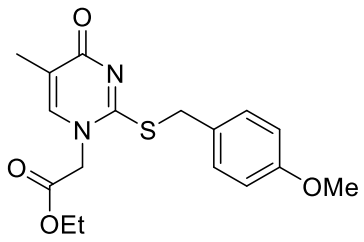


31

To a solution of 25 mL ethanol was added thiothymine (1.00 g, 7.00 mmol) followed by aqueous potassium hydroxide (510 mg, 9.10 mmol) in 15 mL water. The solution was heated to 45 ^{o}C then 4-methoxybenzyl chloride (1.42 mL, 10.5 mmol) was

added to the reaction. The reaction was stirred for 8 h., then evaporated to dryness under reduced pressure. The resultant residue was suspended in a solution of 10 % sodium bicarbonate (20 mL), then the precipitate was collected by filtration, and washed with water, methanol, and diethyl ether to yield a pure white solid (1.3 g, 4.96 mmol, 71 %). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 12.66 (br, 1H), 7.76 (s, 1H), 7.32 (d, *J* = 8.6 Hz, 2H), 6.87 (d, *J* = 8.6 Hz, 2H), 4.32 (s, 2H), 3.73 (s, 3H), 1.87 (s, 3H). **HRMS** (EI) calculated for $C_{13}H_{14}N_2O_2S$ $[M]^+$ 262.0776, found 262.0770.

Ethyl 2-(2-((4-methoxybenzyl)thio)-5-methyl-4-oxo-3,4-dihydropyrimidin-1(2H)-yl)acetate (32)

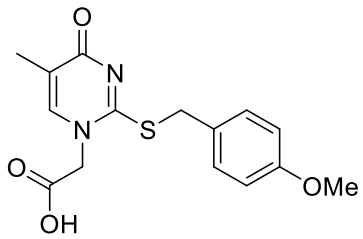


32

To a solution of sodium (210 mg, 9.15 mmol) in ethanol (15 mL) at reflux, was added 2-(4-methyoxybenzyl)thiothymine (1.2 g, 4.57 mmol) followed by ethyl bromoacetate (1.0 mL, 9.15 mmol). The reaction was stirred for 2 h. at reflux, then the solvent was evaporated under reduced pressure. The residue was dissolved in water (15 mL) and extracted with a solution of 3:1 dichloromethane: methanol (2 x 60 mL). The organic layers were combined, dried over sodium sulfate, and evaporated under reduced pressure. The crude oil was washed with a solution of 1:1 ethyl acetate: hexanes (3 x 10 mL), the washes were combined and evaporated under reduced pressure. The resultant

solid was suspended in ice cold ethyl acetate and the product was collected by filtration, yielding a pure white solid (612 mg, 38 %). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 7.62 (s, 1 H), 7.34 (d, *J* = 8.7 Hz, 2 H), 6.88 (d, *J* = 8.7 Hz, 2 H), 4.76 (s, 2 H), 4.37 (s, 2 H), 4.17 (q, *J* = 7.1, 2 H), 3.73 (s, 3 H), 1.82 (s, 3 H), 1.18 (t, *J* = 7.1, 3 H). **^{13}C NMR** (400 MHz, DMSO) δ 167.9, 167.5, 159.2, 142.1, 130.9, 128.6, 117.4, 114.4, 62.2, 55.5, 52.8, 34.9, 14.4, 13.8. **HRMS** (EI) calculate for $C_{17}H_{20}N_2O_4S$ $[M]^+$ 348.1144, found 348.1145.

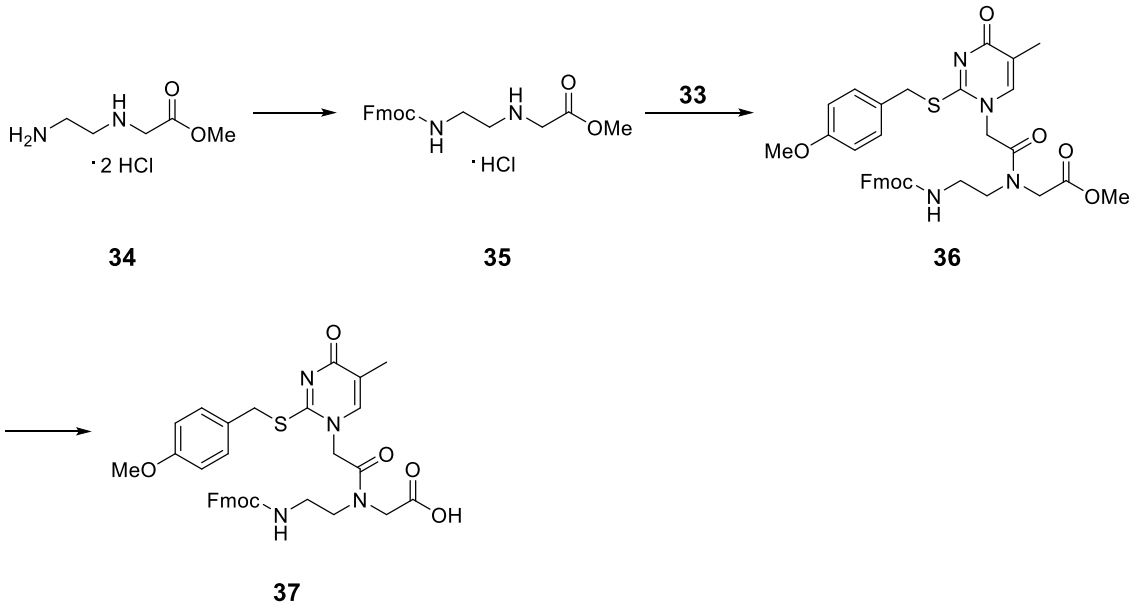
2-(2-((4-Methoxybenzyl)thio)-5-methyl-4-oxo-3,4-dihydropyrimidin-1(2H)-yl)acetic acid (33)



33

To a solution of **32** (250 mg, 0.72 mmol) dissolved in methanol (2.2 mL) and tetrahydrofuran (5 mL) was added 0.72 mL of 2 M lithium hydroxide, then the solution was stirred for 15 min. The solvent was evaporated under reduced pressure, then the residue was dissolved in 5 mL water and the solution was acidified with 2 M hydrochloric acid. The resultant precipitate was filtered and washed with cold ethyl acetate and diethyl ether which yielded a pure white solid (185 mg, 0.38 mmol, 80 %). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 13.53 (s, 1 H), 7.62 (s, 1 H), 7.34 (d, *J* = 8.4 Hz, 2 H), 6.88 (d, *J* = 8.4 Hz, 2 H), 4.64 (s, 2 H), 4.37 (s, 2 H), 3.73 (s, 2 H), 1.82 (s, 3 H). **^{13}C NMR** (400 MHz, DMSO) δ 168.9,

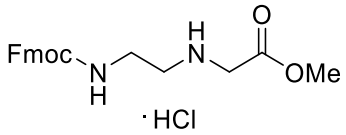
167.9, 161.6, 159.1, 142.3, 130.9, 128.7, 117.5, 114.4, 55.6, 52.9, 34.8, 13.8. **HRMS** (EI) calculated for $C_{15}H_{16}N_2O_4S$ $[M]^+$ 320.0831, found 320.0828.



34 **35** **36** **37**

Scheme 4.6. Synthesis of Fmoc-thiothymine PNA monomer (**37**)

Methyl-*N*-(2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)ethyl)glycinate hydrochloride (35)

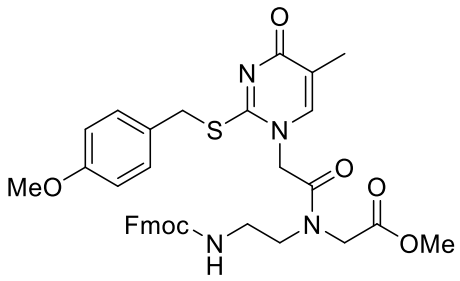


35

To a suspension of **34** (1.00 g, 4.87 mmol) and Fmoc-succinimide (1.80 g, 5.36 mmol) in 16 mL dioxane was added a solution of sodium bicarbonate (1.23 g, 14.61 mmol) in 20 mL water, then the resulting solution was stirred for 50 min. The solution was diluted with 60 mL water, then the white precipitate was extracted with dichloromethane (3 x 50

mL). The dichloromethane layers were combined, dried over sodium sulfate, filtered, and the resultant solution was acidified with ethereal hydrochloric acid. The solution was cooled overnight and the resultant precipitate filtered to yield the product (**35**) as a pure white solid (432 mg, 23 %). **^{1}H NMR** (400 MHz, DMSO-d_6) δ 9.21 (br s, 2H), 7.90 (d, *J* = 7.5 Hz, 2H), 7.70 (d, *J* = 7.5 Hz, 2H), 7.54 (t, *J* = 5.1, 1H), 7.42 (t, *J* = 7.5 Hz, 2H), 7.34 (t, *J* = 7.5 Hz, 2H), 4.34 (d, *J* = 6.7 Hz, 2H), 4.23 (t, *J* = 6.8, 1H), 3.96 (s, 2H), 3.74 (s, 3H), 3.33 (overlapping with water), 3.00 (t, *J* = 6.4 Hz, 2H). **^{13}C** NMR (400 MHz, DMSO-d_6) δ 167.9, 156.8, 144.3, 141.2, 128.1, 127.6, 125.6, 120.6, 66.1, 55.4, 53.0, 47.3, 47.1, 37.3. This closely corresponded to ^{1}H NMR spectra previously reported in the literature.[6]

Methyl-*N*-(2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)ethyl)-N-(2-(2-((4-methoxybenzyl)thio)-5-methyl-4-oxopyrimidin-1(4H)-yl)acetyl) glycinate (36)

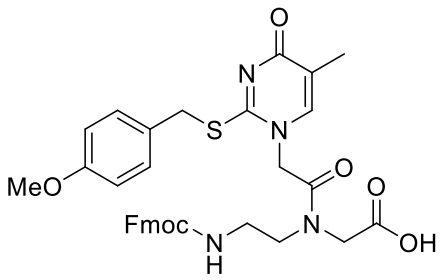


36

To a solution of **33** (280 mg, 0.87 mmol) in 7 mL dimethylformamide at 0 ^{o}C was added HBTU (338 mg, 0.89 mmol) and diisopropylethylamine (0.48 mL, 1.55 mmol) and stirred for 2 min., then **35** (375 mg, 0.96 mmol) was added to the reaction and the mixture was stirred for 2 h., then allowed to warm to room temperature. The reaction mixture was diluted with 15 mL ethyl acetate, poured into 40 mL water, then extracted with ethyl

acetate (3 x 20 mL). The organic layers were combined, washed with 5 % citric acid in water (20 mL), saturated sodium bicarbonate (20 mL), and brine (20 mL), then dried over sodium sulfate, filtered, after which the solvent was evaporated under reduced pressure. This yielded a crude orange oil which was purified by column chromatography on silica gel (1 % methanol-dichloromethane to 5 % methanol-dichloromethane) to yield compound **36** as a pure white solid (312 mg, 0.48mmol, 55 %) HRMS (ESI) calculated for $C_{35}H_{36}N_4O_7SNa$ $[M+Na]^+$ 679.2202, found 679.2205.

***N*-(2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)ethyl)-*N*-(2-(2-((4-methoxybenzyl)thio)-5-methyl-4-oxopyrimidin-1(4H)-yl)acetyl)glycine (37)**



37

To a solution of **36** (312 mg, 0.48 mmol) in 8 mL tetrahydrofuran at 0 ^{O}C was added 8 mL 2 M lithium hydroxide then the solution was stirred for 30 min. The reaction was neutralized with 2 M hydrochloric acid then the solvent was evaporated under reduced pressure. The residue was suspended in 10 mL ice cold water and the resultant precipitate was collected by filtration. The solid was suspended in cold dichloromethane, then the product was collected by filtration to yield the PNA monomer (**37**) as a pure white solid (126 mg, 20 mmol, 41 %) **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.06 (m, 1H), 7.89 (d, *J* = , 2H),

7.67 (m, 2H), 7.43-7.27 (m, 7H), 6.80 (d, *J* = , 2H), 4.70 (s, 2H), 4.28 (s, 2H), 4.16 (s, 2H), 3.69 (s, 3H), 3.64 (s, 2H), 3.59 (s, 1H), 3.38 (m, 2H), 3.15 (m, 2H), 1.73 (s, 3H). **^{13}C NMR** (400 MHz, DMSO) δ 170.2, 168.2, 166.8, 162.0, 159.0, 156.5, 144.4, 142.5, 141.1, 130.8, 129.4, 128.8, 128.1, 127.8, 127.6, 125.9, 121.9, 120.5, 117.1, 114.3, 66.0, 55.5, 53.5, 53.0, 48.4, 47.2, 38.4, 34.9, 13.7. HRMS (ESI) calculated for $C_{34}H_{34}N_4O_7SNa$ $[M+Na]^+$ 665.2046, found 665.2030.

Oligomer synthesis

The PNA oligomers were synthesized using an ABI 433A peptide synthesizer manufactured by Perkin Elmer Applied Biosystems. Oligomerization was carried out using newly synthesized ^{S}U and ^{S}T monomers, commercially available PNA monomers: Fmoc-A(Bhoc)-AEG-OH, Fmoc-G(Bhoc)-AEG-OH, and Fmoc-C(Bhoc)-AEG-OH (purchased from PolyOrg, Inc.), and *N*α-Fmoc-*N*ε-Boc-Lysine (purchased from Chem-Impex Int'l Inc.), using standard *FastMoc* oligomerization methods for Fmoc-based solid-phase peptide synthesis. Fmoc-RAM-PS was used as a solid support resin preloaded with lysine at 0.057 mmol/g. Synthesis was carried out on a 5.0 µmol scale. Monomers were prepared at a concentration of 25 µmol dissolved in 110 µL. Solutions of 0.4 M diisopropylethylamine in *N*-methyl-2-pyrrolidone (NMP) and 0.19 M HBTU in NMP were prepared for monomer coupling. Fmoc deprotection was performed using a solution of 20 % 4-methylpiperidine in dimethylformamide. Unreacted terminal amino groups were capped with acetic anhydride, using a solution of 1:25:25 acetic anhydride: pyridine: NMP. Following automated synthesis, the resin was treated with a solution of 95 % trifluoroacetic acid and 5 % triethylsilane to cleave the oligomer from the resin and remove the protecting

group from the nucleobases (Bhoc) and amino group (Boc). The solvent was then evaporated under a nitrogen stream, the resultant residue was washed twice with cold ether, dissolved in a solution of 0.05% trifluoroacetic acid in water then purified by reverse-phase HPLC. Reverse-phase HPLC was performed on an Agilent Microsorb-MV 100-5 C18 250 x 4.6 mm column heated to 50 ^{o}C. The purified PNA oligomer was eluted using a gradient (water/0.1 % trifluoroacetic acid to acetonitrile/0.1 % trifluoroacetic acid) over 60 min.

Appendix of Spectra

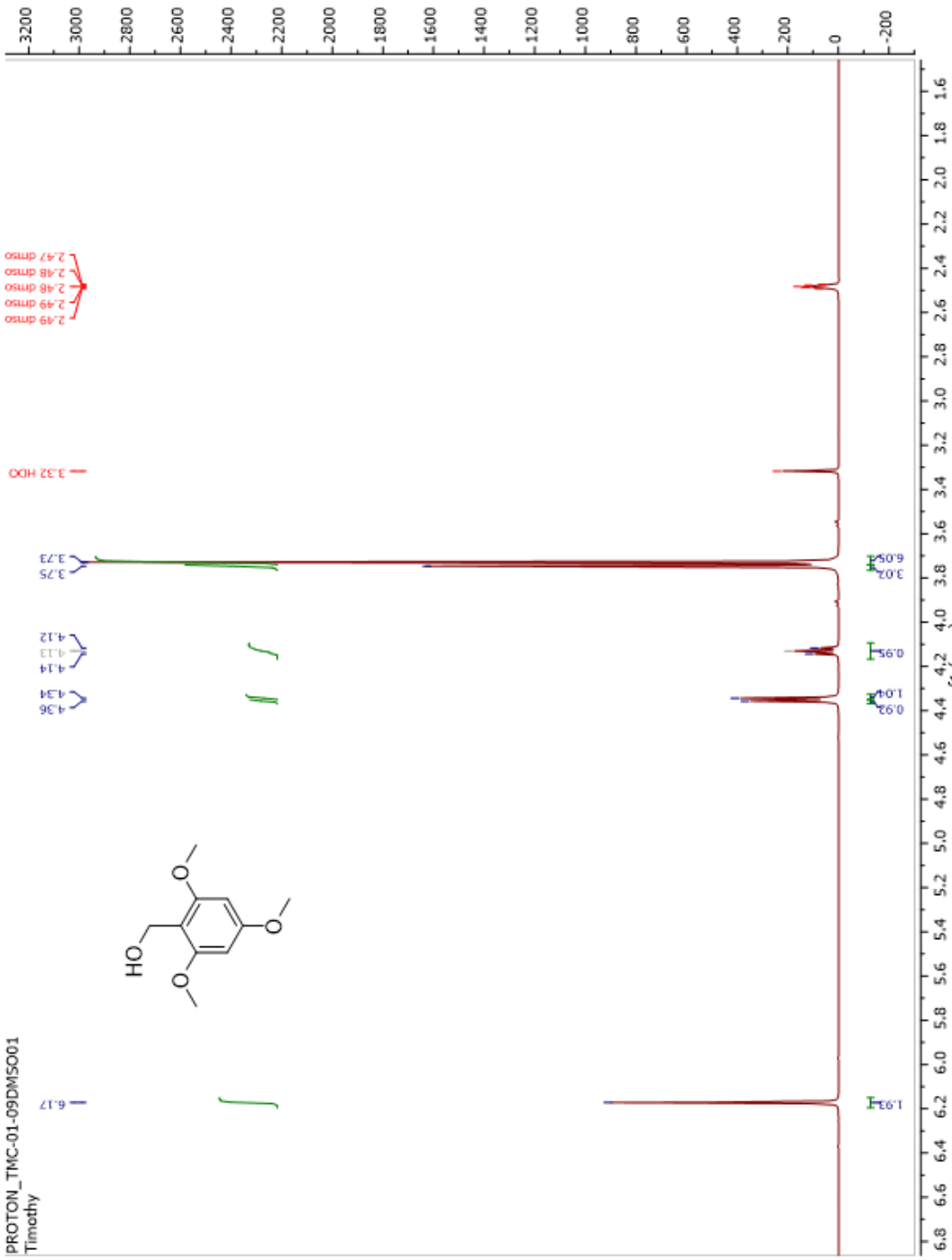


Figure A.1. ^{1}H NMR spectra of **8**

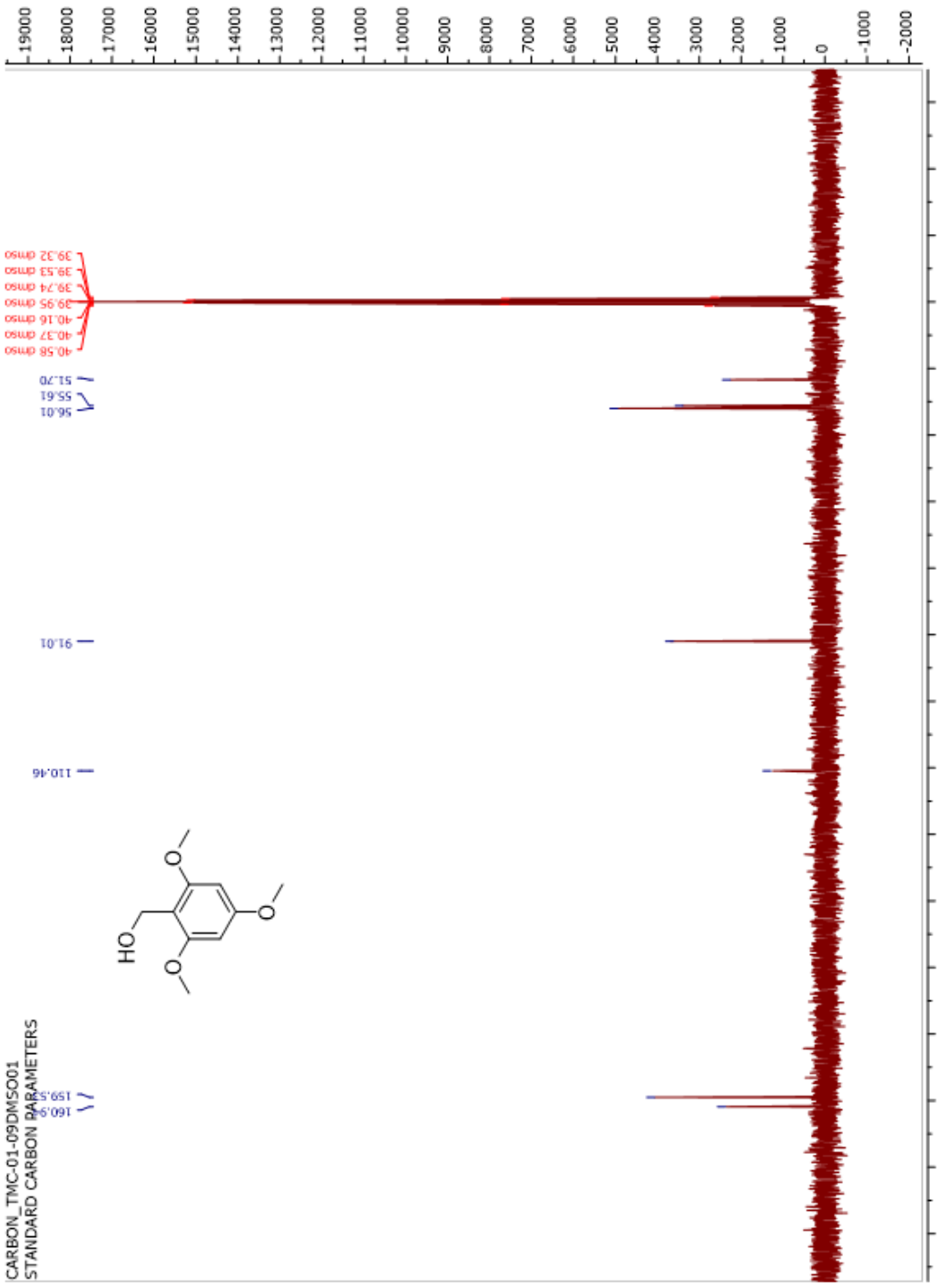


Figure A.2. ^{13}C NMR spectra of **8**

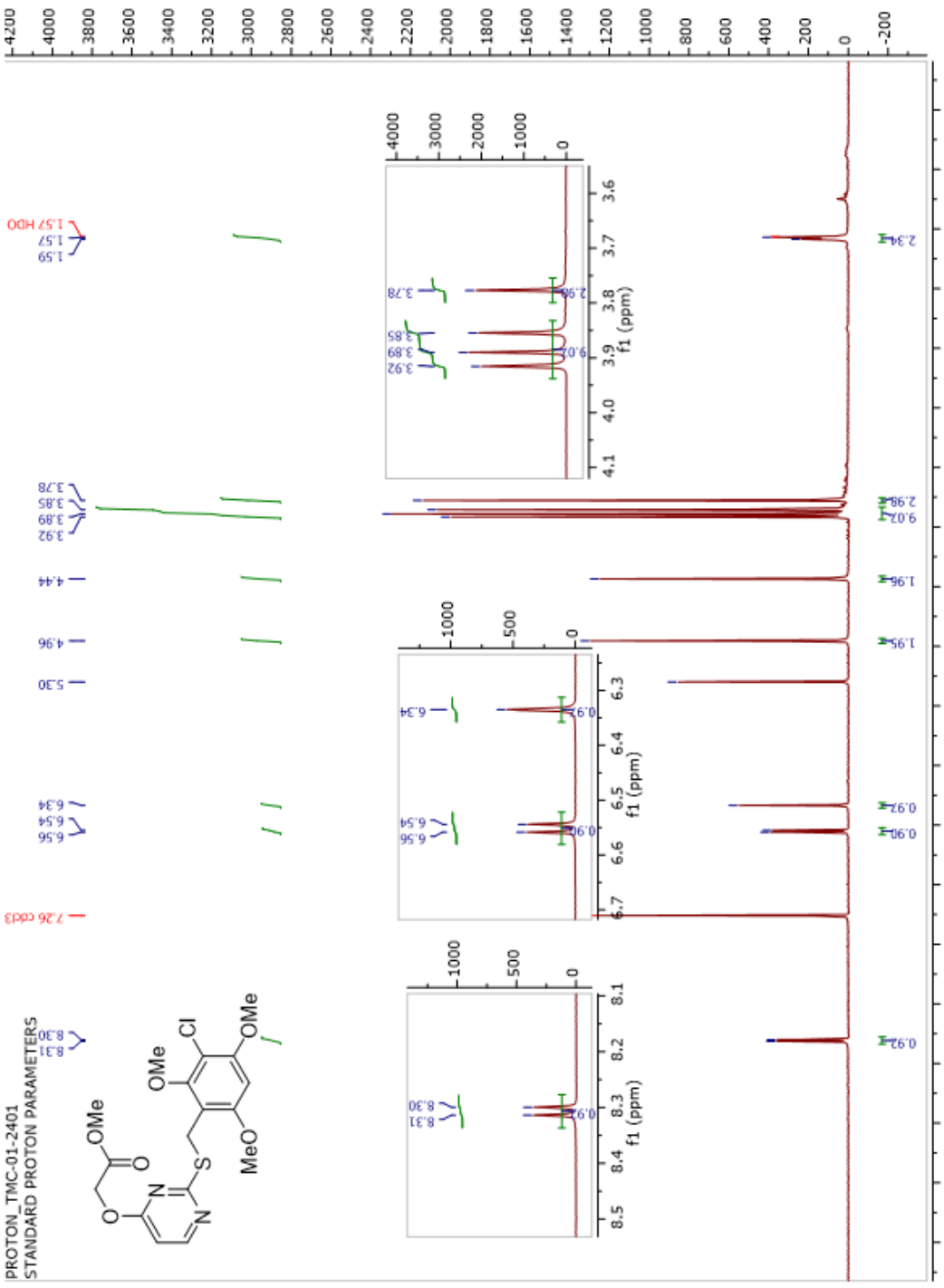


Figure A.3. ^{1}H NMR spectra of **11**

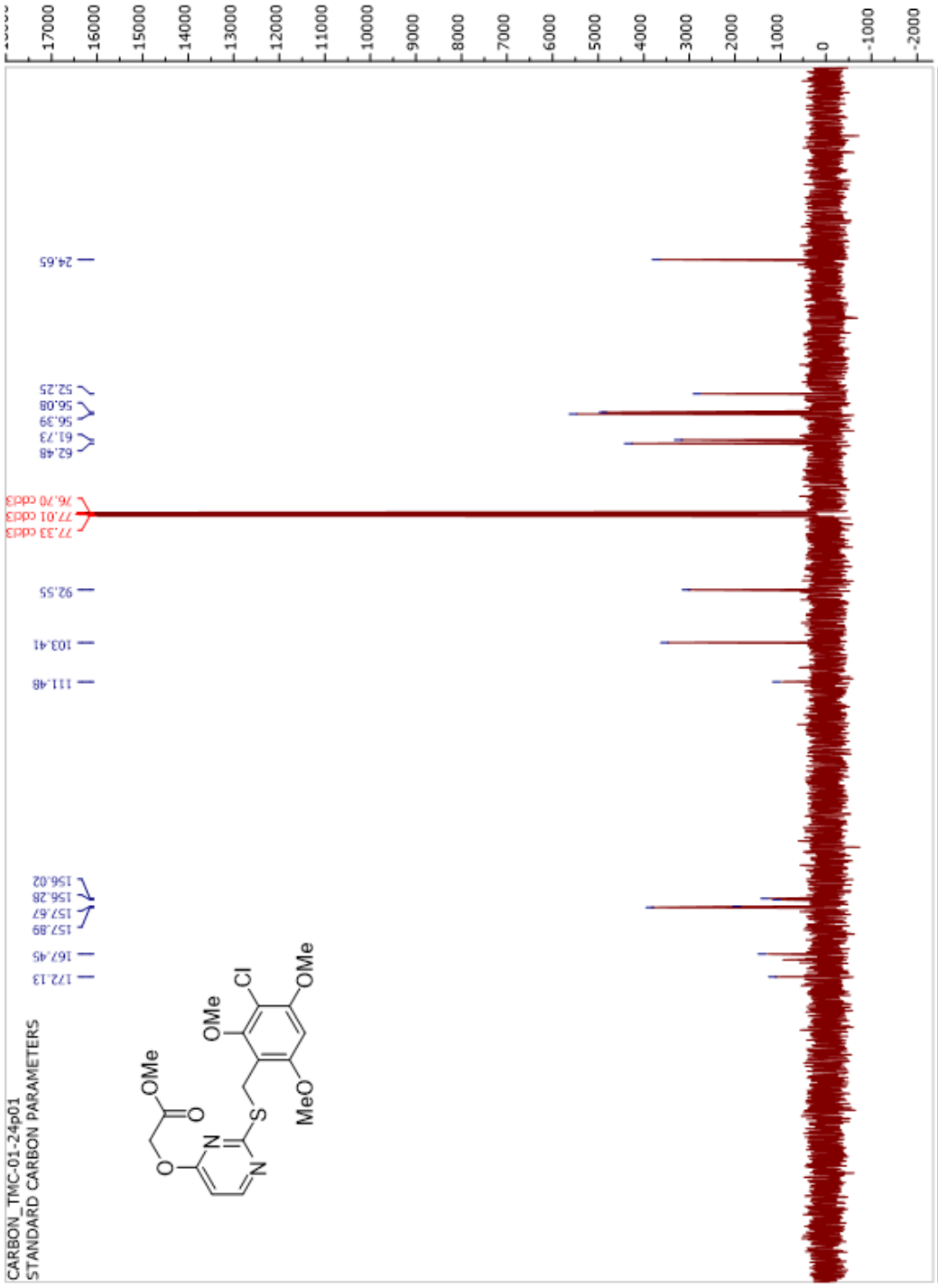


Figure A.4. ^{13}C NMR Spectra of **11**

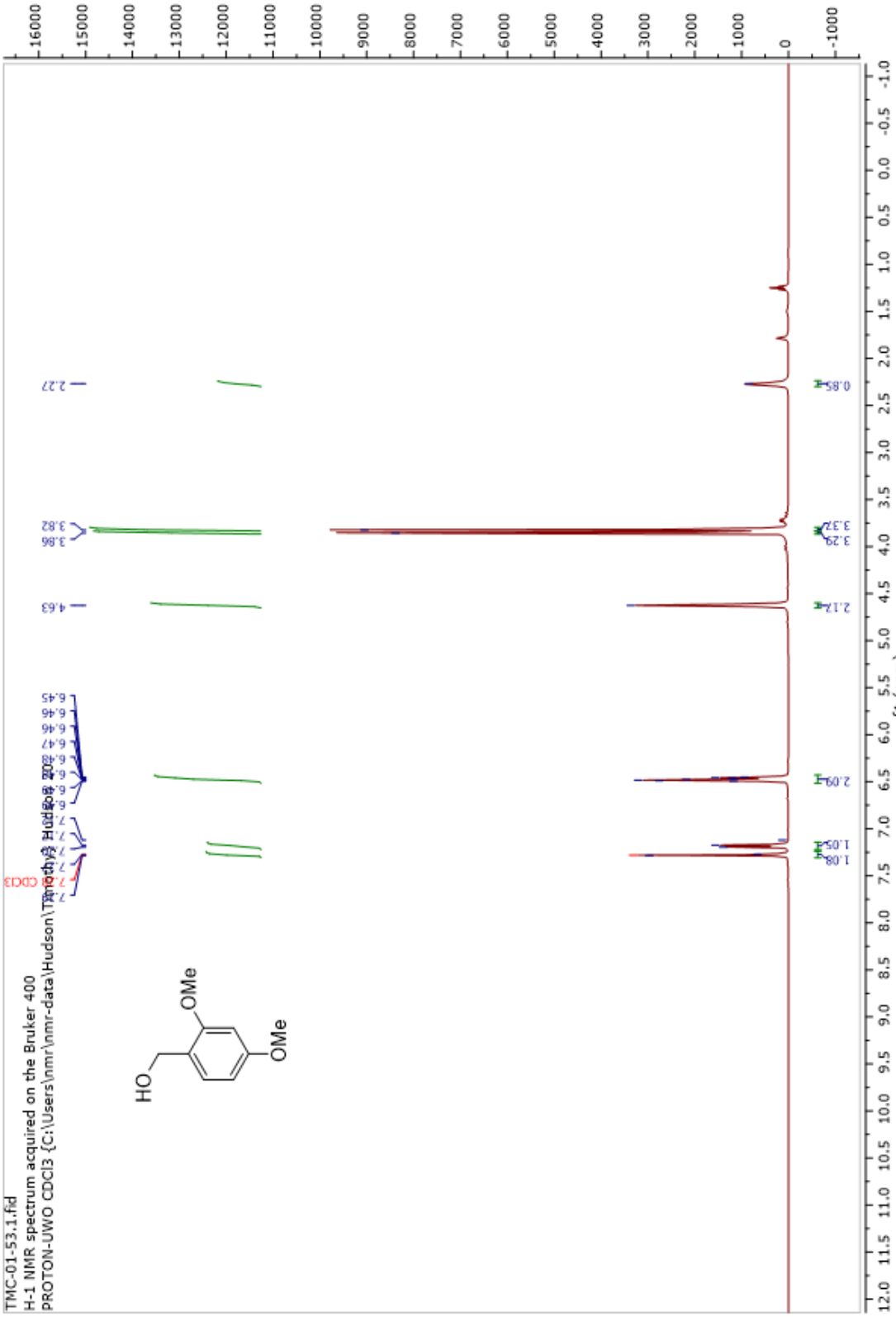


Figure A.5. ^{1}H NMR spectra of **13**

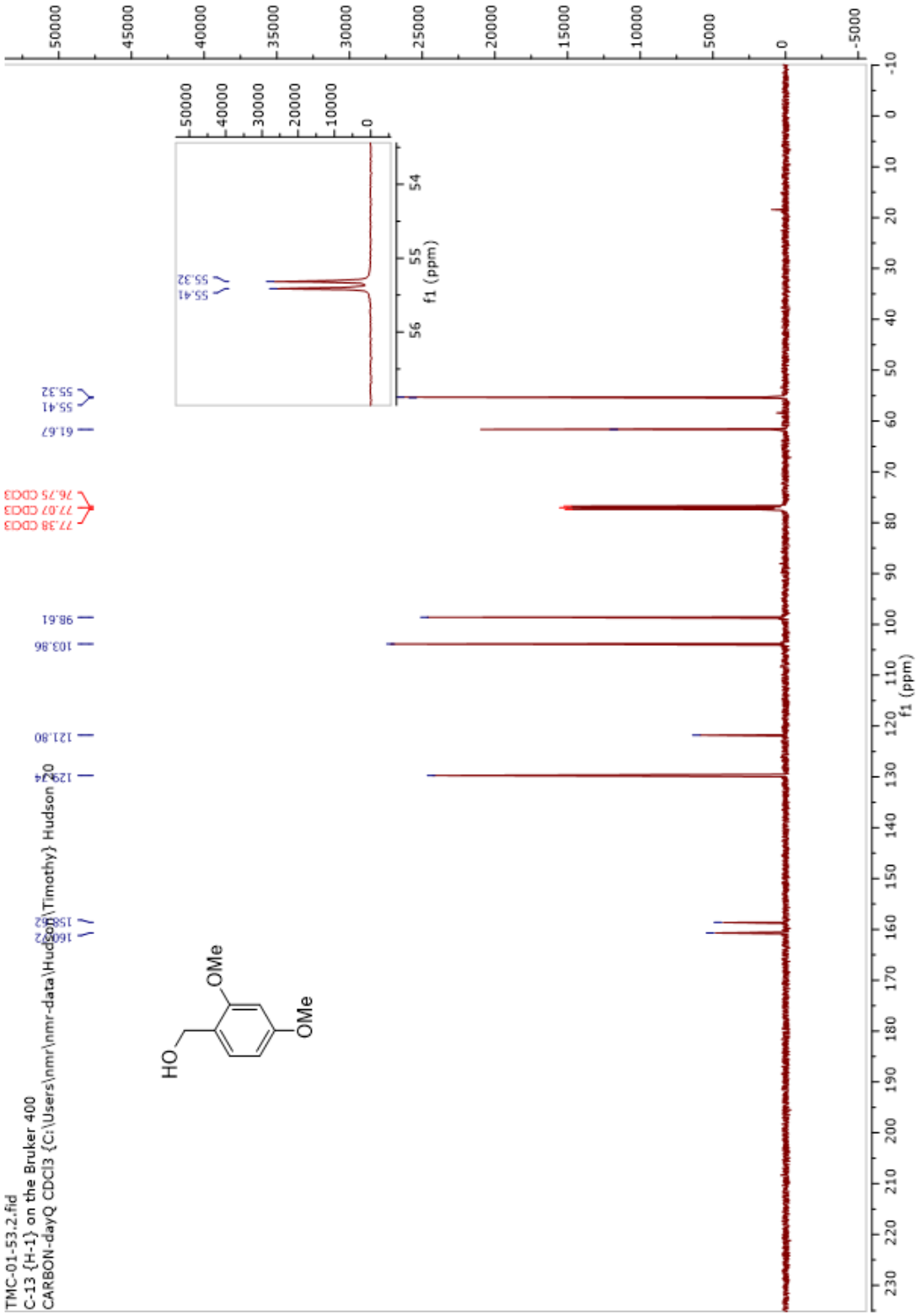


Figure A.6. ^{13}C NMR spectra of **13**

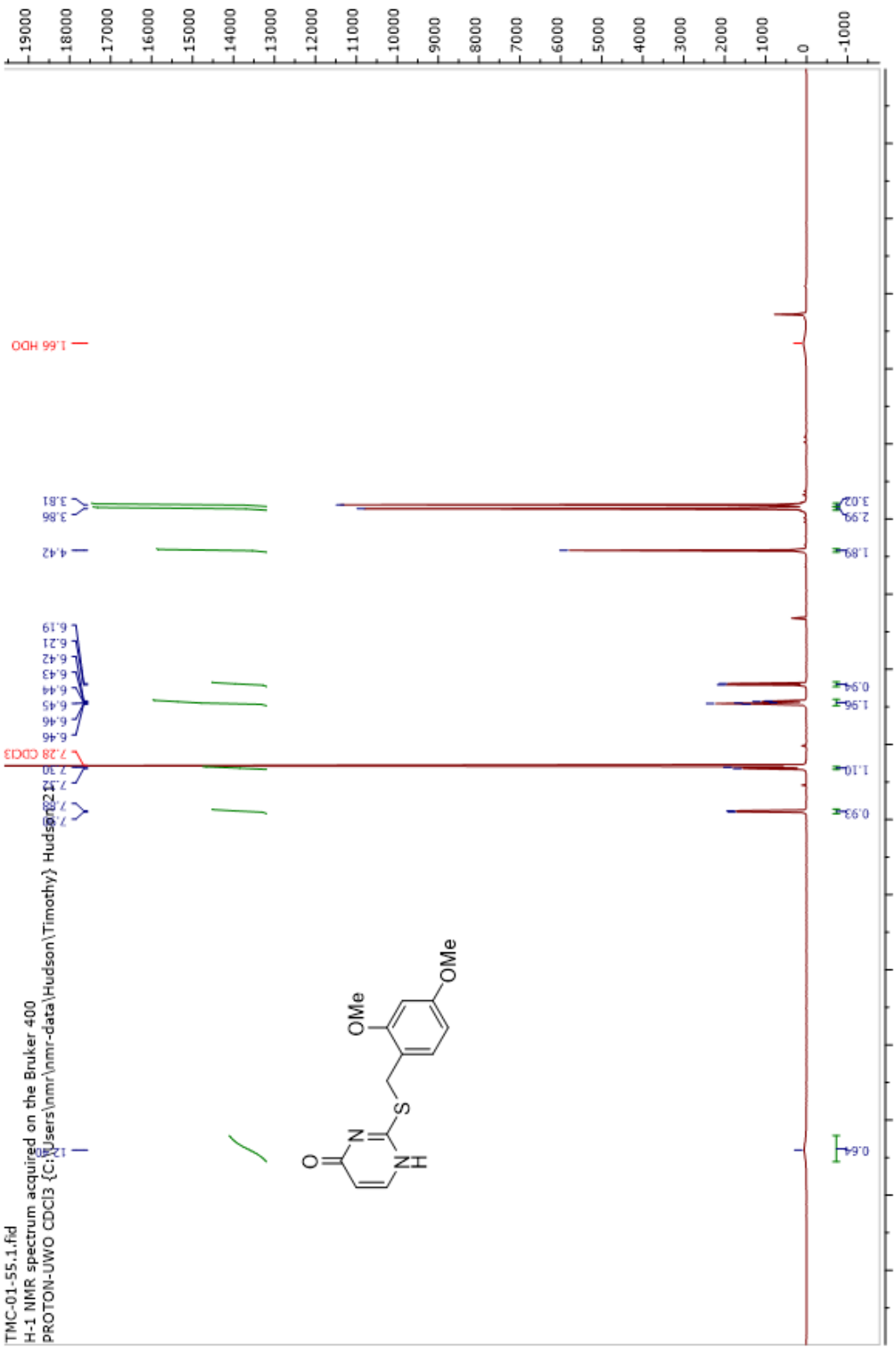


Figure A.7. ^{1}H NMR spectra of **15**

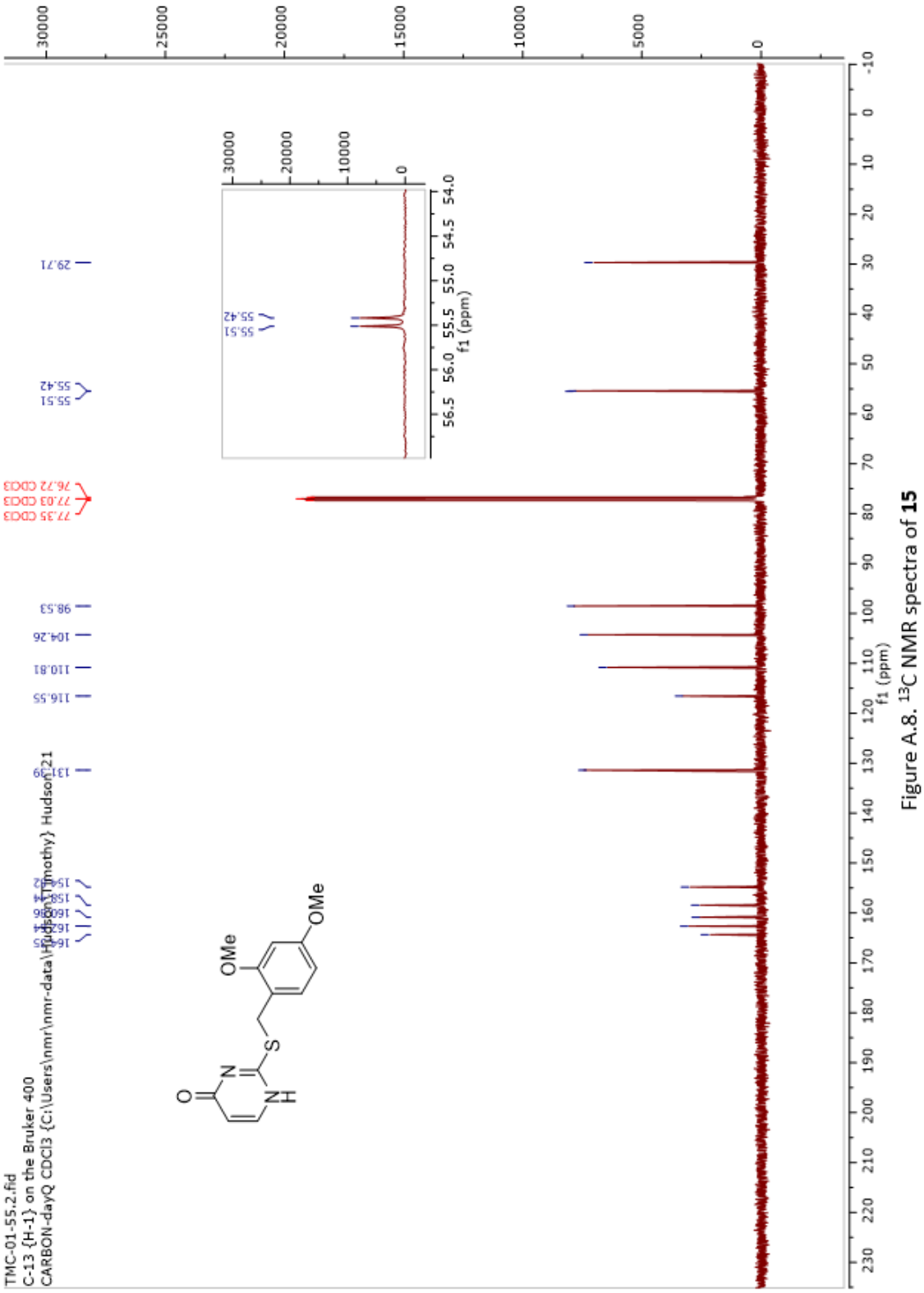


Figure A.8. ^{13}C NMR spectra of **15**

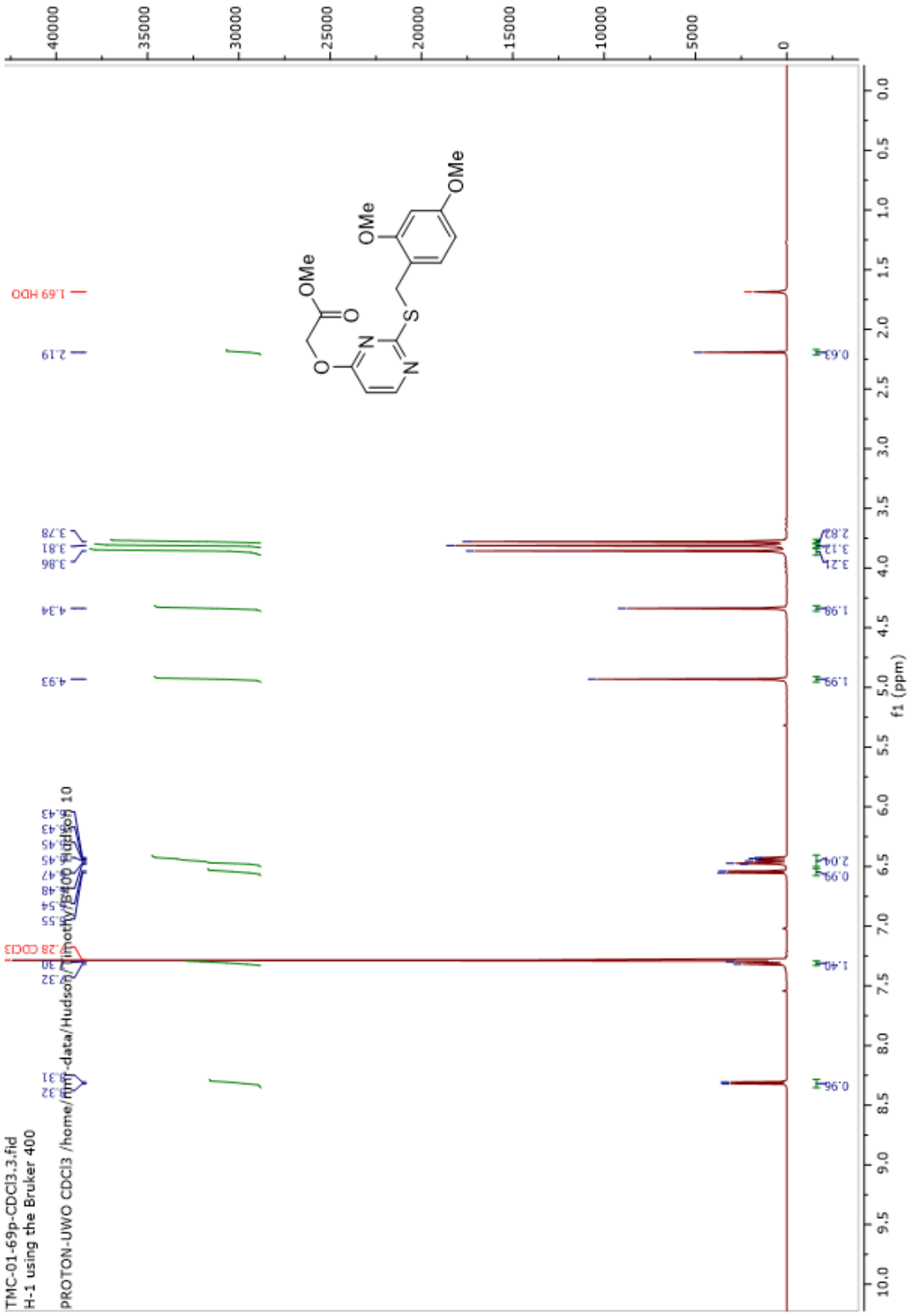


Figure A.9. ^{1}H NMR spectra of **16b**

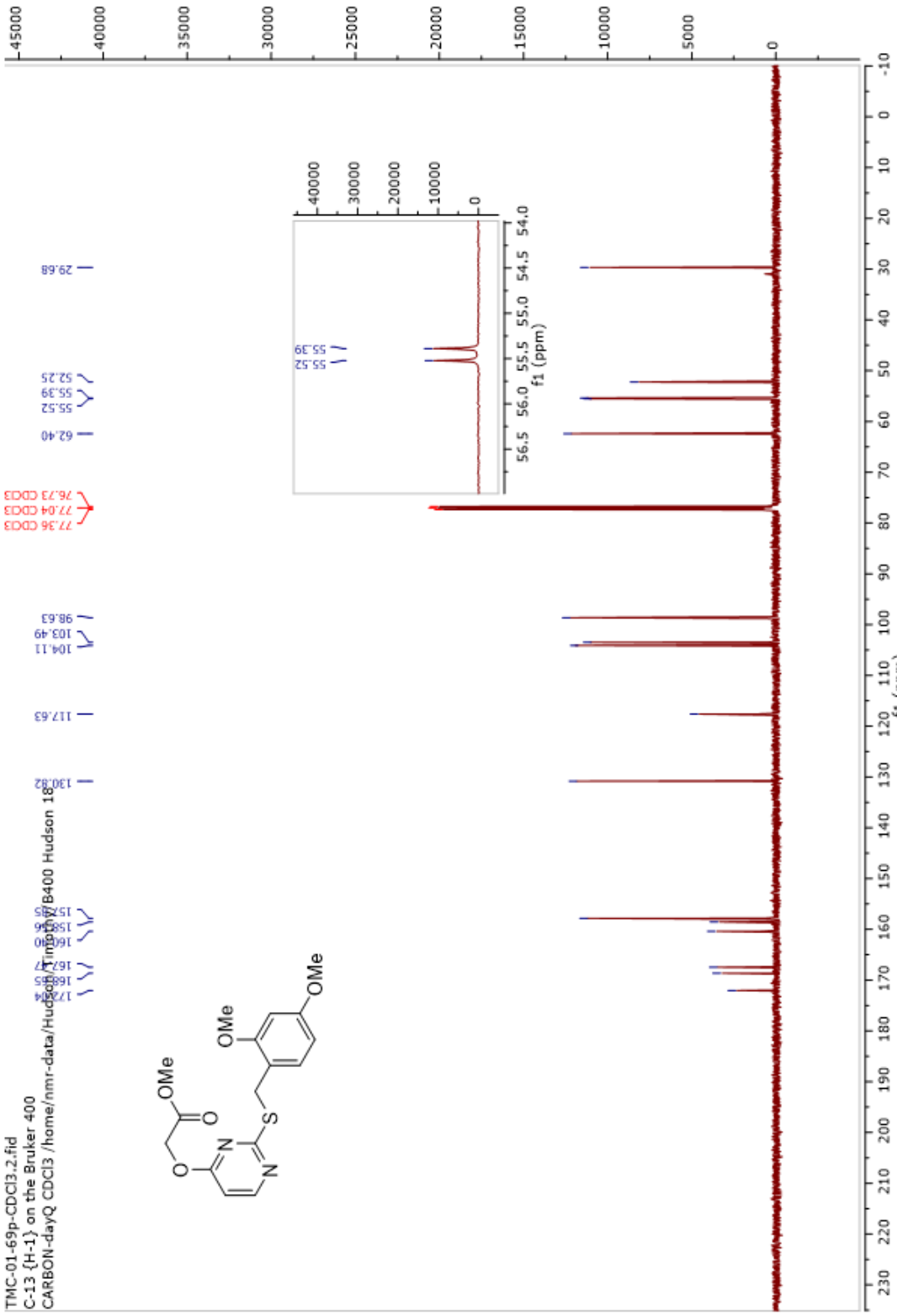


Figure A.10. ^{13}C NMR spectra of **16b**

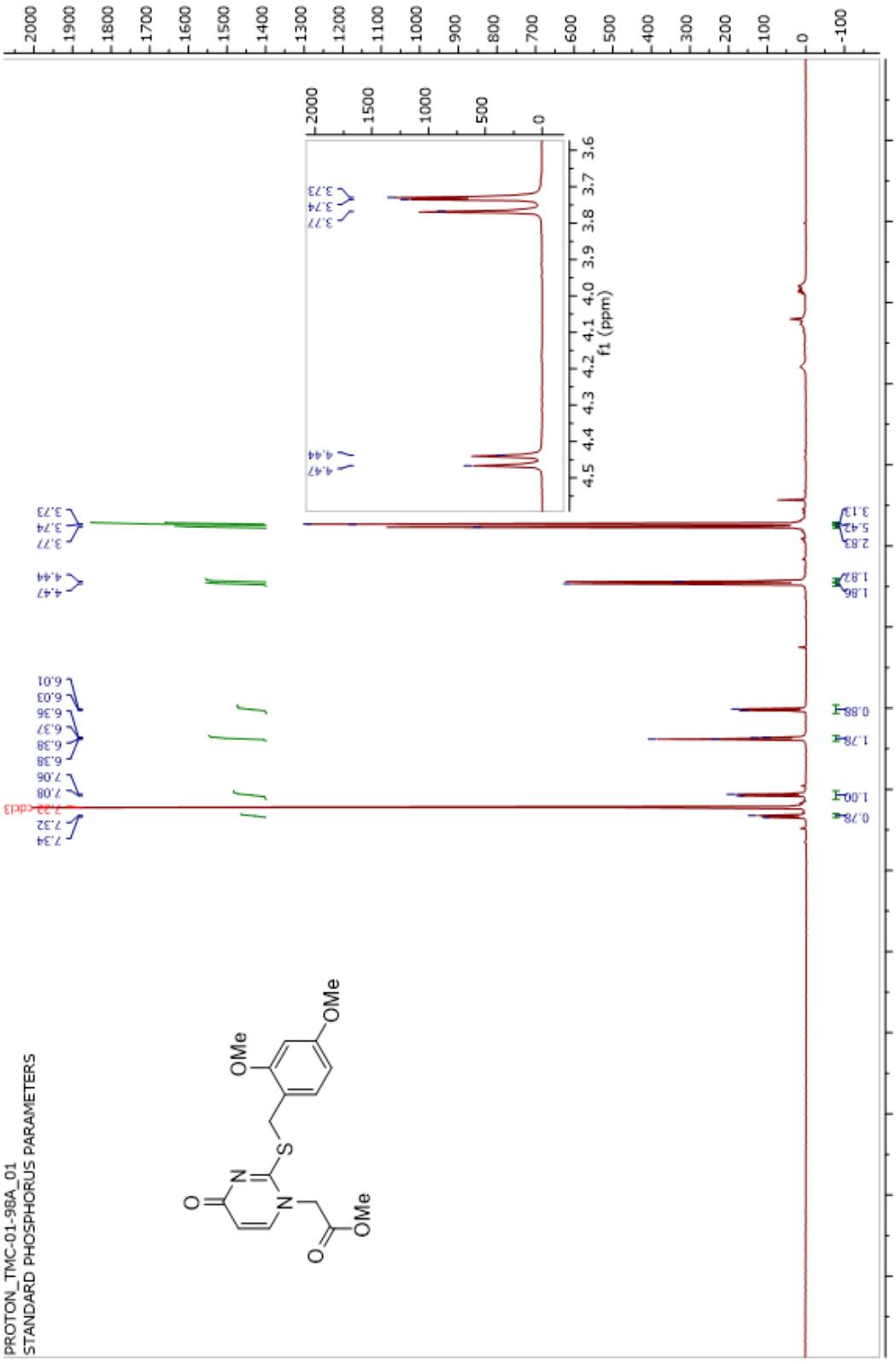


Figure A.11. ^{1}H NMR spectra of **16a**

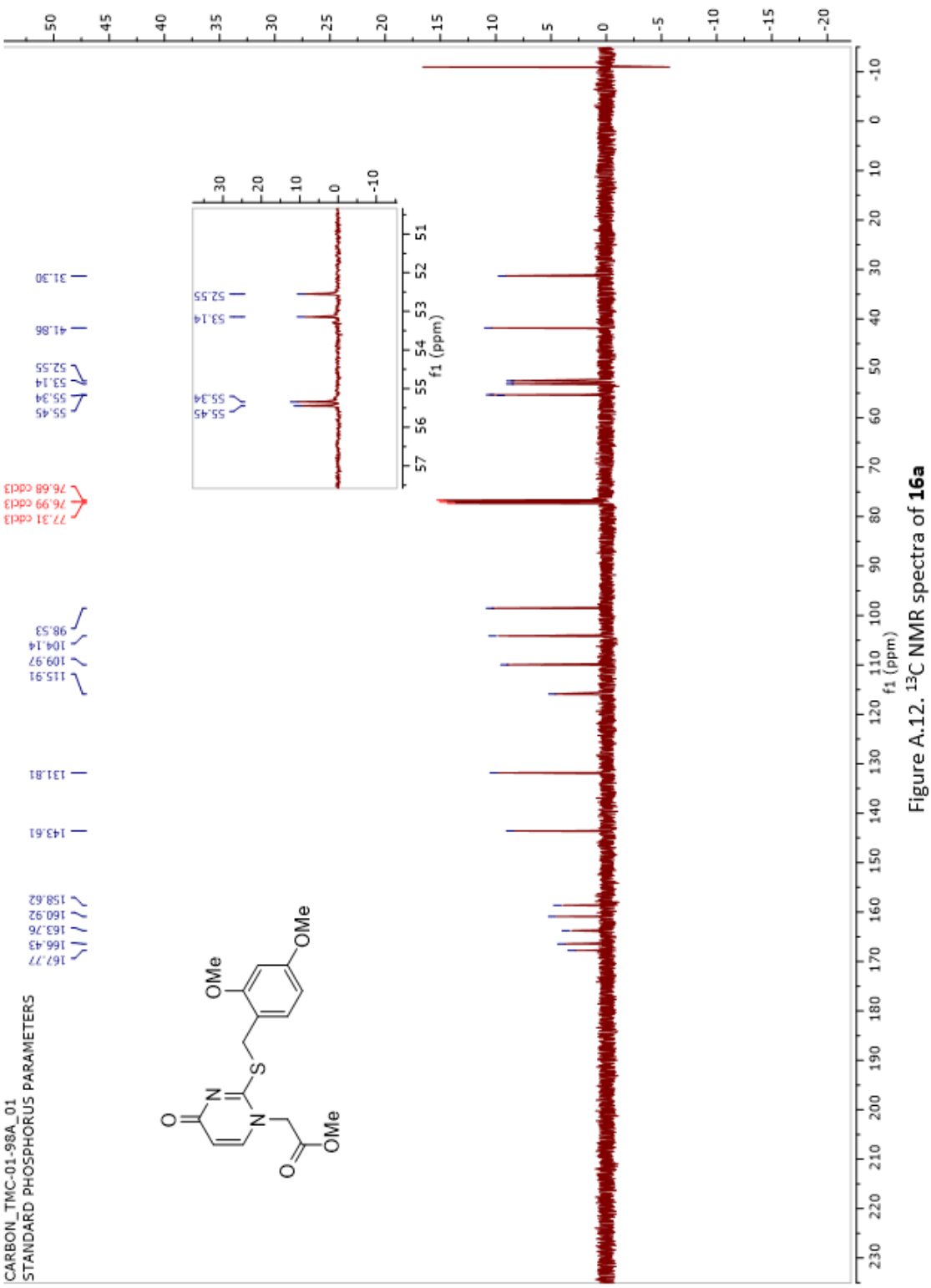


Figure A.12. ^{13}C NMR spectra of **16a**

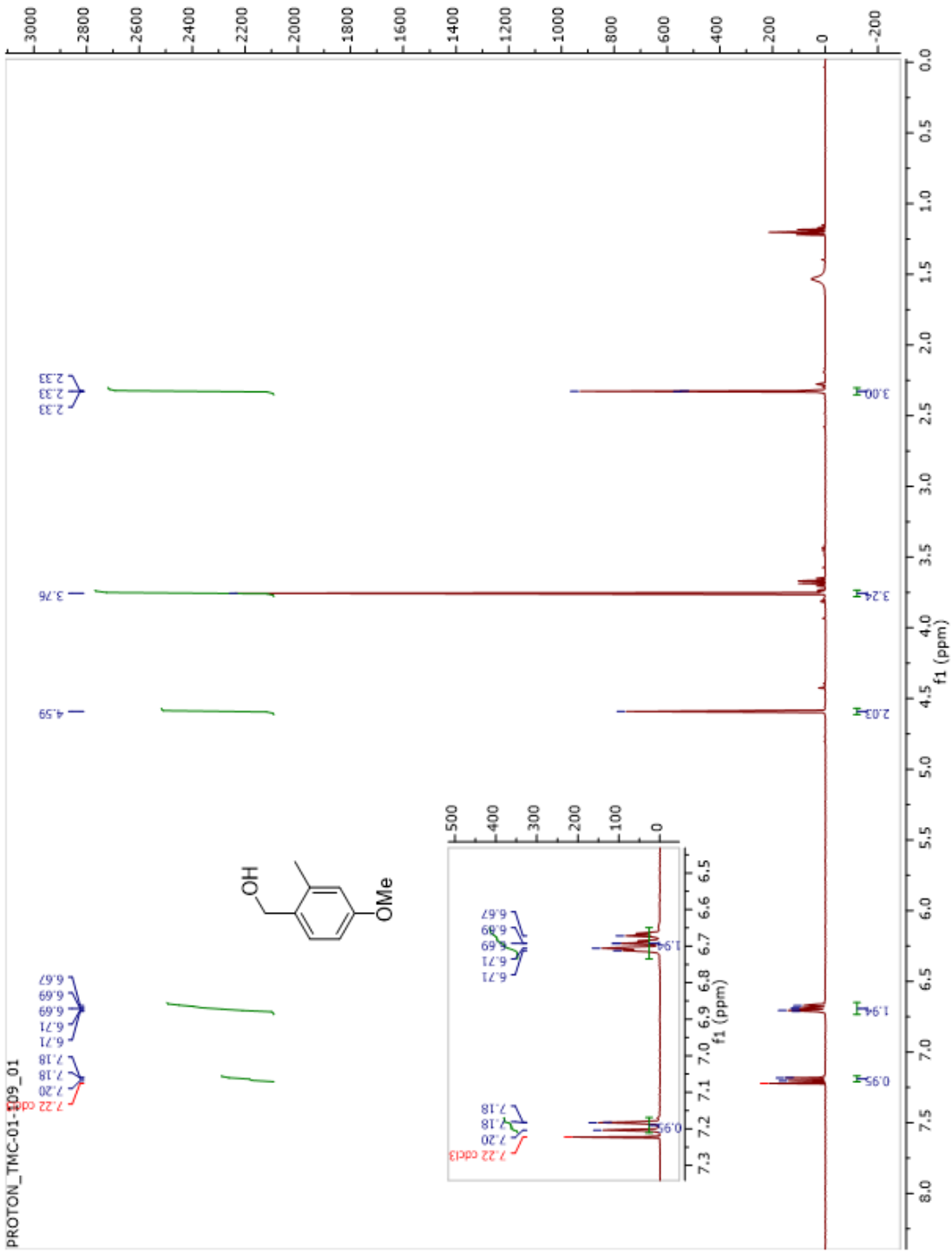


Figure A.13. ^{1}H NMR spectra of **18**

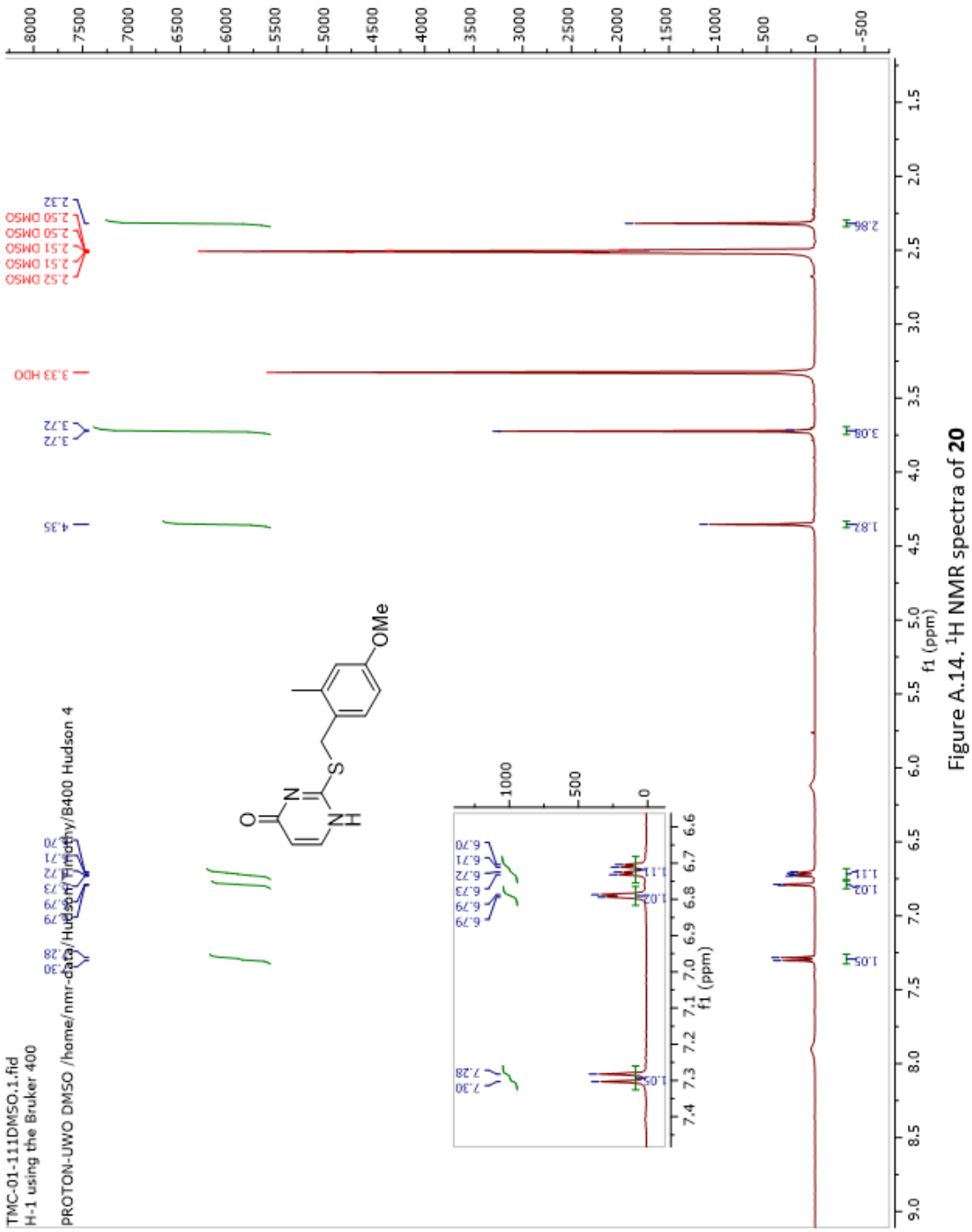


Figure A.14. ^{1}H NMR spectra of **20**

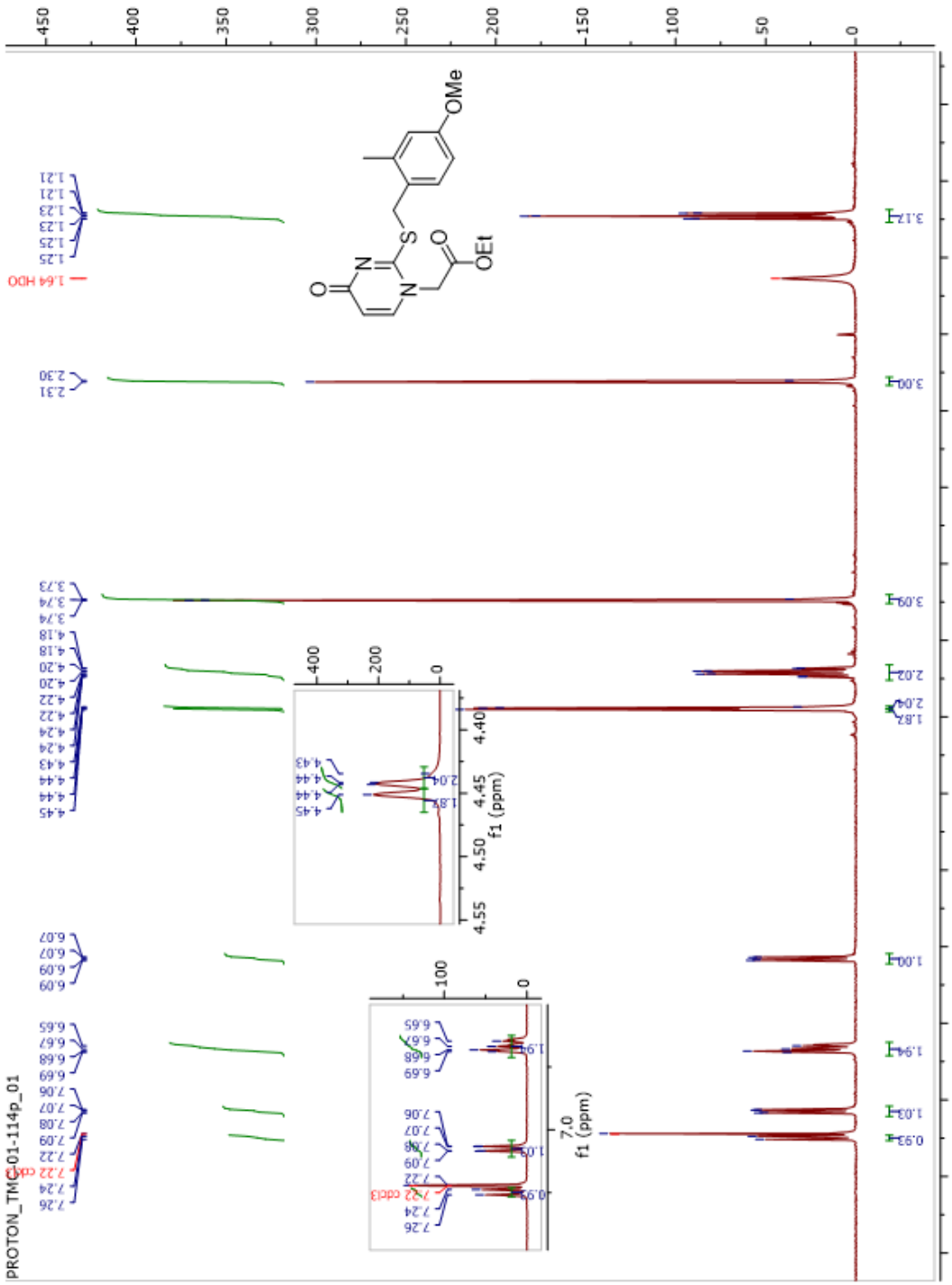


Figure A.15. ^{1}H NMR spectra of **21**

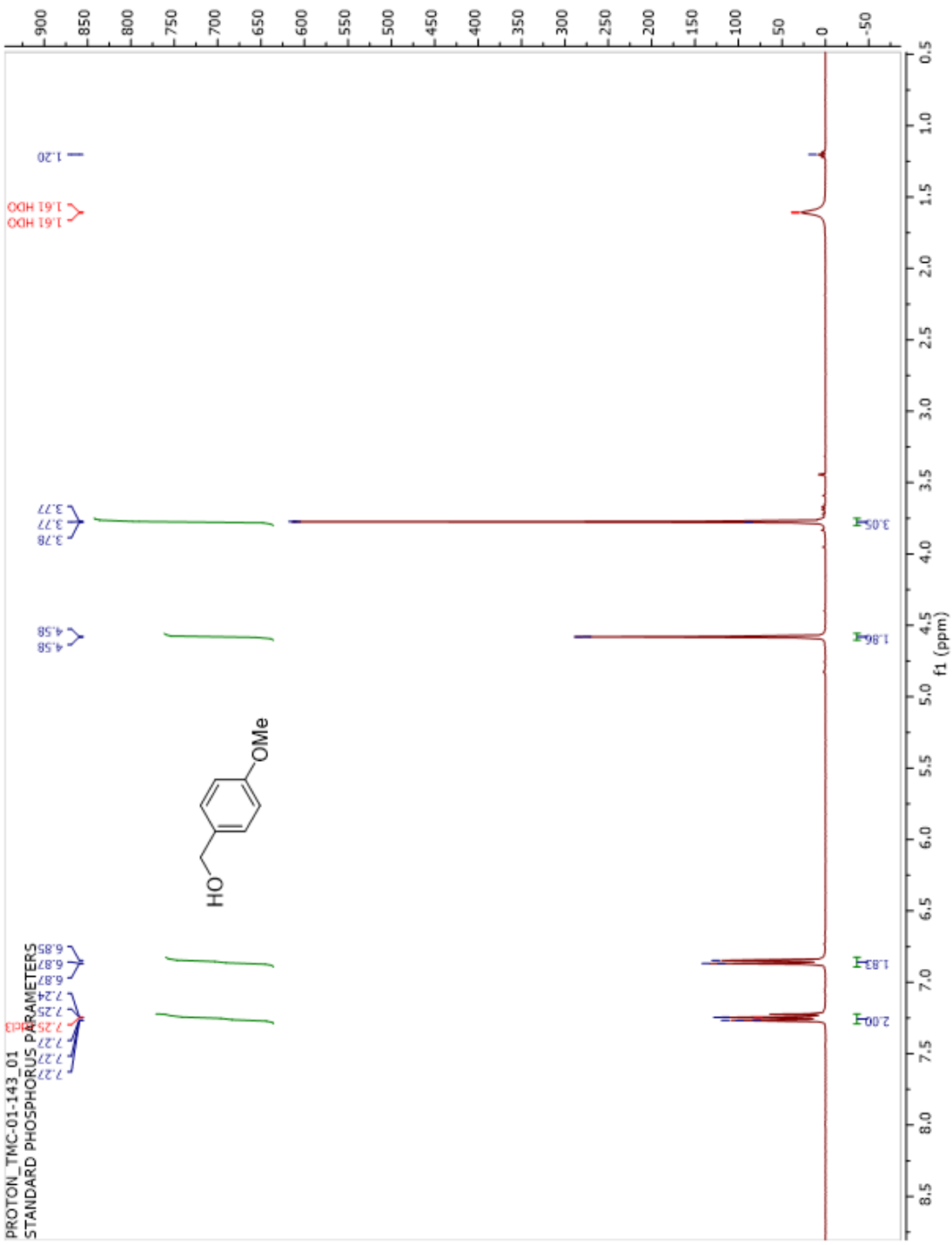


Figure A.16. ^{1}H NMR spectra of **23**

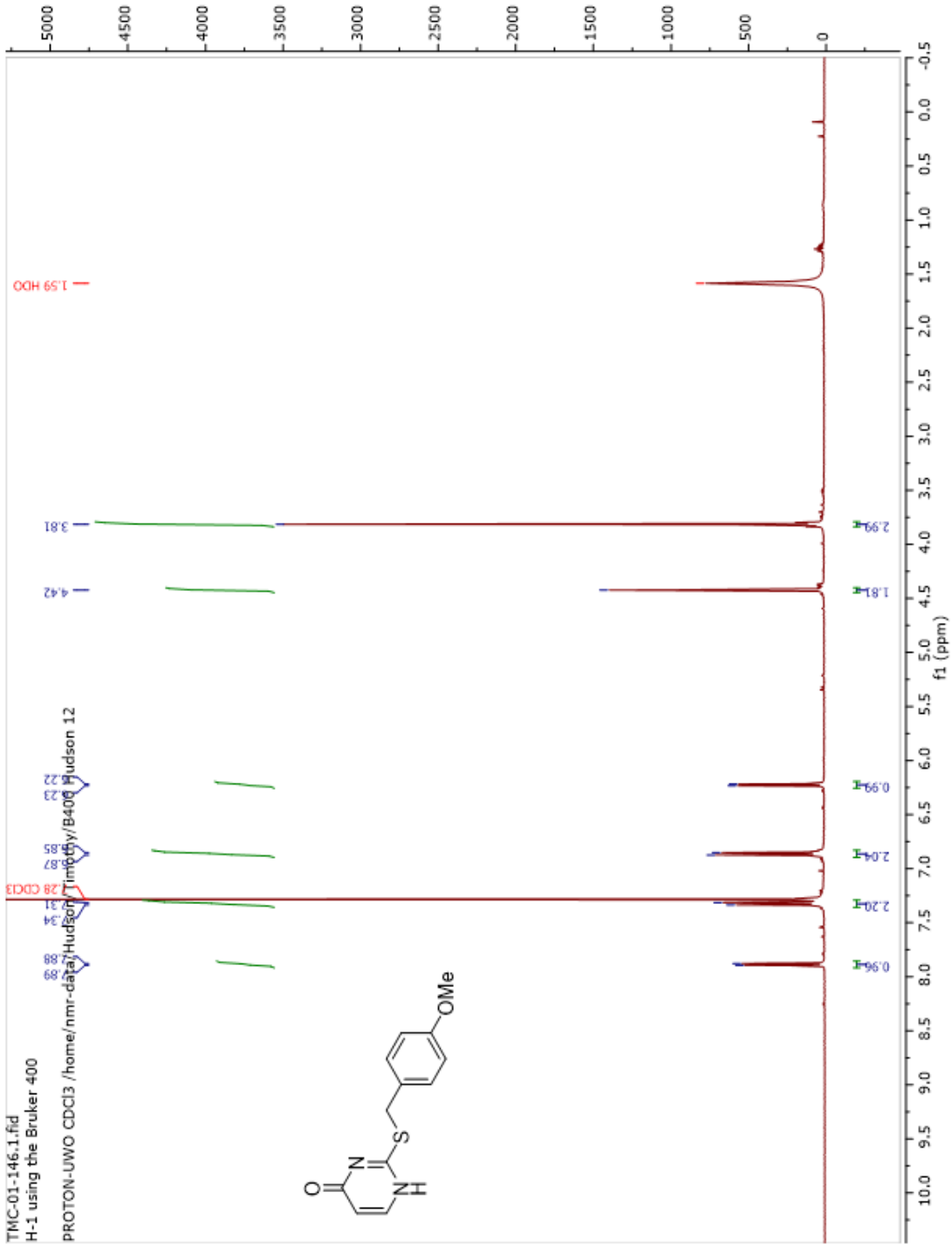


Figure A.17. ^{1}H NMR spectra of **2**

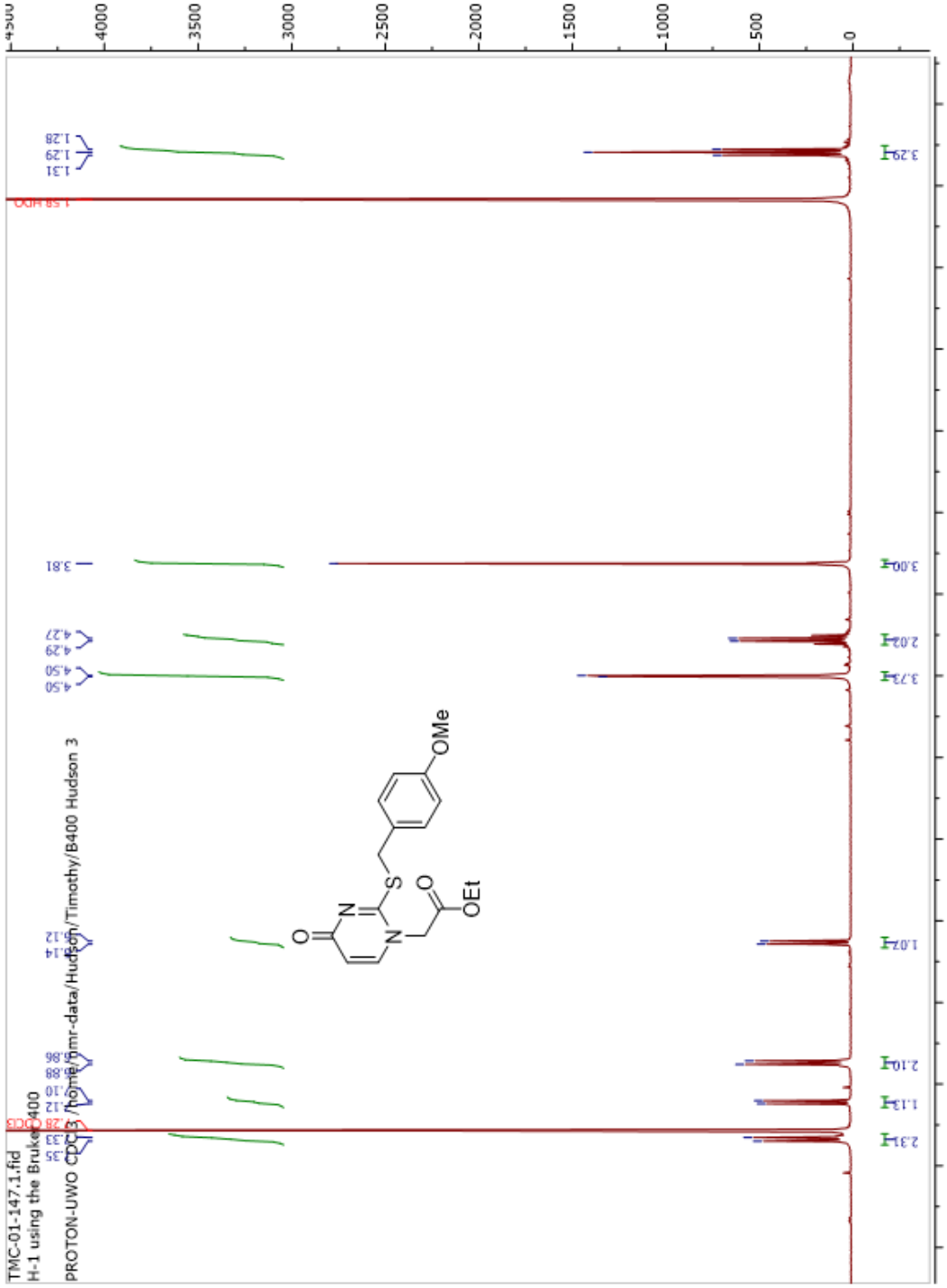


Figure A.18. ^{1}H NMR spectra of **3**

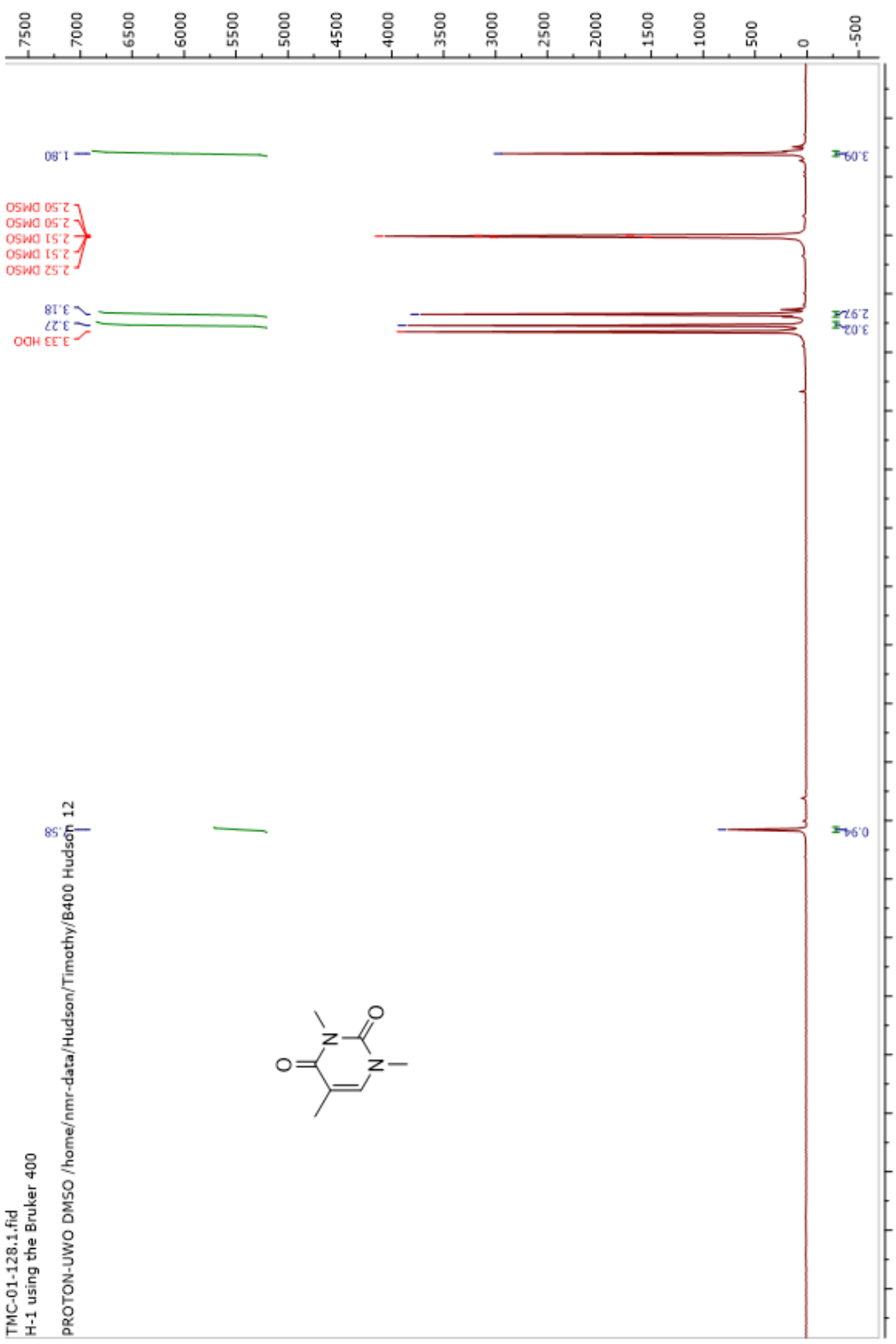


Figure A.19. ^{1}H NMR spectra of **25**

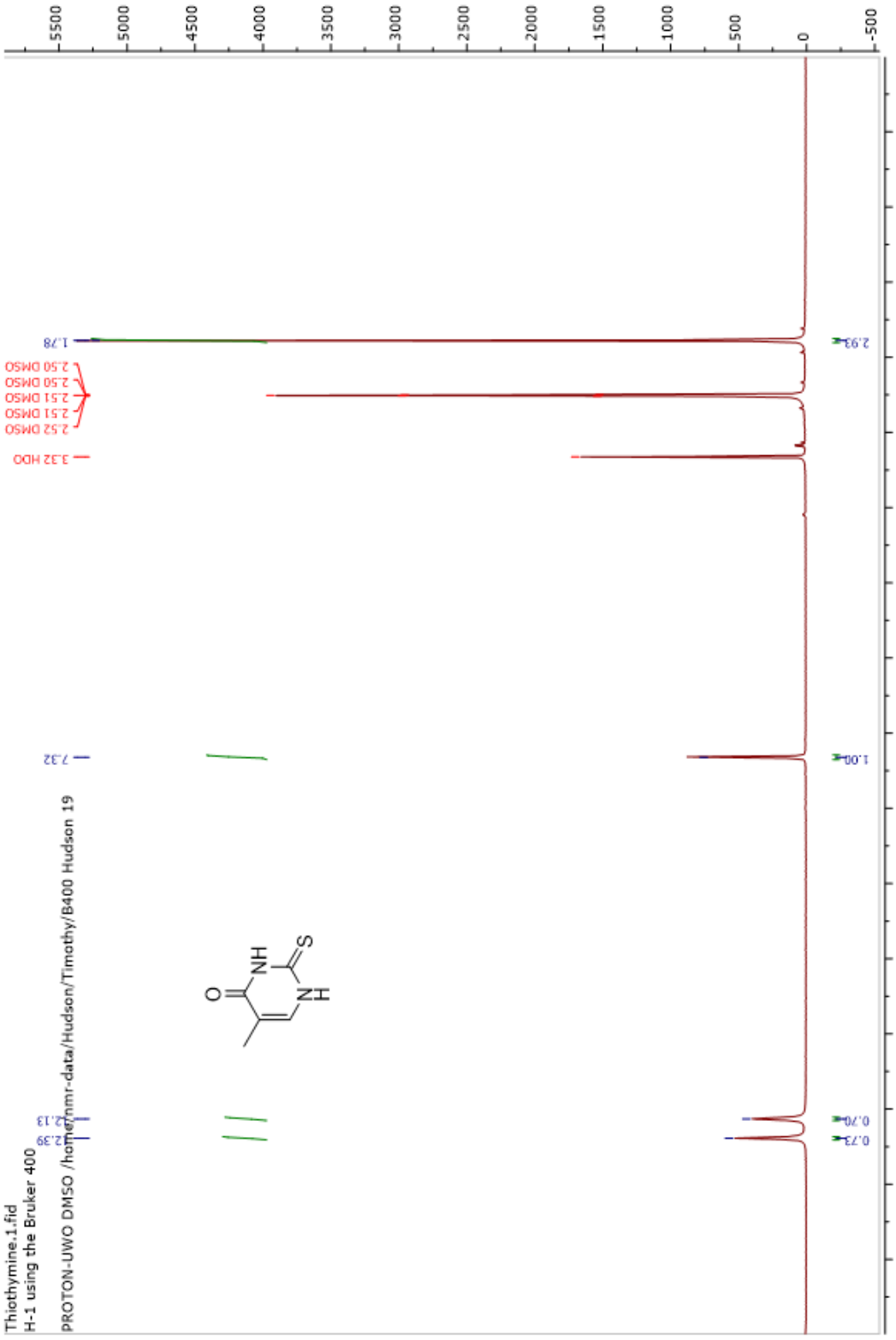


Figure A.20. ^{1}H NMR spectra of **27**

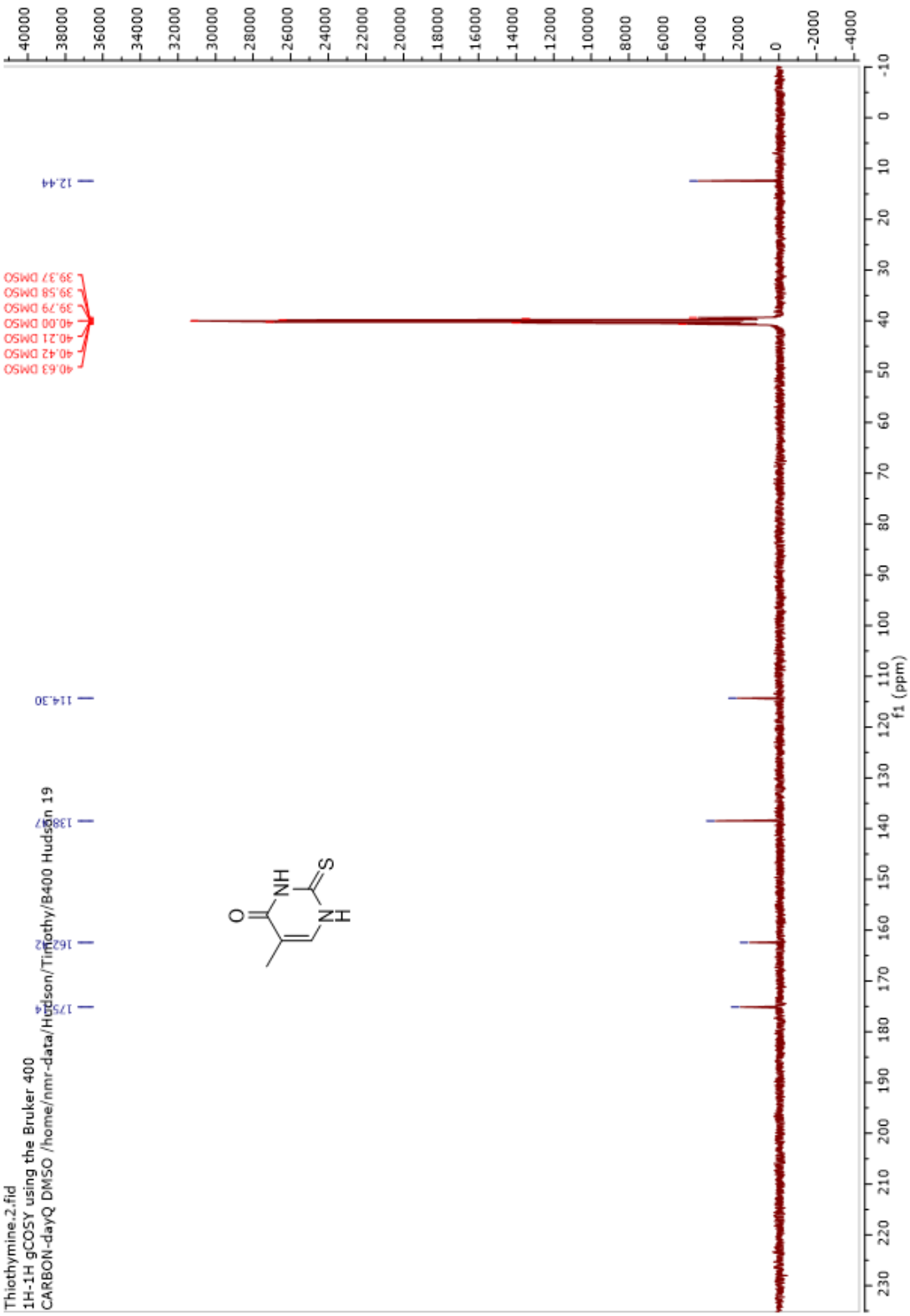


Figure A.21. ^{13}C NMR spectra of **27**

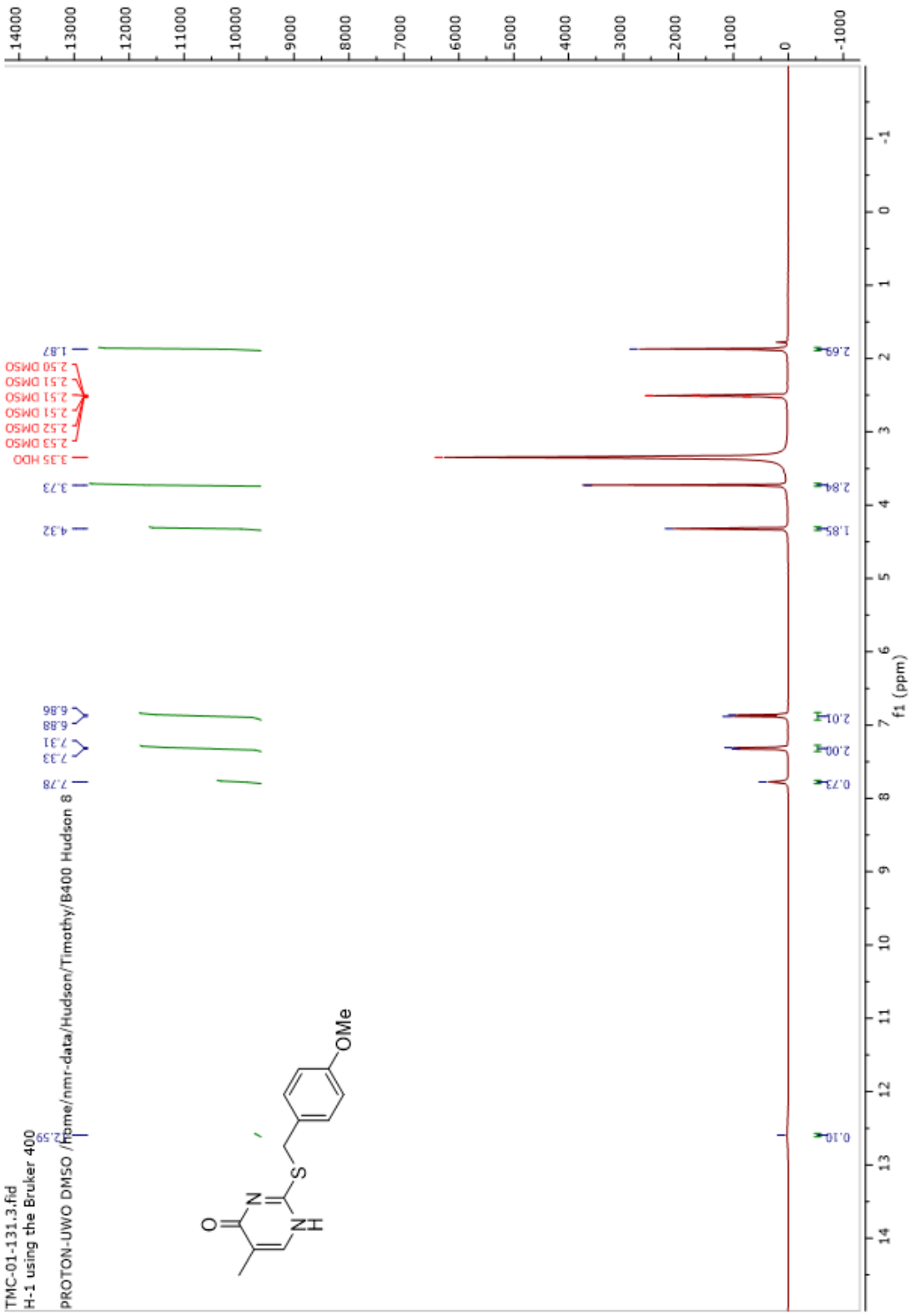


Figure A.22. ^{1}H NMR spectra of **31**

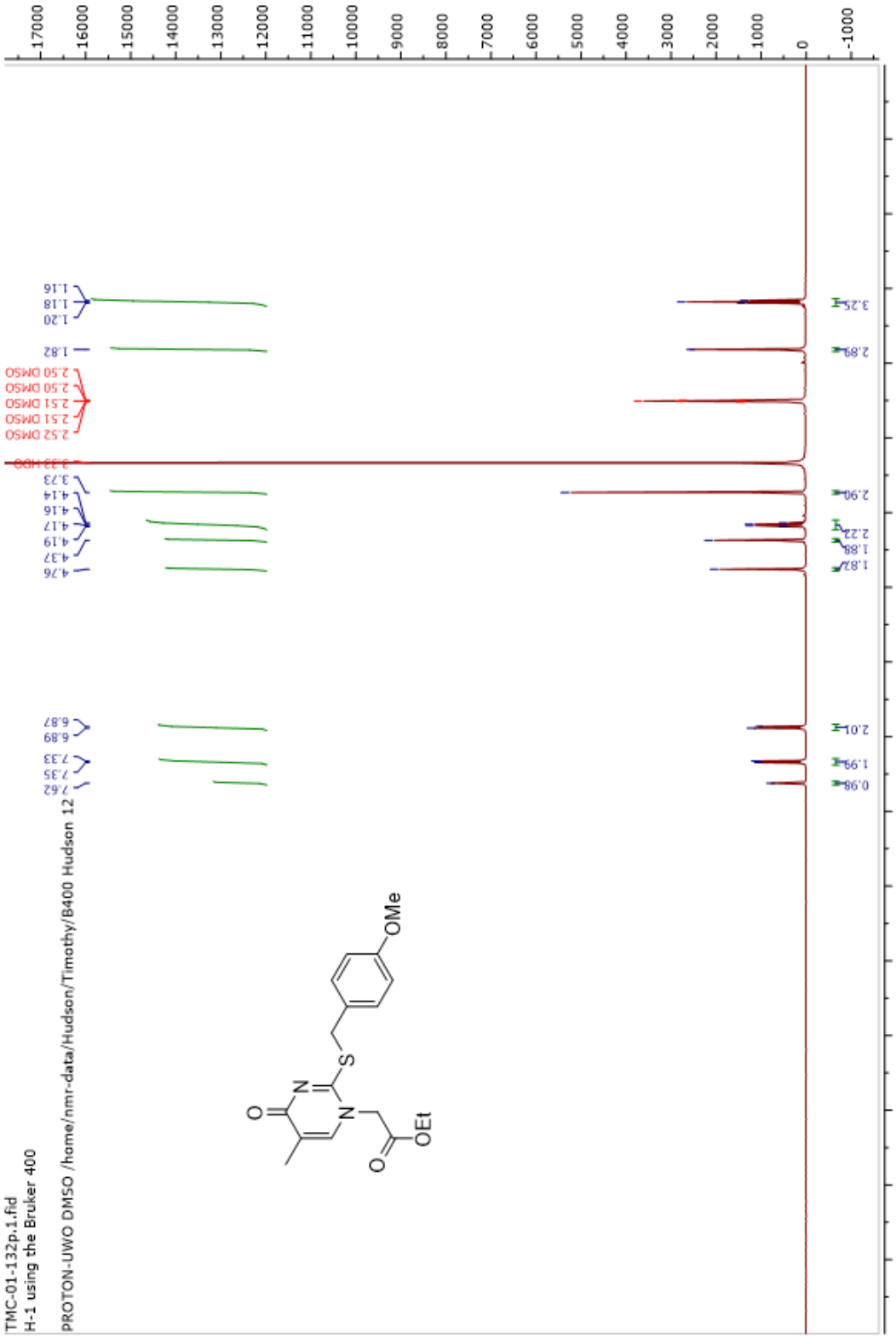


Figure A.23. ^{1}H NMR spectra of **32**

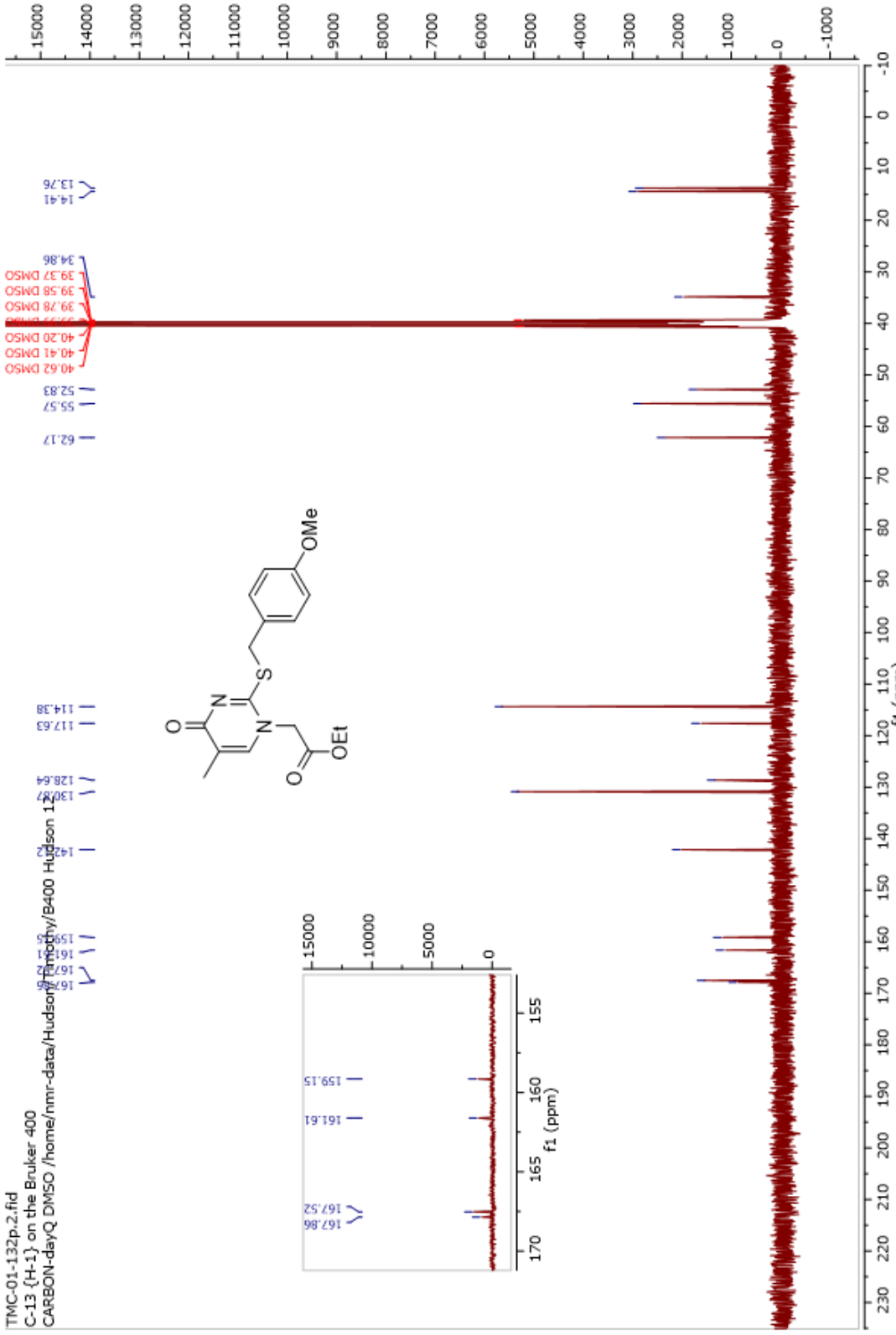


Figure A.24. ^{13}C NMR spectra of **23**

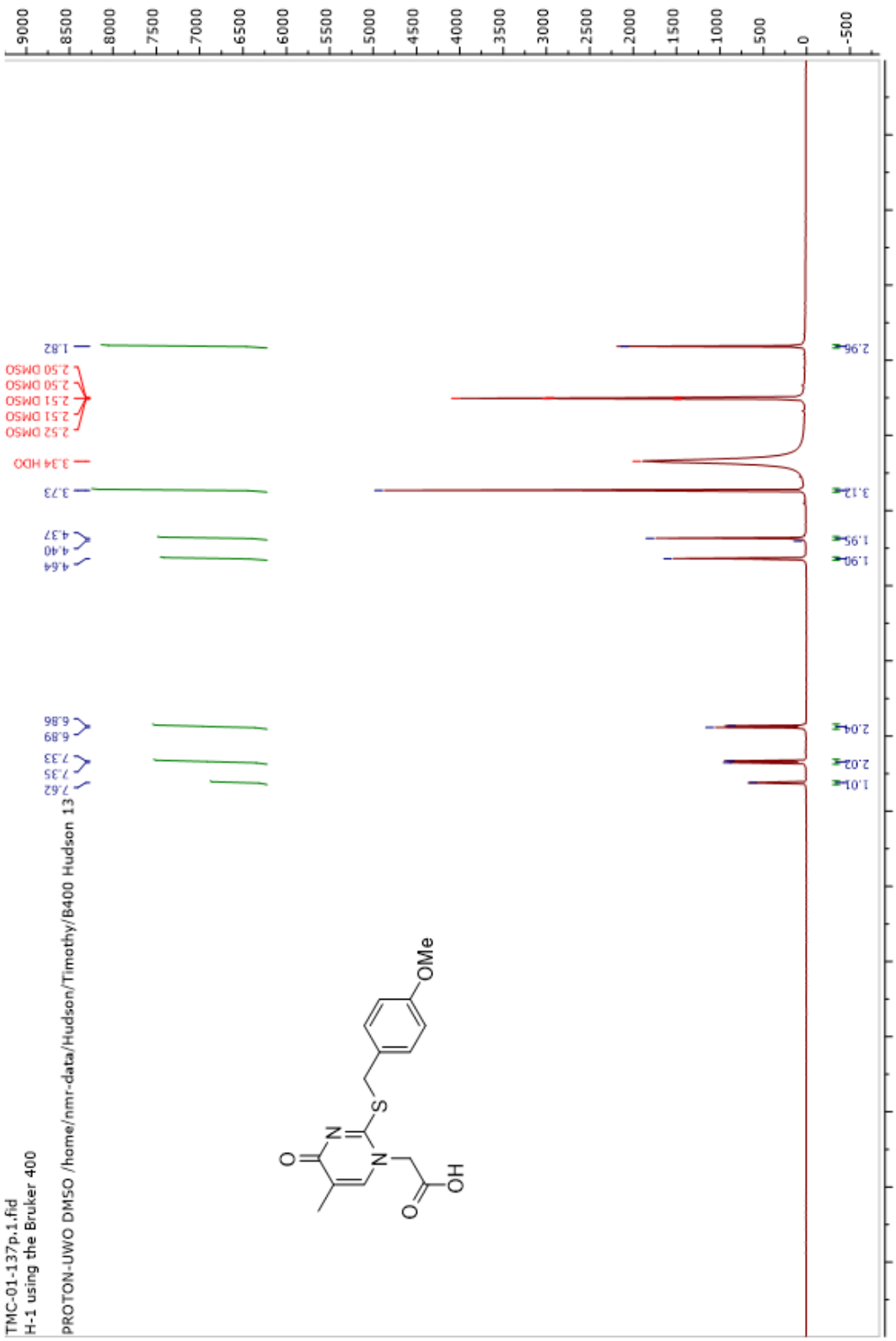


Figure A.25. ^{1}H NMR spectra of **33**

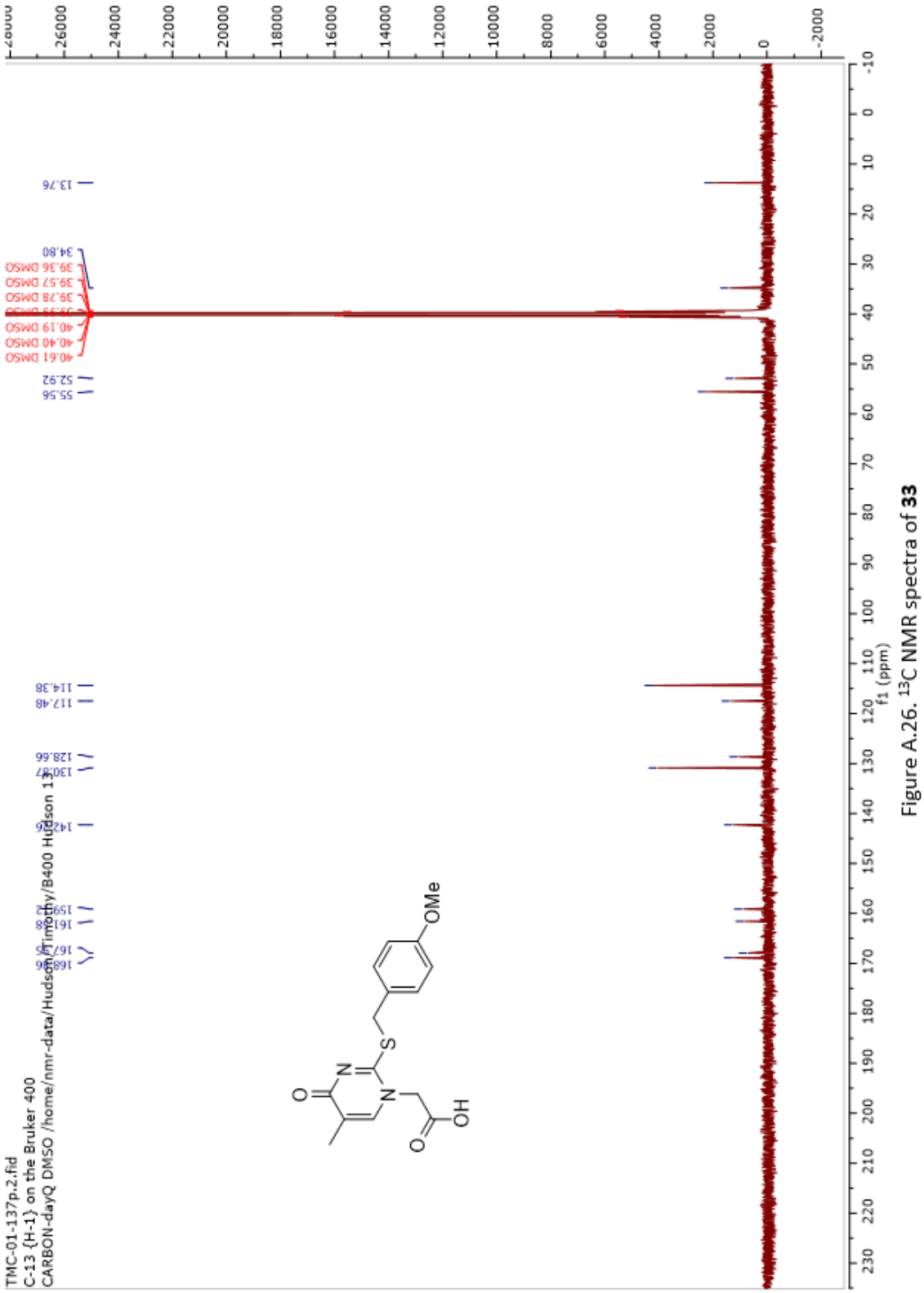


Figure A.26. ^{13}C NMR spectra of **33**

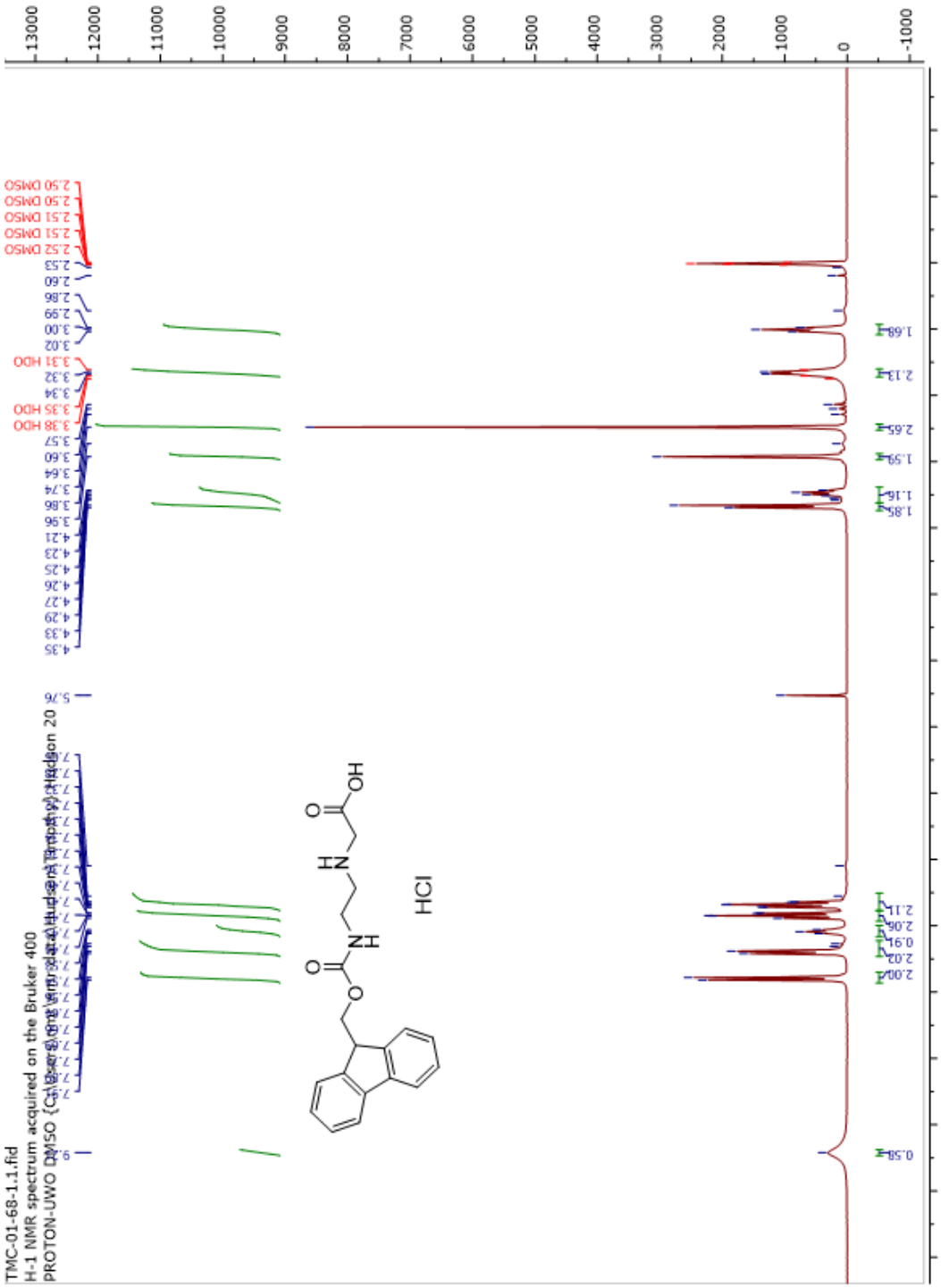


Figure A.27. ^{1}H NMR spectra of **35**

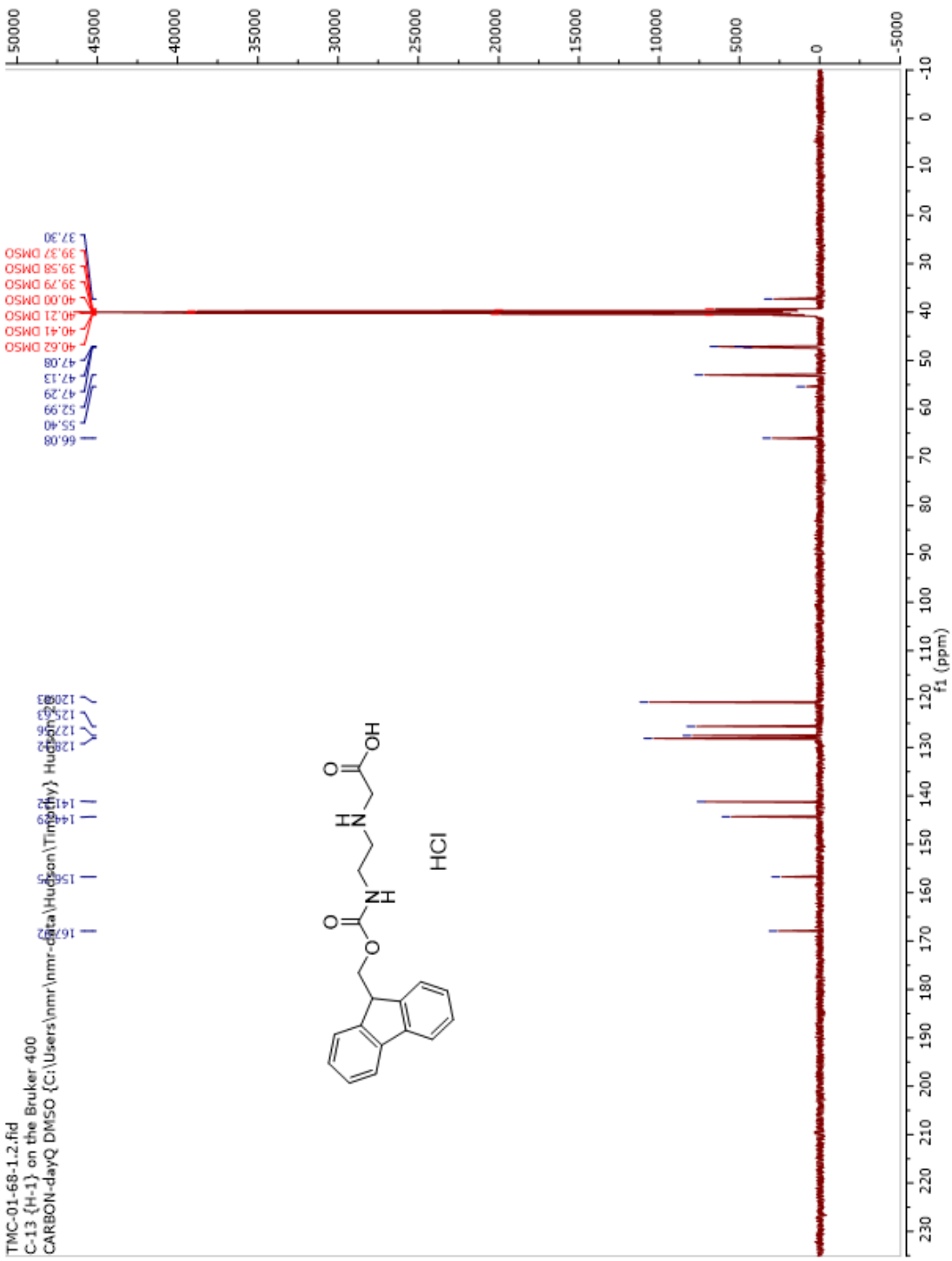


Figure A.28. ^{1}H NMR spectra of **35**

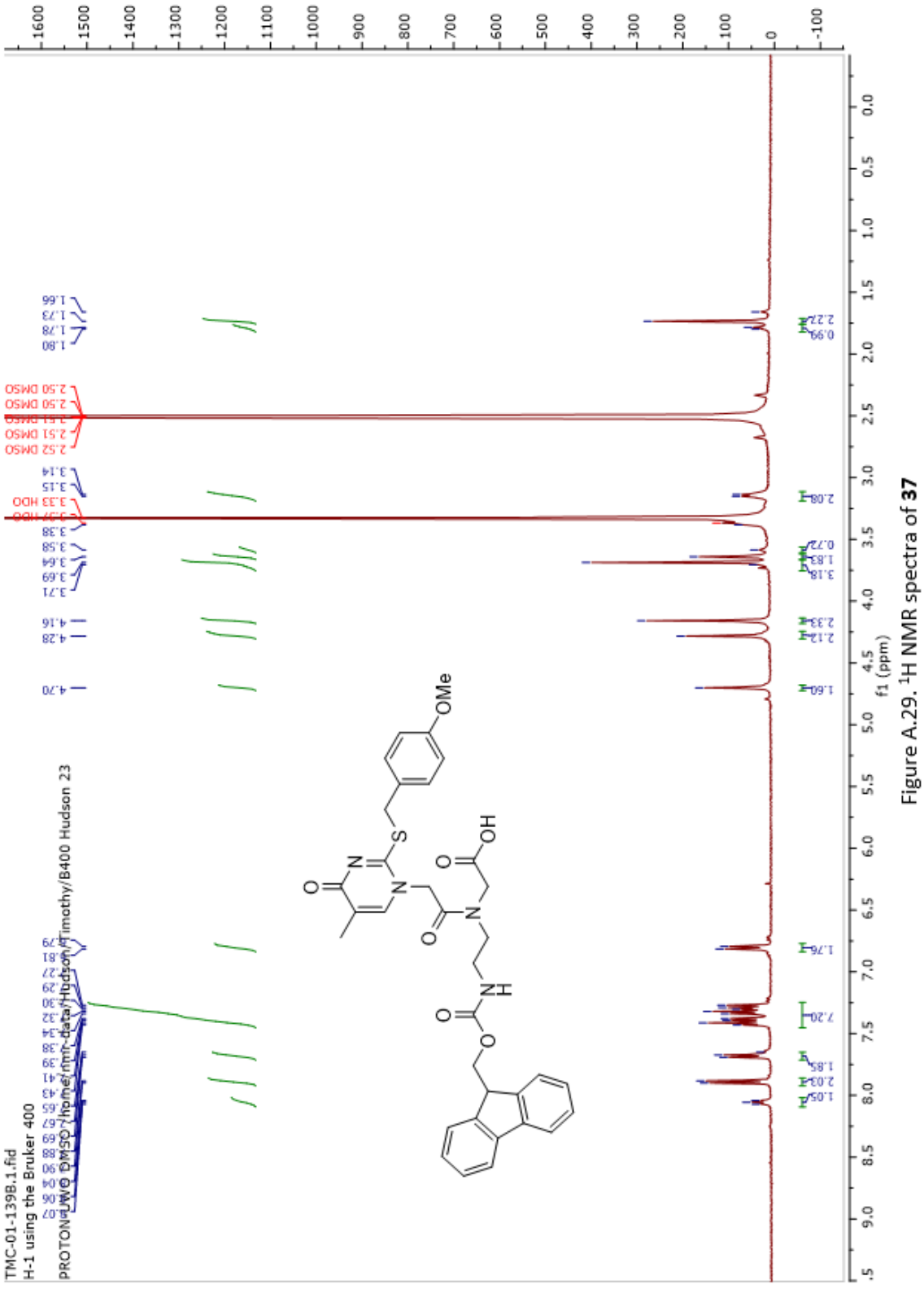


Figure A.29. ^{1}H NMR spectra of **37**

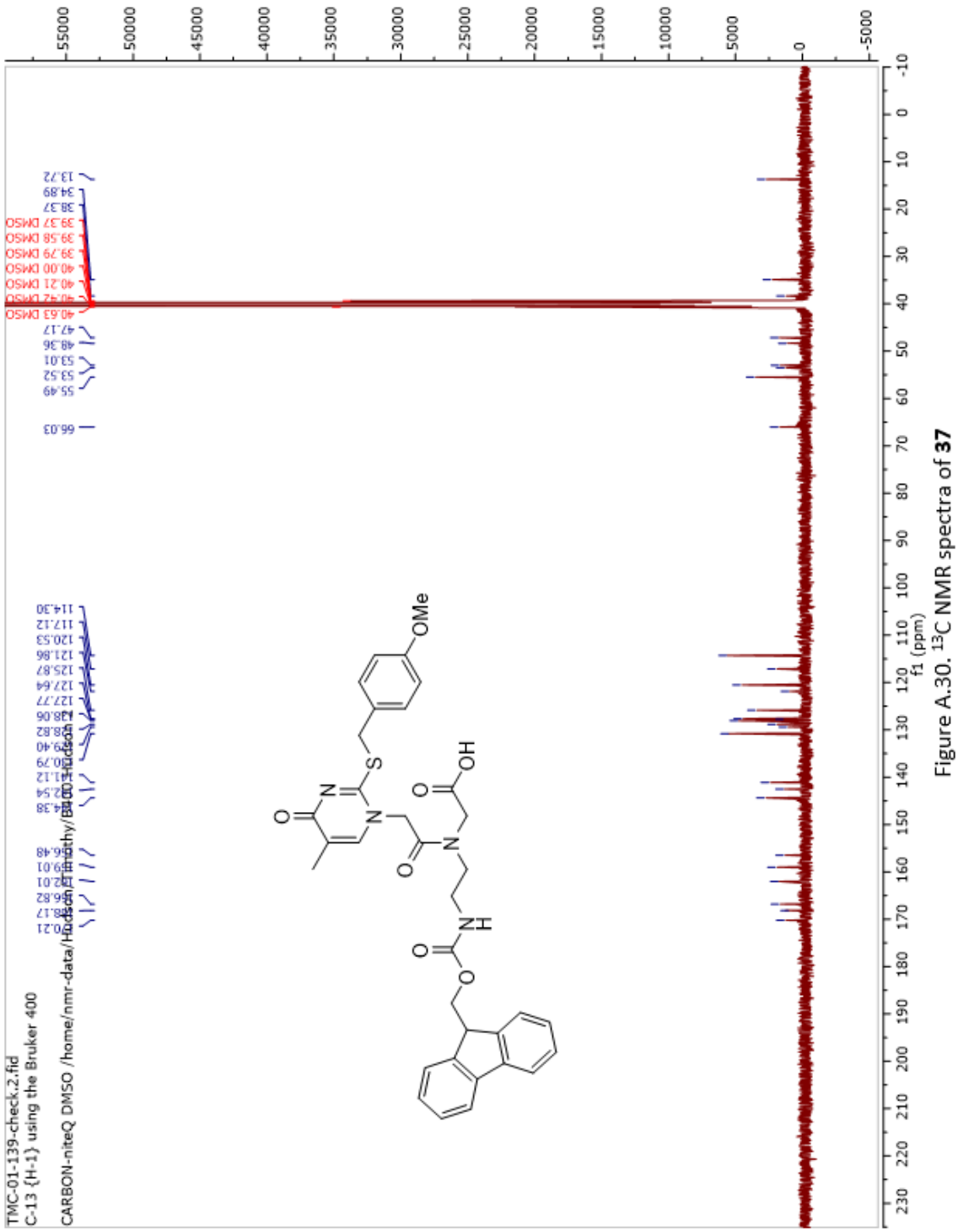


Figure A.30. ^{13}C NMR spectra of **37**

VITA

NAME: Timothy Martin-Chan

Post-Secondary Education

The University of Western Ontario
London, Ontario, Canada
B.Sc. Honors Biochemistry and Chemtry
2013-2016

The University of Western Ontario
London, Ontario, Canada
M.Sc. Chemistry
2016-2018

Related Work Experience

Teaching Assistant
Department of Chemistry, the University of Western Ontario
London, Ontario, Canada
2016-2018

Poster Presentations

Martin-Chan, Timothy and Hudson, Robert H.E. *Synthesis of 2-thiouracil PNA for Fmoc-based synthesis of pseudo-complementary oligomer*. Presented at 100th Canadian Chemistry Conference and Exhibition, Toronto, Canada, May 28-June 1, 2017.

Martin-Chan, Timothy; Park, Gyeongsu; El Samman, Amer; Hudson, Robert H.E. *A Chimeric Nucleobase – Phenylazo derivatives as an intrinsic nucleobase quencher*. Presented at 100th Canadian Chemistry Conference and Exhibition, Toronto, Canada, May 28-June 1, 2017.