PREPARATION OF CLICKABLE MONOMERS COMPATIBLE WITH AUTOMATED PNA SYNTHESIS

by

Grace Newell

Submitted in partial fulfillment of the

requirements for Departmental Honors in

the Department of Chemistry and Biochemistry

Texas Christian University

Fort Worth, Texas

May 6th, 2019

PREPARATION OF CLICKABLE MONOMERS COMPATIBLE WITH AUTOMATED PNA SYNTHESIS

Project Approved:

Supervising Professor: Jean-Luc Montchamp, Ph.D.
Department of Chemistry and Biochemistry

Dr. Eric Simanek, Ph.D.
Department of Chemistry and Biochemistry

Dr. Matthew Chumchal, Ph.D.
Department of Biology

ABSTRACT

Peptide nucleic acids (PNA) are artificially synthesized monomers or polymers that mimic DNA or RNA sequences.[1,2] Due to their stability in biological conditions and their ability to bind complementary to DNA or RNA, PNAs have potential medicinal value since they can be used to block processes like replication, transcription, and protein synthesis.[3,4] Though most PNAs are commercially synthesized, the goal of this project was to introduce a propargyl moiety. This enables the final PNA monomer to have an alkyne which allows functional groups (like a polyamine tail, fluorescent tag, or an alkylating group) to be added at the end or any time throughout the synthesis. The PNA monomer will be made with all five nucleotide bases (adenine, cytosine, guanine, thymine, and uracil). Another importance of this PNA monomer is its ability to undergo click reactions to add functional groups or a charge.[5,6] Click chemistry is a chemical reaction that commonly uses copper-catalyzed coupling to combine an azide with an alkyne. The ability to use click chemistry is beneficial since it can be done in biological conditions, has a near quantitative yield with few byproducts, and is relatively quick to perform. In conclusion, this project is useful since these PNA sequences can be used to modulate processes and treat a variety of diseases while having the ability to add groups that will give the PNA various functionalities.

ACKNOWLEDGEMENTS

I owe sincere and earnest gratitude to Dr. Jean-Luc Montchamp. This thesis would not have been possible without his guidance and continuous help throughout the research process. Dr. Montchamp provided me with the research idea and has allowed me to grow into a patient, creative, and successful researcher. Additionally, I would like to thank Axel Sabourin for the time he spent teaching me the techniques I utilized throughout my synthesis in addition to the expertise he provided along the way. Furthermore, I would like to thank Dr. Matthew Chumchal and Dr. Eric Simanek for their insight into the development and finalization of this thesis. Moreover, I would like to thank the Science and Engineering Research Center for the generous grant which provided the funds that made this research project possible.

TABLE OF CONTENTS

List of Figures and Schemes……..…6

List of Abbreviations……8

Introduction……..…..10

Results and Discussion……...17

Conclusion and Future Work……..28

Experimental Section…….31

References……..40

LIST OF FIGURES AND SCHEMES

Figure 1. Comparison between DNA and PNA……..……...10

Figure 2. Example of a functional PNA polymer……..…….12

Figure 3. Target PNA monomer…….……...13

Scheme 1. General click chemistry reaction…….…….….13

Scheme 2. Example of a click reaction specific for target PNA monomer…….....14

Scheme 3. General synthetic scheme to synthesize the target PNA monomer(s)..……....14

Scheme 4. Automated PNA synthesis……....16

Figure 4. Synthesized thymine-containing PNA monomer……...17

Figure 5. Monoalkylated vs. dialkylated product as a result of the first step…….....18

Scheme 5. First synthetic step: substitution reaction……..18

Figure 6. Structure of Fmoc…….……..19

Scheme 6. Second synthetic step: protecting nitrogen using Fmoc……..……..19

Scheme 7. Third synthetic step: Hydrolysis……...20

Scheme 8. Fourth synthetic step: Reductive amination……..21

Scheme 9. Fifth synthetic step: Addition of the base…….22

Scheme 10. Overall synthesis of the thymine-containing PNA monomer…….22

Figure 7. High Resolution Mass spectrum of the final product……..23

Figure 8. Rotamers of the final compounds……...24

Figure 9. ^{1}H NMR of the aldehyde……25

Figure 10. The two rotamers of the aldehyde compound……...25

Figure 11. ^{1}H NMR of the final compound……26

Figure 12. TLCs of the final product……..27

Scheme 11. Proposed automated PNA synthesis using clickable PNA monomer……29

Figure 13. ^{1}H NMR of the product from the first step……32

Figure 14. ^{1}H NMR after addition of Fmoc……34

Figure 15. ^{13}C NMR after addition of Fmoc……...34

Figure 16. ^{1}H NMR of the aldehyde……..36

Figure 17. ^{13}C NMR of the aldehyde……36

Figure 18. ^{13}C NMR of the final product……38

Figure 19. ^{1}H NMR of the final compound……39

LIST OF ABBREVIATIONS

Abbreviation	Full Name
^{13}C NMR	Carbon Nuclear Magnetic Resonance
^{1}H NMR	Proton Nuclear Magnetic Resonance
abs. ppm	Absolute Parts Per Million
AcOH	Acetic Acid
Base	One of five DNA/RNA bases (adenine, cytosine, guanine, thymine, or uracil)
Boc	*tert*-butyloxycarbonyl protecting group
Br	Bromine
DCM	Dichloromethane
DIPEA	Diisopropylethylamine
DMF	Dimethylformamide
DNA	Deoxyribonucleic Acid
EDC.HCl	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide Hydrochloride
EtOAc	Ethyl Acetate
eq. or equiv.	equivalence
Fmoc	9-fluorenylmethoxycarbonyl
FmocCl	9-fluorenylmethoxycarbonyl Hydrochloride
Gly-t-Bu-ester.HCl	Glycine *tert*-butyl ester hydrochloride
H	Hydrogen
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate
HCOOH	Formic Acid
Hex	Hexanes
hrs	Hours
***i*-Pr$_2$NEt**	N,N-Diisopropylethylamine
K$_2$CO$_3$	Potassium Bicarbonate
M	Molar
m/z	Mass-to-charge ratio
MeOH	Methanol
MgSO$_4$	Magnesium Sulfate
mRNA	Messenger Ribonucleic Acid
N	Nitrogen
N$_2$	Under Nitrogen gas
Na$_2$CO$_3$	Sodium Carbonate
NaBH$_3$CN	Sodium Cyanoborohydride
NaHCO$_3$	Sodium Bicarbonate
NH$_4$Cl	Ammonium Chloride
NMP	N-methyl-2-pyrrolidone
O	Oxygen

t-Bu	*tert*-butyl
TLC	Thin Layer Chromatography
P	Phosphorus
PAL-PEG-PS	Peptide Amide Linker-polyethylene glycol-polystyrene
PG	Protecting Group
PNA	Peptide Nucleic Acid
R	Any group
r.t.	Room Temperature
RNA	Ribonucleic Acid
t-BuOH	*tert*-Butyl alcohol
TFA	Trifluoroacetic Acid
THF	Tetrahydrofuran
XAL-PEG-PS	Xanthenylamide Linker-polyethylene glycol-polystyrene

INTRODUCTION

Peptide nucleic acids (PNAs) are synthetically made molecules that mimic DNA and RNA sequences.[1,2] Because of their structural and functional similarity, PNAs are able to bind to DNA and RNA molecules and as a result can regulate important cellular processes such as DNA replication, transcription, and translation.[3,4] The sequence-specific binding between DNA and PNA or RNA and PNA exhibits traditional Watson-Crick hydrogen-bonding between the bases of the DNA or RNA and the bases of the PNA.[7] This hydrogen-bonding pattern is an important feature of PNA as it further improves the ability of PNA to successfully target certain genes and regulate cellular processes. Figure 1 illustrates the structural similarity between the structure of a PNA and that of DNA or RNA. As seen in Figure 1, PNAs derive their name from the peptide bonds that are connecting PNA monomers together to create the PNA polymer.

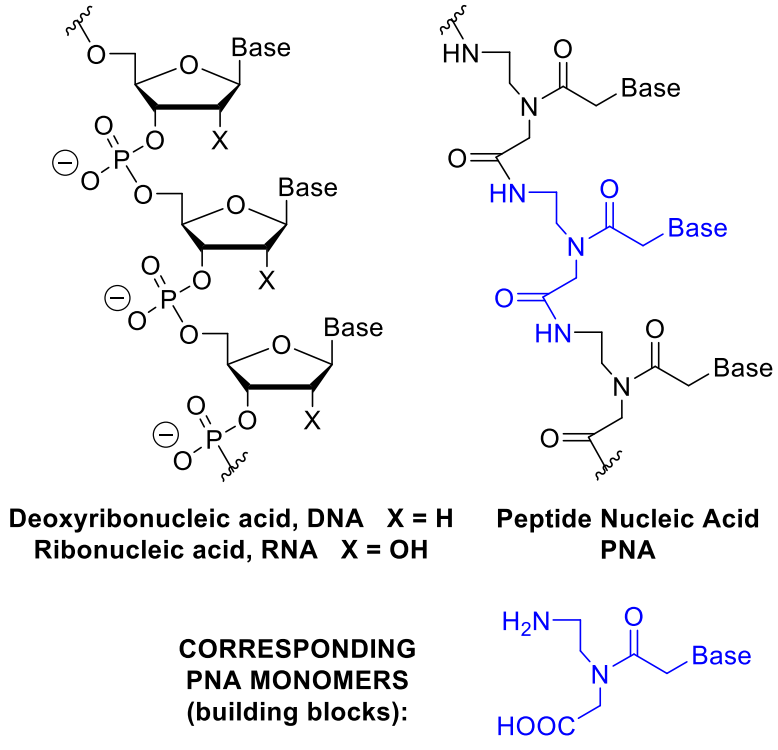


Figure 1: Comparison of the structure of DNA/RNA with the structure of PNA.

There has been much research done on PNAs due to their potential medicinal value based on their complementary binding to DNA or RNA.[3] Since PNAs are able to successfully target certain sequences of DNA or RNA, there is potential for gene or antisense therapy.[4,8-11] In addition, research has shown that PNAs have been successful in antiviral therapy and target-directed anti-cancer therapy.[12-15] Another important advantage of PNA is its resistance to degradation by enzymes like nucleases or proteases due to their non-traditional polyamide backbone.[14, 16] This property makes PNAs even more medicinally valuable as they are less likely to be degraded within the cell, thus increasing their ability to successfully attach to their sequence of interest.

The continued interest in PNAs' medicinal value has led to more research on ways to modify PNA to improve their ability to function in biological conditions. For instance, poor solubility of PNA polymers in water could be problematic for medicinal applications.[17] As a result, researchers have begun to experiment with the solubility of PNA in water by attaching charged groups to the molecule.[17] Additionally, more interest has been placed on functionalizing PNA by adding groups that could enhance their usefulness.[18] For example, attaching a fluorescent tag to the PNA would be beneficial in diagnostics, as biosensors, or as a cell biology probe. Furthermore, the addition of an alkylating group could aid in destruction of a DNA or RNA fragment that is causing disease. Alternatively, attaching folate may be beneficial in targeted anti-cancer therapy.[19-21] Figure 2 shows the structure of a PNA polymer with spheres indicating where functional groups could be attached.

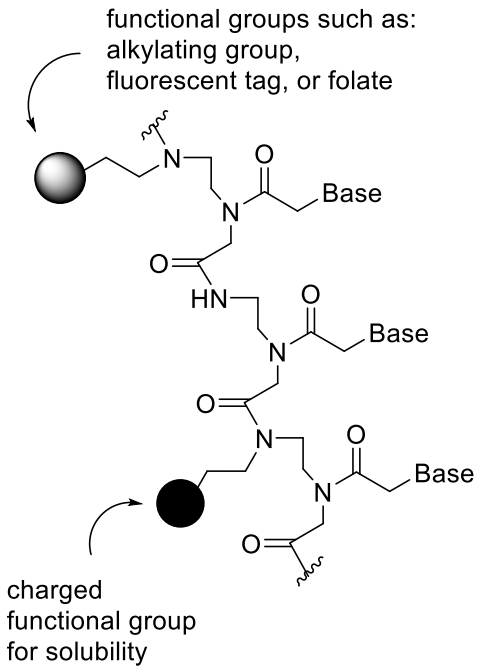
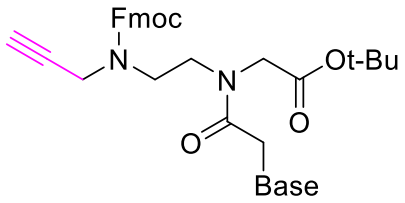


Figure 2. Example of a functional PNA polymer. The grey sphere represents where functional groups could attach, where the black sphere represents a charged group that could be attached.

The goal of this research was to synthesize a PNA monomer that is compatible with click chemistry. Click chemistry is a versatile and orthogonal strategy to install any number of different functional groups either before or after the PNA oligomer is synthesized.

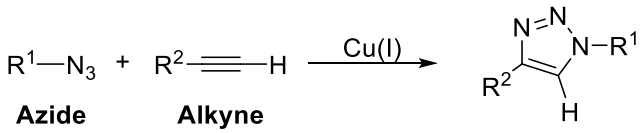
The target monomer is shown below in Figure 3. This PNA monomer has never been made before according to SciFinder searches, however, the monomer has much potential for medicinal use. As seen below in Figure 3, the monomer contains a propargyl group shown in pink, which is beneficial for attaching the functional groups including those groups described earlier.



Target PNA Monomer

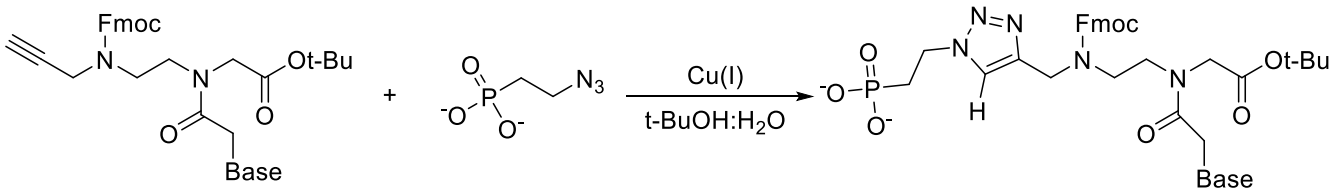
Figure 3: Target PNA monomer with the propargyl group shown in pink.

In general, a click chemistry reaction is any reaction that does not require purification, gives high yields, and is quick and simple to perform. [5,6,22] The most common example of click chemistry is the copper-catalyzed reaction between an azide and an alkyne (Scheme 1).[5,6,23] Click chemistry is advantageous in many ways and may be performed in water which is important for compounds synthesized for biological applications.[5, 22]



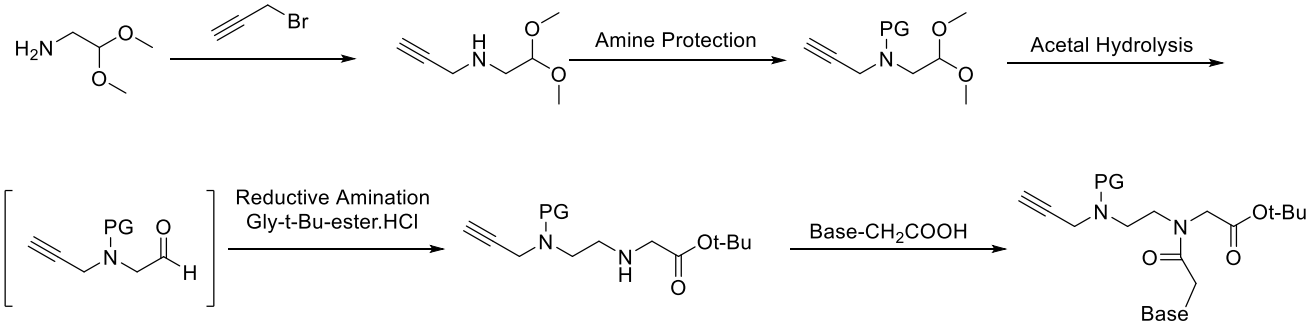
Scheme 1: General copper-catalyzed click chemistry cycloaddition between an azide and an alkyne.

Applying click chemistry to PNA has not been explored broadly.[23, 24] To advance these studies, this research focused on synthesizing a target PNA monomer compatible with chick chemistry (Figure 3). The generality of the monomer is shown in Scheme 2. Here the target PNA monomer undergoes a click reactions with an azide containing a phosphonate group. This charged group should improve PNA solubility in water.



Scheme 2: Possible click reaction of the target PNA monomer with a charged functional group.

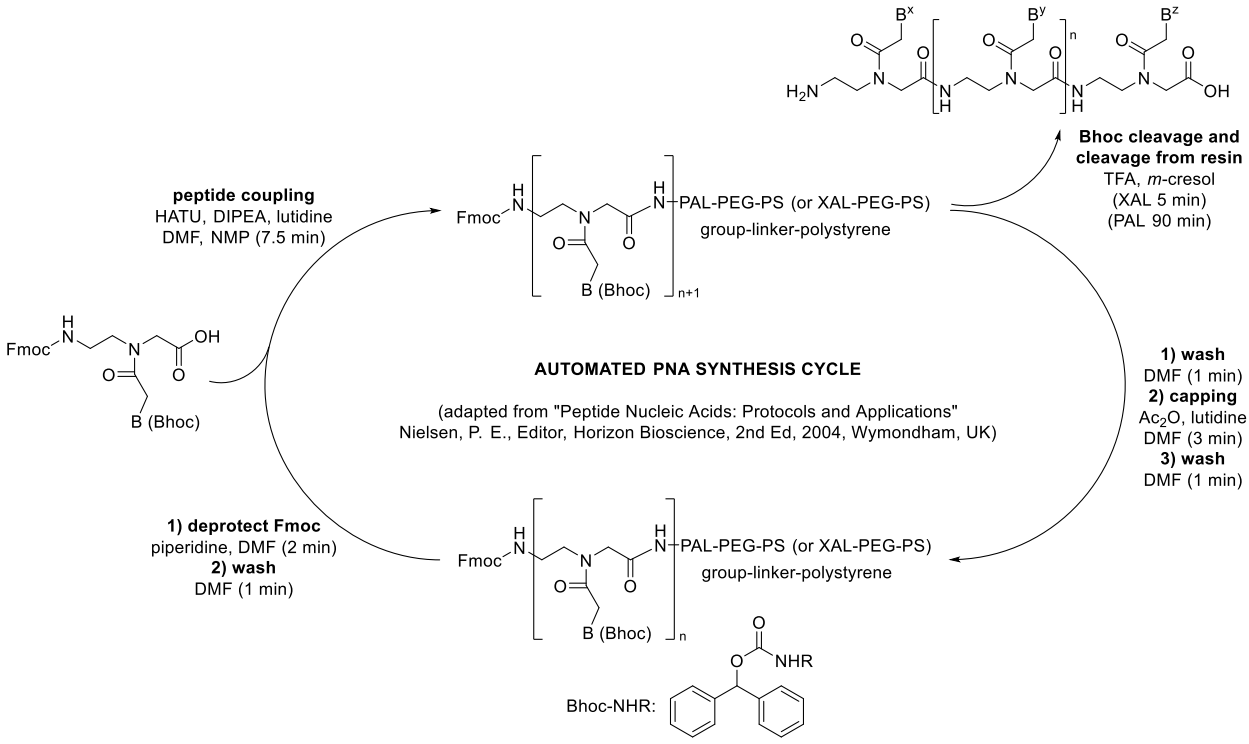
In order to synthesize the target PNA monomer, a general synthetic scheme was proposed, shown in Scheme 3. While the literature describes the synthesis of various PNA monomers,[25-29] this target PNA monomer has never been synthesized before. The design allows for the efficient synthesis of any base-containing monomer (adenine, cytosine, guanine, thymine, and uracil), because only the last step must be modified to introduce a new base. This design is important since there does not have to be a different route for each base. As seen in Scheme 3, the synthesis begins with the addition for propargyl bromide to aminoacetaldehyde dimethyl acetal. The propargyl bromide introduces the propargyl group to the compound which will eventually be important for click chemistry. Next, the addition of a protecting group (PG) is important to ensure the secondary amine will not react in subsequent steps or during automated PNA synthesis. The next two steps involve hydrolysis and reductive amination in order to introduce another nitrogen available for base addition. Finally, the last step involves the addition of the base which will be unique for each of the five bases (adenine, cytosine, guanine, thymine, and uracil).



Scheme 3: General synthetic scheme to synthesize the target PNA monomer(s).

Once the PNA monomers are synthesized they can either be functionalized via click chemistry or undergo automated PNA synthesis in order to create PNA polymers. It is important to note that not all five of the base-containing monomers are needed for applications, and instead just introducing modifications at thymine, for example, could be sufficient. Automated PNA synthesis is very similar to peptide synthesis which is the act of bringing amino acids together to create proteins/polypeptides. Automated PNA synthesis creates a chain of connected monomers, known as a polymer, by a series of coupling steps. The production of PNA polymers is medicinally important in order to synthesize a PNA sequence that matches the sequence of interest in a certain gene or mRNA.

The most common way automated PNA synthesis is performed is shown in Scheme 4.[30, 31] The synthesis begins with attaching a single PNA monomer to a solid, polystyrene (PS) resin via a linker (PAL) and a short polyethylene glycol (PEG) spacer. The PNA monomer is attached to the resin at the C-terminus, which allows for the Fmoc-amine of the monomer to be freed and available for the next step. Next, the Fmoc group is removed from the first PNA monomer allowing for that free amine to condense with the carboxyl group of the incoming PNA monomer. This coupling step uses HATU as the coupling agent. This iterative process continues by attaching one monomer at a time. Once the desired polymer is created, the PNA oligomer is cleaved from the resin in addition to any protecting groups on the bases.

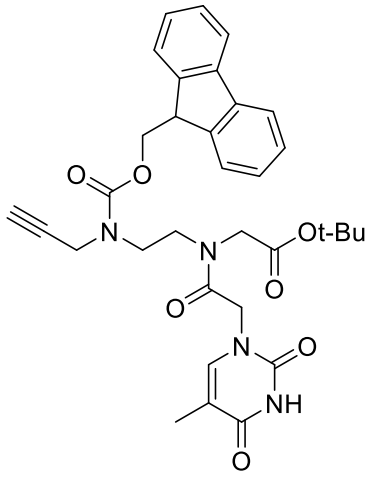


Scheme 4. Automated PNA synthesis.

Because of the specificity, the PNA polymers can be made with only a few monomers containing certain functional groups, or it can be made with a functional group attached to every single monomer if desired. By being able to pick and choose what is attached to which monomer unit, these PNA polymers have the ability to be rationally designed for various applications. As a result, this target PNA monomer can be very beneficial for medicinal applications.

RESULTS AND DISCUSSION

The target PNA monomer was synthesized with a thymine base attached using Fmoc as a protecting group. This thymine-containing PNA monomer can be seen in Figure 4. The overall synthesis consisted of 5 steps and resulted in a 25% overall yield, or 76% yield average per step. This section will explain each step used in order to synthesize this final thymine-containing PNA monomer.



PNA Monomer with Thymine Base

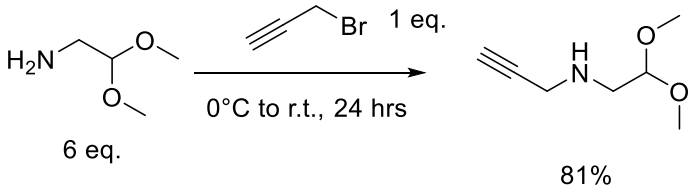
Figure 4. Synthesized thymine-containing PNA monomer with Fmoc as a protecting group.

1. Substitution of the Aminoacetal with Propargyl Bromide

The first step of the synthesis involves a substitution reaction between propargyl bromide and aminoacetaldehyde dimethyl acetal seen in Scheme 5.[25] The purpose of this step was to introduce the propargyl group which will be important for attaching functional groups via click reactions later on. Initially, this step seemed problematic as the resulting product contained both dialkylated and monoalkylated product, shown in Figure 5. Since the monoalkylated product is needed for the next step, the order of the addition of the reactants was crucial.

Figure 5. Monoalkylated product vs. dialkylated product resulting from the first step. The monoalkylated product is the one that is needed for the next step.

As a result, it became clear that in order to maximize the amount of monoalkylated product obtained, propargyl bromide in toluene must be added dropwise to a flask containing aminoacetaldehyde dimethyl acetal. Performing the reaction this way resulted in the majority of the final product to be monoalkylated (^{1}H NMR showed a 94:6 ratio of monoalkylation versus dialkylation). The chemical yield of the monoalkylated product was 81%, which was subsequently used without further purification.



Scheme 5. Step 1: Substitution reaction between propargyl bromide and aminoaldehyde dimethyl acetal.

2. Protection of the Secondary Amine

The next step of the synthesis is the protection of the nitrogen atom in the propargyl amine. This step is very important step for two reasons. First, the amine is reactive. Because of its highly reactive nature, it must be protected so it will not react during any steps further in the synthesis. If the amine was not protected, it would react to create unwanted side products. As a result, it is very important to protect this group to allow for better yields and cleaner reactions following this step. The second reason protection of the nitrogen is important is because it is needed for automated

PNA synthesis shown in Scheme 4. The removal of the protecting group from the nitrogen allows for controlled elongation of the PNA. Thus, to make sure the target PNA monomer can be used in automated PNA synthesis, the addition of this protecting group was crucial.

Several protecting groups are able to protect this nitrogen. However, Fmoc was the protecting group of choice because Fmoc is a commonly used protecting groups for amino groups and is commonly used for automated PNA synthesis. Additionally, Fmoc is easy to attach in high yield. The structure of Fmoc is shown in Figure 6. Scheme 6 shows the conditions that were used in order to protect the nitrogen using Fmoc.[26] This reaction was fast and easy to perform and resulted in a high yield (99%) after purification using chromatography on a silica gel.

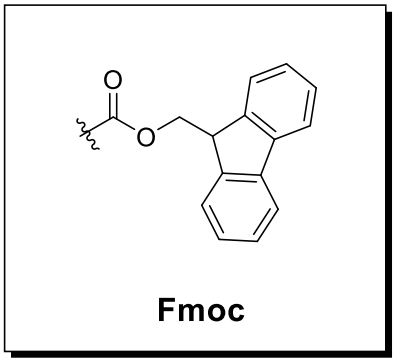
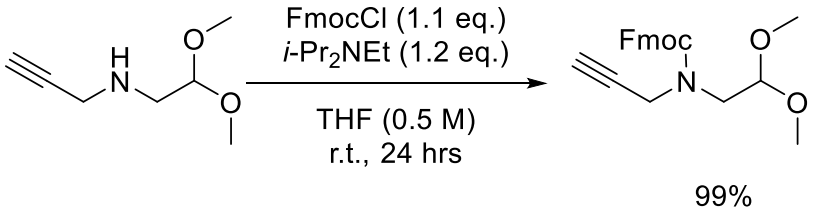


Figure 6: Structure of Fmoc.

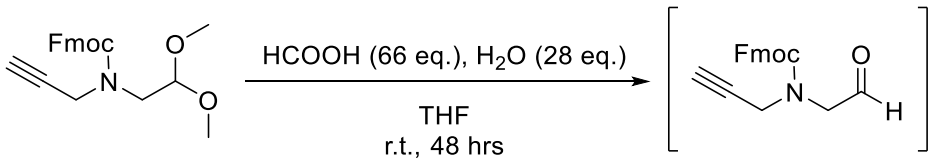


Scheme 6. Step 2: Protection of the nitrogen using Fmoc as the protecting group.

3. Hydrolysis of the Acetal to an Aldehyde

The next step of the synthesis involved hydrolyzing the acetal in order to unmask an aldehyde. The formation of the aldehyde facilitates reductive amination in the subsequent step.

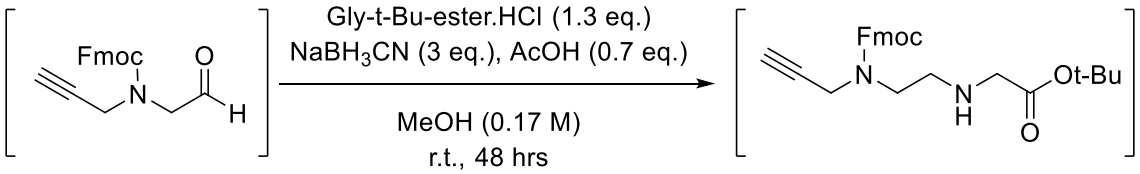
The aldehyde intermediate (shown in the brackets on Scheme 7), was not isolated and was instead used directly in the next step. Scheme 7 shows the conditions that were utilized in the hydrolysis.[27]



Scheme 7. Step 3: Hydrolysis of the acetal to an aldehyde.

4. Reductive Amination

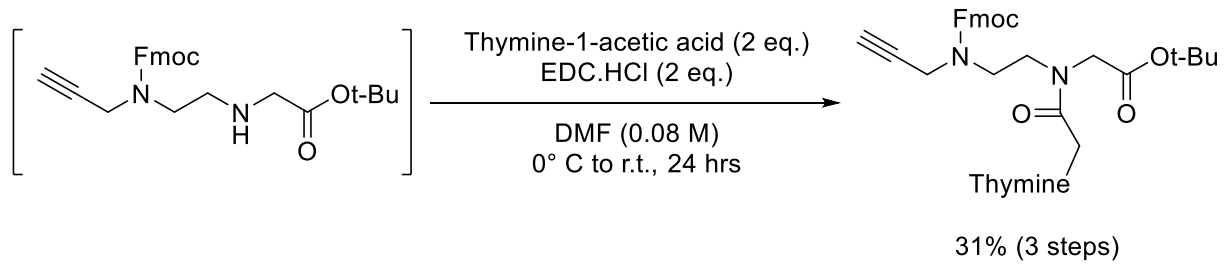
After the aldehyde intermediate was formed, the next step was reductive amination to introduce the remainder of the monomer's backbone where one of the five bases (adenine, cytosine, guanine, thymine, or uracil) will be ultimately attached. Reductive amination produces an amine via an imine intermediate by reacting either an aldehyde or a ketone with an amine. Since the previous step produced an aldehyde, it provides the carbonyl group used in reductive amination. In addition, glycine *tert*-butyl ester hydrochloride is a primary amine that was used in this step as it was commercially available. Thus, the reaction between the aldehyde and the primary amine formed a secondary amine, which will be used in the next step to attach a base. The success of this step relied upon the correct conditions. After multiple trials, using sodium cyanoborohydride emerged as the best choice of reducing agent. Scheme 8 shows the conditions that proved to be most successful for this step.[28] The isolation of the product was difficult because of the polarity of the free amine and was thus avoided.



Scheme 8. Step 4: Reductive amination.

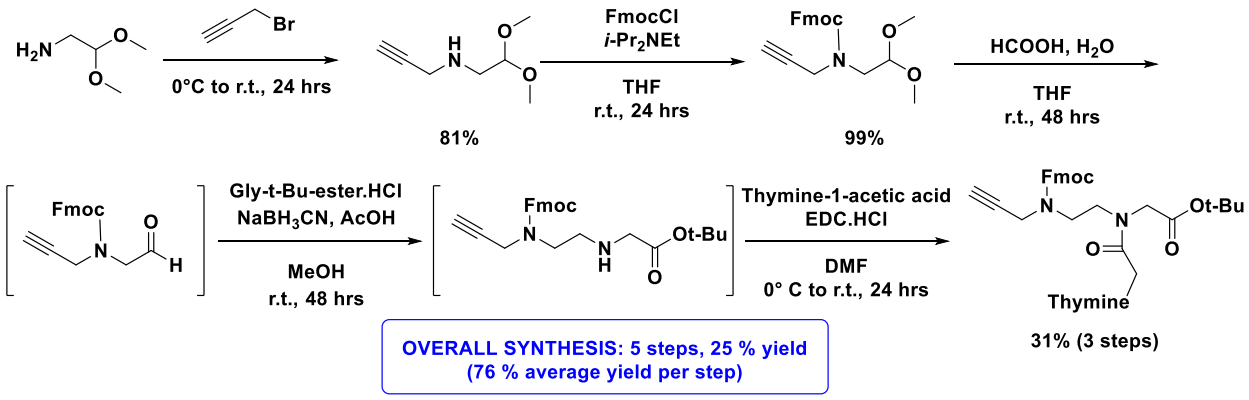
5. Addition of a Nucleobase

The final step of this synthesis was the addition of one of the five nucleotide bases (adenine, cytosine, guanine, thymine, and uracil). In this case, thymine was explored first since it does not require a protecting group like the other three bases do. In addition, the carboxylic acid reagent is commercially available.[32] The conditions used to add the thymine base are shown in Scheme 9.[29] This final reaction completes the synthesis of the thymine-containing PNA monomer and proved to be successful. A TLC and ^{1}H NMR showed some impurities, so the product was purified using silica gel chromatography starting with 2:3 (v/v) ethyl acetate/hexanes as the elution solvent, and eventually ending with 9:1 ethyl acetate/hexanes. After the first purification, ^{1}H NMR showed there was still some impurity, so the product was purified again using a silica column starting at 2:3 ethyl acetate/hexanes and ending with 100% ethyl acetate. After this second purification, TLC and the ^{1}H NMR showed no impurity. Since the products from the previous steps were not isolated, the 31% yield was calculated over three steps which is around a 70% average yield for each step. In addition to ^{13}C NMR and ^{1}H NMR, high resolution mass spectrometry results verified that the product was the intended thymine-containing PNA monomer.



Scheme 9. Step 5: Attachment of the nucleobase.

Scheme 10 shows the overall synthesis that was completed in order to synthesize the final thymine-containing PNA monomer. As noted in scheme 10, the overall yield was 25% over the five-step synthesis, corresponding to average yield of 76% per step.



Scheme 10. Overall synthesis of the thymine-containing PNA monomer.

A high resolution mass spectrum of the final monomer confirmed that the final product was in fact the desired thymine-containing PNA monomer. Figure 7 shows the experimental mass spectrum obtained from the final product. The expected mass of the final product, $C_{33}H_{36}N_4O_7$, was m/z 601.2657 and the mass spectrum found the actual mass to be m/z 601.2565. This data shows a difference of only about 0.3 ppm, a variation that confirms the identity of the final product. In addition to a high resolution mass spectrum of the final product, NMRs were collected throughout the synthesis. These spectra are shown in the experimental section. It is important to

note the appearance of rotamers in these spectra which can make the analysis of spectra complicated.

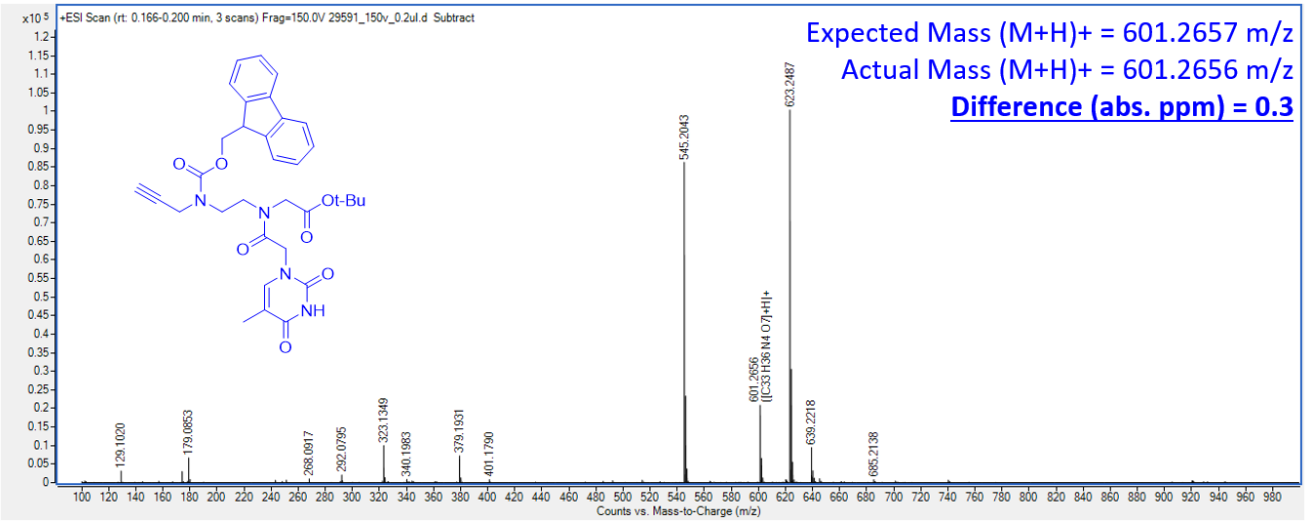


Figure 7. High resolution mass spectrum of the final product.

Rotamers occur due to resonance in the amide structure creating a partial double-bond character. As a result, the rotation around the C-N bond is slow enough that the NMR "sees" two isomers of the carbamate. This behavior is shown in Figure 8 along with the C-N bonds in the final compound that are responsible for the rotamers. In the NMR of the aldehyde, there will be two rotamers since there is only one nitrogen, whereas in the NMR of the final compound there will be four rotamers as a result of two amide nitrogens in the molecule.

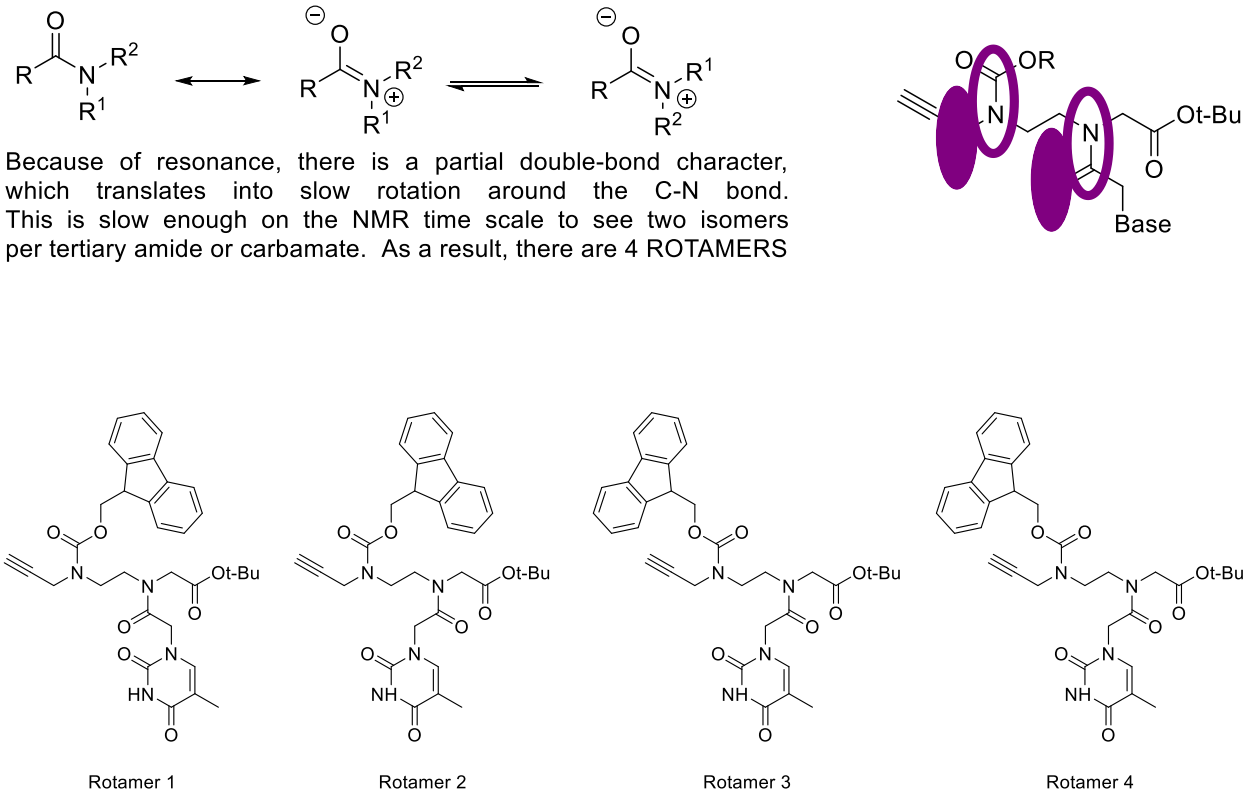


Figure 8. Resonance structures showing the double-bond character of an amide/carbamate moiety and the resulting four isomers that are possible.

Figure 9 shows the ^{1}H NMR of the aldehyde product obtained in step three. As seen in the NMR, there are two aldehyde peaks around 9.5 ppm which are a result of the two rotamers of the compound. Figure 10 shows the two rotamers that are possible for this structure.

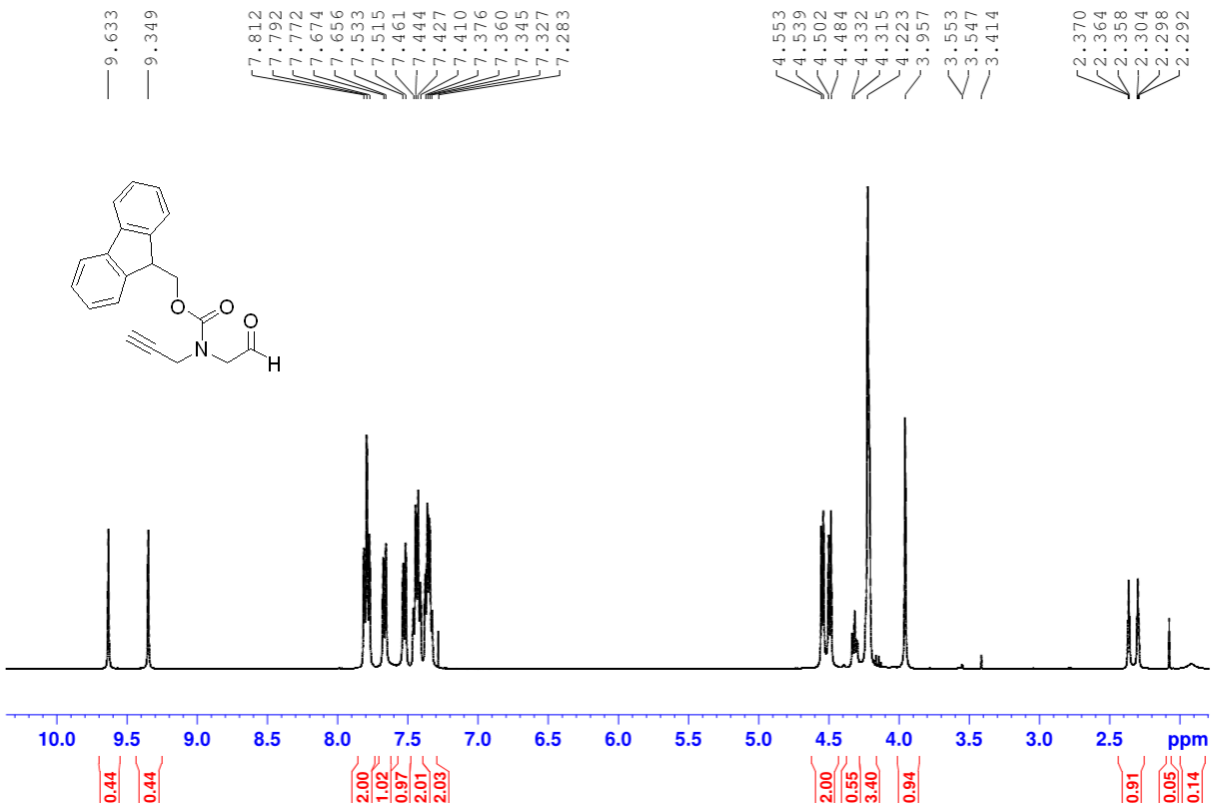


Figure 9. ^{1}H NMR of the aldehyde.

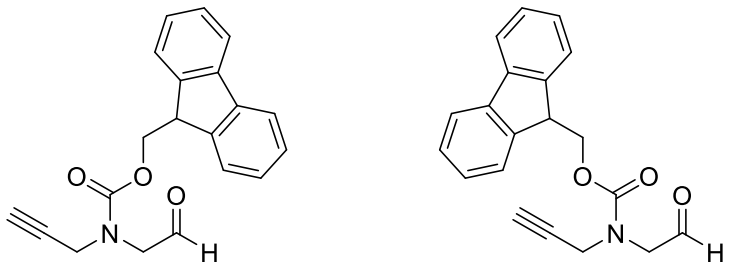


Figure 10. The two rotamers of the aldehyde compound.

In addition, Figure 11 shows a ^{1}H NMR of the final compound. The NMR looks messy, but this is because there are four rotamers that exist (Figure 8).

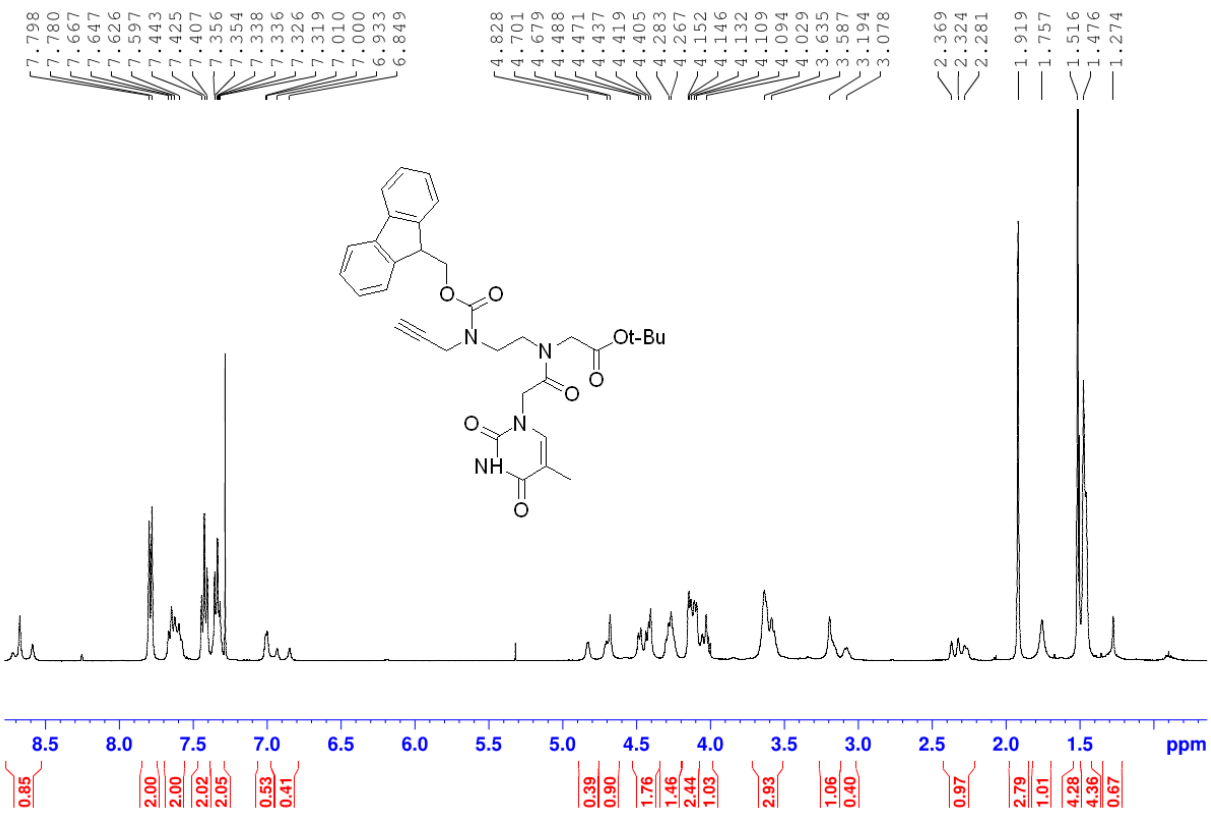


Figure 11. ^{1}H NMR of the final compound.

The ^{1}H-NMR spectrum of the final product is complex and hard to deconvolute. Two-dimensional NMR may be useful in assigning each rotamer and its relative amount. Because of the complexity of the ^{1}H-NMR spectrum it is important to ensure the final product is pure. This can be done with thin layer chromatography (TLC). Four different TLCs were performed, each in a different solvent system to confirm there were no impurities in the final product. Figure 12 shows the results of the TLCs with the solvent ratios and Rf values noted. As seen, there is only one spot that was both UV and $KMnO_4$ active. Because there was only one spot that shows up in all four of the TLCs, it can be concluded the final product is pure. The purity of the product taken with the high resolution mass spectrum establish the identity of the target molecule.

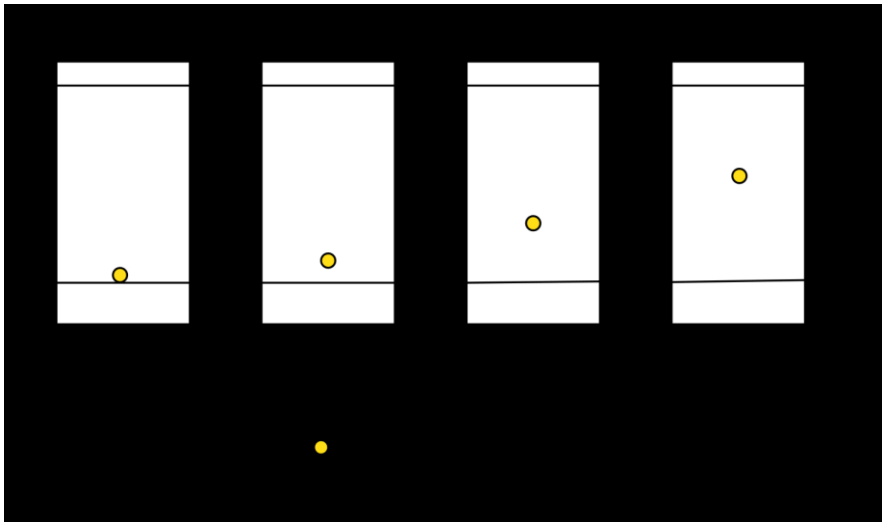


Figure 12. TLCs of the final product.

CONCLUSION AND FUTURE WORK

The synthesis of the thymine-containing PNA monomer was successful. Not only did analysis of ^{13}C NMR, ^{1}H NMR, and high resolution mass spectrum results confirm the final product was the intended thymine-containing PNA monomer, the overall yield was satisfactory. A yield of 25% over five steps means that on average each step has around a 76% yield which is quite good. As a result, this synthetic scheme was proven to be successful.

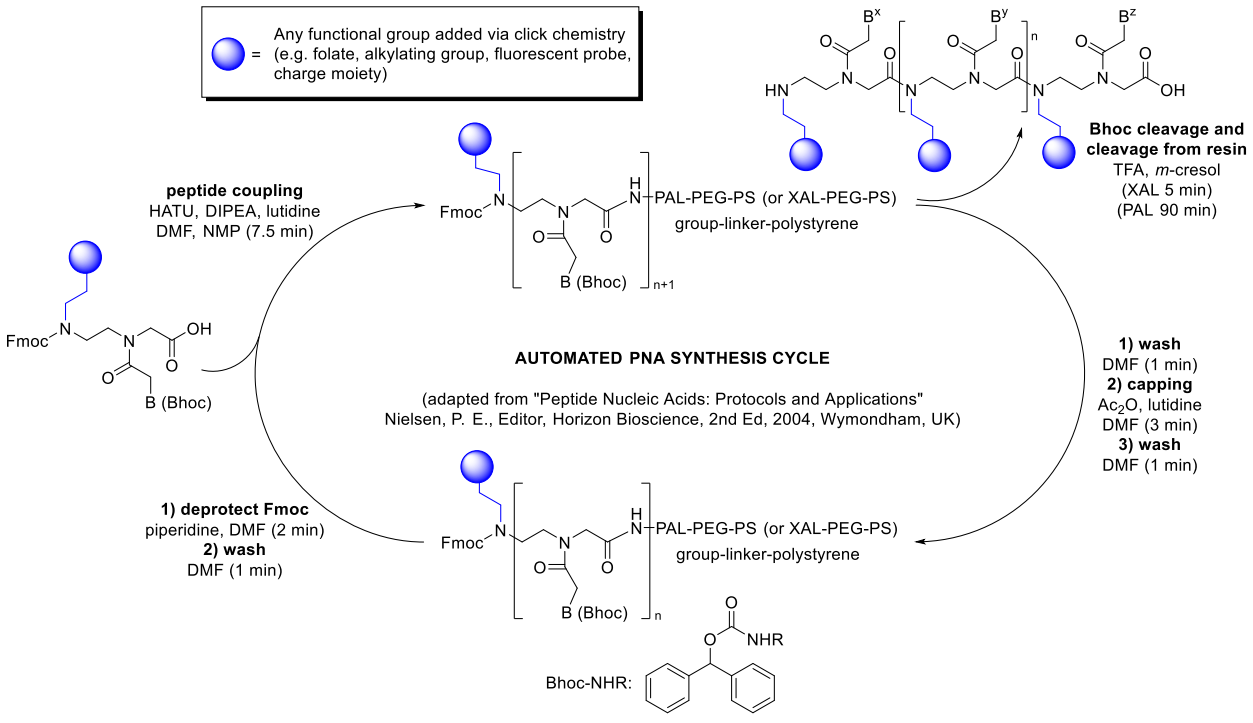
There is a lot to be done in the future on this project. First, the other four base-containing PNA monomers should be made (adenine, cytosine, guanine and uracil). Currently, the synthesis of the base-protected cytosine acetic acid [32, 33] has been completed and is ready to be attached using the final step shown in Scheme 9. However, the only thing that will change is the thymine acetic acid will be exchanged for the cytosine acetic acid. After the remaining four base-containing PNA monomers are made, either automated PNA synthesis can be performed or click chemistry will be used to attach specific functional groups prior to incorporation into the oligomer.

The ability to use click chemistry was the purpose of the synthesized PNA monomer. Future work includes using click chemistry to attach functional groups that are beneficial for medicinal applications. Some important functional groups that may be added in the future include the following:

- Charged moiety (like a phosphonate group) to increase the solubility of PNA polymers in water
- Fluorescent probes could be helpful if the PNAs intended use is to be a biosensor for diagnostics
- Alkylating groups may be useful if the PNAs goal is to destroy a specific sequence, like an mRNA that may be causing disease

- Folate may assist in target-directed anti-cancer therapy.

These functionalized PNA monomers may also be used in automated PNA synthesis to create a polymer. Automated PNA synthesis with these PNA monomers would look much like it does in Scheme 4 with a few modifications. Scheme 11 shows what the automated PNA synthesis would most likely look like. What is highlighted in blue shows the major difference from what was seen in Scheme 4. In this case, because the PNA monomers have the ability to be functionalized using click chemistry, there will be either distinct functional groups or the original propargyl group attached to the nitrogen instead of just a hydrogen which was seen in Scheme 4. This will also allow for specificity on what monomers contain what specific functional group, if any.



Scheme 11. Proposed automated PNA synthesis using PNA monomers synthesized from this research.

Furthermore, these PNA monomers and polymers may be sold to pharmaceutical companies or other researchers that can use them in medicinal applications (e.g. gene or antisense therapy, antiviral therapy, or target-directed anticancer therapy).

EXPERIMENTAL SECTION

General Chemistry:

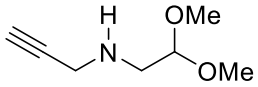
^{1}H NMR spectra were recorded on a 300-MHz Varian INOVA spectrometer or 400-MHz Bruker Avance spectrometer. Chemical shifts for ^{1}H NMR spectra (in parts per million) relative to internal tetramethylsilane (Me_4Si, δ = 0.00 ppm) with deuterated chloroform. ^{13}C NMR spectra were recorded at 75.5 or 101 MHz. Chemical shifts for ^{13}C NMR spectra are reported (in parts per million) relative to $CDCl_3$ (δ = 77.0 ppm). Flash chromatography experiments were carried out on Silica Gel premium Rf grade (40−75 μm). Ethyl acetate/hexane mixtures or dichloromethane/acetone were used as the eluent for chromatographic purifications. TLC plates were visualized by UV or immersion in permanganate potassium (3 g $KMnO_4$, 20 g K_2CO_3, 5 mL 5% aq. NaOH and 300 mL of water) followed by heating. High resolution mass spectra (HRMS) were obtained either by direct probe (EI/CI) and analyzed by magnetic sector, or by electrospray using a TOF analyzer.

Reagent and solvents:

All starting materials were purchased from commercial sources and used as received. The solvents were distilled under N_2 and dried according to standard procedures (DMF from $MgSO_4$; CH_3CN, toluene and dichloromethane from CaH_2).

First Step – Substitution Reaction:[25] A 100 mL round bottom flask containing aminoacetaldehyde dimethyl acetal (6.0 equiv, 32 mL, 300 mmol) was flushed with N_2. Propargyl bromide (80% in toluene, 1.0 equiv, 5.6 mL, 50 mmol) was added dropwise at room temperature. The reaction mixture stirred for 24 hours in room temperature under N_2. The reaction was

quenched with DCM (40 mL) and washed with water, $NaHCO_3$, and finally brine. The organic layer was dried with $MgSO_4$, filtered, and evaporated. ^{1}H NMR showed 94% of the product was the intended monoalkylated product with 6% was the unwanted dialkylated product. The final yield was 81% and the product was used without additional purification. Figure 13 shows the ^{1}H NMR of the product.



no rotamer

^{1}H NMR (400 MHz, $CDCl_3$): δ = 4.51 (t, J = 6 Hz, 1H), 3.47 (d, J = 2 Hz, 2H), 3.41 (s, 6H), 2.81 (d, J = 6 Hz, 2H), 2.25 (t, J = 2 Hz, 1H).

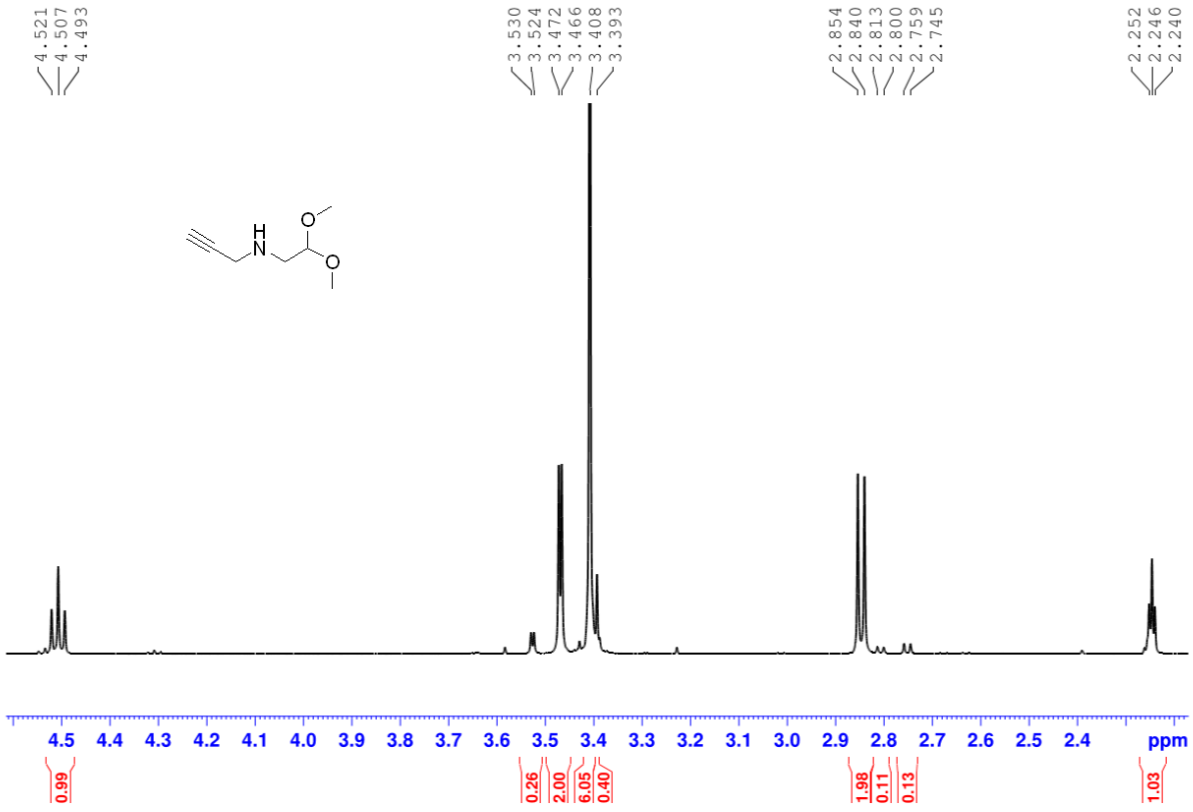
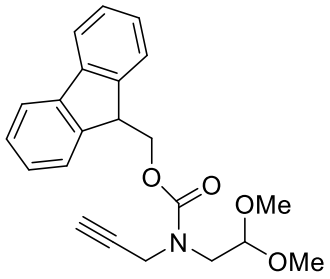


Figure 13. ^{1}H NMR of the product from the first step

Second Step – Protection of the Nitrogen:[26] The product from the first step (1 equiv, 4.06 g, 32.2 mmol) was transferred to a 250 mL round bottom flask in which THF (64.4 mL, 0.5 M) was added in addition to DIPEA (1.2 equiv, 6.73 mL, 38.6 mmol). Then FmocCl (1.1 equiv, 9.16 g, 35.4 mmol) was added and reaction mixture stirred overnight in room temperature and under N_2. The reaction was quenched using ethyl acetate and washed with NH_4Cl and then brine. The organic layer was dried with $MgSO_4$ and was filtered and then evaporated. ^{1}H NMR showed some impurities so the product was purified using a silica column 9:1 then 8:2 hexane/ethyl acetate mixture. ^{1}H NMR of the purified product showed very few impurities to afford 11.27 g (96%). Figure 14 shows the ^{1}H NMR of this product. Figure 15 shows the ^{13}C NMR of this product.



2 rotamers

^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.79 (d, *J* = 8 Hz, 2H), 7.68 (d, *J* = 8 Hz, 1H), 7.62 (d, *J* = 8 Hz, 1H), 7.43 (t, *J* = 8 Hz, 2H), 7.34 (t, *J* = 8 Hz, 2H), 4.63 (d, *J* = 6 Hz, 1H), 4.53 (t, *J* = 6 Hz, 0.5H), 4.45 (d, *J* = 7 Hz, 1H), 4.31 (t, *J* = 7 Hz, 0.5H), 4.26 (t, *J* = 6 Hz, 0.5H), 4.26-4.24 (m, 2H), 4.10 (t, *J* = 5 Hz, 0.5H), 3.52 (d, *J* = 5 Hz, 1H), 3.44 (s, 3 H), 3.26 (d, *J* = 6 Hz, 1H), 3.25 (s, 3H), 2.28 (m, 0.5H), 2.23 (m, 0.5H). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 155.7 (2 peaks), 143.9, 141.4, 127.7, 127.1 (2 peaks), 125.2, 124.7, 120.0, 103.5 (2 peaks), 79.4, 79.2, 71.9, 68.0, 67.0, 54.7, 48.4, 47.6.

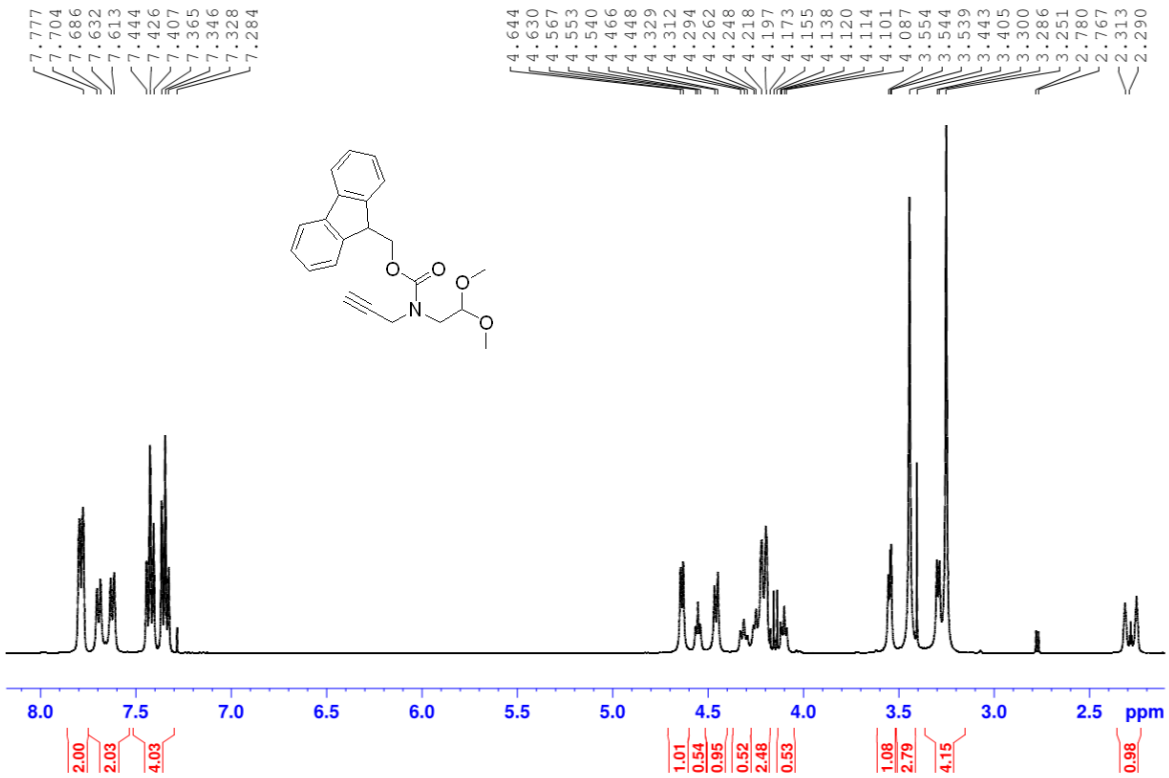


Figure 14. [1]H NMR after addition of Fmoc.

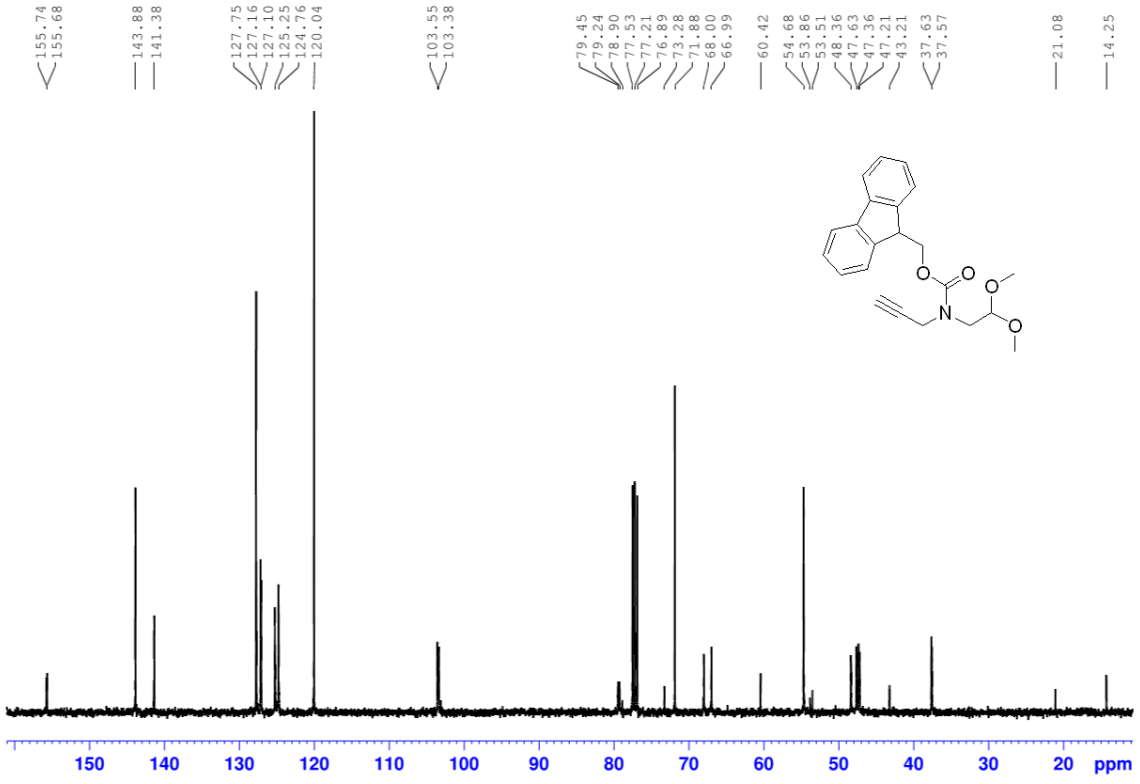
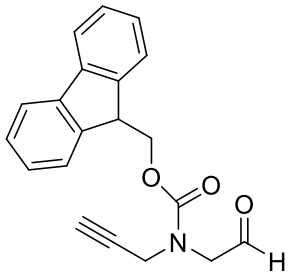


Figure 15. [13]C NMR after addition of Fmoc.

Third Step – Hydrolysis of the Acetal to an Aldehyde:[27] The product from the second step (1 equiv, 6.39 g, 17.5 mmol), was transferred to a 250 mL round bottom flask and was dissolved in formic acid (66 equiv, 43.34 mL, 1,148.7 mmol). Then THF (43 mL) and water (28 equiv, 8.9 mL, 487.3 mmol) was added. The reaction stirred for 48 hours in room temperature under N_2. The reaction mixture was then quenched using DCM and was washed with a solution of NaOH pellets and water to basify the aqueous layer. The organic layer was then washed with brine and the resulting organic layer was dried with $MgSO_4$, filtered, and evaporated. The product was not purified as it was an intermediate and was used directly in the next step. Figure 16 shows the ^1H NMR of this product and Figure 17 shows the ^{13}C NMR of this product.



2 rotamers

^1H NMR (400 MHz, $CDCl_3$): δ = 9.63 (s, 0.5H), 9.35 (s, 0.5H), 7.79 (t, *J* = 8 Hz, 2H), 7.66 (d, *J* = 7 Hz, 1H), 7.52 (d, *J* = 7 Hz, 1H), 7.47-7.32 (m, 4H), 4.54 (d, *J* = 6 Hz, 1H), 4.49 (d, *J* = 7 Hz, 1H), 4.31 (t, *J* = 7 Hz, 0.5H), 4.25-4.18 (m, 3.5H), 3.96 (s, 1H), 2.36 (t, *J* = 2 Hz, 0.5H), 2.30 (t, *J* = 2 Hz, 0.5H). ^{13}C NMR (101 MHz, $CDCl_3$): δ = 197.7 (2 peaks), 155.5 (2 peaks), 143.6, 141.3, 127.9, 127.2, 125.2, 124.7, 120.0, 78.1, 73.5, 68.4, 67.9, 56.2, 55.6, 47.1, 37.6 (2 peaks).

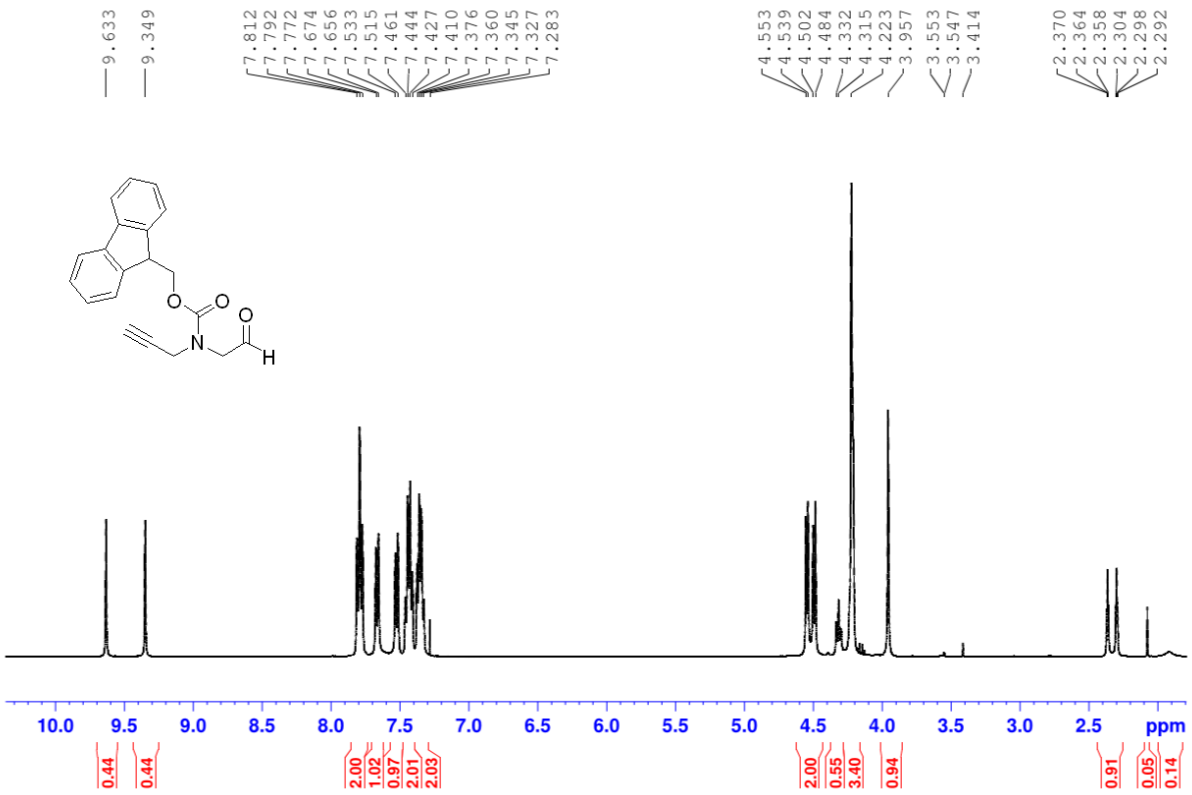


Figure 16. ^{1}H NMR of the aldehyde.

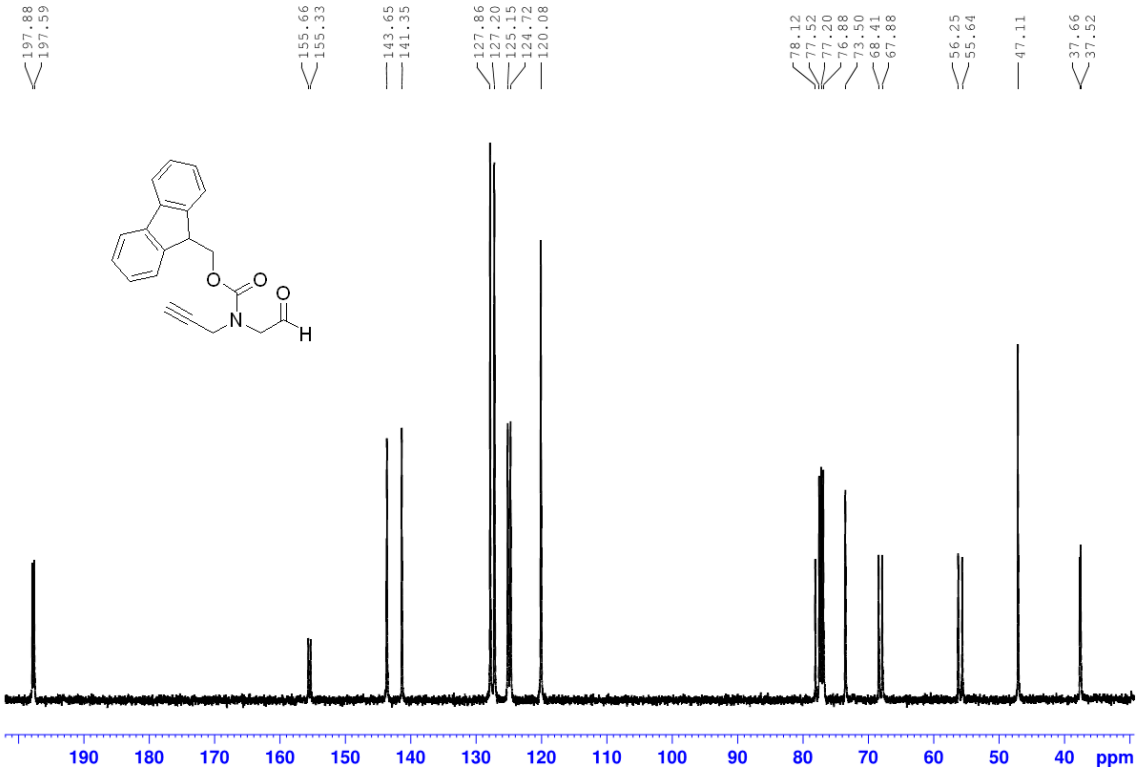


Figure 17. ^{13}C NMR of the aldehyde.

Fourth Step – Reductive Amination:[28] The product from reaction three (1 equiv, 0.705 g, 2.21 mmol) was dissolved in methanol (13 mL, 0.170 M). Glycine *tert*-butyl ester HCl (1.3 equiv, 0.81 g, 2.87 mmol) was added and stirred for 5 minutes. $NaBH_3CN$ (3 equiv, 0.416 g, 6.63 mmol) was then added along with a few drops of acetic acid (0.7 equiv). Reaction stirred at room temperature for 96 hours under N_2. The reaction mixture was quenched with ethyl acetate (50 mL) and washed with saturated K_2CO_3 until aqueous layer was basic. The aqueous layer was then washed with ethyl acetate (50 mLx3) and the organic layers were combined, dried with $MgSO_4$, filtered, and evaporated.

Fifth Step – Addition of the Base:[29] The product from reaction four (1 equiv, 0.90 g, 2.07 mmol) and thymine-1-acetic acid (2 equiv, 0.763 g, 4.14 mmol) was dissolved in anhydrous DMF (23.9 mL, 0.08 M). Then, EDC hydrochloride (2 equiv, 0.784 g, 4.14 mmol) was added at 0°C. Reaction stirred for 48 hours in room temperature under N_2. DMF was removed via the water pump. Ethyl acetate was then added to quench the reaction and the reaction mixture was washed with Na_2CO_3, NH_4Cl, and brine. The organic layer was dried with $MgSO_4$ and was filtered and evaporated. A TLC and ^1H NMR showed some impurities, so the product was purified using a silica column starting at 2:3 ethyl acetate/hexanes and eventually ending with 9:1 ethyl acetate/hexanes. ^1H NMR showed there was still some impurities so the product was purified again using a silica column starting at 2:3 ethyl acetate/hexanes and ending with 100% ethyl acetate. ^1H NMR showed the product was pure and after the two columns there was 0.388 g of product left (31% but over 3 steps). Figure 18 shows the ^{13}C NMR of the final product and Figure 19 shows the ^1H NMR for the final compound.

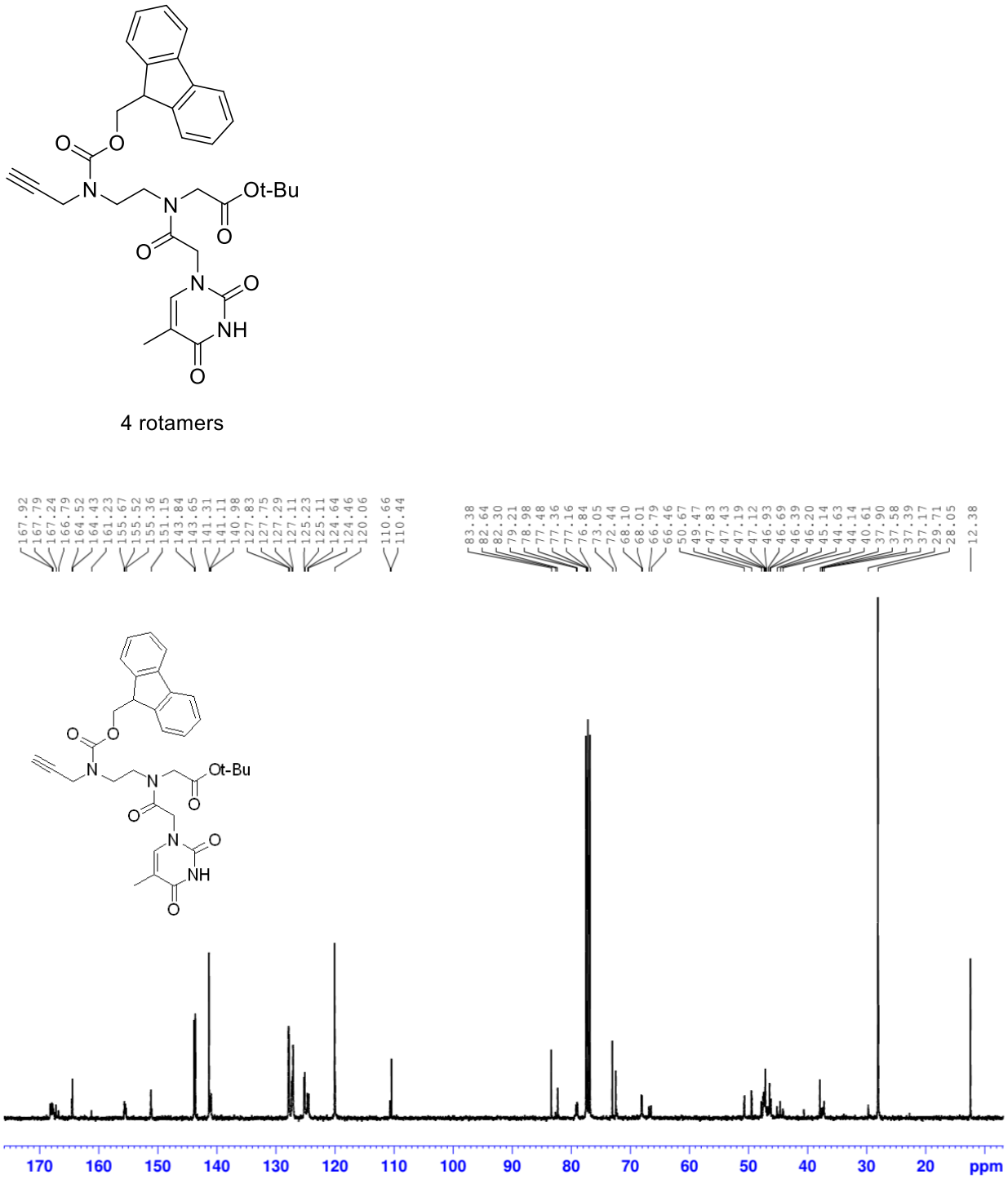


Figure 18. ^{13}C NMR of the final product.

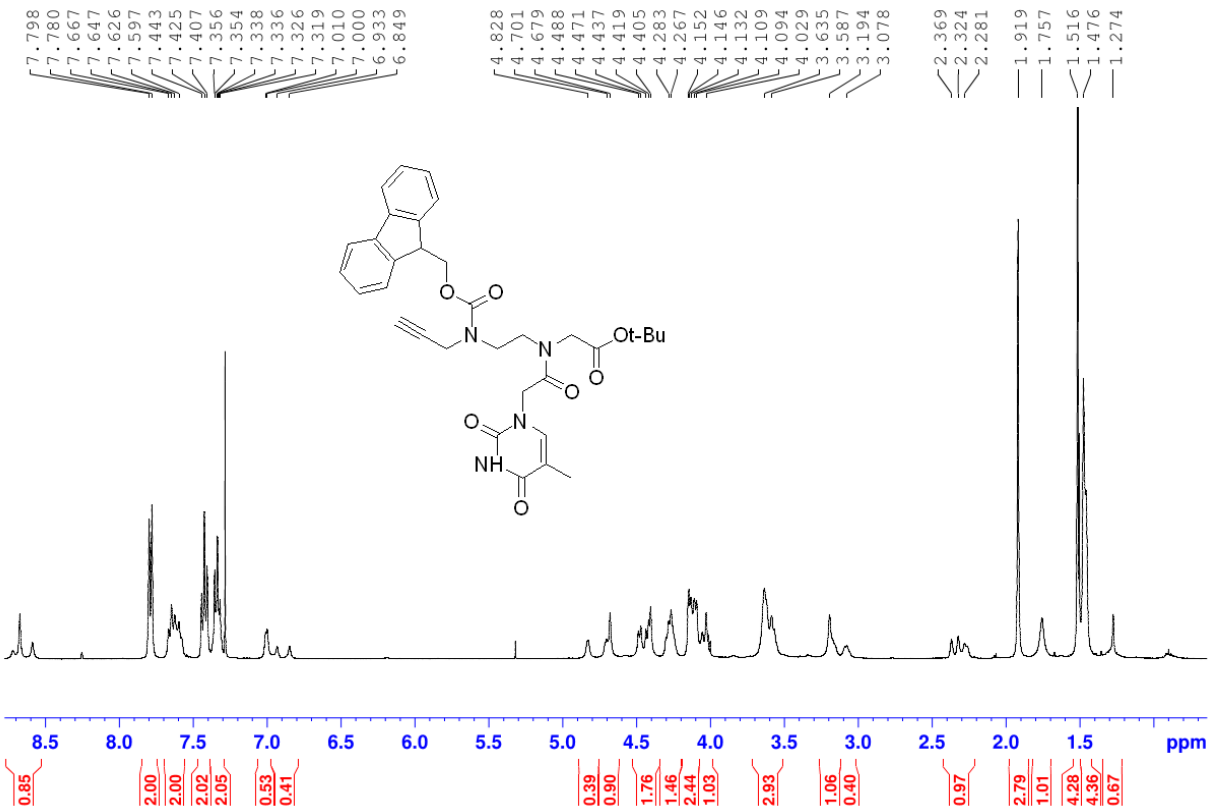


Figure 19. ^{1}H NMR of the final compound.

REFERENCES

1. “Peptide nucleic acid conjugates: Synthesis, properties and applications,” Zhilina, Z. V.; Ziemba, A. J.; Ebbinghaus, S. W. *Curr. Top. Med. Chem.* **2005**, *5*, 1119-1131.

2. “Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide,” Nielsen, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497-1500.

3. “Recent highlights in modified oligonucleotide chemistry”, Cobb, A. J. A. *Org. Biomol. Chem.* **2007**, *5*, 3260-3275.

4. Crooke, S. T. *Antisense Drug Technology* (2nd Edition) CRC Press, Boca Raton, 2008.

5. “Click Chemistry: Diverse Chemical Function from a Few Good Reactions”, Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004-2021.

6. “The growing applications of click chemistry”, Moses, J. E.; Moorhouse, A. D. *Chem. Soc. Rev.* **2007**, *36*, 1249-1262.

7. “PNA hybridizes to complementary oligonucleotides obeying the Watson-Crick hydrogen-bonding rules,” Egholm, M.; Buchardt, O.; Christensen, L.; Behrens, C.; Freier, S. M.; Driver, D. A.; Berg, R. H.; Kim, S. K.; Norden, B.; Nielsen, *P. E. Nature* **1993**, *365*, 566-568.

8. “Inhibiting gene expression with peptide nucleic acid (PNA)—peptide conjugates that target chromosomal DNA,” Hu, J.; Corey, D. R. *Biochemistry* **2007**, *46*, 7581-7589.

9. “Use of peptide nucleic acids to manipulate gene expression in the malaria parasite plasmodium falciparum,” Kolevzon, N.; Narsereddin, A.; Naik, S.; Yavin, E.; Dzikowski, R. *PLoS ONE* **2014**, *9,* 1-9.

10. “Novel antisense and peptide nucleic acid strategies for controlling gene expression,” Braasch, D. A.; Corey, D. R. *Biochemistry* **2002**, *41*, 4503-4510.

11. “Inhibition of gene expression inside cells by peptide nucleic acids: effect of mRNA target sequence, mismatched bases, and PNA length,” Doyle, D. F.; Braasch, D. A.; Simmons, C. G.; Janowksi, B. A.; Corey, D. R. *Biochem*istry **2001**, *40*, 53-64.

12. “Developments in cell-penetrating peptides as antiviral agents and a vehicle for delivery of peptide nucleic acid targeting hepadnaviral replication pathway,” Ndeboko, B.; Hantz, O.; Lemamy, G. J.; Cova, L. *Biomolecules* **2018**, *8*, 1-13.

13. “Antiviral effects of PNA in duck hepatitis B virus infection model,” Chen, Z.; Cheng, A.; Wang, M.; Xu, D.; Zeng, W.; Li, Z. *Acta Pharm. Sinca* **2007**, *10*, 1652-1658.

14. “Antitumor effects of EGFR antisense guanidine-based peptide nucleic acid in cancer models,” Thomas, S. M.; Sahu, B.; Rapireddy, S.; Bahal, R.; Wheeler, S. E.; Procopio, E. M.; Kim, J.; Joyce, S. C.; Contrucci, S.; Wang, Y.; Chiosea, S. I.; Lathrop, K. L.; Watkins, S.; Grandis, J. R.; Armitage, B. A.; Ly, D. H. *ACS Chem. Biol.* **2013**, *8*, 345-352.

15. “Peptide nucleic acids targeting mitochondria enhances sensitivity of lung cancer cells to chemotheraphy,” Chen, S.; Tu, X.; Xie, L.; Xiong, L.; Soung, J.; Ye, X. *Am. J. Transl. Res.* **2018**, *10*, 2940-2948.

16. “Sequence selective double strand DNA cleavage by peptide nucleic acid (PNA) targeting using nuclease S1,” Demidov, V.; Frank-Kamenetskii, M. D.; Egholm, M.; Buchardt, O.; Nielsen, P. E. *Nucleic Acids Res.***1993**, *21*, 2103-2107.

17. “Therapeutic peptide nucleic acids: principles, limitations, and opportunities,” Quijana, E.; Bahal, R.; Ricciardi, A.; Saltzman, W. M.; Glazer, P. M. *Yale J. Biol. Med.* **2017**, *90*, 583-598.

18. “Gamma-substituted peptide nucleic acids constructed from L-lysine are a versatile scaffold for multifunctional display,” Englund, E. A.; Apella D. H. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 1414-1418.

19. “Engineering folate—drug conjugates to target cancer: from chemistry to clinic,” Vlahov, I. R.; Leamon, C. P. *Bioconjugate Chem.* **2012**, *23*, 1357-1369.

20. “Receptor-targeted gene delivery viafolate-conjugated polyethylenimine,” Guo, W.; Lee, R. J. *AAPS PharmSci*. **1999**, *1*, 1-7.

21. “Defined folate-PEG-siRNA conjugates for receptor-specific gene silencing,” Dohmen, C.; Fronhlich, T.; Lachelt, U.; Rohl, I.; Vornlocher, H.; Hadwiger, P.; Wagner, E. *Mol. Ther.- -Nucleic Acids.* **2012**, *1*, 1-6.

22. “Click chemistry with DNA,” El-Sagheer, A. H.; Brown, T. *Chem. Soc. Rev*. **2010**, *39*, 1388-1405.

23. “Synthesis of N-propynyl analogue of peptide nucleic acid (PNA) monomers and their use in the click reaction it prepare N-functionalized PNAs,” Howarth, N. M.; Ricci, J. *Tetrahedron* **2011**, *67*, 9588-9594.

24. “Clickable peptide nucleic acids (cPNA) with tunable affinity,” Chouikhi, D.; Barluenga, S.; Winssinger, N. *Chem. Commun.* **2010**, *46*, 5476-5478.

25. “Formation of substituted pyrroles via an imine condensation/aza-claisen rearrangement/imine – allene cyclization process by MAOS,” Bremner, W. S.; Organ, M. G. *J. Comb. Chem.* **2008**, *10*, 142-147.

26. “Process improvement on the synthesis of lipoglycopeptide antibiotic telavancin,” Liang, Y.; Feng, W. *Hecheng Huaxue* **2011**, *19*, 550-553.

27. “Design, synthesis and evaluation of small molecule CD4-mimics as entry inihibitors possessing broad spectrum anti-HIV-1 activity,” Curreli, F.; Belov, D. S.; Ramesh, R. R.; Patel, N.; Altieri, A.; Kurkin, A. V.; Debnath, A. K. *Bioorg. Med. Chem.* **2016**, *24*, 5988-6003.

28. Tang, H.; Pio, B.; Teumelsan, N.; Pasternak, A.; Dejesus, R. Preparation of 2-oxopiperazines as inhibitors of the renal outer medullary potassium channel. WO Patent 2013062892, May 2, 2013.

29. “Peptide nucleic acid with a lysine side chain at the beta-position: synthesis and application for DNA cleavage,” Sugiyama, T.; Kuwata, K.; Imamura, Y.; Demizu, Y.; Kurihara, M.; Takano, M.; Kittaka, A. *Chem. Pharm. Bull.* **2016**, *64*, 817-823.

30. “Peptide coupling reagents, more than a letter soup,” El-Faham, A.; Albericio, F. *Chem. Rev.* **2011**, *111*, 6557-6602.

31. “Peptide nucleic acids: protocols and applications,” Neilsen, P. E., Editor, Horizon Bioscience, 2nd Ed, 2004, Wymondham, UK.

32. “Synthesis of peptide nucleic acid monomers containing the four natural nucleobases: thymine cytosine, adenine, and guanine and their oligomerization,” Dueholm, K. L.; Egholm, M.; Behrens, C.; Christensen, L.; Hansen, H. F.; Vulpius, T.; Petersen, K. H.; Berg, R. H.; Nielsen, P. E.; Buchardt, O. *J. Org. Chem.* **1994**, *59*, 5767-5773.

33. “Oligonucleotide analogues with integrates bases and backbone. Part 20,” Piefer, M.; Giacomo, F. D.; Schandl, M.; Vasella, A. *Helv. Chim. Acta* **2009**, *92*, 1134-1166.