UNIVERSITY OF CALIFORNIA
RIVERSIDE

Syntheses and Characterizations of Modified Nucleosides and Their Incorporation into Oligodeoxynucleotides

Doctor of Philosophy

in

Chemistry

by

Mario Rolando Vargas

June 2012

Dissertation Committee:

Dr. Yinsheng Wang, Chairperson
Dr. Eric Chronister
Dr. Guy Bertrand

The Dissertation of Mario Rolando Vargas is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I would like to acknowledge my research advisor Dr. Yinsheng Wang for allowing me to join his lab back in 2008, and for helping me obtain my Ph.D. I would like to acknowledge Dr. Eric Chronister for his support and advice throughout the years, and Dr. Guy Bertrand who has been serving in my committee since I arrived at UCR. I would like to especially thank my friend and colleague, Jianshuang Wang, for his help in my research. I also want to extend my gratitude to all my former professors including Dr. Dennis Mitchell, Dr. Matthias Selke, Dr. Wayne Tikkanen, Dr. Carlos Gutierrez, and Dr. Francois Mathey for instilling, through their lectures, a love for chemistry in me.

I would like to say thank you to my friends for their help and support throughout the years. I want to express my appreciation to my girlfriend, Viraphon Pravongviengkham, for her love and support during difficult times. Lastly I would like to express my gratitude to my family for their love and support. To my hero, my father, Cesar Vargas, who is no longer with us but in heaven, to my mother Juana Vargas, to my hardworking sister Vilma Vargas, and to my studious sister Celia Vargas.

ABSTRACT OF THE DISSERTATION

Syntheses and Characterizations of Modified Nucleosides and Their Incorporation into Oligodeoxynucleotides

by

Mario Rolando Vargas

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, June 2012
Dr. Yinsheng Wang, Chairperson

Tobacco-specific nitrosamines, *N*'-nitrosonornicotine (NNN) and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), are found in tobacco and its smoke. Both compounds can generate the same intermediate upon metabolic activation by cytochrome P450, which can then form mutagenic DNA adducts. These adducts have been detected in the DNA of lung cancer patients. One of the most persistent and abundant adduct that has been detected in several tissues of rats that have been treated with NNN, is O^2-(pyridyloxobutyl)-thymine (O^2-POBdT). *N*3-POBdT is a new adduct that has never been detected in genomic DNA. Chapter two of this dissertation reports on the syntheses of O^2-POBdT and *N*3-POBdT and their incorporation into oligodeoxyribonucleotides (ODNs) for conducting biological studies.

5-Hydroxymethyl-2-deoxycytidine (hmdC) is a modified nucleotide that has recently been detected genomic DNA. Tet enzymes oxidize 5-methylcytosine (mC) to generate hmdC. Recently, it has been discovered that hmdC plays a crucial role in stem cell development. Like hmdC, 5-formyl-dC (fdC) has also recently been detected in genomic DNA. Its levels are lower

than hmdC and its exact role remains to be determined. Chapter 3 of this dissertation reports on the syntheses of hmdC and fdC and their incorporation into ODNs.

8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxo-dG) is an oxidized lesion found in genomic DNA. It is weakly mutagenic *in vitro* and *in vivo*, causing G→T transversions. 2'-Deoxyribonolactone (dL) is an oxidized abasic site found in damaged DNA. If left unrepaired it can produce single- and double-strand breaks, and dL is also mutagenic. Generation of the dL in short ODNs can be readily accomplished from insertion of a precursor, 7-nitroindole, d(7-Ni), followed by irradiation. Chapter 4 of this dissertation reports on the syntheses of 8-oxodG and d(7-Ni) and their incorporation into ODNs.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS…….iv

ABSTRACT OF THE DISSERTATION…….v

TABLE OF CONTENTS…….....…..vii

LIST OF ABBREVIATIONS…….……..….xii

LIST OF SCHEMES…….……....xiv

LIST OF FIGURES……..…..……xv

LIST OF TABLES…….……...xvii

CHAPTER 1 Introduction……..……..1

1.1 DNA damage and genomic instability…….……...…...…….1

1.2 Oxidative DNA damage and epigenetic regulation…….……...…1

1.3 DNA damage and tandem lesions……...…2

1.4 Tobacco-specific nitrosamine (TSNAs)-induced DNA modifications……..…….3

1.5 Scope of this thesis……4

References……6

CHAPTER 2 Chemical Syntheses of O^2-POBdT and $N3$-POBdT Phosphoramidites for Incorporation into Oligodeoxyribonucleotides…….....12

2.1 Introduction……...…12

2.1.2. P450 isoforms in humans and other species…….……..16

2.1.3. The abundance and persistence of O^2-POBdT in rodents…….……16

2.1.4. Repair and mutagenicity of O^2-POBdT…….……19

2.1.5. The detection of O^2-POBdT in humans…….…..……...20

2.1.6. $N3$-POBdT, an undetected POB adduct in DNA……..……..…20

2.1.7. The intention behind the synthesis of O^2-POBdT and $N3$-POBdT……21

2.2 Results…….……...…..21

2.3 Discussion……..…….....30

2.4 Conclusions……...…….……..32

2.5 Experimental Procedures……..…….....32

2.5.1. Synthesis of *N*3-[3-{2-(3-Pyridyl)-1,3-dithiane-2-yl}propyl]-3',5'-*O*-[isobutyryl]thymidine……..…….……33

2.5.2. Synthesis of *N*3-[3-{2-(3-Pyridyl)-1,3-dithiane-2-yl}-5'-*O*-[4,4'-dimethoxytrityl]thymidine……..……..…….....34

2.5.3. Synthesis of 4-(1,3-dithian-2-yl)-4-(3-pyridyl)butanol…….....……36

References…….….39

CHAPTER 3 Syntheses of 5-Hydroxymethyl-2'-deoxycytidine and 5-formyl-2'-deoxycytidine and their Incorporation into Oligodeoxyribonucleotides…….....45

3.1 Introduction……..……45

3.2 Detection and Sequencing of hmdC in genomic DNA……..……..…47

3.2.1. HmdC levels in different cells and tissues……..47

3.3 Detection and mutagenicity of fdC…….....48

3.4 The role of hmdC and Tet enzymes during ES cell differentiation…….49

3.5 The intention behind the synthesis of hmdC and fdC……...……...50

3.6 Results and Discussion……..…...50

3.7 Conclusions……..…….56

3.8 Experimental Procedures……..…….....57

References……..66

CHAPTER 4 Dual Incorporation of 8-oxo-7,8-dihydro-2'-deoxyguanosine and 7-Nitroindole into Oligodeoxyribonucleotides…….....70

4.1 Introduction…….…...……..70

4.2 Mutagenicity of 8-oxo-dG *in vitro*……..……72

4.2.1. Mutangenicity of 8-oxo-dG *in vivo*…….....72

4.3 7-Nitroindole as a precursor for 2'deoxyribonolactone……..…..…73

4.3.1. Mutation spectrum and mutation frequencies for 7-nitroindole and 2'-deoxyribonolacton…….……73

4.4 The intention behind the formation of 8-oxo-dG-dL tandem lesion in ODNs…….....74

4.5 Results and Discussion……..……..74

4.6 Experimental Procedures……..……...…83

References…….….91

APPENDIX A…….……..94

APPENDIX B……...…100

APPENDIX C……...…108

LIST OF ABBREVIATIONS

ACN	acetonitrile
AGT	O^6-alkylguanine DNA alkyltransferase
BnOH	benzyl alcohol
ES-cells	embryonic stem cells
ESI-MS	electrospray ionization mass spectroscopy
DEAD	diethyl azodicarboxylate
dheC	5-(1,2-dihydroxyethyl)-2'-deoxycytosine
dL	2'-deoxyribonolactone
d(7-Ni)	7-nitroindole-2'-deoxyribonucleoside
DMSO	dimethyl sulfoxide
DMTrCl	4,4-dimethoxytrityl chloride
fdC	5-formyl-2'-deoxycytosine
GC-MS	gas-chromatography mass spectrometry
HMBC	heteronuclear multiple-bond correlation
hmdC	5-hydroxymethyl-2'-deoxycytosine
HPB	4-hydroxy-1-(3-pyridyl)-1-butanone
$LiAlH_4$	lithium aluminum hydride
$NaIO_4$	sodium periodate
NBS	N-bromosuccinimide
NER	nucleotide excision repair
NNAL	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol
NNK	4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone
NNN	*N*'-nitrosonornicotine
ODN	oligodeoxyribonucleotide

8-oxo-dG	7,8-dihydro-8-oxo-2'-deoxyguanosine
POB	pyridyloxobutyl
$POCl_3$	phosphorus oxychloride
O^2-POBdC	O^2-[4-(3-pyridyl)-4-oxobut-1-yl]-2'-deoxycytidine
O^6-POBdG	O^6-[4-(3-pyridyl)-4-oxobut-1-yl]-2'-deoxyguanosine
O^2-POBT	O^2-[4-(3-pyridyl)-4-oxobut-1-yl]thymidine
*N*3-POBdT	3',5'-*O*-diisobutyryl-*N*3-[4-(3-pyridyl)-4-oxobutyl-1-yl]thymidine
RNS	reactive nitrogen species
ROS	reactive oxygen species
TBDMCl	*tert*-butyldimethylchlorosilane
TEA	triethylamine
Tg	thymidine glycol
THF	tetrahydrofuran
Ts	tosyl
TSNA	tobacco specific nitrosamine

LIST OF SCHEMES

Scheme 2-1 Bioactivation of NNK and NNN by cytochrome P450 enzymes……14

Scheme 2-2 Synthesis of 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanol……..….22

Scheme 2-3 Synthesis of N3-[3-{3-(3-Pyridyl)-1,3-dithiane-2-yl}propyl]-3',5'-*O*-[isobutyryl]thymidine (**2b**)…….……...….22

Scheme 2-4 Synthetic scheme for compound *N*3-[3-{2-(3-Pyridyl)-1,3-dithiane-2-yl}propyl]-3',5'-*O*-[4,4'-dimethoxytrityl]thymidine……..26

Scheme 2-5 Synthesis of dithiane-protected O^2-POBdT lesion……...28

Scheme 3-1 Synthetic scheme for hmdC phosphoramidite……..51

Scheme 3-2 Synthetic scheme for dheC phosphoramidite. Conditions: (a) DMTrCl, DMAP, pyridine, (b) TBDMCl, (c) $Bu_3SnCHCH_2$, $(PH_3P)_3PdCl_2$, DMF, 80 °C, 1.5 h; (d) (1)) 2,4,6-triisopropylbenzenesulfonyl chloride, DMAP, Et_3N, ACN, RT, 28 h; (2) conc. NH_4OH, 0°C-rt, 1 h; (e) OsO_4, 4-methylmorpholine-*N*-oxide, acetone-H_2O-t-BuOH (4:1:1), RT, 4h; (f) Ac_2O, DMAP, py, rt, 20 h; (g) TBAF, THF, 0 °C-RT, 6 h; (h) i-$Pr_2NP(Cl)O(CH_2)_2CN$, i-PrNHEt, CH_2Cl_2, 0 °C-RT, 0……53

Scheme 4-1 Synthetic scheme for 7,8-dihydro-5'-*O*-(4,4'-dimethoxytrityl)-N^2-isobutyryl-8-oxo-2'-deoxyguanosine phosphoramidite…….……...75

Scheme 4-2 Synthetic scheme for 7-nitroindole-phosphoramidite……...…76

LIST OF FIGURES

Figure 2-1 Structures of POB-DNA adducts and HPB……..…15

Figure 2-2 ^{1}H NMR (top) and ^{1}H-^{13}C HMBC spectrum (bottom) for *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-[isobutyryl]thymidine……...…22

Figure 2-3 ^{1}H-^{13}C HMBC spectrum of dithiane-protected O^{2}-POBdT……...…...29

Figure 3-1 Structures of hmdC, dheC, and fdC……..46

Figure 3-2 The negative-ion ESI-MS for the modified 12mer……...54

Figure 3-3 The product-ion spectrum of the ESI-produced $[M-3H]^{3-}$ ion of d(ATGGCGXGCTAT), where "X" represents dheC……...…55

Figure 4-1 Structures of lesions 8-oxo-dG and 2'-deoxyribonolactone (dL), and 7-nitroindole nucleoside……..……71

Figure 4-2 Negative-ion ESI-MS of d(ATGGCLGCGTA)……...…78

Figure 4-3 The product-ion spectrum of the ESI-produced $[M-3H]^{3-}$ ion (*m/z* 1187.7) of d(ATGGCLGCTAT). Shown in the inset is a scheme summarizing the observed fragment ions…….……78

Figure 4-4 LC-MS/MS identification of 5'-dinucleotide formation in calf thymus DNA treated with Cu(II)/H_2O_2/ascorbate. Extracted-ion chromatograms (EICs) for monitoring the *m/z* 540 → 346 transition for pdL-pdG in Fenton reagent-treated calf thymus DNA (a) and calf thymus DNA doped with authentic lesion-containing ODN (b).

Shown in (c) and (d) are the MS/MS averaged from the 21.4 min peak in (a) and the 21.3 min peak in (b)……..……80

Figure 4-5 LC-MS/MS identification of 5'-dL-(8-oxodG)-3' tandem lesion formation in calf thymus DNA treated with Cu(II)/H_2O_2/ascorbate. Extracted-ion chromatograms (EICs) for monitoring the *m/z* → 362 transitions for pdL-p(8-oxodG) in Fenton reagent-treated calf thymus DNA (under reaction conditions described in Table 4-1, a) and calf thymus DNA doped with authentic tandem lesion-containing ODN (b). Shown in (c) and (d) are the MS/MS averaged from the 20.7-min peak in (a) and the 20.8-min peak in (b)……...……...81

Figure 4-6 Quantification of the formation of the 5'-dL-3' and 5'-dL-(8-oxo-dG)-3' lesions in calf thymus DNA exposed with Cu(II) and ascorbate along with H_2O_2 or γ-rays. (a) induction of the 5'-dL-dG-3' and 5'-dL-(8-oxo-dG)-3' lesions in calf thymus DNA by Cu(II)/H_2O_2/ascorbate. The data represent the means ± S.D. of results from three independent treatments and LC-MS/MS quantification experiments……...……..82

LIST OF TABLES

Table 2-1 Adduct levels in NNK-treated rats...……...…….18

Table 4-1 Concentrations of Fenton-type reagents employed for the treatment of calf thymus DNA…….……85

Chapter 1 Introduction

1.1 DNA damage and genomic instability

Maintaining the integrity of the genetic code is vital for proper function and survival of a cell. Although some modifications to the DNA such as cytosine methylation or hydroxymethylation of 5-methylcytosine are vital modifications (1-4), modifications by endogenous or exogenous sources can result in detrimental DNA damage such as deamination and oxidation of nucleo bases, formation of DNA adducts, and single- and double-strand breaks. For example, high levels of reactive nitrogen species (RNS), e.g., nitrous anhydride (N_2O_3), can lead to deamination of nucleobases (5-8). Reactive oxygen species (ROS), such as $^{\bullet}OH$ radical, generated from normal aerobic metabolism, can lead to oxidation of nucleobases and generation of strand breaks. ROS can also be produced exogenously from exposure to irradiation such as X-rays or γ-rays. Molecules that form adducts with DNA can be derived from environmental pollutants such as polycyclic aromatic hydrocarbons (PAHs) or heterocyclic aromatic compounds (HCAs), derived from cooked foods or tobacco smoke (48, 49). If left unrepaired these lesions can cause point mutations or frameshift mutations; some of these lesions can stop the entire DNA replication process, which is lethal to a cell (9). If mutations caused by lesions occur at genes that regulate growth or protect cells against cancer, it can lead to tumor formation.

1.2 Oxidative DNA damage and epigenetic regulation

ROS include superoxide, $O_2^{\bullet-}$, endogenously derived from leakage of the electron transport system chain. $O_2^{\bullet-}$ can then be converted by superoxide dismutase (SOD) to H_2O_2,

which can then react with Cu^{+} or Fe^{++} to yield •OH radical. •OH radical can then react with numerous macromolecules including DNA. For example, •OH radical can abstract hydrogen atoms from numerous positions of the nucleosides to yield different types of modified nucleosides, e.g,, 5-formyl-dC (fdC), 5-formyl-dU (fdU), thymidine glycol (Tg), 8-oxo-dG and tandem lesions. •OH can also lead to the formation 8,5'-cyclopurine-2'-deoxynucleosides such as (5'S)-cdA (9, 10). Some of these lesions can then interfere with cellular processes that use DNA as a template—such as replication and transcription. For example, it has been demonstrated *in vitro* that 8-oxo-dG containing template strand directs the misincorporation of dAMP and dCMP opposite the oxidized lesion (50, 51). Lesions present in genes that are actively transcribed can stall or block RNA elongation, or induce base misincorporation in the growing mRNA strand. This can then lead to compromised gene expression, affecting the spatial and temporal flow of genetic information in a cell, which can lead to growth defects during development (11). Specifically, a lesion in the template strand can block RNA polymerases, leading to truncated transcripts, slow RNA polymerase translocation, and promote the incorporation of mismatched bases into mRNA potentially altering the transcript (12-15).

1.3 DNA damage and tandem lesions

Besides single nucleoside damage, damage to two adjacent nucleotides is another form of DNA damage. Studies conducted with oligodeoxynucleotides (ODNs) or DNA usually involves the generation of •OH exogenously via exposure to ionizing radiation. •OH then reacts with DNA directly or forms other ROS such as peroxyl radical in the presence of molecular oxygen. Using γ-rays to irradiate calf thymus DNA in the presence of oxygen, Cadet et al. (16) discovered that peroxyl radicals form on pyrimidine nucleosides and form tandem lesions with vicinal guanine,

N-(2-deoxy-β-D-erythro-pentofuranosyl)formylamine-8-oxo-7,8-dihydro-2'-deoxyguanosine (dβF-8-oxo-dG and 8-oxo-dG-dβF). However, these types of lesions have never been detected in cellular DNA. γ-irradiation of synthetic duplexes also led to the formation of intrastrand cross-link between the 6-hydroxy-5,6-dihydrocytosin-5-yl radical and a vicinal guanine leading, after dehydration to the generation of the G[8-5]C adduct (17). As a general trend the addition reaction of the pyrimidine radicals is favored in the absence of oxygen and when the purine base is located on its 5'-side. Other tandem lesions formed include: G[8-5]T, T[5m-8]G, A[8-5m]T, and T[5m-8]A (18). Both G[8-5]C and G[8-5m]T have been proposed as degradation products in aerated solutions using Cu(I) or Fe(II) with H_2O_2 and ascorbate (19, 20). Both tandem lesions have been detected in HeLa cells after being exposed to γ-rays. Like single lesions, tandem lesions can also alter DNA replication. G[8-5]C blocked DNA replication and was found to be mutagenic in vivo, giving rise to 8.7% G→T and 1.2% G→C transversion mutations (52). G[8-5]T lesion also stopped DNA synthesis by Pol I after incorporating the correct nucleotide dAMP opposite the lesion (53).

1.4 Tobacco-specific Nitrosamines (TSNAs)-induced DNA modification

Tobacco use in the United States is the largest preventable cause of death and disease in the United States (21). Tobacco use causes heart disease, adverse reproductive effects, and the exacerbation of chronic health conditions. Every year, approximately 443,000 persons in the United States die from smoking-related illnesses. In 2010, 19.3% (45.3 million) of U.S. adults were current cigarette smokers. Among the several carcinogenic compounds found in tobacco smoke are the nicotine and polycyclic aromatic hydrocarbons (PAHs). Although there are several metabolic pathways by which PAHs are activated, one known pathway yields radical cations

which can form depurinating adducts such as B[α]P-6-N7Gua or B[α]P-6-N7Ade (22). Replicative DNA polymerases will then insert a dAMP opposite the abasic site thus giving rise to G→T and A→T when the daughter strand is replicated (23). Such mutations have been detected in K-ras and p53 genes of lung cancer patients [24-29]. Another class of compounds that are found in tobacco and tobacco smoke are tobacco-specific nitrosamines (TSNAs) such as NNK and NNN (30, 31). As with PAHs, these TSNAs must also be activated by P450 to form reactive intermediates that react with DNA bases. Metabolically activated NNN and NNK can also form depurinating adducts or form stable covalently bound adducts that can be mutagenic (32, 33).

1.5 Scope of this thesis

Two tobacco-specific components, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and *N*'-nitrosonornicotine (NNN) can be directly extracted and measured (34, 35, 36). The levels of NNN and NNK are not as high as nicotine levels and their amounts vary depending on the tobacco source and how it is processed (30, 31). Nevertheless, NNK and NNN are deleterious only upon metabolic activation by cytochrome P450, a multifunctional enzyme, that yields a common intermediate from NNK and NNN and forms adducts with DNA nucleotides (37). Some of these pyridyloxobutylated adducts (POB adducts) are hard to repair and persist in the cell's genome. POB adducts have been linked to cancer in rodents and are also present in human-pulmonary tissue of cancer patients. Some of the POB-DNA adducts that have been detected and studied are: 7-POBG, O^2-POBdC, O^2-POBdT and O^6-POBdG. One of the most persistent and abundant is O^2-POBdT. Recently, the synthesis of *N*3-POBdT has been reported (38). Although *N*3-POBdT has never been detected, its synthesis can allow for incorporation into oligodeoxyribonucleotides (ODNs) for *in vitro* replication and mutagenic studies. Chapter 2 is to

report on the synthesis of O^2-POBT and *N*3-POBdT and their incorporation into ONDs for *in vitro* analyses.

Recently two oxidized nucleotides have been detected in genomic DNA. 5-formyl-2'-deoxycytidine (fdC) and 5-hydroxymethyl-2'-deoxycytidine (hmdC). HmdC has only recently been detected in genomic DNA (2, 3). In 2009, it was detected from digested genomic DNA and it was discovered that it is derived from 5-methylcytosine (mC). Since then hmdC has been detected in several tissues and its function is currently being deciphered at an enthusiastically fast pace because evidence shows that it plays a crucial role in the development and differentiation of stem cells (39-42). It has been quantified at 0.6% to 0.23% in nerve cells (1). fdC has only been recently detected in mouse embryonic stem cells (ES cells) at about 0.02% (43). Site-specific mutagenesis studies have been conducted. Although it does cause some mutations, its mutation frequency is quite low, about 0.03-0.28% (44). Chapter 3 reports on the syntheses of fdC and hmdC and the incorporation a precursor of fdC into an ODN.

Chapter 4 of this dissertation reports on the synthesis of two oxidized lesions, 8-oxo-7,8-dihydro-2'-deoxyguanosine (8-oxo-dG) and 7-nitroindole (7-Ni), and their contiguous incorporation into a docecameric ODN. 8-oxo-dG results from oxidation of dG by reactive oxygen species. If the lesion remains unrepaired, it has been shown to cause G→T transversions. 7-nitroindole nucleoside, d(7-Ni), is a synthetic non-natural nucleoside that has never been detected in genomic DNA. It serves as a precursor for the formation of 2'-deoxyribonolactone (dL) lesion (46). dL is a C1'-oxidized abasic site. *In vivo* studies shows that it mainly causes G-to-A and G→T transversions in *E. coli* (47). The tandem lesion 8-oxo-dG/dL was successfully formed and it was found to be induced in calf thymus DNA upon exposure to Fenton reagent.

References

1. Heintz, N. and Kriaucionis, S. (2009) The Nuclear DNA Base 5-hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science*. 324: 929-930.

2. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science. 324: 930-935.

3. Carell, T., Globisch, D. and Münzel, M. (2011) 5-Hydroxymethylcytosine, the Sixth Base of the Gemone. *Angew. Chem. Int. Ed*. 50(29): 6460-6468.

4. Leonhardt, H. and Spada, F. (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res*. 38(19): e181.

5. Lindahl, T. and Barnes, D.E. (2000) Repair of endogenous DNA damage. *Cold Spring Harb. Symp. Quant. Biol*. 65: 127-133.

6. Lindahl, T. and Nyberg, B. (1972) Rate of depurination of native deoxyribonucleic acid. *Biochemistry*. 11: 3610-3618.

7. Conney, A.H., Miller, E.C. and Miller, J.A. (1956) The metabolism of methylated aminoazo dyes. V. Evidence for induction of enzyme synthesis in the rat by 3-methylcholanthrene. *Cancer Res*. 16: 450-459.

8. Luch, A. (2005) Nature and nurture–lessons from chemical carcinogenesis. *Nat. Rev. Cancer*. 5: 113-125.

9. Von Sonntag C. *The Chemical Basis of Radiation Biology*, Taylor and Francis, New York, 1987.

10. Dizdaroglu, M., Jaruga, P. and Rodriguez, H. (2001) Identification and quantification of 8,5'-cyclo-2'-deoxyadenosine in DNA by liquid chromatography mass spectrometry. *Free Radic. Biol. Med*. 30: 774-784.

11. Smith, J.M. and Koopman, P.A. (2004) The ins and outs of transcriptional control: nucleocytoplasmic shuttling in development and disease. *Trends Genet*. 20: 4-8.

12. Scicchitano, D.A., Olesnicky, E.C., and Dimitri, A. (2004) Transcription and DNA adducts: what happens when the message gets cut off? *DNA Repair*. 3: 1537-1548.

13. Tornaletti, S. (2005) Transcription arrest at DNA damage sites. *Mutat. Res*. 577: 131-145.

14. Tornaletti, S. and Hanawalt, P.C. (1999) Effect of DNA lesions on transcription elongation. *Biochimie*. 181: 139-146.

15. Scicchitano, D.A. (2005) Transcription past DNA adducts derived from polycyclic aromatic hydrocarbons. *Mutat. Res*. 577: 146-154.

16. Bourdat, A., Douki, T., Frelon, S., Gasparutto, D. and Cadet, J. (2000) Tandem Base Lesions Are Generated by Hydroxylation Radical within Isolated DNA in Aerated Aqueous Solution. *J. Am. Chem. Soc*. 122: 4549-4556.

17. Gu, C. and Wang. Y. (2004) LC-MS/MS identification and yeast polymerase η bypass of a novel γ-irradiation-induced intrastand cross-link lesion G[8-5]C. *Biochemistry*. 43: 6745-6750.

18. Bellon, S., Ravanat, J.-L, Gasparutto, D. and Cadet, J. (2002) Cross-linked thymidine-purine base tandem lesions: synthesis, characterization, and measurement in γ-irradiated DNA. *Chem. Res. Toxicol*. 15: 598-606.

19. Zhang, Q. and Wang, Y. (2003) Independent generation of 5-(2'-deoxycytinyl)methyl radical and the formation of a novel crosslink lesion between 5-methylcytosine and guanine. *J. Am. Chem. Soc*. 125: 12795-12802.

20. Cao, H. and Wang, Y. (2007) Quantification of oxidative single-base and intrastrand cross-link lesions in unmethylated and CpG-methylated DNA induced by Fenton-type reagents. *Nucleic Acids Res*. 35: 4833-4844.

21. http://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/index.htm

22. Devansan, P.D., Rama-Krishma, N.V.S., Todorovic, R., Rogan, E.G., Cavalieri, E.L., Jeong, H., Jankowiak, R. and Small, G.J. (1992) Identification and quantification of benzo[α]pyrene-DNA adducts fromed by rat liver microsomes *in vitro*. *Chem. Res. Toxicol*., 5, 302-309.

23. Sagher, D. and Strauss, B. (1983) Insertion of nucleotides opposite apurinic/apyrimidinic sites in deoxyribonucleic acid during in vitro synthesis. Uniqueness of adenine nucleotides. Biochemistry. 22: 4512-4526.

24. Yamamoto, F. and Perucho, M. (1984) Activation of human c-K-*ras* oncogene. *Nucleic Acids Res*. 12: 8873-8885.

25. Rodenhuis, S. and Siebos, R.J. (1990) The *ras* oncogenes in human lung cancer. *Am. Rev. Repair. Dis*. 142: S27-S30.

26. Capella, G., Cronauer-Mitta, S., Pienado, M.A. and Perucho, M. (1991) Frequency and spectrum of mutations at codons 12 and 13 of the c-K-*ras* gene in human tumors. *Environ. Health Perspect*. 93: 125-131.

27. Top, B., Mooj. W.J., Klaver, S.G., Berrigter, L., Wisman, P., Elbers, H.R., Viser, S. and Rodenhuis, S. (1995) Comparative analysis of *p53* gene mutations and protein accumulation in human non-small-cell lung cancer. *Int. J. Cancer*. 64: 83-91.

28. Gealy, R., Zhang, L., Siegfried, J.M., Luketich, J.D. and Keohavong, P. (1999) Comparison of mutations in the *p53* gene K-*ras* genes in lung carcinomas from smoking and non-smoking women. *Cancer Epidemiol. Biomarkers Prev*. 8: 297-302.

29. Vahakanga, K.H., Bennett, W.P., Castren, K., Welsh, J.A., Khan, M.A., Bjomek, B., Alavanja, M.C. and Harris, C.C. (2001) p53 and K-*ras* mutations in lung cancers from former and never smoking women. *Cancer Res*. 61: 4350-4356.

30. Idris, A. M., Nair, J., Ohshima, H., Fiesen, M. Brouet, I., Faustman, E. M. and Bartsch, H. (1991) Unusually high levels of carcinogenic tobacco-specific nitrosamines in Sudan snuff (toombak). *Carcinogenesis.* 12(6): 1115-1118.

31. Stepanov, I., Jensen, J., Hatsukami, D. and Hecht, S. S. (2008) New and traditional smokeless tobacco: Comparision of toxicant and carcinogen levels. *Nicotine Tob. Res.* 10(12): 1173-1782.

32. Hecht, S. S., Villalta, P. W., Sturla, S. J., Cheng, G., Yu, N., Upadhyaya, P. and Wang, M. (2004) Identification of O^2-substituted pyridimidine adducts formed in reactions of 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-1-butanone and 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanol with DNA. *Chem. Res. Toxicol.* 17: 588-597.

33. Foiles, P. G., Peterson, L. A., Miglietta, L. M. and Ronai, Z. (1992) Analysis of mutagenic activity and ability to induce replication of polyoma DNA sequences by different model metabolites of the carcinogenic tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Mut. Res.* 279: 91-101.

34. Hecht, S. S., Miglietta, L. M., Foiles, P. G. and Staretz, M. E. (1997) Evidence for an Important Role of DNA Pyridyloxobutylation in Rat Lung Carcinogenesis by 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone: Effects of Dose and Phenethyl Isothiocyanate. *Cancer Res.* 57: 259-266.

35. Peterson, L. A., Kenney, P. M. and Thomson, N. M. (2003) The pyridyloxobutyl DNA adduct, O6-[4-oxo-4-(3-pyridyl)butyl]guanine, is detected in tissues from 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone-treated A/J mice. *Chem. Res. Toxicol.* 16: 1-6.

36. Hecht, S. S., Upadhyaya, P., Villalta, P. W., McIntee, E. J., Shi, Y., Sturla, S. J., Cheng, G. and Wang, M. (2003) Identification of adducts formed by pyridyloxobutylation of deoxyguanosine and DNA by 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanone, a

chemically activated form of tobacco-specific carcinogens. *Chem. Res. Toxicol.* 16: 616-626.

37. Hecht, S. S., Villalta, P. W., Sturla, S. J. and Wang, M. (2004) Identification of O^2-Substituted Pyrimidine Adducts Formed in Reactions of 4-(Acetoxymethyllnitrosamino)-1-(3-pyridyl)-1-butanone and 4-(Acetoxymethylnitrosamio)-1-(3-pyridyl)-1-butanol with DNA. *Chem. Res. Toxicol.* 17: 588-597.

38. Krishnegowda, G., Sharma, A. K., Krzeminski, J. et al. (2011) Facile Syntheses of O^2-[4-(3-Pyridyl-4-oxobutyl-1-yl]thymidine, the Major Adduct Formed by Tobacco Specific Nitrosamine 4-methylnitrosamino-1-(3-pyridyl)-1-butanone (NNK) *in vivo*, and Its Site-Specifically Adducted Oligodeoxynucleotides. *Chem. Res. Toxicol.* 24: 960-967.

39. Leonhardt, H. and Spada, F. (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res*. 38(19): e181.

40. Sowers, L. C., Zhang, Y., Hong, K., Taranova, O. V., D'Alessio, A. C. and Ito, S. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 466: 1129-1135.

41. Koh K. P. and Rao, A. (2011) Tet1 Tet2 Regulates 5-hydroxymethylcytosine Production and Cell Lineage Specification in Mouse Embryonic Stem Cells. *Cell Stem Cell*. 8: 200-231.

42. Spada, F., Leondardt, H., Bultmann, S., Schmidt, C. S., Brachmann, A. and Szwagierczak, A. (2011) Characterization of PvuRts1I endonuclease as a tool to investigate genomic 5-hydroxymethylcytosine. *Nucleic Acids Res*. 39(12): 5149-5156.

43. Carell, T., Hagemeier, C. and Deimi, C. A. (2011) The Discovery of 5-Forymlcytosine in Embryonic Stem Cell DNA. *Angew. Chem*. 50(31): 7008-7012.

44. Kamiya, H., Tsuchiya, H., Karino, N., Ueno, Y., Matsuda, A. and Harashima, H. (2002) Mutagenicity of 5-Formylcytosine, an Oxidation Product of Methylcytosine, in DNA in Mammalian Cells. *J. Biochem*. 132: 551-555.

45. Pursell, Z. F., McDonald, J. T., Mathews, C. K. and Kunkel, T. A. (2008) Trace amounts of 8-oxo-dGTP in mitochondrial dNTP pools reduce DNA polymerase γ replication fidelity. *Nucleic Acids Res*. 36(7): 2174-2181.

46. Kotera, M., Roupioz, Y., Defrancq, E., Bourdat, A., Garcia, C. C. and Lhomme, J. (2000) The 7-nitroindole Nucleoside as a Photochemical Precursor of 2'-Deoxyribonolactone: Access to DNA Fragments Containin This Oxidative Abasic Lesion. *Chem. Eur. J*. 6(22): 4163-4169.

47. Saparbaev, M., Dumy, P., Constant, J. and Faure, V. (2004) 2'-Deoxyribonolactone lesion produces G →A transitions in *Escherichia coli*. *Nucleic Acids Res*. 32(9): 2937-294.

48. Straff, K., Baan, R., Grosse, Y., Secretan, B., El Ghissassi, F. and Cogliano, V. (2005) Carcinogenicity of polyclic aromatic hydrocarbons. Policy Watch. *Lancet Oncol*. 6: 931-932.

49. Mababe, S., Tohyama, K., Wada, O. and Aramaki, T. (1991) Detection of carcinogen, 2-amino-1-methyl-6-phenylimidazo[4,5-*b*]pyridine, in cigarette smoke condensate. Carcnogenesis. 12: 1945-1947.

50. Efrati, E., Tocco, G., Eritja, R., Wilson, S.H. and Goodman, M.F. (1999) "Action-at-a-Distance" Mutagenesis. *J Biol. Chem*. 274(22): 15920-15927.

51. Johnson, K.A, Thai, D.M. and Hanes, J.W. (2006) Incorporation and Replication of 8-oxo-deoxyguanosine by the Human Mitochondrial DNA Polymerase. *J. Biol. Chem*. 281(47): 36241-36248.

52. Hong, H., Cao, H. and Yinsheng. W. (2007) Formation and genotoxicity of a guanine-cytosine instrastrand cross-link *in vivo. Nucleic Acids Res.* 35(12): 7118-7127.

53. Jiang, Y., Hong, H., Cao, H. and Yinsheng, Wang (2007) In Vivo Formation and in Vitro Replication of a Guanine-Thymine Instrastrand Cross-Link Lesion. *Biochemistry*. 46(44): 12757-12763.

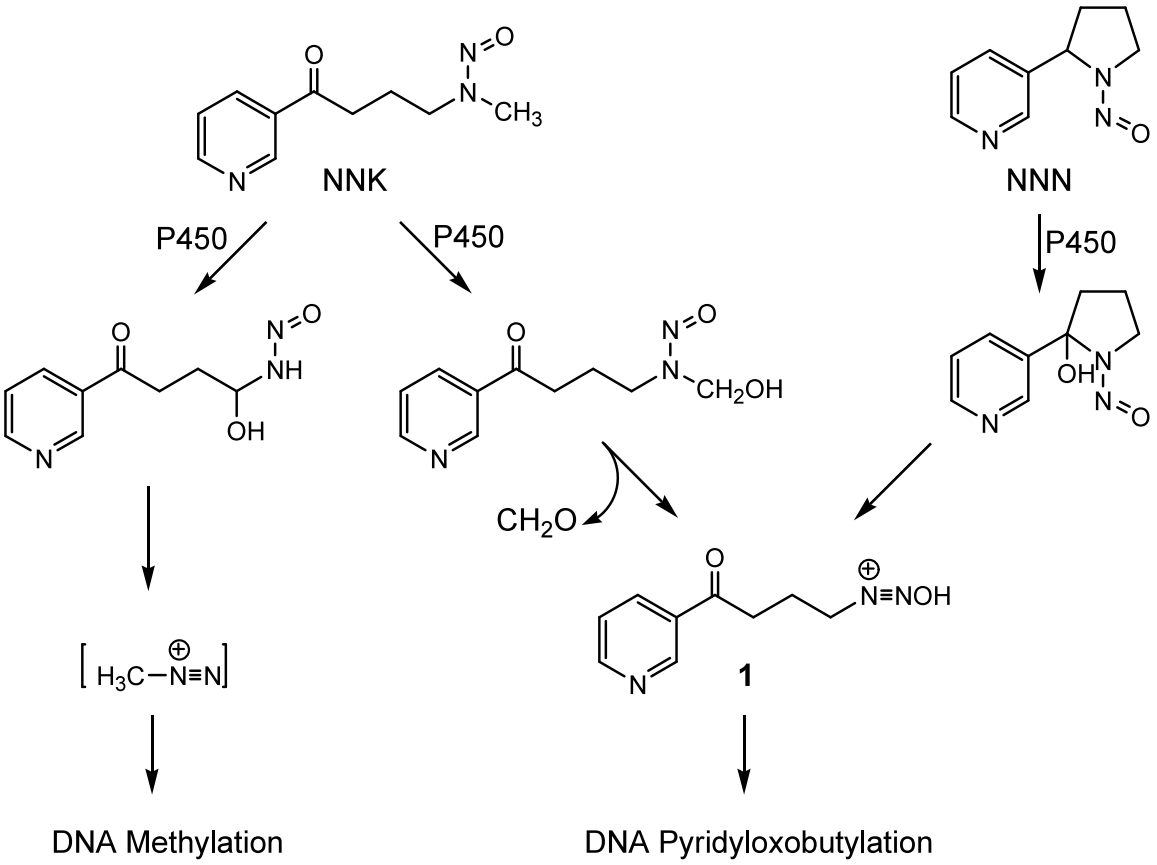
Chapter 2 Chemical Syntheses of O^2-POBdT and $N3$-POBdT Phosphoramidites for Incorporation into Oligodeoxynucleotides

2.1 Introduction

Tobacco-specific nitrosamines (TSNAs) such as 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and *N*'-nitrosonornicotine (NNN) are found in tobacco products. These can simply be extracted directly from tobacco and measured using gas chromatography-thermal energy analysis (GC/TEA) (6, 7). Their levels vary depending on the type of tobacco used and how it is processed. Generally, NNN is found at higher levels than NNK (6, 7, 33). All TSNAs are bioactivated by cytochrome P450 via hydroxylation; upon hydroxylation at the methyl carbon of NNK, a reactive intermediate is generated, **1**, Scheme 2-1, that forms miscoding pyridyloxobutyl adducts (POB adducts) with DNA bases (31). Studies have found a link between POB adducts and tumorigenesis in rodents that were treated for several weeks with NNK, underlining the importance of studying such adducts (2). Similarly, upon 2'-hydroxylation of NNN, the same reactive species that forms from NNK is generated and also forms POB adducts (Scheme 2-1). When rodents were treated for several weeks with both enantiomers of NNN, if was discovered that *S*-NNN gave more total POB adducts in the liver and esophagus and *R*-NNN gave more total adducts in the lungs (24). In general *S*-NNN is the most potent enantiomer.

The POB-DNA adducts that have been detected and studied thoroughly are: *N7*-POBdG, O^2-POBdC, O^2-POBdT and O^6-POBdG. Both *N*7-POBdG and O^2-POBdC readily lose their nucleobases to yield *N*7-POBG and O^2-POBC, respectively (5). Most of these release HPB upon treatment with mild acid; the release of HPB has been estimated at about 50%; a different estimate of 65% has also been published (34). Usually HPB is monitored when investigating the

levels of POB adducts in DNA (Figure 2-1). Interestingly one of the most abundant and persistent adduct is O^2-POBdT. We report here the syntheses of O^2-POBT and *N*3-POBdT nucleosides for incorporation into oligodeoxynucleotides (ODNs).



Scheme 2-1 Bioactivation of NNK and NNN by cytochrome P450 enzyme

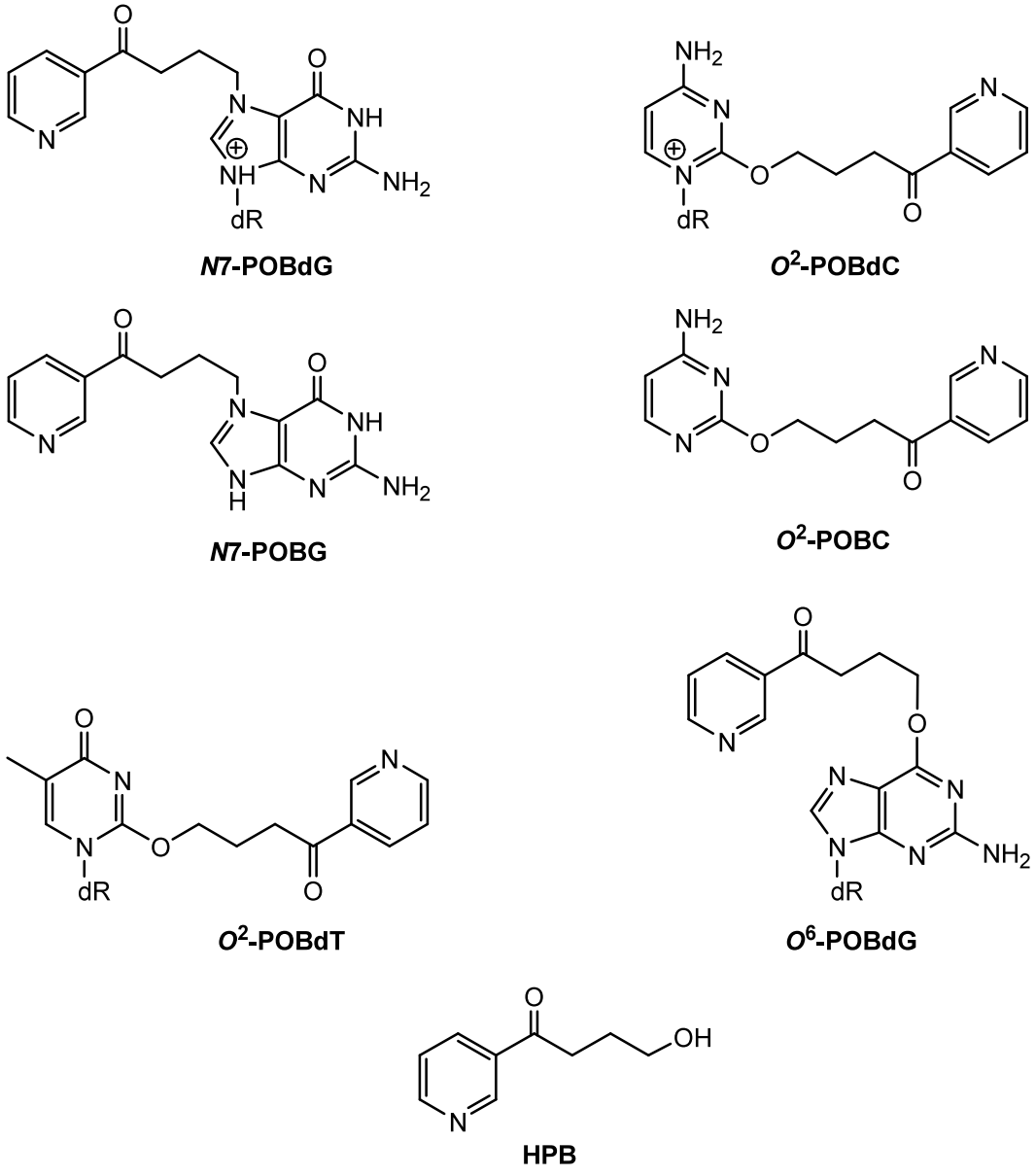


Figure 2-1 Structures of POB-DNA adducts and HPB

2.1.2. P450 isoforms in humans and other species

There are several isoforms of P450, the enzyme that activates TSNAs, each present in different concentrations in different tissues; each isoform of P450 has a different activity towards TSNAs such as NNN or NNK. Some are even inducible in certain tissues (17, 20, 19). P450s used in studies may be obtained from microsomes prepared from tissues or they may be generated as recombinant proteins. Interspecies differences in the activities of P450 also exist (8).

In a recent study, 116 human lung microsomes were analyzed for the amounts of two isoforms of P450: 2A13 and 2A6. It was discovered that 90% had detectable 2A6, whereas only 12% of had detectable 2A13. However, it was discovered that only 2A13 contributes to the metabolism of NNK and levels of CYP2A proteins vary among individuals (10). They concluded that individuals with relatively high level of 2A13 expression will likely have an increased risk of developing smoking-related lung cancer (10). This type of study highlights the interindividual difference in levels of P450 isoforms and the differences in activities towards NNK. Other studies using the same isoforms of P450 also show differences in levels among individuals (11).

2.1.3. The abundance and persistence of O^2-POBdT in rodents

When rodents were treated with different doses of NNK and their hepatic and pulmonary DNA analyzed, it was discovered that O^2-POBdT adduct was present in highest levels among all known POB-DNA adducts (22) (Table 1-1). When a similar experiment was done with NNAL, a metabolite of NNK, O^2-POBdT was also present at highest level in pulmonary and hepatic DNA

of rodents (35). O^2-POBdT was also detected as the most abundant and persistent adduct in extrahepatic tissues when rodents were treated with NNK and both enantiomers of NNAL (32).

Table 2-1 Adduct levels in NNK-treated rats (22)

Mean ± S.D., N = 5 (fmol/mg DNA)					
Tissue	Dose of NNK (mmol/kg)	*N*7-POBG	O^2-POBdT	O^2-POBC	O^6-POBdG
Lung	0.025	933 ± 89	1120 ± 66	483 ± 36	251 ± 26
	0.1	1800 ± 478	2020 ± 483	840 ± 169	487 ± 101
	Saline control	84 ± 7	ND	ND	ND
Liver	0.025	3550 ± 1600	3530 ± 725	2930 ± 521	28 ± 17
	0.1	12200 ± 1600	12300 ± 1690	7800 ± 1680	140 ± 25
	Saline control	ND	ND	ND	ND

When F344 rats were given NNN in their drinking water for several weeks, O^2-POBdT was detected at highest levels in liver and lung DNA. *S*-NNN produced higher levels of POB-DNA adducts in the esophagus than *R*-NNN, supporting previous data that *S*-NNN is more tumorigenic than *R*-NNN (24). Other tissues from rodents have also indicated O^2-POBdT as the most abundant POB-DNA adduct present when treated with NNN, with *S*-NNN being the most potent (25).

2.1.4. Repair and mutagenicity of O^2-POBdT

Some POB-DNA adducts, such as O^6-POBdG, are repaired by AGT with the rate of repair proportional to the size of the cavity of AGT (27). For example, NNKOAc treated bacterial cells expressing hAGT have been shown to contain 80% less mutations relative to cells with no AGT expression (36). Cells that have been treated with NNKOAc have shown an increased mutagenicity and toxicity when a scaffold of BER is lost (27). Interestingly, BER-generated abasic site intermediate causes mutations if left unrepaired with the majority of the mutations being AT-to-GC (37). When DNA from lung tumors where isolated from A/J mice that had been treated with NNKOAc, it was discovered that all the tumors analyzed contained a mutation on codon 12 of Ki *ras* gene with mainly G-to-A and G-to-T mutations (38).

Both NNK and NNN have also been shown to be mutagenic in target tissues in *lac*Z and *lac*I transgenic mice (39, 40-42). NNK induced an increased rate of GC→AT mutations but since NKK both methylates and pyridyloxobutylates DNA, it is difficult to associate specific mutations with specific adducts. Based on the all the data up to now, the two mutagenic pyridyloxobutyl candidates are O^6-POBdG and O^2-POBdT. It has been suggested that if mutations at AT base pairs are required to produce oncogenic proteins, the formation of O^2-

POBdT and its replicative bypass is important for tumor formation. If mutations at the GC base are to cause cancer, O^6-POBdG is most likely responsible.

2.1.5. The detection of O^2-POBdT in humans

In 2007 Holze et al. (26) extracted DNA from human lung tissue. They treated the DNA with mild acid and monitored the release of HPB from POB adducts in DNA using GC-MS. They discovered that the concentration of HPB-releasing lung DNA adducts was significantly higher in 21 self-reported smokers compared to 11 self-reported nonsmokers (404 ± 258 fmol vs. 59 ± 56 fmol HPB/mg DNA). All of the 32 patients had a form of lung cancer except for two (26). Thus, one can say that smoking leads to increased levels of POB adducts and tumorigenesis that may be linked to cancer.

2.1.6. *N*3-POBdT, an undetected POB adduct in DNA

*N*3-POBdT has never been detected in any hepatic or pulmonary tissues of rodents treated with a TSNA such as NNN and NNK (22, 23, 24). It has not been detected in prokaryotic and mammalian cells that have been treated with either NNK or NNN or even NNKOAc. It has not been detected in cell lines with deficient repair mechanism (25). Since it has not been detected, no mutagenic studies have been conducted.

2.1.7. The intention behind the synthesis of O^2-POBdT and *N*3-POBdT

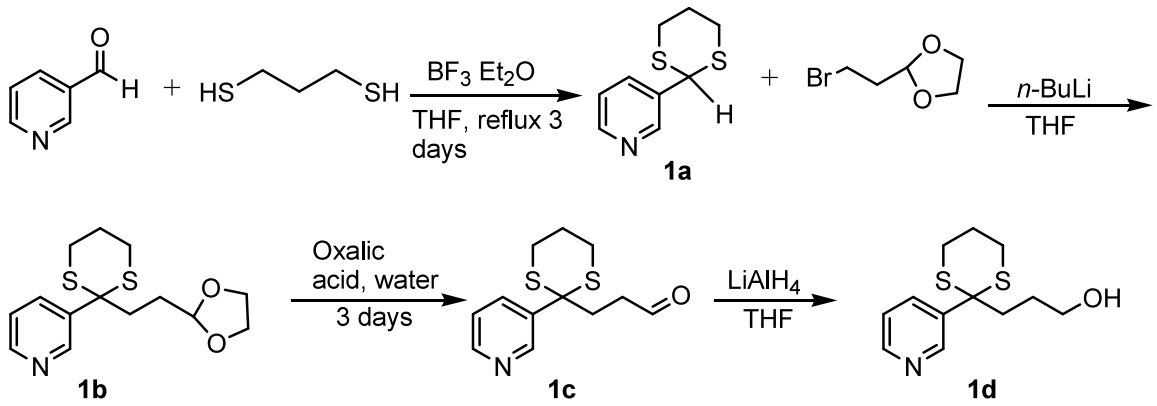
I synthesized 3',5'-*O*-[isobutyryl]-*N*3-POBdT and confirmed its structure with 2D ^{1}H-^{13}C NMR (Scheme 2-2) To our knowledge there was no published papers on the synthesis of such a compound at the time, probably due to its lack of detection in DNA treated with NNK or NNN. Regretfully, no additional syntheses were attempted in our part. A recent paper reported the synthesis of *N*3-POBdT using a similar approach as ours but with a much higher yield (43). Thus we followed the published procedures for the synthesis of *N*3-POBdT (43) (Scheme 2-3).

Although O^2-POBdT has been synthesized, the reported yields are less than 1%. Using a recent published procedure (43) we intended to synthesize O^2-POBdT phosphoramidite for incorporation into ODNs for biological studies.

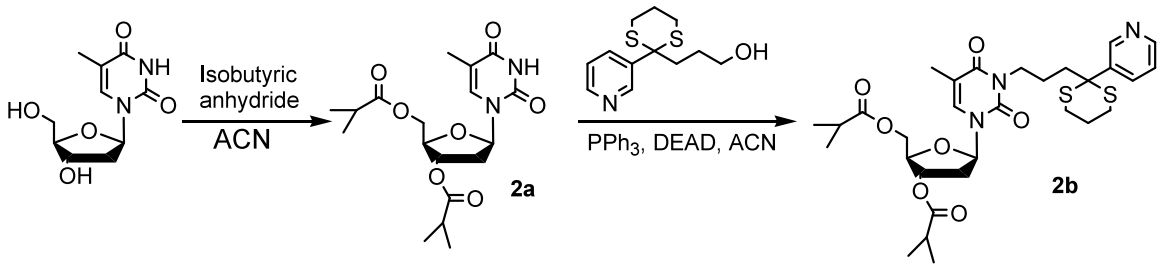
2.2 Results

In order to synthesize *N*3-POBdT, starting material 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanol, **1d**, Scheme 2-2, was synthesized according to the method of Hecht (44). The final structure was verified by ^{1}H NMR and MS, see experimental procedures.

Our initial approach used for the synthesis of *N*3-POBdT was initiated by first generating the starting material 3',5'-*O*-[isobutyryl]thymidine by reacting dT with isobutyric anhydride at room temperature overnight. ^{1}H NMR spectrum of the compound showed the disappearance of both 3' and 5'-hydroxy protons that were present in dT and the appearance of new peaks corresponding to the methyl groups of isobutyl group used to protect 3', 5'-hydroxy groups, **2b**, Scheme 2-3.



Scheme 2-2 Synthesis of 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanol



Scheme 2-3 Synthesis of *N***3**-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-[isobutyryl]thymidine (**2b**)

3',5'-*O*-[isobutyryl]thymidine, **2a**, was then reacted with alcohol, **1d**, via Mitsunobu reaction conditions to yield the isobutyryl-protected *N*3-POBdT, **2b**, Scheme 2-3. Unlike most Mitsunobu reactions that are carried out at room temperature or slightly above ambient temperature, the conditions required refluxing in ACN overnight. The ^{1}H NMR spectrum showed new peaks corresponding to isobutyl-protected *N*3-POBdT including a triplet at 3.70 ppm corresponding to H1 of the butyl part of the molecule. 2D NMR was needed to confirm the exact structure. ^{1}H-^{13}C HMBC showed strong coupling between H1 (butyl) and C2 and C4 carbons of the pyrimidine ring of the adduct. Neither O^2 nor O^4-POBdT structures can have coupling between H1 (butyl) and both C2 and C4 carbons of the pyridine ring (Figure 2-2). Thus, the above spectral features support the successful preparation of compound **2b.**

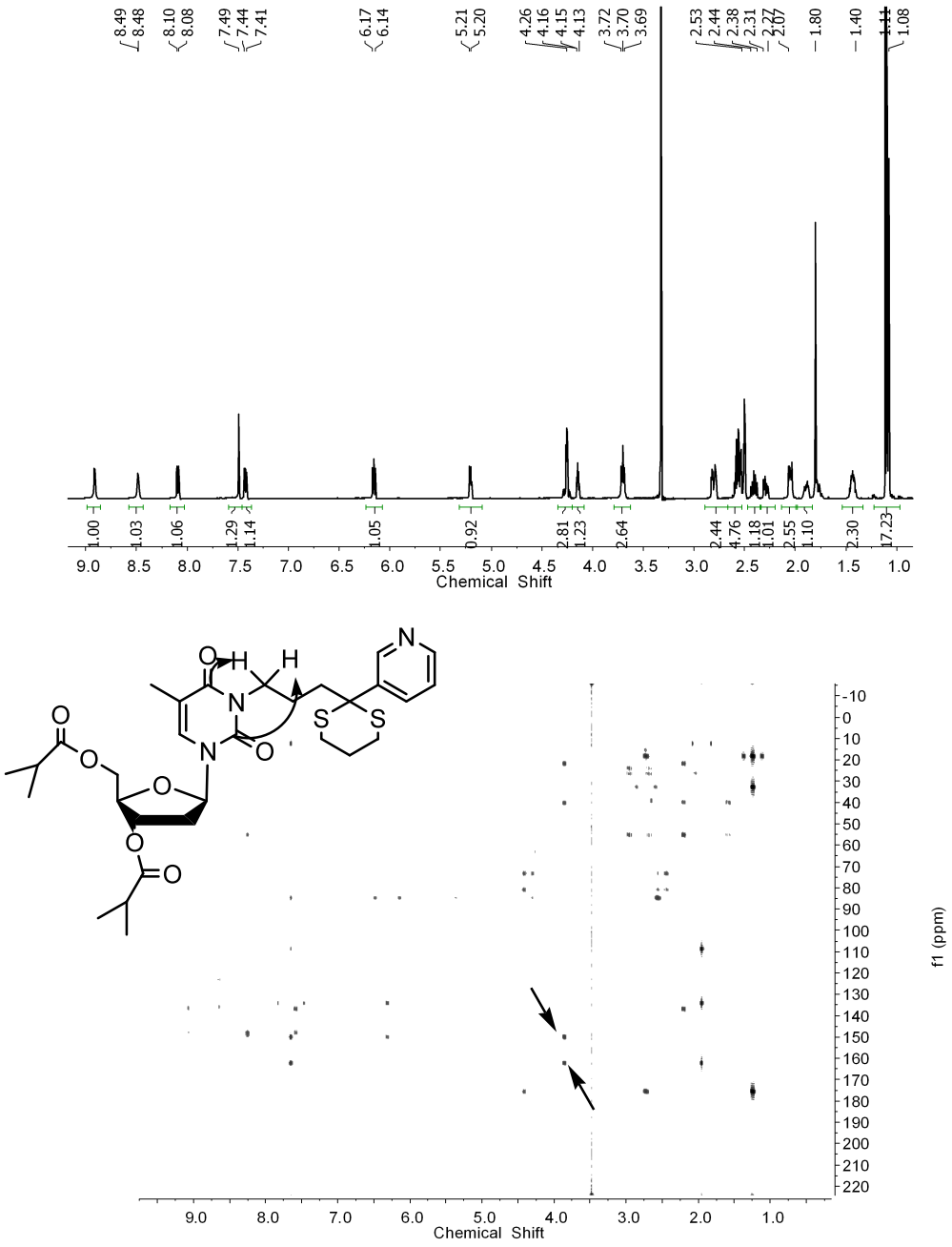
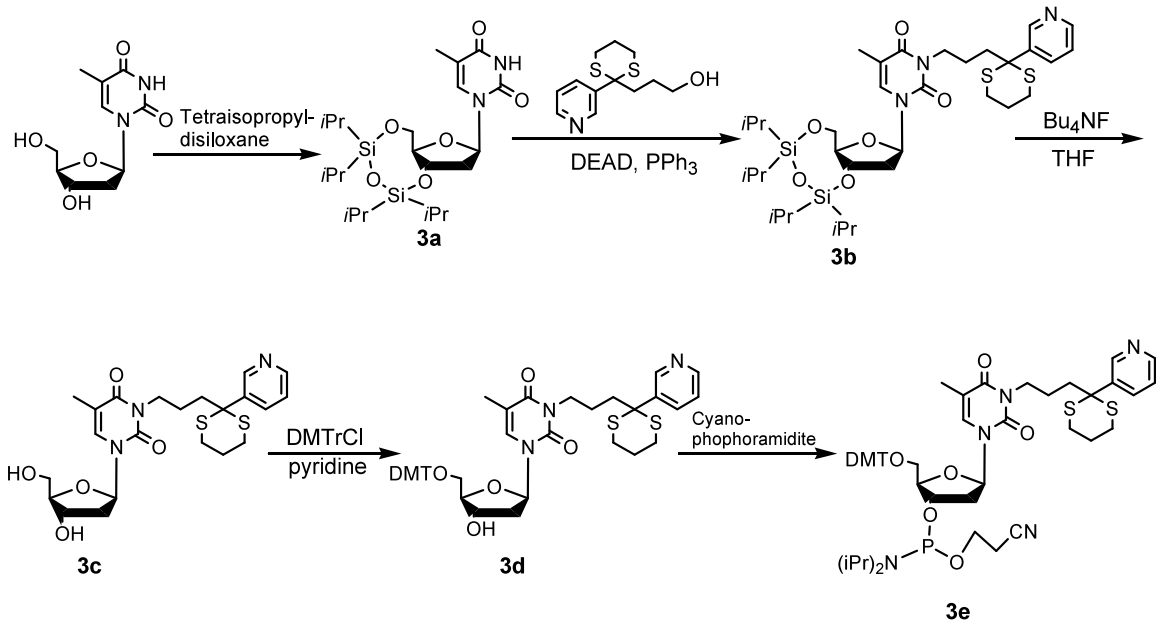


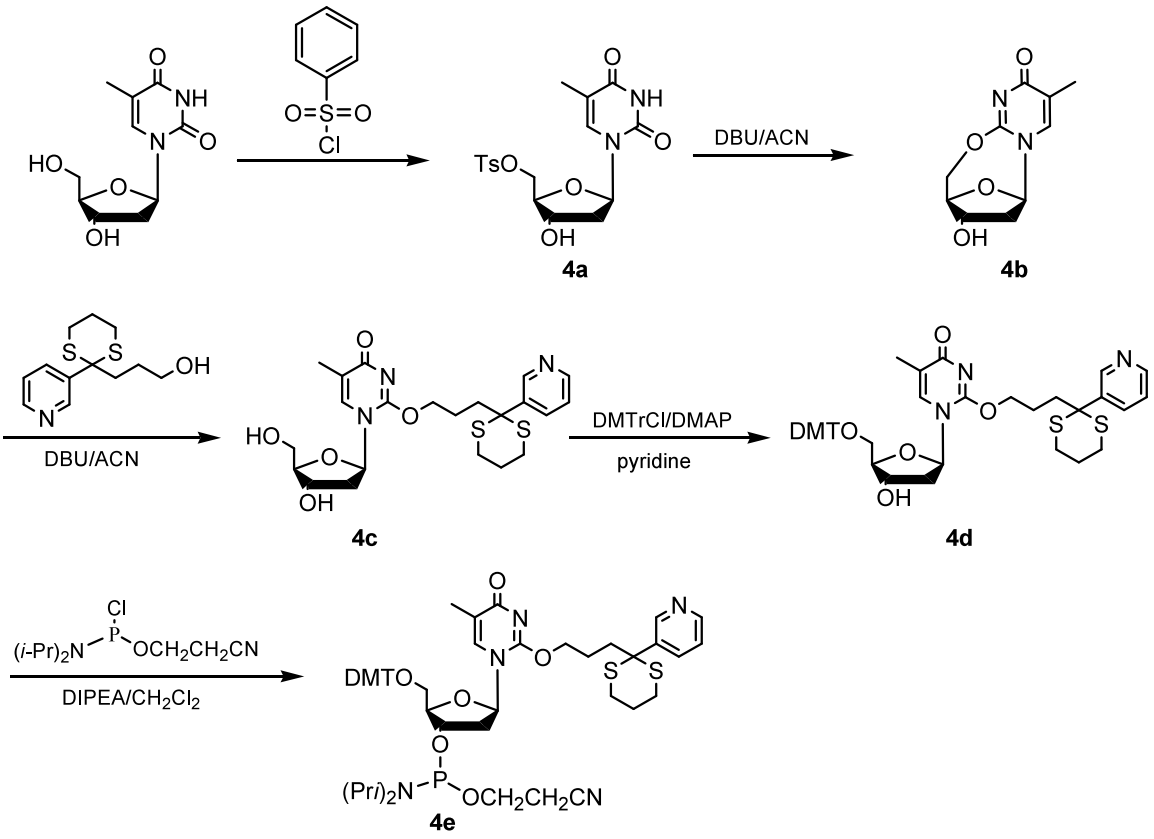
Figure 2-2 ^{1}H NMR (top) and ^{1}H-^{13}C HMBC spectrum (bottom) for *N***3**-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-[isobutyryl]thymidine

Instead of using isobutyric anhydride to protect the 3' and 5' hydroxyl groups of dT, 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane was used to obtain compound **3a**, Scheme 2-4 (43). A Mitsunobu reaction was then carried out to obtain compound **3b**, which was followed by reaction with Bu_4NF and DMTrCl to obtain compound **3d**, Scheme 1-4. Unfortunately the synthesis of the phosphoramidite derivative was not attempted.



Scheme 2-4 Synthetic scheme for compound *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-5'-*O*-[4,4'-dimethoxytrityl]thymidine, **3d**

We synthesized dithiane-protected O^2-POBdT (Scheme 2-5). Generation of O^2-5'-anhydrothymidine, **4b**, was accomplished by refluxing tosylated-thymidine in ACN and DBU. It was crucial to isolate and purify the cyclic intermediate before reacting it with 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanol, **1d,** to yield the dithiane-protected O^2-POBdT (**4c**, Sheme 2-5). The structure of **4c** is supported by ^{1}H-NMR and 2-dimensional HMBC measurements. In particular, the two methylene protons in the pyridyloxobutyl functionality exhibit strong correlation with the C2, but not C4 of the thymine ring in the 2-D HMBC spectrum (Figure 2-3); this distinct spectral feature supports the site of pyridyloxobutylation in this thymidine lesion. Attempts to incorporate the phosphoramidite derivative of **4c** into ODNs failed.



Scheme 2-5 Synthesis of dithiane-protected O^2-POBdT lesion

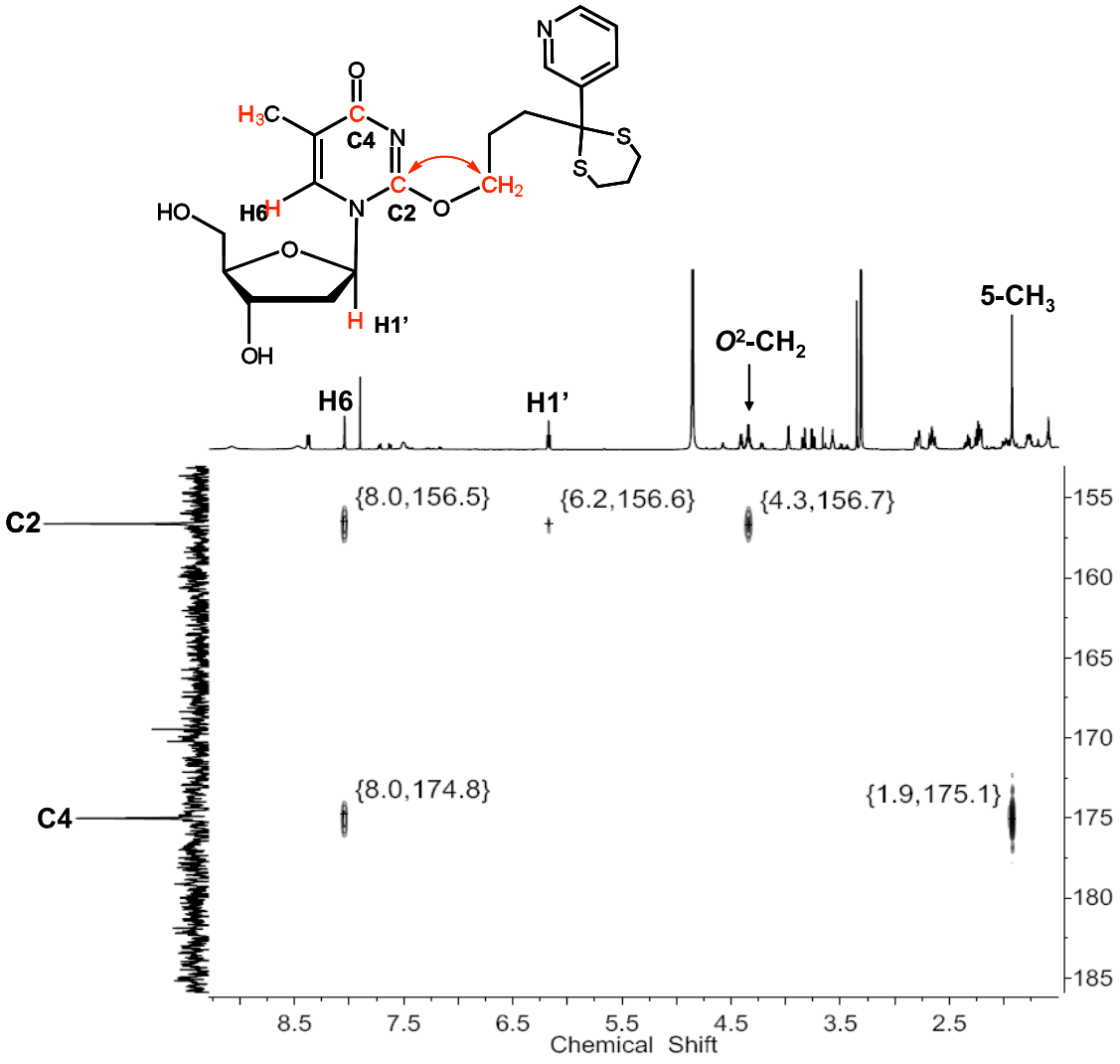


Figure 2-3 ^{1}H-^{13}C HMBC spectrum of dithiane-protected *O*^{2}-POBdT

2.3 Discussion

The synthesis of 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanol was carried out according to the method of Hecht (44). One slight modification was done in the preparation of 2-(3-pyridyl)-1,3-dithiane. Instead of reacting 3-pyridine carboxaldehyde with 1,3-propanedithiol by refluxing for 0.5 h as reported the reaction was refluxed for 3 days to obtain a much better yield. It was discovered that refluxing for only 0.5 h did not generate any product at all as indicated by TLC.

Much milder conditions were used for the synthesis of 3',5'-*O*-diisobutyrylthymidine. Isobutyric anhydride was used to protect both hydroxy groups of dT at room temperature. The new electronic effects of the isobutyl protecting groups on the deoxyribose ring were made apparent on the ^{1}H NMR spectrum as indicated by a downfield shift of 0.64 ppm for the H5' protons. Both 3' and 5' hydroxy proton resonances disappeared; peaks corresponding the methyl protons of the isobutyl group appeared at 1.08 ppm; ESI-MS gave *m/z* 383.2 for the $[M+H]^{+}$ ion, which agreed with the calculated value of 382.7. It was necessary to protect both hydroxy groups to be able to carry out the Mitsunobo reaction in the subsequent reaction for generation of *N*3-[3{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-diisobutyrylthymidine, **2b**, Scheme 2-3. Product yield was quite low and not weighed. However, enough product was obtained for NMR analysis. ^{1}H NMR showed the appearance of a triplet corresponding to H1 protons of the butyl part of the molecule. Additionally, the NMR spectrum was also missing the singlet corresponding to N-H of the pyrimidine ring. ^{1}H-^{13}C HMBC was used to determine the exact structure of the molecule. There were strong cross peaks for H1 of the butyl part with carbons C2 and C4 on the pyrimidine ring of the molecule (Figure 2-2). MS (ESI) gave *m/z* of 620.1 which agreed with the calculated value of 620.2 for the $[M+H]^{+}$ ion. Only enough product was obtained for NMR analysis.

To obtain better yields of the desired *N*3-POBdT, the method of Amin was used (43). 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane was used to protect the 3' and 5' hydroxyl groups. A Mitsunobo reaction was used to obtain the POB adduct, **3b**, Scheme 2-4. This was followed by reaction with Bu_4NF and DMTrCl to obtain compound **3d**, Scheme 2-4. All compounds agreed with NMR and MS data when compared to the published data. No attempt was made to obtain the phosphoramidite derivative.

A previous method had been developed for the synthesis of O^2-methylthymidine (O^2-MedT). The method involves generation of O^2-5'-anhydrothymidine from tosylated-thymidine in dry methanol in the presence of DBU to generate O^2-MedT. A similar approach was used for the synthesis of O^2-POBdT. Upon refluxing tosylated-thymidine in ACN, O^2-5'-anhydrothymidine, **4b**, is formed (Scheme 2-5). It is thought that delocalization of the N3 aromatic electrons causes the carbonyl electrons at C2 of pyrimidine to attack the C5' of the tosylated-thymidine. DBU then removes the exocyclic proton at N3 to generate the thermally stable O^2-5'-anhydrothymidine. However, attempts to react the anhydrothymidine with 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanol, **1d,** without first isolation and purification of anhydrothymidine, failed. Only upon purification of the cyclic intermediate, reaction with 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanol**,** in the presence of DBU, generated dithiane protected O^2-POBdT, **4c**. (Scheme 2-5). Using 2-D HMBC, cross-peaks were detected for O^2-CH_2 and C2 of the pyrimidine ring and O^2-CH_2 and C4 of the pyrimidine ring (Figure 2-3). The yields were low and no optimization reactions were performed.

2.4 Conclusion

*N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-[isobutyryl]thymidine, **2b**, was successfully synthesized and confirmed with 2D NMR. However, the yield was quite low and no optimization reactions were performed. Using a different approach, *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-[4,4'-dimethoxytrityl]thymidine, **3d**, was obtained. Generation of dithiane-protected O^2-POBdT was achieved with low yield.

2.5 Experimental Procedures

Materials. All chemicals, unless otherwise specified, were obtained from Sigma-Aldrich (St. Louis, MO). All anhydrous solvents, unless otherwise specified, were purchased from Across Organics.

Mass Spectrometry (MS). Electrospray ionization-mass spectrometry (ESI-MS) and tandem MS (MS/MS) experiments were carried out on an LCQ Deca XP ion-trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). A mixture of acetonitrile and water (50:50, v/v) was used as solvent for electrospray. The spray voltage was 3.0 kV, and the temperature of the heated capillary was maintained at 275 °C.

NMR. ^{1}H NMR spectra were recorded at 300 Mhz on a Varian Inova 300 NMR spectrometer (Varian Inc., Palo Alto, CA). Two-dimensional NMR spectra were recorded in D_2O using a Variant Unity (Varian, Inc., Palo Alto, CA) spectrometer operating at 500 MHz at 25 °C. Resonance assignments were based on ^{1}H-^{13}C HMBC experiments. The HMBC spectra were

acquired using sweep widths of 5006.3 Hz and 30177.3 Hz for ^{1}H and ^{13}C, respectively. The first delay was set to match a 140-Hz coupling constant, and the second delay was set to match a long-range coupling constant of 8 Hz.

2.5.1. Synthesis of *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-[isobutyryl]thymidine

Synthesis of 3',5'-*O*-[isobutyryl]thymidine (2a, Scheme 2-3) Thymidine (550 mg, 2.06 mmol) was co-evaporated three times with dry pyridine to eliminate residual water. It was then dissolved in 30 ml of ACN. TEA (0.69 ml, 5.0 mmol), 34 mg DMAP, and 0.82 ml (5.0 mmol) isobutyric anhydride were added and the reaction was stirred overnight at room temperature. The crude was purified with silica gel to yield 0.76 g of **2a** (87.4%). ^{1}H NMR (DMSO-d_6) δ 1.01-1.04 (s, 12H, $(CH_3)_2$COO-), 1.71 (s, 3H, H7), 2.17-2.42 (m, 2H, H2', H2'') 3.23 (s, 1H, $(CH_3)_2CH$-), 4.06 (m, 1H, H4')), 4.18-4.22 (m, 2H, H5', H5'')), 5.12-5.14 (m, 1H, H3'), 6.07-6.12 (m, 1H, H1'), 7.39 (s, 1H, H6), 11.29 (s, 1H, N-H); ^{13}C NMR ($(CD_3)_2SO$) δ 12.78, 19.28, 33.79, 36.32, 64.18, 72.01, 81.76, 84.57, 110.40, 136.18, 150.78, 164.44, 176.36; MS (ESI): $[M + H]^+$ calcd *m/z* 383.2, found 382.7.

Synthesis of *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl]-3',5'-*O*-diisobutyryllthymidine (2b, Scheme 2-3). PPh_3 (0.13 g, 0.50 mmol) and 0.142 g (0.56 mmol) of 4-(1,3-dithian-2-yl)-4-(3-pyridyl)butanol were dissolved in 3 ml of ACN. Diethyl azocarboxylate (0.22 ml, 0.49 mmol) and 0.163 g (0.43 mmol) of 3',5'-*O*-diisobutyldT were then added in 3 ml of ACN. The solution was stirred at room temperature for about 7 hrs. An additional 10% of PPh_3 and DEAD and dithiane alcohol was then added and the solution was refluxed overnight. The crude was isolated with silica gel column chromatography (Rf: 0.28). ^{1}H NMR (DMSO-d_6) δ 1.11 (s, 12H), 1.40 (m,

2H, CH_2 dithiane), 1.80 (s, 1H, H7), 1.85 (m, 2H, CH_2 dithiane), 2.07 (m, 2H, CH_2 dithiane), 2.27-2.31 (m, 1H, H2'), 2.31-2.38 (m, 1H, H2''), 2.53 (m, 2H, H3 butyl), 2.60 (m, 2H, H2 butyl), 3.69-3.72 (t, 2H, H1 butyl), 4.13-4.15 (m, 1H, H4'), 4.16-4.26 (m, 2H, H5'), 5.20-5.21 (m, 1H, H3'), 6.14-6.17 (t, 1H, H1'), 7.41-7.44 (m, 1H,), 7.49 (s, 1H, H6), 8.08-8.10 (dd, 1H), 8.48-8.49 (d, 1H), 8.91 (s, 1H); ^{13}C NMR ($(CD_3)_2SO$) δ 13.41, 15.21, 19.27, 22.94, 25.02, 27.49, 33.82, 36.53, 56.21, 61.13, 64.14, 74.19, 81.90, 81.90, 85.75, 109.62, 129.35, 132.22, 133.00, 134.10, 136.89, 150.15, 151.00, 157.21, 163.12, 176.35; MS (ESI): $[M + H]^+$ calcd *m/z* 620.2, found 620.1.

2.5.2. Synthesis of *N*3-[3-{2-(3-Pyridyl)-1,3-dithiane-2-yl}propyl]-5'-*O*-[4,4'-dimethoxytrityl]thymidine

Synthesis of 3',5'-*O*-[1,1,3,3-Tetrakis(isopropyl)-1,3-disiloxanediyl]thymidine (3a, Scheme 2-4) (43). The synthesis of disiloxane thymidine was carried out according to the method of Amin (43). Thymidine (2g, 8.26 mmol) was first dissolved in a small amount of pyridine and evaporated under vacuum. After the second time heat was applied with vacuum for several hours to eliminate any residual amount of water present. It was than dissolved in about 50 ml of dry pyridine. 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane (2.8 ml, 8.88 mmol) was added dropwise under an argon atmosphere. The solution was stirred overnight at room temperature. Afterwards the pyridine was removed and the crude redissolved in CH_2Cl_2 and washed successively with 10% HCl, saturated sodium bicarbonate, and brine. The organic layer was dried, concentrated and purified with silica gel (CH_2Cl_2/MeOH: 99/1) to give 3.5 g of **3a** (88%). ^{1}H NMR (DMSO-d_6) δ 1.01 (s, 24H, $(CH_3)_2$CH-), 1.75 (s, 3H, H7), 2.24-2.42 (m, 2H, H2', H2''), 3.67-3.75 (m, 1H,

H4'), 3.89-4.00 (m, 2H, H5'), 4.51-4.59 (m, 1H, H3'), 5.97-6.01 (t, 1H, H1'), 7.39 (s, 1H, H6), 11.31 (s, 1H, N-H). MS (ESI): $[M + H]^+$ calcd *m/z* 485.2, found 485.1.

Synthesis of *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-[1,1,3,3-tetrakis(isopropyl)-1,3,-disiloxanediyl]thymidine (3b, Scheme 2-4). (43). 3',5'-*O*-[1,1,3,3-Tetrakis(isopropyl)-1,3-disiloxanediyl]-thymidine (180 mg, 0.371 mmol), 104 mg of 4-(1,3-dithian-2-yl)-4-(3-pyridyl)butan-1-ol (0.407 mmol), 107mg (0.408 mmol) of PPh_3, and 0.19 ml (0.417 mmol) of diethyl azodicarboxylate were dissolved in about 3 ml of freshly distilled THF (dried with sodium). The mixture was gently refluxed under argon overnight. The solvent was then evaporated and the crude was purified with silica gel [hexane/EtOAC, 7:3 (v/v)] to yield 188 mg (70.1%) of **3b**. ^{1}H NMR (DMSO-d_6) δ 1.06-1.32 (m, 24H, -CH(CH_3)$_2$, 1.53 (m, 2H, CH_2 dithiane), 1.87 (s, 1H, H7), 1.89-2.04 (m, 4H, CH_2 dithiane), 2.35 (m, 1H, H3 butyl), 2.57-2.79 (m, 4H, H1, H2 butyl), 3.78-3.86 (m, 2H, H5'), 4.58-4.66 (m, 1H, H3'), 5.96-6.00 (m, 1H, H1'), 7.44-7.48 (dd, 1H, H5 pyridyl), 7.51 (s, 1H, H6), 8.31-8.32 (d, 1H, H4 pyridyl), 8.44 (d, 1H, H6 pyridyl), 9.01 (s, 1H, H2 pyridyl). MS (ESI): $[M + H]^+$ calcd *m/z* 722.3, found 722.3.

Synthesis of *N*3-[3-{2-(3-pyridyl)-1,3-dithian-2-yl}propyl]thymidine (3c Scheme 2-4). Compound **3b** (182 mg) was dissolved in 10 ml of THF, to which was added 1 ml of Bu_4NF (10 M) drop by drop at room temperature. After stirring under argon for 1.5 hr, the crude was purified by column chromatography to yield 180 mg of **3c**, above the theoretical yield. MS (ESI): $[M + H]^+$ calcd *m/z* 480.2, found 480.3.

Synthesis of *N*3-[3-{2-(3-pyridyl)-1,3-dithiane-2-yl}propyl]-5'-*O*-[4,4'-dimethoxytrityl]thymidine (3d, Scheme 2-4). Compound **3c** was co-evaporated twice with

pyridine and twice with CH_2Cl_2 prior to the reaction. The compound (121 mg) was subsequently dissolved in 9 ml of dry pyridine to which was added 103 mg of DMTrCl and 2 mg of DMAP. The solution was then stirred at room temperature for 4 h. After 4 h, additional DMTrCl (103 mg) was added and left stirring overnight. The crude was purified with column chromatography to obtain 105 mg of **3d** (53.2). ^{1}H NMR δ 1.00-1.05 (m, 2H, CH_2 dithiane), 1.82 (s, 3H, CH_3), 1.89-1.91 (m, 2H, CH_2 dithiane), 2.00-2.04 (m, 2H, CH_2 dithiane), 2.45-2.49 (m, 2H, H2'), 2.60-2.80 (m, 4H, H2, H3 butyl), 3.77 (m, 9H, -OMe, H5'', H1-butyl), 3.89 (m, 1H, H4'), 4.18 (m, 1H, H5'), 4.50 (m, 1H, 3'OH), 6.35-6.40 (t, 1H, H1'), 6.82 (m, 4H, Ar-H), 7.28-7.45 (m, 9H, Ar-H), 7.50 (s, 1H, H6), 8.19-8.22 (dd, 1H, H5 pyridyl), 8.49 (d, 1H, H4 pyridyl), 8.61 (d, 1H, H6), 9.11 (s, 1H, H2). MS (ESI): $[M + Na]^+$ calcd *m/z*, 804.3, found 804.3.

2.5.3. Synthesis of 4-(1,3-dithian-2-yl)-4-(3-pyridyl)butanol

Synthesis of 2-(3-pyridyl)-1,3-dithiane (1a, Scheme 2-2) (44). 3-pyridine carboxyaldehyde (4.24 ml, 45.2 mmol) was dissolved in 25 ml of freshly distilled THF under an argon atmosphere. To this solution was added 6.01 ml of 1,3-propanedithiol (59.8 mmol) and 1.88 ml of BF_3Et_2O (15 mmol). The reaction was refluxed until most of the starting material was converted into product, as monitored by TLC (3 days). The crude solution was then poured into 100 ml of water and the pH was adjusted to 4.0 with 1 N HCl. The crude was then purified with silica gel (hexane/EtOAc: 2/1) to yield 3.49 g of 2-(3-pyridyl)-1,3-dithiane (34%). ^{1}H NMR δ 1.88-2. 00 (m, 1H, -$SCH_2CH_2CH_2S$-), 2.16-2.24 (m, 1H, -$SCH_2CH_2CH_2S$-), 2.89-3.15 (m, 4H, -$SCH_2CH_2CH_2S$-), 5.18 (s, 1H, C-H), 7.29 (dd, 1H, H5 pyridyl *J*=7.9 Hz, 4.8 Hz), 7.81-7.85 (dt, 1H, H4 pyridyl *J*=7.9 Hz, 1.8 Hz), 8.54-8.54 (dd, 1H, H2 pyridyl, J=4.85 Hz, 1.57 Hz), 8.69 (d, 1H, H6 pyridyl, J=2.05 Hz); MS (ESI): $[M + H]^+$ calcd *m/z* 198.0, found 198.1.

Synthesis of 2-[2-(3-pyridyl)-1,3-dithiane]ethyl-1,3-dioxolane (1b, Scheme 2-2) (44). A solution of 2-(3-pyridyl)-1,3-dithiane (3.49 g, 17.7 mmol) was dissolved in freshly distilled THF (84 ml) under an argon atmosphere at -78 °C. To this solution, *n*-butyllithium (8.5 ml, 2.5 M) was slowly added. The solution was stirred for 1 h at -78 °C. A second solution was then prepared consisting of 2-(2-bromoethyl)-1,3-dioxolane (2.5 ml, 21.3 mmol) in dry THF (9 ml). This solution was then added to the original solution drop by drop at -78 °C and allowed to stir for 2 h at that temperature. The solution was then allowed to warm to room temperature overnight. The solution was subsequently poured into 150 ml of water and extracted with EtOAc several times. The combined organic layers were dried and concentrated to yield the crude product which was purified by silica gel (hexane/EtOAc: 2/1) to give **1b** (3.58 g, 68.1%). ^{1}H NMR ($CDCl_3$) δ 1.63-1.68 (m, 2H, -$SCH_2CH_2CH_2S$-), 1.90-1.99 (m, 2H, -$SCH_2CH_2CH_2S$-), 2.13-2.18 (m, 2H, -$SCH_2CH_2CH_2S$-), 2.60-2.74 (m, 4H, -OCH_2CH_2O-), 3.76-3.91 (m, 4H, -CH_2CH_2-), 4.78 (t, 1H, H2, *J*= 4.38 Hz), 7.31 (dd, 1H, H5 pyridyl, *J*=8.1 Hz, 4.8 Hz), 8.19-8.23 (dq, 1H, H4 pyridyl, *J*=8.11 Hz, 1.63 Hz), 8.50-8.52 (dd, 1H, H2 pyridyl, *J*=4.75 Hz, 1.55 Hz), 9.14 (d, 1H, H6 pyridyl, *J*=2.34 Hz). MS (ESI): $[M + H]^+$ calcd *m/z* 298.1, found, 298.1.

Synthesis of 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanal (1c, Scheme 2-2) (44). Aryl acetal **1b** (3. 58 g, 12.0 mmol) was dissolved in 5 ml of THF and mixed with aqueous oxalic acid solution (0.4M, 1L). The solution was stirred at room temperature for 72 h and extracted with ether (200 ml). The combined organic layers were dried and concentrated to give a yellowish residue. The crude was purified with silica gel (hexane/EtOAc: 1/1) to give **1c** (1.77 g, 58%). ^{1}H NMR ($CDCl_3$) δ 1.92-2.02 (m, 2H, CH_2 dithiane), 2.34-2.38 (m, 2H, CH_2 dithiane), 2.52-2.56 (m, 2H, CH_2 dithiane), 2.60-2.77 (m, 4H, H1 and H2), 7.32 (dd, 1H, H5 pyridyl J=8.12 Hz, 4.72 Hz),

8.17-8.20 (m, 1H, H4 pyridyl), 8.53 (dd, 1H, H2 pyridyl, *J*=4.68 Hz, 1.34 Hz), 9.12 (d, 1H, H6 pyridyl, *J*=2.0 Hz), 9.67 (s, 1H, H1); MS (ESI): $[M + H]^+$ calcd *m/z* 254.1, found 254.2.

Synthesis of 4-(1,3-dithian-2-yl)-4-(3-pyridyl)butan-1-ol (1d, Scheme 2-2) (44). Butanal **1c** (1.77 g, 6.99 mmol) was dissolved in 14 ml of anhydrous THF; a second solution was prepared consisting of $LiAlH_4$ in 7 ml of anhydrous THF. The original solution was then added to $LiAlH_4$ solution at 0 °C, under argon, drop by drop, and stirred at 0 °C for 1 h. H_2O (0.6 ml), 15% NaOH (0.6 ml), and H_2O (1.7 ml) were then added in sequence. The white precipitate was then filtered off and washed with ether. The filtrate was concentrated to yield a yellowish residue. The residue was purified with silica gel (100% EtOAc) to give **1d** (1.29 g, 72.5%). ^1H NMR ($CDCl_3$) δ 1.50-1.60 (m, 2H, CH_2 dithiane), 1.91-2.01 (m, 2H, CH_2 dithiane), 2.04-2.09 (m, 2H, CH_2 dithiane), 2.60-2.74 (m, 4H, H2 and H3), 3.53-3.57 (t, 2H, H1), 7.35 (dd, 1H, H5 pyridyl, *J*=8.07 Hz, 4.71 Hz), 8.22-8.25 (m, 1H, H4 pyridyl), 8.52-8.54 (dd, 1H, H2 pyridyl, *J*=4.82 Hz, *J*=1.6 Hz), 9.16 (d, 1H, H6 pyridyl, *J*=1.7 Hz); MS (ESI): $[H + H]^+$ calcd *m/z* 256.1, found 256.3.

References

1. Belinsky, A. S., Foley, J. F., White, C. M., Anderson, M. W. and Maronpot, R. R. (1990) Dose-Response Relationship between O^6-Methylguanine Formation in Clara Cells and Induction of Pulmonary Neoplasia in the Rat by 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Cancer Res.* 50: 3772-3780.

2. Hecht, S. S., Miglietta, L. M., Foiles, P. G. and Staretz, M. E. (1997) Evidence for an Important Role of DNA Pyridyloxobutylation in Rat Lung Carcinogenesis by 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone: Effects of Dose and Phenethyl Isothiocyanate. *Cancer Res.* 57: 259-266.

3. Peterson, L. A., Kenney, P. M. and Thomson, N. M. (2003) The pyridyloxobutyl DNA adduct, O^6-[4-oxo-4-(3-pyridyl)butyl]guanine, is detected in tissues from 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone-treated A/J mice. *Chem. Res. Toxicol.* 16: 1-6.

4. Hecht, S. S., Upadhyaya, P., Villalta, P. W., McIntee, E. J., Shi, Y., Sturla, S. J., Cheng, G. and Wang, M. (2003) Identification of adducts formed by pyridyloxobutylation of deoxyguanosine and DNA by 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanone, a chemically activated form of tobacco-specific carcinogens. *Chem. Res. Toxicol.* 16: 616-626.

5. Hecht, S. S., Villalta, P. W., Sturla, S. J., Cheng, G., Yu, N., Upadhyaya, P. and Wang, M. (2004) Identification of O^2-substituted pyridimidine adducts formed in reactions of 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-1-butanone and 4-(acetoxymethylnitrosamino)-1-(3-pyridyl)-1-butanol with DNA. *Chem. Res. Toxicol.* 17: 588-597.

6. Idris, A. M., Nair, J., Ohshima, H., Fiesen, M. Brouet, I., Faustman, E. M. and Bartsch, H. (1991) Unsually high levels of carcinogenic tobacco-specific nitrosamines in Sudan snuff (toombak). *Carcinogenesis.* 12(6): 1115-1118.

7. Stepanov, I., Jensen, J., Hatsukami, D. and Hecht, S. S. (2008) New and traditional smokeless tobacco: Comparision of toxicant and carcinogen levels. *Nicotine Tob. Res.* 10(12): 1173-1782.

8. Murphy, S. E., Hecht, S. S. and Jala, J. R. (2005) Cytochrome P450 Enzymes as Catalysts of Metabolism of 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone, a Tobacco Specific Carcinogen. *Chem. Res. Toxicol.* 18(2): 95-110.

9. Smith, T. J., Guo, Z., Gonzalez, F. J., Guengerich, F. P., Stoner, G. D. and Yang, C. S. (1992) Metabolism of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone in human lung and liver microsomes and cytochromes P-450 expressed in hepatoma cells. *Cancer. Res.* 52: 1757-1763.

10. Zhang, X., D'Agostino, J., Wu, H., Zhang, Q., Weymarn, L., Murphy, S. E. and Ding, X (2007) CYP2A13: Variable Expression and Role in Human Lung Microsomal Metabolic Activation of the Tobacco-Specific Carcinogen 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone. *J. Pharmacol. Exp. Ther.* 323(2): 570-578.

11. Brown, P. J., Brdard, L. L., Reid, K. R., Petsikas, D. and Massey, T. E. (2007) Analysis of CYP2A Contributions to Metabolism of 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone in Human Peripheral Lung Microsomes. *Drug Metab. Dispos.* 35(11): 2086-2094.

12. Guengerich, F. P. (1993) Purification and characterization of lung enzymes involved in xenobiotic metabolism. In *Metabolic Activation and Toxicity of Chemical Agents to Lung Tissue and Cells* (Gram, T. E.,) pp 77-87, Pergamon Press Ltd, United Kingdom.

13. Shimada, T., Yun, C. H., Yamazaki, H., Gautier, J. C., Beaune, P. H. and Guengerich, F. P. (1992) Characterization of human lung microsomal cytochrome P-4501A1 and its role in the oxidation of chemical carcinogens. *Mol. Pharmacol.* 41: 856-864.

14. Anttila, S., Hukkanen, J., Hakkola, J., Stjernvall, T., Beaune, P., Edwards, R. J., Boobies, A. R., Pelkonen, O. and Raunio, H. (1997) Expression and localization of CYP3A4 and CYP3A5 in human lung. *Am. J. Respir. Cell Mol. Biol.* 16: 242-249.

15. Hukkanen, J., Pelkonen, O., Hakkola, J. and Raunio, H. (2002) Expression and regulation of xenobiotic-metabolizing cytochrome P450 (CYP) enzymes in human lung. *Crit. Rev. Toxicol.* 32: 391-411.

16. Shimada, T., Yamazaki, H., Mimura, M., Inui, Y. and Guengerich, F. P. (1994) Interindividual variations in human liver cytochrome P450 enzymes involved in the

oxidation of drugs, carcinogens and toxic chemicals: Studies with liver microsomes of 30 Japanese and 30 Caucasians. *J. Pharmacol. Exp. Ther.* 270: 414-423.

17. Ding, X. and Kaminsky, L. S. (2003) Human extrahepatic cytochromes P450: Function in xenobiotic metabolism and tissue-selective chemical toxicity in the respiratory and gastrointestinal tracts. *Annu. Rev. Pharmacol. Toxicol.* 43: 149-173.

18. Shimada, T., Yamazaki, H., Mimura, M., Wakamiya, N., Ueng, Y. F., Guengerich, F. P. and Inui, Y. (1996) Characterization of microsomal cytochrome P450 enzymes involved in the oxidation of xenobiotic chemicals in human fetal livers and adult lungs. *Drug Metab. Dispos.* 24: 515-522.

19. Wei, C., Caccavela, R. J., Weyand, E. H., Chen, S. and Iba, M. M. (2002) Induction of CYP1A1 and CYP1A2 expressions by prototypic and atypical inducers in the human lung. *Cancer Lett.* 178: 25-36.

20. Guengerich, F. P. (1995) Human cytochrome P450 enzymes. *In Cytochrome P450: Structures, Mechanism, and Biochemistry* (Ortiz de Montellano, P.R., Ed.) pp 473-535, Plenum Press, New York.

21. Gervot, L., Rochat, B., Gautier, J. C., Bohnenstengel, F., Kroemer, H., De Berardinis, V., Martin, H., Beaune, P. and De Waziers, I. (1999) Human CYP2B6: expression, inducibility and catalytic activities. *Pharmacogenetics.* 9: 295-306.

22. Wang, M., Cheng, G., Villalta, P. W. and Hecht, S. S. (2007) Development of liquid chromatography electrospray ionization tandem mass spectrometry methods for analysis of DNA adducts of formaldehyde and their application to rats treated with N-nitrosodimethylamine or 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Chem. Res. Toxicol.* 20(8): 1141-1148.

23. Lao, Y., Yu, N., Kassie, F., Villalta, P. W. and Hecht, S. S. (2007) Formation and accumulation of pyridyloxobutyl DNA adducts in F344 rats chronically treated with 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol. *Chem. Res. Toxicol.* 20(2): 235-245.

24. Lao, Y., Yu, N., Kassie, F., Villalta, P. W. and Hecht, S. S. (2007) Analysis of pyridyloxobutyl DNA adducts in F344 rats chronically treated with (R)- and (S)-N'-nitrosonornicotine. *Chem. Res. Toxicol.* 20(2): 246-256.

25. Zhang, S., Wang, M., Villalta, P. W., Lindgren, B. R., Lao, Y. and Hecht, S. S. (2009) Quantitation of pyridyloxobutyl DNA adducts in nasal and oral mucosa of rats treated chronically with enantiomers of N'-nitrosonornicotine. *Chem. Res. Toxicol.* 22(5): 949-956.

26. Hölzle, D., Schlöbe, D., Tricker, A. R. and Richter, E. (2007) Mass spectrometric analysis of 4 hydroxy-1-(3-pyridyl)-1-butanone-releasing DNA adducts in human lung. *Toxicol.* 232(3): 277-285.

27. Mijal, R. S., Thompson, N. M. and Fleischer. (2004) The repair of the tobacco specific nitrosamine derived adduct O^6-[4-oxo-4-(3-pyridyl)butyl]guanine by O^6-alkylguanine-DNA alkyltransferase variants. *Chem. Res. Toxicol.* 17(3): 424-434.

28. Li, L., Perdigao, J., Peff et al. (2009) The influence of repair pathways on the cytotoxicity and mutagenicity induced by the pyridyloxobutylation pathway of tobacco-specific nitrosamines. *Chem. Res. Toxicol.* 22(8): 1464-1472.

29. Pauly, G.T., Peterson, L. A. and Moschel, R. C. (2002) Mutagenesis by O^6-[4-oxo-4-(3-pyridyl)butyl]guanine in *Escherichia coli* and human cells. *Chem. Res. Toxicol*. 15(2): 165-169.

30. Bhanot, O. S., Grevatt, P. C., Donahue, J. M., Gabrielides, C. N. and Solomon, J. J. (1992) In vitro DNA replication implicates O^2-ethyldeoxythymidine in transversion mutagenesis by ethylating agents. *Nucleic Acids Research.* 20(3): 587-594.

31. Hecht, S. S., Villalta, P. W., Sturla, S. J. and Wang, M. (2004) Identification of O^2-Substituted Pyrimidine Adducts Formed in Reactions of 4-(Acetoxymethyllnitrosamino)-1-(3-pyridyl)-1-butanone and 4-(Acetoxymethylnitrosamio)-1-(3-pyridyl)-1-butanol with DNA. *Chem. Res. Toxicol.* 17: 588-597.

32. Zhang, S., Wang, M. and Villalta, P.W. (2009) Analysis of pyridyloxbutyl and pyridylhydroxybutyl DNA adducts in extrahepactic tissues of F344 rats treated chronically with 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and enantiomers of4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol. *Chem. Res. Toxicol.* 22(5): 926-93.

33. Österdahl, B., Jansson, C. and Paccou, A. (2004) Decreased Levels of Tobacco-Specific *N*-Nitrosamines in Moist Snuff on the Swedish Market. *J. Agric. Food Chem.* 52: 5085-5088.

34. Foiles, P. G., Peterson, L. A., Miglietta, L. M. and Ronai, Z. (1992) Analysis of mutagenic activity and ability to induce replication of polyoma DNA sequences by different model metabolites of the carcinogenic tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Mut. Res.* 279: 91-101.

35. Nanxiong, Y., Lao, Y., Kassie, F., Villalta, P. W. and Stephen, S. S. (2007) Formation and Accumulation of Pyridyloxobutyl DNA Adducts in F344 Rats Chronically Treated with 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanone and Enantiomers of Its Metabolite, 4-(Methylnitrosamino)-1-(3-pyridyl)-1-butanol. *Chem. Res. Toxicol.* 20: 235-245.

36. Mijal., R. S., Loktionova, N. A., Vu, C. C., Pegg, A. E. and Peterson, L. A. (2005) O^6-Pyriyloxobutyguanine Adducts Contribute to the Mutagenic Properties of Pyridyloxobutylating Agents. *Chem. Res. Toxicol.* 18: 1619-1625.

37. Simonelli, V., Narisco, L., Dogliotti., E. and Fortini, P. (2005) Base excision repair intermediates ae mutagenic in mammalian cells. *Nucleic Acids Research.* 33(14): 4404-4411.

38. Ronai, Z. A., Gradia, S., Peterson, L. A. and Hecht, S.S. (1993) G to A transitions and G to T transversions in codon 12 of the Ki-ras oncogene isolated from mouse lung tumors induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and related DNA methylating and pyridyloxobutylating agents. *Carcinogeneis.* 14(11): 2419-2422.

39. Mijal, R. S., Hahn, J. N., Li, L. et al. (2008) Mgmt deficiency alters the in vivo mutational spectrum of tissues exposed to the tobacco carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK). *Carcinogenesis.* 29(4):866-874.

40. Pressentin, M. D. M. von., Kosinska, W. and Guttenplan, J. B. (1999) Mutagenesis induced by oral carcinogens in lacZ mouse (MutaMouse) tongue and other oral tissues. *Carcinogenesis.* 20(11): 2167-2170.

41. Pressentin, M. D. M. von., Chen, M. and Guttenplan, J. B. (2001) Mutagenesis induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and N-nitrosonornicotine in lacZ upper aerodigestive tissue and liver and inhibition by green tea. *Carcinogenesis.* 22(1): 203-206.

42. Hashimoto, K., Ohsawa, K.I. and Kimura, M. (2004) Mutations induced by 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in the lacZ and cII genes of MutaMouse. *Mutat. Res.* 560(2): 119-131.

43. Krishnegowda, G., Sharma, A. K., Krzeminski, J. et al. (2011) Facile Syntheses of O^2-[4-(3-Pyridyl-4-oxobutyl-1-yl]thymidine, the Major Adduct Formed by Tobacco Specific Nitrosamine 4-methylnitrosamino-1-(3-pyridyl)-1-butanone (NNK) *in vivo*, and Its Site-Specifically Adducted Oligodeoxynucleotides. *Chem. Res. Toxicol.* 24: 960-967.

44. Lin., J., Amin, S., Murphy. S. E., Solomon, J. J. and Hecht, S. S. (1993) Synthesis of [3, 3-D2] 4-hydroxy-1-(3-pyridyl)-1-butanone, an internal standard for analysis of tobacco-specific nitrosamine hemoglobin and DNA adducts. *J. Label Compd. Radiopharm.* 38(4): 285-292.

CHAPTER 3 Syntheses of 5-Hydroxymethyl-2'-deoxycytidine and 5-formyl-2'-deoxycytidine and their Incorporation into Oligodeoxynucleotides

3.1 Introduction

5-hydroxymethyl-2'-deoxycytidine (hmdC) is a modified nucleoside that has been recently detected in genomic DNA of mammalian tissues (2, 3). Tet enzymes oxidize 5-methylcytosine (mC) in DNA to yield hmdC (2, 3, 6, 7). HmdC plays a crucial role in the development and differentiation of embryonic stem cells (ES cells). The levels of hmdC and of Tet enzymes vary upon differentiation of a stem cell (7).

Another recently discovered nucleoside is 5-formyl-dC (fdC). fdC, like hmdC, has recently been detected in genomic DNA and its exact role is yet unknown. Its levels are even lower than that of hmdC (5). However, because it can be artificially produced *in vitro* upon exposure to ionizing radiation or by reaction with Fenton reagents, from mC, and because it was thought that fdC was responsible for the high rate of mutations observed at CpG sites, site-specific mutagenesis studies have been conducted (23). These studies indicate that fdC is not a potent mutagen (24). We report here the synthesis of fdC and hmdC for incorporation into synthetic oligodeoxynucleotides (Figure 3-1).

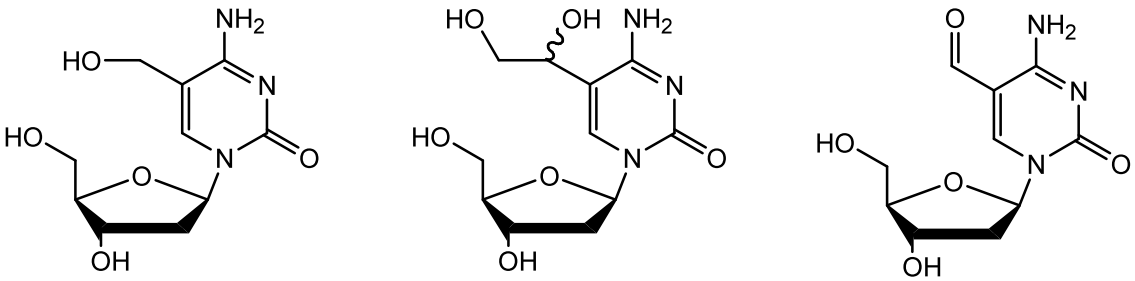


Figure 3-1 Structures of hmdC, dheC, and fdC

3.2 Detection and Sequencing of hmdC in genomic DNA

The first time hmdC was detected was by simply using TLC plates to visualize ^{32}P-dNMPs that had been obtained from digested DNA. The intensity of the new spot was inversely proportional to the mC spot indicating that hmdC was derived from mC (2). Further confirmation on the exact structure was obtained with HPLC and MS (2). Other methods use enzymes that cleaved DNA at specific cytosine sequences (3). A more accurate method that is independent of nucleotide sequence uses β-glucosyltransferase (β-gt) to attach a radioactive labeled UDP-[^{3}H]glucose onto the hydroxy group of hmdC in genomic DNA and that allows for detection and quantification in several tissues (7). A biotin tag can also be attached to a UDP-6-azidoglucose of hmdC for detection, purification, and sequencing (12). J-binding protein coupled to magnetic beads can also be used to capture glycosylated hmdC (13). One method allows for detection and sequencing at the same time. Single DNA polymerase molecules are observed in real time while they catalyze the incorporation of fluorescently labeled nucleotides complementary to a template nucleic acid strand (9). For example, incorporation of a mC or hmdC would give rise to a unique fluorescence trace. Only *in vitro* studies with ODNs have been conducted using this technique (9). Nanopore detection and sequencing has also been used (16).

3.2.1. HmdC levels in different cells and tissues

Very recently, hmdC was discovered in Purkinje neurons and granule cells (2). Although the percentages that were given were estimates of hmdC levels at CpG sites, 0.6% and 0.23%, in Purkinje and neuronal cells respectively, its discovery initiated a race to detect, quantify, and elucidate its role in genomic DNA (2). In human embryonic kidney cells, hmdC was present at 4-

6% at specific cytosine sequences. These short cytosine sequences were obtained using specific restrictive enzymes and thus the percentages reported are likely overestimates (3). Using HPLC-MS, Carell et al. detected hmdC in mouse tissues, mainly in the central nervous system, at levels from 0.3 to 0.6% (4). These detection methods target specific sequences and the percentages reported may not reflect the true abundance of hmdC in genomic DNA.

Other studies that detect hmdC independent of nucleotide sequence indicate hmdC is present from 0.2 to 1.2%. For example, hmdC was detected in the cortex at about 1.2%, cerebellum 0.9%, kidney 0.7%, and testis 0.2% (7). The method used here allows for the detection and quantification of the levels of this modified nucleoside in genomic DNA. Similar methods report highest hmdC levels at 0.4% in mouse cerebellum (12).

3.3 Detection and mutagenicity of fdC

In 2011, fdC was detected in mouse ES cells at very low levels (0.02%) (5). However, long before this, fdC was artificially produced and incorporated into oligodeoxynucleotides (ODNs) for *in vitro* studies (22). The purpose was to examine if fdC was mutagenic. It was concluded that fdC may cause C→T mutations at CpG sites (22). It was found that the Klenow fragment of *E. coli* misincorporated TMP and dAMP opposite fdC (23). Authors speculated that fdC my cause C→A and C→T mutations during replication (23). A different *in vitro* experiment revealed similar results but also revealed base substitutions (mutations) at sites adjacent to fdC. *In vivo* studies using fdC-containing vectors showed that fdC weakly blocked DNA replication in mammalian cells. The mutation spectrum reported was broad (fdC→G, fdC→A, and fdC→T). However, the mutational frequencies were low (0.03–0.28%) (24). Thus, fdC is not at all a potent mutagen. It has been speculated that fdC is derived from hmdC and then removed.

Thymidine DNA glycosylase (TDG) removes fdC at G:fdC base pairs faster than for G:T base pairs (26). However, given the overwhelming evidence that hmdC functions as epigenetic modification in ES cells, oxidizing hmdC to fdC would have deleterious effects in gene expression and development of ES cells.

3.4 The role of hmdC and Tet enzymes during ES cell differentiation

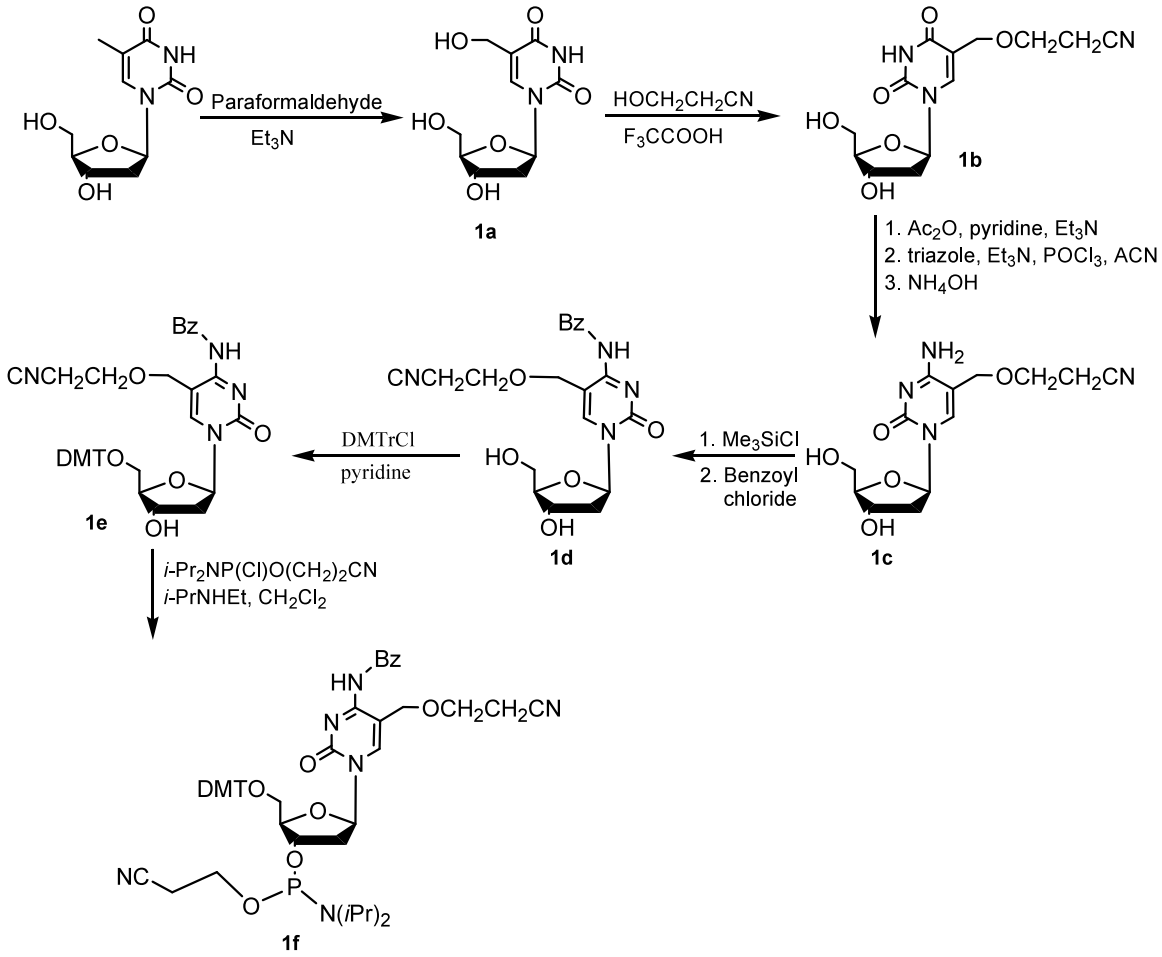
Current research on hmdC is focused in its role in the differentiation of embryonic stem cells (ES cells). Tet enzymes, Tet1, Tet2, and Tet3 oxidize 5-methyl-2'-deoxycytIdine (mdC) to form hmdC. An ES cell maintains high levels of Tet transcripts (Tet1 Tet2) and hmdC, but as the cell differentiates, hmdC levels and Tet1 and Tet2 transcript levels drop while Tet3 transcript levels increase (7). It appears the hmdC levels are correlated with gene activation; in Tet1 knockdown ES cells, increase in methylation at the *Nanog* promoter region was detected (6). When fragmented genomic DNA was glucosylated at hmdC nucleotides and pulled down with J-binding proteins, promoters of the pluripotency genes (*Nanog, Oct4, Klf4*) were enriched in hmdC by 7-14 fold (13). Other studies showed decreases in Tet1, Tet2 and Oct4 mRNA levels while Tet3 levels increased during differentiation; embryonic fibroblasts differentiating into ES cells showed the opposite trend. Oct4 depletion down-regulated Tet1, Tet2 and hmdC levels, while up-regulating Tet3 levels. Prolonged depletion of Tet1 or Tet2 also affected normal ES cell differentiation as it led to abnormal cell tissue such as high levels of immature glandular tissue or scattered giant cells relative to healthy cells (11).

3.5 The intention behind the syntheses of hmdC and fdc

Published methods were used to synthesize both hmdC and fdC (27, 28). Our objective was to synthesize ODNs housing a site-specifically incorporated hmdC and fdC. These ODNs will then be used for future studies on how these modified nucleosides affect DNA replication in mammalian cells and how they are recognized by proteins in mammalian cells.

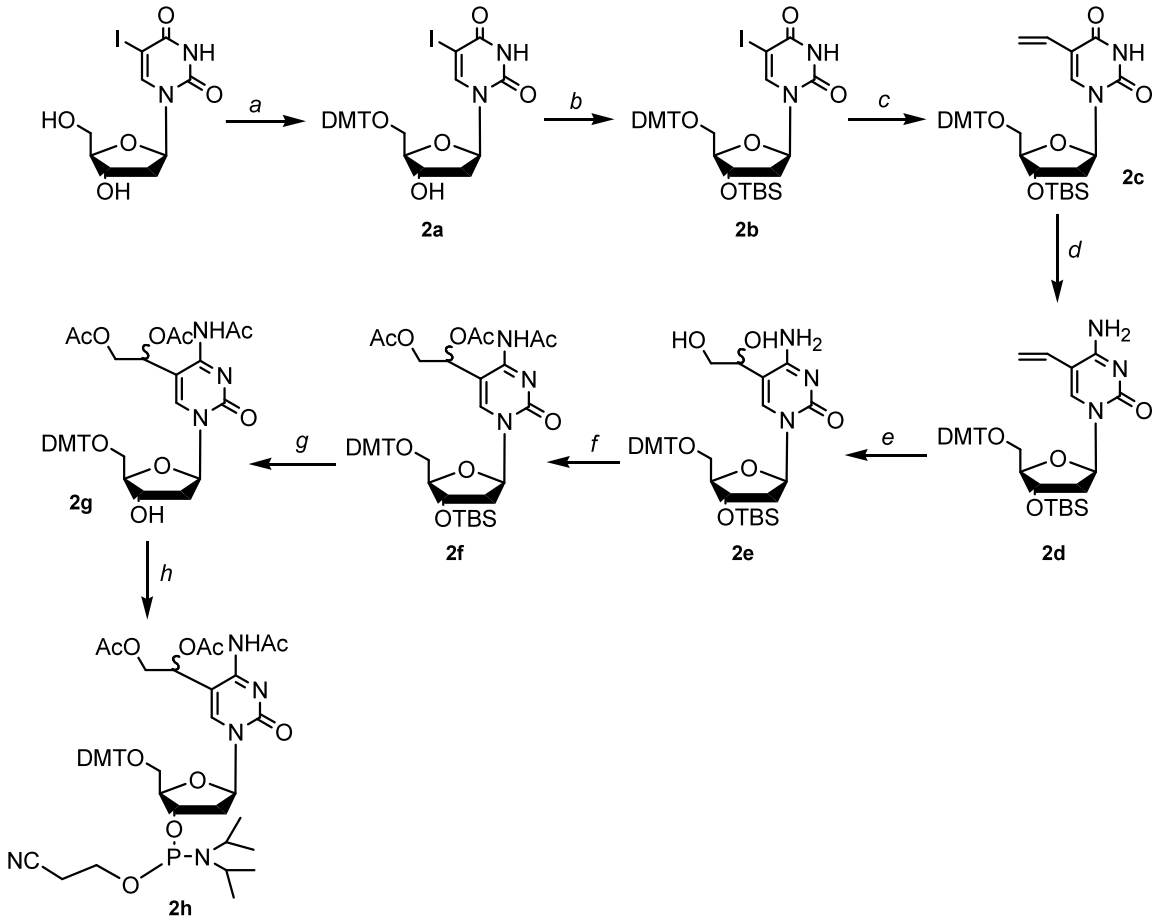
3.6 Results and Discussion

The synthesis of hmdC was conducted using the method of Sowers (27). The key reaction step was the conversion of 5-hydroxymethyl-2'-deoxyuridine (hmdU) **1b**, to the dC analog, **1c**, Scheme 3-1. The yield for this step was modest (39.6%). ^{1}H NMR and MS were used to verify the synthesis of **1f**, Scheme 3-1 and Appendix B. A number of attempts were made to incorporate **1f** into ODNs but without success. Most likely the incorporation did occur but the harsh postsynthetic de-protection of the cyanoethyl group did not work. After ODN synthesis was complete the solid support was placed in concentrated ammonia and heated at 65 °C for 60 h (27). The ODN was then purified using HPLC and the fractions collected were analyzed with MS (ESI) and MS/MS which showed no evidence of the desired ODN harboring an hmdC. When **1b** was de-protected using the same conditions to de-protect the ODN, removal of the cyanoethyl group did not occur. Further heating did not work either. Thus, most likely incorporation did occur but was not de-protected with the conditions reported by Sowers.



Scheme 3-1 Synthetic scheme for hmdC phosphoramidite

For the synthesis of fdC, the method of Matsuda was used (27). The conversion of vinyl-dU (**2c**) to vinyl-dC (**2d**) was very inefficient (Scheme 3-2). In addition, the phosphoramidate derivative that was incorporated into the ODN did not contain the 5-formyl functionality but instead contained 5-(1,2-dihydroxyethyl)-2'-deoxycytidine, (dheC), that required postsynthetic NH_4OH treatment to remove the acetyl groups and reaction with $NaIO_4$ to generate the 5-formyl functionality of fdC. ^1H NMR and MS were used to confirm the structure of **2h** (Scheme 3-2). A 12-mer in the sequence (ATGGCGXGCTAT) was successfully synthesized, where "X" represents dheC, the precursor to fdC, Scheme 3-2, at the 6-position and was confirmed with MS and MS/MS (Figure 3-2 and 3-3).



Scheme 3-2 Synthetic scheme for dheC phosphoramidite. Conditions: (a) DMTrCl, DMAP, pyridine, (b) TBDMCl, (c) $Bu_3SnCH{=}CH_2$, $(PH_3P)_2PdCl_2$, DMF, 80 °C, 1.5 h; (d) (1) 2,4,6-triisopropylbenzenesulfonyl chloride, DMAP, Et_3N, ACN, RT, 28 h; (2) conc. NH_4OH, 0 °C-rt, 1 h; (e) OsO_4, 4-methylmorpholine-*N*-oxide, acetone-H_2O-*t*-BuOH (4:1:1), RT, 4h; (f) Ac_2O, DMAP, py, rt, 20 h; (g) TBAF, THF, 0 °C-RT, 6 h; (h) *i*-$Pr_2NP(Cl)O(CH_2)_2CN$, *i*-PrNHEt, CH_2Cl_2, 0 °C-RT, 0.5 h

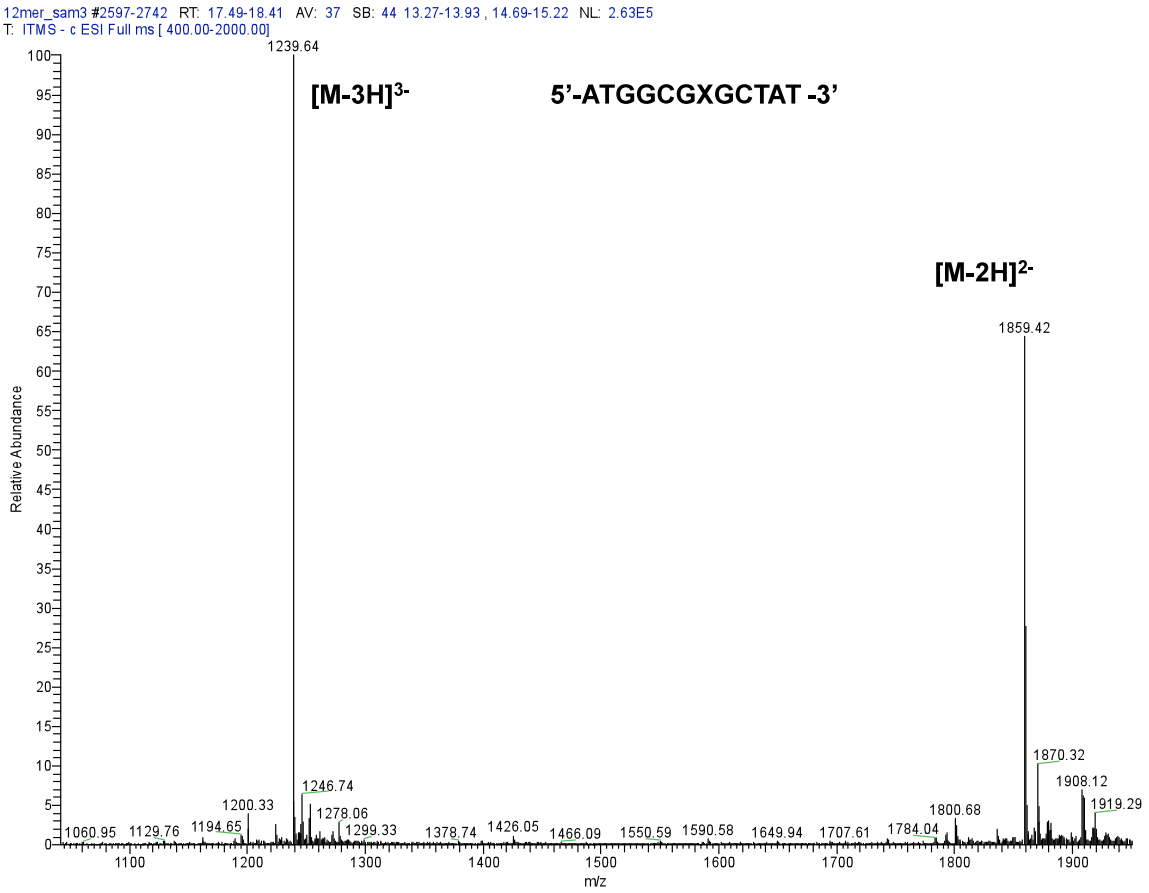


Figure 3-2 The negative-ion ESI-MS for the modified 12mer

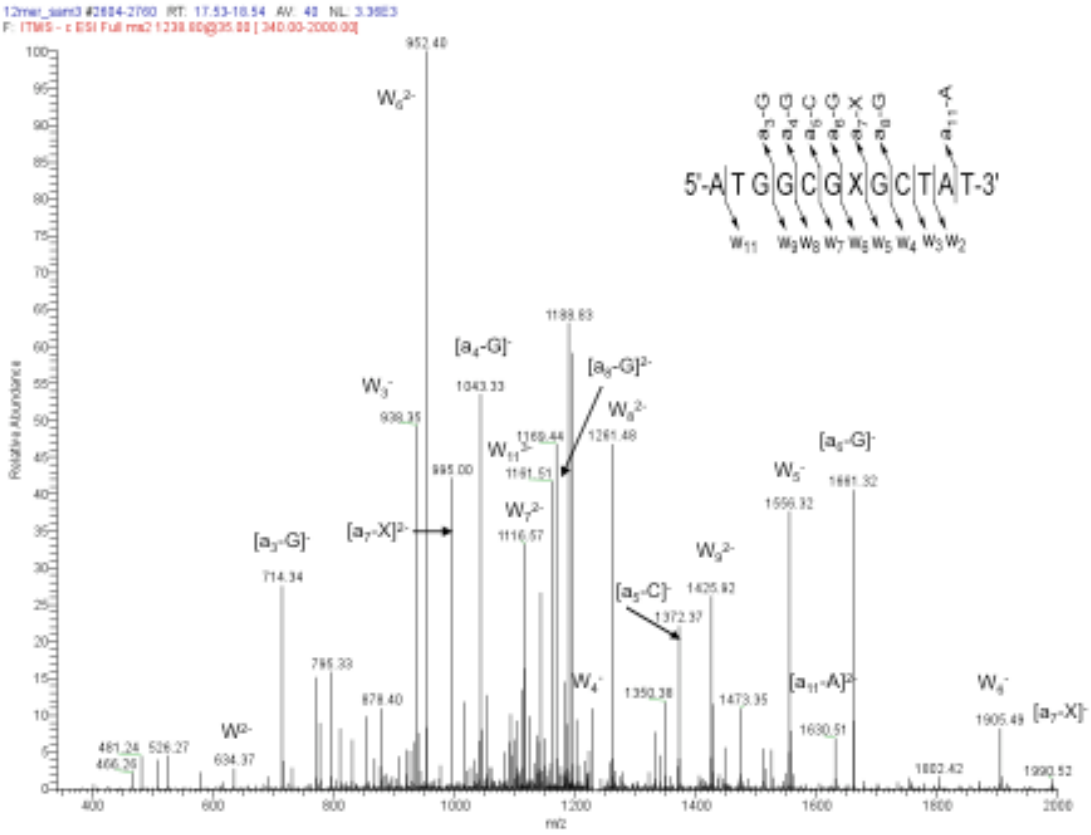


Figure 3-3 The product-ion spectrum of the ESI-produced $[M-3H]^{3-}$ ions of d(ATGGCGXGCTAT), where "X" represents dheC

The ODN (12mer) obtained was characterized by ESI-MS and MS/MS to confirm the sequence and location of dheC. The measured *m/z* values for the multiply deprotonated ions is consistent with the calculated values. MS/MS of the $[M-3H]^{3-}$ ion (*m/z* 1239.64) gives a series of sequence ions, allowing for the absolute confirmation of dheC at the 6th position from the 3' end on the ODN (Figure 3-3). Upon collisional activation, an ODN undergoes cleavages at the *N*-glycosidic bond and the 3' C-O bond of the same nucleoside to form [a_n-Base] and its complementary w_n ions, which carry the 5' and 3' termini of the ODN, respectively. Thymine usually does not undergo these types of cleavages (28). Figure 3-3 shows the [a_3-G], [a_4-G], [a_5-C], [a_7-X], [a_8-G], [a_{11}-A], w_2, w_3, w_4, w_5, w_6, w_8, w_9, and w_{11} ions. The observed masses of w_6, w_8, w_9, and w_{11} are 60 Da higher in masses than the calculated masses of the corresponding fragments for the unmodified ODN, whereas the masses of the w_2, w_3, w_4, and w_5 ions are the same as the calculated masses for the corresponding fragments from the unmodified ODN, indicating the presence of dheC at the 6th position of the 12mer. Analyzing the [a_n-Base] ions gives the same results.

3.7 Conclusion

The syntheses of phosphoramidites, building blocks of hmdC and fdC were successfully synthesized, though the incorporation of hmdC into ODNs failed. Most likely incorporation did occur but the postsynthetic conditions did not remove the cyanoethyl group used for protecting the hydroxymethyl group during ODN synthesis. A 12mer containing the precursor for fdC was successfully obtained and confirmed with MS (ESI) and MS/MS.

3.8 Experimental Procedures

Materials. All chemicals, unless otherwise specified, were obtained from Sigma-Aldrich (St. Louis, MO). All anhydrous solvents, unless otherwise specified, were purchased from Acros Organics.

Mass Spectrometry (MS). Electrospray ionization-mass spectrometry (ESI-MS) and tandem MS (MS/MS) experiments were carried out on an LCQ Deca XP ion-trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). A mixture of acetonitrile and water (50:50, v/v) was used as solvent for electrospray. The spray voltage was 3.0 kV, and the temperature of the ion transport tube was maintained at 275 °C.

ODN Synthesis and Purification. ODN (12mer) was synthesized on a Beckman Oligo 1000S DNA synthesizer (Fullerton, CA) at 1-μmol scale. The synthesized phosphoramidite building block was dissolved in anhydrous acetonitrile at a concentration of 0.07 M. Standard phosphoramidite building blocks (Glen Research Inc., Sterling, VA) of dA, dC, dG, and dT were employed and standard ODN assembly protocol was used without any modification. The ODN containing 5-(1,2-dihydroxyethyl)-2'-deoxycytidine (dheC) linked to the solid support was treated with concentrated NH_4OH at 55 °C for 12 h. The purification of the ODN was performed on a Beckman HPLC System (32 Karat software version 3.0, pump module 125) with a UV detector (module 126) monitoring at 260 nm. A C-18 column was used and a linear gradient of ACN from 0 to 35% in 0.1 M TEAA buffer, pH 7.0, was used for the mobile phase.

NMR. ^{1}H NMR spectra were recorded at 300 MHz on a Varian Inova 300 NMR spectrometer (Varian Inc., Palo Alto, CA).

Synthesis of 5-hydroxymethyl-2'-deoxyuridine (1a, Scheme 3-1). (27) 2'-Deoxyuridine (4g) was dissolved 0.5 M TEA in water (80 ml total). To this solution was added 2.36 g paraformaldehyde and the solution was heated at 60 °C overnight. After day one, 3.43 g of paraformaldehyde was added followed by 2.44 g after day two. Each time additional paraformaldehyde was added, 1 ml TEA and 10 ml water was added to keep the solution basic. The crude was purified by silica gel to yield 2.35 g of 5-hydroxymethyl-2'-deoxyuridine (51.9%). ^{1}H NMR (DMSO-d_6) δ, 11.30 (s, 1H, NH), 7.72 (s, 1H, H6), 6.16 (t, 1H, H1', J=6.0 Hz), 5.24 (d, 1H, 3'OH), 4.90-4.96 (m, 1H, 5'OH), 4.14 (m, 1H, H3'), 4.12 (s, 2H, 5-CH_2O-), 3.87 (m, 1H, H4'), 3.77 (m, 2H, 5H', H5''), 2.07-209 (m, 2H, H2'-H2''), 0.93-0.97 (t, 1H, 5-CH_2OH). MS (ESI): $[M + H]^+$ calcd *m/z* 259.1, found 258.9.

Synthesis of 5-(2-cyanoethoxy)methyl-2'-deoxyuridine (1b, Scheme 3-1). (27) 5-Hydroxymethyl-2'-deoxyuridine (2.35 g, 9.10 mmol) in 2-cyanoethyl alcohol (47 ml) was treated with a catalytic amount of trifluoroacetic acid (470 μl). The solution was stirred at 130 °C under vacuum for 1 h. The product was purified with silica gel (1.5 g, 53.0%); ^{1}H NMR (DMSO-d_6) δ, 11.40 (s, 1H, NH), 7.93 (s, 1H, H6), 6.15 (t, 1H, H1', J=6.0 Hz), 5.23-5.25 (d, 1H, 3'OH), 5.01 (t, 1H, 5'OH), 4.21 (m, 1H, H3'), 4.17 (s, 2H, 5-CH_2O-), 3.77 (m, 1H, H4'), 3.61 (m, 4H, H5', H5'', CH_2CH_2CN), 2.72 (t, 2H, CH_2CH_2C N, J=6.3 Hz), 2.08-2.12 (m, 2H, H2', H2''), 1.14-1.19 (t, 2H, 5-CH_2OH). HRMS (ESI): $[M + H]^+$ calcd *m/z* 334.1, found 334.1.

Synthesis of 3',5'-bis-*O*-Acetyl-5-(2-cyanoethoxy)methyl-N^4-triazoyl-2'-deoxycytidine. (27) 5-(2-Cyanoethoxy)methyl-2'-deoxyuridine (1.5 g) was dried by co-evaporation with dry pyridine. Pyridine (45ml) was than added, followed by dropwise addition of acetic anhydride (3.4 ml). The solution was stirred at room temperature for 3 h. Anhydrous methanol was added and the mixture was stirred for an additional 5 min. Solvents were removed under reduced pressure and the product was dried to a foam by co-evaporation with toluene. The acetylated derivative was purified with silica gel to obtain 1.96 g.

The above acetylated derivative (624 mg) was dissolved in 12.5 ml of CH_3CN, to which solution was added 1,2,4-triazole (2.2 g, 31.9 mmol) followed by addition of Et_3N and $POCl_3$ (0.66 ml, drop by drop). The solution was stirred at room temperature for 1 h. Stirring was continued for 1 h upon which ethyl acetate was added and washed with saturated $NaHCO_3$ and brine. The organic solution was dried over anhydrous Na_2SO_4, concentrated and purified with silica gel to obtain 0.61g (87%). ^{1}H NMR (DMSO-d_6) δ 9.38 (s, 1H, H6), 8.41 (s, 2H, H3 and H5 triazole), 6.14 (t, 1H, H1'), 5.22-5.24 (d, 1H, H3'), 4.77 (s, 2H, 5-CH_2O-), 4.45 (m, 1H, H4'), 4.31 (m, 2H, 5H', H5'), 3.67 (t, 2H, $CNCH_2CH_2O$-), 2.68-2.75 (t, 2H, $CNCH_2CH_2O$-), 2.60-2.66 (m, 1H, H2'), 2.35-2.45 (m, 1H, H2''), 2.08 -2.02 (m, 6H, CH_3-). MS (ESI): $[M + H]^+$ calcd *m/z* 447.2, found 447.1.

Synthesis of 5-(2-cyanoethoxy)methyl-2'-deoxycytidine (1c, Scheme 3-1). (27) After isolation and purification of the acetylated triazolyl derivative, it was dissolved in 1,4-dioxane (10 ml) and treated with 29% aqueous ammonia (2ml). After achieving complete conversion, as monitored by TLC, the solvent was evaporated and the crude was re-dissolved in 0.4M TEA in MeOH an allowed to react overnight at room temperature to obtain the desired 3',5'-deprotected cyanoethoxy protected dC compound, 172 mg (39.6%). ^{1}H (DMSO-d_6) δ 7.86 (s, 1H, H6), 7.44

(s, 1H, NH), 6.68 (s, 1H, NH), 6.13 (t, 1H, H1', *J*=6.0 Hz), 5.18-5.19 (d, 1H, 3'OH), 4.98-5.03 (m, 1H, 5'OH), 4.21 (m, 3H, H3', 5-CH_2O-), 3.75 (m, 1H, H4'), 3.54-3.58 (m, 4H, H5', H5'', -CH_2CH_2CN), 2.74-2.78 (m, 2H, -CH_2CH_2CN), 1.89-2.00 (m, 1H, H2'), 2.08-2.16 (m, 1H, H2''). MS (ESI): $[M + H]^+$ calcd *m/z* 311.1, found 311.1.

Synthesis of N^4-benzoyl-5-(2-cyanoethoxy)-methyl-2'-deoxycytidine (1d, Scheme 3-1). (27) 5-(2-Cyanoethoxy)methyl-2'-deoxycytidine (172 mg, 0.55 mmol) was dried by coevaporation with dry pyridine. Pyridine (3ml) was added, followed by trimethylsilyl chloride (0.35 ml, 2.74 mmol). After 15 min at room temperature, the stirred solution was treated with benzoyl chloride (0.33 ml, 2.84 mmol). The solution was continuously stirred for an additional 3 h. The mixture was then cooled to 0 °C and water (5 ml) was added. After 5 min, 29% aqueous ammonia was added and the mixture was stirred at room temperature for 15 min. Solvents were removed under reduced pressure and the residue was dissolved in dichloromethane and washed with saturated aqueous sodium bicarbonate and brine. The organic phase was dried and filtered. Solvents were removed under reduced pressure and the product was isolated by silica gel chromatography to give 133 mg of **1d** (57.8%); ^{1}H NMR (DMSO-d_6) δ 8.26 (s, 1H, H6) 7.47-8.16 (m, 5H, aromatic), 6.16 (t, 1H, H1', *J*= 6.0 Hz), 5.28-5.29 (d, 1H, 3'OH), 5.08 (t, 1H, 5'OH), 4.45 (s, 2H, 5-CH_2O-), 4.23-4.29 (m, 1H, H3'), 3.85 (m, 1H, H4'), 3.60-3.85 (m, 4H, H5', H5'', CH_2CH_2CN), 2.74 (t, 2H, CH_2CH_2CN), 2.12-2.26 (m, 2H, H2', H2''). MS (ESI): $[M + H]^+$ calcd *m/z* 415.2, found 415.1.

Synthesis of N^4-benzoyl-5-(2-cyanoethoxymethyl)-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine-3'-*O*-(2-cyanoethyl)-*N*,*N*-diisopropylphosphororamidite (1f, Scheme 3-1). (27). Compound **1d** (40 mg, 0.097 mmol) was converted to the protected 5'-dimethoxytrityl

compound by standard methods. Dimethoxytrityl group was installed on to the 5' hydroxyl group using 4,4'-dimethoxytrityl chloride (50 mg, 0.15 mmol) in pyridine. $AgNO_3$ (25 mg, 0.15 mmol) was added and the reaction was left overnight. The product was washed with brine, extracted with dichloromethane and purified by silica gel to yield 63 mg (90.5%). ^{1}H NMR (DMSO-d_6) δ 6.82-8.85 (aromatic and H6), 6.15 (t, H1', J=6 Hz), 5.36-5.37 (d, 3'OH), 4.22 -4.29 (m, H3', 5-CH_2O-), 3.94 (m, H4', CH_2CH_2CN), 3.41 (t, CH_2CH_2CN), 3.23-3.25 (m, H5', H5''), 2.26-2.30 (m, H2', H2''). MS (ESI): $[M + H]^+$ *m/z* calcd 717.3, found 717.2.

The purified dimethoxytrityl derivative was converted to the 3'-phosphoramidite using 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite and diisopropylethylamine in dry acetonitrile and purified using flash column chromatography to yield **1f**. ^{31}P NMR δ 160.2 ppm. MS (ESI): $[M + Na]^+$ calcd *m/z* 939.4, found 939.1.

Synthesis of 5'-*O*-(4,4'-Dimethoxytrityl)-5-iodo-2'-deoxyuridine (2a, Scheme 3-2). (23) 5-Iodo-2'-deoxyuridine (5g, 14.1 mmol) was dissolved in 88 ml of pyridine, to which solution was added 5.74 g (16.9 mmol) of dimethoxytrityl chloride, DMAP (88.5 mg, 0.720 mmol), and TEA (3.55 ml, 24.1 mmol). The reaction was allowed to proceed at room temperature for about 4 hrs. Reaction was quenched by the addition of some methanol and the crude was purified column chromatography to yield 6.27 g (67.6%) of compound **2a**. ^{1}H NMR (DMSO-d_6) δ 8.00 (s, 1H, H6), 6.87-7.40 (m, 10H aromatic), 6.09 (t, 1H, H1'), 5.28-5.30 (d, 1H, 3'OH), 4.20-4.24 (m, 1H, H3'), 3.87-3.92 (d, 1H, H4'), 3.72 (s, 6H, -OCH_3), 3.15-3.18 (m, 2H, H5', H5''), 2.25 (m, 2H, H2', H2''). MS (ESI): $[M + Na]^+$ calcd *m/z* 679.1, found 678.9.

Synthesis of 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine (2b, Scheme 3-2). (23) Compound **2a** (6.27 g, 9.6 mmol) was dissolved in 30 ml of ACN. 2.28

ml *N*-methylimidazole, 1.58 g *tert*-butyldimethylsilyl chloride, and 4.8 g were then added. The solution was allowed to react at room temperature for 1 h. The solvents were evaporated under reduced pressure and the crude was purified by silica gel chromatography to yield 5.76 g of compound **2b** (78.2%). ^{1}H NMR ($CDCl_3$) δ 11.73 (s, 1H, N-H), 8.06 (s, 1H, H6), 6.87-7.41 (m, 13H, Ar-H), 6.04-6.08 (t, 1H, H1'), 4.32-4.39 (m, 1H, H3'), 3.75 (m, 1H, H4'), 3.73 (s, 6H, -OCH_3), 3.12 (m, 2H, H5'), 2.07-2.38 (m, 2H, H2'), 0.78-0.84 (m, 9H, $C(CH_3)_3$), 0.07-0.02 (m, 6H, $SiCH_3$) MS (ESI): $[M + Na]^+$ *m/z* calcd 771.2 found, 792.9.

Synthesis of 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-5-vinyl-2'-deoxyuridine (2c, Scheme 3-2). (23) Compound **2b** (5.76 g, 7.47 mmol), $(Ph_3P)_2PdCl_2$ (525 mg, 0.74 mmol) and tributyl(vinyl)tin (2.84ml, 9.73mmol) were dissolved in DMF (46 ml) and stirred at 80 °C for 1.5 h. The reaction mixture was filtered through a Celite pad and the filtrate was taken in EtOAc. The organic layer was washed with diluted $NaHCO_3$, H_2O and brine, dried ($NaSO_4$) and evaporated under reduced pressure. The residue was purified by column chromatography to give 4.09 g of **2c** (81.6%). ^{1}H NMR ($CDCl_3$) δ 7.77 (s, 1H, H6), 7.28-7.44 (m, 9H, aromatic), 6.88-6.85 (m, 4H, aromatic), 6.36 (t, 1H, H1'), 5.75-5.88 (dd, 1H, H1'', J=17.6, 11.2 Hz), 4.99 (d, 1H, H2''(Z), J=17.6 Hz), 4.95 (d, 1H, H2''(E), J=11.2 Hz), 4.01-4.02 (m, 1H, H4'), 3.82 (s, 6H, CH_3O), 3.45 (m, 1H, H5'a), 3.27 (m, 1H, H5'b), 2.51 (m, 1H, H2'a), 2.31 (m, 1H, H2'b), 0.87 (s, 9H, $(CH_3)_3$C-Si), 0.02-0.05 (m, 6H, CH_3-Si). MS (ESI): $[M + Na]^+$ calcd *m/z* 693.3, found 693.1.

Synthesis of 3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-5-vinyl-2'-deoxycytidine (2d, Scheme 3-2). (23) Compound **2c** (4.09 g, 6.09 mmol) was dissolved in ACN to which Et_3N (2.42ml, 17.4mmol), DMAP (2.13g, 17.4mmol), and 2,4,6-triisopropylbenzenesulfonyl chloride (5.27g, 17.4 mmol) were added and stirred at room temperature for 28 h. The mixture was cooled

in an ice bath, concentrated NH_4OH (100 ml) was added, and then stirred at room temperature for 1 h. The mixture was concentrated under reduced pressure and extracted with EtOAc. The organic layer was washed with H_2O and brine, dried ($NaSO_4$) and evaporated under reduced pressure. The residue was purified by column chromatography (0-4% MeOH in $CHCl_3$) to give 418 mg of **2d** as a foam (10.3%). ^{1}H NMR ($CDCl_3$) δ 7.96 (m, 1H, H6), 7.28-7.40 (m, 9H, aromatic), 6.84-6.81 (m, 4H, aromatic), 6.31 (t, 1H, H1'), 5.90-5.99 (dd, 1H, H1''), 5.11-5.17 (d, 1H, H2''(Z), *J*=17.6 Hz), 4.93-4.46 (d, 1H, H2'(E), *J*=11.2 Hz), 4.39-4.46 (m, 1H, H3'), 3.97 (m, 1H, H4'), 3.80 (s, 6H, CH_3O), 3.47 (dd, 1H, H5'a), 3.20 (dd, 1H, H5'b), 2.47-2.59 (m, 1H, H2'a), 2.12-2.22 (m, 1H, H2'b), 0.81 (s, 9H, $(CH_3)_3$C-Si), 007-0.00 (m, 6H, CH_3-Si).

Synthesis of 3'-*O*-(*t*-Butyldimethylsilyl)-5-(1,2-dihydroxyethyl)-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine (2e, Scheme 3-2). (23) A solution of OsO_4 in *t*-BuOH (5mg/ml, 1.6 ml) was added to a solution of **2d** (418 mg, 0.624 mmol) and *N*-methylmorphorine *N*-oxide (146.4mg, 1.25 mmol) in acetone/H_2O/*t*-BuOH (4:1:1, 9ml) and the resulting mixture stirred at room temperature for 4 h. After the reaction was quenched with saturated aqueous Na_2SO_3 (15 ml) the mixture was extracted with EtOAc. The organic layer was washed with saturated aqueous Na_2SO_3 (30 ml) and brine, dried (Na_2SO_4) and evaporated under reduced pressure. The residue was purified by column chromatography to give **2e** as a foam (221mg, 50.3%). ^{1}H NMR ($CDCl_3$) δ 7.92 (s, 0.33H, H6, isomer A), 7.70 (s, 0.67H, H6 isomer B), 7.49-7.28 (m, 9H, Ar-H), 6.34-6.29 (t, 1H, 1H'), 4.52-4.39 (m, 1H, H-3'), 4.28-4.18 (m, 1H, C*H*OH, isomer B), 4.01-4.00 (m, 1H, H4'), 3.92-3.89 (m, 0.3H, C*H*OH, isomer A), 3.86-3.84 (m, 6H, CH_3O), 3.67-3.42 (m, 2H, CH_2OH), 3.40-3.22 (m, 2H, H5'), 2.54-2.42 (m, 1H, H2'a), 2.21-2.09 (m, 2H, H2'b), 0.80-0.93 (m, 9H, (CH3)3C-Si), 0.07-0.08 (m, 6H, CH_3-Si). MS (ESI): $[M + Na]^+$ calcd *m/z* 726.3, found 726.1.

Synthesis of 5-(1,2-Diacetoxyethyl)-N^4-acetyl-3'-*O*-(*t*-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine (2f, Scheme 3-2). (23) A mixture of **2e** (221 mg, 0.314 mmol), DMAP (4.2 mg, 0.034 mmol), Ac_2O (133.3 ul, 1.4 mmol) in pyridine (4 ml) was stirred at room temperature for 20 h. After EtOH was added the mixture was evaporated under reduced pressure. After successive co-evaporation with toluene the residue was dissolved in EtOAc and washed with H_2O, and brine, dried (Na_2SO_4) and evaporated under reduced pressure. The residue was purified by column chromatography to give **2f** as a foam (202 mg, 77.5%). MS (ESI): $[M + H]^+$ *m/z* calcd 830.4, found 829.9.

Synthesis of 5-(1,2-Diacetoxyethyl)-N^4-acetyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine (2g, Scheme 3-2). (23) A solution of TBAF in THF (1.0 M, 0.484 ml, 1.67 mmol) was added to a solution of **2f** in THF(3.4 ml) in an ice bath and stirred at room temperature for 6 h. The reaction mixture was taken in EtOAc (56 ml), washed with H_2O and brine, dried (Na_2SO_4) and evaporated under reduced pressure. The residue was purified by column chromatography to give **2g** as a foam (116 mg, 66.6%). ^{1}H NMR ($CDCl_3$) δ 7.41-7.28 (m, 9H, Ar-H), 6.82-6.86 (m, 4H, Ar-H), 6.16-6.26 (m, 1H, H1'), 5.59 (m, 1H, C*H*OAc), 4.00-4.45 (m, 4H, CH_2OAc, H3', H4'), 3.32-3.55 (m, 2H, H5'), 2.02 (m, 2H, H2'), 1.56 (s, 9H, CH_3CO). MS (ESI): $[M + Na]^+$ calcd *m/z* 738.3, found 738.1.

Synthesis of 3'-*O*-[(Cyanoethyl)-(*N*,*N*-diisopropylamino)phosphinyl]-5-(1,2-diacetoxyethyl)-N^4-acetyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine (2h, Scheme 3-2). (23) Cyanoethyldiisopropylchlorophosphoramidite was added to a solution of **2g** (116 mg, 0.16 mmol) followed by *N*,*N*-diisopropylethylamine in CH_2Cl_2 at 0 °C and stirred at room temperature for 0.5 h. The mixture was taken in $CHCl_3$ and washed with H_2O, and brine, dried with Na_2SO_4 and

evaporated under reduced pressure. The residue was quickly passed through a silica column to eliminate inorganic salts. ^{31}P NMR δ 150.0 ppm. MS (ESI): $[M + Na]^+$ calcd *m/z* 938.4, found 938.1.

References

1. Lister, R., Pelizzola, M., Dowen, R. H. and Hawkings, R. D. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 462: 315-322.

2. Heintz, N. and Kriaucionis, S. (2009) The Nuclear DNA Base 5-hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science*. 324: 929-930.

3. Tahiliani, M., Koh, K.P., Yinghua, S. et al. (2009) Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. Science. 324: 930-935.

4. Carell, T., Globish, D., Wagner, M, et al. (2010) Tissue distribution of hmdC and search for active demethylation intermediates. *PLoS ONE*. 5(12): e15367.

5. Carell, T., Hagemeier, C. and Deimi, C. A. (2011) The Discovery of 5-Formylcytosine in Embryonic Stem Cell DNA. *Angew. Chem*. 50(31): 7008-7012.

6. Carell, T., Globisch, D. and Münzel, M. (2011) 5-Hydroxymethylcytosine, the Sixth Base of the Gemone. *Angew. Chem. Int. Ed*. 50(29): 6460-6468.

7. Leonhardt, H. and Spada, F. (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res*. 38(19): e181.

8. Sowers, L. C., Zhang, Y., Hong, K., Taranova, O. V., D'Alessio, A. C. and Ito, S. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 466: 1129-1135.

9. Turner, S. W., Korlach, J. and Clark, T. A. (2010) Direct detection of DNA methylation during single molecule, real-time sequencing. *Nature Methods*. 7(6): 461-467.

10. Shock, L. S., Thakkar, P. V., Peterson, E. J., Moran, R. G. and Taylor, S. M. (2011) DNA methyltransferase 1, cytosine methylation, and cytosine hydroxymethylation in mammalian mitochondria. *Proc. Natl. Acad. Sci USA*. 108(9): 3630-3635.

11. Koh K. P. and Rao, A. (2011) Tet1 Tet2 Regulates 5-hydroxymethylcytosine Production and Cell Lineage Specification in Mouse Embryonic Stem Cells. *Cell Stem Cell*. 8: 200-231.

12. Song, C., Szullwach, K. E., Fu, Y., Dai, Q. and Yi, C. (2011) Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol*. 29(1): 68-75.

13. Robertson, A. B., Dahl, J. A. and Vågbo, C. B. (2011) A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res*. 39(8): e55.

14. Kriukiene et al. (2011) Methyltransferase-DirectedDerivatization of 5-hydroxymethylcytosine in DNA. *Angew. Chem. Int. Ed.* 123(9): 2138-2141.

15. Spada, F., Leondardt, H., Bultmann, S., Schmidt, C. S., Brachmann, A. and Szwagierczak, A. (2011) Characterization of PvuRts1I endonuclease as a tool to investigate genomic 5-hydroxymethylcytosine. *Nucleic Acids Res*. 39(12): 5149-5156.

16. Wanunu, M., Karni-Cohen, D., Johnson, R. R. and Fields, L. (2011) Discrimination of Methylcytosine from Hydroxymethylcytosine in DNA Molecules. *J. Am. Chem. Soc*. 133: 487-492.

17. Ficz, G., Branco, M. R., Seisenberger, S., Santos, F., Krueger, F., Hore, T. A. and Marques, C. J. (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*. 473: 398-404.

18. Brown, T., Schofield, C. J., Hansen, A. S. and Thalhammer, A. (2011) Hydroxylation of methylated CpG dinucleotides reverses stabilization of DNA duplexes by cytosine 5-methylation. *Chem. Commun*. 47: 5325-5327.

19. Pastor, W. A., Pape, U. J., Huang, Y., Henderson, H. R. and Lister, R. (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*. 473: 394-397.

20. Zhang, Y., Sun, Y. E., Zhao, K., Cui, K., Wang, Zhibin, Xia, K., Ito, S., D'Alessio, A. C. and Wu, H. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature*. 473: 389-394.

21. Smith et al. (2010) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*. 473: 394-397.

22. Kasai, H., Matsuda, A., Kaji, H., Ueno, Y., Karino, N., Kamiya, H. and Kamiya-Murata, N. (1999) Formation of 5-formyl-2'-deoxycytidine from 5-methyl-2'-deoxycytidine from 5-methyl-2'-deoxycytidine in duplex DNA by Fenton-type reactions and γ-irradiation. *Nucleic Acids Res*. 27(22): 4385-4390.

23. Matsuda, A., Ueno, Y. and Karino, N. (2001) Synthesis and properties of oligonucleotides containing 5-formyl-2'-deoxycytidine: in vitro DNA polymerase reactions on DNA templates containing 5-formyl-2'-deoxycytidine. *Nucleic Acids Res*. 29(12): 2456-2463.

24. Kamiya, H., Tsuchiya, H., Karino, N., Ueno, Y., Matsuda, A. and Harashima, H. (2002) Mutagenicity of 5-Formylcytosine, an Oxidation Product of Methylcytosine, in DNA in Mammalian Cells. *J. Biochem*. 132: 551-555.

25. Drohat, A. C. and Maiti, A. (2011) Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxycytosine. *J. Biol. Chem*. 286(41): 35334-35338.

26. Dai, Q. and He, C. (2011) Syntheses of 5-Forymyl- and 5-Carboxyl-dC Containing DNA Oligos as Potential Oxidation Products of 5-Hydroxymethylcytosine in DNA. *Org. Lett*. 13(13): 3446-3449.

27. Sowers, L. C., Lin, S. S., Fujimoto, J. and Tardy-Planechaud, S. (1997) Solid phase synthesis and restriction endonuclease cleavage of oligodeoxynucleotides containing 5-(hydroxymethyl)-cytosine. *Nucleic Acids Res*. 25(3): 553-558.

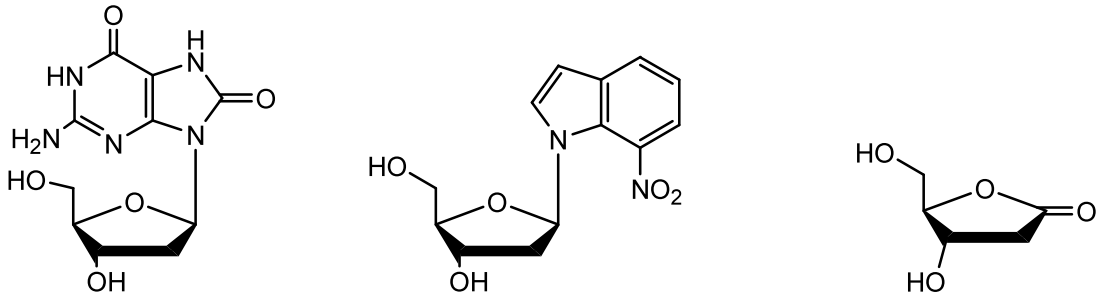
28. Wang, Z., Wan, K.X., Ramanathan, R., Taylor, J.S. and Gross, M.L. (1998) Structure and fragmentation mechanisms of isomeric T-rich oligodeoxynucleotides: a comparison of four tandem mass spectrometric methods. *J. Am. Soc. Mass Spectrum*. 9(7): 683-691.

CHAPTER 4 Dual Incorporation of 8-oxo-7,8-dihydro-2'-deoxyguanosine and 7-Nitroindole into Oligodeoxynucleotides

4.1 Introduction

8-Oxo-7,8-dihydro-2'-deoxyguanosine (8-oxo-dG) is well known and studied oxidatively induced DNA lesion found in genomic DNA. It can be produced from both exogenous and endogenous sources of reactive oxygen species (1,2). It is even used as a marker for genomic DNA damage as it can easily be detected in urine of rodents and man. 8-oxo-dG in template DNA promotes mutagenic insertions, such as dAMP or dCMP, directly opposite the lesion (4,6). It also promotes mutations *in vivo* (2,3,7).

2-Deoxyribonolactone (dL) can be derived from the precursor 7-nitroindole d(7-Ni) via irradiation (11). Usually, an *in vitro* study first incorporates d(7-Ni) into a synthetic oligodeoxyribonucleotide (ODN) strand, which is then irradiated to form the oxidized lesion dL within the synthetic strand. 2-Deoxyribonolactone can undergo both β and δ eliminations to form strand breaks in DNA (11, 14). Like 8-oxo-dG, dL is also mutagenic; and repair studies of the lesion by itself and as part of a tandem lesion have also been conducted (11, 12, 14). We report here the syntheses of 8-oxo-dG and 7-nitroindole nucleoside and their contiguous incorporation into an ODN.



8-oxo-dG **d(7-Ni)** **dL**

Figure 4-1 Structures of lesions 8-oxo-dG and 2'-deoxyribonolactone (dL), and 7-nitroindole nucleoside

4.2. Mutagenicity of 8-oxo-dG in *in vitro*

One of the first *in vitro* studies for 8-oxo-dG showed that incorporation of dCTP is slightly favored over dATP opposite the lesion. Extension beyond the mutagenic pairs was seen with both pol α and the Klenow fragment (1). Using human DNA polymerase, pol β, incorporation of dAMP directly opposite 8-oxo-dG was slightly favored over dCMP, depending on the downstream sequence context beyond 8-oxo-dG. In addition, mismatch incorporations were seen upstream and downstream adjacent to the 8-oxo-dG lesion. However, when pol α replaced pol β no upstream errors were detected. Error rates at sites adjacent to 8-oxo-dG were roughly 1% of the values opposite 8-oxo-dG, indicating the potential of 8-oxo-dG to induce tandem mutations when pol β is involved (4). A different study showed that 8-oxo-dGTP was incorporated opposite dC with a specificity constant of 0.005 $\mu M^{-1}s^{-1}$ and was extended 96% of the time. The specificity constant for incorporation of 8-oxo-dGTP opposite dA was 0.2 $\mu M^{-1}s^{-1}$, 13-fold higher than incorporation opposite dC, with extension 70% of the time (6).

4.2.1. Mutagenicity of 8-oxo-dG in *in vivo*

An *in vivo* study reported that 8-oxodG could induce a G-to-T transversion at frequencies of 0.7% and 0.6% depending on the type of assay used (2). Thus, even though the oxidized lesion induces nucleotide misincorporation *in vitro*, it is not potently mutagenic *in vivo*. A different *in vivo* study also yielded predominant G-to-T mutations (>78%) but with mutation frequencies varying from 2.5 to 4.8% in COS cells and 1.8% in *E. coli* cells (3). An *in vitro* study using *in vivo* lesion levels found in rodents, discovered that as little as 0.06% of the oxidized lesion relative to dGTP is enough to cause A-to-C mutations (7). Pol γ was used in this last study.

4.3 7-Nitroindole is a precursor for 2'-deoxyribonolactone

2'-Deoxyribonolactone (dL) is an oxidized lesion that has never been detected in DNA. Successful incorporation into ODNs has been achieved by incorporating its precursor 7-nitroindole d(7-Ni) followed by irradiation and purification. dL is quite alkali-labile so conditions must be appropriate for its generation and purification. It can form a tandem lesion with thymidine glycol (Tg) and repair mechanisms have been studied for this tandem lesion (14). When d(7-Ni) is irradiated it forms a radical at the NO_2 group. Molecular studies indicate that H1' is 2.27 Å away from the nitro group. X-ray structure gave a distance of 2.29 Å. Thus, a radical is formed that abstracts the H1' proton yielding a C1' radical on the sugar moiety. Through several intermediates the radical yields an oxidized lesion, 2'-deoxyribonolactone (dL) (11).

4.3.1. Mutation spectrum and mutation frequencies for 7-nitroindole and 2'-deoxyribonolactone

Incorporation of dATP opposite 7-Ni, using an exonuclease-deficient Klenow fragment, gave a specificity constant of 0.077 $\mu M^{-1}min^{-1}$, 9-fold higher than for incorporation of dATP opposite T. An *in vivo* study gave mutation spectra induced by d(7-Ni) and dL. For dL, TMP was the most frequently incorporated nucleotide (55%) followed by dAMP (18%) and dGMP (12%). For dNi, TMP (42%) was the most frequent nucleotide mismatched followed by dAMP (40%). However, the mutation frequency for the site-specifically incorporated dL nucleoside in *WT E. coli*, were (5.5-5.7) x 10^{-5}. d(7-Ni) had a mutation frequency of (14.0-2.2) x 10^{-5} in *WT E.*

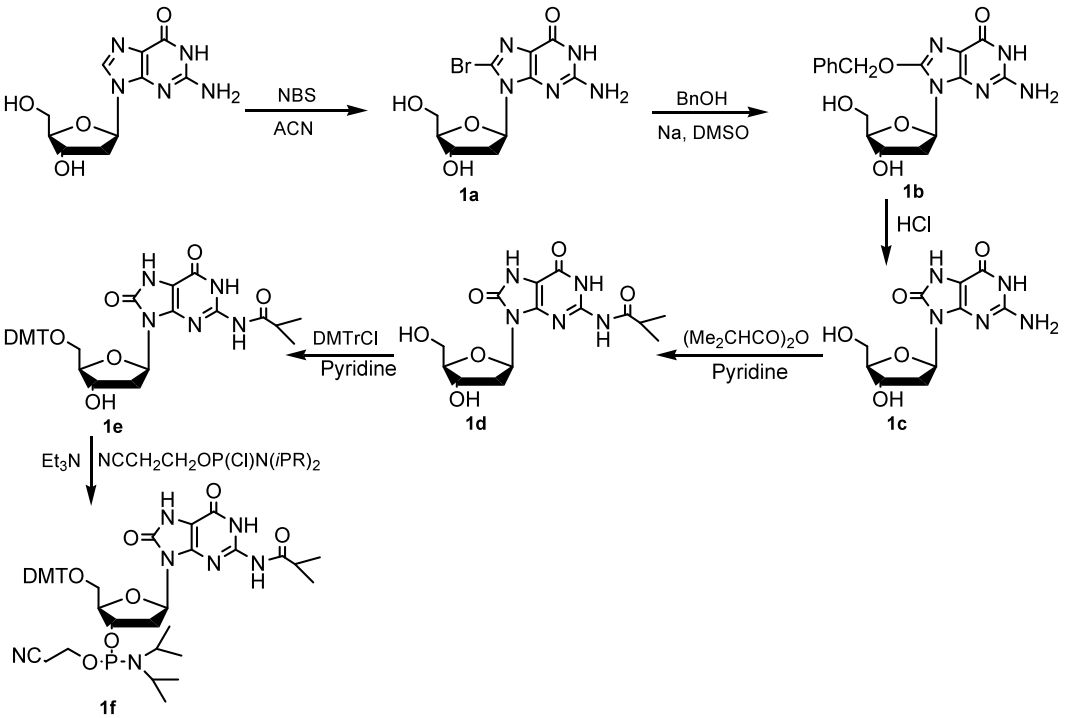
coli. Thus, these two lesions are not mutagenic as indicated by their respective mutation frequencies (12).

4.4 The intent behind the formation of 8-oxo-dG –dL tandem lesion in ODNs

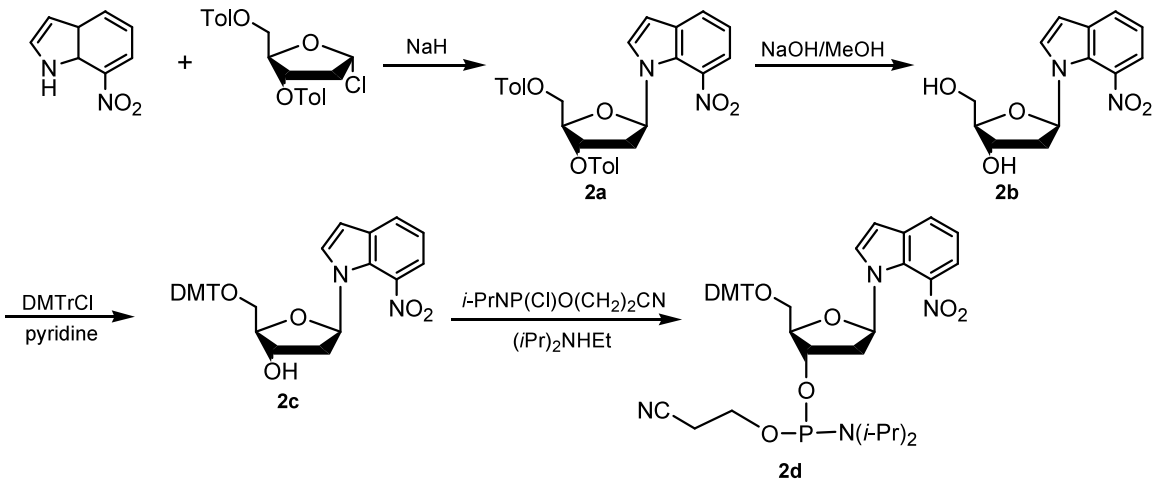
The main focus of this project was to successfully synthesize 8-oxo-dG and 7-nitroindole d(7-Ni) using published methods and to incorporate both lesions side-by-side in a docecameric ODN. Exposure to UV light converts d(7-Ni) to dL, rendering the tandem dL-8-oxo-dG-containing ODN substrate. With the use of this tandem lesion as standard, we can then identify and quantify its formation in calf thymus DNA treated with Fenton reagents.

4.5 Results and Discussion

Several published methods were used to synthesize 8-oxo-dG (15-17) and Scheme 3-1 shows the overall synthetic scheme for 8-oxo-dG-phosphoramidite. All structures were confirmed with ^{1}H NMR and ESI-MS, see experimental procedures and appendix C. 7-Nitroindole nucleoside was synthesized according to published procedures (11, 18, 19). Scheme 4-2 shows the overall synthetic procedure used for the syntheses of 7-nitroindole nucleoside and its phosphoramidite derivative. All structures were confirmed with ^{1}H NMR and ESI-MS, see experimental procedures and appendix C. Both lesions were successfully incorporated into an ODN.



Scheme 4-1 Synthetic scheme for 7,8-Dihydro-5'-*O*-(4,4'-dimethoxytrityl)-N^2-isobutyryl-8-oxo-2'-deoxyguanosine phosphoramidite



Scheme 4-2 Synthetic scheme for 7-nitroindole-phosphoramidite

Lesion dL was incorporated into a 12mer, d(ATGGCLGCGTAT), and characterized by ESI-MS and ESI-MS/MS (Figures 4-2 and 4-3). The deconvuled mass of the ODN was 3565 Da, which was 119 Da lower than the calculated mass of the unmodified d(ATGGCAGCGTAT) (A replaced L), which clearly indicates the presence of the oxidized abasic site. The product-ion spectrum of the $[M - 3H]^{3-}$ ion (m/z 1187.7, Figure 4-3) of the ODN showed the formation of w_3, w_4, w_5, w_6, w_7^{2-}, $w_9^{2-,}$ and w_{11}^{3-}, and $[a_n\text{-Base}]$ ions, that is $[a_3\text{-G}]$, $[a_4\text{-G}]$, $[a_5\text{-C}]$, $[a_6\text{-L}]^{2-,}$ $[a_7\text{-G}]^{2-,}$ $[a_8\text{-G}]^{2-}$ and $[a_9\text{-G}]^{2-}$ ions. The measured masses of the w_3, w_4, w_5, and w_6 ions are the same as the calculated masses for the corrsponding ions of the unmodified ODN, whereas the w_7^{2-}, w_9^{2-}, and w_{11}^{2-} ions exhibited 119 Da lower in mass than the corresponding ions formed from the unmodified d(ATGGCAGCGTAT). In addition, the masses for ions $[a_6\text{-L}]^{2-}$, $[a_7\text{-G}]^{2-}$, $[a_8\text{-C}]^{2-}$ and $[a_9\text{-G}]^{2-}$ ion are 119 Da lower than the calculated masses of the corresponding fragments from the unmodified d(ATGGCAGCGTAT). These results establish the presence of dL at the seventh position, from the 3' end, in the ODN. Data for the tandem lesion dL-8-oxo-dG in not available.

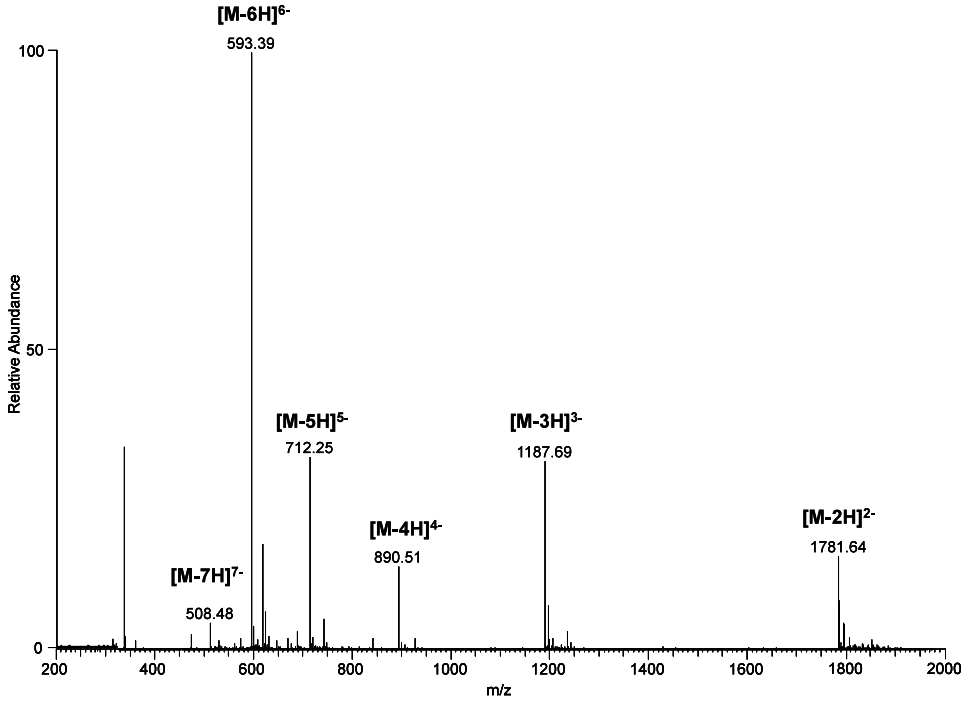


Figure 4-2 Negative-ion ESI-MS of d(ATGGCLGCGTAT)

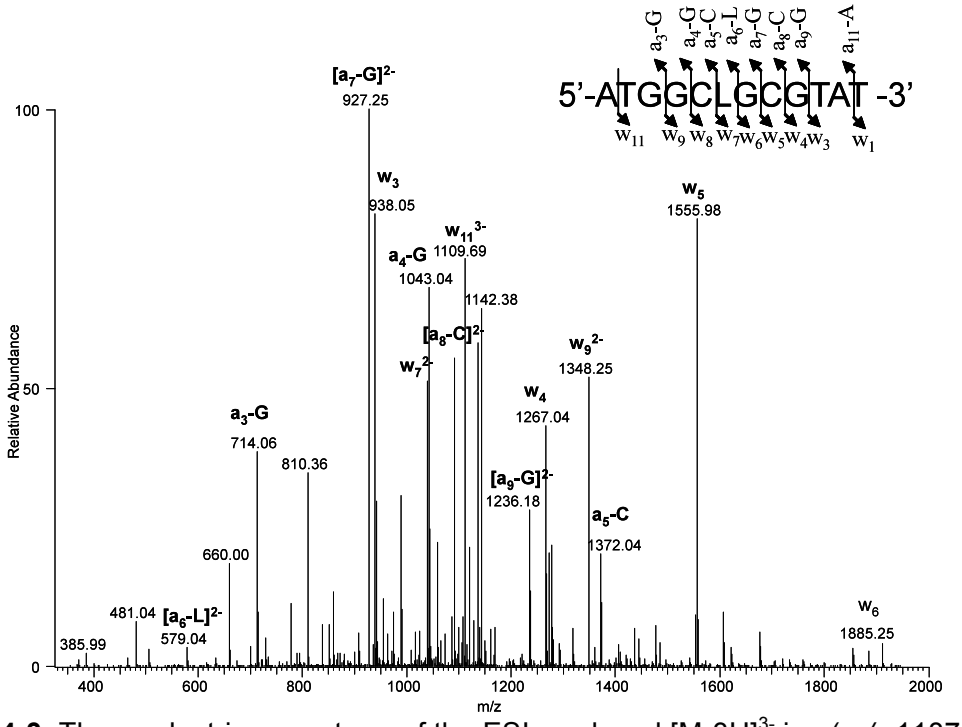


Figure 4-3 The product-ion spectrum of the ESI-produced $[M-3H]^{3-}$ ion (*m/z* 1187.7) of d(ATGGCLGCGTAT). Shown in the inset is a scheme summarizing the observed fragment ions

The tamdem lesion, 5'-dL-(8-oxo-dG)-3' and lesion, 5'-dL-dG-3', where the 2'-deoxyribose lactone is located on the 5' side of an unmodified guanine, can be induced in calf-thymus DNA when treated with Fenton reagent (Figures 4-4 and 4-5). Quantification of the results revealed a significant increase in the formation of both lesions with exposure to elevating concentrations of Fenton reagent (Figure 4-6). We observed that the yield of the tandem lesion, 5'-dL-(8-oxodG)-3', was merely a few folds (2~5 folds) lower than what was found for the dL that is situated on the 5' side of an unmodified dG at the high dose of treatment range (Figure 6). Our previous results had revealed that Fenton type reagents can dramatically induce the formation of another type of tandem lesion: 5'-Tg-(8-oxodG)-3'. In that study, we also observed that copper ions, especially Cu(I), can stimulate the γ ray-induced formation of the 5'-Tg-(8-oxodG)-3' tandem lesion. The amount of the 5'-dL-(8-oxodG)-3' formed in the presence of 200 μM of Cu(II) and 16 mM of ascorbate is approximately 5 times more than 5'-Tg-(8-oxodG)-3' induced under the same conditions (21).

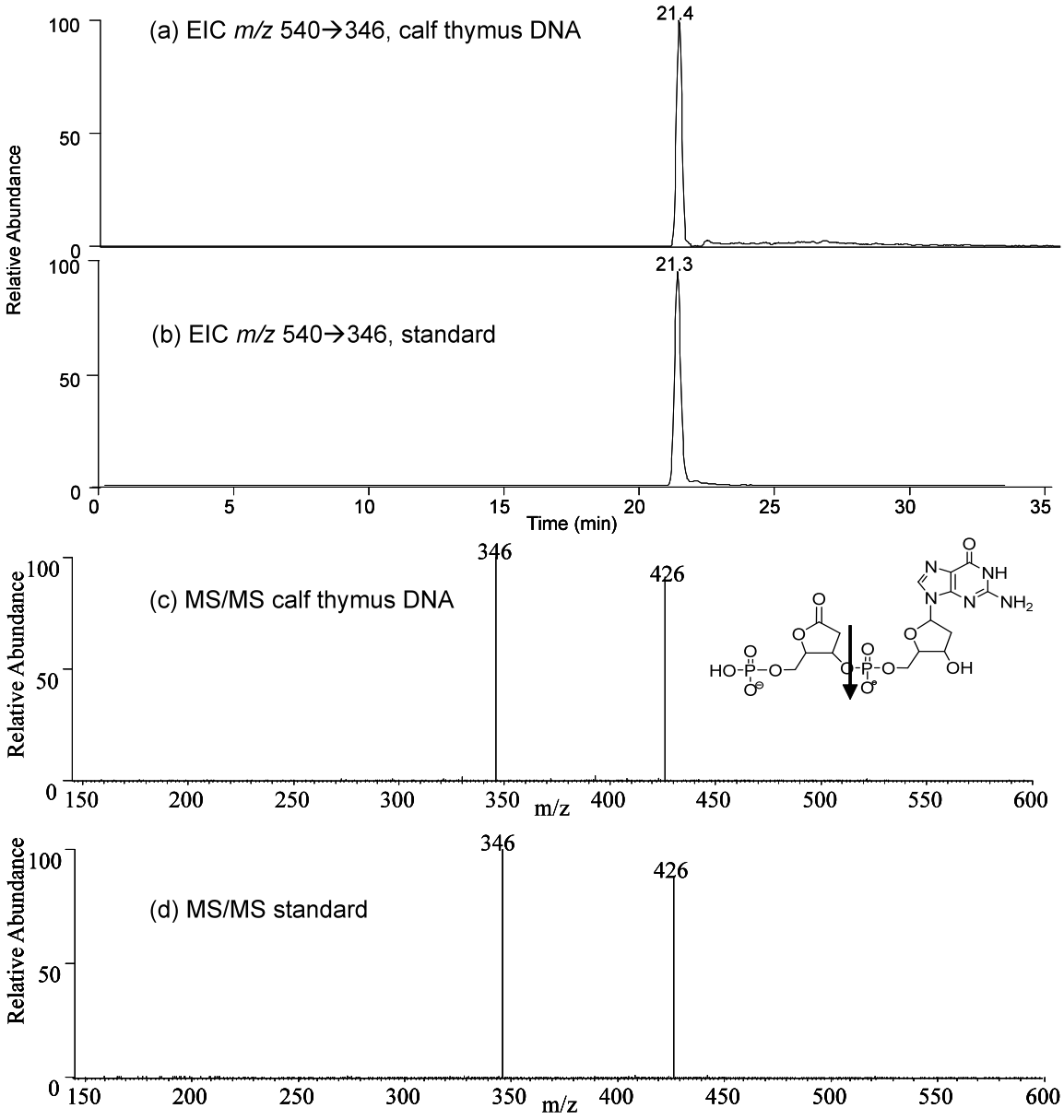


Figure 4-4 LC-MS/MS identification of 5'-dinucleotide formation in calf thymus DNA treated with Cu(II)/H_2O_2/ascorbate. Extracted-ion chromatograms (EICs) for monitoring the *m/z* 540 → 346 transition for pdL-pdG in Fenton reagent-treated calf thymus DNA (a) and calf thymus DNA doped with authentic lesion-containing ODN (b). Shown in (c) and (d) are the MS/MS averaged from the 21.4 min peak in (a) and the 21.3 min peak in (b)

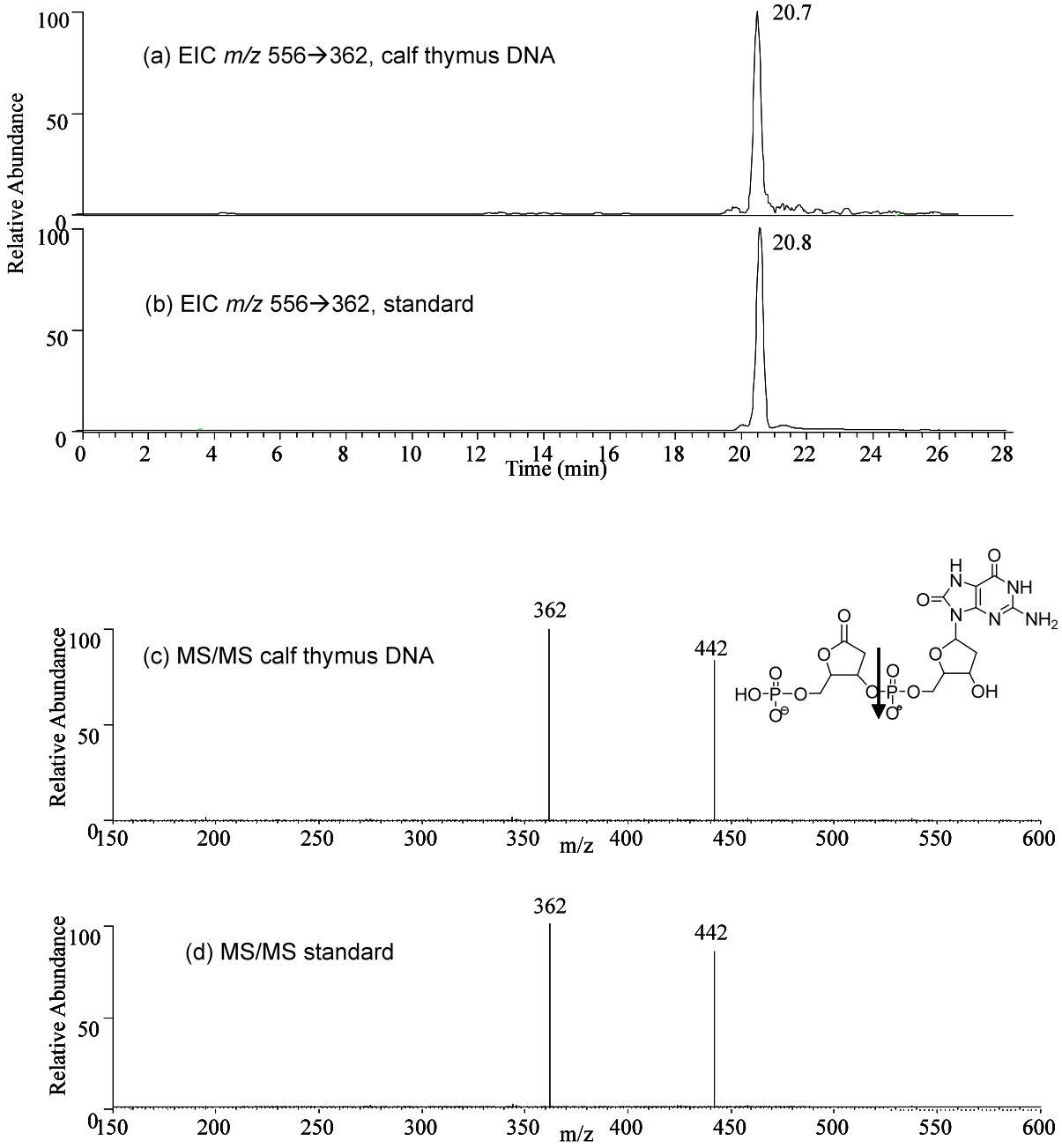


Figure 4-5 LC-MS/MS identification of 5'-dL-(8-oxodG)-3' tandem lesion formation in calf thymus DNA treated with Cu(II)/H_2O_2/ascorbate. Extracted-ion chromatograms (EICs) for monitoring the m/z → 362 transitions for pdL-p(8-oxodG) in Fenton reagent-treated calf thymus DNA (under reaction conditions described in Table 4-1, a) and calf thymus DNA doped with authentic tandem lesion-containing ODN (b). Shown in (c) and (d) are the MS/MS averaged from the 20.7-min peak in (a) and the 20.8-min peak in (b)

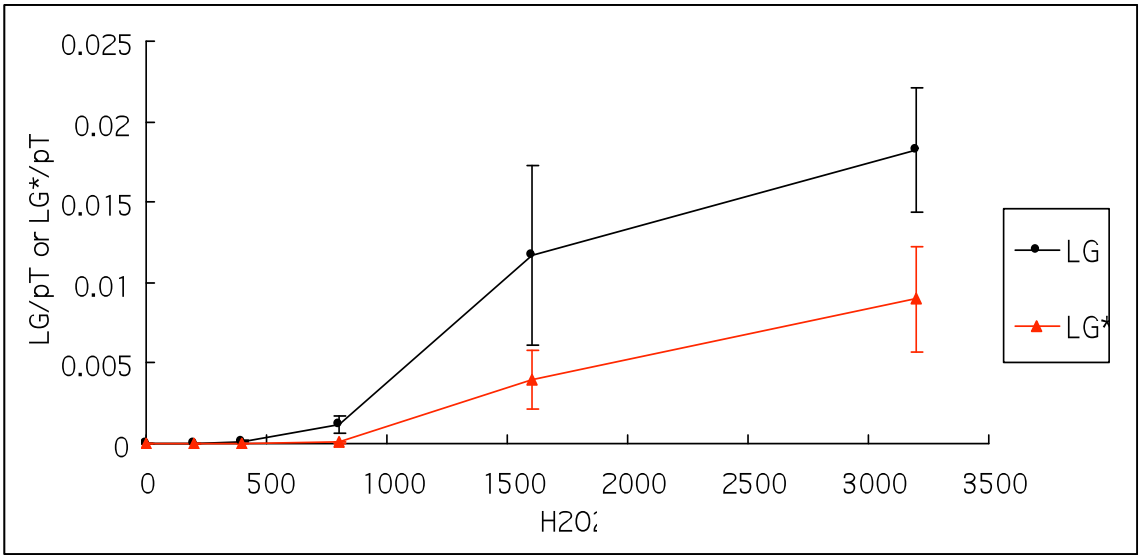


Figure 4-6 Quantification of the formation of the 5'-dL-3' and 5'-dL-(8-oxo-dG)-3' lesions in calf thymus DNA exposed with Cu(II) and ascorbate along with H_2O_2 or γ-rays. (a) induction of the 5'-dL-dG-3' and 5'-dL-(8-oxo-dG)-3' lesions in calf thymus DNA by Cu(II)/H_2O_2/ascorbate. The data represent the means ± S.D. of results from three independent treatments and LC-MS/MS quantification experiments

4.6 Experimental Procedures

Materials. All chemicals, unless otherwise specified, were obtained from Sigma-Aldrich (St. Louis, MO). All anhydrous solvents, unless otherwise specified, were purchased from Acros Organics. 1-chloro-2'-deoxy-3',5'-di-*O*-toluoyl-D-ribofuranose was purchased from Carbosynth. Common reagents for DNA synthesis were obtained from Glen Research. Copper (II) chloride, L-methionine, L-ascorbic acid and calf thymus DNA were from Sigma-Aldrich (St. Louis, MO). Hydrogen peroxide (30%) and nuclease P1 were purchased from Fisher Scientific (Fair Lawn, NJ) and MP Biomedicals (Aurora, OH), respectively. Unmodified ODNs used in this study were purchased from Integrated DNA Technologies (Coralville, IA). 1,1,1,3,3,3-hexafluoro-2-propanol (HFIP) was purchased from TCI America (Portland, OR). Shrimp alkaline phosphatase was obtained from USB Corporation (Cleveland, OH); all other enzymes were from New England Biolabs (Ipswich, MA).

ODN Synthesis and Purification. ODN (12mer) was synthesized on a Beckman Oligo 1000S DNA synthesizer (Fullerton, CA) at 1-μmol scale. The synthesized phosphoramidite building block was dissolved in anhydrous acetonitrile at a concentration of 0.07 M. Standard phosphoramidite building blocks (Glen Research Inc., Sterling, VA) of dA, dC, dG, and dT were employed and standard ODN assembly protocol was used without any modification. The ODN linked to the solid support was treated with concentrated NH_4OH at RT for 3 days. The purification of the ODN was performed on a Beckman HPLC System (32 Karat software version 3.0, pump module 125) with a UV detector (module 126) monitoring at 260 nm. A C-18 column was used and a linear gradient of ACN from 0 to 35% in 0.1 M TEAA buffer, pH 7.0, was used for the mobile phase.

HPLC. A 4.6 mm X 250 mm Apollo C18 column (5 um in particle size, 300 Å in pore size, Alltech Associate Inc., Deerfield, IL) was used for the separation of synthetic ODNs. A solution of 50 mM triethylammonium acetate (TEAA, solution A) and a mixture of 50 mM TEAA and acetonitrile (70/30, v/v, solution B) were used as mobile phases. The flow rate was 0.8 mL/min, and a gradient of 5 min 0-20% B, 45 min 20-50% B, and 5 min 50-100% B was employed.

Mass Spectrometry (MS). Electrospray ionization-mass spectrometry (ESI-MS) and tandem MS (MS/MS) experiments were carried out on an LCQ Deca XP ion-trap mass spectrometer (Thermo Fishcher Scientific, San Jose, CA). A mixture of acetonitrile and water (50:50, v/v) was used as solvent for electrospray. The spray voltage was 3.0 kV, and the temperature of the heated capillary was maintained at 275 °C. LC-MS/MS quantification was performed by using an Agilent 1100 capillary HPLC pump (Agilent Technologies, Santa Clara, CA) interfaced with an LTQ linear ion-trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). A 0.5×250 mm Zorbax SB-C18 column (5 μm in particle size, Agilent Technologies) was used for the separation of the DNA hydrolysates, and the flow rate was 7.0 μL/min. A 10-min gradient of 0-20% methanol in 400 mM HFIP (pH was adjusted to 7.0 by addition of triethylamine), followed by a 30-min gradient of 20-50% methanol in 400 mM HFIP, was used for the separation. The mass spectrometer was set up for monitoring the fragmentation of the $[M-H]^-$ ions of 2'-deoxyadenosine-5'-phosphate (pdA) and dinucleotides pdL-p(8-oxodG) and pdL-pdG. The capillary temperature for the electrospray source of the mass spectrometer was maintained at 300 °C to minimize the formation of the HFIP adducts of nucleotides.

Treatment of calf thymus DNA. Calf thymus DNA was desalted by ethanol precipitation. The DNA pellet was redissolved in a solution containing 100 mM NaCl, 1 mM EDTA and 10 mM Tris-HCl (pH 7.5), and the DNA was annealed by heating the solution to 90 °C and cooling slowly to room temperature. Aliquots of DNA (150 μg) were incubated with $CuCl_2$ (25-400 μM), H_2O_2 (0.2-3.2 mM), and ascorbate (2-32 mM) in a 0.5-mL solution at room temperature for 50 min (The concentrations of individual Fenton reagents are shown in Table 4-1). In this respect, chemicals used in Fenton-type reaction were freshly dissolved in doubly distilled water and the reactions were carried out under aerobic conditions. The reactions were terminated by adding an excess amount of L-methionine, and the resulting DNA samples were desalted by ethanol precipitation and quantified by measuring the UV absorbance at 260 nm.

Table 4-1. Concentrations of Fenton-type reagents employed for the treatment of calf thymus DNA[a]

	Control	#1	#2	#3	#4	#5
$CuCl_2$ (μM)	0	25.0	50.0	100	200	400
H_2O_2 (μM)	0	200	400	800	1600	3200
Ascorbate (mM)	0	2.0	4.0	8.0	16	32

a, All reactions were carried out in a 500 μL solution containing 150 μg of calf thymus DNA.

Enzymatic digestion of calf thymus DNA. To the above treated DNA (20 μg) were added 2 units of nuclease P1 and a 30-μL buffer solution containing 300 mM sodium acetate (pH 5.0) and 10 mM zinc acetate. The digestion was continued at 37 °C for 4 h, and the enzyme in the resulting digestion mixtures was removed by chloroform extraction. The amount of nucleosides in the mixture was quantified by UV absorbance measurements, and aliquots of the nucleoside mixture were subjected directly to C-MS/MS analysis.

NMR. ^{1}H NMR spectra were recorded at 300 Mhz on a Varian Inova 300 NMR spectrometer (Varian Inc., Palo Alto, CA). Two dimensional NMR spectra were recorded in D_2O using a Variant Unity (Varian, Inc., Palo Alto, CA) spectrometer operating at 500 MHz at 25 °C.

8-Bromo-2'-deoxyguanosine (1a, Scheme 4-1). (20) To a suspension of 2'-deoxyguanosine (2.0 g, 7.5 mmol) in a mixture of acetonitrile (80 ml) and water (20 ml) was added *N*-bromosuccinimide (2.0 g, 11.2 mmol) in three portions. The reaction mixture was stirred for 30 min at room temperature. The solvents were removed and the precipitate suspended in acetone (40 ml) stirred for 2 h at room temperature and cooled overnight at -20 °C. The precipitate was collected by filtration, extensively washed with cold acetone and dried under vacuum to give 1.35 g (73.4%) of a slightly yellow powder. ^{1}H NMR (DMSO-d_6) δ 10.78 (s, 1H, N1-H), 6.47 (s, 2H, N2-H2), 6.16 (dd, *J*= 7.3, 7.3, 1H, C1-H), 4.40 (ddd, *J*= 3.0, 3.1, 6.0, 1H, 3'H), 3.80 (ddd, *J* = 3.1, 5.4, 5.5, 1H, H4'), 3.63 (dd, *J* = 5.3, 11.6, 1H, H5'), 3.16 (ddd, J = 6.6, 7.6, 13.7, 1H, H2'), 2.10 (ddd, *J* = 3.0, 6.5, 13.3, 1H, H2'). MS (ESI): $[M+ Na]^+$ calcd *m/z* 368.0, found 367.9.

Synthesis of 8-benzyloxy-2'-deoxyguanosine (1b, Scheme 4-1). (16) BnOH (5.2 ml) was reacted with 1.8 g of Na metal and stirred at room temperature for 5 h under an argon atmosphere. To this solution was added 500 mg of 8-bromo-2'-deoxyguanosine and 3.4 ml of DMSO. The

reaction was stirred at 60 °C overnight. After the reaction was complete, the solution was poured into ether to precipitate the desired compound. The precipitate was filtered and purified by silica gel column chromatography with 15-20% MeOH in $CHCl_3$ to yield 312 mg of **1b** (57.9%). ^{1}H NMR (DMSO-d_6) δ 7.31-7.95 (m, 5H, Ar-H), 6.35 (s, 2H, NH2), 6.06-6.10 (t, 1H, H1'), 5.40 (s, 2H, -CH2-Ph), 5.12 (m, 1H, 3'OH), 4.72-4.83 (m, 1H, 5'OH), 4.20-4.26 (m, 1H, H3'), 3.68-3.74 (m, 1H, H4'), 3.33-3.35 (m, 2H, 5H'), 2.78-2.90 (m, 1H, H2'), 1.96-2.06 (m, 1H, 2H''). MS (ESI): $[M + H]^+$ calcd *m/z* 374.1, found 374.1.

Synthesis of 8-oxo-2'-deoxyguanosine (1c, Scheme 4-1). (16) 8-Benzyloxy-2'-deoxyguanosine (412 mg) was dissolved in 52 ml of MeOH to which was added 5.5 ml of 1.0 M HCl. The reaction was stirred at room temperature overnight. The crude was triturated with 50/50 MeOH/Ether to eliminate most impurities. This gave 302 mg of 8-oxo-dG (96.6%). ^{1}H NMR (DMSO-d_6) δ 6.44 (s, 2H, NH2), 6.00-6.05 (t, 1H, H1'), 5.10-5.12 (d, 1H, 3'OH), 4.78-4.81 (m, 1H, 5'OH), 4.29-4.35 (bs, 1H, 8-OH), 4.07 (m, 1H, H3'), 3.71 (m, 1H, H4'), 3.38-3.54 (m, 2H, H5'), 3.16-3.18 (m, 1H, H2'), 1.87-1.99 (m, 1H, H2'').

Synthesis of 7,8-Dihydro-N^2-isobutyryl-8-oxo-2'-deoxyguanosine (1d, Scheme 4-1). 8-oxo-dG (45 mg, 0.159 mmol) was co-evaporated in anhydrous pyridine and then suspended in 1.6 ml of anhydrous pyridine. Trimethylchlorosilane (86.3 mg, 0.794 mmol) was added slowly to the suspension cooled in an ice-bath. After 30 min, isobutyric anhydride (0.13 ml, 0.794 mmol) was added dropwise and the reaction mixture was stirred for 2.5 h at room temperature. The reaction was then chilled in an ice-bath, 0.30 ml of cold water was added and the solution was stirred for 15 min. Concentrated aqueous NH_4OH was added, and the solution was stirred for another 30 min and evaporated to give an oil with salts. Crude was purified to with silica gel column

chromatography to yield 33 mg of **1d** (58.8%). ^{1}H NMR (DMSO-d_6) δ 6.04 (t, 1H, H1'), 4.34-4.40 (m, 1H, H3'), 3.77 (m, 1H, H4'), 3.53-3.61 (m, 1H, H5'), 3.37-3.34 (m, H5'), 3.00 (m, 1H, H2'), 2.73-2.84 (m, 1H –C*H*$(CH_3)_2$), 1.90-2.01 (m, 1H, H2''), 1.09-1.11 (d, 6H, -CH(C*H*$_3)_2$). MS (ESI): $[M + H]^+$ *m/z* calcd 354.1, found 354.0.

Synthesis of 7,8-Dihydro-5'-*O*-(4,4'-dimethoxytrityl)-N^2-isobutyryl-8-oxo-2'-deoxyguanosine (1e, Scheme 4-1). (15) To a solution of **1d** (33 mg) in dry pyridine cooled in ice-water was added 4,4'-dimethoxytrityl chloride. The cooling bath was then removed, and stirring was continued at room temperature for 15 min. The reaction mixture was cooled in ice water and quenched with water. The solution was extracted with CH_2Cl_2 and the combined organic layers were washed with water and dried with $MgSO_4$. The solvent was evaporated under reduced pressure, and the crude residue was purified by column chromatography over silica gel using EtOAc-hexane containing 2% of triethylamine as the eluting solvent. The yield of **1e** was 60 mg (98.0%).

Synthesis of 1-(2'-Deoxy-3',5'-di-*O*-p-toluoyl-β-D-*erythro*-pentafuranosyl)-7-nitroindole (2a, Scheme 4-2). (18) To a solution of 7-nitroindole (395 mg, 2.4 mmol) in ACN (6 ml) was added NaH (60% dispersion in oil, 0.111 g, 2.78 mmol) and the solution was stirred at room temperature for 30 min. Chloro-3,5-di-*O*-toluoyl-2'-deoxy-α-D-ribofuranose (1.00 g, 2.59 mmol) was dissolved in about 20 ml of ACN. The 7-nitroindole solution was added to the deoxyribosyl chloride solution and the resulting solution was stirred overnight at room temperature. The solution was filtered through celite and the solvent removed *in vacuo*. The oily residue was purified by column chromatography (CH_2Cl_2) affording compound **2a** (233 mg, 16.5%) ^{1}H NMR ($CDCl_3$) δ 2.39 (s, 3H CH_3), 2.42 (s, 3H, CH_3), 2.73 and 2.90 (2m, 2H, H2' and H2''), 4.50 (m,

3H, H4', H5' and H5''), 5.59 (m, 1H, H3'), 6.55 (t, *J*= 6.4 Hz, 1H, H1'), 6.64 (d, *J* = 3.4 Hz, 1H, H2), 7.08-7.28 (m, 5H, 4H-Tol and H5), 7.57 (d, *J* = 3.4 Hz, 1H, H3), 7.77-7.97 (m, 6H, 4H-Tol and H4, H6).

Synthesis of 1-(2'-deoxy-β-D-erythro-pentafuranosyl)-7-nitroindole (2b, Scheme 4-2). Diester **2a** (233 mg, 0.397 mmol) was dissolved in 1% NaOH-methanolic solution and stirred for 90 min at room temperature. The solvent was then evaporated and the free nucleoside **2b** was purified by column chromatography (CH_2Cl_2/MeOH: 90/10, v/v) to give 44 mg (39.8%). ^{1}H NMR (DMSO-d_6) δ 2.40 (m, 2H, H2' and H2''), 3.37 (m, 2H, H5' and H5''), 3.76 (m, 1H, H4'), 4.28 (m, 1H, H3'), 4.87 (t, *J* = 4 Hz, 1H, 5'OH), 5.29 (d, *J* = 4 Hz, 1H, 3'OH), 6.19 (t, *J* = 6 Hz, 1H, H1'), 6.77 (d, *J* = 3.4 Hz, 1H, H2), 7.20 (t, *J* = 8 Hz, 1H, H6).

Synthesis of 5'-Dimethoxytrityl-7-nitroindolenucleoside (2c, Scheme 4-2). (11, 19) Compound **2b** (99 mg, 0.356 mmol) was dissolved in 4 ml of pyridine to which 0.094 ml (0.676 mmol) of TEA was added followed by 0.169 g (0.498 mmol) of DMTrCl and 2 mg (0.0178 mmol) of DMAP. The reaction was kept at room temperature for 1.5 h. After the reaction was complete, as monitored by TLC, the solvent was evaporated under vacuum and the crude was purified by silica gel column chromatography to yield 206.6 mg (80.3 %). ^{1}H NMR ($CDCl_3$) δ 2.57-2.62 (m, 2H, H2'), 3.27-3.28 (d, = 4Hz, 2H, H5', H5''), 3.78 (s, 6H, OCH_3), 4.08 (m, 1H, H4'), 4.48-4.55 (m, 1H, H3'), 6.44-6.59 (t, 1H, H1'), 6.74 (d, 1H, H3), 6.75-6.78 (m, 4H, Ar-H), 7.14 (m, 1H, H5), 7.15-7.34 (m, 9H, Ar-H), 7.61-7.62 (d, 1H, H2), 7.78 and 7.84 (d, *J* = 7 Hz, 2H, H4, H6).

Synthesis of 5'-Dimethoxytrityl-7-nitroindole phosphoramidite (2d, Scheme 4-2). (11, 19) Compound **2c** (130 g, 0.224 mmol) was dissolved in 3 ml of anhydrous CH_2Cl_2. DIEA (98 μl,

0.593 mmol) was added followed by 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (75 μl, 0.336 mmol). The reaction was stirred at room temperature under Argon for 0.5 h. The crude was purified by silica gel flash column chromatography that had been treated with a small amount of TEA. The yield of purified product was 131 mg (74.9%). ^{31}P NMR gave peaks at 147.236 and 147.530 ppm. MS (ESI): $[M + Na]^+$ *m/z* calcd 803.3, found 803.3.

References

1. Grollman, A. P, Takeshita, M. and Shibutani, S. (1991) Insertion of specific bases during DNA synthesis past the oxidation damaged base 8-oxodG. *Nature*. 349: 331-334.

2. Cheng, K. C., Cahhill, D. S., Kasai, H., Nishimura, S. and Loeb, L. A. (1992) 8-Hydroxygaunane, an Abundant Form of Oxidative DNA Damage, Causes G→T and A→C Substitutions. *The J. Biol. Chem*. 267(1): 166-172.

3. Moriya, M. (1993) Single-stranded shuttle phagemid for mutagenesis studies in mammalian cell: 8-oxoguanine in DNA induces targeted G•C →T•A transversions in simian kidney cells. *Pro. Natl. Acad. Sci. USA*. 90: 1122-1126.

4. Efrati, E., Tocco, G., Eritja, R., Wilson, S. H. and Goodman, M. F. (1999) "Action-at-a-Distance" Mutagenesis. *J. Biol. Chem*. 274(22): 15920-15927.

5. Marnett, L. (2000) Oxyradicals and DNA damage. *Carcinogenesis*. 21(3): 361-370.

6. Johnson, K. A., Thai, D. M. and Hanes, J. W. (2006) Incorporation and Replication of 8-oxo-deoxyguanosine by the Human Mitochondrial DNA Polymerase. *J. Biol. Chem*. 281(47): 36241-36248.

7. Pursell, Z. F., McDonald, J. T., Mathews, C. K. and Kunkel, T. A. (2008) Trace amounts of 8-oxo-dGTP in mitochondrial dNTP pools reduce DNA polymerase γ replication fidelity. *Nucleic Acids Res*. 36(7): 2174-2181.

8. Mendelman, L. V., Boosalis, M. S., Petruska, J. and Goodman, M. F. (1989) Nearest Neighbor Influences on DNA Polymerase Insertion Fidelity. *J. Biol. Chem*. 264(24): 14415-14423.

9. Sasaki, S., Kawaguchi, R. and Taniguchi, Y. (2011) Adenosine-1,3-diazaphenoxazine Derivative for Selective Base Pair Formation with 8-Oxo-2'-deoxyguansine in DNA. *J. Am. Chem. Soc*. 133: 7272-7275.

10. Greenberg, M. M. and Chen, T. (1998) Model Studies Indicate That Copper Phenanthroline Induces Direct Strand Breaks via β-Elimination of the 2'-Deoxyribonolactone Intermediate Observed in Enediyne Mediated DNA Damage. *J. Am. Chem. Soc*. 120: 3815-3816.

11. Kotera, M., Roupioz, Y., Defrancq, E., Bourdat, A., Garcia, C. C. and Lhomme, J. (2000) The 7-nitroindole Nucleoside as a Photochemical Precursor of 2'-Deoxyribonolactone: Access to DNA Fragments Containin This Oxidative Abasic Lesion. *Chem. Eur. J*. 6(22): 4163-4169.

12. Saparbaev, M., Dumy, P., Constant, J. and Faure, V. (2004) 2'-Deoxyribonolactone lesion produces G → A transitions in *Escherichia coli*. *Nucleic Acids Res*. 32(9): 2937-2946.

13. Crey-Desbiolles, C., Berthet, N., Kotera, Mitsuharu and Dumy, P. (2005) Hybridization properties and enzymatic replication ofoligonucleotides containing the photocleavable 7-nitroindole base analog. Nucleic Acids *Res. Nucleic Acids Res*. 33(5): 1532-1543.

14. Greenberg, M. M. and Imoto, S. (2008) DNA Tandem Lesion Repair by Strand Displacement Synthesis and Nucleotide Excision Repair. *Biochem*istry. 47: 4306-4316.

15. Johnson, F., Shibutani, S. and Bodepudi, V. (1992) Synthesis of 2'-Deoxy-7,8-dihydro-8-oxoguanosine and 2'-Deoxy-7,8-dihydro-8-oxoadenosine and Their Incorporation into Oligomeric DNA. *Chem. Res. Toxicol*. 5: 608-617.

16. Kumar, S. and Nampalli, S. (2000) Efficient Synthesis of 8-oxo-dGTP: A Mutagenic Nucleotide. *Bioorg. Med. Chem. Lett*. 10: 1677-1679.

17. Bruice, T. and Challa, H. (2004) Deoxynucleic guanidine: synthesis and incorporation of purine nucleosides into positively charged DNG oligonucleotides. *Bioorg. Med. Chem*. 12: 1475-1481.

18. Lhomme, J., Defrancq, E., Bourdat, A. and Kotera, M. (1998) A Highly Efficient Synthesis of Oligodeoxyribonucleotides Containing the 2'-Deoxyribonolactone Lesion. *J. Am. Chem. Soc*. 120: 11180-11811.

19. Saito, I., Tainaka, K. and Okamoto, A. (2002) A facile incorporation of the aldehyde function into DNA: 3-formylindole nucleoside as an aldehyde-containing universal nucleoside. *Tetrah. Lett*. 43: 4581-4583.

20. Ludovic, C. J. and Scharer, O. D. (2002) Preparation of C8-Amine and Acetylamine Adducts of 2'-Deoxyguanosine Suitably Protected for DNA Synthesis. Org. Lett. 4(24): 4205-4208.

21. Yuan, B., Y. Jiang and Wang, Y. (2010) Efficient formation of the tandem thymine glycol/8-oxo-7,8-dihydroguanine lesion in isolated DNA and the mutagenic and cytotoxic properties of the tandem lesions in Escherichia coli cells. *Chem. Res. Toxicol*. 23(1): 9-11.

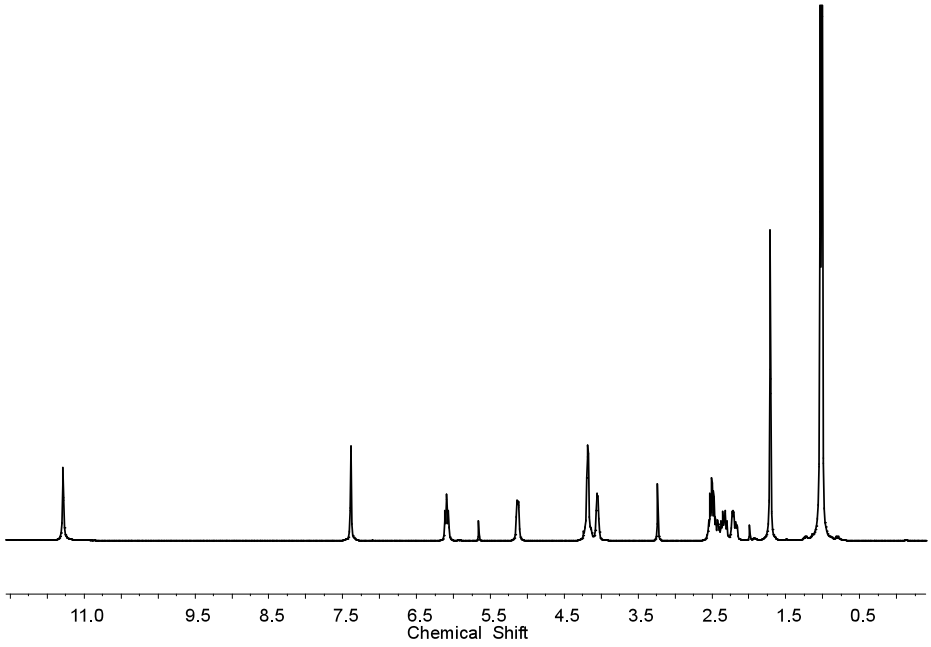


Figure S2-1 ^{1}H NMR spectrum of 3',5'-*O*-diisobutyrylthymidine in DMSO-d_6 (300 MHz, 25 °C)

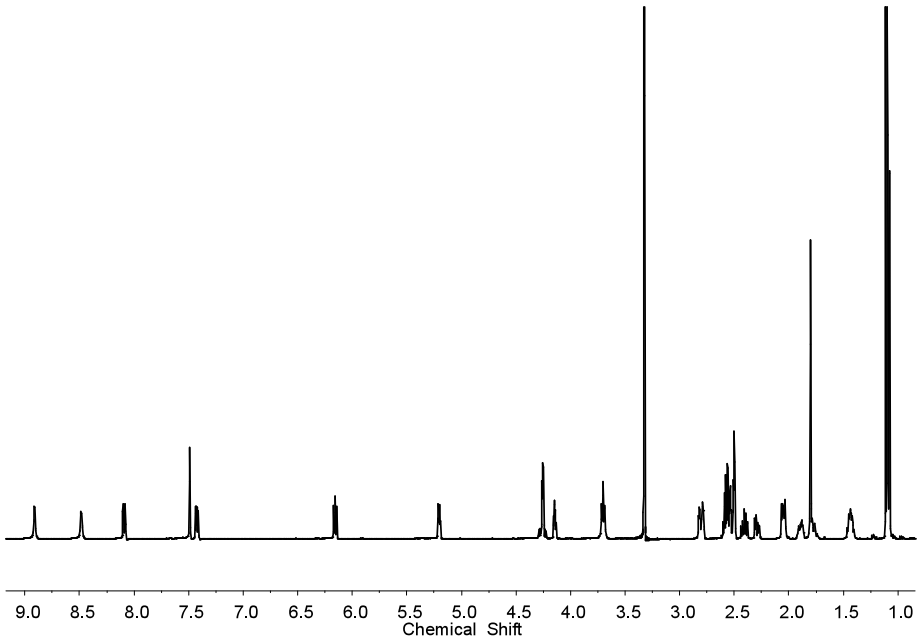


Figure S2-2 ^{1}H NMR spectrum of *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propyl]-3',5'-*O*-[isobutyryll]thymidine in DMSO-d_6 (300 MHz, 25 °C)

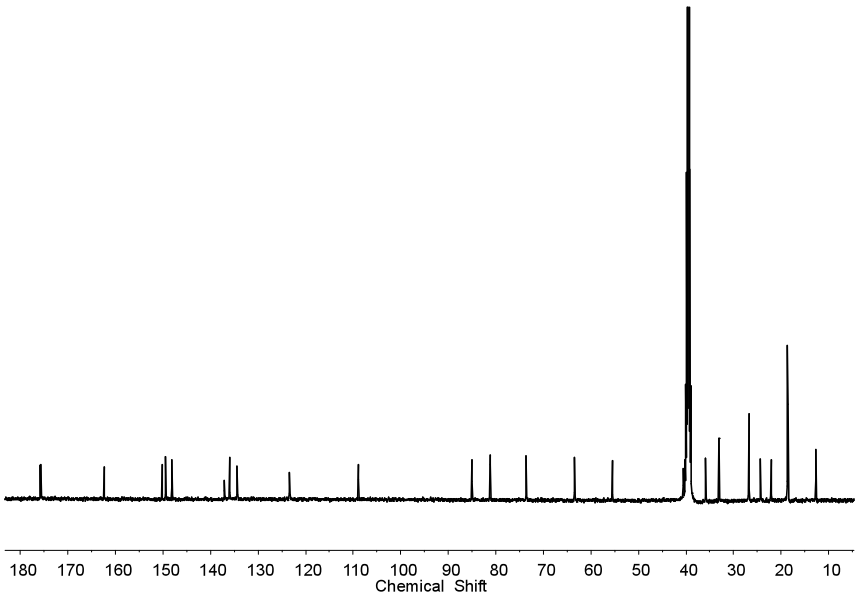


Figure S2-3 ^{13}C NMR spectrum for *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl}propy]-3',5'-*O*-[isobutyryll]thymidine in DMSO-d_6 (300 MHz, 25 °C)

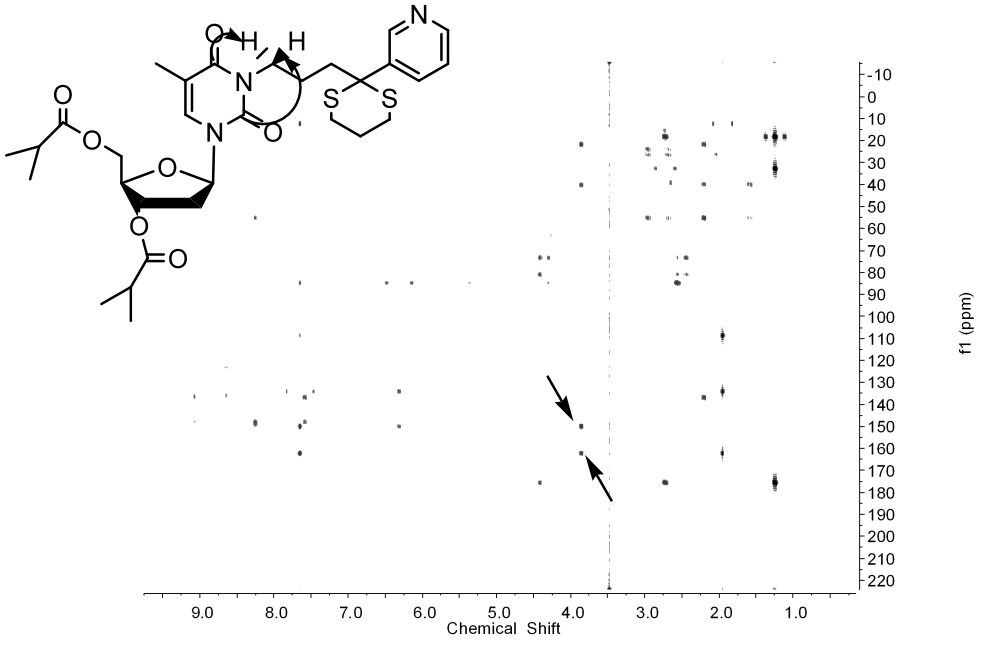


Figure S2-4 ^{1}H-^{13}C HMBC spectrum for *N*3-[3-{2-(3-Pyridyl)-1,3-dithian-2-yl]-3',5'-*O*-diisobutyryllthymidine in DMSO-d_6 (500 MHz, 25 °C)

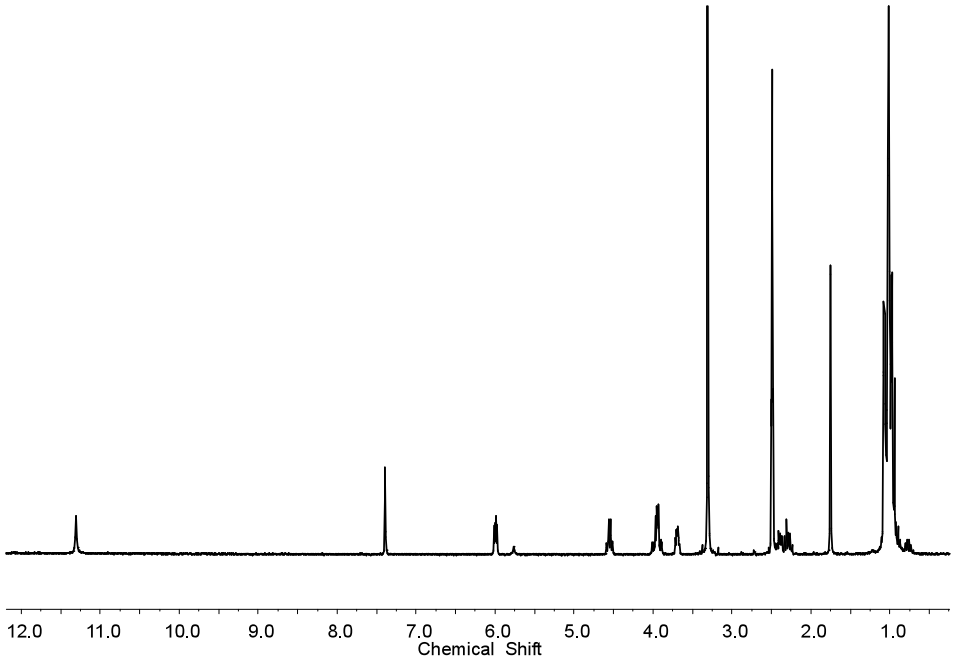


Figure S2-5 ^{1}H NMR spectrum of 3',5'-*O*-[1,1,3,3-tretrakis(isopropyl)-1,3-diisiloxane]-thymidine in DMSO-d_6 (300 MHz, 25 °C)

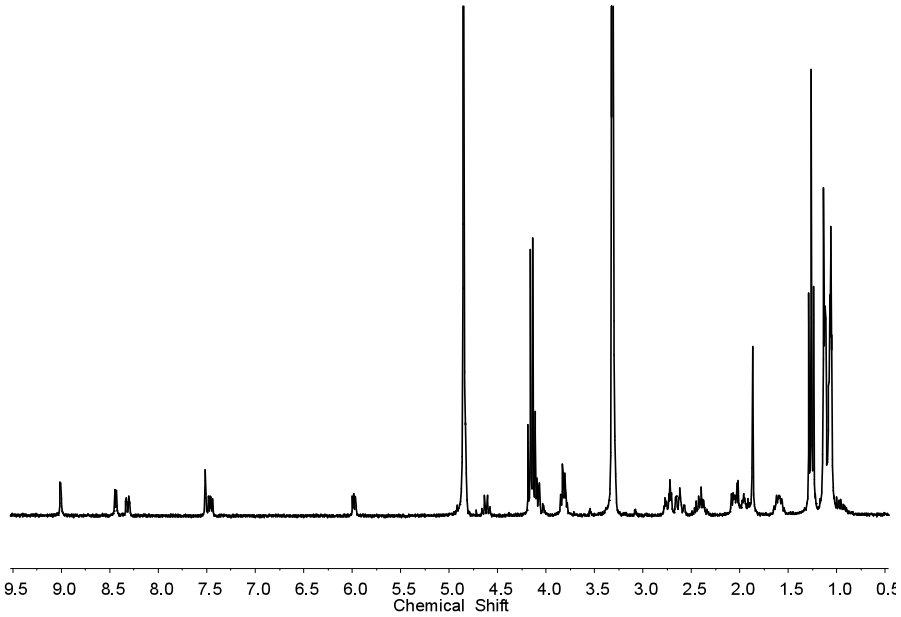


Figure S2-6 ^{1}H NMR spectrum of *N*3-[3-{2-(3-pyridyl)-1,3-dithiane-2-yl}propyl]-3',5'-*O*-[1,1,3,3-tetrakis(isopropyl)-1,3-disiloxanediyl]thymidine in DMSO-d_6 (300 MHz, 25 °C)

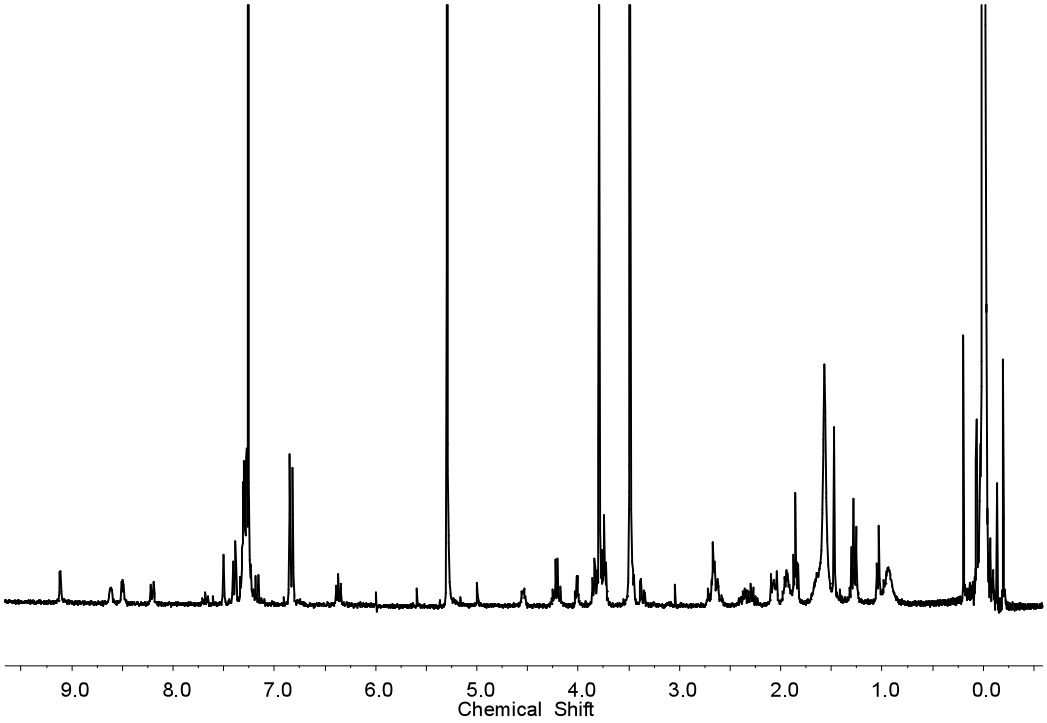


Figure S2-7 ^{1}H NMR spectrum of *N*3-[3-{2-(3-pyridyl)-1,3-dithiane-2-yl}propyl]-3',5'-*O*-[4,4'-dimethoxytrityl]thymidine in DMSO-d_6 (300 MHz, 25 °C)

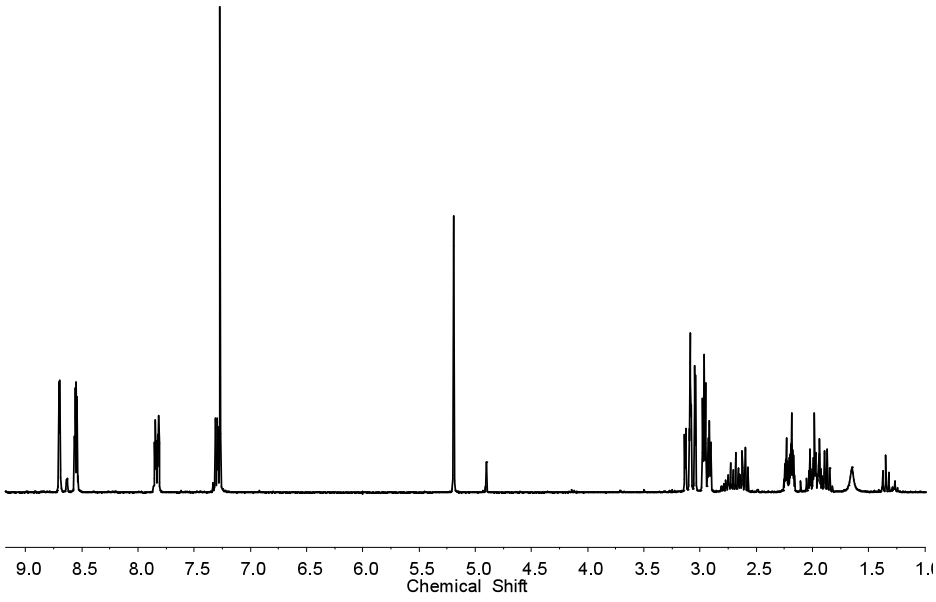


Figure S2-8 ^{1}H NMR spectrum of 2-(3-pyridyl)-1,3-dithiane in $CDCl_3$ (300 MHz, 25 °C)

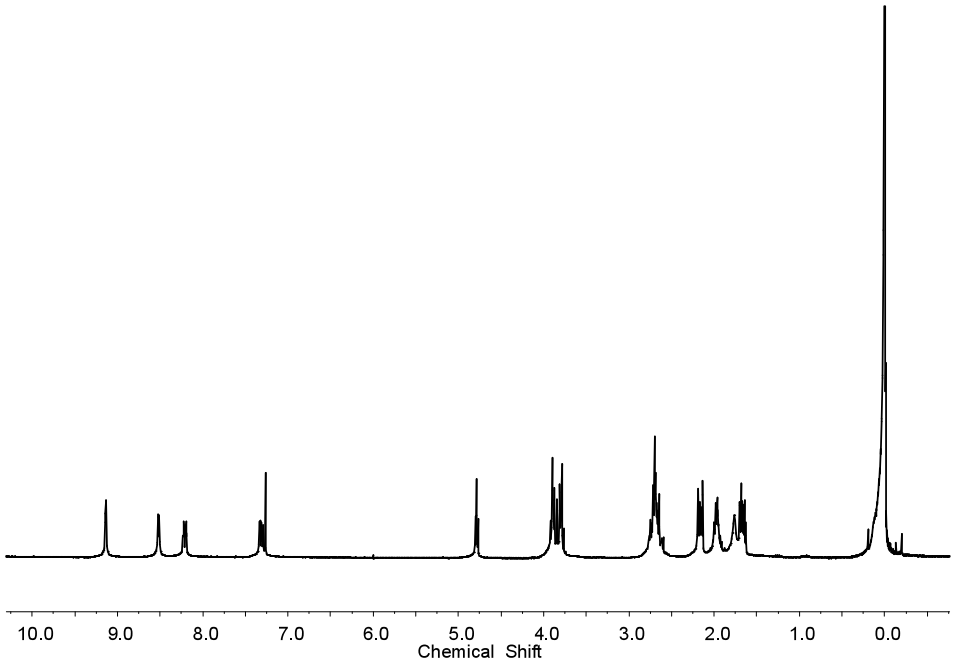


Figure S2-9 ^{1}H NMR spectrum of 2-[2-(3-pyridyl)-1,3-dithiane]ethyl-1,3-dioxolane in $CDCl_3$ (300 MHz, 25 °C)

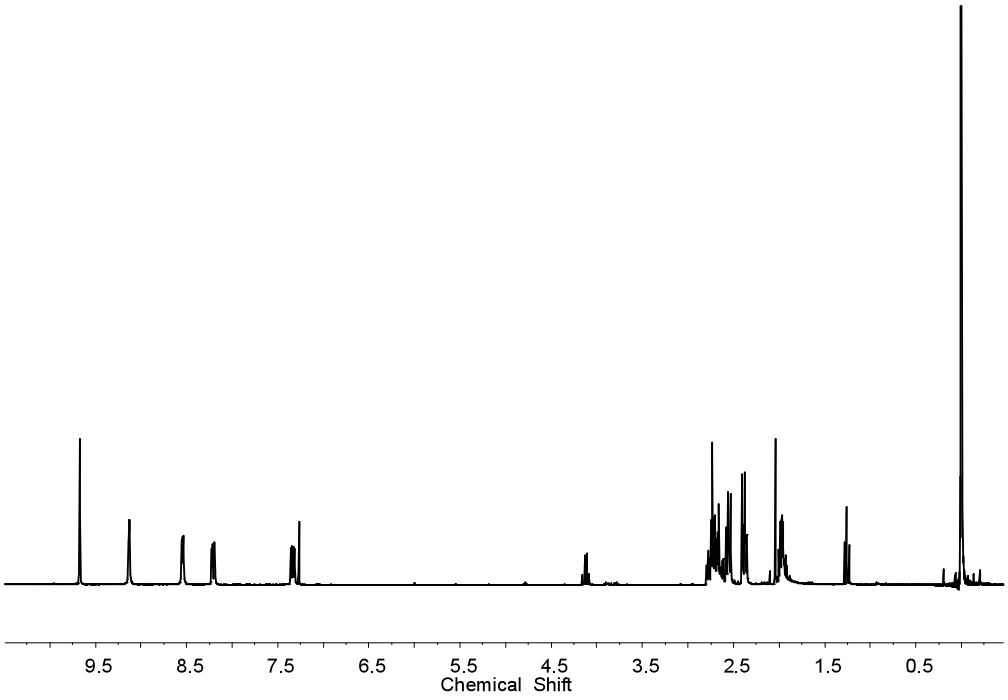


Figure S2-10 ^{1}H NMR spectrum of 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butanal in $CDCl_3$ (300 MHz, 25 °C)

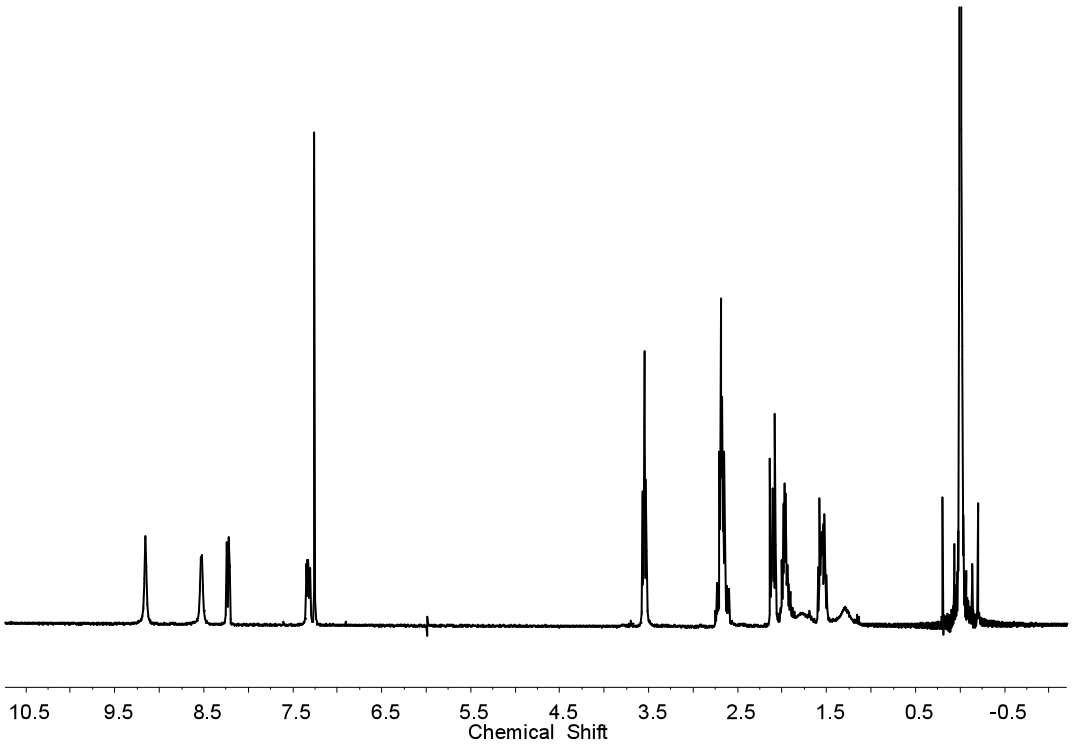


Figure S2-11 ^{1}H NMR spectrum of 4-(1,3-dithiane-2-yl)-4-(3-pyridyl)butan-1-ol in $CDCl_3$ (300 MHz, 25 °C)

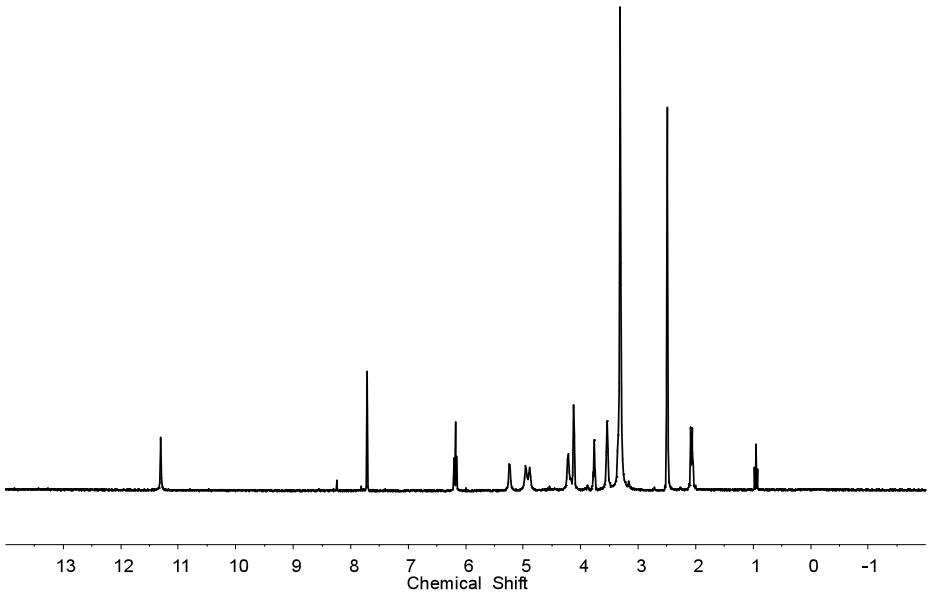


Figure S3-1 ^{1}H NMR of 5-hydroxymethyl-2'-deoxyuridine in DMSO-d_6 (300 MHz, 25 °C)

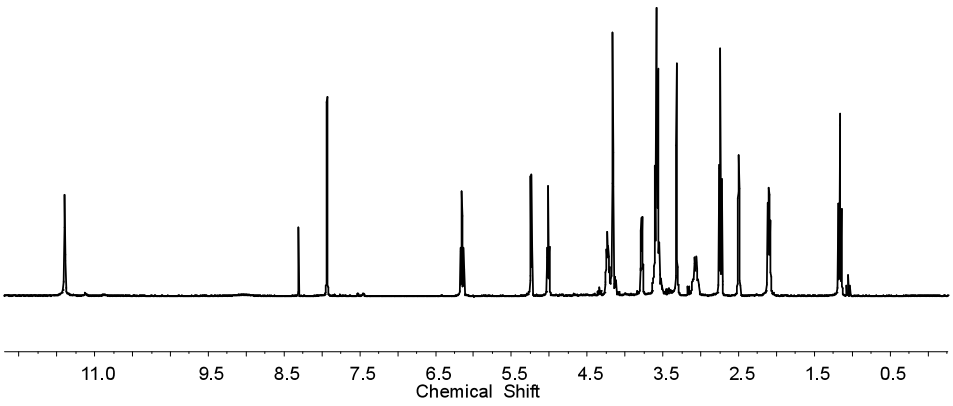


Figure S3-2 ^{1}H NMR spectrum of 5-(2-cyanomethoxy)methyl-2'-deoxyuridine in DMSO-d_6 (300 MHz, 25 °C)

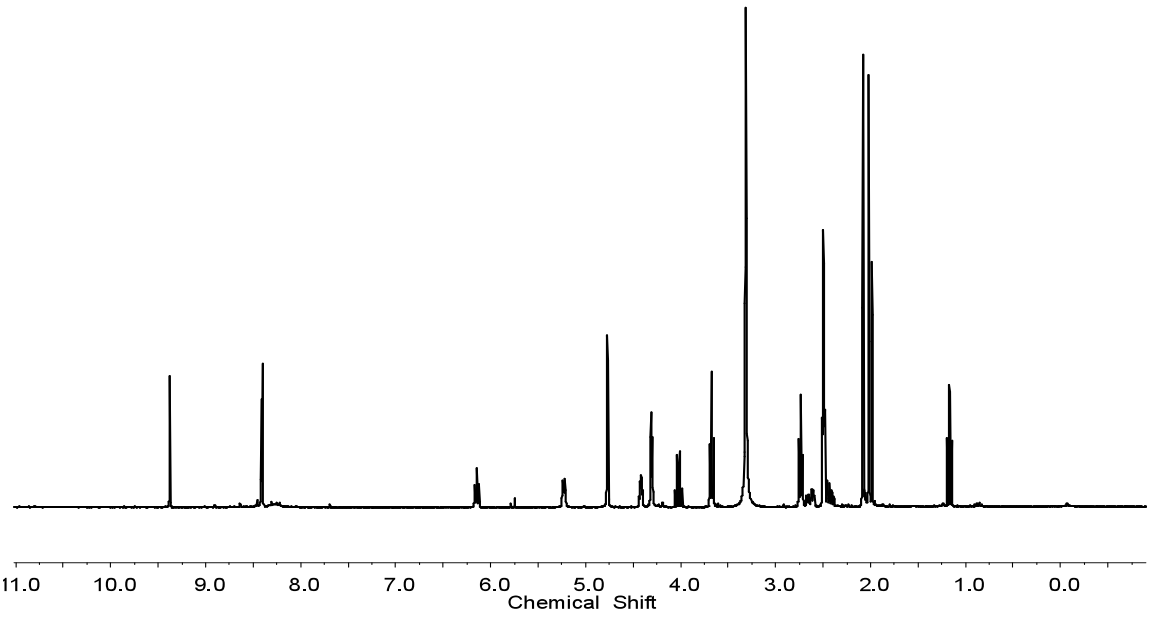


Figure S3-3 ^{1}H NMR spectrum of 3',5'-bis-*O*-Acetyl-5-(2-cyanoethoxy)methyl-N^4-triazol-2'-deoxycytidine in DMSO-d_6 (300 MHz, 25 °C)

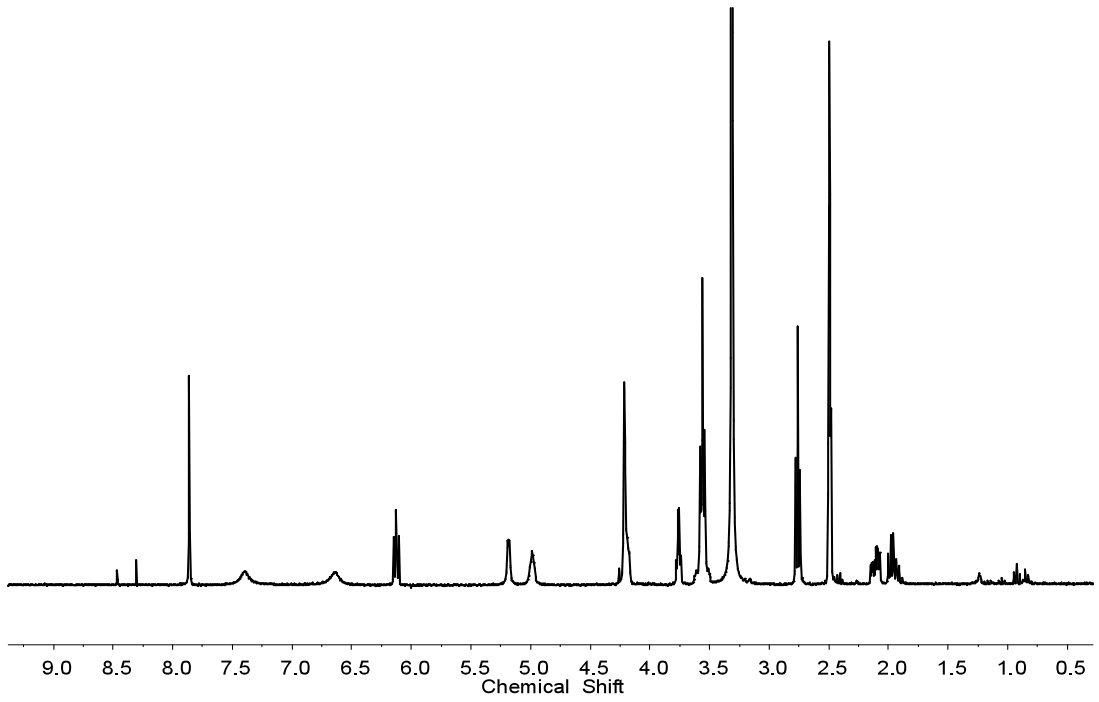


Figure S3-4 ^{1}H NMR spectrum of 5-(2-cyanoethoxy)methyl-2'-deoxycytidine in DMSO-d_6 (300 MHz, 25 °C)

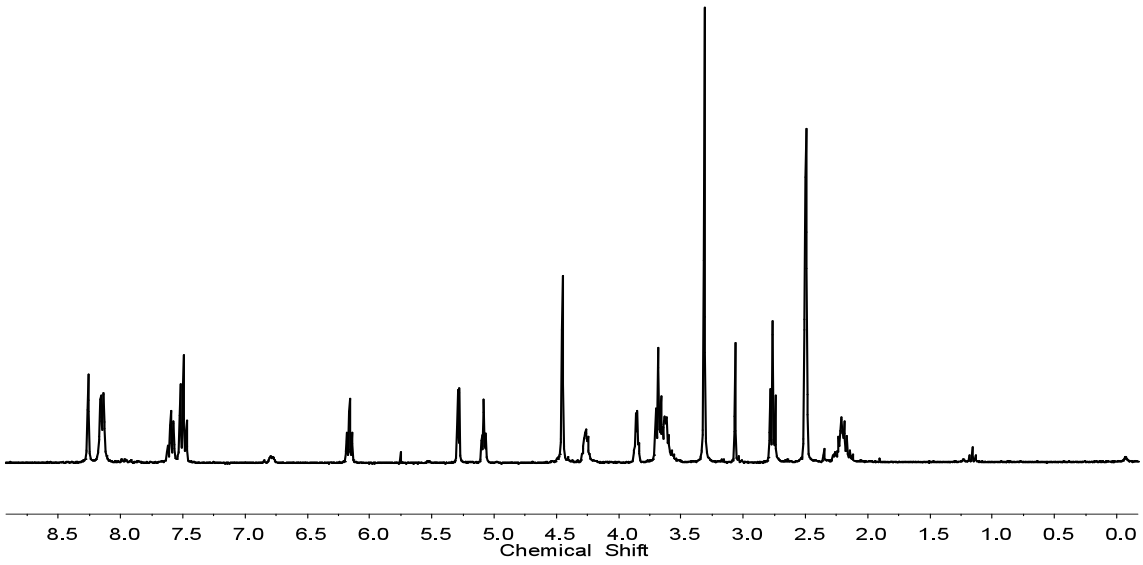


Figure S3-5 ^{1}H spectrum of N^4-benzoyl-5-(2-cyanoethoxy)methyl-2'-deoxycytidine in DMSO-d_6 (300 MHz, 25 °C)

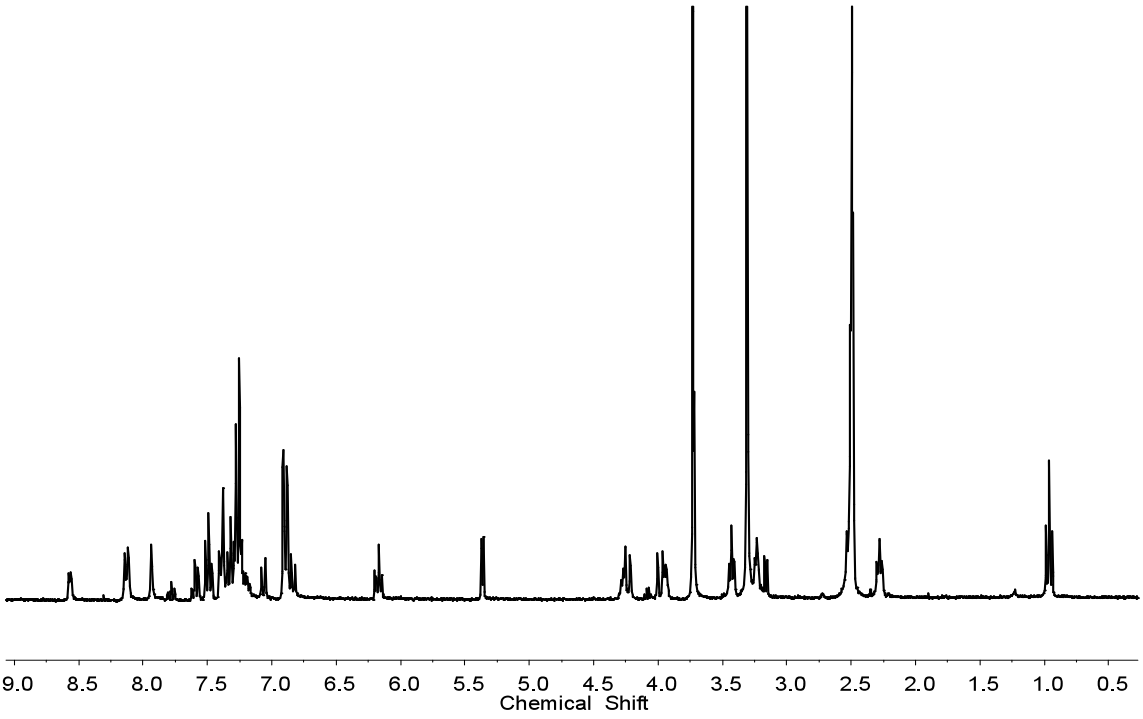


Figure S3-6 ^{1}H NMR spectrum of N^4-benzoyl-5-(2-cyanoethoxy)methyl-5'-(4,4'-dimethoxytrityl)-2'-deoxycytidine in DMSO-d_6 (300 MHz, 25 °C)

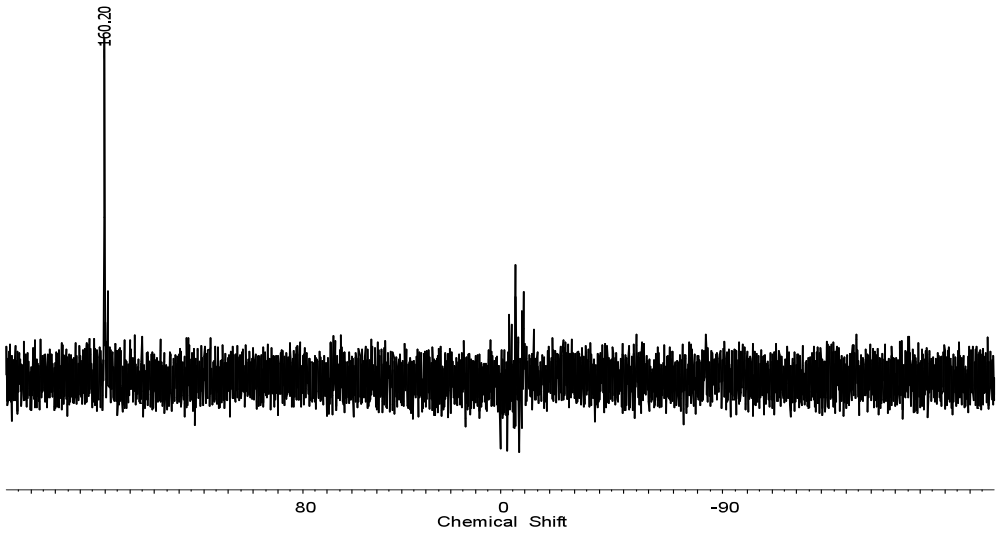


Figure S3-7 ^{31}P NMR spectrum of N^4-benzoyl-5-(2-cyanoethoxy)-methyl-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine-3'-*O*-(2-cyanoethyl)-*N*,*N*-diisopropylphosphoramidite in non-deuterated solvent (300 MHz, 25 °C)

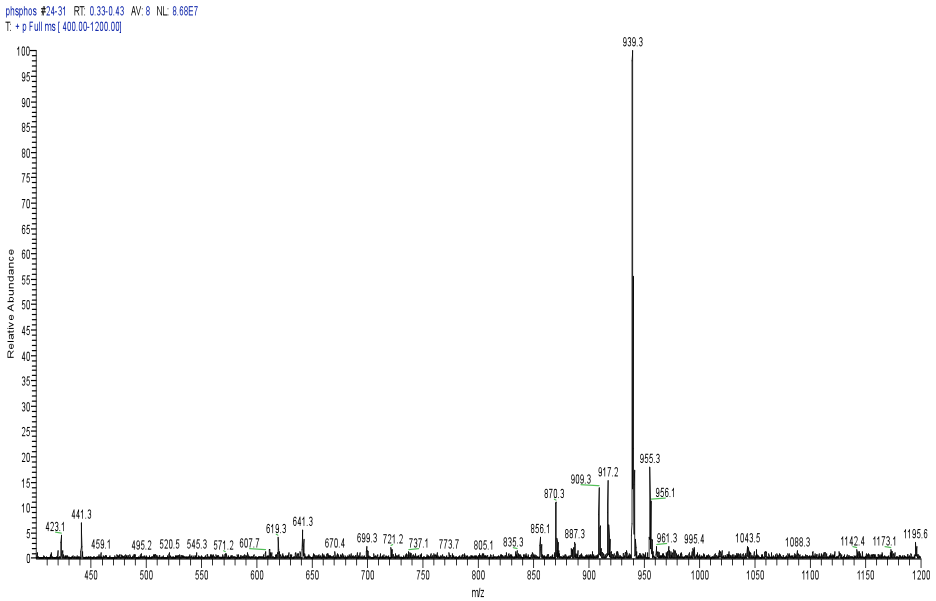


Figure S3-8 Positive-ion MS (ESI) for hmdC-phosphoramidite

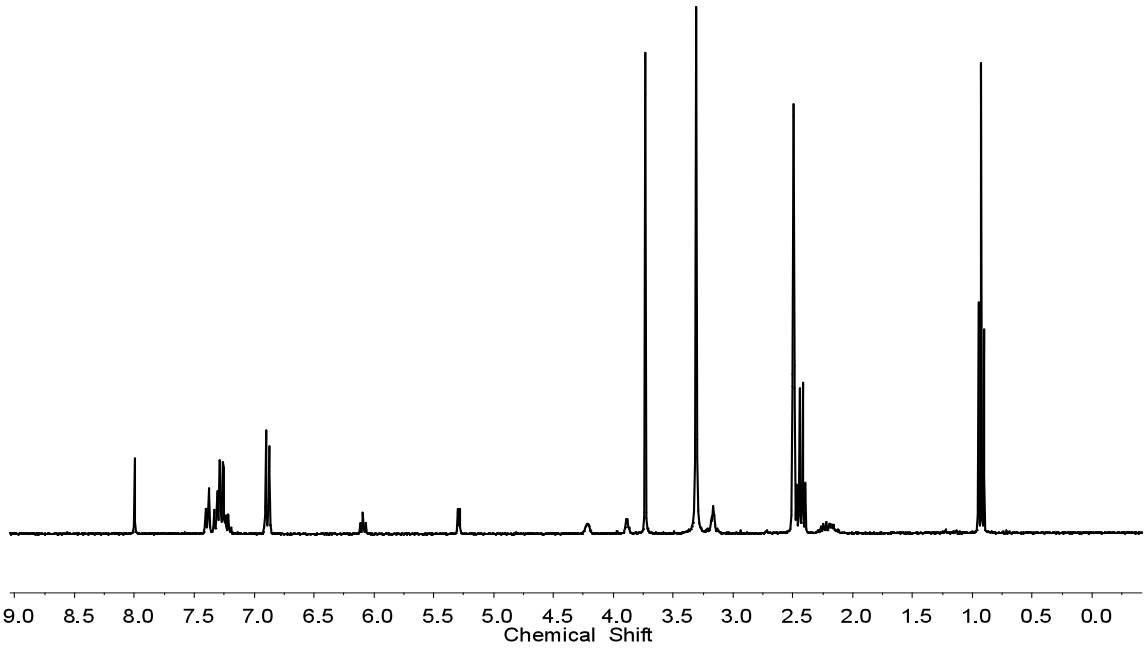


Figure S3-9 ^{1}H NMR spectrum of 5'-*O*-(4,4'-Dimethoxytrityl)-5-iodo-2'-deoxyuridine in DMSO-d_6 (300 MHz, 25 °C)

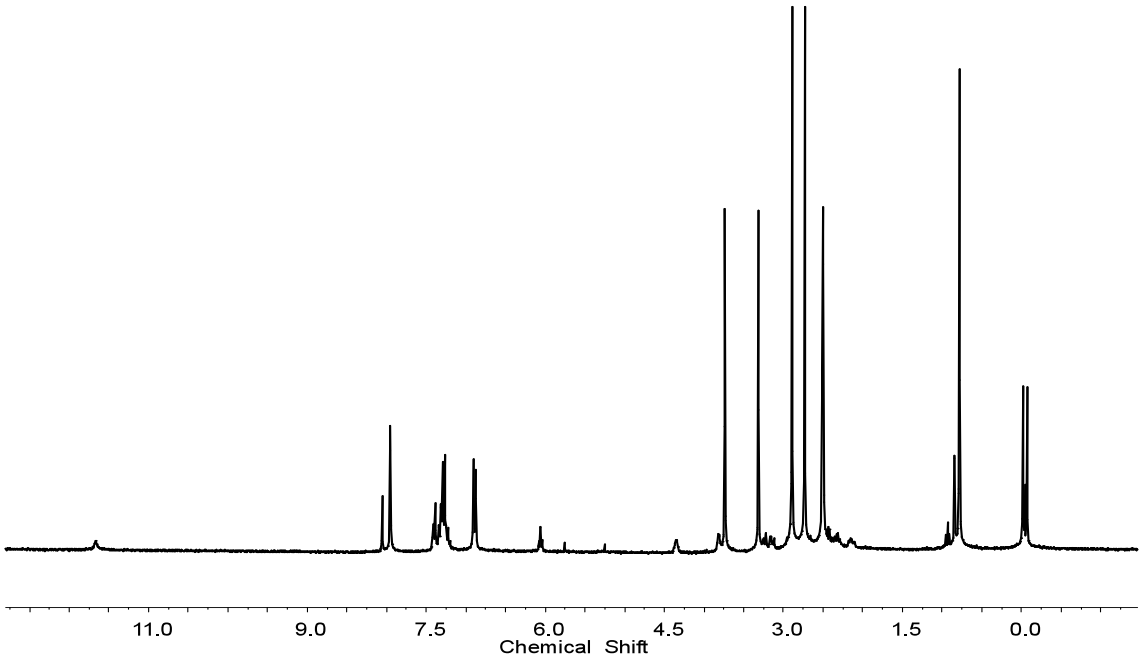


Figure S3-10 ^{1}H NMR spectrum of 3'-*O*-(t-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-5-iodo-2'-deoxyuridine in $CDCl_3$ (300 MHz, 25 °C)

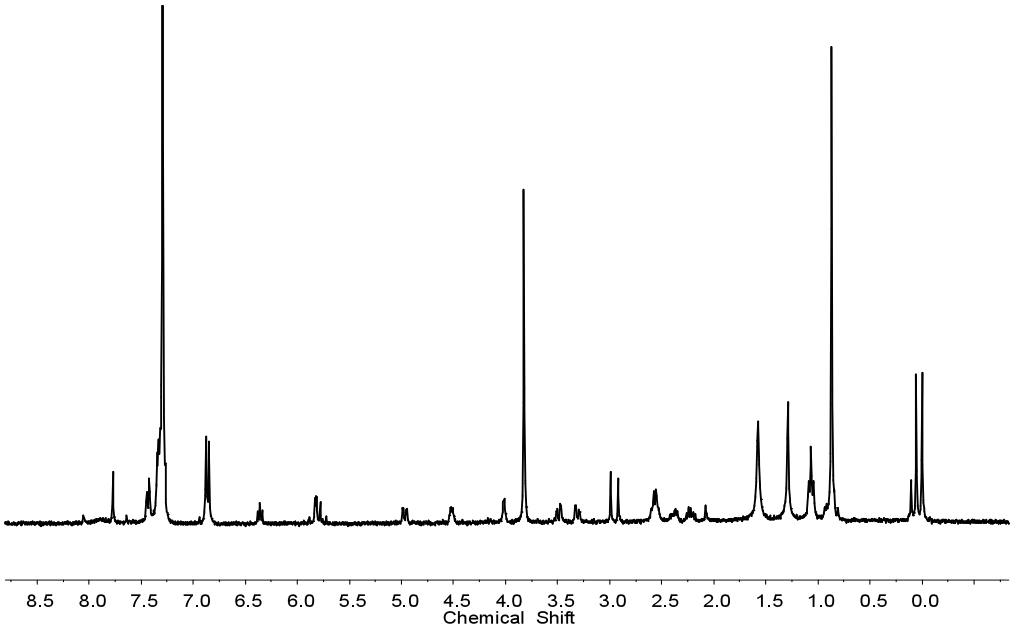


Figure S3-11 ^{1}H NMR spectrum of 3'-*O*-(t-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-5-vinyl-2'-deoxyuridine in $CDCl_3$ (300 MHz, 25 °C)

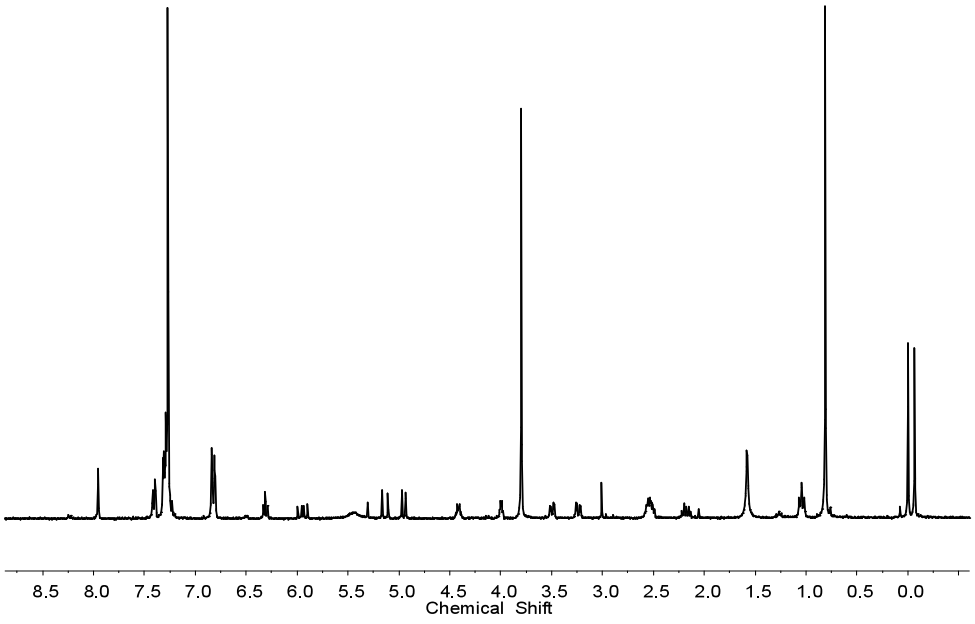


Figure S3-12 ^{1}H NMR for 3'-*O*-(t-butyldimethylsilyl)-5'-*O*-(4,4'-dimethoxytrityl)-5-vinyl-2'-deoxycytosine in $CDCl_3$ (300 MHz, 25 °C)

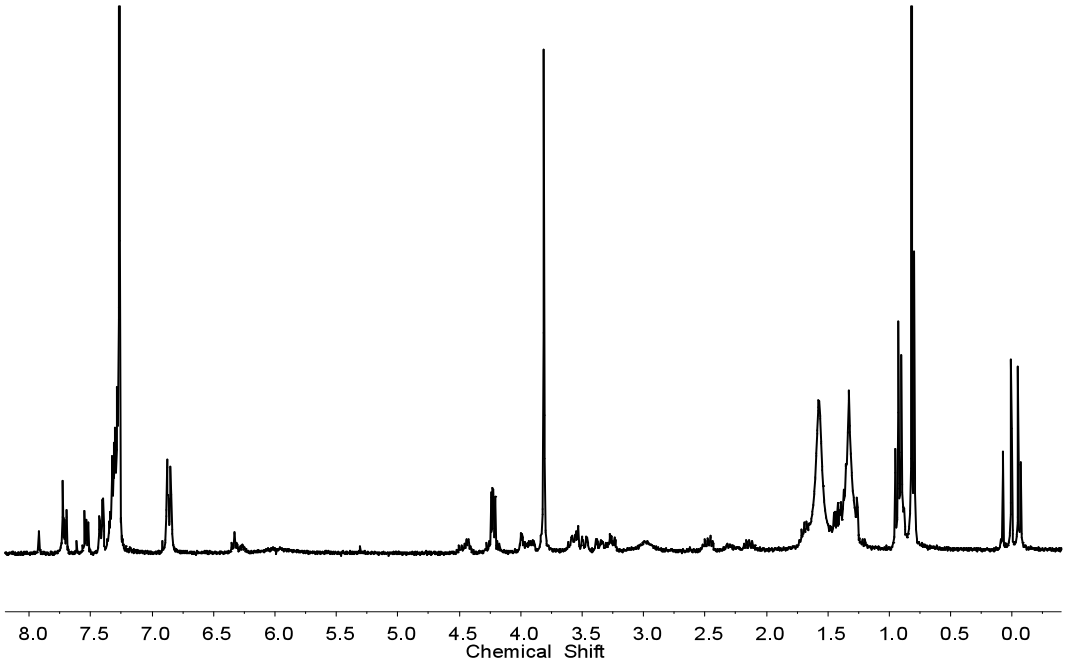


Figure S3-13 ^{1}H NMR spectrum of 3'-*O*-(t-butyldimethylsilyl)-5-(1,2-dihydroxyethyl)-5'-*O*-(4,4'-dimethoxytrityl)-2'-deoxycytidine in $CDCl_3$ (300 MHz, 25° C)

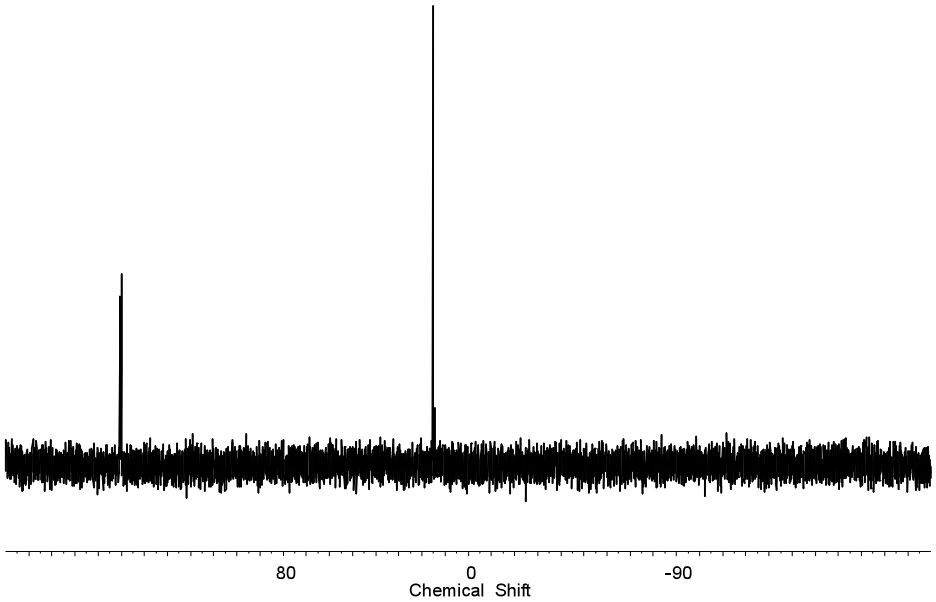


Figure S3-14 ^{31}P NMR spectrum of dheC-phosphoramidite in non- deuterated solvent (300 MHz, 25 °C)

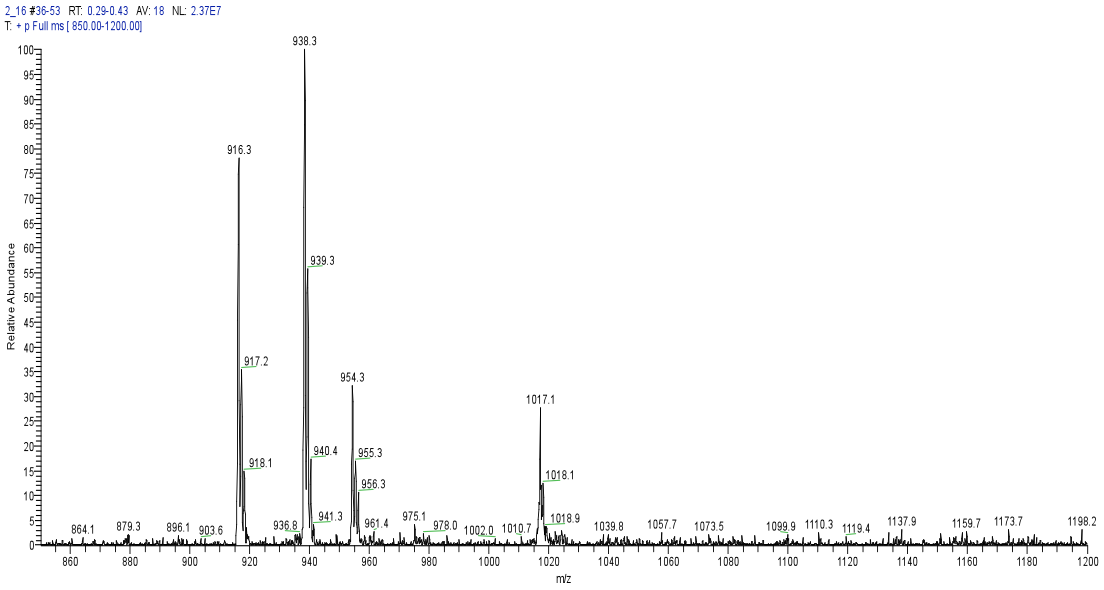


Figure S3-15 Positive-ion MS (ESI) for dheC-phosphoramidite

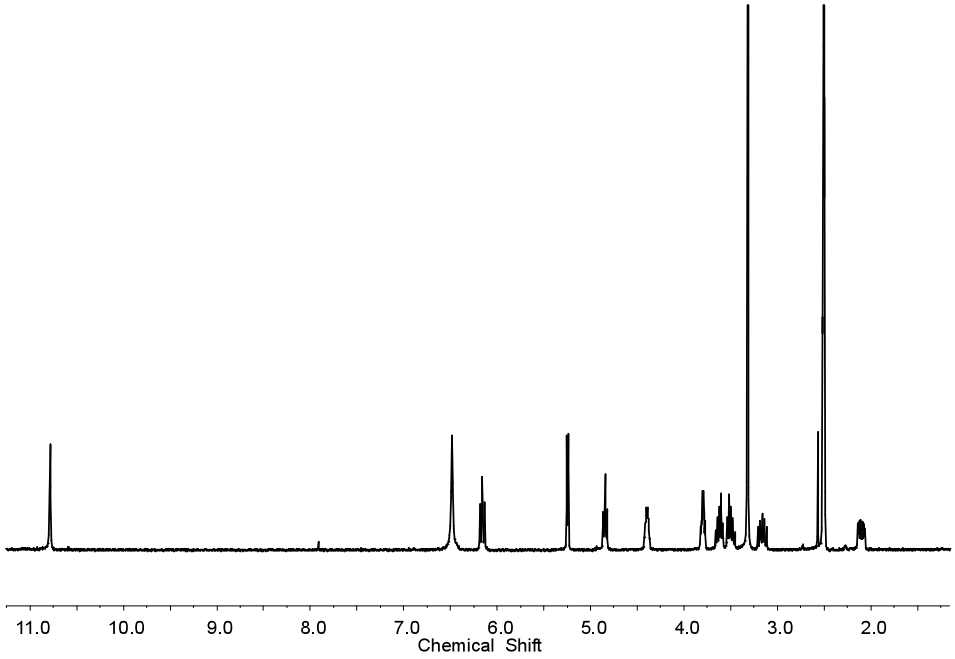


Figure S4-1 ^{1}H NMR spectrum of 8-bromo-2'-deoxyguanosine in DMSO-d_6 (300 MHz, 25 °C)

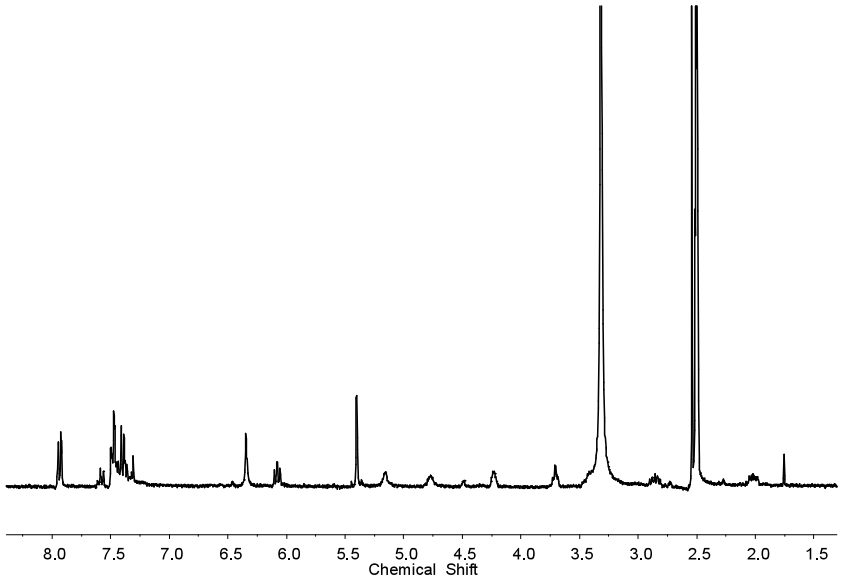


Figure S4-2 ^{1}H NMR spectrum of 8-Benzyloxy-2'-deoxyguanine in DMSO-d_6 (300MHz, 25 °C)

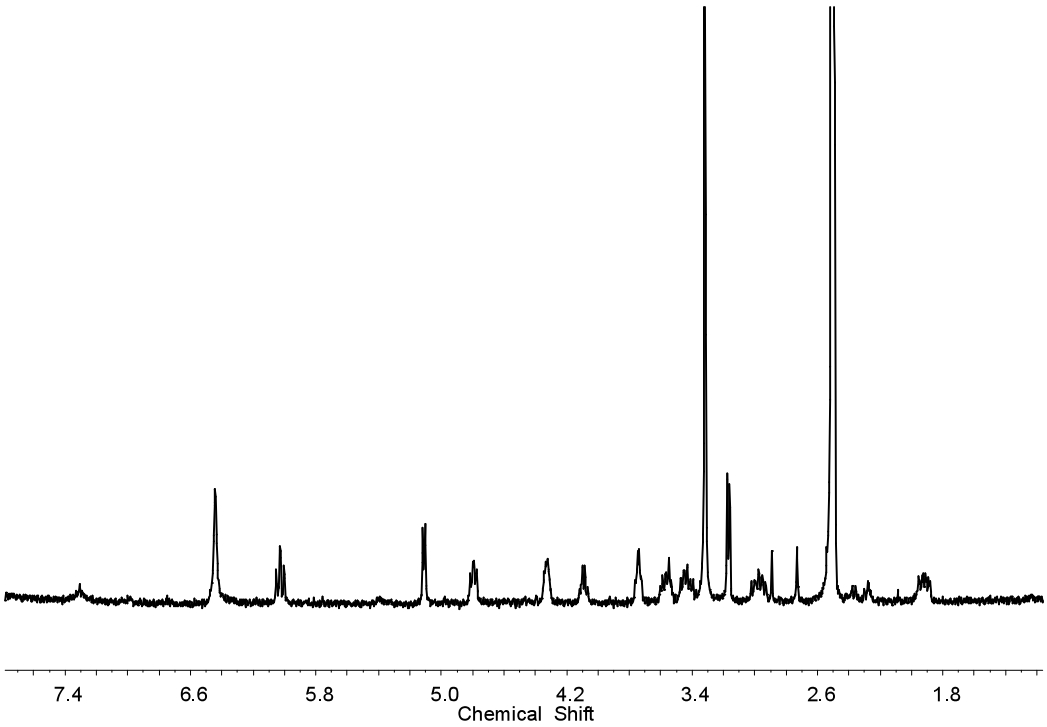


Figure S4-3 ^{1}H NMR spectrum of 8-oxo-7,8-dihydro-2'-deoxyguanosine in DMSO-d_6 (300 MHz, 25 °C)

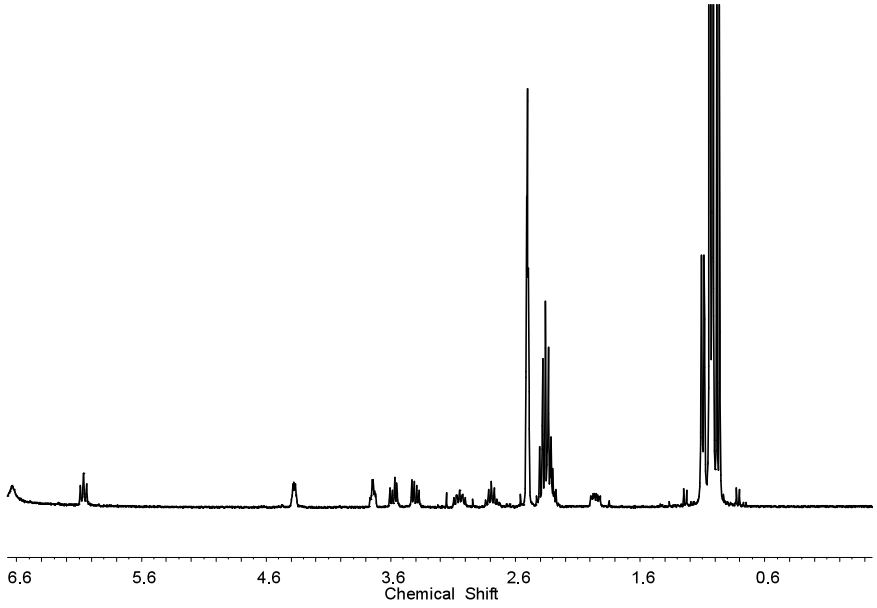


Figure S4-4 ^{1}H NMR spectrum of 7,8-dihydro-N^2-isobutyryl-8-oxo-2'-deoxyguanosine in DMSO-d_6 (300 MHz, 25 °C)

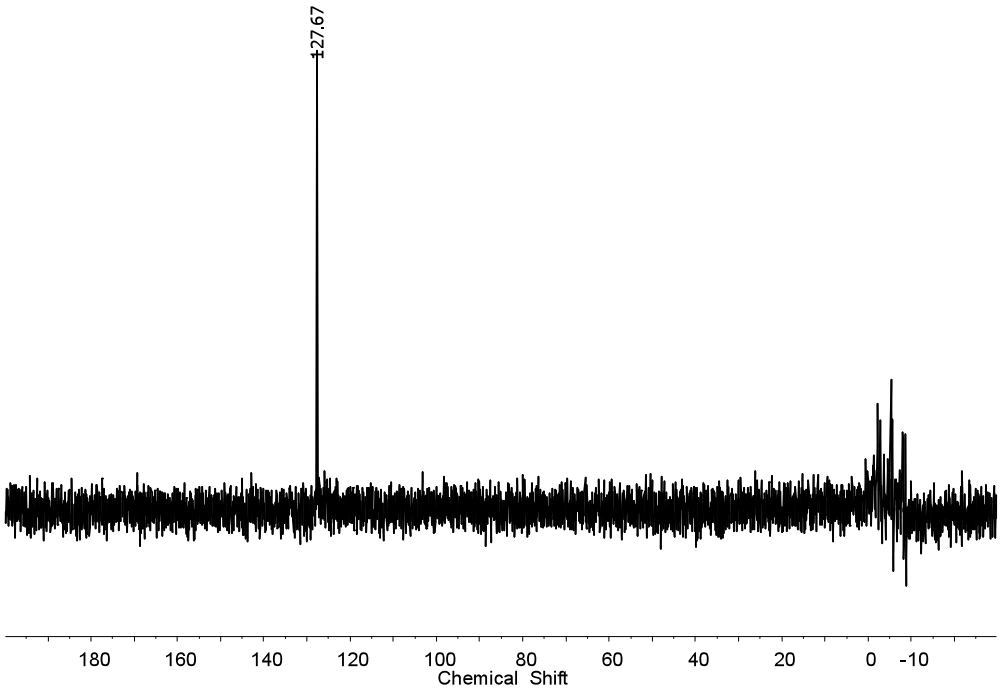


Figure S4-5 ^{31}P for 7,8-dihydro-5'-*O*-(4,4'-dimethoxytrityl)-N^2-isobutyryl-8-oxo-2'-deoxyguanosine phosphoramidite

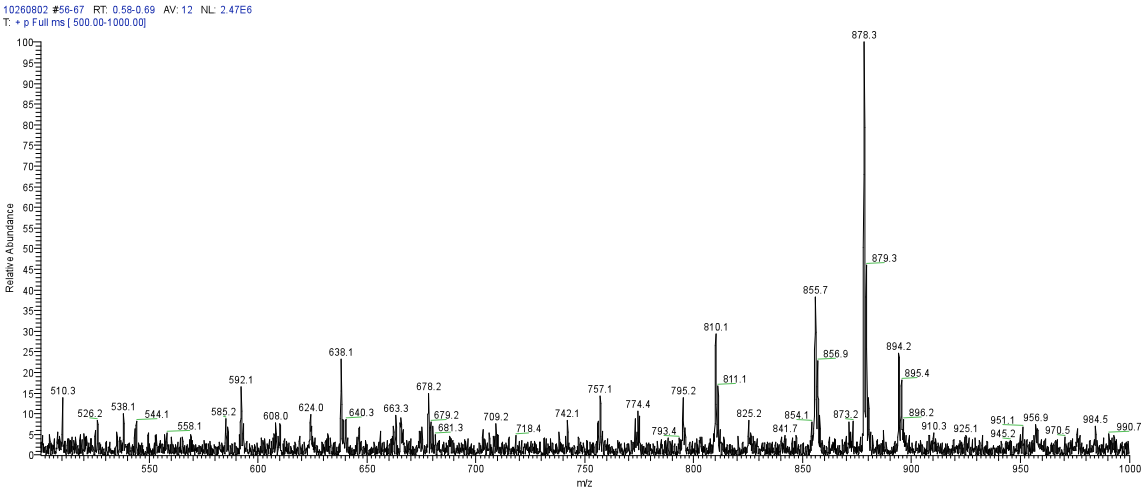


Figure S4-6 Positive-ion MS (ESI) of 8-oxo-dG phosphoramidite

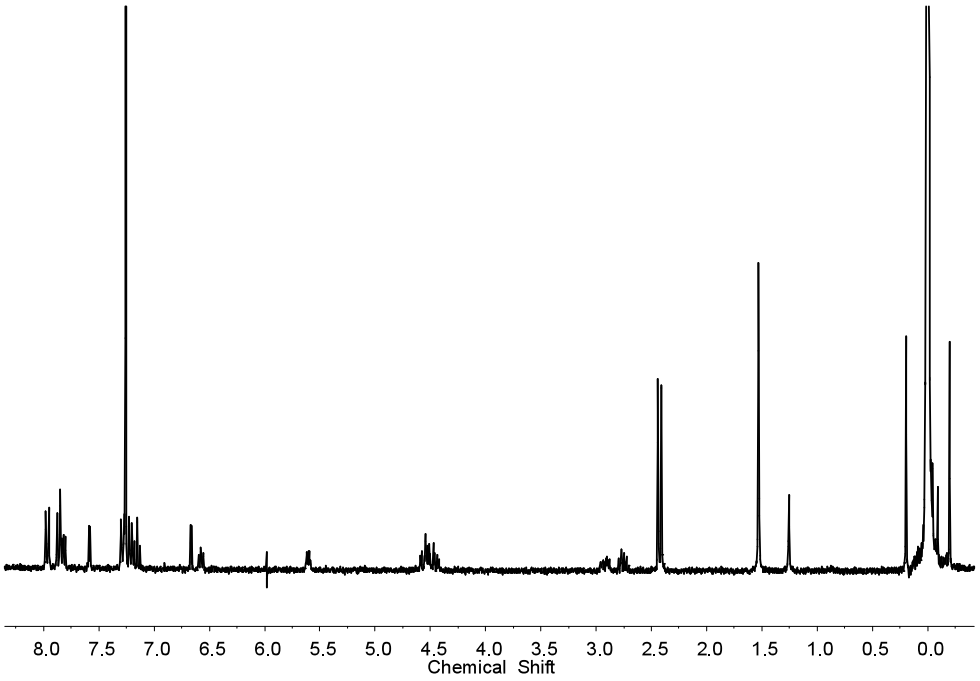


Figure S4-7 ^{1}H NMR spectrum of 1-(2'-Deoxy-3',5'-di-*O*-p-toluoyl-β-D-*erythro*-pentafuranosyl)-7-nitroindole in $CDCl_3$ (300 MHz, 25 °C)

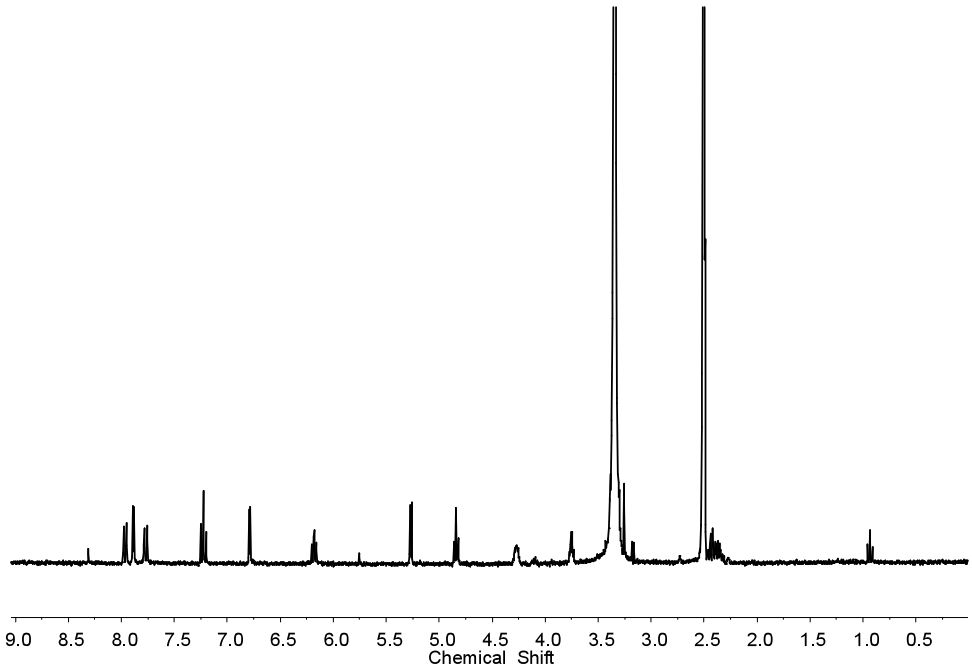


Figure S4-8 ^{1}H NMR spectrum of 1-(2'-deoxy-β-D-*erythro*-pentafuranosyl)-7-nitroindole in DMSO-d_6 (300 MHz, 25°C)

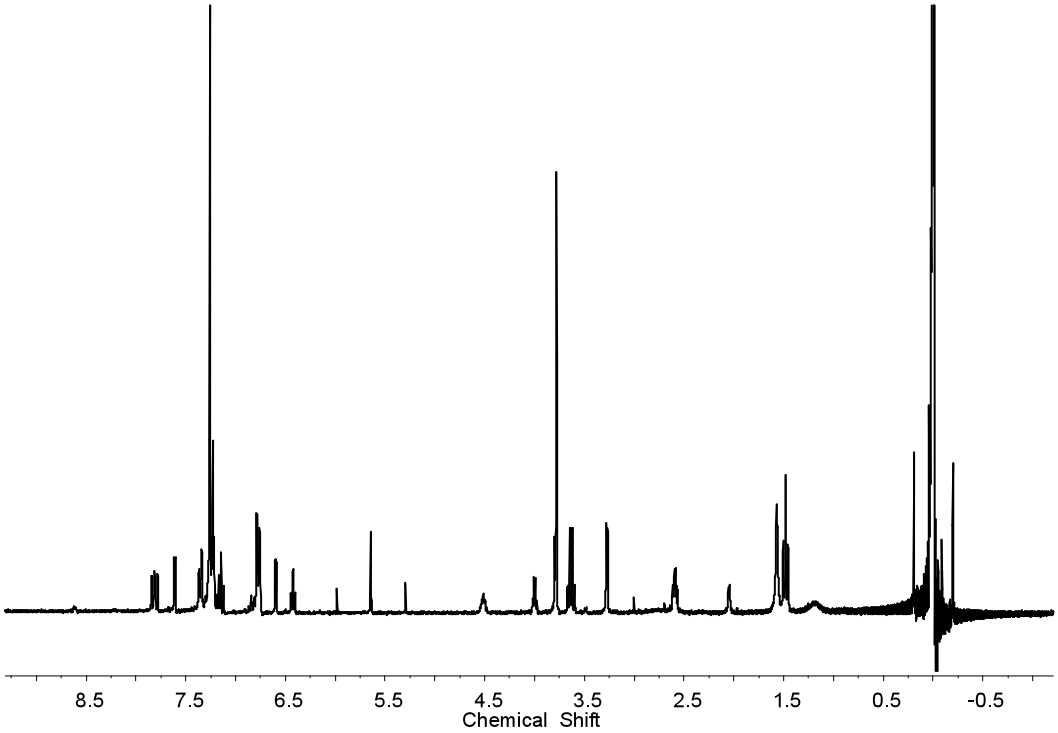


Figure S4-9 ^{1}H NMR spectrum of 5'-dimethoxytrityl-7-nitroindole-2'-deoxynucleoside in $CDCl_3$ (300 MHz, 25 °C)