Synthesis and Study of Rigidified Nucleosides Analogues for Probing the Importance of the Deoxyribose DNA Backbone

Author: Han Yueh

Boston College

The Graduate School of Arts and Sciences

Department of Chemistry

SYNTHESIS AND STUDY OF RIGIDIFIED NUCLEOSIDES ANALOGUES FOR PROBING THE IMPORTANCE OF THE DEOXYRIBOSE DNA BACKBONE

A dissertation

by

HAN YUEH

submitted in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

October 2012

For Nan

ACKNOWLEDGEMENTS

Six years, over two thousand days, the journey of pursuing my Ph.D. has been amazing. Till now, it is still like a dream to me for being at this stage of my life. It is always hard to leave your country and the people you love to a new world to start the life on your own. Words can not express how grateful I am of all the people that have met since I came here.

First and foremost, I would like to express my gratitude to Prof. McLaughlin, my advisor. He gave me the opportunity to study and work in his research laboratory and led me into the world of nucleic acid chemistry. His guidance and advice have proven to be invaluable during my tenure in his laboratory, he showed me the path to become a scientist and a great person.

I have had the opportunity to work with a number of talented post-doc, graduate students, and undergraduates in Prof. McLaughlin's group. I appreciate the help from Dr. Christianson and Dr. Schlegel for revising my papers and thesis. I am indepted to Dr. Horhota who paved the way for the research on the rigidified nucleosides in this lab. I owe my thanks to Dr. Greco, Dr. Chen, and Dr. Arico for kind help and advice. I would like to thank Dr. Theile for the helpful discussions and advice on science and for being a great friend in my life. I would also like to thank my colleagues: Eric, Ayan,

Nick, Alena, Yiran and all the people I have been working with, for making graduate life in this laboratory a truly enjoyable experience.

I would like to thank Dr. Boylan and Dr. Jayasundera for the help on NMR studies, Dr. Li for the help on Crystallography, and Mr. Domin for the Mass. Since the first day I came to U.S., joined Boston College, I got a lot of help from Administrative personal of the Chemistry Department, and staff of the Office of International Students and Scholars made it easier for me to adopt to this environment.

Finally, I would like to express by deepest gratitude to my parents and my sister for their endless love, prayers and encouragement. My mother has provided the unconditional support, and without her I would never have been able to accomplish this. I would also like to thank the rest of my family and friends for all their help and support.

Han

2012 at Boston

Table of Contents

Table of Figures

Table of Schemes

List of Tables

Table of Abbreviations

A	adenosine
ADP	adenosine 5'-biphosphate
AMP	adenosine 5'-phosphate
ATP	adenosine 5'-triphosphate
aXNA	anhydrous xylose nucleic acid
BBr_3	boron tribromide
BH_3-DMS	dimethylsulfide borane
BTN	bridged threose nucleic acid
BzCl	benzoyl chloride
C	cytidine
CAN	cerium (IV) ammonium nitrate
cCMP	cytosine 3',5'-cyclic phosphate
CD	circular dichroism
$CDCl_3$	deuterated chloroform
COSY	correlation spectroscopy
cyclo-dA	cyclo-8,5'-2'-deoxyadenosine
cyclo-dU	cyclo-6,5'-2'-deoxyuridine
dA	2'-deoxyadenosine
dC	2'-deoxycytidine
DCM	dichloromethane
dG	2'-deoxyguanine

DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DMTrCl	4,4'-dimethoxytrityl chloride
DNA	β-D-2'-deoxyribonucleic acid
ds-DNA	double-stranded DNA
dT	2'-deoxythimidine
dU	2'-deoxyuridine
D_2O	deuterium oxide
Et_2O	diethyl ether
Et_3N	triethylamine
G	guanosine
H_2	hydrogen
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
IPA	isopropyl alcohol
IUPAC	international union of pure and applied chemistry
LDA	lithium diisopropylamide
LiCl	lithium chloride
LNA	locked nucleic acid
MeCN	acetonitrile
MeOH	methanol
$MgCl_2$	magnesium chloride

mRNA	messenger RNA
$NaBH_4$	sodium borohydride
NaCl	sodium chloride
$NaHCO_3$	sodium bicarbonate
NH_3	ammonia
NH_4Cl	ammonium chloride
NH_4OH	ammonium hydroxide
NMR	nuclear magnetic resonance
NOESY	nuclear Overhauser effect spectroscopy
Pd	palladium
Pyr	pyridine
R_f	retention factor
SeO_2	selenium dioxide
ss-DNA	single-stranded DNA
T	thymidine
TBAF	tetra-*n*-butylammonium fluoride
TBP	TATA-binding protein
*t*BuOK	potassium-*tert*-butoxide
TEA	triethylamine
TEAA	triethylammonium acetate
THF	tetrahydrofuran
TIPSCl	triisopropylsilyl chloride

TLC	thin layer chromatography
T_M	thermal melting point
TMSCl	trimethylsilyl chloride
TNA	threose nucleic acid
tRNA	transfer RNA
UV	ultraviolet spectroscopy
XNA	xylose nucleic acid
9-BBN	9-borabicyclo[3.3.1]nonane

Abstract

SYNTHESIS AND STUDY OF RIGIDIFIED NUCLEOSIDES ANALOGUES FOR PROBING THE IMPORTANCE OF THE DEOXYRIBOSE DNA BACKBONE

By Han Yueh

Under the direction of Professor Larry W. McLaughlin

Nucleic acids are the only biopolymers capable of encoding and transferring information, this property has placed them at the fundamental core of all living organisms, and made them a topic of intense research for over a century. The former studies in our laboratory on simplified nucleic acid backbones provided insight into how we might rationally alter nucleic acid structure into one that possesses properties not observed in natural DNA and RNA. Here the work began as an investigation into rigidified nucleic acid systems capable of functioning as DNA.

The first rigidified nucleic acid system we designed, the cyclo-2'-deoxynucleic acids, has a linkage between the C5' of the ribose sugar and base to lock the χ angle into the similar angle as the native nucleoside. These rigidified bases show great impact towards the DNA structure, destabilizing

double helix formation. This in fact can also be found in nature to inhibit the TATA binding protein associating with its target region.

The next generation of rigidified nucleosides has an extended 7-membered ring instead of the 6-membered ring that was present in the first generation to push the base closer to the helical center. Both diastereomers of the ring-expanded-cyclo-2'-deoxyadenosine have been successfully synthesized and characterized and are ready to perform further studies.

The third system is the hydroxymethyl-cyclo-nucleosides. The modified nucleosides in this project not only have the same linkage as the cyclo-nucleosides in the first system to restrict the base rotation, but also have an extra carbon (C6') to give the backbone more flexibility which might better stabilize a double helix than the first generation.

Chapter 1

Introduction

1. Introduction

1.1 Discovery of Nucleic Acids

In the late 1800's, Miescher isolated a phosphorus and nitrogen containing substance that was commonly observed in various cell types, and started the research into this substance which he named "nuclein".[1] Research on nuclein has remained steady since then, and resulted in the discovery of ribose sugars linked through phosphodiester bonds forming long polymers (Figure 1.1). Klein and Thannhauser enzymatically cut these large and complex biopolymers down to the monomer level making it possible to determine real structural information. In 1951 Todd synthesized and confirmed the four deoxynucleoside monomers, and established the D-configuration and β-glycosidic linkage successfully[2,3]. The key to success was the development of methods of phosphorylation; for example, the preparation of the 3'- and 5'- phosphates of deoxyadenosine (Figure 1.2)[4]. These results supported the linear structure through phosphodiester linkages and further explained the earlier crystallographic data produced by Astbury and Hammarsten which postulated a 3.34 Å repeating unit composed of flat nucleotides standing out perpendicularly to the rod shaped sugar-phosphate backbone.

Years later after Hammarsten and Astbury, Gullard postulated the presence of hydrogen bonding interactions between the purines (A, G) and pyrimidines (C, T). He suggested that these hydrogen bonds could involve

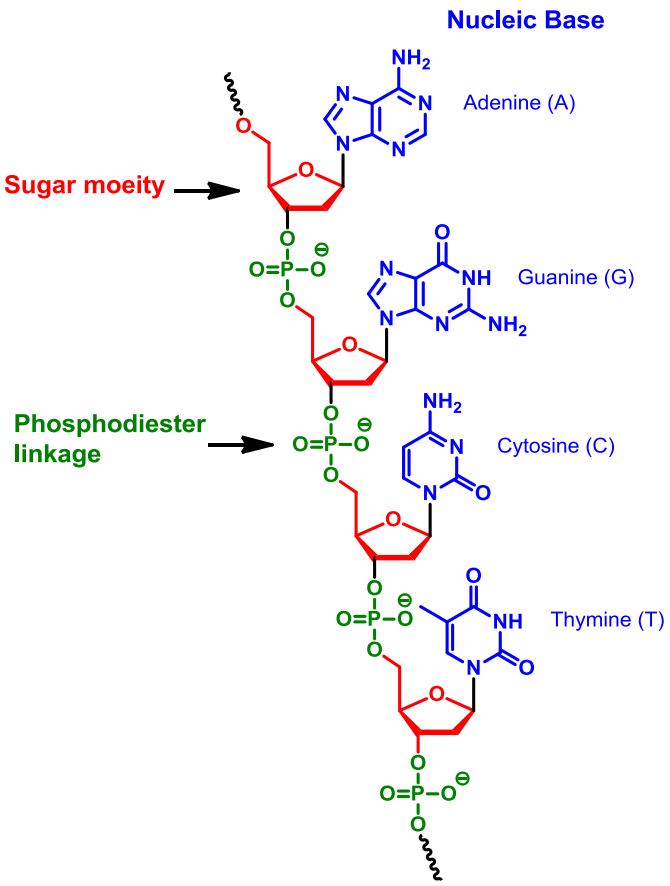


Figure 1.1. The three major components, nucleobase, carbohydrate, and phosphodiester linkage in a single strand of DNA.

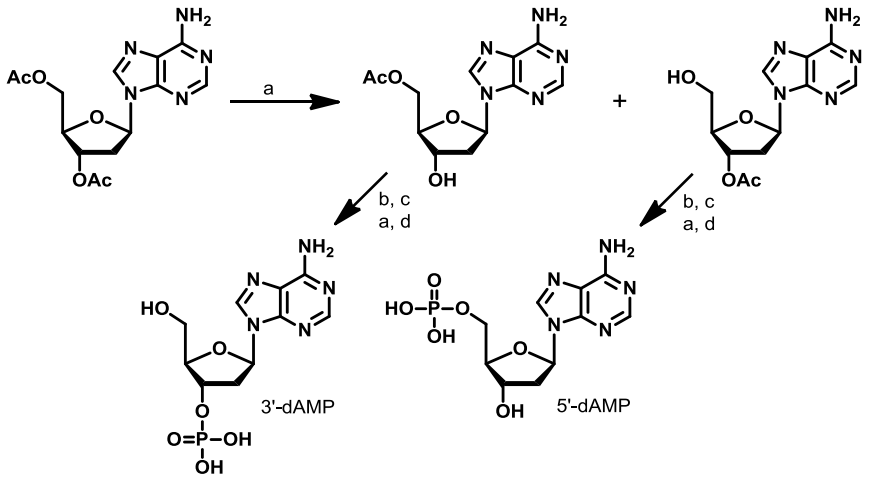


Figure 1.2. Todd's synthesis of deoxyadenosine 3'- and 5'-phosphates. a) MeOH, NH_3; b) $(PhO)_2P(O)OP(H)(O)OCH_2Ph$; c) N-chlorosuccinimide; d) H_2/Pd on Carbon. Image taken from Blackburn[4].

nucleotides either in adjacent chains or within a single chain. Chargaff then began to investigate the base composition of DNA from a variety of sources using the paper chromatography technique to separate the products of hydrolysis of DNA and employing an UV spectrophotometer to quantify their relative abundance.[5] His research proved that the number of purines is always equal to the number of pyrimidines, and therefore supported Gullard's postulation.

In 1953, while Watson and Crick were trying to construct a model of DNA structure whch was supported by previous research, they realized that the enols of T and G could actually tautomerize into the keto form.[6] (Figure 1.3) The most critical information came from the crystal data provided to them by Wilkins and Franklin. Franklin believed their crystal data showed a helix in which the nucleobases were sequestered from water on the inside of the helix with phosphate exposed on the exterior.

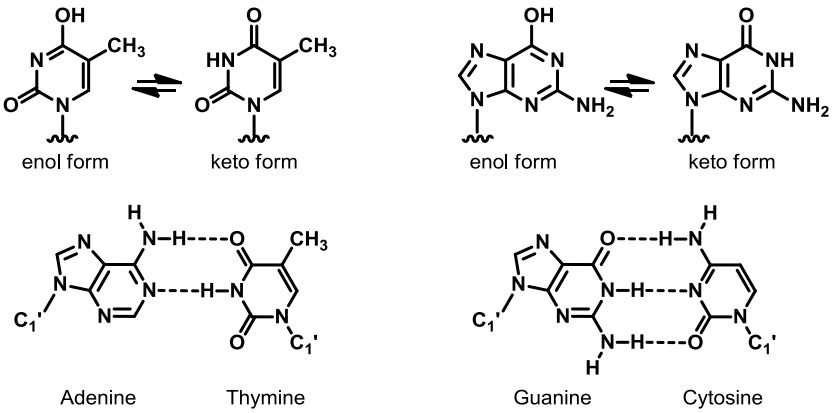


Figure 1.3. Complemetary hydrogen-bonded base-pairs as proposed by Watson and Crick. The thymine and guanine are showing in the revised keto forms.[7]

Watson and Crick thereafter created a model with the following major features (Figure 1.4): 1. Two polynucleotide chains wind around the center axis to form a double helix. 2. Two DNA strands run in opposite directions (antiparallel), each one forms a right-handed helix. 3. The bases occupy the core of the helix, and the surface of the double helix contains two grooves with unequal width, major and minor grooves. 4. Each base forms hydrogen bonds with a base in the opposite strand resulting in a planar base pair. This is the phenomenon know as complementary base pairing, resulting in the specific association of the double helix.

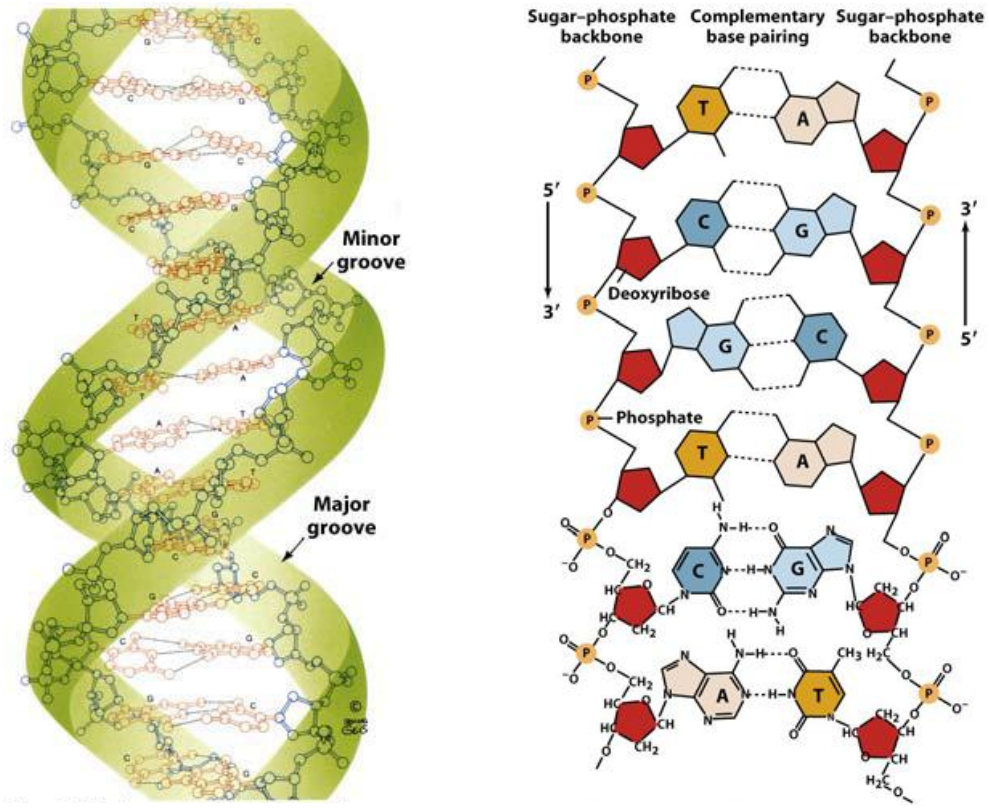


Figure 1.4. Three-dimensional structure of DNA (left), complementary strands of DNA (right). Image taken from Voet.[8]

1.2 DNA Structure

A nucleotide is the phosphate ester of a nucleoside which is composed of a nitrogenous heterocyclic nucleobase, a five-carbon sugar (either ribose or 2-deoxyribose), and one phosphate group; it can also be called a nucleoside monophosphate. The phosphate group forms bonds with either the 2, 3, or 5-carbon of the sugar, with the C5 site most common. They are components of both ribonucleic acid (RNA) and deoxyribonucleic acid (DNA).

The nucleobases are either monocyclic pyrimidines or bicyclic purines. The major purines, adenine (A) and guanine (G), and are found in both DNA and RNA. The major pyrimidines are cytosine (C), thymine (T) found in DNA, cytosine and uracil (U) in RNA (Figure 1.5). In DNA, bases are connected at the anomeric position of 2-deoxy-D-ribose through a ring nitrogen forming the four nucleosides: 2'-deoxyadenosine (dA), 2'-deoxyguanosine (dG), 2'-deoxycytidine (dC), and 2'-deoxythymidine (dT).

The nucleosides functionalized with phosphate esters are called nucleotides, and the simplest of them have a single hydroxyl group of the sugar esterified by a mono-phosphate; for example, adenosine 5'-phosphate is a 5'-ribonucleotides, also abbreviated as AMP (Figure 1.6). A pyrophosphoric ester on the 5'-hydroxyl group is adenosine 5'-biphosphate (ADP), and ATP for adenosine 5'-triphoosphate. Nucleotides containing two phosphate monoesters are called nucleoside bisphosphates. The nucleosides which have

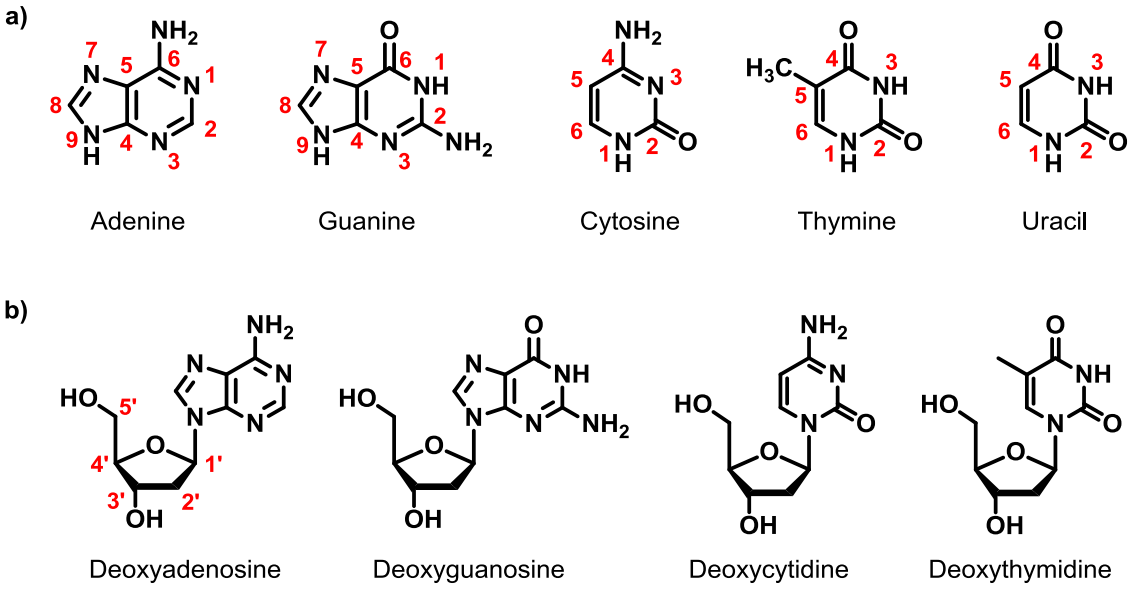


Figure 1.5. a) Structures of the five major nucleobases with the IUPAC numbering systems. **b)** Four major deoxyribonucleosides and the pentose carbon numbering system.

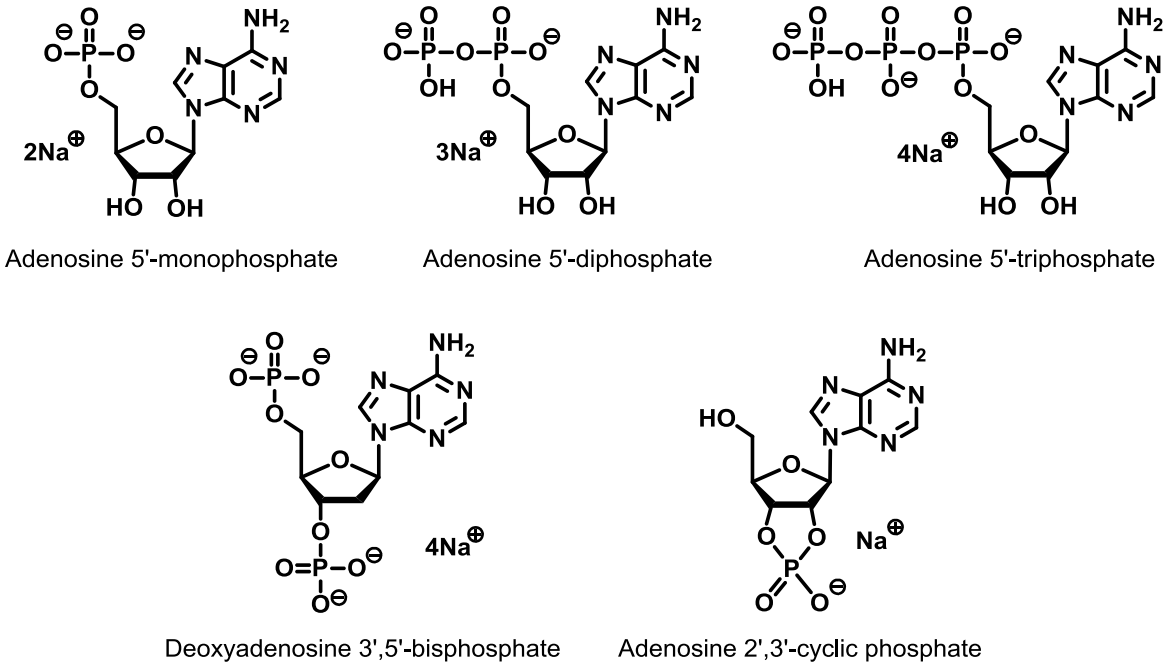


Figure 1.6. Structure of common nucleotides. All present as the sodium salts as observed at neutral pH.

two neighboring hydroxyl groups connected to a single phosphate are called cyclic nucleotides, as cytosine 3',5'-cyclic phosphate, cCMP.

The detailed conformational structure is defined by the torsion angles α, ß, γ, δ, ε, and ζ along the phosphate backbone, θ for the sugar, and χ for the glycosidic bond (Figure 1.7). We can describe the shape of nucleotides more easily by other parameters, like the sugar pucker and the *syn-anti* conformation of the glycosidic bond. The sugar "puckering" describes the major displacement of C-2' and C-3' from the C1'-O4'-C4' plane; the sugar moiety is flexible, but generally exits in two conformations: the 3' endo for RNA and 2' endo for DNA.

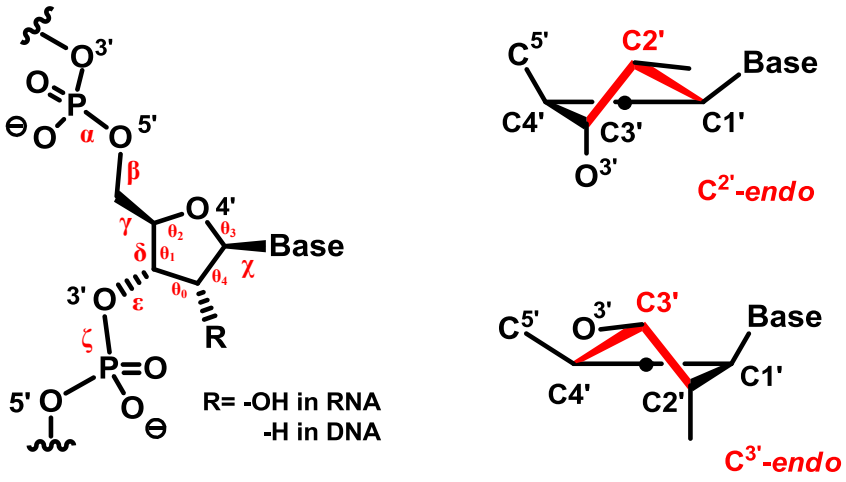


Figure 1.7. Torsion angles for polynucleotide chains in the phosphate backbone and the structures of the C2'-*endo* and C3'-*endo* sugar puckers. Image taken from Blackburn[4].

The plane of the bases is close to perpendicular to the sugars, it allows the bases to adopt two principle orientations (Figure 1.8). The *anti*

conformation has the smaller H-6 of pyrimidine and the H-8 of purine on top of the sugar ring; the *syn* conformation has the larger atoms (O-2 on pyrimidine, N-3 on purine) in that position. While pyrimidines occupy a narrow angle of *anti* conformations; purines may be found in a wider range of *anti* conformations that can extend into high-*anti* conformations for some 8-azapurine nucleosides.

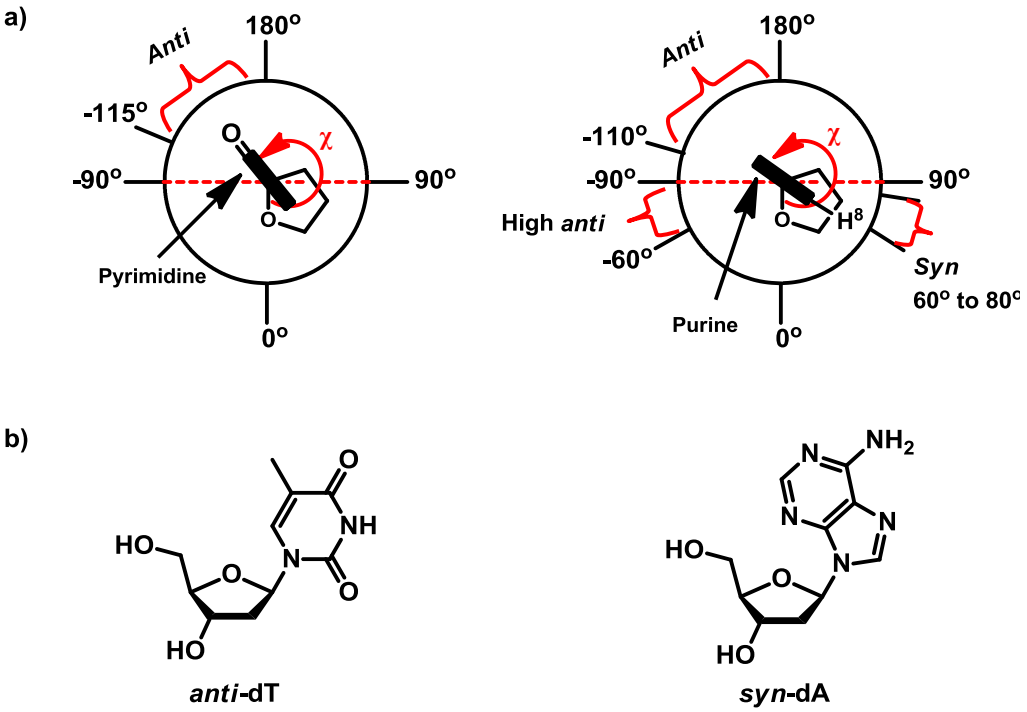


Figure 1.8. a) The range of glycosylic bonds in pyrimidine and purine nucleic acids for the *Anti* and *Syn* conformations. **b)** Examples of the *anti* conformation for deoxythimidine and the *syn* conformation for deoxyadenosine-5'-phosphate.

Crystal structures showed that DNA has a number of different forms. The three main types are the A-form, B-form, and the less common Z-form

(Figure 1.9). The biologically most common form of DNA is known as B-form DNA which was initially described by Watson and Crick. In the ideal B-form, the right-handed helix has 10 base pairs (bp) per turn, a helical twist of 36° per bp, and 34 Å rise per turn.

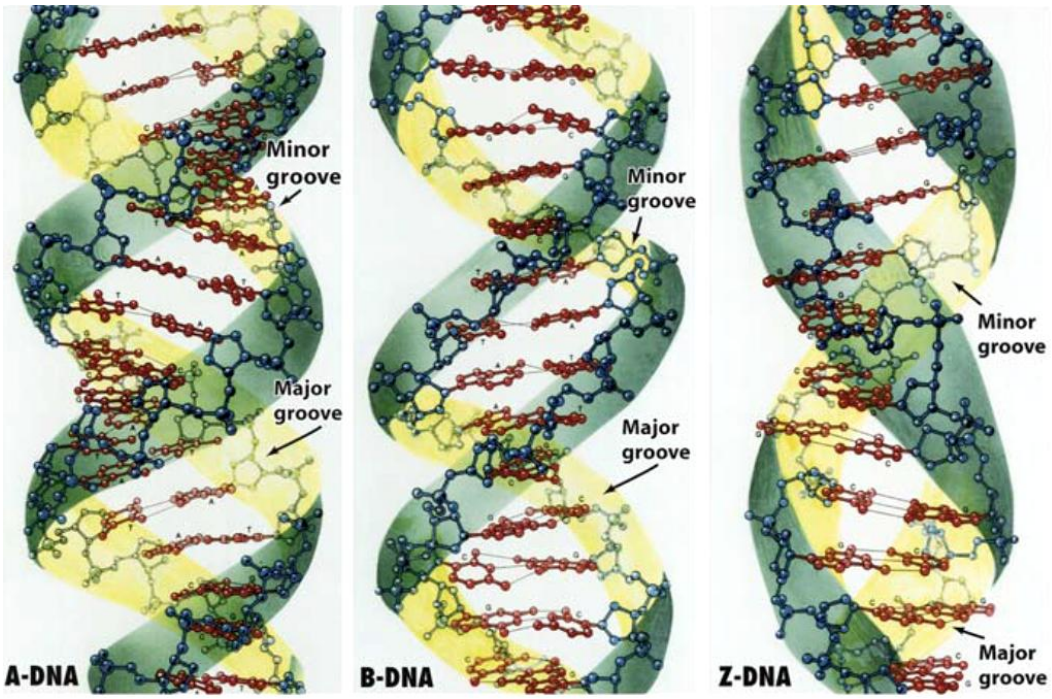


Figure 1.9. Structures of A-, B-, and Z-DNA forms. Image taken from Voet. [8]

Under dehydrating conditions, B-DNA can transform into A-DNA. A-DNA has 11 bp per turn, a pitch of 28 Å, and a wider right-handed helix than in the B-form. One of the most noticeable features in A-DNA is that the planes of base pairs are tilted 20° with respect to the helix axis.

DNA can be converted to the Z-form under certain conditions; such as alternating purine-pyrimidine sequence, especially poly(dGC)$_2$, and high salt

condition. The salt stabilizes Z-DNA relative to the B-form by reducing the electrostatic repulsions between the closest approaching phosphate groups on opposite strands. Z-DNA has a left hand helix that results from an alternating *syn-anti* conformation of the nucleobases along the helix axis.

1.3 Nucleic Acid Functions

DNA carries genetic information in all cells and in many viruses. Its primary function is the faithful storage and transmission of information from one generation to the next. Each strand of a DNA duplex acts as a template during replication. This process is not perfectly accurate, a low baseline level of mutation might occur. However these mutations have presumably allowed the evolution of many diverse forms of life.

According to the central dogma of molecular biology formulated by Crick in 1958, DNA directs its own replication and transcription to form an RNA of complementary sequence. The RNA is then translated into the corresponding amino acids to form proteins. In the ribosome, each set of three nucleotides in the mRNA (messenger RNA, a protein-coding gene) pairs with three complementary tRNA (transfer RNA, a small RNA molecule) nucleotides attached to the corresponding amino acid. The ribosome catalyzes the joining of amino acids, which are the monomeric units of proteins.

The discovery of the ribosome showed that RNA could also have chemical functionality. This discovery led to research on identifying numerous enzymatic RNAs, or so called ribozymes.[9,10,11]

1.4 Rigidified Nucleic Acid

One class of nucleosides that has garnered much attention is the rigidified nucleic acids. These nucleosides contain a chemical modification that reduces the flexibility of the molecule. Our interest in this class of nucleosides is based on the design of a nucleic acid system which is capable of information storage and still has the same functionalities as native RNA and DNA like replication, transcription, and translation. We hypothesize that by controlling the rigidity and the structure of these nucleic acids further into the oligomers, we may be able to direct duplex formation. We believe that a suitable rigidified structure may be able to minimize the entropic cost during duplex formation and may even be capable of performing nonenzymatic template directed polymerization.

In the past few decades, there are a great number of structurally restrained nucleosides that have been synthesized. There are two major categories for these rigidified nucleosides; the first class of these contains a modified sugar moiety, the second class is constructed by creating a linkage between the sugar and the nucleobase to restrain the structure.

In the first class, the modified nucleosides have a bicyclic or tricyclic ring system.[12,13,14] Other than on the pentose sugar, the modification can be realized at a variety of locations through the formation of many unique bridges (Figure 1.10). These conformationally constrained nucleosides may have the ability to interact with natural DNA or RNA, and can potentially inhibit protein synthesis. Inhibition could be realized through a number of different mechanisms, depending on which section is chosen as the target, a mRNA or a non-spliced pre-mRNA.

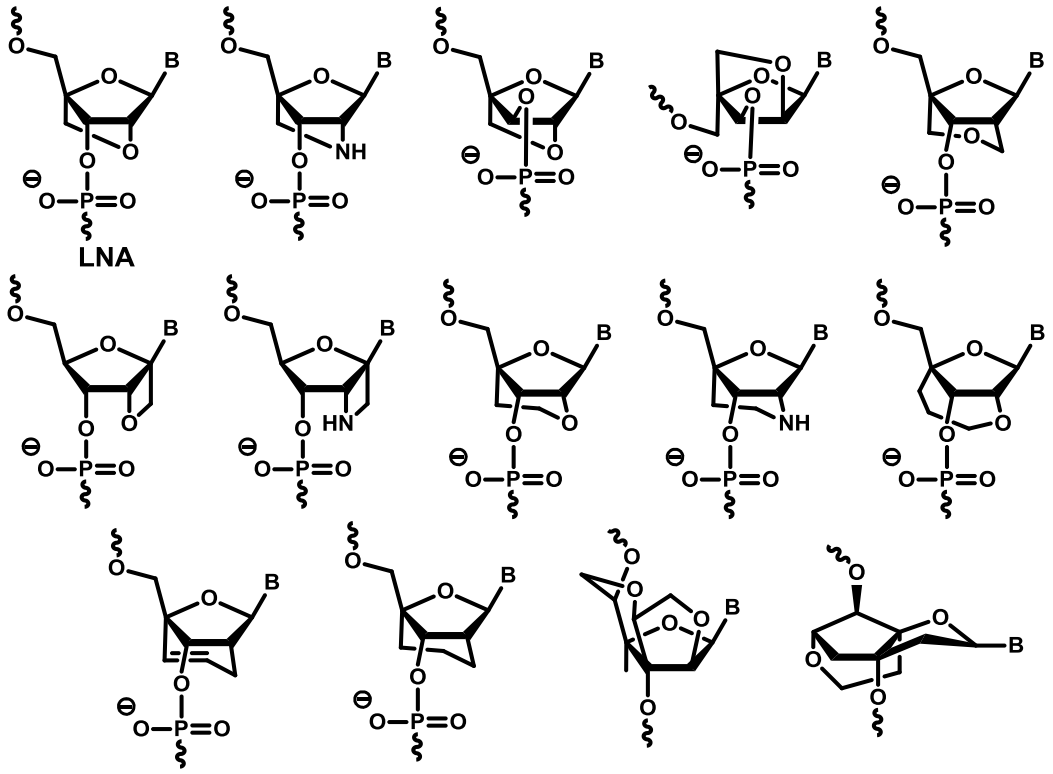


Figure 1.10. Conformationally restrained rigid nucleic acids, bicyclic and tricyclic (last two structures) molecules.[15,16] The most well studied is the original LNA nucleoside.

The two most common inhibition mechanisms are through the arrest of translation by antisense oligonucleotides and/or the ubiquitous cellular enzymes. The advantage of this concept over traditional small molecule protein targeting in drug design is that ideally only minimal structural knowledge about the target RNA is required; i.e. one only needs to know the sequence.

The most well known and studied rigidified nucleosides of this class are the Locked Nucleic acids (LNA). In 1997, Imanishi from Japan initially synthesized the uridine and cytosine derivatives of LNAs. When analyzing the crystal structures, they found the extra linkage in LNA locks the structure into $C_{3'}$-*endo* sugar puckering with an *anti*- nucleobase orientation.[17] A year later, the Wengel group from Denmark published the synthesis of A, G, C, T, U, and 5-Me-C derivatives of LNAs, and incorporated these monomers into DNA strands or homo-polymers. Thermal studies indicated that these molecules not only formed stable duplexes with both DNA and RNA, but they also stabilized the structure by 7 – 8 ºC per modified base.[11] The extra ring on the structure enhances the thermal stability of the duplex and improves the sequence specificity. The mechanisms behind this are the more favorable base stacking and lower entropic cost when the conformationally restrained nucleosides appear within an oligonucleotide duplex.

Another type of rigidified nucleoside contains an extra linkage to connect the nucleobase with the carbohydrate. The structures of these modified nucleosides are usually more rigid than the molecules in the first class, like LNAs. The greater level of rigidity results from both the extra cyclization which helps to rigidify the sugar moiety as in LNA, and limited rotation of the nucleobase around the glycosidic bond. The base orientation in normal nucleosides resides mainly in the *anti* conformation, but the nucleobases are free to rotate around the glycosidic bond, since it is not fixed.

The research of rigidified nucleosides in our lab began with Dr. Allen Horhota. Dr. Horhota had been working on two series of rigidified nucleosides (Figure 1.11), the BTN (Bridged Threose Nucleic Acid) and aXNA (Anhydrous Xylose Nucleic Acid).[18] BTN, which contains an extra methylene group to link the nucleobase with the $C_{4'}$ of the threose sugar, is derived from the simplified nucleic acid TNA (Threose Nucleic Acid). The aXNA has an anhydrous bond between the base and 3'OH of the xylose sugar and shifts the connection of the sugar backbone from 3'OH-5'OH (in XNA) to 2'OH-5'OH (in aXNA). As mentioned above, the extra linkages in these two cyclized nucleosides not only rigidifies the entire structure, but locks the glycosidic bond in a fixed position.

These rigid nucleosides were designed to decrease the flexibility and restrict the conformational rotation which we hoped it would result in a better thermal stability of duplexes. Unfortunately, the result of the sequence

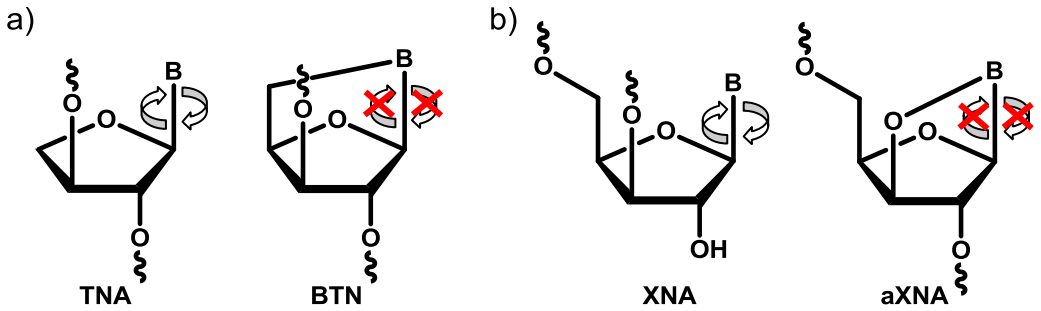


Figure 1.11. The structures of: a) Threose Nucleic Acid (TNA) and Bridged-Threose Nucleic Acid (BTN); b) Xylose Nucleic Acid (XNA) and Anhydrous Xylose Nucleic Acid (aXNA). The rigidified nucleic acids, BTN and aXNA, have extra linkages to connect the base with sugar moiety and lock the χ torsion angle.

studies showed that these molecules do not stabilize the helical structure, since there is no double helix formation when they were incorporated into a duplex. The reason for this could be due to the very rigid structure of these molecules, and also that the backbones are not in the natural orientation (3' to 5'). In BTN, the backbone is 2' to 3'; the backbone is 2' to 5' in aXNA.

This thesis describes the synthesis and properties of three rigidified nucleic acid systems: the cyclo-2'-deoxynuclosides, ring expanded-cyclo-2'-deoxynucleosides, and cyclo-2'-deoxy-5'-hydroxymethylnucleosides. The first rigidified nucleic acid system, cyclo-nucleoside, has an extra bond to lock the base with sugar moiety and maintains the 3'- to 5'-hydroxyl backbone linkage.

The next generation of rigidified nucleosides has an extended 7-membered ring instead of the 6-membered ring present in the first generation. We hoped this would push the base closer to the helix center and

improve the decreased thermal stability in cyclo-nucleosides. The third system not only has the same linkage as the cyclo-nucleosides in the first system to restrict the base rotation, but also have an extra carbon (C6') to give the backbone more flexibility which might better stabilize a double helix than the first generation.

1.5 References

1. F. Miescher, *Die Histochemischen und Physiologischen Arbeitn.* 1897.
2. D.H. Hayes, A.M. Michelson, A.R. Todd, Mononucleotides derived from deoxy-adenosine and deoxyguanosine. *J. Chem. Soc.* **1955,** 808-815.
3. J.G. Buchanan, A.R. Todd, *Adv. Carbohydr. Chem.*, **2000,** *55*, 1-13.
4. G.M. Blackburn, M.J. Gait, D. Loakes, D.M. Williams, *Nucleic Acids in Chemistry and Biology.* 3rd Edition, Royal Society of Chemistry: Cambridge, UK, 2006.
5. E. Chargaff, Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia*, **1950,** *6*, 201-209.
6. J.D. Watson, F.H.C. Crick, A structure for deoxyribose nucleic acid. *Nature,* **1953,** *171*, 737-738.
7. Saenger, *Principles of Nucleic Acid Structure.* Springer-Verlag: New York, 1984.

8. D. Voet, J.G. Voet, C.W. Pratt, *Fundamentals of Biochemistry*. John Wiley & Sons: New York, 1999.

9. T.R. Cech, *J. Am. Med. Assoc.*, **1988**, *260*, 3030-3034.

10. R.F. Ketting, T.H. Haverkamp, H.G. van Luenen, R.H. Plasterk, *Cell,* **1999**, *99*, 133-41.

11. J.A. Esteban, N.G. Walter, G. Kotzorek, J.E. Heckman, J.M. Burke, *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6091-6096

12. P. Srivastava, J. Barman, W. Pathmasiri, O. Plashkevych, M. Wenska, J. Chattopadhyaya, *J.Am. Chem. Soc.* **2007,** *129*, 8362-8379.

13. C.J. Leumann, *Bioorg. Med. Chem.* **2002,** *10,* 841-854.

14. A.A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen, J. Wengel, *Tetrahedron* **1998**, *54* , 3607-3630.

15. P. Nielson, M. Peterson, J.P. Jacobsen, ***J****. Chem. Soc., Perkin. Trans. 1.* **2000,** *24,* 3706-3713.

16. B.M. Keller, C.J. Leumann, *Angew. Chem. Int. Ed.* **2000**, *39*, 2278-2281.

17. S. Obika, D. Nanbu, Y. Hari, K. Morio, Y. In, T. Ishida, T. Imanishi, *Tetrahedron Letter*, **1997**, *38*, 8735-8738.

18. A. Horhota, The Synthesis and Study of Modified Nucleic Acid Backbone to Probe The Requirements of Funtional Nucleic Acid Systems. Ph.D. Thesis, Boston College, Chestnut Hill, MA, November 2007.

Chapter 2

Synthesis and Properties of DNA Containing Cyclo-2'-Deoxynucleic Acids*

* Han Yueh, Hongchuan Yu, Christopher S. Theile, Ayan Pal, Allen Horhota, Nicholas Greco, and Larry W. McLaughlin. *Nucleosides, Nucleotides, and Nucleic Acids*. **2012,** *31*, 661-679.

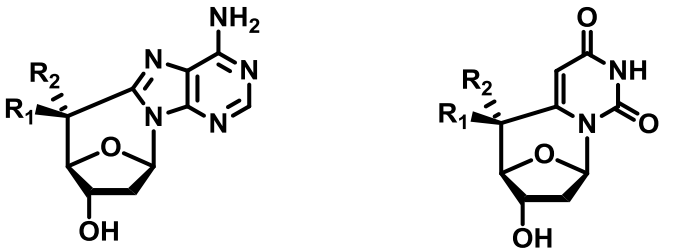
2. Synthesis and Properties of Cyclo-2'-Deoxynucleic Acids

2.1. Introduction

DNA damage is a ubiquitous part of the cellular life cycle. Oxidative stress is particularly problematic for nucleic acids. In many cases this stress is initiated by a free radical, commonly the hydroxyl radical (•OH). Hydroxyl radicals are generated by both exogenous agents such as carcinogenic molecules and ionizing radiation, as well as endogenous sources.[1] DNA damage by free radicals is commonly termed as "oxidatively induced damage to DNA" in living cells. Such damage has been shown to result in mutagenesis, carcinogenesis, and aging.[2] In terms of specific reactivity between these reactive oxygen species and DNA, two main adducts are likely: the addition of oxygen to the unsaturated bonds of the heterocyclic bases and abstraction of hydrogen atoms from the sugar backbone. H-abstraction has also been shown to occur at the methyl group of thymine, generating thymine dimers. As a result of these various damage pathways, a number of nucleobase and deoxyribose perturbations lead to single and double strand breaks, formation of abasic sites, DNA–protein and carbohydrate-nucleobase cross-links.[3-7]

In this study we focus on 8,5'-cyclo-2'-deoxyadenosine and 6,5'-cyclo-2'-deoxyuridine nucleosides (Figure 2.1.). 8,5'-cyclo-2'-deoxyadenosine represents a unique class of naturally occurring, helix-distorting DNA lesions that result in a second covalent linkage between the sugar and base moieties

of a single nucleoside.[8] Unnatural 6,5'-cyclo-2'-deoxyuridine derivatives were also investigated, as they have similar structural features compared to the cyclo-deoxyadenosine moieties.



1*R*: R_1 = H, R_2 =OH
1*S*: R_1 = OH, R_2 =H

2*R*: R_1 = H, R_2 =OH
2*S*: R_1 = OH, R_2 =H

Figure 2.1. The *R* and *S* diastereomers of 8,5'-cyclo-2'-deoxyadenosine and 6,5'-cyclo-2'-deoxyuridine.

These lesions are initiated by H-abstraction at the C5' by the hydroxyl radical species. This is followed by attack of the C5'-centered radical at the C8 of the purine base, leading to C8–C5'-cyclization. Subsequent oxidation of the resulting N7-centered radical results in the intramolecular cyclized product with a covalent bond between the C5' and C8 positions of the purine nucleoside.[9]

Both *R* and *S* diastereomers (Figure 2.1) are possible and although they result in similar puckering and distortion, the unique position of the 5' OH causes them to have different properties.[8, 10-13] It has been shown that the *R*-diastereomer is more prevalent under γ-radiation of an anaerobic water

solution of ss-DNA (single stranded-DNA), whereas the *R*/*S* ratio is reversed in favor of the *S*-diastereomer in ds-DNA (double stranded-DNA).[14] Moreover, it is postulated that the *R*- and *S*-diastereomers of cyclo-dA may be recognized differently during nucleoside excision repair due to the stereospecificity of the repair enzymes. Studies have shown that the repair synthesis was more efficient in the plasmid substrate containing the *R*-diastereomer than in the substrate with the *S*-diastereomer lesion.[15, 16] These differences demand that each diastereomer needs to be prepared separately, which lead to our synthetic efforts in this chapter.

Cyclonucleoside lesions have major biological implications. These include blocking transcription and mutagenesis.[8] Studies have proven that the lesions can alter the binding of transcription factors to their cognate DNA recognition elements.[16] Cyclonucleosides can block the TATA binding protein (TBP) from associating with its target region, causing a transcriptional roadblock, which in turn shuts off gene expression. The exact mechanism of inhibition is not yet known.[8] TBP binds in the minor groove, causing a major local conformational distortion.[17] One of the reasons that nature may have chosen the TATA box instead of a GCGC box for transcriptional control is the ease of A·T minor groove deformation rather than the stronger G·C base-pairs.[17] Since cyclonucleosides block TBP binding, several factors may be operating: the rigidity imparted by the extra covalent bond could deform the secondary structure of the oligonucleotide duplex, the cyclonucleoside could

have more favorable entropic effect during the helix formation, or some combination of the two.

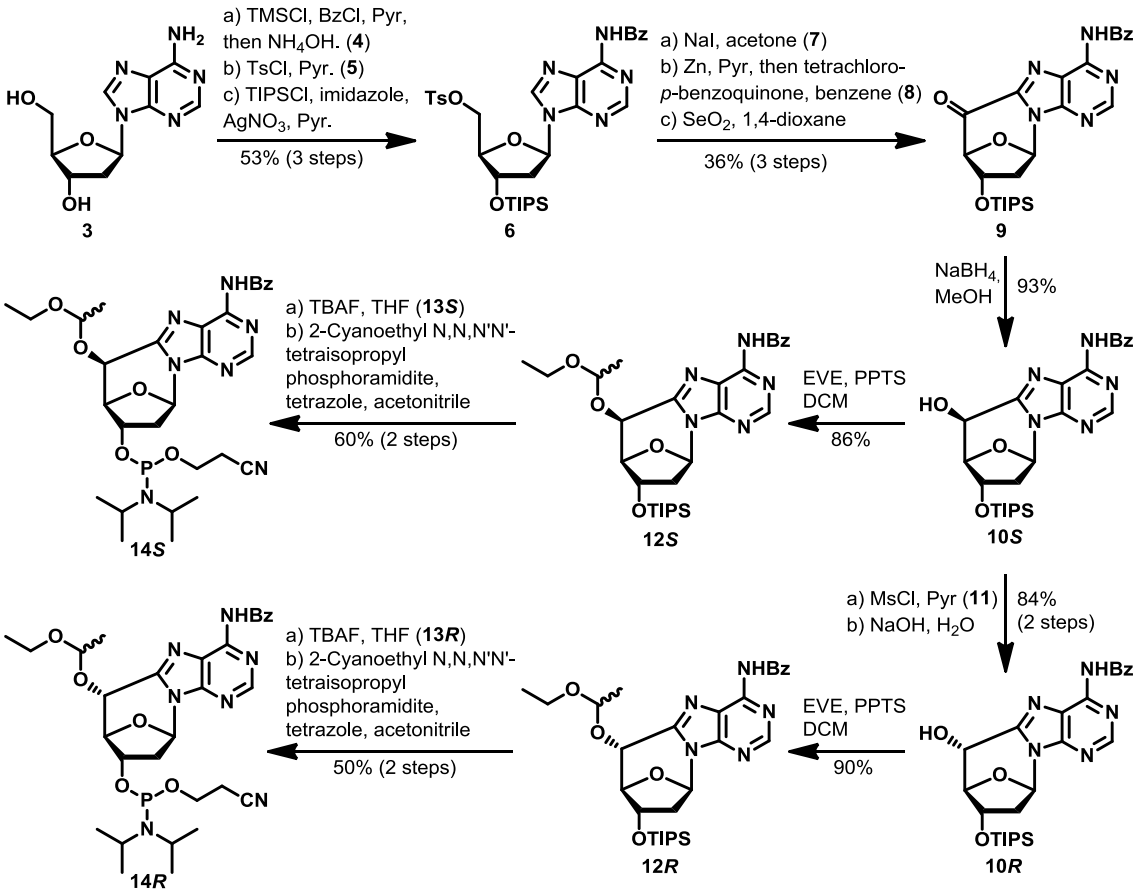
With stereochemically pure samples of these cyclic nucleosides, we hoped to address the question of whether or not this lesion results in stronger hydrogen bond interactions with its Watson-Crick pair. The restricted flexibility of the C5'-C8 covalent bond might help to pre-organize and thereby reduce entropic loss while forming a base pair. On the other hand, if these lesions destabilize the duplex formation, it will confirm that the deformation caused by the extra bond overpowers the effect of the weak A•T base pair of the TATA box, and results in TBP inhibition. Since 6,5'-cyclo-2'-deoxyuridine features a similar lesion, it was synthesized by Dr. Hongchuan Yu to aid in the study of duplex stability. Due to synthetic constraints, the deoxyuridine monomer was produced instead of thymidine in order to retain aromaticity after introducing the C6-C5' bond.

This chapter presents efficient syntheses of the R and S diastereomers of 8,5'-cyclo-2'-deoxyadenosine. Then we incorporated these interesting nucleosides into DNA to study how the cyclo-linkage affects the stability of duplex formation.

2.2. Synthesis of 8,5'-Cyclo-2'-Deoxyadenosine

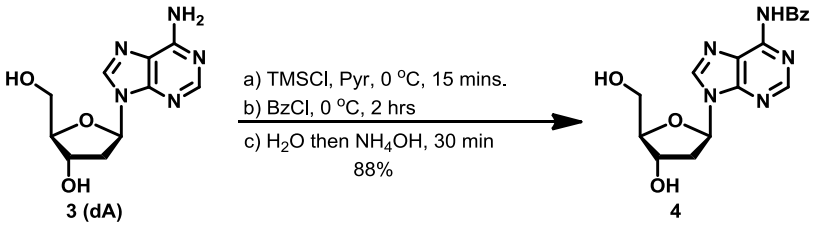
The focus of this study was to prepare both diastereomers based upon dA by introducing a second bond between the sugar and the heterocycle to

rigidify the nucleosides. We then wanted to incorporate the modified bases into the strands to investigate their behavior. These cyclonucleosides have a second bond between C5' and C8 for the 2'deoxyadenosine and C5' and C6 for 2'-deoxyuridine. The cross-links at these locations would result in glycosidic bond angles similar to those observed in double-stranded nucleic acids. The complete synthetic scheme of both *R* and *S* diastereomers of 8,5'-cyclo-2'-deoxyadenosine phosphoramidites is shown below (Scheme 2.1).



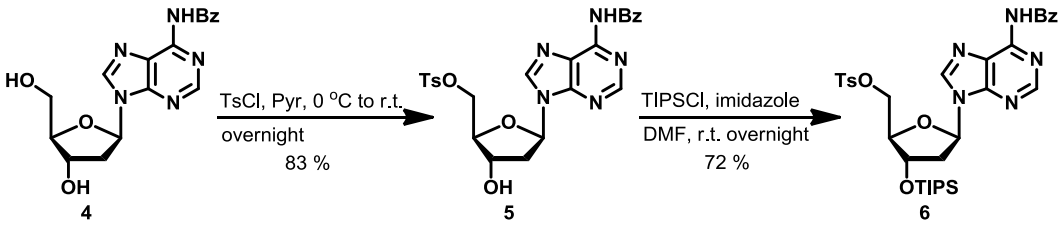
Scheme 2.1. Synthesis of the *R* and *S* diastereomers of 8,5'-cyclo-2'-deoxyadenosine phosphoramidites.

The cyclo-2'-deoxyadenosine monomer was initiated from 2'-deoxyadenosine (**3**) with the transient protection to selectively protect the N^6-amino group by benzoyl chloride (BzCl).[18] 2'-deoxyadenosine was treated with TMSCl first to protect 3'-OH and 5'-OH, then put on the benzoyl group on the N^6 position, followed by NH_4OH to remove the silyl groups. The product precipitated out from the aqueous solution after diethyl ether wash, no further purification needed. This is a well known and established step with a good yield.



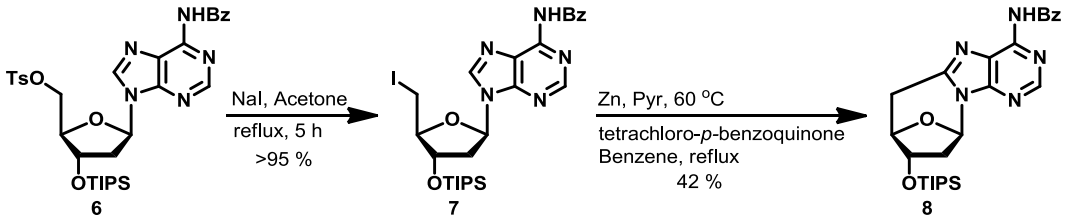
Scheme 2.2. Transient protection for 2'-deoxyadenosine.

Compound 4 was then treated with *p*-toluenesulfonyl chloride to protect the primary hydroxyl group on the 5' position, followed by protecting the 3'-OH with a triisopropylsilyl group. We established the conditions for silyl protection without using expensive silver nitrate as a catalyst. The high boiling point solvent, DMF, in this reaction could be easily removed by brine wash.



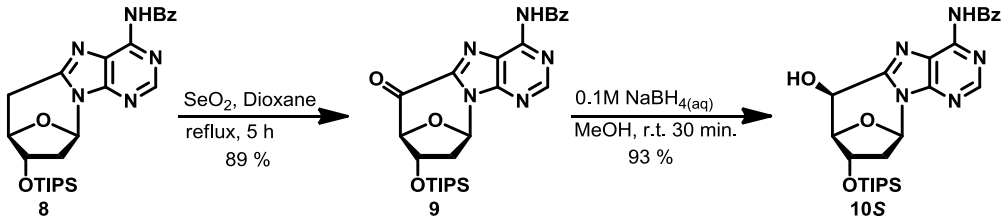
Scheme 2.3. Synthesis of compound **5** and compound **6**.

The synthesis of compound **8**, 8,5'-cyclo-2',5'-dideoxyadenosine, was carried out in three steps. Substitution of the 5'-tosyl by iodide to obtain the 5'-iodo compound **7** permitted a zinc-mediated cyclization to generate the cyclized product. After the cyclization, the product had to be treated with tetrachloro-*p*-benzoquinone to regenerate the aromaticity on the nucleobase to get the rigidified cyclo-nucleoside **8**.



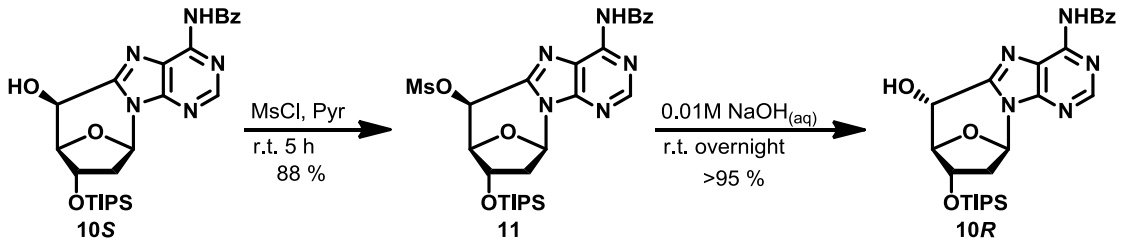
Scheme 2.4. Synthesis of compound **7** and compound **8**.

The next step was to generate the 5'-OH to afford the 3' to 5' linkage as found in nature. Compound **8** was oxidized at the C5' position with SeO_2 in dioxane to yield the ketone compound **9**, which upon reduction with $NaBH_4$ formed exclusively the *S*-diastereomer (compound **10*S***). We proposed this was due to the steric hinderance of the cyclized structure that only allowed $NaBH_4$ to react from the less hindered position.



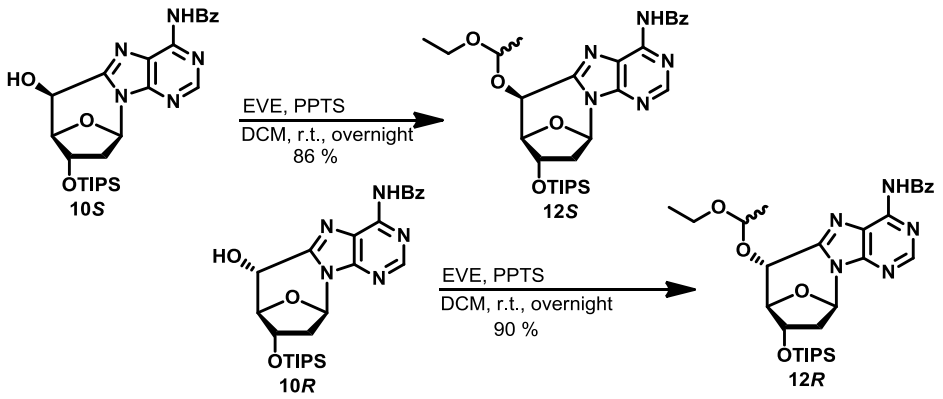
Scheme 2.5. Synthesis of compound **9** and compound **10*S***.

To obtain the *R* diastereomer, **10*S*** was first treated with mesyl chloride to put on a good leaving group on C5' position and then inverted with sodium hydroxide to obtain **10*R*** in high yields. Here, it was important to carefully control the concentration of the sodium hydroxide to avoid the de-benzoylation on the N^6 position.



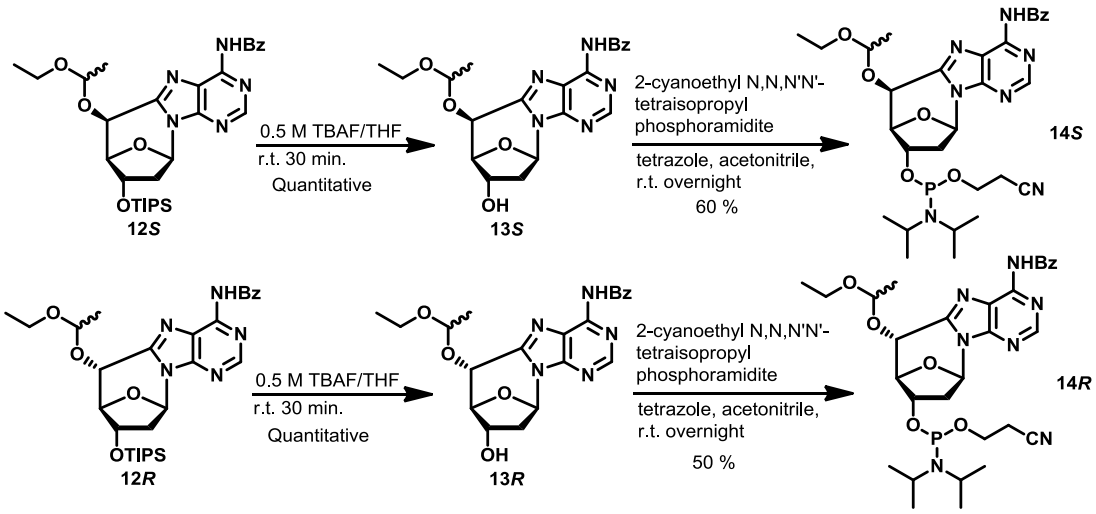
Scheme 2.6. Synthesis of compound **11** and compound **10*R***.

From this point forward the two diastereomers were processed in parallel. The newly formed 5' hydroxyls of both the *R* and *S* diastereomers were unable to be protected with the standard 4,4'-dimethoxytrityl chloride (DMTrCl) protecting group due to the steric hindrance in these rigidified nucleosides. Instead we opted to use ethyl vinyl ether to form 1-ethylethoxy protected **12*S*** and **12*R***. The 1-ethylethoxy is an acid labile protecting group which has similar deprotecting conditions as DMTrCl on the DNA synthesizer. The ethyl vinyl ether protection resulted in two additional diastereomers at the newly formed methyl position, which co-migrated during silica gel chromatography.



Scheme 2.7. Synthesis of compound **12*S*** and **12*R***.

The final steps involved removal of the 3' silyl protecting group by fluoride ion (tetra-*n*-butylammonium fluoride, TBAF) and conversion to the 3′-phosphoramidites (**14*S***, **14*R***). The presence of the two phosphorus diastereomers and the two 1-ethylethoxy protected diastereomers could be confirmed by the presence of four phosphorus resonances in the final *R* and *S*. The overall yield for the 10-step synthesis of the *S* diastereomer is 9.0 %, and for the 12-step synthesis of the *R* diastereomer is 7.5 %; it is around 80 % yield in terms of average stepwise for both diastereomers.



Scheme 2.8. Synthesis of compound **13*S***, **13*R*** and compound **14*S***, **14*R***.

The synthesis of the 3′-phosphoramidites for the cyclo-deoxyuridine (cyclo-dU) derivatives were completed by Dr. Hongchuan Yu using an established protocol[19] with the exception of the chlorination reaction on the C5 position, which was more efficiently accomplished using CAN (cerium (IV) ammonium nitrate)/LiCl in 88% yield. The details of that synthesis can be found in his thesis.[20]

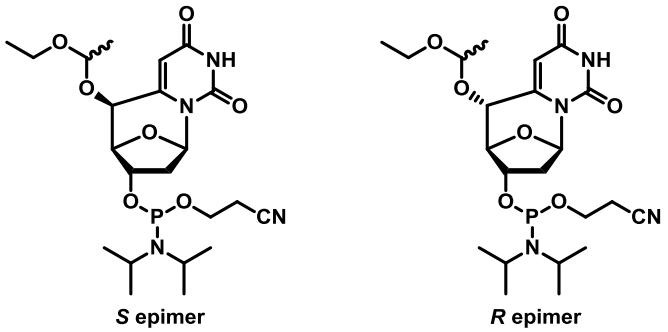


Figure 2.2. The *R* and *S* diastereomers of 6,5'-cyclo-2'-deoxyuridine phosphoramidites.

2.3. Crystal Structures of nucleosides and Thermal Denaturation Studies of the Sequences Containing Modified Nucleosides

The four cyclonucleosides, were incorporated into 12-mer or 14-mer sequences containing a core d(AAAA)/d(TTTT) sequence. In the purified oligos, the presence of the cyclonucleosides had no effect on the activity of nuclease enzymes, allowing us to digest the oligos and confirm the incorporation by HPLC.

After synthesizing these compounds and obtaining crystal structural information (Figure 2.3), it became clear that these nucleosides would not enhance duplex stability. The reasoning behind this is that the C5'-C8 (dA) and C5'-C6 (dU) bonds "pulls" the heterocycle away from the center of the DNA double helix preventing it from forming a Watson-Crick base pair.

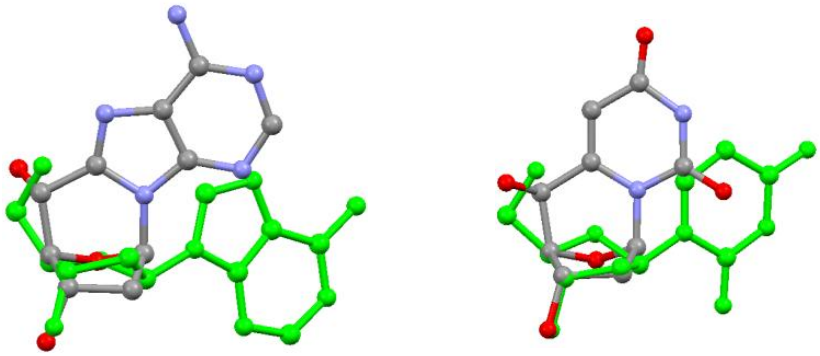


Figure 2.3. The crystal structures of **10*R***, with protecting groups hidden, and deprotected ***S*** – diastereomer for cyclo-dU overlayed with 2'-dA and 2'-dU (green structures).[21, 22]

Torsion angles can help to explain the nature of the duplexes containing rigid nucleosides. By introducing the bond between C8 and C5' in cyclo-2'-deoxyadenosine and C6 and C5' in cyclo-2'-deoxyuridine the glycosidic torsion angles, χ, became fixed at -146.7° and -152.1° respectively. The dA-dT core in the Dickerson Dodecamer has χ values ranging from near -80° to near -115° while the corresponding RNA sequence exhibits χ values around -160°. These values suggest that the fused nucleosides might be better accommodated in an RNA A-like duplex; however even these RNA/DNA heteroduplexes lack stability (see Table 2.1, entry 16 and 17).

Table 2.1. Thermal denaturation experiments of the native sequences (entry 1, 6, and 15) and modified sequences. Entry 15 – 17 are DNA/RNA heteroduplexes pairs. Cyclonuclosides are denoted by underline and bold text. Isotherms are attached in Chapter 2.7, Figure 2.8 – Figure 2.21.

Entry	Sequence	T_M (°C)
1	5'-d(C C G G A A A A C G C C)/5'-d(G G C G T T T T C C G G)	49
2	5'-d(C C G G **A*S***[1] A **A*S*** A C G C C)/5'-d(G G C G T T T T C C G G)	40
3	5'-d(C C G G **A*R***[2] A **A*R*** A C G C C)/5'-d(G G C G T T T T C C G G)	36
4	5'-d(C C G G A A A A C G C C)/5'-d(G G C G **U*S***[3] T **U*S*** T C C G G)	28
5	5'-d(C C G G A A A A C G C C)/5'-d(G G C G **U*R***[4] T **U*R*** T C C G G)	36
6	5'-d(C C G G A A A A A A C G C C)/5'-d(G G C G T T T T T T C C G G)	51
7	5'-d(C C G G A **A*S*A*S*A*S*A*S*** A C G C C)/5'-d(G G C G T T T T T T C C G G)	17
8	5'-d(C C G G A **A*R*A*R*A*R*A*R*** A C G C C)/5'-d(G G C G T T T T T T C C G G)	28
9	5'-d(C C G G A A A A A A C G C C)/5'-d(G G C G T **Us Us Us Us** T C C G G)	28
10	5'-d(C C G G A A A A A A C G C C)/5'-d(G G C G T **U*R*U*R*U*R*U*R*** T C C G G)	16
11	5'-d(C C G G A **A*S*A*S*A*S*A*S*** A C G C C)/5'-d(G G C G T **U*S*U*S*U*S*U*S*** T C C G G)	14
12	5'-d(C C G G A **A*S*A*S*A*S*A*S*** A C G C C)/5'-d(G G C G T **U*R*U*R*U*R*U*R*** T C C G G)	*
13	5'-d(C C G G A **A*R*A*R*A*R*A*R*** A C G C C)/5'-d(G G C G T **U*S*U*S*U*S*U*S*** T C C G G)	*
14	5'-d(C C G G A **A*R*A*R*A*R*A*R*** A C G C C)/5'-d(G G C G T **U*R*U*R*U*R*U*R*** T C C G G)	*
15	5'-d(C C G G A A A A C G C C)/5'-G G C G U U U U C C G G	43
16	5'-d(C C G G **A*S*** A **A*S*** A C G C C)/5'-G G C G U U U U C C G G	29
17	5'-d(C C G G **A*R*** A **A*R*** A C G C C)/5'-G G C G U U U U C C G G	30

[1] =8,5'(*S*)-cyclo-2'-deoxyadenosine; [2] =8,5'(*R*)-cyclo-2'-deoxyadenosine;
[3] =6,5'(*S*)-cyclo-2'-deoxyuridine; [4] =6,5'(*R*)-cyclo-2'-deoxyuridine
*=Duplex formation was not observable.

The torsion angle for the C4'-C5' bond (γ) is also problematic. The γ torsion angle prefers a value close to 70° for RNA and a value near 40° for DNA. In these conformations the 5'-O is positioned above the sugar and angled toward the heterocycle. In the case of the modified nucleosides, that position is occupied by the C8-C5' bond in cyclo-2'-deoxyadenosine or the C6-C5' bond in cyclo-2'-deoxyuridine. The fused 5'-OH is left in a position normally inhabited by the C5' hydrogens. The *S* diastereomer results in a γ torsion angle near -50° while the *R* produces a torsion angle of about -170° .

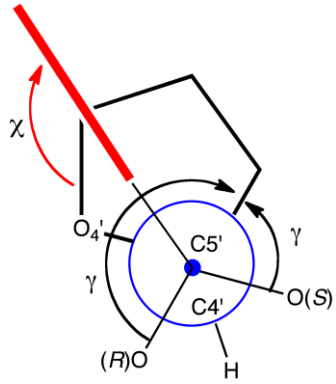


Figure 2.4. An overhead view of a cyclonucleoside highlighting the relevant torsion angles.

The results of thermal denaturation studies fit in with our hypothesis. In table 2.1, we firstly incorporated two modified bases into the middle of the sequence alternatively (entry 2 – 5), and found destabilization. Then we incorporated four modified bases into the middle of the sequence consecutively (entry 7 – 14) to investigate if there was cooperative

stabilization between these cyclo-nucleic acids. The results showed that these consecutive bases distorted the structure more and in some modified base – modified base pairs (entry 12 – 14) the T_M could not be found.

2.4. Conclusions

Our effort towards preparing pure diastereomers of nucleosides of cyclo-2'-deoxyadenosine and cyclo-2'-deoxyuridine in appreciable yields has been successful. We incorporated these nucleosides into DNA sequences and found that the T_M in most of the cases is reduced. These results are in accordance with theoretical studies performed in the labs of John Miller.[12] Our data proves that cyclonucleoside lesions destabilize duplex formation. This supports the hypothesis that duplex deformation caused by this lesion overpowers the effects of the weaker TATA box base pairing, thus inhibiting TBP binding. Our crystal structures show that the nucleobase is "pulled" back from its Watson and Crick position, explaining why these molecules form weaker hydrogen bonds as a base pair. These compounds are of continued interest due to the fact that they are resistant towards enzymatic repair and may be valid drug targets.[23]

2.5. Experimental

2.5.1. General Information

Reagents were purchased from Sigma-Aldrich, Acros, Oakwood, Glen Research, Lancaster, MP Biomedical, Chem-Impex International, Fisher, Chem Genes and Molekula. Flash column were performed using Dynamic Adsorbents silica gel (60 Å, particle size 32-63 μm) and TLC monitoring with TLC Silica Gel with F-254 Indicator (Dynamic Adsorbents). TLCs were visualized by 260nm UV light and stained by 10% sulfuric acid. All reactions were carried out under a nitrogen atmosphere with dry solvent under anhydrous conditions unless indicated otherwise. Dry tetrahydrofuran (THF), diethyl ether (Et_2O), *N*,*N*-dimethylformamide (DMF), pyridine (pyr), acetonitrile (MeCN), and dichloromethane (DCM) were obtained by passing commercially available pre-dried, oxygen-free formulations through activated alumina columns. Dry ethanol was purchased from Sigma-Aldrich and used directly. NMR spectra were taken by Varian VNMRS400, VNMRS500, or INOVA 500 instruments and calibrated using residual undeuterated solvent (CDCl3: δ H = 7.24 ppm, δ C = 77.23 ppm, DMSO-*d*6: δ H = 2.50 ppm, δ C = 39.51 ppm). Abbreviations of multiplicities were designated as follow: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet b = broad. High-resolution mass spectra (HRMS) were recorded on a Waters LCT or JEOL AccuTOF mass spectrometer using ESI (electrospray ionization) or DART (direct analysis in real time). UV spectra were acquired by Beckman

(Fullerton, CA) DU650 spectrophotometer. HPLC was performed on Waters (Milford, MA) Delta 600 controllers with a 2487 dual wavelength detector. RNA oligomers were purchased from Integrated DNA Technologies (Coralville, IA). Polyacrylamide Gel Electrophoresis was performed by using a Hoefer Scientific (San Francisco, CA) apparatus. Gel imaging was performed on Bio-Rad (Hercules, CA) FXpro Molecular Imager with ethidium bromide staining or Kodak-K radioisotope screens. The X-ray diffraction data were collected using a Bruker Kappa Apex Duo fully automated single crystal diffractometer, duo wavelength system with sealed molybdenum tube and high brightness copper source.

2.5.2. DNA Synthesis and Purification

The native and modified DNA oligomers were prepared by Applied Biosystems 394 DNA/RNA synthesizer using standard solid phase phosphoramidite techniques and the reagents purchased from Glenn Research (Sterling, VA). The native DNA strands were performed under standard 1 μmole synthesis protocol with 2 minutes waiting time during the base coupling; the sequences with modified nucleosides incorporated were using the same protocol with 1 hour waiting time. All the native sequences were giving over 95% average stepwise yield; the yield for the syntheses with modified bases were not be able to be monitored since the protecting group on the 5'-O is not UV active.

All oligos were deprotected and cleaved off from the solid supports by 12 hours treatment with concentrated ammonia at 55 °C water bath. The solutions were decanted and evaporated by speedvac, then redissolved in 5-8 mL of DI water for HPLC purification. The HPLC purification was accomplished by self-packed Oligo R3 column (0.46 cm × 10 cm), and the mobile phase buffer A: 50 mM TEAA (triethylammonium acetate, pH 7.0) 5% acetonitrile, Buffer B: 50 mM TEAA (pH 7.0) 70% acetonitrile. The program of the running gradient is showing below in Table 2.2.

Table 2.2. Gradient for purifying oligos on Oligo-R3 column.

Time (min)	Flow Rate (mL/min)	Buffer A (%)	Buffer B (%)
0	3 (Initial)	100	0
25	3	50	50
30	3	0	100
40	3	0	100
41	3	100	0
61	3	100	0

The next step was removing the trityl protecting group on the 5'-end by treating the oligos in 80% acetic acid for 30 minutes then desalted through Sephade G-10 column and lyophilized after removing the acetic acid by rotary evaporator. The oligonucleotides were redissolved in D.I. water and determined the concentration by UV absorbance. Further purification of the oligos was accomplished by 20% denaturing acrylamide gel.

2.5.3. Enzymatic Digestion

In order to prove the incorporation of the modified nucleotides was successful, we firstly synthesized the Dickerson dodecamer[24], d(CGC GAA TTC GCG), with two *R*- diastereomer of cyclo-dA. Then we performed the enzymatic digestion experiments to clarify the composition of the nucleotides.

A mixture of 3 μL of alkaline phosphatase, 3 μL of phosphodiesterase I, 4 μL of nuclease P1, 10 μL of 10X alkaline phoaphatase buffer, 10 μL of 1.0 M MgCl2, and 70 μL of water was added to a eppendorf tube containing 2 nmoles of the DNA oligomer. The mixture was left at 37 ^{o}C overnight, then cooled to room temperature. The digested nucleoside mixture was analyzed by reverse-phase analytical HPLC with a self-packed C18 column (0.46 cm × 25 cm, 5 μm 100 Å resin). Mobile phase buffer A: 100 mM TEAA buffer (pH 7.0). Mobile phase buffer B: acetonitrile. Flow rate: 1 mL/min. Gradient of the elution: 0-25% in 30 min and 25-100% in 10 min. The mixtures were monitored by UV at 260 nm. HPLC traces were attached in Chapter 2.7, Figure 2.7. In modified sequence, there is no dA on the strand but a new peak with the integral area represented 2 bases. Assuming the cyclo-dA possesses a similar extinction coefficient as the natural base, the result represented the phosphoramidite of the cyclo-dA is working perfectly on the DNA synthesizer (Table 2.3).

Table 2.3. Enzymatic digestion results.

Control Sequence 5'-d(CGC GAA TTC GCG)-3'

Nucleoside	Area	ε at 260 nm	Corrected Area	Normalized Area	total
dC	79	7300	0.011	1.000	2.0
dG	134	11700	0.011	1.058	2.1
dT	50	8800	0.006	0.525	1.1
dA	77	15400	0.005	0.462	0.9
Cyclo-dA	0	15400	0	0	0

Modified Sequence 5'-d(CGC G$A_R A_R$ TTC GCG)-3'

Nucleoside	Area	ε at 260 nm	Corrected Area	Normalized Area	total
dC	110	7300	0.015	1.000	2.0
dG	171	11700	0.015	0.970	1.9
dT	89	8800	0.010	0.671	1.3
dA	0	15400	0	0	0
Cyclo-dA	126	15400*	0.008	0.543	1.1

*Using the same extinction coefficient as dA.

2.5.4. Thermal Denaturation Studies

All UV melting experiments were carried out in 20 mM sodium phosphate buffer at pH 7.0 with 100 mM or 1 M sodium chloride using AVIV Biomedical Inc. (Lakewood, NJ) spectrophotometer model 14DS UV-Vis with a temperature controller and five-sample holder. Samples (0.5 μM for each complementary sequence, 1.0 μM for total) were heated under 95 °C water bath for 5 min and the water bath allowed to cool to room temperature then cooled to 4 °C in the coldroom. Samples were place in a 1 cm path length quartz cell with stopper and the denaturation experiments were performed between 4 °C and 95 °C. Data was collected every degree under 260 nm. The

T_m values were determined by utilizing the first derivative from the melting profile with Origin and Microsoft Excel. The T_M isotherms are attached in chapter 2.7, figure 2.7 to figure 2.20.

2.5.5. Thermal Circular Dichroism studies

All Circular Dichroism experiments were carried out in 20 mM sodium phosphate buffer at pH 7.0 with 100 mM or 1 M sodium chloride using AVIV Biomedical Inc. (Lakewood, NJ) Circular Dichroism spectrophotometer model 420 with a temperature controller. Samples preparation: single-stranded samples (concentration of oligomers: 8 μM) were performed the experiment directly. Double-stranded samples (total concentration of oligomers: 8 μM) were heated under 95 ^{o}C water bath for 5 min and the water bath allowed to cool to roomtemperature then cooled to 4 ^{o}C in the coldroom. Samples were place in a 1 cm path length quartz cell with stopper and the denaturation experiments were performed at 25 ^{o}C. CD spectra are attached in Chapter 2.7, Figure 2.21 and Figure 2.22.

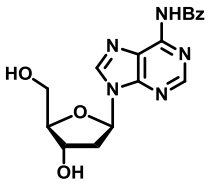
2.5.6. Crystallization of Compound 10*R*

50 mg of compound **10*R*** was firstly dissolved in 0.25 mL of dichloromethane in a 6 X 50 mm disposable culture tube at room temperature. The culture tube was then sitting in a capped 20 mL sample

vial which was filled with 2 mL of hexane. The whole assembly was kept undisturbed to obtain the crystals. The X-ray data and crystal structure were collected and solved by Dr. Bo Li, Director of X-ray Crystallography Center at Boston College. The Structural data and information are attached in Chapter 2.7, Figure 2.5 - 2.6 and Table 2.4 – 2.10.

2.5.6. Synthesis

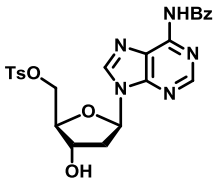
Compound 4: N-benoyl-2'-deoxyadenosine



2'-deoxyadenosine (10.0 g, 40.0 mmol) was dissolved in dry pyridine (400 mL) and cooled by ice bath. Chlorotrimethylsilane (26.4 mL, 208 mmol) was added to the solution and the mixture was kept on ice for 30 minutes. Benzoyl chloride (30.0 mL, 260 mmol) was added at 0 °C and the reaction mixture warmed to ambient temperature and stirred for 2 hours. The flask was cooled back to 0 °C and the reaction quenched with water (80 mL). After 15 min concentrated aqueous ammonium hydroxide (80 mL) was added and the mixture was kept on ice for another 30 minutes. Solvents were removed under reduced pressure and the residue dissolved in water (300 mL) and washed twice with diethyl ether (200 mL each). Compound **4** precipitated from the aqueous solution as a white solid after sitting at 4 °C overnight (12.5 g, 88%). TLC R_f : 0.17

(DCM/MeOH 83:17). ^{1}H NMR (400 MHz; CD_3OD): δ 2.49 (m, 1H), 2.84 (m, 1H), 3.74-3.86 (m, 2H), 4.07 (d, 1H), 6.55 (t, 1H), 7.54 (t, 2H), 7.63 (t, 1H), 8.07 (d, 2H), 8.63 (s, 1H), 8.69 (m, 1H). ^{13}C NMR (100MHz; CD_3OD): δ 41.6, 63.4, 72.8, 86.8, 89.9, 125.6, 129.6, 129.9, 134.0, 135.1, 144.7, 151.3, 153.0, 168.3. ESI-MS (pos.): 356.1353 $[M+H]^+$ (HSMS calc. 356.1359).

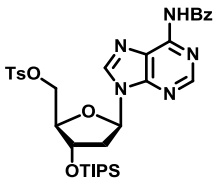
Compound 5: N-Benzoyl-5'-tosyl-2'-deoxyadenosine



Compound **4** (10.2 g, 28.7 mmol) was dissolved in dry pyridine (140 mL) and cooled to 0 °C. *p*-tosylsulfonyl chloride (8.20 g, 43.0 mmol) was added and the reaction mixture was allowed to warm to ambient temperature and stir overnight. The reaction was quenched with water at 0 °C and solvents were removed under reduced pressure. The residue was dissolved in DCM and washed with water 3 times and once with brine. The organic layer was dried over sodium sulfate and concentrated in a rotary evaporator. The crude product was purified by flash chromatography (DCM/MeOH 95:5) to yield the **5** as a white solid (12.1 g, 83%). TLC R_f: 0.32 (DCM/MeOH 90:10). ^{1}H NMR (400 MHz; $CDCl_3$): δ 2.43 (s, 3H), 2.49 (d, 1H), 2.58 (m, 1H), 2.97 (m, 1H), 4.23-4.3 (m, 3H), 4.83 (bm, 1H), 6.49 (t, 1H), 7.31 (d, 2H), 7.54 (t, 2H), 7.63 (t, 1H), 7.74 (d, 2H), 8.03 (dd, 2H), 8.18 (s, 1H), 8.75 (s, 1H), 8.98 (s, 1H). ^{13}C NMR (100 MHz; CD_3OD): δ 41.6, 63.4, 72.8, 86.8, 89.9, 125.6, 129.6, 129.6, 129.9, 129.9, 134.0, 134.2,

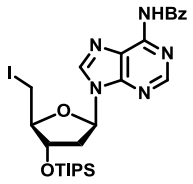
134.3, 135.1, 144.7, 146.5, 151.3, 153.0, 153.1, 168.3. ESI-MS (pos.): 510.1455 $[M+H]^+$ (HSMS calc. 510.1447).

Compound 6: N-Benzoyl-3'-triisopropylsilyl-5'-tosyl-2'-deoxyadenosine



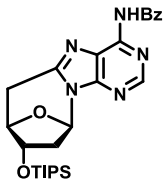
Compound **5** (5.60 g, 11.0 mmol), imidazole (2.30 g, 33.4 mmol), and silver nitrate (3.78 g, 22.2 mmol) were dissolved in dry pyridine (56 mL) and cooled to 0 ^oC. Triisopropylsilyl chloride (4.70 ml, 22.0 mmol) was added and the reaction was warmed to ambient temperature. After stirring for two days the reaction was quenched with water while on an ice bath. The solvents were removed under reduced pressure and the residue dissolved in dichloromethane and washed with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was purified by flash chromatography (DCM/MeOH 97:3) to yield **6** as a white (5.2 g, 72%). TLC R_f: 0.65 (DCM/MeOH 97:3). ^1H NMR (400 MHz; $CDCl_3$): δ 1.07-1.12 (m, 21H), 2.42 (s, 3H), 2.47-2.51 (m, 1H), 2.93 (m, 1H), 4.22 (s, 1H), 4.19-4.3 (m, 2H), 4.73 (t, 1H), 6.47 (t, 1H), 7.29 (d, 2H), 7.52 (t, 2H), 7.6 (q, 1H), 7.74 (d, 2H), 8.03 (d, 2H), 8.20 (s, 1H), 8.73 (s, 1H). ^{13}C NMR (100 MHz; $CDCl_3$): δ 11.9, 17.9, 21.6, 40.3, 68.8, 73.0, 85.0, 85.4, 123.4, 125.8, 127.9, 128.8, 129.9, 132.2, 132.8, 133.4, 141.9, 145.2, 149.4, 151.4, 152.1, 164.7. ESI-MS (pos.): 666.2762 $[M+H]^+$ (HSMS calc. 666.2782).

Compound 7: N-Benzoyl-3'-triisopropylsilyl-5'-iodo-2'-deoxyadenosine



Compound **6** (4.26 g, 6.40 mmol) and sodium iodide (3.84 g, 2.56 mmol) were dissolved in dry acetone (32 mL) and refluxed for four hours. The solid suspension filtered through Celite and the filtrate was concentrated under vacuum. The crude was purified by flash chromatography (DCM/MeOH 97.5/2.5) to yield **7** as a white foam (3.9 g, 98%) TLC R_f : 0.42 (DCM/MeOH 97.5/2.5). ^{1}H NMR (400 MHz; $CDCl_3$): δ 1.04-1.14 (m, 21H), 2.23 (s, 1H), 2.47 (m, 1H), 3.08 (m, 1H), 3.36-3.53 (m, 2H), 4.07 (m, 1H), 4.70 (t, 1H), 6.44 (t, 1H), 7,47 (t, 2H), 7.56 (t, 1H), 7.99 (d, 2H), 8.24 (s, 1H), 8.74 (s, 1H), 9.22 (s, 1H). ^{13}C NMR (100 MHz; $CDCl_3$): δ 6.3, 12.0, 18.0, 39.9, 75.4, 85.2, 86.7, 123.7, 127.8, 128.8, 132.7, 133.6, 142.2, 149.6, 151.4, 152.5, 164.6. ESI-MS (pos.): 622.1715 $[M+H]^+$ (HSMS calc. 622.1710).

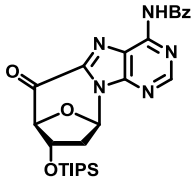
Compound 8: N-Benzoyl-3'-triisopropylsilyl-8,5'-cyclo-2',5'-dideoxyadenosine



Compound **7** (1.25 g, 2.00 mmol) and zinc powder (2.63 g, 40.2 mmol) were dissolved in dry pyridine (70 mL) and heated to 80 °C for 4 hours. The zinc powder was removed by filtration and the filtrate was concentrated under reduced pressure. The residue and tetrachloro-1,4-benzoquinone (0.490 g, 2.00 mmol) were dissolved in benzene (40 mL) and refluxed for 2 hours. The solvent was then evaporated under vacuum and the crude product was purified by flash chromatography

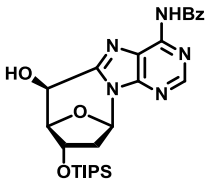
(DCM/MeOH 98:2) to yield the cyclized product as a yellow foam (0.42 g, 42%). TLC R_f: 0.32 (DCM/MeOH 97:3). ^{1}H NMR (400 MHz; $CDCl_3$): δ 0.98-1.07 (m, 21H), 2.36 (m, 1H), 2.66 (m, 1H), 3.11 (d, 1H), 3.58 (m, 1H), 4.51 (m, 1H), 4.80 (d, 1H), 6.58 (d, 1H), 7.50 (t, 2H), 7.59 (t, 1H), 7.99 (t, 2H), 8.68 (s, 1H), 8.74 (s, 1H). ^{13}C NMR (100 MHz; $CDCl_3$): δ 11.8, 17.8, 31.2, 48.1, 75.3, 83.2, 84.8, 127.8, 128.8, 132.8, 133.5, 146.8, 147.4, 148.1, 148.3, 151.8, 164.5. ESI-MS (pos.): 516.2425 $[M+Na]^+$ (HSMS calc. 516.2407).

Compound 9: N-Benzoyl-3'-triisopropylsilyl-5'-keto-8,5'-cyclo-2'-deoxyadenosine



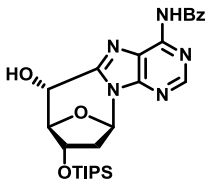
Compound **8** (0.186 g, 0.377 mmol) and selenium dioxide (90.0 mg, 0.814 mmol) were dissolved in 1,4-dioxane (50 mL) and refluxed for 1 hour. The reaction mixture was filtered through Celite and the solvent removed under vacuum. The product was purified by flash chromatography (DCM/MeOH 97:3) to yield **9** as a yellow foam (0.170 g, 89%). TLC R_f: 0.23 (DCM/MeOH 95:5). ^{1}H NMR (400 MHz; $CDCl_3$): δ 1.02-1.07 (m, 21H), 2.14 (s, 1H), 2.56 (m, 1H), 2.72 (m, 1H), 4.76 (dd, 1H), 4.96 (s, 1H), 6.83 (d, 1H), 7.52 (t, 2H), 7.61 (t, 1H), 7.97 (d, 2H), 8.92 (s, 1H), 9.19 (s, 1H). ^{13}C NMR (100 MHz; $CDCl_3$): δ 11.8, 17.8, 44.8, 72.7, 85.9, 92.3, 122.9, 127.9, 129.0, 133.2, 140.8, 147.9, 148.3, 152.3, 156.1, 164.3, 184.6. ESI-MS (pos.): 530.2186 $[M+Na]^+$ (HSMS calc. 530.2200).

Compound 10*S*: N-Benzoyl-3'-triisopropylsilyl-8,5'(*S*)-cyclo-2'-deoxyadenosine



Compound **9** (0.454 g, 0.890 mmol) was dissolved in methanol (27 mL) and cooled to 0 °C, 0.1 M sodium borohydride aqueous solution (5 mL) was added and the stirring reaction was allowed to warm to ambient temperature. After 1.5 hours the reaction was neutralized with 1N hydrochloric acid and the solvents were removed under vacuum. The crude product was purified by flash chromatography (DCM/MeOH 95:5) to yield **10*S*** as a white powder (0.425 g, 93%). TLC R_f: 0.22 (DCM/MeOH 97:3). ^{1}H NMR (400 MHz; $CDCl_3$): δ 0.95-1.03 (m, 21H), 2.31 (m, 1H), 2.55 (s, 1H), 4.56 (d, 1H), 4.88 (q, 1H), 5.03 (s, 1H), 5.46 (d, 1H), 6.50 (d, 1H), 7.50 (t, 2H), 7.59 (t, 1H), 7.97 (d, 2H), 8.69 (s, 1H), 8.93 (s, 1H). ^{13}C NMR (100 MHz; $CDCl_3$): δ 11.8, 17.9, 47.0, 64.8, 69.3, 85.5, 86.4, 122.6, 127.8, 128.9, 132.9, 133.4, 148.6, 149.9, 151.5, 152.5, 164.8. ESI-MS (pos.): 510.2534 $[M+H]^+$ (HSMS calc. 510.2537).

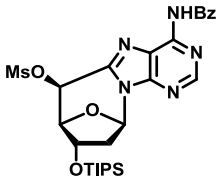
Compound 10*R*: N-Benzoyl-3'-triisopropylsilyl-8,5'(*R*)-cyclo-2'-deoxyadenosine



Compound **11** (0.370 g, 0.630 mmol) was dissolved in a dimethylformamide/water mixture (50:17.6 mL) and 0.1 M sodium hydroxide (7.55 mL) was added over 5 hours and after complete addition stirred for an additional hour. The reaction mixture was neutralized with 1 N hydrochloric acid and the solvents were removed under

vacuum. The crude product was purified by flash chromatography (DCM/MeOH 97:3) to yield **10*R*** as a white powder (0.26 g, 80%). TLC R_f: 0.17 (dichloromethane/methanol = 30/1). ^{1}H NMR (400 MHz; $CDCl_3$): δ 1.01-1.08 (m, 21H), 2.35 (m, 1H), 2.54 (m, 1H), 4.42 (q, 1H), 4.67 (s, 1H), 4.94 (d, 1H), 6.19 (s, 1H), 6.61 (d, 1H), 7.53 (dt, 2H), 7.59 (dt, 1H), 8.07 (dd, 2H), 8.75 (s, 1H), 9.32 (s, 1H). ^{13}C NMR (100 MHz; $CDCl_3$): δ 11.8, 17.9, 45.9, 65.8, 71.7, 85.2, 89.0, 121.6, 128.2, 128.8, 132.9, 133.5, 148.0, 149.1, 149.5, 153.1, 165.0. ESI-MS (pos.): 532.2351 $[M+Na]^+$ (HSMS calc. 532.2356).

Compound 11: N-Benzoyl-3'-triisopropylsilyl-5'-mesyl-8,5'(*S*)-cyclo-2'-deoxyadenosine

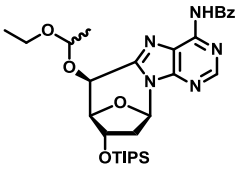


Compound **10*S*** (0.570 g, 0.570 mmol) was dissolved in pyridine (5.7 mL) and methanesulfonyl chloride (110 µL, 1.43 mmol) was added at 0 °C. The reaction was slowly warmed ambient temperature and stirred overnight before being quenched with water. The solvents were removed under vacuum and the residue was dissolved in DCM and washed with water 3 times and once with brine. After removing the DCM under reduced pressure, the crude product was purified by flash chromatography (DCM/MeOH 97.5:2.5) to yield the mesylated product as a white powder (0.31 g, 91%). TLC R_f: 0.38 (DCM/MeOH 97.5:2.5). ^{1}H NMR (400 MHz; $CDCl_3$): δ 1.01-1.10 (m, 21H), 2.40 (m, 1H), 2.66 (m, 1H), 3.57 (s, 3H), 4.89-4.93 (m, 2H), 6.12 (d, 1H), 6.55 (d, 1H), 7.52 (t, 2H), 7.61 (t,

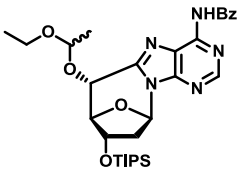
1H), 7.94 (d, 2H), 8.68 (s, 1H), 8.73 (s, 1H). ^{13}C NMR (100 MHz; $CDCl_3$): δ 11.8, 17.9, 40.0, 46.7, 70.1, 70.4, 85.1, 85.5, 122.8, 127.8, 129.0, 133.0, 133.4, 145.2, 149.2, 150.0, 152.9, 164.5. ESI-MS (pos.): 610.2136 $[M+Na]^+$ (HSMS calc. 610.2132).

Compound 12*S* and 12*R*

N-Benzoyl-3'-triisopropylsilyl-5'-(1-ethoxyethyl)-8,5'(*S*)-cyclo-2'-deoxyadenosine and N-Benzoyl-3'-triisopropylsilyl-5'-(1-ethoxyethyl)-8,5'(*R*)-cyclo-2'-deoxyadenosine



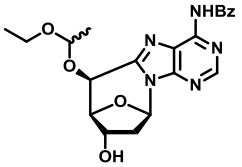
Compound **10*S*** (0.150 g, 0.300 mmol) and pyridinium-*p*-toluenesulfonate (75.0 mg, 0.300 mmole) were dissolved in dichloromethane (3 mL) followed by the addition of ethyl vinyl ether (0.175 mL, 1.80 mmol) and stirred overnight at ambient temperature. The solvents were removed under vacuum and the crude product was purified by flash chromatography (DCM/MeOH/Et_3N 96:3:1) to yield **12*S*** as a yellow foam (0.12 g, 86%). TLC R_f: 0.3 (DCM/MeOH 97:3). ^{1}H NMR (400 MHz; $CDCl_3$): δ 1.04-1.26 (m, 21H), 1.25-1.30 (bm, 3H), 1.50 (dd, 3H, J = 40.2, 5.4 Hz), 2.34 (m, 1H), 2.60 (m, 1H), 3.62-3.88 (bm, 2H), 4.74 (dd, 1H, J = 18.0, 6.4 Hz), 4.95 (m, 1H), 5.21-5.56 (dq, 1H, J = 132.4, 5.2 Hz), 5.33 (dd, 1H, J = 14.4, 6.0 Hz), 6.52 (q, 1H), 7.52 (m, 2H), 7.61 (dt, 1H), 7.97 (m, 2H), 8.74 (s, 1H). ESI-MS (pos.): 582.3085 $[M+H]^+$ (HSMS calc. 582.3112).



12*R* was prepared using the same procedure as **12*S*** with the exception of chromatographic conditions (DCM/MeOH/Et_3N 96.4/2.5/1) to yield the product as a yellow foam (0.12 g, 90%). TLC R_f : 0.45 (DCM/MeOH 97:3). ^1H NMR (400 MHz; $CDCl_3$): δ 1.01-1.08 (m, 21H), 1.21-1.28 (bm, 3H), 1.40 (dd, 3H, J = 60.8, 5.6 Hz), 2.32 (m, 1H), 2.48 (m, 1H), 3.59-3.77 (bm, 2H), 4.38 (q, 1H), 4.82 (d, 1H), 4.87 (s, 1H), 5.14-5.44 (dq, 1H, J = 109.6, 5.6 Hz), 6.60 (d, 1H), 7.52 (dt, 2H), 7.60 (dt, 1H), 7.97 (dd, 2H), 8.76 (s, 1H), 8.86 (d, 1H). ESI-MS (pos.): 604.2913 $[M+Na]^+$ (HSMS calc. 604.2931).

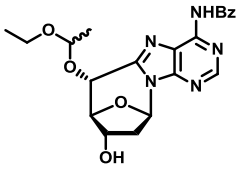
Compound 13*S* and 13*R*

N-Benzoyl -5'-(1-ethoxyethyl)-8,5'(*S*)-cyclo-2'-deoxyadenosine (13*S*) and N-Benzoyl-5'-(1-ethoxyethyl)-8,5'(*R*)-cyclo-2'-deoxyadenosine (13*R*)



Tetra-*n*-butylammonium fluoride (1.0 M solution in THF, 0.50 mL) was added to a solution of **12*S*** (0.132 g, 0.227 mmol) in THF (1.5 mL) at ambient temperature and stirred for one hour. The solvent was removed under reduced pressure and the crude product was purified by flash chromatography (DCM/MeOH/Et_3N 96.5:2.5:1) to yield the **13*S*** as a white powder in quantitative yield (95 mg). TLC R_f : 0.22 (DCM/MeOH 97:3). ^1H NMR (400 MHz; $CDCl_3$): δ 1.07-1.27 (dt, 3H), 1.25-1.48 (dd, 3H), 2.32 (m, 1H), 2.59-2.66 (bm, 3H), 3.62 (m, 1H), 3.81 (m, 1H), 4.84 (d, 1H), 4.93 (q, 1H), 5.28 (d, 1H), 5.33 (q, 1H), 6.50 (d, 1H), 7.52

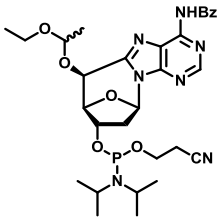
(dt, 2H), 7.60 (dt, 1H), 7.97 (dd, 2H), 8.74 (s, 1H). ESI-MS (pos.): 426.1780 $[M+H]^+$ (HSMS calc. 426.1777).



13*R* was prepared using the same procedure as the *S* diastereomer to yield the product as a white powder in quantitative yield (92 mg). TLC R_f : 0.17 (DCM/MeOH 97.5:2.5). ^1H NMR (400 MHz; $CDCl_3$): δ 1.19-1.30 (bm, 3H), 1.41 (dd, 3H, *J* = 47.6, 5.6 Hz), 2.28 (m, 1H), 2.49 (m, 1H), 2.96 (q, 2H), 3.59-3.78 (bm, 2H), 4.42 (m, 1H), 4.86 (s, 1H), 4.87 (s, 1H), 5.14-5.36 (dq, 1H, *J* = 82.8, 5.6 Hz), 6.59 (d, 1H), 7.50 (dt, 2H), 7.59 (t, 1H), 7.97 (d, 2H), 8.73 (s, 1H). ESI-MS (pos.): 426.1772 $[M+H]^+$ (HSMS calc. 426.1777).

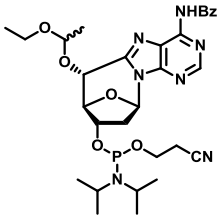
Compound 14*S* and 14*R*

N-Benzoyl-3'-(2-Cyanoethyl N,N-diisopropyl phosphorodiamidite)-5'-(1-ethoxyethyl)-8,5'(*S*)-cyclo-2'-deoxyadenosine (14*S*) and N-Benzoyl-3'-(2-Cyanoethyl N,N-diisopropyl phosphorodiamidite) -5'-(1-ethoxyethyl)-8,5'(*R*)-cyclo-2'-deoxyadenosine (14*R*)



Compound **13*S*** (66.0 mg, 0.155 mmol) was dissolved in acetonitrile (6.2 mL). 1H-tetrazole in acetonitrile (0.45 M solution in MeCN, 0.21 mL) was added followed by 2-Cyanoethyl N,N,N',N'-tetraisopropyl phosphorodiamidite (61.0 µL, 0.186 mmol) and the mixture was stirred overnight at ambient temperature. The

reaction was quenched with 5% trietheylamine in methanol and the solvents were removed under vacuum. The residue was dissolved in DCM and washed with saturated sodium bicarbonate three times and once with brine. The organic layer was dried over sodium sulfate. After removing the sodium sulfate, hexanes were added to effect precipitation. The precipitate was collected by filtration and washed with hexanes several times to yield **14*S*** as a white powder (60 mg, 60%). TLC R_f: 0.39 (DCM/MeOH 95:5). ^{31}P NMR (162 MHz; $CDCl_3$): δ 146.7, 146.8, 147.0, 147.2. ESI-MS (pos.): 648.2671 $[M+Na]^+$ (HSMS calc. 648.2675).



Compound **14*R*** was obtained by a similar procedure as **14*S*** to yield a white powder (22 mg, 50%). TLC R_f: 0.42 (DCM/MeOH 95:5). ^{31}P NMR (162 MHz; $CDCl_3$): δ 147.4, 147.6, 147.7, 147.9. ESI-MS (pos.): 648.2665 $[M+Na]^+$ (HSMS calc. 648.2675).

2.6. References

1. B. Halliwell and J. M. C. Gutteridge, *Free radicals in biology and medicine*, Oxford University Press, Oxford ; New York, 2007.
2. E. C. Friedberg, *DNA repair and mutagenesis*, ASM Press, Washington, D.C., 2006.
3. C. V. Sonntag, *Free-radical-induced DNA damage and its repair : a chemical perspective*, Springer, Berlin, 2006.
4. M. Dizdaroglu, *Mutation Research*, **1992**, *275*, 331-342.
5. A. P. Breen and J. A. Murphy, *Free Radical Biology and Medicine*, **1995**, *18*, 1033-1077.
6. M. D. Evans, M. Dizdaroglu and M. S. Cooke, *Mutation Research-Reviews in Mutation Research*, **2004**, ***567***, 1-61.
7. P. Jaruga and M. Dizdaroglu, *DNA Repair*, **2008**, *7*, 1413-1425.
8. C. Marietta, H. Gulam and P. J. Brooks, *DNA Repair*, **2002**, *1*, 967-975.
9. B. Pullman and A. Pullman, *Nature*, **1961**, *189*, 725-727.
10. T. P. Haromy, J. Raleigh and M. Sundaralingam, *Biochemistry*, **1980**, *19*, 1718-1722.
11. G. I. Birnbaum, M. Cygler, L. Dudycz, R. Stolarski and D. Shugar, *Biochemistry*, **1981**, *20*, 3294-3301.
12. K. Miaskiewicz, J. H. Miller and A. F. Fuciarelli, *Nucleic Acids Research*, **1995**, *23*, 515-521.

13. J. A. Theruvathu, P. Jaruga, M. Dizdaroglu and P. J. Brooks, *Mechanisms of Ageing and Development*, **2007**, *128*, 494-502.

14. M. L. Dirksen, W. F. Blakely, E. Holwitt and M. Dizdaroglu, *International Journal of Radiation Biology*, **1988**, *54*, 195-204.

15. P. J. Brooks, D. S. Wise, D. A. Berry, J. V. Kosmoski, M. J. Smerdon, R. L. Somers, H. Mackie, A. Y. Spoonde, E. J. Ackerman, K. Coleman, R. E. Tarone and J. H. Robbins, *Journal of Biological Chemistry*, **2000**, *275*, 22355-22362.

16. I. Kuraoka, C. Bender, A. Romieu, J. Cadet, R. D. Wood and T. Lindahl, *Proceedings of the National Academy of Sciences of the United States of America*, **2000**, *97*, 3832-3837.

17. Z. S. Juo, T. K. Chiu, P. M. Leiberman, I. Baikalov, A. J. Berk and R. E. Dickerson, *Journal of Molecular Biology*, **1996**, *261*, 239-254.

18. M.J. Gait, *Oligonucleotide Synthesis: A Practical Approach*. Oxford University Press: New York, 1984.

19. Y. Suzuki, A. Matsuda and T. Ueda, *Chemical & Pharmaceutical Bulletin*, **1987**, *35*, 1085-1092.

20. H. Yu, Novel Cyclo Deoxynucleosides: Synthesis and Evaluation. Ph.D. Thesis, Boston College, Chestnut Hill, MA, May 2012.

21. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proceedings of the National Academy of Sciences*

of the United States of America-Biological Sciences, **1981**, *78*, 2179-2183.

22. B. L. Partridge and S. A. Salisbury, *NDB ID: BDL075,* **1996**.

23. A. Romieu, D. Gasparutto, D. Molko and J. Cadet, *Journal of Organic Chemistry*, **1998**, *63*, 5245-5249.

24. R. Wing, H. Drew, T. Takano, C. Broka, S. Tanaka, K. Itakura, R. E. Dickerson, *Nature*, **1980**, *287*, 755-758.

2.7. Spectral and Crystal Data

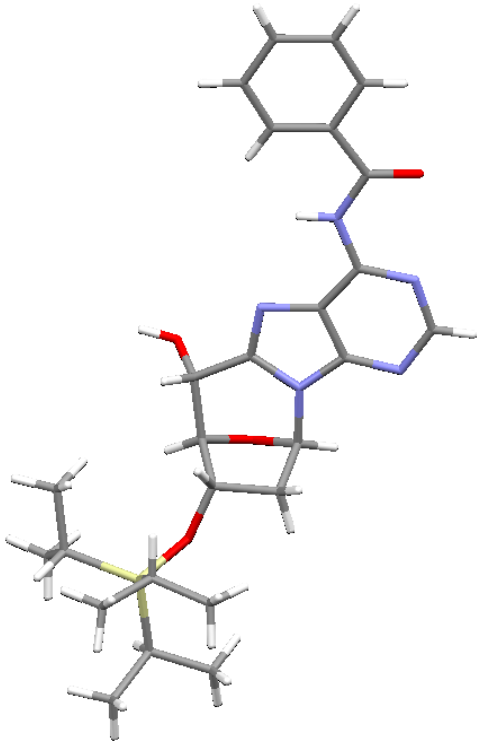


Figure 2.5. Crystal Structure of compound **12*R*** used to confirm the stereochemistry of the cyclo-2'-deoxyadenosine diastereomers.

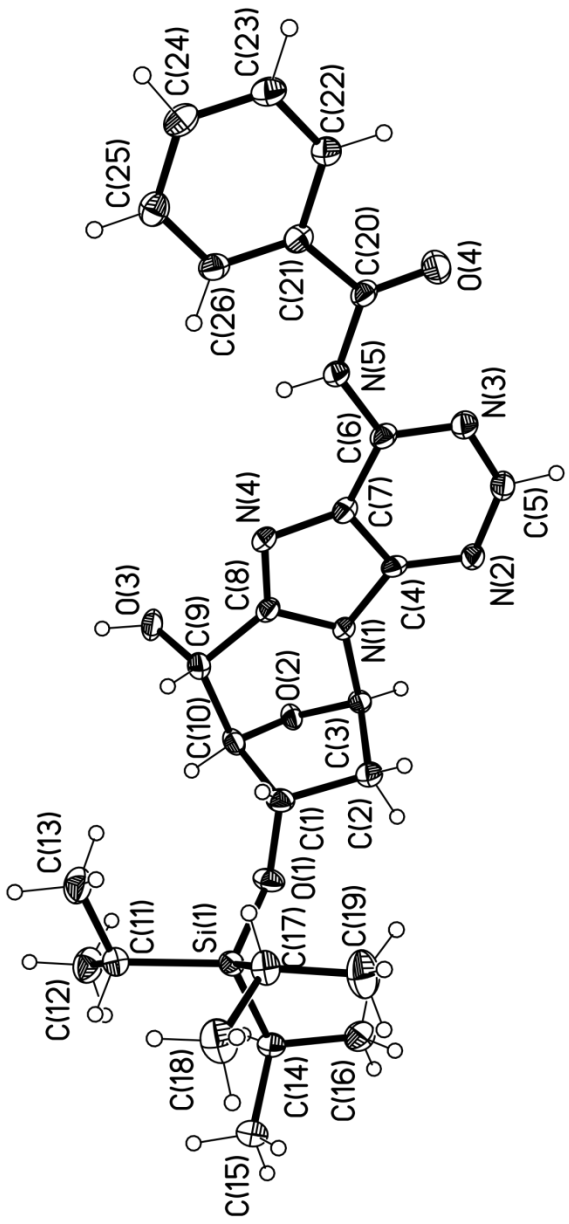


Figure 2.6. Thermal ellipsoid plot of **12*R***.

Table 2.4. Crystal data and structure refinement for sad.

Identification code	C26H35N5O4Si	
Empirical formula	C26 H35 N5 O4 Si	
Formula weight	509.68	
Temperature	100(2) K	
Wavelength	0.71073 Å	
Crystal system	Monoclinic	
Space group	C2	
Unit cell dimensions	a = 13.7148(19) Å	α= 90°.
	b = 12.3887(17) Å	β= 92.012(2)°.
	c = 15.352(2) Å	γ = 90°.
Volume	2606.8(6) $Å^3$	
Z	4	
Density (calculated)	1.299 Mg/m^3	
Absorption coefficient	0.132 mm^{-1}	
F(000)	1088	
Crystal size	0.18 x 0.07 x 0.03 mm^3	
Theta range for data collection	2.22 to 28.35°.	
Index ranges	-18<=h<=17, -16<=k<=16, -20<=l<=20	
Reflections collected	14888	
Independent reflections	6260 [R(int) = 0.0301]	
Completeness to theta = 28.35°	98.0 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.9961 and 0.9767	
Refinement method	Full-matrix least-squares on F^2	
Data / restraints / parameters	6260 / 36 / 431	
Goodness-of-fit on F^2	1.051	
Final R indices [I>2sigma(I)]	R1 = 0.0406, wR2 = 0.0959	
R indices (all data)	R1 = 0.0477, wR2 = 0.1015	
Absolute structure parameter	0.01(9)	
Extinction coefficient	na	
Largest diff. peak and hole	0.417 and -0.170 $e.Å^{-3}$	

Table 2.5. Atomic coordinates (x 10^4) and equivalent isotropic displacement parameters ($Å^2$x 10^3) for sad. U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
Si(1)	8095(1)	-633(1)	1410(1)	18(1)
O(1)	8221(1)	-764(1)	2481(1)	22(1)
O(2)	8596(1)	-384(1)	4455(1)	17(1)
O(3)	8622(1)	1962(1)	4630(1)	24(1)
O(4)	3926(1)	1782(1)	7592(1)	38(1)
N(1)	7084(1)	-40(1)	5019(1)	16(1)
N(2)	6086(1)	-983(1)	6055(1)	19(1)
N(3)	5028(1)	225(1)	6806(1)	23(1)
N(4)	6529(1)	1651(1)	5210(1)	18(1)
N(5)	4992(1)	2108(1)	6520(1)	20(1)
C(1)	7658(1)	-291(2)	3138(1)	20(1)
C(2)	7268(1)	-1180(2)	3734(1)	20(1)
C(3)	7691(1)	-871(1)	4623(1)	17(1)
C(4)	6389(1)	-115(1)	5633(1)	16(1)
C(5)	5414(1)	-740(2)	6627(1)	23(1)
C(6)	5349(1)	1080(2)	6370(1)	18(1)
C(7)	6057(1)	940(1)	5748(1)	18(1)
C(8)	7132(1)	1033(1)	4791(1)	17(1)
C(9)	7892(1)	1370(2)	4156(1)	19(1)
C(10)	8343(1)	360(2)	3757(1)	19(1)
C(11)	8728(1)	631(2)	1051(1)	22(1)
C(12)	9833(2)	589(2)	1214(2)	31(1)
C(13)	8312(2)	1640(2)	1475(2)	29(1)
C(14)	8764(2)	-1848(2)	1015(1)	22(1)
C(15)	8675(2)	-1981(2)	20(1)	31(1)
C(16)	8503(2)	-2892(2)	1471(1)	29(1)
C(17)	6758(1)	-538(2)	1097(1)	25(1)
C(18)	6566(2)	-216(2)	138(2)	36(1)
C(19)	6177(2)	-1545(2)	1316(2)	40(1)
C(20)	4277(1)	2410(2)	7088(1)	20(1)
C(21)	3953(1)	3564(2)	7037(1)	19(1)
C(22)	3258(1)	3875(2)	7638(1)	22(1)
C(23)	2899(1)	4915(2)	7632(1)	25(1)
C(24)	3223(1)	5658(2)	7036(1)	24(1)
C(25)	3908(1)	5363(2)	6439(1)	23(1)
C(26)	4266(1)	4313(2)	6439(1)	20(1)

Table 2.6. Bond lengths [Å] and angles [°] for sad.

Si(1)-O(1)	1.6547(13)
Si(1)-C(14)	1.874(2)
Si(1)-C(11)	1.883(2)
Si(1)-C(17)	1.8831(19)
O(1)-C(1)	1.418(2)
O(2)-C(3)	1.411(2)
O(2)-C(10)	1.447(2)
O(3)-C(9)	1.421(2)
O(3)-H(3O)	0.859(16)
O(4)-C(20)	1.210(2)
N(1)-C(4)	1.366(2)
N(1)-C(8)	1.377(2)
N(1)-C(3)	1.469(2)
N(2)-C(5)	1.330(2)
N(2)-C(4)	1.331(2)
N(3)-C(6)	1.336(2)
N(3)-C(5)	1.340(3)
N(4)-C(8)	1.311(2)
N(4)-C(7)	1.384(2)
N(5)-C(20)	1.386(2)
N(5)-C(6)	1.386(2)
N(5)-H(5N)	0.862(16)
C(1)-C(10)	1.539(3)
C(1)-C(2)	1.539(3)
C(1)-H(1)	1.000(15)
C(2)-C(3)	1.513(2)
C(2)-H(2A)	0.983(15)
C(2)-H(2B)	0.996(16)
C(3)-H(3)	0.978(15)
C(4)-C(7)	1.397(2)
C(5)-H(5)	0.941(16)
C(6)-C(7)	1.397(3)
C(8)-C(9)	1.510(2)
C(9)-C(10)	1.534(3)
C(9)-H(9)	0.988(15)
C(10)-H(10)	0.991(15)
C(11)-C(12)	1.528(3)
C(11)-C(13)	1.529(3)
C(11)-H(11)	1.002(15)
C(12)-H(12A)	0.985(17)
C(12)-H(12B)	0.975(17)
C(12)-H(12C)	0.989(17)
C(13)-H(13A)	1.010(17)
C(13)-H(13B)	1.004(16)
C(13)-H(13C)	0.999(17)
C(14)-C(16)	1.521(3)
C(14)-C(15)	1.537(3)
C(14)-H(14)	0.999(15)
C(15)-H(15A)	0.987(17)
C(15)-H(15B)	0.982(17)
C(15)-H(15C)	0.974(17)

C(16)-H(16A)	0.984(17)
C(16)-H(16B)	1.008(17)
C(16)-H(16C)	1.001(16)
C(17)-C(19)	1.524(3)
C(17)-C(18)	1.540(3)
C(17)-H(17)	0.998(16)
C(18)-H(18A)	0.992(17)
C(18)-H(18B)	1.001(18)
C(18)-H(18C)	0.987(17)
C(19)-H(19A)	0.982(17)
C(19)-H(19B)	0.999(17)
C(19)-H(19C)	0.971(18)
C(20)-C(21)	1.498(3)
C(21)-C(26)	1.385(3)
C(21)-C(22)	1.403(3)
C(22)-C(23)	1.379(3)
C(22)-H(22)	0.944(16)
C(23)-C(24)	1.383(3)
C(23)-H(23)	0.973(16)
C(24)-C(25)	1.384(3)
C(24)-H(24)	0.931(16)
C(25)-C(26)	1.390(3)
C(25)-H(25)	0.953(16)
C(26)-H(26)	0.933(15)

O(1)-Si(1)-C(14)	102.01(8)
O(1)-Si(1)-C(11)	109.87(8)
C(14)-Si(1)-C(11)	109.70(8)
O(1)-Si(1)-C(17)	109.04(8)
C(14)-Si(1)-C(17)	116.80(9)
C(11)-Si(1)-C(17)	109.13(9)
C(1)-O(1)-Si(1)	128.76(11)
C(3)-O(2)-C(10)	102.58(12)
C(9)-O(3)-H(3O)	110.0(17)
C(4)-N(1)-C(8)	106.34(14)
C(4)-N(1)-C(3)	131.09(15)
C(8)-N(1)-C(3)	122.56(14)
C(5)-N(2)-C(4)	111.96(16)
C(6)-N(3)-C(5)	117.75(16)
C(8)-N(4)-C(7)	103.74(15)
C(20)-N(5)-C(6)	127.89(16)
C(20)-N(5)-H(5N)	120.2(16)
C(6)-N(5)-H(5N)	111.9(16)
O(1)-C(1)-C(10)	108.68(14)
O(1)-C(1)-C(2)	109.63(15)
C(10)-C(1)-C(2)	103.00(14)
O(1)-C(1)-H(1)	114.6(13)
C(10)-C(1)-H(1)	108.6(13)
C(2)-C(1)-H(1)	111.6(13)
C(3)-C(2)-C(1)	103.01(15)
C(3)-C(2)-H(2A)	114.0(12)
C(1)-C(2)-H(2A)	116.7(13)
C(3)-C(2)-H(2B)	108.1(13)

C(1)-C(2)-H(2B)	109.8(13)
H(2A)-C(2)-H(2B)	105.1(18)
O(2)-C(3)-N(1)	106.91(13)
O(2)-C(3)-C(2)	104.76(13)
N(1)-C(3)-C(2)	110.22(14)
O(2)-C(3)-H(3)	106.6(12)
N(1)-C(3)-H(3)	107.4(12)
C(2)-C(3)-H(3)	120.1(13)
N(2)-C(4)-N(1)	128.90(16)
N(2)-C(4)-C(7)	125.86(16)
N(1)-C(4)-C(7)	105.23(15)
N(2)-C(5)-N(3)	128.75(19)
N(2)-C(5)-H(5)	116.5(14)
N(3)-C(5)-H(5)	114.6(14)
N(3)-C(6)-N(5)	121.41(16)
N(3)-C(6)-C(7)	119.52(16)
N(5)-C(6)-C(7)	119.07(17)
N(4)-C(7)-C(4)	110.98(15)
N(4)-C(7)-C(6)	132.85(17)
C(4)-C(7)-C(6)	116.16(16)
N(4)-C(8)-N(1)	113.70(15)
N(4)-C(8)-C(9)	128.02(16)
N(1)-C(8)-C(9)	118.14(15)
O(3)-C(9)-C(8)	107.60(14)
O(3)-C(9)-C(10)	109.80(15)
C(8)-C(9)-C(10)	109.21(15)
O(3)-C(9)-H(9)	114.1(13)
C(8)-C(9)-H(9)	111.7(12)
C(10)-C(9)-H(9)	104.3(13)
O(2)-C(10)-C(9)	108.23(13)
O(2)-C(10)-C(1)	104.43(14)
C(9)-C(10)-C(1)	115.37(15)
O(2)-C(10)-H(10)	108.1(13)
C(9)-C(10)-H(10)	110.2(14)
C(1)-C(10)-H(10)	110.1(13)
C(12)-C(11)-C(13)	109.95(17)
C(12)-C(11)-Si(1)	112.83(14)
C(13)-C(11)-Si(1)	111.88(14)
C(12)-C(11)-H(11)	107.4(13)
C(13)-C(11)-H(11)	109.0(14)
Si(1)-C(11)-H(11)	105.5(14)
C(11)-C(12)-H(12A)	112.4(16)
C(11)-C(12)-H(12B)	109.5(17)
H(12A)-C(12)-H(12B)	108(2)
C(11)-C(12)-H(12C)	110.1(15)
H(12A)-C(12)-H(12C)	108(2)
H(12B)-C(12)-H(12C)	109(2)
C(11)-C(13)-H(13A)	108.0(15)
C(11)-C(13)-H(13B)	111.5(16)
H(13A)-C(13)-H(13B)	111(2)
C(11)-C(13)-H(13C)	111.5(17)
H(13A)-C(13)-H(13C)	109(2)
H(13B)-C(13)-H(13C)	106(2)

C(16)-C(14)-C(15)	110.76(17)
C(16)-C(14)-Si(1)	114.08(14)
C(15)-C(14)-Si(1)	112.68(15)
C(16)-C(14)-H(14)	104.5(13)
C(15)-C(14)-H(14)	109.1(12)
Si(1)-C(14)-H(14)	105.1(13)
C(14)-C(15)-H(15A)	111.4(16)
C(14)-C(15)-H(15B)	113.7(16)
H(15A)-C(15)-H(15B)	107(2)
C(14)-C(15)-H(15C)	111.6(16)
H(15A)-C(15)-H(15C)	111(2)
H(15B)-C(15)-H(15C)	102(2)
C(14)-C(16)-H(16A)	112.5(16)
C(14)-C(16)-H(16B)	109.4(16)
H(16A)-C(16)-H(16B)	111(2)
C(14)-C(16)-H(16C)	110.1(16)
H(16A)-C(16)-H(16C)	105(2)
H(16B)-C(16)-H(16C)	109(2)
C(19)-C(17)-C(18)	110.46(18)
C(19)-C(17)-Si(1)	113.83(15)
C(18)-C(17)-Si(1)	113.12(14)
C(19)-C(17)-H(17)	104.8(14)
C(18)-C(17)-H(17)	103.4(13)
Si(1)-C(17)-H(17)	110.4(14)
C(17)-C(18)-H(18A)	108.2(17)
C(17)-C(18)-H(18B)	115.3(18)
H(18A)-C(18)-H(18B)	107(2)
C(17)-C(18)-H(18C)	113.7(17)
H(18A)-C(18)-H(18C)	109(2)
H(18B)-C(18)-H(18C)	103(2)
C(17)-C(19)-H(19A)	113.1(19)
C(17)-C(19)-H(19B)	116.1(18)
H(19A)-C(19)-H(19B)	104(2)
C(17)-C(19)-H(19C)	111.0(19)
H(19A)-C(19)-H(19C)	106(3)
H(19B)-C(19)-H(19C)	106(3)
O(4)-C(20)-N(5)	122.31(17)
O(4)-C(20)-C(21)	121.57(16)
N(5)-C(20)-C(21)	116.12(16)
C(26)-C(21)-C(22)	119.11(17)
C(26)-C(21)-C(20)	125.11(16)
C(22)-C(21)-C(20)	115.75(17)
C(23)-C(22)-C(21)	120.16(18)
C(23)-C(22)-H(22)	121.1(14)
C(21)-C(22)-H(22)	118.7(14)
C(22)-C(23)-C(24)	120.21(18)
C(22)-C(23)-H(23)	119.1(15)
C(24)-C(23)-H(23)	120.7(15)
C(23)-C(24)-C(25)	120.27(18)
C(23)-C(24)-H(24)	116.4(15)
C(25)-C(24)-H(24)	123.4(15)
C(24)-C(25)-C(26)	119.69(18)
C(24)-C(25)-H(25)	120.2(14)

C(26)-C(25)-H(25)	120.0(14)
C(21)-C(26)-C(25)	120.55(16)
C(21)-C(26)-H(26)	121.6(15)
C(25)-C(26)-H(26)	117.8(15)

Symmetry transformations used to generate equivalent atoms:

Table 2.7. Anisotropic displacement parameters ($Å^2 x\ 10^3$)for sad. The anisotropic displacement factor exponent takes the form: $-2\pi^2[\ h^2 a^{*2} U^{11} + ... + 2\ h\ k\ a^*\ b^*\ U^{12}\ ]$

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
Si(1)	21(1)	17(1)	16(1)	1(1)	3(1)	1(1)
O(1)	28(1)	22(1)	16(1)	0(1)	5(1)	9(1)
O(2)	18(1)	15(1)	19(1)	4(1)	4(1)	1(1)
O(3)	28(1)	15(1)	27(1)	3(1)	1(1)	-6(1)
O(4)	41(1)	19(1)	57(1)	3(1)	30(1)	5(1)
N(1)	19(1)	12(1)	18(1)	-1(1)	4(1)	0(1)
N(2)	20(1)	14(1)	24(1)	-1(1)	5(1)	0(1)
N(3)	22(1)	18(1)	31(1)	-2(1)	8(1)	0(1)
N(4)	21(1)	12(1)	21(1)	-2(1)	2(1)	1(1)
N(5)	20(1)	15(1)	24(1)	-2(1)	6(1)	1(1)
C(1)	23(1)	18(1)	18(1)	-1(1)	4(1)	5(1)
C(2)	21(1)	18(1)	21(1)	-2(1)	4(1)	1(1)
C(3)	18(1)	13(1)	19(1)	1(1)	6(1)	3(1)
C(4)	15(1)	15(1)	18(1)	-4(1)	1(1)	2(1)
C(5)	23(1)	16(1)	31(1)	0(1)	8(1)	-2(1)
C(6)	16(1)	16(1)	24(1)	-4(1)	2(1)	1(1)
C(7)	17(1)	14(1)	21(1)	-2(1)	0(1)	0(1)
C(8)	20(1)	12(1)	18(1)	-1(1)	0(1)	0(1)
C(9)	21(1)	16(1)	19(1)	3(1)	2(1)	-1(1)
C(10)	20(1)	17(1)	19(1)	5(1)	6(1)	1(1)
C(11)	25(1)	20(1)	21(1)	1(1)	2(1)	-1(1)
C(12)	25(1)	27(1)	41(1)	-1(1)	5(1)	-4(1)
C(13)	30(1)	19(1)	39(1)	3(1)	6(1)	1(1)
C(14)	25(1)	21(1)	21(1)	-4(1)	7(1)	0(1)
C(15)	43(1)	30(1)	22(1)	-4(1)	8(1)	-6(1)
C(16)	35(1)	21(1)	30(1)	-2(1)	4(1)	3(1)
C(17)	23(1)	24(1)	28(1)	4(1)	0(1)	-2(1)
C(18)	30(1)	42(1)	36(1)	10(1)	-8(1)	-3(1)
C(19)	29(1)	42(2)	50(2)	12(1)	-2(1)	-10(1)
C(20)	16(1)	15(1)	28(1)	-5(1)	5(1)	-1(1)
C(21)	16(1)	18(1)	24(1)	-3(1)	0(1)	0(1)
C(22)	22(1)	20(1)	26(1)	-2(1)	5(1)	0(1)
C(23)	25(1)	22(1)	30(1)	-4(1)	6(1)	5(1)
C(24)	24(1)	19(1)	29(1)	-2(1)	-3(1)	5(1)
C(25)	25(1)	20(1)	23(1)	2(1)	-3(1)	1(1)
C(26)	19(1)	20(1)	22(1)	-4(1)	2(1)	1(1)

Table 2.8. Hydrogen coordinates (x 10^4) and isotropic displacement parameters (Å^{2}x 10^3) for sad.

	x	y	z	U(eq)
H(3O)	8752(18)	2549(16)	4359(16)	35
H(5N)	5251(15)	2582(17)	6189(14)	24
H(1)	7124(14)	196(17)	2916(14)	24
H(2A)	6559(11)	-1291(19)	3717(13)	24
H(2B)	7549(15)	-1890(14)	3572(14)	24
H(3)	7822(15)	-1426(15)	5064(12)	20
H(5)	5210(16)	-1306(16)	6986(13)	28
H(9)	7598(14)	1762(18)	3652(12)	22
H(10)	8948(13)	551(19)	3459(13)	23
H(11)	8602(16)	670(20)	405(10)	27
H(12A)	10156(19)	1265(17)	1049(17)	47
H(12B)	10100(20)	3(19)	875(17)	47
H(12C)	9988(18)	460(20)	1839(11)	47
H(13A)	8614(18)	2292(17)	1196(16)	44
H(13B)	8434(19)	1640(20)	2124(11)	44
H(13C)	7588(13)	1680(20)	1383(17)	44
H(14)	9463(12)	-1720(19)	1191(14)	27
H(15A)	9098(17)	-2568(18)	-182(18)	47
H(15B)	8852(19)	-1330(18)	-303(16)	47
H(15C)	8000(13)	-2090(30)	-178(17)	47
H(16A)	8969(17)	-3477(19)	1362(17)	43
H(16B)	7819(14)	-3110(20)	1285(16)	43
H(16C)	8537(19)	-2790(20)	2118(11)	43
H(17)	6445(16)	56(16)	1426(14)	30
H(18A)	5853(13)	-120(20)	39(18)	54
H(18B)	6784(18)	-750(20)	-301(16)	54
H(18C)	6910(19)	446(19)	-33(18)	54
H(19A)	5471(13)	-1450(30)	1216(19)	61
H(19B)	6250(20)	-1800(20)	1931(13)	61
H(19C)	6360(20)	-2150(20)	955(18)	61
H(22)	3047(16)	3360(16)	8044(13)	27
H(23)	2418(15)	5120(20)	8054(14)	30
H(24)	2959(16)	6349(15)	7067(15)	29
H(25)	4117(16)	5869(17)	6017(13)	27
H(26)	4737(13)	4141(19)	6036(12)	24

Table 2.9. Torsion angles [°] for sad.

C(14)-Si(1)-O(1)-C(1)	-159.98(16)
C(11)-Si(1)-O(1)-C(1)	83.71(17)
C(17)-Si(1)-O(1)-C(1)	-35.85(18)
Si(1)-O(1)-C(1)-C(10)	-122.85(14)
Si(1)-O(1)-C(1)-C(2)	125.27(15)
O(1)-C(1)-C(2)-C(3)	120.80(15)
C(10)-C(1)-C(2)-C(3)	5.25(17)
C(10)-O(2)-C(3)-N(1)	-69.48(15)
C(10)-O(2)-C(3)-C(2)	47.50(16)
C(4)-N(1)-C(3)-O(2)	-146.68(17)
C(8)-N(1)-C(3)-O(2)	34.3(2)
C(4)-N(1)-C(3)-C(2)	100.0(2)
C(8)-N(1)-C(3)-C(2)	-79.00(19)
C(1)-C(2)-C(3)-O(2)	-32.18(17)
C(1)-C(2)-C(3)-N(1)	82.51(16)
C(5)-N(2)-C(4)-N(1)	178.14(17)
C(5)-N(2)-C(4)-C(7)	-0.6(3)
C(8)-N(1)-C(4)-N(2)	-179.45(17)
C(3)-N(1)-C(4)-N(2)	1.4(3)
C(8)-N(1)-C(4)-C(7)	-0.48(18)
C(3)-N(1)-C(4)-C(7)	-179.61(16)
C(4)-N(2)-C(5)-N(3)	0.6(3)
C(6)-N(3)-C(5)-N(2)	-0.7(3)
C(5)-N(3)-C(6)-N(5)	-179.64(17)
C(5)-N(3)-C(6)-C(7)	0.6(3)
C(20)-N(5)-C(6)-N(3)	-1.6(3)
C(20)-N(5)-C(6)-C(7)	178.17(17)
C(8)-N(4)-C(7)-C(4)	-0.48(19)
C(8)-N(4)-C(7)-C(6)	178.27(19)
N(2)-C(4)-C(7)-N(4)	179.62(16)
N(1)-C(4)-C(7)-N(4)	0.62(19)
N(2)-C(4)-C(7)-C(6)	0.6(3)
N(1)-C(4)-C(7)-C(6)	-178.36(15)
N(3)-C(6)-C(7)-N(4)	-179.29(18)
N(5)-C(6)-C(7)-N(4)	0.9(3)
N(3)-C(6)-C(7)-C(4)	-0.6(2)
N(5)-C(6)-C(7)-C(4)	179.64(15)
C(7)-N(4)-C(8)-N(1)	0.17(19)
C(7)-N(4)-C(8)-C(9)	-175.49(17)
C(4)-N(1)-C(8)-N(4)	0.2(2)
C(3)-N(1)-C(8)-N(4)	179.43(15)
C(4)-N(1)-C(8)-C(9)	176.33(15)
C(3)-N(1)-C(8)-C(9)	-4.5(2)
N(4)-C(8)-C(9)-O(3)	67.1(2)
N(1)-C(8)-C(9)-O(3)	-108.35(17)
N(4)-C(8)-C(9)-C(10)	-173.73(16)
N(1)-C(8)-C(9)-C(10)	10.8(2)
C(3)-O(2)-C(10)-C(9)	79.96(16)
C(3)-O(2)-C(10)-C(1)	-43.44(15)
O(3)-C(9)-C(10)-O(2)	70.15(17)
C(8)-C(9)-C(10)-O(2)	-47.62(18)

O(3)-C(9)-C(10)-C(1)	-173.34(14)
C(8)-C(9)-C(10)-C(1)	68.90(19)
O(1)-C(1)-C(10)-O(2)	-93.88(16)
C(2)-C(1)-C(10)-O(2)	22.35(17)
O(1)-C(1)-C(10)-C(9)	147.47(15)
C(2)-C(1)-C(10)-C(9)	-96.29(17)
O(1)-Si(1)-C(11)-C(12)	65.01(16)
C(14)-Si(1)-C(11)-C(12)	-46.35(16)
C(17)-Si(1)-C(11)-C(12)	-175.48(14)
O(1)-Si(1)-C(11)-C(13)	-59.59(16)
C(14)-Si(1)-C(11)-C(13)	-170.96(15)
C(17)-Si(1)-C(11)-C(13)	59.91(16)
O(1)-Si(1)-C(14)-C(16)	47.07(16)
C(11)-Si(1)-C(14)-C(16)	163.51(15)
C(17)-Si(1)-C(14)-C(16)	-71.68(17)
O(1)-Si(1)-C(14)-C(15)	174.49(14)
C(11)-Si(1)-C(14)-C(15)	-69.08(16)
C(17)-Si(1)-C(14)-C(15)	55.73(17)
O(1)-Si(1)-C(17)-C(19)	-62.88(18)
C(14)-Si(1)-C(17)-C(19)	52.02(19)
C(11)-Si(1)-C(17)-C(19)	177.11(16)
O(1)-Si(1)-C(17)-C(18)	169.96(16)
C(14)-Si(1)-C(17)-C(18)	-75.15(19)
C(11)-Si(1)-C(17)-C(18)	49.95(19)
C(6)-N(5)-C(20)-O(4)	6.4(3)
C(6)-N(5)-C(20)-C(21)	-173.00(16)
O(4)-C(20)-C(21)-C(26)	-175.15(19)
N(5)-C(20)-C(21)-C(26)	4.3(3)
O(4)-C(20)-C(21)-C(22)	3.1(3)
N(5)-C(20)-C(21)-C(22)	-177.50(16)
C(26)-C(21)-C(22)-C(23)	-0.3(3)
C(20)-C(21)-C(22)-C(23)	-178.67(17)
C(21)-C(22)-C(23)-C(24)	-0.2(3)
C(22)-C(23)-C(24)-C(25)	0.2(3)
C(23)-C(24)-C(25)-C(26)	0.3(3)
C(22)-C(21)-C(26)-C(25)	0.8(3)
C(20)-C(21)-C(26)-C(25)	179.00(17)
C(24)-C(25)-C(26)-C(21)	-0.8(3)

Symmetry transformations used to generate equivalent atoms:

Table 2.10. Hydrogen bonds for sad [Å and °].

D-H...A	d(D-H)	d(H...A)	d(D...A)	<(DHA)
O(3)-H(3O)...N(2)#1	0.859(16)	1.942(17)	2.788(2)	168(3)

Symmetry transformations used to generate equivalent atoms:
#1 -x+3/2,y+1/2,-z+1

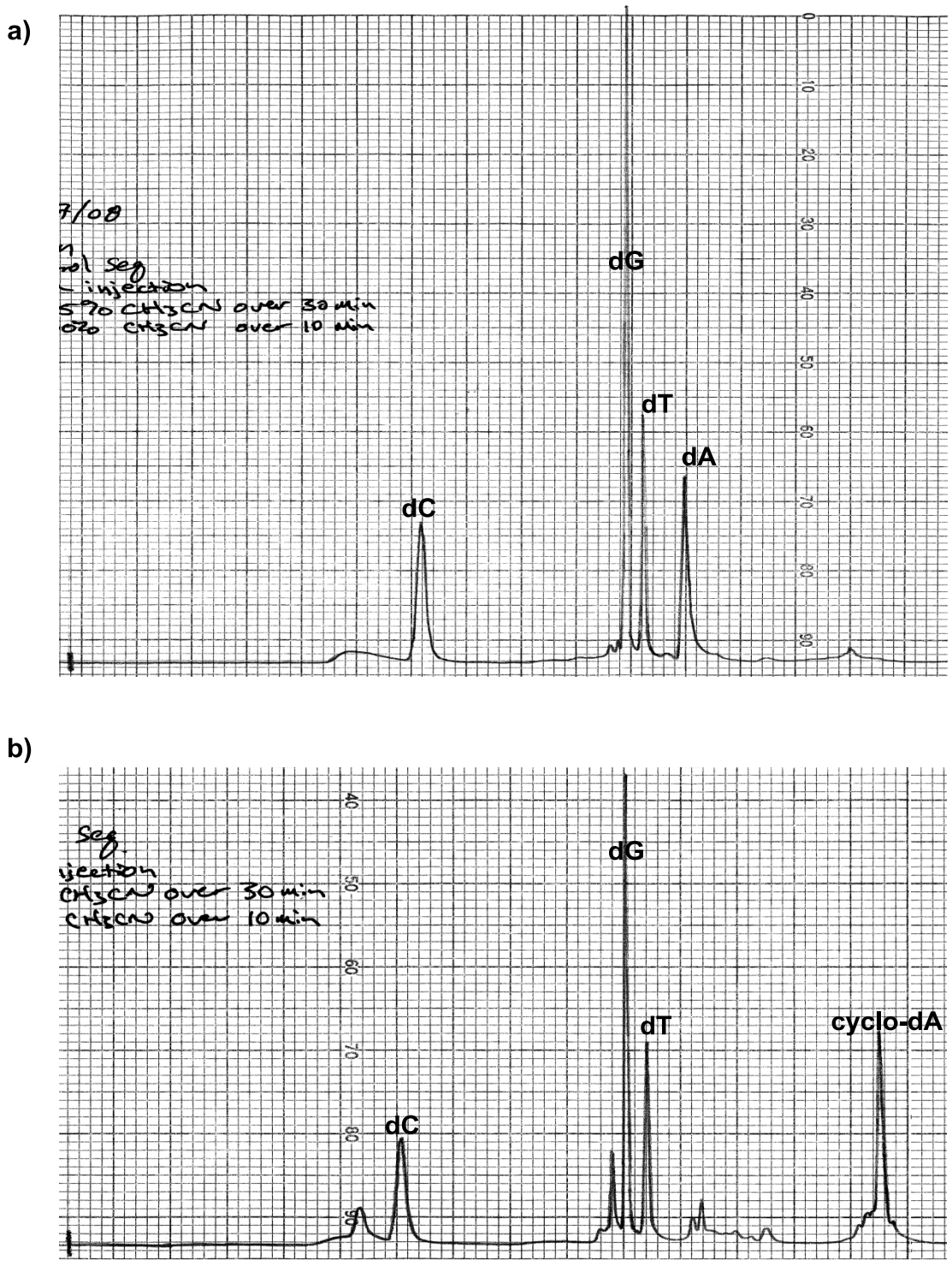


Figure 2.7. HPLC traces for Table 2.3. a) Control sequence. b) Modified sequence.

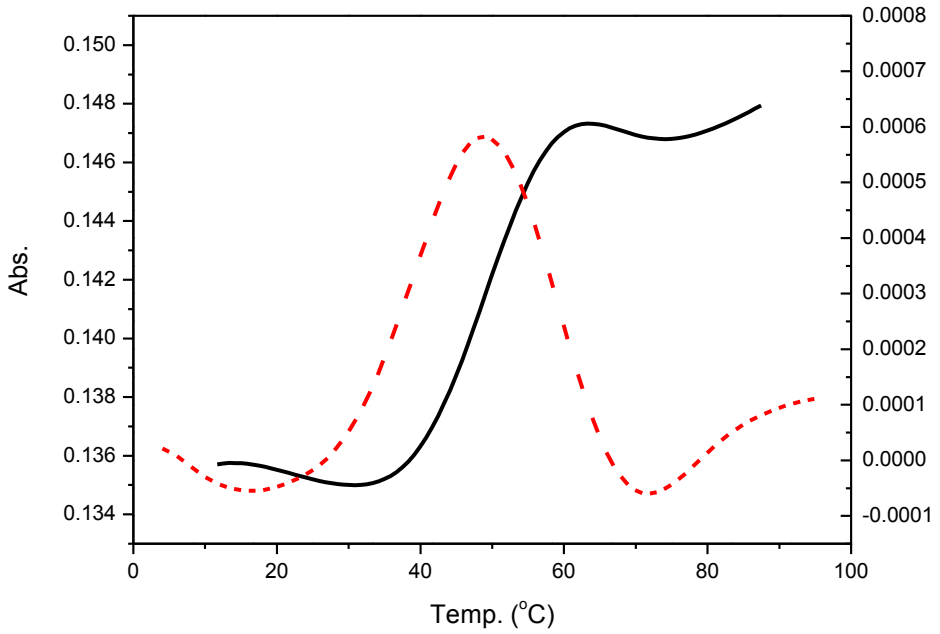


Figure 2.8. T_M isotherm for entry 1 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

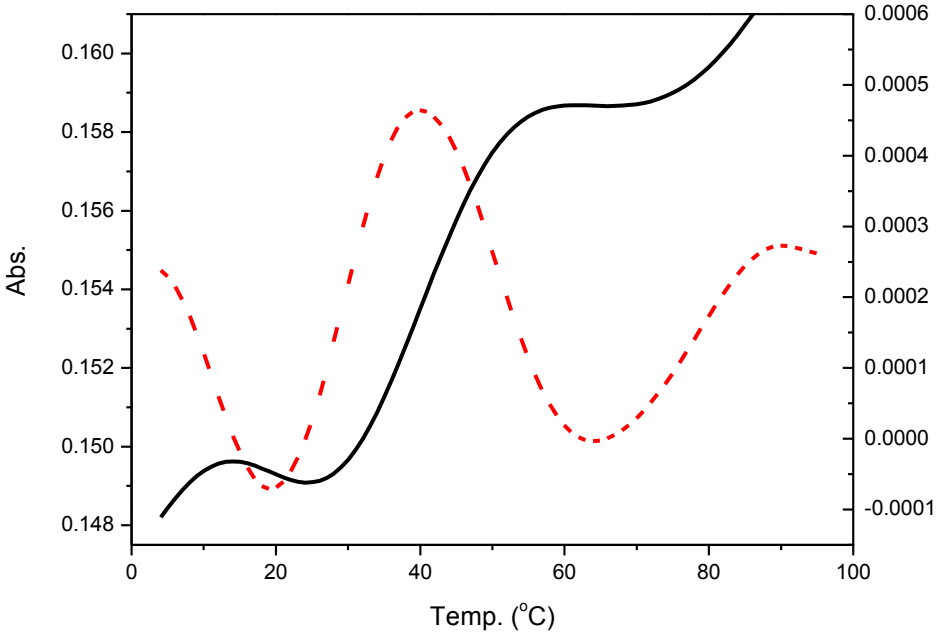


Figure 2.9. T_M isotherm for entry 2 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

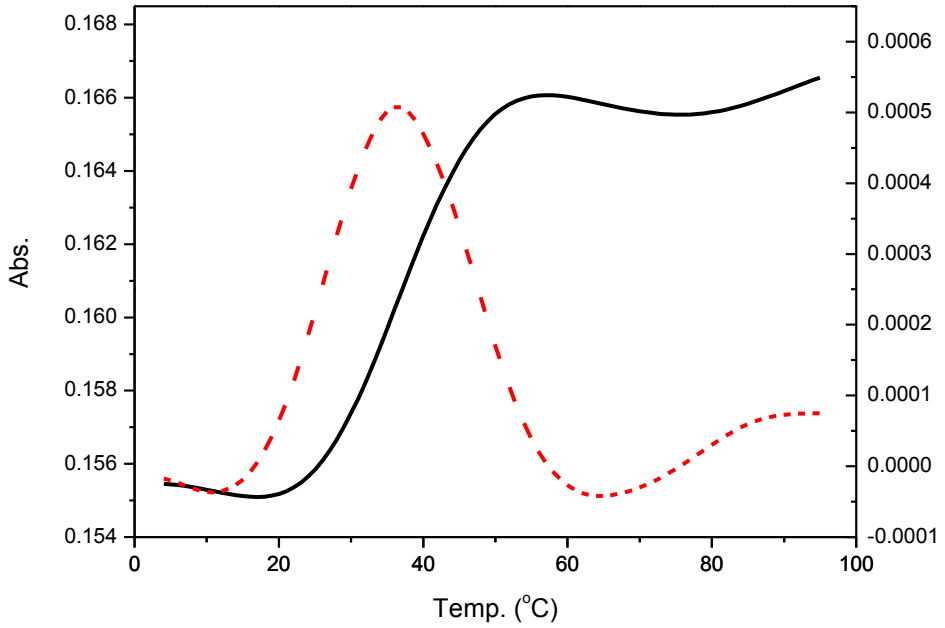


Figure 2.10. T_M isotherm for entry 3 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

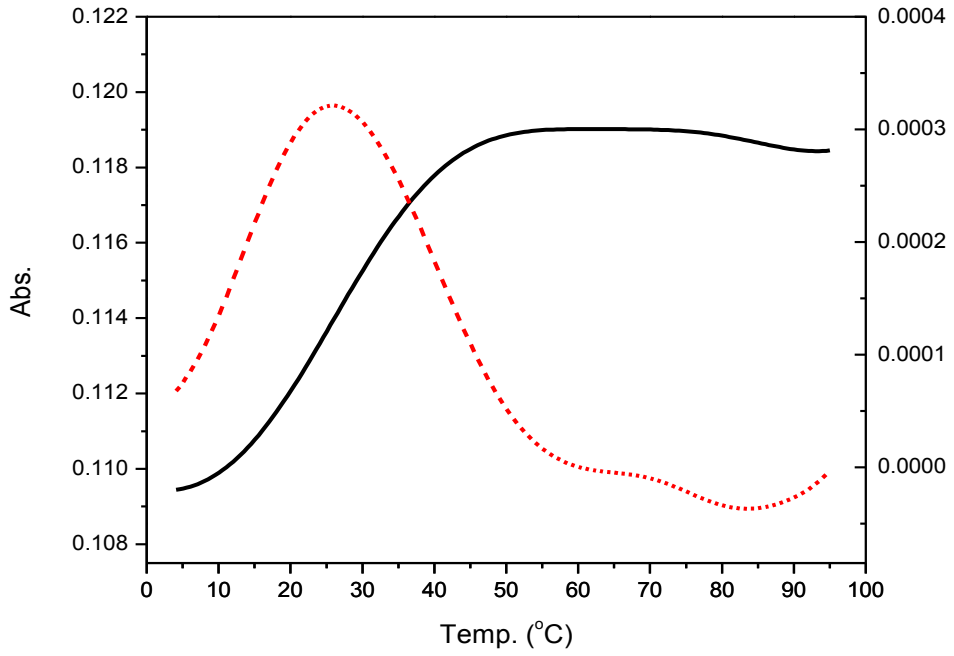


Figure 2.11. T_M isotherm for entry 4 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

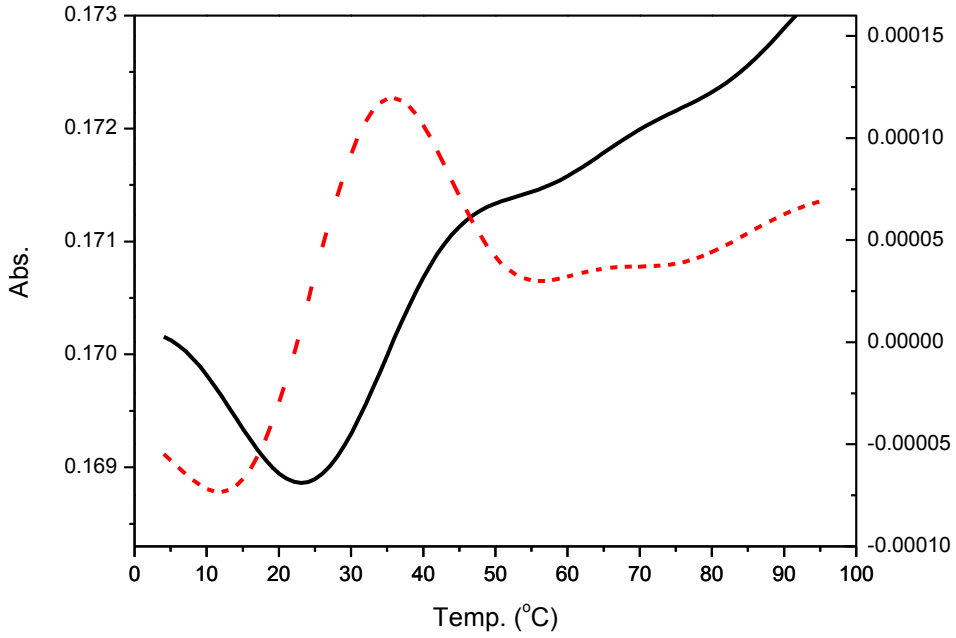


Figure 2.12. T_M isotherm for entry 5 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

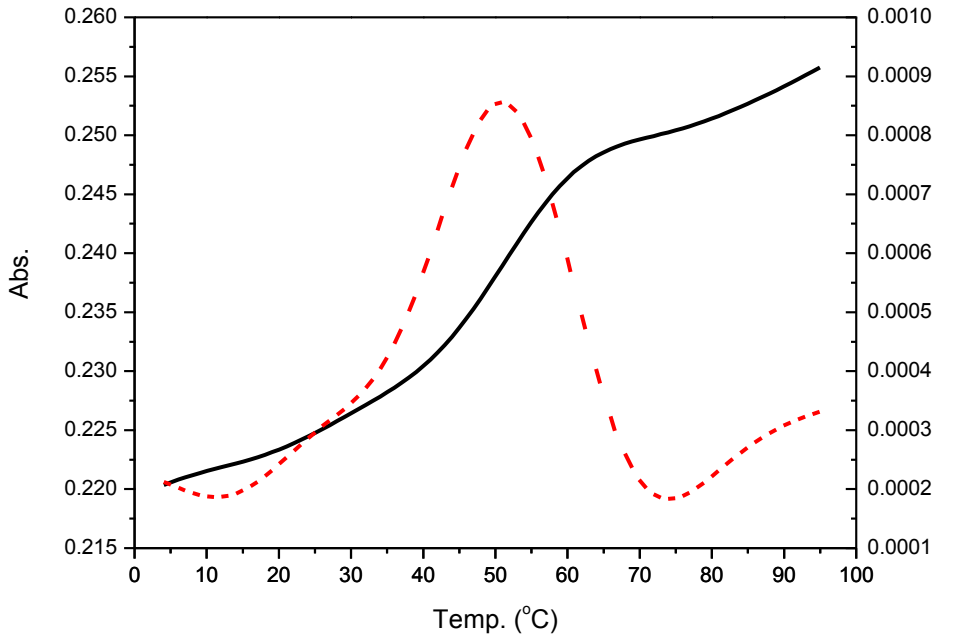


Figure 2.13. T_M isotherm for entry 6 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

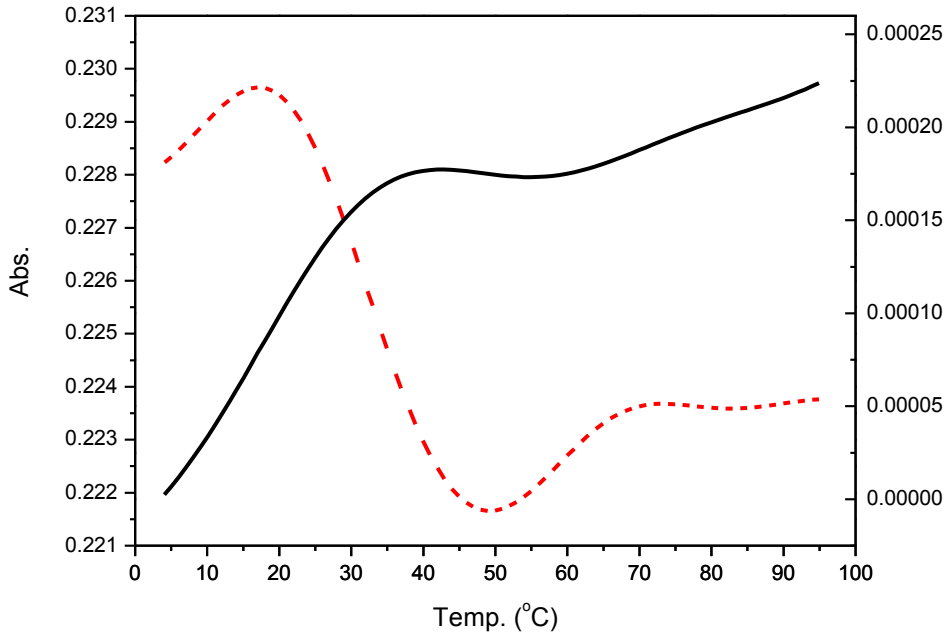


Figure 2.14. T_M isotherm for entry 7 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

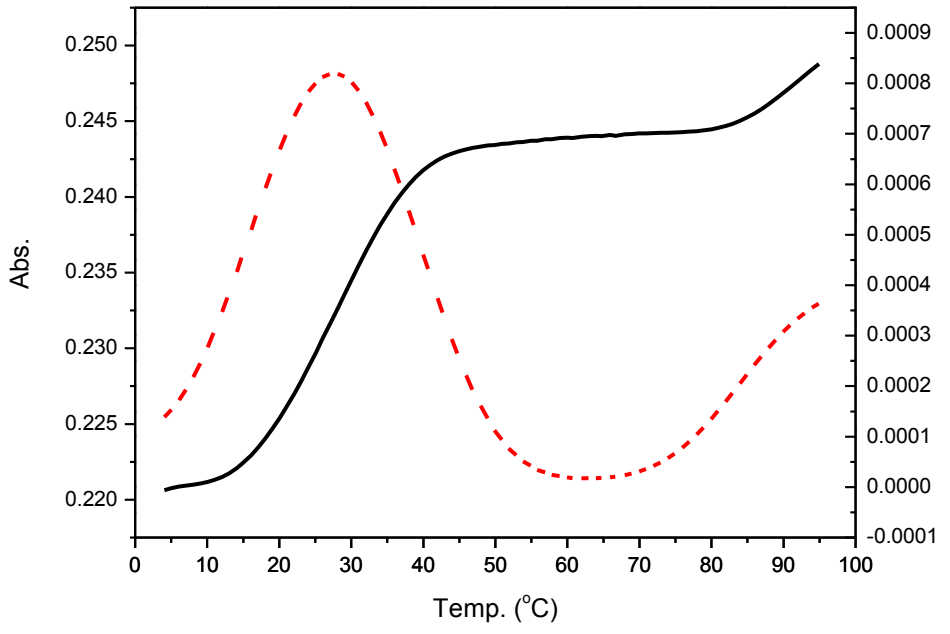


Figure 2.15. T_M isotherm for entry 8 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

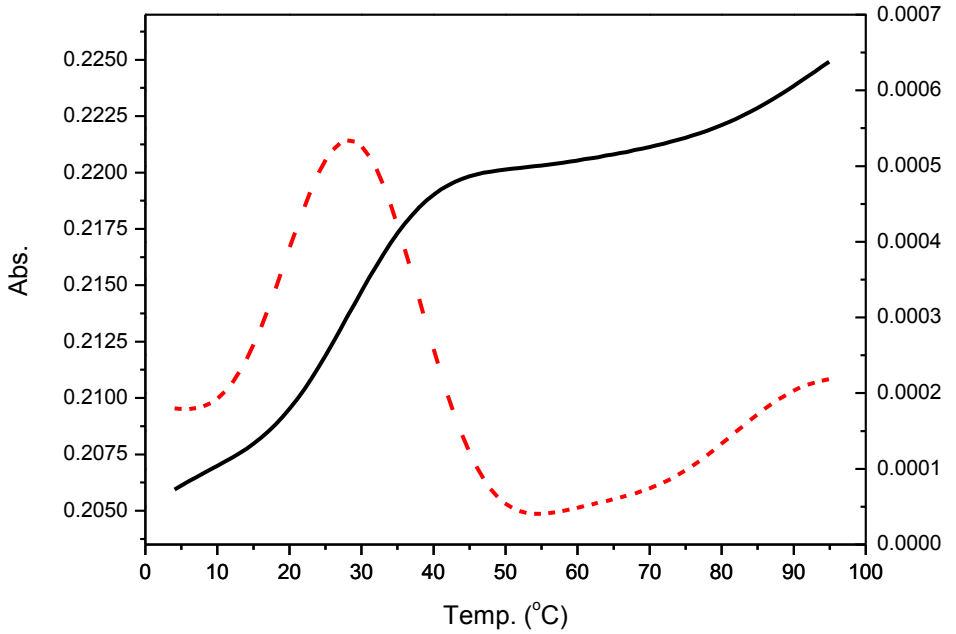


Figure 2.16. T_M isotherm for entry 9 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

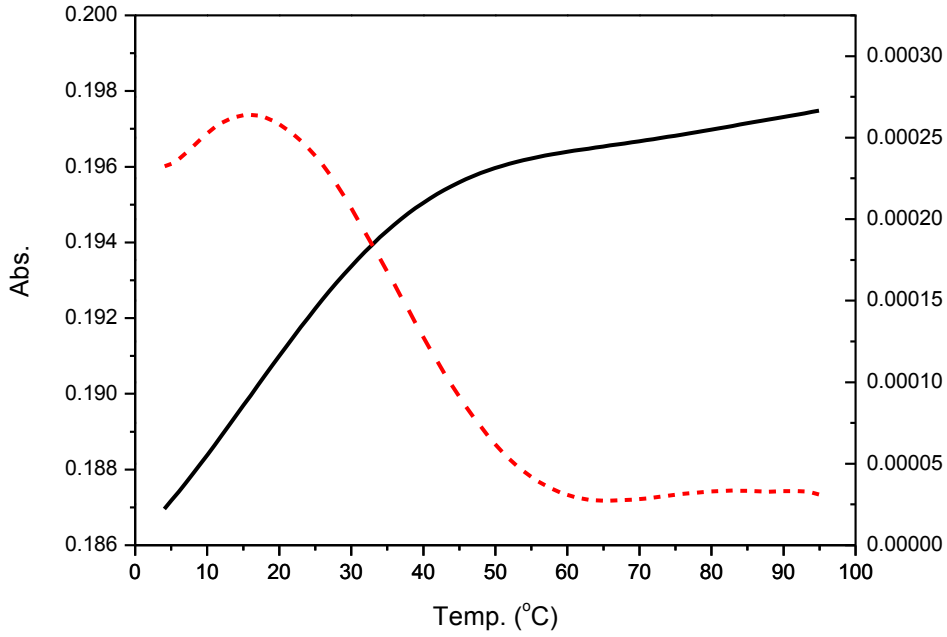


Figure 2.17. T_M isotherm for entry 10 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

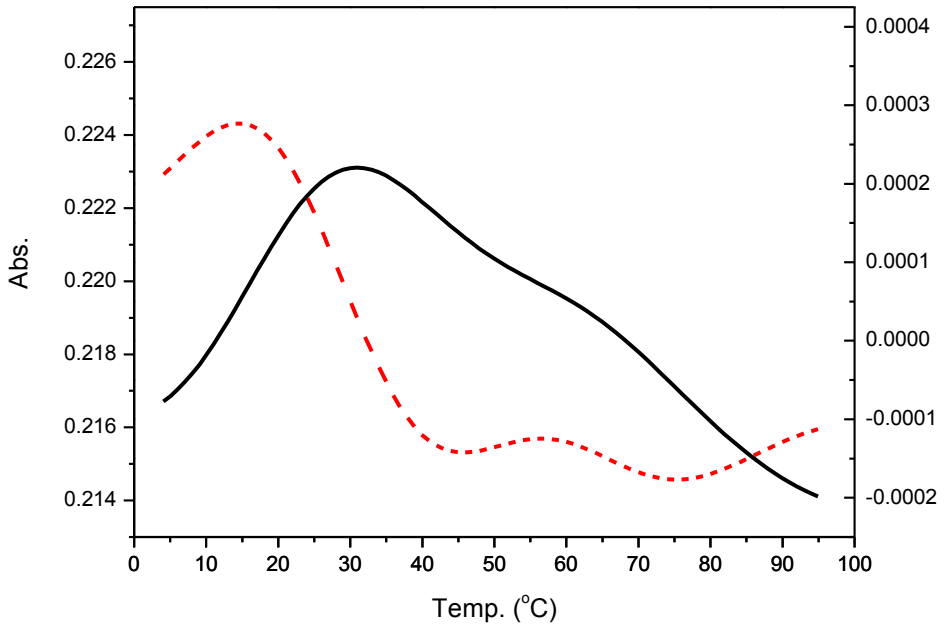


Figure 2.18. T_M isotherm for entry 11 in Table 2.1. DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

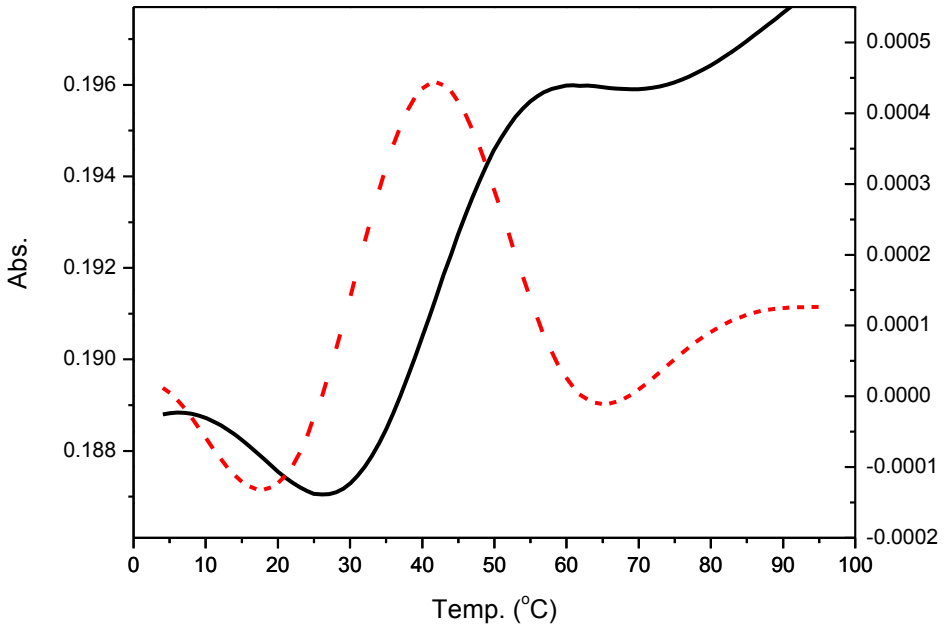


Figure 2.19. T_M isotherm for entry 15 in Table 2.1. RNA, DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

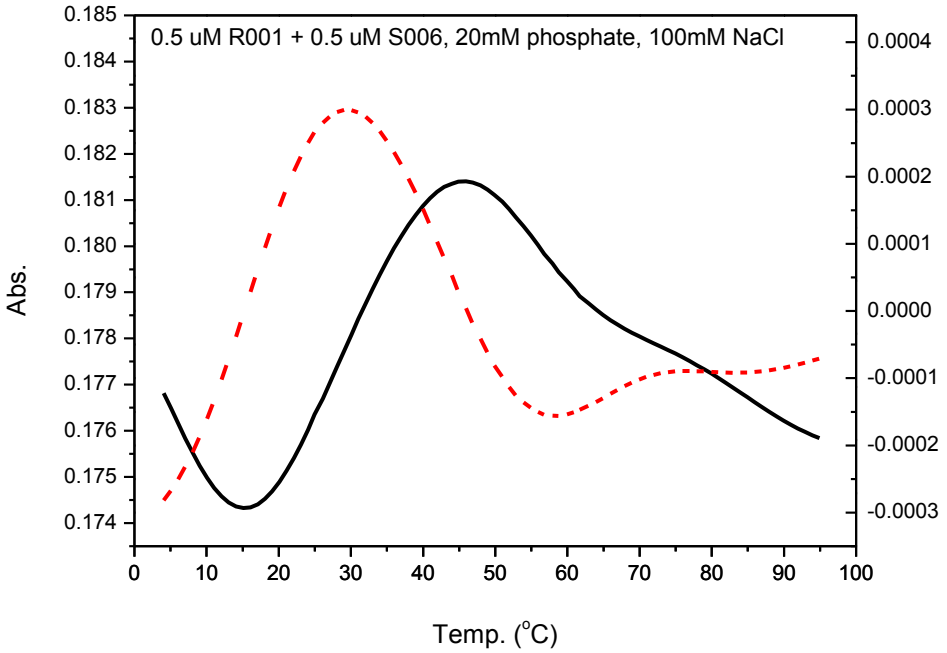


Figure 2.20. T_M isotherm for entry 16 in Table 2.1. RNA, DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

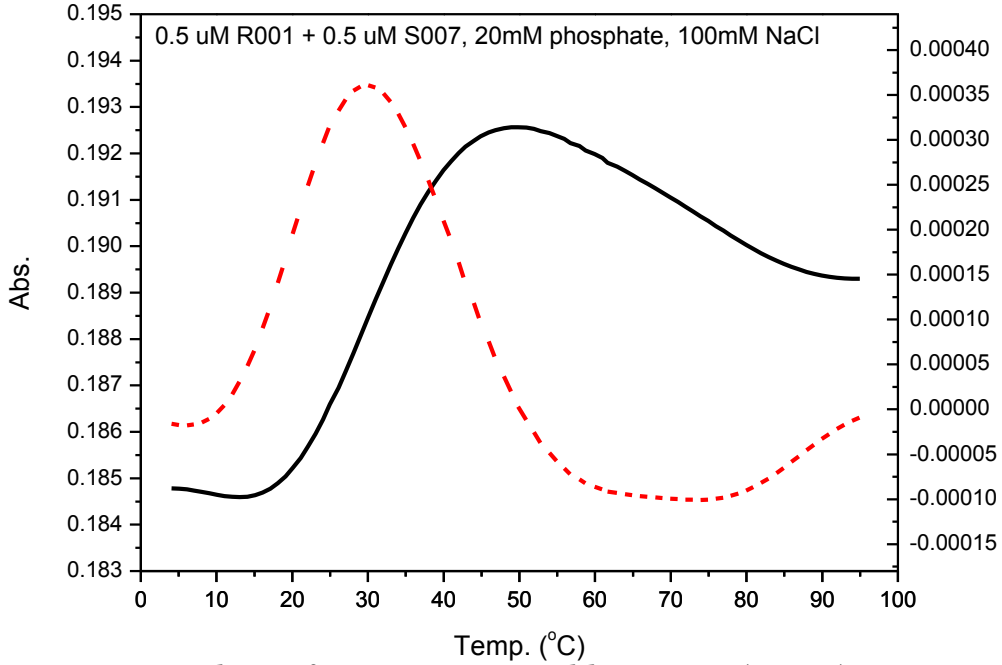


Figure 2.21. T_M isotherm for entry 17 in Table 2.1. RNA, DNA concentration: 0.5 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

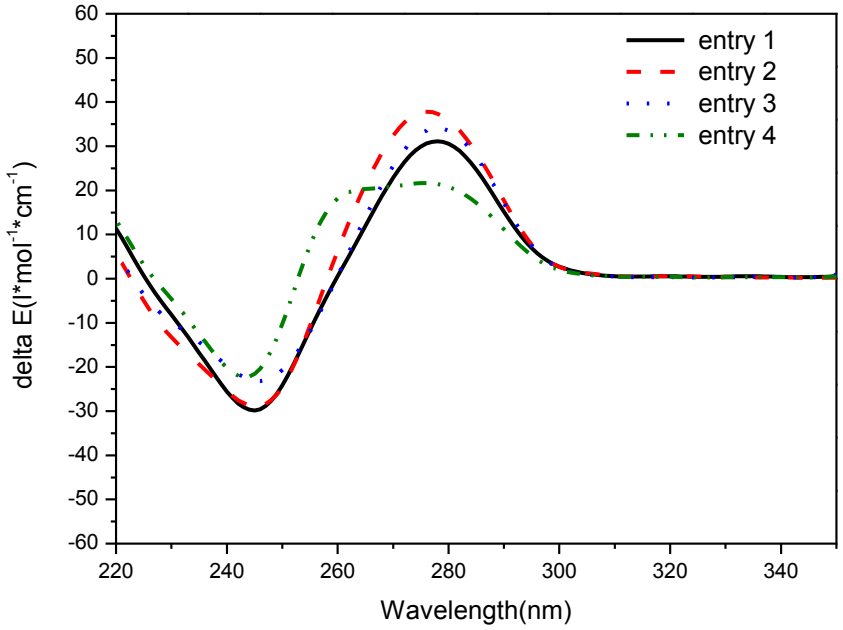


Figure 2.22. CD Spectra for entry 1 to entry 4 in Table 2.1. DNA concentration: 4.0 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

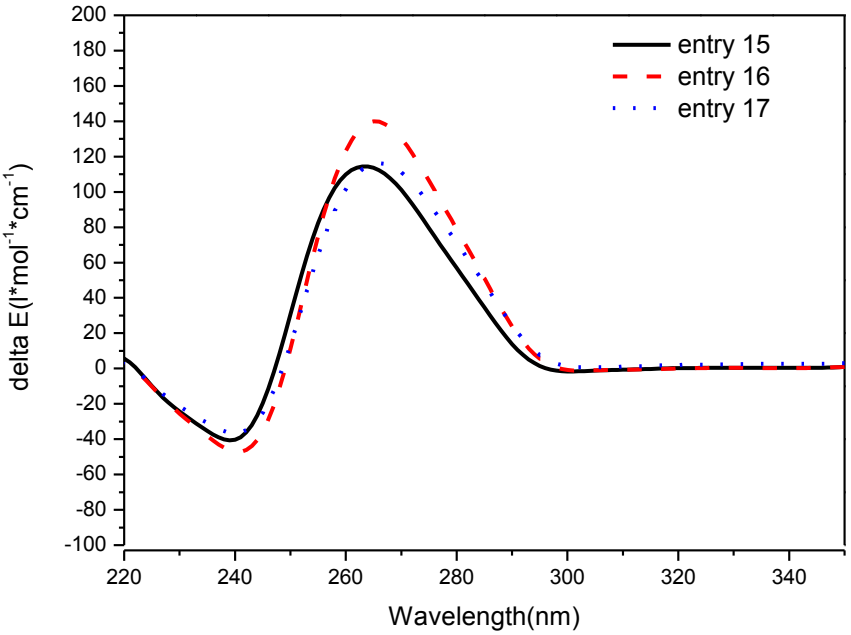
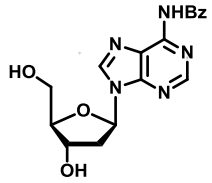
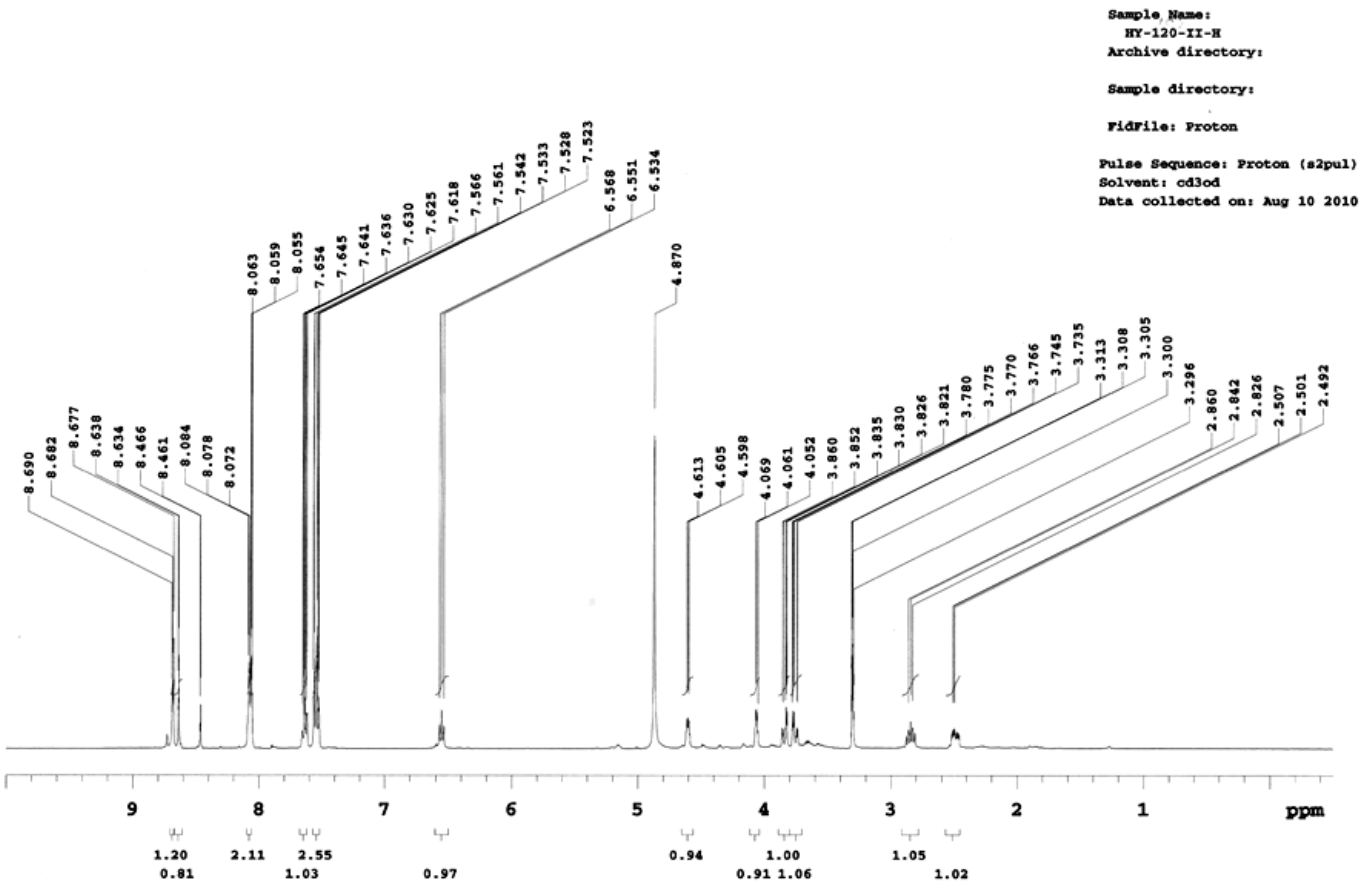


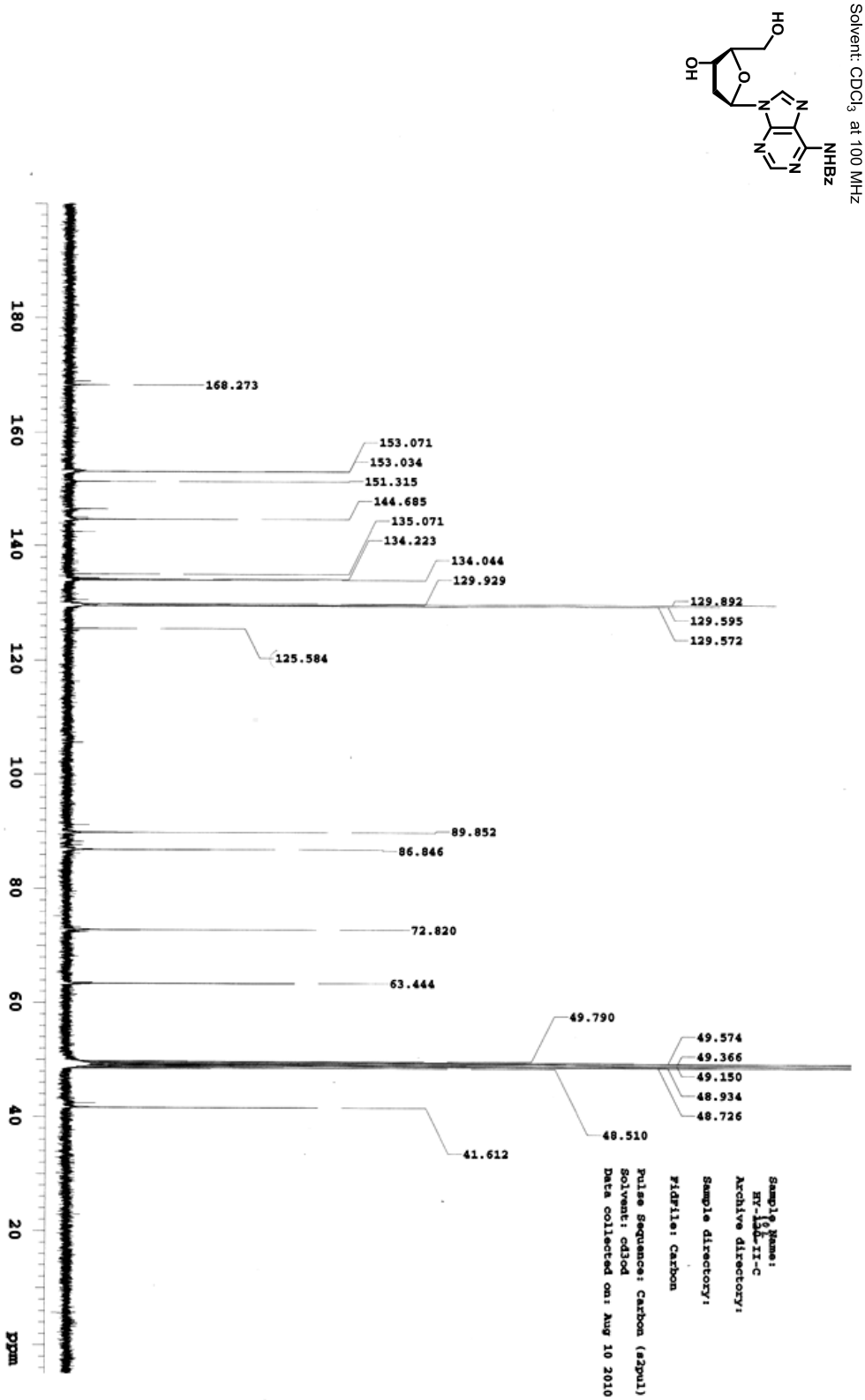
Figure 2.23. CD Spectra for entry 15 to entry 17 in Table 2.1. RNA and DNA concentration: 4.0 μM for each sequence. Buffer condition: 20 mM phosphate pH 7.0, 100 mM NaCl.

^{1}H NMR of Compound **4**
Solvent: $CDCl_3$ at 400 MHz

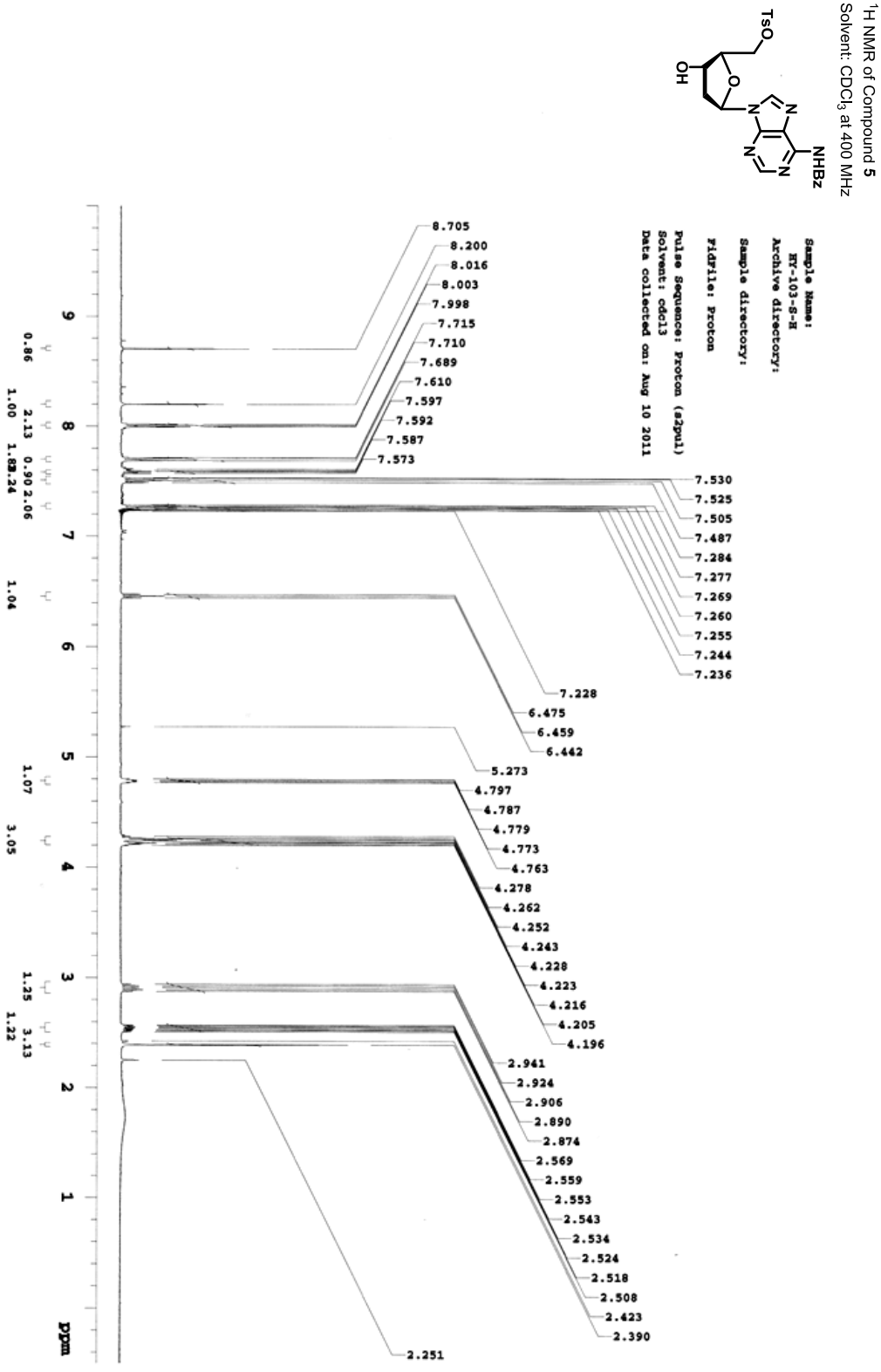




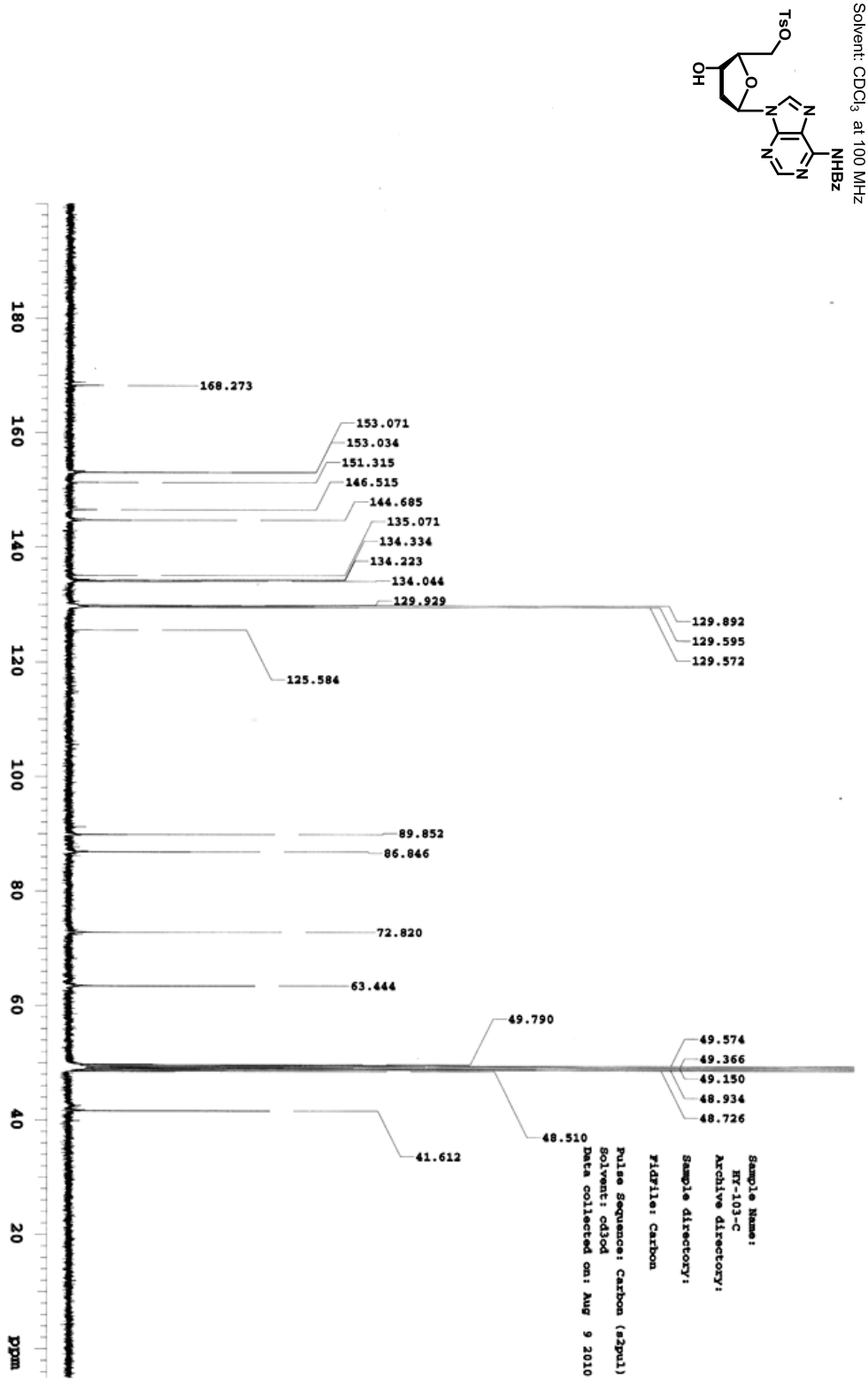
^{13}C NMR of Compound **4**
Solvent: $CDCl_3$ at 100 MHz



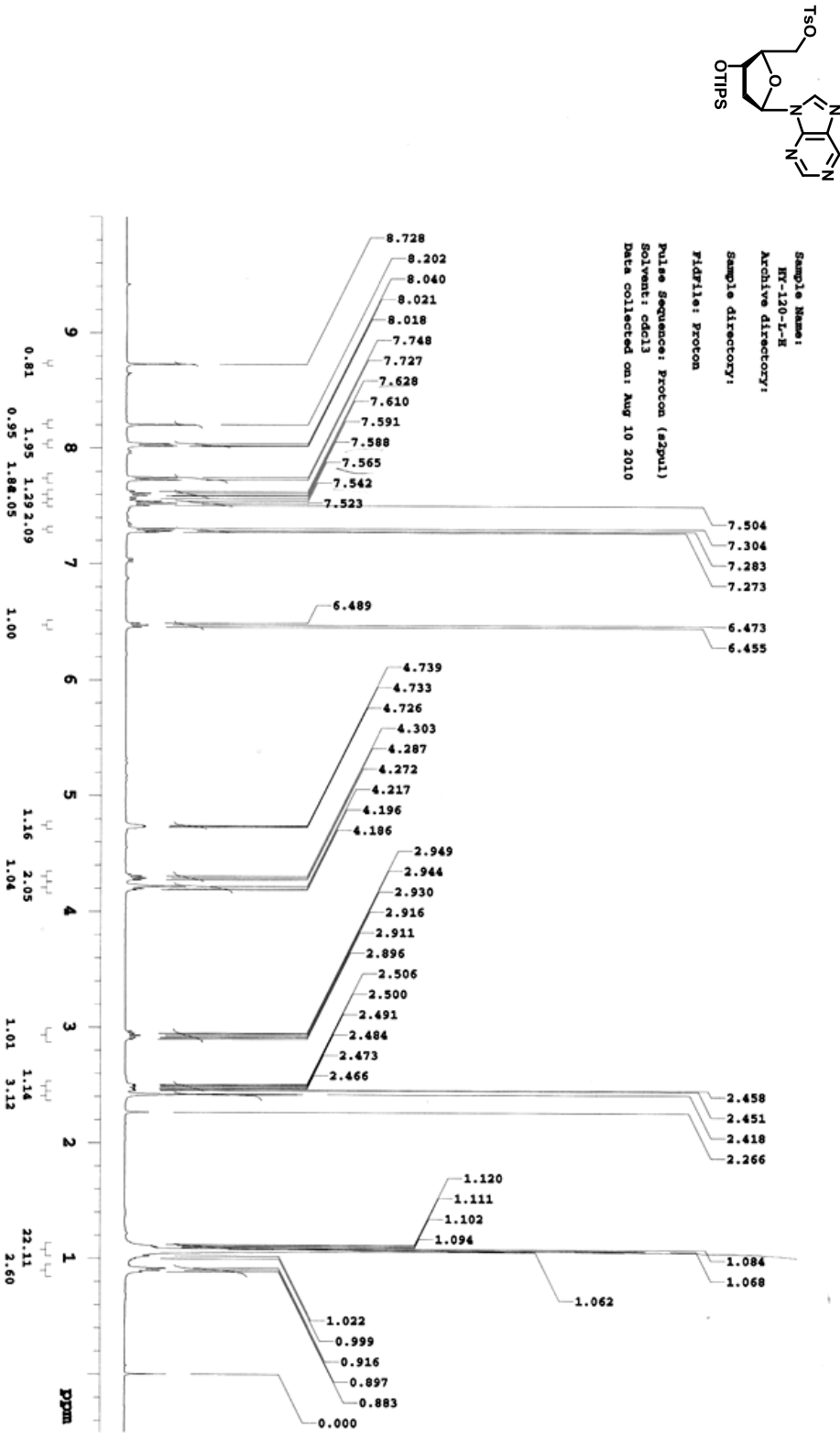
^{1}H NMR of Compound **5**
Solvent: $CDCl_3$ at 400 MHz



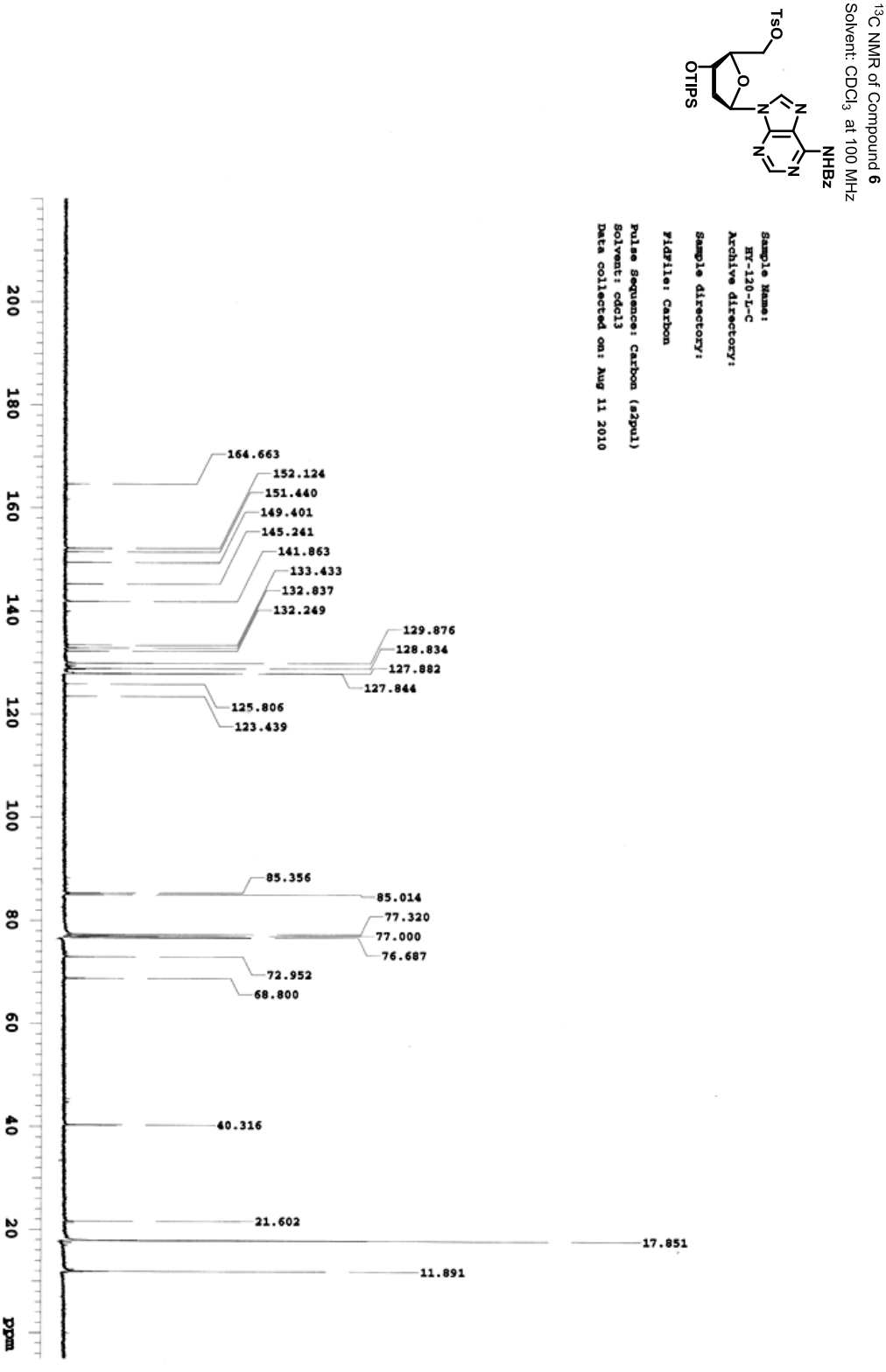
^{13}C NMR of Compound **5**
Solvent: $CDCl_3$ at 100 MHz



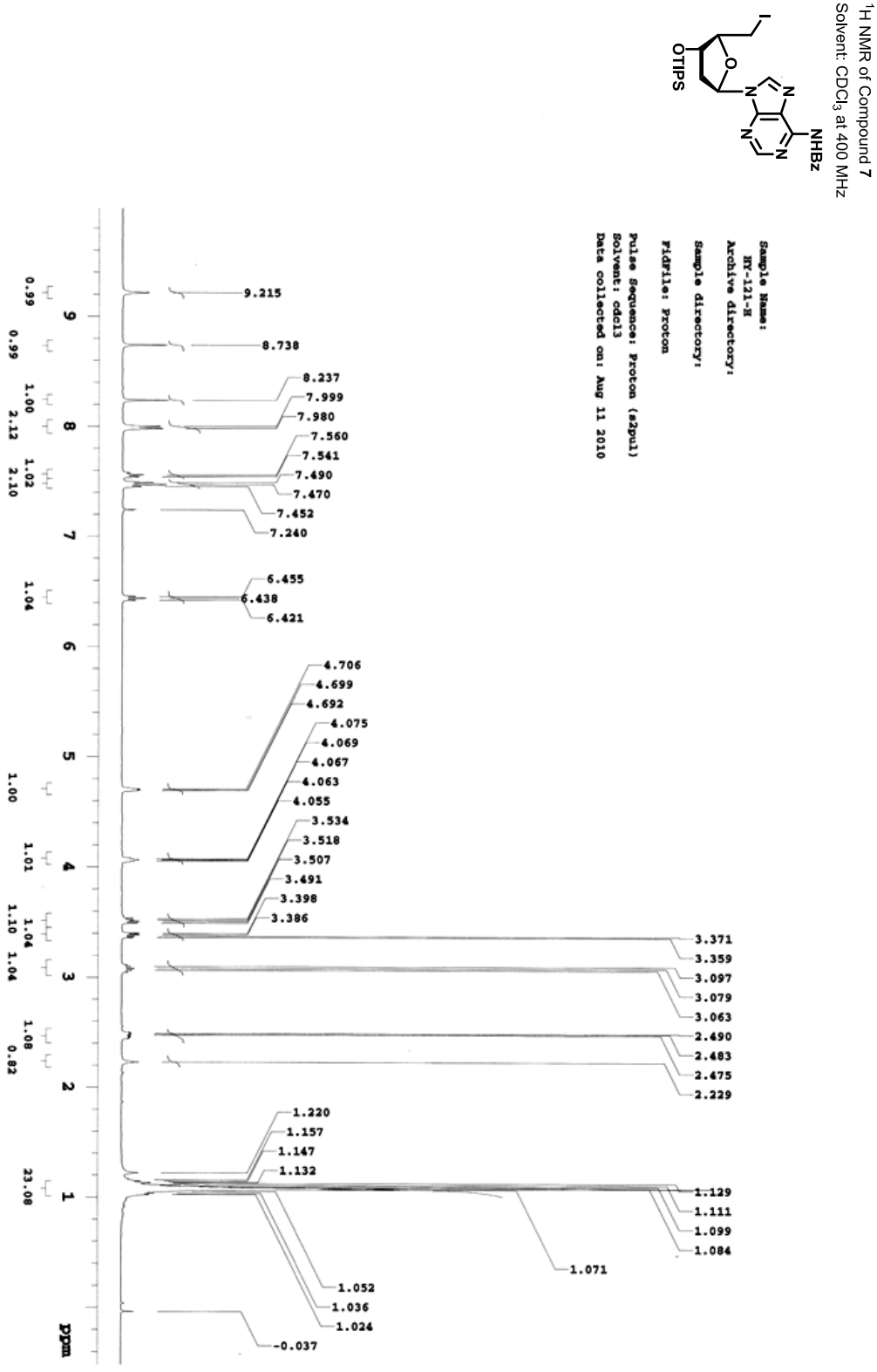
^{1}H NMR of Compound **6**
Solvent: $CDCl_3$ at 400 MHz



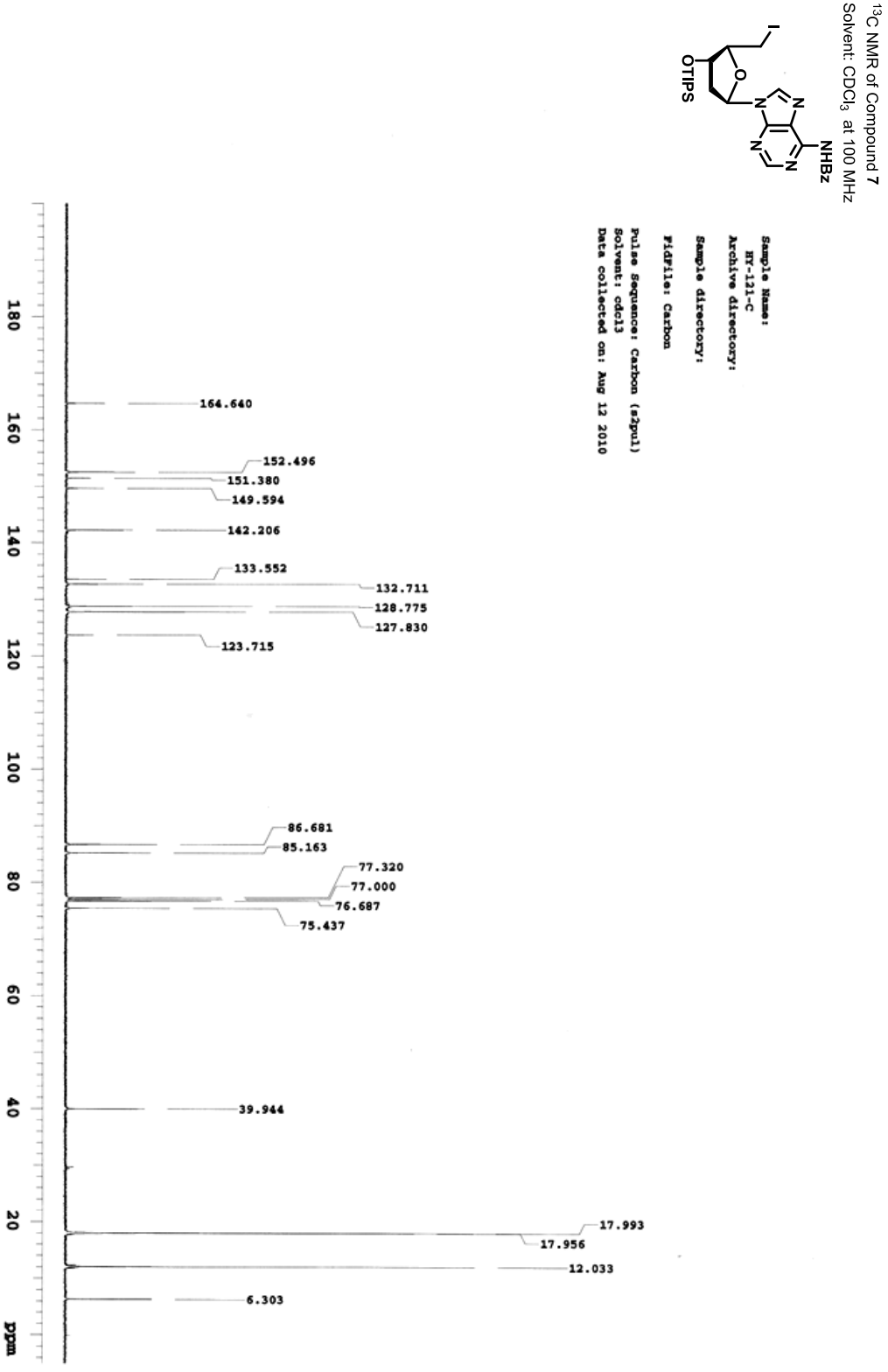
^{13}C NMR of Compound **6**
Solvent: $CDCl_3$ at 100 MHz



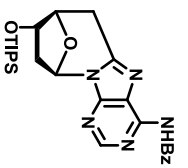
^{1}H NMR of Compound **7**
Solvent: $CDCl_3$ at 400 MHz

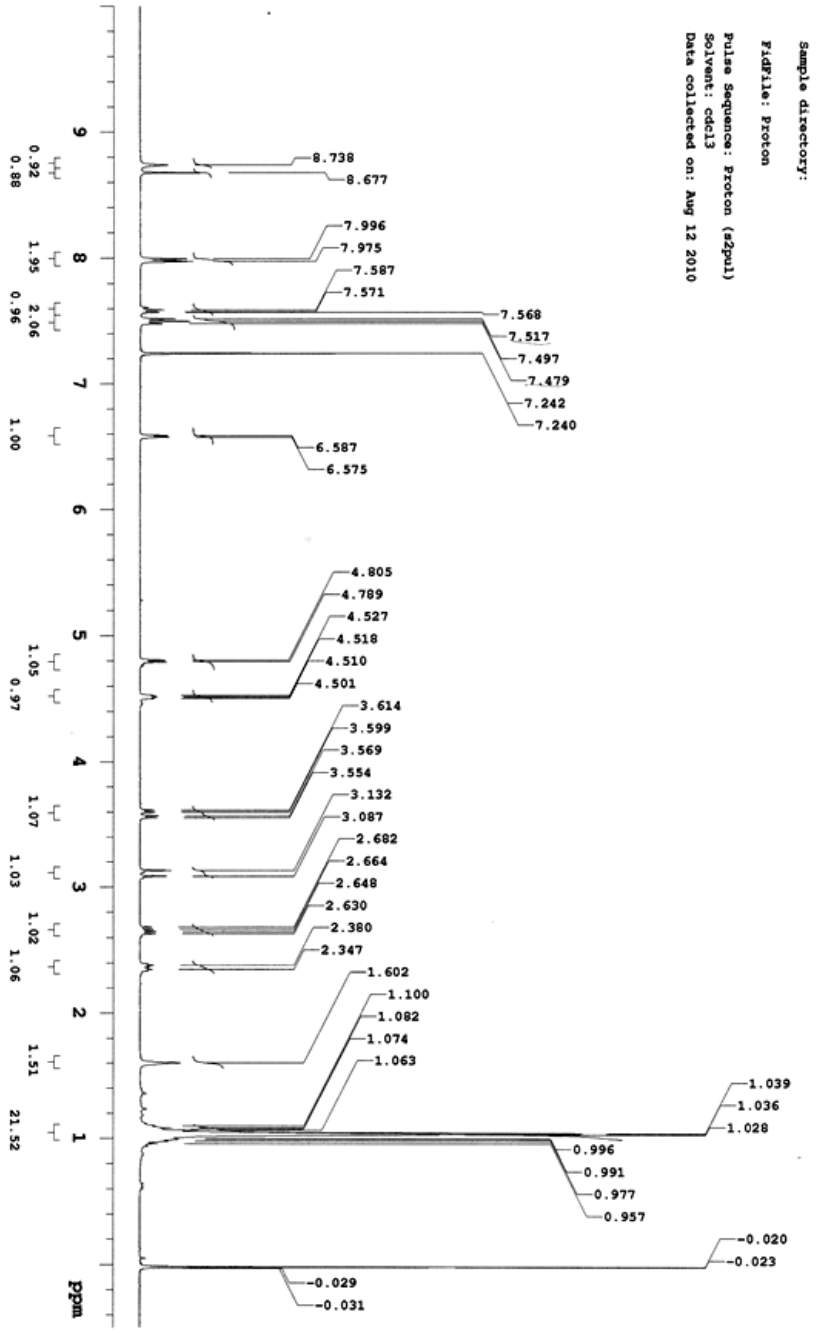


^{13}C NMR of Compound **7**
Solvent: $CDCl_3$ at 100 MHz

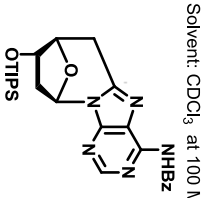


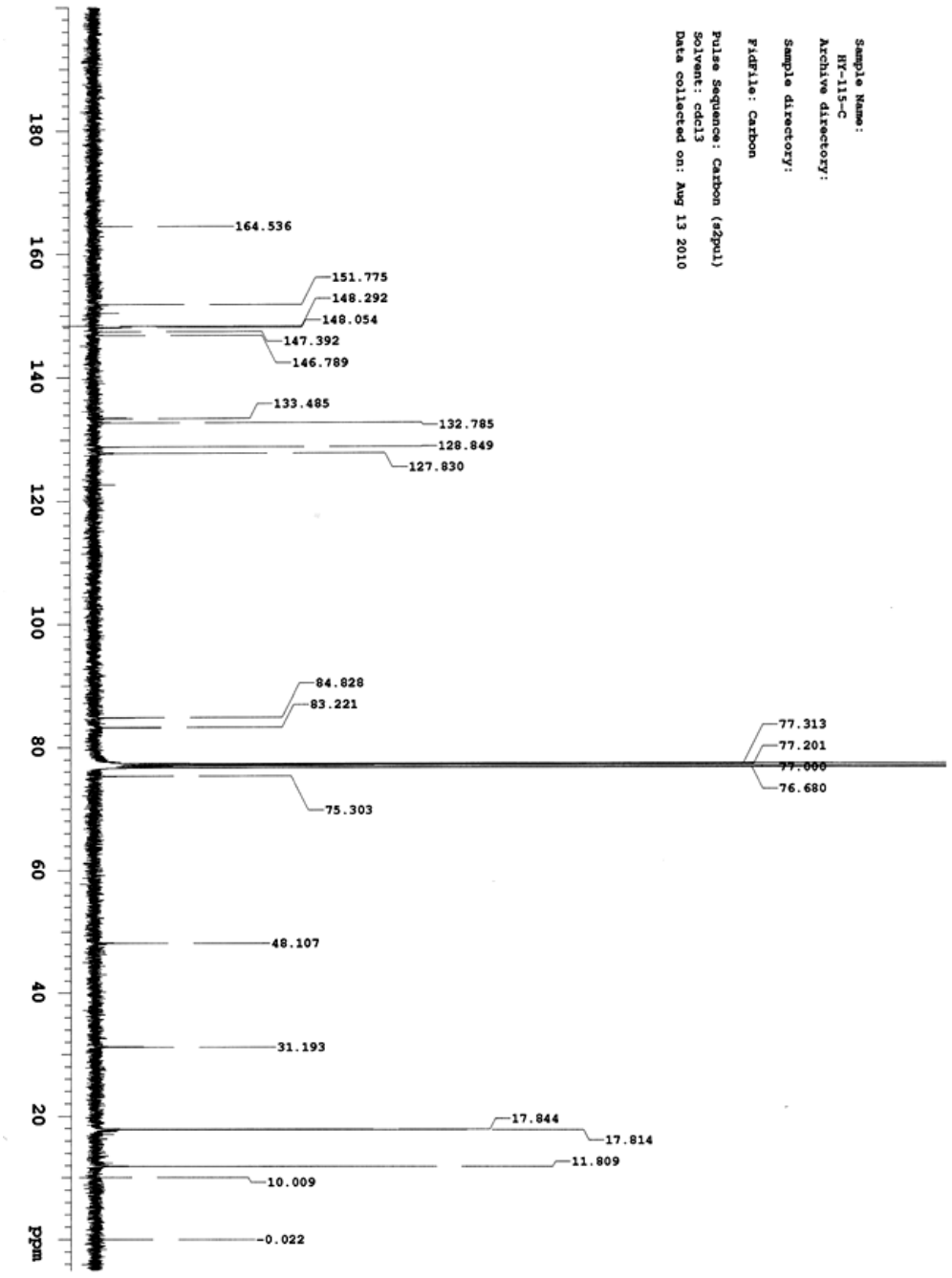
^{1}H NMR of Compound **8**
Solvent: $CDCl_3$ at 400 MHz



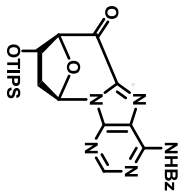


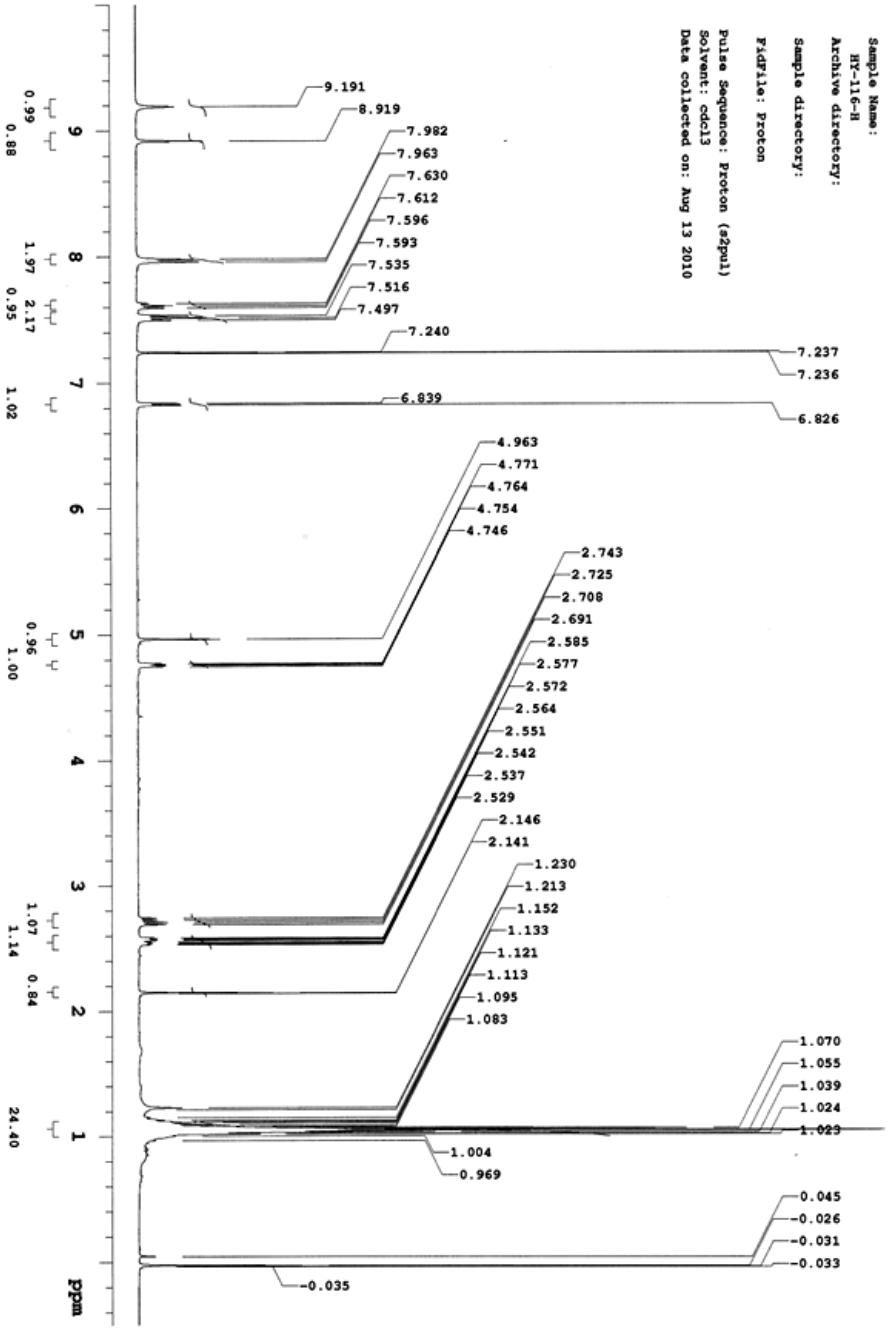
^{13}C NMR of Compound **8**
Solvent: $CDCl_3$ at 100 MHz



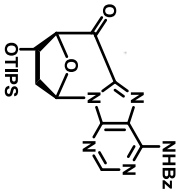


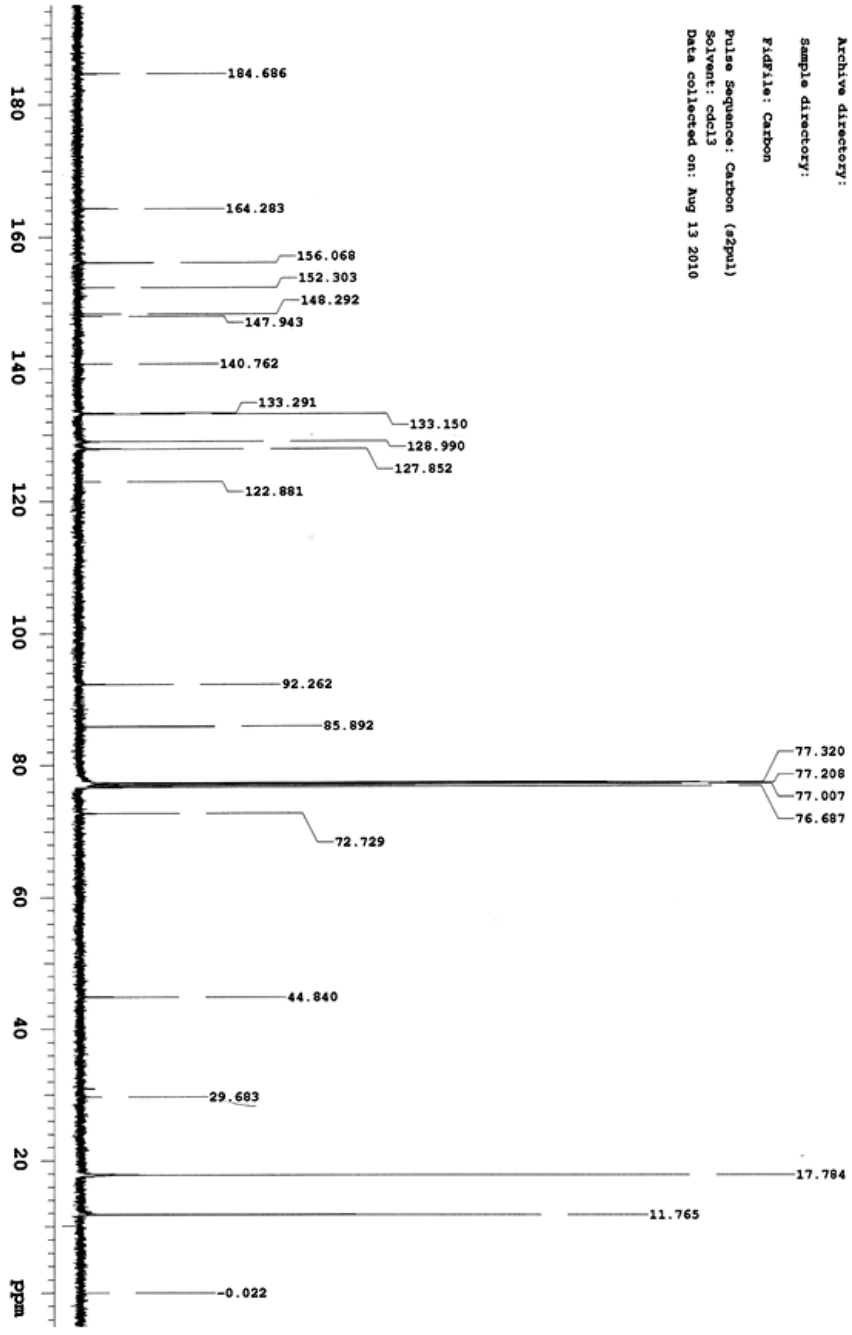
^{1}H NMR of Compound **9**
Solvent: $CDCl_3$ at 400 MHz



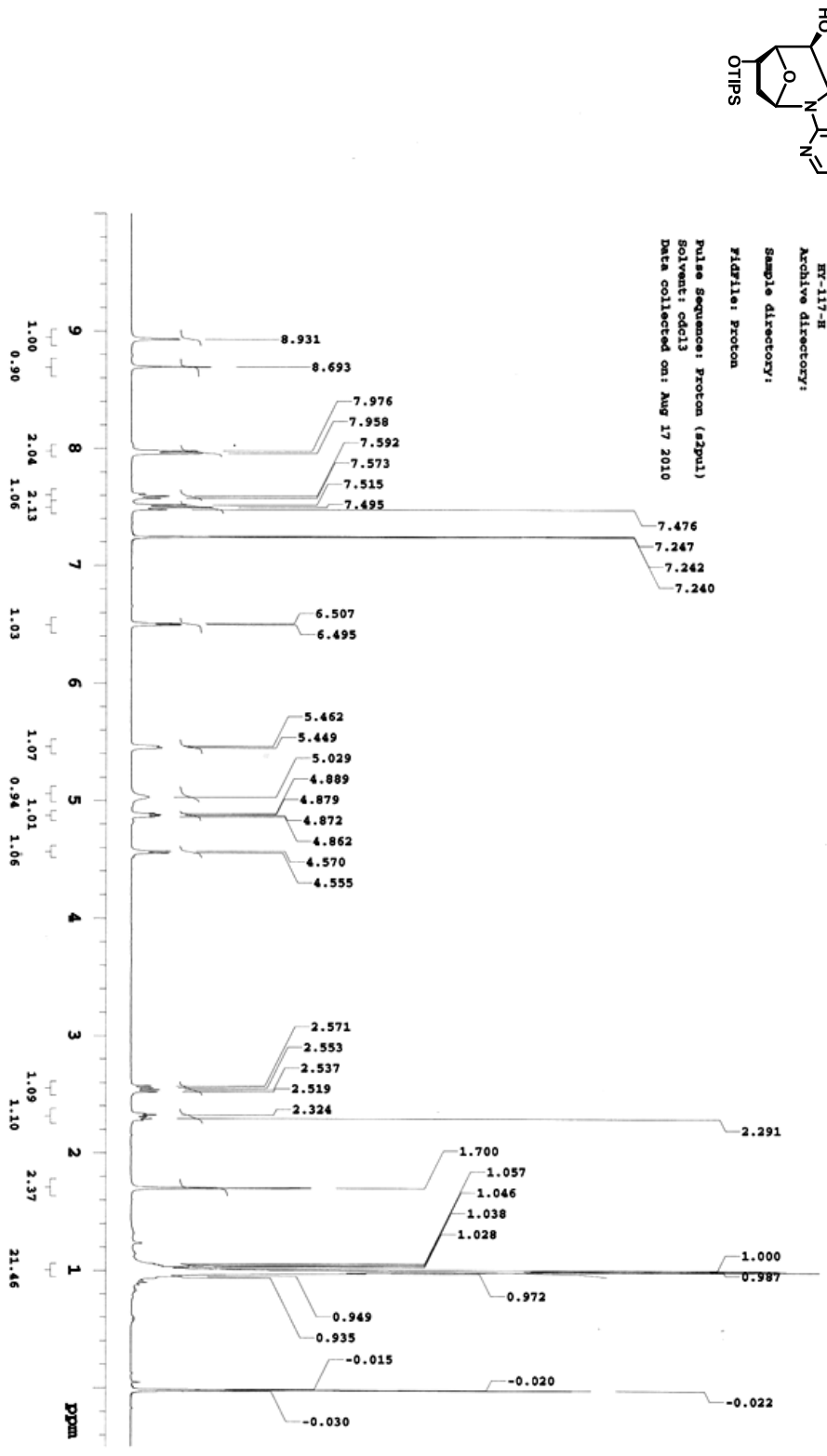


^{13}C NMR of Compound **9**
Solvent: $CDCl_3$ at 100 MHz

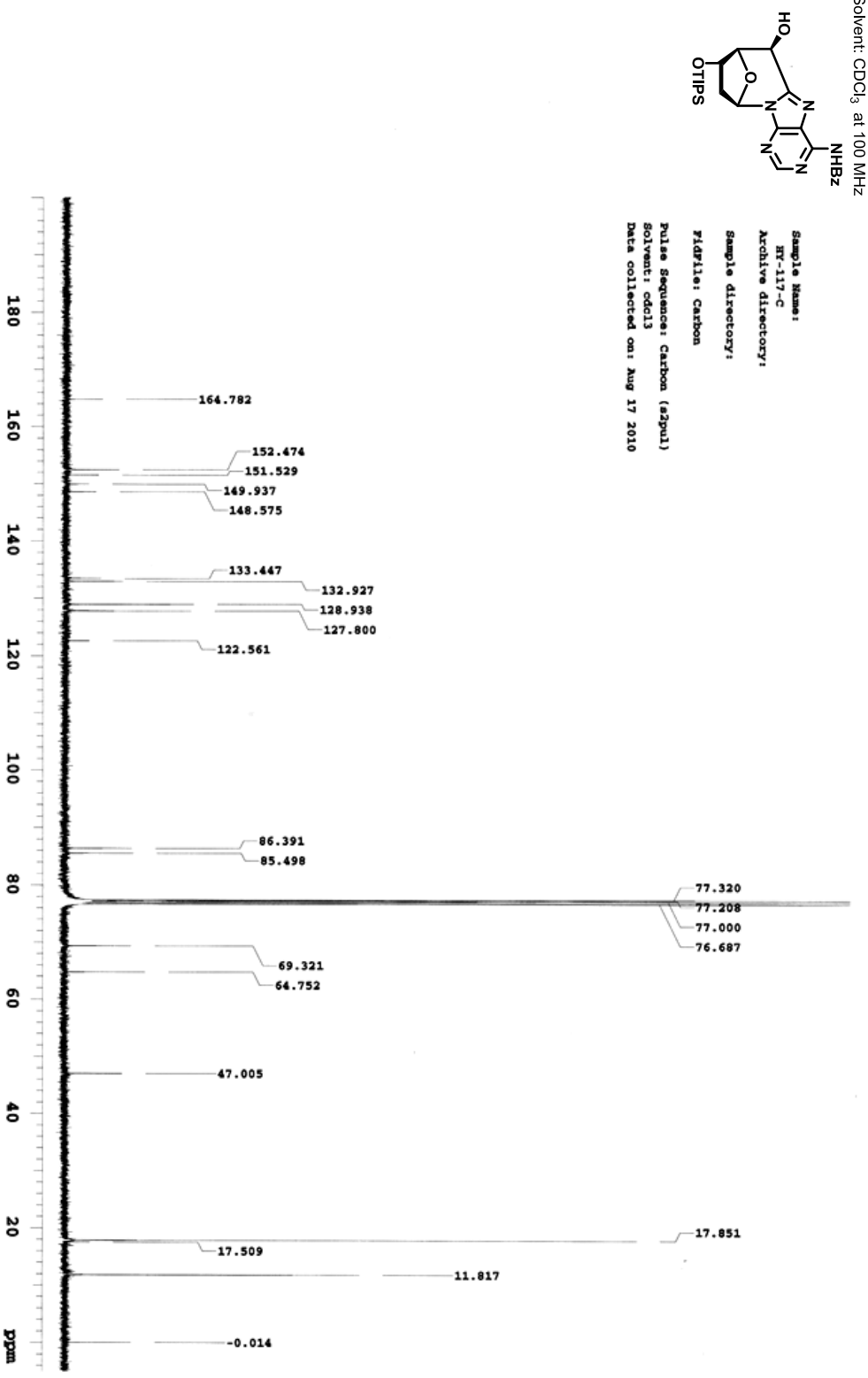




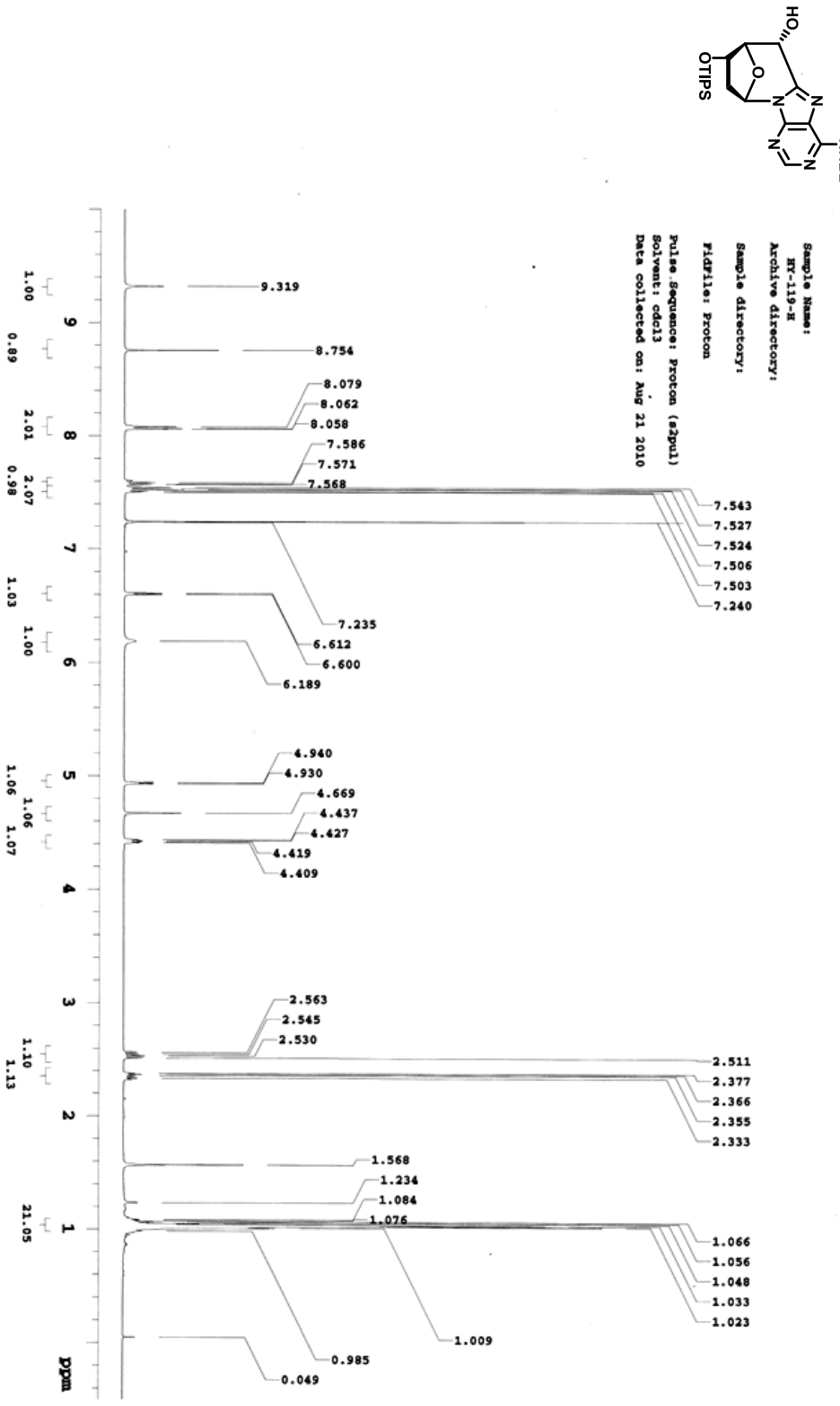
^{1}H NMR of Compound **10S**
Solvent: $CDCl_3$ at 400 MHz



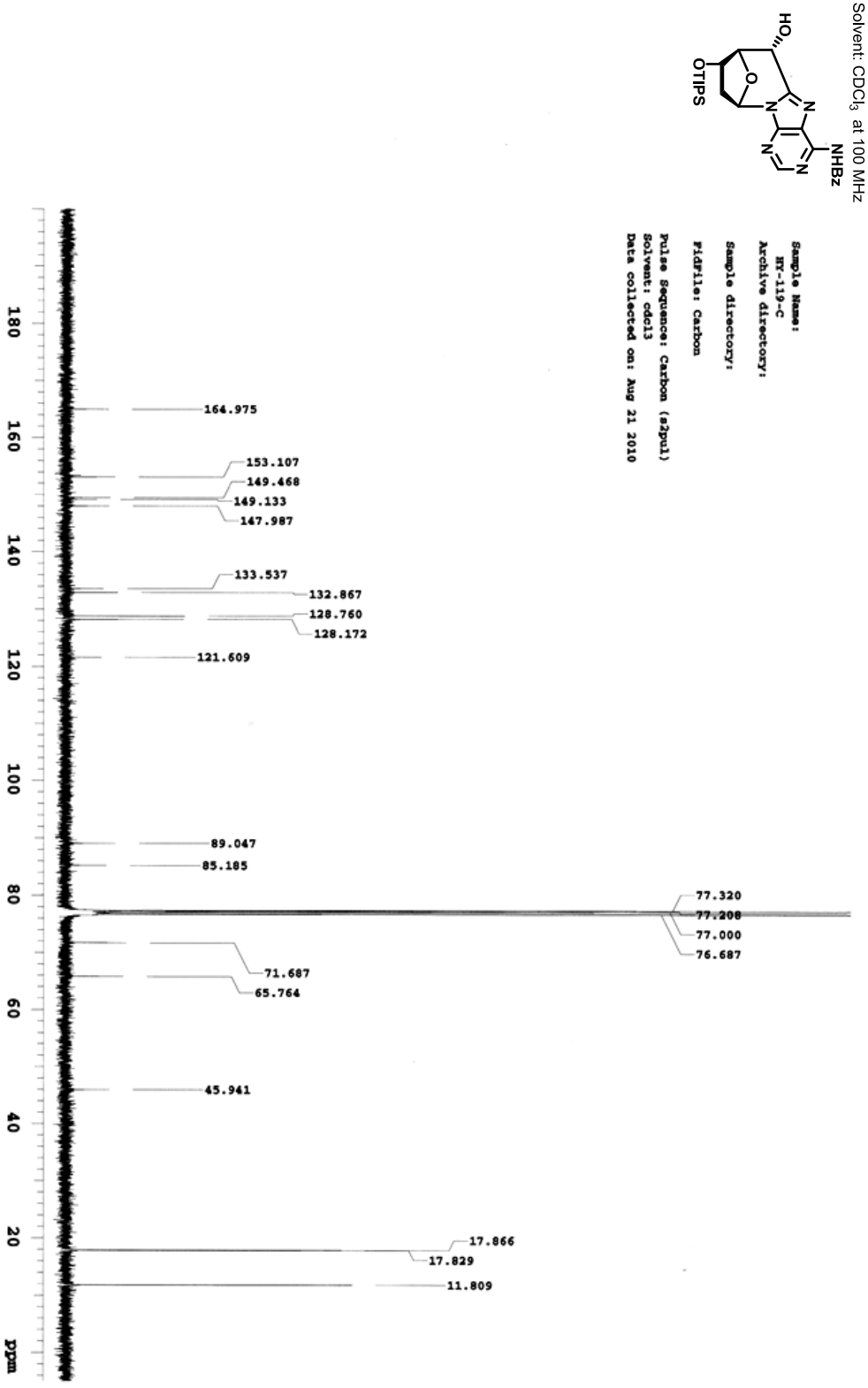
^{13}C NMR of Compound **10S**
Solvent: $CDCl_3$ at 100 MHz



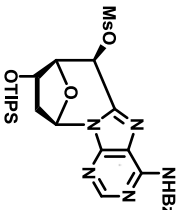
^{1}H NMR of Compound **10*R***
Solvent: $CDCl_3$ at 400 MHz

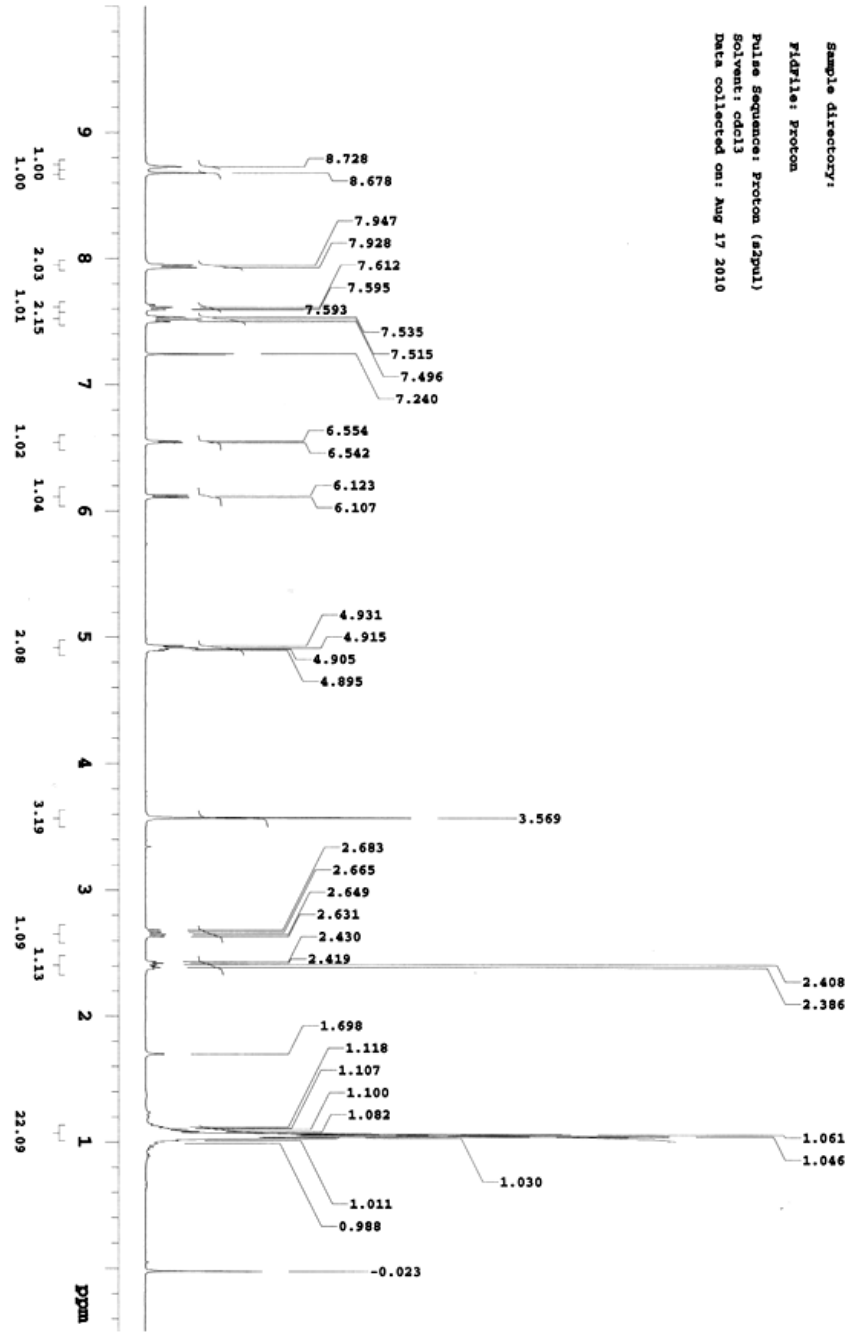


^{13}C NMR of Compound **10*R***
Solvent: $CDCl_3$ at 100 MHz

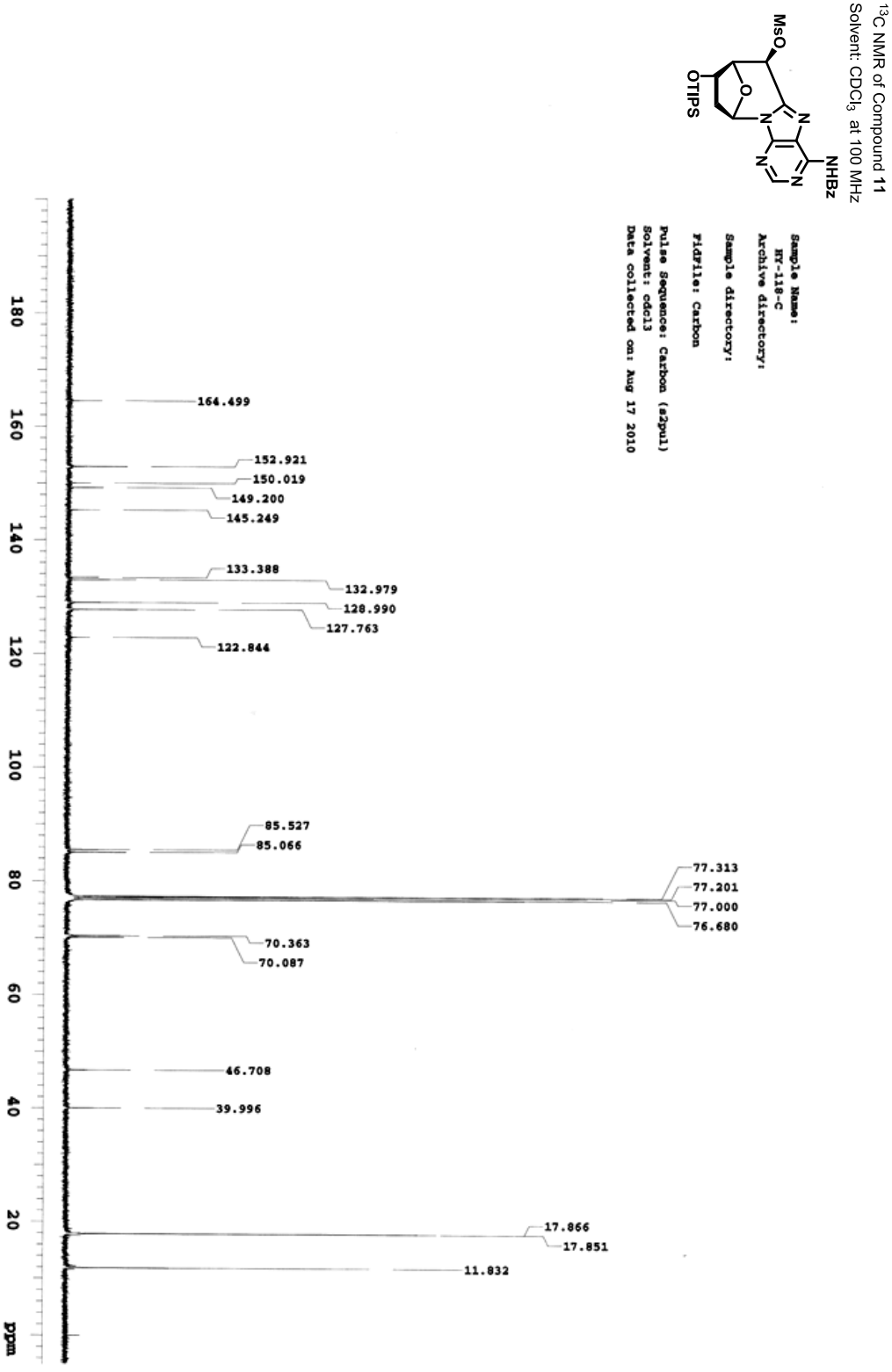


^{1}H NMR of Compound **11**
Solvent: $CDCl_3$ at 400 MHz

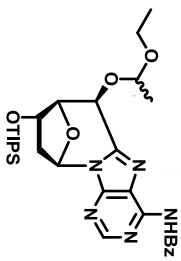


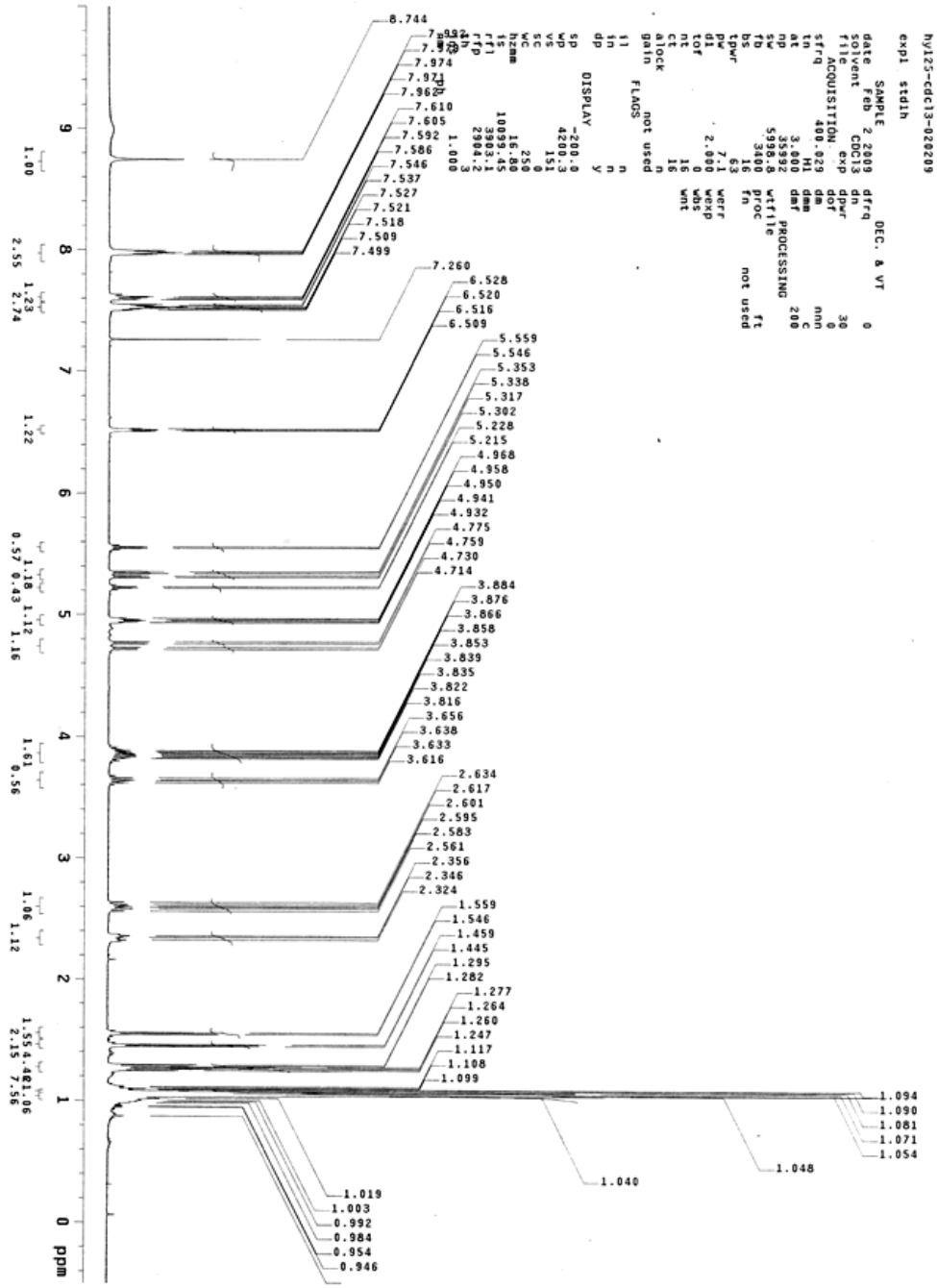


^{13}C NMR of Compound **11**
Solvent: $CDCl_3$ at 100 MHz

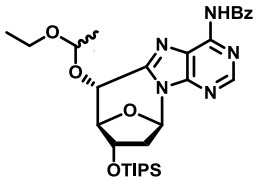


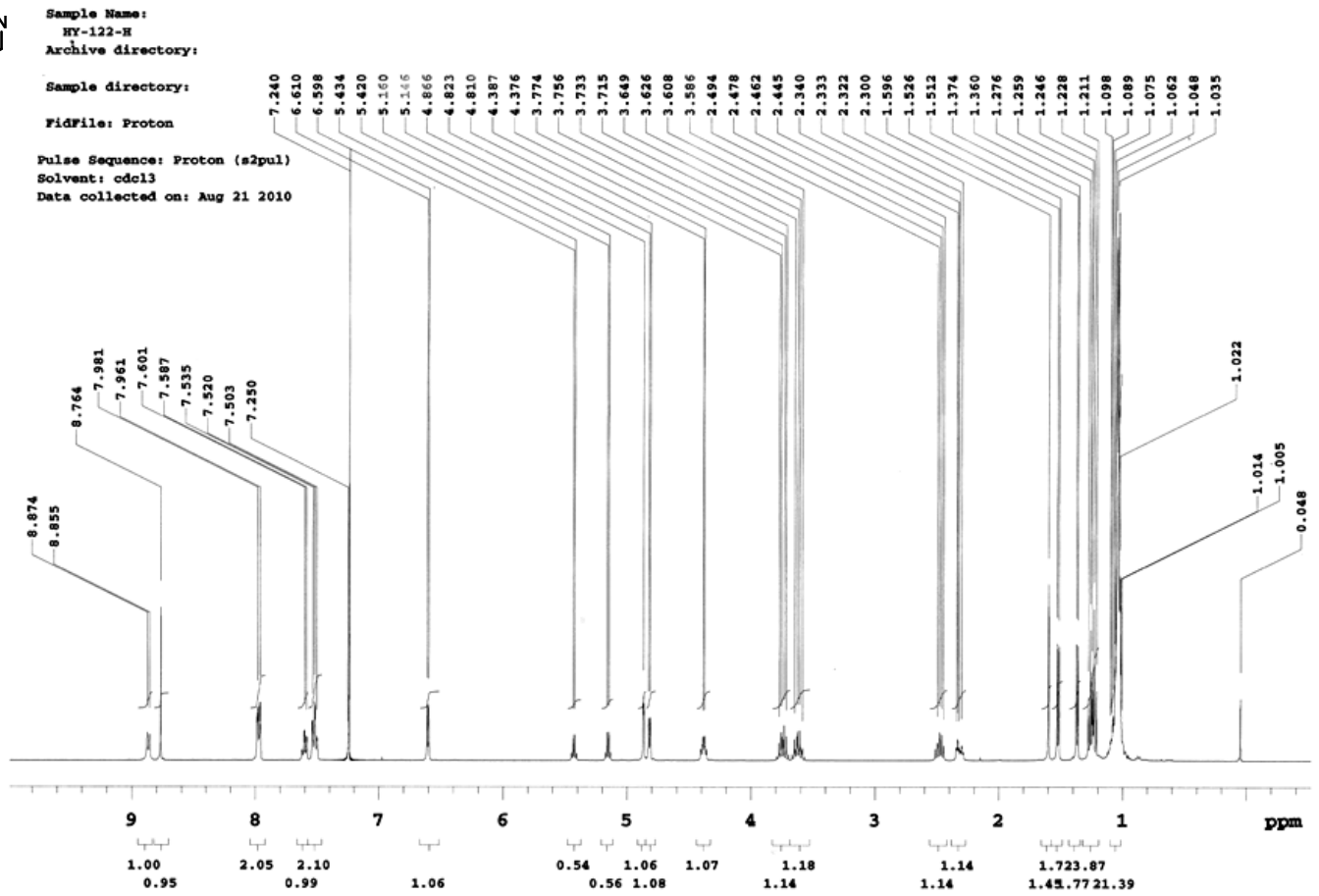
^{1}H NMR of Compound **12S**
Solvent: $CDCl_3$ at 400 MHz



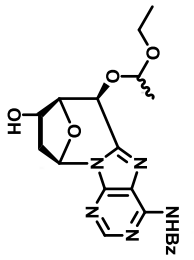


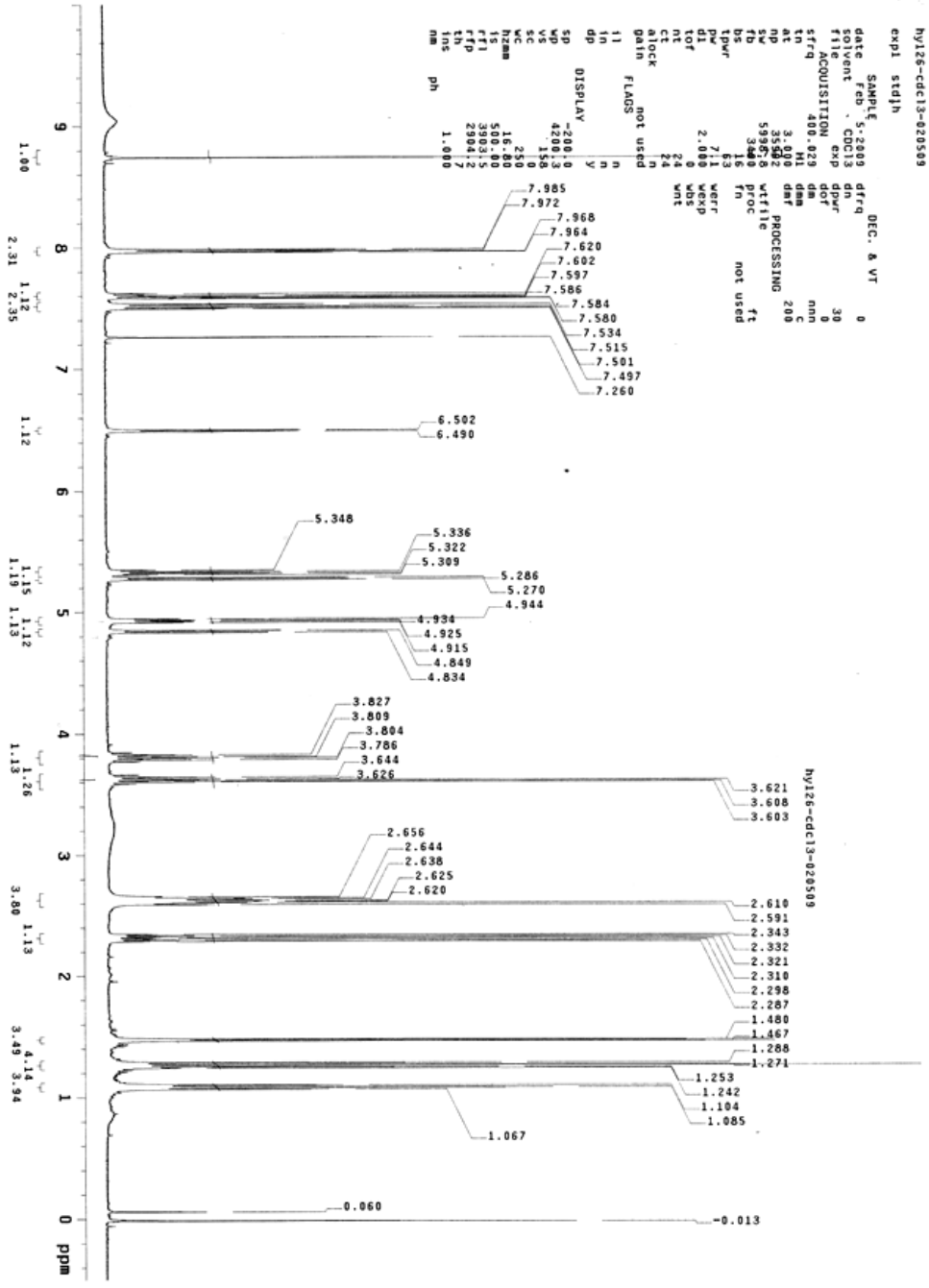
^{1}H NMR of Compound **12*R***
Solvent: $CDCl_3$ at 400 MHz



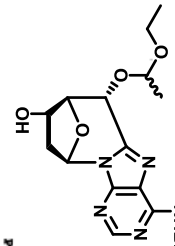


^{1}H NMR of Compound **13S**
Solvent: $CDCl_3$ at 400 MHz

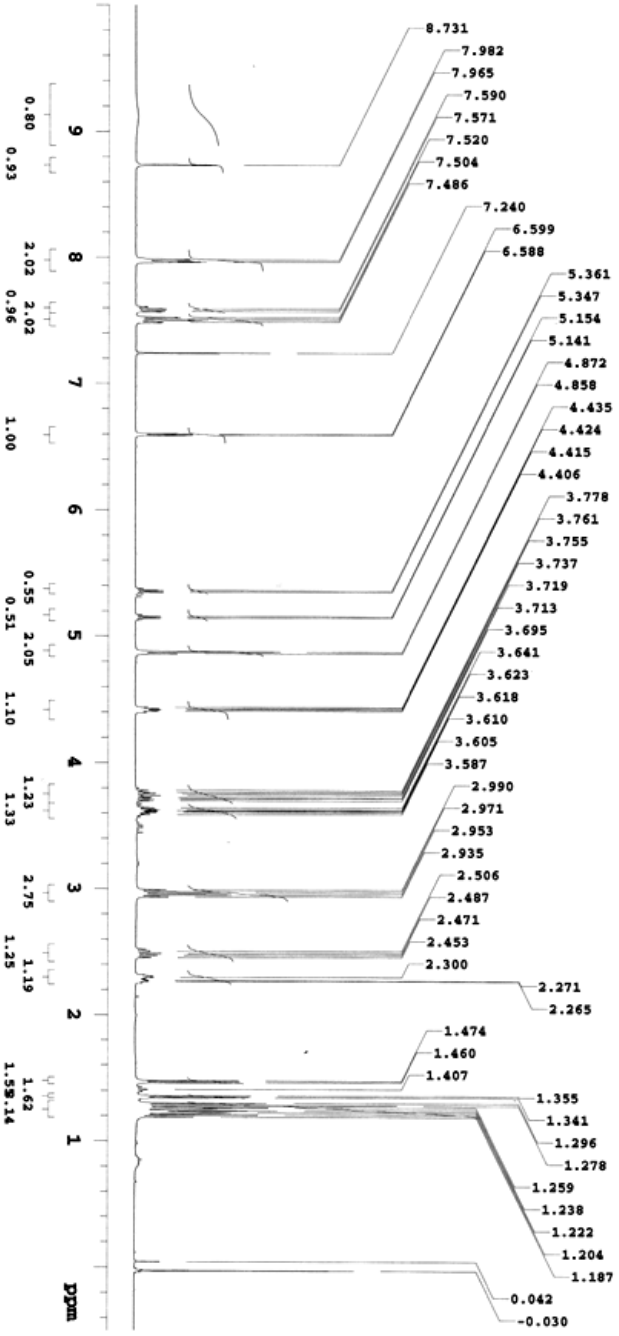




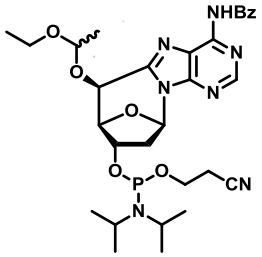
^{1}H NMR of Compound **13*R***
Solvent: $CDCl_3$ at 400 MHz



Sample Name:
RY-123-
Archive directory:
Sample directory:
FidFile: Proton
Pulse Sequence: Proton (s2pul)
Solvent: cdcl3
Data collected on: Aug 23 2010



^{31}P NMR of Compound **14*S***
Solvent: $CDCl_3$ at 162 MHz

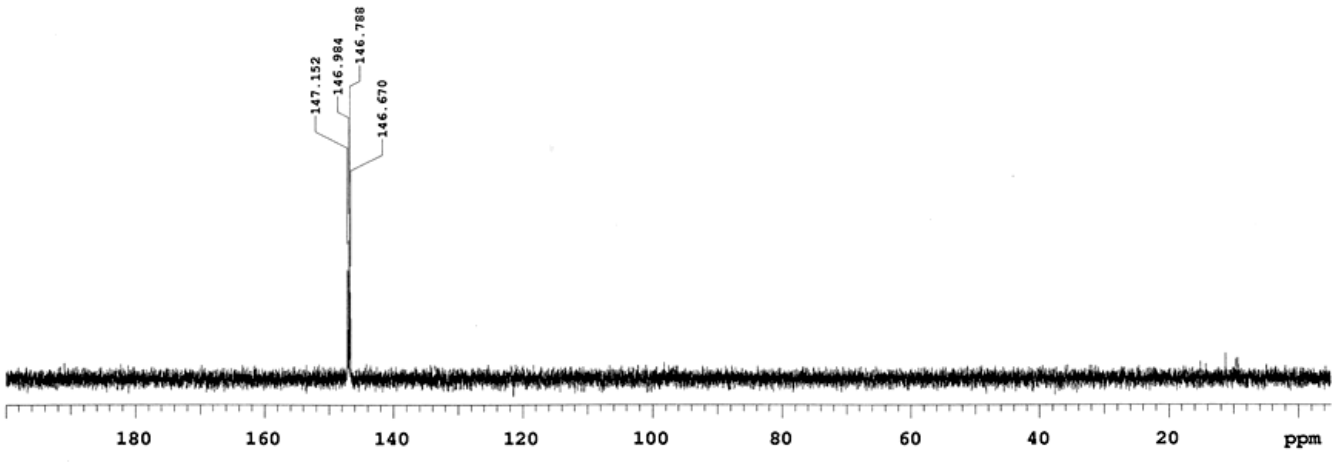


Sample Name:
HY-~~103-e~~ 127-P
Archive directory:

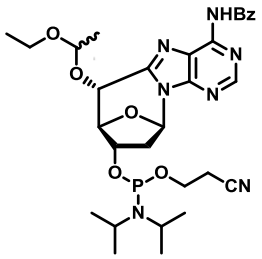
Sample directory:

FidFile: Phosphorus

Pulse Sequence: Phosphorus (s2pul)
Solvent: cdcl3
Data collected on: Aug 27 2010



^{31}P NMR of Compound **14*R***
Solvent: $CDCl_3$ at 162 MHz

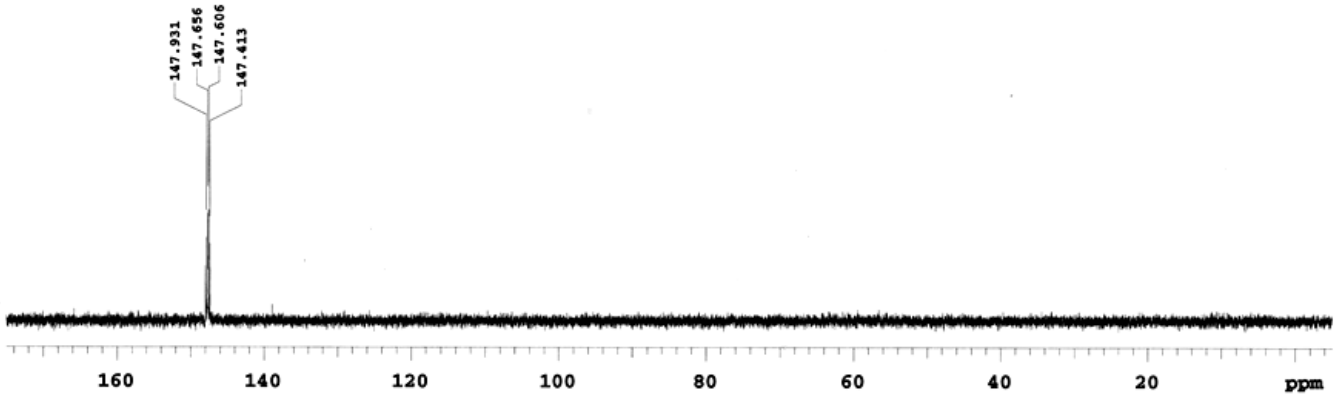


Sample Name:
HY-124-P-2
Archive directory:

Sample directory:

FidFile: Phosphorus

Pulse Sequence: Phosphorus (s2pul)
Solvent: cdcl3
Data collected on: Aug 26 2010



Chapter 3

Synthesis and Characterization of Ring-Expanded Cyclo-2'-Deoxynucleic Acids

3. Synthesis and Characterization of Ring-Expanded Cyclo-2'-Deoxynucleic Acids

3.1. Introduction

Chemical modification of nucleic acids has been drawing a lot of attention for many decades due to the potential for treatment of serious diseases such as viral infections, cancer and genetic disorders. The modified nucleic acid analogues are designed with the goal of effectively and selectively targeting single stranded DNA and RNA. [1-5]

DNA composed of the nucleosides containing rigidified sugars (e.g. Locked Nucleic Acids[6,7]) is generally more stable than the corresponding native sequences.[8,9] This effect likely reflects a more favorable entropic effect in the energetics of helix formation since the rigidified sugars have some degree of preorganization toward the conformation present in the DNA duplex. Previously we synthesized 8,5'-cyclo-2'-deoxyadenosine and 6,5'-cyclo-2'-deoxyuridine (1*S*, 1*R* and 2*S*, 2*R* in Figure 3.1).[10] These cyclo-nucleosides can be formed through oxidative damage to DNA. It is known that cyclo-nucleosides inhibit binding of proteins such as the TATA binding protein.[11] These rigidified cyclo-nucleosides were incorporated into DNA oligomers to study whether such species inhibited protein binding by deforming the backbone, or by enhanced binding to the complimentary strand. Thermal melting studies showed that the sequences with these rigid nucleosides did

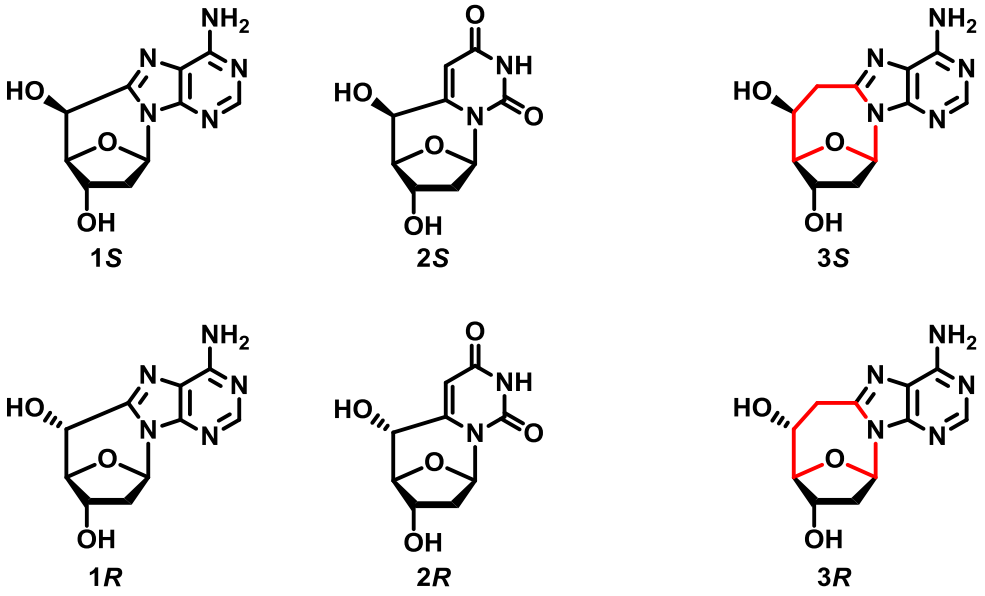


Figure 3.1. The first generation cyclo-deoxynucleosides: 8,5'-cyclo-2'-deoxyadenosines (**1*S***, **1*R***) and 6,5'-cyclo-2'-deoxyuridines (**2*S***, **2*R***); the novel ring-expanded-8,5'-cyclo-deoxynucleosides (**3*S*** and **3*R***).

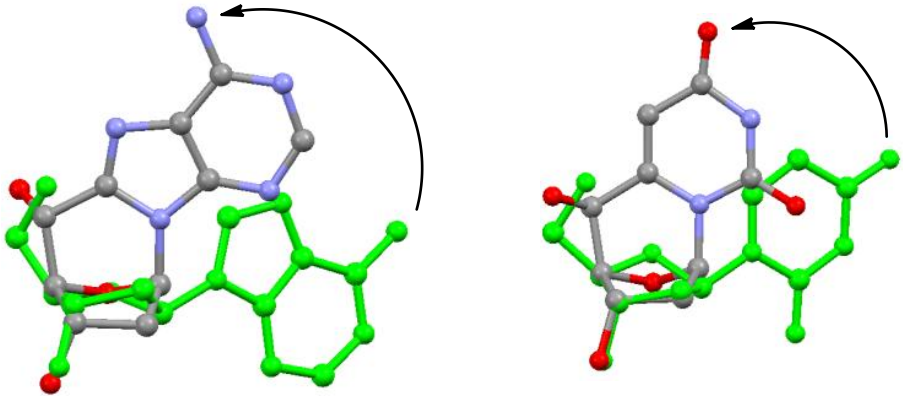


Figure 3.2. The crystal structures of **1*R***, with protecting groups hidden, and 2*S* overlayed with 2'-dA and 2'-dU (green structures).[10,12,13] The extra linkage pulls the base away from the hydrogen bonding center.

not form stable duplexes, likely the reasoning behind the reduced binding affinity of the TBP.

The crystal structures of the cyclo-nucleosides (cyclo-dA and cyclo-dU) are one crucial piece of evidence to explain this phenomenon.[10] The extra linkage between the sugar and base "locked" the modified nucleoside into a pre-organized structure and restricted the χ-angle into a fixed position as we designed, but it also "pulled" the base away from its natural position, causing duplex destabilization (Figure 3.2).

Here we present the synthesis of a new class of cyclo-deoxynucleosides: the ring-expanded cyclo-deoxynucleosides. Our theory is that by introducing an extra methylene onto the cyclized linkage, we could "push" the base back towards its natural position. The linkage should also decrease entropic rotation around the glycosidic bond. The ribose version of the ring-expanded adenosine analogues has been synthesized before, but the chirality on the C5' position has never been solved.[14] In this chapter we not only successfully synthesized both diastereomers of ring-expanded cyclo-deoxyadenosine but clarified the stereochemistry.

3.2. Synthesis of Ring-Expanded-8,5'-Cyclo-2'-Deoxyadenosine

The goal of this study was to prepare both diastereomers of Ring-Expanded-8,5'-Cyclo-2'-Deoxyadenosine based upon Cyclo-Deoxyadenosine.

Introducing an extra methylene group into the bridge between the sugar and the heterocycle of cyclo-dA would rigidify the nucleosides and "push" the nucleobase back to the helical center. Model building studies suggested that the *S*- diastereomer would result in similar glycosidic bond angles and closer space orientation on the sugar and base to the dA observed in double-stranded nucleic acids[15] (Figure 3.3).

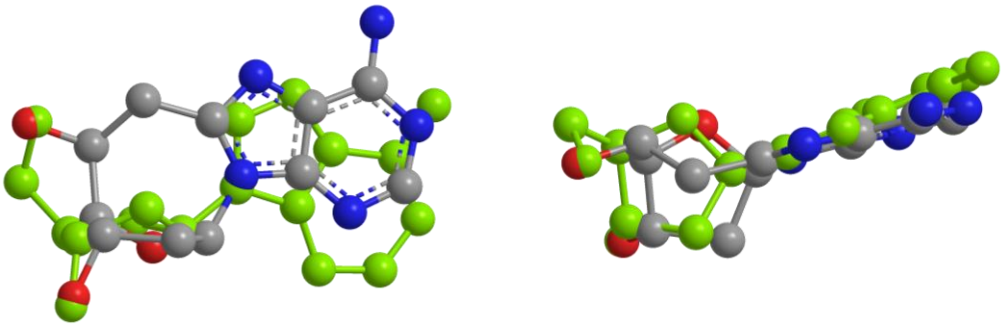
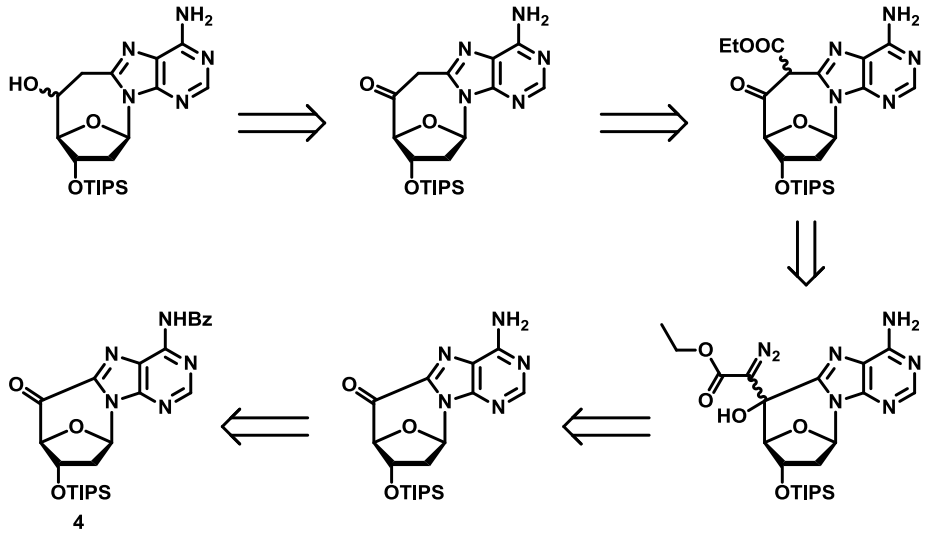


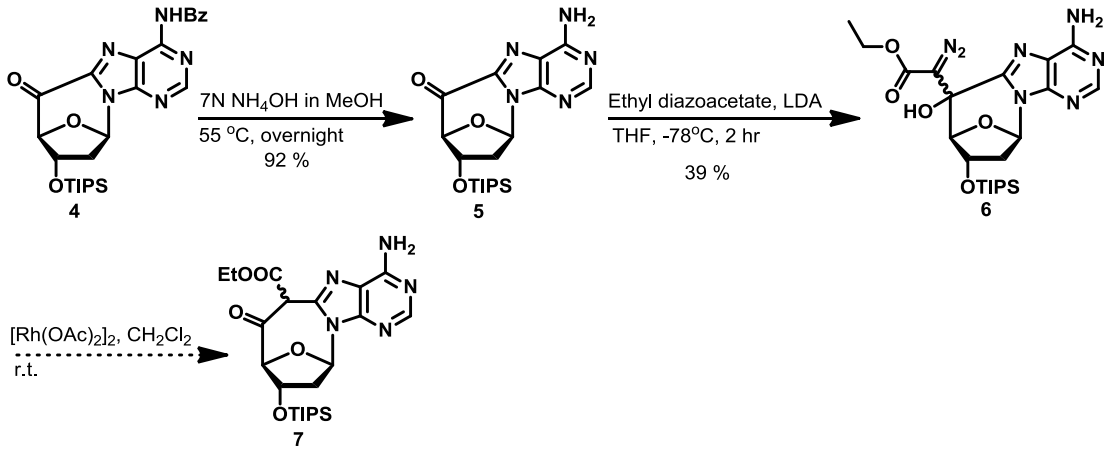
Figure 3.3. Molecular modeling of **3*S*** and dA (green structure).[11] The extra methylene linkage restricts the base rotation against the glycosidic bond and allows the hydrogen binding sites to move closer to the helical center.

The first synthetic strategy we envisioned was going through the rhodium-catalyzed dediazotization/rearrangement studied by Pellicciari and Nagao[16,17]. The ring-expanded cyclo-nucleosides could be formed by reducing the ring-expanded cyclized-ketone compound. The ketone compound was generated by rhodium-catalyzed rearrangement on an important intermediate (**4**) from the synthesis of cyclo-dA then followed with decarboxylation (Scheme 3.1).[10,18]

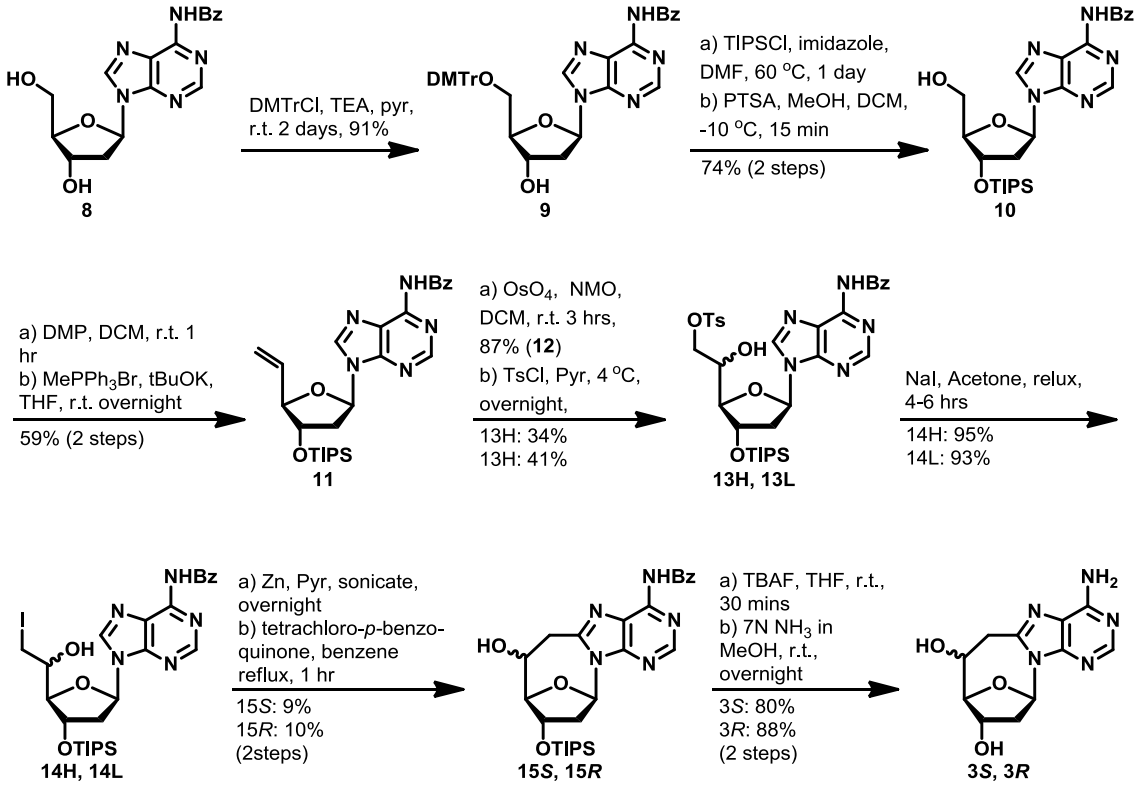


Scheme 3.1. Retro-synthesis of the ring-expanded 8,5'-cyclo-2'-deoxy-adenosine by rhodium-catalyzed rearrangement.

The synthesis started with compound **4** (ref. 10) that was treated with 7N ammonium hydroxide in methanol to remove the *N*-benzoyl group, thus avoiding a side reaction in which the ethyl diazoacetate would react with the benzoylamide in the next step. Then ethyl diazoacetate was reacted with freshly made LDA (Lithium diisopropylamide) at -78 °C to generate the lithiated ethyl α-diazoacetate intermediate. This was added to compound **5** which resulted in diastereomeric alcohols, compounds **6*S*** and **6*R*** which were characterized by proton NMR and high-resolution mass. Unfortunately, the rearrangement never succeeded, even when the amino group was re-protected as an imino group prior to the rhodium-catalyzed reaction (Scheme 3.2).



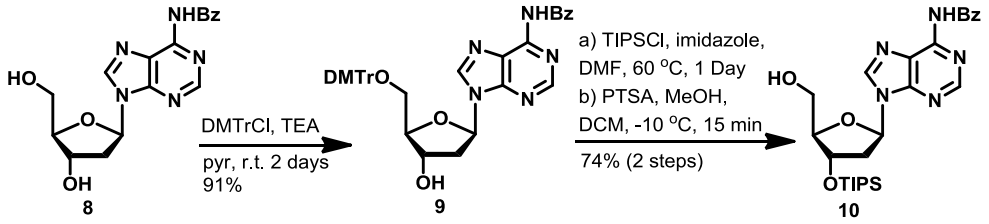
Scheme 3.2. Synthetic scheme following the rhodium-catalyzed rearrangement strategy.



Scheme 3.3. Synthesis of the *R* and *S* diastereomers of ring expanded-8,5'-cyclo-2'-deoxyadenosine.

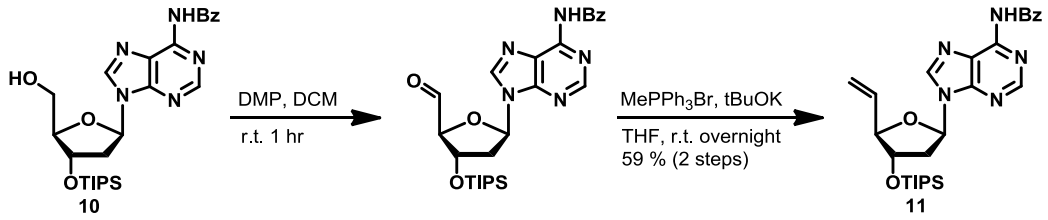
We therefore revised our strategy toward the 8,5'-cyclo-2'-deoxyadenosine synthesis.[10] The idea was to insert an iodomethylene group onto the 5' carbon of the sugar, followed by a zinc mediated cyclization reaction to generate the ring-expanded nucleoside.

The synthesis (Scheme 3.3) started with *N*-benzoylated 2'-deoxyadenosine[10] (compound **8**) to perform a "transient protection", protecting the 3'-OH with a silyl protecting group but leaving the 5'-OH free. Firstly we treated compound **8** with DMTrCl to selectively protect the primary 5'-OH with a bulky, acid labile group to get compound **9** with a very good yield. The trityl group could fall off during the silica gel chromatography, so the elution solvent had to contain 1% TEA. The next step was to protect the 3'-OH with the triisopropylsilyl (TIPS) protecting group. Then using the mild acid, *p*-toluenesulfonic acid, at low temperature we could remove the acid labile DMTr group from the 5'-O position while avoiding de-purination to obtain compound **10**.



Scheme 3.4. Synthesis of compound **9** and compound **10**.

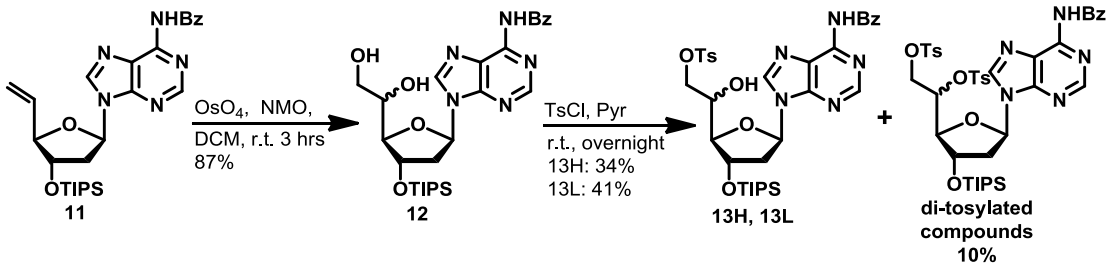
The next step was to attach the 6' carbon on to the sugar. Compound **10** was oxidized by the self-prepared Dess-Martin periodinane[19] reagent to generate the 5'-aldehyde compound. This aldehyde compound did not have effective UV absorption, did not char well, and was found unstable under column purification conditions. We therefore performed the next Wittig reaction right after work-up to install the terminal olefin (compound **11**) with a 59% yield in 2 steps.



Scheme 3.5. Synthesis of compound **11**.

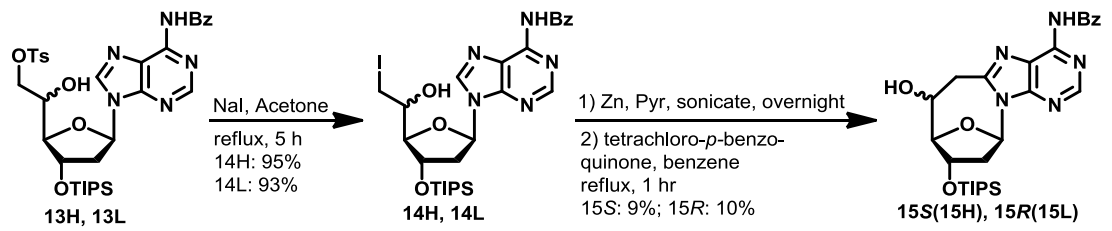
Compound **11** was then treated with a catalytic amount of osmium tetroxide to oxidize the terminal alkene into the racemic dihydroxy compound (**12**). These racemic compounds were inseparable, but the two diastereomers could be separated by column chromatography in the next step after putting the *p*- tosyl group on the 6'-OH. It was very hard to avoid the di-tosylated side products during this tosylation reaction; the low temperature condition (4 ^{o}C) helped to suppress the formation of these side products to about 10%. The separation on a silica gel column was also a big challenge. We were able to efficiently separate them using 2% IPA in DCM as an elution solvent. At

this stage, we named these two diastereomers as **13H** and **13L** according to their movement on the TLC plate, Higher or Lower.



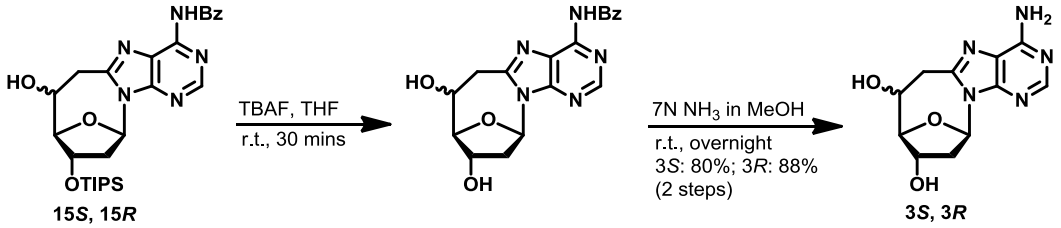
Scheme 3.6. Synthesis of compound **12** and compound **13**.

The next step was to convert the 6'-*O* –tosyl into a 6'-iodo group with very good yields for both diastereomers (compound **14H** and **14L**) using sodium iodide in acetone under reflux condition. The zinc-mediated cyclization was at first not performing well under the conditions used in the cyclo-dA synthesis.[6] We therefore performed the sonication induced cyclization reaction followed by tetrachloro-benzoquinone to regenerate the aromaticity of the purine base and get compounds **15*S*** and **15*R***. The chirality of the ring-expanded cyclo-dA could now be solved using 2D-NMR studies that will be discussed in section 3.3.



Scheme 3.7. Synthesis of compound **14** and compound **15**.

The free ring-expanded cyclo-nucleosides (**3*S*** and **3*R***) were obtained from **15*S*** and **15*R*** using with tetra-n-butylammonium fluoride (TBAF) in tetrahydrofuran (THF) to remove the triisopropylsily (TIPS) protecting group on the 3'-OH, followed by 7N ammonia in methanol to remove the benzoyl group on the 6-amino position. The final compounds (**3*S*** and **3*R***) were very polar, making it hard to purify them on a silica gel column, but they could be purified by HPLC using a reversed phase C_{18} column with water and acetonitrile as elution solvents.



Scheme 3.8. Synthesis of compound **3**.

3.3. Characterization of the Chirality

Since we introduced an extra methylene group on to the bridge of nucleobase and 5' carbon of sugar, there was a new chiral center generated, the C5'. It was a great challenge to solve and characterize these two diastereomers. Due to the synthetic strategy, two diastereomers were generated after the osmium tetroxide reaction (compound **12**). The two diastereomers could be separated after installing a tosyl group on the 6'-OH,

the stereochemistry remained unsolved until we generated and characterized the identity of cyclized compound **15**.

From the top view of the ring-expanded cyclo-deoxyadenosines (Figure 3.4) we knew that the distance between the 3'-H and 5'-H in *S* and *R* diastereomers would be very different, 3.75 Å in the *S*- diastereomer and 2.58 Å in the *R*- diastereomer. This difference allowed us to perform 2D NMR experiments to clarify the stereochemistry of these two compounds.

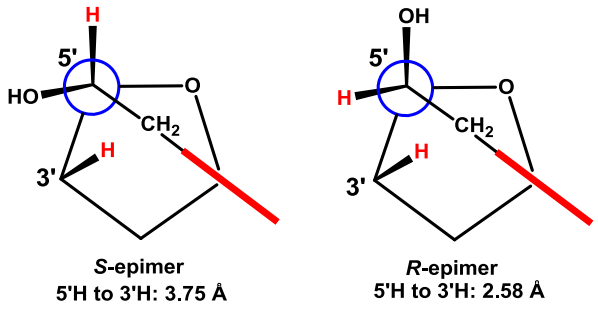


Figure 3.4. The top view for the two diastereomers of ring-expanded cyclo-deoxyadenosines. The distance between the 3'-H and 5'-H in *S* and *R* diastereomers are very different, 3.749 Å in *S*- diastereomer and 2.578 Å in *R*- diastereomer.

Firstly, we performed two-dimensional homonuclear correlation spectroscopy (COSY) experiments to identify the position of the protons on the sugar for both diastereomers (the COSY spectra are in Chapter 3.7). We then carried out through-space nuclear Overhauser effect spectroscopy (NOESY) experiments. Results of the NOESY experiments are showed in figure 3.5. The 2D spectrum of the ring-expanded cyclo-dA synthesized from **compound 13H** (**13H-->14H-->15H**) showed that the 5'-H only had interaction

with the 4'-H, but no interaction with 3'-H (Figure 3.5.a.); in the NOESY spectrum of the ring-expanded cyclo-dA from **compound 13L (13L-->14L-->15L)**, 5'-H had interactions with both 4'-H and 3'-H (Figure 3.5.b.).

The NOESY studies provided a clear indication that compound **15H** is the ***S* - diastereomer (15*S*)**, which has a longer distance between 3' and 5'-H and no interaction between these protons in the spectrum. Compound **15L** is the ***R* - diastereomer (15*R*)**, and has a shorter distance between 3' and 5'-H, showing an interaction in the NOESY experiment. The NOESY results also explained the movement of the compounds on the TLC plates. The 5'-OH of ***R* - diastereomer** is more accessible to the silica gel on the TLC which is more polar so moves slower (R_f: 0.24 ;DCM:MeOH = 91:9). It is the **compound 15L (15*R*)**. **Compound 15H (15*S*)** is the ***S* – diastereomer**, which has a less accessible 5'-OH and is less polar than ***R* - diastereomer**, so it moves faster on the TLC (R_f: 0.27 ;DCM:MeOH = 91:9).

After the last two deprotection steps (Scheme 3.8), the final compounds (**3*S*** and **3*R***) were purified on HPLC using a reversed phase C_{18} column. The mobile phase solvent A was water and solvent was B acetonitrile. The running gradient started with 100% A, in 25 minutes changed to 80% A and 20% B, then went back to 100% A in 2 minutes and the column was washed with pure solvent A (water) for 11 minutes to prepare it for the next run. The separation was very successful with the difference in retention time between two diastereomers being 3 minutes (Figure 3.6).

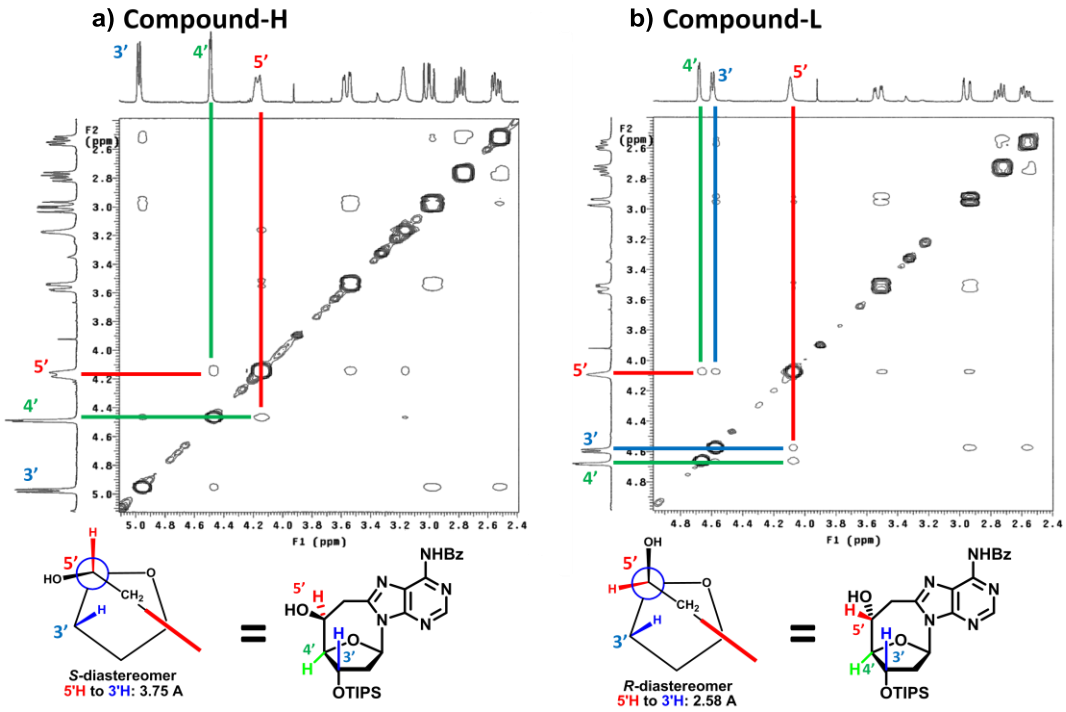


Figure 3.5. NOESY spectra for the ***S*** and ***R*** diastereomers of compound **16**. a) the compound H, which is the ***S*** diastereomer, with a longer distance between 3' and 5' protons; b) the compound L, the *R* diastereomers, with a shorter distance between 3' and 5' protons.

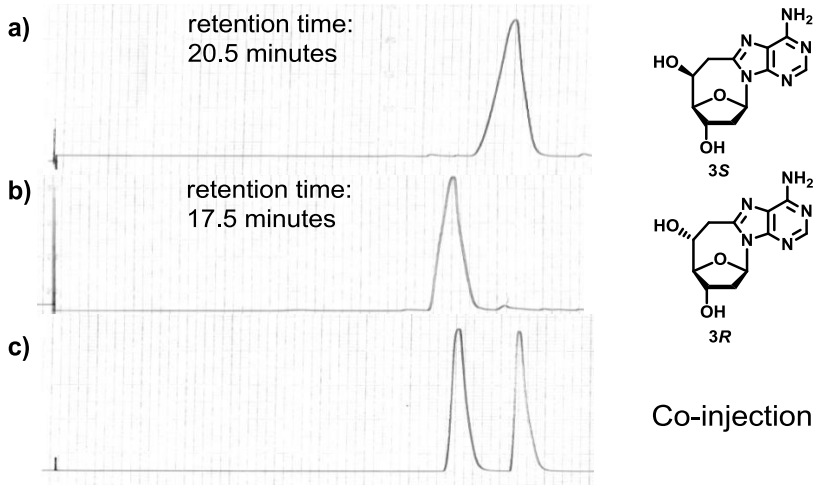


Figure 3.6. HPLC traces for compound **3*S*** and **3*R***. a) the ***S*** - diastereomer, b) the ***R*** – diastereomer, c) co-injection of two diastereomers. The ***S*** - diastereomer has lower polarity than the ***R*** - diastereomer which has longer retention time on the C-18 reversed phase column. The co-injection is mixing two diastereomers after purification.

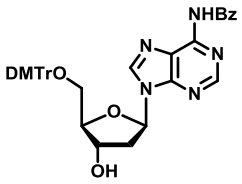
3.4. Conclusions

In this chapter, we presented an alternate design of the cyclo-nucleosides: the ring-expanded cyclo-deoxyadenosine. The synthetic strategy successfully inserted an extra methylene group onto the bridge of the cyclo-deoxyadenosine and obtained two novel ring-expanded cyclo-nucleosides. Using 2D NMR, we fully characterized these two diastereomers. The two free nucleosides of the ring-expanded cyclo-deoxyadenosine were purified by reversed phase HPLC and clarified the polarity difference correlated to the new generated chiral center on the C5' position. In the future, there will be more experiments carried out with these novel rigidified cyclo-nucleosides. They could have better thermal stability than the first generation of cyclo-nucleosides[10] in the DNA strands, or have the potential to perform enzymatic/non-enzymatic polymerization.

3.5. Experimental

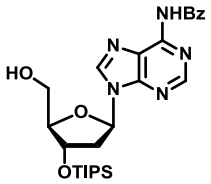
Reagents were purchased from Sigma-Aldrich, Acros, Oakwood, Glen Research, Lancaster, MP Biomedical, Chem-Impex International, and Fisher. Flash column were performed using Dynamic Adsorbents silica gel (60 Å, particle size 32-63 μm) and TLC monitoring with TLC Silica Gel with F-254 Indicator (Dynamic Adsorbents). TLCs were visualized by 260nm UV light and stained by 10% sulfuric acid. Dry tetrahydrofuran (THF), diethyl ether (Et_2O), *N*,*N*-dimethylformamide (DMF), pyridine (pyr), acetonitrile (MeCN), and dichloromethane (DCM) were obtained by passing commercially available pre-dried, oxygen-free formulations through activated alumina columns. Dry acetone was purchased from Sigma-Aldrich and used directly. NMR spectra were taken by Varian VNMRS400, VNMRS500, or INOVA 500 instruments and calibrated using residual undeuterated solvent ($CDCl_3$: δ H = 7.24 ppm, δ C = 77.00 ppm, D_2O: δ H = 4.80 ppm, Pyridine-D5: δ C = 135.91 ppm). Abbreviations of multiplicities were designated as follow: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. High-resolution mass spectra (HRMS) were recorded on a Waters LCT or JEOL AccuTOF mass spectrometer using ESI (electrospray ionization) or DART (direct analysis in real time). HPLC was performed on Waters (Milford, MA) Delta 600 controllers with a 2487 dual wavelength detector.

Compound 9



Compound 8 (7.10 g, 20.0 mmol) and 4,4'-dimethoxytritylchloride (8.13g, 24 mmol) were dissolved in dry pyridine (100 mL). Triethylamine (4.2 mL, 30 mmol) was added to the solution and the reaction mixture was kept under room temperature for 2 days. After the reaction had gone to completion (monitoring by TLC) water (30 mL) was added and the mixture was kept on ice for another 30 minutes. Solvents were removed under reduced pressure and the residue dissolved in dichloromethane and washed with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was purified by flash chromatography (DCM:MeOH:Et_3N = 97:2:1) to yield compound **9** as a white (12.0 g, 91%). TLC R_f: 0.41 (DCM:MeOH = 97:3). ^1H NMR (400 MHz, $CDCl_3$): δ 2.46 (m, 1H), 2.76 (m, 1H), 3.31 (d, 2H), 3.68 (s, 6H), 4.11 (q, J = 4.1 Hz, 1H), 4.61 (m, 1H), 6.41 (t, J = 6.4 Hz, 1H), 6.70 (dd, J = 8.8 Hz, 1.2 Hz, 4H), 7.10-7.21 (m, 7H), 7.30 (dd, J = 6.8 Hz, 1.6 Hz, 2H), 7.42 (t, J = 7.6 Hz, 2H), 7.51 (t, J = 7.2 Hz, 1H), 7.94 (dd, J = 7.2 Hz, 1.6 Hz, 2H), 8.09 (s, 1H), 8.63 (s, 1H). ^{13}C NMR (100MHz, $CDCl_3$): δ 40.4, 55.2, 63.7, 71.9, 84.8, 86.4, 113.1, 123.3, 126.9, 127.8, 127.9, 128.1, 128.8, 130.0, 132.7, 133.7, 135.6, 141.5, 144.5, 149.4, 151.3, 152.4, 158.5, 164.8. . HRMS (pos.): 658.2648 $[M+H]^+$ (HSMS calc. 658.2666).

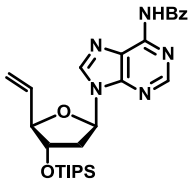
Compound 10



Compound **9** (5.60 g, 8.51 mmol), and imidazole (2.30 g, 33.4 mmol) were dissolved in dry DMF (56 mL) and cooled to 0 °C. Triisopropylsilyl chloride (4.70 ml, 22.0 mmol) was added and the reaction was warmed to ambient temperature. After stirring for 1 day the reaction was quenched with water while on an ice bath. 200 mL of ethyl acetate was added into the mixture, which was extracted with brine solution three times and once with water. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was maintained under high vacuum overnight, then dissolved in dichloromethane (85 mL) and incubated on an ice bath. The methanol (20 mL) solution of *p*-toluenesulfonic acid (0.69 g, 10.0 mmol) was added into the reaction mixture by drops. After stirring on the ice bath for half an hour, the reaction was quenched by saturated $NaHCO_{3(aq)}$ solution and extracted with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was purified by flash chromatography (DCM:MeOH = 97:3) to yield compound **10** as a white solid (3.5 g, 74%, two steps). TLC R_f: 0.25 (DCM:MeOH = 97:3). ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.03-1.10 (m, 21H), 2.29 (dd, J = 12.8 Hz, 5.6 Hz, 1H), 3.01 (m, 1H), 3.74 (d, J = 11.6 Hz, 1H), 3.93 (dd, J = 12.8 Hz, 1.6 Hz, 1H), 4.18 (s, 1H), 4.75 (d, J = 4.8 Hz, 1H), 6.34 (dd, J = 9.6 Hz, 5.6 1H), 7.46 (t, J = 7.2 Hz, 2H), 7.55 (t, J = 7.2 Hz, 1H), 7.98 (dd, J = 7.2 Hz,1.6 Hz, 2H), 8.08 (s, 1H),

8.70 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 11.9, 17.9, 41.7, 63.3, 74.1, 87.8, 90.6, 124.5, 127.9, 128.8, 132.8, 133.5, 142.6, 150.2, 150.7, 152.0, 164.7. HRMS (pos.): 512.2675 $[M+H]^+$ (HSMS calc. 512.2693).

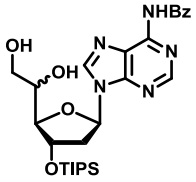
Compound 11



Compound **10** (5.60 g, 10.9 mmol), Dess-Martin periodinane (6.95 g, 16.4 mmol) were dissolved in dichloromethane (110 mL). The reaction mixture was stirring under room temperature for one hour then quenched by $Na_2S_2O_{3(aq)}$ and $NaHCO_{3(aq)}$ then extrcteded with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum to get the crude aldehyde compound. The methyl triphenyl phosphonium bromide (14.29 g, 40.0 mmol) and potassium tert-butoxide were dissolved in THF (140 mL) and stirred under room temperature for an hour. The crude aldehyde from the Dess-Martin oxidation was dissolved in THF (40 mL) then added into the Wittig reagents mixture and stirred for overnight. The reaction was quenched by $NH_4Cl_{(aq)}$ then extracted with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum to get the crude alkene compound. The crude product was purified by flash chromatography (Ethyl Acetate:Hexane = 60:40) to yield compound **11** as a white solid (3.28 g, 59%). TLC R_f: 0.38 (DCM:MeOH = 97:3). ^1H NMR (400 MHz, $CDCl_3$): δ 1.00-1.14 (m, 21H), 2.53

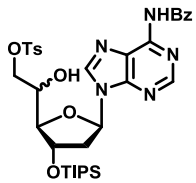
(m, 1H), 2.87 (m, 1H), 4.48 (m, 1H), 4.60 (m, 1H), 5.24 (d, J = 10.0 Hz 1H), 5.34 (d, J = 16.8 Hz 1H), 5.98 (m, 1H), 6.47 (t, J = 6.6Hz, 1H), 7.50 (t, J = 7.6 Hz, 2H), 7.58 (t, J = 7.4 Hz, 1H), 8.01 (d, J = 7.6 Hz, 2H), 8.15 (s, 1H), 8.78 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 12.0, 17.9, 40.2, 76.0, 84.9, 88.9, 118.0, 123.7, 127.8, 128.7, 132.6, 133.6, 135.5, 141.7, 149.5, 151.5, 152.5, 164.7. HRMS (pos.): 508.2730 $[M+H]^+$ (HSMS calc. 508.2744).

Compound 12



Compound **11** (4.08 g, 8.0 mmol), and *N*-methyl morpholine-*N*-oxide (2.49 g, 20.4 mmol) were dissolved in DCM (80 mL), then 1% osmium tetroxide aqueous solution (4.0 mL, 0.16 mmol) was added into the solution. After stirring under room temperature for 3 hours, the reaction was quenched by $Na_2S_2O_{3(aq)}$ and extracted with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was purified by flash chromatography (DCM:MeOH = 97:3) to yield the inseparable racemic compound **12** as a white solid (3.78 g, 87%). TLC R_f: 0.42 (DCM:MeOH = 97:3). HRMS (pos.): 542.2801 $[M+H]^+$ (HSMS calc. 542.2793).

Compound 13



Compound **12** (3.48 g, 6.4 mmol), and 4-dimethylaminopyridine (0.94 g, 7.7 mmol) were dissolved in

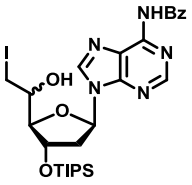
dry pyridine (30 mL) on ice bath, then 5 mL of *p*-toluenesulfonyl chloride (1.59 g, 8.3 mmol) pyridine solution was added into the mixture. The reaction was stirring in cold room (4 ^{o}C) for one day and checked by TLC. The reaction was quenched by water when it went completion, and extracted with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was purified by flash chromatography (DCM:IPA = 98:2) to yield the compound **13H** (1.51 g, 34%) and **13L** (1.836 g, 41%) as white solid.

13H: TLC R_f: 0.47 (DCM:IPA = 97.5:2.5). ^{1}H NMR (400 MHz, $CDCl_3$): δ 0.97-1.01 (m, 21H), 2.21 (dd, *J* = 12.8 Hz, 5.6 Hz, 1H), 2.32 (s, 3H), 2.85 (m, 1H), 3.95 (m, 3H), 4.16 (s, 1H), 4.67 (d, *J* = 4.8Hz, 1H), 6.27 (m, 2H), 7.20 (d, *J* = 8.0 Hz, 2H), 7.41 (t, *J* = 8.0 Hz, 2H), 7.51 (t, *J* = 7.6 Hz, 1H), 7.64 (d, *J* = 8.0 Hz, 2H), 7.93 (d, *J* = 7.6 Hz, 2H), 8.04 (s, 1H), 8.50 (s, 1H), 9.26 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 11.8, 17.8, 21.5, 41.2, 69.6, 69.7, 74.6, 87.7, 88.4, 124.5, 127.7, 127.8, 128.7, 129.6, 132.6, 132.8, 133.2, 142.5, 144.8, 150.3, 150.4, 151.6, 164.6. HRMS (pos.): 696.2873 $[M+H]^+$ (HSMS calc. 696.2887).

13L: TLC R_f: 0.38 (DCM:IPA = 97.5:2.5). ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.01-1.06 (m, 21H), 2.27 (dd, *J* = 13.2 Hz, 5.6 Hz, 1H), 2.39 (s, 3H), 3.01 (m, 1H), 4.04-4.13 (m, 3H), 4.72 (d, *J* = 4.4Hz, 1H), 6.08 (s, 1H), 6.29 (dd, *J* = 9.6 Hz, 5.6 Hz, 1H), 7.28 (d, *J* = 8.0 Hz, 2H), 7.46 (t, *J* = 7.6 Hz, 2H), 7.55 (t, *J* = 7.2

Hz, 1H), 7.74 (dd, J = 6.8 Hz, 1.6Hz, 2H), 7.98 (d, J = 7.2 Hz, 2H), 8.07 (s, 1H), 8.66 (s, 1H), 9.31 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 12.0, 17.9, 21.6, 40.8, 69.9, 70.1, 72.7, 87.1, 89.6, 124.4, 127.8, 127.9, 128.7, 129.8, 132.4, 132.8, 133.3, 142.7, 145.0, 150.2, 150.6, 151.9, 164.6. HRMS (pos.): 696.3272 $[M+H]^+$ (HSMS calc. 696.2887).

Compound 14



Compound **14H** and **14L** shared the same procedure as described below:

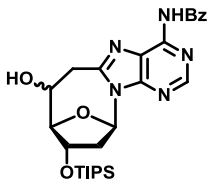
Compound **13** (**13H** or **13L,** 1.39 g, 2.0 mmol), and sodium iodide (1.50 g, 10.0 mmol) were dissolved in dry acetone (25 mL), then heated to reflux for 4 to 6 hours. The solid suspension was removed by celite filtration, and the filtrate was concentrated under rotary evaporator. The crude product was purified by flash chromatography (DCM:MeOH = 98:2) to yield the compound **14** (**14H:** 1.24 g, 95%; **14L**: 1.21 g, 93%) as white solid.

14H: TLC R_f: 0.28 (DCM:MeOH = 97:3). ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.06-1.11 (m, 21H), 2.28 (dd, J = 12.8 Hz, 5.6 Hz, 1H), 3.04 (m, 1H), 3.17 (t, J = 9.6 Hz, 1H), 3.27 (m, 1H), 3.97 (m, 1H), 4.64 (s, 1H), 4.74 (d, J = 4.4 Hz, 1H), 6.37 (dd, J = 5.6 Hz, 9.6 Hz, 1H), 6.56 (d, J = 12.4 Hz, 1H), 7.52 (t, J = 7.2 Hz, 2H), 7.61 (t, J = 7.2 Hz, 1H), 8.01 (d, J = 8.4 Hz, 2H), 8.07 (s, 1H), 8.75 (s, 1H), 8.97 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 6.1, 12.0, 18.0, 41.0, 72.7, 75.4, 88.2,

89.0, 124.7, 127.9, 129.0, 133.0, 133.4, 142.6, 150.4, 150.5, 151.9, 164.3. HRMS (pos.): 652.1815 $[M+H]^+$ (HSMS calc. 652.1811).

14L: TLC R_f: 0.25 (DCM:IPA = 97:3). ^1H NMR (400 MHz, $CDCl_3$): δ 1.09-1.14 (m, 21H), 2.33 (dd, *J* = 13.2 Hz, 5.2 Hz, 1H), 3.09 (m, 1H), 3.30 (m, 2H), 4.08 (m, 1H), 4.24 (d, *J* = 4.4 Hz, 1H), 5.93 (d, *J* = 2.8 Hz, 1H), 6.36 (m, 1H), 7.48 (t, *J* = 7.6 Hz, 2H), 7.57 (t, *J* = 7.6 Hz, 1H), 8.02 (d, *J* = 7.2 Hz, 2H), 8.12 (s, 1H), 8.74 (s, 1H), 9.57 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 7.1, 12.0, 17.9, 40.7, 72.7, 75.4, 87.0, 90.8, 124.3, 127.9, 128.6, 132.6, 133.3, 142.7, 150.1, 150.6, 151.9, 164.8. HRMS (pos.): 652.1826 $[M+H]^+$ (HSMS calc. 652.1811).

Compound 15



Compound **15*S* (15H)** and **15*R* (15L)** shared the same procedure as described below:

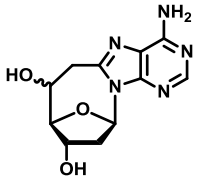
Compound **14** (**14H** or **14L,** 0.8 g, 1.2 mmol) and zinc powder (1.6 g, 24.0 mmol) were dissolved in dry pyridine (40 mL) and sonicated for overnight. The zinc powder was removed by filtration and the filtrate was concentrated under reduced pressure. The residue and tetrachloro-1,4-benzoquinone (0.27 g, 1.1 mmol) were dissolved in benzene (60 mL) and refluxed for 1 hour. The solvent was then evaporated under vacuum and the crude product was purified by flash chromatography (DCM/IPA 97:3) to yield

the cyclized product as a pale-yellow foam (**15*S***: 57.0 mg, 9%; **15*R***: 63.0 mg, 10%).

15*S*: TLC R_f: 0.16 (DCM:MeOH = 97:3). ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.05-1.11 (m, 21H), 2.52 (m, 1H), 2.77 (m, 1H), 2.98 (m, 1H), 3.15 (s, 1H), 3.54 (dd, *J* = 15.6 Hz, 3.2 Hz, 1H), 4.15 (d, *J* = 12.0 Hz, 1H), 4.47 (d, *J* = 3.6 Hz, 1H), 4.95 (d, *J* = 6.4 Hz, 1H), 6.85 (dd, *J* = 8.0 Hz, 2.0Hz, 1H), 7.47 (t, *J* = 7.6 Hz, 2H), 7.56 (t, *J* = 7.6 Hz, 1H), 7.97 (d, *J* = 7.6 Hz, 2H), 8.68 (s, 1H), 8.99 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 11.9, 17.9, 29.7, 35.0, 44.8, 66.6, 71.7, 82.7, 92.0, 121.7, 127.9, 128.8, 132.7, 133.5, 148.2, 149.7, 152.1, 152.7, 164.8. HRMS (pos.): 524.2835 $[M+H]^+$ (HSMS calc. 524.2693).

15*R*: TLC R_f: 0.16 (DCM:MeOH = 97:3). ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.06-1.15 (m, 21H), 2.56 (m, 1H), 2.73 (m, 1H), 2.94 (dd, *J* = 16 Hz, 2.4 Hz, 1H), 3.51 (dd, *J* = 16Hz, 4.0 Hz,1H), 4.07 (s, 1H), 4.57 (d, *J* = 6.0 Hz, 1H), 4.66 (d, *J* = 2.8 Hz, 1H), 6.81 (dd, *J* = 8.4 Hz, 2.0Hz, 1H), 7.44 (t, *J* = 7.6 Hz, 2H), 7.50 (t, *J* = 7.6 Hz, 1H), 8.12 (d, *J* = 7.6 Hz, 2H), 8.47 (s, 1H), 9.03 (s, 1H). ^{13}C NMR (100 MHz, $CDCl_3$): δ 11.9, 17.9, 29.7, 34.6, 43.8, 66.9, 73.5, 82.2, 93.1, 121.9, 128.3, 128.6, 132.5, 133.5, 147.6, 151.9, 152.0, 152.3, 165.0. HRMS (pos.): 524.2821 $[M+H]^+$ (HSMS calc. 524.2693).

Compound 3



Compound **3*S*** and **3*R*** shared the same procedure: Compound **15** (**15*S*** or **15*R*,** 30.0 mg, 55.0 μmol) was dissolved in dry THF (3 mL) and added in 1M of TBAF/THF solution (1 mL). The mixture was stirred under room temperature for 30 minutes, then the solvent was removed by rotary evaporator. The residue was treated with 7N ammonia in methanol (10 mL) under room temperature for overnight. The solution was then dried under vacuum and the crude product was purified by reverse phase C18 column on high-performance liquid chromatography (Water/acetonitrile) to yield the free ring-expanded cyclo-nucleoside as a white powder (**3*S*:** 12.0 mg, 80%; **3*R*** : 13.2 mg, 88%).

3*S*: TLC R_f: 0.27 (DCM:MeOH = 91:9). ^{1}H NMR (400 MHz, D_2O:Pyr-d5 = 7:3): δ 2.20 (m, 1H), 2.52 (m, 1H), 2.70 (dd, *J* = 12.4, 9.6 Hz, 1H), 3.10 (dd, *J* = 12.8, 2.4 Hz, 1H), 3.73 (dt, *J* = 9.2, 3.6 Hz, 1H), 4.34 (d, *J* = 2.4 Hz, 1H), 4.71 (d, J = 5.2 Hz, 1H), 6.26 (dd, *J* = 6.8, 1.6 Hz, 1H), 7.79 (s, 1H). ^{13}C NMR (100 MHz, D_2O:Pyr-d5 = 7:3): δ 32.4, 41.4, 65.2, 69.5, 81.7, 90.3, 115.6, 146.4, 147.7, 151.2, 153.6. HRMS (pos.): 264.1098 $[M+H]^+$ (HSMS calc. 264.1097).

3*R* : TLC R_f: 0.24 (DCM:MeOH = 91:9). ^{1}H NMR (400 MHz, D_2O:Pyr-d5 = 7:3): δ 2.27 (m, 1H), 2.50 (m, 1H), 2.83 (dd, *J* = 13.6, 1.6 Hz, 1H), 3.12 (m, 1H), 4.00 (m, 1H), 4.45 (d, *J* = 5.2 Hz, 1H), 4.47 (d, J = 2.0 Hz, 1H), 6.34 (dd, *J* =

6.4, 2.0 Hz, 1H), 7.73 (s, 1H). ^{13}C NMR (100 MHz, D_2O:Pyr-d5 = 7:3): δ 32.2, 40.7, 65.2, 71.2, 81.7, 91.2, 115.8, 147.9, 149.2, 151.0, 153.4. HRMS (pos.): 264.1107 $[M+H]^+$ (HSMS calc. 264.1097).

3.6. References

1. S. T. Crooke, C. F. Bennett, *Annu. Rev. PharmacoL Toxicol.* **1996**, *36,* 107-129.
2. S. Agrawal, *Trends Biotechnol.* **1996**, *14*, 376-387.
3. S. M. Freier, K.-H. Altmann, *Nucleic Acids Res.* **1997**, *25*, 4429-4443.
4. R. Yarchoan, H. Mitsuya, C. E. Myers, S. Broder, *N. Engl. J. Med.* **1989**, *321*, 726-738.
5. M. E. Gleave1, B. P. Monia, *Nat. Rev. Cancer* **2005,** *5*, 468-479.
6. S. Obika, D. Nanbu, Y. Hari, K. Morio, Y. In, T. Ishida, T. Imanishi, *Tetrahedron Lett*. 1997, 38, 8735–8738.
7. A. A. Koshkin, S. K. Singh, P. Nielsen, V. K. Rajwanshi, R. Kumar, M. Meldgaard, C. E. Olsen, J. Wengel, *Tetrahedron* **1998**, *54*, 3607–3630.
8. C. Wahlestedt, P. Salmi, L. Good, J. Kela, T. Johnsson, T. Hökfelt, C. Broberger, F. Porreca, J. Lai, K. Ren, M. Ossipov, A. Koshkin, N. Jakobsen, J. Skouv, H. Oerum, M. H. Jacobsen, J. Wengel, *Proc. Natl. Acad. Sci.* **2000**, *97*, 5633-5638.
9. P. Srivastava, J. Barman, W. Pathmasiri, O. Plashkevych, M. Wenska, J. Chattopadhyaya, *J. Am. Chem. Soc.* **2007**, *129*, 8362-8379.
10. H. Yueh, H. Yu, C. S. Theile, A. Pal, A. Horhota, N. J. Greco, L. W. McLaughlin, *Nucleosides, Nucleotides, and Nucleic Acids*. **2012**, *31*, 661-679.

11. L. Kuraoka, C. Bender, A. Romieu, J. Cadet, R. D. Wood, T. Lindahl, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3832-3837.

12. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*, **1981**, *78*, 2179-2183.

13. B. L. Partridge and S. A. Salisbury, *NDB ID: BDL075,* **1996**.

14. A. Matsuda, T. Ueda, *Chem. Pharm. Bull.* **1986**, *34*, 1573–1578.

15. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura and R. E. Dickerson, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences*, **1981**, *78*, 2179-2183.

16. R. Pellicciari, R. Fringuelli, E. Sisani, M. Curini, *J. Chem. Soc. Perkin Trans. 1.* **1981**, 2566-2569.

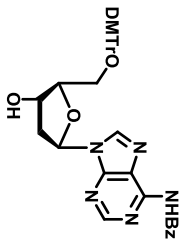
17. K. Nagao, M. Chiba, S. W. Kim, *Synthesis*, **1983**, *3*, 197-199.

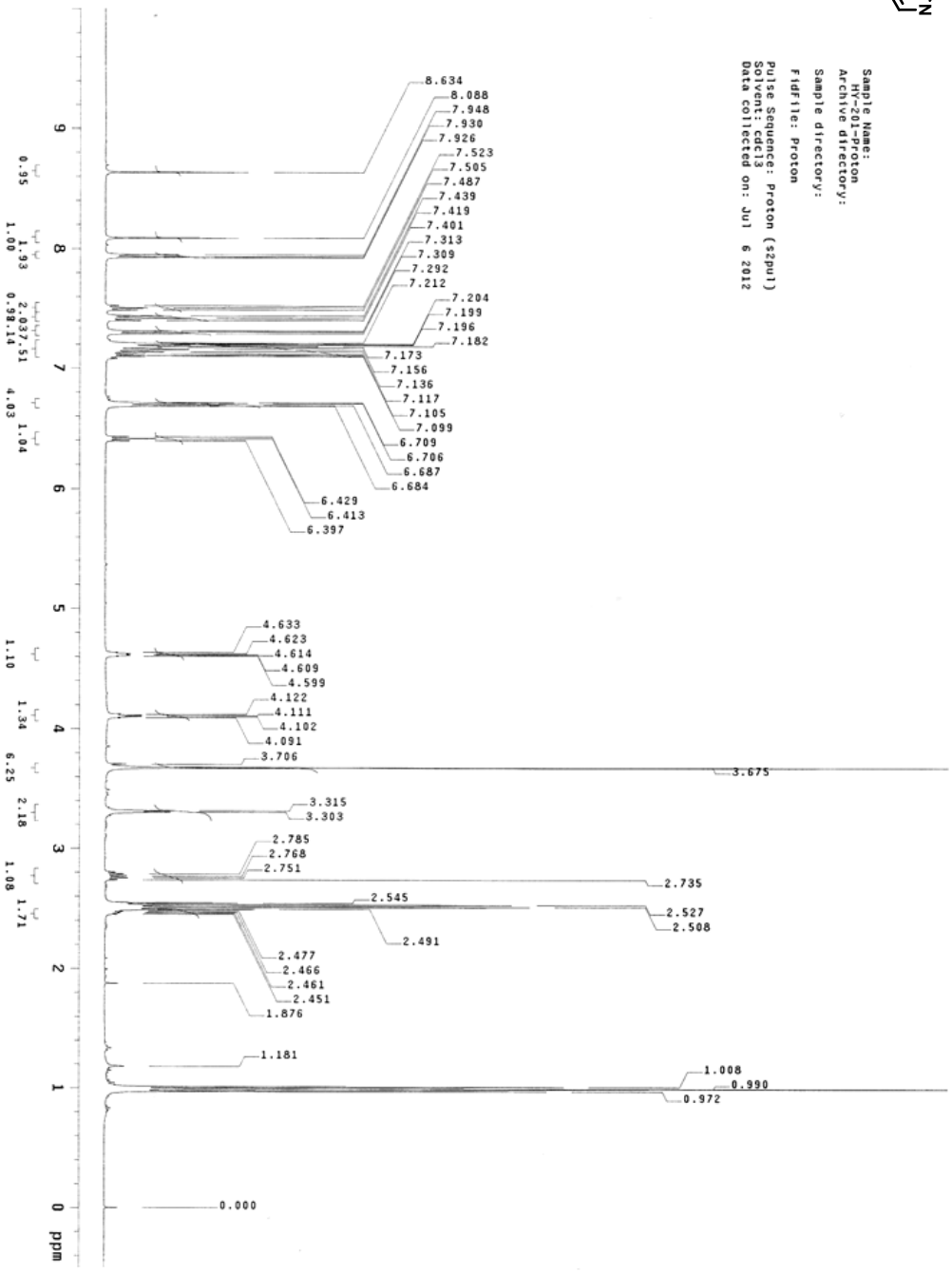
18. I. Braun, F. Rudroff, M. D. Mihovilovic, T. Bach, *Synthesis*, **2007**, *24*, 3896-3906

19. R. K. Boeckman, Jr., P. Shao, J. J. Mullins, *Organic Synthesis, Coll.* **2004,** *10*, 696-701.

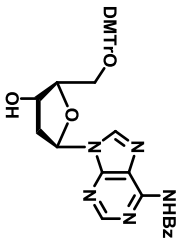
3.7. NMR Spectral Data

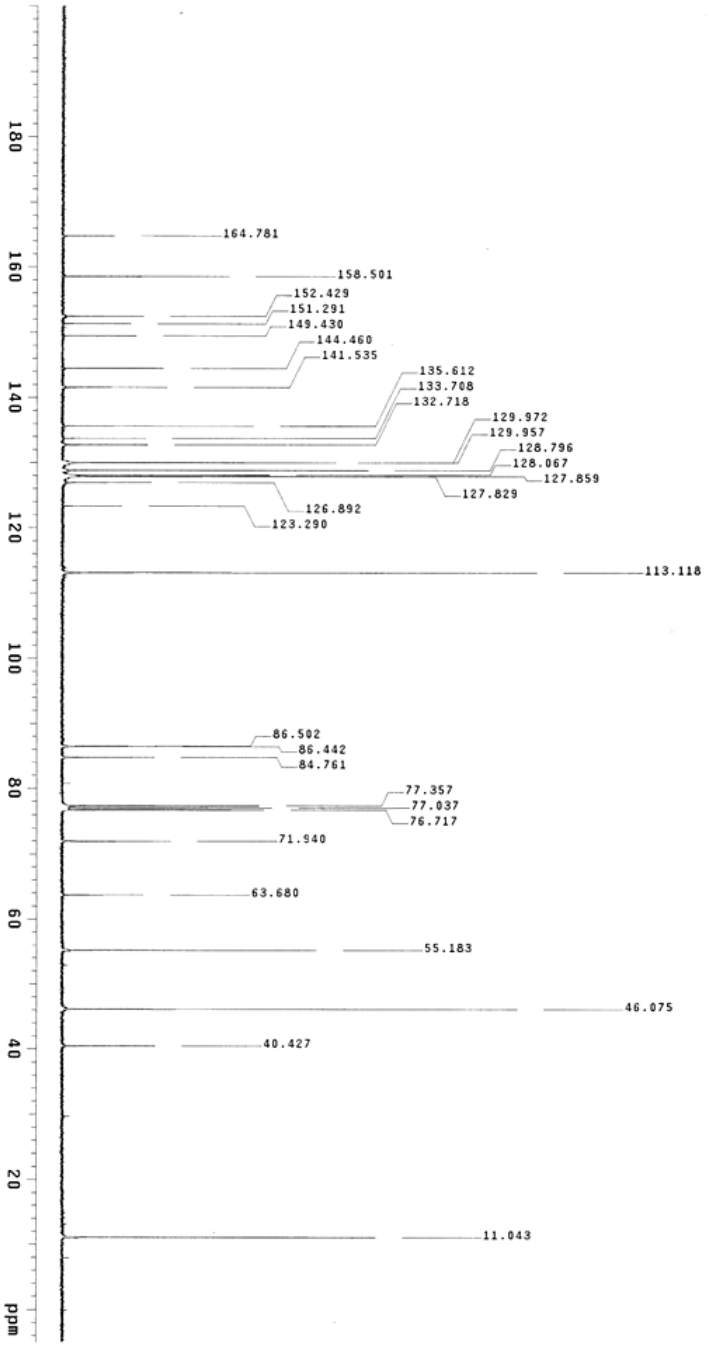
^{1}H NMR of Compound **9**
Solvent: $CDCl_3$ at 400 MHz



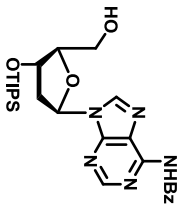


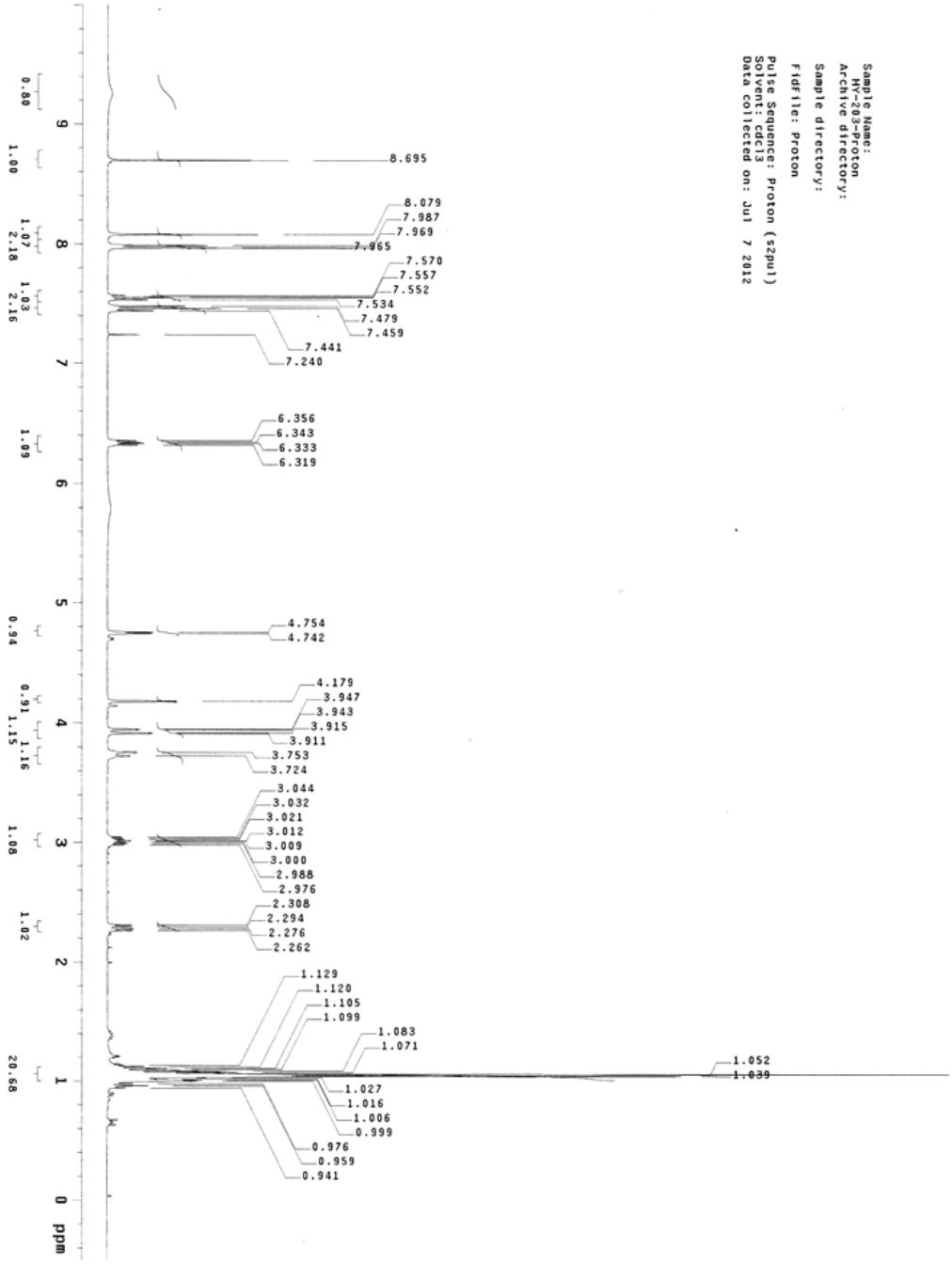
^{13}C NMR of Compound **9**
Solvent: $CDCl_3$ at 100 MHz



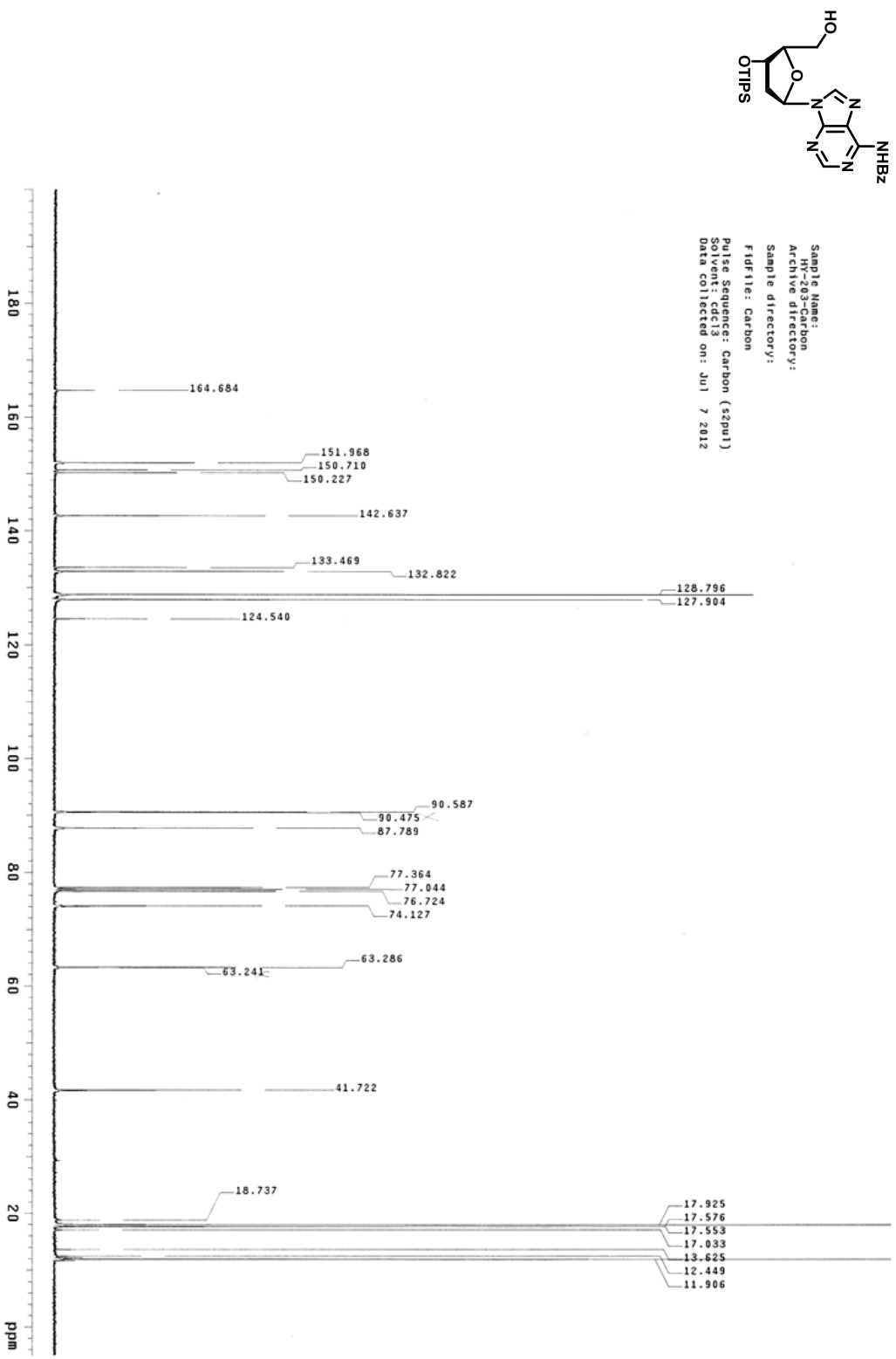


^{1}H NMR of Compound **10**
Solvent: $CDCl_3$ at 400 MHz

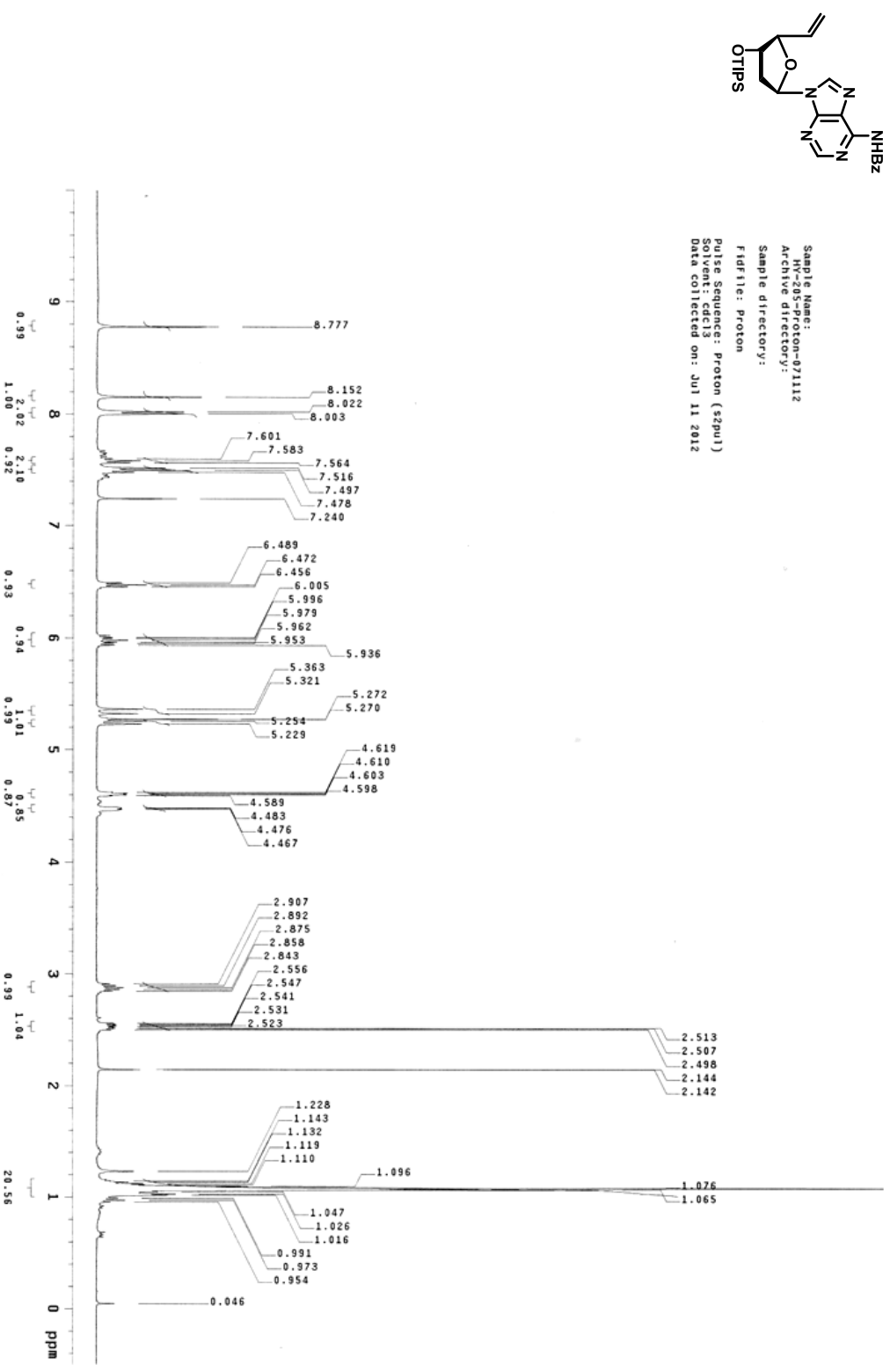




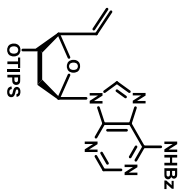
^{13}C NMR of Compound **10**
Solvent: $CDCl_3$ at 100 MHz

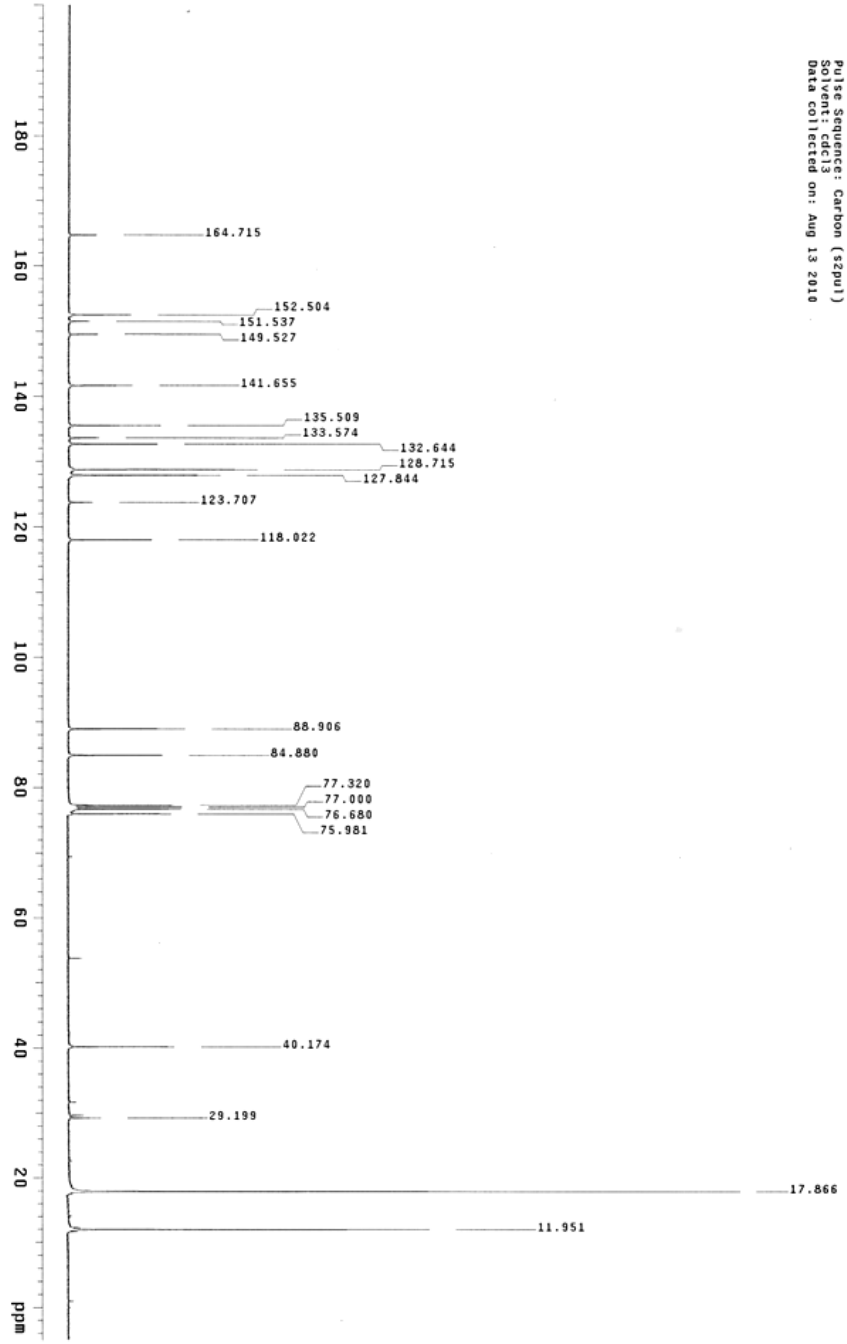


^{1}H NMR of Compound **11**
Solvent: $CDCl_3$ at 400 MHz

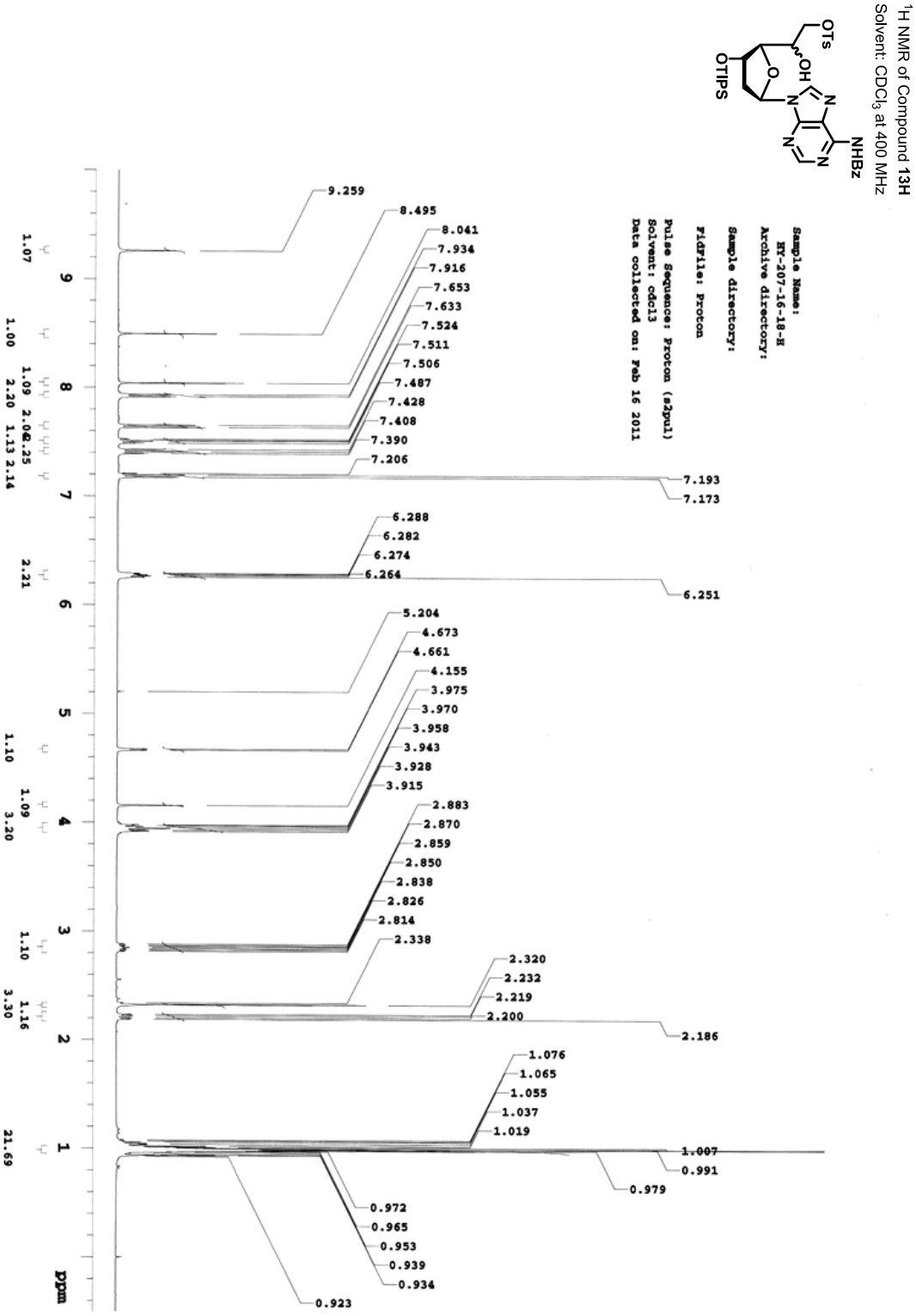


^{13}C NMR of Compound **11**
Solvent: $CDCl_3$ at 100 MHz

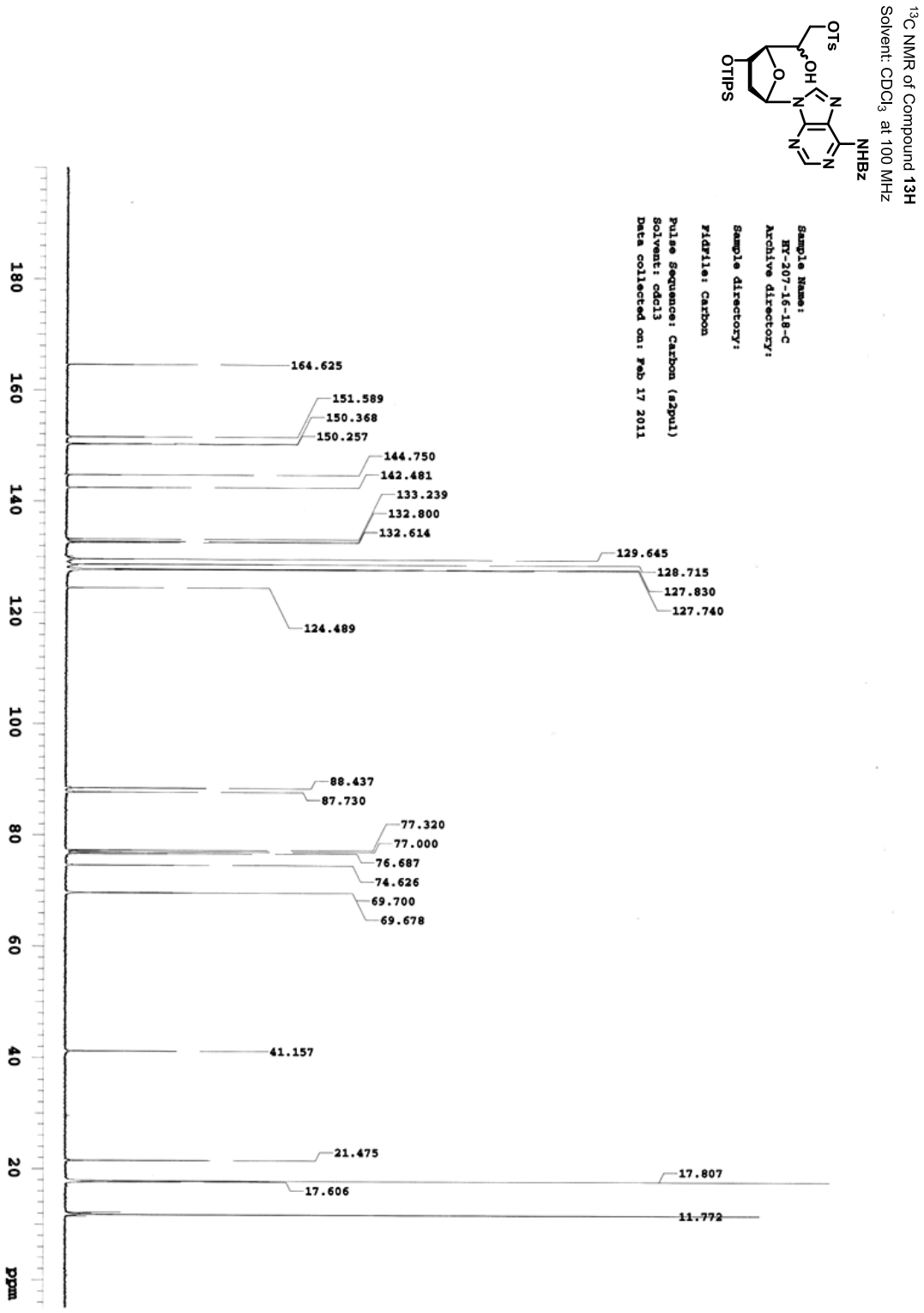




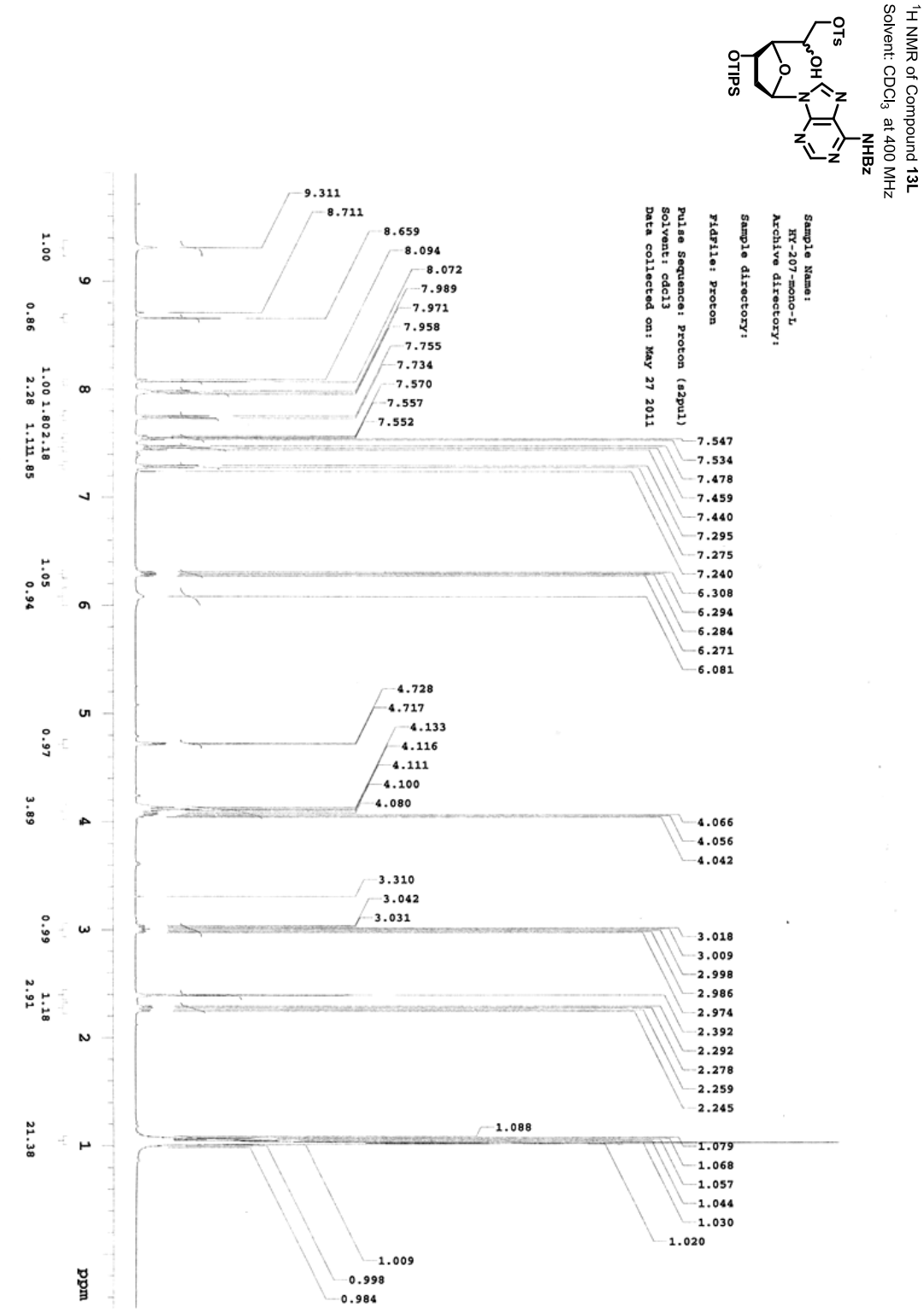
^{1}H NMR of Compound **13H**
Solvent: $CDCl_3$ at 400 MHz



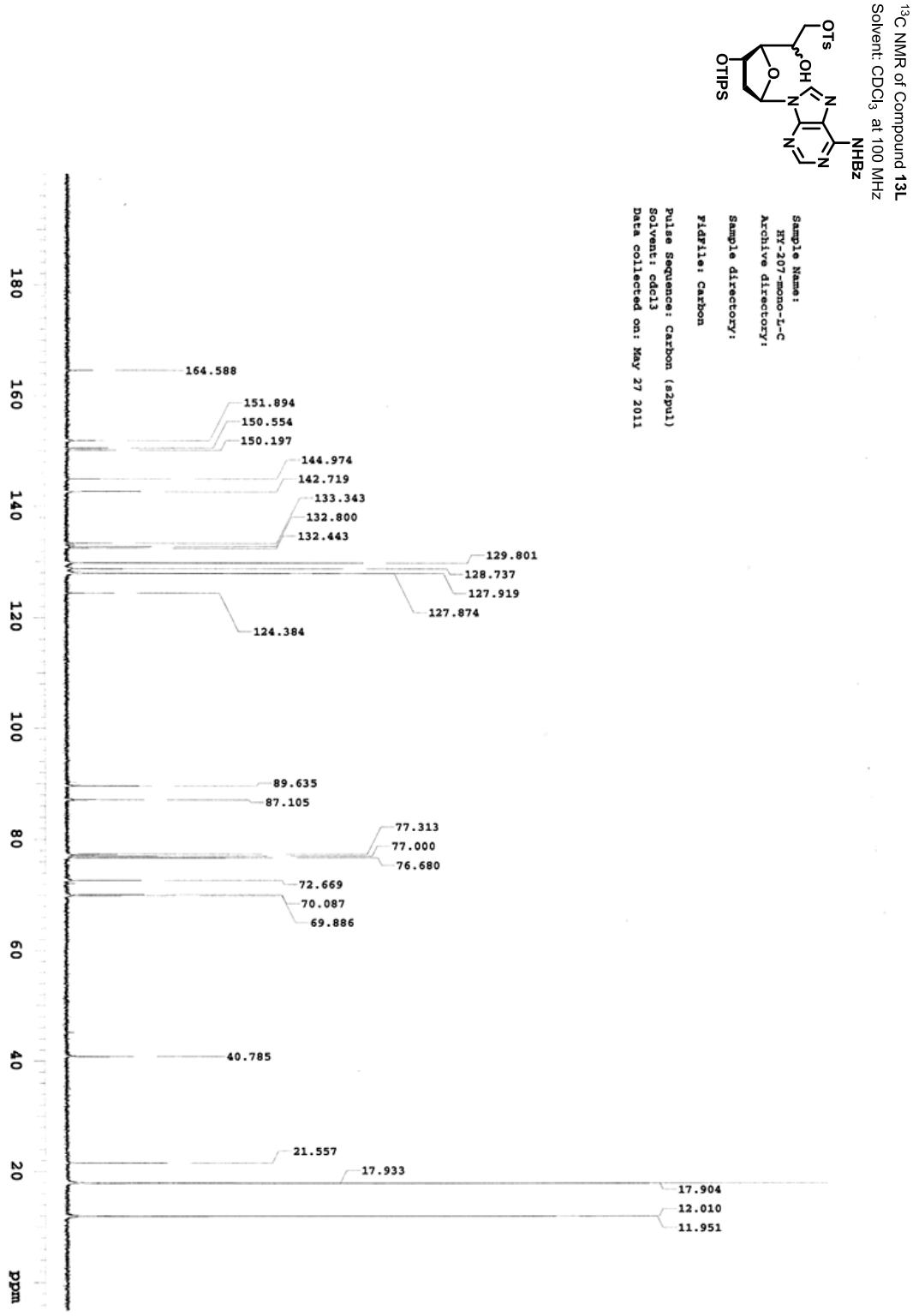
^{13}C NMR of Compound **13H**
Solvent: $CDCl_3$ at 100 MHz



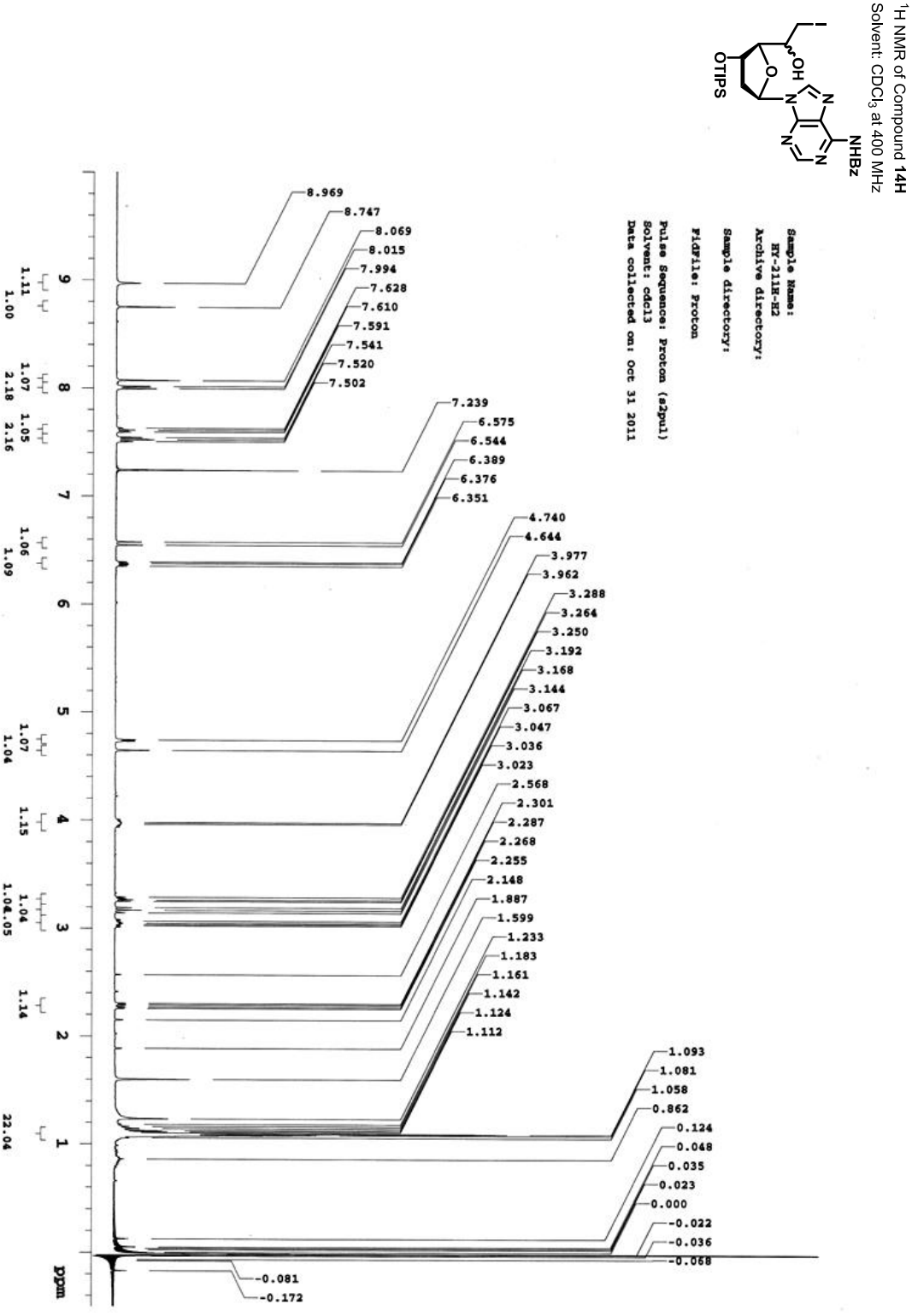
^{1}H NMR of Compound **13L**
Solvent: $CDCl_3$ at 400 MHz



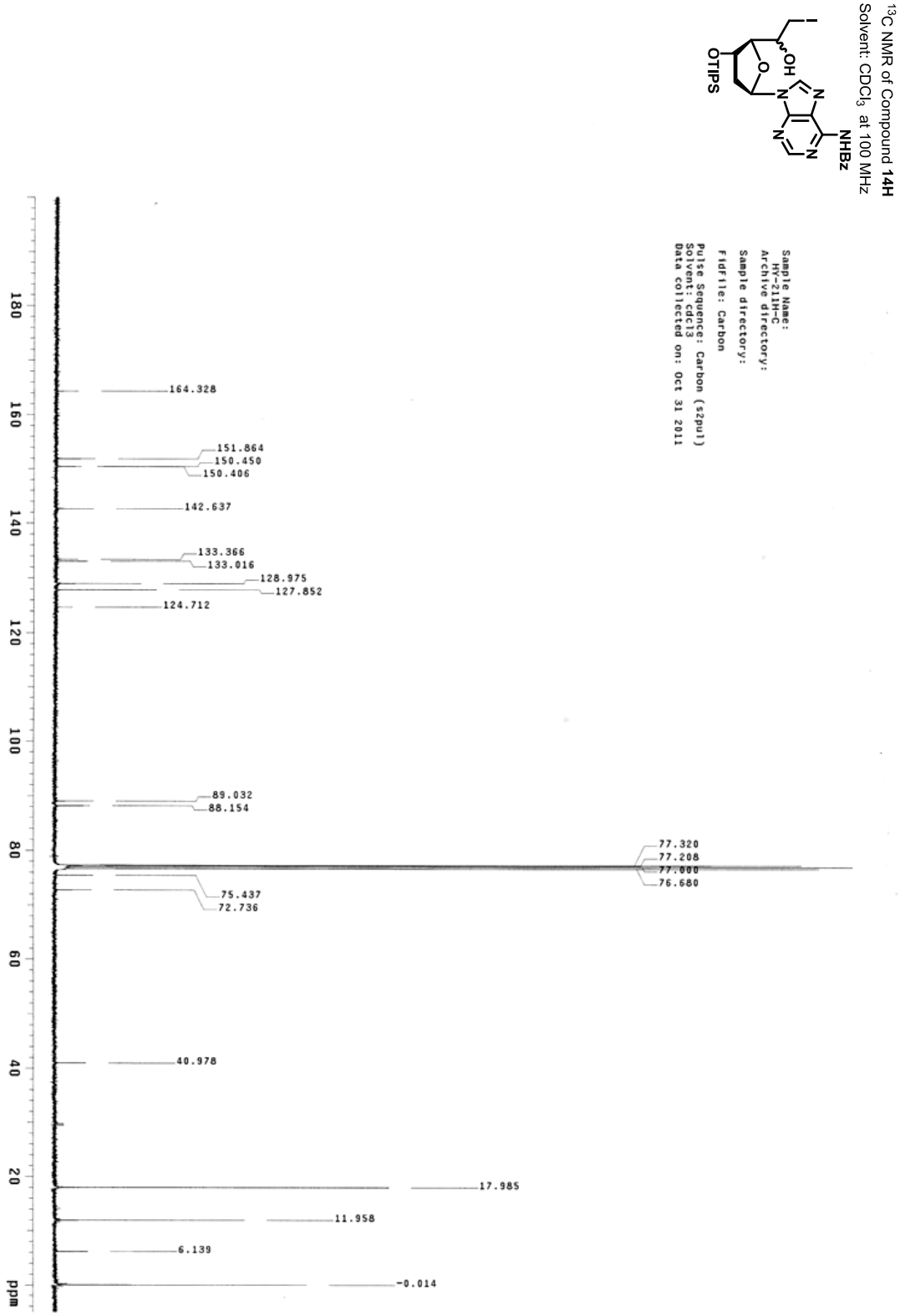
^{13}C NMR of Compound **13L**
Solvent: $CDCl_3$ at 100 MHz



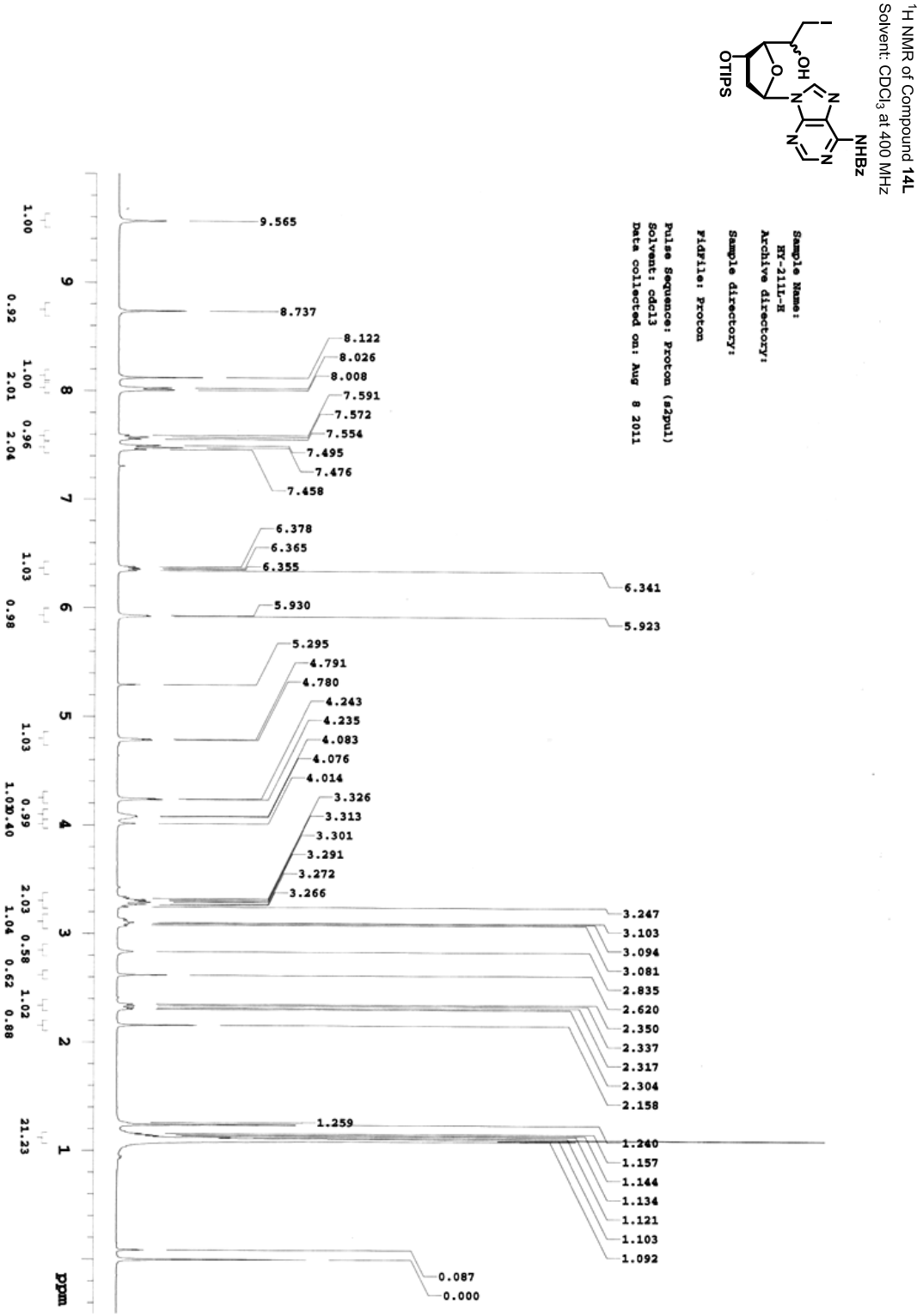
^{1}H NMR of Compound **14H**
Solvent: $CDCl_3$ at 400 MHz



^{13}C NMR of Compound **14H**
Solvent: $CDCl_3$ at 100 MHz

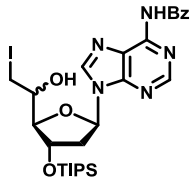


^{1}H NMR of Compound **14L**
Solvent: $CDCl_3$ at 400 MHz



^{13}C NMR of Compound **14L**

Solvent: $CDCl_3$ at 100 MHz

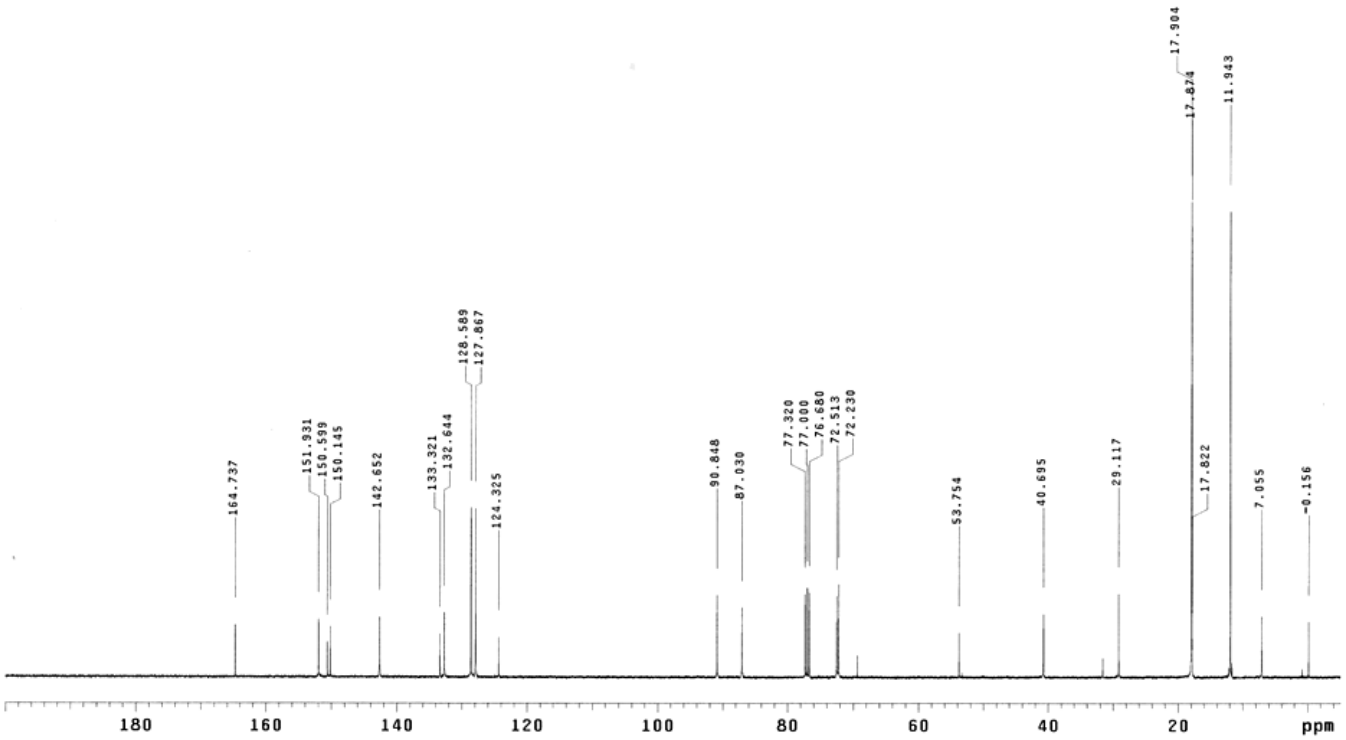


Sample Name:
HY-211L-C
Archive directory:

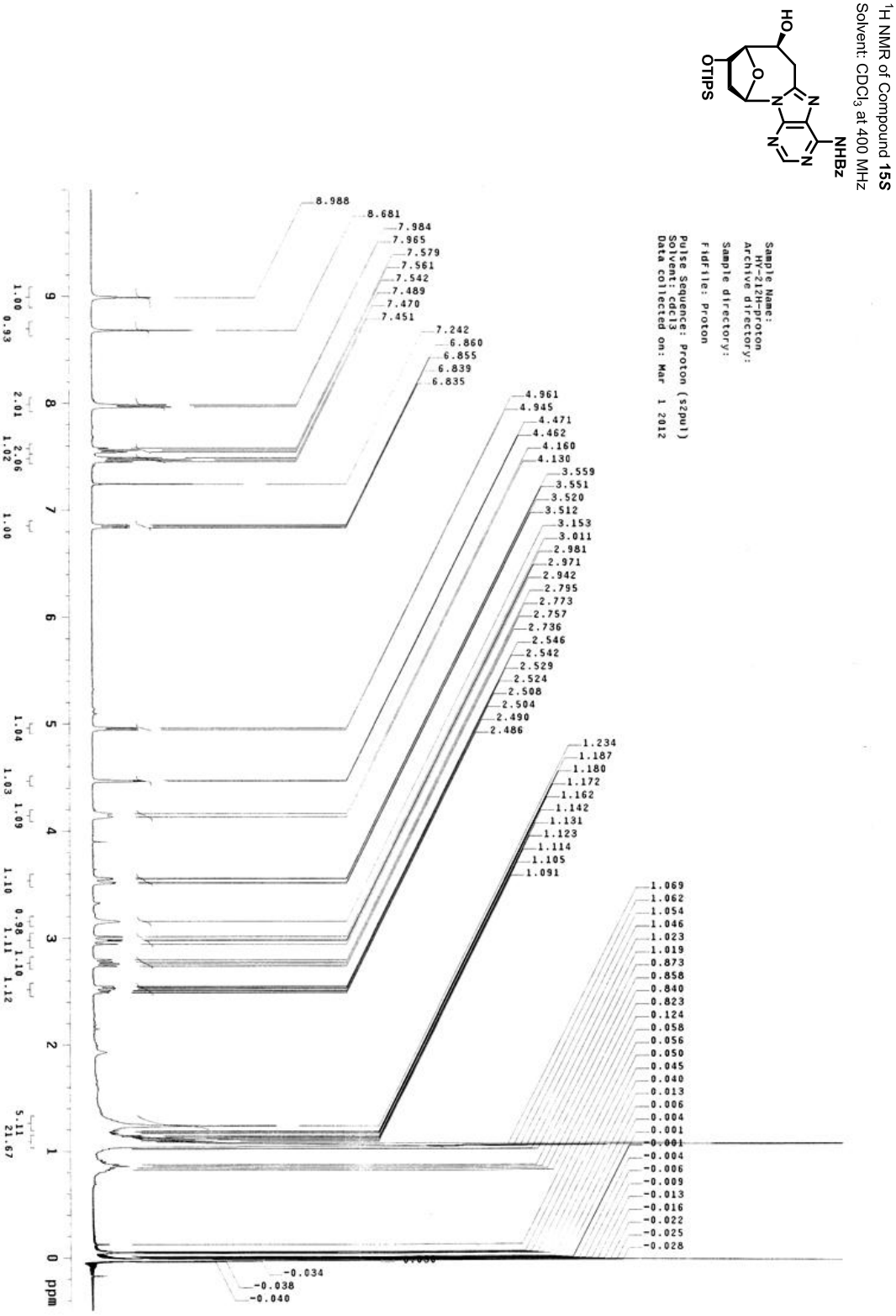
Sample directory:

FidFile: Carbon

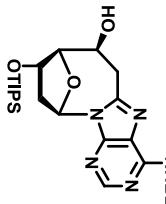
Pulse Sequence: Carbon (s2pul)
Solvent: cdcl3
Data collected on: Aug 8 2011

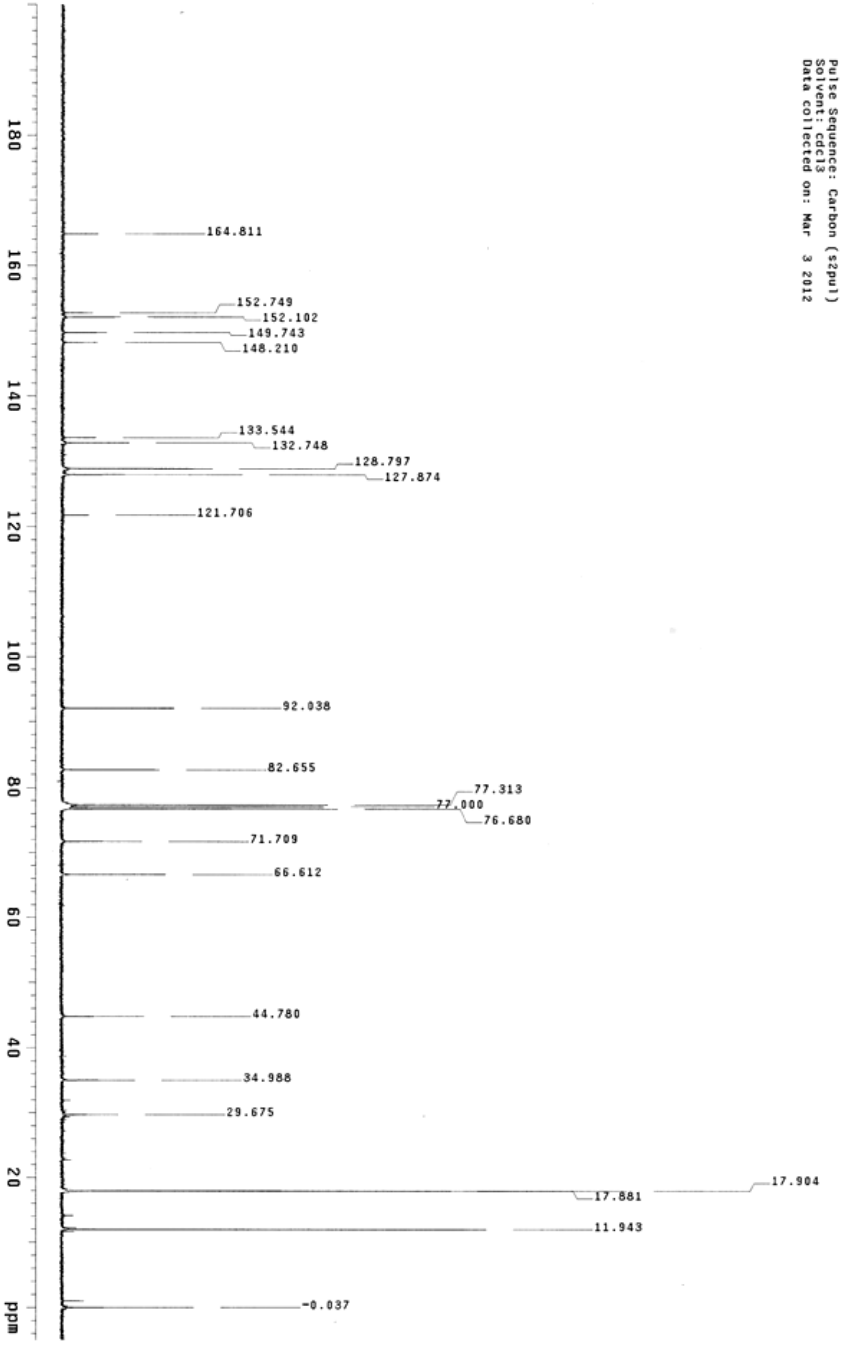


^{1}H NMR of Compound **15S**
Solvent: $CDCl_3$ at 400 MHz

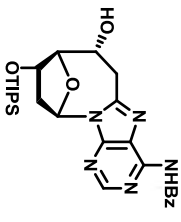


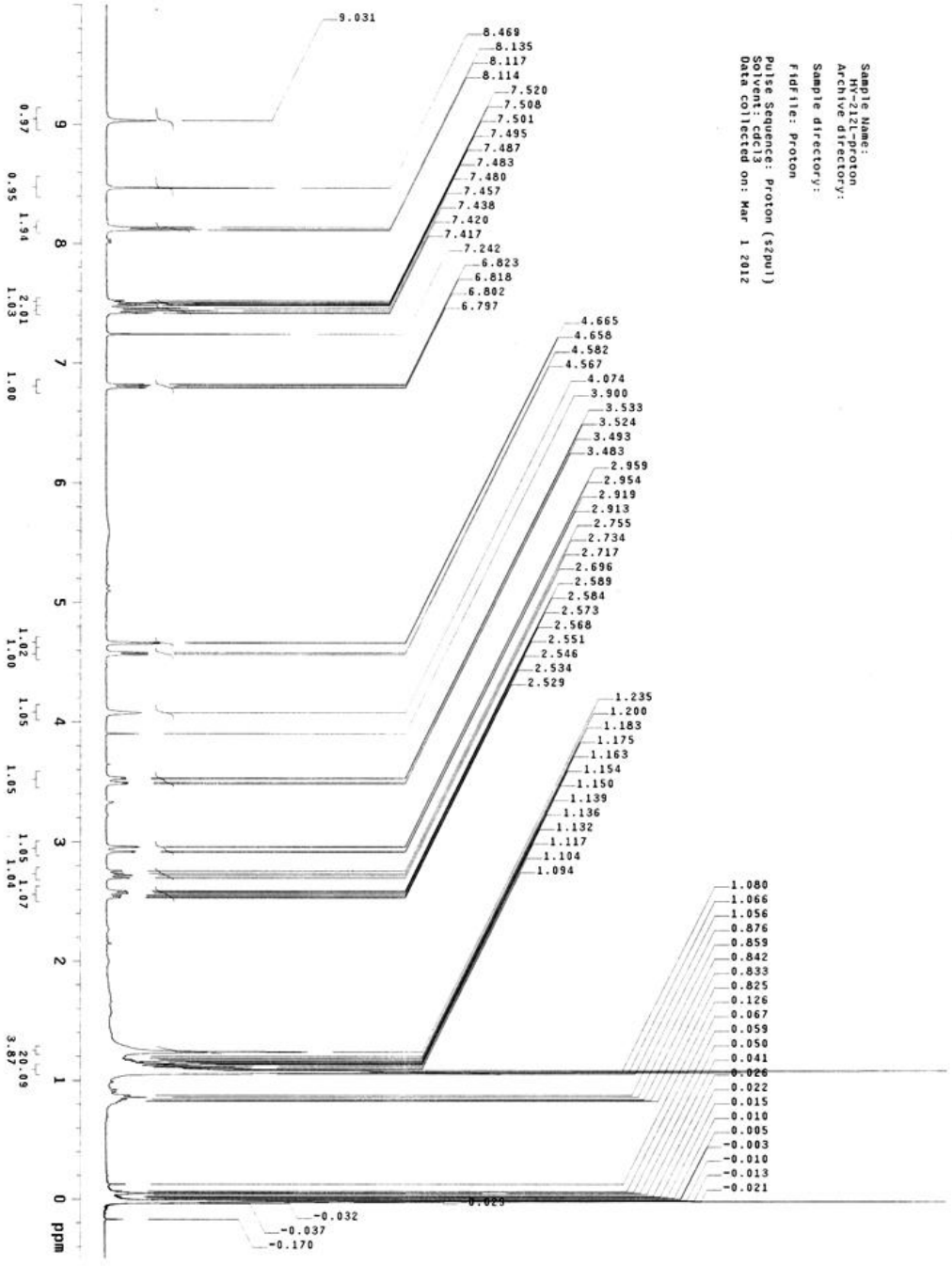
^{13}C NMR of Compound **15S**
Solvent: $CDCl_3$ at 100 MHz



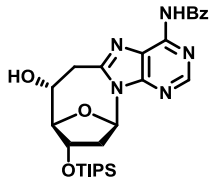


^{1}H NMR of Compound **15*R***
Solvent: $CDCl_3$ at 400 MHz





^{13}C NMR of Compound **15*R***
Solvent: $CDCl_3$ at 100 MHz

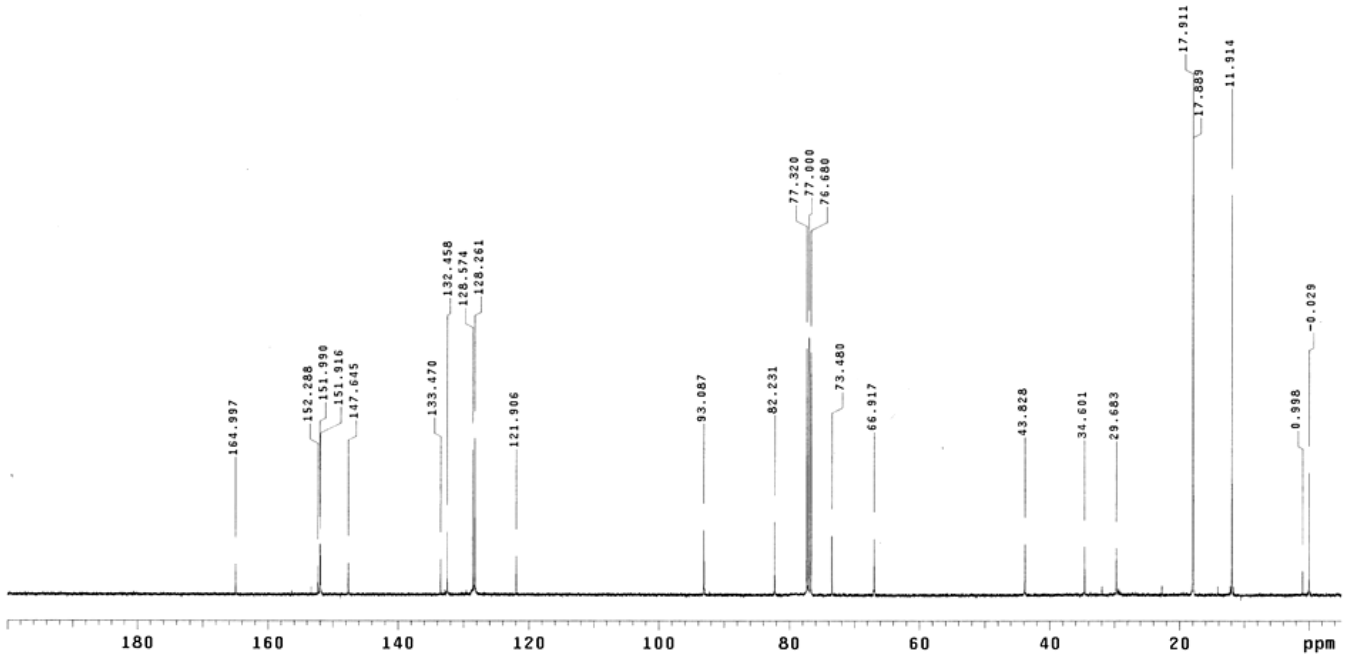


Sample Name:
HY-212L-Carbon
Archive directory:

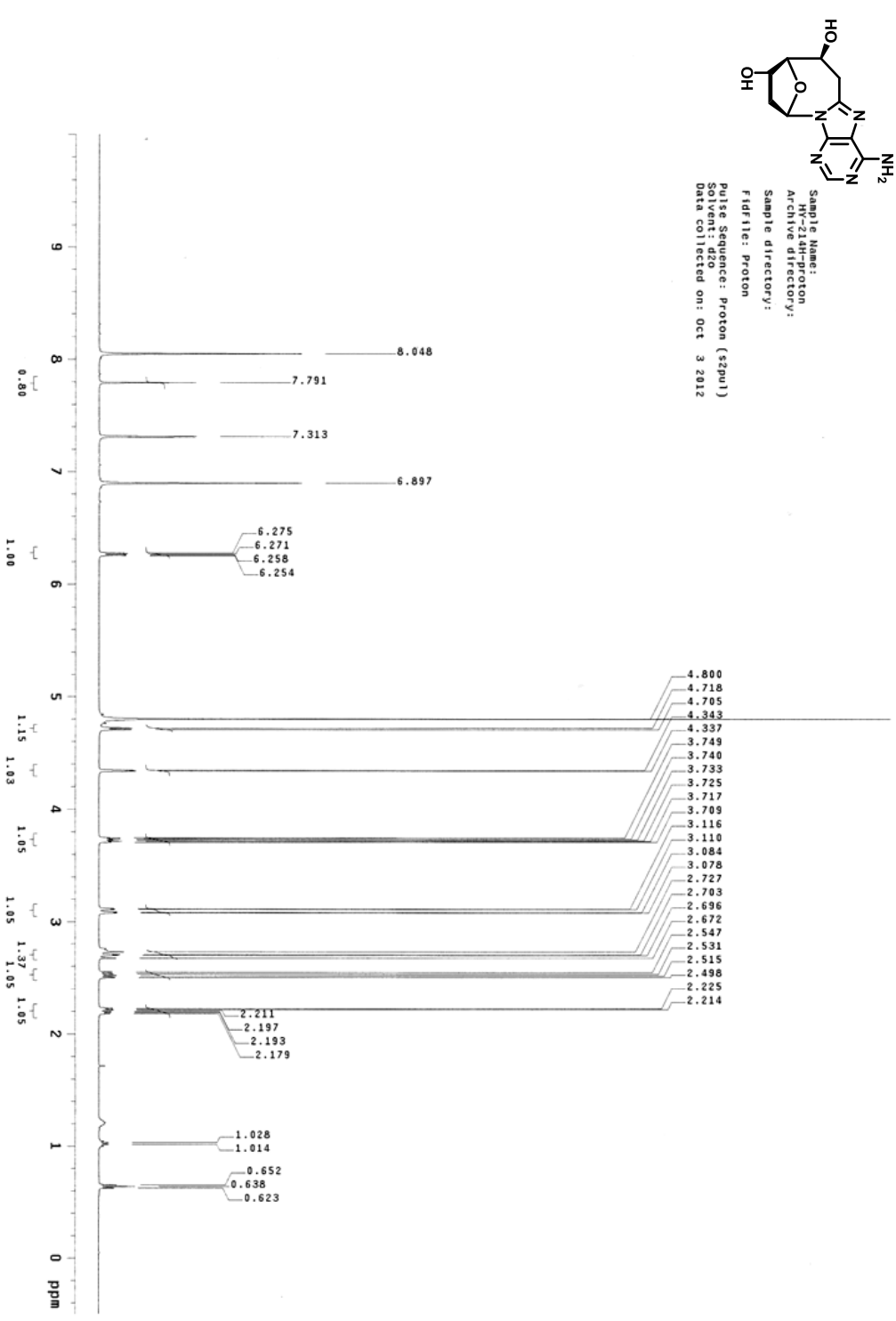
Sample directory:

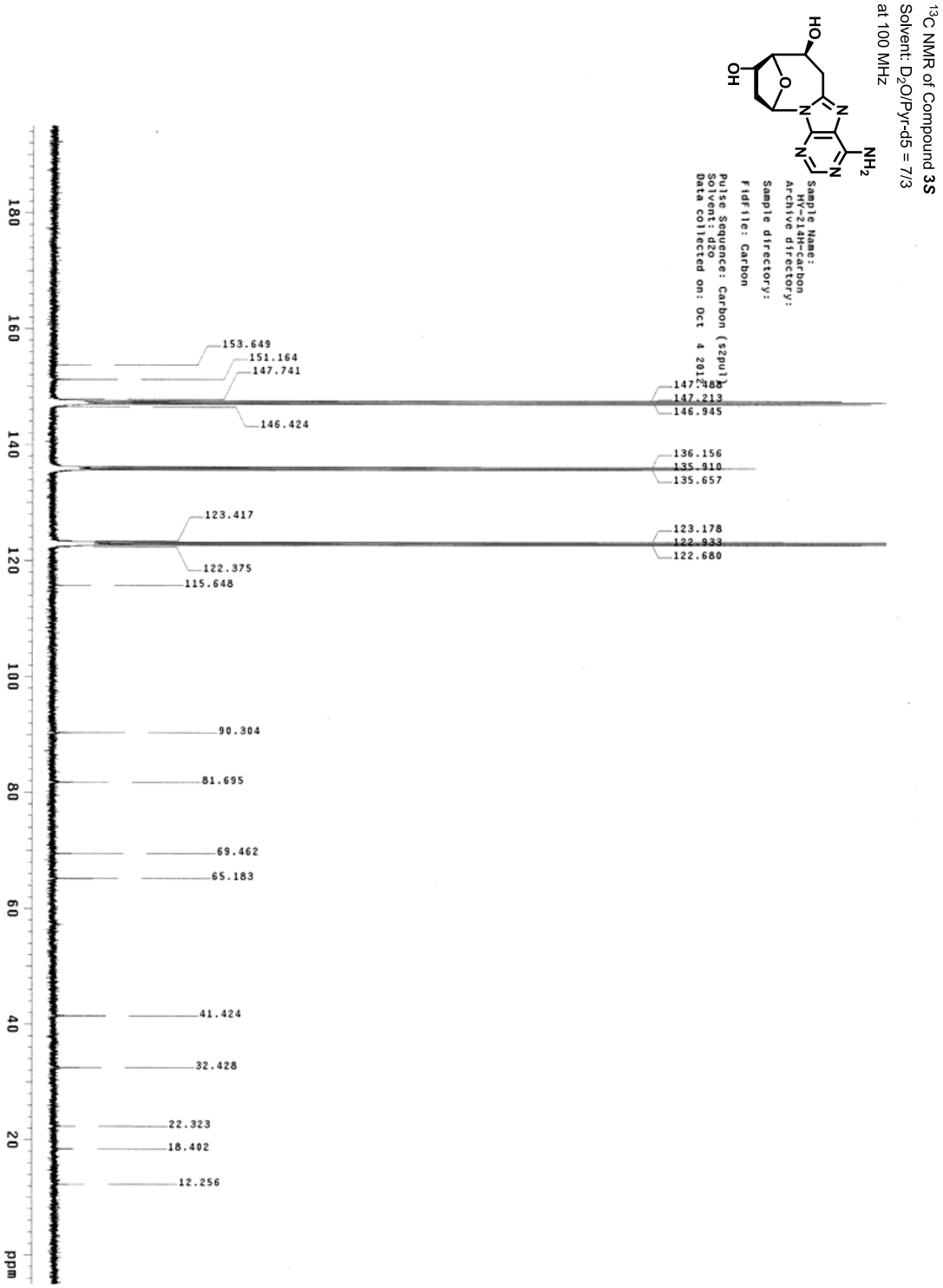
FidFile: Carbon

Pulse Sequence: Carbon (s2pul)
Solvent: cdcl3
Data collected on: Mar 4 2012

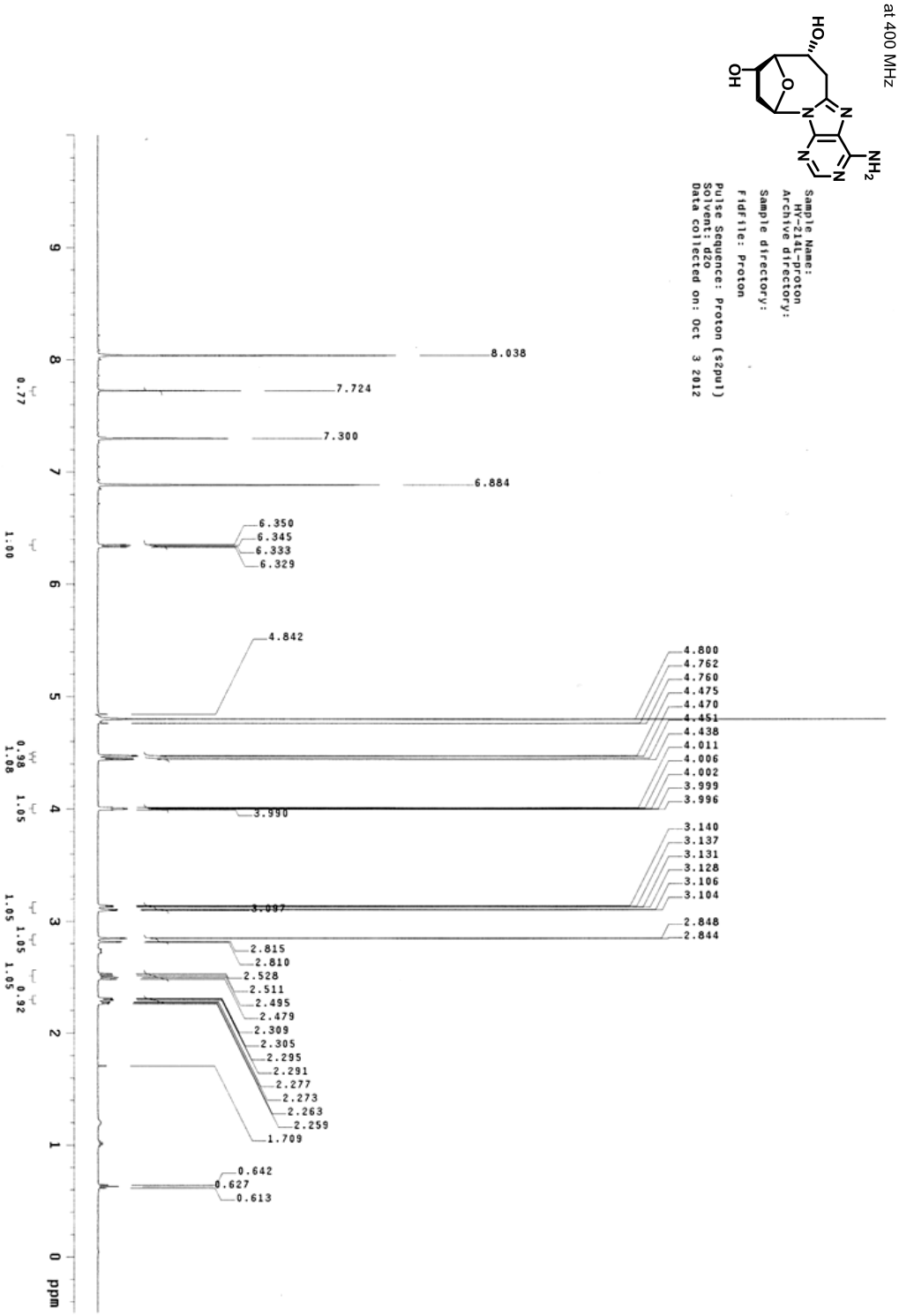


^{1}H NMR of Compound **3S**
Solvent: D_2O/Pyr-d_5 = 7/3
at 400 MHz

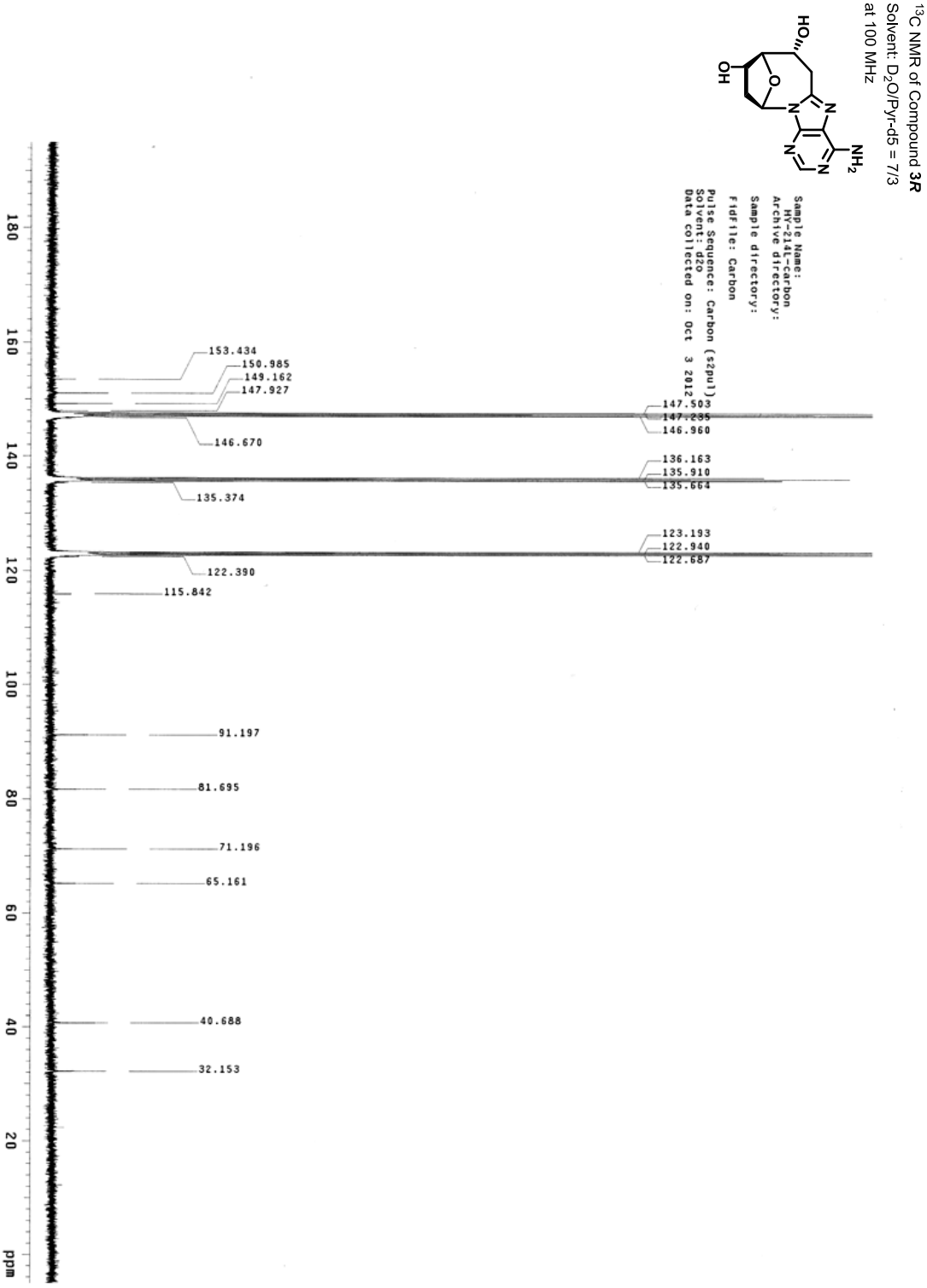



13C NMR of Compound 3S
Solvent: D2O/Pyr-d5 = 7/3
at 100 MHz
HO
OH
O
N
N
N
N
NH2
Sample Name:
HY-214H-carbon
Archive directory:
Sample directory:
Fidfile: Carbon
Pulse Sequence: Carbon (s2pul)
Solvent: d2o
Data collected on: Oct 4 2012
153.649
151.164
147.741
147.188
147.213
146.945
146.424
136.156
135.910
135.657
123.417
123.178
122.933
122.680
122.375
115.648
90.304
81.695
69.462
65.183
41.424
32.428
22.323
18.402
12.256
180
160
140
120
100
80
60
40
20
ppm

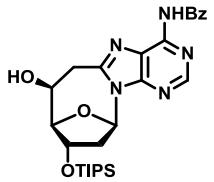
^{1}H NMR of Compound **3*R***
Solvent: D_2O/Pyr-d_5 = 7/3
at 400 MHz



^{13}C NMR of Compound **3R**
Solvent: D_2O/Pyr-d5 = 7/3
at 100 MHz

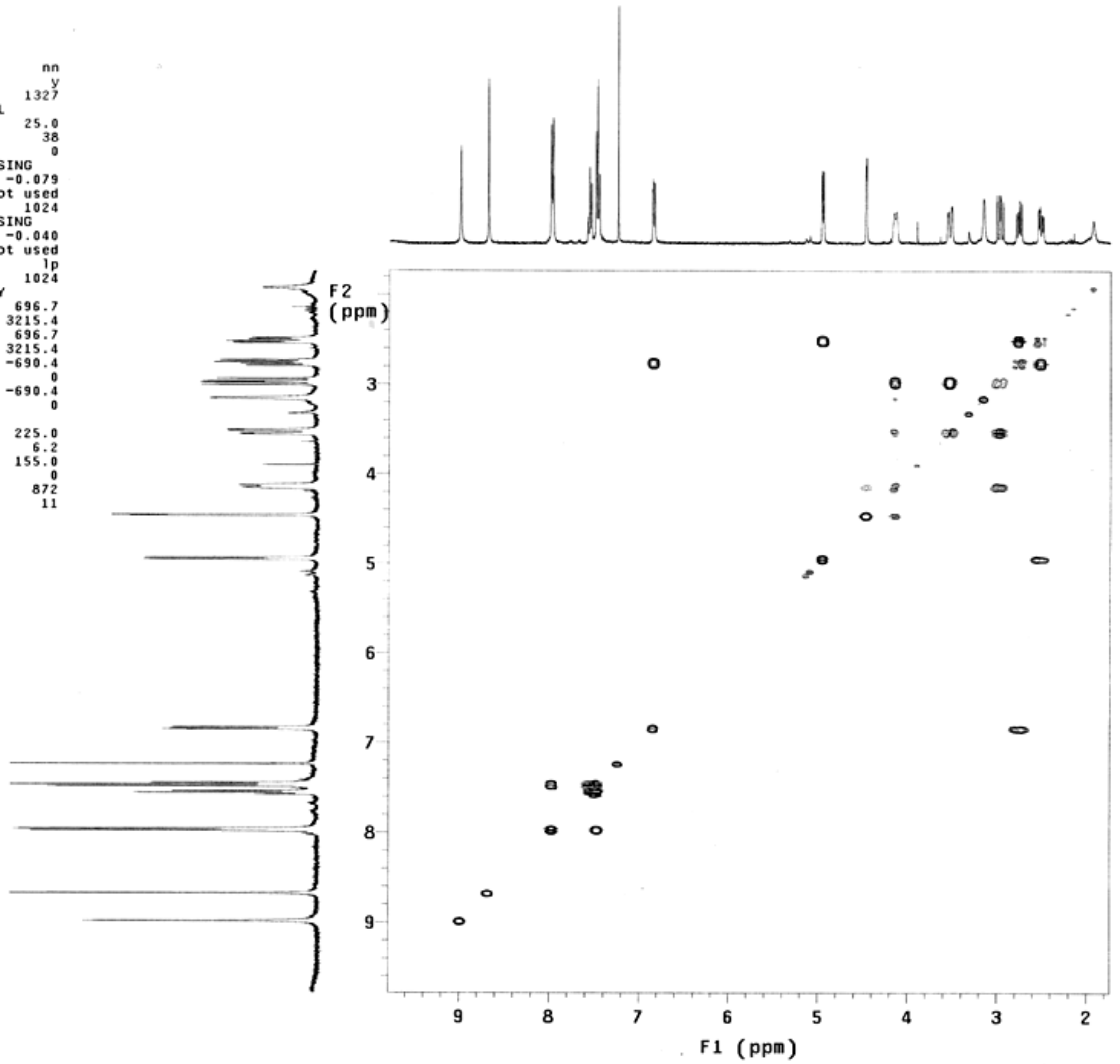


COSY of Compound **15*S***
Solvent: $CDCl_3$ at 500 MHz

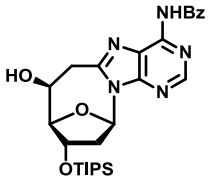


exp2 Cosy

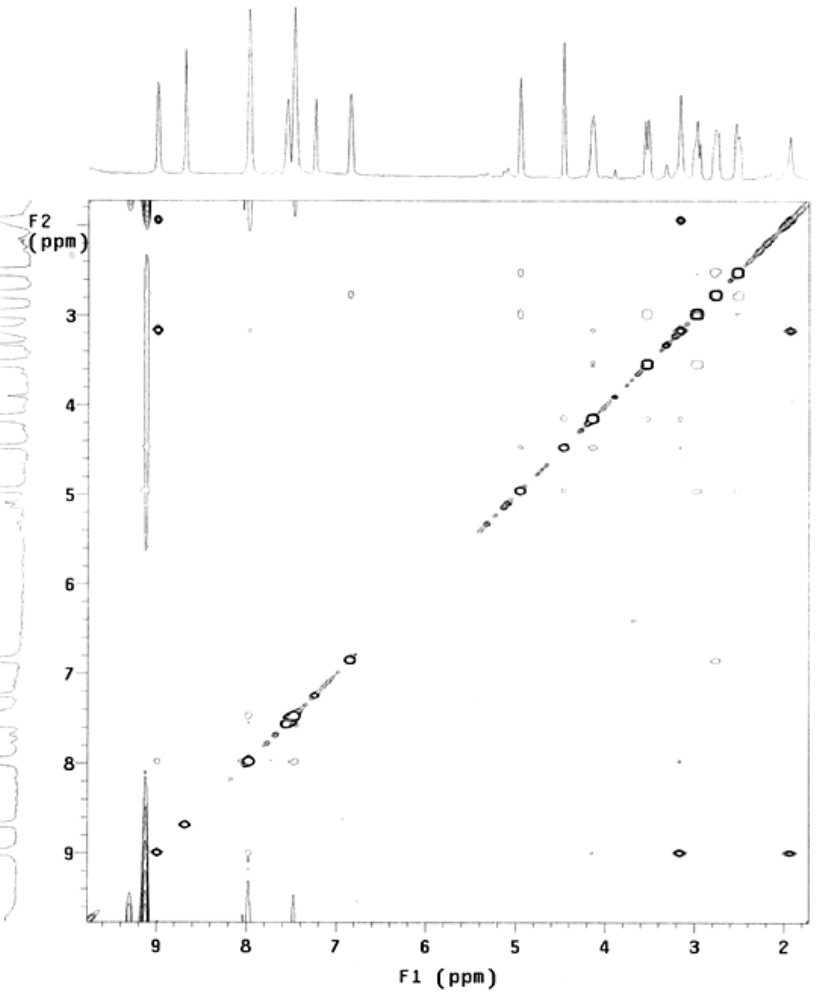
SAMPLE		FLAGS	
date	Mar 2 2012	hs	nn
solvent	cdcl3	sspul	y
sample		hsglvl	1327
ACQUISITION		SPECIAL	
sw	3221.6	temp	25.0
at	0.159	gain	38
np	1024	spin	0
fb	4000	F2 PROCESSING	
ss	16	sb	-0.079
d1	1.000	sbs	not used
nt	8	fn	1024
2D ACQUISITION		F1 PROCESSING	
sw1	3221.6	sb1	-0.040
ni	128	sbs1	not used
PRESATURATION		proc1	lp
satmode	nnn	fn1	1024
satdly	0	DISPLAY	
satfrq	499.8	sp	696.7
satpwr	-13	wp	3215.4
TRANSMITTER		sp1	696.7
tn	H1	wp1	3215.4
sfrq	399.769	rf1	-690.4
tof	302.4	rfp	0
tpwr	60	rfl1	-690.4
pw	10.375	rfp1	0
tpwr_cf	0.977	PLOT	
GRADIENTS		wc	225.0
gzlvl1	1769	sc	6.2
gt1	0.001000	wc2	155.0
gstab	0.000500	sc2	0
DECOUPLER		vs	872
dn	C13	th	11
dm	nnn	ai cdc av	



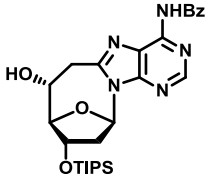
NOESY of Compound **15*S***
Solvent: $CDCl_3$ at 400 MHz




```
expl  Noesy

      SAMPLE
date    Mar  2 2012
solvent      cdcl3
file /home/ALL/lwm~
/Han Yueh/200/2D/H~
Y-212H-NOESY_4_212~
          _02.fid
   ACQUISITION
sfrq        399.769
tn               H1
at             0.159
np              1024
sw            3221.6
fb              4000
bs                32
ss                16
tpwr              60
pw              10.4
compH         0.9774
d1             1.000
mix            0.500
tof            296.3
nt                48
ct                48
gain              52
sspul              y
gt1           0.0025
gzlvl1          1769
hsgt          0.0020
hsglvl          1327
satmode          nnn
satpwr           -13
satdly             0
scuba              n
zqflg              y
alt_grd            n
      FLAGS
il                 n
in                 n
dp                 y
hs                nn
sspul              y
  2D ACQUISITION
sw1           3221.6
ni               128
phase       arrayed
     DISPLAY
sp             690.6
wp            3215.4
vs              6109
sc                 6
wc               225
hzmm           14.29
is            744.69
rfl           -684.3
rfp                0
th                 4
ins          100.000
ai    ph

      2D DISPLAY
sp1            690.6
wp1           3215.4
sc2                0
wc2              155
rfl1          -684.3
rfp1               0
      DEC. & VT
dfrq          100.531
dn                C13
dpwr               31
dof                0
dm                 n
dmm                c
dmf            29412
dseq
dres             9.0
homo               n
temp            25.0
     PROCESSING
gf             0.073
gfs        not used
wtfile
proc              ft
fn              1024
math               f
ssfilter   not used
ssorder            3
ssntaps            3
  2D PROCESSING
gf1            0.037
gfs1       not used
wtfile1
proc1             tp
fn1             1024

ACQUISITION ARRAYS
array          phase
arraydim         256

  i            phase
  1                1
  2                2
```

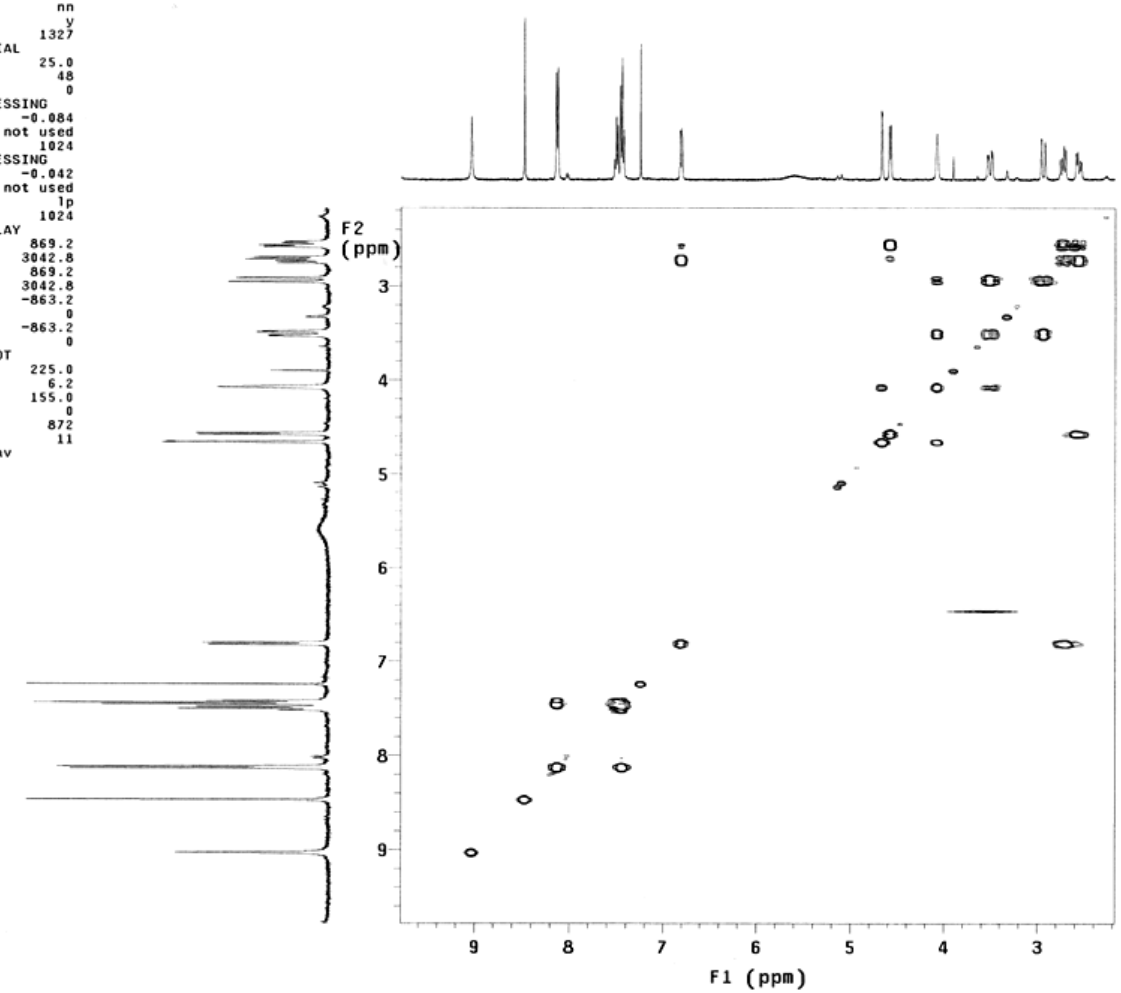


COSY of Compound **15*R***
Solvent: $CDCl_3$ at 500 MHz

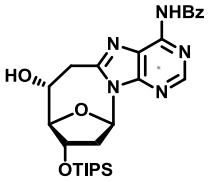


exp2 Cosy

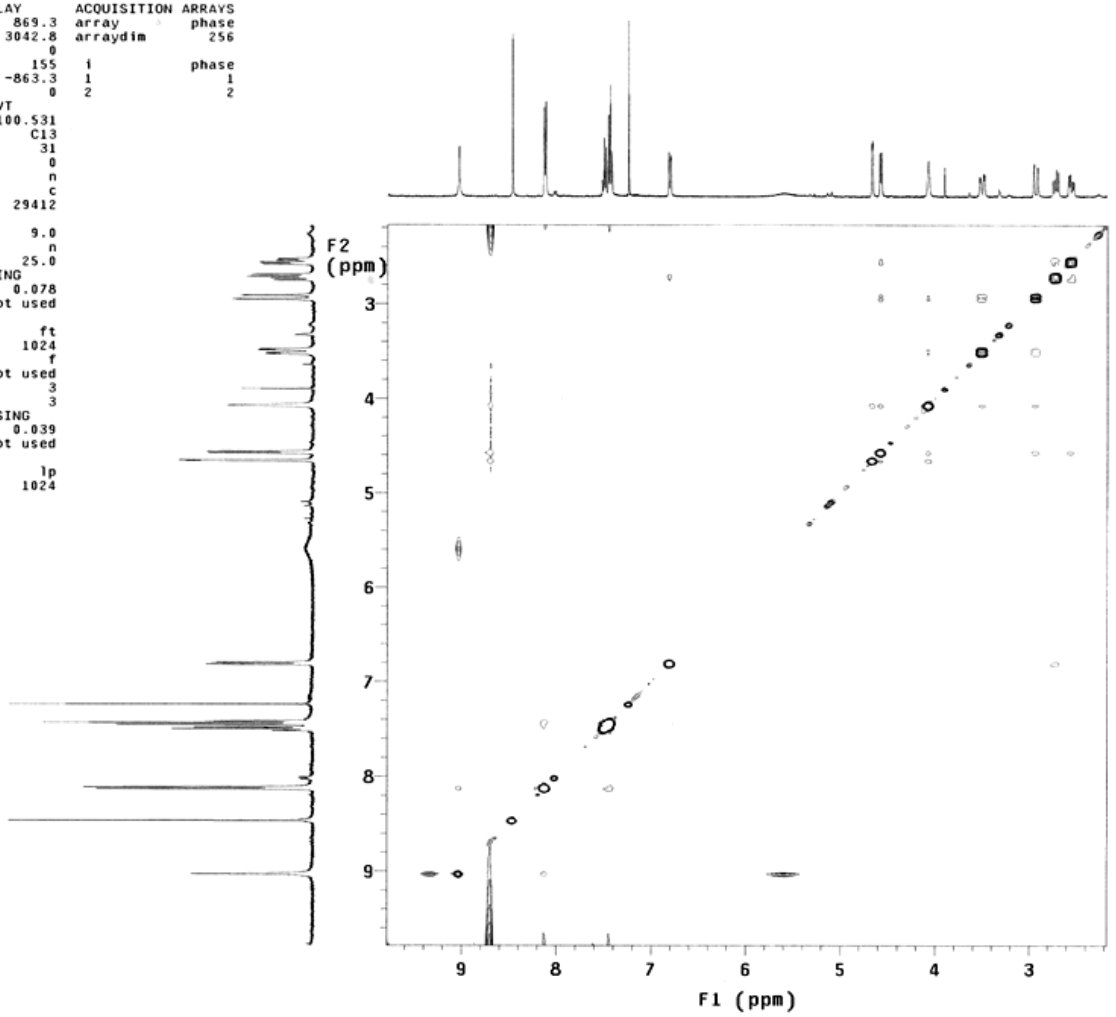
SAMPLE		FLAGS	
date	Mar 2 2012	hs	nn
solvent	cdcl3	sspul	y
sample		hsglvl	1327
ACQUISITION		SPECIAL	
sw	3048.8	temp	25.0
at	0.168	gain	48
np	1024	spin	0
fb	4000	F2 PROCESSING	
ss	16	sb	-0.084
d1	1.000	sbs	not used
nt	8	fn	1024
2D ACQUISITION		F1 PROCESSING	
sw1	3048.8	sb1	-0.042
ni	128	sbs1	not used
PRESATURATION		proc1	lp
satmode	nnn	fn1	1024
satdly	0	DISPLAY	
satfrq	499.8	sp	869.2
satpwr	-13	wp	3042.8
TRANSMITTER		sp1	869.2
tn	H1	wp1	3042.8
sfrq	399.769	rf1	-863.2
tof	388.8	rfp	0
tpwr	60	rf11	-863.2
pw	10.375	rfp1	0
tpwr_cf	0.977	PLOT	
GRADIENTS		wc	225.0
gzlvl1	1769	sc	6.2
gt1	0.001000	wc2	155.0
gstab	0.000500	sc2	0
DECOUPLER		vs	872
dn	C13	th	11
dm	nnn	ai cdc av	



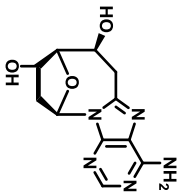
NOESY of Compound **15*R***
Solvent: $CDCl_3$ at 400 MHz

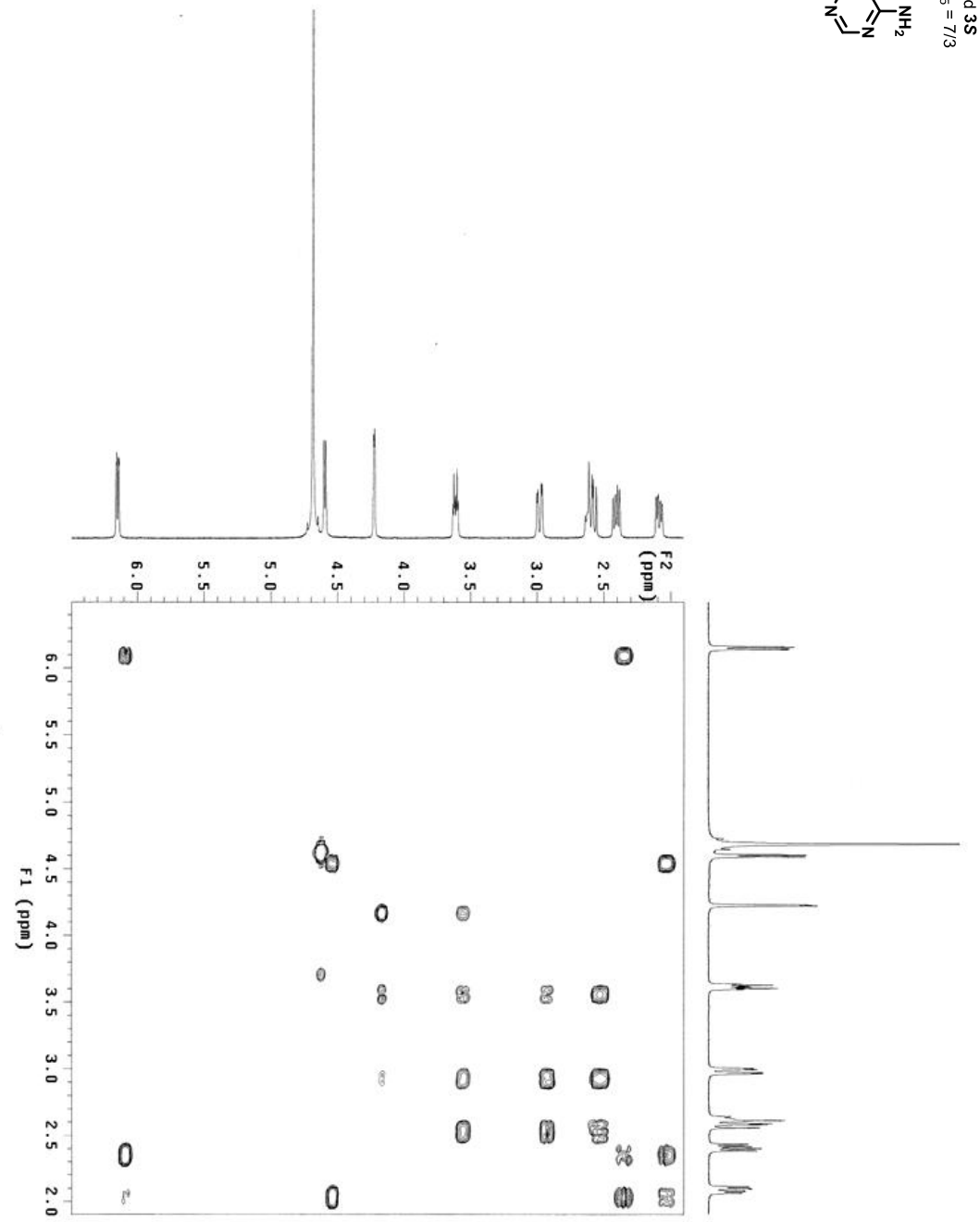



```
expl  Noesy

        SAMPLE
date    Mar  3 2012
solvent       cdcl3
file /home/ALL/lwm~
/Han Yueh/200/2D/H~
Y-212L-NOESY_4_212~
             _02.fid
    ACQUISITION
sfrq         399.769
tn                H1
at            0.168
np             1024
sw           3048.8
fb             4000
bs               32
ss               16
tpwr             60
pw             10.4
compH        0.9774
d1           1.000
mix          0.500
tof            388.9
nt               48
ct               48
gain             52
sspul             y
gt1          0.0025
gzlvl1         1769
hsgt         0.0020
hsglvl         1327
satmode         nnn
satpwr          -13
satdly            0
scuba             n
zqflg             y
alt_grd           n
        FLAGS
il                n
in                n
dp                y
hs               nn
sspul             y
  2D ACQUISITION
sw1          3048.8
ni              128
phase       arrayed
      DISPLAY
sp            869.3
wp           3042.8
vs             5750
sc                6
wc              225
hzmm          13.52
is           744.69
rfl          -863.3
rfp               0
th                4
ins          100.000
ai    ph

        2D DISPLAY
sp1           869.3
wp1          3042.8
sc2               0
wc2             155
rfl1         -863.3
rfp1              0
      DEC. & VT
dfrq         100.531
dn               C13
dpwr              31
dof               0
dm                n
dmm               c
dmf           29412
dseq
dres            9.0
homo              n
temp           25.0
     PROCESSING
gf            0.078
gfs       not used
wtfile
proc             ft
fn             1024
math              f
ssfilter  not used
ssorder           3
ssntaps           3
  2D PROCESSING
gf1           0.039
gfs1      not used
wtfile1
proc1            lp
fn1            1024

ACQUISITION ARRAYS
array         phase
arraydim        256

 i            phase
 1                1
 2                2
```

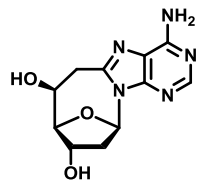


COSY of Compound **3S**
Solvent: D_2O/Pyr-d_5 = 7/3
at 400 MHz

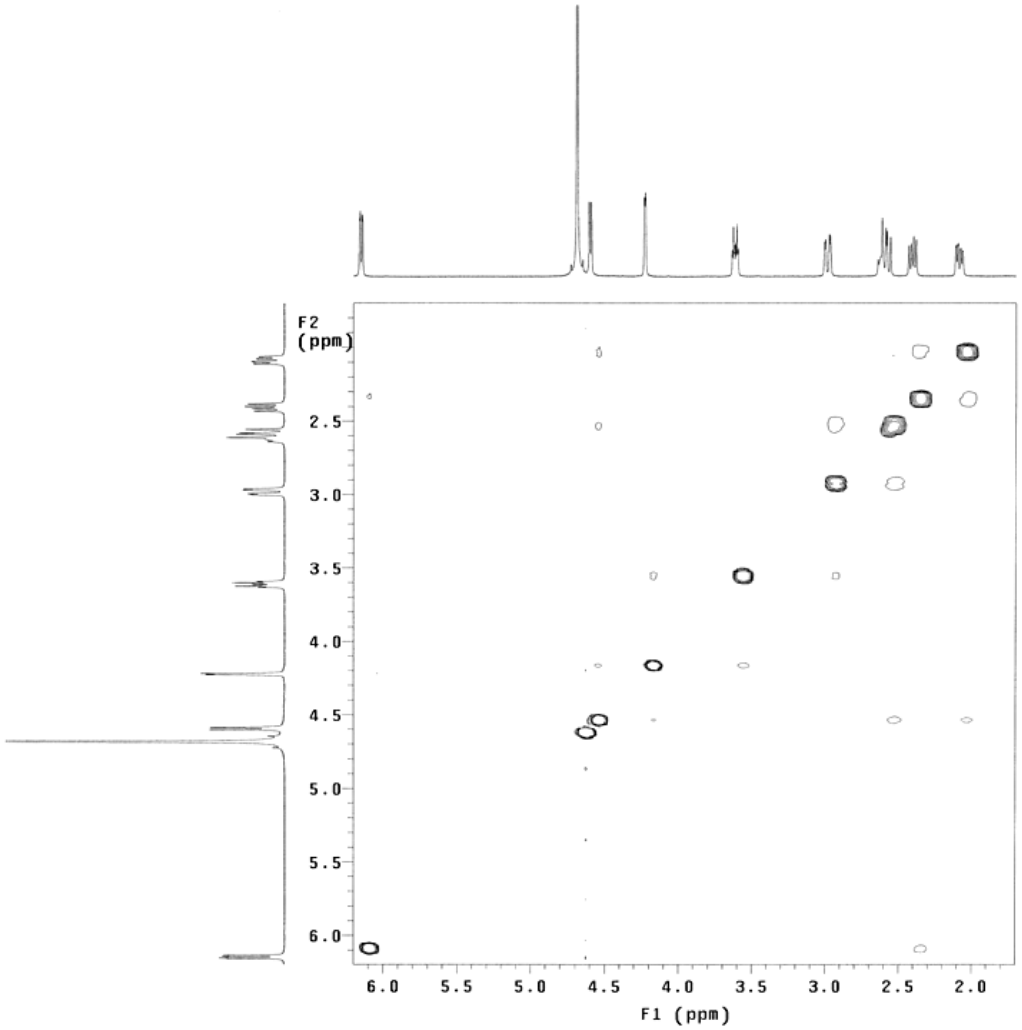




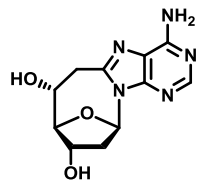
NOESY of Compound **3*S***
Solvent: D_2O/Pyr-d_5 = 7/3
at 500 MHz

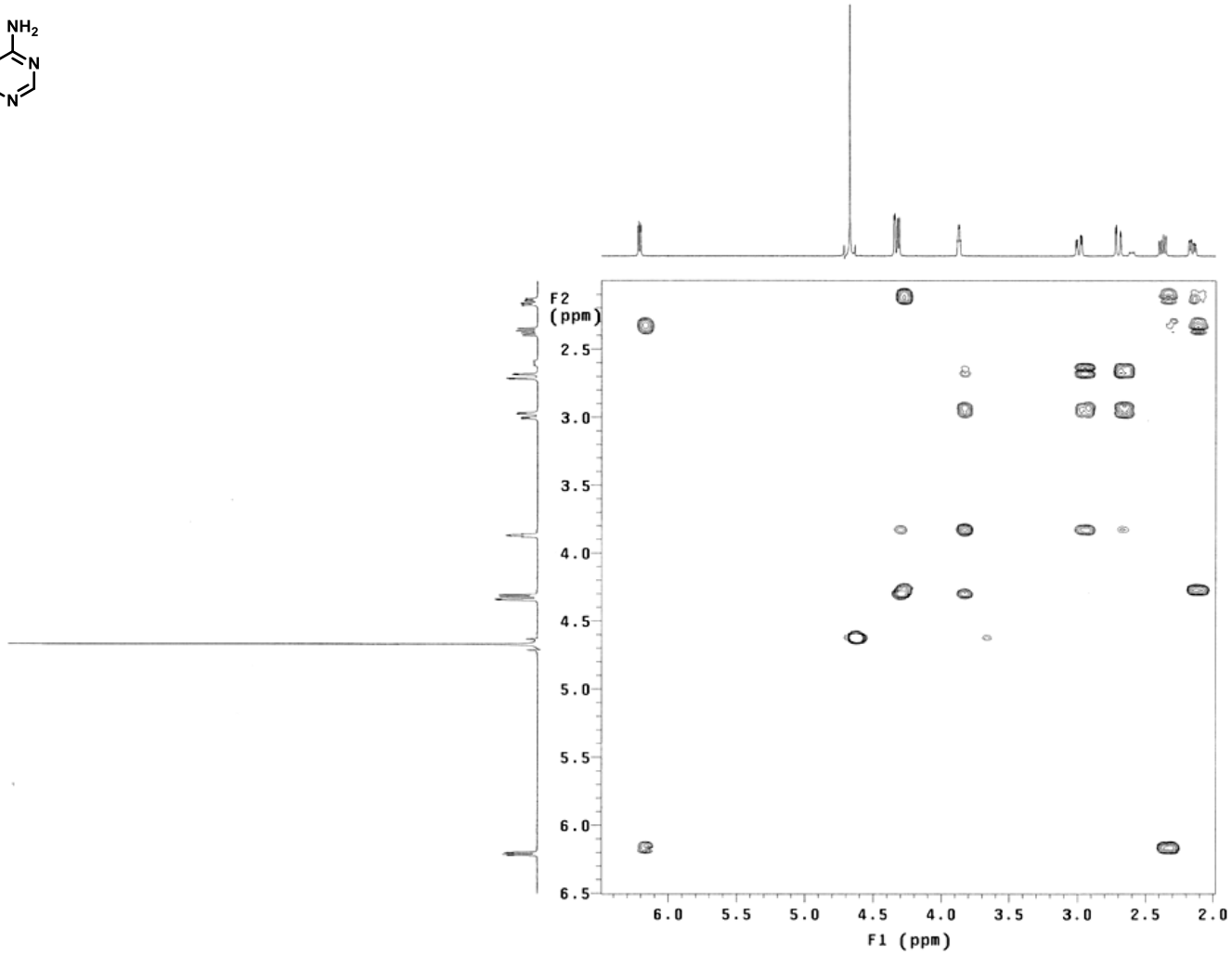



```
 Sample Name:
   HY-214H-NOESY-552
 Archive directory:

 Sample directory:

 FidFile: Noesy

Pulse Sequence: Noesy (NOESY)
Solvent: d2o
Data collected on: Oct 16 2012

 Temp. 25.0 C / 298.1 K
Sample #22, Operator: walkup
INOVA-500 "nmr16"

 Relax. delay 1.000 sec
 Mixing 0.400 sec
 Acq. time 0.135 sec
 Width    3787.9 Hz
 2D Width 3787.9 Hz
 8 repetitions
 2 x 552 increments
OBSERVE   H1, 399.7673042 MHz
DATA PROCESSING
 Gauss apodization 0.062 sec
F1 DATA PROCESSING
 Gauss apodization 0.031 sec
FT size 1024 x 1024
Total time 2 hr, 59 min
```

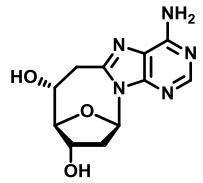


COSY of Compound **3*R***
Solvent: D_2O/Pyr-d_5 = 7/3
at 400 MHz

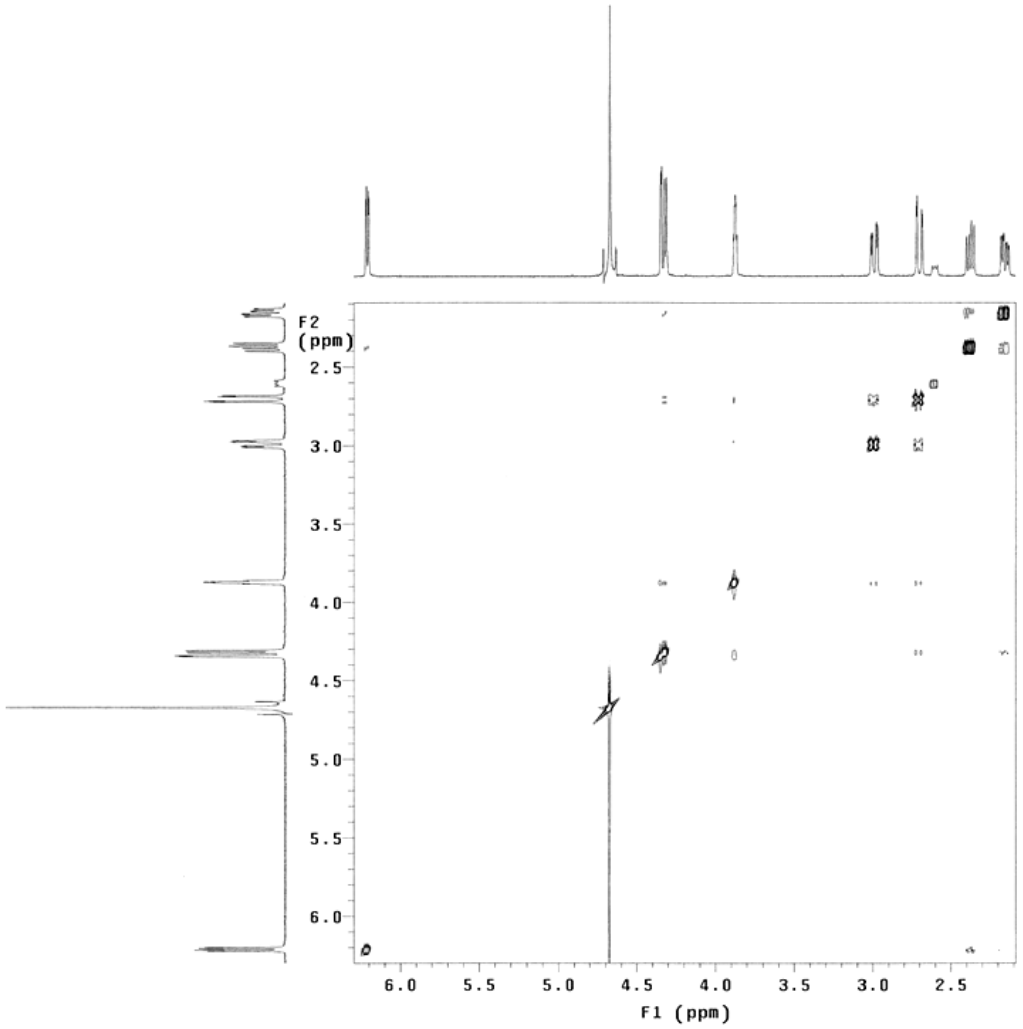




NOESY of Compound **3*R***
Solvent: D_2O/Pyr-d_5 = 7/3
at 500 MHz




```
Sample Name:
  HY214LNOESY256
Archive directory:

Sample directory:

FidFile: Noesy

Pulse Sequence: Noesy (NOESY)
Solvent: d2o
Data collected on: Oct 20 2012

 Temp. 25.0 C / 298.1 K
Operator: lwm
INOVA-500  "nmr16"

 Relax. delay 1.000 sec
 Mixing 0.400 sec
 Acq. time 0.128 sec
 Width    7996.0 Hz
 2D Width 7996.0 Hz
 8 repetitions
 2 x 1240 increments
OBSERVE   H1, 499.7732968 MHz
DATA PROCESSING
 Gauss apodization 0.059 sec
F1 DATA PROCESSING
 Gauss apodization 0.287 sec
FT size 16384 x 16384
Total time 6 hr, 41 min
```



Chapter 4

Alternate Design – 5'-Hydroxymethyl-8,5'-Cyclo-2',5'-Dideoxy Nucleic Acids

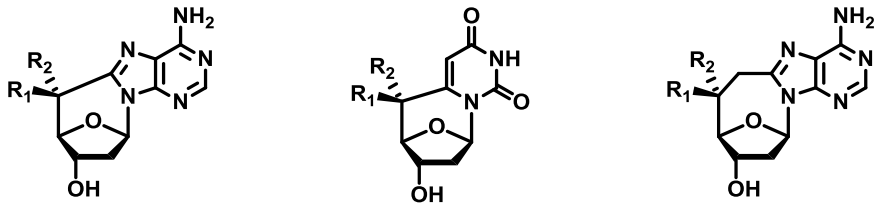
4. Alternate Design – 5'-Hydroxymethyl-Cyclo-2',5'-Dideoxynucleic Acids

4.1. Introduction

In the last two chapters, we successfully synthesized the cyclo-2'-deoxynucleosides and ring-expanded cyclo-2'-deoxynucleosides (Figure 4.1). In chapter 2, we presented the synthesis of the *S* and *R* diastereomers for both 8,5'-cyclo-2'-deoxyadenosines and 6,5'-cyclo-2'deoxyuridines (Figure 4.1, **1*R***, **1*S***, **2*R***, and **2*S***), then incorporated these rigidified nucleosides into a DNA sequence to study their thermal behaviour.[1] The cyclo-2'-nucleosides possessed a second covalent linkage between the sugar and base moieties which restricted the base rotation around the glycosidic bond. This linkage not only rigidified the structure, but "pulled" the nucleobase away from the helical center, which resulted in the destabilization of the double helix structure. In nature, 8,5'-cyclo-2'-deoxyadenosines can be formed through oxidative damage to DNA. It is known that these cyclo-nucleosides inhibit binding of proteins such as the TATA binding protein.[2,3]

In chapter 3, we discussed the synthesis of a new generation of rigidified nucleosides, the ring-expanded cyclo-nucleosides. When compared to the cyclo-nucleosides in chapter 2, we inserted an extra carbon onto the bridge between the base and the sugar to expand the cyclized ring on these modified nucleosides (Figure 4.1, **3*R*** and **3*S***). By introducing an extra

methylene into the cyclized linkage, we could "push" the base back towards its natural position and the linkage should still decrease the entropic penalty associated with rotation around the glycosidic bond. The stereochemistry of the newly generated chiral center on the C5' position was characterized by 2D-NMR experiements.



1*R*: R_1 = H, R_2 =OH
1*S*: R_1 = OH, R_2 =H

2*R*: R_1 = H, R_2 =OH
2*S*: R_1 = OH, R_2 =H

3*R*: R_1 = H, R_2 =OH
3*S*: R_1 = OH, R_2 =H

Figure 4.1. The structures of 8,5'-cyclo-2'-deoxyadenosines (**1*R***, **1*S***), 6,5'-cyclo-2'deoxyuridines (**2*R***, **2*S***), and ring-expanded 8,5'-cyclo-2'-deoxynucleosides (**3*R*** , **3*S***).

Here we discuss three different strategies to synthesize another class of rigidified nucleosides, 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosines (Figure 4.2.a). The goal of synthesizing these modified nucleosides was to generate nucleosides with better thermal stability during duplex formation by having better base-paring accessibility and minimized entropic cost. The new design has an extra carbon atom (C6') to elongate the sugar backbone. The primary 6'-OH can adopt a variety of conformations, not like the rigidified structure in cyclo-nucleosides (**1*R***, **1*S***, **2*R***, and **2*S***) which have secondary hydroxyls on the C5' position fixing their orientation (Figure 4.2.b). It gives the structure more flexibility during duplex hybridization, which

might result in better accessibility to from hydrogen bonds with their base-pairing partners than the first generation of cyclo-nucleosides.[4] This primary 6'-OH also allows us to utilize the conventional acid-labile trityl protecting group, DMTr. Conventionally available phosphoramidite reagents (chapter 2) contain an acid labile DMTr group at the O5', while for the 8,5'-cyclo-2'-deoxyadenosines and 6,5'-cyclo-2'deoxyuridines, we needed to utilize EVE instead of the bulky DMTrCl to protect the O5' position due to the steric hindrance on the cyclized structure.[1] One drawback of this 1-ethylethoxy group was its lack of UV activity, therefore the coupling yield could not be

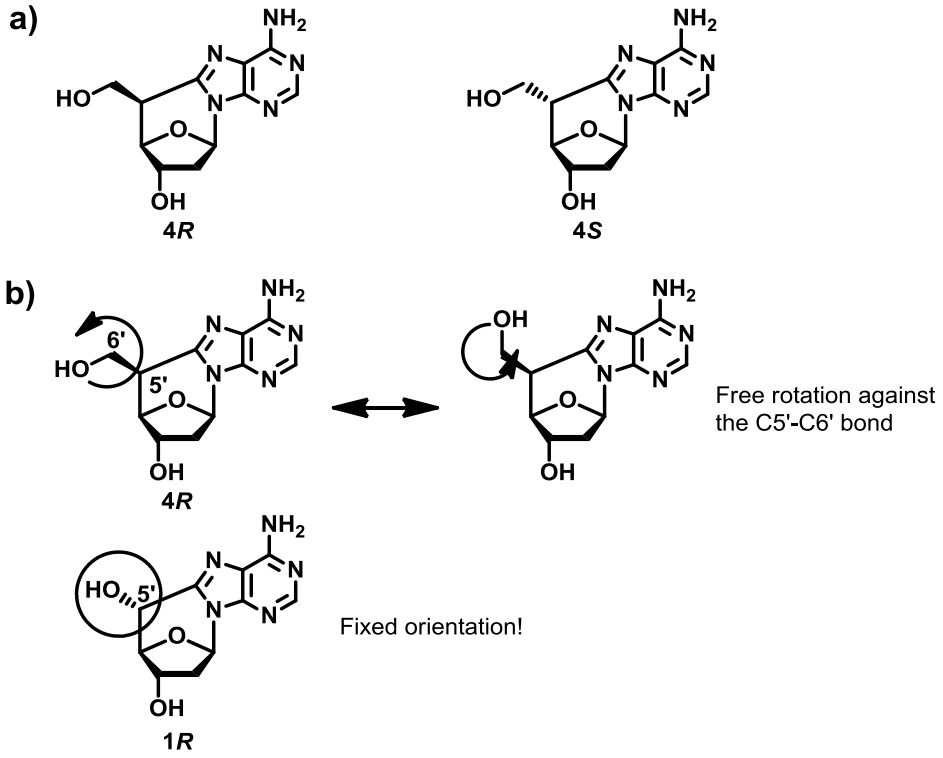


Figure 4.2. a) The structures of 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosines. b) 6'-OH rotates around the C5' – C6' bond in **4*R***, and the fixed orientation of the 5'-OH on **1*R***.

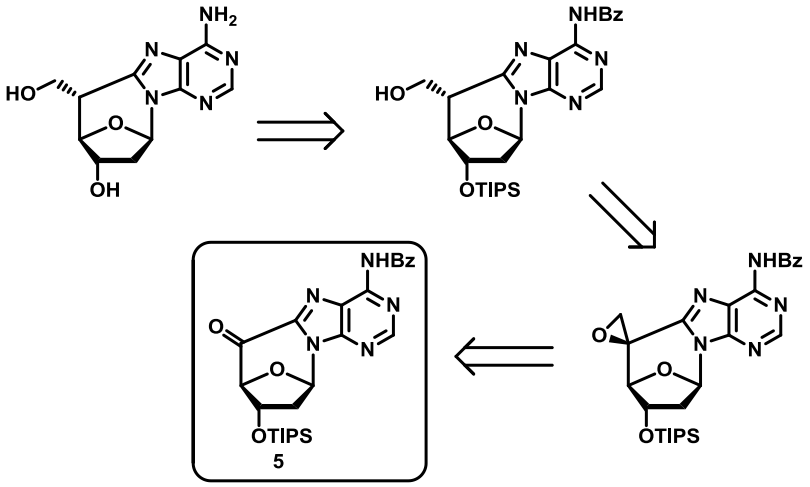
monitored during DNA synthesis. Another feature of these new rigidified nucleosides is that they still possess the secondary linkage to restrict the nucleobase rotation around the glycosidic bond, which reduces the entropic effect compared to the native nucleosides.

4.2. Evaluation of the Synthetic Strategy

There were several studies on the syntheses and biological characterization of the C5'-hydroxymethyl nucleosides in the past few decades. In 1960's, the 5'-hydroxymethyl-5'-deoxyadenosine and the 5'-hydroxymethyl-2',5'-dideoxythymidine had already been synthesized.[5,6] In 2005, the 5'-hydroxylmethyl-5'-deoxyadenosine (5'-homoaristeromycin) had been made by the Schneller group who also investigated the antiviral activity against orthopox viruses.[7] However a practical synthetic strategy for 5'-hydroxymethyl-cyclo-nucleic acids was still needed. Here we discuss three different strategies of introducing an extra methylene group onto the C5' position of the 8,5'-cyclo-2'-deoxynucleosides to discover the possibility of synthesizing these novel rigidified nucleosides.

The first synthetic strategy that we investigated started with one important intermediate from the 8,5'-cyclo-2'-deoxyadenosines synthesis, the

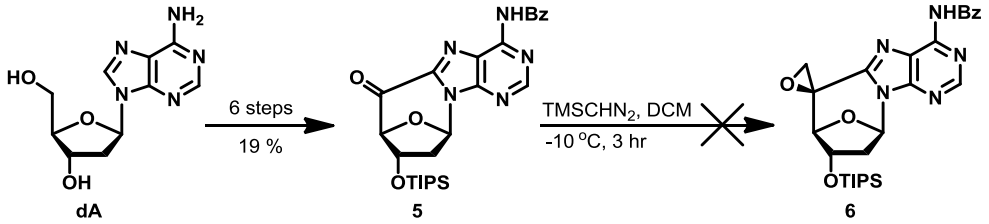
5'-keto-8,5'-cyclo-2'-deoxyadenosine (compound **5**).[1] We first wanted to generate an epoxide on the C5' position of compound **5** toward the carbonyl group, then followed by reducing the 5'-spiro-epoxide into 5'-hydroxymethyl (Sheme 4.1).[8,9] It might give two diastereomers after the epoxidation reaction, but due to the steric hindrance of the structure the nucleophile could only attack the C5' from the less hindered position to give one diastereomer exclusively.[8]



Scheme 4.1. Retrosynthesis of 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxy-adenosine. The synthesis started with compound **5**, an intermediate of the 8,5'-cyclo-2'-deoxyadenosine synthesis.

Compound **5** was obtained by a six-step synthesis from deoxyadenosine with 19% yield, detail of the procedures were described in chapter 2.[1] To perform the epoxidation on the carbonyl position, we treated compound **5** with trimethylsilyldiazomethane in dichloromethane at low temperatures. Trimethylsilyldiazomethane is a widely used methylation reagent as a less-

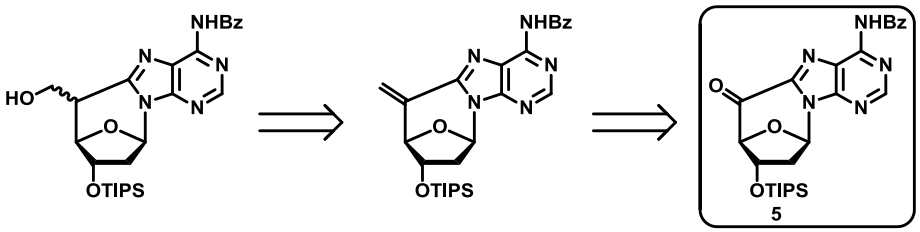
explosive replacement for diazomethane.[10] Unfortunately, we could not install the epoxide on the C5' position (Scheme 4.2) to get to compound **6**, although this pathway had been established for the cyclo-uridine derivatives.[11] The reason could be the exo-cyclic amide group on the C6 position was interfering with the reagent.[12]



Scheme 4.2. First synthetic strategy for 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosine through epoxidation on the carbonyl group of compound **5**.

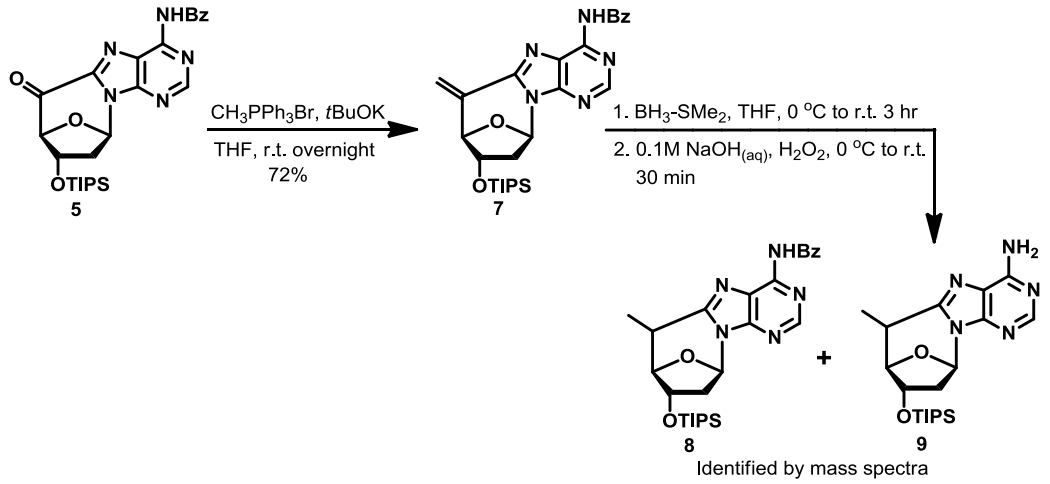
In the second strategy, we tried to insert the extra carbon and oxygen in different steps. First, the extra methylene group was installed onto the C5' of compound **5**, then hydroxylated at the C6' position to obtain the hydroxymethyl compound (Scheme 4.3.). The benefit of this route was that both diastereomers would be obtained at the end of the synthesis and it also started with the same intermediate, compound **5**. This strategy has also been shown as practical on furanyl scaffolds.[13]

The reaction went through the Wittig olefination to insert the extra carbon on C5' position (Scheme 4.4). The Wittig reagent was prepared from the phosphonium salt, methyltriphenylphosphonium bromide, reacting with



Scheme 4.3. Retrosynthesis of the second strategy for synthesizing 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosine.

potassium *tert*-butoxide (*t*BuOK) to form the triphenyl phosphonium ylide. This was then reacted with the carbonyl group on compound **5** to get compound **7**. The next step was to oxidize the terminal olefin into the hydroxymethyl group by hydroboration-oxidation reaction. Hydroboration-oxidation is a two-step reaction that converts an alkene group into alcohol. It is a syn-addition reaction leading to *cis*-conformation with the hydroxyl group attaching to the less-substituted carbon.[14] In the first attempt, we used 9-borabicyclo[3.3.1]nonane (9-BBN), a more stable borane derivative, to possibly give only one dominant diastereomer due to the highly stereoselective nature.[15] Unfortunately, the 9-BBN hydroboration did not work due to the very rigid structure of the cyclo-nucleoside (**7**). We then tried a less hindered, non-stereoselective borane reagent, dimethylsulfide borane (BH_3-DMS).[16] The reaction was unable to give the desired product but generated compound **8** and **9** instead (Scheme 4.4). Compound **8** and **9** were identified from the doublet -CH_3 peak on the H-NMR of the crude product and the mass spectra.

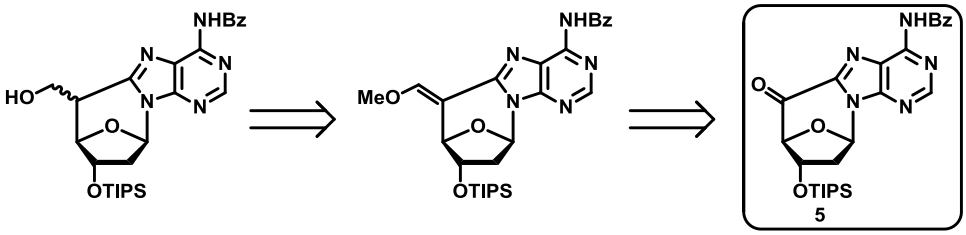


Scheme 4.4. Second synthetic strategy for synthesizing 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosine.

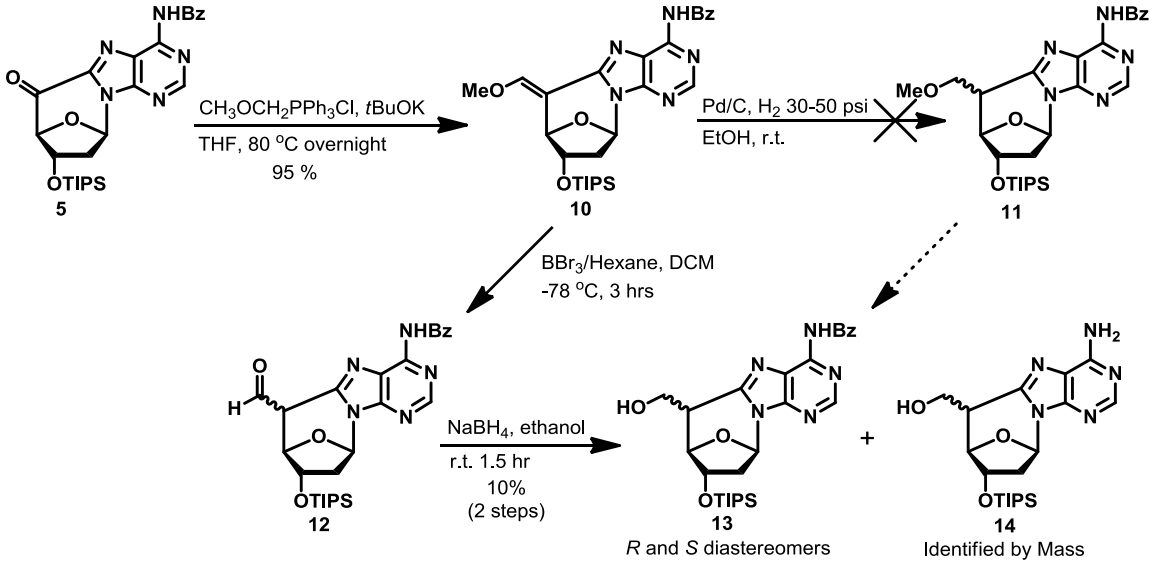
Since the Wittig reaction was shown to be successfully applied to our second strategy, in the third synthetic strategy we chose to use another Wittig reaction to attach both C6' and O6' in one reaction (Scheme 4.5). After the installation of the methoxyethylene group, the 5'-hydroxymethyl-cyclo-nucleosides could be generated by the reduction of the C-C double bond and the removal of the methoxy group. Both diastereomers (5'-hydroxymethyl-8,5'(*R*)-cyclo-2',5'-dideoxyadenosine and 5'-hydroxymethyl-8,5'(*S*)-cyclo-2',5'-dideoxyadenosine) could be obtained from this synthetic strategy.

(Methoxymethyl)triphenylphosphonium bromide was first reacted with *t*BuOK to generate the Wittig reagent. Then the THF solution of compound **5** was added into the previous mixture and reacted under elevated temperature to produce compound **10** in high yield. The Wittig reaction exclusively gave a

single isomer of compound **10**, as shown using NMR. After getting the methoxyethylene compound (**10**), there were two different routes to get the target 5'-hydroxymethyl-cyclo-nucleosides: a) first, the reduction of the C5'-C6' double bond, then deprotection of the 6'-methoxy; b) second, removal of the protecting group first, then the reduction.



Scheme 4.5. Retrosynthesis of the third strategy for synthesizing 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosine with a different Wittig reaction to insert C6' and O6' in one step.



Scheme 4.6. The synthetic scheme of the third strategy for synthesizing 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosine.

We first tried to reduce the C5'-C6' double bond with hydrogen catalyzed by platinum on charcoal. The reaction was started with 30 psi of hydrogen pressure, there was no significant change after reacting overnight and monitoring by TLC. Then the hydrogen pressure was increased to 50 psi for another period of 12-18 hours generating a new spot which had a slightly higher polarity then the starting material. This was the debenzoylated side product as identified by the mass spectrum of the crude reaction mixute, but the majority was remaining starting material, compound **10**. Then we performed an alternate route b to remove the methoxy group first. The methoxy deprotection was carried out with boron tribromide (BBr_3), a strong Lewis acid, in hexane.[17] The resulting compound **10** was first dissolved in DCM, then the solution was brought to -78 ^oC in a dry ice/acetone bath. BBr_3 was then added as a solution in hexane into the cold DCM solution. After 3 hours TLC indicated completion of the reaction. The resulting compound **12** was not purified since the aldehyde compound was found to be unstable under silica gel column conditions. After the workup, the crude mixture was treated with sodium borohydride in ethanol to reduce the aldehyde group generating both *R* and *S* diastereomers. The combined yield for **13*R*** and **13*S*** was not optimal, and there were some debenzoylated side products (compound **14**) generated which could be identified by mass spectrometry. These two diastereomers could be barely separated by TLC, and since they were unable to be purified by column chromatography, we utilized

preparative TLC to separate them instead. According to the relation between the structure and the polarity, we assumed that the compound developing lower on the TLC, which had higher polarity, should be the *S* – diastereomer due to the more accessible hydroxyl group to the silica gel on the 6' position; the higher spot on the TLC should be the *R* diastereomer (Figure 4.3).

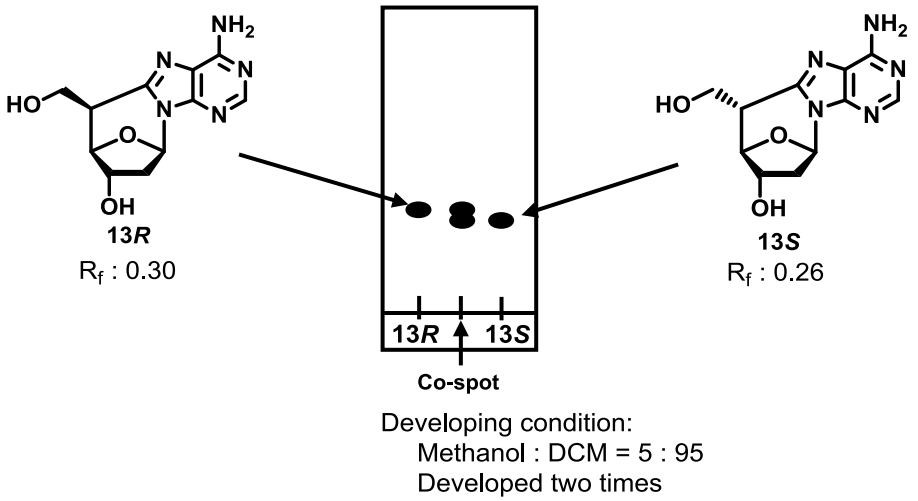


Figure 4.3. TLC of the compound **13*R*** and **13*S***. The plate was developed in 5% methanol in DCM, two times in order to observe the polarity difference of these two diastereomers.

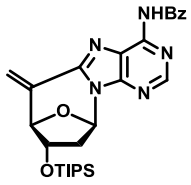
4.3. Conclusions

In this chapter we presented the possible synthetic strategies for synthesizing the novel rigidified nucleosides: the 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosines. We discussed and performed three different strategies and 4 synthetic routes to investigate the possibility of synthesizing both diastereomers. The first strategy showed that the epoxidation with diazomethane based reagent on the cyclo-deoxyadenosine was not feasible. In the second strategy, we learned the hydroboration-oxidation could not happen on the cyclo-deoxyadenosine derivative, but we learned that the Wittig reaction was a practical way to install varied functional groups onto the C5' position of the cyclo-nucleoside. Utilizing the versatile intermediate, compound **5**, we established a workable pathway to synthesize the 5'-hydroxymethyl-8,5'-cyclo-2',5'-dideoxyadenosine derivatives of both diastereomers. It could be a valuable reference for people who are interested in investigating the total synthesis and the biophysical properties of these rigidified nucleosides.

4.4. Experimental

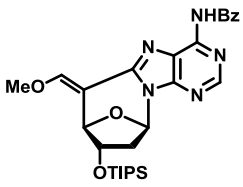
Reagents were purchased from Sigma-Aldrich, Acros, Oakwood, MP Biomedical, Chem-Impex International, Fisher. Flash column were performed using Dynamic Adsorbents silica gel (60 Å, particle size 32-63 μm), Preparative TLC was from Analtech (1000 microns with F-254 Indicator), TLC monitoring with TLC Silica Gel with F-254 Indicator (Dynamic Adsorbents). TLCs were visualized by 260nm UV light and stained by 10% sulfuric acid. Dry tetrahydrofuran (THF) and dichloromethane (DCM) were obtained by passing commercially available pre-dried, oxygen-free formulations through activated alumina columns. Dry ethanol was purchased from Sigma-Aldrich and used directly. NMR spectra were taken by Varian VNMRS400, VNMRS500, or INOVA 500 instruments and calibrated using residual undeuterated solvent (CDCl3: δ H = 7.24 ppm, δ C = 77.00 ppm). Abbreviations of multiplicities were designated as follow: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. High-resolution mass spectra (HRMS) were recorded on a Waters LCT or JEOL AccuTOF mass spectrometer using ESI (electrospray ionization) or DART (direct analysis in real time).

Compound 7



Methyltriphenylphosphonium bromide (0.35 g, 0.98 mmol) was dissolved in dry THF (3.5 mL) under room temperature, and *t*BuOK (0.10 g, 0.86 mmol) was slowly adde into the solution. The mixture was allowed to react for an hour then a THF solution of Compound **5** (0.125 g, 0.25 mmol in 1 mL THF) was added into the reaction mixture. After stirring for overnight under ambient temperature, the reaction was quenched with concentrated $NH_4Cl_{(aq)}$ (2mL). DCM (10 mL) was added and the solution was extracted extracted with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was purified by flash chromatography (DCM:MeOH = 97.5:2.5) to yield compound **7** as a white solid (0.89 g, 72%). TLC R_f: 0.40 (DCM:MeOH = 97.5:2.5). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.03-1.12 (m, 21H), 2.43 (d, J = 12.5 Hz, 1H), 2.68 (m, 1H), 4.57 (d, J = 4.0 Hz, 1H), 5.08 (s, 1H), 5.54 (s, 1H), 6.44 (s, 1H), 6.66 (d, J = 5.5 Hz, 1H), 7.52 (t, J = 7.0 Hz, 2H), 7.59 (t, J = 7.0 Hz, 1H), 8.00 (d, J = 6.5 Hz, 2H), 8.74 (s, 1H), 9.03 (s, 1H). HRMS (pos.): 506.2591 $[M+H]^+$ (HSMS calc. 506.2587).

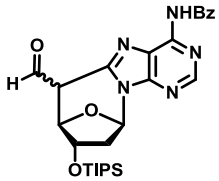
Compound 10



Methoxymethyltriphenylphosphonium bromide (1.51 g, 4.4 mmol) was dissolved in THF (16 mL) under room

temperature, and *t*BuOK (0.43 g, 3.85 mmol) was slowly adde into the solution. The mixture was allowed to react for an hour then a THF solution of Compound **5** (0.56 g, 1.10 mmol in 4 mL THF) was added into the reaction mixture. After stirring for overnight under ambient temperature, the reaction was quenched with concentrated $NH_4Cl_{(aq)}$ (10mL). DCM (25 mL) was added and the solution was extracted extracted with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was purified by flash chromatography (DCM:MeOH = 97.5:2.5) to yield compound **10** as a white solid (0.58 g, 95%). TLC R_f: 0.25 (DCM:MeOH = 97:3). ^{1}H NMR (500 MHz, $CDCl_3$): δ 1.06-1.13 (m, 21H), 2.43 (m, 1H), 2.63 (m, 1H), 3.91 (s, 3H), 4.52 (dd, *J* = 14.0, 6.5 Hz, 1H), 5.40 (s, 1H), 6.61 (d, *J* = 5.5 Hz, 1H), 7.46 (td, *J* = 7.5, 3.0 Hz, 1H), 7.51 (m, 2H), 7.59 (t, *J* = 7.5 Hz, 1H), 7.66 (m, 1H), 8.00 (d, *J* = 7.5 Hz, 2H), 8.63 (s, 1H), 8.98 (s, 1H). HRMS (pos.): 536.2702 $[M+H]^+$ (HSMS calc. 536.2693).

Compound 12



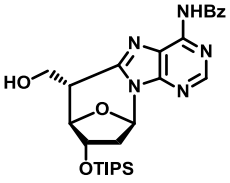
Compound **10** (0.1 g, 0.2 mmol) was dissolved in dry DCM (52 mL) and cooled to -78 °C in dry ice/acetone bath. Boron tribromide/hexan solution (1 M, 2.0 mL) was added and the reaction was kept stirring at -78 °C. After reacting for 3 hours the reaction was quenched with diethyl ether (5 mL) on dry ice/acetone bath for

10 minutes. The mixture was diluted with DCM (50 mL) and extracted with concentrated $NaHCO_{3(aq)}$ once, water three times and once with brine soltion. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was staying on high vacuum for overnight then directly went onto next step.

Compound 13*S* and **13*R***

The crude compound **12** (started from 0.5 mmol of compound **10**) and sodium borohydride (19 mg, 0.5 mmol) were dissolved in dry ethanol (100 mL). The reaction mixture stirred under room temperature for one and half hours then quenched with water. The mixture was diluted with DCM and extracted with water three times and once with brine. The organic layer was dried over sodium sulfate and concentrated under vacuum. The crude product was purified by preparative TLC (DCM:MeOH = 95:5) to yield compound **13*S*** and **13*R*** as a white powder (**13*S*** : 15 mg, **13*R*** : 10 mg, total yield: 10%). TLC R_f: **13*S*** : 0.26, **13*R*** : 0.31 (DCM:MeOH = 95:5, developed twice).

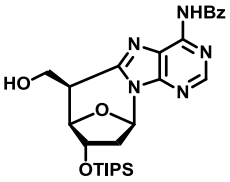
13*S*



^{1}H NMR (400 MHz, $CDCl_3$): δ 1.06-1.10 (m, 21H), 1.94 (dd, *J* = 16.0, 3.5 Hz, 1H), 2.12 (m, 2H), 2.47 (m, 1H), 2.57 (d, *J* = 11.5 Hz, 1H), 3.86 (d, *J* = 9.5 Hz, 1H), 4.33 (d, *J* = 10.0

Hz, 1H), 4.67 (d, J = 9.5 Hz, 1H), 7.57 (t, J = 9.5 Hz, 2H), 7.67 (t, J = 9.5 Hz, 1H), 7.98 (td, J = 9.5, 1.0 Hz, 2H), 8.71 (s, 1H), 8.89 (s, 1H). ^{13}C NMR (100MHz, $CDCl_3$): δ 12.8, 18.0, 29.7, 42.2, 47.6, 50.4, 70.2, 74.1, 108.0, 113.4, 127.7, 129.3, 133.6, 146.3, 152.2. HRMS (pos.): 524.2690 $[M+H]^+$ (HSMS calc. 524.2693).

13*R*



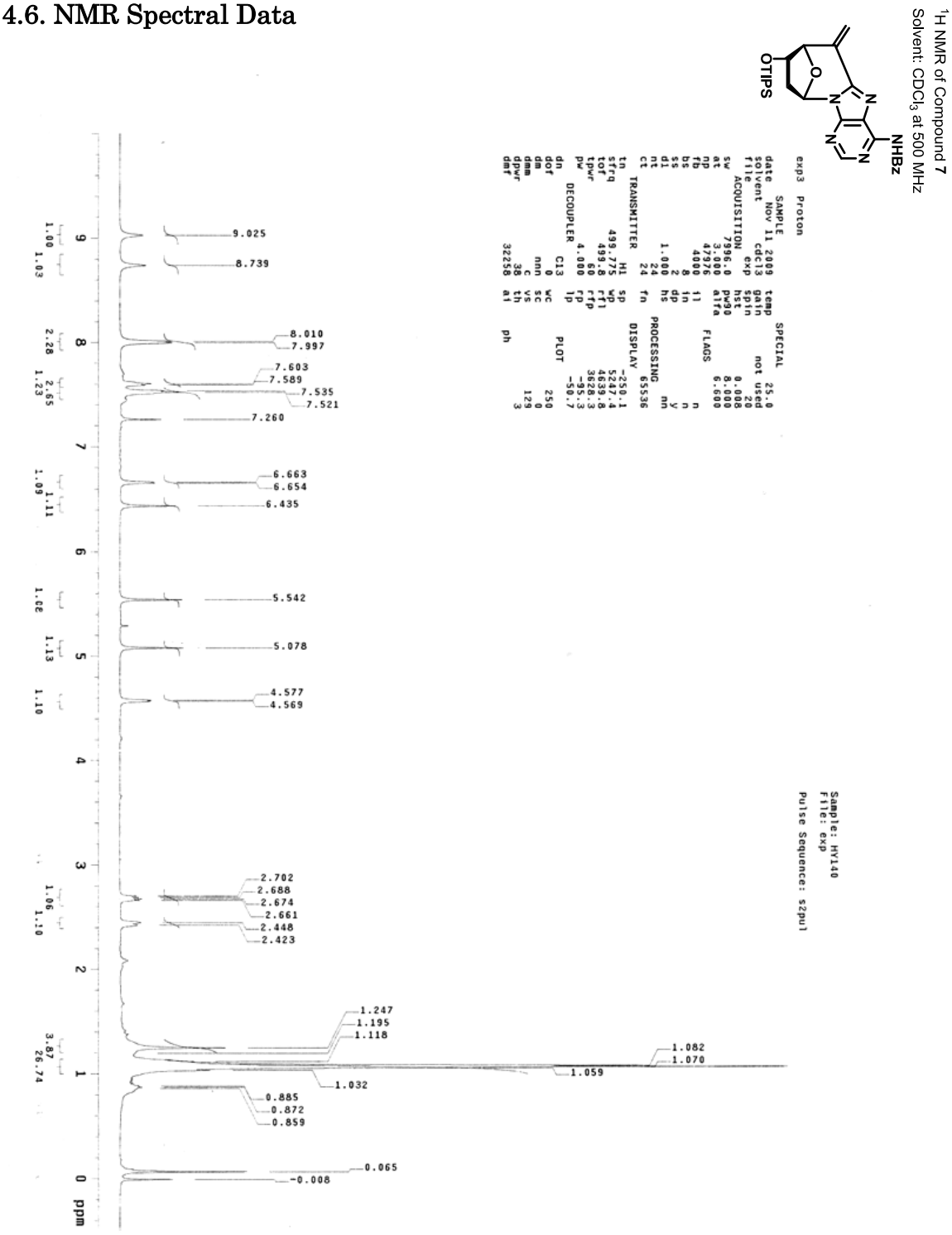
δ 1.06-1.09 (m, 21H), 1.94 (dt, J = 16.0, 5.0 Hz, 1H), 2.13-2.17 (m, 2H), 2.39 (dd, J = 16.0, 13.0 Hz, 1H), 2.50 (dd, J = 12.0, 4.5 Hz, 1H), 4.02 (dd, J = 9.0, 2.0 Hz, 1H), 4.70-4.75 (m, 2H), 7.57 (td, J = 9.0, 2.0 Hz, 2H), 7.67 (tt, J = 9.0, 1.5 Hz, 1H), 7.99 (dt, J = 9.0, 1.5 Hz, 2H), 8.72 (s, 1H), 8.90 (s, 1H). ^{13}C NMR (100MHz, $CDCl_3$): δ 12.8, 18.0, 29.7, 44.1, 44.8, 50.6, 68.9, 73.0, 108.0, 113.4, 127.7, 129.3, 132.1, 133.6, 142.3, 152.2, 157.0, 162.6, 166.5. HRMS (pos.): 524.2669 $[M+H]^+$ (HSMS calc. 524.2693).

4.5. References

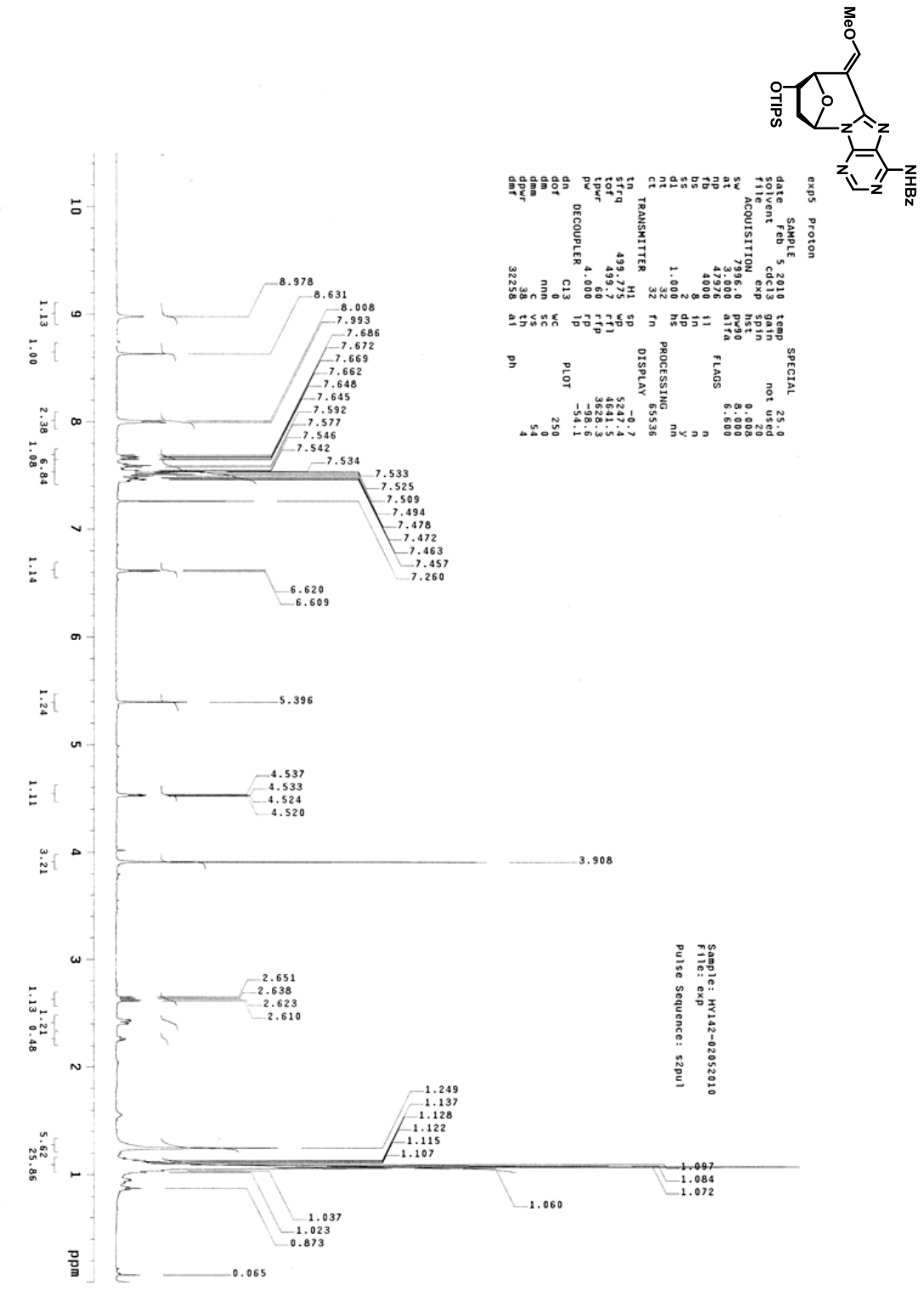
1. H. Yueh, H. Yu, C. S. Theile, A. Pal, A. Horhota, N. J. Greco, L. W. McLaughlin, *Nucleosides, Nucleotides, and Nucleic Acids*. **2012**, *31*, 661-679.
2. L. Kuraoka, C. Bender, A. Romieu, J. Cadet, R. D. Wood, T. Lindahl, *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3832-3837.
3. C. Marietta, H. Gulam, P. J. Brooks, *DNA Repair*, **2002**, *1*, 967-975.
4. R. Zhang, L. A. Eriksson, *Chem. Phy. Lett.* **2006**, *417*, 303-308.
5. K. J. Ryan, H. Arzoumanian, E. M. Acton, L. Goodman, *J. Am. Chem. Soc.* **1964**, *86*, 2503-2508.
6. G. Etzold, G. Kowollik, O. Langen, *Chem. Comm.* **1968**, 422.
7. M. Yang, S. W. Schneller, *Bioorg. Med. Chem. Lett.* **2005**, *15*, 149-151.
8. B. A. Otter, E. A. Falco, *Tet. Lett.* **1978**, *20*, 4383-4386.
9. I. M. Sasson, B. A. Otter, *J. Heterocyclic Chem.* **1987**, *24*, 1439-1444.
10. D. Seyferth, A. W. Dow, H. Menzel, T. C. Flood, *J. Am. Chem. Soc.* **1968**, *90*, 1080-1082.
11. I. M. Sasson, B. A. Otter, *J. Org. Chem.* **1981**, *46*, 1114-1120.
12. H. Nishiyama, H. Nagase, K. Ohno, *Tet. Lett.* **1979**, *20*, 4671-4674.
13. S. Bera, V. Nair, *Tetrahedron*, **2002**, *58*, 4865-4871.
14. H. C. Brown, G. Zweifel, *J. Am. Chem. Soc.* **1959**, *81*, 247.
15. J. A. Soderquist, *J. Org. Chem.* **1981**, *46*, 4599-4600.
16. J. A. Marshall, A. M. Mikowski, *Org. Lett.* **2006**, *8*, 4375-4378.

17. R. Dharanipragada, G. Fodor, *J. Chem. Soc. Perkin Trans. I* **1986**, 545-550.

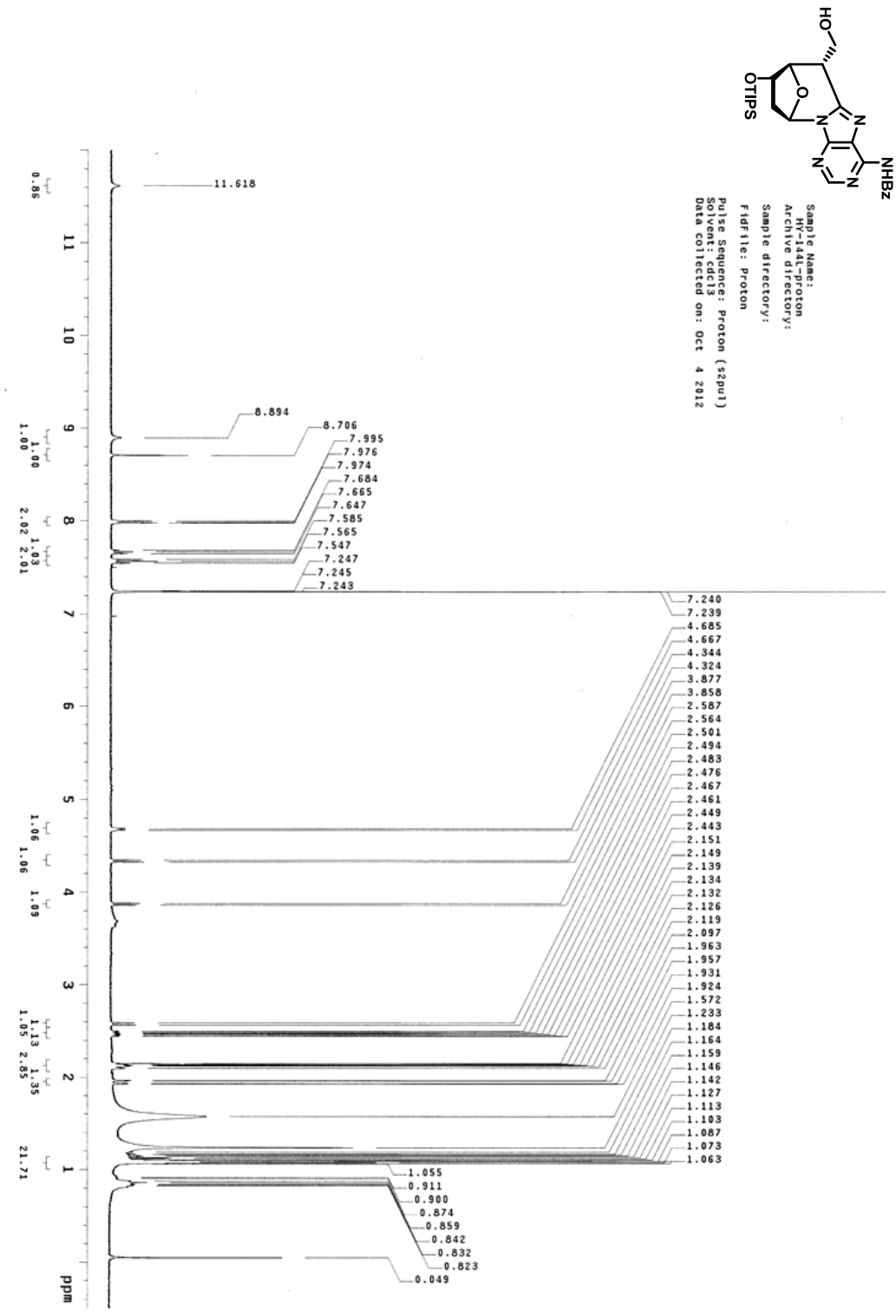
4.6. NMR Spectral Data



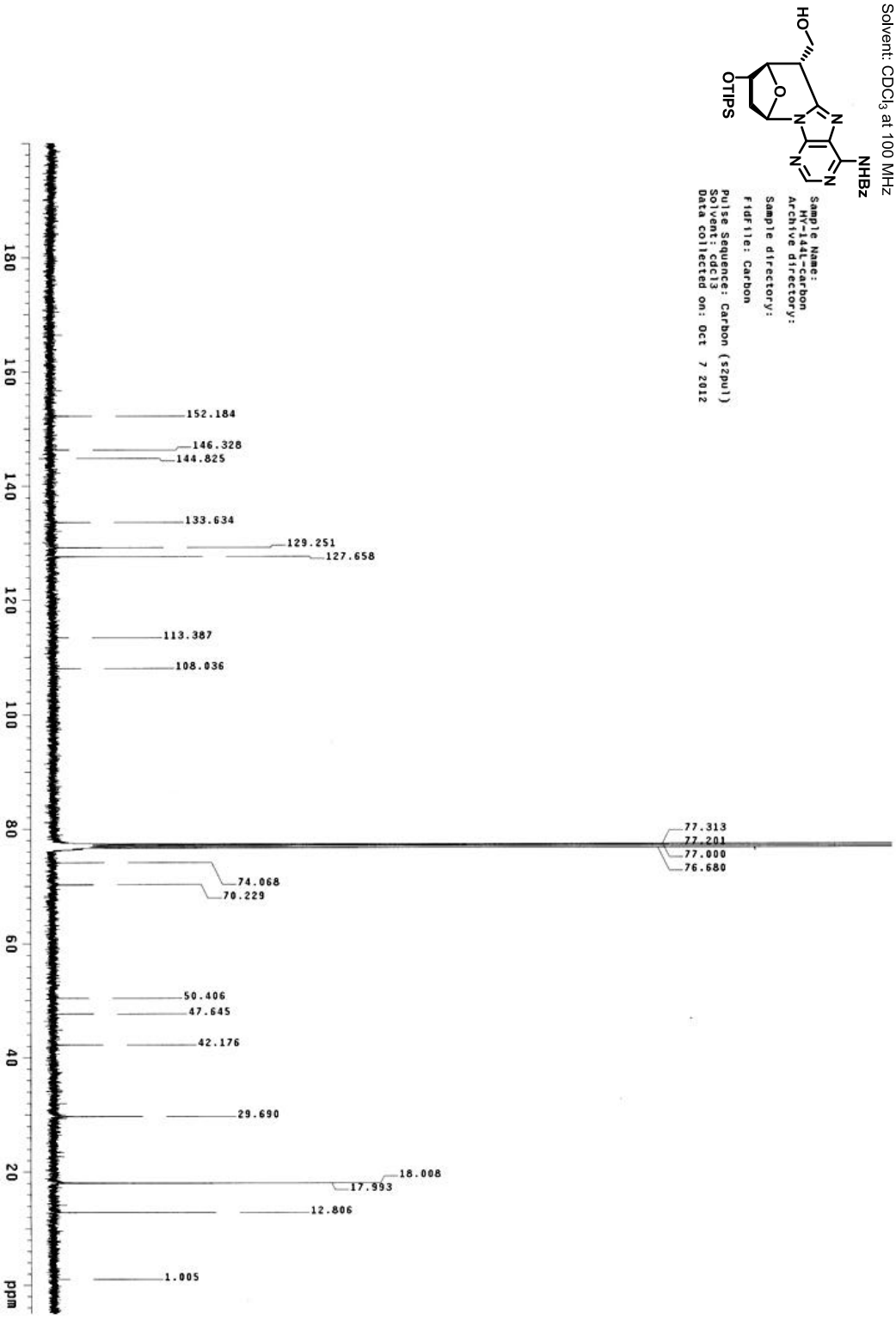
^{1}H NMR of Compound **10**
Solvent: $CDCl_3$ at 500 MHz



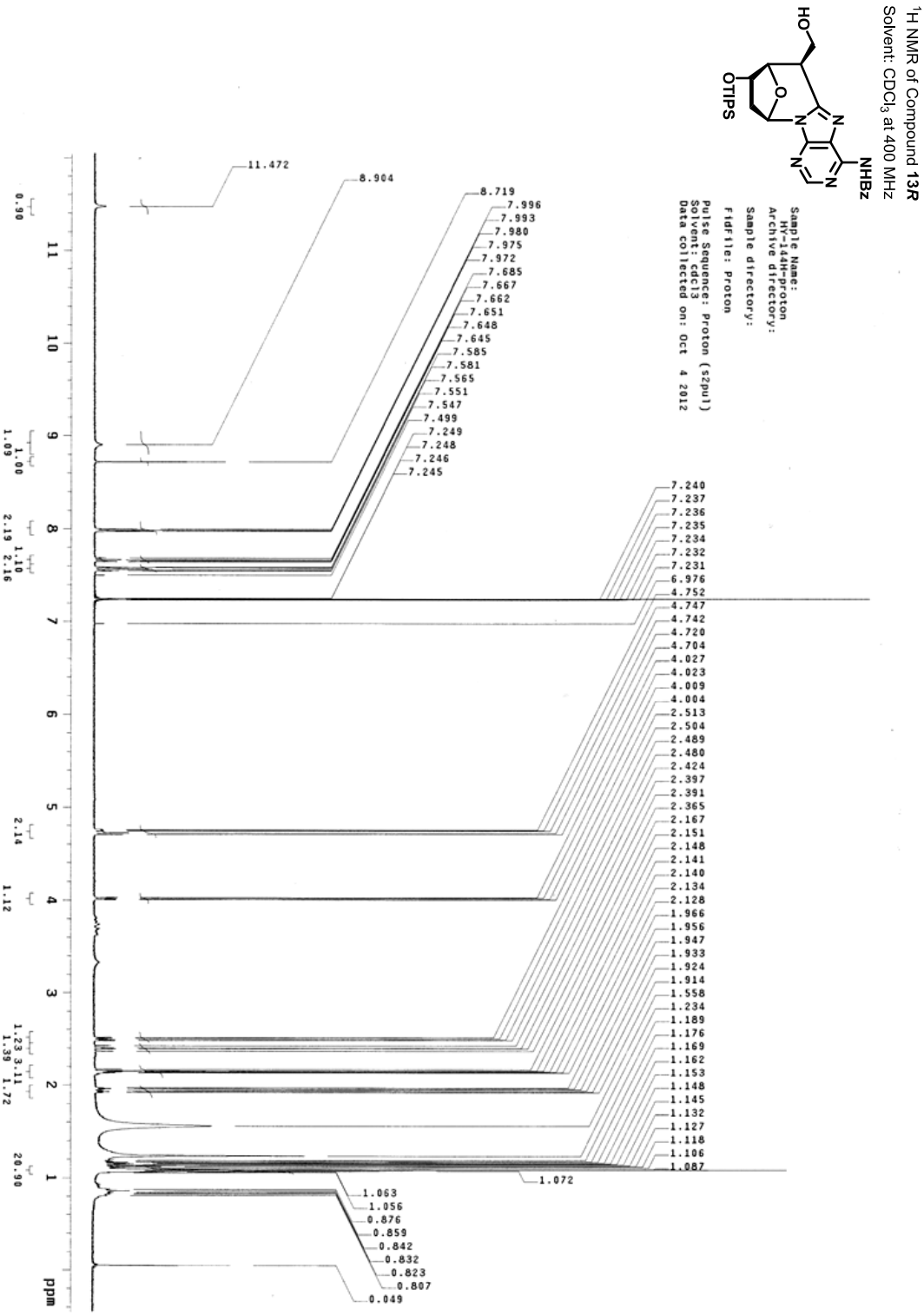
^{1}H NMR of Compound **13S**
Solvent: $CDCl_3$ at 400 MHz



^{13}C NMR of Compound **13S**
Solvent: $CDCl_3$ at 100 MHz



^{1}H NMR of Compound **13*R***
Solvent: $CDCl_3$ at 400 MHz



^{13}C NMR of Compound **13*R***
Solvent: $CDCl_3$ at 100 MHz