

Georgia State University

ScholarWorks @ Georgia State University

Chemistry Theses

Department of Chemistry

5-8-2020

An Innovative Way To Synthesize 2'-SeMeANA-U AND 2'-SeMeANA-C

Mohammad Sazid Hassan
hassansazid@gmail.com

Follow this and additional works at: https://scholarworks.gsu.edu/chemistry_theses

Recommended Citation

Hassan, Mohammad Sazid, "An Innovative Way To Synthesize 2'-SeMeANA-U AND 2'-SeMeANA-C." Thesis, Georgia State University, 2020.
doi: <https://doi.org/10.57709/17551729>

This Thesis is brought to you for free and open access by the Department of Chemistry at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Chemistry Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

AN INNOVATIVE WAY TO SYNTHESIZE 2'-SeMeANA-U AND 2'-SeMeANA-C

by

MOHAMMAD SAZID HASSAN

Under the Direction of Zhen Huang, PhD

ABSTRACT

Nucleic acids are complex macromolecules that can store and transfer information for generations; they are an integral part of the central dogma of molecular biology. Structural determination of nucleic acids is crucial to learn about their characteristics and functions. However, there are two problems in determining the structure of a nucleic acid: the phasing problem and the crystallization problem. Selenium modification of nucleic acids is an ideal way to counter both issues. This thesis paper focuses on the synthesis of 2'-SeMe-arabino modification of uridine and cytidine as they both may enhance the crystallization capabilities of B-form DNA. Advances in these studies may lead to the discovery of new drugs and therapeutics designed to target diseases with minimum dosage and highest precision possible.

INDEX WORDS: Nucleic acid, uridine, cytidine, synthesis, mechanism, analysis, selenium, scheme, structure, central dogma.

AN INNOVATIVE WAY TO SYNTHESIZE 2'-SeMeANA-U AND 2'-SeMeANA-C

by

MOHAMMAD SAZID HASSAN

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2020

Copyright by
Mohammad Sazid Hassan
2020

AN INNOVATIVE WAY TO SYNTHESIZE 2'-SeMeANA-U AND 2'-SeMeANA-C

by

MOHAMMAD SAZID HASSAN

Committee Chair: Jun Yin

Committee: Dabney Dixon
Kathryn Grant

Electronic Version Approved:

Office of Graduate Services

College of Arts and Sciences

Georgia State University

May 2020

DEDICATION

I would like to think that it was nothing other than my hard work and effort that brought me this far, but I would be lying to myself. While I did have to work hard to come this far, I cannot in my right mind deny the contribution and support of some of the most special people in my life, without whom I don't know what I would have done. Firstly, I would like to dedicate this thesis and all the effort that went into it to my lovely mother, Nadia Sultana. My mother raised me with strictness, discipline and tough love but also with kindness and compassion. If it wasn't for my mother, I would not be studying here in the USA, trying to realize my true potential.

I would also like to dedicate this thesis to my father, Rear Admiral Nazmul Hassan. Even though you were initially skeptical of sending me away to the USA seven years ago, you have been nothing but supportive of me and my ambitions. Despite the distance, you have been an amazing father to me, supporting me in both my good times and bad in any way that you have known.

To both of my parents, I am both lucky and proud to be your son and I am very thankful to Almighty that I have been blessed with you as my parents. I hope I have made both of you proud and happy.

I would also like to dedicate my thesis to my little sister, Tanzim Nanjiba Hassan. While you are not very vocal about it, deep down I know that you have been very supportive of me and you want me to be happy and successful in life. As your elder brother, I wish the same for you.

I would also like to dedicate this thesis to my maternal grandfather, Professor Brigadier General Nazrul Islam (Retired). You have been very supportive of my efforts since the day I first

arrived in the USA and you have been giving me life lessons that I am able to reflect upon now and apply to my own life.

Lastly, I would like to dedicate this thesis to those special people in my life who are unfortunately no longer amongst us, my late maternal grandmother, Ilme Akhtar Jahan, my late paternal grandmother, Bilkis Begum and my late paternal grandfather, Colonel Abdus Samad (Retired). I wish all of you were here today to see how far I have come and everything that I have achieved so far. I hope I have made all of you proud.

ACKNOWLEDGEMENTS

When pursuing a master's degree in chemistry, it is vitally important to receive professional support, be it in the form of theoretical and/or practical guidance, funding, advise or assistance. Throughout my time as a master's student in Chemistry, I have been fortunate and thankful to receive any and every form of support.

I would like to sincerely thank Dr. Zhen Huang for accepting me into his research lab and guiding me through his research. His contribution in helping me transition from a lecture room to a research lab and doing hardcore chemistry research has been invaluable to me. I am very thankful to him for helping me grow as a student and a researcher. Attending his lecture classes and meetings have really been a good learning experience for me.

I am very grateful to Dr. Jun Yin for accepting me into his research family and agreeing to be my committee chair. I am thankful to him for taking an interest in my professional development and his offer to let me work in his lab as a PhD student. I appreciate and look forward to the opportunity of working in his research lab. I am also thankful for his valuable input and guidance in helping me write this thesis paper.

I would also like to thank my defense committee members, Dr. Dabney Dixon and Dr. Kathryn Grant. They have given me valuable advice during me time as a master's student that has helped me navigate through this graduate program.

I would like to extend my sincerest gratitude to Dr. Cen Chen. While working towards his own Ph.D., he has been an incredible mentor to me. He has guided me through projects, answered my questions, critiqued my work and steered me in the right direction. Working under him has been an honor and privilege.

I would also like to thank Shuo Liu. When I was in the beginning stages of this research and did not know what to do or how to proceed, Shuo Liu patiently taught and trained me. His training and guidance helped me to gain the knowledge and core skills necessary to carry out research projects in a chemistry lab.

I would also like to thank Lin Qin. While working as TA's in a chemistry lab, he suggested I come and join Dr. Huang's lab. He was the one who assisted me in taking the first transitioning step from a classroom to a research lab. I am also thankful for his guidance in helping me refine my core skills as a chemistry researcher.

I am also thankful to my lab mates Edwin Ogbonna, Bello Babatunde and Dr. Ziyuan Fang for their valuable insights and assistance in helping me complete my research and creating a work environment for me that helped me complete my research. I am thankful towards all my professors at Georgia State University for their guidance and wisdom. I am also thankful for all the friends I have made here who have made my time here worthwhile.

I would also like to recognize and thank the Bangladeshi Community here in Georgia who have supported me and made me feel at home in Atlanta, GA, thousands of miles away from my family and home country.

Lastly, I would like to thank the Department of Chemistry at Georgia State University for financial, technical and academic support without which I wouldn't be able to complete my master's education.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	VI
LIST OF TABLES	X
LIST OF FIGURES	XI
1 GENERAL INTRODUCTION	1
1.1 Nucleic acid and central dogma.....	1
1.2 Challenges faced in the structure determination of nucleic acids	2
<i>1.2.1 Crystallization.....</i>	<i>2</i>
<i>1.2.2 Phasing.....</i>	<i>3</i>
1.3 Selenium derivatization of nucleic acids	4
1.4 Recent studies on the selenium derivatization at specific locations.....	6
<i>1.4.1 Selenium modification on the sugar and phosphate.....</i>	<i>6</i>
<i>1.4.2 Selenium modifications on the nitrogenous base</i>	<i>7</i>
1.5 Selenium modified nucleic acids in drug design and therapeutics	7
1.6 Purpose of the study.....	9
2 SYNTHESIS OF 2'-SEMEANA-U	10
2.1 Introduction	10
<i>2.1.1 Synthesis of 2'SeMe-U and its oligonucleotides.....</i>	<i>10</i>
<i>2.1.2 Synthesis of 2'-SeMeANA-dT.....</i>	<i>10</i>
2.2 Results and Discussion.....	12

2.3	Conclusion	17
2.4	Experimental section.....	17
2.4.1	<i>General</i>	17
2.4.2	<i>Synthetic steps and characterization</i>	18
3	SYNTHESIS OF 2'-SEMEANA-C	27
3.1	Introduction	27
3.1.1	<i>Selenium modification on cytidine</i>	27
3.1.2	<i>Problems with direct derivatization of 2'-SeMeANA-C</i>	28
3.2	Results and Discussion	29
3.3	Conclusion	31
3.4	Experimental section.....	32
3.4.1	<i>General</i>	32
3.4.2	<i>Synthetic steps and characterization</i>	32
	REFERENCES.....	39
	APPENDICES.....	41
	Appendix A. Mechanism for Key Reactions.....	41
	Appendix B. Characterization data for key compounds.....	42
	<i>Appendix B.1. NMR and MS Spectra for 2'-SeMeANA-U</i>	42
	<i>Appendix B.2. NMR and MS Spectra for 2'-SeMeANA-C</i>	54

LIST OF TABLES

Table 1.1: Possible methods to address the phasing problem	3
--	---

LIST OF FIGURES

Figure 1.1: The central dogma ¹	2
Figure 1.2: The different positions of selenium modification in NA ²	5
Figure 1.3: Insertion of 2'-SeMe groups from B-form helix into the groove of the A-form helix ¹⁶	6
Figure 1.4: Synthesis scheme of the NTP α Se Analogs ⁷	7
Figure 2.1: Synthesis of the 2'-SeMe-U phosphoramidite ⁵	10
Figure 2.2: Initial attempted synthesis of 2'-SeMeANA-dT	11
Figure 2.3: Final, simplified synthesis scheme for 2'-SeMeANA-dT	12
Figure 2.4: Final synthesis scheme for 2'-SeMeANA-U	13
Figure 2.5: BOM deprotection; quenching the reaction with a mixture of TEA and iPrOH	14
Figure 2.6: Mass spectrum focusing on selenium isotopic distribution; [M+Na] ⁺ : 707.2105 (calc. 707.2063).	16
Figure 2.7: TIPDS protection at the 3' and 5' positions.....	18
Figure 2.8: BOM protection on the amino group	19
Figure 2.9: Activating the 2' OH group with methanesulfonyl chloride.....	20
Figure 2.10: Methyl selenium addition at the 2' position.....	21
Figure 2.11: BOM deprotection from the amino group	22
Figure 2.12: TIPDS deprotection from the 3' and 5' positions (U).....	24
Figure 2.13: Tritylation of the 5' OH group (U)	24
Figure 2.14: Phosphoramidite synthesis in preparation for solid-phase oligonucleotide synthesis (U).....	26
Figure 3.1: Synthetic scheme of 2' α -MeSe-C/U ²⁸	28

Figure 3.2: Direct incorporation of selenium into cytidine	28
Figure 3.3: Final synthesis scheme for 2'-SeMeANA-C	30
Figure 3.4: Conversion of base from U to C	33
Figure 3.5: Protection of the 4-N amino group on the cytidine base	34
Figure 3.6: TIPDS deprotection from the 3' and 5' positions (C)	35
Figure 3.7: Tritylation of the 5' OH group (C)	36
Figure 3.8: Phosphoramidite synthesis in preparation for solid-phase oligonucleotide synthesis (C)	37

1 GENERAL INTRODUCTION

1.1 Nucleic acid and central dogma

Nucleic acids are complex macromolecules that store information and transfers it from one generation to the next. Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are composed of long string of nucleotides, with one hydroxy group per sugar being the main difference between them. . Each nucleotide is split into three groups: a nitrogenous base, a five-carbon sugar and a phosphate group. DNAs contain the phosphate-deoxyribose sugar backbone and any one of the nitrogenous bases adenine (A), guanine (G), cytosine (C) and thymine (T) while RNAs contain the phosphate-ribose sugar backbone with any one of the same nitrogenous bases except thymine. RNA has the nitrogenous base uracil (U) instead of thymine (T).

The storage and transformation of genetic information can be explained via the central dogma. Put forth originally by Francis Crick, the central dogma has been and still is a subject of intense debate and scrutiny. However, some of the main factors behind the survival of the central dogma are how it gives a simplified explanation to the storage and transfer of genetic information from DNA to RNA to proteins,. The central dogma is not a chemical concept, but rather an informational one. The central dogma, like other scientific theories, is subject to modification once new discoveries are made. The central dogma postulates that there is no transfer of information from protein to nucleic acid. However, some exceptions have been discovered that may devalue the central dogma as an absolute principle. For example: screening of prions, agents of analog, protein conformation-based inheritance that can confer beneficial phenotypes to cells, has revealed that in eukaryotic organisms such as fungi, the flow of information is from proteins to the genome, a direct violation of the central dogma¹. Regardless, the central dogma does show the principal route of genetic information transfer.



Figure 1.1: The central dogma¹

Such is the importance of the central dogma that it has helped scientists and researchers understand the biology of living beings from birth to death. The central dogma also helped to understand how certain diseases affect human beings, e.g., various forms of cancer. Initially, drugs were used to target proteins in hopes of disease intervention. With advances in technology and medicine, the focus is gradually shifting from targeting proteins to targeting nucleic acid activities.

1.2 Challenges faced in the structure determination of nucleic acids

Structural determination of nucleic acids is an on-going field of research, which has led to the discovery of several nucleic-acid targeting disease-inhibiting drugs. Currently, the most successful method to determine the structure of a macromolecule is X-ray crystallography. However, there are two problems associated with the X-ray crystallography of nucleic acids: crystallization and phasing.

1.2.1 Crystallization

X-ray crystallography necessitates that the macromolecules to form crystals before the structure can be determined. Formation of good crystals is a time-consuming process. For a native nucleic acid, it can take weeks and sometimes months to form a crystal capable of generating good diffraction data. Crystallization can take many trials and before a crystal yields good data; this can involve months of trials and waiting. The surface of a nucleic acid is full of repetitive phosphate

groups with negative charges on each². Moreover, the buffer conditions for crystallography usually have a high concentration of salt. These buffer conditions favor A-form DNA duplexes much more than B-form DNA duplexes. B-form DNAs are more naturally available than A-form DNAs, which is why the crystallization studies of B-form DNAs are of high importance.

1.2.2 Phasing

Assuming the crystallography of nucleic acid molecules leads to the formation of a viable crystal, it is then analyzed for diffraction data. During that analysis, an electron density map is constructed. This is usually done with the application of the Fourier transformation. In the Fourier transformation, there are two major variables which are necessary to accurately calculate the electron density, the structure factor amplitude F and phase of the reflection α . The diffraction experiment can usually measure the amplitude. However, the phase of the reflection α is not directly available from that experiment. Phase is a descriptive term for electromagnetic waves, such as X-ray. When two waves of same or similar wavelengths add or cancel each other, interference occurs. The interference can either be constructive or destructive. There can be constructive or destructive interference among the diffracted X-ray beams by nucleic acid crystals. Due to the interference, the electron density map calculation does not generate the correct crystal structure². To address the issue, several strategies have been applied, as listed in Table 1.1.

Table 1.1: Possible methods to address the phasing problem

Methods	Requirement	Techniques
Direct method	Short oligonucleotides and high-resolution data	
Molecular replacement	Good homolog model	
Isomorphous replacement	Heavy-atom derivatization	Heavy metal cation soaking or co-crystallization
MAD/SAD phasing	Heavy-atom derivatization	Bromine derivatization Indirect selenium derivatization (Se-Met-U1A method) Direct selenium derivatizations

With the presence of a similar model, molecular replacement can be a simple method to solve the phase problem. However, the molecular replacement method is more difficult to apply on nucleic acids because most nucleic acids structures are novel and tightly stacked³. Which is why, nucleic acids require a method that utilizes heavy atom derivatization. Isomorphous replacement, an ideal method for the phasing of protein molecules, caused random hydrolyzation of the phosphate backbone, breaking down the nucleic acid structure into several units. To create a complete data set, several isomorphous native crystals and heavy atom derivatives were necessary, making the isomorphous replacement method costly and time consuming. In this respect, multi-wavelength anomalous diffraction (MAD) phasing is simpler, because it only requires one native crystal, as opposed to several native crystals in the isomorphous replacement method.

1.3 Selenium derivatization of nucleic acids

There are several heavy atoms that can be used for MAD phasing. One of those atoms is selenium. Selenium has the K absorption edge (0.9795 Å). This wavelength is an ideal wavelength for X-ray diffraction experiments at most synchrotron radiation facilities. Moreover, nucleic acids derivatized with a selenium atom show minimal structural perturbation. Huang, Egli, first described the concept of a selenium modified nucleic acid, where selenium is substituted with oxygen at various key points within the molecule of a nucleic acid⁴⁻⁶. Through repeated experimentation, it has been found that the selenium modification of nucleic acid, combined with multiple wavelength anomalous diffraction (MAD) phasing, addresses the crystallization and phasing problem of nucleic acid X-ray crystallography. Moreover, compared with the 5-halogen derivatized nucleic acid molecules, Se-derivatized molecules are comparatively unaffected by radiation damage⁷, this is advantageous especially as selenium provides the same phasing power

as bromine⁶. Moreover, as selenium and oxygen are both group VI-A elements, the use of selenium provides more options for derivatization sites compared to the halogen atoms, as showcased in figure 1.2. The literature shows that selenium facilitates the crystallization and structure determination of the derivatized nucleic acids without significant perturbations to the overall nucleic acid structure⁸⁻¹⁵.

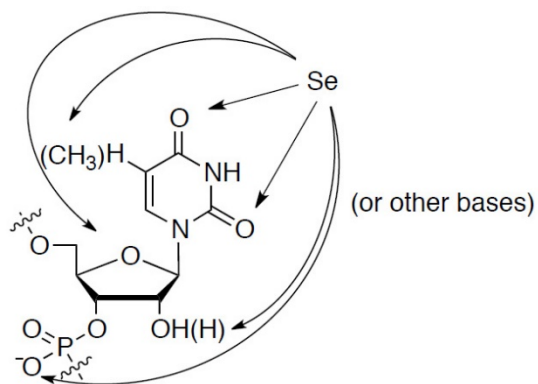


Figure 1.2: The different positions of selenium modification in NA²

Studies have been done on nucleic acid molecules with selenium derivatization at the 2' position on the sugar. These studies have shown that this specific modification has led to enhanced crystallization of the B-form DNA's¹⁶. This is a theoretical study showcasing how the 2'-SeMe modification can facilitate crystal growth of A-form DNA by destabilizing the B-form helix. Due to steric hindrance with neighboring residues, the large methylseleno group cannot tolerate the B-helix geometry. The destabilization fits well in the minor groove of the A-form helix. The groove acts as the origin of the B-form to A-form conversion. This conversion facilitates crystallization.

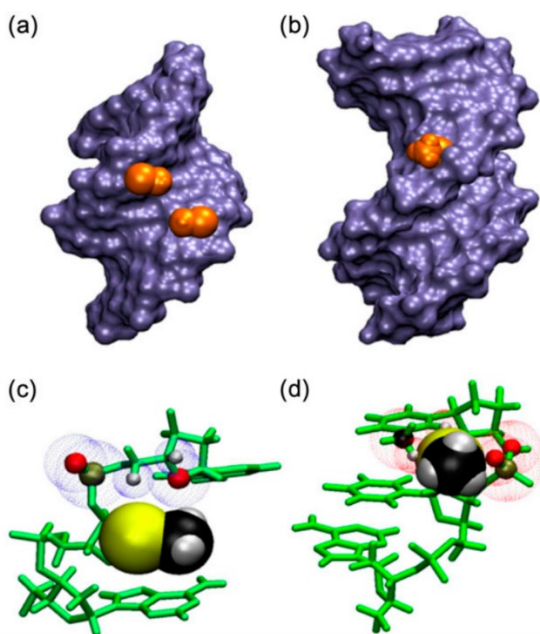


Figure 1.3: Insertion of 2'-SeMe groups from B-form helix into the groove of the A-form helix¹⁶

1.4 Recent studies on the selenium derivatization at specific locations

As discussed in 1.3, selenium and oxygen are both group VI A elements. Due to functional similarities between the two, selenium derivatization provides more options for derivatization targets on the nucleic acid molecule compared to the traditional halogen derivatization.

1.4.1 Selenium modification on the sugar and phosphate

This study utilizes a protection-free one-pot triphosphate synthesis strategy⁷. Normally, the synthesis pathway for most selenium derivatization studies calls for the introduction of protecting groups so that the selenium modification can take place on a specific target on the nucleic acid. However, using the protection free strategy, Lin et al. successfully synthesized nucleoside 5'-(α -P-seleno)-triphosphates (NTP α Se) and incorporated them into RNAs by T7 RNA polymerase¹⁷. During the triphosphate synthesis, the phosphorylating reagent was generated in situ. No purification was required. The reaction was highly regioselective at the 5'-hydroxyl group of

nucleosides containing no protecting group on the sugar or the nucleobases. The selenium was introduced after being treated with the phosphitylating reagent by adding 3H-1,2-benzothaselenol-3-one (BTSe) at room temperature. Hydrolysis was the next step. Instead of traditional HPLC, the NTP α Se analogs were purified via the boronate affinity method.

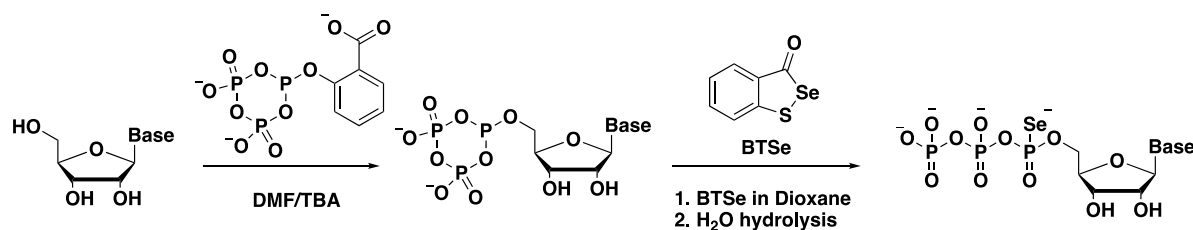


Figure 1.4: Synthesis scheme of the NTP α Se Analogs⁷

1.4.2 Selenium modifications on the nitrogenous base

In the year 2013, Salon et al. successfully synthesized the 6-Se-G phosphoramidite¹⁸. It was used as a probe to study the RNA secondary structure. Compared to the duplex structure, the 6-Se modification fit better in the bulge and wobble structure. Salon's group carried out crystallization of the 6-Se-G-modified-RNA/DNA/RNase H complex. Analysis revealed that the 6-Se modification significantly increased the quality of the generated crystals, even though the 6-Se modifications were removed naturally via hydrolysis. The native RNA complex with the DNA and RNase H gave resolution of 2.70 Å, while the 6-Se modified RNA complex with DNA and RNase H gave a resolution of 1.60 Å.

1.5 Selenium modified nucleic acids in drug design and therapeutics

Structure-based drug design (SBDD) has been very popular with proteins but less so with nucleic acids. SBDD requires the accurate structure of the macromolecule being used to be predetermined. By facilitating crystallization and improving the quality of crystals, selenium-modified nucleic acids can enhance structure based drug design of nucleic acids. As the 3D

structures become more detailed, researchers can better understand the drug target structure and the interactions between the drug targets and small molecule ligands.

The major oligonucleotides therapeutics containing antisense oligonucleotides (ASOs), aptamers, ribozymes and siRNAs, were studied extensively for over 30 years. Unfortunately, there are only 6 nucleic acid-based drugs that are FDA approved since 2017¹⁹. There are several barriers when it comes to structure-based drug design using nucleic acids including low affinity, poor delivery, vulnerability to nucleases and off-target effects. To overcome these barriers, various chemical modification of nucleic acids have been synthesized, including phosphorothioate (PS), 2'-O-methyl (2'-OMe), locked nucleic acids (LNAs), phosphorodiamidate morpholino oligomer (PMO)²⁰ and so on. By altering the structure and charge of the oligonucleotides, these nucleic acid chemical modifications can increase affinity, nuclease resistance, delivery and reduce off-target effects. One study of Se-modified nucleic acids showed that the replacement of the non-bridging oxygens on the phosphate backbone could protect the oligonucleotide from nuclease degradation²¹,²². Another study on 2'-SeMe modified oligonucleotides showed that the modification reduced multiple conformations by destabilizing unfavorable structures, potentially increasing the affinity and specificity of those selenium modified nucleic acids¹⁶. Anticancer activity of selenium modified nucleotides was also reported.²³

These results tell us that studies to accurately determine the structure of nucleic acids is vitally important for nucleic acids to be used in structure-based drug design and therapeutics. The unique properties of selenium-modified nucleic acids show great potential in drug development and clinical therapeutics. However, selenium-modified nucleic acid studies are costly, a primary factor for the studies being limited. This may prevent the broader application of selenium-modified

nucleic acids. With time and progress in these studies, application of selenium-modified nucleic acids may be achieved in SBDD and clinical therapeutics.

1.6 Purpose of the study

Prior to the structure determination of any selenium-modified nucleic acid molecule, it is vitally important to design, develop and improve upon a synthetic scheme for the selenium modified nucleotide, around which the DNA oligonucleotide can be synthesized. The purpose of the synthetic pathway is the regioselective introduction of the selenium atom and introducing protecting groups at target points of the nucleotide to synthesize the target product during the oligonucleotide synthesis. It is also important to maximize the yield at each step of the synthesis scheme and to ensure optimum purity at each step.

This study primarily focused on the 2'-arabino modification of uridine and cytidine, where commercially available uridine is the starting material for both of the synthesis studies. Studies were done previously where the 2'-alpha modification was tested to enhance the crystallization capabilities of A-form DNA^{8, 9}. Studies to test the 2'-arabino modification to enhance the crystallization capabilities of B-form DNAs have been done. One of these studies shall be briefly discussed in the upcoming sections. Those studies served as a valuable reference to develop the synthetic strategies for this study. This thesis studies the synthesis steps and eventual results of 2-SeMeANA-U and 2-SeMeANA-C.

2 SYNTHESIS OF 2'-SeMeANA-U

2.1 Introduction

To better understand how to proceed with the synthesis of the 2'-SeMeANA-U, previous studies were explored. The synthesis schemes established in those studies helped were used as the foundation to establish the final synthesis scheme for 2'-SeMeANA-U.

2.1.1 Synthesis of 2'-SeMe-U and its oligonucleotides

There have been studies showcasing the successful addition of selenium on the uridine nucleotide, followed by the corresponding structure and function studies of its oligonucleotides. In the year 2002, Du and Teplova et al. carried out a study where selenium was introduced onto the 2' position of the uridine nucleotide, followed by the synthesis, structure and function study of its DNA oligonucleotide^{5, 6}. The selenium at the 2' position was introduced as an α modification and it retained the 3'-endo sugar puckering of the A-form DNA and RNA molecules. Comparison of the crystal structures of the modified A-form oligonucleotide with the native oligonucleotide showed very little structural perturbation.

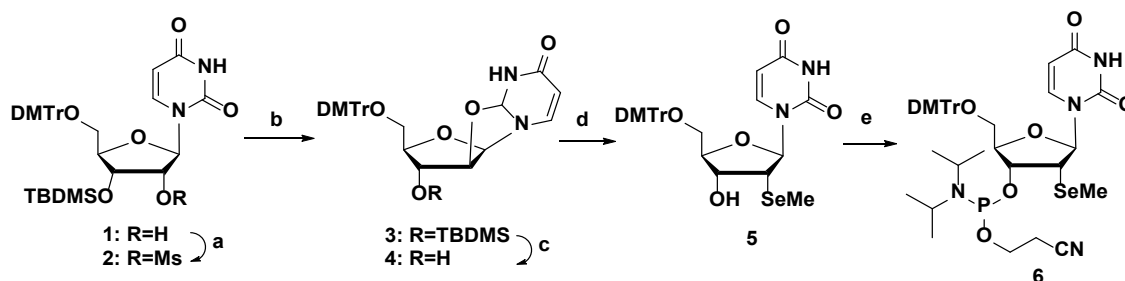


Figure 2.1: Synthesis of the 2'-SeMe-U phosphoramidite⁵

2.1.2 Synthesis of 2'-SeMeANA-dT

In 2018, Lingrui Zheng attempted the synthesis of 2'-SeMeANA-dT as part of his master's thesis. The initial attempted synthesis scheme is shown in Figure (2.2)

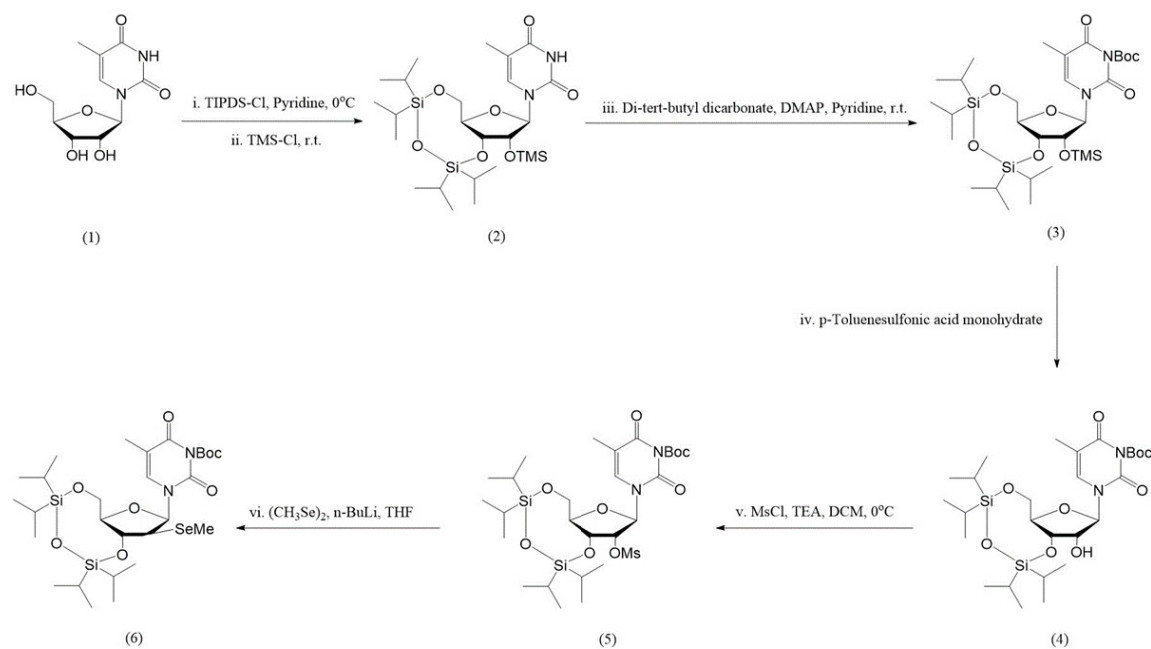


Figure 2.2: Initial attempted synthesis of 2'-SeMeANA-dT

In the scheme, Zheng attempted to protect the 2'-OH group and then introduce Boc protection on the amino group at the base. Once the amino group was protected, the 2' position was deprotected, and underwent mesylation to activate the 2' position for subsequent introduction of the Se-methyl.

In the current work, the synthesis scheme was simplified to maximize the yield. In the simplified reaction scheme, the 2' position was no longer protected while the protections on the 3' and 5' positions and the amino group on the base were introduced like before. Omitting the one protection step made no significant change in the overall purity of the reaction product while giving a higher yield. The final, simplified scheme is shown in figure 2.3.

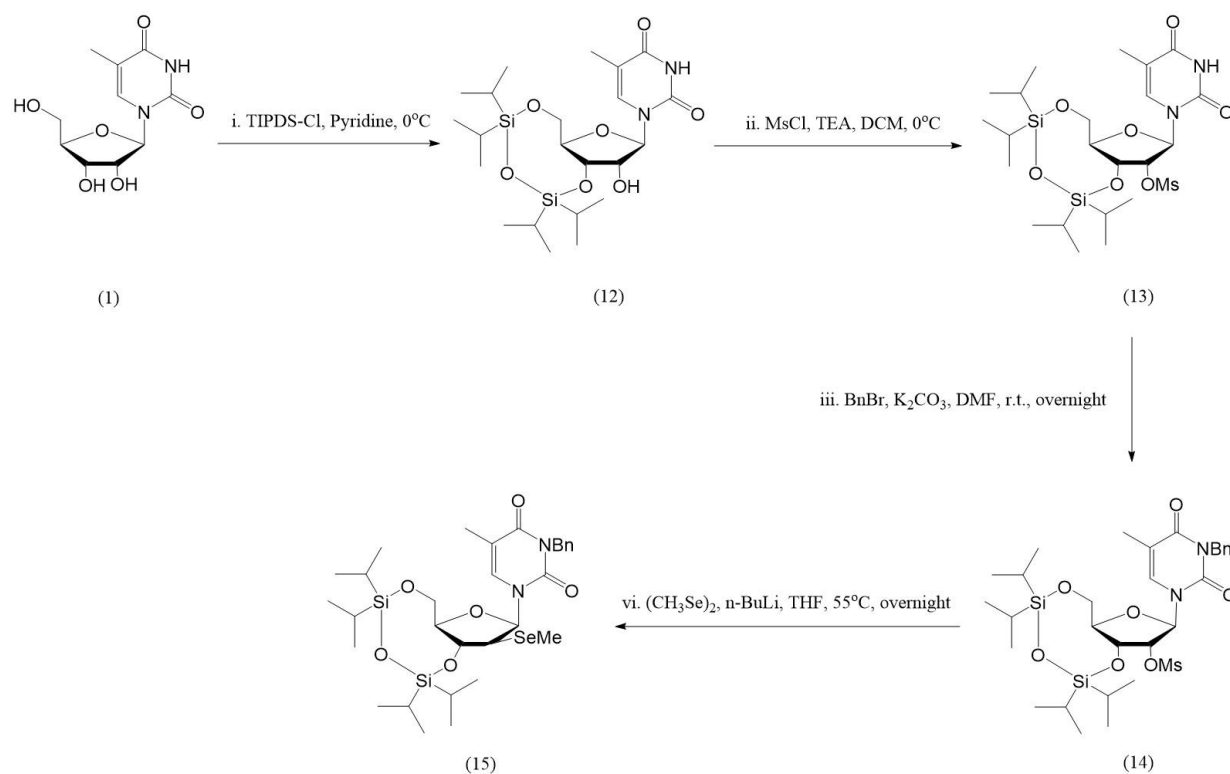


Figure 2.3: Final, simplified synthesis scheme for 2'-SeMeANA-dT

This synthetic scheme (figure 2.3) was used as a reference to develop a final synthetic scheme for 2'-SeMeANA-U.

2.2 Results and Discussion

Commercially available uridine (**compound 1**) was treated with tetraisopropylidisilylene (TIPDS) to protect the 3' and 5' hydroxyl groups. Then the product (**compound 2**) was treated with benzyloxymethyl acetal (BOM) to protect the 3-N amino group. To activate the 2'-OH group the product (**compound 3**) was treated with methanesulfonyl chloride. To generate the Se-Me nucleophile, dimethyl diselenide was dissolved in anhydrous THF, treated with n-butyllithium solution at -78°C. This solution was introduced to the activated **compound 4** at 60°C. After the incorporation of selenium, the BOM protecting group from **compound 5** dissolved in anhydrous THF was deprotected at -78°C by treatment with BBr₃. This reaction was quenched with a mixture

of triethylamine (TEA) and isopropanol in a 1:1 v/v ratio. It is important that the isopropanol is anhydrous to prevent protic solvents such as methanol or water from displacing the Se-Me from the 2' position (Figure 2.5). **Compound 6** was treated with 3HF.TEA to remove the TIPDS group and generate **Compound 7**. The 5' hydroxyl group on **compound 7** was selectively protected with the DMTr protecting group. Lastly, the 2'SeMeANA-U (**compound 8**) was converted to its phosphoramidite. This final step gave an overall 90% yield (Figure 2.4).

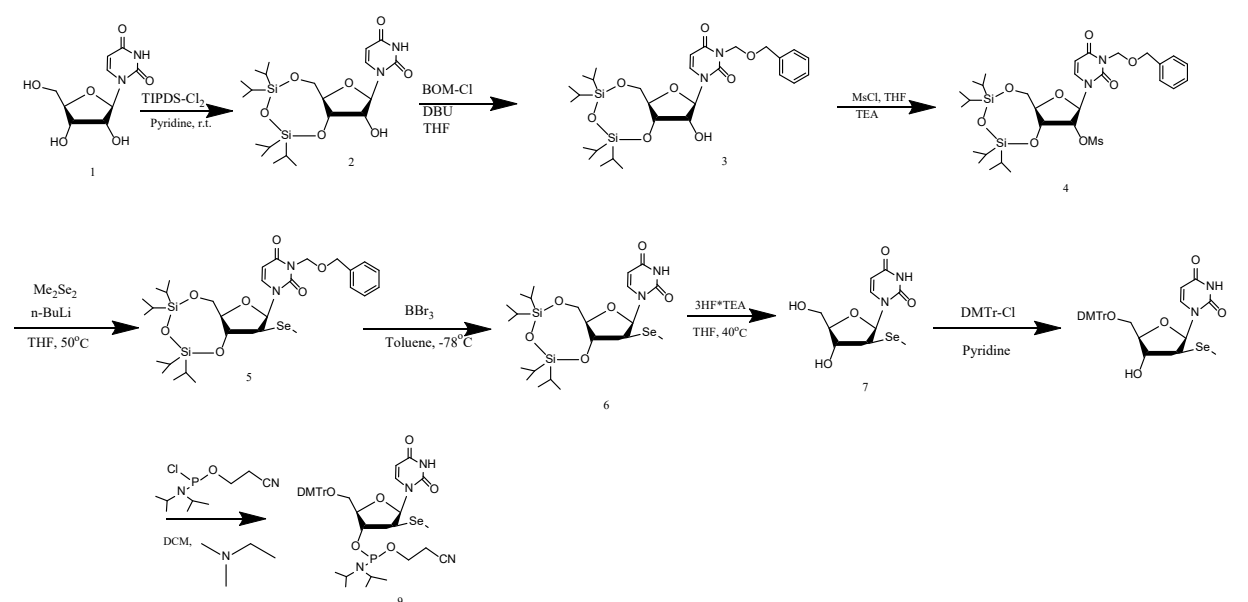


Figure 2.4: Final synthesis scheme for 2'-SeMeANA-U

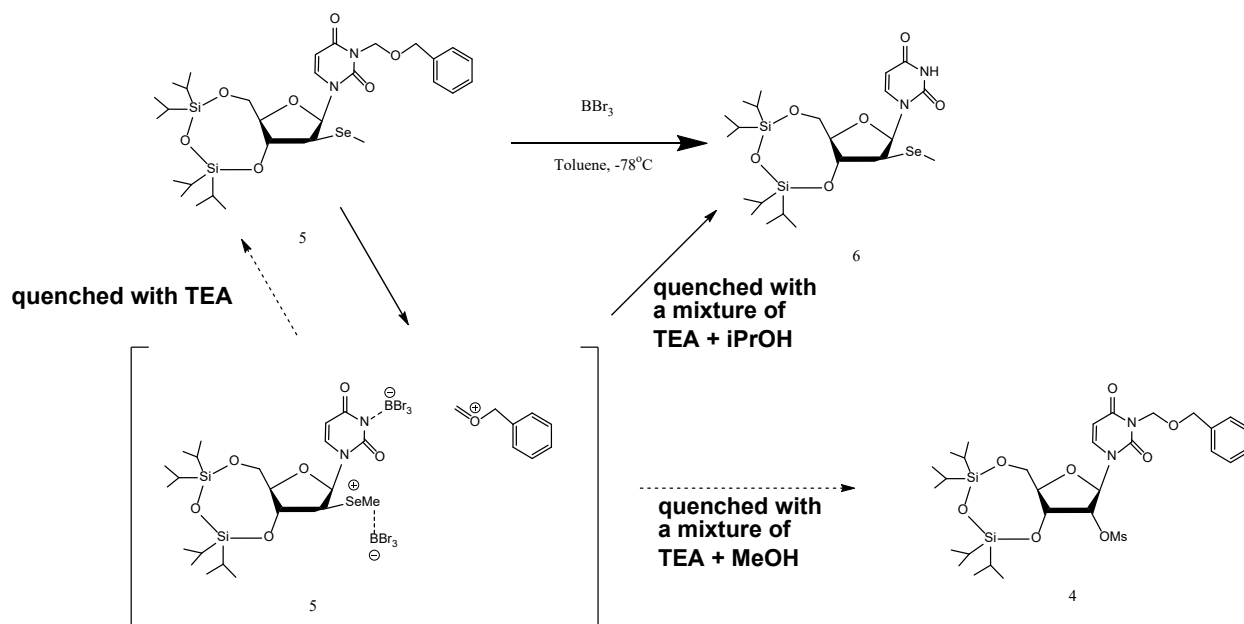


Figure 2.5: BOM deprotection; quenching the reaction with a mixture of TEA and $i\text{PrOH}$

NMR and MALDI-TOF analysis were performed at key steps of the novel synthesis scheme to ensure that the desired product was synthesized. In the ^1H NMR spectra of compound 3, the tall peaks with a chemical shift between 1 ppm to 1.10 ppm indicated that the 3' and 5' hydroxy groups were protected by TIPDS. After BOM protection, multiple peaks with chemical shifts between 4.5 ppm to 5.5 ppm. These peaks corresponded to the hydrogen protons within the BOM protecting group. The hydrogen protons closer to the aromatic ring shifted more to the right (lower ppm) while the hydron protons closer to the base of the nucleotide shifted more to the left (higher ppm). There were also peaks in the aromatic region of the NMR spectra (7.27 ppm to 7.37 ppm), which came from the aromatic ring of the BOM protecting group. After the activation of the 2'-OH with methanesulfonyl chloride, there was a peak observed, which had a chemical shift of 3.28 ppm while changes in coupling patterns could be seen on adjacent peaks (^1H NMR, compound 4). After selenium addition at the 2' position, a singlet was observed with a chemical shift of 2.08 ppm (^1H NMR, compound 5). Prior to the introduction of selenium, the

hydrogen proton at the 1' position of the sugar had maintained singlet coupling with chemical shifts between 5.70 ppm to 5.75 ppm (^1H NMR, compounds 3 and 4). After selenium addition, the 1'H proton coupling changed from singlet coupling to doublet coupling, with chemical shifts between 6.40 ppm to 6.55 ppm (^1H NMR, compound 5). This confirmed that the selenium introduction was in the β position and not the α position. MS analysis after selenium introduction also showcased the predicted selenium isotopic distribution (Figure 2.6). BOM deprotection, followed by TIPDS deprotection removed the respective peaks from the NMR spectra whilst changing the coupling patterns of the peaks that were still present in the NMR. After the DMTr group was introduced to the 5' position, a singlet was observed with chemical shifts at approximately 3.80 ppm corresponding to the hydrogen protons on the methoxy groups in the DMTr. The hydrogen protons in the DMTr aromatic rings had chemical shifts between 7.25 ppm to 7.46 ppm. MS analysis was done on the phosphoramidite to confirm the final product. The detailed characterization data can be found in the experimental section while the relevant NMR and MS spectra can be found in Appendix B.1.

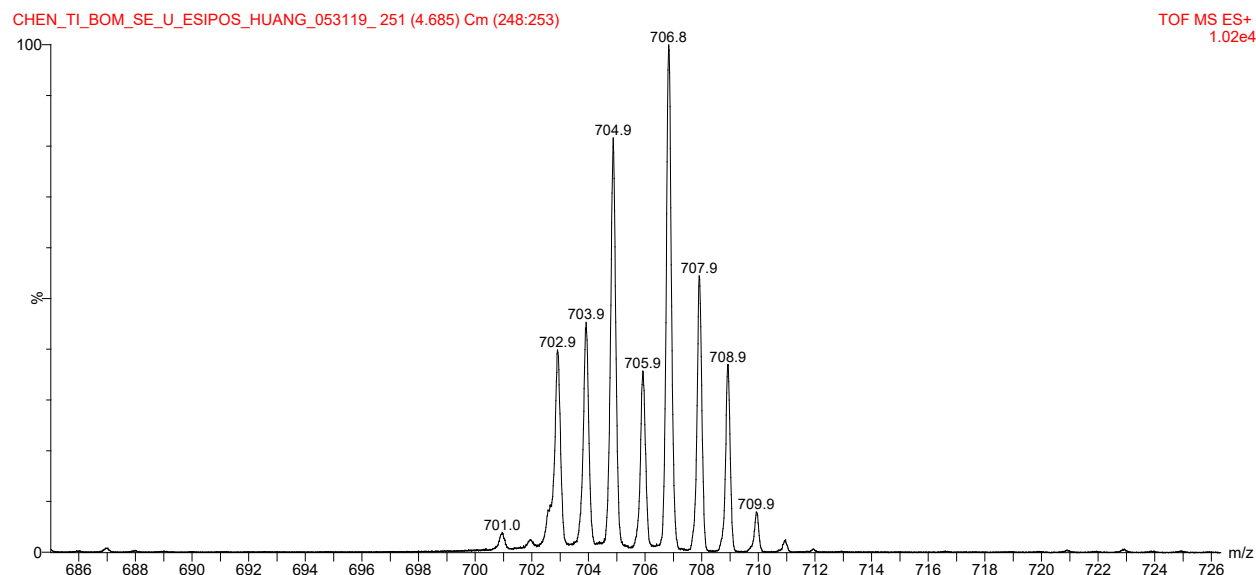


Figure 2.6: Mass spectrum focusing on selenium isotopic distribution; $[M+Na]^+$: 707.2105 (calc. 707.2063).

It was observed that the reactions to generate the products prior to selenium incorporation were clean, did not form by-products. These synthesis steps are quite similar to the synthesis steps in previous studies for the 2'Se-modified guanosine⁸ and 2'Se-modified adenosine⁹. Since this study focused on the β -modification instead of the α , the protecting groups involved are different and the synthesis scheme is novel. Particularly with the introduction of the methyl selenium at the 2' position and with the steps going forward, flash chromatography became vitally important. Because thin-layer chromatography (TLC) would sometimes show multiple spots, instead of one. If there are by-products formed during a reaction and they have UV-Absorption, it will show on the TLC plate, along with the target product. Another problem with these reactions would be that all the starting material would not react, and this would also show during the TLC. Silica gel chromatography would help to isolate the target product from the unreacted starting material and the unwanted by-products and the recovered product containing solvents would then be dried under pressure under vacuum to obtain the pure product.

Besides the selenium incorporation step, the BOM deprotection step and the TIPDS deprotection step presented a unique challenge in comparison with the rest of the synthesis scheme for 2'-SeMeANA-U. Deprotection of the protecting groups from the nitrogenous base and the 3' and 5' positions on the sugar made the compound increasingly polar and it was evidenced by TLC. It was during these two steps that moisture prevention from the reaction was vitally important and argon purging helped to prevent moisture from interacting with the reaction and forming by-products.

2.3 Conclusion

In conclusion, the final synthetic scheme was devised for the formation of 2'-SeMeANA-U (Figure 2.4). This synthetic scheme ensured the successful synthesis of the 2'-SeMe-arabino uridine phosphoramidite. The phosphoramidite product generated from the synthesis was utilized by Dr. Cen Chen to carry out the synthesis of several DNA oligonucleotides for structure and function studies. The synthetic scheme developed for the synthesis of the Se-modified uridine phosphoramidite can be used to synthesize several oligonucleotides to carry out structure and function studies. The results from these studies can help us understand the behavior of selenium modified arabino uridine oligonucleotides for possible applications in SBDD or even therapeutics.

2.4 Experimental section

2.4.1 General

Most solvents and reagents were purchased from Sigma, Fluka, or Aldrich (PA) and used without purification, unless otherwise specified. Solid reagents were dried under high vacuum when it was necessary to do so. Reactions with compounds sensitive to air or moisture were performed under argon purging. Solvent mixtures are indicated as volume/volume ratios. Thin-layer chromatography (TLC) was performed on Merck 60 F254 plates (0.25 mm thick). TLC spots

were visualized under UV light. Column purification was performed using Fluka silica gel 60 (mesh size 0.040-0.063 mm) using a silica gel and crude compound weight ratio of ca. 30:1. ^1H spectra were recorded using Bruker-300 or 400 (300 or 400 MHz). All chemical shifts (δ) are in ppm relative to tetramethylsilane and all coupling constants (J) are in Hz. High resolution (HR) MS were either obtained with electrospray ionization (ESI) on a Q-TOFTM Waters Micromass at Georgia State University.

2.4.2 Synthetic steps and characterization

2.4.2.1 Tetraisopropylidisiloxanylidene (TIPDS) protection

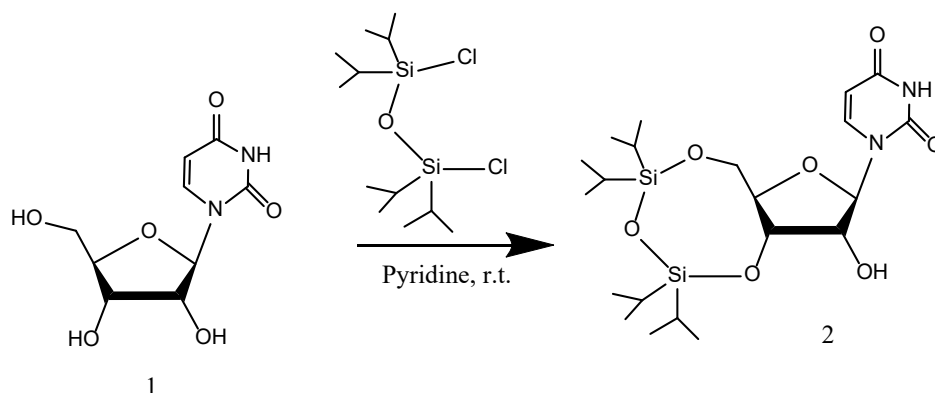


Figure 2.7: TIPDS protection at the 3' and 5' positions

3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-uridine (compound 2): Compound **1** (11 g, 49.1 mmol) was dissolved in 250 mL anhydrous pyridine and the solution was purged with dry argon. The reaction was placed in an ice-bath. 1,3-dichloro-1,1,3,3-tetraisopropylidisiloxane (12.8 mL, 58.9 mmol) was added dropwise. After the addition, the reaction was taken off the ice-bath and stirred at room temperature for 5 h. TLC was performed with 50% ethyl acetate in hexane as the eluent. 5 mL of methanol was added to quench the reaction. The solvent was evaporated under reduced pressure. Upon evaporation of the solvents, the product was dissolved in 300 mL methylene chloride. The organic solution was washed with 7 mL HCl (3 M), 75 mL

saturated sodium bicarbonate, 100 mL water and 100 mL brine. The aqueous layer was separated from the organic layer. The organic layer was dried over anhydrous MgSO_4 , filtered to remove the MgSO_4 and then concentrated in vacuum. The final product was a dry, white solid. The product was used in the next step without further purification. ^1H -NMR (400 MHz, CDCl_3) δ (ppm): 1.03-1.09 (m, 28H, 4*×*iPrSi), 3.28 (br, s, 1H, OH), 4.01 (dd, $J_1=2.2$ Hz, $J_2=13.1$ Hz, 1H, H-5'), 4.09 (d, $J=8.4$ Hz, 1H, H-4'), 4.18 (m, 2H, H-2', H-5'), 4.38 (dd, $J_1=4.7$ Hz, $J_2=8.6$ Hz, 1H, H-3'), 5.67 (d,) 5.71 (d, $J=8.1$ Hz, 1H, H-1'), 7.66 (d, $J=8.1$ Hz 1H, H-6), 8.52 (br, s, 1H, NH). ^1H -NMR spectrum is identical to the literature²⁴.

2.4.2.2 Benzyloxymethyl acetal (BOM) protection

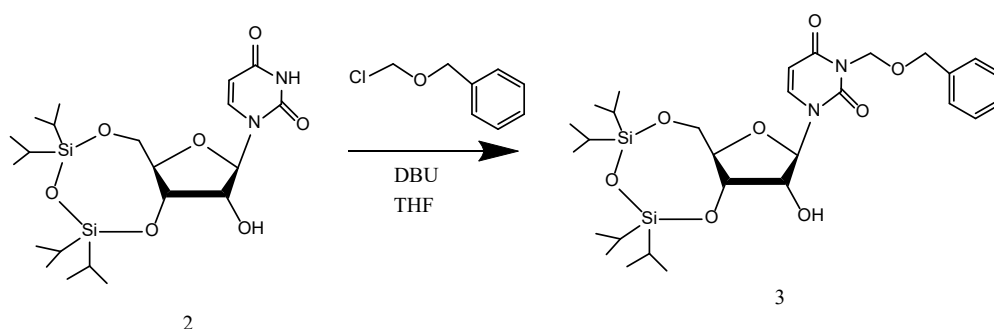


Figure 2.8: BOM protection on the amino group

3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-N3-(benzyloxymethylacetal)-uridine (compound 3): compound 2 (6.5 g, 13.4 mmol) was dissolved in 50 mL THF. 1,8-diazabicyclo(5.4.0)undec-7-ene (DBU, 4 mL, 26.8 mmol) was added to the mixture. The mixture was treated with benzyl chloromethyl ether (2.8 mL, 20.1 mmol) and stirred at room temperature for 2 h. TLC was done with 50% ethyl acetate in hexane as the eluent. The solvent was evaporated under reduced pressure. The dried product was then washed with 50 mL water and 70 mL brine. The organic layer was separated from the aqueous layer, dried over anhydrous MgSO_4 , filtered and then concentrated in vacuum. The residue was purified by silica gel

chromatography (25 % ethyl acetate in hexane) to give about the desired product with 80% yield.

$^1\text{H-NMR}$ (400 MHz, CDCl_3) δ (ppm): 1.02-1.10 (m, 28H, 4 \times iPrSi), 2.90 (br, s, 1H, OH), 4.00 (dd, $J_1=2.8$ Hz, $J_2=13.2$ Hz, 1H, H-5'), 4.10 (dt, $J_1=3.7$ Hz, $J_2=8.5$ Hz, 1H, H-4'), 4.13 (d, $J=4.9$ Hz, 1H, H-2'), 4.21 (dd, $J_1=1.4$ Hz, $J_2=13.3$ Hz, 1H, H-5'), 4.34 (dd, $J_1=4.9$ Hz, $J_2=8.8$ Hz, 1H, H-3'), 4.72 (s, 2H, $\text{CH}_2\text{OCH}_2\text{Ph}$), 5.47 (d, $J=2.4$ Hz, 2H, $\text{CH}_2\text{OCH}_2\text{Ph}$), 5.71 (d, $J=8.2$ Hz, 1H, H-5), 5.73 (s, 1H, H-1'), 7.38-7.26 (m, 6H, aromatic, H-6), 7.64 (d, $J=8.2$ Hz, 1H, H-6). HRMS (ESI): $\text{C}_{29}\text{H}_{46}\text{N}_2\text{O}_8\text{Si}_2$; $[\text{M}+\text{Na}]^+$: 629.1387 (calc. 629.2690)

2.4.2.3 Activation of the 2' OH group

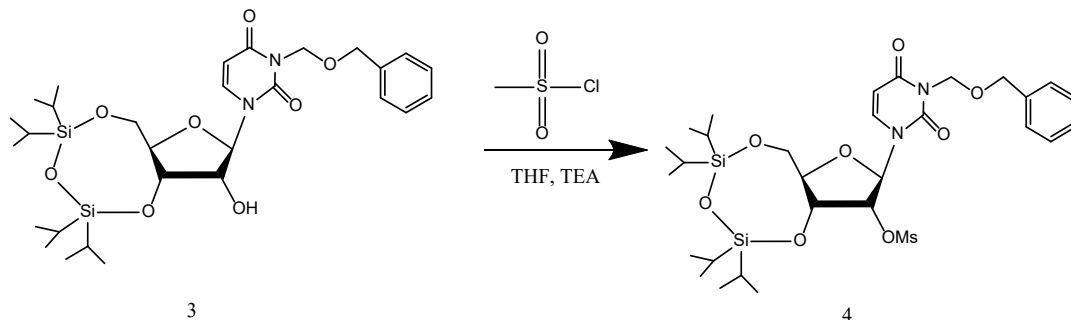


Figure 2.9: Activating the 2' OH group with methanesulfonyl chloride

2'-O-Mesyl-3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-N3-(benzyloxymethyl acetal)-uridine (compound 4): Compound 3 (12.5 g, 20.6 mmol) was dissolved in 150 mL THF. Triethylamine (11.5 mL, 82.4 mmol) was then added to the mixture. To this solution, methanesulfonyl chloride (3.2 mL, 41.2 mmol) was added dropwise. The reaction was stirred at room temperature for 3 hours. TLC was performed with 50% ethyl acetate in hexane as the eluent. The solvent was evaporated under reduced pressure. The dried product was dissolved in 150 mL methylene chloride. The organic solution was washed with 50 mL water and 50 mL brine. After washing, the aqueous layer and organic layer were separated. The organic layer was dried over anhydrous MgSO_4 , filtered and then concentrated in vacuum. The derived residue was

purified by silica gel chromatography (15 % ethyl acetate in hexane) to give the desired product.

This reaction had a 90% yield. $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ (ppm): 1.04-1.11 (m, 28H, 4*×*iPrSi), 3.28 (s, 3H, O_3SCH_3), 3.99 (dd, $J_1=2.3$ Hz, $J_2=13.7$ Hz, 1H, H-5'), 4.10 (dd, $J_1=1.9$ Hz, $J_2=9.7$ Hz, 1H, H-4'), 4.25-4.30 (m, 2H, H-3', H-5'), 4.71 (s, 2H, $\text{CH}_2\text{OCH}_2\text{Ph}$), 4.97 (d, $J=4.4$ Hz, 1H, H-2'), 5.48 (m, 2H, $\text{CH}_2\text{OCH}_2\text{Ph}$), 5.72 (d, $J=8.2$ Hz, 1H, H-5), 5.79 (s, 1H, H-1'), 7.26-7.37 (m, 5H, aromatic), 7.74 (d, $J=8.2$ Hz, 1H, H-6). HRMS (ESI): $\text{C}_{30}\text{H}_{48}\text{N}_2\text{O}_{10}\text{SSi}_2$; $[\text{M}+\text{H}]^+$: 685.1212 (calc. 685.2646).

2.4.2.4 Selenium introduction at the 2' position

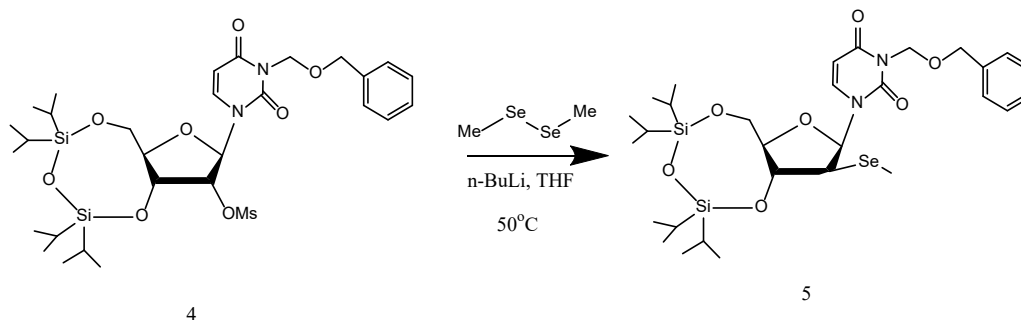


Figure 2.10: Methyl selenium addition at the 2' position

3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-N3-(benzyloxymethylacetal)-2'-methylseleno-arabinouridine (compound 5): Dimethyl diselenide (Me_2Se_2 , 2.4 mL, 25.3 mmol) was dissolved in anhydrous tetrahydrofuran (25 mL). This reaction is very sensitive to moisture and so the reaction was purged under argon. This reaction is also sensitive to temperature and so the solution was cooled to -78°C by immersing the reaction flask in an acetone/dry ice bath. In the ice bath, the reaction was stirred for about 10-15 min to allow the temperature to lower. At this stage, n-butyllithium solution (5 mL, 2.5 M in hexane, 12.6 mmol) was added dropwise using a syringe. The reaction was warmed up to room temperature slowly. In a separate round-bottom flask, **Compound 4** (4.3 g, 6.3 mmol) was dissolved in 25 mL anhydrous THF. After the starting

material was dissolved, it was then transferred to the reaction flask containing the dimethyl diselenide using a syringe. The reaction mixture was heated up to 65°C and stirred overnight. To verify completion of the reaction, thin-layer chromatography was performed with 50% ethyl acetate in hexane as the eluent. The reaction solution was evaporated under reduced pressure. The concentrated solution was then re-dissolved in 100 mL ethyl acetate. The organic solution was washed with 30 mL water and 40 mL brine. The organic and aqueous layers were separated. The organic layer was dried over anhydrous MgSO_4 , filtered and then evaporated to dryness. The crude product was then purified by silica gel chromatography (10% ethyl acetate in hexane) to give desired product as a white solid. This reaction had a yield of 60%. ^1H -NMR (400 MHz, CDCl_3) δ (ppm): 1.03-1.11 (m, 28H, 4*×*iPrSi), 2.08 (s, 3H, SeCH₃), 3.64 (dd, $J_1=7.0$ Hz, $J_2=10.1$ Hz, 1H, H-2'), 3.76 (d, $J=8.1$ Hz, 1H, H-4'), 4.04 (dd, $J_1=2.7$ Hz, $J_2=13.2$ Hz, 1H, Ha-5'), 4.13 (dd, $J_1=1.7$ Hz, $J_2=13.2$ Hz, 1H, Hb-5'), 4.18 (dd, $J_1=8.6$ Hz, $J_2=9.7$ Hz, 1H, H-3'), 4.68 (d, $J=3.1$ Hz, 2H, CH₂OCH₂Ph), 5.51 (s, 2H, CH₂OCH₂Ph), 5.75 (d, $J=8.2$ Hz, 1H, H-5), 6.40 (d, $J=6.9$ Hz, 1H, H-1'), 7.28-7.37 (m, 5H, Ph), 7.58 (d, $J=8.2$ Hz, 1H, H-6). HRMS (ESI): $\text{C}_{30}\text{H}_{48}\text{N}_2\text{O}_7\text{SeSi}_2$; $[\text{M}+\text{Na}]^+$: 707.2105 (calc. 707.2063).

2.4.2.5 Benzyloxymethyl acetal (BOM) deprotection

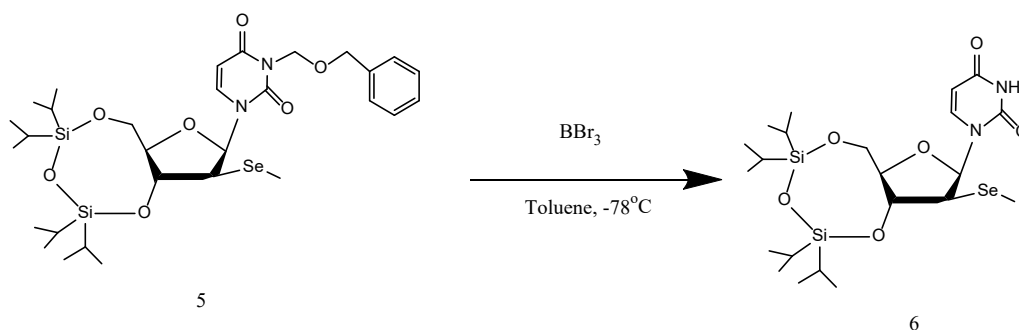


Figure 2.11: BOM deprotection from the amino group

3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'-methylseleno-arabinouridine

(compound 6): Compound 5 (1.7 g, 2.5 mmol) was placed in a three neck round bottom flask. The flask was purged with argon. The starting material was dissolved in 15 mL anhydrous toluene. The toluene was injected into the round bottom flask with a syringe. The solution was placed in a dry ice/acetone bath and stirred for 15 min. After the temperature of the solution reached -78°C , a solution of boron tribromide (3.8 mL, 1 M in hexane, 3.8 mmol) was injected. The reaction was stirred at -78°C for 1 h. Thin layer chromatography was done with 50% ethyl acetate in hexane as the eluent. The reaction was quenched by adding a mixture of triethylamine and anhydrous isopropanol in a 1:1 volume by volume ratio (2 mL). The solution was removed from the dry-ice bath. After quenching the reaction, the reaction solution was stirred for 1 h to warm up to room temperature. The solvents were then evaporated under reduced pressure, until about 2 mL of the solution was left. The remainder of the solution was then diluted with 50 mL of ethyl acetate. The organic solution was washed with 30 mL water. The water layer was then extracted with ethyl acetate 3 times. Each time, 10 mL of ethyl acetate was used. The combined organic layer was washed with 40 mL brine. The organic layer was then separated from the brine, dried over anhydrous MgSO_4 filtered and then evaporated to dryness. The residue was then purified by silica gel chromatography (15% ethyl acetate in hexane) to give the desired product in 70% yield. ^1H -NMR (400 MHz, CDCl_3) δ (ppm): 1.05-1.12 (m, 28H, 4*×*iPrSi), 2.13 (s, 3H, SeCH₃), 3.64 (dd, $J_1=7.0$, $J_2=10.1$ Hz, 1H, H-2'), 3.77 (dt, $J_1=2.4$, $J_2=8.1$ Hz, 1H, H-4'), 4.05 (dd, $J_1=2.8$, $J_2=13.2$ Hz, 1H, Ha-5'), 4.13 (dd, $J_1=1.9$, $J_2=13.2$ Hz, 1H, Hb-5'), 4.21 (dd, $J_1=8.3$, $J_2=10.0$ Hz, 1H, H-3'), 5.72 (dd, $J_1=2.0$, $J_2=8.1$ Hz, 1H, H-5), 6.39 (d, $J=7.0$, 1H, H-1'), 7.59 (d, $J=8.2$ 1H, H-6), 9.10 (br, 1H, NH). HRMS (ESI): $\text{C}_{22}\text{H}_{40}\text{N}_2\text{O}_6\text{SeSi}_2$; $[\text{M}-\text{H}]^-$: 563.1525 (calc. 563.1512).

2.4.2.6 Tetraisopropylidisiloxanylidene (TIPDS) deprotection

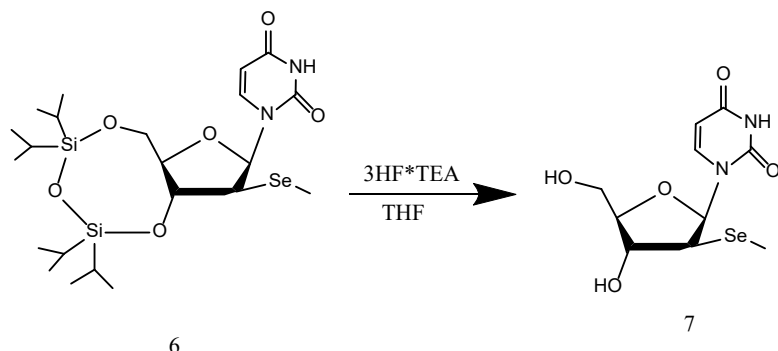


Figure 2.12: TIPDS deprotection from the 3' and 5' positions (U)

5-Methyl-2'-methylseleno-arabinouridine (compound 7): Compound 6 (760 mg, 1.35 mmol) was first dissolved in 5 mL anhydrous THF. The solution was then treated with 0.22 mL of triethylamine trihydrofluoride ($3\text{HF}\cdot\text{Et}_3\text{N}$). After addition of the $3\text{HF}\cdot\text{Et}_3\text{N}$ the solution was heated to 40°C and stirred for 2 hours. TLC was done with 10% methanol in dichloromethane as the eluent. The solvent was evaporated to dryness. The crude product was then subjected to silica gel chromatography with 5% methanol in dichloromethane to obtain the pure product. HRMS (ESI): $\text{C}_{10}\text{H}_{14}\text{N}_2\text{O}_5\text{Se}$; $[\text{M}-\text{H}]^-$: 320.9982 (calc. 320.9990).

2.4.2.7 Introduction of the dimethoxytrityl group (DMTr-) at the 5' position

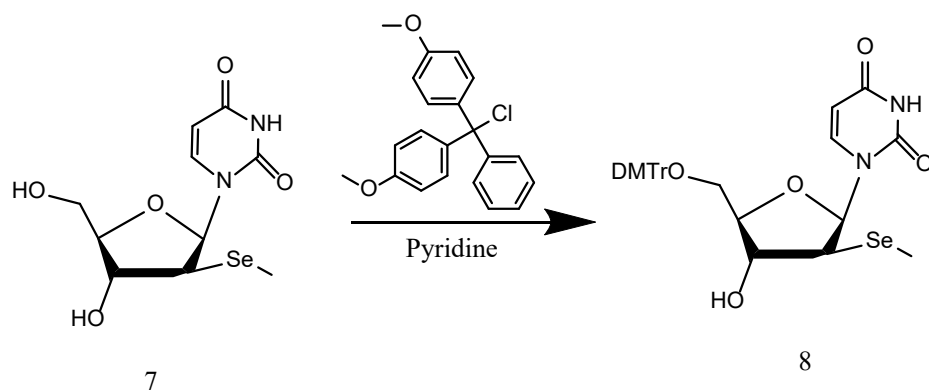


Figure 2.13: Tritylation of the 5' OH group (U)

5'-O-(4,4'-dimethoxytrityl)-5-methyl-2'-methylseleno-arabinouridine (compound 8):

Compound 7 (450 mg, 1.40 mmol) was dried over high vacuum and co-evaporated with anhydrous pyridine (2 x 10 mL). Then the starting material was dissolved in 15 mL anhydrous pyridine. The reaction was then cooled in an ice-bath. While the reaction was cool, it was then treated with 522 mg (1.54 mmol) of the dimethoxytrityl chloride. The reaction was then taken off the ice bath and allowed to reach room temperature. The reaction was stirred for 2 h. Thin layer chromatography was done with 50% ethyl acetate in hexane as the eluent. The solvent was evaporated under reduced pressure. The concentrated product was dissolved in 30 mL methylene chloride, then washed with 20 mL water and 30 mL brine. The aqueous layer and organic layer were separated from each other. The organic layer was dried over anhydrous MgSO_4 , filtered and then evaporated to dryness. The crude product underwent silica gel chromatography with 3% MeOH in methylene chloride with 1% Et_3N as the eluent. The final product for this step was a white solid with 90% yield. $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ (ppm): 2.13 (s, 3H, SeCH_3), 2.66 (d, $J=4.1$ Hz, 1H, 3'-OH), 3.48 (dd, $J_1=3.8$ Hz, $J_2=10.7$ Hz, 1H, H-5'), 3.57 (dd, $J_1=3.3$ Hz, $J_2=10.8$ Hz, 1H, H-5'), 3.61 (dd, $J_1=7.3$ Hz, $J_2=8.3$ Hz, 1H, H-2'), 3.81 (s, 6H, OCH_3), 3.88 (dt, $J_1=3.6$ Hz, $J_2=6.1$ Hz, 1H, H-4'), 4.30 (td, $J_1=4.0$ Hz, $J_2=7.9$ Hz, 1H, H-3'), 5.72 (dd, $J_1=2.0$, $J_2=8.1$ Hz, 1H, H-5), 6.41 (d, $J=7.2$ Hz, 1H, H-1'), 6.85-6.87 (m, 4H, Ph), 7.25-7.46 (m, 9H, Ph), 7.59 (d, $J=8.2$ 1H, H-6), 8.95 (br, 1H, NH). HRMS (ESI): $\text{C}_{31}\text{H}_{31}\text{N}_2\text{O}_7\text{Se}$; $[\text{M-H}]^-$: 623.1298 (calc. 623.1296).

2.4.2.8 Phosphoramidite conversion for solid-phase oligonucleotide synthesis

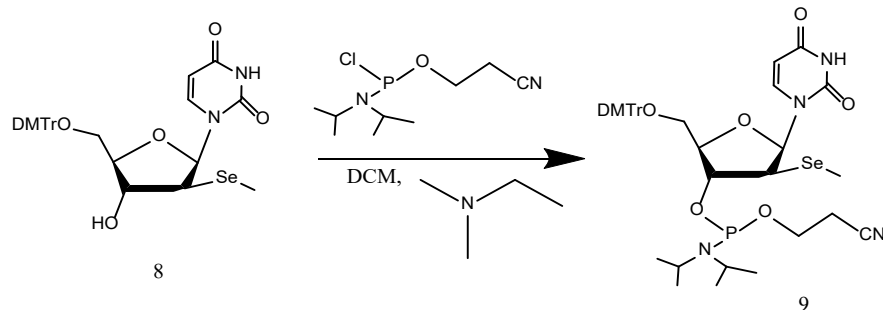


Figure 2.14: Phosphoramidite synthesis in preparation for solid-phase oligonucleotide synthesis (U)

3'-O-(2-Cyanoethyl-*N,N*-diisopropylphosphoramidite)-5'-O-(4,4'-dimethoxytrityl)-5-methyl-2'-methylseleno-arabinouridine (compound 9): Compound 8 (680 mg, 1.1 mmol) was dried under high vacuum and dissolved in 5 mL anhydrous methylene chloride (CH_2Cl_2). To the solution, dimethylethanolamine (0.7 mL, 6.6 mmol) was injected. Lastly, the phosphoramidite reagent 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (0.2 mL, 0.8 mmol) was added to the reaction and stirred for about 1 h under dry argon. The reaction mixture was evaporated under reduced pressure. The concentrated product was then dissolved in 2 mL CH_2Cl_2 . The reaction mixture was then precipitated under vigorous stirring in a 400 mL hexane solution (400 mL). After carefully decanting the hexane solution, the crude product was purified by silica gel chromatography (30% ethyl acetate in CH_2Cl_2 containing 1% dimethylethylamine). The purified product was precipitated again in the 400 mL hexane solution. The precipitate was then dissolved in 2 mL CH_2Cl_2 . The solvent was then evaporated. The derived product (659 mg, 80%) was a white foam. HRMS (ESI): $\text{C}_{40}\text{H}_{49}\text{N}_4\text{O}_8\text{PSe}$; $[\text{M}+\text{Na}]^+$: 847.2376 (calc. 847.2351).

3 SYNTHESIS OF 2'-SeMeANA-C

3.1 Introduction

Another synthesis scheme that was devised for this study was for the synthesis of 2'-SeMeANA-C. Up to a point in the synthesis scheme, the schemes for this compound and the modified uridine are similar. Initially, an independent synthesis scheme was designed for the modified synthesis. However, that was unsuccessful; the reasons for this were explored in this study.

3.1.1 Selenium modification on cytidine

In one study, 5-methylselenyl-cytidine and its corresponding DNA and RNA nucleotides were synthesized to study the effect of selenium modification on cytidine²⁵. The 5-SeMe modified cytidine nucleotides were similar to the native nucleotides in terms of the crystal structure. The results showcased that selenium modification in that manner did not cause structural perturbations.

In another study, the α -modification of the 2'-SeMe-C was synthesized along with its corresponding oligonucleotides for structure and function studies²⁶⁻²⁸. To synthesize the nucleotide prior to oligonucleotide synthesis, an indirect synthesis scheme was employed, where the cytidine was derived from the conversion of its corresponding uridine derivative. This route was adopted because the direct synthesis of 2'-SeMe-C gave a very low yield. After synthesis of the modified DNAs and RNAs, the structure was determined through X-ray diffraction with MAD phasing. The modified nucleotides exhibited the same 3'-endo sugar pucker as native DNA and the substitution had no effect on the stability of the duplexes in a UV melting study. This demonstrated that the 2'-selenium functionalities were suitable for RNA and A-DNA derivatization for X-ray crystallography.

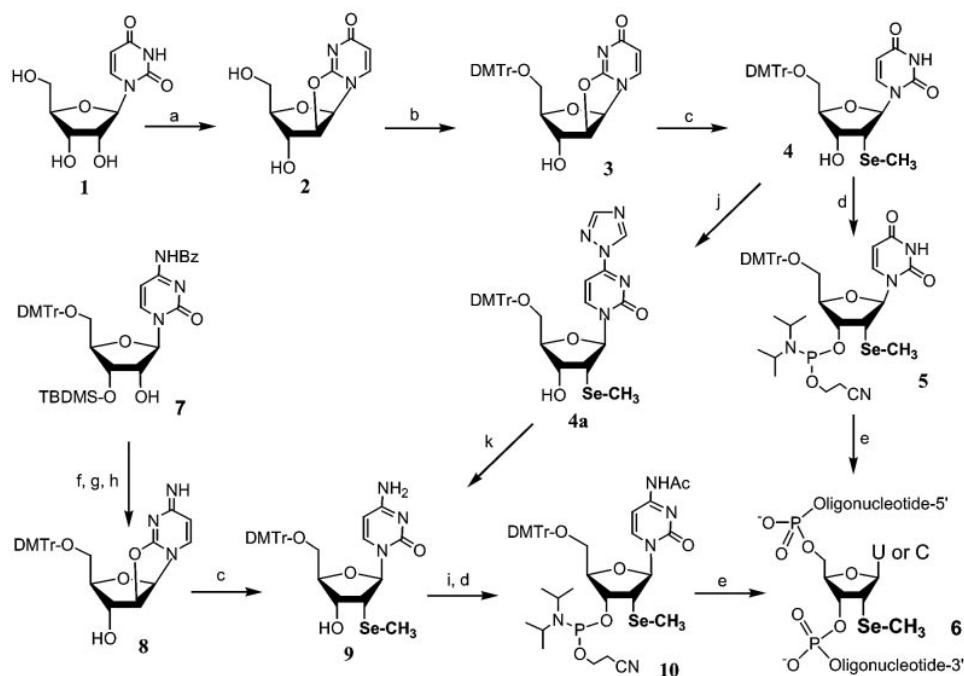


Figure 3.1: Synthetic scheme of 2'- α -MeSe-C/U²⁸

3.1.2 Problems with direct derivatization of 2'-SeMeANA-C

As part of his dissertation, Dr. Cen Chen initially attempted to synthesize 2'-SeMeANA-C with a direct synthesis scheme shown in figure 3.2

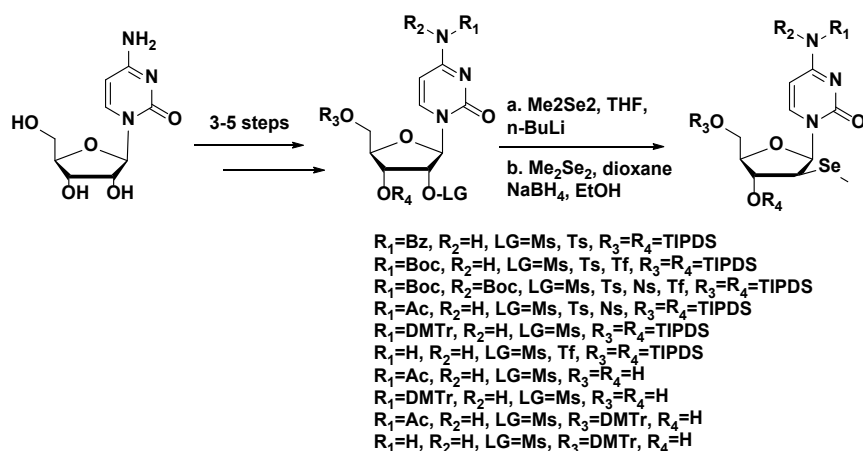


Figure 3.2: Direct incorporation of selenium into cytidine

This reaction scheme was unsuccessful. A computational study by Dr. Chen revealed that the selenium incorporation step has a high energy barrier of 3.2 kcal/mol, significantly higher than

the uridine derivative with the same leaving group. One of the two ways to ensure success of this synthesis scheme would be to conduct the reaction at high temperature with extremely stable protecting groups. However, this will create problems with deprotection in the subsequent steps. The other way to ensure success of this scheme is to use highly active leaving groups. However, that resulted in the formation of numerous side products. It was concluded that the synthesis scheme was not practical.

As an alternative, we explored the synthesis of 2'-SeMeANA-C indirectly by synthesizing its corresponding uridine derivative and converting it to cytidine after deprotection of the N3 amino group on the uridine base.

3.2 Results and Discussion

Commercially available uridine (**compound 1**) was treated with tetraisopropylidisilene (TIPDS) to protect the 3' and 5'-hydroxyl groups. The product (**compound 2**) was treated with benzyloxymethyl acetal (BOM) to protect the 3-N amino group. To activate the 2'OH group, the product (**compound 3**) was treated with mesylate. Dimethyl diselenide was dissolved in anhydrous THF, and treated with n-butyllithium solution at -78°C. This generated the Se-Me nucleophile. This solution was introduced to the activated **compound 4** at 60°C. After the incorporation of selenium, the BOM protecting group from **compound 5** dissolved in anhydrous THF was deprotected at -78°C by treatment with BBr₃. This reaction was quenched with a mixture of triethylamine (TEA) and isopropanol in a 1:1 v/v ratio. It is important that the isopropanol is anhydrous to prevent protic solvents such as methanol or water from displacing the Se-Me from the 2' position (Figure 2.5). **Compound 6** was then treated with phosphoryl chloride and triazole. This activated the 4-O on the base for the uridine to be converted to cytidine. The activated intermediate was then treated with ammonia to complete the conversion of uridine to cytidine

(**compound 10**). This compound then went through acetylation to give **compound 11**. **Compound 11** was treated with 3HF.TEA to remove the TIPDS group and generate **compound 12**. The 5'-hydroxyl group on **compound 12** was selectively protected with the DMTr protecting group. Lastly, the 2'-SeMeANA-U (**compound 13**) was converted to its phosphoramidite (**compound 14**) by treatment with N,N-diisopropylchlorophosphoramidite in the presence of dimethylamine in dry CH_2Cl_2 (Figure 3.3).

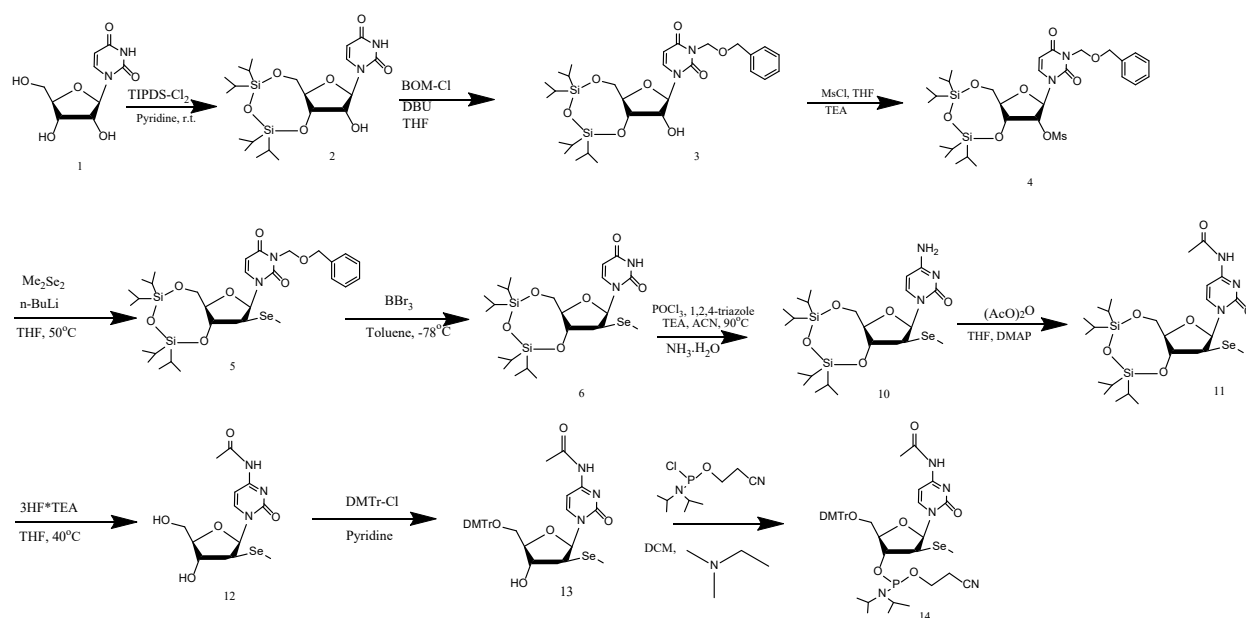


Figure 3.3: Final synthesis scheme for 2'-SeMeANA-C

The first 5 steps of this synthetic scheme is the same as the synthetic scheme for the modified uridine. After conversion of the uridine to cytidine and protection of the 4-N amino group with acetic anhydride, NMR analysis revealed a singlet peak amongst the other peaks, with a chemical shift of 2.2 ppm corresponding to the hydrogen proton of the methyl group at the protecting group of the 4-N (^1H NMR, compound 11). TIPDS deprotection removed its respective peaks from the NMR spectra (^1H NMR, compound 12). After the DMTr group was

introduced to the 5' position, a singlet was observed with chemical shifts at approximately 3.80 ppm corresponding to the hydrogen protons on the methoxy groups in the DMTr. The hydrogen protons in the DMTr aromatic rings had chemical shifts between 7.27 ppm to 7.36 ppm, and it could be seen in the NMR spectra (^1H NMR, compound 13). MS analysis was done on the phosphoramidite to confirm the final product. The detailed characterization data can be found in the experimental section while the relevant NMR and MS spectra can be found in Appendix B.2.

As previously mentioned, a direct synthesis scheme was devised for the synthesis of 2'-SeMeANA-C. Due to calculated high energy barrier that synthesis scheme was pursued. Therefore, the final product was indirectly synthesized using the already devised synthesis scheme for 2'-SeMeANA-U. After BOM deprotection for the formation of **Compound 6**, uridine was converted to cytidine. Various protecting groups were tested in a minute scale reaction before proceeding with the reaction at a gram scale. If the reaction was conducted at a temperature that was too high, it would disintegrate the starting material before the formation of the final product, but if the temperature was not high enough, the reaction would not even take place. The reaction time was vitally important because if the reaction was prolonged, then many by-products would form.

The amino group at N4 in **compound 10** is a strong nucleophile and it could hamper with the solid-phase oligonucleotide synthesis after the phosphoramidite was synthesized. Therefore, it was necessary to subject **compound 10** to acetylation to protect the N4 amino group. The subsequent steps were similar to the synthesis scheme showcased in figure 2.4.

3.3 Conclusion

In conclusion, a successful synthetic scheme was devised for the formation of 2'-SeMeANA-C (Figure 3.3). This scheme resulted in the successful synthesis of the 2'-SeM-arabino

cytidine phosphoramidite. The phosphoramidite product generated from the synthesis was utilized by Dr. Cen Chen to carry out the synthesis of several DNA oligonucleotides for structure and function studies. The synthetic schemes for the phosphoramidites of all the nucleotides, particularly cytidine and uridine, are very similar. However, structure and function studies of their respective oligonucleotides should reveal their uniqueness and set them apart from one another. Following the synthetic scheme in Figure 3.3, several oligonucleotides can be synthesized and each of them can be studied for potential application in SBDD and therapeutics.

3.4 Experimental section

3.4.1 General

Most solvents and reagents were purchased from Sigma, Fluka, or Aldrich (PA) and used without purification, unless otherwise specified. Solid reagents were dried under high vacuum when it was necessary to do so. Reactions with compounds sensitive to air or moisture were performed under argon purging. Solvent mixtures are indicated as volume/volume ratios. Thin layer chromatography (TLC) was done on Merck 60 F254 plates (0.25 mm thick). TLC spots were visualized under UV light. Column purification was performed using Fluka silica gel 60 (mesh size 0.040-0.063 mm) using a silica gel, crude compound weight ratio of ca. 30:1. ¹H spectra were recorded using Bruker-300 or 400 (300 or 400 MHz). All chemical shifts (δ) are in ppm relative to tetramethylsilane and all coupling constants (J) are in Hz. High resolution (HR) MS were either obtained with electrospray ionization (ESI) on a Q-TOFTM Waters Micromass at Georgia State University.

3.4.2 Synthetic steps and characterization

The first 5 steps for this reaction are very similar to the first 5 steps for the synthesis scheme of 2'-SeMeANA-U. Therefore, only the synthesis steps from step 6 have been shown here.

3.4.2.1 Conversion of uridine to cytidine

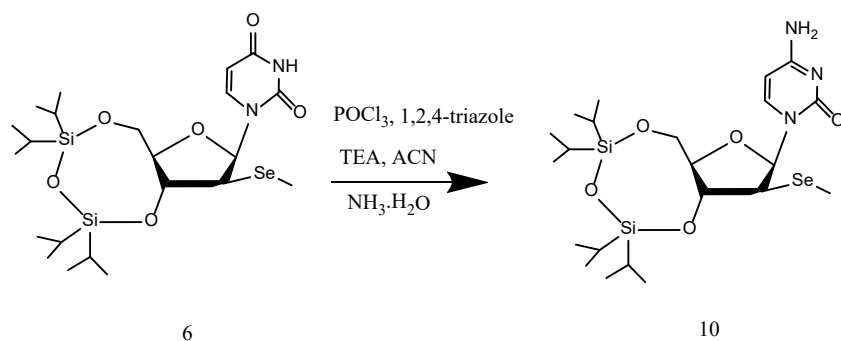


Figure 3.4: Conversion of base from U to C

3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'-methylseleno-arabinocytidine

(compound 10): For this step, phosphorus oxychloride (1.2 mL, 13.2 mmol) was added to an argon purged solution of 1,2,4-triazole (3.6 g, 52.8 mmol) and anhydrous acetonitrile (30 mL). The reaction was stirred at room temperature for 1 h. After 1 h, anhydrous triethylamine (15 mL, 105.6 mmol) was added, and the reaction was stirred for another hour. The reaction was then transferred into the round-bottomed flask containing the starting material **compound 6** (2.5 g, 4.4 mmol) dissolved in 10 mL anhydrous acetonitrile purged with argon. The reaction was then heated to 90°C and stirred overnight. The solution was cooled to 50°C and then treated with $\text{NH}_3\cdot\text{H}_2\text{O}$ (10 mL, 15 M). The solution was stirred for 2 h. TLC was done with 4% methanol in dichloromethane confirmed the completion of the reaction. Afterwards, the reaction mixture was evaporated reduced under pressure to approximately 15 mL. The crude product was extracted with ethyl acetate (3 x 100 mL). The combined organic layer was washed with 100 mL brine and dried over anhydrous MgSO_4 before evaporation. The crude product was subjected to silica gel chromatography with 2% methanol in dichloromethane as the eluent, which gave us the desired product (1.9g, 75%). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ (ppm): 1.01-1.11 (m, 28H, 4*×*iPrSi), 2.03 (s, 3H, SeCH_3), 3.70 (dd, $J_1=6.8$ Hz, $J_2=8.1$ Hz, 1H, H-2'), 3.76 (dt, $J_1=7.2$ Hz, $J_2=3.3$ Hz, 1H, H-4'),

4.05 (m, 2H, H-5'), 4.25 (t, $J=7.7$ Hz, 1H, H-3'), 5.74 (d, $J=7.4$ Hz, 1H, H-5), 6.44, (d, $J=6.6$ Hz, 1H, H-1'), 7.58 (d, $J=7.4$, 1H, H-6). HRMS (ESI): $C_{22}H_{41}N_3O_5SeSi_2$; $[M+Na]^+$: 586.2769 (calc. 586.1648).

3.4.2.2 Protection of the 4-N amino group with acetic anhydride

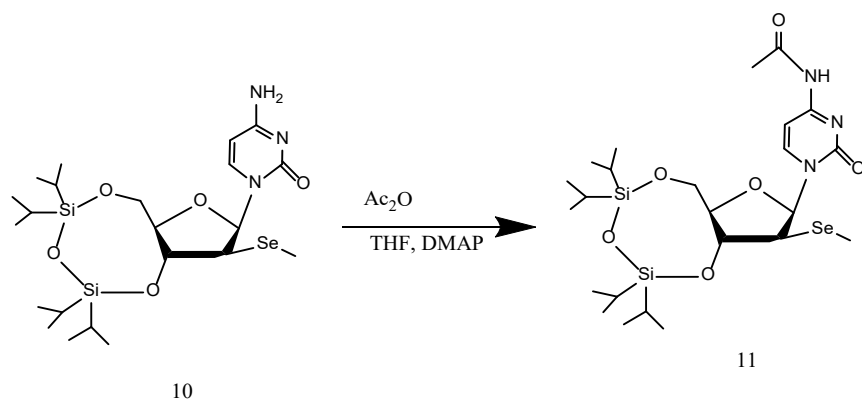


Figure 3.5: Protection of the 4-N amino group on the cytidine base

N⁴-Acetyl-3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'-methylseleno-arabinocytidine (Compound 11): The starting material **Compound 10** (1.3 g, 2.3 mmol) was dissolved in 20 mL tetrahydrofuran in a pressure vessel. Then triethylamine (2.6 mL, 18.4 mmol) was added. The reaction mixture was then treated with acetic anhydride (1.3 mL, 13.8 mmol) and N,N-dimethylaminopyridine (DMAP, 281 mg, 2.3 mmol). The reaction was then heated to 85°C and left to stir overnight. TLC was done with 4% methanol in dichloromethane as the eluent. About 2 mL of MeOH was added and the mixture was stirred for 20 min to consume excess acetic anhydride. The reaction mixture was evaporated under reduced pressure. The crude product was re-dissolved in 100 mL ethyl acetate. The organic solution was washed with 30 mL water and then 30 mL of brine. The organic layer and water layer were separated. The organic layer was dried over anhydrous MgSO₄ before evaporation. The crude product underwent silica gel chromatography with 1% methanol in dichloromethane as the eluent to give the desired product

(1.3 g, 95%). $^1\text{H-NMR}$ (400 MHz, CDCl_3) δ (ppm): 1.05-1.13 (m, 28H, $4\times\text{iPrSi}$), 2.07 (s, 3H, SeCH_3), 2.27 (s, 3H, CH_3), 3.73 (dd, $J_1=6.7$ Hz, $J_2=8.7$ Hz, 1H, H-2'), 3.76 (dt, $J_1=7.5$ Hz, $J_2=3.0$ Hz, 1H, H-4'), 4.09 (d, $J=3.0$, 2H, H-5'), 4.24 (t, $J=8.1$ Hz, 1H, H-3'), 6.47, (d, $J=6.6$ Hz, 1H, H-1'), 7.45 (d, $J=7.5$ Hz, 1H, H-5), 7.97 (d, $J=7.5$, 1H, H-6), 10.27 (br, 1H, NH). HRMS (ESI): $\text{C}_{24}\text{H}_{43}\text{N}_3\text{O}_6\text{SeSi}_2$; $[\text{M}+\text{H}]^+$: 606.1866 (calc. 605.1934).

3.4.2.3 Tetraisopropylidisiloxanylidene (TIPDS) deprotection

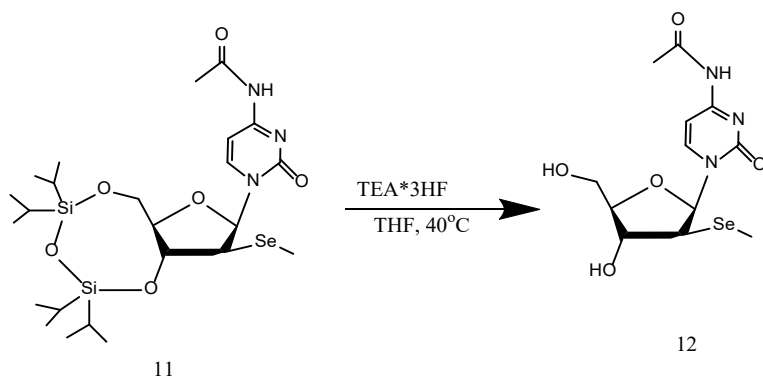


Figure 3.6: TIPDS deprotection from the 3' and 5' positions (C)

N⁴-Acetyl-3',5'-O-(1,1,3,3-tetraisopropylidisiloxane-1,3-diyl)-2'-methylseleno-

arabinocytidine (compound 12): Compound 11 (1 g, 1.7 mmol) was first dissolved in 10 mL anhydrous THF. The solution was then treated with 0.26 mL (1.7 mmol) of triethylamine trihydrofluoride ($3\text{HF}\cdot\text{Et}_3\text{N}$). After addition of the $3\text{HF}\cdot\text{Et}_3\text{N}$ the solution was heated to 40°C and stirred for 2 hours. TLC was done with 10% methanol in dichloromethane as the eluent. The solvent was evaporated to dryness. The crude product was then subjected to silica gel chromatography with 5% methanol in dichloromethane to obtain the pure product as a white solid (609 mg, 99% yield).

3.4.2.4 Introduction of the dimethoxytrityl group (DMTr-) at the 5' position

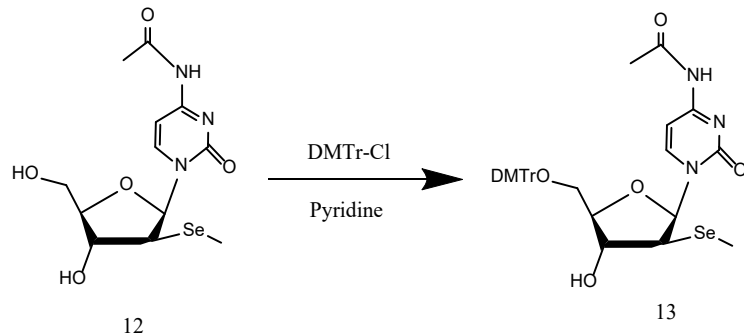


Figure 3.7: Tritylation of the 5' OH group (C)

N⁴-Acetyl-5'-O-(4,4'-dimethoxytrityl)-2'-methylseleno-arabinocytidine (compound 13): Compound **12** (420 mg, 1.2 mmol) was dried over high vacuum and co-evaporated with anhydrous pyridine (2 x 10 mL). Then the starting material was dissolved in 15 mL anhydrous pyridine. The reaction was then cooled in an ice-bath. While the reaction was cool, it was then treated with 432 mg (1.3 mmol) of the dimethoxytrityl chloride. The reaction was then taken off the ice bath and allowed to reach room temperature. The reaction was stirred for 2 h. Thin layer chromatography was done with 50% ethyl acetate in hexane as the eluent. The solvent was evaporated under reduced pressure. The concentrated product was dissolved in 30 mL methylene chloride, then washed with 20 mL water and 30 mL brine. The aqueous layer and organic layer were separated from each other. The organic layer was dried over anhydrous MgSO₄, filtered and then evaporated to dryness. The crude product underwent silica gel chromatography with 3% MeOH in methylene chloride with 1% Et₃N as the eluent. The final product for this step was a white solid with 90% yield (718mg). ¹H-NMR (400 MHz, CDCl₃) δ (ppm): 1.97 (s, 3H, SeCH₃), 2.25 (s, 3H, CH₃), 3.50 (dd, *J*₁=4.4 Hz, *J*₂=10.7 Hz, 1H, Ha-5'), 3.57 (dd, *J*₁=3.4 Hz, *J*₂=10.7 Hz, 1H, Hb-5'), 3.83 (s, 6H, OCH₃), 3.94 (t, *J*=6.4 Hz, 1H, H-2'), 4.02-4.05 (m, 1H, H-4'), 4.28 (t, *J*=6.2, 1H, H-3'), 6.54, (d, *J*=6.2 Hz, 1H, H-1'), 6.88 (d, *J*=8.8, 4H, Ph), 7.27-7.36 (m, 8H, Ph, H-

5), 7.45 (d, $J=7.4$ Hz, 2H, Ph), 8.16 (d, $J=7.5$, 1H, H-6), 9.32 (br, 1H, NH). HRMS (ESI): $C_{33}H_{35}N_3O_7Se$; $[M+Na]^+$: 688.1511 (calc. 688.1538).

3.4.2.5 Phosphoramidite addition for solid-phase oligonucleotide synthesis

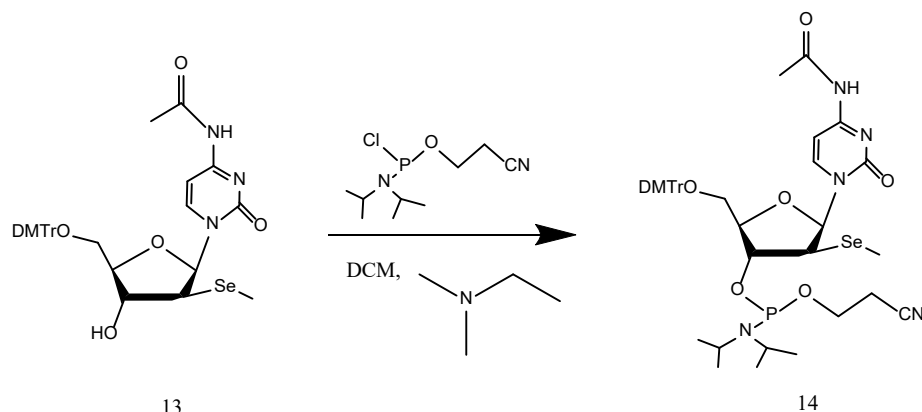


Figure 3.8: Phosphoramidite synthesis in preparation for solid-phase oligonucleotide synthesis (C)

N⁴-Acetyl-3'-O-(2-cyanoethyl-N,N-diisopropylphosphoramidite)-5'-O-(4,4'-dimethoxytrityl)-2'-methylseleno-arabinocytidine (compound 14): Compound 13 (240 mg, 0.4 mmol) was dried under high vacuum and dissolved in 2 mL anhydrous methylene chloride (CH_2Cl_2). To the solution, dimethylethanolamine (0.2 mL, 2.4 mmol) was injected. Lastly, the phosphoramidite reagent 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite (80 μ L, 0.4 mmol) was added to the reaction and stirred for about 1 h under dry argon. The reaction mixture was evaporated under reduced pressure. The concentrated product was then dissolved in 2 mL CH_2Cl_2 . The reaction mixture was then precipitated under vigorous stirring in a 400 mL hexane solution (400 mL). After carefully decanting the hexane solution, the crude product was purified by silica gel chromatography (30% ethyl acetate in CH_2Cl_2 containing 1% dimethylethylamine). The purified product was precipitated again in the 400 mL hexane solution. The precipitate was then

dissolved in 2 mL CH_2Cl_2 . The solvent was then evaporated. The derived product (260 mg, 75%) was a white foam. HRMS (ESI): $\text{C}_{42}\text{H}_{52}\text{N}_5\text{O}_8\text{PSe}$; $[\text{M}+\text{Na}]^+$: 887.8305 (calc. 887.8285).

REFERENCES

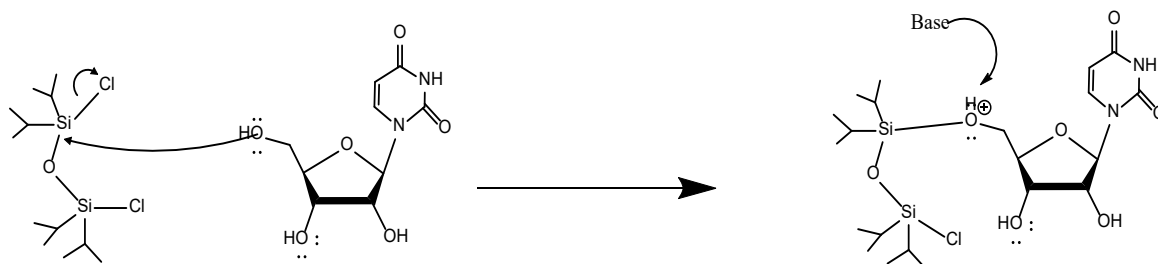
1. Koonin, E. V., Does the central dogma still stand? *Biology Direct* **2012**, 7 (1), 27.
2. Lin, L.; Sheng, J.; Huang, Z., Nucleic acid X-ray crystallography via direct selenium derivatization. *Chemical Society Reviews* **2011**, 40 (9), 4591-4602.
3. Campbell, N. H.; Parkinson, G. N., Crystallographic studies of quadruplex nucleic acids. *Methods (San Diego, Calif.)* **2007**, 43 (4), 252-63.
4. Carrasco, N.; Ginsburg, D.; Du, Q.; Huang, Z., Synthesis of selenium-derivatized nucleosides and oligonucleotides for X-ray crystallography. *Nucleosides, nucleotides & nucleic acids* **2001**, 20 (9), 1723-34.
5. Du, Q.; Carrasco, N.; Teplova, M.; Wilds, C. J.; Egli, M.; Huang, Z., Internal derivatization of oligonucleotides with selenium for X-ray crystallography using MAD. *Journal of the American Chemical Society* **2002**, 124 (1), 24-5.
6. Teplova, M.; Wilds, C. J.; Wawrzak, Z.; Tereshko, V.; Du, Q.; Carrasco, N.; Huang, Z.; Egli, M., Covalent incorporation of selenium into oligonucleotides for X-ray crystal structure determination via MAD: proof of principle. Multiwavelength anomalous dispersion. *Biochimie* **2002**, 84 (9), 849-58.
7. Olieric, V.; Rieder, U.; Lang, K.; Serganov, A.; Schulze-Briese, C.; Micura, R.; Dumas, P.; Ennifar, E., A fast selenium derivatization strategy for crystallization and phasing of RNA structures. *RNA* **2009**, 15 (4), 707-715.
8. Salon, J.; Sheng, J.; Gan, J.; Huang, Z., Synthesis and Crystal Structure of 2'-Se-Modified Guanosine Containing DNA. *The Journal of Organic Chemistry* **2010**, 75 (3), 637-641.
9. Sheng, J.; Salon, J.; Gan, J.; Huang, Z., Synthesis and crystal structure study of 2'-Se-adenosine-derivatized DNA. *Science China Chemistry* **2010**, 53 (1), 78-85.
10. Hassan, A. E. A.; Sheng, J.; Zhang, W.; Huang, Z., High Fidelity of Base Pairing by 2-Selenothymidine in DNA. *Journal of the American Chemical Society* **2010**, 132 (7), 2120-2121.
11. Hassan, A. E. A.; Sheng, J.; Jiang, J.; Zhang, W.; Huang, Z., Synthesis and Crystallographic Analysis of 5-Se-Thymidine DNAs. *Organic Letters* **2009**, 11 (12), 2503-2506.
12. Jiang, J.; Sheng, J.; Carrasco, N.; Huang, Z., Selenium derivatization of nucleic acids for crystallography. *Nucleic Acids Res* **2007**, 35 (2), 477-485.
13. Moroder, H.; Kreutz, C.; Lang, K.; Serganov, A.; Micura, R., Synthesis, Oxidation Behavior, Crystallization and Structure of 2'-Methylseleno Guanosine Containing RNAs. *Journal of the American Chemical Society* **2006**, 128 (30), 9909-9918.
14. Salon, J.; Sheng, J.; Jiang, J.; Chen, G.; Caton-Williams, J.; Huang, Z., Oxygen Replacement with Selenium at the Thymidine 4-Position for the Se Base Pairing and Crystal Structure Studies. *Journal of the American Chemical Society* **2007**, 129 (16), 4862-4863.
15. Sheng, J.; Jiang, J.; Salon, J.; Huang, Z., Synthesis of a 2'-Se-thymidine phosphoramidite and its incorporation into oligonucleotides for crystal structure study. *Organic letters* **2007**, 9 (5), 749-752.
16. Thompson, R. A.; Spring, A. M.; Sheng, J.; Huang, Z.; Germann, M. W., The importance of fitting in: conformational preference of selenium 2' modifications in nucleosides and helical structures. *J Biomol Struct Dyn* **2015**, 33 (2), 289-297.
17. Lin, L.; Caton-Williams, J.; Kaur, M.; Patino, A. M.; Sheng, J.; Punetha, J.; Huang, Z., Facile synthesis of nucleoside 5'-(alpha-P-seleno)-triphosphates and phosphoroselenoate RNA transcription. *RNA* **2011**, 17 (10), 1932-8.
18. Salon, J.; Gan, J.; Abdur, R.; Liu, H.; Huang, Z., Synthesis of 6-Se-guanosine RNAs for structural study. *Organic letters* **2013**, 15 (15), 3934-3937.

19. Stein, C. A.; Castanotto, D., FDA-Approved Oligonucleotide Therapies in 2017. *Molecular therapy : the journal of the American Society of Gene Therapy* **2017**, 25 (5), 1069-1075.
20. Sharma, V. K.; Sharma, R. K.; Singh, S. K., Antisense oligonucleotides: modifications and clinical trials. *MedChemComm* **2014**, 5 (10), 1454-1471.
21. Carrasco, N.; Huang, Z., Enzymatic Synthesis of Phosphoroselenoate DNA Using Thymidine 5'-(α -P-seleno)triphosphate and DNA Polymerase for X-ray Crystallography via MAD. *Journal of the American Chemical Society* **2004**, 126 (2), 448-449.
22. Brandt, G.; Carrasco, N.; Huang, Z., Efficient substrate cleavage catalyzed by hammerhead ribozymes derivatized with selenium for X-ray crystallography. *Biochemistry* **2006**, 45 (29), 8972-8977.
23. Lin, L.; Sheng, J.; Momin, R. K.; Du, Q.; Huang, Z., Facile synthesis and anti-tumor cell activity of Se-containing nucleosides. *Nucleosides, nucleotides & nucleic acids* **2009**, 28 (1), 56-66.
24. Bobkov, G. V.; Mikhailov, S. N.; Van Aerschot, A.; Herdewijn, P., Phosphoramidite building blocks for efficient incorporation of 2'-O-aminoethoxy(and propoxy)methyl nucleosides into oligonucleotides. *Tetrahedron* **2008**, 64 (27), 6238-6251.
25. Zhang, W.; Sheng, J.; Hassan, A. E.; Huang, Z., Synthesis of 2'-deoxy-5-(methylselenyl)cytidine and Se-DNAs for structural and functional studies. *Chemistry, an Asian journal* **2012**, 7 (3), 476-9.
26. Carrasco, N.; Buzin, Y.; Tyson, E.; Halpert, E.; Huang, Z., Selenium derivatization and crystallization of DNA and RNA oligonucleotides for X-ray crystallography using multiple anomalous dispersion. *Nucleic Acids Res* **2004**, 32 (5), 1638-46.
27. Hobartner, C.; Micura, R., Chemical synthesis of selenium-modified oligoribonucleotides and their enzymatic ligation leading to an U6 SnRNA stem-loop segment. *Journal of the American Chemical Society* **2004**, 126 (4), 1141-9.
28. Buzin, Y.; Carrasco, N.; Huang, Z., Synthesis of Selenium-Derivatized Cytidine and Oligonucleotides for X-ray Crystallography Using MAD. *Organic Letters* **2004**, 6 (7), 1099-1102.

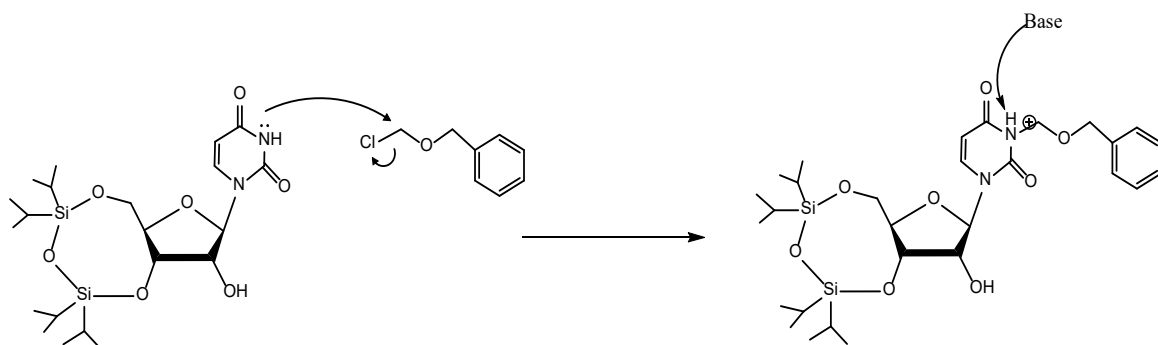
APPENDICES

Appendix A. Mechanism for Key Reactions

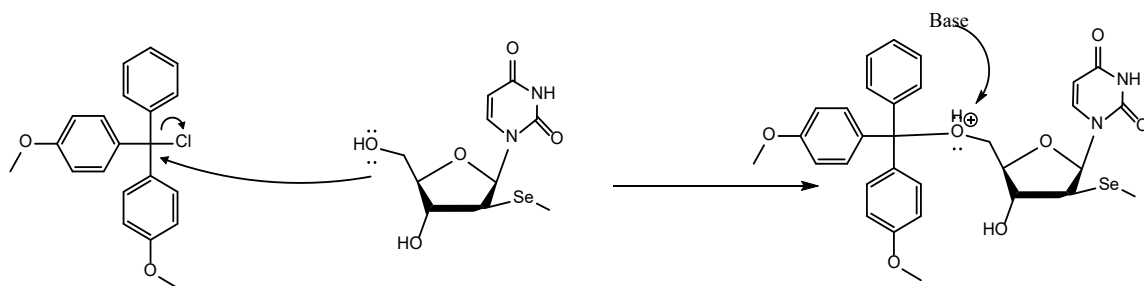
a. 3', 5' – TIPDS protection:



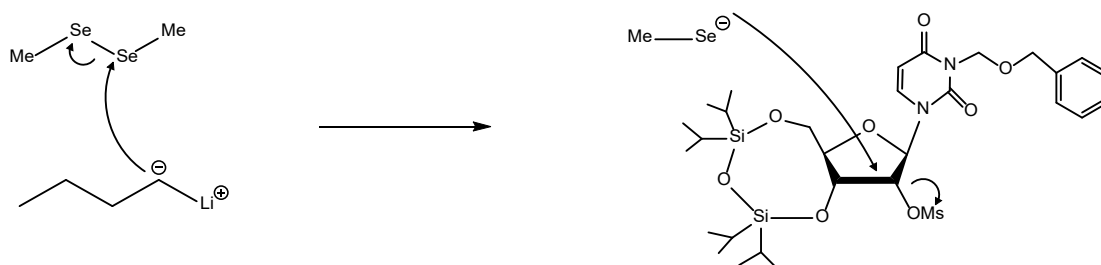
b. BOM protection:

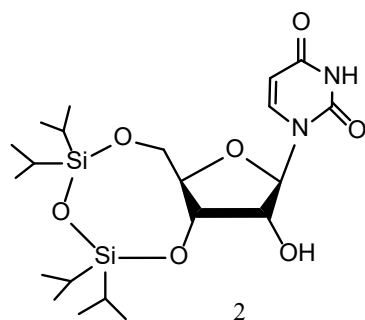


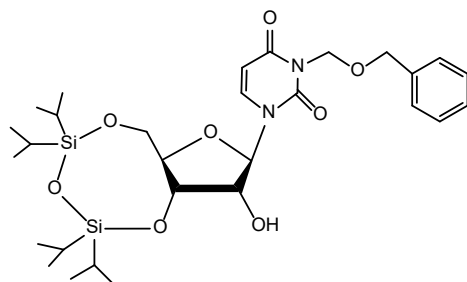
c. 5'-DMTr protection:



d. 2' – MeSe introduction:

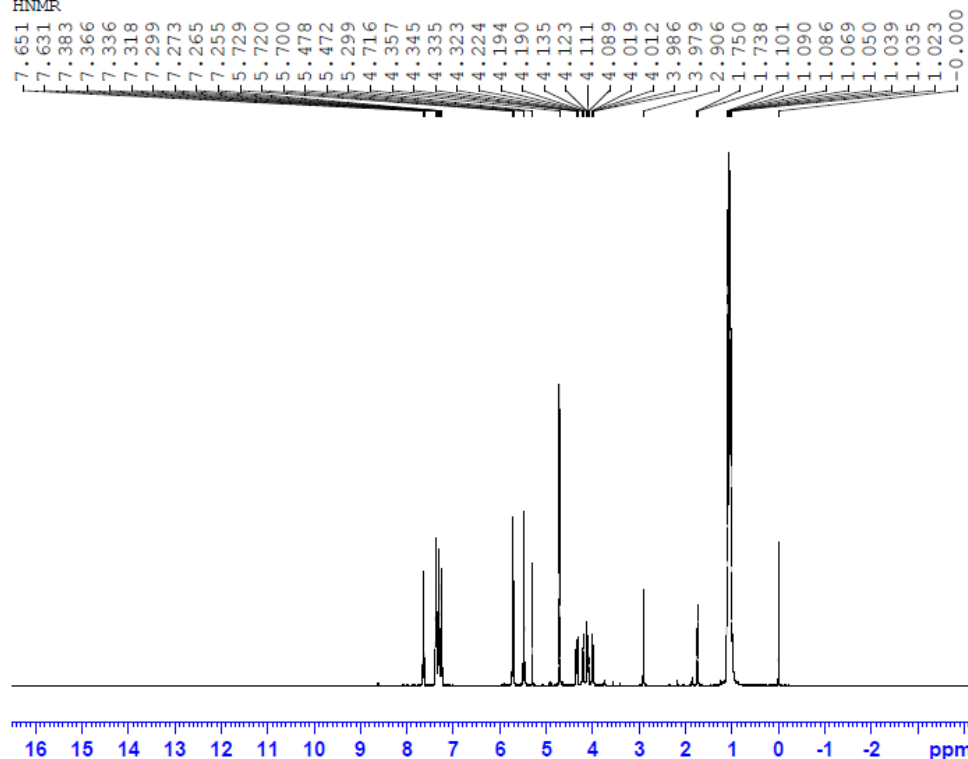


Appendix B. Characterization data for key compounds***Appendix B.1. NMR and MS Spectra for 2'-SeMeANA-U***



3

3N-BOM-3',5'-TIPDS-U
CDCl₃
HNMR



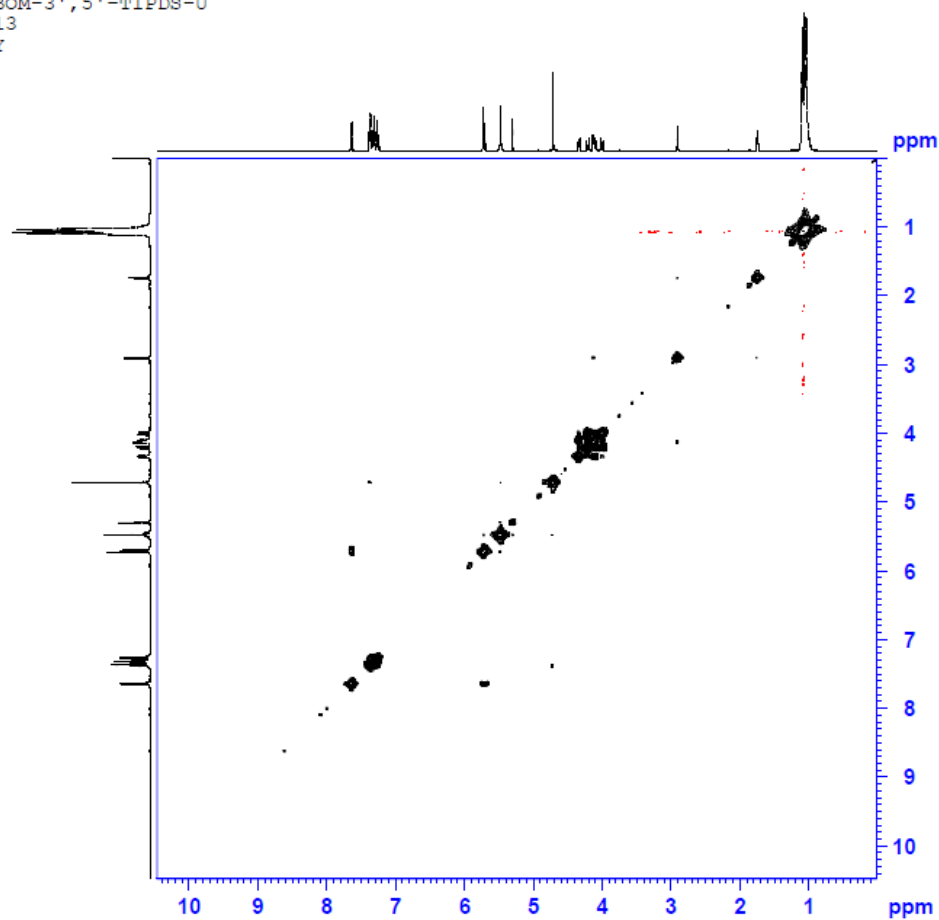
Current Data Parameters
NAME 20180708_BOMT1-U
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20180708
Time 10.49
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 65536
SOLVENT CDCl₃
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.126314 Hz
AQ 3.9583748 sec
RG 71.8
DW 60.400 usec
DE 6.50 usec
TE 295.3 K
D1 1.00000000 sec
TD0 1

===== CHANNEL f1 =====
NUC1 1H
P1 11.00 usec
PL1 0 dB
SFO1 400.1324710 MHz

F2 - Processing parameters
SI 32768
SF 400.1300078 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

3N-BOM-3',5'-TIPDS-U
CDCl₃
COSY



Current Data Parameters
NAME 20180708_BOM3'-U
EXNO 4
PROCNO 1

F2 - Acquisition Parameters
Date_ 20180708
Time_ 11.18
INSTRUM spect
PROBHD 5 mm FASBO BB-
PULPROG cosyzgpgf
TD 65536
SOLVENT CDCl₃
NS 1
DS 8
SWH 4789.272 Hz
FIDRES 0.2328512 Hz
AQ 0.21381112 sec
RG 114
LW 104.400 usec
TE 6.50 usec
TE 285.3 K
DO 0.00000300 sec
D1 1.48689198 sec
D12 0.00009400 sec
D16 0.00020000 sec
IN0 0.00020825 sec

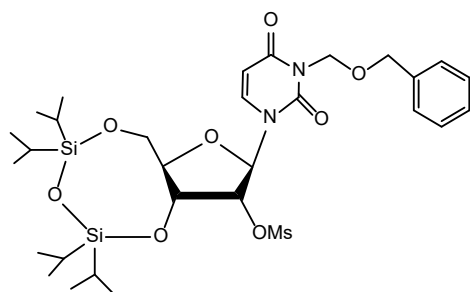
===== CHANNEL f1 =====
NUC1 1H
PC 11.00 usec
P1 11.00 usec
PL1 0 dB
SFO1 400.1318006 MHz

===== GRADIENT CHANNEL =====
GPHAM(1) SINE.100
GPE1 10.00 %
P16 1000.00 usec

F1 - Acquisition parameters
TD 128
SFO1 400.1318 MHz
FIDRES 78.024712 Hz
SW 12.000 ppm
F1MODE QF

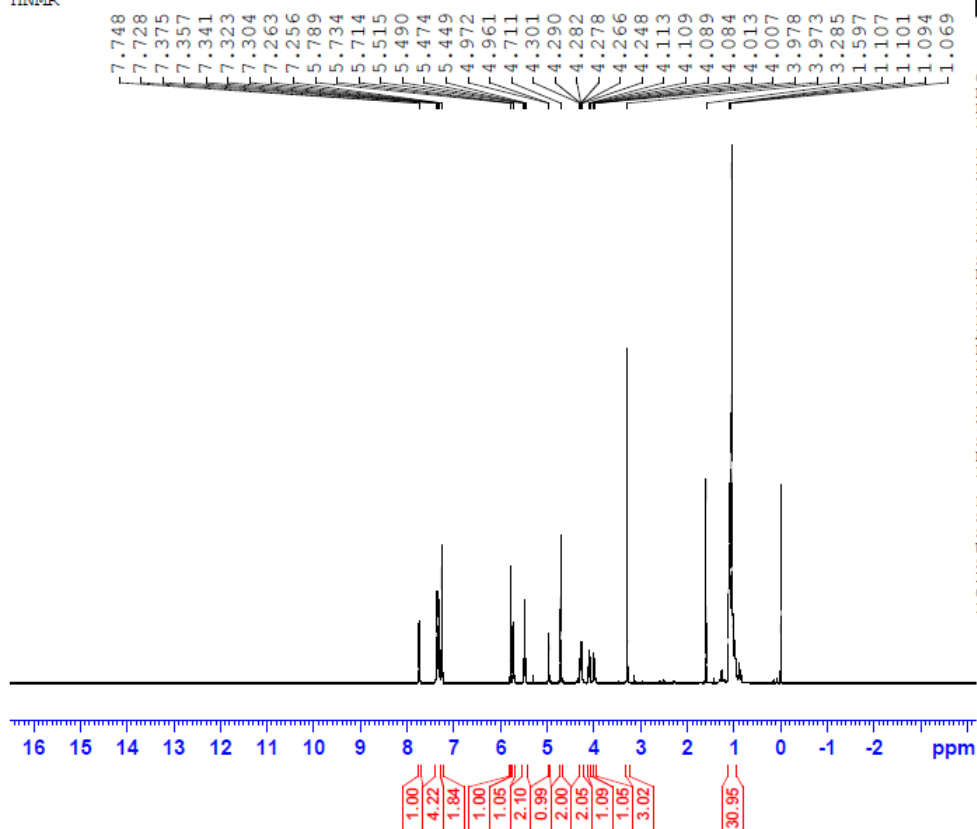
F2 - Processing parameters
SI 1024
SF 400.1300078 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0
PC 1.40

F1 - Processing parameters
SI 1024
MC2 QF
SF 400.1300078 MHz
WDW SINE
SSB 0
LB 0 Hz
GB 0



4

3N-BOM-3',5'-TIPDS-2'-Ms-U
CDCl₃
HNMR

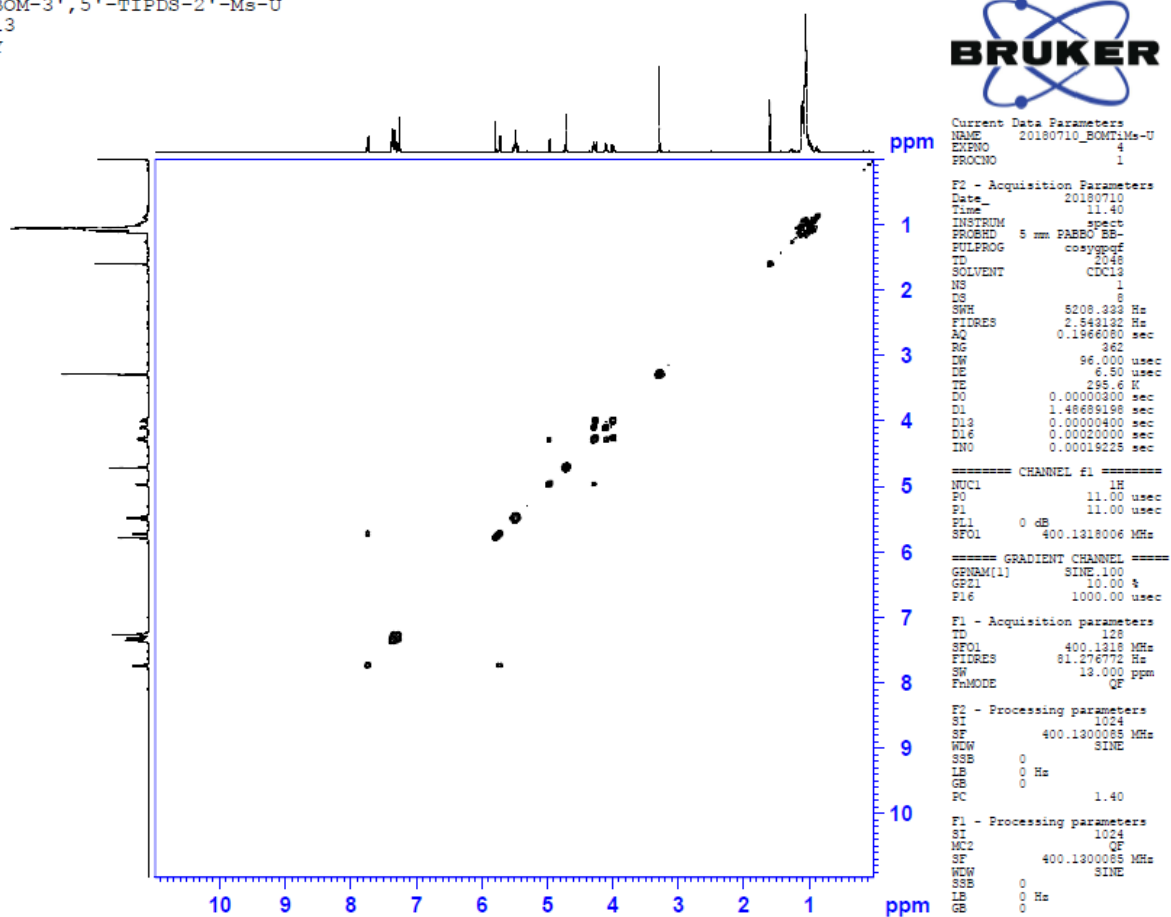


Current Data Parameters
NAME 20180710_BOMTiMs-U
EXPNO 1
PROCNO 1

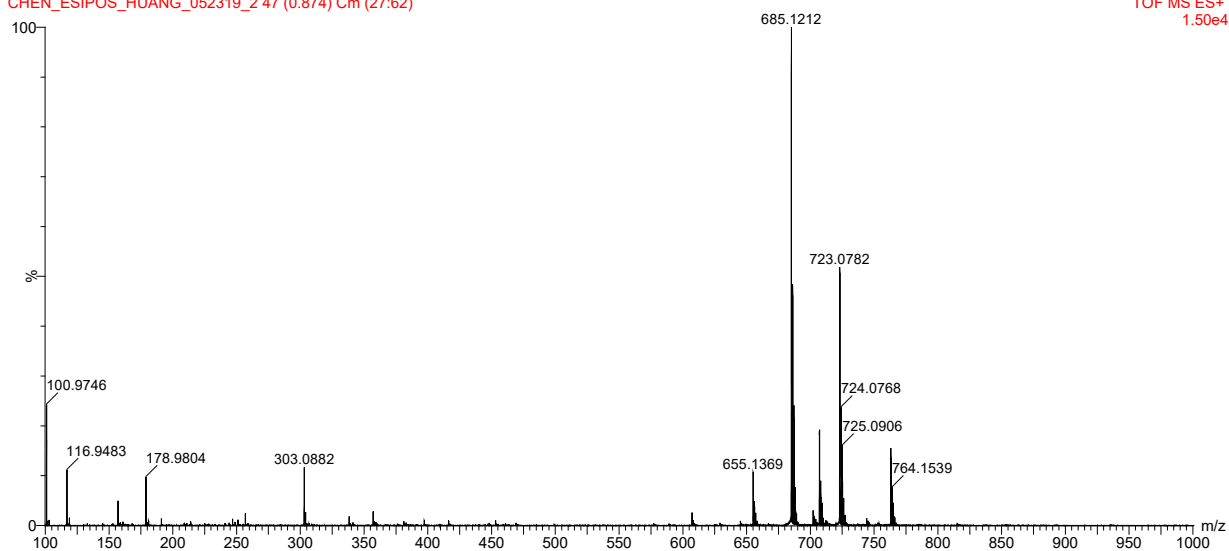
F2 - Acquisition Parameters
Date_ 20180710
Time_ 10.27 h
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 65536
SOLVENT CDCl₃
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.252629 Hz
AQ 3.9583746 sec
RG 181
DW 60.400 usec
DE 6.50 usec
TE 295.4 K
D1 1.00000000 sec
TD0 1
SF01 400.1324710 MHz
NUC1 1H
P1 11.00 usec

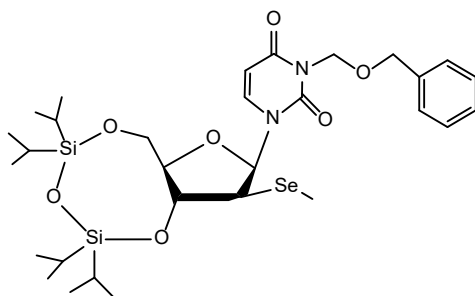
F2 - Processing parameters
SI 32768
SF 400.1300085 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

3N-BOM-3',5'-TIPDS-2'-Ms-U
CDCl₃
COSY



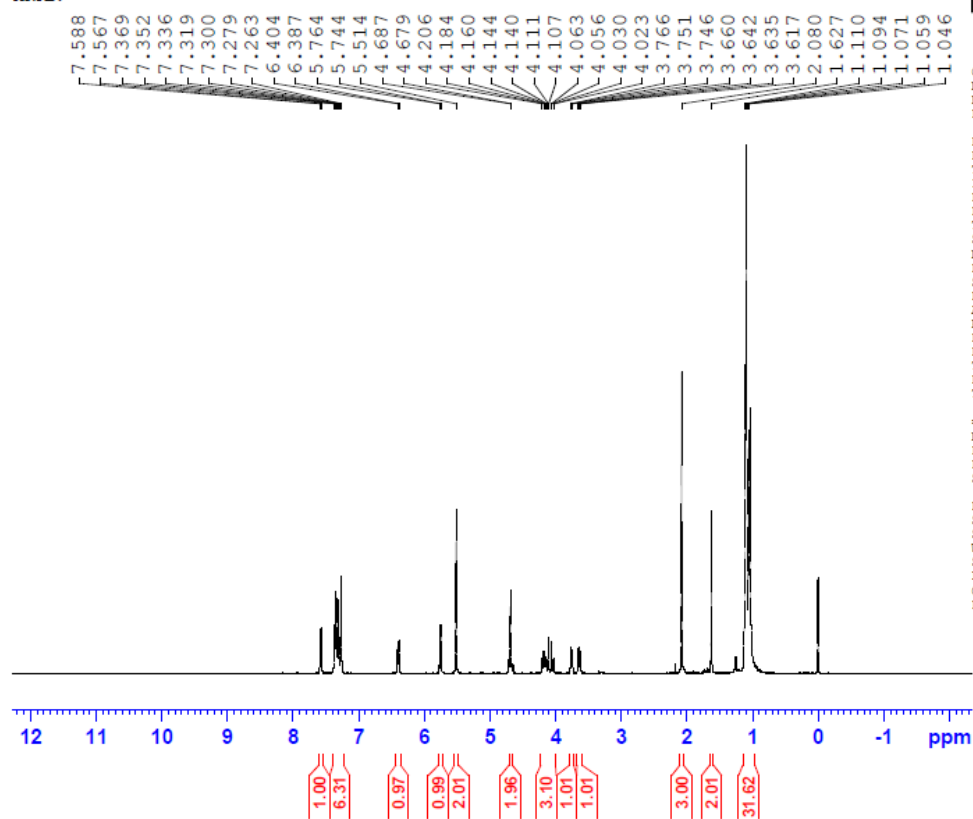
CHEN_ESIPOS_HUANG_052319_2 47 (0.874) Cm (27:62)





5

3N-BOM-3',5'-TIPDS-2'-beta-MeSe-U
CDCl₃
HNMR



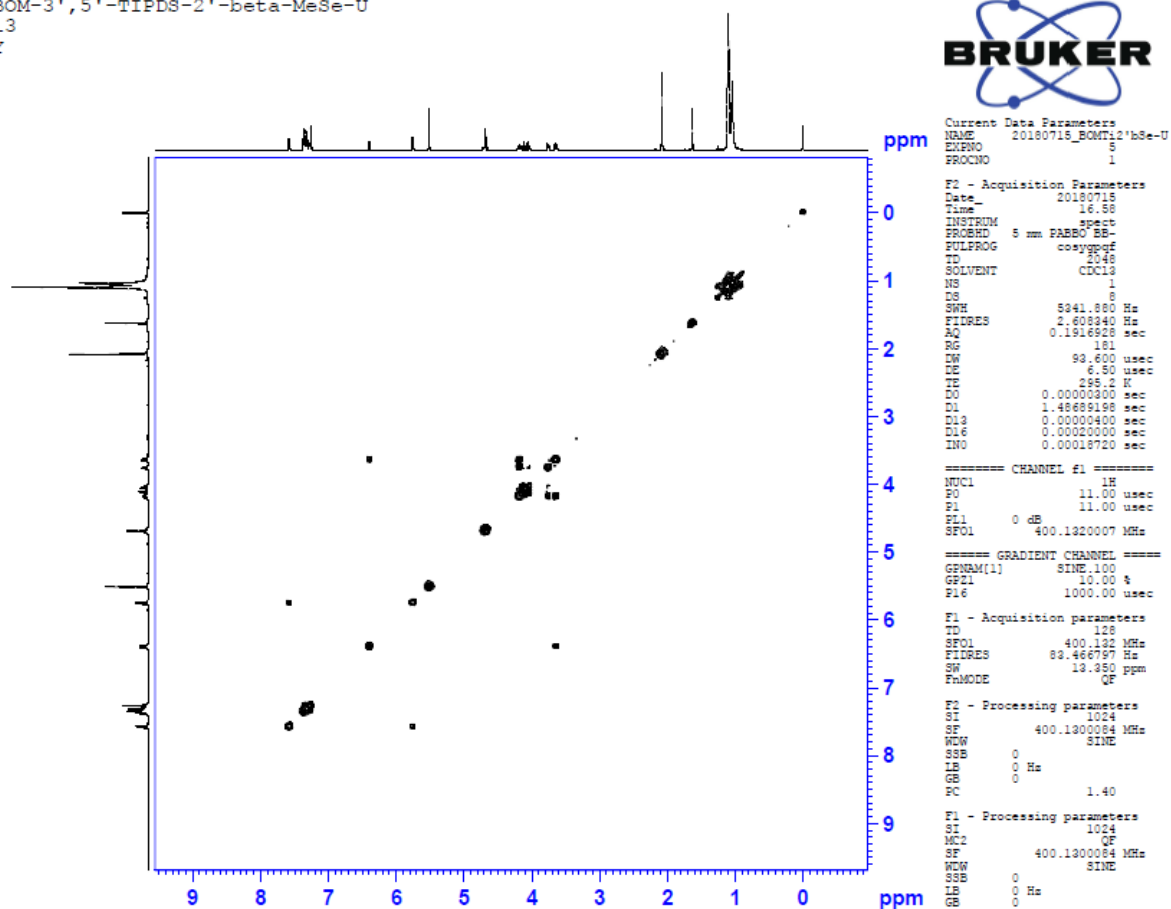
Current Data Parameters
NAME 20180715_BOMT12'bSe-U
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20180715
Time 18.36
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zgpg30
TD 65536
SOLVENT CDCl₃
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.126814 Hz
AQ 3.9883745 sec
RG 90.5
DW 60.400 usec
DE 6.50 usec
TE 298.2 K
D1 1.00000000 sec
TDO 1

===== CHANNEL f1 =====
NUC1 1H
P1 11.00 usec
PL1 0 dB
SFO1 400.1324710 MHz

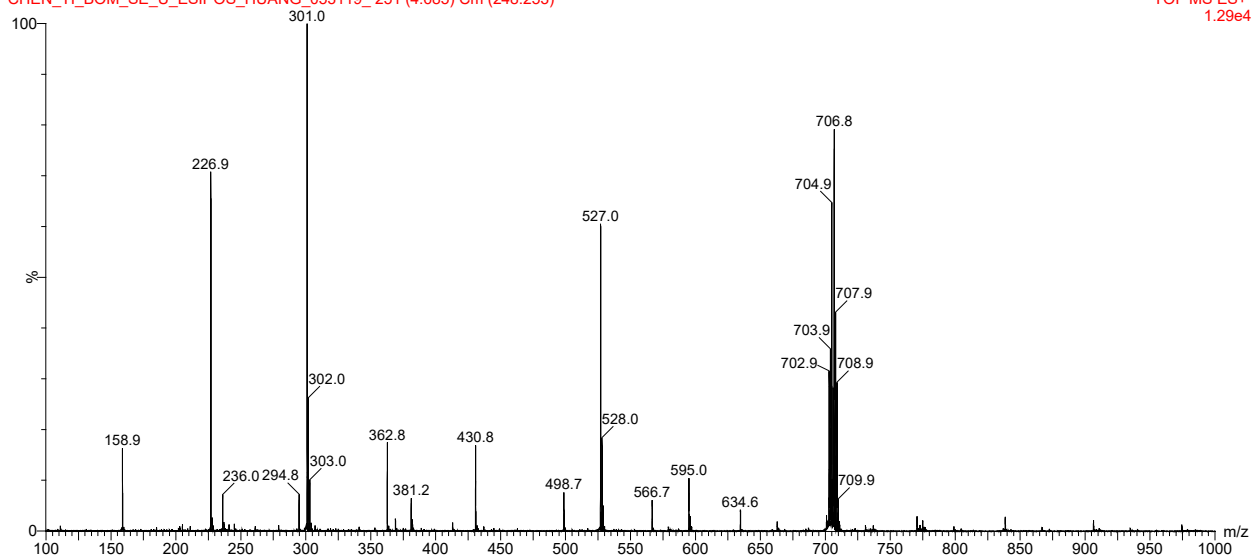
F2 - Processing parameters
SI 32768
SF 400.1300084 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

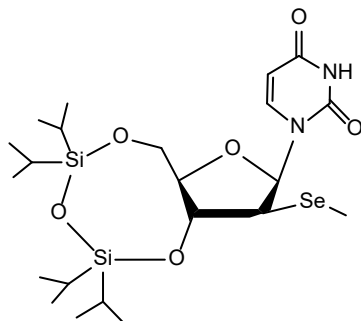
3N-BOM-3',5'-TIPDS-2'-beta-MeSe-U
CDCl₃
COSY



CHEN_TI_BOM_SE_U_ESIPOS_HUANG_053119_251 (4.685) Cm (248:253)

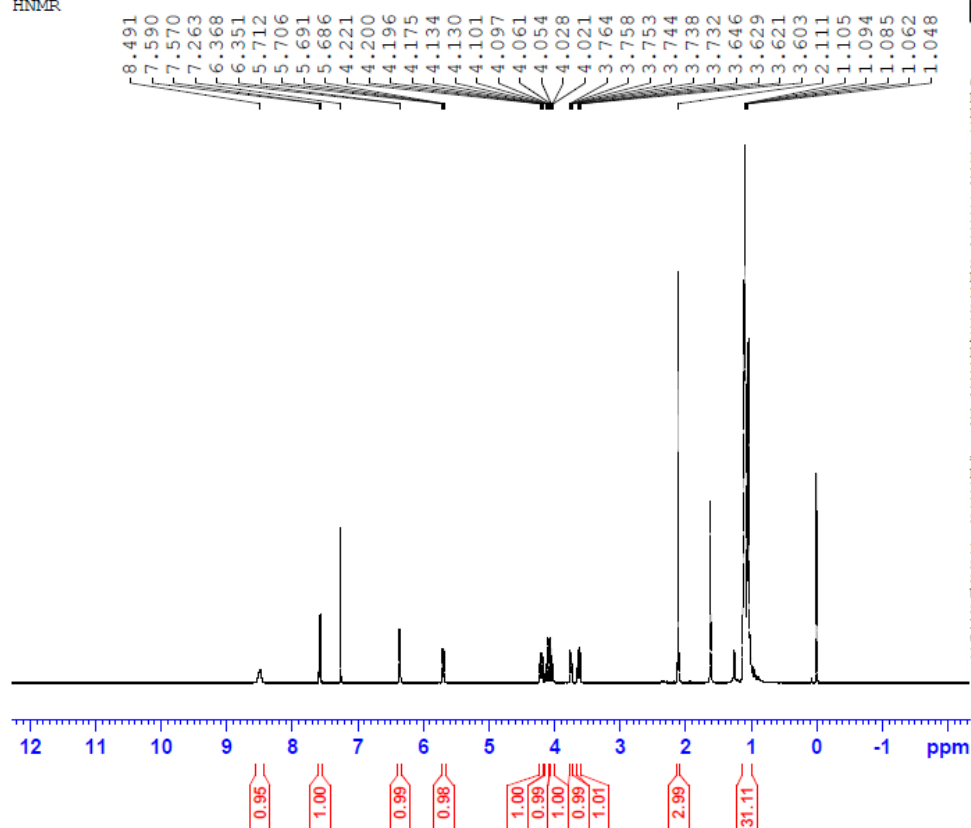
TOF MS ES+
1.29e4





6

3',5'-TIPDS-2'-ara-MeSe-U
CDCl₃
HNMR



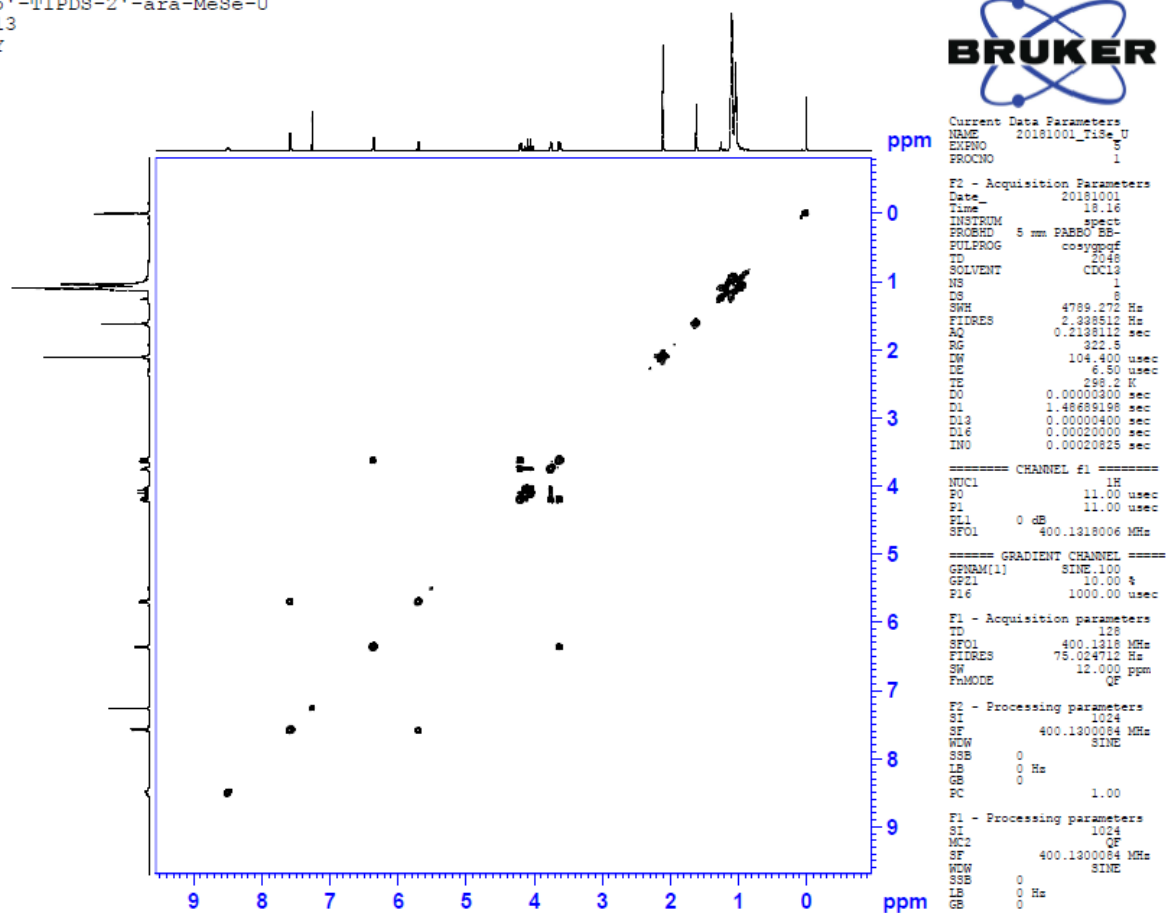
Current Data Parameters
NAME 20181001_TiSe U
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20181001
Time_ 16.28
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 65536
SOLVENT CDCl₃
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.126314 Hz
AQ 3.9583745 sec
RG 228.1
DW 60.400 usec
DE 6.90 usec
TE 298.2 K
D1 1.00000000 sec
TD0 1

===== CHANNEL f1 =====
NUC1 1H
P1 11.00 usec
PL1 0 dB
SFO1 400.1324710 MHz

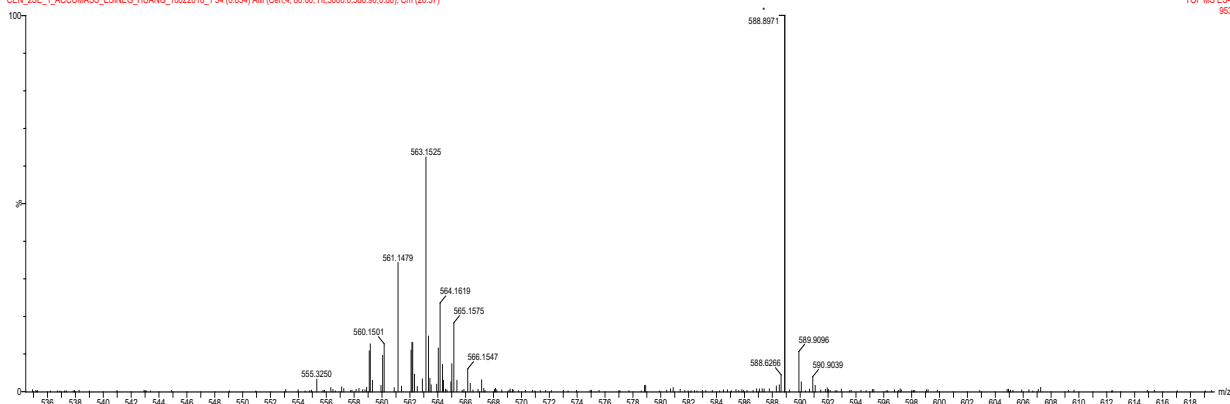
F2 - Processing parameters
SI 32768
SF 400.1300083 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

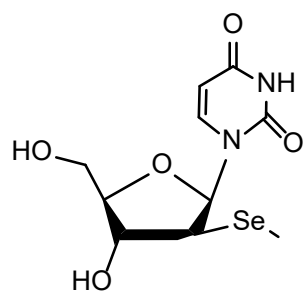
3',5'-TIPDS-2'-ara-MeSe-U
 CDCl₃
 COSY



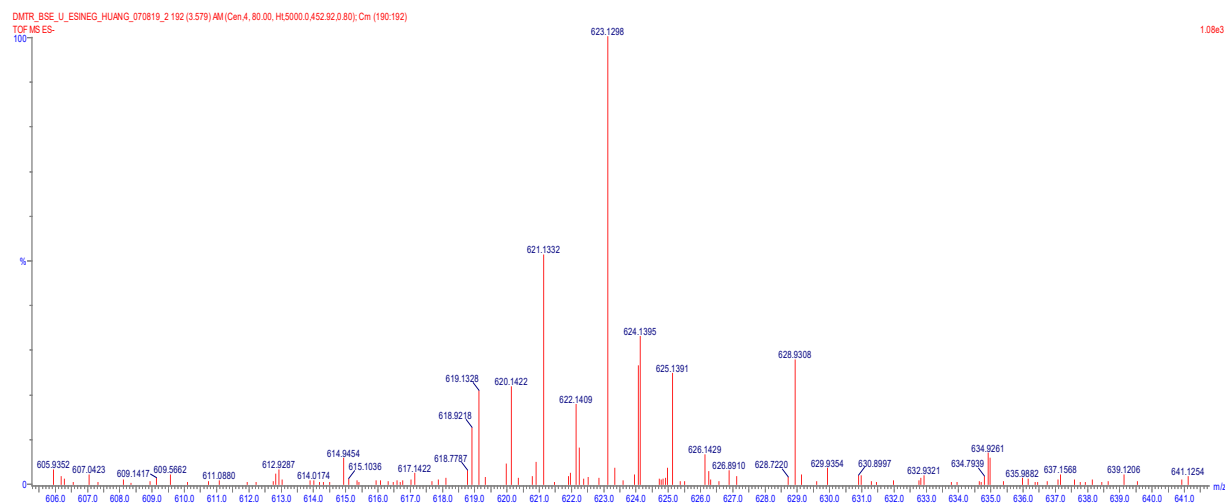
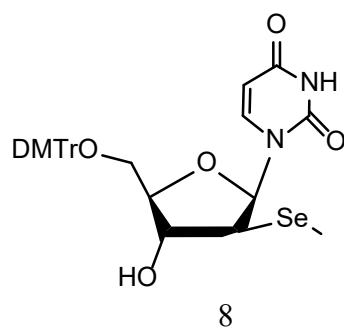
In MeOH+0.5%NH₃ 588.8971 as ITSD

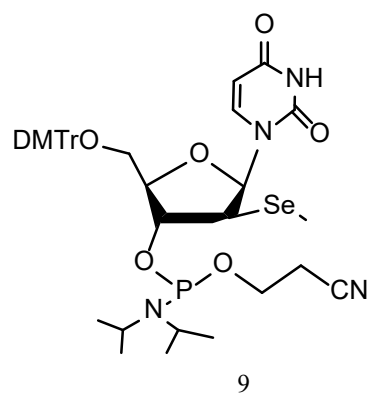
CEN_2SE_1_ACCUMASS_ESINEG_HUANG_10022018_1_34 (0.634) AM (Cen:4, 80.00, HL:5000.0,588.90,0.80); Cm (28.37)



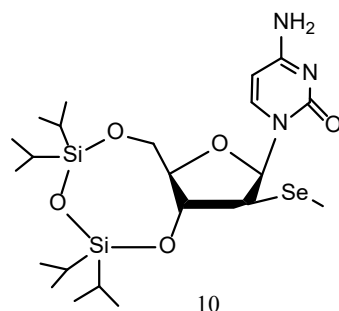


7

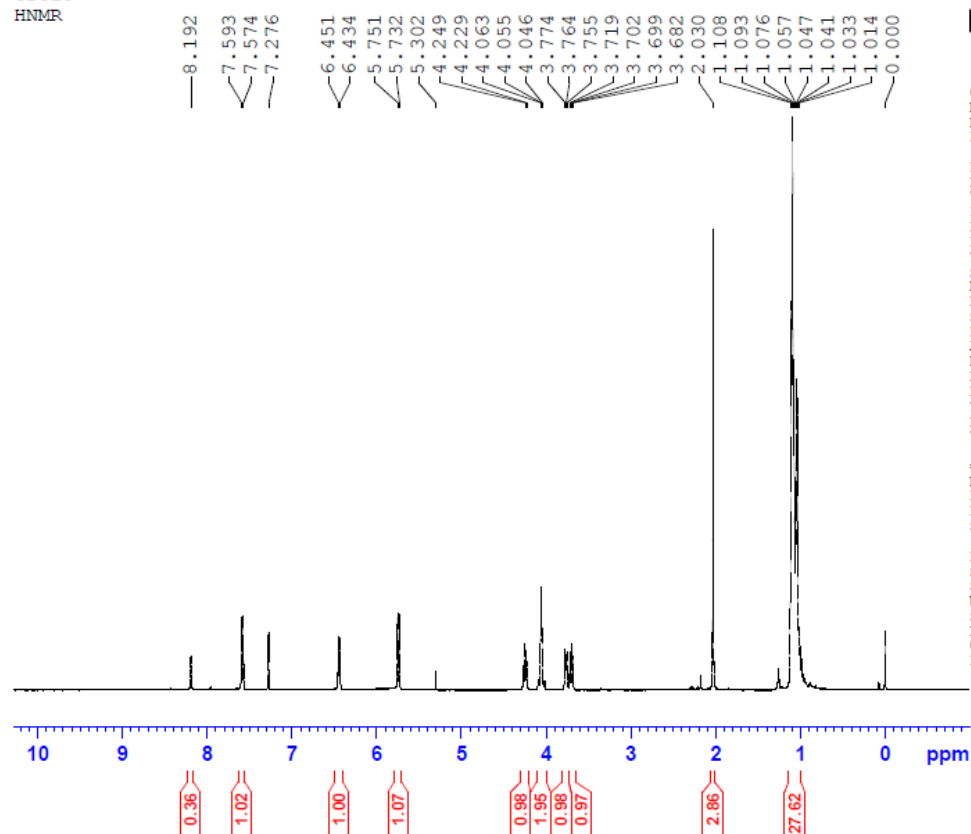




Appendix B.2. NMR and MS Spectra for 2'-SeMeANA-C



Ti2'bSe-C
CDCl₃
HNMR



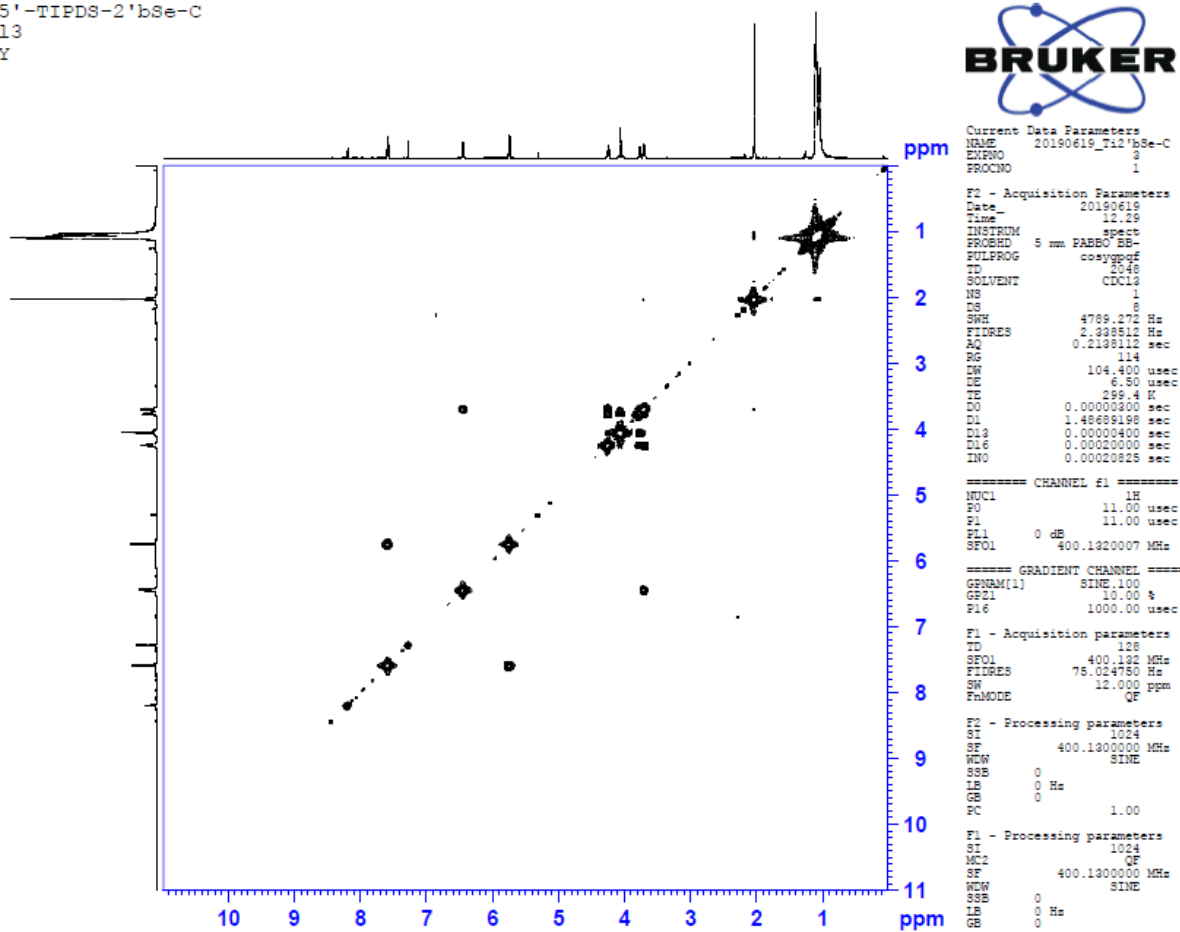
Current Data Parameters
NAME 20190619_Ti2'bSe-C
EXPNO 1
PROCNO 1

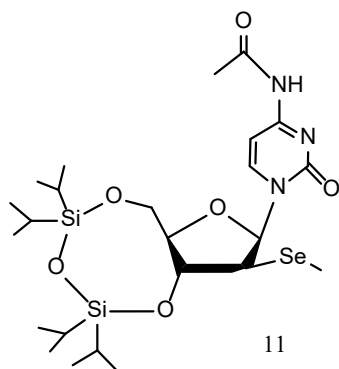
F2 - Acquisition Parameters
Date_ 20190619
Time 12.03
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 65536
SOLVENT CDCl₃
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.126314 Hz
AQ 3.9583745 sec
RG 64
DW 60.400 usec
DE 6.60 usec
TE 299.2 K
D1 1.00000000 sec
TDO 1

===== CHANNEL f1 =====
NUC1 1H
P1 11.00 usec
PL1 0 dB
SFO1 400.1324710 MHz

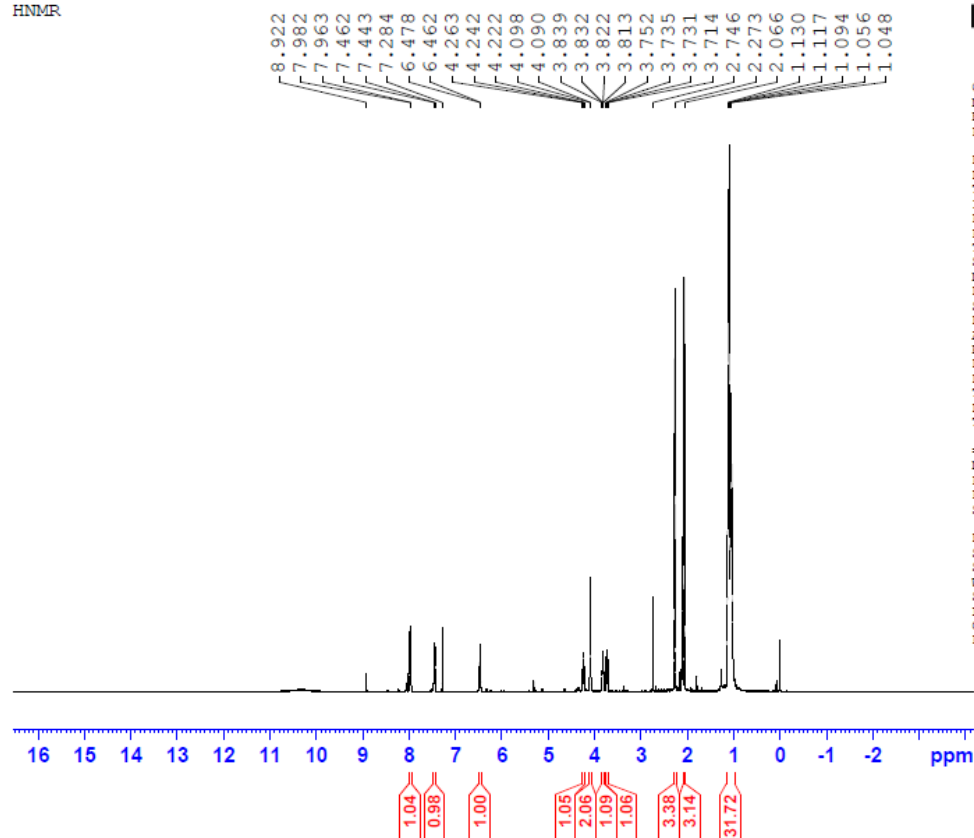
F2 - Processing parameters
SI 32768
SF 400.1300036 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

3',5'-TIPDS-2'bSe-C
CDCl₃
COSY





3',5'-TIPDS-4-N-Ac-2'-MeSe-arabinocytidine
CDCl₃
HNMR

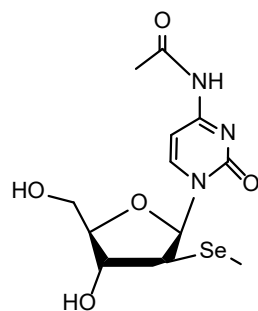


Current Data Parameters
NAME 20190813_AcTi2'bSe-C
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190813
Time_ 16.48
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 65536
SOLVENT CDCl₃
NS 16
DS 2
SWH 8278.146 KHz
FIDRES 0.126314 KHz
AQ 3.9883745 sec
RG 90.5
DW 60.400 usec
DE 6.50 usec
TE 300.5 K
D1 1.00000000 sec
TD0 1

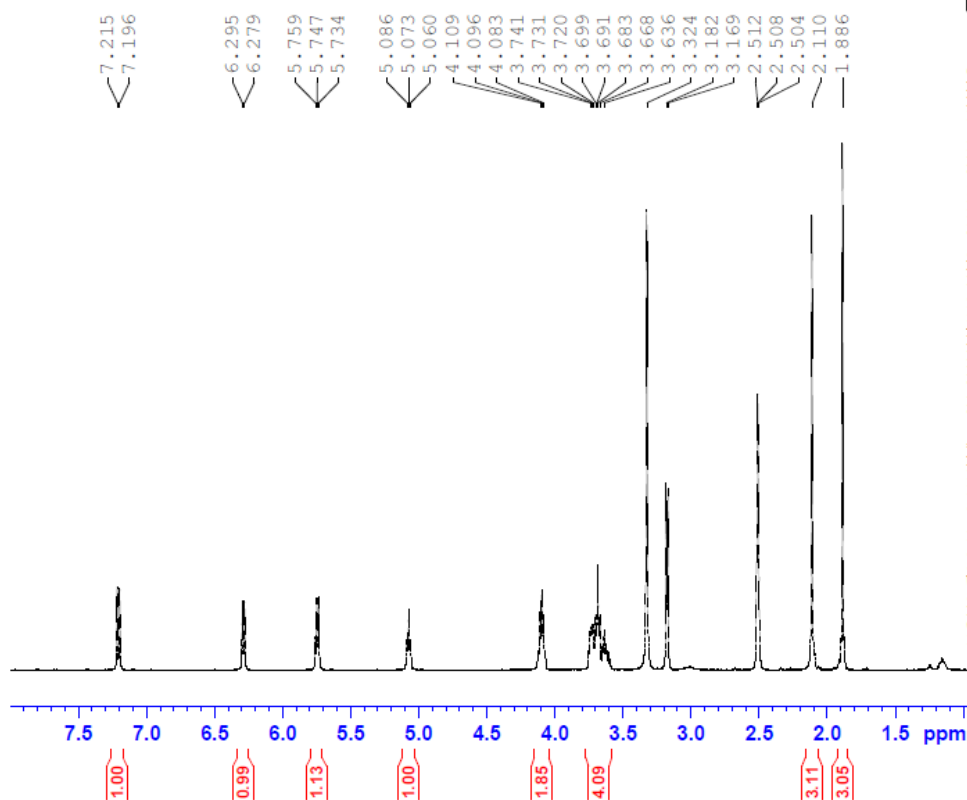
===== CHANNEL f1 =====
NUC1 1H
P1 11.00 usec
PL1 0 dB
SFO1 400.1324710 MHz

F2 - Processing parameters
SI 32768
SF 400.1300000 MHz
WDW EM
SSB 0
LB 0.30 KHz
GB 0
PC 1.00



12

4-N-Ac-2'-MeSe-arabinocytidine
DMSO
HNMR



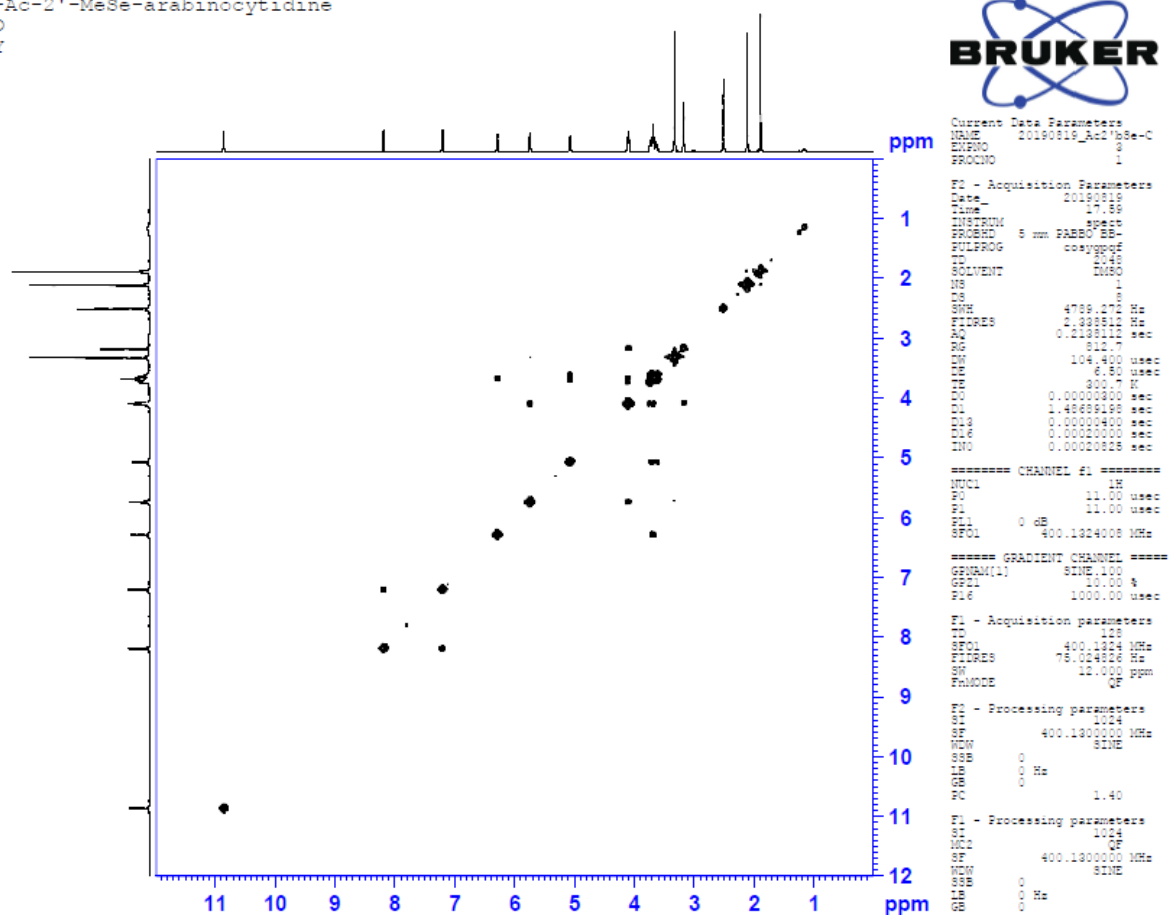
Current Data Parameters
NAME 20190819_Ac2'MeSe-C
EXPNO 1
PROCNO 1

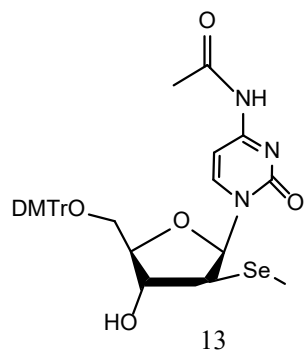
F2 - Acquisition Parameters
Date_ 20190819
Time 16.54
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zgpg30
TD 65536
SOLVENT DMSO
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.126314 Hz
AQ 3.9883748 sec
RG 322.8
DW 60.400 usec
DE 6.80 usec
TE 300.4 K
D1 1.00000000 sec
TD0 1

===== CHANNEL f1 =====
NUC1 1H
P1 11.00 usec
PL1 0 dB
SFO1 400.1824710 MHz

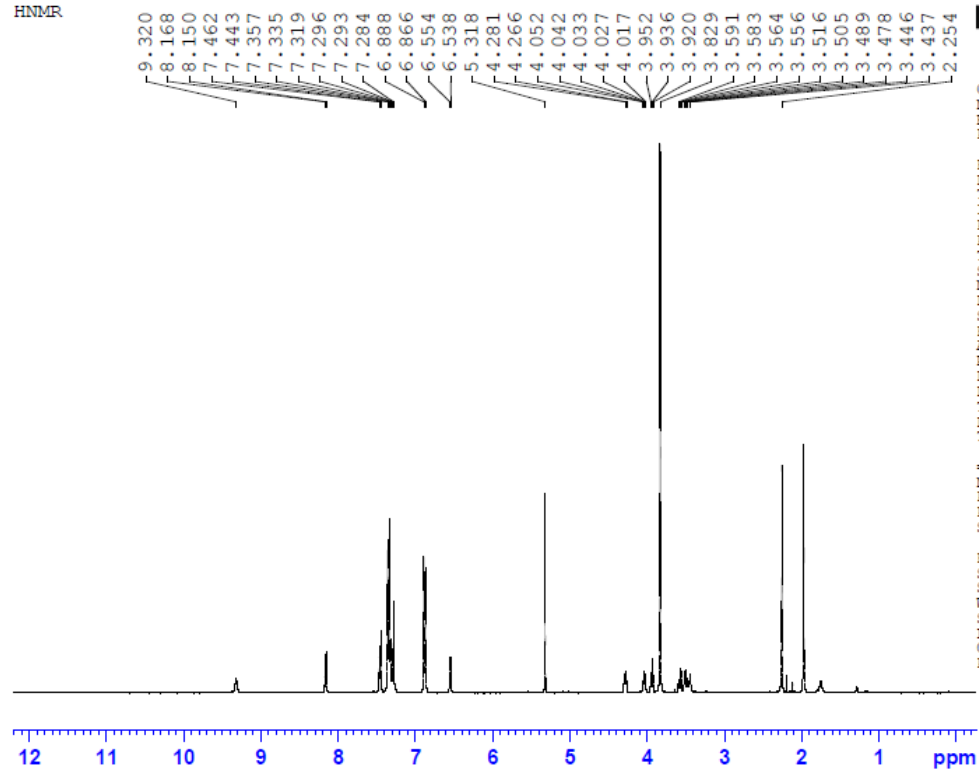
F2 - Processing parameters
SI 32768
SF 400.1800000 MHz
WDW EM
SSB 0
LBS 0.30 Hz
GB 0
PC 1.00

4-N-Ac-2'-MeSe-arabinocytidine
DMSO
COSY





5'-DMTr-4-N-Ac-2'-MeSe-arabinocytidine
CDCl₃
HNMR



Current Data Parameters
NAME 20190819_Ac2'bSeD-C
EXPNO 1
PROCNO 1

F2 - Acquisition Parameters
Date_ 20190826
Time_ 16.15
INSTRUM spect
PROBHD 5 mm PABBO BB-
PULPROG zg30
TD 65536
SOLVENT CDCl₃
NS 16
DS 2
SWH 8278.146 Hz
FIDRES 0.126314 Hz
AQ 3.9583745 sec
RG 181
DW 60.400 usec
DE 6.50 usec
TE 300.0 K
D1 1.00000000 sec
TD0 1

===== CHANNEL f1 =====
NUC1 1H
P1 11.00 usec
PL1 0 dB
SFO1 400.1324710 MHz

F2 - Processing parameters
SI 32768
SF 400.1300000 MHz
WDW EM
SSB 0
LB 0.30 Hz
GB 0
PC 1.00

