Nucleoside Analogues that Mimic Water in the Minor Groove

Author: Alena Carlson

Boston College

The Graduate School of Arts and Sciences

Department of Chemistry

NUCLEOSIDE ANALOGUES THAT MIMIC WATER IN THE MINOR GROOVE

a thesis

by

ALENA CARLSON

submitted in partial fulfillment of the requirements

for the degree of

Master of Science

December 2012

Nucleoside Analogues that Mimic Water in the Minor Groove

Alena Carlson

Advisor: Prof. Larry W. McLaughlin

Abstract

Hydration in the minor groove of B-form DNA influences the stability and conformation of duplex DNA. An analogue nucleoside that attempts to mimic the N3-hydrated adenine was designed and synthesized. Obtaining the modified adenine was a signifcant synthetic effort. The attempted incorporation of the new analogue in a DNA strand yielded inconclusive results, and further tests are ongoing. The incorporation of the new modified purine in a DNA duplex may be useful for a better understanding of the hydration of B-form DNA.

Acknowledgements

I am grateful to Professor Larry W. McLaughlin for accepting me into his lab and for his support over the past two years. His scientific insight and life wisdom inspired and guided me.

I would like to thank the members of my committee, Professors Mary Roberts and Jianmin Gao for taking time to consider my thesis and to attend my oral examination.

I would like to express my gratitude to all the group members of the McLaughlin lab for their generous help in the lab and emotional support. I would like to thank Mark Schlegel for carefully reviewing my thesis. I also appreciate the suggestions made by Carl Christianson.

I thank all the students and employees in the Boston College Chemistry Center. It has been an unforgettable experience.

Table of Contents

Abbreviations and Acronyms

A	Adenine
Ac	Acetyl
AIBN	Azobis(isobutyronitrile)
Bn	Benzoyl
Calcd	Calculated
CD	Circular Dichroism
dA	2’-Deoxyadenosine
dC	2’-Deoxycytidine
dc^3A	3-Deaza-2’-deoxyadenosine
dhm^3c^3A	3-deaza-3-hydroxymethyl-2’-deoxyadenosine
dG	2’-Deoxyguanosine
dm^3c^3A	3-Deaza-3-methyl-2’-deoxyadenosine
DBU	1,8-Diazabicyclo[5.4.0]undue-7-ene
DCM	Dichloromethane
DIPEA	Diisopropylethylamine
DMAP	4-(Dimethylamino)pyridine
DMF	N,N-Dimethylformamide
DMT	4,4’-Dimethoxytrityl
DMSO	Dimethylsulfoxide

DNA	β-D-2’-Deoxyribonucleic Acid
dNTP	2’-Deoxyribonucleoside 5’-Triphosphate
Et_2O	Diethyl Ether
EtOAc	Ethyl Acetate
Equiv	Equivalents
G	Guanine
h	Hour(s)
Hex	Hexanes
HPLC	High Performance Liquid Chromatography
HRMS	High Resolution Mass Spectrum
M_4SA	2,2,3,3,-Tetramethylsuccinic Anhydride
M_4SI	2,2,3,3,-Tetramethylsuccinimide
M	Molar
min	Minute(s)
NMR	Nuclear Magnetic Resonance
NBS	N-Bromosuccinimide
PAGE	Polyacrylamide Gel Electrophoresis
Pyr	Pyridine
R_f	Retention Factor
rt	Room Temperature
T	Thymine

TBAF	*n*-Tetrabutylammonium Fluoride
TBS, TBDMS	*tert*-Butyldimethylsilyl
TCA	Trichloroacetic Acid
TEA	Triethylamine
T_f	Triflate (Trifluoromethanesulfonyl)
TFA	Trifluoroacetic Acid
TLC	Thin Layer Chromatography
TMS	Trimethylsilyl
Tol	*para*-Toluoyl
RNA	β-D-Ribonucleic Acid
U	Uridine
UV	Ultraviolet

1.1. The Roles of Nucleic Acids

Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are essential to the central dogma of molecular biology introduced by Francis Crick.[1] It describes how information is transferred and stored in living organisms.[2] The study of the information carriers – nucleic acids – has become a journey consisting of combined efforts of many scientists and ultimately changed the way we live.

From the discovery[3-5] of helical structure of DNA to the human genome sequencing project, the path has been very controversial and many ethical questions about the consequences of the findings have been raised. DNA technology not only transformed the understanding of diseases and treatments, but also crossed the boundaries of science and medicine affecting all areas of our life. Changes in the world of legal matters and forensic science[6] have come as an immediate application of the new discoveries. Scientists are constantly searching for new ways of making sequencing cheaper and faster. Next-generation sequencing platforms have emerged and enabled a wide variety of applications, "allowing researchers to ask virtually any question of the genome, transcriptome, and epigenome of any organism."[7] We are looking at the future where the sequencing of a patient's genome in a doctor's office, "while-you-are-waiting", is a routine procedure. Possible uses of this sensitive information raise the concerns of privacy and moral.

In addition to such important practical applications of DNA technology as the one towards the diagnosis of diseases and production of new pharmaceuticals, agricultural, environmental, and engineering opportunities are hard to overestimate. The DNA duplex structure has a great potential in the construction of nanodevices.[8] The self-assembly and sequence specific recognition of DNA can lead to miniature electronics and optical devices, biosensors and probes. Genetic advances in agriculture increased the production of food and helped in the fight against hunger on the planet with the population of seven billion people. Yet the long-term effects of genetically modified produce are unclear. Some of the risks amount to "potential human health impacts, including allergens, transfer of antibiotic resistance markers, and unknown effects."[9]

Another reason to study DNA is to understand evolution. Analysis of a genetic sequence has become a much more powerful method than examination of physical traits.[10-12] It turns out that millions and millions of species on the Earth are all related and are made from the same four building blocks of DNA.

The understanding of DNA and RNA three-dimensional structures has advanced dramatically since their initial discovery.[13,14] This progress has been largely due to the increased power and resolution capabilities of X-ray crystallography.[15] But more recently other important methods have emerged and play important roles in providing structural information: NMR spectroscopy,[16] molecular modeling, and chemical/biochemical probe techniques.[17]

2.2. Structural Features of Nucleic Acids

A detailed description of structural parameters in nucleic acids can be found in many sources;[14,18,19] here, only the elements that are essential to the title project are considered. Fundamentally, the four nucleosidic bases in DNA and RNA that are responsible for coding the genetic information are adenine, cytosine, guanine, and thymine for DNA (**Figure 1A**) and adenine, cytosine, guanine, and uracil for RNA (**Figure 1B**).

The sugar-phosphate backbone bridges the nucleosides creating a long polymer that finds another, complementary strand to pair with. The right-handed duplex of natural DNA forms due to hydrogen bonding between the nucleobases, modulation of electrostatic repulsion by cations, and hydrophobic static interactions being the major factors. The asymmetry in base pairs results in two parallel types of grooves: major and minor (**Figure 2**).

DNA has an overall negative charge due to the phosphodiesters in the backbone. Phosphoric acid is unique[2] due to several important factors. Specifically, it can "link the two nucleotides and still ionize," stabilizing the diesters against hydrolysis and retaining the DNA molecule within a lipid membrane. The negative charge metal ions that manipulate the processes of transcription and replication.

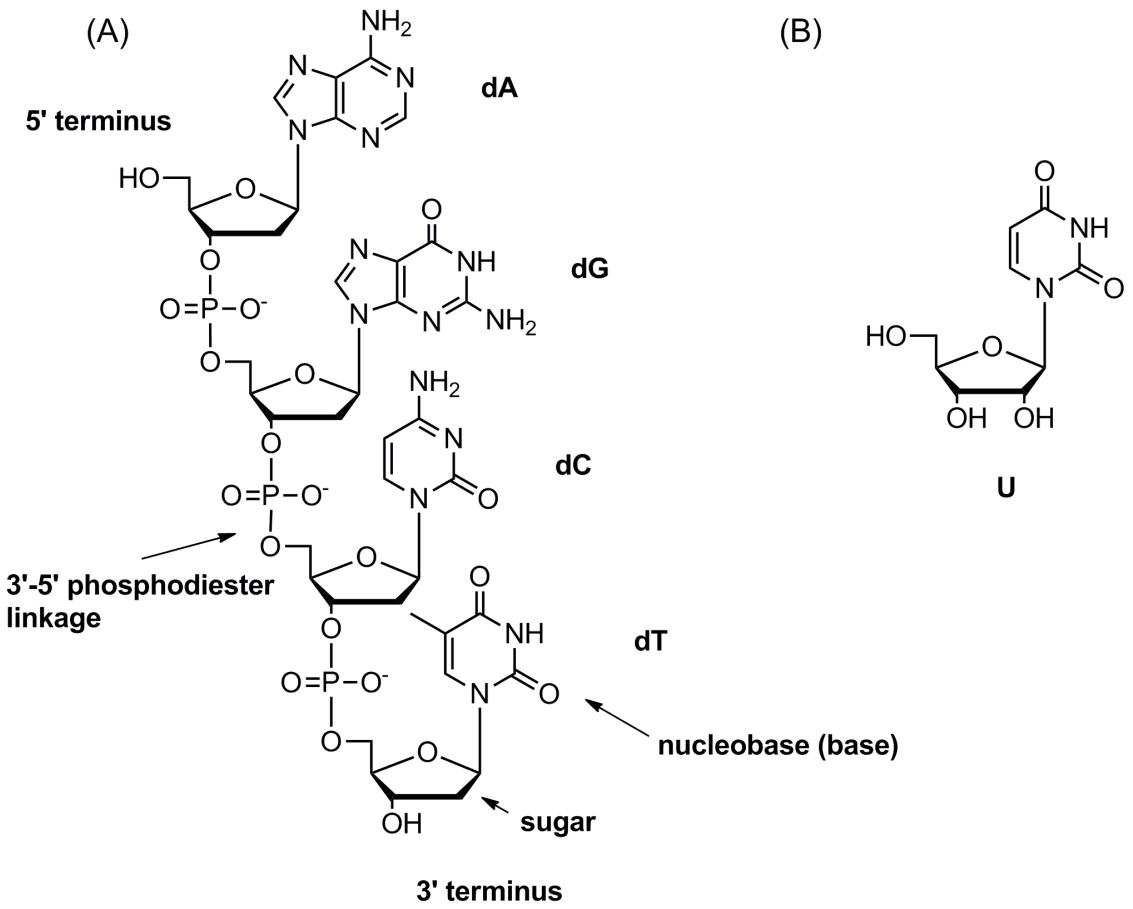


Figure 1. The Nucleosidic Bases in DNA (A) and RNA (B)

In spite of its high level of stability, DNA is still fairly flexible. The duplex can attain several conformations – a phenomenon termed polymorphism. The most common forms are the A-form and B-form DNA, with the B-form the most common at physiological conditions.

The favored structure of a duplex is determined by many forces. DNA duplexes are not static structures; they are equilibrium states of dynamic structures. DNA can

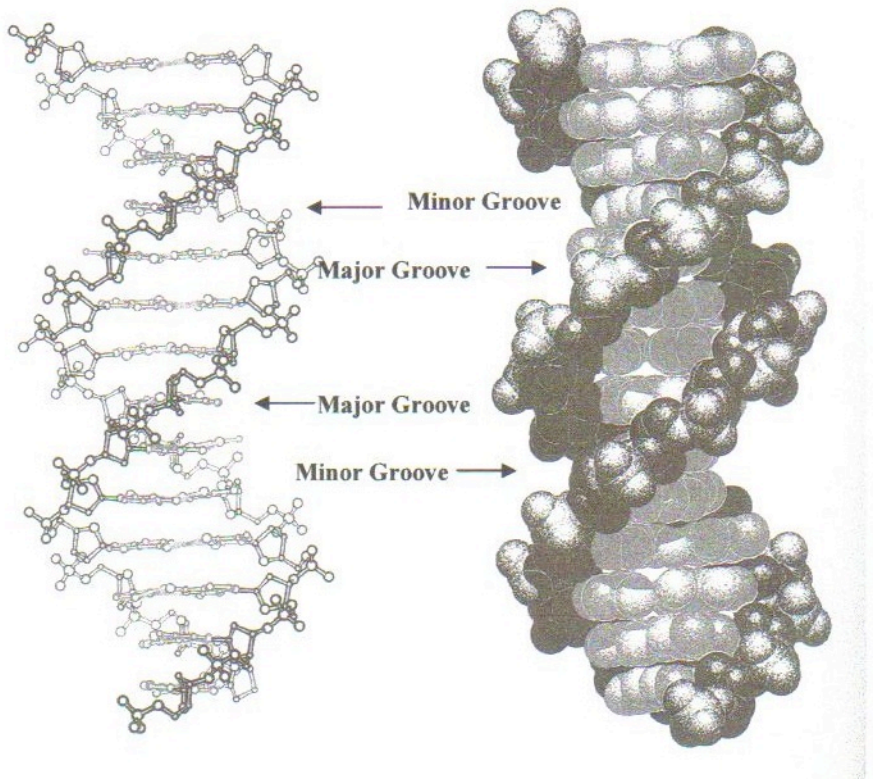


Figure 2. DNA duplex shown in ball-and-stick and space filling representations. Images adapted from Sanger[14].

compress, stretch, bend under the influence of many factors: DNA sequence, salt content, metal ion binding, temperature, protein interactions, and hydration. Base-complementarity is one of the most important factors for the formation and stability of the duplex. Base-stacking or pi-stacking is an important stabilizing factor when p-orbitals of the aromatic rings of the heterocycles interact. Since DNA has an overall negative charge, the repulsion of the phosphate groups can be mitigated by binding cations. This project aims to study hydration of a DNA duplex as a major factor determining its form and stability.

Since structure dictates function, the overall shape of a duplex directs its interactions with proteins and growth factors. In general, the major groove of B-DNA is more accessible and binds a wider range of molecules compared to the minor groove. The major and minor grooves differ in functional groups, electrostatic potential, sterics, and hydration. The minor groove binding ligands often contain aromatic rings that become isohelical with the groove (repeating the curve of the helix), for example, netropsin, distamycin, berenil.[14b]

DNA structure and conformation is highly dependent on the sequence. A-DNA is rich in G-C base pairs (**Figure 3**); A-T rich sequences demonstrate B-DNA parameters. If a DNA duplex has a run of several consecutive A-T base pairs, called A-tracts,[21] the minor groove becomes unusually narrow, creating a special electrostatic binding pocket for positively charged ligands. Several crystallographic studies have reported that the width of the minor groove in A-T regions is about 2.8-3.2 Å compared to 6.0 Å in average B-DNA. Since the DNA minor groove is narrower in A-T rich sequences, aromatic rings in the minor groove binding ligands fit better and make optimal van der Waals contacts with the walls of the groove. Additional specificity is guaranteed by the functional groups: specifically the C2 carbonyl oxygen of thymine and the N3 nitrogen of adenine.

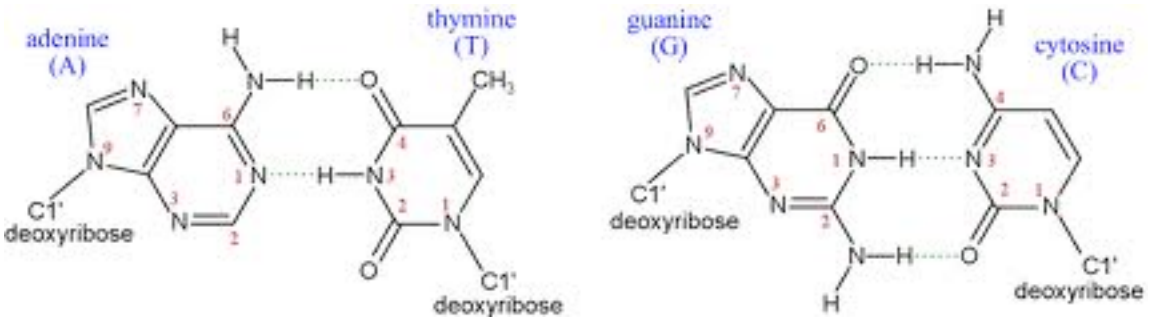


Figure 3. The Watson-Crick base pairs

In a G-C base pair, the amino group of guanine presents a steric block for the binding of small molecules. An adenine modification synthesized in this project continues the study of the minor groove A-T rich sequences.

Nucleoside Analogues

Modified nucleosides and nucleotides are found in nature and can also be obtained synthetically. While biological molecules participate in regulation, defense, information transfer,[2,14] etc. in organisms, synthetic nucleosides, in addition to medicinal purposes, are widely used as probes[22] to study biological mechanisms.

Modified bases occur naturally mostly in transfer RNAs (tRNAs). The numerous modifications[23] are involved in diversifying the genetic code, synthetase recognition, and structure stabilization. A simple example of a modification is DNA methylation.

Natural DNA contains 5-methylcytosine, N6-methyladenine, and others. There are many biological functions of methylation including gene expression, the discrimination of self-DNA from foreign-DNA, and base-base mismatch repair.

There are many applications of nucleobase modifications in research and medicine. A duplex containing 5-propynyl-dU is relatively more stable because stacking with the neighboring bases is improved due to an extended π-structure (**Figure 4**).[24,14b] A greater understanding of protein and drug binding in the Hoogsteen face of guanine has improved by the synthesis and analysis of a 7-deaza-dG residue and testing it in a duplex. 2-Aminopurine is a fluorescent base and can be used as a probe of nucleic acid conformations. 5-Bromo-dU and 5-iodo-dU undergo photolytic cross-linking and are used to cross-link DNA and proteins.

Backbone and sugar modifications have been important in antisense therapies which involve steric blocking or degradation of the coding strand of RNA through the formation of a base paired duplex and subsequent inhibition of the gene expression. A lot of anti-viral and anti-HIV compounds are chain terminators where the 3'-OH group is absent so that phosphorylation cannot take place. Some examples[25] of these biologically active modified nucleosides are zidovudine, zalcitabine, and abacavir. Other nucleoside modifications (2'-modifications[26], PNA[27], LNA[28], phosphorothioates,[29] etc.) have become potent antivirals and antisense agents for the control of gene expression.

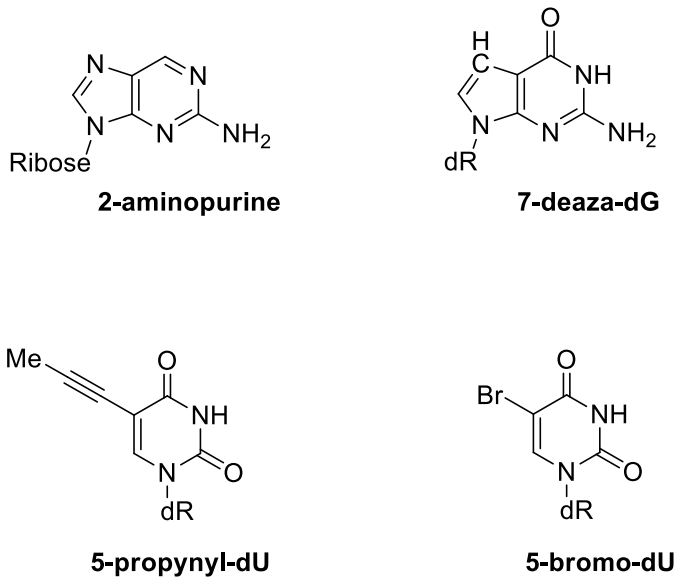


Figure 4. Nucleoside analogues used in structural studies

3-Deazapurine are the analogues used in this project. The molecules with the modification occur in nature and have antiviral and antibiotic activity.[30] In the lab, 3-deazapurines have been used as various biological probes. For example, 3-deaza-3-

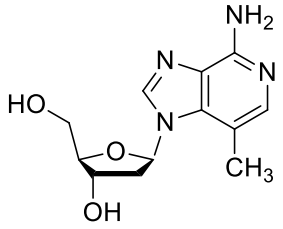


Figure 5. Anti-fixed purine analogue, 3-deaza-3-m3thyladenosine

methyladenosine (**Figure 5**) has been used as an anti-fixed purine analogue,[31] and as an analogue of N3 methylated adenosine, a cytotoxic lesion in DNA.[32] This project explores the synthesis of a new 3-deaza-3-substituted adenine and its ability to probe the structure of a DNA duplex.

1.3. Steps in Chemical Synthesis of Nucleosides and DNA, and the Use of Tetramethylsuccinimide Protecting Group in Glycosylation

Chemical synthesis of DNA[33-36] is usually done automatically using phosphoramidite chemistry on a solid support (**Figure 6**). The resulting single-stranded chain consists of a number of building blocks—nucleosides—connected with a phosphodiester bridge[14] (see section 1.2). Many types of nucleosides, both natural and unnatural, are commercially available and are ready to be incorporated into oligonucleotides. The challenge of this project, as well as several other projects in the McLaughlin group, was to synthesize and use novel modified nucleosides. The new modifications can be designed so that the DNA duplexes have new and useful properties.

There are several prevalent strategies[14,37] in the synthesis of modified nucleosides. The glycosydic bond can be created by a nucleophilic attack on the sugar by a nitrogen or a carbon atom of a modified nucleobase. Alternatively, the base can be constructed starting at the prebuilt nitrogen or carbon on the sugar. When the former method A is chosen (as it was in this project, see Chapter 2) three major stages of synthesis take place. First, the required modification is installed on the nucleobase, then the glycosydic bond is formed, and finally the formation of the corresponding phosphoramite.

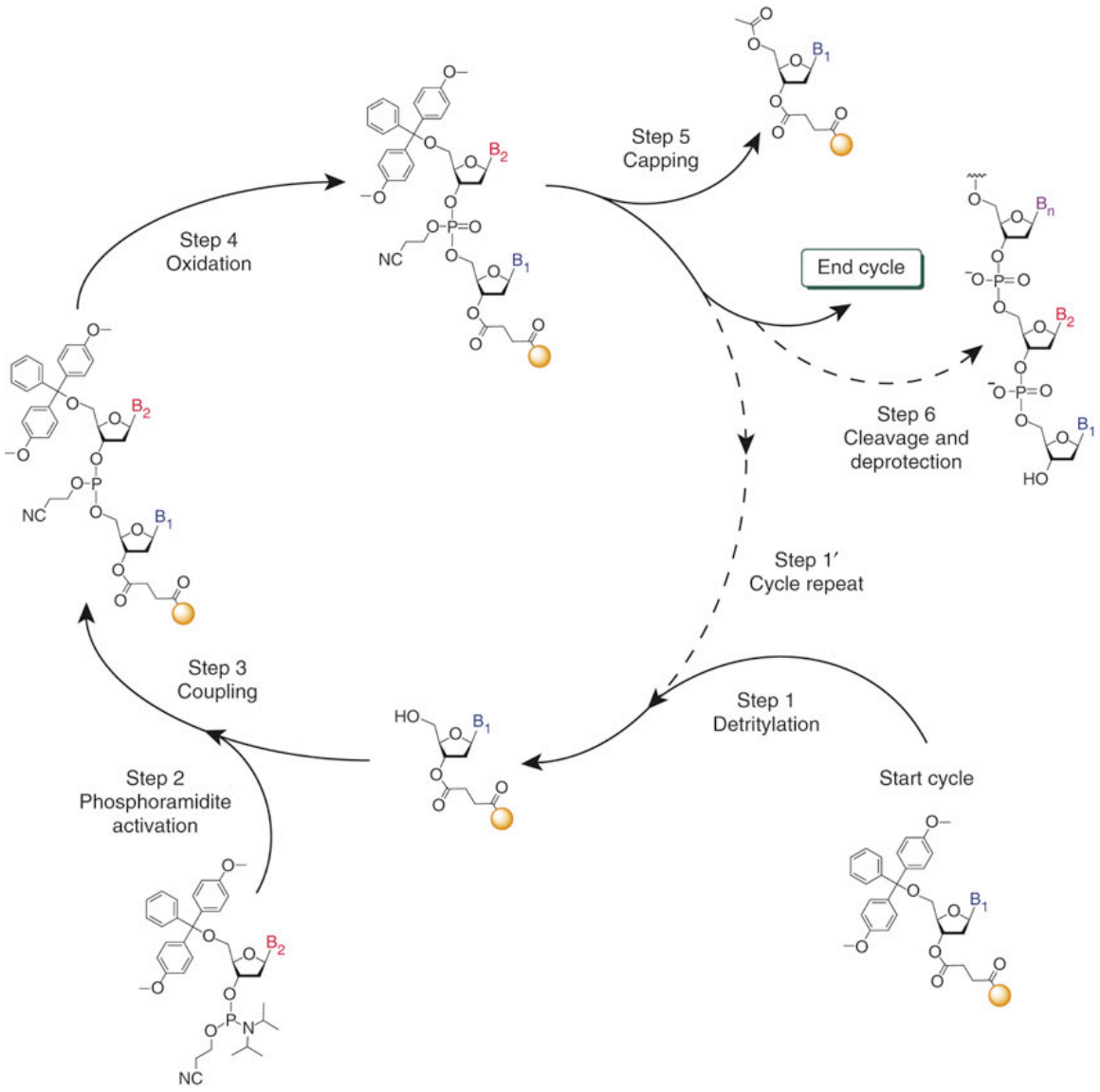
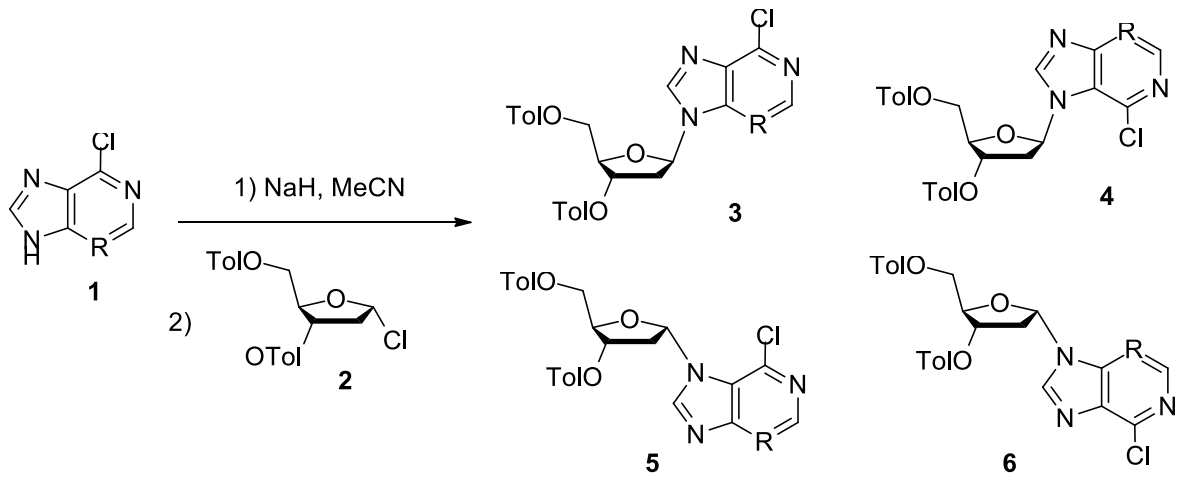


Figure 6. Steps in solid phase DNA phosphoramidite synthesis cycle[36].

Formation of the glycosydic bond presents a number of challenges. Natural nucleosides have a β-configuration at C1 carbon, and in most of the synthetic methods at least some of the unwanted α-isomer is formed. On the other hand nitrogen atoms on the heterocyclic bases are similarly nucleophilic, which means that glycosylation can take place at both N9 and N7 atoms (in purines). Obtaining the desired nucleoside with

high regio- and stereoselectivity, and in good yields became the goal of many researchers including those in McLaughlin group.

Scheme 1 illustrates how in a purine glycosylation four different products can be obtained. The amount of the desired isomer varies depending on the type of substrate.



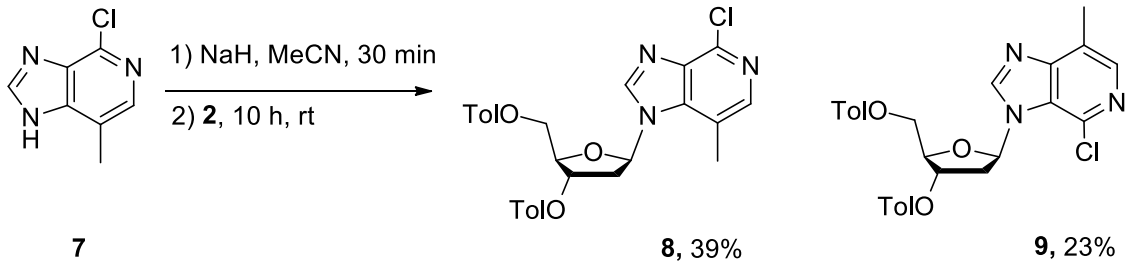
Scheme 1. Possible products in the glycosylation using sodium salt method[14]

In addition to the negative effects of the formation of multiple isomers on the overall yields, the difficulty of glycosylation is emphasized at the chromatographic stage as the products are very hard to separate.

One of the most used methods in the glycosylation of purines and deazpurines is the sodium salt method.[38] In this method, the nucleobase anion is formed in the reaction with sodium hydride (or potassium hydride) in acetonitrile. The resulting salt reacts with chlorosugar (2-deoxy-3,5-di-O-(4-toluoyl)-D-ribofuranosyl chloride) via an S_N2 reaction. The major product is the desired β-2'-deoxyribonucleoside with the correct

configuration and regioselectivity. The high yield is the consequence of a much faster nucleophilic attack of the purinyl anion and subsequent displacement of chloride in comparison with the anomerization of the chlorosugar. The small amount of oxonium ion that does form presents an opportunity for nucleophilic attack from above and below the plane of the sugar which leads to the formation of the unwanted α-isomer.

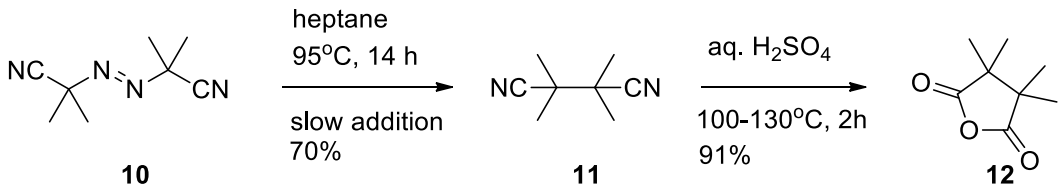
The glycosylation of nucleoside analogue **1** (**Scheme 1**) proceeds with the formation of the desired β-N9 isomer as the major product due to favorable electronics and sterics. In the case of 3-deaza-3-substituted purines, such as **7** (**Scheme 2**), the electronics and sterics are not favorable for providing the correct selectivity. A series of attempts were undertaken in the McLaughlin lab in order to synthesize 3-deaza-3-substituted purine nucleosides[39] (**Scheme 2**). Amination of **8** proved to be a more difficult task in comparison with ordinary nucleosides. Usually, heating with ammonia installs the amino group at the 6 position, simultaneously removing Tol protecting groups. For 3-deaza-3-substituted purine nucleosides, the conditions must be much harsher: an overnight treatment with hydrazine followed by a 6 h reaction with Raney nickel.[31]



7 **8,** 39% **9,** 23%

Scheme 2. Glycosylation of 3-deaza-3-methyl-6-chloroadenine[39]

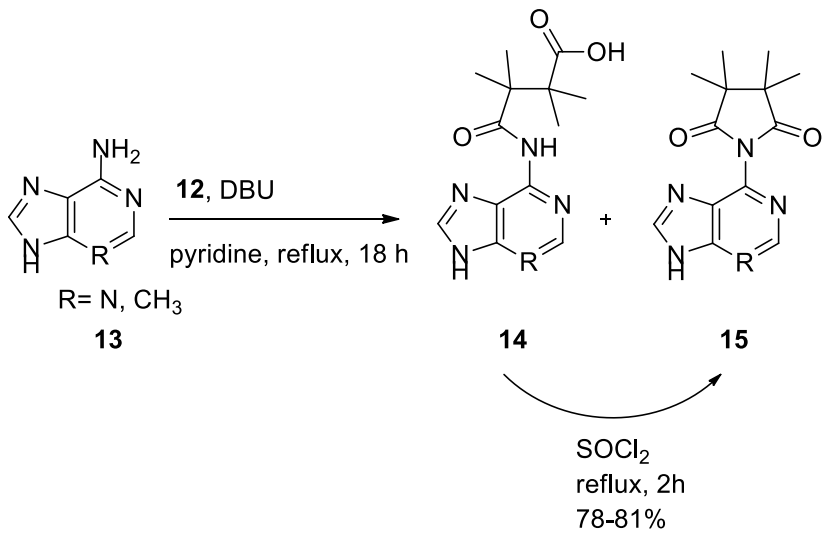
In the McLaughlin lab, the improvements to the Irani procedure[31] were made and resulted in higher yields of the N9 product. Alternate coupling conditions were examined, as well as modifications to **7**. The amine was installed before coupling and protected in order to make the 6 position more bulky. A directing/protecting group with sufficient steric bulk in the 6 position was found to be the tetramethylsuccinimide (M_4SI) group.[40]



Scheme 3. The synthesis of 2,2,3,3-tetramethylsuccinic anhydride

The synthesis[41] of the M_4SI group is shown in **Scheme 3**. The starting material for this procedure is AIBN **10**, a reagent commonly used as a radical initiator. Slow addition of AIBN to hot heptane affords dinitrile **11**, which is suspended in aqueous H_2SO_4 and heated for two hours to obtain tetramethylsuccinic anhydride M_4SA **12**. This compound is not available commercially, but its preparation is a relatively easy procedure because the crude product does not require chromatography and can be used immediately for the next step.

M_4SA reacts with adenine and other purines (including deaza-purines) in refluxing pyridine in the presence of DBU (**Scheme 4**). A small amount of the resulting amide can be converted to the imide with refluxing $SOCl_2$ in 78-81% yields of the M_4SI-protected



Scheme 4. Installing the M_4SI protecting group[41]

purine. The crystal structure of M_4SI-protected adenine shows that the methyl groups are positioned above N7 of adenine and block the access of this position for the chlorosugar.

M_4SI performed well with 3-deaza-3-substitued purines, which is why this protecting group was chosen for this project. The M_4SI protecting group is compatible with DNA solid-phase synthetic procedures and can be cleanly removed with ammonium hydroxide at 55°C. An important drawback is that a substrate like 3-deaza-3-methyladenine protected with M_4SI is a more hindered nucleophile, therefore the reaction is slower and larger amounts of α-nucleosides are formed. In spite of this

disadvantage, M_4SI successfully fulfills the roles of a directing/protecting group in case of 3-deaza-3-substituted purines allowing for the synthesis of β-purines in good yields.

1.4. Hydration as a Structural Feature of B-DNA

Water is not just a medium to keep the solutes dissolved. It is also responsible for stabilization of overall structure. In nucleic acids, phosphate—phosphate electrostatic repulsion is diminished due to high dielectric constant of water. Bases self-assemble into ordered structures at least partially due to hydrophobic forces, and as a consequence of entropy reduction driven by water molecules.

Many studies,[18, 42-44] including X-ray crystallography, NMR, and IR have shown that DNA duplex is hydrated, and that the hydration is not uniform around the molecule. The distribution of water molecules depends on the helical type and the characteristics of the grooves.[15b]

The hydration level can greatly influence a nucleic acid's conformation.[45] In an environment with relatively high humidity, DNA prefers to be in B form; A-type DNA occurs under conditions of low humidity or high salt concentration. As studies on Dickerson-Drew dodecamer have shown,[14,18] water molecules are found around the phosphates, and around nitrogen and oxygen atoms in the major groove. The minor groove contains two shells of hexagonally arranged water molecules that directly

contact the bases. The zigzag pattern is named the spine of hydration and is considered to be a strong stabilizing structural component of B-DNA (**Figure 7**).

The spine of hydration is a characteristic found in sequences rich in dA-dT base pairs. A molecule of water bridges the N3 atom of adenine and O2 carbonyl of the thymine residue directly below. A-DNA has the grooves of completely different size and electronic characteristics, thus, does not have the same hydration pattern. A-DNA typically consists of long runs of C-G base pairs. The amino group of guanine is bulky and interferes with the formation of the "regular" spine of the molecules.

The spine of hydration in A-T rich sequences stabilizes the B conformation. For this reason complex poly(dA)*poly(dT) does not undergo transformation from B to A form. Polynucleotides with guanosine replaced by inosine (lacking the N2 amino group) behave as A-T polymers and occur in B-family double helices. Theoretical calculations suggested that the presence of the cross-strand hydrogen bonding may stabilize the duplex by inhibiting base pair opening. Another study suggested that the propeller twist of the bases was sub-optimal to accommodate the spine of hydration, which compensated for the lost energy.[46]

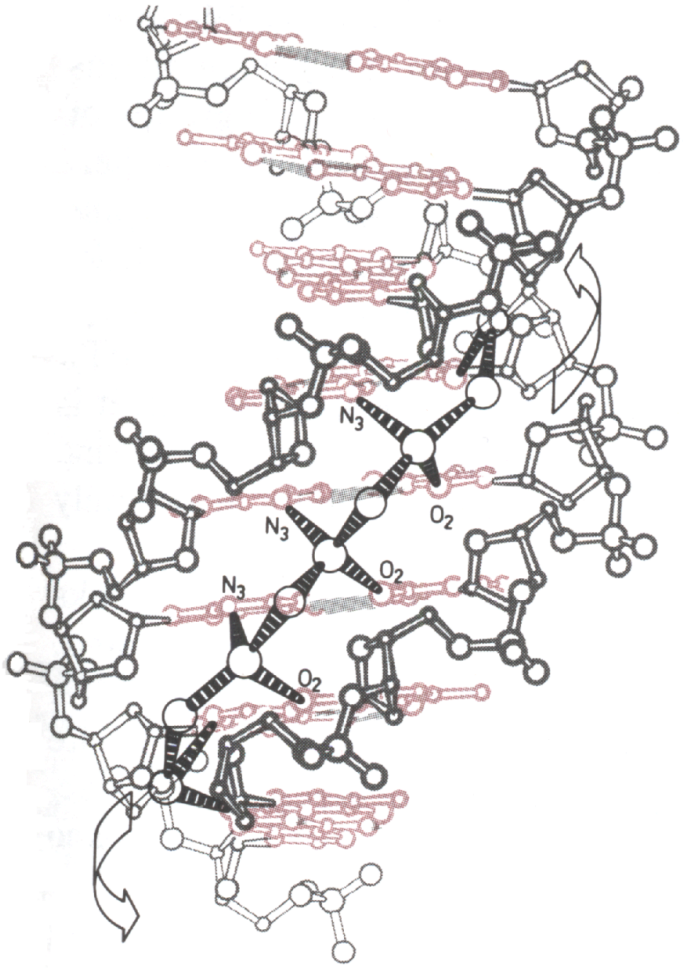


Figure 7. Drawing shows schematically how water molecules are accommodated in the minor groove of sequence AATT of the B-DNA-type dodecamer. The first hydration layer is hydrogen bonded with pyrimidine O2 and purine N3.Image adapted from Sanger.[14]

In addition to the highly ordered hydration, duplexes containing long runs of A-T base pairs have another interesting feature: the formation of so-called A-tracts.[47] In a recent review,[21e] Haran and Mohanty investigate the biological importance of these repeated runs of 4-6 adenine-thymine base pairs that create "global curvature of the

DNA duplex, which can be macroscopically characterized by anomalously slow migration on polyacrylamide gels".[21e] DNA A-tracts are defined as four or more consecutive A-T base pairs without a T-A step. Repeats of A_4T_4 are bent, whereas T_4A_4 are straight, the fact that became the topic of study for many researchers.

It has been established from many oligonucleotide crystal structures that A-T-rich minor grooves tend to be narrower than average. It has been observed[49] that the width of the minor groove determines the hydration pattern. Privé[50] correlated the high content of A-T base pairs with a narrow minor groove and the resistance of A-T-rich DNA to undergo the transition from B to A-type, and concluded that the spine brought an additional stabilization due to the greater isolation of the spine from the bulk solvent.

The exact role of hydration of DNA on its stability, geometry, ligand binding, packaging of DNA is yet to be understood. Stability of duplex DNA and rules that change it are central for understanding the many processes involving DNA. Although the existence of the spine of hydration has been proven by X-ray crystallography,[14,48] this project probes the stability of the duplex DNA indirectly.

1.5. Previous Studies of the Hydration in DNA

Most biological processes rely on molecular recognition and reversible interactions. The structure of the DNA duplex and, consequently, replication and gene expression, depend on interaction with proteins as well as ions and water. Water molecules can act as both acceptors and donors of hydrogen bonds creating extended networks. The hydration of DNA is essential for the maintenance of its structural integrity.

The role of "structural" waters in the major groove of DNA has been examined by Richard Gumport.[51] His team synthesized a nucleoside analogue depicted on **Figure 8**. The structure was characterized by X-ray diffraction and proton NMR. The crystal structure shows an intramolecular hydrogen bond between an amino proton on N6 of adenine and the oxygen of the hydroxymethyl group. When superimposed onto particular adenines in the structure of the tryptophan (trp) repressor/operator complex, the analogue places the oxygen of the hydroxymethyl group very close to the oxygen contributed by the water in the complex.

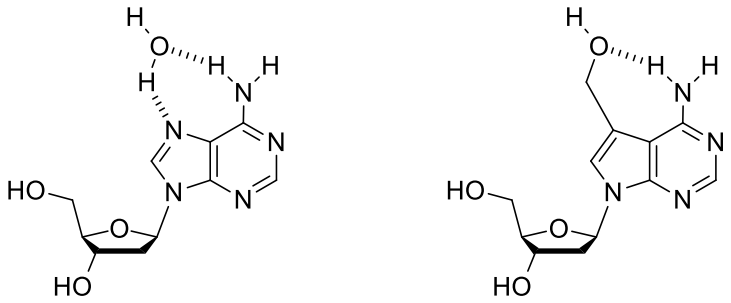
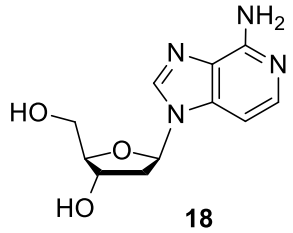


Figure 8. Water mimic by Gumport[51]

Eliminating interactions in the minor groove that are involved in the spine of hydration allows measuring its contribution to the stability of duplex. With this in mind, Richard Cosstick[53] and co-workers prepared 2'-deoxy-3-deazaadenosine (dc^3A) **18**, **Figure 9**. The incorporation of the analogue into a duplex DNA destabilized the structure.



18

Figure 9. 2'-deoxy-3-deazaadenosine (d3cA)[53]

Understanding the contribution of the hydration to duplex conformation and stability has been one of the subjects of studies in the McLaughlin lab. The next chapter discusses the results of the series of experiments including the findings of this thesis that were conducted throughout several years.

2. Results and Discussion

The McLaughlin group has a long-standing interest in studying the hydration of the minor groove. In order to probe the spine of hydration, two analogue bases **1** and **2** were synthesized[54] (**Figure 1**). 3-Deazaadenine **1** is a derivative of adenine from which N3 has been deleted and 3-methyl-2-pyridone **2** is a C-nucleoside that mimics thymine but lacks the O2 carbonyl. This base pair is similar to dA-dT but eliminates the polar

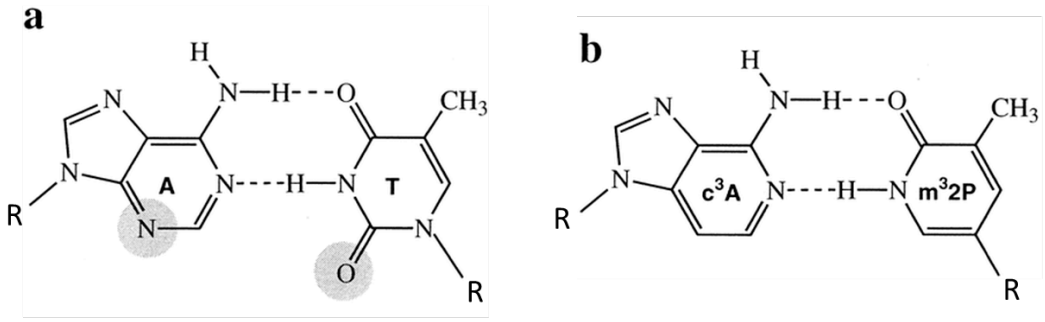


Figure 1. T. Lan analogues.[54]

functional groups in the minor groove. Ultraviolet melting temperature studies indicated that the presence of this base pair in dA-dT rich sequence results in a destabilized duplex. Moreover, circular dichroism studies have shown that sequences with four analogue base pairs structurally resemble A-form helices rather than B-form helices. On the contrary, when a single substitution is made in G-C rich sequences, the duplex stability is retained. This proves that the spine of hydration is disrupted while the incorporated analogues do not affect the fundamental duplex properties of base pairing and base stacking[54].

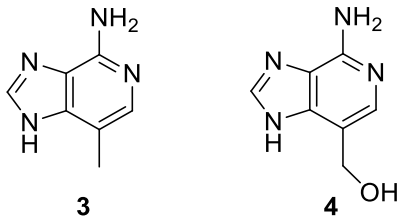


Figure 2. The modification of adenine 3-deaza-3-methyladenine (m^3c^3A) and 3-deaza-3-hydroxymethyladenine ($h^3m^3c^3A$) adds bulk into the minor groove. **4** increases the stability of a duplex compared to **3**.

In order to better understand the contribution of water molecules on the stability of B-form DNA, analogues **3**, 3-deaza-3-methyladenine, and **4**, 3-deaza-3-hydroxymethyladenine were designed, synthesized[55] (**Figures 2** and **3**), and incorporated into a DNA duplex. Both of these modified nucleobases are "missing" the N3 nitrogen and, in addition, introduce steric bulk into the minor groove.

Three types of oligonucleotides were prepared: 1) those containing 3-deazaadenine **1**, 2) those containing 3-deaza-3-methyladenine **3**, and 3) those containing 3-deaza-3-hydroxymethyladenine **4**. Thermodynamic parameters were obtained for the formation of all duplexes (see **Table 1**)[55]. The elimination of the N3 nitrogen from a dA residue

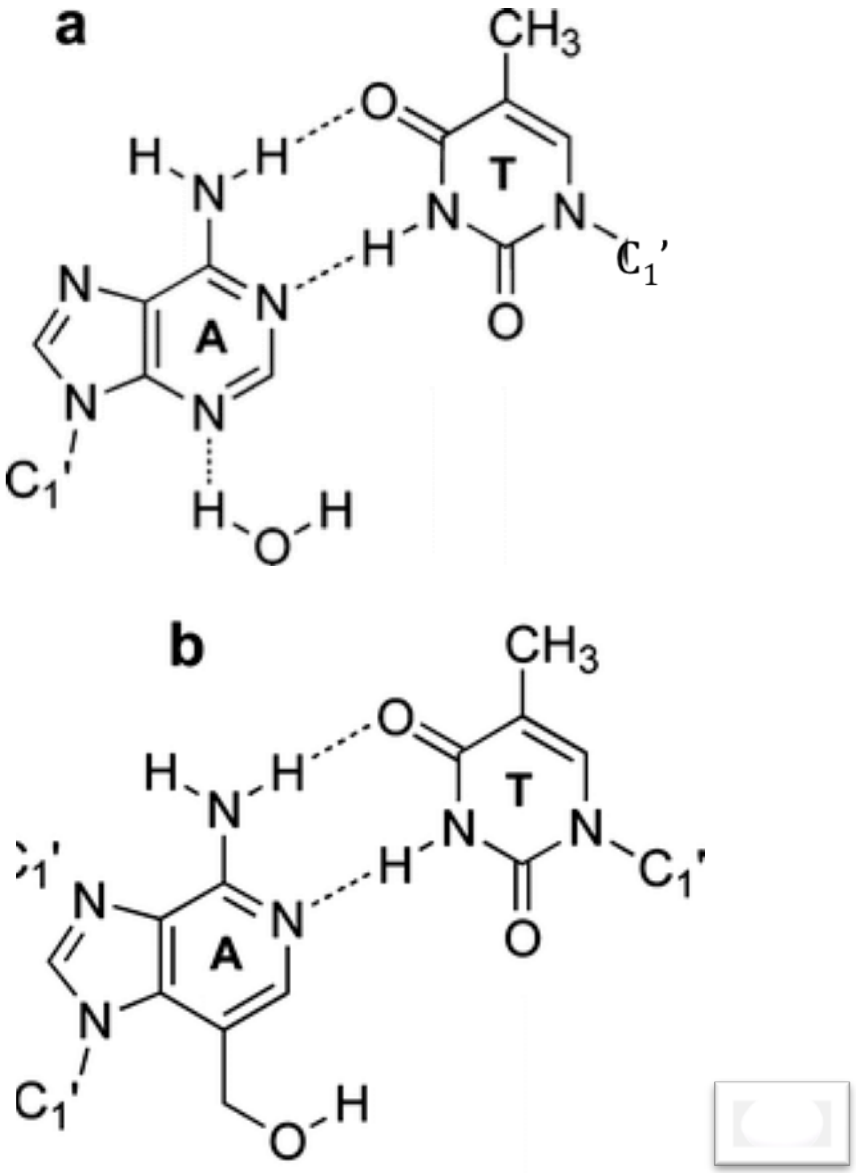


Figure 3. (a) Water molecule hydrogen-bonded to the N3 of dA as part of the minor-groove spine of hydration; (b) Covalently bound water mimic[55]

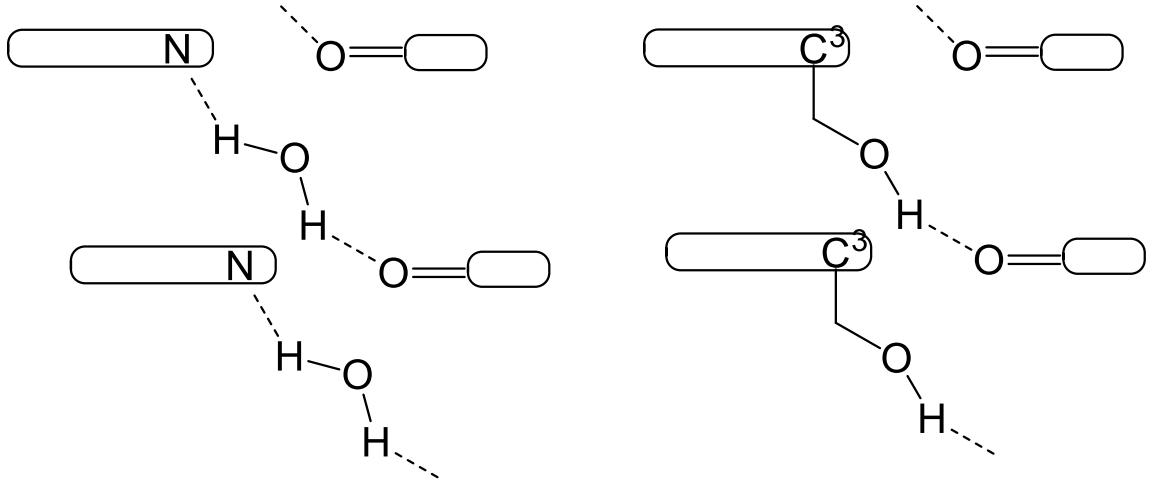
(duplex **2**) resulted in a 2.4°C decrease in the melting temperature (T_M). The introduction of four deazaadenine residues (duplex **5**) reduced the T_M by 12.3°C. The 3-deaza-3-methyl analogues destabilized the duplex (duplexes **3** and **6**, **Table 1**) to a

greater extent due to the increased steric bulk in the minor groove. Although the hydroxymethyl group is even bulkier, its incorporation in a duplex (duplex **4** and **7**) increased the T_M relative to the methyl analogue.

duplex	sequence at 5 μM duplex	T_M (°C)	Δ*G* (kcal/mol)
1	d(CCGG AAAA CGCC)	63.4	−19.1 ± 0.6
2	d(CCGG AAA^{H}A CGCC)	61.0	−16.4 ± 0.4
3	d(CCGG AAA^{CH_3}A CGCC)	58.3	−14.8 ± 0.3
4	d(CCGG AAA^{CH_2OH}A CGCC)	60.9	−16.4 ± 2.3
5	d(CCGG $(A^{H})_4$ CGCC)	51.1	−13.9 ± 0.2
6	d(CCGG $(A^{CH_3})_4$ CGCC)	44.4	−10.8 ± 0.4
7	d(CCGG $(A^{CH_2OH})_4$ CGCC)	54.7	−15.2 ± 0.7

Table 1. Thermodynamic stabilities of serlf-complementary duplexes[55]

These results show that the absence of hydrogen-bonding acceptors in the minor groove destabilize the duplex as a result of disruption of the spine of hydration. Methyl groups also disrupt the spine of hydration, and introduce additional steric bulk. Hydroxymethyl substituents, although the bulkiest, provide stabilization through a modified water structure. A possible interaction mimicking water molecule is illustrated in **Figure 4**.

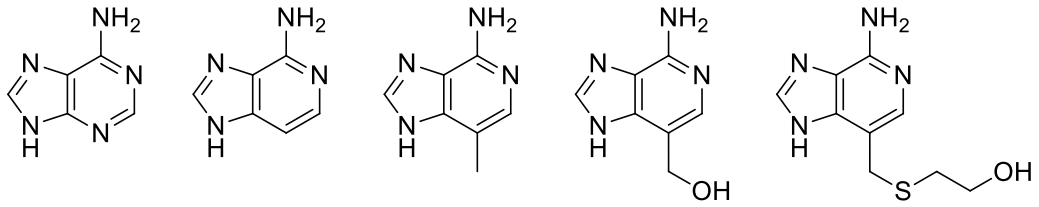


a **b**

Figure 4. (a) Representation of water molecules in the minor groove. Long ovals denote adenine, short ovals denote thymine. The N3 nitrogen of adenine and O2 of thymine are shown. (b) Structural water mimic provided by 3-deaza-3-hydroxymethyladenine.

Although substitution of dA with hydroxymethyl analogue **4** relatively stabilizes the duplex in comparison to **3**, the T_M (duplex **7**, Table **1**) is still significantly lower than that of the native sequence (duplex **1**). This thesis presents the synthesis of another analogue of adenine **5 (Figure 4)** that may become a better mimic of water in the minor groove. The hydroxyl group is positioned on an extended linker and may be more effective in hydrogen bonding in the minor groove of B-form DNA.

Building on the work done by J. Arico[55] and T. Lan,[54] this project becomes a logical extension of the previous research and investigates the synthesis of a novel adenine analogue that may become a good mimic of hydration in the minor groove.



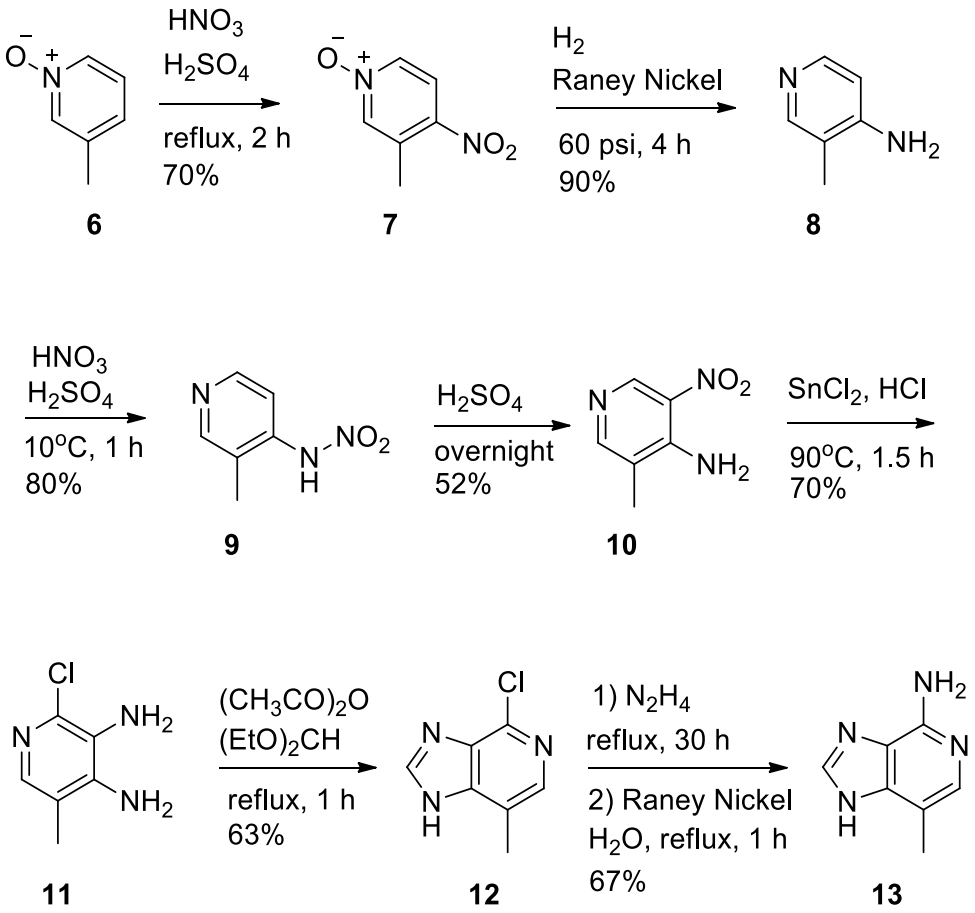
A **c^3A** **m^3c^3A** **hm^3c^3A** **$hEtThm^3c^3A$**

1 **3** **4** **5**

Fugure 5. Adenine and its modifications. Nucleobase **5** is the target analogue that may stabilize a B-form DNA duplex via water mimic in the minor groove. Note, that alternatively these analogues may be named using standard IUPAC nomenclature as the imidazo[4,5-c]pyridine ring system.

The idea behind the design of (2-(((4-amino-1H-imidazo[4,5-c]pyridin-7-yl)methyl)thio)ethanol) **5** is that the molecule may increase the stability of the DNA duplex compared with a duplex containing **4**. The thioethanol linker would provide extra length, postulated to result in a better hydrogen bonding to the neighboring T.

The synthesis of the new nucleobase **5** builds on previous[31,40,55] synthetic procedure. The first part of the synthesis (**Scheme 1**) consists of a series of reactions that results in 3-deazaadenine functionalized with the bromide group at the 3 position (compound **13**). The original synthesis was proposed by Irani,[31] but later was improved by Arico[40,55] in the McLaughlin group.



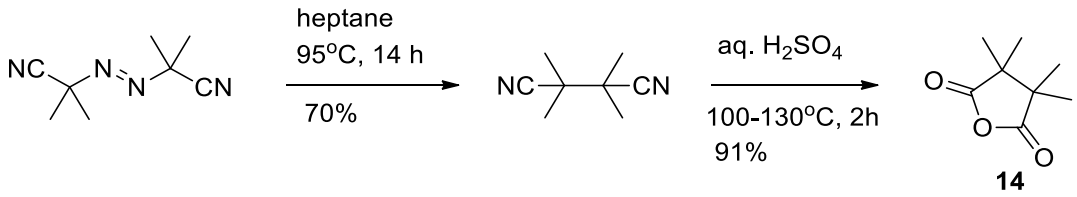
Scheme 1. Synthesis of 3-deaza-3-methyladenine

Refluxing commercially available 3-picoline-N-oxide **6** in a mixture of concentrated nitric and sulfuric acids gives the first nitro compound **7** in 70% yield. Hydrogenation with Raney Nickel at 60 psi using the Parr shaker reduces the nitro group to the amino group within **4** hours. The amino compound **8** is quickly used in the next step (or stored at 4°C), because it oxidizes easily. Another nitration, but at a lower temperature— at 10°C—is followed by the rearrangement of the nitro group in concentrated sulfuric acid. Attempts were made to improve the yields in the latter reaction. The procedure[55] calls for stirring **9** in sulfuric acid at room temperature for 18-48 hours until complete. During

these 18-48 hours the formation of product is monitored by removing aliquots, neutralizing with NH_4OH, filtering and analyzing the precipitate by ^1H NMR. This method of monitoring the reaction is very time consuming and proved to be unnecessary. Contrary to the procedure description,[55] in all of the preparations of the rearranged compound **10** during this project, the expected yield of 50-52% of the target compound was invariably observed after only 18 hours without any increase in the yield over the next 30 hours. The reaction time can be significantly reduced by altering the conditions. First, the starting compound **9** is slowly dissolved in sulfuric acid, then the mixture is gradually heated, and finally allowed to reflux for 30 minutes. The solution is allowed to cool to room temperature, then neutralized with NH_4OH and filtered. Refluxing the solution for 30 min instead of stirring at room temperature overnight provides the product faster, but unfortunately doesn't give any significant increase in the yields.

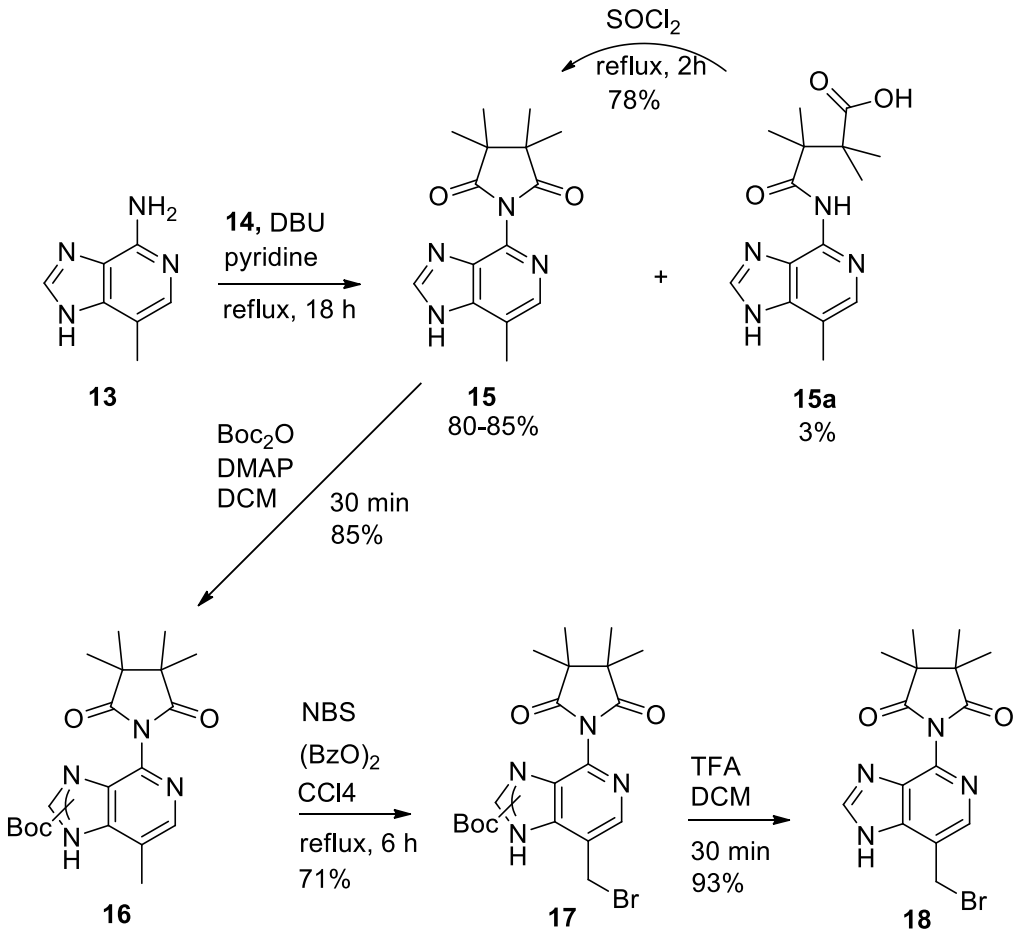
In spite of the low yields obtained for compound **10**, the 50% yields are sufficient for the goals of the project. The reduction of compound **10** with stannous chloride in concentrated hydrochloric acid is followed by refluxing the diamino product **11** in triethylorthoformate and acetic anhydride. The cyclized compound **12** is transformed into 3-methyl-3-deazaadenine **13** in two steps: by first refluxing **12** in anhydrous hydrazine, then in water with Raney Nickel. Compounds **7-12** can be purified in order to obtain the best characterization data, but routinely, for faster results, product **13** becomes the first one in this series to be purified by column chromatography.

The next important step in the synthesis is the protection of the 6-amino group in preparation for glycosylation. The tetramethylsuccinimide (M_4SI) protecting group (**Scheme 2**) was developed in Prof. McLaughlin's group and fulfills the requirements of DNA synthesis protocol (see Chapter 1). It is stable under mildly acidic and basic conditions, but can be removed with ammonium hydroxide at 55 ^oC. Moreover, it has also been shown to be an efficient directing group during the glycosylation step[40]. The bulky methyl groups restrict glycosylation at N7 position of the heterocycle, therefore increasing the yield of the desired N9 isomer.



Scheme 2. The synthesis of 2,2,3,3-tetramethylsuccinic anhydride[41]

M_4SI is installed at N6 position of 3-deaza-3-methyladenine **13** by refluxing it with **14** in pyridine in the presence of DBU (**Scheme 3**). The original procedure[55] calls for an additional step (see Chapter 1, **Scheme 4**): a small amount of the amide, resulting in the reaction with pyridine and DBU, is converted to the imide in refluxing $SOCl_2$. In this research project the thionyl chloride step has been excluded due to its being impractical. On one hand, the amount of the unwanted amide is very small (less than 3%). On the other hand, there are significant losses of the desired imide during column chromatography purification, so the product yield is higher if fewer steps are used.

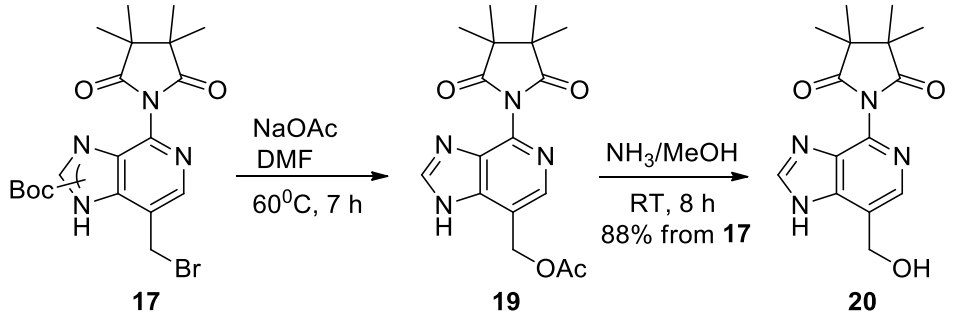


Scheme 3. The synthetic route to the functionalized (bromo) analogue

The resulting 3-methyl-3-deazaadenine **15** with the installed M_4SI is further protected at N7/N9 with a Boc protecting group. This is done in order to increase the solubility of **15** in carbon tetrachloride in the next step. Radical bromination is achieved by the treatment of **15** with N-bromosuccinimide and benzoyl peroxide as a radical initiator resulting in the brominated/ functionalized molecule **17**. Finally, the Boc group is removed by treatment with trifluoroacetic acid in DCM. Nucleobase **18** is a novel

analogue that is suitable for further modification: bromide can be easily displaced by a good nucleophile in an S_N2 reaction.

Previously,[55] the Boc group has been left on during further modification of molecules. Specifically, the treatment of compound **17** with NaOAc in DMF at 60°C displaces the bromide as well as the Boc group (**Scheme 4**). The acetate **19** is further hydrolyzed with NH_3/MeOH at room temperature for 8 hours to afford the hydroxymethyl analogue **20**.



Scheme 4. Boc protecting group removal at elevated temperatures

However, the removal of the Boc group at an elevated temperature may not be ideal for some of the more sensitive reactions. For example, a substitution with a strong nucleophile, a kinetically controlled process, may quickly form additional unwanted isomers or lead to decomposition at higher temperatures. The sensitivity of each specific nucleophile to temperature should be taken into account as well. In this project, the installation of the long thiol linker (*vide infra*) onto molecule **17** did not succeed, most likely, due to the steric hindrance between the bulky nucleophile and Boc group. On the other hand, after the removal of Boc, molecule **18** became a perfect

substrate for the nucleophilic attack of the thio nucleophile **21** (**Schemes 5, 6**). Thus compound **18** becomes a useful starting material for further modification of the purine analogue.

21

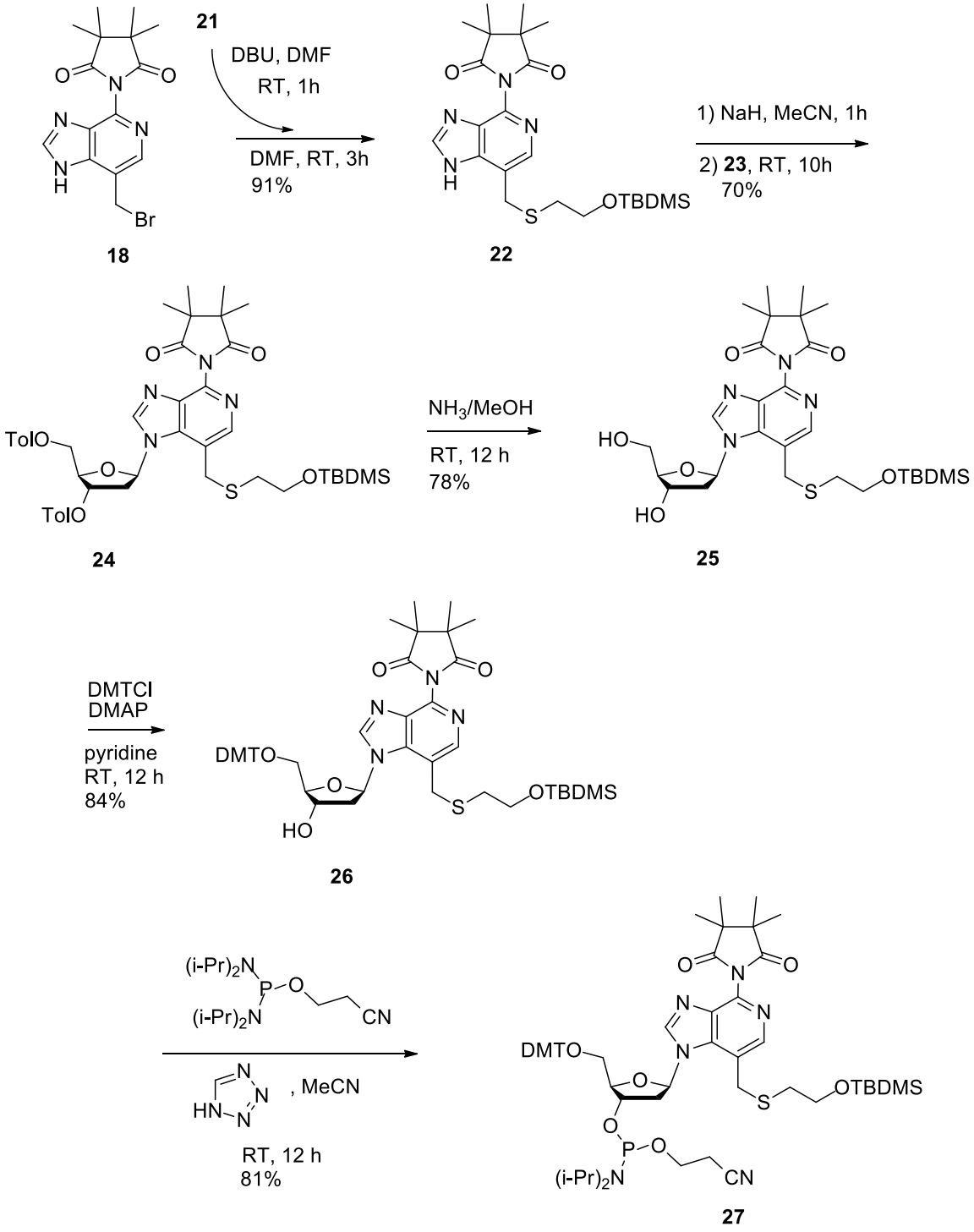
Scheme 5. 2-mercaptoethanol is protected with TBDMS group

For the synthesis of the reactant **21**, 2- mercaptoethanol is protected with tetrabutyldimethylsilyl group[56] (**Scheme 5**). The reaction proceeds quantitatively without any undesired protection on the sulfur because of the great affinity of silicon to oxygen. The presence of thiol in **21** can be checked by Elman's reagent, or DTNB (5,5'-Dithiobis-(2-Nitrobenzoic Acid)). Elman's reagent is commonly used to quantitate thiols in proteins, cells and plasma by absorption measurements. It readily forms a mixed disulfide with thiols, liberating the chromophore 5-mercapto-2-nitrobenzoic acid (absorption maximum 410 nm). In this project, there is no need to quantitate thiols, but simply their presence can be detected by mixing a small amount of DTNB and **21** affording a solution of a bright red color.

The next step installs the hydroxyethanol thio linker: **21** is first treated with DBU in DMF for 1 hour, then stirring the mixture with the starting nucleobase for 3 hours gives modified nucleobase **22** (**Scheme 6**).

Glycosylation is a particularly difficult reaction in the case of modified 2'-deoxy purines (see Chapter 1) because it must be both regio- and diastereoselective. The formation of four different isomers is possible and the separation is difficult. The sodium salt method[14] used in this project gave good yields of the desired β-N9 nucleoside. Nucleobase **22** is treated with NaH in MeCN (**Scheme 6**) followed by addition of chlorosugar[57] **23.** After stirring for about 10 hours, TLC shows complete conversion of the starting material. The desired β-N9 isomer is obtained in 70% yield. After the work up, it is separated from the rest of the products by flash chromatography in 1:1 hexane-ethyl acetate solvent system. The identity of **24** was confirmed by examining 2D COSY and NOESY NMR spectra (see **Table 2**).

The M_4SI group stays on when the tolouyl groups are removed by the treatment with a solution of 7M ammonia in methanol at room temperature to yield 78% of **25**. Protection of the 5'-OH with the acid-labile DMT group must be performed in very dry reaction conditions. Starting material **25** is co-evaporated with pyridine at least twice to insure there is no water present. Water is a better nucleophile compared to the



Scheme 6. The formation of the novel phosphoramidite **27**

hydroxyl groups in **25**, therefore reacts faster with DMT-Cl reagent. A straightforward phosphitylation of **26** afforded cyanoethyl phosphoramidite **27** in 81% yield.

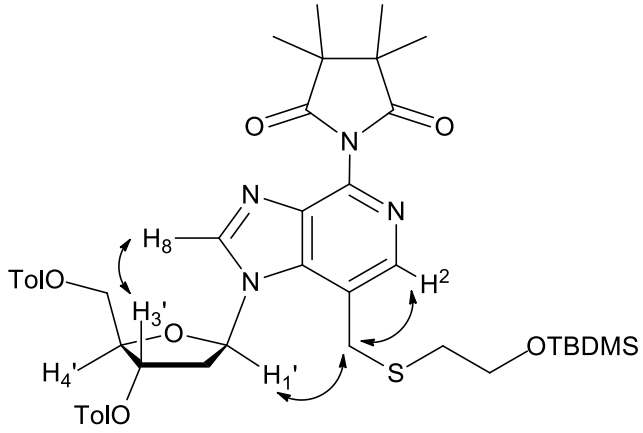


Table 2. 2D NOESY. Selected important NOESY signals are shown with arrows.

	Proton δ, ppm	NOESY
	3-CH_2(4.4-4.1)	$H^{1'}$, H^{2}
	$H^{1'}$(6.8)	3-CH_2, $H^{2'}$
	$H^{2'}$(2.9)	$H^{1'}$
	$H^{2''}$(2.8)	$H^{3'}$, $H^{2'}$, $H^{1'}$
	$H^{3'}$ (5.6)	H^{8}, $H^{5'/5''}$,$H^{2''}$

In order to investigate the effect of the new modified purine on duplex stability, phosphoramidite **27** must be incorporated into an oligonucleotide. A modified oligonucleotide is then annealed with its complementary oligonucleotide and the duplex's stability is measured by thermal denaturation (T_M) studies.

There are several factors that have to be considered when incorporating **5** into a DNA duplex. The new side chain on the modified adenine is bulky, and it's unclear how it will be tolerated in an oligonucleotide strand and in a duplex. If this modification is successfully installed in an oligonucleotide, how will the linker with many degrees of freedom arrange itself in space? Where would the new hydrogen bonds form? Answers to these questions would clarify the importance of hydration in stabilizing A-T tracts, and the ability of the minor groove to accommodate non-natural functional groups.

The synthesis of a modified DNA sequence containing **5** has been attempted (see chapter 3, experimental section for the details). Sequences **1-3** (**Table 3**) were prepared by solid-phase DNA synthesis. The DNA synthesizer showed 100 % coupling efficiency for the analogue **5** with the overall oligonucleotide yield at 70%. Inspired by the seemingly easy incorporation of the new analogue into an oligonucleotide sequence, we

Oligonucleotide	Sequence	MALDI-TOF MS
1	5'-d(CCGGAAAACGCC)-3'	3623.35
2	5'-d(GGCCTTTTGCGG)-3'	3668.51
3	5'-d(CCGG**5**AAACGCC)-3'	3311.70

Table 3. Oligonucleotides' sequences: **1** and **3** are complementary to **2**, but in **3** one of the adenines is replaced by **5**.

continued with the deprotection and purification steps. The treatment with ammonium hydroxide overnight at 55^{o}C deprotected and cleaved the oligonucleotides from the CPG

beads. The purification of the native and modified sequences was accomplished by HPLC Oligo R3 reverse-phase C18 column, trityl on.

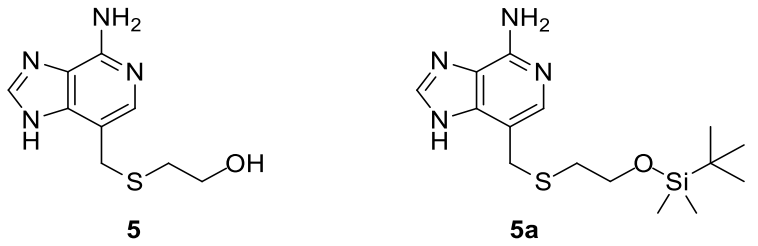


Figure 6. The new adenine analogue hEtThm^{3}c^{3}A and TBDMS-protected hEtThm^{3}c^{3}A

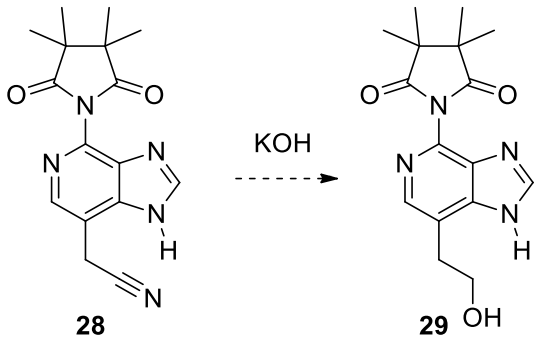
The DMT-protected 12-mers had retention times of about 17.5 minutes, except for sequence **3**, which had a retention time of 24 min. We assumed this was due to the presence of the TBS group (molecule **5a**, **Figure 6**) remaining on the modified residue. The collected oligonucleotide solutions were then reduced in volume and detritylated in 80% acetic acid for 30 min.

The resulting native oligonucleotides **1** and **2** were desalted (Sephadex G-10) and lyophilized. The entire product containing sequence **3** was divided into two parts. One half was desalted and lyophilized just like the native sequences (TBDMS protecting group stayed on). The other half was dissolved in 1 M TBAF in THF solution for 3 hours in order to remove the TBDMS protecting group first, then desalted.

In spite of the promising results shown by the DNA synthesizer and the HPLC (complete incorporation of the novel phosphoramidite and longer retention times due to TBDMS group on), MALDI-TOF spectra showed that the modified nucleobase might not have been incorporated. For the reasons that are still unclear, the mass of the modified sequence did not have the correct mass. The product subjected to the treatment with TBAF was decomposed.

Conclusions and Outlook

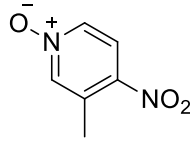
Analogue **5** is thought to be a good candidate for probing the hydration in the minor grove because of its long side chain containing a hydrogen bond donor postulated to result in more favorable hydrogen bonding to the neighboring T, therefore increasing the stability of the DNA duplex. Other molecules may become good mimics of water as well, for example, analogue **29** (**Scheme 8**). The modified purine **28** has been successfully synthesized (see chapter 3, experimental details) and the synthesis of molecule **29** is ongoing.



Scheme 8. A possible synthesis of a new purine analogue

In summary, a novel purine analogue has been synthesized and tested in an oligonucleotide strand with the goal of probing the hydration in the minor groove. Although the multi-step synthesis of the new adenine analogue **5** turned out to be a time-consuming process, the incorporation of this molecule in a DNA duplex may be useful for a better understanding of the hydration of B-form DNA. The attempted incorporation of the new analogue in a DNA strand yielded inconclusive results, and further tests are ongoing.

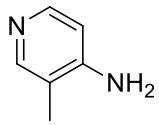
3. Experimental Details and Supporting Information



7

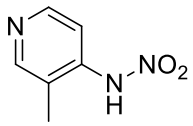
4-Nitro-3-picoline N-oxide (7).[55] To a flask containing 3-picoline N-oxide **6** (20 g, 183 mmol) was slowly added 100 mL of concentrated H_2SO_4. While keeping the

temperature below 80 °C, 40 mL of 90% fuming HNO_3 was added. The reaction mixture was then carefully heated to 100°C until an exothermic reaction was observed with evolution of a brown gas. The flask was removed from heat until the bubbling had subsided, then heated at 120 °C for 2 h, after which it was cooled to room temperature and poured over 300 g of crushed ice. This mixture was placed in an ice bath and titrated with concentrated NH_4OH until a yellow precipitate formed at basic pH. The mixture was filtered and the solid collected and dried to yield 19.74 g (70%) of **7**. This compound could be used directly for the next step; recrystallization from methanol yielded analytically pure material. ^{1}H NMR (400 MHz, acetone-d6): δ 8.29 (s, 1H), 8.19 (m, 1H), 8.11(d, J=12 Hz, 1H), 2.59 (s, 3H); MS (ESI): m/z 154 (M^+).



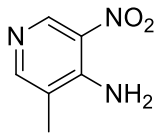
8

4-Amino-3-methylpyridine (8).[55] Raney nickel 2800 (8.0 g; **CAUTION**: pyrophoric; weighed under tared volume of water) was added to 100 mL of methanol and pre-reduced in a Parr shaker with H_2 at 60 psi for 5 min. Compound **10** (8.0 g, 52 mmol) was added and reduced at 65 psi for 4 h. The mixture was then filtered through Celite, washed thoroughly with methanol and dried *in vacuo* to obtain 5.05 g (90%) of a slightly green solid. Compound **8** is oxidized easily; therefore it was purged with argon and stored at 4 °C until used. ^{1}H NMR (400 MHz, acetone-*d*6): δ 7.96 (s, 1H), 7.94 (d, *J* = 4.8 Hz, 1H), 6.55 (d, *J*=4.8 Hz), 5.24 (bs, 2H), 2.08 (s, 3H); MS (ESI): *m/z* 108 (M^+).



9

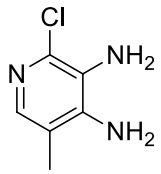
4-Nitramino-3-methylpyridine (9).[55] Compound **8** (6.80 g, 63 mmol) was dissolved in 51 mL of concentrated H_2SO_4 and cooled to 5 °C in a dry ice/ethanol bath. 70% HNO_3 (21 mL) was added slowly to maintain the reaction temperature below 10 °C. The reaction mixture was allowed to stir at 10 °C for 1 h and then poured onto 200 g of crushed ice. Concentrated NH_4OH was then added until a creamy off-white precipitate was formed at pH 7. The mixture was filtered and the precipitate recrystallized from hot water to yield 7.72 g (80%) of **9**. ^{1}H NMR (400 MHz, DMSO-*d*6): δ 8.15 (d, 2H), 8.00 (d, *J* = 6.6 Hz, 1H), 2.08 (s, 1H); MS (ESI): *m*/*z* = 153 (M^+).



10

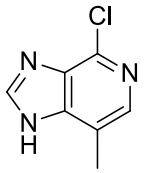
3-methyl-5-nitropyridin-4-amine (10).[55] Compound **9** (2.00 g, 13 mmol) was added to 19 mL of concentrated H_2SO_4 in small portions to avoid decomposition of the starting material. The mixture was allowed to stir at room temperature for 18-48 h until complete (reaction was monitored by removing aliquot, neutralizing with NH_4OH, filtering, and analyzing the precipitate by ^{1}H NMR) and then quenched by pouring over 200 g of crushed ice. This mixture was then slowly titrated with concentrated NH_4OH

while cooling in an ethanol/dry ice bath. When the mixture reached pH 6.5-7 a yellow precipitate formed. This was filtered and dried to yield 1.02 g (52%) of **10**. ^{1}H NMR (400 MHz, DMSO-*d*6): δ 8.87 (s, 1H), 8.09 (s, 1H), 7.66 (bs, 2H), 2.12 (s,3H). MS (ESI): *m*/*z* 153 (M^{+}).



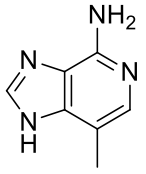
11

2-chloro-5-methylpyridine-3,4-diamine (11).[55] Anhydrous stannous chloride (12.64 g, 66.7 mmol,) was added to 51 mL of concentrated HCl followed by **10** (3.4 g, 22 mmol) was added and the reaction mixture heated at 90 °C for 1.5 h. TLC (9:1 DCM:MeOH) of an aliquot (pH adjusted to 12 and extracted with EtOAc) showed complete consumption of starting material. The reaction mixture was allowed to cool to room temperature and poured over 475 g of crushed ice. The pH was adjusted to 12 with 7 M NaOH and the reaction mixture extracted with EtOAc (4 x 200 mL). The combined organic layers were dried with $MgSO_4$, filtered, and dried *in vacuo* to afford 2.46 g (70%) of white solid. In most cases the crude product was pure enough to be used directly in the next step; otherwise, flash column chromatography eluting with 9:1 DCM:MeOH gave pure **11**. ^{1}H NMR (400 MHz, DMSO-*d*6): δ 7.20 (s, 1H), 5.47 (bs, 2H), 4.63 (bs, 2H), 1.95 (s, 3H). MS (ESI): *m*/*z* 157 (M^{+}).



12

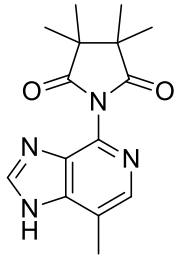
4-chloro-7-methyl-1*H*-imidazo[4,5-*c*]pyridine (12)[55]**.** Under argon atmosphere, 20 mL of anhydrous triethylorthoformate (distilled from Na) and 20 mL acetic anhydride (fractionally distilled) was added to 4.0 g (25.4 mmol) of **11** and refluxed for 1 h. The reaction was cooled to room temperature and dried *in vacuo.* Flash chromatography eluting with 9:1 DCM:MeOH yielded 2.52 g (63%) of **12**. ^{1}H NMR (400 MHz, DMSO-d_6): δ = 13.3 (bs, 1H) 8.45 (s, 1H), 7.91 (s, 1H), 2.44 (s, 3H). MS (ESI)**:** *m*/*z* 167 (M^+).



13

7-methyl-1*H*-imidazo[4,5-*c*]pyridin-4-amine (13). To a flask containing 4-chloro-7-methyl-1*H*-imidazo[4,5-*c*]pyridine **12** (1.50 g, 7.49 mmol) was added 25 mL of anhydrous hydrazine. The mixture was heated to reflux for 30 h. Excess hydrazine was removed *in vacuo* and co-evaporated with EtOH (3 x 15 mL). Water (45 mL) and Raney nickel (2.25 g, CAUTION: pyrophoric, weighed under tared volume of water) were added and the mixture heated to reflux with vigorous stirring to prevent the Raney Ni from sticking to the stir bar. After 3 h the mixture was filtered, while still hot, through a pad of Celite and the pad washed with 200 mL of hot water (CAUTION: prevent leftover

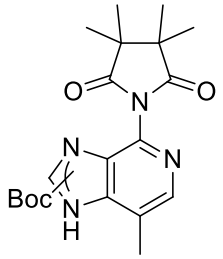
Raney Ni from drying). Evaporation of volatiles *in vacuo* afforded a brown solid. This was re-dissolved in minimal MeOH and 9 g of silica gel were added. After thorough drying of the mixture the resulting fine powder was loaded onto a flash column and purified by elution with 9:1 DCM:MeOH containing 0.7 M NH_3. Evaporation of the appropriate fractions afforded **13** as a white amorphous powder (743 mg, 67%) and starting material (400 mg, 13%). ^{1}H NMR (DMSO-d_6, 400 MHz): δ 12.55 (brs, 1H), 8.08 (s, 1H), 7.42 (s, 1H), 5.86 (brs, 2H), 2.26 (s, 3H) ppm; ^{13}C NMR (DMSO-*d*6, 101 MHz) δ 150.79, 140.69, 139.39, 126.57, 108.00, 14.55 ppm; HRMS (DART-TOF) calcd for $C_7H_9N_4H^+$ $[M + H]^+$ 149.0827, found: 149.0834.



15

To a flask containing **13** (0.69 g, 4.66 mmol) and anhydride **14** (1.5 g, 9.6 mmol) was added pyridine (50 mL) and DBU (2.1 mL). The mixture was heated to reflux for 18 h. Volatiles were removed in vacuo and the residue co-evaporated with toluene (3 x 10 mL) to remove residual pyridine. The resulting brown oil was purified by flash column chromatography eluting with 95:5 DCM:MeOH to obtain 1.1 g (85%) of white solid. ^{1}H NMR (acetone-d6, 400 MHz) δ 8.27 (s, 1H), 8.12 (s, 1H), 2.57 (s, 3H), 1.33 (s,

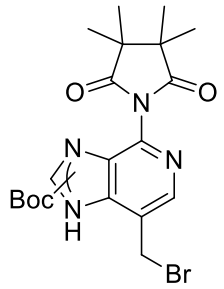
12H); ^{13}C NMR (acetone-d6, 101 MHz) δ 182.63, 145.10, 142.18, 49.06, 22.47, 14.94; HRMS (DART-TOF) calcd for $C_{15}H_{19}N_4O_2H^+$ $[M + H]^+$ 287.1508, found: 287.1508.



16

tert-butyl 7-methyl-4-(3,3,4,4-tetramethyl-2,5-dioxopyrrolidin-1-yl)-1H-imidazo[4,5-c]pyridine-1-carboxylate and tert-butyl 7-methyl-4-(3,3,4,4-tetramethyl-2,5-dioxopyrrolidin-1-yl)-3H-imidazo[4,5-c]pyridine-3-carboxylate (16). To a flask containing **15** (2.85 g, 9.95 mmol) was added Boc2O (3.04 g, 13.9 mmol) and DMAP (243 mg, 1.99 mmol). DCM (52 mL) was added andt he mixture stirred at rt for 30 min (anhydrous conditions not needed). The starting material gradually dissolved to produce a yellow homogeneous solution. TLC (97:3 DCM: MeOH) indicated that a small amount of starting material remained. More Boc2O (1.1 g, 5 mmol) and DMAP (122 mg, 1 mmol) were added to drive the reaction to completion. After allowing the mixture to stir for another 15 min, TLC indicated that no starting material remained. Volatiles were removed in vacuo and the residue purified by flash column chromatography eluting with 97.5:2.5 DCM:MeOH to afford the desired product as a white foam (3.27 g, 85%, mixture of N7/N9 regioisomers). 5: Rf = 0.29, 0.36 (silica gel, 97:3 DCM:MeOH); 1H NMR (CDCl3, 400 MHz): δ 8.41 (s, 0.5H), 8.39 (d, J = 0.8 Hz, 0.5H), 8.38 (s, 1H), 8.32 (d, J = 0.8

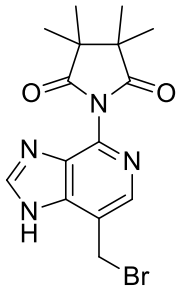
Hz, 1H), 2.73 (d, J = 0.8 Hz, 3H), 2.63 (d, J = 0.8 Hz, 1.5H), 1.64 (s, 9H), 1.60 (s, 4.4H), 1.39 – 1.29 (m, 17.8H) ppm; 13C NMR (CDCl3, 101 MHz) δ 181.62, 181.46, 151.92, 146.67, 146.31, 145.86, 145.27, 144.91, 143.76, 138.26, 137.65, 137.47, 130.87, 126.75, 121.55, 86.96, 86.66, 81.05, 47.98, 46.73, 28.05, 27.98, 27.96, 27.81, 22.63, 21.66, 21.51, 18.87, 13.67 ppm; HRMS (ESI-TOF) calcd for C20H27N4O4H+ [M + H]+ 387.20323, found: 387.20368.



17

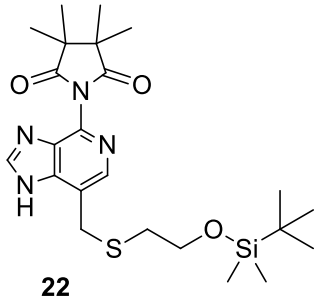
***tert*-butyl 7-(bromomethyl)-4-(3,3,4,4-tetramethyl-2,5-dioxopyrrolidin-1-yl)-1*H*imidazo[4,5-*c*]pyridine-1-carboxylate** and ***tert*-butyl 7-(bromomethyl)-4-(3,3,4,4-tetramethyl-2,5-dioxopyrrolidin-1-yl)-3*H*-imidazo[4,5-*c*]pyridine-3-carboxylate (17).** To a flask containing **16** (3.27 g, 8.46 mmol) was added *N*-bromosuccinimide (1.28 g, 7.2 mmol, recrystallized from hot H_2O), benzoyl peroxide (131 mg, 0.54 mmol), and 175 mL of CCl_4 (reagent grade, not anhydrous). The mixture was heated to reflux with stirring. After 3 h additional NBS (301 mg, 1.69 mmol) and $(BzO)_2$ (31 mg, 0.127 mmol) were added. The reaction was carefully monitored by TLC (97:3 DCM:MeOH) afterwards.

After 6.5 h total the reaction was virtually complete, with only a trace of starting material and a trace of undesired dibrominated product. The reaction was allowed to cool to rt and filtered to removed succinimide. Volatiles were removed *in vacuo* and the residue purified by flash column chromatography eluting with DCM→97:3 DCM:MeOH to afford the desired product as a white foam (2.78 g, 71%, mixture of N7/N9 regioisomers), as well as the dibrominated product as an off-white solid (357 mg, 7.7%, mixture of N7/N9 regioisomers). **6**: R_f = 0.36, 0.43 (silica gel, 97:3 DCM:MeOH); $_1$H NMR ($CDCl_3$, 400 MHz) δ 8.61 (s, 1H), 8.50 (s, 1H), 8.49 (s, 0.5H), 8.41 (s, 0.5H), 5.13 (s, 1H), 4.90 (s, 2H), 1.70 (s, 4.5H), 1.62 (s, 9H), 1.42 – 1.30 (m, 18H) ppm; $_{13}$C NMR (101 MHz, $cdcl_3$ δ 181.27, 181.08, 151.19, 146.01, 145.47, 144.84, 143.69, 139.70, 132.91, 126.07, 125.29, 121.49, 87.70, 87.03, 48.02, 46.78, 29.46, 27.92, 27.88, 27.85, 23.84, 22.45, 21.51, 21.37, 18.80 ppm; HRMS (DART-TOF) calcd for $C_{20}H_{25}Br_1N_4O_4H_+$ [M - Boc + $H]_+$ 467.11170, found: 467.11272.



18 **1-(7-(bromomethyl)-1H-imidazo[4,5-c]pyridin-4-yl)-3,3,4,4-tetramethylpyrrolidine-2,5-dione (12).** Compound **17** (1.11 g, 2.38 mmol) was dissolved in 20 ml of dichloromethane. While stirring, trifluoroacetic acid (4 ml) was added drop

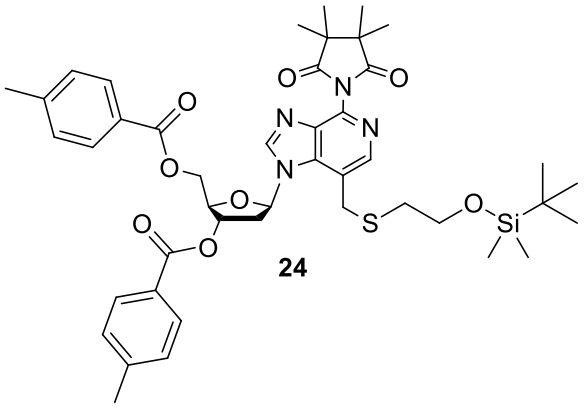
wise. The mixture was allowed to stir at room temperature for 30 min, at which time TLC (5:95 MeOH:DCM) indicated that all starting material had been consumed. The reaction was quenched by the addition of saturated solution of sodium bicarbonate. The resulting mixture was extracted with dichloromethane, washed with water and saturated solution of sodium chloride. The organic layer was dried with anhydrous sodium sulfate, filtered, and dried *in vacuo*. The crude product was purified by column chromatography (3:97 MeOH:DCM) to afford the title compound as a white foam (810 mg, 93%). ^{1}H NMR (400 MHz, Chloroform-d) δ 8.28 (s, 1H), 7.74 (s, 1H), 4.59 (s, 2H), 1.34 (s, 12H). ^{13}C NMR (101 MHz, Chloroform-d) δ 182.16 , 143.60 , 141.05 , 76.91 , 48.25 , 25.78 , 21.73 . HRMS (ESI-TOF) calcd for $C_{15}H_{18}BrN_4O_2$ 365.06131, found: 365.06199.



22

1-(7-(((2-((tert-butyldimethylsilyl)oxy)ethyl)thio)methyl)-1H-imidazo[4,5-c]pyridin-4-yl)-3,3,4,4-tetramethylpyrrolidine-2,5-dione (22). To a dry flask containing 2-((tert-butyldimethylsilyl)oxy)ethanethiol **21** (95mg, 0.41 mmol) and 2 ml of DMF under inert atmosphere was added DBU (85 µL, 0.57 mmol). The mixture was stirred at room temperature for 30 min. Then compound **18** dissolved in 2 ml of DMF was added to the mixture drop wise and stirred for 3 hrs. The reaction was quenched

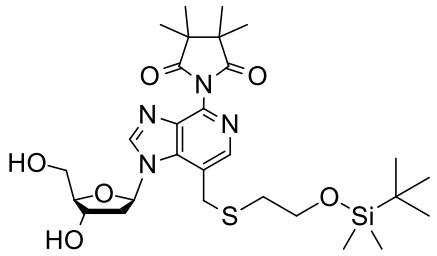
with saturated solution of sodium bicarbonate, extracted with DCM, washed with water and saturated solution of sodium chloride. The organic layer was dried with anhydrous sodium sulfate and filtered. Volatiles were removed *in vacuo* and the residue purified by column chromatography (97:3 DCM:MeOH) to afford 178 mg (91%) of **22** as a white solid. ^{1}H NMR (400 MHz, Chloroform-d) δ 8.14 (s, 1H), 7.72 (s, 1H), 3.91 (s, 2H), 3.77 (t, J = 6.2 Hz, 2H), 2.48 (t, J = 6.2 Hz, 2H), 1.31 (s, 12H), 0.86 (s, 9H), 0.07 (s, 6H). ^{13}C NMR (101 MHz, Chloroform-d) δ 182.11 , 143.32 , 140.70 , 120.81 , 77.55 , 63.71 , 48.06 , 33.96 , 30.46 , 26.08 , 21.65 , 18.53 , -5.09. HRMS (ESI-TOF) calcd for $C_{23}H_{37}N_4O_3SSi$ $[M+H]^+$ 477.23556, found: 477.23668.



24

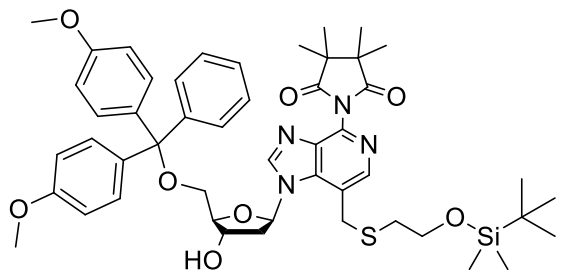
(2R,3S,5R)-5-(7-(((2-((tert-butyldimethylsilyl)oxy)ethyl)thio)methyl)-4-(3,3,4,4-tetramethyl-2,5-dioxopyrrolidin-1-yl)-1H-imidazo[4,5-c]pyridin-1-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl 4-methylbenzoate (24). Compound **22** (200 mg, 0.42 mmol) was combined with NaH (42 mg dispersion in mineral oil, 0.88 mmol) under N_2 atmosphere. Dry acetonitrile (3 ml) was added and the mixture was stirred at room temperature for 1 hr. Then α-

chloro sugar **23** (342 mg, 0.88 mmol) dissolved in 2 ml of MeCN was added dropwise to the flask. The mixture was stirred overnight, after which the reaction was quenched with ammonium chloride. The resulting mixture was extracted with ethyl acetate, washed with water and saturated solution of sodium chloride. The organic layer was dried with anhydrous sodium sulfate, filtered, and dried *in vacuo*. The crude product was purified by column chromatography (1:1 Hexane:EtOAc) to afford **14** as a yellow foam (238 mg, 70%). ^{1}H NMR (500 MHz, Chloroform-d) δ 8.22 (d, J = 4.9 Hz, 2H), 7.99 – 7.92 (m, 2H), 7.87 – 7.81 (m, 2H), 7.24 (dd, J = 23.7, 8.0 Hz, 4H), 6.86 (dd, J = 8.3, 5.3 Hz, 1H), 5.70 (dt, J = 5.7, 2.4 Hz, 1H), 4.66 – 4.51 (m, 3H), 4.30 (d, J = 13.7 Hz, 1H), 4.11 – 4.01 (m, 1H), 3.80 (qt, J = 10.5, 6.2 Hz, 2H), 2.92 (ddd, J = 13.7, 5.3, 2.5 Hz, 1H), 2.83 (ddd, J = 14.0, 8.3, 6.0 Hz, 1H), 2.69 – 2.52 (m, 2H), 2.40 (d, J = 14.8 Hz, 7H), 2.15 (s, 1H), 1.38 – 1.18 (m, 18H), 0.87 (s, 12H), 0.07 (d, J = 1.8 Hz, 6H). ^{13}C NMR (101 MHz, Chloroform-d) δ 181.49 (d, J = 16.1 Hz), 166.32 , 166.04 , 145.06 , 144.81 , 144.41 , 143.56 , 142.27 , 139.36 , 139.07 , 137.96 , 130.38 – 129.96 (m), 129.81 (d, J = 3.6 Hz), 129.67 – 129.21 (m), 126.75 , 126.50 (d, J = 3.8 Hz), 119.54 , 85.98 , 84.32 (d, J = 11.3 Hz), 82.89 , 82.63 , 77.55 , 74.78 , 74.51 (d, J = 12.8 Hz), 64.15 , 63.95 , 63.56 , 56.18 , 53.53 , 53.16 , 48.18 (d, J = 5.8 Hz), 38.86 , 36.83 , 33.92 , 32.15 (d, J = 3.4 Hz), 31.23 , 29.89 (d, J = 4.4 Hz), 29.57 , 26.13 , 23.03 – 22.77 (m), 22.05 – 21.47 (m), 18.52 (d, J = 3.0 Hz), 14.37 (d, J = 7.8 Hz), -4.98 – -5.42 (m). HRMS (ESI-TOF) calcd for $C_{44}H_{56}N_4O_8NaSiS$ $[M+Na]^+$ 851.3486, found: 851.3498.



25

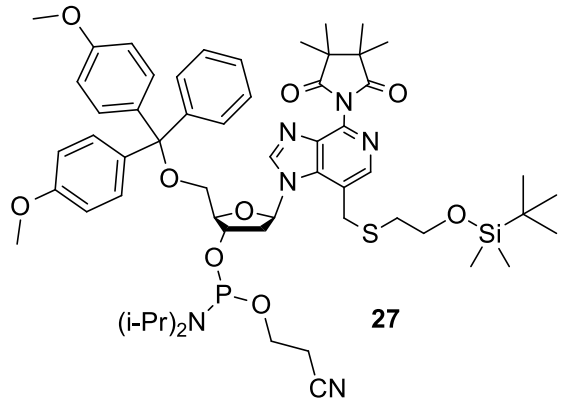
1-(7-(((2-((tert-butyldimethylsilyl)oxy)ethyl)thio)methyl)-1-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-1H-imidazo[4,5-c]pyridin-4-yl)-3,3,4,4-tetramethylpyrrolidine-2,5-dione (25). To a dry flask containing **24** (268 mg, 0.32 mmol) was added 4 mL of MeOH with 7 M NH_3. The flask was tightly stoppered and the mixture allowed to stir at room temperature overnight. TLC indicated complete conversion to **25**. Volatiles were removed *in vacuo* and the residue purified by column chromatography eluting with (97:3 DCM:MeOH) to afford **25** as a white foam (147 mg, 78%). ^{1}H NMR (400 MHz, Methanol-d4) δ 8.68 (s, 1H), 8.17 (s, 1H), 6.76 (t, J = 6.3 Hz, 1H), 4.52 (dt, J = 6.0, 4.0 Hz, 1H), 4.34 (d, J = 13.6 Hz, 1H), 4.11 (d, J = 13.6 Hz, 1H), 4.00 (q, J = 3.6 Hz, 1H), 3.80 – 3.55 (m, 4H), 2.79 – 2.51 (m, 4H), 1.34 – 1.29 (m, 13H), 0.85 (s, 9H), 0.07 (d, J = 1.8 Hz, 6H). ^{13}C NMR (101 MHz, Methanol-d4) δ 183.35 , 146.01 , 143.55 , 140.61 , 139.18 (d, J = 11.9 Hz), 122.32 , 89.58, 87.67 , 72.19 , 64.80 , 64.58 , 62.84 , 49.79 , 42.01 , 35.10 , 31.57 , 30.87 , 26.56 , 22.01 (d, J = 2.3 Hz), 19.30 , -4.98 (d, J = 2.1 Hz). HRMS (DART-TOF) calcd for $C_{28}H_{45}N_4O_6SSi$ $[M+H]^+$ 593.2829, found: 593. 2842.



26

1-(1-((2R,4S,5R)-5-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-4-hydroxytetrahydrofuran-2-yl)-7-(((2-((tert-butyldimethylsilyl)oxy)ethyl)thio)methyl)-1H-imidazo[4,5-c]pyridin-4-yl)-3,3,4,4-tetramethylpyrrolidine-2,5-dione (26). Compound **25** (200 mg, 0.34 mmol) was co-evaporated with pyridine twice to ensure dry reaction conditions. Then 5 ml of pyridine was added to the flask, followed by DMAP (10 mg, 0.08 mmol) and DMTCl (138 mg, 0.41 mmol). The mixture was allowed to stir overnight, after which the reaction was quenched with water. The resulting mixture was extracted with DCM, washed with water and saturated solution of sodium chloride. The organic layer was dried with anhydrous sodium sulfate, filtered, and dried *in vacuo*. The crude product was purified by column chromatography (silica gel washed with 1% TEA in 98:2 DCM:MeOH, then eluted with two column volumes of solvent system without TEA before loading compound) eluting with 98:2 DCM:MeOH to afford **26** (252 mg, 84%) as a white foam. ^{1}H NMR (400 MHz, Chloroform-d) δ 8.21 (s, 1H), 8.13 (s, 1H), 7.39 – 7.32 (m, 2H), 7.30 – 7.14 (m, 7H), 6.83 – 6.75 (m, 4H), 6.70 (t, J = 6.2 Hz, 1H), 4.52 (q, J = 5.0 Hz, 1H), 4.27 (d, J = 13.3 Hz, 1H), 4.14 – 3.99 (m, 2H), 3.91 – 3.74 (m, 8H), 3.41 – 3.31 (m, 1H), 3.25 (dd, J = 10.1, 5.3 Hz, 1H), 2.75 – 2.54 (m, 4H), 1.31 (s, 12H), 0.90 (d, J = 1.2 Hz,

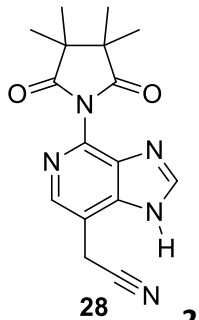
9H), 0.07 (d, J = 1.8 Hz, 6H). ^{13}C NMR (101 MHz, Chloroform-d) δ 221.72 , 181.63 , 178.71 , 158.82 , 145.36 , 144.56 , 143.28 , 142.58 , 139.23 , 138.86 , 138.04 , 135.66, 130.18, 128.20, 127.23 , 119.53 , 113.50, 87.96 , 86.94 , 85.90 , 85.58 , 76.91 , 72.57 , 63.96 , 63.58 , 55.44 , 48.14, 41.07 , 34.19 , 31.23 , 26.18, 21.88 , 21.53 , 18.59 , -5.02 . HRMS (DART-TOF) calcd for $C_{49}H_{63}N_4O_8SSi$ $[M+H]^+$ 895.4131, found: 895.4137.



27

(2R,3S,5R)-2-((bis(4-methoxyphenyl)(phenyl)methoxy)methyl)-5-(7-(((2-((tert-butyldimethylsilyl)oxy)ethyl)thio)methyl)-4-(3,3,4,4-tetramethyl-2,5-dioxopyrrolidin-1-yl)-1H-imidazo[4,5-c]pyridin-1-yl)tetrahydrofuran-3-yl (2-cyanoethyl) diisopropylphosphoramidite (27). To a flask containing **26** (30 mg, 0.034 mmol) dissolved in MeCN (0.5 mL) was added tetrazole (67µL of 0.45 M solution in MeCN, 0.031 mmol) followed by 2-cyanoethyl N,N,N,N-tetraisopropylphosphane (21 µL, 0.068

mmol). The mixture was allowed to stir at room temperature overnight, after which TLC indicated that the reaction was complete. MeOH was added and stirring continued for 5 min. The mixture was extracted with DCM, washed with saturated $NaHCO_3$ and brine. The organic layer was dried with anhydrous Na_2SO_4 and filtered. Volatiles were removed *in vacuo* and the residue purified by column chromatography (silica gel washed with 1% TEA in 99.5:0.5 DCM:MeOH, then eluted with two column volumes of the solvent system without TA before loading the compound) eluting with 95.5:0.5 DCM:MeOH to afford **27** as a white foam (28 mg, 81%, 1:1 mixture of diastereomers). ^{31}P NMR (162 MHz, Chloroform-d) δ 149.269, 149.710 ppm. HRMS (DART-TOF) calcd for $C_{58}H_{80}N_6O_9PSSiK^+$ $[M+K]^+$ 1133.4773, found : 1133.4183.



28 **2-(4-(3,3,4,4-tetramethyl-2,5-dioxopyrrolidin-1-yl)-1H-imidazo[4,5-c]pyridin-7-yl)acetonitrile (28)**

To the flask containing **18** (200 mg, 0.55 mmol), KCN (54 mg, 0.82 mmol), and 18-crown-6 ether (217 mg, 0.82 mmol) was added 5.5 ml DMF. The mixture was allowed to stir at room temperature for 2 hours. Solvent was evaporated and the crude product was

purified by column chromatography (MeOH:DCM 2:98). ^{1}H NMR (500 MHz, DMSO-d6) δ 13.52 (s, 1H), 8.51 (s, 1H), 8.24 (s, 1H), 4.35 (s, 2H), 1.34 (s, 12H). MS calculated for $C_{16}H_{18}N_5O_2$ 312.1455, found: 312.1516.

DNA Synthesis

Sequences **1-3** (**Table 1**) were prepared by solid-phase DNA synthesis. The conditions for the incorporation of analogue nucleotide containing modified base **5** (**Figure 1**) were changed: the wait time during coupling cycle was increased to 2 hours from standard 2 minutes used for the native sequences. The DNA synthesizer showed 100 % coupling efficiency for the

Oligonucleotide	Sequence	MALDI-TOF MS
1	5'-d(CCGGAAAACGCC)-3'	3623.35
2	5'-d(GGCCTTTTGCGG)-3'	3668.51
3	5'-d(CCGG**5**AAACGCC)-3'	3311.70

Table 1. Oligonucleotides' sequences: **1** and **3** are complementary to **2**, but in **3** one of the adenines is replaced by **5**.

analogue **5** with the overall oligonucleotide yield at 70%. Inspired by the seemingly easy incorporation of the new analogue into an oligonucleotide sequence, we continued with the following steps. After the DNA synthesizer reactions were complete, the beads with were transferred into separate vials. The vials were filled with ammonium hydroxide and placed in 55^{o}C water bath overnight in order to deprotect the oligos and cleave them from the CPG beads. Then the cooled supernatant solutions was dried *in vacuo*, then re-dissolved in 4-6 mL of water for purification.

The purification of the native and modified sequences was accomplished by HPLC

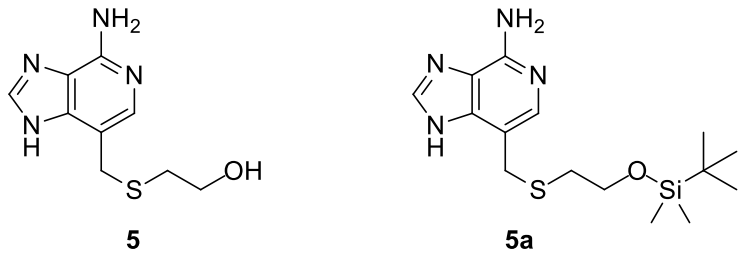


Figure 1. The new adenine analogue hEtThm^{3}c^{3}A and TBDMS-protected hEtThm^{3}c^{3}A

(Oligo R3 reverse-phase C18 column, trityl on), starting with 100% A, using a linear gradient from 0 to 50% B over 30 min (A: 1 M TEAA (pH 7) in 5% acetonitrile; B: 1 M TEAA (pH 7) in 70% acetonitrile). The DMT-protected 12-mers had retention times of about 17.5 minutes, except for sequence **3**, which had a retention time of 24 min. We assumed this was due to the presence of a TBS group remaining on the modified residue. The collected oligonucleotides were then reduced in volume and detritylated in 80% acetic acid for 30 min.

The resulting native oligonucleotides **1** and **2** were desalted (Sephadex G-10) and lyophilized. The entire product containing sequence **3** was divided into two parts. One half was desalted and lyophilized just like the native sequences (TBDMS protecting group stayed on). The other half was dissolved in 1 M TBAF in THF solution for 3 hours in order to remove the TBDMS protecting group first, then desalted.

In spite of the promising results shown by the DNA synthesizer and the HPLC, MALDI spectra showed that the modified nucleobase might not have been incorporated. For the reasons that are still unclear, the mass of the modified sequence did not match the expected number. Expected mass: 3713 (TBDMS-deprotected), 3827 (TBDMS group on), found: 3311.7.

References

(1) Crick, F.H.C. *Symp. Soc. Exp. Biol.* **1958***, XII*, 139-163.

(2) Westheimer, F.H., *Science, New Series*, **1987**, *235, 4793*, 1173-1178.

(3) Miescher, F. in *Die hisochemischen und Physioloischem Arbeiten*, **1897**.

(4) Watson, J. D.; Crick, F. H. *Nature* **1953**, *171 (4356),* 737–738.

(5) Franklin, R. E.; Gosling, R. G. *Nature*, **1953**, *172 (4369),* 156–157.

(6) Gill et al, *Nature*, **1985**, *318*, 577.

(7) Illumina.com

(8) (a) Abu-Salah, K. M., Ansari, A. A., Alrokayan, S. A.. *Journal of Biomedicine and Biotechnology*, **2010** (2010), Article ID 715295.

(b) Ogasawara, S., K. Fujimoto. CheBioChem, **2005**, *6(10),* 1756-1760.

(9) U.S. Department of Energy Office of Science, ORNL.gov.

(10) Ecker, J.R., Davis, R.W. *Proc Natl Acad Sci USA.* **1986**, *83 (15):* 5372–5376.

(11) Napoli, C., Lemieux, C., Jorgensen, R. *Plant Cell*, **1990**, *2 (4):* 279–289.

(12) Stram, Y., Kuznetzova, L. *Virus Genes*, **2006**, *32 (3):* 299–306.

(13) (a) Arnott, S. *Trends Biochem. Sci*., **2006**, *31*, 349-354.

(b) Fuller, W. *Nature*, **2003**, *424*, 876-878.

(c) Klug, A. *J. Mol. Biol*., **2004**, *335*, 3-26.

(d) Maddox, B., *Rosalind Franklin*. **2002**, Harper Collins, London.

(e) Watson, J. D. *The Double Helix*. **1968**, Weidenfeld and Nicolson, London.

(14) (a) Saenger, W. *Principles of Nucleic Acid Structure*. **1984**, Springer-Verlag, Berlin.

(b) Blackburn, G. M., Gait, M. J., Loakes, D., and Williams, D. M. *Nucleic Acids in Chemistry and Biology*. **2006**, 3rd Edition, The Royal Society of Chemistry, Cambridge, UK.

(c) Bloomfield, V. A., Crothers, D. M., and Tinoco, I., Jr. *Nucleic Acids. Structures, Properties and Functions*. **2000**, University Science Books, Sausalito, California.

(15) (a) McPherson, A. *Introduction to Macromolecular Crystallography*. **2003**, Wiley-Liss, New Jersey.

(b) Neidle, S. *Principles of Nucleic Acid Structure*. **2008**, Elsevier.

(16) (a) Lane, A. N. *Methods in Enzymology*, **1995**, *261*, 413-435.

(b) Wemmer, D. E. *Curr. Opin. Struct. Biol*., **1991**, *1*, 452-458.

(17) Wilson, W. D. Science, 2002, 295, 2103-2104.

(18) Drew, H. R., Dickerson, R. E. *J. Mol. Biol.* **1981**, *151*, 535-556.

(19) Hartmann, B., Lavery, R. *Quat. Rev. Biophys*., **1996**, *29*, 309-368.

(20) (a) Morales, J. C. and Kool, E. T. *JACS* **2000**, *122*, *6*, 1001-1007.

(b) Guckian, K. M., Morales, J. C. & Kool, E. T. *J. Org. Chem.* **1998**, 3, 9652-9656.

(c) Morales, J. C. & Kool, E. T. *Nature Structural Biology* **1998**, 5, 11, 950-954.

(d) Kool, E. T. *Biopolymer,* **1998**, 48, 1, 3-17.

(21) (a) Leroy, J. L., Charretier, E., Kochoyan, M. & Gueron,M. *Biochemistry* **1988**, *27*, 8894–8898.

(b) Nadeau, J.G. and Crothers, D. M. *Proceedings of the National Academy of Sciences USA* **1989**, *86*, 2622–2626.

(c) MacDonald, D., Herbet, K., Zhang, X., Polgruto, T. and Lu, P. *Journal of Molecular Biology* **2001**, *306*, 1081–1098.

(d) Stefl, R., Wu, H., Ravindranathan, S., Sklenar, V. and Feigon, J. *Proceedings of the National Academy of Sciences USA,* **2004**, *101*, 1177–1182.

(e) Haran, T. E., Mohanty, U. *Quarterly Reviews of Biophysics* **2009**, *42*, 1, pp. 41–81.

(22) (a) Summers, J. S., Shaw, B. R. *Curr. Med. Chem*., **2001**, *8*, 1147-1155.

(b) Leumann, C. *Bioorg. Med. Chem*., **2002**, *10*, 841-854.

(c) Iyer, R. P., Roland, A., Zhou, W., Ghosh, K. *Curr. Opin. Mol. Ther*., **1999**, *1*, 344-358.

(23) Agris, P.F. *Nucleic Acids Res*. **2004**, *32*, 223-238.

(24) Eckstein, F. Ann. *Rev. Biochem*., **1985**, *54*, 367-402.

(25) (a) Chen, B., Quinlan, S. L., & Reid, J. G. *Tetrahedron Letters,* **1995**, *36, 44,* 7961-7964.

(b) Chen, C. & Cheng, Y. *J. Biol. Chem*. **1992**, *267*, 2856–2859.

(c) Huang, P., Farquhar, D. & Plunkett, W. *Journal of Biological Chemistry* **1992**, *267, 4*, 2817-2822.

(26) Frier, S. M., Altmann, K.-H. *Nucleic Acids Res*. **1997**, *25*, 4429-4443.

(27) Nielsen, P. E., Egholm, M., Berg and O. Buchardt. *Science*, **1991**, *5037*, 1497-1500.

(28) Obika, S., Nanbu, D., Hari, Y, Morio, K. *Tetrahedron Lett*. **1997**, *38*, 8735-8738.

(29) Stec, W.J., Wilk, A. *Angew. Chem. Int. Ed*. **1994**, *33*, 709-722.

(30) (a) Chibnall, A*. J. Biol. Chem*. **1924**, *61*, 303.

(b) Ackermann, D.; Mohr, M. Z. *Biol. Chem*. **1936**, *98*, 73.

(c) Waksman, S. A.; Woodruff, H. B. *Proc. Soc. Exp. Biol. Med*. **1942**, *49*, 207.

(d) Gordon, R. K.; Ginalski, K.; Rudnicki, W. R.; Rychlewski, L.; Pankaskie, M. C.; Bujnicki, J. M.; Chiang, P. K. *Eur. J. Biochem.* **2003**, *270*, 3507–3517.

(e) Yamada, T.; Komoto, J.; Lou, K.; Ueki, A.; Hua, D. H.; Sugiyama, K.; Takata, Y.; Ogawa, H.; Takusagawa, F. *Biochemical Pharmacology*, **2007**, *73*, 981–989.

(31) Irani, R. J.; SantaLucia, J. Jr. *Nucleosides, Nucleotides & Nucleic Acids,* **2002**, *21 (11–12)*, 737–751.

(32) Plosky, B. S.; Frank, E. G.; Berry, D. A.; Vennall, G. P; McDonald, J. P.; Woodgate, R. *Nucleic Acids Res*. **2008**, *36 (7),* 2152–2162.

(33) Gait, M. J., *Oligonucleotide Synthesis: A Practical Approach.* **1984**, IRL Press, Oxford, UK.

(34) Sinha N. D., Biernat, J., and Köster, H. *Nucleic Acids Res*. **1984**, *12*, 4539-4557.

(35) Itakura, K., Rossi, J.J., and Wallace, R. B. *Annu. Rev. Biochem*. **1984**, *53*, 323-356.

(36) Greco, N.J., Tor, Y. *Nature Protocols*, **2007**, *2*, 305 - 316.

(37) Vorbruggen, H., Ruh-Pohlenz, C. *Handbook of Nucleoside Synthesis*. **2001**, Wiley & Sons.

(38) Kazimierczuk, Z., Cottam, H.B., Revankar,G.R., Robins, R.K. *J. Am. Chem. Soc.,* **1984**, *106*, 6379-6382

(39) Arico, J. *3-Substituted purines*. **2010**, Boston College.

(40) Arico, J. W.; Calhoun, A. K.; Salandria K. J.; McLaughlin, L. W. *Org. Lett*. **2010**, *12 (1),* 120-122.

(41) Whitesides, G. M.; Brainard, R. L.; Lavalleye, J. P. P. M.; Sowinski, A. F.; Izumi, A. N.; Moore, S. S.; Brown, D. W.; Staudt, E. M. *Organometallics* **1985**, *4*, 1819-1830.

(42) Kubinec, M.G., Wemmer, D.E., *J. Amer. Chem. Soc*., **1992**, *114*, 8739-8740.

(43) Vlieghe, D., Turkenberg,J.P., Van Meervelt, L., *Acta Crystallogr*., **1999**, D55, 1495-1502.

(44) Rueda, M., Cubero, E., Laughton, C.A., Orozco, M. *Biophys. J.,* **2004**, *87*, 800-811.

(45) Westhof, E., Beveridge, D.L., *Water Sci. Rev*., **1990**, *5*, 24-136.

(46) Lipanov, A. A., Chuprina, V. P. *Nucleic Acids Research*, **1987**, *15*, 5833–5844.

(47) (a) Wing, R.; Drew, H.; Takano, T.; Broka, C.; Takana, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *287*, 755–758.

(b) Alexeev, D. G.; Lipanov, A. A.; Skuratovskii, I. Y. *Nature* **1987**, *325*, 821–823.

(c) Burkhoff, A. M.; Tullius, T. D. *Cell* **1987**, *48*, 935–943.

(48) (a) Berman, H.M. *Curr. Opin. Struct. Biol*., **1991**, *1*, 423-427.

(b) Berman, H. M., *Curr. Opin. Struct. Biol*. **1994**, *4*, 345-350.

(49) (a) Quintana, J. R., Grzeskowiak, K., Yanagi, K, Dickerson, R. E. *J. Mol. Biol.* **1992**, *225*, 379-395.

(b) Lipanov, A., Kopka, M.L., Kaczor-Grzeskowiak, M., Quintana, J., Dickerson, R.E. *Biochemistry*, **1993**, *32*, 1373-1389.

(50) Privé, G.G., Yanagi, K., Dickerson, R.E. *J. Mol. Biol*., **1991**, *217*, 177-199.

(51) Jason K. Rockhill, Scott R. Wilson, and Richard I. Gumport *J. Am. Chem. Soc*. **1996**, *118*, 10065-10068.

(52) (a) Wang, R.-W.; Gold, B. *Org. Lett*. **2009**, *11*, 2465–2468.

(b) Ganguly, M; Wang, R-W.; Marky, L. A.; Gold, B*. J. Am. Chem. Soc*. **2009**, *131*, 12068–12069.

(53) Lever, C., Li, X., Cosstick, R., *Nucleic Acids Research*, **1993**, *21, 8,* 1743-1746.

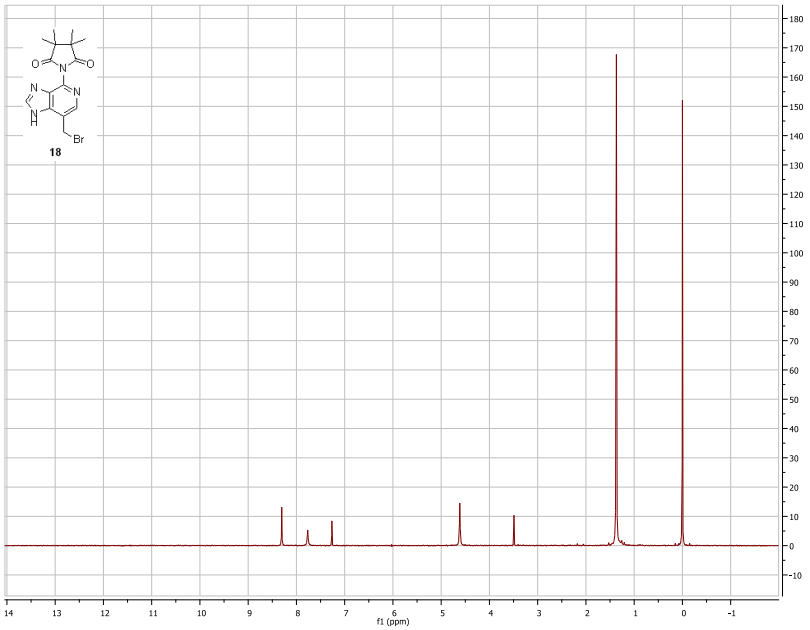
(54) (a) Lan, T. and McLaughlin, L. W. *J.A.C.S*., **2000**, *122*, 6512.

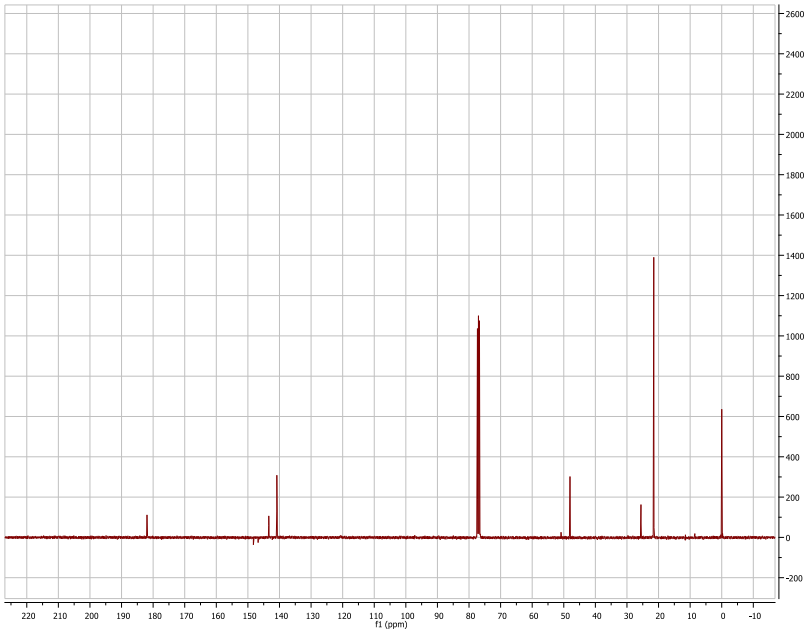
(b) Lan, T. and McLaughlin, L. W. *Biochemistry*, **2001**, *40*, 968-976.

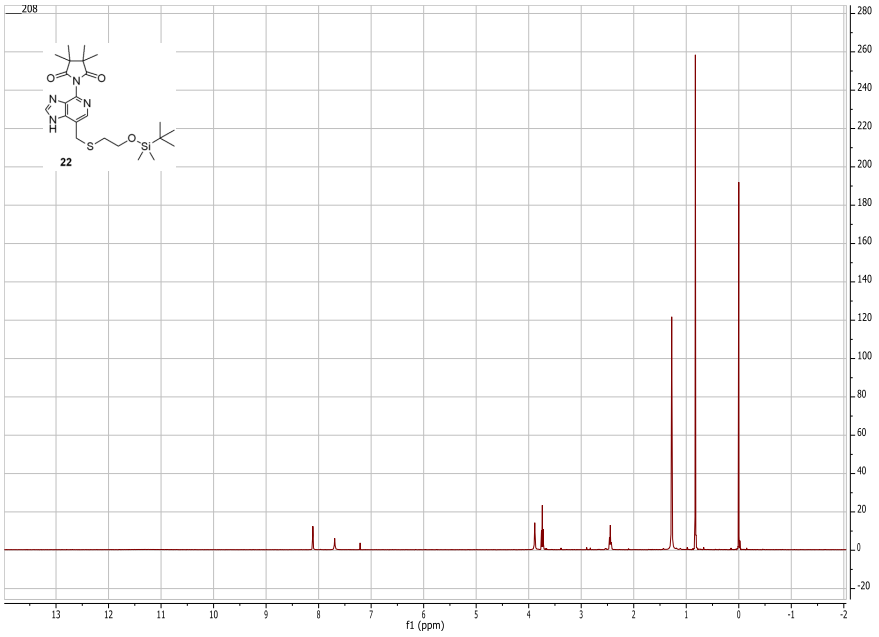
(55) Salandria, K. J., Arico J. W., Calhoun, A. K., and Larry W. McLaughlin. *J.A.C.S*., **2011**, *133*, 1766-1768.

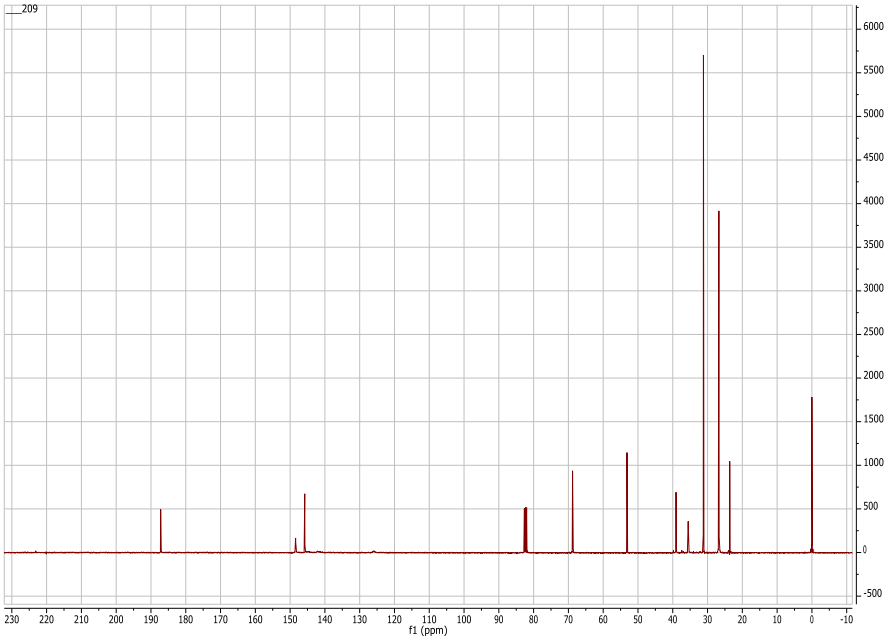
(56) Alpegiani, M., Bedeschi, A., Perrone, E., Zarini, F., Franceschi, G. *Heterocycles*. **1985**, *23, 9,* 2255.

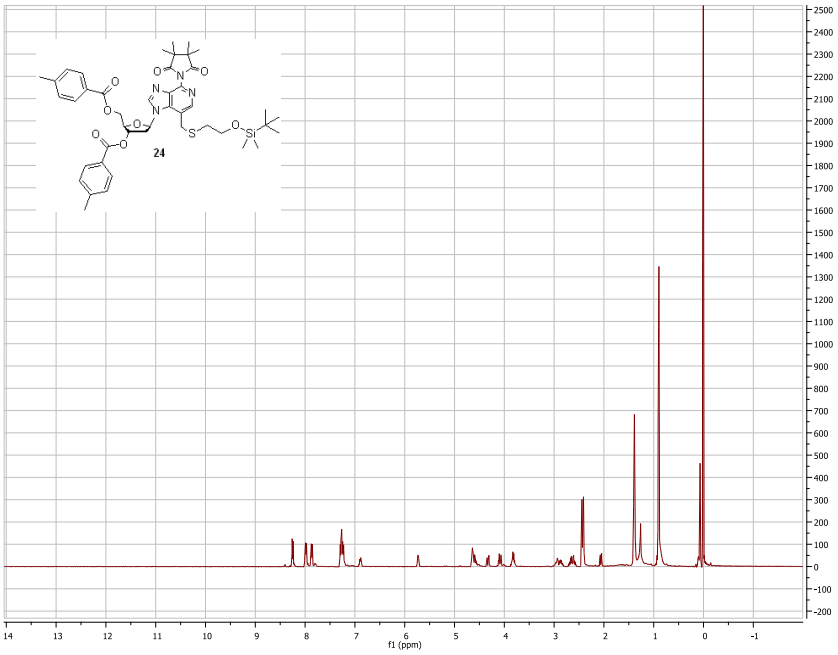
(57) Rolland, V., Kotera, M., Lhomme, J. *Syn. Comm*., **1997**, *27 (20),* 3505-3511.

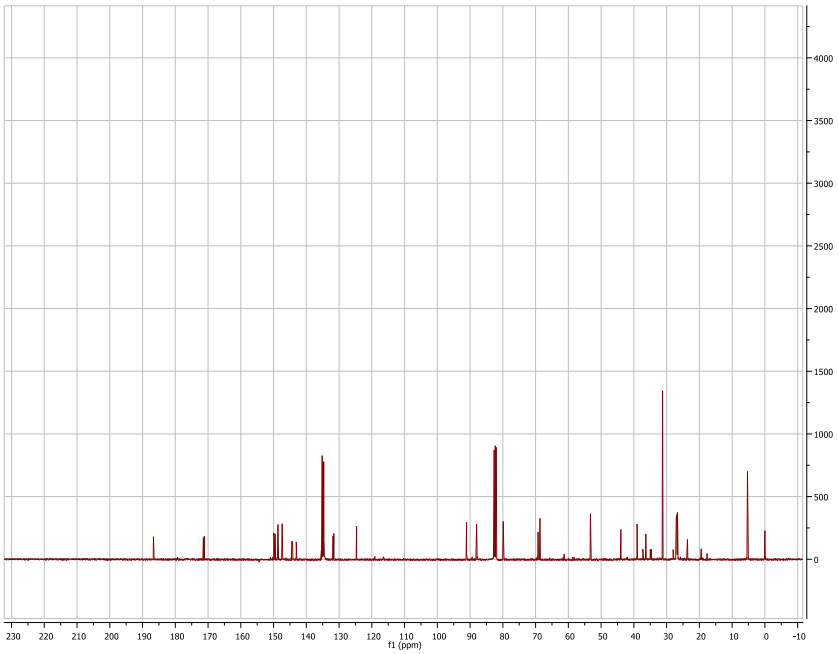

O
N
O
N
N
H
Br
18
f1 (ppm)

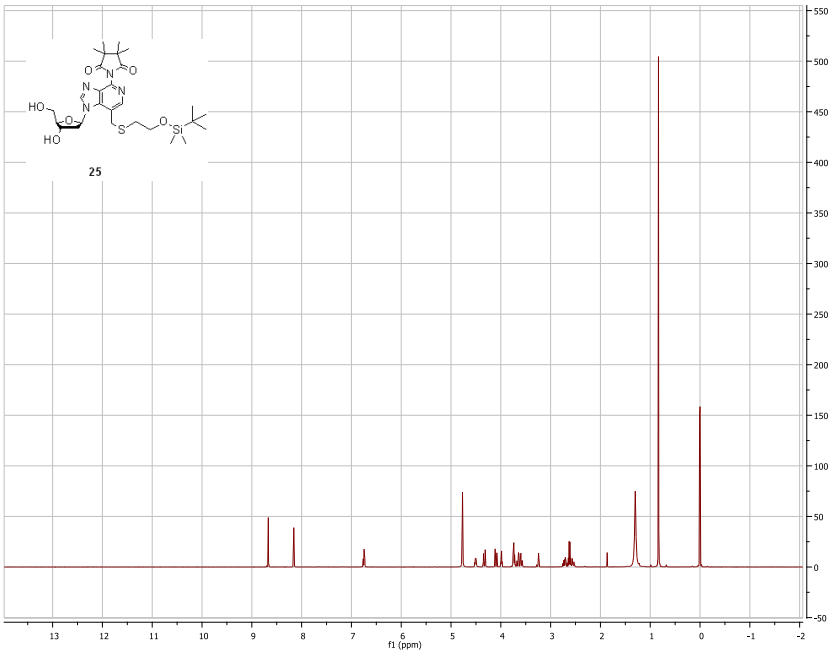

f1 (ppm)

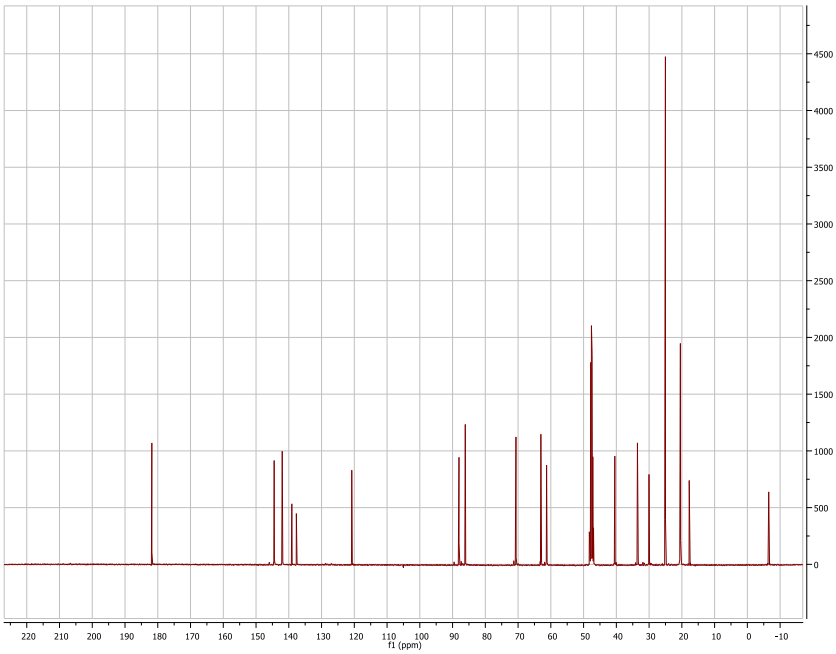

208
22
f1 (ppm)

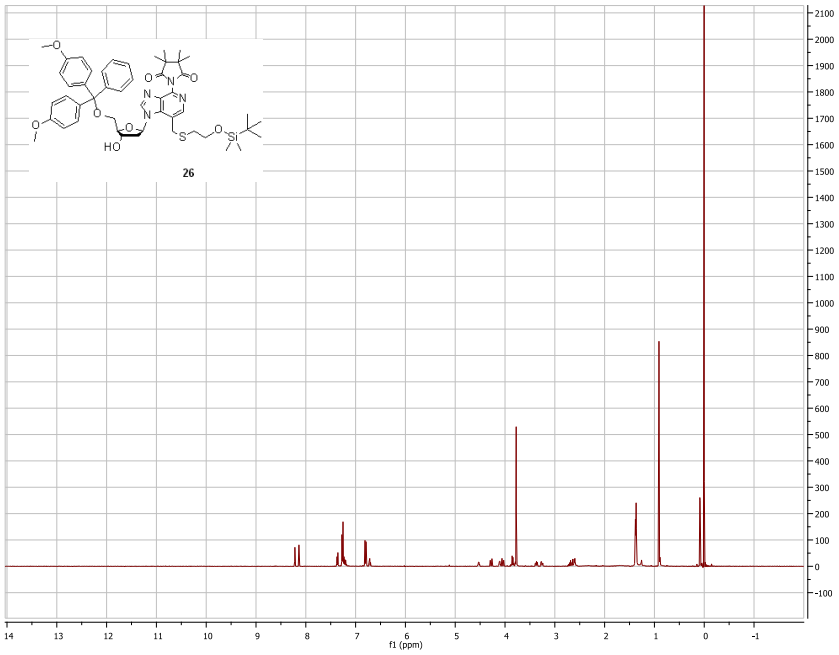

209
f1 (ppm)

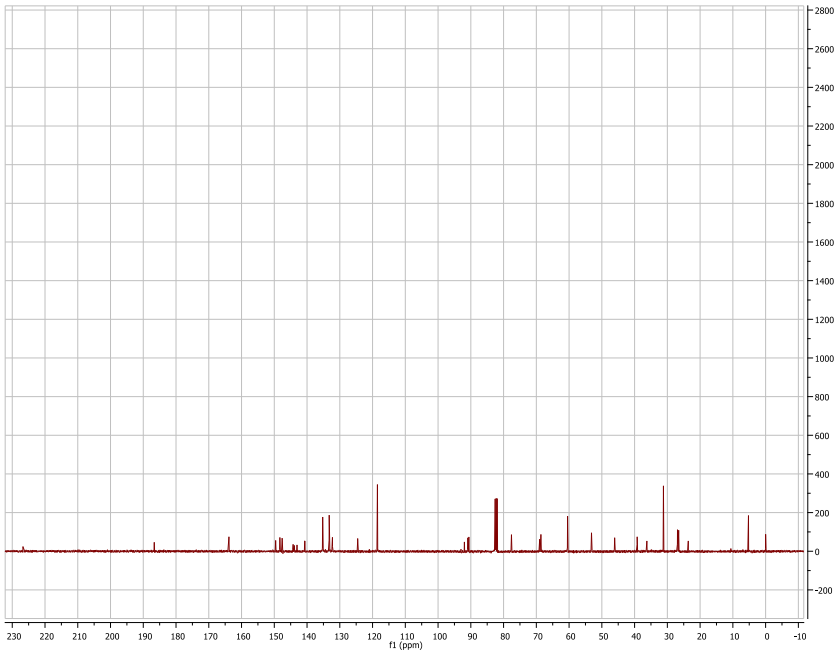

24
f1 (ppm)

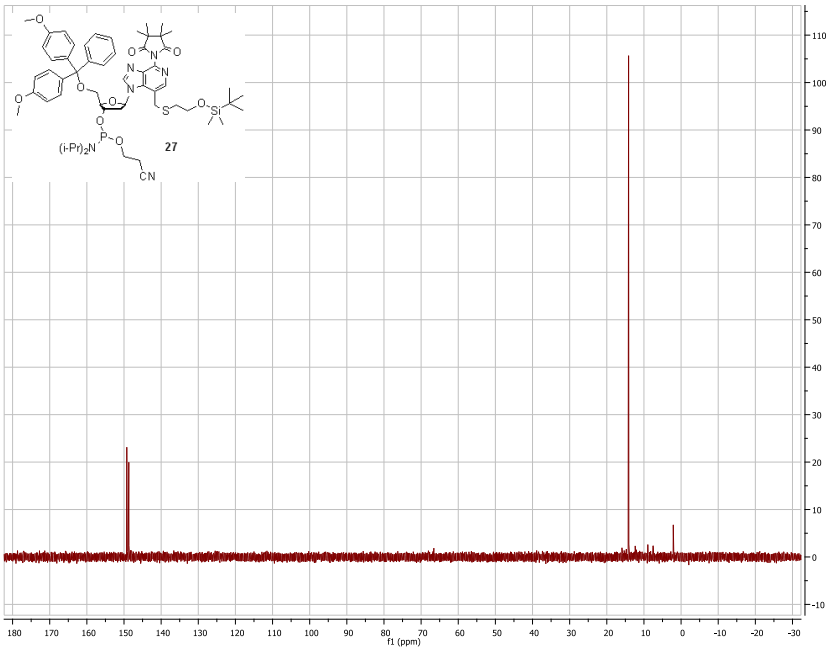

f1 (ppm)


25
f1 (ppm)


f1 (ppm)


26
f1 (ppm)


f1 (ppm)


27
f1 (ppm)