# SYNTHESIS AND *IN VITRO* REPLICATION STUDIES OF $N^5$-ALKYLATED FORMAMIDOPYRIMIDINE (FAPy-dGuo) ADDUCTS IN DNA

By

Plamen Petkov Christov

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY
in
Chemistry

December, 2007
Nashville, Tennessee

Approved:

Professor Carmelo J. Rizzo

Professor Ned A. Porter

Professor Frederick .P. Guengerich

Professor Michael Stone

Professor Brian O. Bachmann

ACKNOWLEDGEMENTS

I have always considered my studies at Vanderbilt University as a privilege and a great honor. When I look back at the years spent at Vanderbilt University I see years of very productive and intriguing research. I now see myself as having grown mentally and spiritually.

First, I would like to express my deepest gratitude to my advisor Dr. Carmelo J. Rizzo, who has been an excellent teacher and mentor to me. I thank you for allowing me to work in your lab, for always being available to discuss the progress of my research work and for helping me write this thesis. Thank you very much for giving me independence to work on my projects, for listening to my ideas and accepting my suggestions. I also thank you for always being concerned about me and my family and for the numerous cook outs, which I enjoyed very much. I owe a lot to you.

I also would like to thank Drs. Connie and Thomas Harris. Their valuable pieces of advice and excellent scientific discussions over my projects were greatly appreciated.

I also need to thank the members of my Ph.D. committee: Dr. Frederick Guengerich, Dr. Ned Porter, Dr. Michael Stone, and Dr. Brian Bachman. I had all of you as my teachers in different courses and it was a pleasure to listen to your lectures. All of you have contributed greatly to the development of my thesis, by giving good pieces of advice and always willing to find time for my ARM or Ph.D. meetings, although your schedules were extremely busy.

My thanks also go to Dr. Ivan D. Kozekov, who has been helpful in and out of lab. My acquaintance with Dr. Ivan Kozekov dates back to the time when I was an undergraduate student at Sofia University. He was the first one to see something in me and to invite me to do a summer research work in his research group.

I also thank Dr. Mariana Palamareva at Sofia University for all her encouragements and good words while I was an undergraduate student in her lab.

Thank you to all of the current and past members of the Rizzo and Harris groups: Dr. Jimmy Stover, Dr. Eric Elmquist, Dr. Angela Goodenough, Dr. Jozsef Szekely, Dr. Hao Wang, Craig Garmedia, Pam Tamura, Albena Kozekova, and Dr. Katya Petrova, your smiles, support, and kind attitude along the way were appreciated.

I am also thankful to many collaborators here at Vanderbilt: Dr. David Hachey, Dr. Wade Calcutt, Lisa Manier, Dawn Overstreet, Dr. Donald Stec, Markus Voehler, and Kyle Brown.

My special thanks go to my church family at Greater Christ Temple Church with pastor Bishop Sherman L. Merritt. Thank you for being my spiritual home away from my home.

Finally, I would like to thank my mother and father for their care and love. My parents supported me emotionally and financially in Bulgaria, especially during my higher education. Later on, when I told them that I wanted to go

abroad and get my Ph.D. degree, they did accept it regardless of the emotional pain from me being overseas. I am really very blessed to have you in my life.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

Table Page

LIST OF ABBREVIATIONS

| | |
|---|---|
| A | adenosine |
| aq | aqueous |
| $AFB_1$ | aflatoxin $B_1$ |
| $AFB_1$-FAPy-dGuo | aflatoxin-$B_1$-formamidopyrimidine deoxyguanosine lesion |
| BD | butadiene |
| bp | base pair |
| BER | base excision repair |
| C | cytosine |
| CGE | capillary gel electrophoresis |
| COSY | correlation spectroscopy |
| cFAPy-dGuo | carbocyclic analogue of unsubstituted formamidopyrimidine deoxyguanosine lesion |
| CID | collision induced dissociation |
| DNA | deoxyribonucleic acid |
| dNTP | 2'-deoxynucleotide triphosphate |
| dATP | 2'-deoxyadenosine triphosphate |
| dCTP | 2'-deoxycytosine triphosphate |
| dGTP | 2'-deoxyguanosine triphosphate |
| dCyt | deoxycytosine |
| dGuo | deoxyguanosine |
| ds DNA | double-stranded DNA |
| DMTr | 4,4'-dimethoxytrityl |

| | |
|---|---|
| DMSO | dimethyl sulfoxide |
| dNMP | deoxynucleotide monophospahte |
| equiv | equivalent |
| ES-MS | electrospray mass spectrometry |
| $\varepsilon$ | etheno |
| FAPy-dGuo | unsubstituted formamidopyrimidine deoxyguanosine lesion |
| FAPy-dGTP | unsubstituted formamidopyrimidine deoxyguanosine triphosphate |
| FPG | formamidopyrimidine glycosylase |
| GSH | glutathione |
| Guo | guanosine |
| HPLC | high performance liquid chromatography |
| HMBC | heteronuclear multiple-bond correlation (spectroscopy) |
| HCA | heterocyclic aromatic amine |
| LC-ES/MS/MS | liquid chromatography – electrospray mass spectrometry |
| MALDI-TOF | matrix assisted laser desorption ionization – time of flight |
| MS | mass spectrometry |
| MeFAPy-dGuo | methyl-formamidopyrimidine deoxyguanosine lesion |
| NMR | nuclear magnetic resonance |
| NMO | *N*-methylmorpholine-*N*-oxide |
| NER | nucleotide excision repair |
| 8-oxo-dGuo | 8-oxo-deoxyguanosine lesion |
| 8-oxo-dGTP | 8-oxo-deoxyguanosine triphosphate |

| | |
|---|---|
| PAGE | polyacrylamide gel electrophoresis |
| PAH | polycyclic aromatic hydrocarbon |
| PCNA | proliferating cell nuclear antigen |
| PCR | polymerase chain reaction |
| ROS's | reactive oxygen species |
| ss DNA | single-stranded DNA |
| SAM | s-adenosyl-L-methionine |
| TLC | thin-layer chromatography |
| $t_{1/2}$ | half-life |
| Tyd | thymidine |
| TLS | translesion synthesis |
| UV | ultraviolet |
| UDG | uracil DNA glycosylase |

# CHAPTER I

# INTRODUCTION

## Background

Cancer has been affecting humans for thousands of years. The first description of cancer was discovered in an Egyptian papyrus dated between 3000-1500 BC (*1*). It referred to "*tumours*" of the breast. In Ancient Greece in about 400 BC Hippocrates, the "Father of Medicine", was the first to recognize the difference between benign and malignant tumors (*1*).

The understanding, diagnosis and the fight against the cancer was progressing very slowly through the ages. The first cancer hospital was founded in Reims, France, in the 18th century with the belief that cancer was a contagious disease. The French gynecologist Recamier described the invasion of the bloodstream by cancer cells in 1829, coining the word metastasis (cancer spread) (*2*).

It was not until 1953 when Francis Crick and James Watson unraveled the structure of DNA that allowed for a rapid development of the cellular and molecular biology (*3*). Since then the fight against the cancer has been based on a molecular level and has allowed for the development of anticancer therapies.

In modern understanding, cancer is a group of diseases characterized by an uncontrolled growth and spread of abnormal cells. Usually cancer cells form a lump of tissue called a tumor, which can affect the organism severely, for

example by blocking blood vessels, pressing against nerves, or by releasing hormones that can affect the normal functions of the body. If cancerous cells break away from the main tumor and start growing and developing in other parts of the tissue this leads to its spread and the cancer becomes malignant. In order for the cancer cells to spread they must be able to leave their environment and travel through the blood or lymph system, a process called invasion. When cancer cells reach their new location, they must be able to make new blood vessels grow around them and supply them with oxygen and nutrients, a process known as angiogenesis. The whole process of a tumor gaining new mutations as it divides and then spreading to other parts of the body is called metastasis. If the cancer cells divide uncontrollably without being able to spread, they form a benign tumor, which can be removed with surgery.

Cancer is caused by many factors including environment, lifestyle, and heredity. These factors may act together or in a sequence to promote the appearance of cancer. The uncontrolled cell division is usually a result of defects or damage in one or more of the genes involved in cell growth. If the damage in the genes is not repaired it leads to miscoding events, such as mis-pairing, deletion or translocation, which may result in permanent mutations. If these permanent mutations persist in critical regions of genes that are involved in cell division, this leads to multiple events that result in the appearance of cancer.

There are four main types of gene involved in cell division - oncogenes, tumor suppressor genes, suicide genes and DNA-repair genes. Oncogenes are genes that start and accelerate the growth of a cell. If these genes are damaged,

the cells are permanently instructed to divide and can become cancerous. The tumor suppressor genes have an opposite function to the oncogenes, they stop the cell division. One of the most important tumor suppressor genes is p53, which is involved in signaling apoptosis. Most cancer cells have mutation in this gene. The suicide genes control apoptosis, or cell suicide. Cells are usually able to commit suicide whenever something goes wrong in order to prevent damage to the neighboring cells or passing on damaged genetic material. If these suicide genes become damaged, then a faulty cell can keep dividing and become cancerous. DNA-repair genes are encoded proteins whose function is to repair damaged DNA. If these genes are damaged, cell loses its ability to repair damaged DNA. DNA damage can then accumulate in the cell, leading to further mutations and ultimately cancer.

## DNA Damage and Its Consequences

### *Environmental and Chemical Carcinogenesis*

Humans are continuously exposed to exogenous and endogenous chemicals that have been shown to induce DNA mutations and cancer. Exposure to these chemicals occurs because they are present in food, air or water or they are products of cell metabolism.

In 1761, Dr. John Hill first linked the development of nasal cancer in people who used tobacco snuff excessively (*4*). It was 14 years later when Sir Percival Pott related the appearance of scrotal cancer in chimney sweeps to their occupation (*5*). In 1915, Yamagiwa and Ichikawa were the first to conduct experimental carcinogenesis (*6*). They repeatedly painted coal tar on the ears of rabbits and succeeded in producing multiple squamous cell carcinomas in the painted areas. Their experiments led to the discovery of a pure carcinogenic chemical, 1,2,5,6-dibenzanthracene. Several years later PAH's were discovered as the active carcinogens in tar by two independent research groups (*7, 8*).

By the end of 1970's many chemicals were established as carcinogens and a question was raised of what was the mechanism of action of these carcinogens in genotoxicity. Of particular importance was the work by the Millers (*9*) and others (*10*), who showed that most of the carcinogens were metabolized or bio-activated to more potent species (metabolite), which were called the "ultimate carcinogens". The "ultimate carcinogen", in its initial perception, was an electrophile that was formed *in vivo,* had finite stability, and could diffuse limited

distances to covalently modify DNA. This early perception of ultimate carcinogen led to the idea that carcinogens can be only activated *in vivo* by nuclear enzymes (*11*). Years later, it was shown that metabolites that react with DNA could be generated in hepatocytes and even trapped outside the cells (*12*). Moreover, it was established that enzymes that are involved in the detoxication of cells could activate chemical carcinogens. For example, gluthathione transferase can activate haloalkenes and 1,2-dihaloethanes, and P450 enzymes can epoxidize vinyl halides and polycyclic hydrocarbons by converting them to more reactive electrophiles (*13*).

In the last 25 years the field of chemical toxicology has developed tremendously due to the development of more sophisticated experiments and methods for analysis of the action of carcinogens *in vivo* and *in vitro*. Today's science relies on the use of genome modified systems, molecular biology, chemical synthesis of carcinogens, nuclear magnetic resonance spectroscopy (NMR), mass spectrometry (MS), protein crystallography, and other techniques to answer the questions of how the carcinogens actually induce cancer. It is considered that the reaction of carcinogens with DNA is the one of the earliest events of development of cancer. Many carcinogens react with DNA to form different covalent complexes referred as DNA adducts. If these DNA adducts are not repaired prior replication, they can cause mutations in the genes involved in important cellular functions as was discussed earlier.

### *Formation of Formamidopyrimidine-dGuo (FAPy-dGuo) Lesions via Oxidation*

DNA damage can be generated by the reaction of DNA with reactive oxygen species (ROS's), such as superoxide radical anion, hydrogen peroxide, and hydroxyl radical (*14-17*). These ROS's are products of normal cellular metabolism. Electron transport chains have a potential to "leak" electrons to oxygen which results in the formation of superoxide (*18*). Superoxide is also formed by certain enzymes. For example, phagocytes release ROS's in order to destroy cells infected with viruses or bacteria. The ROS's can react with the DNA resulting in the formation of strand breaks, DNA-protein cross-links, abasic sites and more then 20 oxidized base lesions including 8-oxo-dGuo (**1.03**) and FAPy lesions derived from dAdo and dGuo.

The FAPy-dGuo (**1.04**) lesion is formed from the addition of hydroxyl radical to the double bond of C8 position of dGuo moieties (**1.01**) in DNA giving an intermediate guanyl radical (**1.02**) (Figure **1-01**). One-electron oxidation of guanyl radical (**1.02**) leads to the formation of 8-oxo-dGuo (**1.03**). A competitive pathway is the one-electron reduction of this intermediate guanyl radical (**1.02**) followed by ring-opening resulting in a FAPy-dGuo (**1.04**) lesion. In general, 8-oxo-dGuo (**1.03**) is the most abundantly formed DNA lesion; however, there is at least one report concerning human leukemia cells in which the ratio FAPy-dGuo (**1.04**) : 8-oxo-dGuo (**1.03**) is almost 3 (*19*).

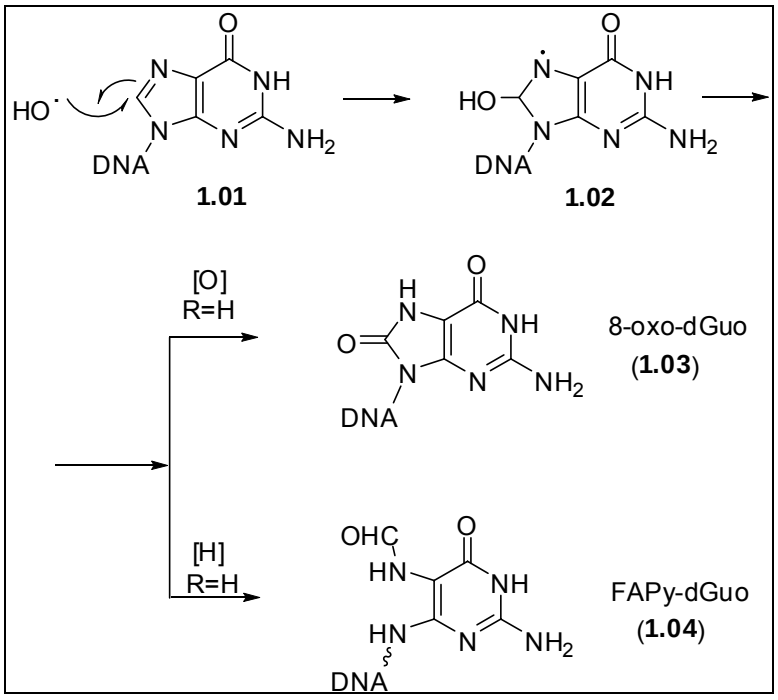



**Figure 1-01**. Formation of 8-oxo-dGuo (**1.03**) and FAPy-dGuo (**1.04**) lesions via oxidation of dGuo moieties (**1.01**) in DNA.

Studies on the mutagenicity and the repair of both 8-oxo-dGuo (**1.03**) and FAPy-dGuo (**1.04**) lesions have been conducted. The 8-oxo-dGuo (**1.03**) lesion is miscoding and results in GC→TA transversions, which have been observed *in vivo* in the *ras* oncogene and the *p53* tumor suppressor gene in lung and liver cancers (*18*). The reason for the GC→TA transversions is that 8-oxo-dGuo (**1.03**) can adopt a *syn* conformation during replication or transcription of DNA, which can cause mis-pairing with dATP through Hoogsteen base pairing (Figure **1-02**) (*20*).

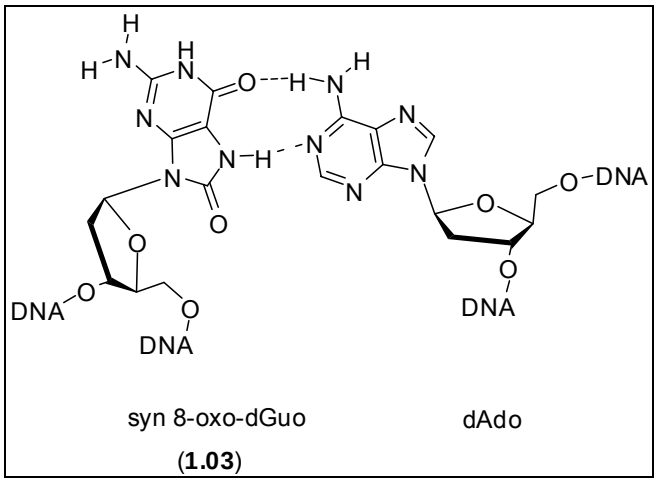



**Figure 1-02**. Formation of Hoogsteen base pairing between 8-oxo-dGuo (**1.03**) and dAdo.

Many *in vitro* replication studies of the FAPy-dGuo adduct (**1.04**) have been conducted using Klenow fragment from *Escherichia coli* DNA polymerase I exo$^-$ (Kf$^-$). The FAPy-dGuo adduct (**1.04**) was efficiently bypassed and dATP was preferentially mis-incorporated opposite of the adduct (*21*).

Interestingly, repair studies of FAPy-dGuo (**1.04**) showed that this adduct is excised by formamidopyrimidine DNA glycosylase (FPG) much more efficiently opposite dCyd than when it mis-pairs with dAdo (*22*). Several reasons were given to explain this observation. First, it was observed that FAPy-dGuo:dCyd lesion is twenty-fold more tightly bound than the FAPy-dGuo:dAdo lesion to the FPG protein, which suggested that the FAPy-dGuo lesion is preferentially recognized when it is paired with dCyd rather then with dAdo. Second, if the FAPy-dGuo:dAdo is repaired by FPG, this would allow the retention of the mis-incorporated dAdo resulting in dGuo→Tyd transversion, an outcome that does not preserve the genomic integrity of the organism.

Structural insight in the binding and recognition of FAPy-dGuo's by the FPG protein was provided by the work of Coste and Casting (*23*). Their work revealed that cFAPy-dGuo (G*), a carbocyclic analogue of FAPy-dGuo (Figure **1-03)** that does not undergo anomerization, is bound to FPG protein in the *anti* conformation in contrast to the 8-oxo-dGuo (**1.03**), which is bound in a *syn* conformation (*20*). The binding of the cFAPy-dGuo (G*) is established by interaction of all the functional groups of the cFAPy-dGuo base with the FPG protein (Figure **1-03**).

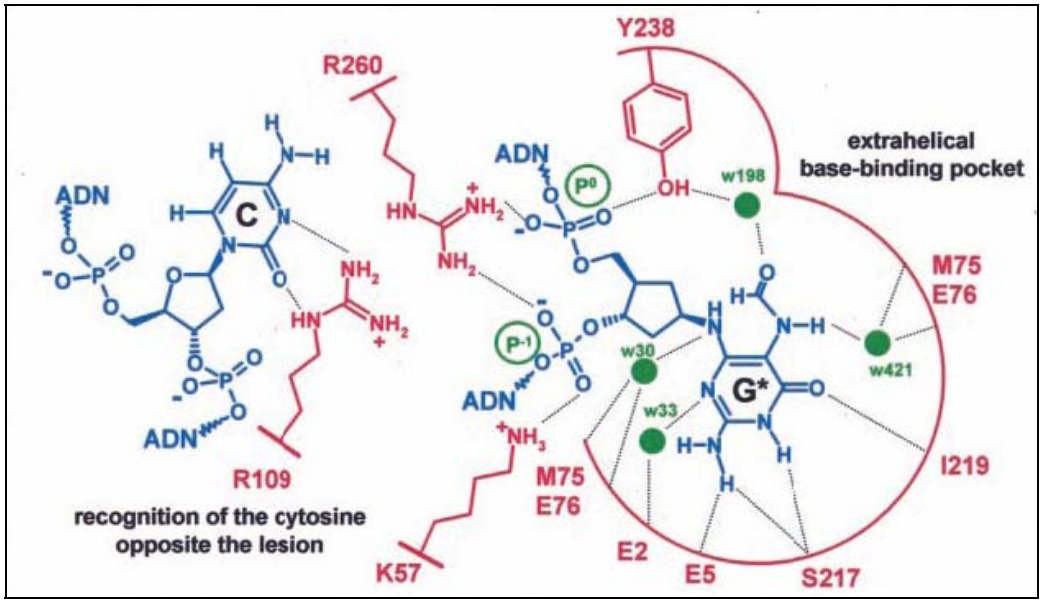



**Figure 1-03**. Interactions between the extrahelical *c*FAPy-dGuo (G*) and FPG residues inside the active site binding pocket: amino acid residues of *LI*-FPG involved in the recognition are shown in *red* and the DNA are shown in *blue*; *C*, the cytosine opposite *c*FapydGuo (*G**); $p^0$ and $p^{-1}$ indicate the phosphate groups bordering the lesion; *small green circles* represent the water mediated interactions; *black dashed lines* represent hydrogen bond interactions (*23*).

The pyrimidine moiety of the cFAPy-dGuo (G*) is involved in the Watson-Crick face recognition and contribute to the proper orientation of the pyrimidine ring in the binding pocket. The functional groups, derived from the ring-opening of the imidazole ring of dGuo, are suggested to play an important role in the rotameric selection of FAPy-dGuo's.

The first repair studies of FAPy-dGuo (**1.04**) in mammalian cells were reported by Bohr and Greenberg (*24*). The source of the mammalian cells were mitochondrial and nuclear extracts from the wild type and knock-out mice lacking one of the major DNA glycosylases for repair of oxidative DNA damage - OGG1. The repair studies of FAPy-dGuo (**1.04**) with the mitochondrial and nuclear extract from the wild type mouse showed that this lesion is efficiently repaired by OGG1, while the repair studies of FAPy-dGuo (**1.04**) with the mitochondrial and nuclear extract from the knock-out mice showed that FAPy-dGuo (**1.04**) was still repaired but with a low efficiency. It was suggested that in the absence of OGG1, which repairs 8-oxo-dGuo (**1.03**) in mouse liver mitochondria (*25*), another enzyme NEIL1 (human homologue of *E. coli* endonuclease VIII) may function as a backup glycosylase for FAPy-dGuo (**1.04**). Interestingly, NEIL1 does not participate in the repair 8-oxo-dGuo (**1.03**) in mouse liver mitochondria (*26*).

The repair of both 8-oxo-dGTP (**1.05**) and FAPy-dGTP (**1.06**) was investigated with the MTH1 protein, which is a mammalian homologue of *E. coli* MutT and is involved in the hydrolysis of 8-oxo-dGTP (**1.05**) in the nucleotide pools (Figure **1-04**). The experiments showed that the FAPy-dGTP (**1.06**) was hydrolyzed more then ten-fold less efficiently then 8-oxo-dGTP (**1.05**). The poor

recognition of FAPy-dGTP (**1.06**) by MTH1 might be due to the loss of the two-ring system (purine), which is necessary for the stacking interaction between the substrate and the enzyme, and the presence of several FAPy-dGuo (**1.06**) isomers one of which could be an inhibitor (*27*).

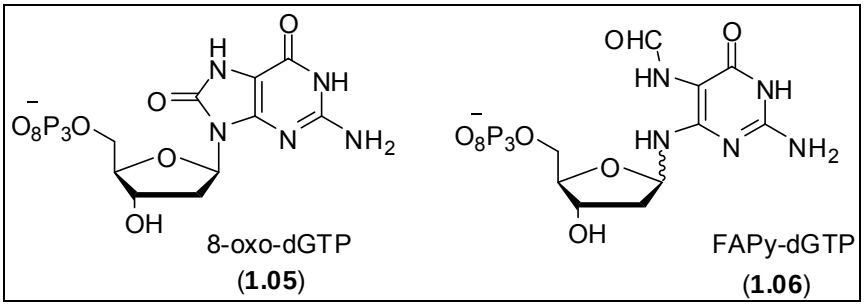



**Figure 1-04**. Structures of the 8-oxo-dGTP (**1.05**) and FAPy-dGTP (**1.06**).

Since FAPy-dGTP (**1.06**) is a poor substrate for MutT, a hypothesis was proposed that FAPy-dGTP (**1.06**) could be produced in the nucleoside triphosphate pool as a result of oxidative stress and become a source of FAPy-dGuo (**1.04**) adduct in duplex DNA (*28*). Experiments conducted with Klenow fragment showed that FAPy-dGTP (**1.06**) is incorporated opposite dC 1000 times less efficiently than is dGTP, but calculations were made that a FAPy-dGTP (**1.06**) concentration of 1% that of the dGTP concentration would yield several FAPy-dGuo lesions (**1.04**) per million base pairs of DNA, which is sufficient to give the levels of detected FAPy-dGuo (**1.04**) in cellular DNA (*29, 30*).

In mammalian cells, (cos-7), FAPy-dGuo (**1.04**) is mutagenic and induces FAPy-dGuo→T transversions (*31*). The levels of the mutational frequency were sequence dependant. For example, Fapy-dGuo (**1.04**) gave 30 % mutation in 5'-T-(FAPy-dGuo)-T context compared to 8% in 5'-T-(FAPy-dGuo)-A context.

### *Formation of $N^5$-Alkylated Formamidopyrimidine-dGuo Lesions by Alkylating Agents*

DNA can also be damaged by alkylating agents (electrophiles), which react with the nucleophilic centers, such as oxygen or nitrogen atoms of the DNA base or backbone. When these agents react with the oxygen atoms of the DNA phosphodiester backbone, a phosphotriesters is formed (*32, 33*) and when they react with the DNA bases, DNA adducts are formed. The DNA adduct together with the opposite base is usually called a lesion.

The N7 position of the dGuo is generally regarded as the most nucleophilic site of DNA (**1.01**) and upon alkylation results in the formation of an N7-alkyl cationic dGuo intermediate (**1.07**). The N7-alkyl cationic species (**1.07**) can readily depurinate resulting in an abasic sites (**1.08**) in DNA (*34*). Abasic sites are strong blocks to replication, but when bypassed they induce the misincorporation of dAdo (*35, 36*). Alternatively, addition of hydroxide to the C8-position of the N7-cationic dGuo intermediate (**1.07**) followed by ring-opening gives raise to 2,6-diamino-4-hydroxy-$N^5$-alkyl-formamidopyrimidine ($N^5$-alkylated FAPy-dGuo) (**1.09**) lesion (Figure **1-05**).

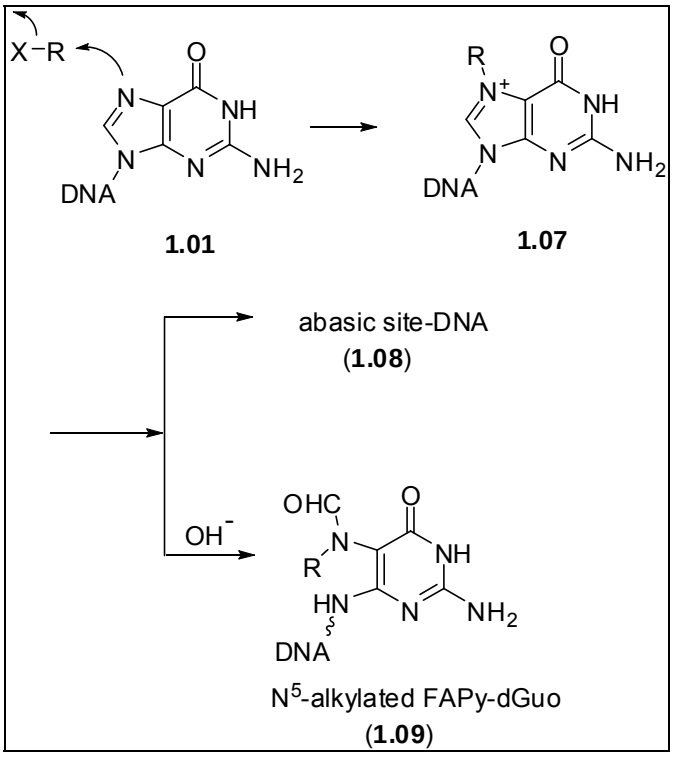



**Figure 1-05**. Alkylation of dGuo moieties (**1.01**) in DNA and formation of abasic site (**1.08**) and $N^5$-alkyated FAPy-dGuo lesion (**1.09**).

An example of a $N^5$-alkylated FAPy-dGuo adduct (**1.09**) that forms as a results of the N7 dGuo alkylation is aflatoxin $B_1$ (**1.10**), which is produced by several fungal species that contaminate food supplies (*37*). Aflatoxin $B_1$ is metabolized by a cytochrome P450 to aflatoxin $B_1$-8,9-exoepoxide (**1.11**), the ultimate carcinogenic species (*38*). This aflatoxin $B_1$ epoxide intercalates into DNA and then alkylates the N7 position of dGuo moieties to form an aflatoxin $B_1$ N7-cationic dGuo species (*39*). Base catalyzed ring-opening of the imidazole ring produces trans-8,9-dihydro-8-(2,6-diamino-4-oxo-3,4-dihydropyrimid-5-yl-formamido)-9-hydroxyaflatoxin $B_1$ ($AFB_1$-FAPy-dGuo) as a mixture of isomers

(**1.12** and **1.13**) (Figure **1-06**) (*40*). Site-specific mutagenesis showed that the $AFB_1$-FAPy-dGuo adduct(s) was more mutagenic that the N7-cationic dGuo species when replicated in *Escherichia coli* using a single-stranded DNA replication vector giving high levels of dGuo → dTyd transversions (*34*).

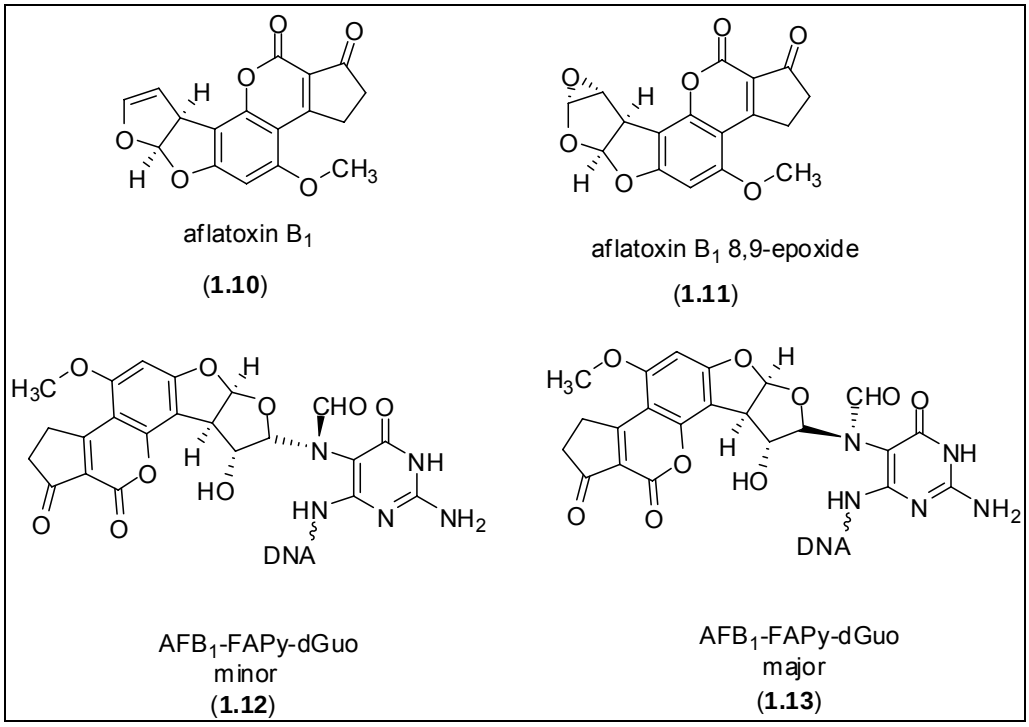



**Figure 1-06**. Structures of aflatoxin $B_1$ (**1.10**), aflatoxin $B_1$ 8-9-exoepoxide (**1.11**) and the two $AFB_1$-FAPy-dGuo isomers (**1.12** and **1.13**).

$N^5$-Alkylated FAPy-dGuo adduct could be also formed from butadiene (BD) (**1.14**), which is a well-established carcinogen in rodents and humans (*41*). BD (**1.14**) is epoxidized in two steps by cytochrome P-450 to form 3,4-epoxybutane (**1.15**) and 1,2,3,4-diepoxybutane (**1.16**), which are the ultimate carcinogenic species (Figure **1-07**). These epoxides react with DNA to form several adducts of dGuo and dAdo moieties in DNA (*42*). In addition to those adducts, Elfara and coworkers reported the formation of an adduct **1.19** that could be formed via an $N^5$-alkylated FAPy-dGuo intermediate (**1.18**) (*43*) (Figure **1-08**).

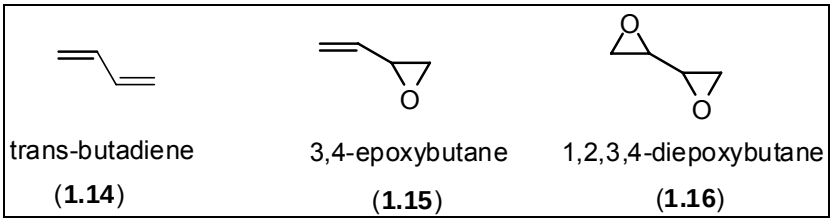



**Figure 1-07**. Structures of trans-butadiene (**1.14**), 3,4-epoxybutadiene (**1.15**) and 1,2,3,4-diepoxybutadiene (**1.16**).

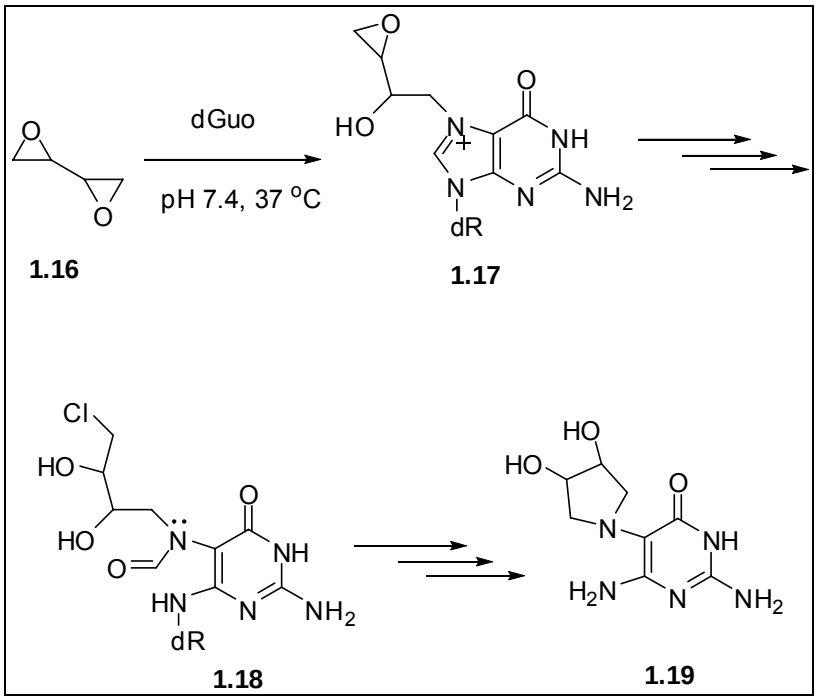



**Figure 1-08**. Formation of adduct **1.19** via the $N^5$-alkylated FAPy-dGuo intermediate **1.18**.

There are alkylating agents that can react directly (without metabolic activation) with dGuo moieties in DNA to form $N^5$-alkylated FAPy-dGuo lesions. In this category of alkylating agents are nor-nitrogen mustard (**1.20**) (*44*), aziridine (**1.21**) (*45, 46*), alkyl halides (**1.22**-**1.26**) (*47-51*) and mitomycin C (**1.27**) (Figure **1-09**) (*52*). The most biological important $N^5$-alkylated FAPy-dGuo lesions formed by the above mentioned alkylating agents are the mitomycin FAPy-dGuo (**1.28**) nor-nitrogen mustard FAPy-dGuo (**1.29**), and MeFAPy-dGuo (**1.30**) lesions, formed from mitomycin C (**1.27**), nor-nitrogen mustard (**1.20**) and methyl halides (**1.24**-**1.26**), respectively (Figure **1-10**).

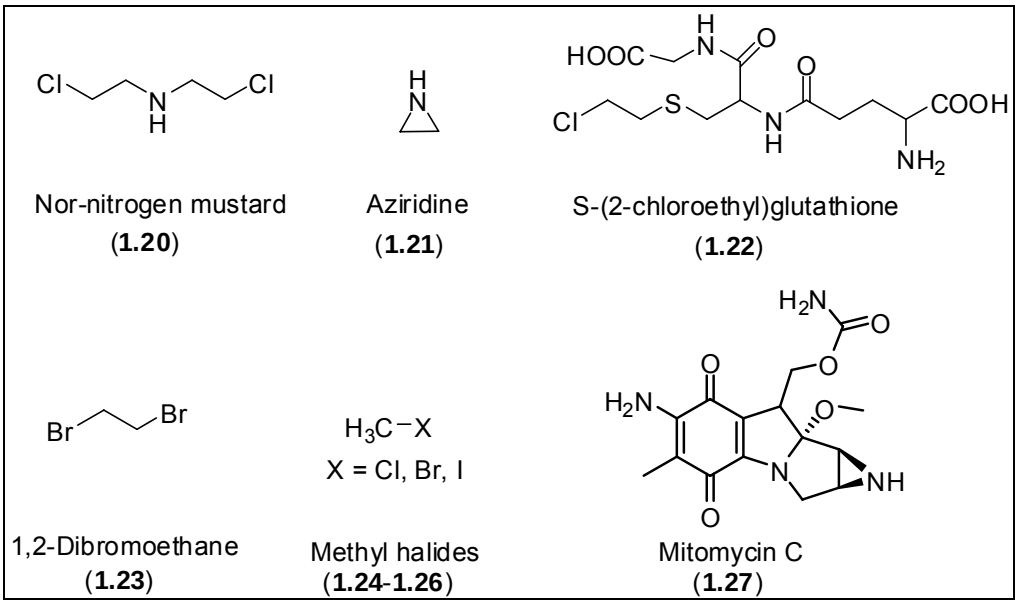



**Figure 1-09**. Structures of nor-nitrogen mustard (**1.20**), aziridine (**1.21**), alkyl halides (**1.22** - **1.26**) and mitomycin C (**1.27**) that are known to form $N^5$-alkylated FAPy-dGuo adducts.

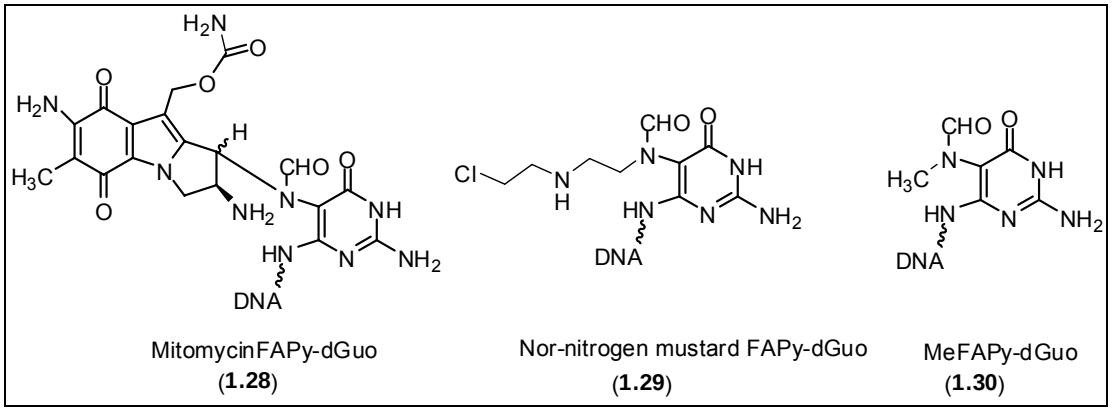



**Figure 1-10**. Structures of mitomycin FAPy-dGuo (**1.28**), not-nitrogen mustard FAPy-dGuo (**1.29**) and MeFAPy-dGuo (**1.30**) adducts.

Mitomycin C (**1.27**), isolated from *Streptomyces caespitosus*, is a clinical used antitumor agent used for the treatment of bladder, rectal, pancreatic, lung and breast cancers. It reacts with DNA *in vivo* and in *vitro* resulting in the formation of several DNA adducts including mitomycin FAPy-dGuo (**1.28**) as a minor product. Interestingly this FAPy-dGuo adduct (**1.28**) is formed in slightly acidic conditions (*52*).

The other important biological $N^5$-alkylated FAPy-dGuo lesion can be formed via alkylation of dGuo moieties in DNA by nor-nitrogen mustard (**1.20**). This alkylating agent is a metabolite of a family of nitrogen mustard chemotherapeutic agents, such as cyclophosphamide (**1.31**). The cyclophosphamide (**1.31**), used to treat a variety of cancers, is metabolized to the cytotoxic alkylating agent phosphoramide mustard (**1.34**), which can be dephosphoramidated to give nor-nitrogen mustard (**1.20**) (Figure **1-11**).

The DNA alkylation chemistry of nor-nitrogen mustard (**1.20**) is complex. It forms monoalkylated adducts at a number of different sites on DNA including the N7 position of dGuo, which can undergo hydroxide ring-opening to form the nor-nitrogen mustard FAPy-dGuo adduct (**1.29**).

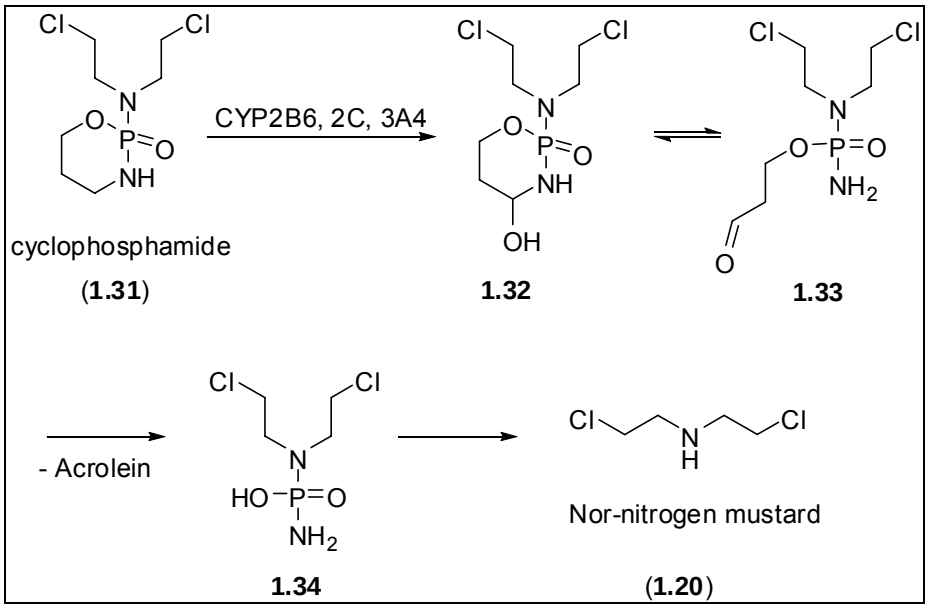



**Figure 1-11**. Formation of nor-nitrogen mustard (**1.20**) from the metabolic activation of cyclophosphamide (**1.31**).

Methyl halides (**1.24** - **1.26**) and other methylating agents, such as dimethyl sulfate (**1.35**), methylmethane sulfonate (**1.36**), N-methyl-nitrosourea (**1.37**), N,N-dimethyl-N-nitrosoamine (**1.38**) (Figure **1-12**), react with dGuo moieties of DNA to form preferentially N7 and $O^6$-guanine adducts. The $O^6$-methylguanine adduct incorrectly pairs with dTTP during DNA replication, leading to a GC→AT base transition mutation (*53*). The N7-methyl dGuo cationic species undergo imidazole ring-opening, analogues to N7-aflatoxin $B_1$ dGuo cationic adduct, in neutral or basic conditions to generate 2,6-diamino-4-hydroxy-$N^5$-methyl-formamidopyrimidine (MeFAPy-dGuo) lesions (**1.30**) (*54*). This important DNA lesion will be a main focus in Chapter II of this dissertation.

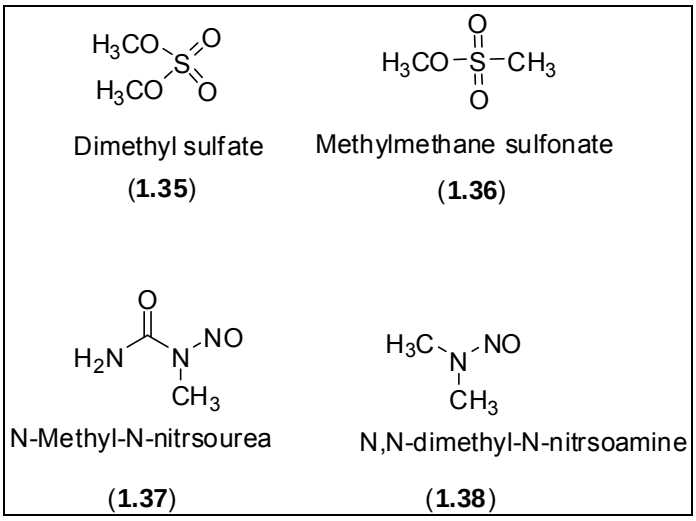



**Figure 1-12**. Structures of dimethyl sulfate (**1.35**), methylmethane sulfonate (**1.36**), N-methyl-N-nitrosourea (**1.37**) and N,N-dimethyl-N-nitrosoamine (**1.38**).

The repair of the $N^{5-}$alkylated FAPy-dGuo lesions has not been well studied due to the lack of convenient pathways for the preparation of oligonucleotides that contain these lesions at define location. Thus far, there have been several reports regarding the repair of $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) and MeFAPy-dGuo lesions (**1.30**) in which ss-DNA containing a single dGuo residue was alkylated with aflatoxin $B_1$ 8,9-epoxide (**1.11**) or dimethyl sulfate (**1.35**), respectively (*55, 56*). The MeFAPy-dGuo (**1.30**) containing oligonucleotides also were obtained via site-specific incorporation of N7-methyl dGTP by PCR (*57, 58*). The syntheses of oligonucleotides containing $AFB_1$-FAPy-dGuo (**1.12**, **1.13**) and MeFAPy-dGuo (**1.30**) lesions by the afore-mentioned methods are not convenient and have many restrictions, which will be discussed in Chapter II of this dissertation.

The repair studies on the MeFAPy-dGuo (**1.30**) lesion usually with respect to the repair of 8-oxo-dGuo (**1.03**), have been performed (*58*). In general, the DNA repair enzymes that repair the 8-oxo-dGuo (**1.03**) adduct recognize and repair FAPy-dGuo (**1.04**) and MeFAPy-dGuo (**1.30**) adducts. The DNA repair enzymes that are used in these studies are the *E. coli* FPGMutM, a glycosylase removing 8-oxo-dGuo (**1.03**) from 8-oxo-dGuo:dCyt pairs; MutY, a monofunctional glycosylase that excises dAdo form 8-oxo-dGuo:dAdo mispairs; and MutT, a phosphatase that degrades the triphosphate of 8-oxo-dGuo (**1.05**). Human homologues of the FPG, MutY, and MutT have been cloned and also used for the repair studies of MeFAPy-dGuo (**1.30**). Experiments with human MutT and FPG showed distinct preferences towards 8-oxo-dGuo (**1.03**) of 20 and 12-fold, respectively, between the most (dGuo) and the least preferred base (dCyd) opposite the lesion. Such preferences were not observed for the MeFAPy-dGuo (**1.30**) adduct, which suggested that MeFAPy-dGuo (**1.30**) is most likely not subjected to complementary base-dependant repair. In contrary to this suggestion, experiments with prokaryotic and mammalian thymine glycol (Tg) glycosylases including *E. coli* endonuclease III (endonuclease III), endo VIII (endonuclease VIII) and the mouse endo III homologue showed that the recognition of MeFAPy-dGuo (**1.30**) was pair dependant, the highest being with dGuo and the lowest with dCyt (*57*). An explanation for this observation is the formation of hydrogen bonds between the 8-oxo-dGuo (**1.03**) or MeFAPy-dGuo (**1.30**) and dCyt in analogy to this in the G:C base pair, which does not cause flipping out of the corresponding DNA adducts. The flipping out of the DNA

adducts is suggested to be one of the reasons how the repair enzymes recognize the DNA damage.

There are several reports on the repair of $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) lesion in comparison to either the $AFB_1$-$N^7$-dGuo (N7-afaltoxin $B_1$ alkylated dGuo) or MeFAPy-dGuo (**1.30**) lesions. It was demonstrated that *E. coli* uvrABC enzyme, which exhibits nucleotide excision repair (NER), can incise both $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) and $AFB_1$-$N^7$-dGuo adducts with equal efficiency *in vitro (56)*; however in another study the FPG protein, which exhibits BER as well, was only able to remove $AFB_1$-FAPy-dGuo (**1.12**, **1.13**) adduct but not the cationic $AFB_1$-$N^7$-dGuo adducts (*59*).

Essigmann and coworkers investigated the *in vitro* and *in vivo* repair of $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) lesion. The *in vitro* repair of $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) lesion was compared to the *in vitro* repair of 8-oxo-dGuo (**1.03**) and MeFAPy-dG (**1.30**) lesions by MutM and hOGG1, and the *in vivo* repair of $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) lesion was compared to the *in vivo* repair of $AFB_1$-$N^7$-dGuo lesion (*55*). The *in vitro* repair experiments showed that the 8-oxo-dGuo (**1.03**) and MeFAPy-dGuo (**1.30**) adducts were excised, whereas the $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) adduct was comparatively less efficiently excised. This observation was contrary to what was previously observed for $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) by Chetsanga (*59*) and Oleykowski (*56*) probably due to the use of shorter oligonucleotides, which may not have been optimal for FPG and hOGG1 enzymes.

The *in vivo* repair experiments with the $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) and $AFB_1$-$N^7$-dGuo lesions were conducted with wild-type, nucleotide excision repair (NER) deficient, BER-deficient (MutM) and NER/BER double deficient mutants of *E. coli*. The repair experiments showed that both $AFB_1$-FAPy-dGuo (**1.12** and **1.13**) and $AFB_1$-$N^7$-dGuo lesions are preferentially repaired by NER (*55*).

## DNA Replication

DNA replication is needed to pass on genetic information from the parental cells to daughter cells. The DNA replication occurs in a semi-conservative fashion where each strand of the parental duplex serves as a template for the newly generated strands. The DNA replication begins with the binding of T-antigen (helicase) to the replication origin and causes a partial unwinding of the double helix at an area known as the “replication fork” (Figure **1-13**) (*60, 61*). As the DNA is unwound specific single stranded DNA binding proteins (replication factor A) prevent the strands from re-annealing. The replication fork is a Y-shaped structure that moves steadily along the parental DNA duplex. The replication of DNA is performed by DNA polymerases, which catalyze the synthesis of DNA in the 5’ to 3’ direction. The two strands of the DNA double helix are not replicated in the same manner, because they run anti-parallel to one another. The lagging strand is synthesized in the 5’ to 3’ direction by a series of short Okazaki fragments. These fragments are stretches of 100 to 200 nucleotides in humans (1000 to 2000 in bacteria) and are stitched together by a DNA ligase to create a continuous strand.

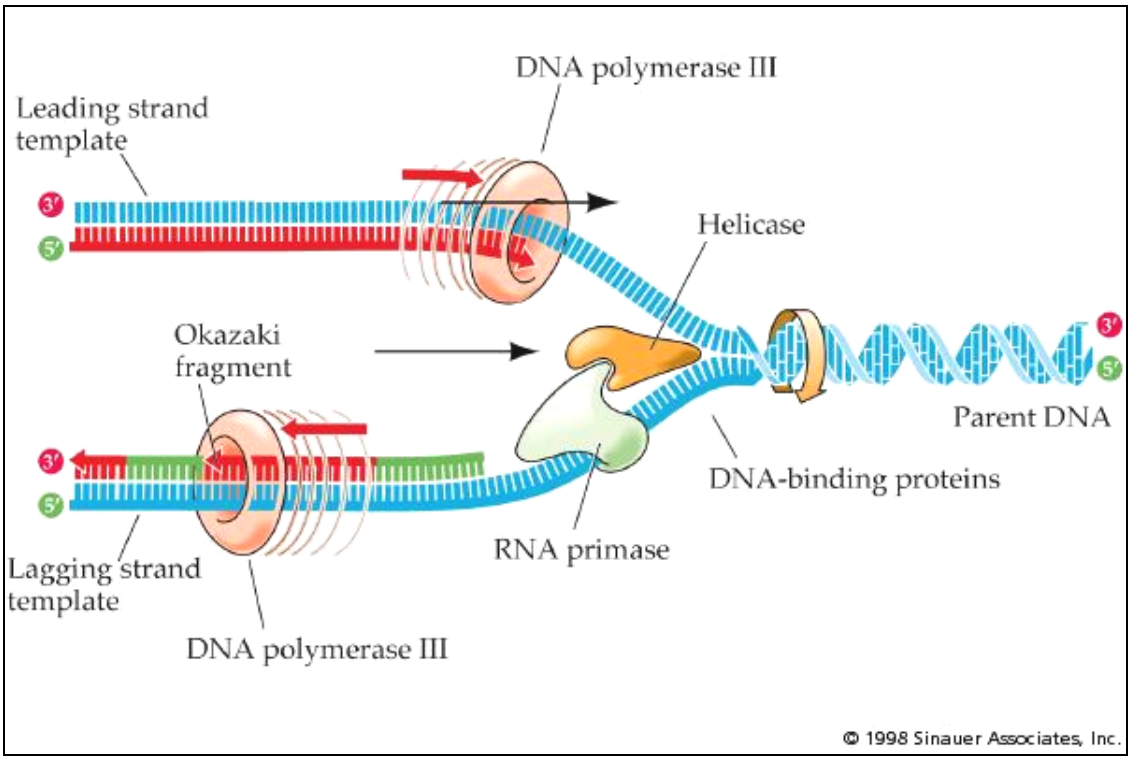



**Figure 1-13**. Illustration of DNA replication fork (*37*).

## *DNA Polymerases*

*Replicative Enzymes*

DNA replication is catalyzed by replicative enzymes that form a bond between the 5'-phosphate of the incoming nucleotide (the other two phosphate groups from the nucleotide triphosphate serve as a leaving group during phosphodiester formation) and the 3'-OH group of the nucleotide at the end of the growing DNA chain with high fidelity. The error rate of these enzymes is about one in $10^8$ – $10^{12}$ nucleotide copying events. This low rate of mis-

incorporation is achieved by several factors. The active sites of these enzymes are tight and do not allow incorporation of an incorrect nucleotide. Another hypothesis states that DNA polymerase may also discriminate against the incorrect triphosphate by an induced-fit conformational change (*62, 63*). After incorporation of the correct base, the replicative enzyme undergoes conformational changes in the active site to bring certain residues into alignment to provide transition-state stabilization for the catalysis of phosphodiester bond formation. Following a mis-incorporation event, the enzyme selects against the resultant mismatched DNA by a slow, rate-limiting chemistry step (Figure **1-14**). This process allows the polymerase to release the mismatched DNA from its active site by excision. In addition, most replicative enzymes possess this intrinsic 3' to 5'-exonuclease activity, or proofreading capabilities.

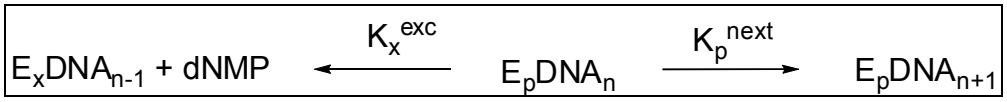



**Figure 1-14**. Illustration of the induced-fit mechanism for exclusion of mis-incorporated dNTP: $E_pDNA_n$ - enzyme-DNA complex, $E_pDNA_{n+1}$ - enzyme-DNA complex with the next correct dNTP, $E_xDNA_{n-1}$ - enzyme-DNA complex with the excluded mismatched base.

Based of sequence similarities, DNA polymerases are classified in three groups: type A, type B and type C, which have homology to *pol A* (pol I), *pol B* (pol II) and *pol C* (pol III) from *E. coli,* respectively. The eukaryotic polymerase $\delta$ belongs to type B. The role of polymerase $\delta$ in replication however is not clear

although there is evidence that it participates in post-replicational DNA repair (*64*).

*Translesion Synthesis (Bypass) Polymerases*

The translesion (bypass) DNA polymerases are able to synthesize past DNA lesions that block DNA replication by replicative DNA polymerases (*65, 66*). These enzymes are classified in two groups - the Y-family and B-family. The human Y-family consists of human pol $\eta$, $\kappa$ and $\iota$ and Rev1, while human pol $\xi$ is a part of the B-family. Several other bypass DNA polymerases were also discovered, but their role in translesion synthesis has not been very well established. The bypass DNA polymerases have low fidelity ($10^3$-$10^4$) that results in the mis-incorporation of nucleotides. One of the reasons for this is that the active site is larger than those of replicative polymerases. For example, the active site of the Y-family polymerase Dpo4, isolated from archaeal aerobic thermophile *Sulfolobus solataricus* can allow the accommodation of two nucleotides, which results in the misincorporation of the same nucleotide across the adduct and the first base past the adduct, respectively. In contrast to the replicative polymerases, the Y-family DNA polymerases lack any intrinsic exonucleolase activity and replicate DNA in a distributive manner where the enzyme dissociates from the DNA after the addition of only a few nucleotides. It was observed that proliferating cell nuclear antigen (PCNA) stimulates the efficiency of nucleotide incorporation by Y-family polymerases opposite undamaged as well as damaged templates (*67*). The exact mechanism of the

action of PCNA is not known, but it was suggested that monoubiquitination of PCNA disrupts the binding of the replicative polymerases to the template-primer complex, which enables the binding of the TLS polymerase to PCNA and to the other components of the replication ensemble (*68, 69*).

## Dissertation Aims

The aim of this dissertation is to study the chemistry of two important DNA lesions, MeFAPy-dGuo (**1.30**) and 2-oxoethyl-FAPy-dGuo (**1.39**) (Figure **1-15**).

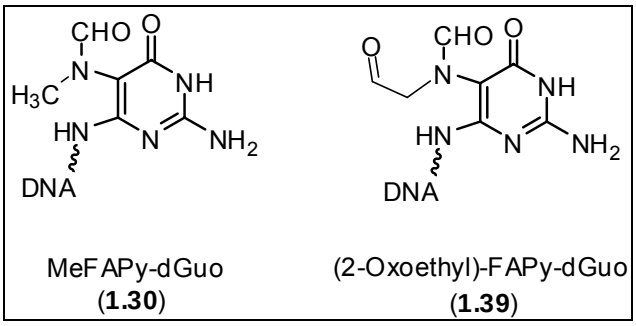



**Figure 1-15**. Structures of the MeFAPy-dGuo (**1.30**) and 2-oxoethyl-FAPy-dGuo (**1.39**) lesions.

In Chapter II, an efficient four step synthesis of the MeFAPy-dGuo phosphoroamidite was developed, which allowed the preparation of oligonucleotides containing the MeFAPy-dGuo lesion (**1.30**). The aims of Chapter II are:

1. Synthesis of MeFAPy-dGuo phosphoroamidite reagent required for solid-phase synthesis of MeFAPy-dGuo (**1.30**) containing oligonucleotides.
2. NMR structural determination of the furanose and pyranose forms of MeFAPy-dGuo adduct (**1.30**) in trinucleotides.
3. Examination of the mis-incorporation profile of MeFAPy-dGuo (**1.30**) catalyzed by prokaryotic (Klenow fragment exo$^-$ of *E. coli*, pol II$^-$, and Dpo4) and eukaryotic (human pol $\eta$, $\kappa$ and $\iota$) DNA polymerases.
4. Determine the steady-state kinetics for the base incorporation opposite the MeFAPy-dGuo adduct (**1.30**) catalyzed by prokaryotic and eukaryotic DNA polymerases.
5. Examination of the full-length extension products past the MeFAPy-dGuo lesion (**1.30**) catalyzed by prokaryotic and eukaryotic DNA polymerases. The sequence of the full-length extension products was determined by an LC-ESI/MS/MS method.

In Chapter III, it was hypothesized that (2-oxoethyl)-FAPy-dGuo (**1.39**) adduct could be formed from the N7-dGuo alkylation of DNA with 2-chlorooxirane and related oxiranes, such as the other halooxiranes, 2-acetoxyoxirane and 2-cyanooxirane, followed by hydroxide ring-opening of the corresponding N7-alkylated dGuo cationic species. The (2-oxoethyl)-FAPy-dGuo adduct (**1.39**) was synthesized, its spectral properties characterized, and used as a standard to identify its formation in the reaction of 2-acetoxyoxirane with dGuo and DNA.

This epoxide was used as model for the reactivity of 2-chlorooxirane and the related oxiranes listed above. The aims of Chapter III are:

1. Synthesis and spectral characterization of (2-oxoethyl)-FAPy-dGuo (**1.39**).
2. Examination of the reactivity of 2-acetoxyoxirane towards dGuo.
3. Examination of the formation of (2-oxoethyl)-FAPy-dGuo adduct (**1.39**) in oligonucleotides treated with 2-acetoxyoxirane.

## References


(1) Shimkin, M. B. (1977) In *Contrary to Nature* United States Government Printing Office, Washington, D.C.

(2) Faguet, G. B. (2005) *The War on Cancer*, Part II ed., Springer Netherlands, Amsterdam, Nederlands.

(3) Watson, J. D. and Crick, F. H. C. (1953) A structure for deoxyribose nucleic acid. *Nature 171*, 737-738.

(4) Redmond, D. E. J. (1970) Tobacco and cancer: the first clinical report. *N. Engl. J. Med. 282*, 18-23.

(5) Potter, M. (1963) Percival Pott's contribution to cancer research. *NCI Monograph No. 10*, 1-13.

(6) Yamagiwa, K. and Ichikawa, K. (1918) Experimental study of the pathogenesis of carcinoma. *J. Cancer Res. 3*, 1-29.

(7) Kennaway, E. L. and Hieger, I. (1930) Carcinogenic substances and their fluorescence spectra. *Brit. Med. J. 1*, 1044-1046.

(8) Cook, J. W., Hewett, C. L. and Hieger, I. (1933) Isolation of a cancer-producing hydrocarbon from coal tar. II. Isolation of 1,2- and 4,5-benzopyrenes, perylene and 1,2-benzoanthracene. *J. Chem. Soc.* , 396-398.

(9) Miller, J. A. (1970) Carcinogenesis by chemicals: an overview. G. H. A. Clowes Memorial Lecture. *Cancer Res. 30*, 559-576.

(10) Ames, B. N., Durtson, W. E., Yamasaki, E. and Lee, F. D. (1973) Carcinogens and mutagens: A simple test system combining liver homogenates for activation and bacteria for detection. *Proc. Natl. Acad. Sci. USA 70*, 2281-2285.

(11) Bresnick, E. (1979) Nuclear activation of polycyclic hydrocarbons. *Drug Metab. Rev. 10*, 209-223.

(12) Miller, R. E. and Guengerich, F. P. (1983) Metabolism of trichloroethylene in isolated hepatocytes, microsomes, and reconstituted enzyme systems containing cytochrome P-450. *Cancer Res. 43*, 1145-1152.

(13) Guengerich, F. P. (2000) Metabolism of chemical carcinogens. *Carcingenesis 21*, 345-351.

(14) Gajaewski, E., Rao, G., Nackerdian, Z. and Dizdaroglu, M. (1990) Modification of DNA bases in mammalian chromatin by radiation-generated free radicals. *Biochemistry 29*, 7876-7882.

(15) Doetsch, P. W., Zastawny, T. H., Martin, A. M. and Dizdaroglu, M. (1995) Monomeric base damage products from adenine, guanine, and thymine induced by exposure of DNA to ultraviolet radiation. *Biochemistry 35*, 737-742.

(16) Aruoma, O. I., Halliwell, B., Gajewski, E. and Dizdaroglu, M. (1989) Damage to the bases in DNA induced by hydrogen peroxide and ferric ion chelates. *J. Biol. Chem. 264*, 20509-20512.

(17) Dizdaroglu, M., Aruoma, O. I. and Halliwell, B. (1990) Modification of bases in DNA by copper ion-1,10-phenanthroline complex. *Biochemistry*, 8447-8451.

(18) Cooke, M. S., Evans, M. D., Dizdaroglu, M. and Lunec, J. (2003) Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J. 17*, 1195 - 1214.

(19) Pouget, J.-P., Douki, T., Richard, M.-J. and Cadet, J. (2000) DNA damage induced in cells by $\gamma$ and UVA radiation as measured by HPLC/GC-MS and HPLC-EC and comet assay. *Chem. Res. Toxicol. 13*, 541-549.

(20) Brieba, L. G., Eichman, B. F., Kokoska, E. J., Doublie, S., Kunkel, T. A. and Ellenberger, T. (2004) Structural basis for the dual coding potential of 8-oxoguanosine by a high-fidelity DNA polymerase. *EMBO J. 23*, 3452-3461.

(21) Weiderholt, C. J. and Greenberg, M. M. (2002) Fapy-dG instructs Klenow exo- to misincorporate deoxyguanosine. *J. Am. Chem. Soc. 124*, 7278-7279.

(22) Weiderholt, C. J., Delaney, M. O., Pope, M. A., David, S. S. and Greenberg, M. M. (2003) Repair of DNA containing FAPy.dG and its $\beta$-C-nucleoside analogue by formamidopyrimidine DNA glycosylase and mutY. *Biochemistry 42*, 9755-9760.

(23) Coste, F., Ober, M., Carell, T., Boiteux, S., Zelwer, C. and Castaing, B. (2004) Structural basis for the recognition of the FapydG Lesion (2,6-

diamino-4-hydroxy-5-formamidopyrimidine) by formamidopyrimidine-DNA glycosylase. *J. Biol. Chem. 279*, 44074-44083.

(24) Hu, J., de Souza-Pinto, N. C., Haraguchi, K., Hogue, B. A., Jaruga, P., Greenberg, M. M., Dizdaroglu, M. and Bohr, V. A. (2005) Repair of formamidopyrimidines in DNA involves different glycosylases: Role of the OGG1, NTH1, and NEIL1 enzymes. *J. Biol. Chem. 280*, 40544-40551.

(25) Souza-Pinto, N. C., Eide, L., Hogue, B. A., thydo, T., Stevnsner, T., Seeberg, E., Klungland, A. and Bohr, V. A. (2001) Repair of 8-oxo-deoxyguanosine lesions in mitochondrial DNA depends on the oxoguanine DNA glycosylase (OGG1) gene and 8-oxoguanine accumulates in the mitochondrial DNA of OGG1-defective mice. *Cancer Res. 61*, 5378-5381.

(26) Jaruga, P., Birincioglu, M., Rosenquist, T. A. and Dizdaroglu, M. (2004) Mouse NEIL1 protein is specific for excision of 2,6-diamino-4-hydroxy-5-formamidopyrimidine and 4,6-diamino-5-formamidopyrimidine from oxidatively damaged DNA. *Biochemistry 43*, 15909-15914.

(27) Kamiya, H., Cadena-Amaro, C., Dugue, L., Yahushiji, H., MInakawa, N., Matsuda, A., Pochet, S., Nakabeppu, Y. and Harashima, H. (2006) Recognition of nucleotide analogs containing the 7,8-dihydro-8-oxo structure by human MTH1 protein. *J. Biochem. 140*, 843-849.

(28) Imoto, S., Patro, J. N., Jiang, Y. L., Oka, N. and Greenberg, M. M. (2006) Synthesis, DNA polymerase incorporation and enzymatic hydrolysis of formamidopyrimidine nucleoside. *J. Am. Chem. Soc. 128*, 14606-14611.

(29) Cadet, J., Douki, T., Gasparutto, D. and Ravanat, J. L. (2003) Oxidative damage to DNA: formation, measurement and biochemical features. *Mutat. Res. 531*, 5-23.

(30) Dizdaroglu, M., Jaruga, P., Birincioglu, M. and Rodriguez, H. (2002) Free radical-induced damage to DNA: mechanisms and measurement. *Free Radic. Biol. Med. 32*, 1102-1115.

(31) Kalam, M. A., Haraguchi, K., Chandani, S., Loechler, E. L., Moriya, M., Greenberg, M. M. and Basu, A. K. (2006) Genetic effects of oxidative DNA damages: comparative mutagenesis of the imidazole ring-opened formamidopyrimidines (Fapy lesions) and 8-oxo-purines in simian kidney cells. *Nucleic Acids Res. 34*, 2305-2315.

(32) Ruddon, R. W. (1995) *Cancer Biology*, Oxford University Press, New York.

(33) Guengerich, F. P. (1992) Metabolic activation of carcinogens. *Pharmacol. Ther. 54*, 17-61.

(34) Smela, M. E., Hamm, M. L., Henderson, P. T., Harris, C. M., Harris, T. M. and Essigmann, J. M. (2002) The aflatoxin $B_1$ formamidopyrimidine adduct plays a major role in causing the types of mutations observed in human hepatocellular carcinoma. *Proc. Nat. Acad. Sci. USA 99*, 6655-6660.

(35) Loeb, L. A. (1985) Apurinic sites as mutagenic intermidiates. *Cell 40*, 483-484.

(36) Gates, K. S., Nooner, T. and Dutta, S. (2004) Biologically relevant chemical reactions of N7-alkylguanine residues in DNA. *Chem. Res. Toxicol. 17*, 839-856.

(37) Smela, M. E., Currier, S. S., Bailey, B. A. and Essigmann, J. M. (2001) The chemistry and biology of aflatoxin $B_1$: from mutational spectrometry to carcinogenesis. *Carcinogenesis*, 535-545.

(38) Johson, W. W., Haris, T. T. and Guengerich, F. P. (1996) Kinetics and mechanism of hydrolysis of aflatoxin $B_1$ exo-8,9-epoxide and rearrangement of the dihydrodiol. *J. Am. Chem. Soc. 118*, 8213-8220.

(39) Essigmann, J. M., Croy, R. G., Nadzan, A. M., Busby, W. F. J., Reinhold, V. N., Biuchi, G. and Wogan, G. N. (1977) Structural identification of the major DNA adduct formed by aflatoxin $B_1$ *in vitro*. *Proc. Natl. Acad. Sci. USA 74*, 1870-1874.

(40) Groopman, J. D., Croy, R. G. and Wogan, G. N. (1981) *In vitro* reactions of aflatoxin $B_1$-adducted DNA. *Proc. Natl. Acad. Sci. USA 78*, 5445-5449.

(41) International Agency for Research on Cancer. (1992) IARC Monographs on the evaluation of carcinogenic risks in humans.1,3-butadiene, pp 237-285, Lyon.

(42) Tretyakova, N. Y., Chaig, S. Y., Walker, V. E. and Swenberg, J. A. (1998) Quantitative analysis of 1,3 butadine-induced DNA addutcs *in vivo* and *in vitro* using liquid chromatography electrospray ionization tandem mass spectrometry. *J. Mass Spectrom. 33*, 363-376.

(43) Zhang, X.-Y. and Elfarra, A. A. (2004) Characterization of the reaction products of 2'-deoxyguanosine and 1,2,3,4-diepoxybutane after acid hydrolysis: formation of novel guanine and pyrimidine adducts. *Chem. Res. Toxicol. 14*, 521-528.

(44) Hemminki, K. (1987) DNA-binding products of nornitrogen mustard, a metabolite of cyclophosphamide. *Chem. Biol. Interact. 61*, 75-88.

(45) Hemminki, K. (1984) Reactions of ethyleneimine with guanosine and deoxyguanosine. *Chem. Biol. Interact. 48*, 249-260.

(46) Cussac, C. and Laval, F. (1996) Reduction of the toxicity and mutagenicity of aziridine in mammalian cells harboring the *Escherichia coli* fpg gene. *Nucleic Acids Res. 24*, 1742-1746.

(47) Lawley, P. D. and Brookes, P. (1963) The alkylation of nucleic acids and their constituent nucleotides. *Biochem. J. 89*, 127-138.

(48) Moschel, P. C., Hudgins, W. R. and Dipple, A. (1984) Substituent-induced effects on the stability of benzylated guanosines: model systems for the factors influencing the stability of carcinogen-modified nucleic acids. *J. Org. Chem. 49*, 363-372.

(49) Koga, N., Inskeep, P. B., Harris, T. M. and Guengerich, F. P. (1986) S-[2-(N7-guanyl)ethyl]glutathione, the major DNA adduct formed from 1,2-dibromoethane. *Biochemistry 25*, 2192-2198.

(50) Humphreys, W. G. and Guengerich, F. P. (1991) Structure of formamidopyrimidine adducts as determined by NMR using specifically $^{15}$N-labeled guanosine. *Chem. Res. Toxicol. 4*, 632-636.

(51) Eistetter, K. and Pfleiderer, W. (1974) Uber die umwandlung von guanosin in isoxanthopterin-N-8-ribosid. *Chem. Ber. 107*, 575-583.

(52) Tomasz, M., Lipman, R., Lee, M. S., Verdine, G. L. and Nakanishi, K. (1987) Reaction of acid-activated mitomycin C with calf thymus DNA and model guanines: Elucidation of the base-catalyzed degradation of N7-alkylguanine nucleosides. *Biochemistry 26*, 2010-2027.

(53) Eadie, J. S., Conrad, M., Toochen, D. and Topal, M. D. (1984) Mechanism of mutagenesis by $O^6$-methylguanine. *Nature 308*, 201-203.

(54) Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P. and Nishimura, S. (1991) 8-Oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proc. Natl. Acad. Sci. USA 88*, 4690-4694.

(55) Alekseyev, Y. O., Hamm, M. L. and Essigmann, J. M. (2004) Aflatoxin $B_1$ formamidopyrimidine adducts are preferentially repaired by the nucleotide excision repair pathway *in vivo*. *Carcinogenesis 25*, 1045-1051.

(56) Oleykowski, C. A., Mayernik, J. A., Lim, S. E., Groopman, J. D., Grossman, L., Wogan, G. N. and Yeung, A. T. (1993) Repair of aflatoxin $B_1$ DNA adducts by the UvrABC endonuclease of *Escherichia coli*. *J. Biol. Chem. 268*, 7990-8002.

(57) Asagoshi, K., Yamada, T., Okada, Y., Terato, H., Ohyama, Y., Seki, S. and Ide, H. (2000) Recognition of formamidopyrimidine by *Escherichia coli*

and mammalian thymine glycol glycosylases. Distinctive paired base effects and biological and mechanistic implications. *J. Biol. Chem. 275*, 24781-24786.

(58) Asagoshi, K., Yamada, T., Terato, H., Ohyama, Y., Monden, Y., Arai, T., Nishimura, S., Aburatani, H., Lindahl, T. and Ide, H. (2000) Distinct repair activities of human 7,8-dihydro-8-oxoguanine DNA glycosylase and formamidopyrimidine DNA glycosylase for formamidopyrimidine and 7,8-dihydro-8-oxoguanine. *J. Biol. Chem. 275*, 4956-4964.

(59) Chetsanga, C. J. and Frenette, G. P. (1983) Excision of aflatoxin $B_1$-imidazole ring opened guanine adducts from DNA by formamidopyrimidine-DNA glycosylase. *Carcingenesis 4*, 997-1000.

(60) Adams, R. L. P. (1991) *DNA Replication*, Oxford University Press, New York.

(61) http://www.mie.utoronto.ca/labs/lcdlab/biopic/fig/11.16.jpg.

(62) Wong, I., Patel, S. S. and Johnson, K. A. (1991) An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. *Biochemistry 30*, 526-537.

(63) Estabrook, A. R. and Reich, N. (2006) Observing an induced-fit mechanism during sequence-specific DNA methylation. *J. Biol. Chem 281*, 37205-37214.

(64) Alba, M. M. (2001) Replicative DNA polymerases. *Genome Biol. 2*.

(65) Prakash, S., Johnson, R. E. and Prakash, L. (2005) Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. *Annu. Rev. Biochem. 74*, 317-353.

(66) Lehmann, A. R., Niimi, A., Ogi, T., Brown, S., Sabbioneda, S., Wing, J. F., Kannouche, P. L. and Green, C. M. (2007) Translesion synthesis: Y-family polymerases and the polymerase switch. *DNA Repair 6*, 891-899.

(67) Jonathan, F., Irmisch, A., Green, C. M., Neiss, A., Trickey, M., Ulrich, H. D., Furuya, K., Watts, F. Z., Carr, A. M. and Lehmann, A. R. (2006) Postreplication repair and PCNA modification in *Schizosaccharomyces pombe*. *Mol. Biol. Cell. 17*, 2976-2985.

(68) Haracska, l., Torres-Ramos, C. A., Johnson, R. E., Prakash, S. and Prakash, L. (2004) Opposing effects of ubiquitin conjugation and SUMO modification of PCNA on replicational bypass of DNA lesions in *Saccharomyces cerevisiae*. *Mol. Cell. Biol. 24*, 4267-4274.

(69) Kannouche, P. L., Wing, J. and Lehmann, A. R. (2004) Interaction of human DNA polymerase η with monoubiquitinated PCNA: A possible mechanism for the polymerase switch in response to DNA damage. *Mol. Cell. 14*, 491-500.

CHAPTER II

# SYNTHESIS, SITE-SPECIFIC INCORPORATION AND TRANSLESION SYNTHESIS PAST THE MeFAPy-dGuo ADDUCT IN DNA

## Introduction

### *Formation of MeFAPy-dGuo (2.05) Lesion, Endogenous and Exogenous Methyl Sources*

MeFAPy-dGuo (**2.05**) is not only an important biological lesion, but also a useful probe for investigating specific protein-DNA interactions (*1, 2*). It is formed upon initial methylation at the N7-position of dGuo moieties (**2.01**) in DNA by methylating agents followed by hydroxide ring-opening of the imidazole ring as described in Chapter I (Figure **2-01**). The major endogenous source of the cationic N7-methyl-dGuo species (**2.02**) is from the reaction of dGuo moieties (**2.01**) of DNA and *S*-adenosyl-L-methionine (**2.06**) (SAM) (Figure **2-02**); SAM is present in high concentrations in the nucleous and is used to methylate dCyt and dAdo in DNA, which play an important role in the regulation of gene expression (*3-5*). Other endogenous methyl sources are betaine (**2.07**), choline (**2.08**), and N-nitroso compounds (Figure **2-02**) (*6*).

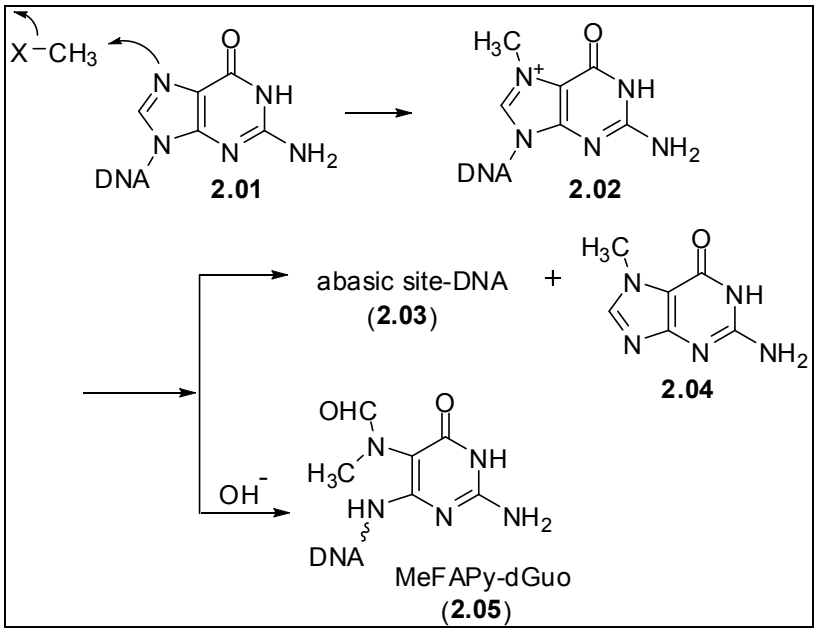



**Figure 2-01**. Alkylation of dGuo moieties (**1.01**) in DNA and formation of abasic site (**2.03**) and MeFAPy-dGuo lesion (**2.05**).

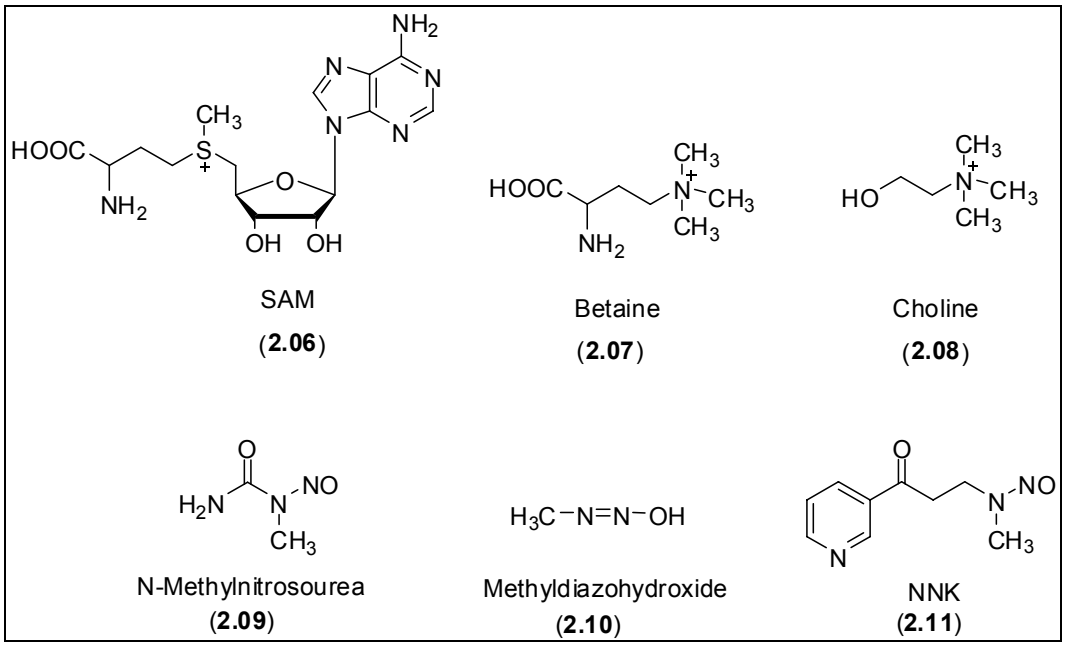



**Figure 2-02**. Structures of endogenous and exogenous methylating agents: SAM (**2.06**), betaine (**2.07**), choline (**2.08**), N-methylnitrosourea (**2.09**), methyldiazohydroxide (**2.10**) and NNK (**2.11**).

For example, N-methylnitrosourea (**2.09**) (Figure **2-02**) can be formed by nitrosation of methylurea, which is a product of the reaction of the catabolite methylamine and carbamyl phosphate, the latter is a precursor of pyrimidines (*7, 8*).

The major exogenous methyl sources are the N-nitrsoamines that are present in food, tobacco smoke, or are part of the environmental pollutants. The N-nitrosoamines undergo metabolic activation to yield methyldiazohydroxide (Figure **2-02**) (**2.10**), which is the methylating agent (*9, 10*). Analysis of laboratory animals treated with 4-(methylnitrosoamino)-1-3-pyridyl)-1-butanone (**2.11**) (NNK) (Figure **2-02**), the most abundant and the strongest carcinogen in tobacco (*11, 12*), showed the following distribution of adducts in their lungs: N7-methylguanine, $O^6$-methylguanine and to lesser extend of $O^4$-methylthymine.

### *Formation of MeFAPy-dGuo (2.05) Lesions and Abasic Sites (2.03)*

There are a number of factors that govern the partitioning between the abasic site (**2.03**) and the MeFAPy-dGuo (**2.05**) from the cationic N7-methyl-dGuo intermediate (**2.02**) (Figure **2-01**). In general, the formation of abasic sites (depurination) (**2.03**) is favored at neutral or acidic pH, and the formation of MeFAPy-dGuo (**2.05**) lesion in basic conditions. For example, the depurination of the cationic N7-methyl-dGuo species (**2.02**) occurs with $t_{1/2}$ =

4.4 h at pH 4.4 and 24 °C, whereas the ring-opening occurs with $t_{1/2}$ = 4.3 h at pH 10 and 24 °C (*13*).

The structure of duplex DNA also plays role in the depurination and the ring-opening of the cationic N7-methyl-dGuo species (**2.02**) (Figure **2-01**). Usually depurination is faster in single-stranded (ss) than double stranded (ds) DNA, while similar dependence has not been clearly observed for the ring-opening (*14*). In ds-DNA, under physiological conditions, the depurination of **2.02** is predominant and the formation of MeFAPy-dGuo (**2.05**) is very slow (*15, 16*). In spite of this slow formation of MeFAPy-dGuo (**2.05**) *in vitro*, experiments with rats treated with methylating agents showed that MeFAPy-dGuo (**2.05**) was formed in significant amounts (*17, 18*).

The structural environment of the cationic N7-methyl-dGuo (**2.02**) and the N7-methyl-Guo (**2.12**) species affects the rate of formation of the corresponding MeFAPy-dGuo (**2.05**) and MeFAPy-Guo (**2.13**) lesions. In polymeric systems, such as poly-7-methyl-dGuo and poly-7-methyl-dGuo-dCyt sequences, the ring-opening occurs very slowly and these systems were described as “remarkably resistant to ring opening at pH 8.9 and 37 °C (*19*). It was also found that in RNA, the depurination of **2.12** is slow and the formation of MeFAPy-Guo (**2.13**) lesion is about 2-3 times greater then in DNA (*19, 20*). Presumably, the reason why the depurination is very slow in RNA is the affect of the hydroxyl group at the 2’-position of the sugar moiety. The electron-withdrawing hydroxyl group destabilizes the oxocarbenium ion-like transition state (**2.14**) of the depurination reaction (Figure **2-03**).

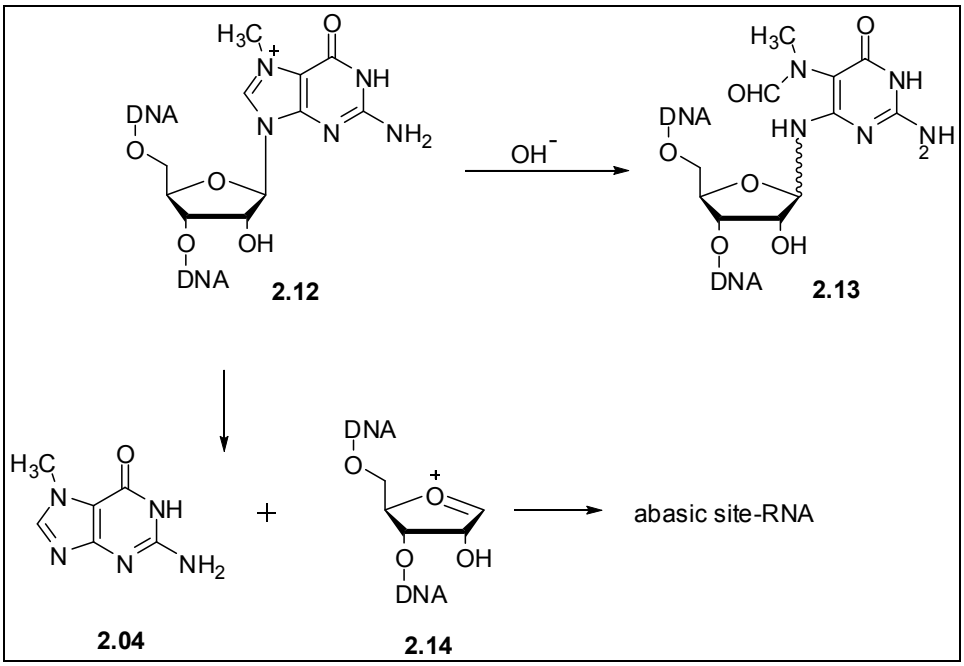



**Figure 2-03**. Effect of the structural context on the rate of ring-opening: Formation of MeFAPy-Guo (**2.13**) lesion and oxocarbenium ion **2.14**.

At the nucleoside level, the rate of formation of FAPy-dGuo is dependent on the substituent at the 5'-position. For example, the presence of a 5'-phosphate group of N7-methyl-Guo decreases the rate of ring-opening, presumably by electrostatic repulsion of the upcoming hydroxide anion (*19, 21*).

### *Anomerization and Stability of $N^5$-Alkylated FAPy-dGuo Adducts*

MeFAPy-dGuo nucleoside exists as a complex mixture of isomers. The MeFAPy-dGuo nucleoside labeled as the $\beta$-anomer (**2.15**) can undergo ring opening of the ribose to form C1'-N9 imine intermediate (**2.16**); re-closure of **2.16** can occur to give the $\beta$ (**2.15**) or $\alpha$-anomer (**2.17**) (*22*) . Additionally, ring closure reaction can also occur with the 5'-hydroxyl group to give the pyranose form as either the $\beta$ (**2.18**) or $\alpha$-anomer (**2.19**) (Figure **2-04**) (*23*).

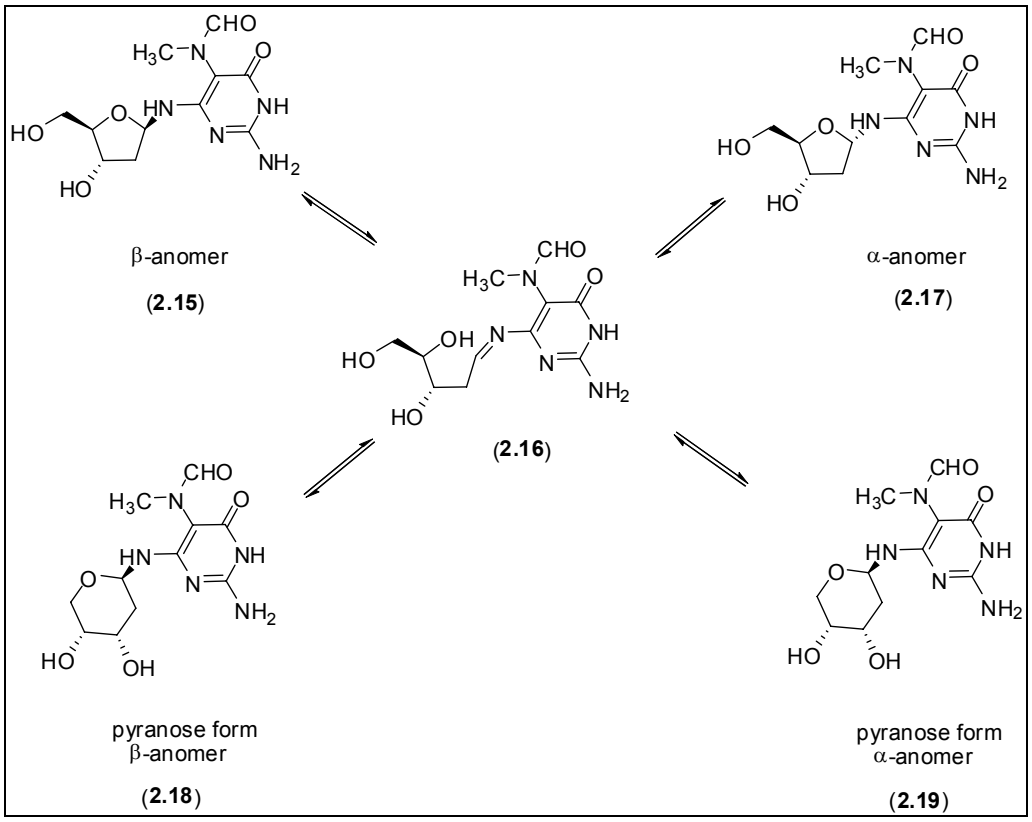



**Figure 2-04**. Anomerization of MeFAPy-dGuo nucleoside (**2.15**).

The six-membered pyranose form of the sugar is favored (*24, 25*). In DNA the formation of the six-membered ring is not possible, because of the 5'-hydroxyl group is involved in a phosphodiester bond (*26*). However, the anomerization is possible in DNA leading to the formation of $\alpha$ or $\beta$-furanose, which are separable by HPLC (*24, 27, 28*). MeFAPy-dGuo nucleoside (**2.15**) and other $N^5$-alkylated FAPy-dGuo's can also exist as atropisomers due to restricted bond rotations about formyl group and sometime the C5-$N^7$ bond. For example, isomers of the $\beta$-ribofuranose triacetate (**2.20 a-d**) derivative exhibit rotation about $N^7$-$C^8$ bond (*24*), and the MeFAPy-Gua (**2.21a** and **2.21b**) exhibit slow rotation about C5-$N^7$ bond (*29, 30*) (Figure **2-05**).

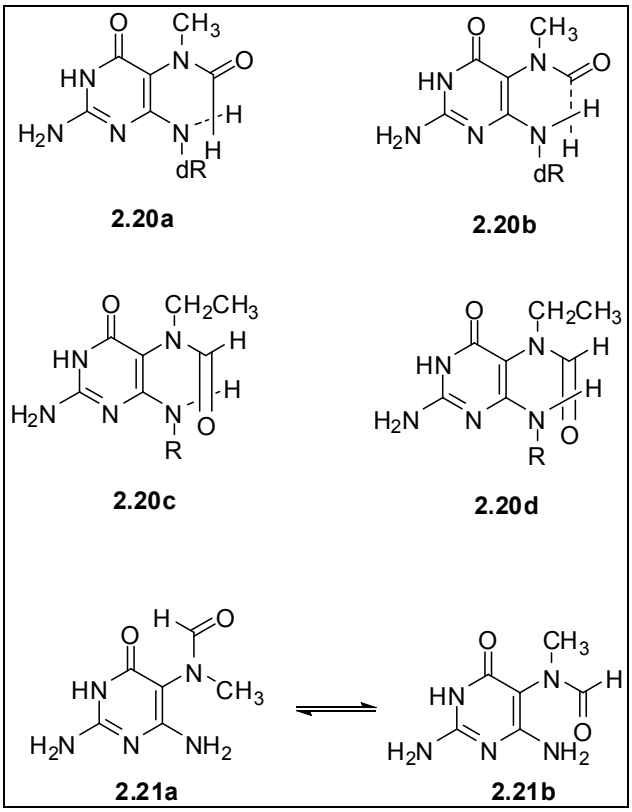



**Figure 2-05**. Formation of atropisomers of $\beta$-ribofuranose triacetate (**2.20 a**-**d**) and the MeFAPy-Gua (**2.21a** and **2.21b**).

The $N^5$-alkylated FAPy-dGuo adducts, as well as the unsubstituted FAPy-dGuo adduct, have been reported to be stable at physiological conditions and unstable in acidic and basic conditions (*31, 32*). This means that FAPy-dGuo's are not spontaneously lost from cellular DNA. It was found that FAPy-dGuo's are more stable in duplex DNA. For example, Chetsanga and coworkers reported that only 4% of the MeFAPy-Gua (**2.21a** and **2.21b**) is lost from duplex DNA after 240 h in pH 7.4 at 37 °C (*31*), whereas Greenberg and coworkers showed that the declycosylation of the unsubstituted FAPy-dGuo in single stranded DNA has a half-life time of lost

about 1500 h at 37 °C (*22*). The lost of $N^5$-alkylated FAPy-dGuo's from DNA or nucleosides is accelerated in acidic conditions, such as 0.1 N HCl or 75 % formic acid 24 -100 °C, 15 min to 72 h (*31, 33*). Interestingly under slightly acidic conditions some $N^5$-alkylated FAPy-dGuo's may undergo re-closure of the imidazole ring to yield the corresponding substituted guanine (34). Acid conditions can also deformylate $N^5$-alkylated FAPy-dGuo lesions (*31, 35, 36*). Deformylation of $N^5$-alkylated FAPy-dGuo's is also observed when they are treated with strong bases. For example, when MeFAPy-dGuo (**2.15**) nucleoside was treated with 0.2 N NaOH at 37 °C for 4 h, this led to the formation of the corresponding deformylated MeFAPy-dGuo (**2.15**) nucleoside (*31*).

### *Synthesis of FAPy-dGuo and $N^5$-alkylated FAPy-dGuo Oligonucleotides*

The development of synthetic strategies for $N^5$-alkylated FAPy-dGuo adducts and their subsequent incorporation into oligonucleotides is very challenging because of the $N^5$-alkylated FAPy-dGuo chemistry discussed above. Oligonucleotides containing $N^5$-alkylated FAPy-dGuo lesions, including the MeFAPy-dGuo (**2.05**) lesion, can be synthesized by treatment of ss-DNA with the corresponding alkylating agent followed by treatment with base (*19, 37-41*). However, there are several restrictions for such a synthesis of MeFAPy-dGuo (**2.05**) containing oligonucleotides. First, the ss-DNA must have only one dGuo residue and second, methylating agents can also react

with other bases. For example, ss-DNA was methylated by dimethyl sulfate (DMS) and then treated with 0.2 M NaOH to perform the ring-opening of the cationic N7-methyl-dGuo (**2.02**) adduct. The methylation occurred also at the N1 and N3-positions of dAdo giving the following distribution of products: MeFAPy-dGuo (**2.05**) 59%, 1-methyladenine (**2.22**) 21% , 3-methyladenine (**2.23**) 6% and other as minor lesions (Figure **2-06**) (*42*).

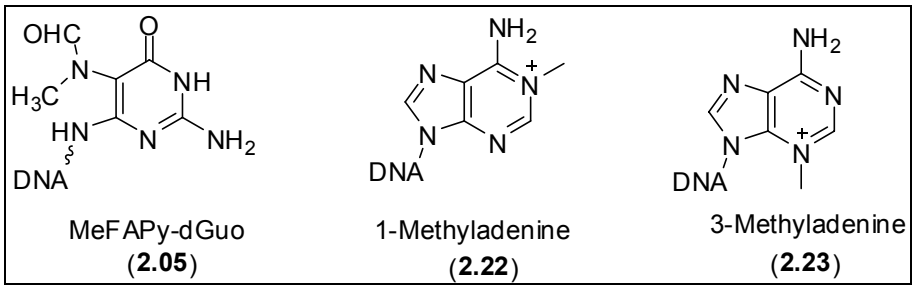



**Figure 2-06**. Structures of the DNA lesions - MeFAPy-dGuo (**2.05**), 1-methyladenine (**2.22**) and 3-methyladenine (**2.23**).

Oligonucleotides containing MeFAPy-dGuo (**2.05**) adduct could be prepared by enzymatic ligation of N7-methyl-dGTP (m7 dGTP) into duplex oligonucleotide (*43*) followed by base treatment in order to perform the hydroxide ring-opening of the cationic N7-methyl-dGuo species. The strategy to utilize ligation has not been used to obtain oligonuleotides containing MeFAPy-dGuo (**2.05**) adduct due to several drawbacks. The ligation strategy requires the preparation of complimentary strand for each of the target oligonucleotides containing the cationic N7-methyl-dGuo (**2.02**) species

(Figure **2-07**). If the reaction had occurred, ideally in good yield, the oligonucleotide containing the cationic N7-methyl-dGuo (**2.02**) species had to be purified from the mixture of other oligonucleotides, which would result in partial depurination of the cationic N7-methyl-dGuo (**2.02**) species as discussed earlier. If the isolation the oligonucleotide containing the cationic N7-methyl-dGuo (**2.02**) adduct had been successful, the modified oligonucleotide had to be subjected to base treatment which subsequently required one more step of purification.

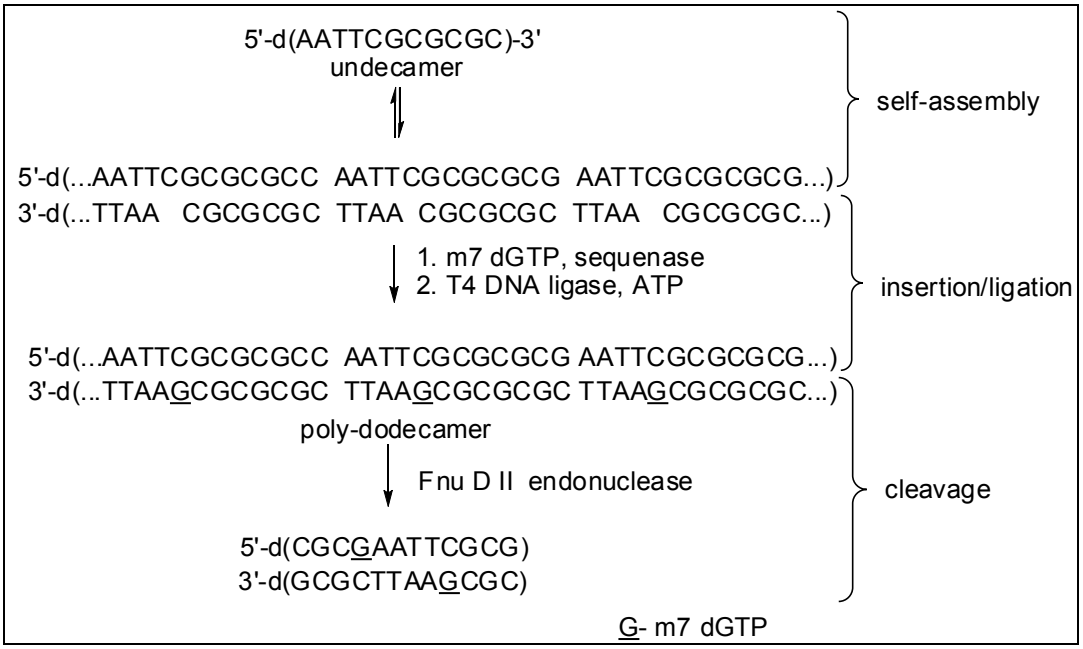



**Figure 2-07**. Synthesis of a Dickerson/Drew dodecamer containing the cationic N7-methyl-dGuo (**2.02**) species by ligation (33).

Oligonucleotides containing the MeFAPy-dGuo (**2.05**) lesion can be obtained by use of a DNA polymerase chain reaction (PCR). This method was developed by Asagoshi and coworkers, who introduced MeFAPy-dGuo (**2.05**) into DNA by incorporation of N7-methyl-dGTP followed by base treatment in order to perform the hydroxide ring-opening of the cationic N7-methyl-dGuo moiety (*44-46*). The utilization of this method for the synthesis of MeFAPy-dGuo (**2.05**) containing oligonucleotides has several complications, which makes it inconvenient. The main complications are the multiple incorporation of N7-methyl-dGTP's unless the template has one dCyd and the potentially difficult purification of the end product. In addition, this method does not allow the preparation of large amounts of MeFAPy-dGuo (**2.05**) containing oligonucleotide.

The best method for obtaining oligonucleotides containing MeFAPy-dGuo (**2.05**) adduct or any other FAPy's (dGuo derivatives) adducts, is by solid-phase synthesis, which requires a suitable phosphoroamidite reagents. This aproach allows for the site-specific incorporation of the corresponding FAPy-dGuo adducts into oligonucleotides in any desired sequence and scale. Thus far, Greenberg and our group have achieved syntheses of the FAPy-dGuo (**2.30**) (*26, 47, 48*) and MeFAPy-dGuo (**2.36**) phosphoroamidites, respectively. The syntheses of these two phosphoroamidites differ in many aspects. Greenberg and coworkers synthesized a FAPy-dGuo (**2.30**) phosphoroamidite as a dinucleotide in which the 3′-nucleotide is the desired FAPy-dGuo moiety. The use of this dinucleotide **2.30** prevents the

isomerization of the desired furanose form of the FAPy-dGuo to the undesired pyranose form during DNA synthesis. Although a well-conceived strategy, the need for a dinucleotide phosphoroamidite also represents a synthetic drawback. The synthesis of FAPy-dGuo dinucleotide phosphoroamidite (**2.30**) is a twelve step synthesis and involves several very low yield reactions (Figure **2-08** and **2-09**). Additionally, in order to synthesize a full range of potential DNA sequences there is a need for four different dinucleotide phosphoroamidite reagents.

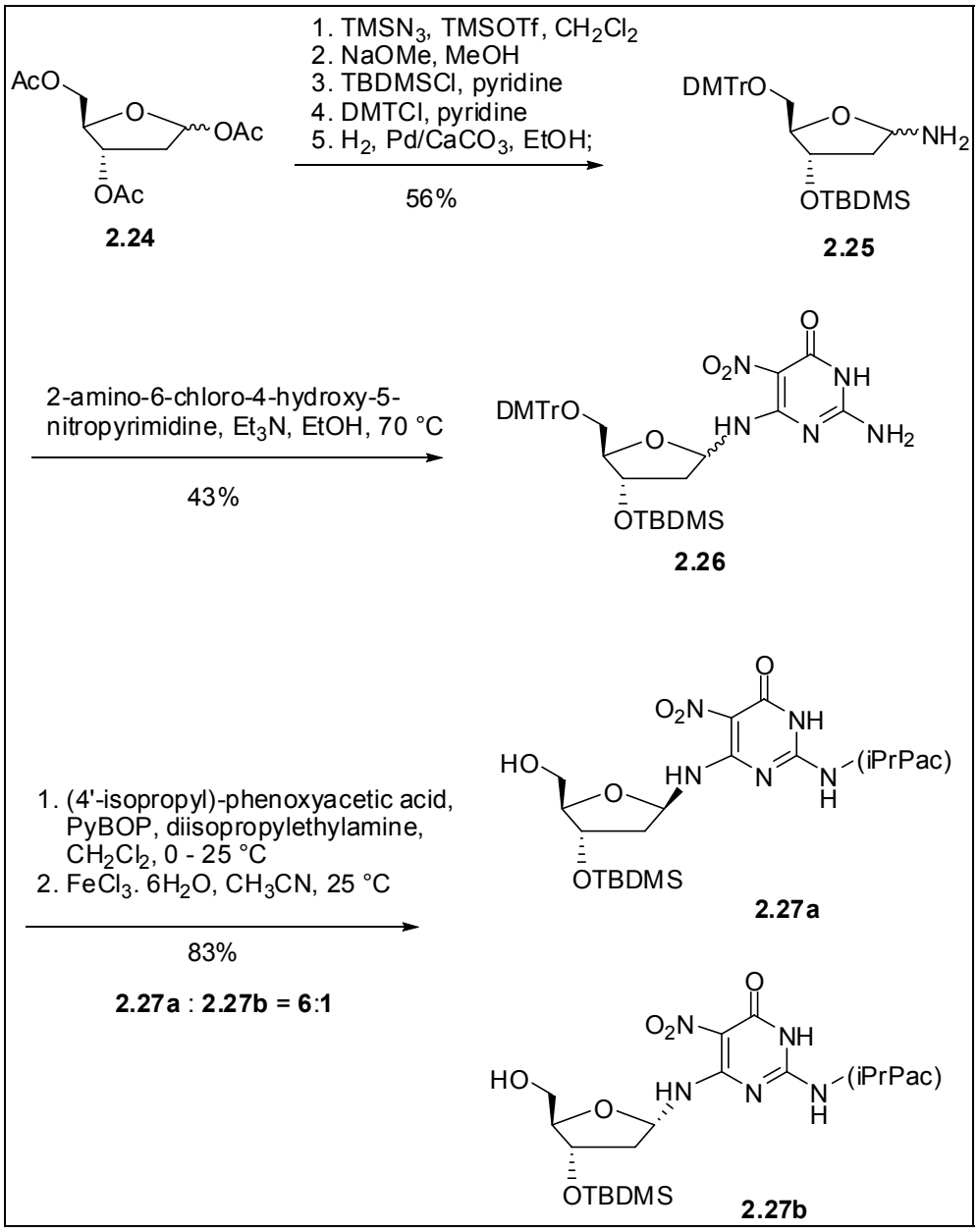



**Figure 2-08**. Greenberg's synthesis of **2.27a** and **2.27b** (*48*).

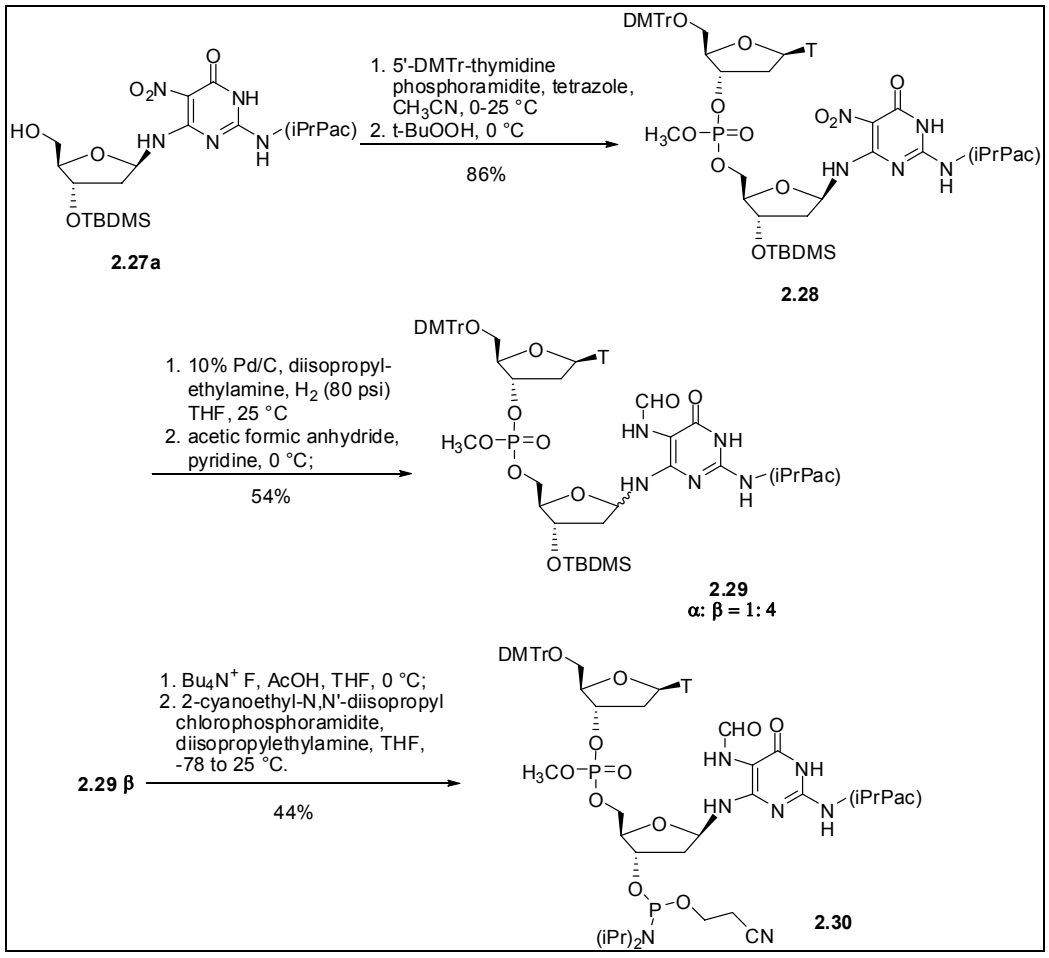



**Figure 2-09**. Synthesis of FAPy-dGuo phosphoroamidite dinucleotide (**2.30**) (*48*).

### *Biological Consequences of the Formation of MeFAPy-dGuo (2.05) Lesion*

The mutagenicity of the MeFAPy-dGuo (**2.05**) lesion has not been well studied due to the lack of suitable method for site-specific incorporation of the MeFAPy-dGuo (**2.05**) adduct into oligonucleotides. Generally, it was accepted that MeFAPy-dGuo (**2.05**) was a non-mutagenic, but lethal lesion (*40, 46, 49, 50*). Asagoshi and coworkers, who used PCR to incorporate N7-methyl-dGTP into oligonucleotides, reported that MeFAPy-dGuo (**2.05**) was a strong, but not absolute, block to DNA replication catalyzed by Klenow fragment exo$^-$ (*46*). The blocking of the extension was due to the mis-incorporation of dATP and dGTP opposite the MeFAPy-dGuo (**2.05**) adduct. The preferred insertion opposite this lesion was dCTP, which represents a non-mutagenic bypass. It was also found that the efficiency of the bypass was dependant on the position of the MeFAPy-dGuo (**2.05**) adduct; when MeFAPy-dGuo (**2.05**) adduct was placed in a terminal position the full-length extension was more efficient (*46*).

Laval and coworkers reported *in vitro* DNA polymerase and transfection studies with DNA containing MeFAPy-dGuo (**2.05**) adduct (*49, 50*). The modified DNA was obtained by methylation with dimethyl sulfate (DMS) of single-stranded M13mp18 DNA, followed by treatment with 0.2 M NaOH to perform the ring-opening of the imidazole ring of the cationic N7-methyl-dGuo moiety (**2.02**). When Klenow fragment exo$^-$ or T4 DNA polymerases were used, the DNA synthesis stopped one base prior to the MeFAPy-dGuo (**2.05**) adduct as well as the other methyl-dAdo and methyl-

dCyt generated lesions. The transfection studies with the MeFAPy-dGuo (**2.05**) containing DNA showed that in the progeny phage the most frequent mutation was observed at adenine sites (A→G transitions), rather than the MeFAPy-dGuo (**2.05**) sites, which was attributed to the presence of coexisting base lesions that are formed during the treatment of DNA with DMS.

## Results and Discussion

### *Site-specific Synthesis of Oligonucleotides Containing MeFAPy-dGuo (2.05) Lesion*

The synthesis of a MeFAPy-dGuo phosphoroamidite is probably possible by adopting Greenberg's strategy for the synthesis of the FAPy-dGuo (**2.30**) phosphoroamidite, but will have the same drawbacks as described above. The strategy we utilized to obtain the MeFAPy-dGuo phosphoroamidite **2.36** was to methylate at the N7-position a dGuo derivative **2.33** that was suitably protected for solid-phase oligonucleotide synthesis (Figure **2-10**).

The synthesis of **2.33** from dGuo (**2.31**) was accomplished using a previously described method (*51*). Protected dGuo derivative **2.33** was treated with an excess of methyl iodide in dry degassed DMSO for 45 min at room temperature to produce the cationic N7-methyl-dGuo (**2.34**), which was not isolated. After removals of the excess methyl iodide, the solution containing **2.34** was briefly treated with 1M NaOH, and then neutralized with

0.1 M HCl to afford pure **2.35** as a white precipitate. The protected MeFAPy-dGuo nucleoside (**2.35**) was isolated by filtration and converted to the desired MeFAPy-dGuo phosphoroamidite (**2.36**) without further purification using standard conditions (*52*).

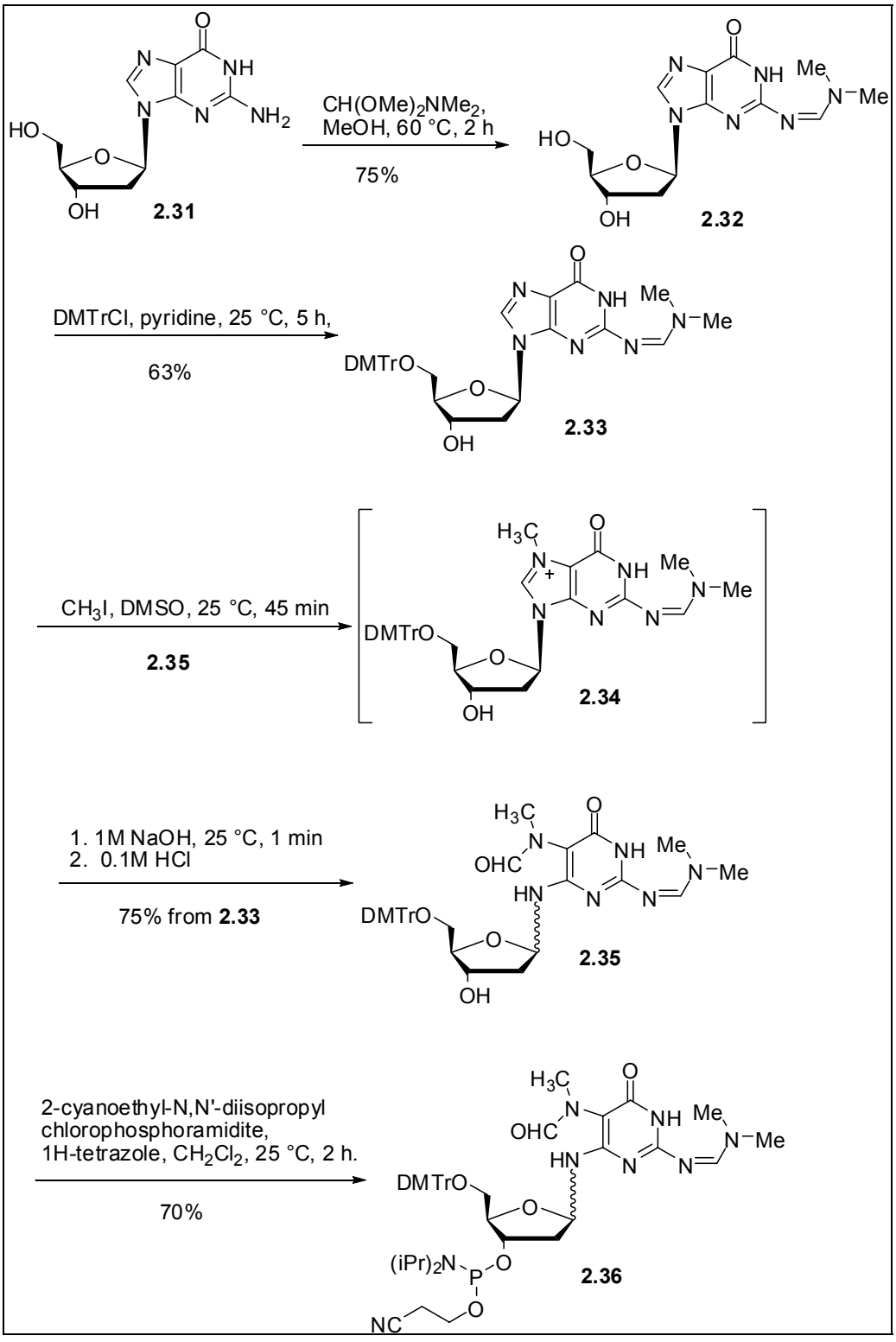



**Figure 2-10**. Synthesis of MeFAPy-dGuo phosphoroamidite (**2.36**).

The synthesis of the MeFAPy-dGuo phosphoroamidite (**2.36**) was accomplished in only four steps and 25% percent overall yield, compared to twelve steps and 0.17% (without considering the preparation of **2.24** and 2-amino-6-chloro-4-hydroxy-5-nitropyrimidine in Figure **2-08**) for the FAPy-dGuo phosphoroamidite (**2.30**). In addition, **2.36** is the only phosphoroamidite required for a synthesis a full range of potential DNA sequences compared to four different FAPy-dGuo dinucleotide phosphoroamidites from the Greenberg's synthesis.

Phosphoroamidite **2.36** was used for the preparation of oligonucleotides containing the MeFAPy-dGuo (**2.05**) lesion at a defined location. Incorporation of the modified nucleotide was performed off-line, using a manual coupling protocol. The crucial step of the oligonucleotide synthesis was the deprotection of the 5′-hydroxy group of the MeFAPy-dGuo nucleotide, since the ribose ring can undergo a rearrangement to the pyranose form under acidic conditions. Initially a 6-mer, 5′-CTT-(MeFAPy-dGuo)-TT-3′ (**2.37a, b**) was synthesized at two deprotection times. When the 5′-DMTr group was removed using a "short" deprotection cycle that consisted of using 160 µL of Deblock reagent (3% $Cl_3CCO_2H$ in methylene chloride) for 20 sec, followed by incorporation of the subsequent nucleotide under standard conditions (on line), fives products were observed by HPLC analysis (Figure **2-11**, Panel A). Each component was separated, and then re-analyzed by HPLC. Component 1 (**2.37a**) did not change composition upon re-analysis. However, when components 2-5 (**2.37b**) were individually re-

analyzed, it was found that they had all equilibrated to the same four component mixture (Figure **2-11**, Panel D). The MALDI-TOF-MS analysis of component 1 (**2.37a**) and components 2-5 (**2.37b**) showed that they all possess the correct mass for the desired oligonucleotide. We tentatively assigned component 1 (**2.37a**) as the pyranose forms of the MeFAPy-dGuo (**2.05**) and components 2-5 (**2.37b**) as the furanose forms from this limited information. In addition to the furanose and pyranose form of the ribose unit, $\alpha$- and $\beta$-anomers and antropisomers of each due to restricted rotation about C5-$N^7$ bond are also possible for $N^5$-alkylated FAPy-dGuo adducts. We reasoned that the MeFAPy-dGuo (**2.05**) would exist largely in the equatorial conformation of the pyranose form leaving two possible atropisomers. We subsequently found that component 1 (**2.37a**) actually consisted of two products, consistent with this prediction. The furanose form of the MeFAPy-dGuo adduct (**2.05**) should exist as a mixture of two atropisomers of the $\alpha$- and $\beta$-anomers, giving a total of four interconverting isomers. A similar behavior was observed for the more extensively studied FAPy-dGuo adduct of aflatoxin $B_1$ ($AFB_1$-FAPy-dGuo) and is in accord with the present observations as well. We also predicted that more of the pyranose form would be produced during longer deprotection cycles.

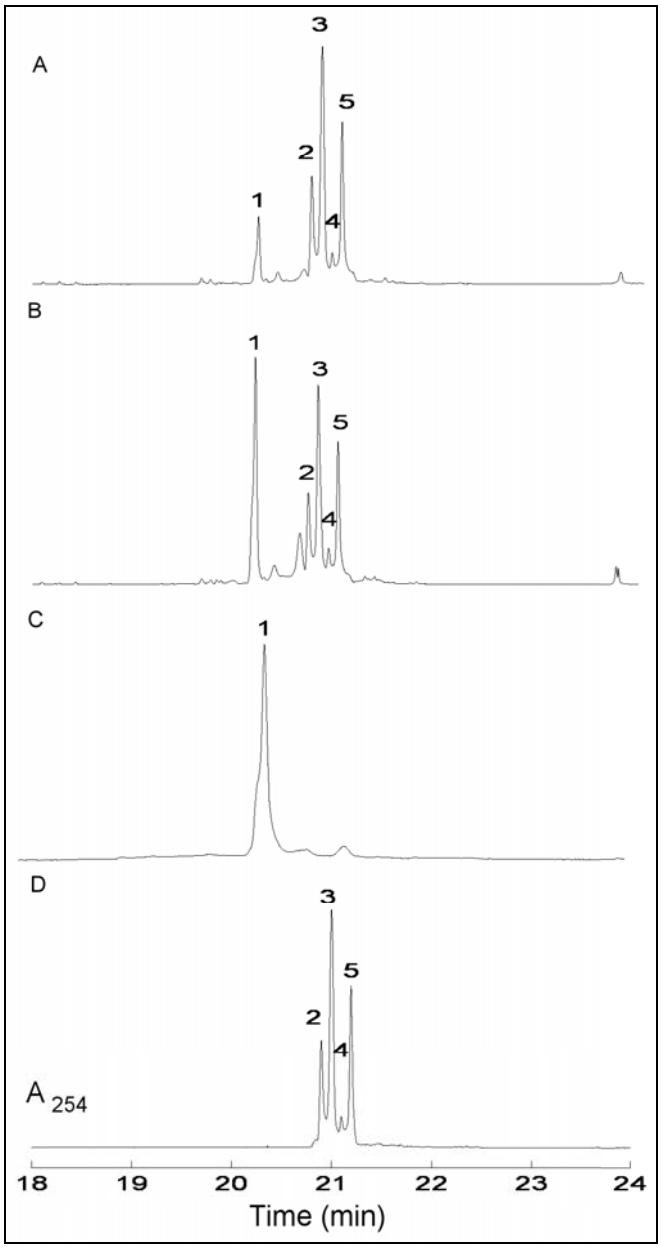



**Figure 2-11**. HPLC analysis of the 5′-CTT-(MeFAPy-dGuo)-TT-3′ (**2.37a** and **b**) oligonucleotides. **A**. Analysis of the oligonucleotide synthesis with a "short" deprotection cycle. **B**. Analysis of the oligonucleotide synthesis with a "long" deprotection cycle. **C**. Analysis of peak 1 (**2.37a**), identified as the pyranose form of the MeFAPy-dGuo lesion, after initial purification. **D**. Analysis of interconverting peaks 2-5 (**2.37b**), identified as the furanose form of the MeFAPy-dGuo lesion, after each peak was individually purified.

When a longer deprotection time cycle (320 µL of Deblock reagent for 50 sec plus an additional 3 min in standby mode) for removing the 5'-DMTr group nucleotide was utilized, a larger percentage of the pyranose form was observed (Figure **2-11**, Panel B), (Table **2-01**).

**Table 2-01**. Yields of the 5′-CTT-(MeFAPy-dGuo)-TT-3′ (**2.37a** and **b**) oligonucleotides depending on the deprotection cycle: "short"-160 µL of Deblock reagent (3% $Cl_3CCO_2H$ in methylene chloride) for 20 sec, "mediate"- 240 µL of Deblock reagent for 50 sec, and "long"- 320 µL of Deblock reagent for 50 sec plus an additional 3 min in standby mode.

| **Deprotection cycle** | **Total yield (OD's)** | **2.27a** (%) | **2.27b** (%) |
|---|---|---|---|
| "short" | 20 | 9 | 91 |
| "mediate" | 28 | 16 | 84 |
| "long" | 18 | 90 | 10 |

### *Synthesis and NMR Studies of the Trinucleotide 5'-A-(MeFAPy-dGuo)-C-3' (2.38a and 2.38b)*

The NMR spectrum of the proposed furanose and pyranose forms of the 5′-CTT-(MeFAPy-dGuo)-TT-3′ oligonucleotides (**2.37a** and **2.37b**) was very complex due to the presence of many inter-converting species. We therefore synthesized the trinucleotide 5′-A-(MeFAPy-dGuo)-C-3′ (**2.38a** and **2.38b**) using the "short" and "long" deprotection cycles. The trinucleotide (**2.38a**) obtained via the "short" deprotection protocol consisted of three interconverting species, while the trinucleotide (**2.38b**) obtained via the "long" deprotecion cycle consisted of two interconverting species (Figure **2-12**).

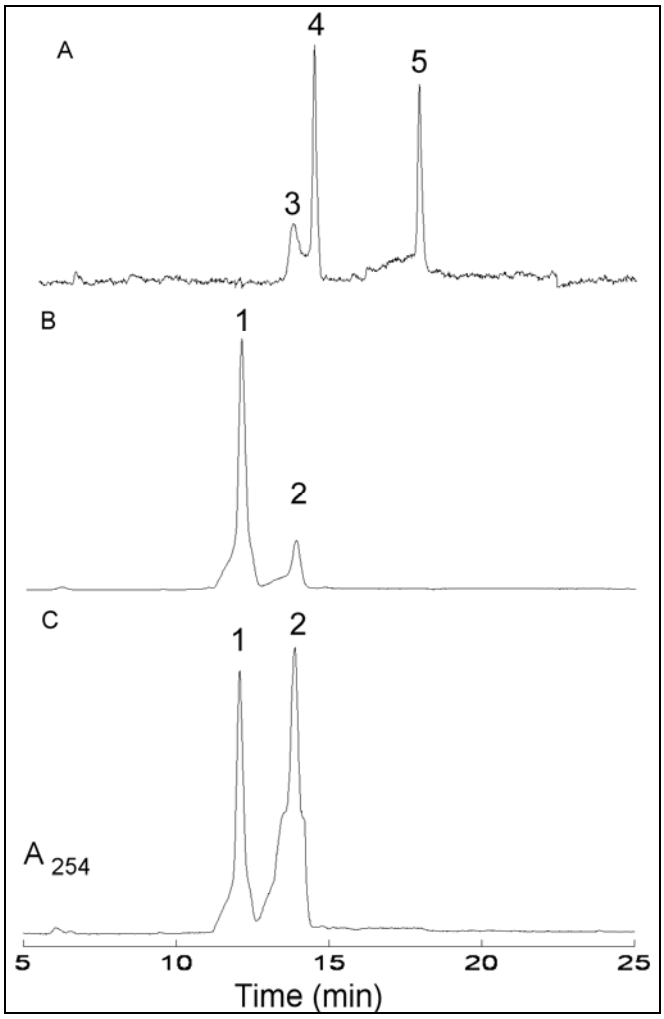



**Figure 2-12**. HPLC analysis of the 5′-A-(MeFAPy-dGuo)-C-3′ oligonucleotides (**2.38a** and **b**). **A**. Analysis of inter-converting peaks 3-5 (**2.38a**), identified as the furanose form of the MeFAPy-dGuo lesion, after each peak was individually purified. **B**. Analysis of peak 1 and 2 (**2.38b**), identified as the pyranose form of the MeFAPy-dGuo lesion, after initial purification. **C**. Analysis of peaks 1 and 2 (**2.38b**) after 24 h.

The high mobility of the trinucleotide prevented the evolution of inter-residue NOE's for use in sequence specific resonance assignment. The spin systems of the individual 2′-deoxyribose units were identified from TOCSY spectra; the assignment of the 2′-deoxyribose proton resonances was based

on connectivity and chemical shift. The carbon resonance of the 2′-deoxyribose units were correlated from the proton assignments by HSQC spectra. Inter-residue connectivity was established by $^{3}J_{PH3'}$ and $^{3}J_{PH5''}$ scalar couplings. Multiple conformations were observed for the furanose (**2.38a**) and pyranose (**2.38b**) samples precluding a full characterization the trinucleotides. However, the pyranose (**2.38b**) and furanose (**2.38a**) forms could be distinguished from the available data. Multiplicity edited HSQC data showed that the 4′-position of the MeFAPy-dGuo residue is a methine (CH) for the pyranose (**2.38b**) form and a methylene ($CH_2$) for the furanose (**2.38a**) form. A three-bond $^{31}P$-$^{1}H$ coupling is observed from the phosphate that bridges the 5′-dAdo and the MeFAPy-dGuo residue. A one-bond $^{1}H$-$^{13}C$ HSQC identified the MeFAPy-dGuo carbon resonance involved in the phosphate bridge as a methylene (furanose form, **2.38a**) or methane (pyranose form, **2.38b**). The carbon resonances of the MeFAPy-dGuo 2′-deoxyribose unit were also compared to the model 2′-deoxyribosides (Figure **2-13**). We observed a significant change in chemical shift for the 4′-carbon resonance of the MeFAPy-dGuo residue depending on the ring size. The 4′-carbon chemical shifts for the MeFAPy-dGuo residue were 82.4 ppm for the furanose (**2.38a**) form and 69.6 ppm in the pyranose (**2.38b**) form in the 5′-A-(MeFAPy-dGuo)-C-3′ trinucleotide. The furanose value is compared to that of the 3′,5′-dibutarate of the $AFB_1$-FAPy-dGuo nucleoside (**2.39**) (*53*), and methyl 2′-deoxyribofuranoside 3,5-dimethyleher (**2.41**) (*54*), which have C4 carbon shift of 83.7 and 82.4 ppm, respectively. The C4′ carbon chemical shifts of the

MeFAPy-dGuo pyranose (**2.38b**) form compares favorably with the pyranose form of the $AFB_1$-FAPy-dGuo nucleoside (**2.40**) (*55*) and methyl 2′-deoxyribopyranose 3,5-diacetae (**2.42**) (*56*) (66.6 and 69.0, respectively).

The NMR experiments with the trinucleotides **2.38a** and **2.38b** showed that our initial tentative assignments, based on the FAPy-dGuo's conversions were correct. This means that through the use of the "short" or "long" deprotection cycle, we were able to synthesize MeFAPy-dGuo containing oligonucleotides having predominantly furanose or pyranose form, respectively.

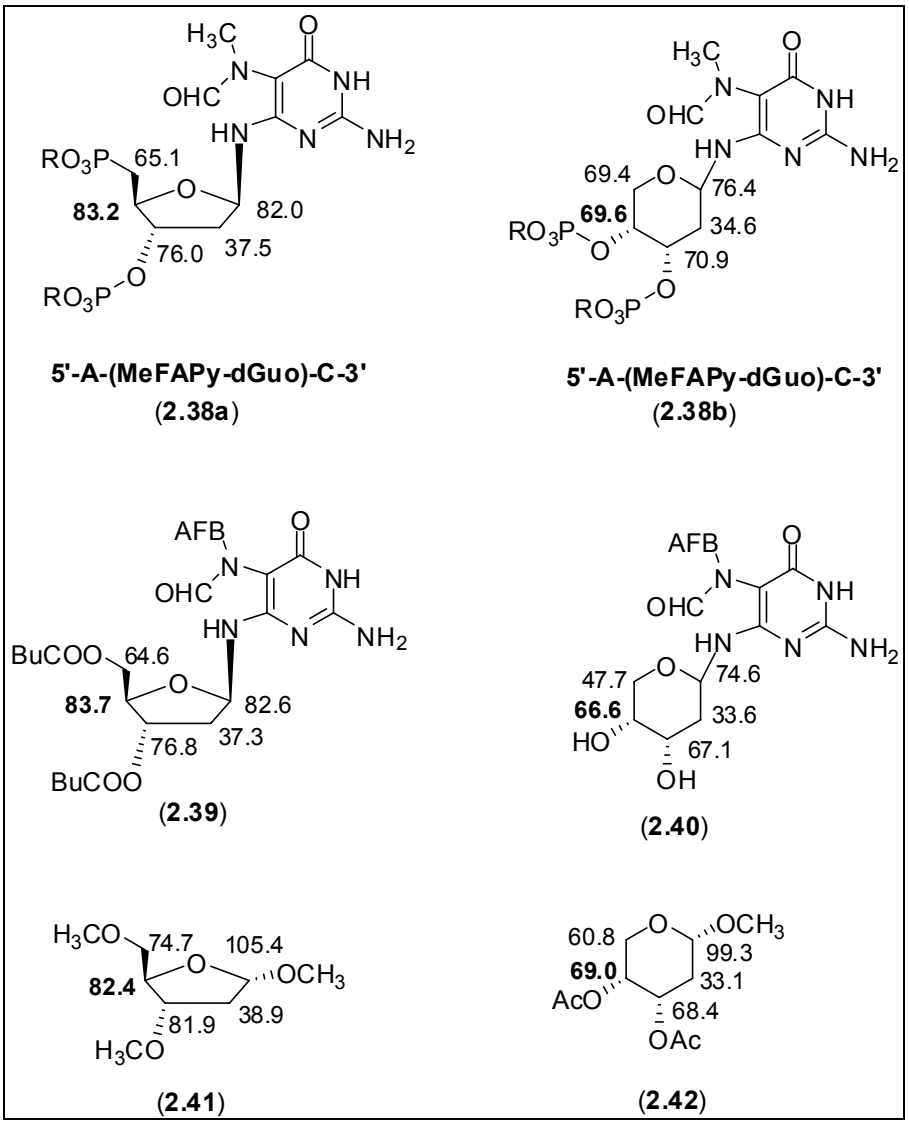



**Figure 2-13**. Comparison of the $^{13}C$ chemical shifts for furanose and pyranose forms of 2′-deoxyribose units.

### *Oligonucleotide Sequences Containing the MeFAPy-dGuo (2.05) Lesion*

Phosphoroamidite **2.36** was used for the preparation of several oligonucleotides containing the MeFAPy-dGuo (**2.05**) lesion at a defined location (Table **2-02**). The modified oligonucleotides were characterized by MALDI-TOF-MS spectrometry. The presence of the MeFAPy-dGuo (**2.05**) was confirmed by enzymatic digestion and analysis of the modified nucleoside by HPLC (Appendix); the modified oligonucleotides were also sequenced by MALDI-TOF analysis using controlled exonuclease digestions with nucleases PI and PII (Appendix).

**Table 2-02**. Characterization of the oligonucleotides containing the MeFAPy-dGuo (furanose and pyranose forms) (**2.05**) lesion by MALDI-TOF-MS spectrometry; a-positive LC-ESI/MS.

| Oligonucleotide | Ring | M/z (Da) | |
|---|---|---|---|
| | | Observed | Calculated |
| 3'-C-(MeFAPy-dGuo)-A-5' | furanose | 900.2[a] | 899.2 |
| | pyranose | 900.2[a] | |
| | | | |
| 3'-TT-(MeFAPy-dGuo)-TTC-5' | furanose | 1803.8 | 1803.6 |
| | pyranose | 1803.1 | |
| | | | |
| 3'-CTCTC-(MeFAPy-dGuo)-CTTCTCC-5' | furanose | 3842.9 | 3841.6 |
| | pyranose | 3843.4 | |
| | | | |
| 3'-CCCCCTTCCTAG-(MeFAPy-dGuo)-TACT-5' | furanose | 5401.9 | 5401.2 |
| | pyranose | 5101.3 | |
| | | | |
| 3'-CCCCCGCTACGAGCATTCCTAAG-(MeFAPy-dGuo)-TACT-5' | furanose | 8498.1 | 8497.3 |
| | pyranose | 8495.2 | |

### *Stability of the Oligonucleotides Containing the MeFAPy-dGuo (2.05) Lesion*

The stability of ss-DNA containing the Me-FAPy-dGuo (**2.05**) adduct was investigated in water and buffer at pH 6.5 and 7.0. The ss-DNA's used to assess the stability of MeFAPy-dGuo (**2.05**) adduct were the 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2.43a** and **b**) oligonucleotides, which were obtained via “short” (**2.43a**) and “long” (**2.43b**) deprotection cycles (see supplementary for HPLC and CZE traces). We initially used MALDI-TOF MS spectrometry to perform the stability assays due to the length of the oligonucleotides. The oligonucleotides **2.43a** and **2.43b** were dissolved in water or pH 7.0 phosphate buffer at room temperature and 95 °C. Aliquots were then taken periodically and analyzed by MALDI-TOF MS spectrometry. The MS analysis showed that oligonucleotides **2.43a** and **2.43b** were stable in pH 7.0 phosphate buffer, but underwent degradation in water by releasing the MeFAPy-Gua (**2.21a** and **2.21b**) followed by cleavage of the resulting abasic oligonucleotides to yield oligonucleotides **2.44a, 2.45a**, **2.44b**, **2.45b** and **2.46** (Figure **2-14**). The percentage of decomposition of oligonucleotides **2.43a** and **2.43b** was determined by the ratio between the corresponding abasic sites (**2.44a** and **2.45b**) and the starting oligonucleotides (**2.43a** and **2.43b**) in the MALDI-TOF MS spectra (see Appendix for MALDI-TOF MS spectra) (Table **2.03**).

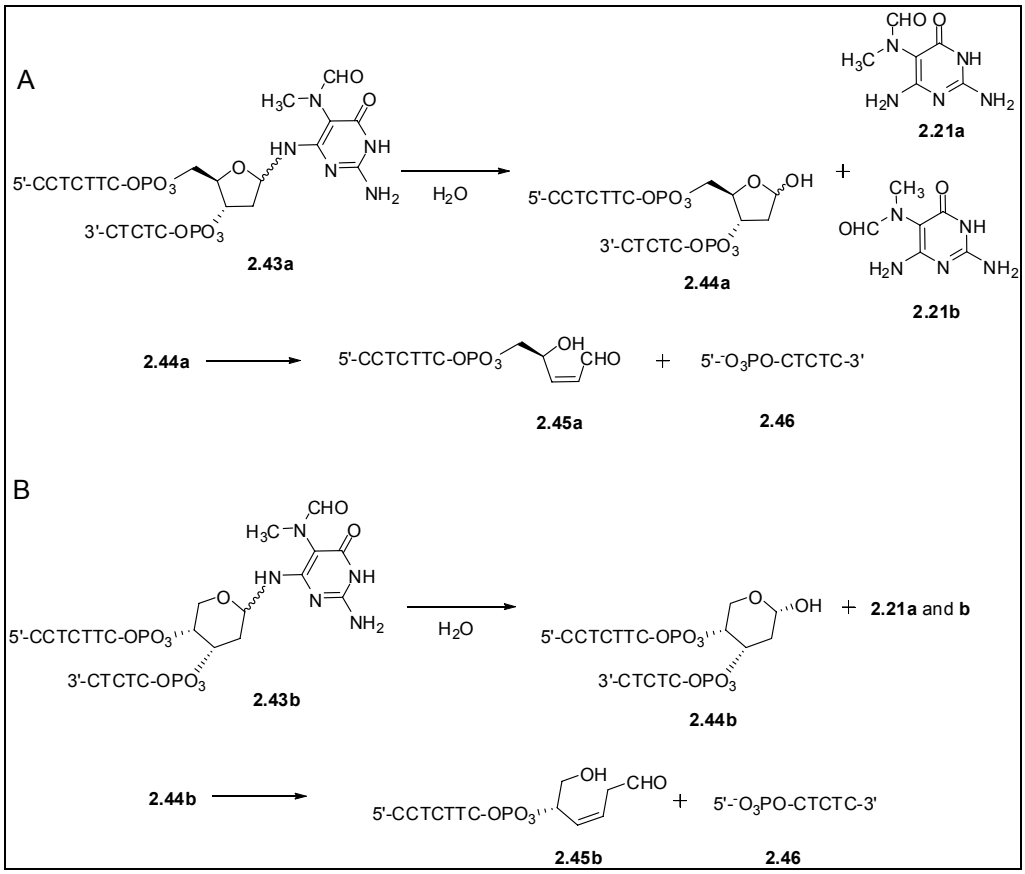



**Figure 2-14**. Stability of the MeFAPy-dGuo oligonucleotide **2.43a** and **2.43b** in water. **A**. Oligonucleotide **2.43a** undergoes lost of MeFAPy-Gua (**2.21a** and **2.21b**) to give the abasic site oligonucleotide **2.44a**, which decomposes to oligonucleotides **2.45a** and **2.46**. **B**. Oligonucleotide **2.43b** undergoes lost of MeFAPy-Gua (**2.21a** and **2.21b**) to give the abasic site oligonucleotide **2.44b**, which decomposes to oligonucleotides **2.45b** and **2.46**.

**Table 2-03**. Percent of decomposition of the oligonucleotides **2.42a** and **2.42b** in water and phosphate buffer pH 7.0 at 25 °C and 95 °C.

| Temperature (°C) | Medium | Time (days) | Decomposition (%) | |
|---|---|---|---|---|
| | | | 2.42a | 2.42b |
| 25 | water | 2 | 50 | 30 |
| | | 5 | 90 | 50 |
| | | | | |
| 90 | water | 1 hour | 100 | 100 |
| | | | | |
| 25 | pH 7 | 2 | stable | stable |
| | | 5 | stable | stable |
| | | | | |
| 90 | pH 7 | 10 min | stable | — |
| | | 30 min | 10% | — |
| | | 1 hour | 20% | — |

The stability of the oligonucleotides **2.43a** and **2.43b** in pH 7.0 buffer was important observation to establish, because it was critical for translesion DNA polymerase studies, which require annealing of the MeFAPy-dGuo containing oligonucleotides with their complementary strands by heating at 95 °C for 5 min followed by slow cooling to room temperature.

We also performed assays on the stability of several other MeFAPy-dGuo containing oligonucleotides in pH 7.0 buffer and the results confirmed that MeFAPy-dGuo containing oligonucleotides were stable under these conditions even at high temperature.

A reasonable explanation for the degradation of oligonucelotides **2.43a** and **2.43b** in deionized water is the lack of any buffer capacity of the deionized water, which may allow the pH to vary depending on the analyte; the analyte may cause the pH to become acidic or basic. We hypothesized

that most likely the oligonucleotides **2.43a** and **2.43b** caused the pH of water to become acidic, since the release of $N^5$-alkylated FAPy-Gua from $N^5$-alkylated FAPy-dGuo oligonucleotides is known to occur in acidic conditions. We also hypothesized that the release of the MeFAPy-Gua (**2.21a** and **2.21b**) occurs via formation of carbinolamine species, which are formed in acidic conditions. In order to explore our hypothesis, we investigated the stability of oligonucleotide **2.43a** in pH 6.5 phosphate buffer. The LC-ESI/MS analysis of oligonucleotide **2.43a** after 24 h at pH 6.5 showed that MeFAPy-dGuo (**2.05**) adduct was completely lost resulting in abasic site oligonucleotide **2.44a** (Figure **2-15,** Panel C). After additional 24 hours the LC-ESI/MS analysis showed full conversion of the abasic site **2.44a** to the oligonucleotides **2.45a**, **2.46** and the MeFAPy-Gua (**2.21a** and **2.21b)** (Figure **2.14**, Panel A) (Figure **2-15**).

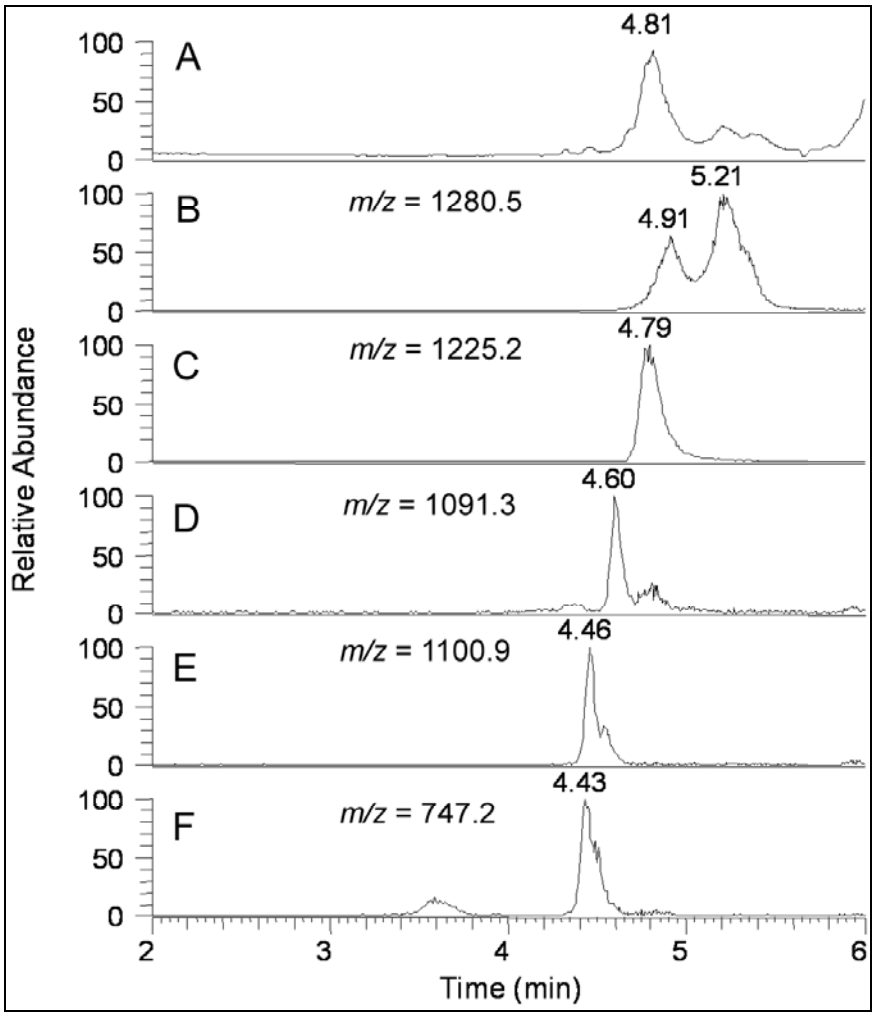



**Figure 2-15**. LC-ESI/MS analysis of the stability of oligonucleotide 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2.43a)** in phosphate buffer pH 6.5. **A**. Total ion current (TIC). **B**. Select ion profile of **2.43a** with *m/z* =1280.5 ([M-3H]). **C**. Select ion profile of **2.44a** with *m/z* =1225.2 ([M-3H]). **D**. Select ion profile of **2.45a** with *m/z* = 1091.3 ([M-2H]). **E**. Select ion profile of **2.44a** as hydrated form with *m/z* =1100.9 ([M-2H]). **F**. Select ion profile of **2.46** with *m/z* =747.2 ([M-2H]).

Oligonucleotide **2.43a** was then heated at 90 °C for 1 h and the changes in its composition were monitored by HPLC (Figure **2.16**). After all of the starting material was converted to the abasic site oligonucleotide **2.44a** (Figure **2-16**, Panel C), an aliquot was taken and added to a solution of T4-pdg (Endo V), which cleaves abasic sites. The T4-pdg reaction mixture was set at 37 °C for 24 h, while the remaining solution of oligonucleotide **2.44a** was heated at 90 °C for additional 1.5 h. The HPLC analysis of both of the above reaction mixtures showed (HPLC traces of the T4-pdg (Endo V) reaction mixture are not shown) that oligonucleotide **2.44a** decomposed to oligonucleotides **2.45a** and **2.46** (Figure **2-16**, Panel D). In a control experiment, oligonucleotide **2.43a** was incubated with T4-pdg (Endo V) in pH 7.0 buffer and after 24 h aliquot was taken and analyzed by LC-ESI/MC, which showed that no degradation of oligonucleotide **2.43a**. This further confirms that the MeFAPy-dGuo (**2.05**) containing oligonucleotides are stable in pH 7.0.

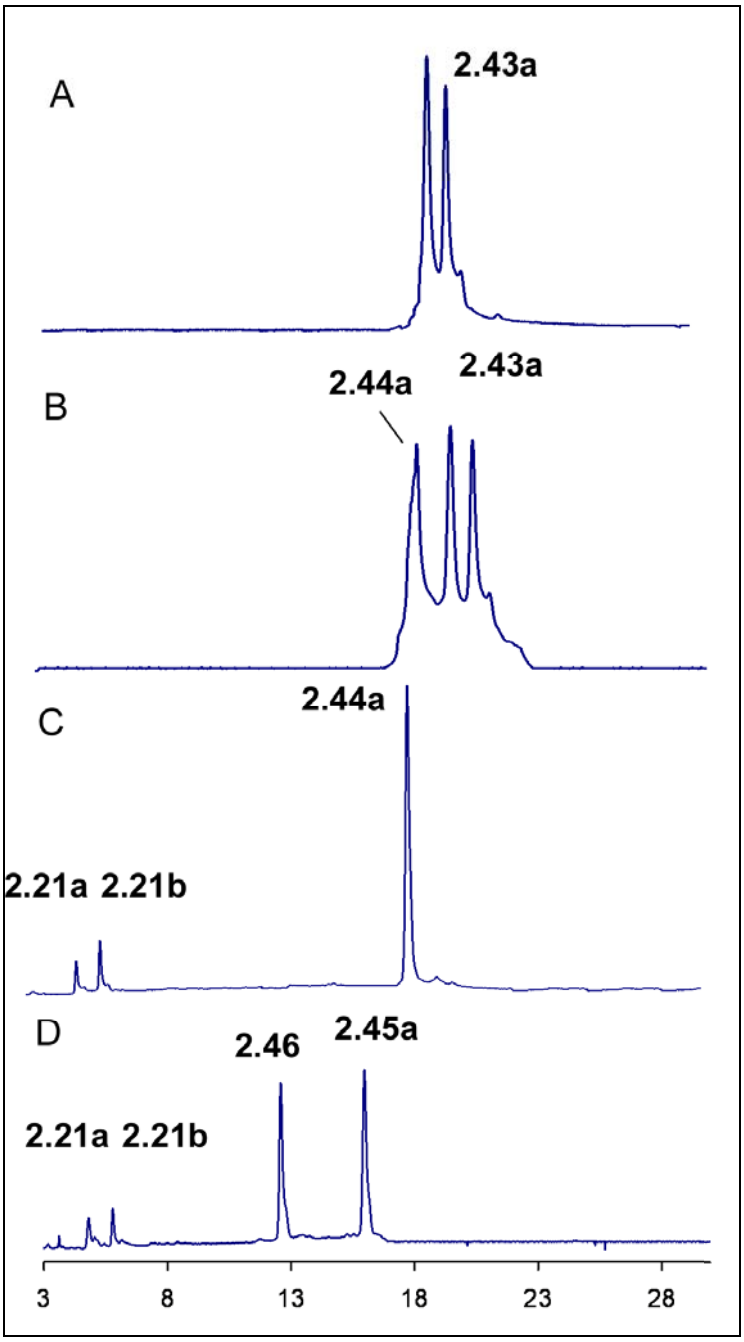



**Figure 2-16**. HPLC analysis of the stability of MeFAPy-dGuo oligonucleotide **2.43a** in water at 90 °C. **A**. Analysis of the oligonucleotide **2.43a** as pure material; **B**. Analysis of the oligonucleotide **2.43a** after 2 days at 25 °C, **2.44a** is the abasic site oligonucleotide; **C**. Analysis of the oligonucleotide **2.43a** after 1 h heating at 90 °C, **2.44a** is the abasic site oligonucleotide, **2.21a** and **2.21b** are the MeFAPy-Gua. **D**. Analysis of the oligonucleotide **2.43a** after 2.5 h heating at 90 °C, **2.44a** is the abasic site oligonucleotide, **2.21a** and **2.21b** are the MeFAPy-Gua**.**

Oligonucleotide **2.43a** was incubated in deionized water and treated with $Na(CN)BH_3$ in order to trap the imine of the ring-opened MeFAPy-dGuo nucleotide **2.47**, a likely intermediate in the degradation of the MeFAPy-dGuo (**2.05**) containing nucleotides (Figure **2-17**).

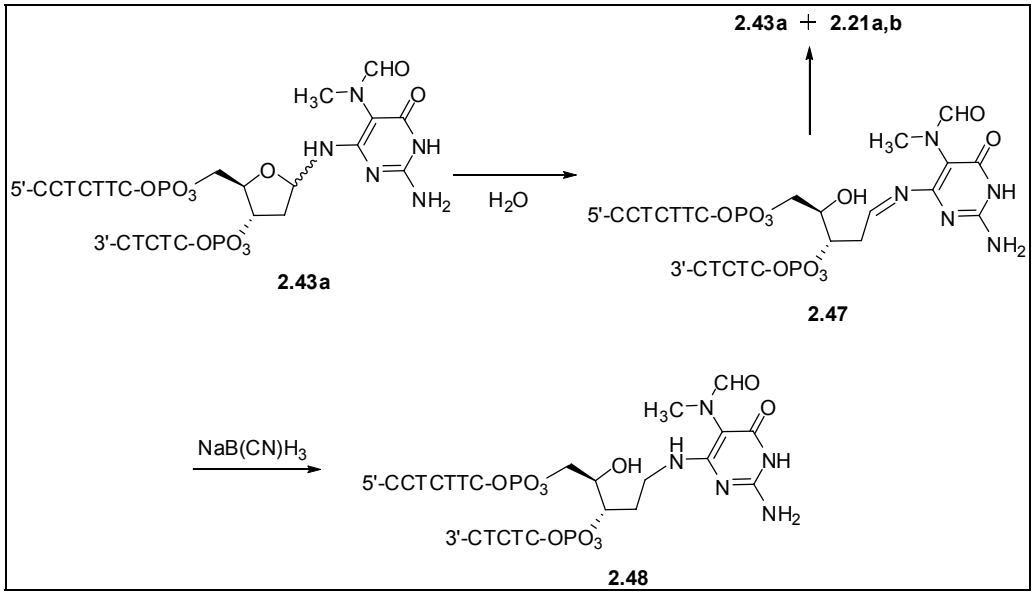



**Figure 2-17**. Formation of the oligonucleotide **2.48** via reduction of the imine oligonucleotide **2.47** with $Na(CN)BH_3$.

After 48 h, the reaction mixture was analyzed by MALDI-TOF-MS and then subjected to acid hydrolysis and alkaline enzyme digest. The MOLDI-TOF-MS analysis of the reaction mixture did not show the formation of **2.45a** and **2.46**, which was otherwise observed when **2.43a** was kept in water for 48 h (see Appendix for MALDI-TOF MS spectra). Acid hydrolysis of the $Na(CN)BH_3$ reaction mixture (containing the oligonucleotide **2.48**) was

performed with 0.1 M HCl at 37 °C for overnight, which are the optimal conditions for hydrolysis of $N^5$-alkylated FAPy-dGuo oligonucleotides (*31, 33*). The HPLC analysis of the acid hydrolysis did not show the release of MeFAPy-Gua (**2.21a** and **2.21b**) (Figure **2-18**, Panel B), which are products of acid hydrolysis of MeFAPy-dGuo (**2.05**) containing oligonucleotides. In a control experiment, the acid hydrolysis of oligonucleotide **2.43a** resulted in the formation of **2.21a** and **2.21b** (Figure **2-18**, Panel A).

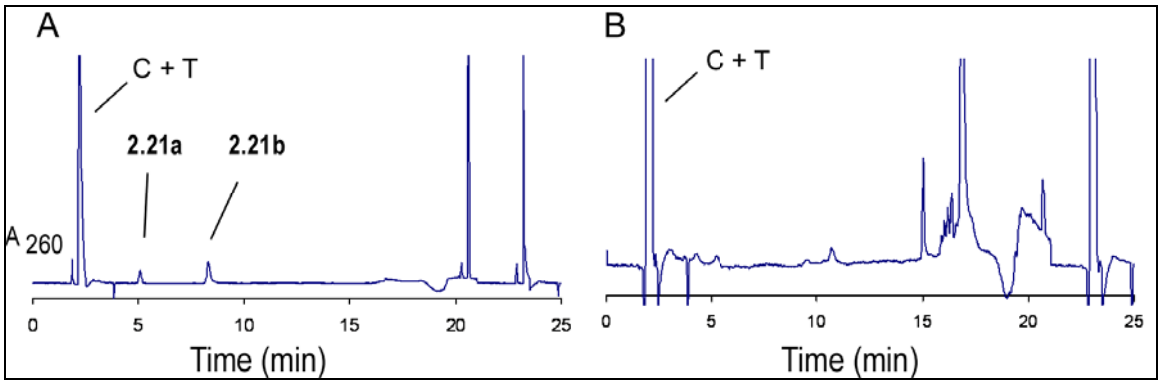



**Figure 2-18**. HPLC analysis of the acid hydrolysis of: **A**. oligonucleotide **2.43a**. B. the oligonucleotide **2.48** formed via reduction of **2.43a** with $Na(CN)BH_3$.

The enzyme digestion of the $Na(CN)BH_3$ reaction mixture was analyzed by LC-ESI/MS/MS. The MS analysis showed the formation of dTyd, dCyt, one unknown species, but not a formation of the MeFAPy-Gua (**2.21a** and **2.21b**) which are otherwise formed when **2.43a** is subjected to enzyme digest (*30*) (Figure **2-19)**. The unknown species, *m/z* 591.1, was identified as

a dinucleotide (**2.49**) being the reduced MeFAPy-dGuo (**2.05**) nucleoside and dCyd (Figure **2-20**). The lack of digestion of this nucleotide by alkaline phosphatase is probably due to the enzymes not being able to recognize the ring-opened reduced sugar and consequently unable to digest the dinucleotide.

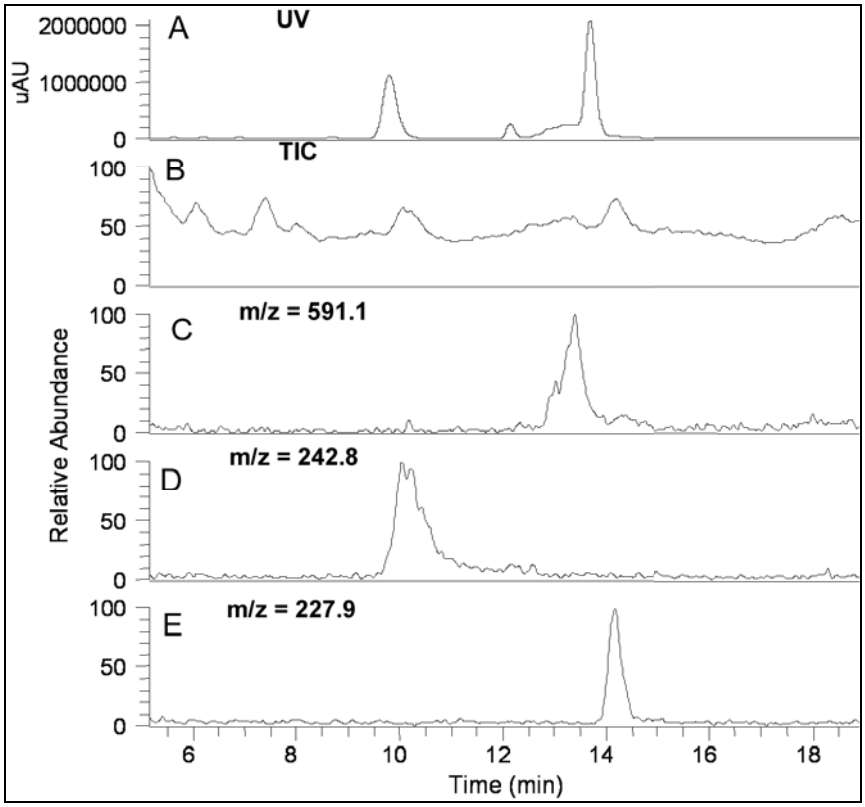



**Figure 2-19**. LC-ESI/MS analyses of the enzyme digest of oligonucleotide **2.48**. **A**. UV traces **B**. Total ion current. **C**. Select ion profile of dinucleotide **2.49** with *m/z* 591.1 ([M-H]). **D**. Select ion profile of dTyd with *m/z* 242.8 ([M-H]). **E**. Select ion profile of dCyt with *m/z* 227.9 ([M-H]).

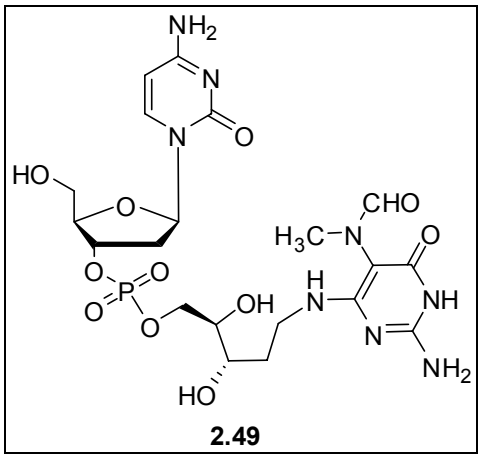



**Figure 2-20**. The structure of dinucleotide **2.49**.

Based on the analysis of the reduction reaction mixture by MALDI-TOF-MS, acid hydrolysis and enzyme digest, it was concluded that the ring-opened imine of the MeFAPy-dGuo containing oligonucleotide **2.47** was successfully reduced to oligonucleotide **2.48**.

Our success in trapping the oligonucleotide imine **2.47** motivated us to attempt to trap the imine **2.16,** which is involved in the anomerization of MeFAPy-dGuo nucleoside **2.15** (Figure **2-04**). A mixture of six- and five-membered MeFAPy-dGuo nucleoside was dissolved in water and 100 eq. $Na(CN)BH_3$ was added in equal portions (33 times of 3.03 eq.) for a period of one month. The resulting mixture was purified by HPLC to afford compound **2.50**, which arises from reduction of imine **2.16** (Figure **2-21**). Interestingly, the MS fragmentations of the mixture of six- and five-membered MeFAPy-dGuo nucleoside **(2.15**) and the compound **2.50** were very different, which also help to identify the structure of **2.50** (Figure **2-22**). Fragmentation of the

MeFAPy-dGuo nucleoside (**2.15**), *m/z* of 299.9, at 25 % relative energy (RE) gave a product ion with *m/z* 184.2, which is assigned as lost of sugar. Furthermore, product ion with *m/z* 184.2 at 25 % RE gave a product ion *m/z* 156.2, which is assigned as loss of CO. The compound **2.50**, *m/z* of 302.0, at 27 % relative energy gave a product ion with *m/z* 274.1, which is assigned as a los of CO. Product ion with *m/z* 274.1 at 25 % RE gave a product ion with *m/z* 256.2, which is assigned as a loss of $H_2O$.

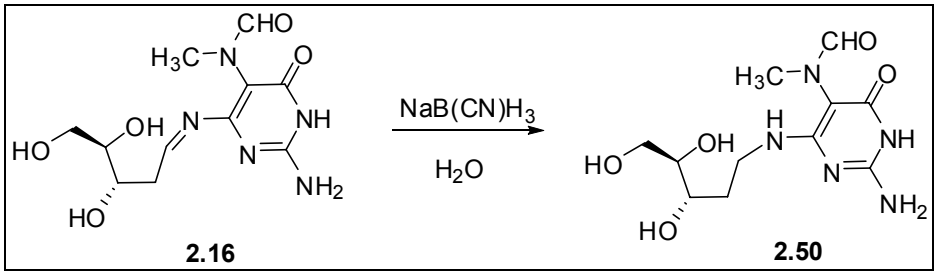



**Figure 2-21**. Formation of **2.50** by reduction of imine **2.16** with $Na(CN)BH_3$.

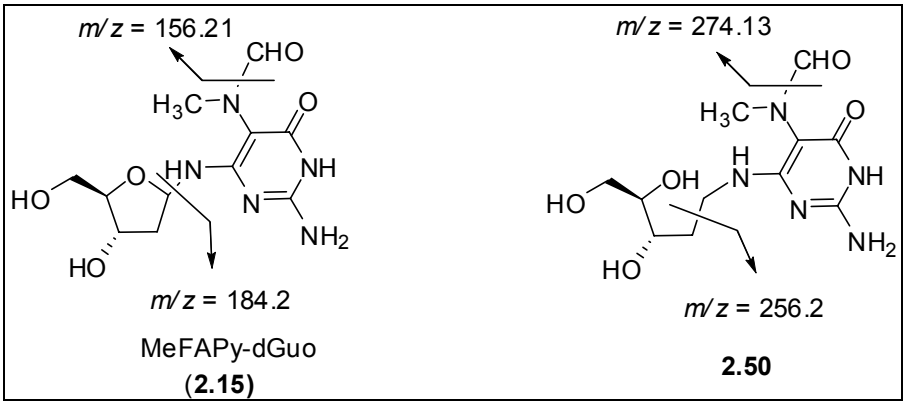



**Figure 2-22**. MS fragmentation of a mixture of 5- and 6-membered ring MeFAPy-dGuo (**2.15**) nucleoside and compound **2.50**.

## Polymerase Bypass of the MeFAPy-dGuo (2.05) Adduct in Oligonucleotides

The *in vitro* bypass and extension of the MeFAPy-dGuo (**2.05**) adduct was examined with prokaryotic and eukaryotic DNA polymerases. The prokaryotic enzymes were *E. coli* DNA polymerase I Klenow fragment exo$^-$ (Kf$^-$), *E. coli* DNA polymerase II exo$^-$ (Pol II$^-$) and *S. solfataricus* P2 DNA polymerase IV (Dpo4). Kf$^-$ and Pol II$^-$ are highly conserved proteins that have roles in translesion synthesis across specific DNA adducts in the bacterial chromosome. Klenow fragment has two distinct enzymatic activities, a 5'→3' polymerase and a 3'→5' exonuclease activity. Kf$^-$ has significantly diminished 3'→5' exonuclease activity through a mutations in the 3'→5'-exonuclease active site. DNA polymerase II is a prototype of the B-family of polymerases and is SOS-inducible (*57, 58*). Dpo4 is phylogenetically related to DinB polymerases and is a useful model for Y-family (lesion bypass) polymerase (*59, 60*). It is a thermostable enzyme and is active up to 70 °C, but can also synthesize DNA at 37 °C (*60*). Dpo4 exhibits misinsertion fidelities in the range of 8 x $10^{-3}$ to 3 x $10^{-4}$ and can facilitate translesion replication on damaged DNA templates containing an abasic site, a *cis-syn* thymine–thymine dimer, the C-8 acetylaminofluorene adduct and cis-platin-guanine (*60*).

The eukaryotic DNA polymerases examined were the human DNA polymerases pol $\delta$, $\eta$, $\kappa$ and $\iota$. Pol $\delta$ is a replicative DNA polymerase that replaces pol $\alpha$, which starts the DNA replication, and completes leading synthesis. It consists of three subunits of 125, 55 and 40 kDa. The 125 kDa

subunit exhibits 3'→5' exonuclease activity and the 40 kDa subunit interacts with PCNA (*61*).

Human pol *η* replicates through a *cis-syn d*Tyd-dTyd dimer by inserting two dATP's opposite the two dTyd's of the dimer with nearly the same efficiency and fidelity as opposite the two undamaged dTyd's (*62*). Defects in the gene coding the human pol *η* are responsible for a cancer predisposition disorder, the *xeroderma pigmentosum* variant (XPV) (*63, 64*). Human pol *η* also plays an important role in efficient and accurate replication through the 8-oxoguanine (8-oxo-dGuo) lesion (*65, 66*), incorporating a dCTP opposite the lesion; however, a low but significant level of dATP is also inserted.

Human pol *κ* misincorporates nucleotides with a frequency of about $10^{-3}$ to $10^{-4}$ opposite all four template bases (*67*). Human pol *κ* inserts dATP opposite an 8-oxo-dGuo lesion more efficiently than dCTP, and it can extend with an about 20-40-fold reduction in efficiency (*68*). Interestingly, human pol *κ* can extend from a dCyt incoporated opposite past $N^2$- dGuo adducts by human pol *ι (69)*.

Human pol *ι* incorporates nucleotides opposite template purines with a much higher efficiency and fidelity than opposite template pyrimidines (*70-72*). In some cases, pol *ι* incorporates nucleotides opposite the DNA adduct, but is unable to carry out the subsequent extension step (*70*). Human pol *ι* plays an important role in the bypass of γ-HO-PdGuo (*69*), an $N^2$-dGuo adduct that arise from the reaction of dGuo with acrolein, which is an end product of lipid peroxidation. Human pol *ι* incorporates dCTP opposite the γ-

HO-PdGuo adduct, but is unable to extend past the adduct; the extension is performed by pol $\kappa$. The sequential action of pols $\iota$ and $\kappa$ may provide cells with an efficient pathway for the error-free bypass of the various $N^2$-dG adducts that are formed by products of lipid peroxidation.

The MeFAPy-dGuo (**2.05**) adduct was incorporated at position-5 of 28mer oligonucleotide (**2.52a**); this oligonucleotide was annealed to a complementary 23mer (**2.53**) (-1) primer strand which was 5′-labeled with $^{32}$P. An oligonucleotide (28mer) (**2.52b**) containing the pyranose form of the MeFAPy-dGuo was also synthesized using the "long" deprotection cycle (Figure **2-23**).

| |
|---|
| 3'-CCCCCGCTACGAGCATTCCTAAGGTACT-5' unmodified oligonucleotide (**2.51**) |
| 3'-CCCCCGCTACGAGCATTCCTAAG-(MeFAPy-dGuo)-TACT-5' (**2.52a**) (furanose form) |
| 3'-CCCCCGCTACGAGCATTCCTAAG-(MeFAPy-dGuo)-TACT-5' (**2.52b**) (pyranose form) |
| 5'-GGGGGCGATGCTCGTAAGGATTC-3' primer (-1) (**2.53**) |

**Figure 2-23**. The oligonucleotide sequences used for running *in vitro* DNA bypass and extension of the MeFAPy-dGuo (**2.05**) lesion with $^{32}$P imaging.

The steady-state insertion rates for each individual dNTP's were determined. Reaction times, enzyme concentrations and dNTP concentrations were chosen so that the maximium product formation would be ~ 20% of the substrate concentration. The steady-state rates versus dNTP concentration were fit using nonlinear regression (hyperbolic fits) in GraphPad

Prism (version 4.0, GraphPad, San Diego, CA) for the estimation of $k_{cat}$ and $K_m$ values (Figure **2-24**). The nonlinear regression (hyperbolic fits) has the equation: $Y=k_{cat} \times X/(k_m + X)$, which describes the binding of a ligand to a receptor that follows the law of mass action. $k_{cat}$ is the turnover number and $K_m$ is the concentration of ligand required to reach half-maximal rate.

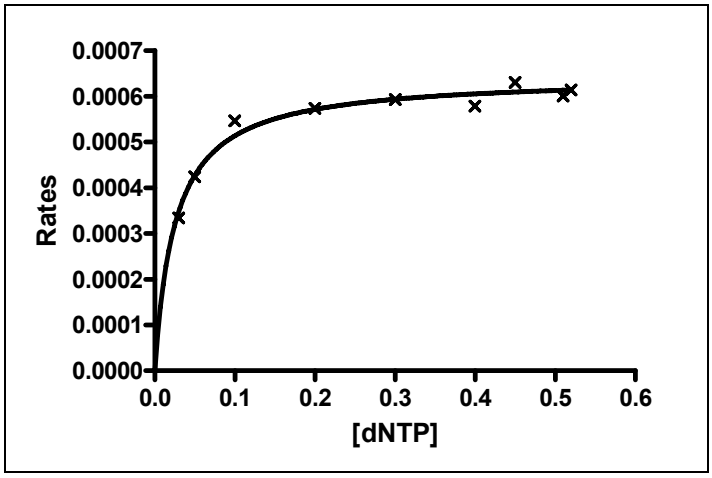



**Figure 2-24**. Example of steady-state kinetic curve: the incorporation of dCTP opposite the MeFAPy-dGuo adduct (**2.05**) in the oligonucleotide **2.52a** catalyzed by Dpo4.

.

### *LC-ESI/MS/MS Sequencing of the Full-length Extension Products Past the MeFAPy-dGuo Lesion (2.05) in Oligonucleotides Using Biotinated Primers*

The full-length extension assays were performed in a similar manner to a previously published procedure with the following modifications (*73*). The MeFAPy-dGuo oligonucleotide (**2.52a**) was annealed to a complementary 23mer (**2.54**) (-1) primer, which was 5′-biotinated. Ten dTyd's were used as a spacer between the biotin and the complementary sequence (Figure **2-25**); a dUrd was incorporated at position 21. After the extension reaction took place, the extended primer was added to a solution of streptavidin coated beads. After gentle shaking for 2 h, the streptavidin beads were thoroughly washed to remove the other components of the reaction mixture. The streptavidin coated beads were incubated with uridine DNA glycosylase (UDG) for three hours, than washed extensively. The UDG treatment results in the removal of the dUrd and the formation of abasic site of the extended primer. The depurinated oligonucleotides were treated with piperidine to form two strands, one being the 3′-end of the primer with 5′-phosphate and the other being the 5′-end of the primer. The 3′-fragment contains the sequence information of interest from the extension reaction. The beads were removed from the liquid containing the extension product(s) and washed several times with water, which was added to the filtrate. The collected solution was lyophilized to dryness, dissolved in water and analyzed by LC-ESI/MS/MS. The 3′-fragment of the extended primer was sequenced by collision-induced dissociation (CID) and mass analysis of the fragments.

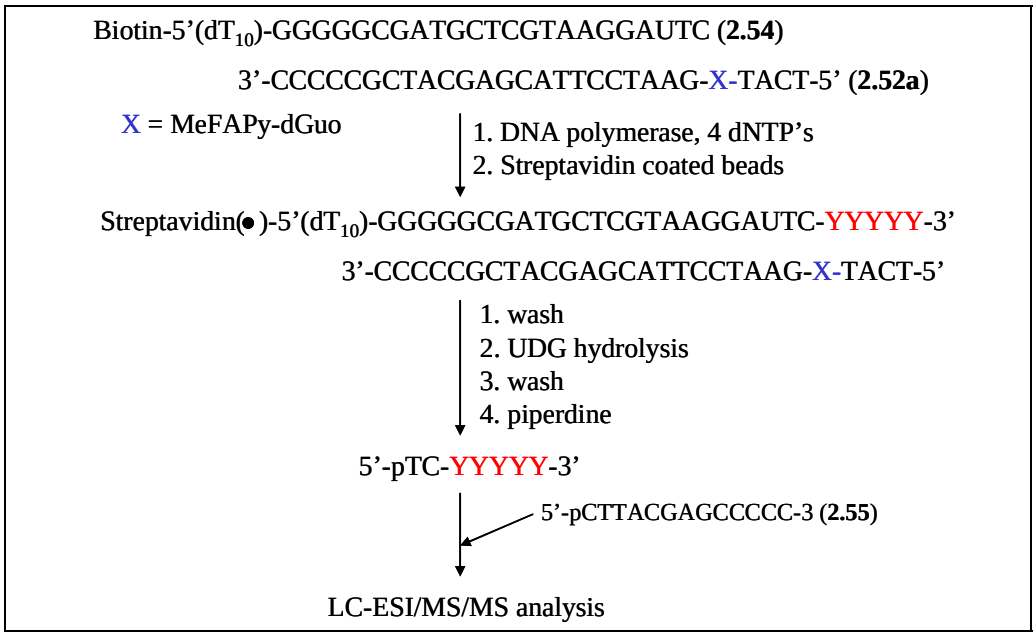



**Figure 2-25**. LC-ESI/MS/MS sequencing of the full-length extension adduct past the MeFAPy-dGuo adduct (**2.05**) in oligonucleotides.

After the fragments were identified, the proposed oligonucleotide sequence was purchased from Midland Certified Reagents as an authentic standard and subjected to the same LC-ESI/MS/MS analysis. The CID's of the commercially obtained products were compared with those of the reaction product(s) to show the same pattern of fragmentation. To establish the percentage of formation of the full-length extension product(s) another oligonucleotide, 5′-pCTTACGAGCCCCC-3′ (**2.55)**, was used as an internal standard. This standard was added prior to the LC-ESI/MS/MS analysis. The standard had molecular ion peak with *m*/*z* 1310.5 ([M-3H]). The standard calibration curves were constructed using seven concentrations of the

corresponding analyte (purchased from Midland Certified Reagents) versus a fixed amount of the standard.

The use of biotinated primers allowed the clean up of the extension products from other components, such as salts, dNTP's and enzymes, and improved the sensitivity of the LC-ESI/MS/MS method. The use of the 13mer (**2.55**) allowed the calculation of the absolute percentage of formation of the full-length extension products.

### *DNA Bypass of MeFAPy-dGuo Adduct (2.05) Catalyzed by Prokaryotic DNA Polymerases*

The MeFAPy-dGuo containing oligonucleotide (**2.52a**) was annealed to a complimentary 23mer (**2.54**) (-1) primer strand, which was 5'-labeled with $^{32}P$. The single nucleotide insertion and the full length extension reactions in the presence of all four dNTP's using prokaryotic DNA polymerases were run at fixed amount of the enzyme and increasing dNTP(s) concentration(s). The single nucleotide incorporation and the full-length extension reactions catalyzed by prokaryotic DNA polymerases with the unmodified sequence (**2.51**) were performed, in the same manner as the adducted sequence, but with two-fold less concentration of the enzyme and the reaction time.

Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in the Oligonucleotide **2.52a** Catalyzed by $Kf^{-}$

The Klenow fragment $exo^{-}$ ($Kf^{-}$) inserted all four dNTP's opposite the MeFAPy-dGuo adduct (**2.05**) (Figure **2-26**, Panel A), which is in accordance with previously published data (*45, 49, 50*). In contrast, $Kf^{-}$ incorporated only dCTP in the unmodified oligonucleotide (Figure **2-26**, Panel B). To measure the mis-incorporation frequencies by $Kf^{-}$ of the MeFAPy-dGuo containing oligonucleotide **2.52a** relative to the unmodified oligonucleotide, a steady-state kinetic assay was performed. While the insertion efficiencies for dCTP and dATP opposite the MeFAPy-dGuo adduct were about 8 and 1.4 times lower, respectively, the insertion efficiencies for dGTP, dTTP were about 3.7 and 25 times higher, respectively, than for the unmodified template (Table **2-04**).

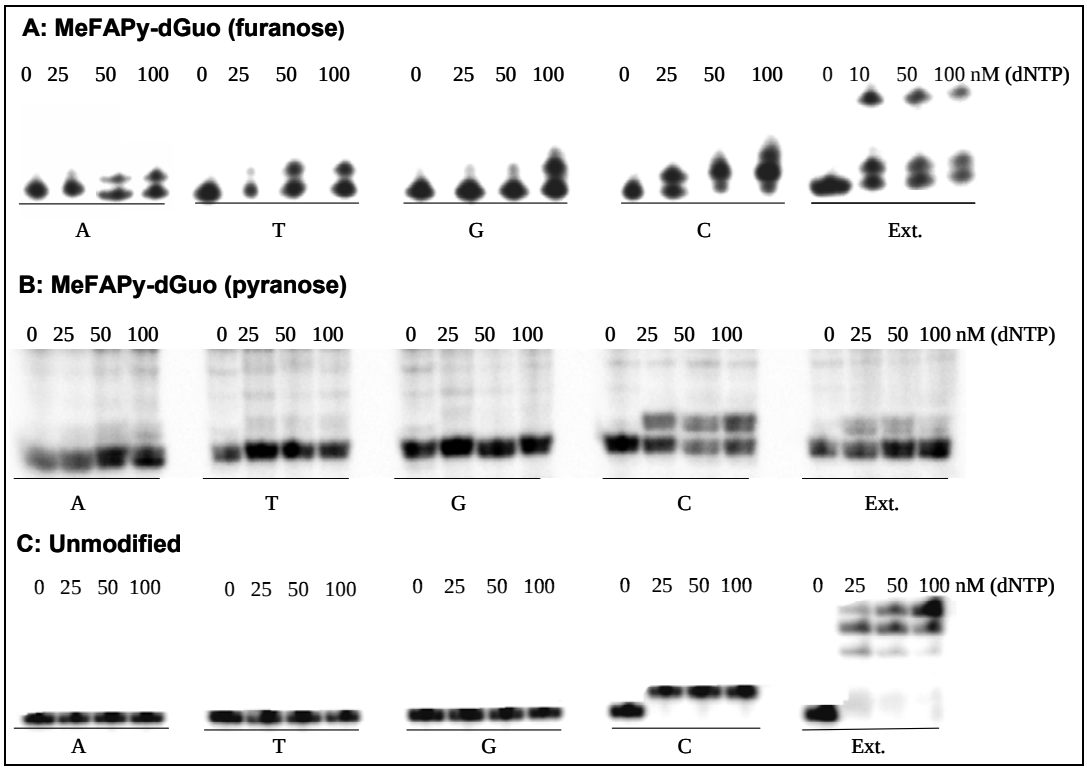



**Figure 2-26**. Single and full-length incorporation assays catalyzed by Kf⁻: **A**. Incorporation across the MeFAPy-dGuo adduct (furanose form) in oligonucleotide **2.52a**; **B**. Incorporation across the MeFAPy-dGuo adduct (pyranose form) in oligonucleotide **2.52b**; **C**. Incorporation of dNTP(s) in the unmodified oligonucleotide (**2.51**).

**Table 2-04**. Steady-state kinetics parameters for Kf⁻.

| Template | dNTP | Km (μM) | $k_{cat}$ ($s^{-1}$ x $10^{-3}$) | $k_{cat}/k_m$ ($\mu M^{-1}$ x $s^{-1}$x $10^{-3}$) | f[a] |
|---|---|---|---|---|---|
| MeFAPy-dGuo | C | 0.07 ± 0.02 | 2.0 ± 0.1 | 30 | 1 |
| | T | 0.33 ± 0.09 | 2.8 ± 0.3 | 8 | 0.28 |
| | G | 0.19 ± 0.02 | 1.7 ± 0.1 | 9 | 0.29 |
| | A | 1.1 ± 0.7 | 1.6 ± 0.2 | 1.5 | 0.05 |
| dGuo | C | 0.013 ± 0.001 | 3.3 ± 0.1 | 260 | 1 |
| | T | 3.4 ± 0.2 | 1.2 ± 0.3 | 0.4 | 0.002 |
| | G | 2.11 ± 0.02 | 5.2 ± 0.3 | 2.4 | 0.009 |
| | A | 2.8 ± 0.9 | 5.5 ± 0.5 | 2 | 0.007 |

[a]**f** - misincorporation frequency = $(k_{cat}/K_m)_{\text{incorporated dNTP}}/(K_{cat}/K_m)_{\text{correct dNTP (dCTP)}}$

In the presence of all four dNTP's, Kf⁻ extended past the MeFAPy-dGuo adduct, although there were a significant pause after incorporation of one nucleotide past the lesion. Experiments with 0-primers with Ado (**2.56**), Tyd (**2.57**) and Guo (**2.58**) showed that the full-length extension was blocked following the mis-incorporation of dGTP, dTTP and dATP opposite the adduct (Figure **2-27**). The inhibition of the extension after mis-incorporation of dATP and dGTP opposite the MeFAPy-dGuo adduct was also reported by Asagoshi and coworkers (*46*).

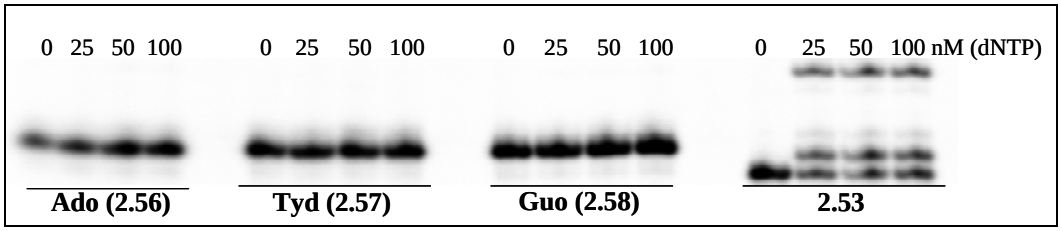



**Figure 2-27**. Full-length incorporation assays with the MeFAPy-dGuo containing oligonucleotide (**2.52a**) catalyzed by Kf⁻ and 0-primers with Ado (**2.56**), Tyd (**2.57**), Guo (**2.58**) and -1 primer **2.53** in the presence of the four dNTP's; 5'-GGGGGCGATGCTCGTAAGGATTCA-3' (**2.56**); 5'-GGGGGC GATGCTCGTA-AGGATTCT-3' (**2.57**); 5'-GGGGGCGATGCTCGTAAGG ATTCG-3' (**2.58**); 5'-GGGGGCGATG-CTCGTAAGGAT TC-3' (**2.53**).

In contrast to the polymerase reactions with the MeFAPy-dGuo (furanose form) oligonucleotide (**2.52a**), Kf⁻ incorporated only dCTP opposite to the MeFAPy-dGuo (pyranose form) adduct in oligonucleotide **2.52b** but further extension was completely blocked (Figure **2-26**, Panel B).

The product(s) of the extension reaction catalyzed by Kf⁻ was analyzed by the LC-ESI/MS/MS method described above. The ESI mass spectrum of the reaction mixture showed a molecular ion peaks with *m*/*z* of 719.2 and 1078.8 (Figure **2-28**). This two molecular ion peaks correspond to M-3H and M-2H ions of the same oligonucleotide, of *m/z* 2159.6 Da (calculated for M-2H). The possible composition of the *m/z* 1078.8 ± 2 Da (M-2H) oligonucleotide was 1 phosphate, 2 dCyd's, 2 dTyd's, 2 Ado's and 1 dGuo. Given the composition and the sequence of the template, the most likely extension product was 5′-pTCCATGA-3′.

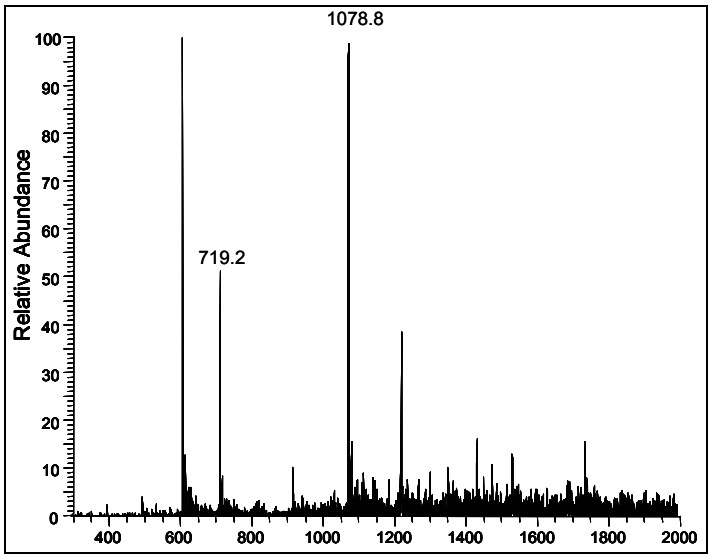



**Figure 2-28**. TIC spectrum of the LC-ESI/MS/MS analysis of the Kf⁻ extension reaction product past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** and the biotinated primer **2.54**; molecular ion peaks with *m/z* 719.2 and *m/z* 1078.8 correspond to M-3H and M-2H, respectively, of the extension product with mass 2159.6 Da identified as the 5′-pTCCATGA-3′.

The CID of the molecular ion peak *m*/*z* 1078.8 matched well with the predicted CID spectrum for the sequence 5′-pTCCATGA-3′, which represents an error-free bypass and extension (Figure **2-29**, Table **2-05**). The predicted CID spectrum was obtained from the "Mongo Oligo Mass Calculator" v2.06 maintained by the Department of Medicinal Chemistry at the University of Utah, which is available online at http://library.med.utah.edu/masspec/ mongo .htmpthe.

An authentic standard of the oligonucleotide 5′-pTCCATGA-3′ was purchased from Midland Certified Reagents and subjected to LC-ESI/MS/MS analysis under similar conditions as the sample from the extension reaction (Appendix).

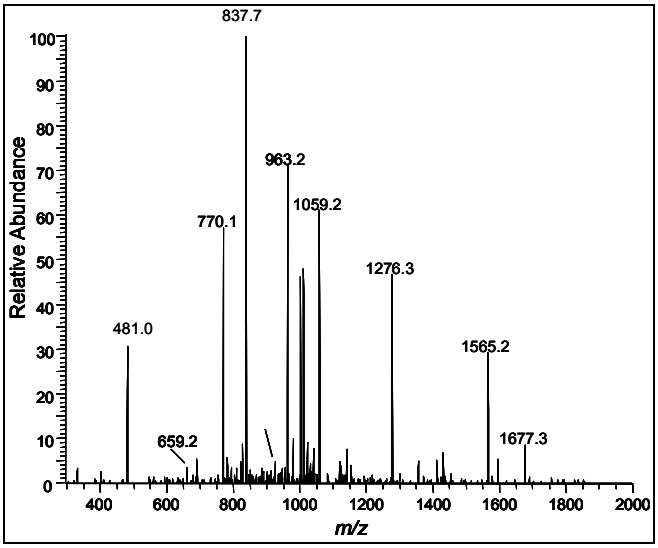



**Figure 2-29**. CID spectrum of the molecular ion peak *m/z* 1078.8 of the $Kf^{-}$ extension product, identified as the 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleoitde **2.52a**.

**Table 2-05**. Observed and theoretical fragmentation for the Kf⁻ extension product *m/z* 1078.83, identified as the 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in the oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481 | 481.0 |
| 5'-pTC ($a_3$-B) | 770.1 | 770.8 |
| 5'-pTCC ($a_4$-B) | 1059.2 | 1059.1 |
| 5'-pTCCAT ($a_6$-B) | 1677.3 | 1676.2 |
| ($a_6$-B, -2) | 837.7 | 837.6 |
| pCCATGA-3' ($w_6$, -2) | 926.8 | 926.7 |
| pCATGA-3' ($w_5$) | 1565.2 | 1565.2 |
| pATGA-3' ($w_4$) | 1276.3 | 1276.2 |
| pTGA-3' ($w_3$) | 963.2 | 963.1 |
| pGA-3' ($w_2$) | 659.2 | 659.1 |

The CID of the sequence was identical to the one found in the Kf⁻ full-length extension, thereby proving its sequence (Table **2-05**).

To establish the yield of the error-free bypass product the oligonucleotide, 5′-pCTTACGAGCCCCC-3′ (**2.55**), was used as internal standard, which was added prior to MS analysis. The TIC spectrum is shown in Figure **2.30** with *m/z* 982.9 and 1310.5 being the M-4H and M-3H, respectively, of the internal standard. A calibration curve for the 5′-pTCCATGA-3′ and the standard **2.55** was constructed and showed good linearity ($R^2$ = 0.997) (Figure **2-31**).

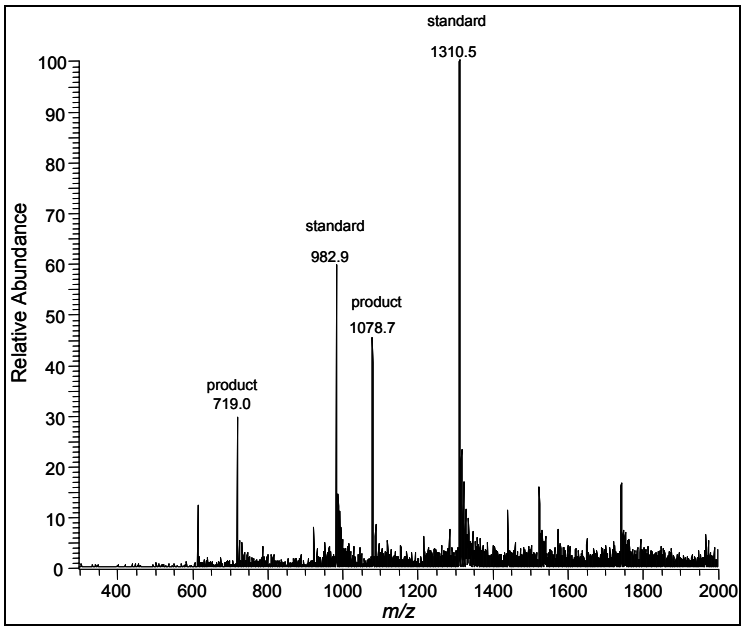



**Figure 2-30** TIC spectrum of the LC-ESI/MS/MS analysis of the $Kf^{-}$ extension product, identified as the 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in the oligonucleotide **2.52a** in the presence of internal standard 5′-pCTTACGAGCCCCC-3′ (**2.55**); molecular ion peaks with *m/z* 719.0 and *m/z* 1078.7 correspond to M-3H and M-2H, respectively, of the extension product with mass 2159.5 Da; molecular ion peaks with *m/z* 982.9 and *m/z* 1310.5 correspond to M-4H and M-3H, respectively, of the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.55**).

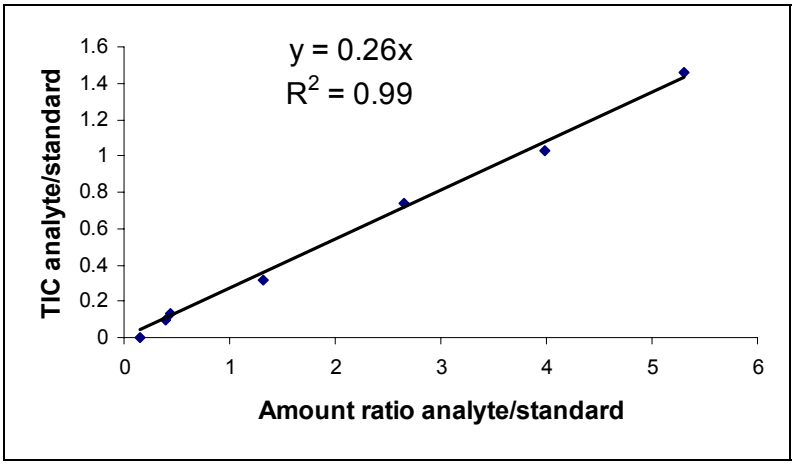



**Figure 2-31**. The calibration curve for the 5′-pTCCATGA-3′ and the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.55**).

The estimated yield of the error-free bypass product 5′-pTCCATGA-3′ was 26%. This low yield is due to the inefficient extension of the product after initial insertion of dCTP and the fact that the extension is blocked after the initial mis-incorporation of dATP, dTTP and dGTP.

## Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in the Oligonucleotide **2.52a** Catalyzed by Pol II$^{-}$

Pol II$^{-}$ incorporated all four dNTP's opposite the MeFAPy-dGuo lesion in the oligonucleotide **2.52a** while only dCTP was inserted for the unadducted oligonucleotide (Figure **2-32**).

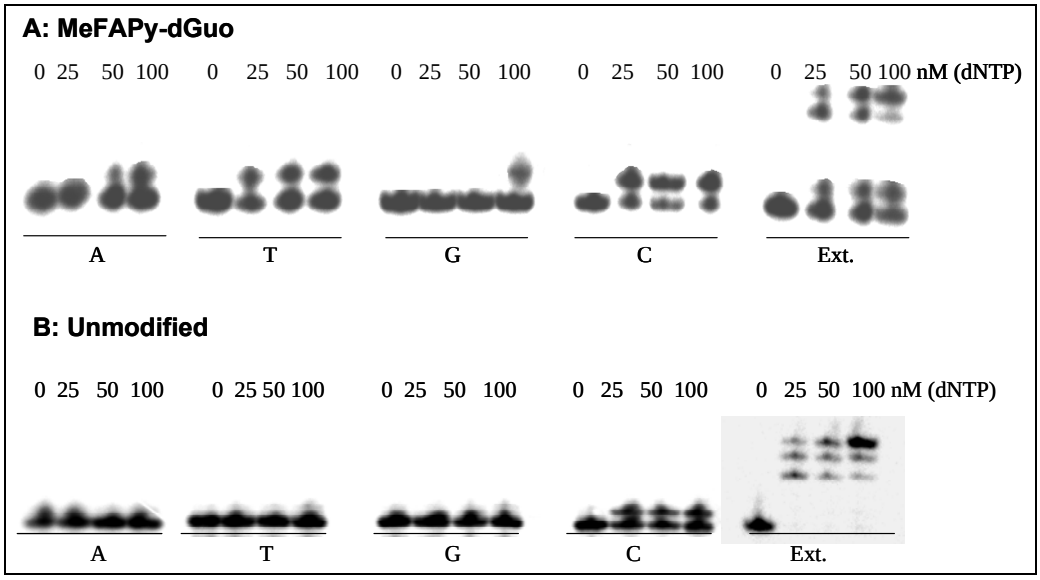



**Figure 2-32**. Single and full-length incorporation assays catalyzed by Pol II$^{-}$: **A**. Incorporation across the MeFAPy-dGuo adduct in oligonucleotide **2.52a**; **B**. Incorporation of dNTP(s) in the unmodified oligonucleotide **2.51**.

Steady-state kinetic assays were performed to obtain more quantitative information about the tendency for pol II$^{-}$ to misincorporate nucleotides opposite the MeFAPy-dGuo lesion. The insertion efficiency for dCTP opposite the MeFAPy-dGuo was nearly the same as the unmodified template, whereas the insertion efficiencies for dTTP, dGTP and dATP were 7, 14 and 80 times higher, respectively, than the unmodified template (Table **2-06**).

**Table 2-06**. Steady-state kinetics parameters for Pol II$^{-}$.

| Template | dNTP | Km (μM) | $k_{cat}$ ($s^{-1}$ x $10^{-3}$) | $k_{cat}/k_m$ ( $M^{-1}$ x $s^{-1}$x $10^{-3}$) | f[a] |
|---|---|---|---|---|---|
| MeFAPy-dGuo | C | 0.016 ± 0.005 | 1.1 ± 0.3 | 69 | 1 |
| | T | 4.8 ± 0.3 | 7.7 ± 0.8 | 2 | 0.02 |
| | G | 7.2 ± 0.1 | 5.1± 0.5 | 0.7 | 0.01 |
| | A | 3.9 ± 0.4 | 6.4 ± 0.5 | 1.6 | 0.02 |
| dGuo | C | 0.012 ± 0.002 | 0.91 ± 0.07 | 76 | 1 |
| | T | 3.3 ± 0.2 | 0.82 ± 0.07 | 0.2 | 0.003 |
| | G | 13.7 ± 0.2 | 0.75 ± 0.09 | 0.05 | 0.001 |
| | A | 15.0 ± 0.6 | 0.29 ± 0.03 | 0.02 | 0.0002 |

[a]**f** - misincorporation frequency = $(k_{cat}/K_m)_{\text{incorporated dNTP}}/(K_{cat}/K_m)_{\text{correct dNTP (dCTP)}}$

Pol II$^{-}$ was able to extend past the MeFAPy-dGuo in the oligonucleotide **2.52a** in the presence of all four dNTP's, but there was blocking product(s) after incorporation of the base opposite the MeFAPy-dGuo adduct. The full-length extension was examined with the 0-primers Ado (**2.56**), Tyd (**2.56**), and Guo (**2.58)** opposite the MeFAPy-dGuo adduct and showed that the extension was completely blocked (Figure **2-33**).

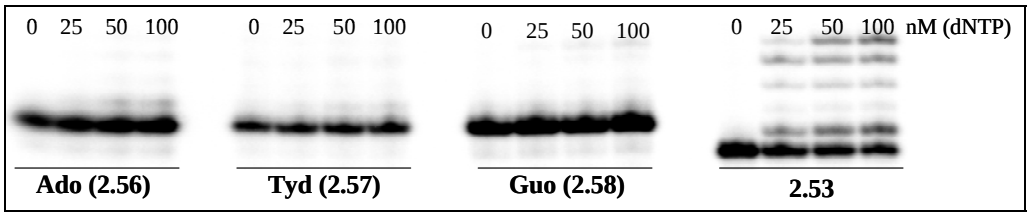



**Figure 2-33**. Full-length incorporation assays with MeFAPy-dGuo containing oligonucleotide (**2.52a**) by pol II⁻ and 0-primers Ado (**2.56**), Tyd (**2.57**),Guo (**2.58**) and -1 primer **2.53** in the presence of the four dNTP's; 5'-GGGGGCGAT-GCTCGTAAGGATTCA-3' (**2.56**); 5'-GGGGGCGATGCTCG TAAGGATTCT-3' (**2.57**); 5'-GGGGGCGATGCTCGTAAGGATTCG-3' (**2.58**); 5'-GGGGGC-GATGCTCGTAAGGATTC-3' (**2.53**).

The nucleotide sequences of the full-length extension product(s) of pol II⁻ were determined using the LC-ESI/MS/MS method described for the extension with Kf⁻. The ESI mass spectrum of the full-length extension reaction showed a molecular ion peaks with *m*/*z* 719.1 and 1079.0 (Figure **2-34**), which correspond to M-3H and M-2H ions of the same oligonucleotide. The molecular mass of the oligonucleotide calculated based on the mass of the molecular ion peak M-2H was 2160.1 Da. The analysis of the oligonucleotide with *m*/*z* 1079.0, performed in a similar manner as that for the Kf⁻ extension reaction product, suggested that the pol II⁻ full-length extension product was 5′-pTCCATGA-3′. The CID of molecular ion peak *m/z* 1079.0 is shown in Figure **2-35** and the comparison of the molecular mass of the observed and theoretical values of the fragments is shown in Table **2-07**.

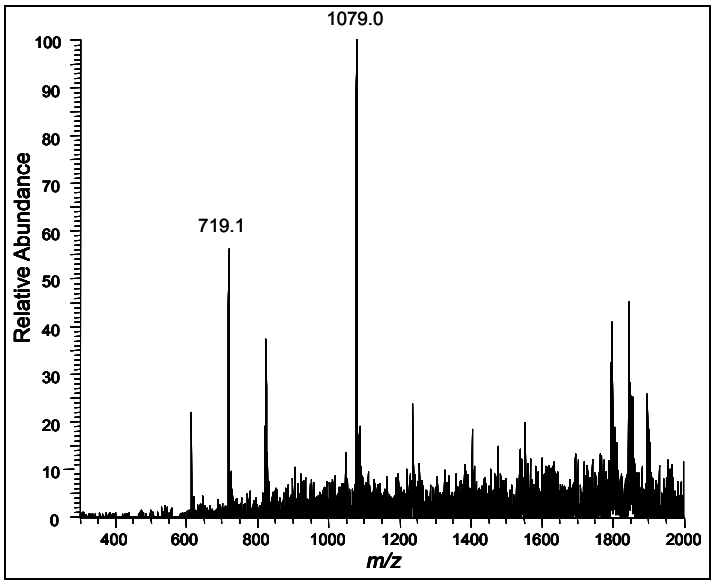



**Figure 2-34**. TIC spectrum of the LC-ESI/MS/MS analysis of the pol II$^{-}$ extension reaction product, past the MeFAPy-dGuo lesion of the oligonucleotide **2.52a** and the biotinated primer **2.54**; molecular ion peaks with *m/z* 719.1 and *m/z* 1079.0 correspond to M-3H and M-2H, respectively, of the extension product with mass 2160.1 Da identified as the 5″-pTCCATGA-3′.

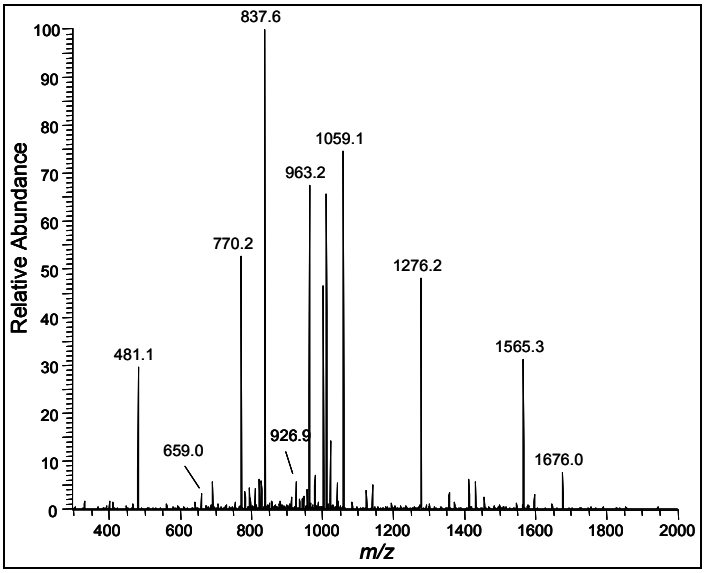



**Figure 2-35**. CID spectrum of the molecular ion peak *m/z* 1079.0 of the pol II$^{-}$ extension product, identified as the 5″-pTCCATGA-3′, past the MeFAPy-dGuo lesion in the oligonucleotide **2.52a**.

**Table 2-07**. Observed and theoretical fragmentation for the pol II$^{-}$ extension product *m/z* 1079.0, identified as the 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481.1 | 481.0 |
| 5'-pTC ($a_3$-B) | 770.2 | 770.8 |
| 5'-pTCC ($a_4$-B) | 1059.1 | 1059.1 |
| 5'-pTCCAT ($a_6$-B) | 1676.0 | 1676.2 |
| ($a_6$-B, -2) | 837.6 | 837.6 |
| pCCATGA-3' ($w_6$, -2) | 926.9 | 926.7 |
| pCATGA-3' ($w_5$) | 1565.3 | 1565.2 |
| pATGA-3' ($w_4$) | 1276.2 | 1276.2 |
| pTGA-3' ($w_3$) | 963.2 | 963.1 |
| pGA-3' ($w_2$) | 659.0 | 659.1 |

The yield for the formation of the free error bypass product, 5′-pTCCATGA-3′, was determined using the internal standard **2.55** and the calibration curve shown in Figure **2-31**. The yield of error-free bypass was 18%. Possible reasons for this low yield are the inefficient extension of the product after initial insertion of dCTP and the fact that the extension from the initial mis-incorporation of dATP, dTTP and dGTP are completely inhibited.

Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in Oligonucleotide **2.51a** Catalyzed by Dpo4.

The ability of the Y-family of DNA polymerases to bypass the MeFAPy-dGuo adduct was examined with Dpo4. Dpo4 inserted all four dNTP's opposite the MeFAPy-dGuo adduct in oligonucleotide **2.52a** (Figure **2-36**, Panel A). At higher dNTP concentrations incorporations of second dATP, dTTP and dCTP were observed. Dpo4 inserted all four dNTP's in unmodified oligonucleotide **2.51** and at higher dNTP(s) concentration an incorporation of second dNTP was observed. The incorporation of second dNTP is not unusual for Dpo4 because the enzyme has a larger active site compared to the $Kf^{-}$ and pol $II^{-}$ and may allow the accommodation of a second molecule of dNTP (*74*).

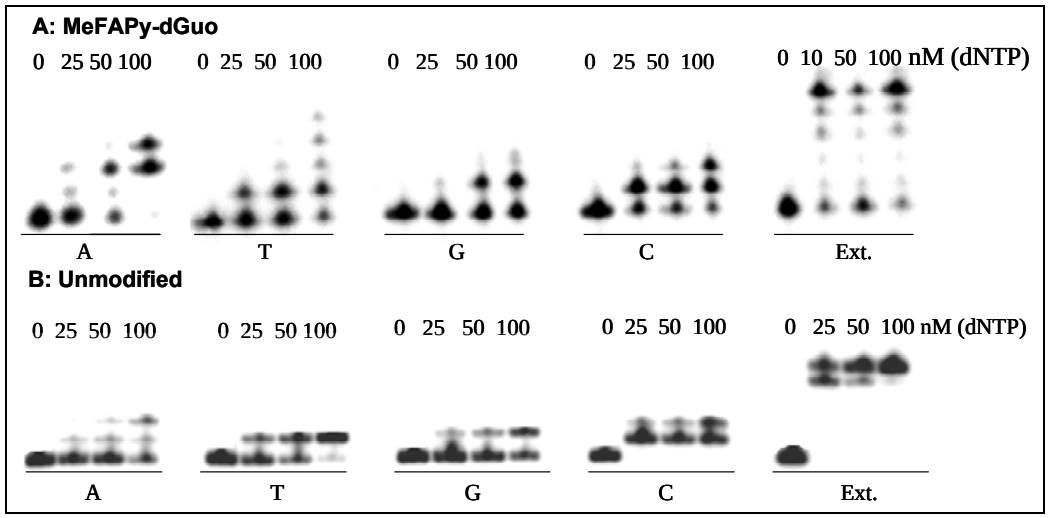



**Figure 2-36**. Single and full-length incorporation catalyzed by Dpo4: **A**. Incorporation across the MeFAPy-dGuo adduct in oligonucleotide **2.52a**; **B** Incorporation of dNTP(s) in the unmodified oligonucleotide **2.51**.

Steady-state kinetic experiments were run in order to investigate the rate of incorporation of the four dNTP's by Dpo4. While the insertion efficiencies for dCTP and dATP opposite the MeFAPy-dGuo were 2.5 and 10 times less, respectively, than the unmodified template, while the insertion efficiencies for dGTP and dTTP were the same as the unmodified template (Table **2-08**).

**Table2-08**. Steady-state kinetics parameters for Dpo4.

| Template | dNTP | Km (μM) | $k_{cat}$ ($s^{-1}$ x $10^{-3}$) | $k_{cat}/k_m$ ( $M^{-1}$ x $s^{-1}$x $10^{-3}$) | f[a] |
|---|---|---|---|---|---|
| MeFAPy-dGuo | C | 0.015 ± 0.002 | 0.6 ± 0.1 | 40 | 1 |
| | T | 0.105 ± 0.09 | 0.34 ± 0.06 | 3 | 0.08 |
| | G | 0.39 ± 0.04 | 1.5 ± 0.4 | 4 | 0.1 |
| | A | 0.52 ± 0.02 | 0.21 ± 0.09 | 0.4 | 0.01 |
| dGuo | C | 0.09 ± 0.01 | 9.6 ± 0.3 | 100 | 1 |
| | T | 2.0 ± 0.3 | 8.12 ± 0.05 | 4 | 0.04 |
| | G | 1.7 ± 0.4 | 10.3 ± 1.2 | 6 | 0.06 |
| | A | 2.4 ± 0.8 | 9.3 ± 0.4 | 4 | 0.04 |

[a]**f** - misincorporation frequency = $(k_{cat}/K_m)_{\text{incorporated dNTP}}/(K_{cat}/K_m)_{\text{correct dNTP (dCTP)}}$

Full-length extension product was observed for Dpo4 when all four dNTP's were added. The extension was efficient and there were no significant pause sites as observed with $Kf^{-}$ and Pol $II^{-}$. Subsequently, full-length extensions were examined with 0-primers Ado (**2.56**), Tyd (**2.56**), and Guo (**2.58)** were run in the presence of all four dNTP's in order to investigate if the initial mis-incorporation products could be extended (Figure **2-37**). The extension assays showed that the 0-primers were extended with low efficiency.

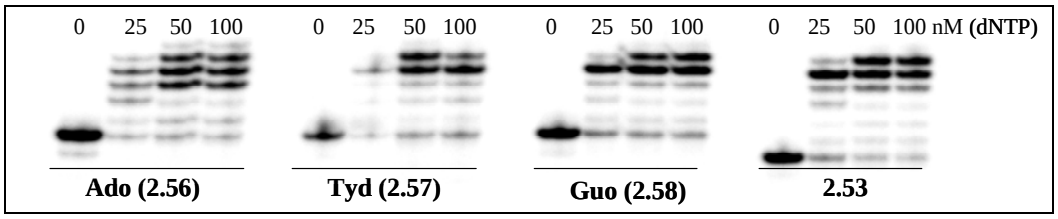



**Figure 2-37**. Full-length incorporation assays with MeFAPy-dGuo containing oligonucleotide (**2.52a**) by Dpo4 and 0-primers Ado (**2.56**), Tyd (**2.57**), Guo (**2.58**) and -1 primer **2.53** in the presence of the four dNTP's; 5'-GGGGGCGATGCTCG-TAAGGATTCA-3' (**2.56**); 5'-GGGGGCGATGCTCGTAAGGATTCT-3' (**2.57**); 5'-GGGGGCGATGCTCGTAAGGATTCG-3' (**2.58**); 5'-GGGGGCGATGCT-CGTAAGGATTC-3' (**2.53**).

Full-length extension reaction with the corresponding biotinated -1 primer **2.54** and the MeFAPy-dGuo containing oligonucleotide **2.52a** was run and analyzed by LC-ESI/MS/MS. The ESI spectrum of the reaction product(s) obtained from the biotinated primer **2.54** showed two molecular ion peaks with *m*/*z* 718.8 and 1078.5 (Figure **2-38**), which correspond to M-3H and M-2H ions of the same oligonucleotide. The molecular peak with *m/z* 1078.8 was used to calculate the molecular mass of the extension product, which is 2160.1 Da. This molecular mass was close to the molecular mass of the identified extension products of $Kf^{-}$ and pol $II^{-}$, which suggested that probably the Dpo4 extension product was the oligonucleotide 5′-pTCCATGA-3′. The CID of molecular ion peak 1078.5, shown in Figure **2-39**, matched very well with the CID spectrum of 5′-pTCCATGA-3′ (Table **2-09**), which represents an error-free bypass and extension.

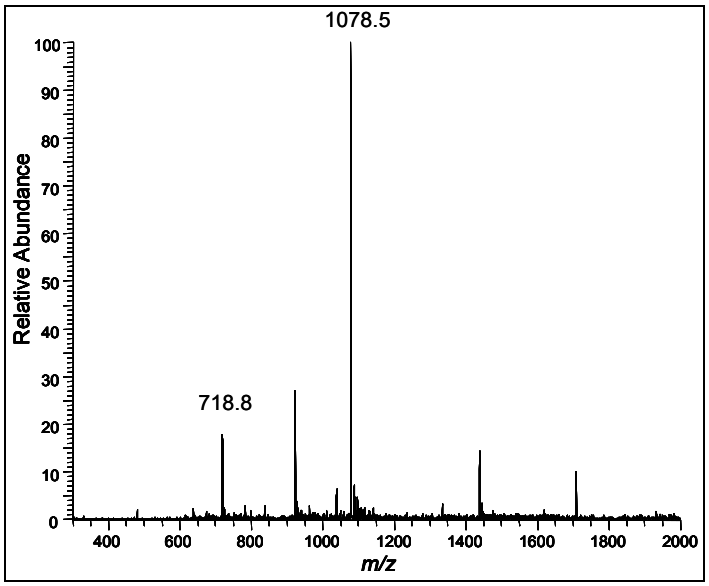



**Figure 2-38**. TIC spectrum of the LC-MS/MS/MS analysis of the Dpo4 extension reaction product past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** using primer **2.53**; molecular ion peaks with *m/z* 718.8 and *m/z* 1078.5 correspond to M-3H and M-2H, respectively, of the extension product with mass 2159.1 Da, identified as the 5′-pTCCATGA-3′.

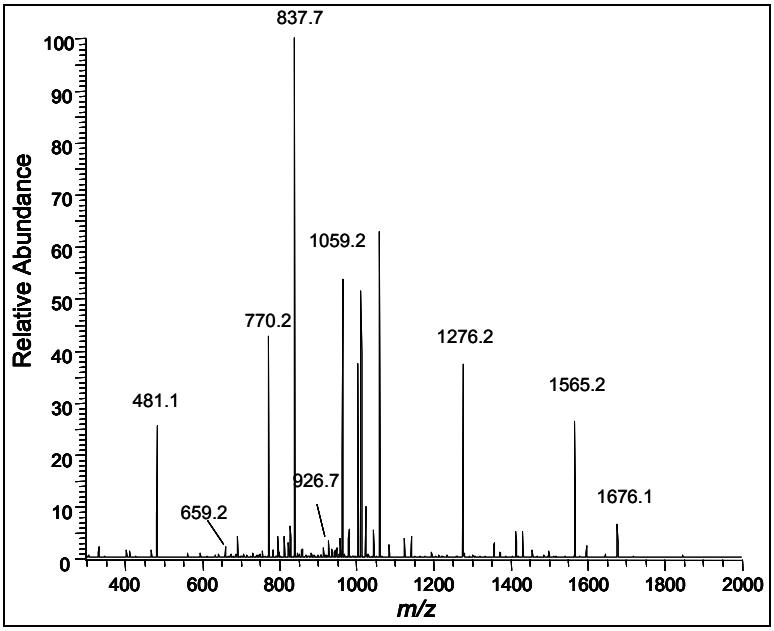



**Figure 2-39**. CID spectrum of the molecular ion peak *m/z* 1078.5 of the Dpo4 extension product, identified as the 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-09**. Observed and theoretical fragmentation for the Dpo4 extension product *m/z* 1078.5, identified as the 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481.1 | 481.0 |
| 5'-pTC ($a_3$-B) | 770.2 | 770.8 |
| 5'-pTCC ($a_4$-B) | 1059.2 | 1059.1 |
| 5'-pTCCAT ($a_6$-B) | 1676.1 | 1676.2 |
| ($a_6$-B, -2) | 837.7 | 837.6 |
| pCCATGA-3' ($w_6$, -2) | 926.7 | 926.7 |
| pCATGA-3' ($w_5$) | 1565.2 | 1565.2 |
| pATGA-3' ($w_4$) | 1276.2 | 1276.2 |
| pTGA-3' ($w_3$) | 963.2 | 963.1 |
| pGA-3' ($w_2$) | 659.2 | 659.1 |

The yield of formation of the free error bypass product, 5′-pTCCATGA-3′, was determined by using the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.56**) and the calibration graph curve shown in Figure **2-31**. The yield of the error-free bypass was 72%, indicating the more efficient bypass and extension of the primer after initial incorporation of dCyt opposite the MeFAPy-dGuo adduct by Dpo4 then $Kf^-$ and pol $II^-$.

Although the extension assays with primers **2.56**, **2.57** and **2.58** showed that Dpo4 can extend products that are formed after initial mis-incorporation of dGTP, dTTP, and dATP, the full-length extension products 5′-pTCGATGA-3′, 5′-pTCTATGA-3′ and 5′-pTCAATGA-3′, respectively, were not detected for the bypass of oligonucleotide **2.52a**. This suggests that the rate of mis-incorporation of the dATP, dTTP and dGTP is low compared to the rate

of incorporation and extension past dCTP opposite the MeFAPy-dGuo adduct (Table **2-08**).

Full-length extension reactions were run with the corresponding biotinated primers of Ado (**2.56**), Tyd (**2.57**) and Guo (**2.58**). The LC-ESI/MS/MS analyses of the extension products showed that the biotinated primers were extended in an error-free manner. The CID spectra and fragmentation tables of the corresponding extension products can be found in the Appendix.

### *DNA Bypass of MeFAPy-dGuo Adduct (2.05) Catalyzed by Eukaryotic DNA Polymerases*

The MeFAPy-dGuo containing oligonucleotide (**2.52a**) was annealed to a complementary 23mer (**2.54**) (-1) primer strand, which was 5'-labeled with $^{32}P$. The single nucleotide insertion and the full length extension reactions in the presence of all four dNTP's using eukaryotic DNA polymerases were run at fixed amount of the enzymes and increasing dNTP(s) concentration(s). The same reactions with unmodified oligonucleotide **2.51** were run at two-fold less concentration of eukaryotic DNA polymerases but with the same reaction time as the adducted oligonucleotide **2.52a**

Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in Oligonucleotide **2.51a** Catalyzed by Human Pol $\delta$.

Human pol $\delta$ inserted only dCTP opposite the MeFAPy-dGuo adduct in oligonucleotide **2.52a** (Figure **2-40**, Panel A) and only dCTP in the unmodified oligonucleotide **2.51** (Figure **2-40**, Panel B). The full-length extension past the MeFAPy-dGuo adduct in oligonucleotide **2.52a** catalyzed by human pol $\delta$ was blocked, which was expected since the replicative human pol $\delta$ does not replicate through DNA damaged lesions.

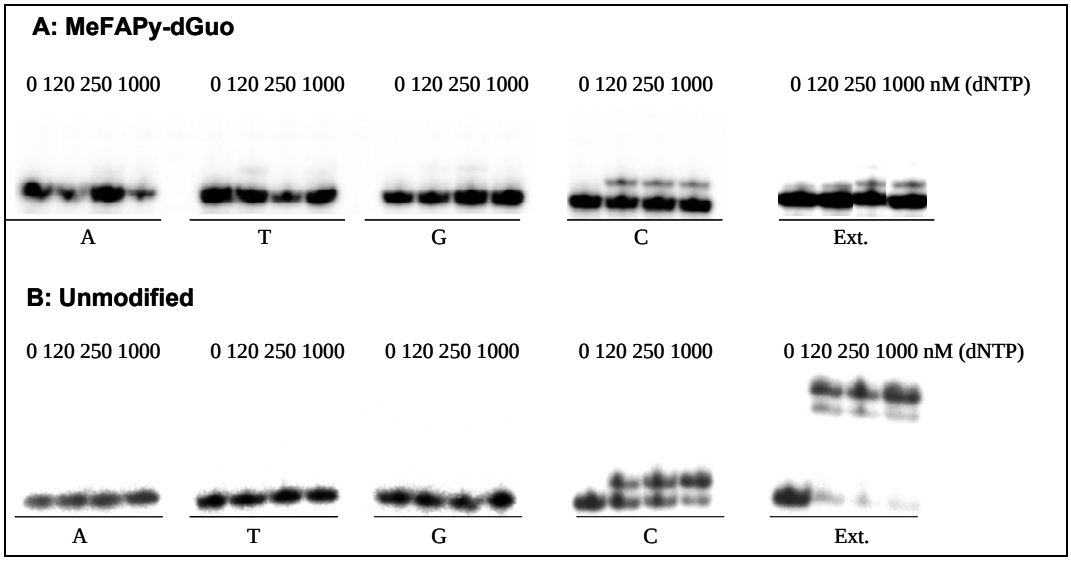



**Figure 2-40**. Single and full-length incorporation assays catalyzed by a human pol $\delta$: **A**. Incorporation across the MeFAPy-dGuo adduct in oligonucleotide **2.52a**; **B** Incorporation of dNTP(s) in the unmodified oligonucleotide **2.51**.

Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in Oligonucleotide **2.51a** Catalyzed by Human Pol $\eta$.

Human pol $\eta$ inserted all four dNTP's opposite the MeFAPy-dGuo adduct in oligonucleotide **2.52a** (Figure **2-41**, Panel A). At higher dNTP concentrations incorporations of second dATP, dGTP and dCTP were also observed.

The gel analysis of the single nucleotide insertion in unmodified oligonucleotide **2.51** showed that the human pol $\eta$ preferentially inserted dCTP, over dATP, dGTP and dTTP, and no subsequent incorporations of a second dNTP was observed (Figure **2-41**, Panel B).

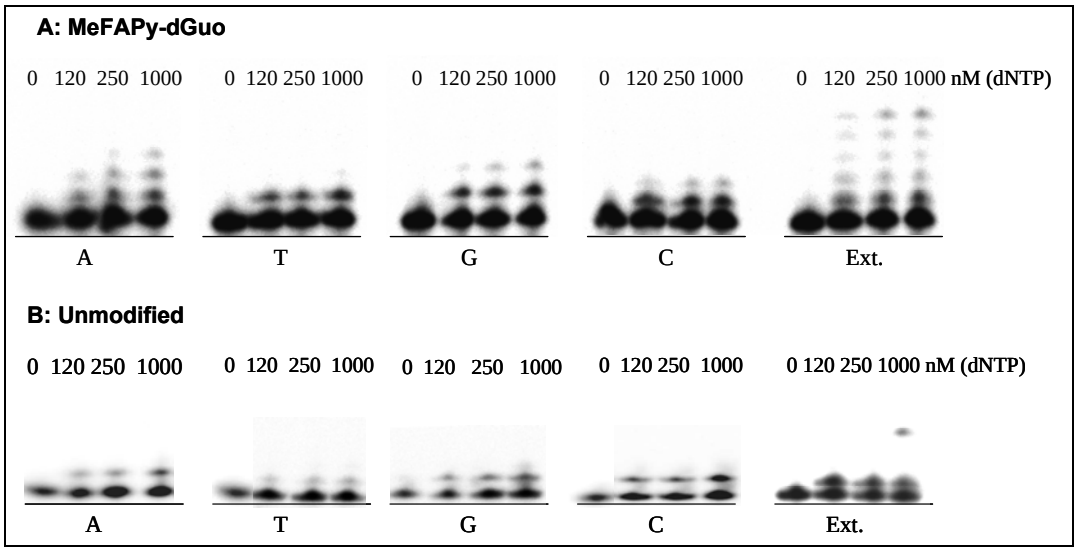



**Figure 2-41**. Single and full-length incorporation assays catalyzed by a human pol $\eta$: **A**. Incorporation across the MeFAPy-dGuo adduct in oligonucleotide **2.52a**; **B** Incorporation of dNTP(s) in the unmodified oligonucleotide **2.51**.

The rates of nucleotide incorporation opposite the MeFAPy-dGuo lesion catalyzed by human pol $\eta$ were determined and the steady-state kinetics parameters are listed in Table **2-10**. The insertion efficiency for dCTP opposite the MeFAPy-dGuo was 6.4 times less than the unmodified template, whereas the insertion efficiencies for dTTP, dGTP and dATP were 1.5, 3.6 and 3 times higher, than the unmodified template, respectively.

**Table 2-10**. Steady-state kinetics parameters for pol $\eta$.

| Template | dNTP | Km (μM) | $k_{cat}$ ($s^{-1}$ x $10^{-3}$) | $k_{cat}/k_m$ ($\mu M^{-1}$ x $s^{-1}$x $10^{-3}$) | f[a] |
|---|---|---|---|---|---|
| MeFAPy-dGuo | C | 8 ± 2 | 26 ± 1 | 3 | 1 |
| | T | 23 ± 3 | 17 ± 2 | 0.7 | 0.7 |
| | G | 9 ± 1 | 3.8 ± 0.3 | 0.4 | 0.1 |
| | A | 29 ± 8 | 14 ± 1 | 0.5 | 0.1 |
| dGuo | C | 1.9 ± 0.5 | 40 ± 3 | 21 | 1 |
| | T | 4 0± 10 | 20 ± 3 | 0.5 | 0.02 |
| | G | 110 ± 40 | 13 ± 1 | 0.1 | 0.005 |
| | A | 140 ± 30 | 20 ± 2 | 0.1 | 0.005 |

[a]f - misincorporation frequency = $(k_{cat}/K_m)_{\text{incorporated dNTP}}$ / $(K_{cat}/K_m)_{\text{correct dNTP (dCTP)}}$

The full-length extension reaction of MeFAPy-dGuo oligonucleotide and the biotinated primer **2.54** catalyzed by pol $\eta$ was run and analyzed by LC-ESI/MS/MS. The ESI spectrum of the extension reaction showed three molecular ion peaks with *m*/*z* 1078.8, 1086.8 and 1099.2 (Figures **2-42** and **2-43**), which correspond to oligonucleotides with masses 2159.6, 2175.2 and 2200.4 Da (M-2H), respectively. The oligonucleotides with masses 2159.6

and 2175.2 were identified as the 5′-pTCCATGA-3′ and 5′-pTCTATGA-3′ based on possible oligonucleotide compositions, the sequence of the template and the corresponding CID's (Figures **2-43** and **2-44**, Tables **2-11** and **2-12**).

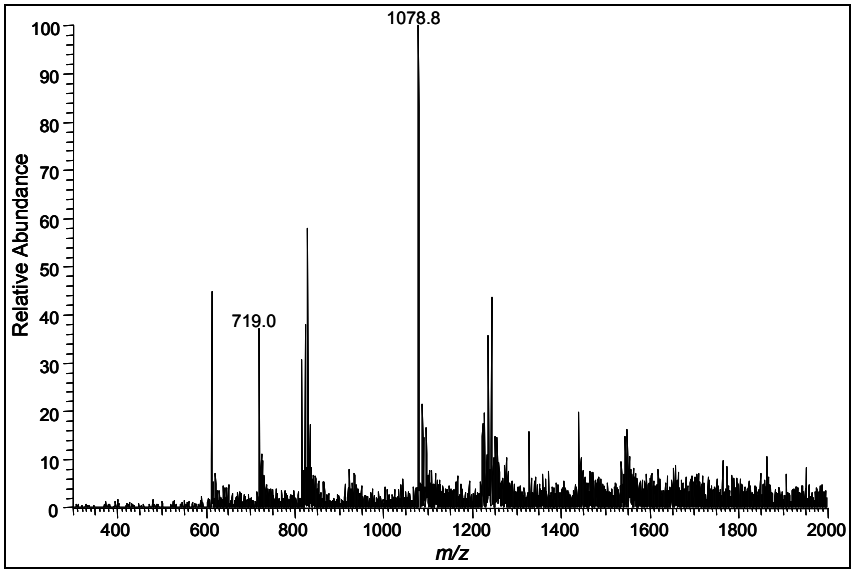



**Figure 2-42**. TIC spectrum of the LC-MS/MS/MS analysis of the pol $\eta$ extension reaction product(s) past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** using the primer **2.54**; molecular ion peaks with *m/z* 719.0 and *m/z* 1078.8 correspond to M-3H and M-2H, respectively, of the extension product with mass 2159.6 Da identified as 5′-pTCCATGA-3′.

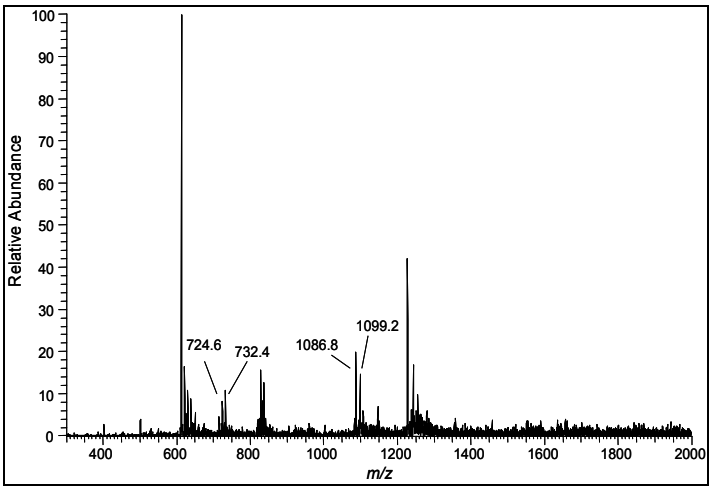



**Figure 2-43**. TIC spectrum of the LC-MS/MS/MS analysis of the pol $\eta$ extension reaction product(s) past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** using the primer **2.54**; molecular ion peaks with *m/z* 724.6 and *m/z* 1086.9 correspond to M-3H and M-2H, respectively, of the extension product with mass 2175.8 Da identified as 5′-pTCTATGA-3′; molecular ion peaks with *m/z* 732.4 and *m/z* 1099.2 correspond to M-3H and M-2H, respectively, of the extension product with mass 2200.4 Da identified as 5′-pTCGATGA-3′.

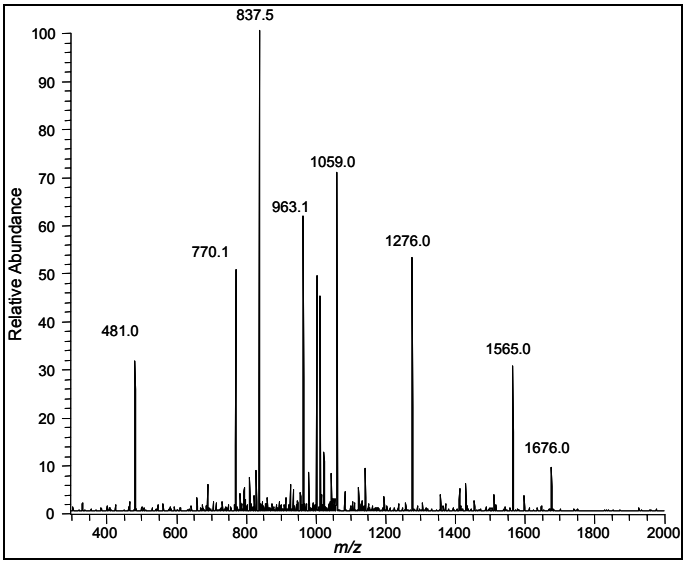



**Figure 2-44**. CID spectrum of the molecular ion peak *m/z* 1078.8 of the pol $\eta$ extension product, identified as 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-11**. Observed and theoretical fragmentation for the pol $\eta$ extension product *m/z* 1078.8, identified as 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481 | 481 |
| 5'-pTC ($a_3$-B) | 770.1 | 770.8 |
| 5'-pTCC ($a_4$-B) | 1059 | 1059.1 |
| 5'-pTCCAT ($a_6$-B) | 1676 | 1676.2 |
| ($a_6$-B, -2) | 837.5 | 837.6 |
| pCATGA-3' ($w_5$) | 1565 | 1565.2 |
| pATGA-3' ($w_4$) | 1276 | 1276.2 |
| pTGA-3' ($w_3$) | 963.1 | 963.1 |

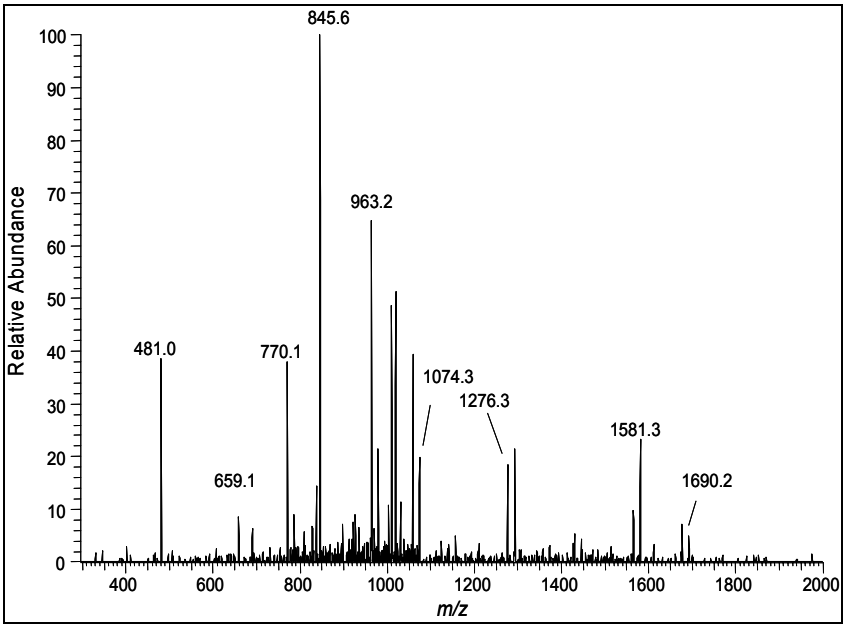



**Figure 2-45**. CID spectrum of the molecular ion peak *m/z* 1086.8 of the pol $\eta$ extension product, identified as 5′-pTCTATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-12**. Observed and theoretical fragmentation for the pol $\eta$ extension product *m/z* 1086.8, identified as 5′-pTCTATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481 | 481 |
| 5'-pTC ($a_3$-B) | 770.1 | 770.8 |
| 5'-pTCT ($a_4$-B) | 1074.2 | 1074.1 |
| 5'-pTCTAT ($a_6$-B) | 1690.2 | 1691.2 |
| ($a_6$-B, -2) | 845.6 | 845.1 |
| pTATGA-3' ($w_5$) | 1581.3 | 1580.2 |
| pATGA-3' ($w_4$) | 1276.3 | 1276.2 |
| pTGA-3' ($w_3$) | 963.2 | 963.1 |
| pGA-3' ($w_2$) | 659.1 | 659.1 |

The possible composition of the oligonucleotide with *m/z* 2200.46 (M-2H) was 1 phosphate, 1 dCyd, 2 dTyd's, 2 dAdo's, and 2 dGuo's. Analysis of the CID spectrum suggests that the sequence was 5′-pTCGATGA-3′, which represents a mis-insertion of dGTP followed by an error-free extension. The CID of the molecular ion *m*/*z* 2200.4 matched very well with the predicted CID (Figure **2-46**, Table **2-13**).

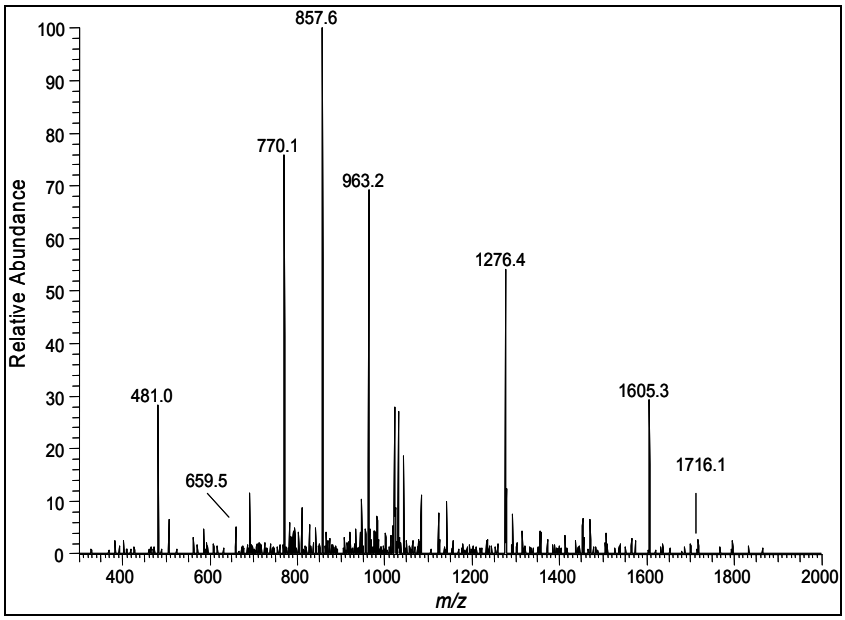



**Figure 2-46**. CID spectrum of the molecular ion peak *m/z* 1099.2 of the pol $\eta$ extension product, identified as 5′-pTCGATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-13**. Observed and theoretical fragmentation for the pol $\eta$ extension product *m/z* 1099.2, identified as 5′-pTCGATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481 | 481 |
| 5'-pTC ($a_3$-B) | 770 | 770.8 |
| 5'-pTCGAT ($a_6$-B) | 1716.1 | 1716.2 |
| ($a_6$-B, -2) | 857.6 | 587.6 |
| pGATGA-3' ($w_5$) | 1605.3 | 1605.2 |
| pATGA-3' ($w_4$) | 1276.4 | 1276.2 |
| pTGA-3' ($w_3$) | 963.2 | 963.1 |
| pGA-3' ($w_2$) | 659.5 | 659.1 |

The yields of the 5′-pTCCATGA-3′, 5′-pTCTATGA-3′ and 5′-pTCGATGA-3′ extension products were determined by using the internal standard 5’-pCTTACGAGCCCCC-3′ (**2.55**) and the calibration curves shown in Figures **2-31**, **2-47** and **2-48**. The yields for the extension products 5′p-TCCATGA-3′, 5′-pTCTATGA-3′ and 5′-pTCGATGA-3′ were 54%, 5% and 8%, respectively.

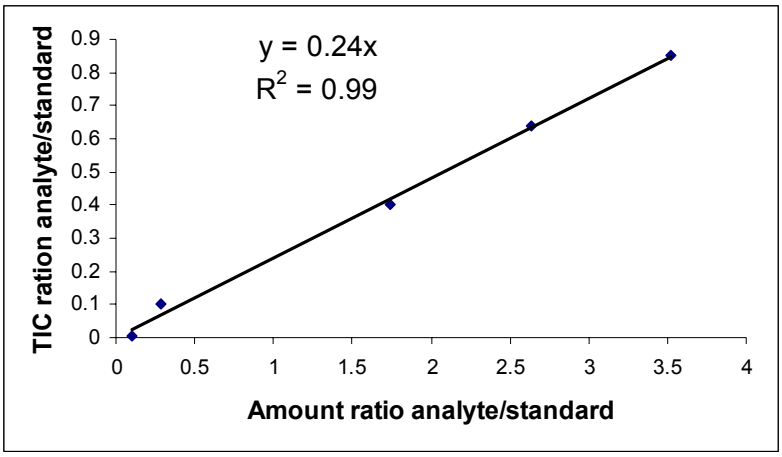



**Figure 2-47**. The calibration curve for the 5′-pTCTATGA-3′ and the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.55**).

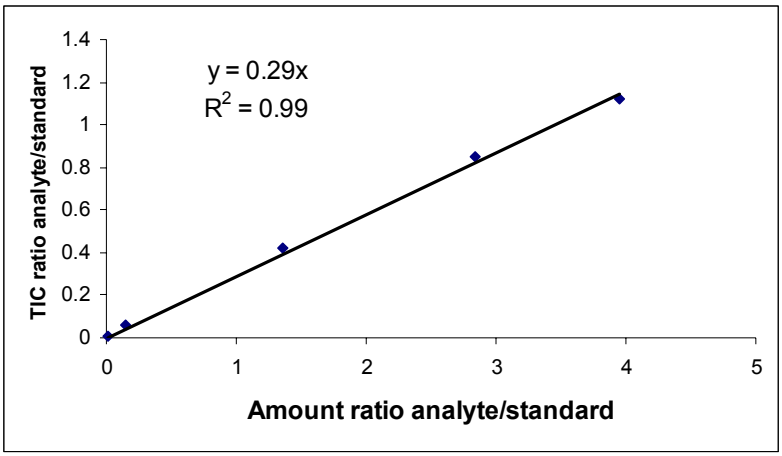



**Figure 2-48**. The calibration curve for the 5′-pTCGATGA-3′ and the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.56**).

Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in Oligonucleotide **2.52a** Catalyzed by Human Pol *ι*.

Human pol *ι* did not catalyze any dNTP incorporation across the MeFAPy-dGuo adduct with good efficiency (Figure **2-49**, Panel A). The gel analysis of the single nucleotide insertion in unmodified oligonucleotide **2.51** showed that the human pol *ι* did not incorporate any dNTP (Figure **2-49**, Panel B).

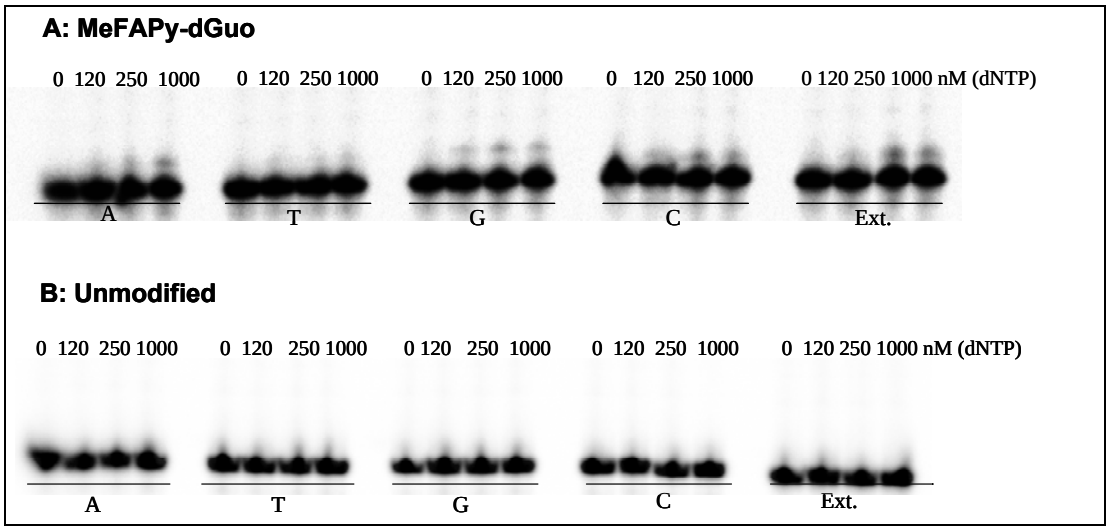



**Figure 2-49**. Single and full-length incorporation assays catalyzed by human pol $\iota$: **A**. Incorporation across the MeFAPy-dGuo adduct in oligonucleotide **2.52a**; **B** Incorporation of dNTP(s) in the unmodified oligonucleotide **2.51**.

Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in Oligonucleotide **2.52a** Catalyzed by Human Pol $\kappa$.

Human pol $\kappa$ inserted dTTP, dGTP and dCTP opposite the MeFAPy-dGuo adduct in oligonucleotide **2.52a** (Figure **2-50**, Panel A), while dCTP and at higher concentration dGTP was inserted in the unmodified oligonucleotide (Figure **2-50**, Panel B).

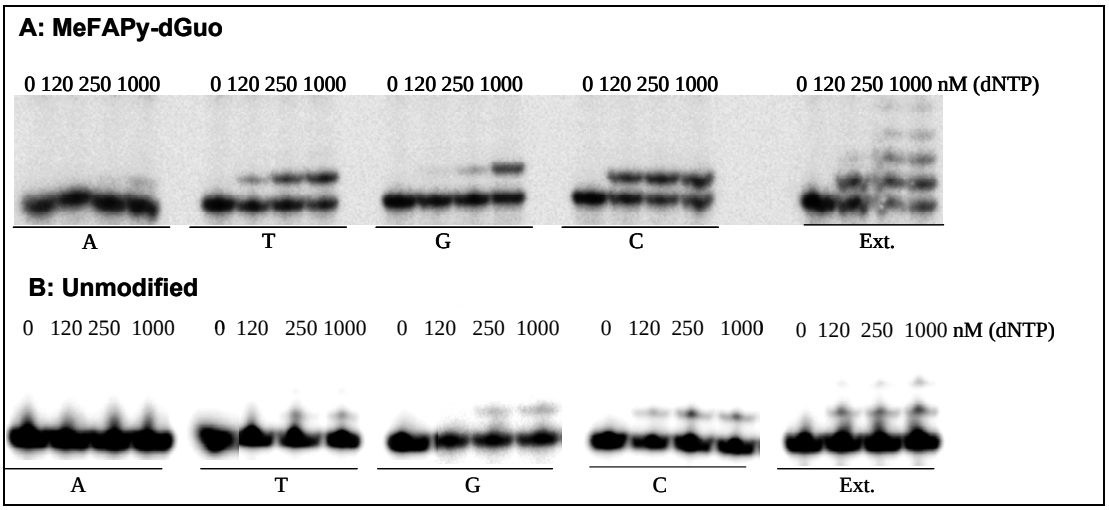



**Figure 2-50**. Single and full-length incorporation assays catalyzed by human pol $\kappa$: **A**. Incorporation across the MeFAPy-dGuo adduct in oligonucleotide **2.52a**; **B** Incorporation of dNTP(s) in the unmodified oligonucleotide **2.51**.

The rates of incorporation of the three dNTP's by human pol *k* were determined by steady-state kinetic experiments. The insertion efficiency for dCTP opposite the MeFAPy-dGuo was 6 times less than the unmodified template, whereas the insertion efficiencies for dGTP and dTTP were the same and 2.8 higher than the unmodified template, respectively (Table **2-14**).

**Table 2-14**. Steady-state kinetics parameters for pol $\kappa$.

| Template | dNTP | Km (μM) | $k_{cat}$ ($s^{-1}$ x $10^{-3}$) | $k_{cat}/k_m$ ($\mu M^{-1}$ x $s^{-1}$x $10^{-3}$) | f[a] |
|---|---|---|---|---|---|
| MeFAPy-dGuo | C | 9 ± 2 | 49 ± 2 | 5.5 | 1 |
| | G | 220 ± 70 | 27 ± 5 | 0.1 | 0.02 |
| | T | 45 ± 9 | 27 ± 2 | 0.6 | 0.11 |
| dGuo | C | 2.1 ± 0.8 | 75 ± 6 | 35 | 1 |
| | G | 300 ± 89 | 42 ± 4 | 0.1 | 0.003 |
| | T | 150 ± 20 | 32 ± 2 | 0.2 | 0.006 |

[a]f - misincorporation frequency = $(k_{cat}/K_m)_{\text{incorporated dNTP}}/(K_{cat}/K_m)_{\text{correct dNTP (dCTP)}}$

Full-length extension reaction with biotinated (-1)-primer **2.54** and the MeFAPy-dGuo oligonucleotide **2.52a** catalyzed by human pol $\kappa$ was run and analyzed by LC-ESI/MS/MS. The ESI spectrum of the extension reaction showed two molecular ion peaks with *m*/*z* 1078.5 and 1086.9 (Figure **2-51** and **2-52**), which were assumed to correspond to oligonucleotides with masses 2159.6 and 2175.8 Da (M-2H), respectively. The possible composition of the 2159.6 Da oligonucleotide was 1 phosphate, 2 dCyd's, 2 dTyd's, 2 dAdo's and 1 dGuo, whereas the possible composition of the 2175.8 Da oligonucleotide was 1 phosphate, 1 dCyt, 2 dTyd's, 3 dAdo's and 1 dGuo. Considering these compositions and the sequence of the template, the only possible extensions products were 5′p-TCCATGA-3′, which represents an error free bypass and extension and 5′-pTCAATGA-3′, which represents mis-insertion of dATP opposite the adduct followed by error-free extension. The CIDs of the molecular ions peaks *m*/*z* 1078.5 (Figure **2-53**) and *m*/*z* 1086.9 (Figure **2-54**) matched very well with the predicted CID spectra of an

authentic (Appendix), commercially purchased oligonucleotides (Table **2-15**, **2-16**).

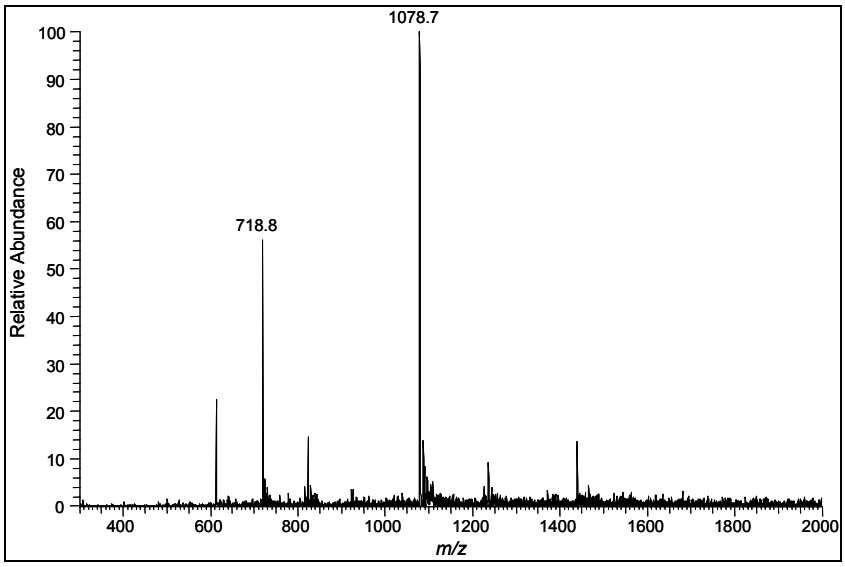



**Figure 2-51**. TIC spectrum of the LC-MS/MS/MS analysis of the pol $\kappa$ extension reaction product(s) past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** using primer **2.54**. Molecular ion peaks with *m/z* 718.8 and *m/z* 1078.7 correspond to M-3H and M-2H, respectively of oligonucleotide with mass 2159.6 Da, identified as 5′-pTCCATGA-3′.

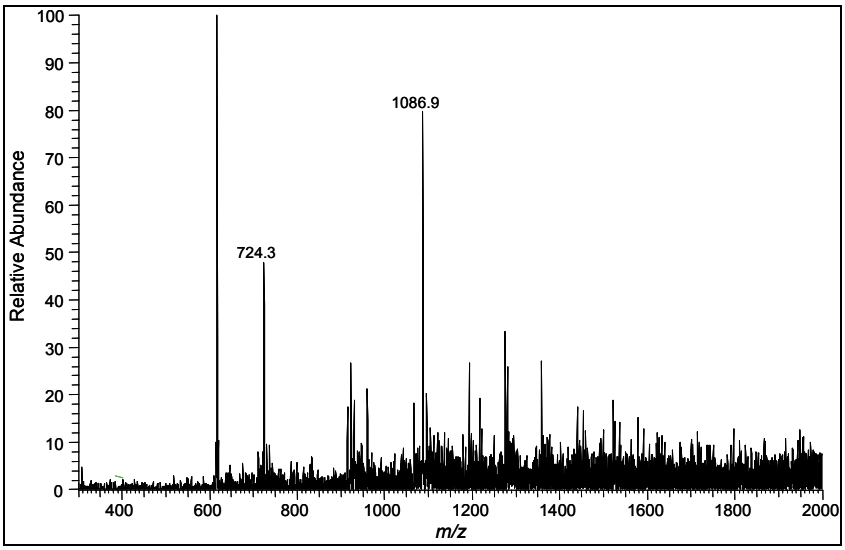



**Figure 2-52**. TIC spectrum of the LC-MS/MS/MS analysis of the pol $\kappa$ extension reaction product(s) past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** using primer **2.54**. Molecular ion peaks with *m/z* 724.3 and *m/z* 1089.9 correspond to M-3H and M-2H, respectively of oligonucleotide with mass 2175.8 Da, identified as 5′-pTCTATGA-3′.

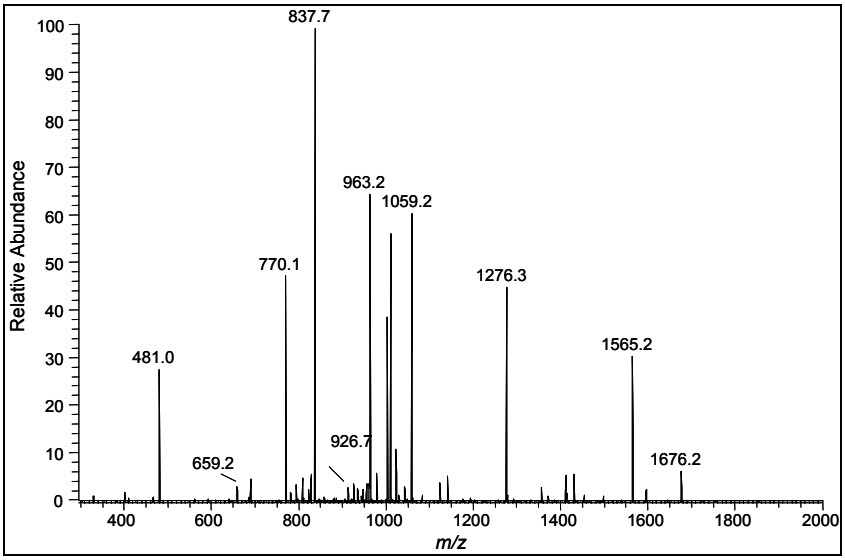



**Figure 2-53**. CID spectrum of the molecular ion peak *m/z* 1078.8 of the pol $\kappa$ extension product, identified as 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

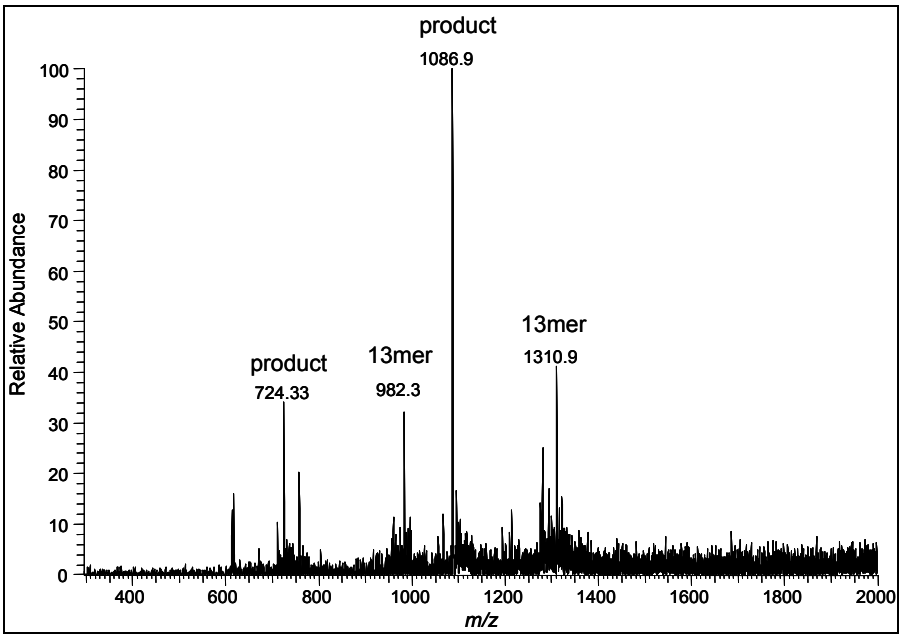



**Figure 2-54**. CID spectrum of the molecular ion peak *m/z* 1086.9 of the pol $\kappa$ extension product, identified as 5'-pTCTATGA-3', past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-15**. Observed and theoretical fragmentation for the pol $\kappa$ extension product *m/z* 1078.8, identified as 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481 | 481 |
| 5'-pTC ($a_3$-B) | 770.1 | 770.8 |
| 5'-pTCC ($a_4$-B) | 1059.2 | 1059.1 |
| 5'-pTCCAT ($a_6$-B) | 1676.2 | 1676.2 |
| ($a_6$-B, -2) | 837.7 | 837.6 |
| pCCATGA-3' ($w_6$, -2) | 926.7 | 926.7 |
| pCATGA-3' ($w_5$) | 1565.2 | 1565.2 |
| pATGA-3' ($w_4$) | 1276.3 | 1276.2 |
| pTGA-3' ($w_3$) | 963.2 | 963.1 |
| pGA-3' ($w_2$) | 659.2 | 659.1 |

**Table 2-16**. Observed and theoretical fragmentation for the pol κ extension product *m/z* 1086.9, identified as 5′-pTCTATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481 | 481 |
| 5'-pTC ($a_3$-B) | 770 | 770.8 |
| 5'-pTCT ($a_4$-B) | 1074.2 | 1074.1 |
| 5'-pTCTAT ($a_6$-B) | 1692.2 | 1691.2 |
| ($a_6$-B, -2) | 845.1 | 845.1 |
| pCTATGA-3' ($w_6$, -4) | 466 | 466.5 |
| pTATGA-3' ($w_5$) | 1580.2 | 1580.2 |
| pATGA-3' ($w_4$) | 1276.1 | 1276.2 |
| pTGA-3' ($w_3$) | 963.1 | 963.1 |
| pGA-3' ($w_2$) | 659.2 | 659.1 |

The yields for the formation of the 5′-pTCCATGA-3′ and 5′-pTCTATGA-3′ extension products were determined using the internal standard **2.55** and the calibration graphs shown in Figures **2-31** and **2.47**. The yield of formation of the error-free product 5′p-TCCATGA-3′ was 55%, whereas the yield of formation of error-prone product 5′-pTCTATGA-3′, was 5%.

Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in Oligonucleotide **2.52a** Catalyzed by a Combination of Human Pols κ/ι.

The human pols κ/ι were used in equimolar amounts and their concentration was kept fixed for the single nucleotide insertion assays.

The single nucleotide insertion catalyzed by a combination of human pols κ/ι seemed to occur at higher dNTP concentrations (Figure **2-55**). The preferable nucleotide incorporated opposite the MeFAPy-dGuo lesion was dCTP, which represents the correct incorporation. Human pols κ/ι also inserted dGTP and dTTP at higher dNTP's concentrations. In contrast to pol κ, which favored the misinsertion of dTTP over dGTP, the combination of pols κ/ι favored the misinsertion of dGTP more than dTTP (Figures **2-50** and **2-55**). This change in the preference for the incorporation of the second nucleotide when a combination of pols κ/ι is used is perhaps due to the tight binding of human pol ι to the template/primer with T more than to the template/primer with G; if the binding is tight, the human pol ι does not dissociate, which prevents pol κ from extending the products. Human pol ι alone did not catalyzed any nucleotide incorporation across the MeFAPy-dGuo adduct with a good efficiency (Figure **2.49**, Panel A).

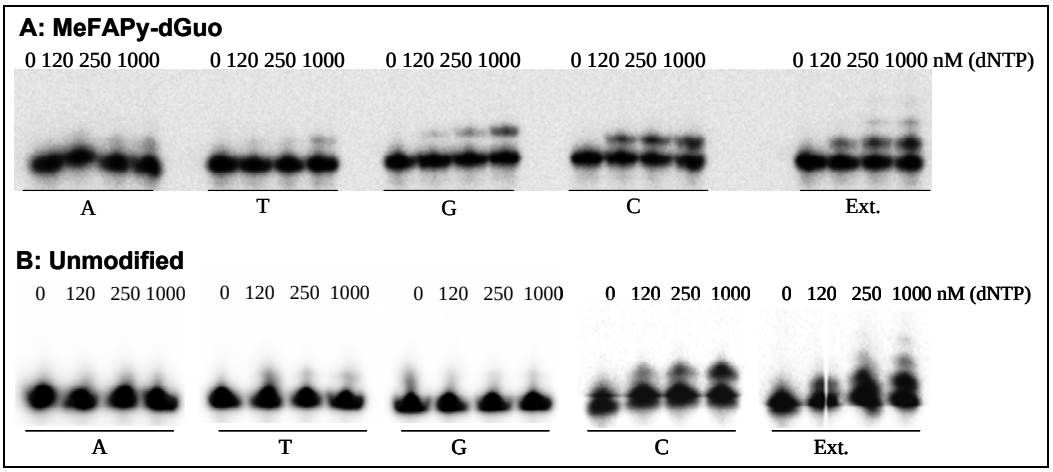



**Figure 2-55**. Single and full-length incorporation assays catalyzed by a combination human pols $\kappa/\iota$: **A**. Incorporation across the MeFAPy-dGuo adduct in oligonucleotide **2.52a**; **B** Incorporation of dNTP(s) in the unmodified oligonucleotide **2.51**.

The single nucleotide incorporation assays with the combination of human pols $\kappa/\iota$ and the unmodified sequence **2.51** showed that only dCTP is incorporated in contrast to human pol $\kappa$, which not only incorporated dCTP but also dGTP at higher dNTP concentration (Figure **2-55**, Panel B and Figure **2-50**, Panel B).

The full-length extension reaction with the MeFAPy-dGuo containing oligonucleotide **2.52a** and the biotinated primer **2.54** catalyzed by a combination of pols $\kappa/\iota$ was run and analyzed by LC-ESI/MS/MS. The ESI spectrum of the extension reaction showed two molecular ion peaks with *m*/*z* 1078.8 and 1086.9 (Figures **2-56** and **2-57**), which correspond to oligonucleotides with masses 2159.6 and 2175.8 Da, respectively. Since these two oligonucleotides had the same molecular masses as the extension

products of pol $\kappa$, 5′-pTCCATGA-3′ and 5″-pTCTATGA-3′, it was initially assumed that they are identical to the afore-mentioned sequences. This was tested by examining the MS/MS (CID) spectra of the molecular ions with *m*/*z* 1078.8 and 1086.9 (Figures **2-58** and **2-59**). The CID spectra of the molecular ions with masses 2159.6 and 2175.8 Da showed fragmentation patterns indicative for the 5′-pTCCATGA-3′ and 5′-pTCTATGA-3′ (Tables **2-17** and **2-18**).

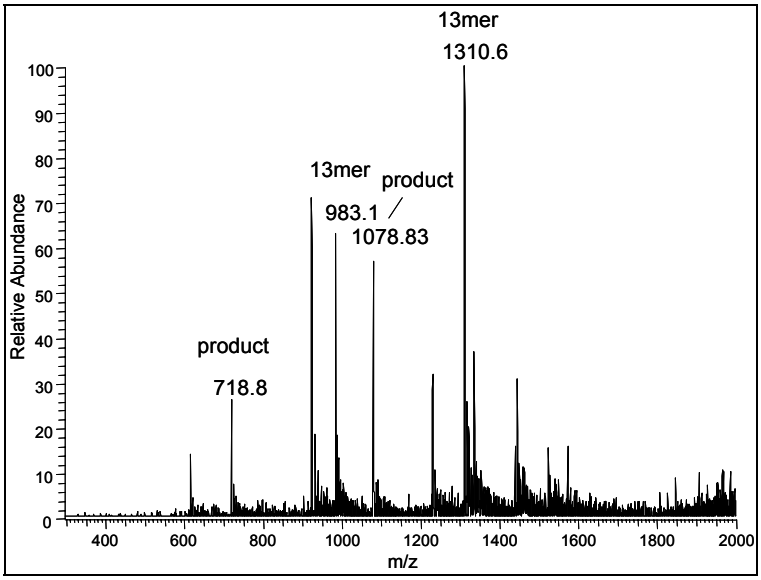



**Figure 2-56**. TIC spectrum of the LC-ESI/MS/MS analysis of the pols $\kappa$/$\iota$ extension reaction product(s) past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** using primer **2.54**; molecular ion peaks with *m/z* 718.8 and *m/z* 1078.8 correspond to M-3H and M-2H, respectively, of the extension product with mass 2159.6, identified as 5′-pTCCATGA-3′; molecular ion peaks with *m/z* 983.1 and *m/z* 1310.6 correspond to M-4H and M-3H, respectively, of the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.55**).

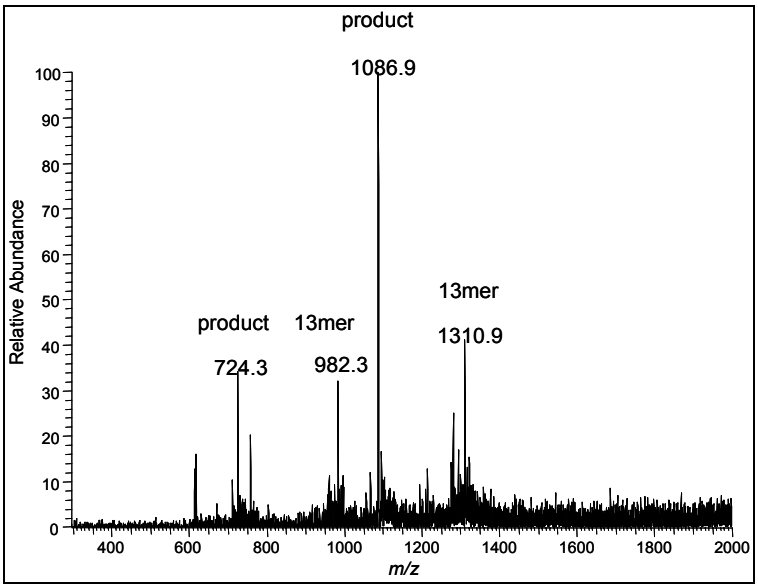



**Figure 2-57**. TIC spectrum of the LC-ESI/MS/MS analysis of the pols $\kappa$/$\iota$ extension reaction product(s) past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** using primer **2.54**; molecular ion peaks with *m/z* 724.3 and *m/z* 1086.9 correspond to M-3H and M-2H, respectively of oligonucleotide with mass 2175.8, identified as 5'-pTCTATGA-3'; molecular ion peaks with *m/z* 982.3 and *m/z* 1310.9 correspond to M-4H and M-3H, respectively, of the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.55**).

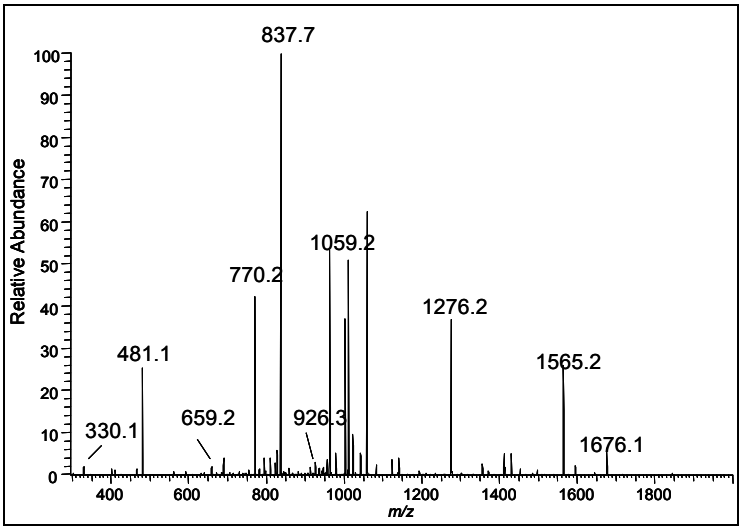



**Figure 2-58**. CID spectrum of the molecular ion peak *m/z* 1078.8 of the pols $\kappa$/$\iota$ extension product, identified as 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-17**. Observed and theoretical fragmentation for the pol κ/ι extension product *m/z* 1078.8, identified as 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481.1 | 481.0 |
| 5'-pTC ($a_3$-B) | 770.2 | 770.8 |
| 5'-pTCC ($a_4$-B) | 1059.2 | 1059.1 |
| 5'-pTCCAT ($a_6$-B) | 1676.1 | 1676.2 |
| ($a_6$-B, -2) | 837.7 | 837.6 |
| pCCATGA-3' ($w_6$, -2) | 926.3 | 926.7 |
| pCATGA-3' ($w_5$) | 1565.2 | 1565.2 |
| pATGA-3' ($w_4$) | 1276.2 | 1276.2 |
| pTGA-3' ($w_3$) | 963.2 | 963.1 |
| pGA-3' ($w_2$) | 659.2 | 659.1 |
| pA-3' ($w_1$) | 330.1 | 330 |

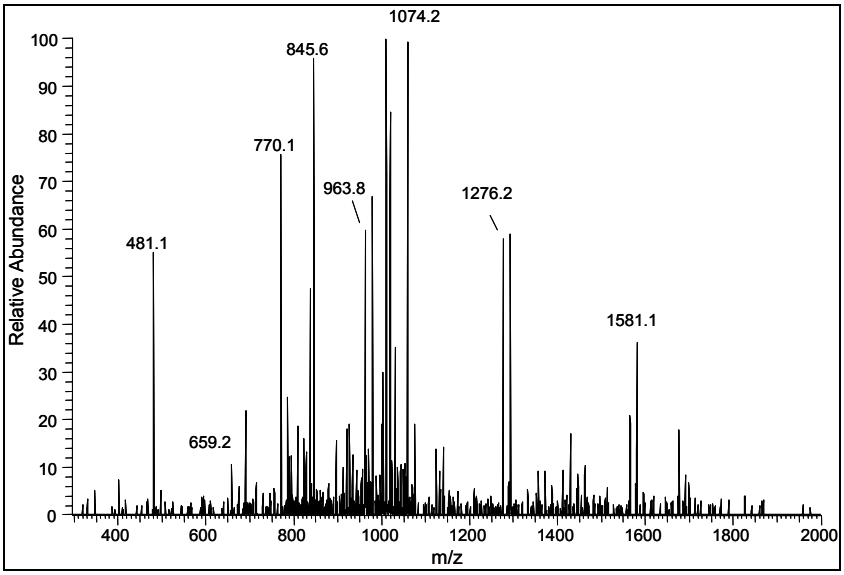



**Figure 2-59**. CID spectrum of the molecular ion peak *m/z* 1086.9 of the pols κ/ι extension product, identified as 5′-pTCTATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.51a**.

**Table 2-18**. Observed and theoretical fragmentation for the pols κ/ι extension product *m/z* 1086.9, identified as 5′-pTCTATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481.1 | 481.0 |
| 5'-pTC ($a_3$-B) | 770.1 | 770.8 |
| 5'-pTCT ($a_4$-B) | 1074.2 | 1074.1 |
| 5'-pTCTAT ($a_6$-B, -2) | 845.6 | 845.1 |
| pTATGA-3' ($w_5$) | 1581.1 | 1580.2 |
| pATGA-3' ($w_4$) | 1276.2 | 1276.2 |
| pTGA-3' ($w_3$) | 963.8 | 963.1 |
| pGA-3' ($w_3$) | 659.2 | 659.1 |

The yields of for the extension products 5′-pTCCATGA-3′ and 5′-pTCTATGA-3′ by the combination of human pols κ/ι were determined using the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.55**) and the calibration curve shown in Figures **2-31** and **2.47**. The yield of formation of the oligonucleotide 5′-pTCCATGA-3′ was 20%, whereas the yield of formation of the oligonucleotide 5′-pTCAATGA-3′ was 5%. The yield of formation of the error free bypass product, 5′-pTCCATGA-3′, was by a factor of two lower than the yield of the same extension product formed by human pol κ, while the yield of the error prone bypass, 5′-pTCTATGA-3′, was the same as the yield of the same extension product formed by human pol κ.

A possible explanation for the difference in the yields of the reaction products formed by a combination of human pols κ/ι and human pol κ is alone that pol κ is primarily involved in the extension past the MeFAPy-dGuo

adduct and that pol $\iota$ is affecting the rate with which pol $\kappa$ incorporates/misincoporates dNTP across the MeFAPy-dGuo lesion (**2.05**).

Polymerase Bypass of the MeFAPy-dGuo Adduct (**2.05**) in Oligonucleotide **2.51a** Catalyzed by a Combination of Human Pols $\eta/\iota$.

The human pol $\eta/\iota$ were used in equimolar amounts and their concentration was kept fixed for the single nucleotide insertion assays.

The combination of human pols $\eta/\iota$ inserted all four dNTP's opposite the MeFAPy-dGuo adduct in oligonucleotide **2.52a** (Figure **2-60**, Panel A) with less efficiency than the pol $\eta$ alone (Figure **2-41**, Panel A). At higher dATP concentration incorporation of second dATP was also observed.

The gel analysis of nucleotide insertion in unmodified sequence **2.51** showed that the human pols $\eta/\iota$ preferentially inserted dCTP over dGTP (Figure **2-60**, Panel B); in the analogous reactions with only human pol $\eta$ all four dNTP's were incorporated (Figure **2-41**, Panel B).

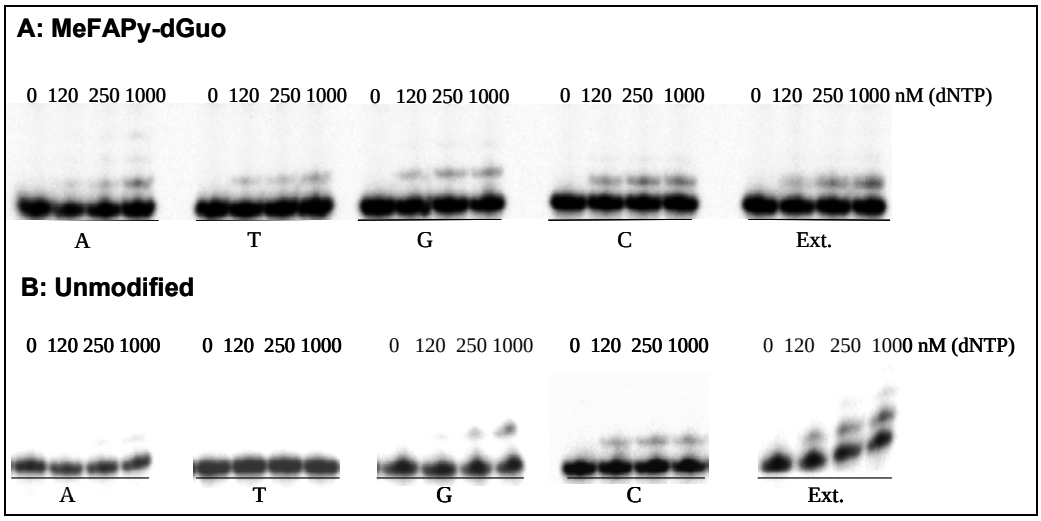



**Figure 2-60**. Single and full-length incorporation assays catalyzed by a combination of human pols $\eta/\iota$: **A**. Incorporation across the MeFAPy-dGuo adduct in oligonucleotide **2.52a**; **B** Incorporation of dNTP(s) in the unmodified oligonucleotide **2.51**.

The product(s) of the full-length extension reaction with MeFAPy-dGuo oligonucleotide **2.52a** and -1 biotinated primer **2.54** catalyzed by a combination of pols $\eta/\iota$ was examined by LC-ESI/MS/MS. Similarly to the extension reaction catalyzed by pol $\eta$, the ESI spectrum of the extension reaction catalyzed by pols $\eta/\iota$ showed three molecular ion peaks with *m*/*z* 1078.5, 1086.9 and 1099.2 (Figures **2-61** and **2-62**). These three molecular ion peaks correspond to three oligonucleotides with masses 2159.1, 2175.8 and 2200.5 Da. It was assumed that these three oligonucleotides have the same composition as those identified as the extension products of pol $\eta$. The CID spectra for the molecular ions with *m*/*z* 1078.5, 1086.9 and 1099.2 (Figures **2-63**, **2-64** and **2-65**) were examined and showed fragmentation patterns indicative for the 5′-pTCCATGA-3′, 5′-pTCTATGA-3′, 5′-pTCGATGA-

3′, respectively, (Tables **2-19**, **2-20** and **2-21**), which confirmed our assumption.

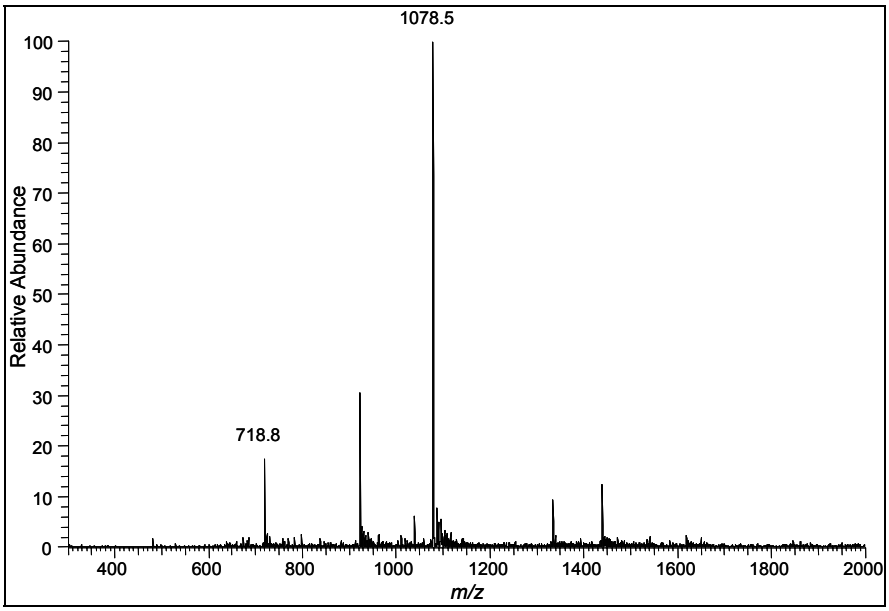



**Figure 2-61**. CID spectrum of the molecular ion peak *m/z* 1086.8 of the combination pols $\eta/\iota$ extension product(s) past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** and the primer **2.54**; molecular ion peaks with *m/z* 718.8 and *m/z* 1078.5 correspond to M-3H and M-2H, respectively, of the extension product with mass 2159.1, identified as 5′-pTCCATGA-3′.

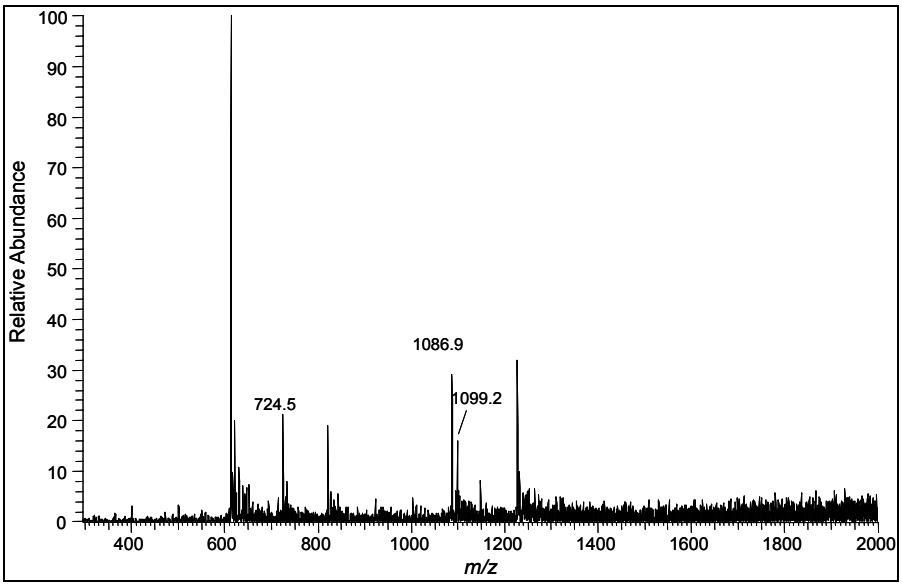



**Figure 2-62**. TIC spectrum of the LC-ESI/MS/MS analysis of the pols $\eta/\iota$ extension reaction product(s) past the MeFAPy-dGuo lesion in oligonucleotide **2.52a** using primer **2.54**; molecular ion peaks with *m/z* 724.6 and *m/z* 1086.9 correspond to M-3H and M-2H, respectively, of the extension product with mass 2175.8, identified as 5′-pTCTATGA-3′, molecular ion peaks with *m/z* 1099.2 correspond to M-2H of the extension product with mass 2200.4, identified as 5′-pTCGATGA-3′.

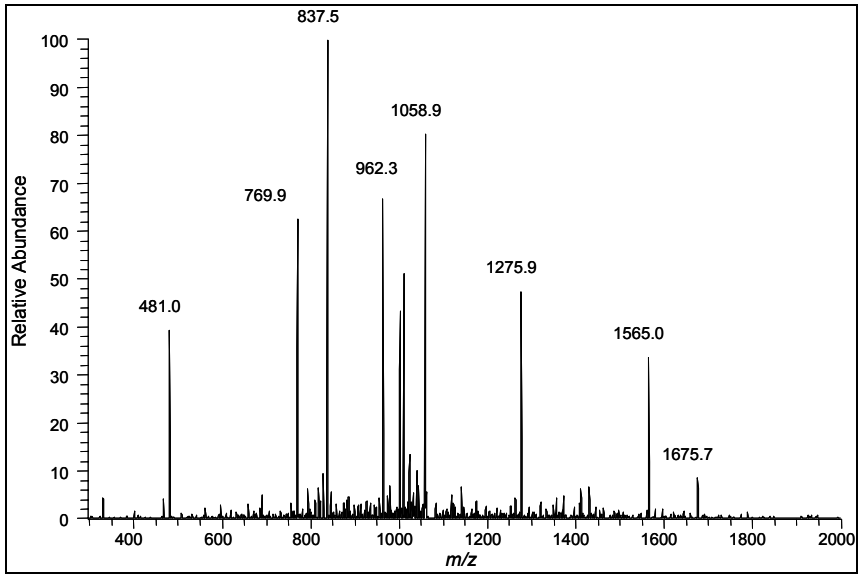



**Figure 2-63**. CID spectrum of the molecular ion peak *m/z* 1078.8 of the combination pols $\eta/\iota$ extension product, identified as 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-19**. Observed and theoretical fragmentation for the combination pols $\eta/\iota$ extension product *m/z* 1078.83, identified as 5′-pTCCATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481.0 | 481.0 |
| 5'-pTC ($a_3$-B) | 769.9 | 770.8 |
| 5'-pTCC ($a_4$-B) | 1058.9 | 1059.1 |
| 5'-pTCCAT ($a_6$-B) | 1675.7 | 1676.2 |
| ($a_6$-B, -2) | 837.5 | 837.6 |
| pCATGA-3' ($w_5$) | 1565 | 1565.2 |
| pATGA-3' ($w_4$) | 1275.9 | 1276.2 |
| pTGA-3' ($w_3$) | 962.3 | 963.1 |

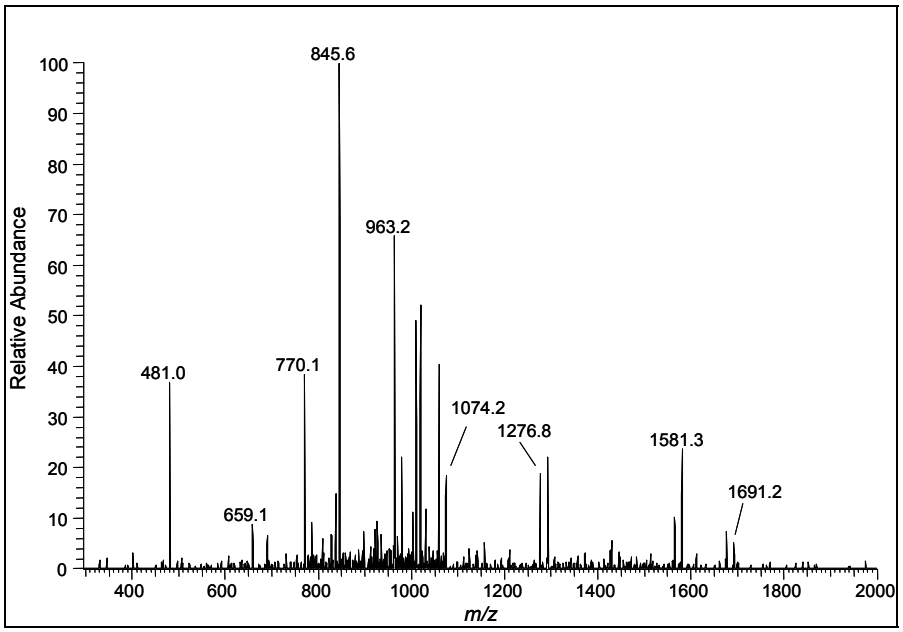



**Figure 2-64**. CID spectrum of the molecular ion peak *m/z* 1086.8 of the combination pols $\eta/\iota$ extension product, identified as 5′-pTCTATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-20**. Observed and theoretical fragmentation for the combination pols $\eta/\iota$ extension product *m/z* 1086.8, identified as 5′-pTCTATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481.0 | 481.0 |
| 5'-pTC ($a_3$-B) | 770.1 | 770.8 |
| 5'-pTCT ($a_4$-B) | 1074.2 | 1074.1 |
| 5'-pTCTAT ($a_6$-B) | 1691.2 | 1691.2 |
| ($a_6$-B, -2) | 845.6 | 845.1 |
| pTATGA-3' ($w_5$) | 1581.3 | 1580.2 |
| pATGA-3' ($w_4$) | 1276.8 | 1276.2 |
| pTGA-3' ($w_3$) | 963.2 | 963.1 |
| pGA-3' ($w_2$) | 659.1 | 659.1 |

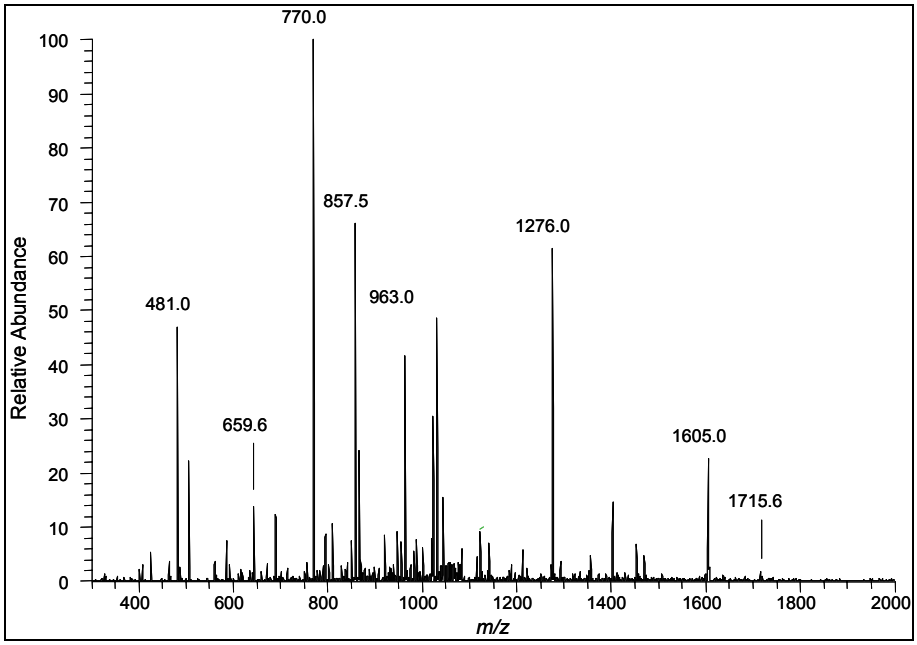



**Figure 2-65**. CID spectrum of the molecular ion peak *m/z* 1099.2 of the combination pols $\eta/\iota$ extension product, identified as 5′-pTCGATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

**Table 2-21**. Observed and theoretical fragmentation for the combination pols $\eta/\iota$ extension product *m/z* 1099.2, identified as 5′-pTCGATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide **2.52a**.

| Fragment assignment | Observed | Theoretical |
|---|---|---|
| 5'-pT ($a_2$-B) | 481.0 | 481.0 |
| 5'-pTC ($a_3$-B) | 770.0 | 770.8 |
| 5'-pTCGAT ($a_6$-B) | 1715.6 | 1716.2 |
| ($a_6$-B, -2) | 857.5 | 587.6 |
| pGATGA-3' ($w_5$) | 1605.0 | 1605.2 |
| pATGA-3' ($w_4$) | 1276.0 | 1276.2 |
| pTGA-3' ($w_3$) | 963.0 | 963.1 |
| pGA-3' ($w_2$) | 659.6 | 659.1 |

The yields of formation of the extension products 5′-pTCCATGA-3′, 5′-pTCTATGA-3′ and 5′-pTCGATGA-3′ were determined by using the internal standard 5′-pCTTACGAGCCCCC-3′ (**2.55**) and the calibration curves shown in Figures **2-31**, **2-47** and **2-48**. The yields of formation of the extension products 5′-pTCCATGA-3′, 5′-pTCTATGA-3′ and 5′-pTCGATGA-3′ were 14%, 0.6% and 0.1%, respectively. The yields of the extension products by the combination of pols $\eta/\iota$ are low compared to the yields of the corresponding extension products of pol $\eta$. A possible explanation for this difference is the inability of pol $\iota$ to efficiently bypass and extend past the MeFAPy-dGuo adduct in oligonucleotide **2.52a** (Figure **2-49**, Panel A).

If pol $\iota$ binds tightly to DNA before pol $\eta$ , this would prevent a dNTP incorporation across the MeFAPy-dGuo adduct, or if pol $\iota$ binds to DNA right

after pol $\eta$ have incorporated a certain dNTP across the MeFAPy-dGuo adduct this would prevent further extension of the products.

## Summary

We have synthesized a MeFAPy-dGuo phosphoroamidite for the first time. This phosphoroamidite was synthesized in four steps and 25% overall yield and was used for the preparation of oligonucleotides containing the MeFAPy-dGuo lesion (**2.05**) at a defined location. The crucial step of the oligonucleotide syntheses was the acid deprotection of the 5'-hydroxy group of the MeFAPy-dGuo nucleotide, since such conditions can cause the ribose ring to undergo a rearrangement to the pyranose form. The NMR experiments of a trinucleotide containing the MeFAPy-dGuo adduct (**2.05**), established that the removal of the 5'-hydroxyl group of the MeFAPy-dGuo nucleotide under the "short" and "long" deprotection cycles results in oligonucleotides containing the MeFAPy-dGuo furanose and pyranose forms of the MeFAPy-dGuo adduct (**2.05**), respectively.

Examination of the stability of MeFAPy-dGuo containing oligonucleotides showed that MeFAPy-dGuo adduct (**2.05**) persists in ss and ds-DNA at neutral pH but deglycosylation of the MeFAPy-dGuo oligonucleotides was observed in water or acidic buffer followed by cleavage of the resulting abasic site oligonucleotides. The deglycosylation of MeFAPy-

dGuo nucleotides occurs through a ribose ring-opened imine intermediate, which can be trapped by reduction with $Na(CN)BH_3$.

The *in vitro* bypass and full length-extension of the MeFAPy-dGuo adduct (**2.05**) was examined with prokaryotic and eukaryotic DNA polymerases. The MeFAPy-dGuo adduct (**2.05**) was found to be highly miscoding; in addition to the correct insertion of dCTP opposite the lesion, all three prokaryotic DNA polymerases, $Kf^-$, pol $II^-$ and Dpo4, mis-inserted dATP, dGTP and dTTP with varying efficiencies. The $k_{cat}/k_m$ values for dCTP insertion opposite the MeFAPy-dGuo lesion (**2.05**) with $Kf^-$ was 8 times lower than the insertion opposite an unmodified dGuo, while the misincorporation frequencies (*f*) for the other dNTP's ranged from ~ 5 to 30 percent (Table **2-22**).

Table **2-22**. Comparison of the for $Kf^-$ insertion efficiencies for the four dNTP's opposite the MeFAPy-dGuo (**2.05**), FAPy-dGuo (**2.59**) and 8-oxo-dGuo adducts (**2.60**).

| dNTP | Insertion efficiency (modified/dGuo) | | |
|---|---|---|---|
| | MeFAPy-dGuo | MeFAPy-dGuo[(a)] | FAPy-dGuo [(b)] |
| C | < 8 | < 66 | < 50 |
| T | > 25 | < 3.4 | < 6.7 |
| G | > 3.7 | < 0.0003 | < 1.4 |
| A | < 1.4 | >12.5 | > 17 |

a) ref.46; b) ref. 75, 76

The insertion efficiency for the dCTP opposite the MeFAPy-dGuo lesion, obtained by Asagoshi and co-workers using $Kf^-$ (*46*), was 66 times lower than the insertion efficiency opposite an unmodified dGuo base (Table **2-21**). In contrast to our observations that the insertion efficiency for dTTP opposite MeFAPy-dGuo adduct (**2.05**) was the highest (>25), Asagoshi and coworkers reported that the insertion efficiency for dATP was the highest; however, the extension of this product was not efficient. Similarly, we observed that the extension followed the misincoporation of dATP was blocked. Although in our bypass experiments with $Kf^-$ the insertion effciencies for dGTP and dTTP were high, the full-length extension of these products were not observed, suggesting the misinsertion of dGTP and dTTP completely inhibits the extension reaction.

Greenberg and coworkers (*75, 76*) reported that dATP was preferentially incorporated opposite the unsubstituted FAPy-dGuo (**2.59**) with an insertion efficiency 17 times higher, than the insertion opposite the unmodified sequence for $Kf^-$ (Figure **2-66**, Table **2-22**). Moreover, the preferential insertion of dATP was observed in mammalian cells, which resulted in a high level of FAPydGuo→dTyd conversions (*66*).

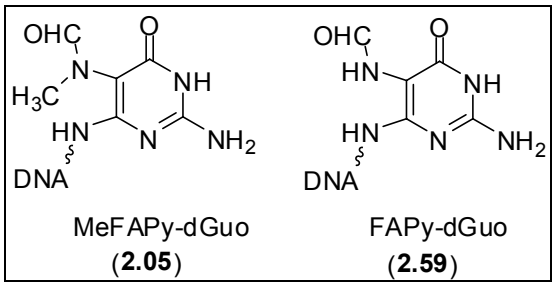



**Figure 2-66**. Structures of the MeFAPy-dGuo (**2.05**), FAPy-dGuo (**2.59**) and 8-oxo-dGuo (**2.60**) lesions.

The $k_{cat}/k_m$ values for dCTP insertion opposite the MeFAPy-dGuo lesion (**2.05**) with Dpo4 was 2.5 times lower than the insertion opposite an unmodified dGuo, while insertion of dCTP by pol II$^-$ was nearly equally efficient opposite the modified and unmodified base. Misincorporation frequencies (*f*) for the other dNTP's ranged from 1-2 and 11-10 percent for pol II$^-$ and Dpo4, respectively.

The MeFAPy-dGuo adduct (**2.05**) was also found to be a complete block to replication catalyzed by human replicative DNA polymerase pol $\delta$ and highly miscoding to the human DNA polymerases pols $\kappa$, $\eta$, and $\iota$, and the combinations of human pols $\eta$/ι and pols $\eta$/$\iota$. Human pol $\kappa$ inserted dCTP, dGTP and dTTP, while human pol $\eta$ inserted all four dNTP's opposite the MeFAPy-dGuo adduct (**2.05**). The same trends of dNTP incorporation of opposite the MeFAPy-dGuo adduct (**2.05**) were observed for the combinations of human pols $\eta$/ι and $\eta$/$\iota$ but with lower efficiency. The $k_{cat}/k_m$ values for dCTP insertion opposite the MeFAPy-dGuo lesion (**2.05**) with pol $\kappa$ and pol $\eta$ were 6 and 6.4 times lower than the insertion opposite an

unmodified dGuo. Misincorporation frequencies (*f*) for the other dNTP's ranged from ~ 2-11 and 13-73 percent for pol $\kappa$ and pol $\eta$, respectively.

The ability of the prokaryotic and eukaryotic translesion DNA polymerases to extend the MeFAPy-dGuo adduct (**2.05**) was also examined in the presence of all four dNTPs. Starting with the -1 primer, full-length extension products were observed for all polymerases. Kf$^-$ and pol II$^-$ showed significant pauses, which are due to the mis-incorporation of dGTP, dTTP and dATP opposite the MeFAPy-dGuo adduct (**2.05**). Dpo4 appeared to be the most efficient prokaryotic enzyme. The human polymerases extended past the MeFAPy-dGuo adduct (**2.05**) with different efficiency; the efficiency of pol $\iota$ was very low and the efficiencies of human pol $\kappa$ and pol $\eta$ alone were greater then the combinations of human pols $\eta/\iota$ or human pols $\eta/\iota$. The products of the extension reactions catalyzed by the prokaryotic and eukaryotic DNA polymerases were sequenced by LC-ESI/MS/MS. The primer used for the full-length extension reaction had dUrd incorporated at position 21 and a biotin attached at the 5'-end through a linker of ten dTyd's. The extension reaction of the biotinated primer was carried out in the presence of all four dNTP's and each of the DNA polymerases. The extended primer was bound to streptavidin coated beads, which allowed the clean up of the extension products and the UDG reaction from undesirable components, such as salts and enzymes. The use of biotinated primers improved greatly the sensitivity of the LC-ESI/MS/MS method.

The yields of formation of the full-length extension products were determined using an internal standard (Table **2-23**). The prokaryotic DNA polymerases bypassed and extended past the MeFAPy-dGuo (**2.05**) in error-free manner. Dpo4 was the most efficient polymerase with 72% formation of the free-error extension product. The eukaryotic DNA polymerases, pol $\kappa$, pol $\eta$, and combinations of pols $\eta/\iota$ and pols $\kappa/\iota$, bypassed and extended the MeFAPy-dGuo lesion (**2.05**) in both error-free and error-prone manner. Human pol $\kappa$ and pol $\eta$ bypassed and extended the MeFAPy-dGuo (**2.05**) lesion more efficiently than the combinations of pols $\eta/\iota$ and pols $\kappa/\iota$.

The error-prone products of the bypass and extension of MeFAPy-dGuo lesion (**2.05**) were identified as the 5′p-TC**G**ATGA-3′ and 5′p-TC**T**ATGA-3′. We did not detect the formation of the full-length extension product 5′p-TC**A**ATGA-3′, which was expected to be formed, since in mammalian cells the FAPy-dGuo (**2.59**) induced high levels of FAPy-dG→dTyd transversions with a mutation frequency 30% (*66*). A reasonable explanation is either that the extension following the mis-insertion of A is not efficient or the efficiency of misinsertion of A is sequence dependant as was observed for FAPy-dGuo (**2.59**) in mammalian cells (*66*).

**Table 2-23**. Summary of the yields of formation of the full-length extension product(s) of the MeFAPy-dGuo (**2.05**) lesion in oligonucleotide **2.52a** by prokaryotic and eukaryotic DNA polymerases.

| Enzyme | Inserted dNTP | Extension Products | Yield (%) | Comments |
|---|---|---|---|---|
| Kf⁻ | C | 5'-pTC**C**ATGA-3' | 26 | error free manner |
| | A, G, T | — — — | — — — | blocking |
| | | | | |
| Pol II⁻ | C | 5'-pTC**C**ATGA-3' | 18 | error free manner |
| | A, G, T | — — — | — — — | blocking |
| | | | | |
| Dpo4 | C | 5'-pTCCATGA-3' | 72 | error free manner |
| | T | — — — | — — — | — — — |
| | G | — — — | — — — | — — — |
| | A | — — — | — — — | — — — |
| | | | | |
| Pol $\kappa$ | C | 5'-pTC**C**ATGA-3' | 55 | error free manner |
| | T | 5'-pTC**T**ATGA-3' | 5 | error prone manner |
| | G | — — — | — — — | — — — |
| | A | — — — | — — — | — — — |
| | | | | |
| Pols $\kappa$/$\iota$ | C | 5'-pTC**C**ATGA-3' | 20 | error free manner |
| | T | 5'-pTC**T**ATGA-3' | 5 | error prone manner |
| | G | — — — | — — — | — — — |
| | A | — — — | — — — | — — — |
| | | | | |
| Pol $\eta$ | C | 5'-pTC**C**ATGA-3' | 54 | error free manner |
| | T | 5'p-TC**G**ATGA-3' | 5 | error prone manner |
| | G | 5'-pTC**T**ATGA-3' | 8 | error prone manner |
| | A | — — — | — — — | — — — |
| | | | | |
| Pols $\eta$/$\iota$ | C | 5'-pTC**C**ATGA-3' | 14 | error free manner |
| | T | 5'p-TC**G**ATGA-3' | 0.6 | error prone manner |
| | G | 5'-pTC**T**ATGA-3' | 0.1 | error prone manner |
| | A | — — — | — — — | — — — |
| | | | | |
| Pol $\iota$ | G, A, C | — — — | — — — | blocking |

## Experimental Procedures

### *General Methods*

$^1$H NMR spectra were recorded at 400 or 500 MHz on a Brucker AM Series NMR spectrometers in DMSO-$d_6$. All chemicals were the best available quality and used as received. Thin-layer chromatography was performed on silica gel glass plates (Merck, Silica Gel 60 $F_{254}$, layer thickness 250 $\mu$m). The chromatograms were visualized under UV light (254 nm) or by staining with an anisaldehyde/$H_2SO_4$ solution, followed by heating. Column chromatography was performed using silica gel (Merck, *70-230* mesh).

### *Mass Spectrometry*

Low and high-resolution FAB mass spectra were obtained at the University of Notre Dame Mass Spectrometry Facility (Notre Dame, IN) using a matrix of nitrobenzyl alcohol. Mass spectra (MALDI-TOF) of oligonucleotides were obtained at the Vanderbilt University MS Resource Laboratory on a Voyager STR instrument (Perseptive Biosystem). The system was operated in negative ion mode using a matrix of 3-hydroxypicolinic acid and ammonium hydrogen citrate.

## *Experimental Procedures for the Synthesis of MeFAPy-dGuo phosphoroamidite (2.36)*

*N*$^2$-*(*(Dimethylamino)methylene)-2'-deoxyguanosine (**2.32)** (*51*)

dGuo.$H_2O$ (3 g, 0.01 mol) was co-evaporated with dry pyridine (3 x 150 mL) and then dried overnight under high *vacuum*. The dry dGuo was suspended in dry methanol (100 mL) and N,N-dimethylformamide dimethyl acetal (3.43 mL, 0.02 mmol) was added. The reaction mixture was heated at 60 °C for 2 h. After cooling to ambient temperature, a white powder was precipitated. The powder was collected by suction filtration and washed with dry methanol to afford the pure **2.32** (2.4 g, 75%). $^1$H NMR (DMSO-$d_6$) 8.58 (s, 1H**,** N=CH), 8.12 (s, 1H, H-8), 6.25 (t, *J* = 8 Hz, 1H, H-1'), 5.35 (d, *J* = 3 Hz, 1H, OH-3'), 4.38-4.30 (m, 1H, H-3'), 4.20-4.00 (m, 1H, H-4'), 3.32-3.25 (m, 2H, H-5'), 3.15 (s, 3H, N-$CH_3$), 3.06 (s, 3H, N-$CH_3$), 2.82-2.68 (m, 1H, H-2'), 2.48-2.32 (m, 1H, H-2').

5'-O-Dimethoxytrityl-N$^2$-[(dimethylamino)methylidene]-2'-deoxyguanosine (**2.33**) (*51*).

Compound **2.32** (2 g, 0.006 mol) was suspended in dry pyridine (100 mL) in cooled to 10 °C; dimethoxytrityl chloride (2.23 g, 0.0066 mol) was added in three equal portions over a period of 1 h. After the addition of the last portion, the cooling bath was removed and the reaction mixture was stirred at room temperature for 5 h. After this time, the reaction was quenched

by addition of 1 mL of methanol and then the pyridine removed in *vacuo* with rotary evaporator. The resulting gummy product was dissolved in solvent, then silicagel was added and the solvent was removed. The dry silicagel was applied on the top of a packed column and the product was then eluted with a mixture of $CH_2Cl_2$ : methanol : pyridine 97 : 2 : 1, followed by an eluent mixture with 1% increasing gradient of methanol. The product **2.33** (2.35 g, 63%) was obtained as a white powder. $^1$H NMR (DMSO-$d_6$) 8.58 (s, 1H**,** N=CH), 8.12 (s, 1H, H-8), 7.40-6.83 (m, 13H, ArH), 6.28 (t, *J* = 6 Hz, 1H, H-1'), 5.37 (d, *J* = 3 Hz, 1H, OH-3'), 4.41-4.30 (m, 1H, H-3'), 4.00-3.90 (m, 1H, H-4'), 3.72 (s, 6H, 2-$OCH_3$), 3.20-3.00 (m, 2H, H-5'), 3.12 (s, 3H, N-$CH_3$), 3.02 (s, 3H, N-$CH_3$), 2.80-2.60 (m, 1H, H-2'), 2.40-2.20 (m, 1H, H-2').

N-3-[(Dimethylamino)methylene]-(5'-O-[bis(4-methoxyphenyl)phenylmethyl-2'-deoxy-β-D-*erythro*-pentofyranosyl)amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-N-(methyl)-formamide (**2.35**)

Methyl iodide (1.5 mL, 24 mmol) was added dropwise to a stirred solution of **2.32** (1 g, 1.60 mmol) in dry, degassed dimethyl sulfoxide (DMSO) (8 mL) at room temperature. After 45 min, the excess methyl iodide was removed in *vacuo* using rotary evaporator, the remaining DMSO was dilluted with 10 mL of distilled water and aqueous NaOH (1 mL, 1 M) was added dropwise over 1 min. The resulting clear solution was immediately neutralized by the dropwise addition of aqueous HCl (0.1 M). The change in the pH was monitored by pH paper. A precipitate was formed upon neutralization, which

was collected by suction filtration to give pure **2.35** (0.79 g, 75%). $^1$H NMR (DMSO-$d_6$) mixture of isomers: δ 11.10 (broad s, 0.5H, NHCO), 10.95 (broad s, 0.5H, NHCO), 8.37 (s, 0.5 H, CHO), 8.29 (s, 0.5H, CHO), 7.82 (s, 0.5H, N=CH), 7.74 (s, 0.5H, N=CH), 7.77-7.19 (m, 9H, ArH), 6.87-6.80 (m, 5H, NH + ArH), 6.32-6.29 (m, 1H, H'-1), 5.39 (d, 0.5H, OH-3', *J* = 10 Hz) 5.32 (d, 0.5H, OH-3', *J* = 10 Hz ), 4.31-4.22 (m, 1H, H'-3), 3.95-3.85 (m, 1H, H'-4), 3.71 (s, 6H, $2CH_3O$), 3.03-2.95 (m, 1H, H-5'), 2.80-2.75 (m, 1H, H-5'), 2.89 (s, 1.5H, $CH_3N$-CHO), 2.86 (s, 1.5H, $CH_3$-N-CH=), 2.84 (s, 1.5H, $CH_3N$-CHO), 2.79 (s, 3H, $CH_3$-N-CH=), 2.08-1.87 (m, 1H, H-2'), 2.33-2.15 (m, 1H, H-2'). HRMS (FAB$^+$) *m/z* calcd for $C_{35}H_{40}N_6O_7$ $[M + H]^+$ 656.2985, found 656.2990.

N-3-[(Dimethylamino)methylene]-3'O-[(2-cyanoethyl)-(N,N-diisopropyl)phosphoroamidite-(5'-O-[bis(4-methoxyphenyl)phenylmethyl -2' -deoxy-β-D-*erythro*-pentofyranosyl)amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-N-(methyl)-formamide (**2.36**)

Compound **2.35** (100 mg, 0.15 mmol) was dried by co-evaporation with anhydrous pyridine (3 x 10 mL) and dried overnight at high vacuum. The gummy residue was dissolved in dry methylene chloride (10 mL) and a solution of anhydrous 1H-tetrazole (12.60 mg, 0.18 mmol) was added, followed by the addition of 2-cyanoethyl-N,N,N',N'-tetraisopropylphosphoro-diamidite (63.29 mg, 0.21 mmol). This reaction mixture was stirred at room temperature for 2 h. After this time, the solvent was removed *in vacuo* with a rotary evaporator. The crude product was purified by flash chromatography

on silica gel eluting with methylene chloride : methanol : pyridine (97 : 2 : 1) to give **2.36** as a mixture of stereoisomers (100 mg, 78%). $^{1}$H NMR ($CD_2Cl_2$) mixture of isomers: δ 11.69 (bs, 0.5H, NHCO), 11.52 (bs, 0.5H, NHCO), 8.58 (s, 0.5H,CHO), 8.56 (s, 0.5H, CHO), 8.00 (s, 0.5H, N=CH), 7.95 (s, 0.5H, N=CH), 7.38-7.15 (m, 9H, aromatic), 6.90-6.75 (m, 4H, aromatic), 6.51-6.29 (m, 1H, H-1'), 6.01 (d, 0.5H, NH, *J* = 12 Hz), 5.92 (d, 0.5H, NH, *J* = 12 Hz), 4.74-4.51 (m, 1H, H-3'), 4.31-4.15 (m, 1H, H-4'), 3.78 (s, 6H, 2-$CH_3O$), 3.70-3.63 (m, 2H, $POCH_2$), 3.63-3.43 (m, 2H, isopropyl CH), 3.20-3.03 (m, 2H, H-5'), 3.04 (s, 3H, $CH_3$-N-CH=), 3.02 (s, 3H, $CH_3$-N-CH=), 3.00 (s, 1.5H, $CH_3$N-CHO), 2.83 (s, 1.5H, $CH_3$N-CHO), 2.65 (t, 2H, $OCH_2$-CN), 2.17-2.02 (m, 1H, H-2'), 2.02-1.89 (m, 1H, H-2'), 2.84 5.89), 1.22-1.01 (m, 12H, isopropyl $CH_3$). $^{31}$P NMR ($CD_2Cl_2$ 121 MHz) δ 150.47, 150.22, 149.74, 149.62, 149.59, 148.80. HRMS ($FAB^+$) *m/z* calcd for $C_{44}H_{58}N_8O_8P$ $[M + H]^+$ 857.4115, found 857.4143.

Oligonucleotide Synthesis

The oligonucleotides were synthesized on a Perseptive Biosystems Model 8909 DNA synthesizer on a 1-μmol scale using manufacturer's Expedite reagents with the standard synthetic protocol for the coupling of the unmodified bases. The coupling of the MeFAPy-dGuo phosphoroamidite (**2.36**) was performed off-line manually. The column was removed from the DNA synthesizer and sealed with two syringes, one of which contained the 1*H*-tetrazole activator solution (100 μL, 1.9–4.0% in $CH_3CN$, w/v, Applied

Biosystems, Foster City, CA) and the other contained the MeFAPy-dGuo (**2.36**) phosphoroamidite (7 mg in 100 μL of amidite diluent (Applied Biosystems, Foster City, CA)). The 1*H*-tetrazole and the phosphoramidite solutions were sequentially drawn through the column (1*H*-tetrazole first), and this procedure was repeated periodically over 30 min. After this time, the column was washed with the manufacturer's amidite diluent and returned to the instrument for the capping, oxidation, and detritylation steps (*52, 77*). The DMTr group of the MeFAPy-dGuo was removed on line with protocol 10/20 (160 μL of $Cl_3CCOOH$ for 20 s, short deprotection) for MeFAPy-dGuo furanose form and with protocol 20/50 (320 μL of $Cl_3CCOOH$ for 50 s, long deprotection) plus an additional 3 min at standby mode for MeFAPy-dGuo pyranose form. The remainder of the synthesis was performed on-line using standard protocols. The modified oligonucleotides were cleaved from the solid support and the exocyclic amino groups were deprotected in a single step using aqueous NaOH (1.2 mL, 0.1 M) at room temperature for overnight.

Nucleoside and Oligonucelotide Purifications

For reaction monitoring and purification, a YMC ODS-AQ column (250 x 4.6 mm, flow rate 1.5 mL/min, 250 x 4.6 mm, flow rate 10 mL/min) or Phenomenex Gemini-C18 column (250 x 4.6 mm, flow rate 1.5 mL/min, 250 x 4.6 mm, flow rate 5 mL/min) monitored at 254 nm were used. HPLC analyses and purifications were carried out on a HPLC (Beckman Instruments; System

Gold Software) equipped with pump module 125 and photodiode array detector module 168.

For the purification of nucleosides a mobile phase consisting of $H_2O$ (A) and $CH_3CN$ (B) was used. For oligonucleotide purification a mobile phase consisting of 100 mM aqueous ammonium formate (A) and $CH_3CN$ (B) was used. The following HPLC gradients were employed.

Gradient 1: initially 1% B; 15 min linear gradient to 10% B; 5 min linear gradient to 20% B; isocratic at 20% B for 5 min; 2 min linear gradient to 80% B; isocratic at 80% B for 3 min; 3 min linear gradient to 1% B (initial conditions).

Gradient 2: initially 1% B; 5 min linear gradient to 5% B; 15 min linear gradient to 12% B; 2 min linear gradient to 80% B; isocratic at 80% B for 2 min; 3 min linear gradient to 1% B (initial conditions).

Gradient 3: 1% initially B; 10 min linear gradient to 5% B; 5 min linear gradient to 10% B; 2 min linear gradient to 80% B; isocratic at 80% B for 2 min; 3 min linear gradient to 1% B (initial conditions).

The gel purification of the oligonucleotides was conducted on denaturing gel containing 8.0 M urea and 16% acrylamide (w/v) (from a 19:1 acrylamide/bis-acrylamide solution (AccuGel, National Diagnostics, Atlanta, GA) with 80 mM Tris borate buffer (pH 7.8) containing 1 mM EDTA.

Enzymatic Analysis of the MeFAPy-dGuo (**2.05**) Containing Oligonucleotides

Enzymatic hydrolysis of the MeFAPy-dGuo (**2.05**) containing oligonucleotides was carried out in one step as follows: oligonucleotide (0.5 $A_{260}$ units) was dissolved in 70 $\mu$L of buffer (pH 7, 0.01 M Tris-HCl, 0.01 M $MgCl_2$). DNase I (5 units), alkaline phosphatase (1.7 units), and snake venom phosphodiesterase I, *Type II* (0.02 units) were added and the solution was incubated at 37 °C for 1.5 h. HPLC analysis was carried out on a HPLC using gradient 1.

Synthesis of 5′-CTT-(MeFAPy-dGuo)-TT-3′ (**2.37a** and **2.37b**)

Purified by HPLC gradient 1. MALDI-TOF MS (HPA) *m/z* calcd for **2.37a** [M-H] 1803.3, found 1803.8; *m/z* calcd for **2.37b** [M-H] 1803.3, found 1803.1.

Synthesis of 5′-A-(MeFAPy-dGuo)-C-3′ (**2.38a** and **2.38b**)

These oligonucleotides were synthesized by the “trityl-on” method in which the 5′-end dimethoxytrityl (DMTr) group was not removed during solid-phase oligonucleotide synthesis. The cleavage from the solid support and the removal of the protecting groups were performed with conc. ammonia at room temperature overnight. After filtration through filter (Millex$^{TM}$, Milipore Corporation, Bedford, MA) and drying *in vacuo*, the oligonucleotide was passed through a Poly-Pak cartridge (Glen Research Co., Ltd.) with 2% trifluoroacetic acid (TFA) to remove the 5’-end DMTr group. The

oligonucleotide was eluted with 20% acetonitrile in water and purified by reverse phase chromatography using gradient 1. ESI/MS for **2.38a** calcd *m/z* 900.2; found 900.1; for **2.38b** cacld *m/z* 900.2, found 900.1.

Synthesis of 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2.42a** and **2.42b**)

Purified by HPLC gradient 2. MALDI-TOF MS (HPA) *m/z* calcd for **2.42a** (M-H) 3841.6, found 3842.9; *m/z* calcd for **2.42b** (M-H) 3841.6, found 3843.4.

Synthesis of 5′-TCAT-(MeFAPy-dGuo)-GAATCCTTACGAGCATCGCCCCC-3′ (**2.52a** and **2.52b**)

Purified by gel electrophoresis. MALDI-TOF MS (HPA) m/z calcd for **2.51a** (M-H) 8496.4, found 8498.1; m/z calcd for **2.51b** (M-H) 8496.4, found 8495.2.

Synthesis of 5′-TCAT-(MeFAPy-dGuo)-GATCCTTCCCCC-3′ **(2.61a** and **2.62b)**

Purified by gel electrophoresis. MALDI-TOF MS (HPA) m/z calcd for **2.56a** (M-H) 5401.2, found 5401.9, m/z calcd for **2.56b** (M-H) 5401.2, found 5401.3.

Capillary Zone Electrophoresis (CZE)

Electrophoretic analyses were carried out on a Beckman P/ACE MDQ Capillary Electrophoresis System 5500 Series monitored at 260 nm. The P/ACE instrument used a 31.2 cm x 100 μm eCAP capillary with samples applied at 10 KV and run at 9 KV. The column was packed with the manufactured 100-R gel using a tris-borate buffer system containing 7.0 M urea.

MALDI-TOF MS Sequencing of Oligonucleotides

The modified oligonucleotides (0.3 $A_{260}$ units) were treated with 2 milliunits of phosphodiesterase I (PI) in ammonium hydrogen citrate buffer (pH 9.4, 24 μl, 20 mM $MgSO_4$) at 37 °C. Aliquots of 4 μL were taken before enzyme addition and at 1, 8, 18, 28 and 38 min time points after enzyme addition. The aliquots were combined in the same, which was kept frozen on dry ice. A complementary experiment was performed in which the modified oligonucleotides (0.3 $A_{260}$ units) were incubated with 2 milliunits of phosphodiesterase II (PII) in ammonium acetate (20 mM, pH 6.6). The aliquots were taken in at the same manner. The two digested mixtures were desalted using Millipore $C_{18}$ Ziptips and eluted onto a MALDI plate on a matrix of 3-hydroxypicolinic acid and ammonium hydrogen citrate.

*NMR Spectroscopy of Trinucleotide 5'-A-(MeFAPy-dGuo)-C-3' (**2.38a** and **2.38b**)*

Two-dimensional techniques ($^1$H-$^1$H double quantum filtered correlated spectroscopy (DQF-COSY) (*78*), $^1$H-$^1$H nuclear Overhauser effect spectroscopy (NOESY) (*79, 80*), $^1$H-$^1$H total correlated spectroscopy (TOCSY) (*81*), and multiplicity edited $^1$H-$^{13}$C heteronuclear single quantum correlation spectroscopy (HSQC) (*82*) experiments) were performed at a $^1$H frequency of 500 MHz on a Bruker Avance spectrometer with a 5 mm CP-TCI-Z cryoprobe. Heteronuclear $^1$H-$^{31}$P COSY (*83*) experiments were performed at a $^1$H frequency of 600 MHz on Bruker Avance spectrometer with a QXI-XYZ probe. Experiments were conducted at 15 ± 0.5 ºC. $^1$H and $^{31}$P spectra were referenced to internal 3-(trimethylsilyl)propionic-2,2,3,3-$d_4$ acid, sodium salt (3-TMSP) and external 85% $H_3PO_4$ (capillary in $H_2O$) respectively. Typical acquisition parameters for homonuclear experiments were as follows: 2K complex data points with 512 increments, 32 scans per FID, sweep width of 5250 Hz in both dimensions, relaxation delay of 2.0 s, States-TPPI mode. Presaturation was sufficient for water suppression. NOESY and TOCSY spectra were recorded with mixing times of 300 and 80 ms respectively. Acquisition parameters for heteronuclear $^{13}$C-$^1$H HSQC experiments were as follows: 1K complex data points (sweep width = 5250 Hz), 360 increments (sweep width = 90000 Hz), 32 scans per FID, garp decoupling, relaxation delay of 1.8 s, Echo-Antiecho acquisition mode, pulses were optimized for $^1J_{CH}$ coupling constants (140 Hz). Acquisition parameters

for heteronuclear $^{31}P$-$^{1}H$ COSY experiments were as follows: 1K complex data points (sweep width = 6000 Hz) with 256 increments (sweep width = 3000 Hz), 128 scans per FID, relaxation delay of 1.8 seconds, STATES mode, pulses were optimized for $^{3}J_{PH}$ coupling constants (12 Hz). Carrier frequencies were set at 4.7 ppm for $^{1}H$, 80 ppm for $^{13}C$, and 0 ppm for $^{31}P$. XWINNMR (v 3.5 patch level 6, Bruker Inc., Karlsruhe, Germany) was used for data processing. Apodization was attained using a skewed sinebell-squared function with a 90 degree shift in both dimensions (180 degrees for COSY experiments); zero-filling and linear prediction was applied in the indirect dimension. Resonance assignments and peak integration was performed using the program SPARKY (*84*).

Stability of 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2**.**42a** and **b**) in Water and Phosphate Buffer pH 7.0

The oligonucleotides **2**.**42a** and **2.42 b** (5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′) (0.5 $A_{260}$ units) were dissolved in water (350 µL) or 100 mM phosphate buffer pH 7.00 (350 µL) and stirred at room temperature or heated at 95 °C. Aliquots (10 µL) were taken periodically for analyses. The oligonucleotides dissolved in the water were analyzed by MALDI-TOF-MS without purification, whereas the oligonucleoitdes dissolved in phosphate buffer were desalted through a Millipore $C_{18}$ Ziptips and then eluted onto a MALDI plate using 3-hydroxypicolinic acid and ammonium hydrogen citrate as the matrix.

Stability of 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2.42a**) in Phosphate Buffer pH 6.5

The oligonucleotide 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2.42a**) (0.5 $A_{260}$ units) was dissolved in 100 mM phosphate buffer pH 6.5 (350 µL) and stirred at room temperature. Aliquots (10 µL) were taken after 24 h and 48 h, and analyzed by the LC-ESI/MS/MS method as described for the analysis of the full-length extension products from the DNA polymerases bypass studies.

Stability of 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2.42a**) in water at 95 °C, Followed by Treatment with T4-pdg (Endo V)

The oligonucleotide 5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (**2.42a**) (1 $A_{260}$ units) was dissolved in water (350 µL). After heating at 95 °C for 1 h, the reaction mixture was analyzed by HPLC using gradient 2. The HPLC analysis showed that all of the starting material was converted to the abasic site oligonucleotide **2.43a**. The reaction mixture was spilt in two portions. The first portion was heated at 95 °C for additional 1.5 h and analyzed by HPLC, while the second portion was lyophilized. The lyophilized sample was dissolved in buffer containing 100 mM NaCl, 10 mM EDTA, 10 mM Tris (pH 8) and 100 µg BSA, followed by addition of 1:10 diluted solution of T4-pdg (Endo V). This reaction mixture was set at 37 °C for 24 h and then analyzed by HPLC using gradient 2.

Reduction of 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2.42a**) in Water with NaB (CN)$H_3$

The oligonucleotide 5′-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3′ (**2.42a**) (1 $A_{260}$ units) was dissolved in water (500 µL) and Na(CN)$BH_3$ (9.4 mg, 0.15 mmol) was added in 5 equal portions over 48 h. Aliquot (15 µL) was taken, desalted through a Millipore $C_{18}$ Ziptips, and then eluted onto a MALDI plate using matrix of 3-hydroxypicolinic acid and ammonium hydrogen citrate as a matrix. Another aliquot (125 µL) was taken and aqueous HCl (25 µL, 0.5 N) was added. After stirring at room temperature for overnight, the acidic reaction mixture was analyzed by the HPLC method using gradient 2.

An aliquot (125 µL) of the Na(CN)$BH_3$ reaction mixture was desalted and lyophilized. The lyophilized sample was subjected to enzyme digestion for 24 h according to the procedure used for the digestion of MeFAPy-dGuo containing oligonucleotides. The enzyme digestion was analyzed by the LC-ESI-MS/MS method described for the analysis of the full-length extension products from the DNA polymerases.

Synthesis of the Reduced MeFAPy-dGuo (**2.50**): N-[(1,2-dideoxy-D-*erythro*-pentitol)2-amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-6-amino-N-(methyl)-formamide

MeFAPy-dGuo nucleoside (**2.16**) (2 mg, 0.0066 mmol) was dissolved in water and Na(CN)$BH_3$ (41 mg, 0.66 mmol) were added in equal portions for a period of 30 days. The reaction was monitored and the product was purified

by HPLC with Phenomenex Gemini-C18 column (250 x 4.6 mm, flow rate 1.5 mL/min, 250 x 4.6 mm, flow rate 5 mL/min) using gradient 3. The HPLC purification afforded pure **2.49** (0.5 mg, 25%). $^{1}$H NMR (DMSO-$d_6$) δ *10.18* (brs, 1h, NH) 7.68 (s, 1H, CHO), 6.50 (bs, 2H, $NH_2$), 6.37 (m, 1H, NH), 4.55 (m, 1H, OH-4'), 4.46 (m, 1H, OH-5'), 4.32 (m, 1H, OH-3' ), 3.49-3.47 (m, 2H, H-3'), 3.45-3.30 (m, 4H, H-1', H-4'), 3.24-3.23 (m, 2H, H-5'), 2.76 (m, 3H, $CH_3$), 1.80-1.78 (m, 1H, H-2'), 1.39-1.36 (m, 1H, H-2'), $^{13}$C NMR (125 MHz, DMSO-$d_6$): $\delta$ 165.66, 163.32, 160.38, 154.51, 94.01, 74.86, 69.56, 63.36, 37.67, 33.11, 31.22; HRMS (FAB$^+$) *m/z* calcd for $C_{11}H_{20}N_5O_5$ $[M + H]^+$ 302.1464, found 302.245

***Experimental Procedures for DNA Polymerase Bypass Assays***

Enzymes and Materials

Kf$^-$, Pol II$^-$, and Dpo4 were expressed and purified as previously described (*85, 86*). Human DNA polymerases pol $\delta$, $\eta$, $\kappa$, and $\iota$ were purchased from Enzymax (Lexington, KY). Uracil DNA glycosylase (UDG) was obtained from Sigma Chemical Co. (St. Louis, MO). Piperidine was purchased from Aldrich and used as received. dNTP solutions (100 mM) were purchased from GE Healthcare (formerly Amersham Biosciences, Piscataway, NJ). Unmodified oligonucleotides and primers were purchased from Midland Certified Reagents (Midland, TX). Streptavidin coated beads

(Streptavidin Sepharose™ High Performance) were purchased from GE Healthcare (Piscataway, NJ, USA).

Labeling and Annealing of Oligonucleotides

The primer was 5'-end-labeled with [$\gamma$-$^{32}$P]-ATP (specific activity 3000, PerkinElmer Life Sciences) using T4 polynucleotide kinase (New England Biolabs, Ipswich, MA) according to the manufacturer's instructions and purified on a Biospin column (BioRad, Hercules, CA). Template and $^{32}$P-labeled primer (1:1 molar ratio) were annealed in 50 mM Tris-HCl buffer (pH 7.8) by heating at 90 °C for 5 min and then slowly cooling to 30 °C.

Single-Nucleotide Incorporation Assays

$^{32}$P-labeled primers were annealed to either the unmodified or the modified (adducted) template, and extension reactions were then carried out in the presence of single dNTP's. The reactions with Kf$^-$, pol II$^-$, and Dpo4 were initiated by the addition of the dNTP (final dNTP concentration was 25, 50, and 100 nM) to preincubated enzyme/DNA mixtures giving a final reaction volume of 10 $\mu$L. The polymerase reactions with pols $\delta$, $\eta$, $\kappa$, pol $\iota$, and the combinations of pols $\kappa/\iota$ and pols $\eta/\iota$ were initiated by the addition of dNTP with final concentrations of 120, 250, and 1000 nM. The final concentrations for the DNA polymerases were 24 nM of Kf$^-$, 24 nM of pol II$^-$, 40 nM of Dpo4, 2 nM of pol $\delta$, 2 nM of pol $\eta$, 2 nM of pol $\kappa$, 2 nM of pol $\iota$, 2 nM of each pol $\eta$ and pol $\iota$ for the combination of pols $\eta/\iota$ and 2 nM of each pol $\kappa$ and pol $\iota$ for

the combination of pols $\kappa/\iota$. The final concentrations of the components for the enzyme reactions were 100 nM DNA duplex, 50 mM Tris-HCl (pH 7.8), 1 mM dithiothreitol of (DTT), 50 $\mu$g bovine serum albumin (BSA) $mL^{-1}$, 50 mM NaCl, and 5 mM $MgCl_2$. The $Kf^-$ and pol $II^-$ reactions were run at room temperature for 10 min, the Dpo4 reactions were run at 37 °C 30 min and pol $\eta$, pol $\kappa$, pol $\iota$ and the combinations of pols $\eta/\iota$ and pol $\kappa/\iota$ reactions were run at 37 °C for 10 min. The polymerase reactions with $Kf^-$, pol $II^-$ and Dpo4 with primer/unadducted template were run at the same reaction conditions as the modified sequence, but with the two-fold less enzyme concentrations and reaction time. The polymerase reactions with pols $\delta$, $\eta$, $\kappa$, pol $\iota$ , and the combinations of pols $\kappa/\iota$ and pols $\eta/\iota$ with primer/unadducted template were run at the same reaction conditions as the modified sequence, but with the two-fold less enzyme concentrations. The polymerase reactions were quenched with 70 $\mu$L of 20 mM EDTA in 95% formamide (v/v) containing xylene cyanol and bromophenol blue dyes and heated at 95 °C for 10 min. Aliquots (6 $\mu$L) were separated by electrophoresis on a denaturing gel containing 8.0 M urea and 16% acrylamide (w/v) (from a 19:1 acrylamide/bisacrylamide solution, AccuGel, National Diagnostics, Atlanta, GA) with 80 mM Tris borate buffer (pH 7.8) containing 1 mM EDTA. The gel was exposed to a PhosphorImager screen (Imaging Screen K, Bio-Rad) overnight. The bands were visualized with a PhosphorImaging system (Bio-Rad, Molecular Imager FX) using the manufacturer's Quantity One software, version 4.3.0.

Full-Length Extension Assay with All Four dNTP's

A $^{32}$P-labeled primer was annealed to either an unmodified or an modified template and extended in the presence of all four dNTP's. The final dNTP concentration for the reactions with Kf$^-$, pol II$^-$, and Dpo4 was 10, 50, and 100 nM each. The reactions with pols $\delta$, $\kappa$, $\eta$ and the combinations of pols $\kappa/\iota$ and pols $\eta/\iota$ were initiated by the addition of dNTP with final concentrations of 120, 250, and 1000 nM. Each reaction was initiated by adding the mixture of dNTPs to preincubated enzyme/DNA mixtures in a mixture of 100 nM DNA duplex, 50 mM Tris-HCl (pH 7.8), 1 mM DTT, 50 $\mu$g BSA mL$^{-1}$, 50 mM NaCl, and 5 mM $MgCl_2$, giving a final reaction volume of 10 $\mu$L. The enzymes final concentrations were 24 nM Kf$^-$, 24nM pol II$^-$, 40 nM Dpo4, 2 nM pol $\delta$, 2 nM pol $\eta$, 2 nM pol $\kappa$, 2 nM pol $\iota$, 2 nM of each pol $\eta$ and pol $\iota$ for the combination of pols $\eta/\iota$ and 2 nM of each pol $\kappa$ and pol $\iota$ for the combination of pols $\kappa/\iota$. The Kf$^-$ and pol II$^-$ polymerase reactions were run at room temperature for 10 min, the Dpo4 reactions were run at 37 °C for 30 min and the polymerase reactions with pols $\delta$, $\eta$, $\kappa$, $\iota$ and the combination of pols $\eta/\iota$ and pols $\kappa/\iota$ reactions were run at 37 °C for 10 min. The polymerase reactions with Kf$^-$, pol II$^-$ and Dpo4 with primer/unadducted template were run at the same reaction conditions as the modified sequence, but with the two-fold less enzyme concentrations and reaction time. The polymerase reactions with pols $\delta$, $\eta$, $\kappa$, pol $\iota$, and the combinations of pols $\kappa/\iota$ and pols $\eta/\iota$ with primer/unadducted template were run at the same reaction conditions as the modified sequence, but with the two-fold less enzyme concentrations. The

polymerase reactions were quenched by the addition of 70 μL of 20 mM EDTA in 95% formamide (v/v) containing xylene cyanol and bromophenol blue dyes and heated at 95°C for 10 min. Aliquots (6 μL) were separated by electrophoresis on a denaturing gel containing 8.0 M urea and 16% acrylamide (w/v) (from a 19:1 acrylamide/bisacrylamide solution, AccuGel, National Diagnostics, Atlanta, GA) with 80 mM Tris borate buffer, pH 7.8, containing 1 mM EDTA. Gels were exposed to a PhosphorImager screen (Imaging Screen K, Bio-Rad) overnight. The bands were visualized with a PhosphorImaging system (Bio- Rad, Molecular Imager FX) using the manufacturer's Quantity One software, version 4.3.0.

Steady-State Kinetics

A $^{32}$P-labeled primer, annealed to either an unmodified or a modified template, was extended in the presence of a single dNTP. Reaction times varied from 5 to 30 min at 37 °C for Dpo4, pol $\eta$, pol $\kappa$, and room temperature for Kf$^-$ and pol II$^-$. The molar ratio of primer-template/enzyme varied from 20:1 to 5:1 for modified oligonucleotides and up to 40:1 for unmodified oligonucleotides. Reaction times, enzyme concentrations, dNTP concentrations were chosen so that the maximium product formation would be ~ 20% of the substrate concentration. All reactions (10 μL final volume) were run at eight dNTP concentrations (in duplicate) and quenched with 70 μL of 20 mM EDTA in 95% formamide (v/v) containing xylene cyanol and bromophenol blue dyes and heated at 95 °C for 10 min. Aliquots (6 μL) were separated by electrophoresis on denaturing gels containing 8.0 M urea and

16% acrylamide (w/ v) (from a 19:1 acrylamide/bisacrylamide solution, AccuGel, National Diagnostics, Atlanta, GA) with 80 mM Tris borate buffer (pH 7.8) containing 1 mM EDTA. The gel was exposed to a PhosphorImager screen (Imaging Screen K, Bio-Rad) overnight. The bands were visualized with a PhosphorImaging system (Bio-Rad, Molecular Imager FX) using the manufacturer's Quantity One software, version 4.3.0. The steady-state rates versus dNTP concentration were fit using nonlinear regression (hyperbolic fits) in GraphPad Prism (version 4.0, GraphPad, San Diego, CA) for the estimation of $k_{cat}$ and $K_m$ values.

LC-ESI/MS/MS Analysis of Oligonucleotide Products from Kf$^-$, Pol II$^-$, Dpo4, Pol $\eta$, Pol $\kappa$ and the Combinations of Pols $\eta/\iota$ and $\kappa/\iota$.

Kf$^-$, pol II$^-$, Dpo4, pol $\eta$, pol $\kappa$, and the combinations of pols $\eta/\iota$ and $\kappa/\iota$ reactions were performed for 6 h in 50 mM of Tris-HCl (pH 7.8), 100-500 of nM DNA duplex containing the biotinated primer (**2.54**), 1 mM DTT, 50 $\mu$g BSA mL$^{-1}$, 50 mM NaCl, and 5 mM $MgCl_2$ (200 $\mu$L total reaction mixture). The reactions were performed with all four dNTP's at 1 mM each and polymerase concentrations of 100 nM of Kf$^-$, pol II$^-$ and Dpo4, and 40 nM of pol $\eta$, pol $\kappa$, and the combinations of pols $\eta/\iota$ and $\kappa/\iota$. To this reaction mixtures streptavidin coated beads (0.5 ml of the streptavidin solution was centrifuged and subsequently washed 3 x 500 µL 100 mM of phosphate buffer (PS) pH 7) and 600 µL of PS pH 7 were added and the resulting suspension was set at a rotating shaker for 2 h. The streptavidin coated beads were washed with 1

mL $H_2O$ (3 x 333 µL). The UDG hydrolysis was conducted with 20 units of UDG in 500 µL solution of 50 mM of Tris-HCl, 1 mM of EDTA, and 1 mM of DTT for 4 h at 37 °C. The streptavidin coated beads were washed with $H_2O$ (3 x 333 µL) and then heated in the presence of piperidine (final concentration of 0.25 M) at 95 °C for 1 h. The liquid was decanted and the beads were washed with 200 $H_2O$ (3 x 200 µL). The combined water washes were lyophilized and then the residue was dissolved in water (70 µL). Aliquot (20 µL) was taken and 0.0068 $A_{260}$ units of the oligonucleotide 5′-pCTTACGAGCCCCC-3′ (**2.56**) (standard) was added. MS analysis was performed in the Vanderbilt University facility on a Waters Acquity UPLC system (Waters, Milford, MA) connected to a Finnigan LTQ mass spectrometer (ThermoElectron) using an Acquity UPLC BEH C18 column (1 *µ*m, 1.0 mm x 100 mm). LC conditions were as follows: buffer A contained 10 mM $NH_4CH_3CO_2$ plus 2% $CH_3CN$ (v/v) and buffer B contained 10 mM $NH_4CH_3CO_2$ plus 95% $CH_3CN$ (v/v). The following gradient program was used with a flow rate of 150 *µ*L: initially 0% B; 3 min linear gradient to 3% B; 1.5 min linear gradient to 20% B; 0.5 min linear gradient 100% B; isocratic at 100% B for 0.5 min; 1 min linear gradient to 0% B; isocratic at 0% B for 3 min. The temperature of the column was maintained at 50 °C and the samples (10 *µ*L) were infused with an auto-sampler. The electrospray conditions were as follows: source voltage 4 kV, source current 100 *µ*A, auxiliary gas flow-rate setting 20, sweep gas flow-rate setting 5, sheath gas flow setting 34, capillary voltage -49 V, capillary temperature 350 °C, and tube lens voltage -90 V.

MS/MS conditions were as follows: normalized collision energy 35%, activation *Q* 0.250, and activation time 30 ms. Product ion spectra were acquired over the range *m*/*z* 345-2000. The ions were selected for CID analysis and the calculations of the CID fragmentations of the candidate oligonucleotide sequence were done using the Mono Oligo Mass Calculator (v. 2.6) from the Mass Spectrometry Group of Medicinal Chemistry at the University of Utah (http://library.med.utah.edu/masspec/). After the oligonucleotide sequence was identified, the proposed sequence was purchased from Midland Certified Reagents (Midland, TX) and subjected to the same LC-ESI/MS/MS analysis in order to compare the CID spectra.

Construction of the Calibration Curves for the Measurements of the Percentage of Formation of the Full-length Extension Products

The standard calibration curves were constructed using 5 to 7 concentrations of the corresponding oligonucleotide (analyte) and a constant amount (0.0034 $A_{260}$ units) of the 5'-pCTTCACGAGCCCCC-3' (**2.55**) (standard). The analytes and the standard were purchased from Midland Certified Reagents (Midland, TX). The yield of formation of the corresponding analyte was calculated based on ratio amount ($A_{260}$ units) of the full-extension product and the amount ($A_{260}$ units) the biotinated primer used for the full-length extension reaction.

## References


(1) Siebenlist, U. and Gilbert, W. (1980) Contacts between *Escherichia coli* RNA polymerase and an early promotor of phage T7. *Proc. Natl. Acad. Sci. USA 77*, 122-126.

(2) Hayashibara, K. C. and Verdine, G. L. (1991) Template-directed interference footprinting of protein-guanine contacts in DNA. *J. Am. Chem. Soc. 113*, 5104-5106.

(3) Rydberg, B. and Lindahl, T. (1982) Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J. 1*, 211-216.

(4) Barrows, L. R. and Magge, P. N. (1982) Nonenzymatic methylation of DNA by S-adenosylmethionine *in vitro*. *Carcingenesis 3*, 349-351.

(5) Holliday, R. and Ho, T. (1998) Gene silencing and endogenous DNA methylation in mammalian cells. *Mutat. Res. 400*, 361-368.

(6) Bartsch, H., Ohsima, H., Shuker, D. E. G., Pignatelli, B. and Calmes, S. (1990) Exposure of humans to endogenous N-nitroso compounds: implications in cancer etiology. *Mutat. Res. 238*, 255-267.

(7) Kodama, M. and Saito, H. (1980) Formation of methylurea from methylamine and carbamyl phosphate: a possible environmental hazard. *Cancer Lett. 10*, 319-324.

(8) Vaughan, P. and Sedgwick, B. (1991) A weak adaptive response to alkylation damage of nitrosation. *Environ. Mol. Mutagen. 15*, 69-70.

(9) Hecht, S. S. (1998) Biochemistry, biology, and carcinogenicity of tobacco-specific N-nitrosamines. *Chem. Res. Toxicol. 11*, 559-603.

(10) Smith, G. B., Catonguay, A., Donnelly, P. J., Reid, K. R., Petsikas, D. and Massey, T. E. (1999) Biotransformation of the tobacco-specific carcinogen 4-(methylnitrosoamino)-1-(3-pyrridyl)-1-butanone (NNK) in freshly isolated human lung cells. *Carcingenesis 20*, 1809-1818.

(11) Hecht, S. S. (1999) Tobacco smoke carcinogens and lung cancer. *J. Natl. Cancer Inst. 91*, 1194-1210.

(12) Wogan, G. N., Hecht, S. S., Felton, J. S., Conney, A. H. and Loeb, L. A. (2004) Environmental and chemical carcinogenesis. *Semin. Cancer Biol. 14*, 473-486.

(13) Vodicka, P. and Hemminki, K. (1988) Depurination and imidazole ring-opening in nucleosides and DNA alkylated by styrene oxide. *Chem. Biol. Interact. 68*, 117-126.

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

(15) Laval, J., Pierre, J. and Laval, F. (1981) Release of 7-methylguanine residues from alkylated DNA by extracts of *Micrococcus luteus* and *Escherichia Coli*. *Proc. Natl. Acad. Sci. USA 78*, 852-855.

(16) Chetsanga, C. J. and Lindahl, T. r. (1979) Release of 7-methylguanine residues whose imidazole rings have been opened from damaged DNA by a DNA glycosylase from *Escherichia coli. Nucleic Acids Res. 1979*, 3676-3684.

(17) Beranek, D. T., Weiss, C. C., Evans, F. E., Chetsanga, C. J. and Kadlubar, F. F. (1983) Identification of $N^5$-methyl-$N^5$-formyl-2,5,6-triamino-4-hydroxypyramidine as a major adduct in rat liver DNA after treatment with the carcinogens, N,N-dimethylnitrosoamine or 1,2-dimethylhydrazine. *Bioch. Biophys. Res. Comm. 110*, 625-631.

(18) Kadlubar, F. F., Beranek, D. T., Weiss, C. C., Evans, F. E., Cox, R. and Irving, C. C. (1984) Characterization of the purine ring-opened 7-methylguanine and its persistence in rat bladder epithelial DNA after treatment with the carcinogen N-methylhitrosourea. *Carcingenesis 5*, 587-592.

(19) Hendler, S., Furer, E. and Srinivasan, P. R. (1970) Synthesis and chemical properties of monomers and polymers containing 7-methylguanine and an investigation of their substrate or template properties for bacterial deoxyribonucleic acid or ribonucleic acid polymerases. *Biochemistry 9*, 4141-4153.

(20) Barbarella, G., Tugnoli, V. and Zambianchi, M. (1991) Imidazole ring-opening of 7-methylguanosine at physiological pH. *Nucleosides Nucleotides 10*, 1759-1769.

(21) Darzynkiewicz, E., Labadi, I., Haber, D., Burger, K. and Lonneberg, H. (1988) 7-Methylguanine nucleotides and their structural analogs; protolytic equlibria, complexing with magnesium(II) ion and kinetics for alkaline opening of the imidazole ring. *Acta Chem. Scand. B42*, 86-92.

(22) Greenberg, M. M., Hantosi, Z., Wiederholt, C. J. and Rithner, C. D. (2001) Studies on N4-(2-deoxy-D-pentofuranosyl)-4,6-diamino-5-formamidopyrimidine (Fapy:dA) and N6-(2-deoxy-D-pentofuranosyl)-6-diamino-5-formamido-4-hydroxypyrimidine (Fapy:dG). *Biochemistry 40*, 15856-15861.

(23) Raoul, S., Bardet, M. and Cadet, J. (1995) Gamma irradiation of 2'-deoxyadenosine in oxygen-free aqueous solutions: identification and conformational features of formamidopyrimidine nucleoside derivatives. *Chem. Res. Toxicol. 8*, 924-933.

(24) Tomasz, M., Lipman, R., Lee, M. S., Verdine, G. L. and Nakanishi, K. (1987) Reaction of acid-activated mitomycin C with calf thymus DNA and model guanines: Elucidation of the base-catalyzed degradation of N7-alkylguanine nucleosides. *Biochemistry 26*, 2010-2027.

(25) Berger, M. and Cadet, J. (1985) Isolation and characterization of the radiation-induced degradation products of 2'-deoxyguanisine in oxygen-free aqeouos solutions. *Z. Naturforsch. 40B*, 1519-1531.

(26) Haraguchi, K. and Greenberg, M. M. (2001) Synthesis of oligonucleotides containing Fapy.dG (N6-(2-deoxy-alpha,beta-D-*erythro*-pentofuranosyl)-2,6-diamino-4-hydroxy-5-formamidopyrimidine). *J. Am. Chem. Soc. 123*, 8636-8637.

(27) Bergdorf, L. T. and Carrel, T. (2002) Synthesis, stability and conformation of the formamidopyrimidine G DNA lesion. *Chem. Eur. J. 8*, 293-301.

(28) Chetsanga, C. J. and Makaroff, C. (1982) Alkaline opening of imidazole ring of 7-methylguanosine. 2. Further studies on reaction mechanisms and products. *Chem. Biol. Interact. 41*, 235-249.

(29) Box, H. C., Ligam, K. T., French, J. B., Potienko, G. and Alderfer, J. L. (1981) $^{13}C$ NMR characterization of alkyalted derivatives of guanosine, adenosine, and cytidine. *Carbohydr., Nucleosides, Nucleotides 8*, 189-195.

(30) Boiteux, S., Belleney, J., Roques, B. P. and Laval, J. (1984) Two rotameric forms of open ring 7-methylguanine are present in alkylated polynucleotides. *Nucleic Acids Res. 12*, 5429-5439.

(31) Chetsanga, C. J., Bearie, B. and Makaroff, C. (1982) Alkaline opening of imidazole ring of 7-methylguanosine. 1. Analysis of the resulting pyrimidine derivatives. *Chem. Biol. Interact. 41*, 217-233.

(32) Hemminki, K., Peltonen, K. and Vodicka, P. (1989) Depurination from DNA of 7-methylguanine, 7-(2-amonoethyl)-guanine and ring-opened 7-methylguanines. *Chem. Biol. Interact. 70*, 289-303.

(33) Lin, J.-K., Miller, J. A. and Miller, E. C. (1977) 2,3-Dihydro-2-(guan-7-yl)-3-hydroxy-aflatoxin $B_1$ a major acid hydrolysis product of aflatoxin $B_1$-DNA or -ribosomal RNA adducts formed in hepatic microsome-mediated reactions and in rat liver microsomes. *Cancer Res. 37*, 4430-4438.

(34) Chetsanga, C. J. and Mavunga, I. (1986) Chemical reclosure of opened imidazole ring of guanine. *Chem. Biol. Interact. 58*, 117-123.

(35) Kohn, K. W. and Spears, C. L. (1967) Stabilization of nitrogen-mustard alkylations and inter-strand crosslinks in DNA by alkali. *Biochim. Biophys. Acta 145*, 734-741.

(36) Barak, R., Vincze, A., Bel, P., Dutta, S. P. and Chedda, G. B. (1993) Mass spectrometric investigation of the presence of 7-methyl ring-opened guanine derivatives in urine. *Chem. Biol. Interact. 86*, 29-40.

(37) Oleykowski, C. A., Mayernik, J. A., Lim, S. E., Groopman, J. D., Grossman, L., Wogan, G. N. and Yeung, A. T. (1993) Repair of

aflatoxin $B_1$ DNA adducts by the UvrABC endonuclease of *Escherichia coli*. *J. Biol. Chem. 268*, 7990-8002.

(38) Smela, M. E., Hamm, M. L., Henderson, P. T., Harris, C. M., Harris, T. M. and Essigmann, J. M. (2002) The aflatoxin $B_1$ formamidopyrimidine adduct plays a major role in causing the types of mutations observed in human hepatocellular carcinoma. *Proc. Nat. Acad. Sci. USA 99*, 6655-6660.

(39) Bailey, E. A., Iyer, R. S., Stone, M. P., Harris, T. M. and Essigmann, J. M. (1996) Mutational properties of the primary aflatoxin $B_1$-DNA adduct. *Proc. Natl. Acad. Sci. USA 93*, 1535-1539.

(40) Boiteux, S. and Laval, J. (1983) Imidazole open ring 7-methylguanine: an inhibitor of DNA synthesis. *Bioch. Biophys. Res. Comm. 110*, 552-558.

(41) Tchou, J., Kasai, H., Shibutani, S., Chung, M. H., Laval, J., Grollman, A. P. and Nishimura, S. (1991) 8-Oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proc. Natl. Acad. Sci. USA 88*, 4690-4694.

(42) Tudek, B., Grąziewicz, M., Kazanova, O., Zastawny, T. H., Obtułowicz, T. and Laval, J. (1999) Mutagenic specificity of imidazole ring-opened 7-methylpurines in m13mp18 phage DNA. *Acta Biochim. Pol. 46*, 785-800.

(43) Ezza-Nikpay, K. and Verdine, G. V. (1992) Aberrantly methylated DNA: site-specific introduction of N7-methyl-2'deoxyguanosine into the Dickerson/Drew dodecamer. *J. Am. Chem. Soc. 114*, 6562-6563.

(44) Asagoshi, K., Yamada, T., Okada, Y., Terato, H., Ohyama, Y., Seki, S. and Ide, H. (2000) Recognition of formamidopyrimidine by *Escherichia coli* and mammalian thymine glycol glycosylases. Distinctive paired base effects and biological and mechanistic implications. *J. Biol. Chem. 275*, 24781-24786.

(45) Asagoshi, K., Yamada, T., Terato, H., Ohyama, Y., Monden, Y., Arai, T., Nishimura, S., Aburatani, H., Lindahl, T. and Ide, H. (2000) Distinct repair activities of human 7,8-dihydro-8-oxoguanine DNA glycosylase and formamidopyrimidine DNA glycosylase for formamidopyrimidine and 7,8-dihydro-8-oxoguanine. *J. Biol. Chem. 275*, 4956-4964.

(46) Asagoshi, K., Terato, H., Ohyama, Y. and Ide, H. (2002) Effects of a guanine-derived formamidopyrimidine lesion on DNA replication. *J. Biol. Chem. 277*, 14589–14597.

(47) Haraguchi, K., Delaney, M. O., Wiederholt, C. J., Sambandam, A., Hantosi, Z. and Greenberg, M. M. (2002) Synthesis and characterization of oligodeoxynucleotides containing formamidopyrimidine lesions and nonhydrolyzable analogues. *J. Am. Chem. Soc. 124*, 3263-3269.

(48) Jiang, Y. L., Wiederholt, C. J., Patro, J. N., Haraguchi, K. and Greenberg, M. M. (2005) Synthesis of oligonucleotides containing Fapy·dG ($N^5$-(2-deoxy-,-D-*erythro*pentofuranosyl)-2,6-diamino-4-hydroxy-5-formamidopyrimidine) using a 5'-dimethoxytrityl dinucleotide phosphoramidite. *J. Org. Chem. 70*, 141-149.

(49) Tudek, B., Boiteux, S. and Laval, J. (1992) Biological properties of imidazole ring-opened N7-methylguanine in M13mp18 phage DNA. *Nucleic Acids Res. 20*, 3079-3084.

(50) O'Connor, T. R., Boiteux, S. and Laval, J. (1988) Ring-opened 7-methylguanine residues in DNA are a block to *in vitro* DNA synthesis. *Nucleic Acids Res. 16*, 5879-5894.

(51) McBride, L. J., Kierzek, R., Beaucage, S. L. and Caruthers, M. H. (1986) Nucleotide chemistry. 16. Amidine protecting groups for oligonucleotide synthesis. *J. Am. Chem. Soc. 108*, 2040-2048.

(52) Elmquist, C. E., Stover, J. S., Wang, Z. and Rizzo, C. J. (2004) Site-specific synthesis and properties of oligonucleotides containing C8-deoxyguanosine adducts of the dietery mutagen IQ. *J. Am. Chem. Soc. 126*, 11189-11201.

(53) Brown, K. B., Deng, Z. J., Iyer, S. R., Iyer, G. L., Voehler, M. W., Stone, M. P., Harris, C. M. and Harris, T. M. (2006) Unraveling the aflatoxin-FAPy conundrum: structural basis for differential replicative

processing of isomeric forms of the foramidopyrimidine-type DNA adduct of aflatoxin $B_1$. *J. Am. Chem. Soc. 128*, 15188-15199.

(54) Birk, C., Voss, J. and Wirsching, J. (1997) Preparation, structural elucidation and reactions of benzyl 2'-deoxy-3,5-di-O-methyl-1,4-dithio-L-*treo*-pentofuranoside and synthesis of the corresponding 2'-deoxy-4'-thionucleosides. *Carbohydr. Res. 304*, 239-247.

(55) Harris, C. M. and Harris, T. M. unpublished data.

(56) Mastihubova, M. and Biely, P. (2004) Deoxy and deoxyfluoro analogues of acetylated methyl β-D-xylopyranoside substrates for acetylxylan esterases. *Carbohydr. Res. 339*, 2101-2110.

(57) Napolitano, R., Janel-Bintz, R., Wagner, J. and Fuchs, P. P. (2000) All three SOS-inducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. *EMBO J. 19*, 6259-6265.

(58) Goodman, M. F., Woodgate, R. and Rangarajan, S. (1999) A phenotype for enigmatic DNA polymerase II in replication restart in UV-radiated *E. coli*. *Proc. Natl. Acad. Sci.* USA, *96*, 9224-9229.

(59) Yang, W. (2003) Damage repair DNA polymerases Y. *Curr. Opin. Struct. Biol. 13*, 23-30.

(60) Boudsocq, F., Iwai, S., Hanaoka, F. and Woodgate, R. (2001) *Sulfolobus solfataricus* P2 DNA polymerase IV (Dpo4): an archaeal

DinB-like DNA polymerase with lesion-bypass properties akin to eukaryotic pol η. *Nucleic Acids Res. 29*, 4607-4616.

(61) Maga, G., Villani, G., Tillement, V., Stucki, M., Locateli, G. A., Frouin, I., Spadari, S. and Hubscher, U. (2001) Okazaki fragment processing: modulation of the strand displacement activity of DNA polymerase δ by the concerted action of replication protein A, proliferating cell nuclear antigen, and flap endonuclease-1. *Proc. Natl. Acad. Sci. USA 98*, 14298-14303.

(62) Washington, M. T., Johson, R. E., Prakash, S. and L., P. (2000) Accuracy of thymine-thymine dimer bypass by *Saccharomyces cerevisiae* DNA polymerase η. *Proc. Natl. Acad. Sci. USA 97*, 3094-3099.

(63) Johnson, R. E., Kondratick, C. M., Prakash, S. and Prakash, L. (1999) hRAD30 mutations in the variant form of *xeroderma pigmentosum*. *Science 285*, 263-265.

(64) Masutani, C., Kusumoto, R., Yamada, A., Dohmoe, N. and Yokoi, M. (1999) The XPV (*xeroderma pigmentosum* variant) gene encodes human DNA polymerase η. *Nature 399*, 700-704.

(65) Haracska, L., Yu, S. L., Johnson, R. E., Prakash, L. and Prakash, S. (2000) Efficient and accurate replication in the presence of 7,8-dihydro-8-oxoguanine by DNA polymerase η. *Nat. Genet. 25*, 458-461.

(66) Kalam, M. A., Haraguchi, K., Chandani, S., Loechler, E. L., Moriya, M., Greenberg, M. M. and Basu, A. K. (2006) Genetic effects of oxidative DNA damages: comparative mutagenesis of the imidazole ring-opened formamidopyrimidines (Fapy lesions) and 8-oxo-purines in simian kidney cells. *Nucleic Acids Res. 34*, 2305-2315.

(67) Johnson, R. E., Prakash, S. and Prakash, L. (2000) The human DINB1 gene encodes the DNA polymerase Pol η. *Proc. Natl. Acad. Sci. USA 97*, 3838-3843.

(68) Haracska, L., Prakash, L. and Prakash, S. (2002) Role of human DNA polymerase κ as an extender in translesion synthesis. *Proc. Nat. Acad. Sci. USA 99*, 16000-16005.

(69) Washington, M. T., Minko, I. G., Johnson, R. E., Wolfle, W. T., Harris, T. M., Lloyd, R. S., Prakash, S. and Prakash, L. (2004) Efficient and error-free replication past a minor-groove DNA adduct by the sequential action of human DNA polymerases ι and κ. *Mol. Cell Biol. 24*, 5687-5693.

(70) Johnson, R. E., Washington, M. T., Haracska, L., Prakash, S. and Prakash, L. (2000) Eukaryotic polymerases ι and ζ act sequentially to bypass DNA lesions. *Nature 406*, 1015-1019.

(71) Haracska, L., Johnson, R. E., Unk, I., Phillips, B. B. and Hurwitz, J. (2001) Targeting of human DNA polymerase *ι* to the replication

machinery via interaction with PCNA. *Proc. Natl. Acad. Sci. USA 98*, 14256-14261.

(72) Washington, M. T., Johnson, R. E., Prakash, L. and Prakash, S. (2004) Human DNA polymerase ι utilizes different nucleotide incorporation mechanisms dependent upon the template base. *Mol. Cell. Biol. 24*, 936–943.

(73) Stover, J. S., Chowdhury, G., Zang, H., Guengerich, F. P. and Rizzo, C. J. (2006) Translesion synthesis past the C8- and $N^2$-deoxyguanosine adducts of the dietary mutagen 2-amino-3-methylimidazo[4,5-f]quinoline in the NarI recognition sequence by prokaryotic DNA polymerases. *Chem. Res. Toxicol. 19*, 1506-1517.

(74) Ling, H., Boudsocq, F., Woodgate, R. and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication. *Cell 107*, 91-102.

(75) Oka, N. and Greenberg, M. M. (2005) The effect of the 2-amino group of 7,8-dihydro-8-oxo-2'-deoxyguanosine on translesion synthesis and duplex stability. *Nucleic Acids Res. 33*, 1637-1643.

(76) Wiederholt, C. J. and Greenberg, M. M. (2002) Fapy.dG instructs Klenow exo(-) to misincorporate deoxyadenosine. *J. Am. Chem. Soc. 124*, 7278-7279.

(77) Choi, J.-Y., Stover, J. S., Angel, K. C., Chowdhury, G., Rizzo, C. J. and Guengerich, F. P. (2006) Biochemical basis of genotoxicity of heterocyclic arylamine food mutagens: Human DNA polymerase η selectively produces a two-base deletion in copying the $N^2$-guanyl adduct of 2-amino-3-methylimidazo[4,5-f]quinoline but not the C8-adduct at the NarI G3 site. *J. Biol. Chem. 281*, 25297-25306.

(78) Derome, A. E. and Williamson, M. P. (1990) Rapid-pulsing artifacts in double-quantum-filtered COSY. *J. Magn. Reson. 88*, 177-185.

(79) Jeener, J., Meier, B. H., Bachmann, P. and Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys. 71*, 4546-4553.

(80) Wagner, R. and Berger, S. (1996) Gradient-selected NOESY- a fourfold reduction of the measurement time for the NOESY experiment. *J. Magn. Reson. 123*, 119-121.

(81) Bax, A. and Davis, D. G. (1985) MLEV-17-based two-dimensional homonuclear magnetization transfer spectroscopy. *J. Magn. Reson. 65*, 355-360.

(82) Willker, W., Leibfritz, D., Kerssebaum, R. and Bermel, W. (1993) Gradient selection in inverse heteronuclear correlation spectroscopy. *Magn. Reson. Chem. 31*, 287-292.

(83) Wang, H., Zuiderweg, E. R. R. and Glick, G. D. (1995) Solution structure of a disulfide cross-linked DNA hairpin. *J. Am. Chem. Soc. 117*, 2981-2991.

(84) Goddard, T. D. and Kneller, D. G. SPARKY 3.11, University of California, San Francisco.

(85) Zang, H., Goodenough, A. K., Choi, J.-Y., Irimia, A., Loukachevitch, L. V., Kozekov, I. D., Angel, K. C., Rizzo, C. J., Egli, M. and Guengerich, F. P. (2005) DNA adduct bypass polymerization by *Sulfolobus solfataricus* DNA polymerase Dpo4. Analysis and crystal structures of multiple base-pair substitution and frameshift products with the adduct 1,$N^2$-ethenoguanine. *J. Biol. Chem. 280*, 29750-29764.

(86) Lowe, L. G. and Guengerich, F. P. (1996) Steady-state and pre-steadystate kinetic analysis of dNTP insertion opposite 8-oxo-7, 8-dihydroguanine by *Escherichia coli* polymerases I exo- and II exo-. *Biochemistry 35*, 9840-9849.

CHAPTER III

# FORMATION OF (2-OXOETHYL)-FAPy-dGuo DNA ADDUCT

## Introduction

### *Reactivity of Vinyl Chloride (3.01) Towards DNA*

Vinyl chloride (**3.01**) is an industrial chemical that is used widely for the preparation of polyvinyl chloride and copolymers. It is a hepatocarcinogen epidemiologically linked to hepatic angiosarcomas in humans and experimentally linked to the same tumors in rodents (*1-4*). Vinyl chloride (**3.01**) is epoxidized by cytochrome P450 2E1 to 2-chlorooxirane (**3.02**), the ultimate carcinogenic species, which rapidly rearranges to 2-chloroacetaldehyde (**3.03**) (Figure **3-01**) (*5*).

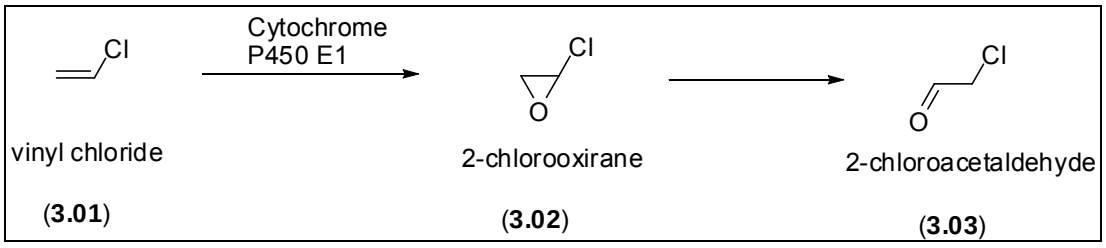



**Figure 3-01**. Epoxidation of vinyl chloride (**3.01**) to 2-chlorooxirane (**3.02**) by cytochrome P450 E1.

2-Chlorooxirane (**3.02**) reacts extensively with DNA whereas 2-chloroacetaldehyde (**3.03**) reacts mainly with proteins. 2-Chlorooxirane (**3.02**) alkylates DNA at N1, $N^2$, N3, and N7 positions of dGuo moieties (**3.04**) in DNA to form the corresponding 2-oxoethyl adducts **3.05** - **3.08**. The oxoethyl adducts on N1 and N3 adducts cyclize to give the 1,$N^2$- and $N^2$,3-etheno adducts **3.09** and **3.11**, respectively; the $N^2$ adduct cyclizes to form the 8-hydroxy-5,6,7,8-tetrahydropyrimido[1,2-*a*]purin-10(3*H*)-one (**3.10**); the N7 oxoethyl adduct **3.08** undergoes depurination resulting in N7-(2-oxoethyl)-Gua (**3.12**) (Figure **3-02**).

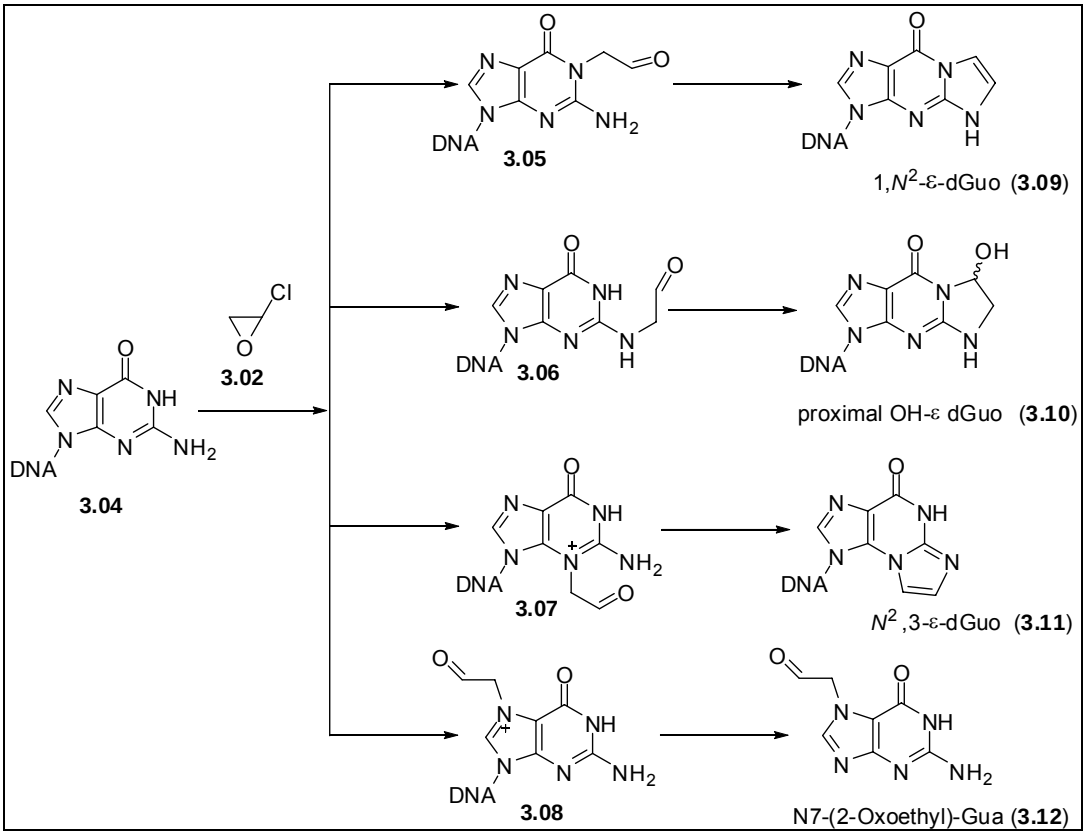



**Figure 3-02**. 2-Oxoethyl alkylation products at the N1, $N^2$, N3 and N7 positions of the dGuo moieties (**3.04**) in DNA.

Alkylation reaction also occurs at N3 position of dCyd (**3.13**) and at N1 position of dAdo (**3.14**) to give 2-oxoethyl derivatives **3.15** and **3.16**; these adducts also cyclize to form 3,$N^4$-ε-dCyd (**3.17**) and 1,$N^6$-ε-dAdo (**3.18**) adducts (Figure **3-03**).

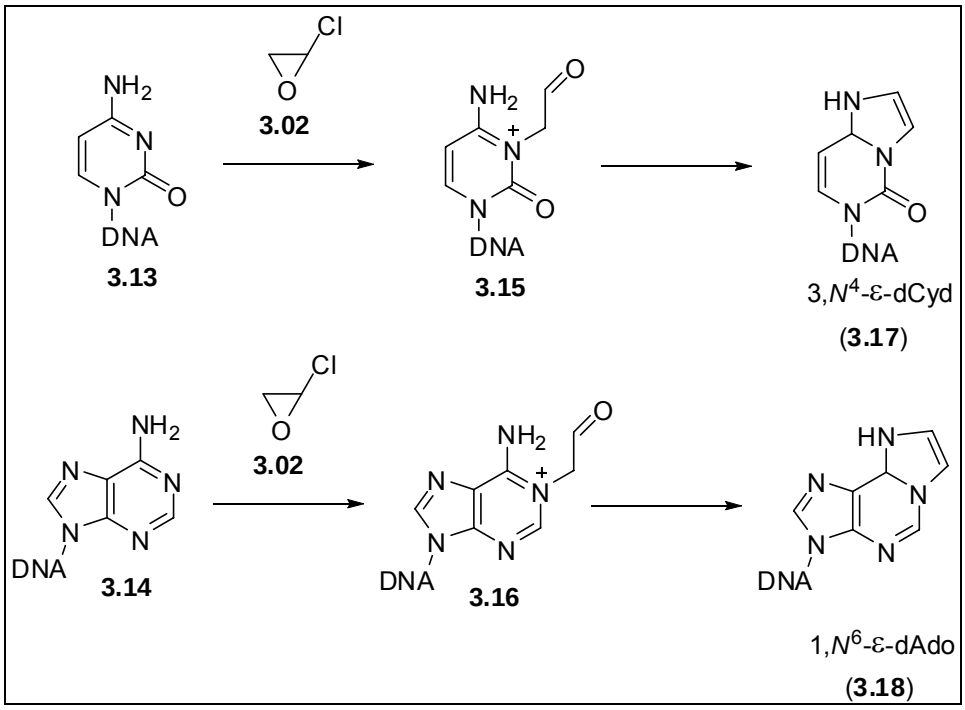



**Figure 3-03**. 2-Oxoethyl alkylation products at the N3 positions of the dCyt (**3.13**) and at N1 position of dAdo (**3.14**) moieties in DNA.

### *Mutagenicity of the Vinyl Chloride (3.01) Induced DNA Adducts*

In the reaction of 2-chlorooxirane (**3.02**) with DNA, N7-(2-oxoethyl)-Gua adduct (**3.12**) predominates, representing ~99% of the product mixture. Nevertheless, the etheno adducts (**3.09** - **3.11**, **3.17** and **3.18**) are generally

considered to be biologically more important than the N7-(2-oxoethyl)-Gua adduct (**3.12**), because the latter has been reported to be non-mutagenic (*6*), whereas the etheno species are miscoding *in vitro* (*7-10*) and both mutagenic and highly persistent *in vivo* (*11*). Although the carcinogenicity of vinyl chloride (**3.01**) is usually ascribed to the etheno adducts, a confounding factor is that the etheno adducts are present in the cells of laboratory animals and humans that have not been exposed to vinyl chloride (**3.01**) or other vinyl monomers that might be capable of forming etheno adducts (*12*). The endogenous etheno derivatives are believed to arise from oxidative degradation products of unsaturated lipids, in particular from the epoxides of $\alpha,\beta$-unsaturated aldehydes (*13*). In support of this hypothesis, the levels of etheno derivatives rise significantly in liver samples in which oxidative stress is present, for example patients with alcoholic fatty liver, Wilson's disease, and hemochromatosis. Thus, a central question is why relatively low levels of exposure to vinyl chloride (**1.01**) would generate sufficiently high concentrations of etheno adducts to create a substantial risk of inducing malignancies when background levels of these adducts are already present in normal hepatic cells. One needs to examine the possibility that some other, as yet unexamined, 2-chlorooxirane adduct is the primary cause of the angiosarcomas resulting from vinyl chloride (**3.01**) exposure, with the main qualification being that the adduct could not also be formed by reactions of epoxidized enals and thus not present in unexposed cells.

### *Formation of (2-Oxoethyl)-FAPy-dGuo Adduct (3.19)*

We have hypothesized that formamidopyrimidine **3.19** ((2-oxoethyl)-FAPy-dGuo) may be formed via basic hydrolysis of the imidazole ring of the N7 adduct **3.08** and could contribute to the mutagenicity of vinyl chloride (**3.01**) (Figure **3-04**). Detection of the (2-oxoethyl)-FAPy-dGuo lesion (**3.19**) in biological samples has never been reported. Neither has the synthesis of the nucleoside nor the base been reported. However, FAPy-dGuo lesions involving other N7-alkyllated guanine derivatives, discussed in Chapter I and II, have been characterized in biological samples and prepared by chemical routes.

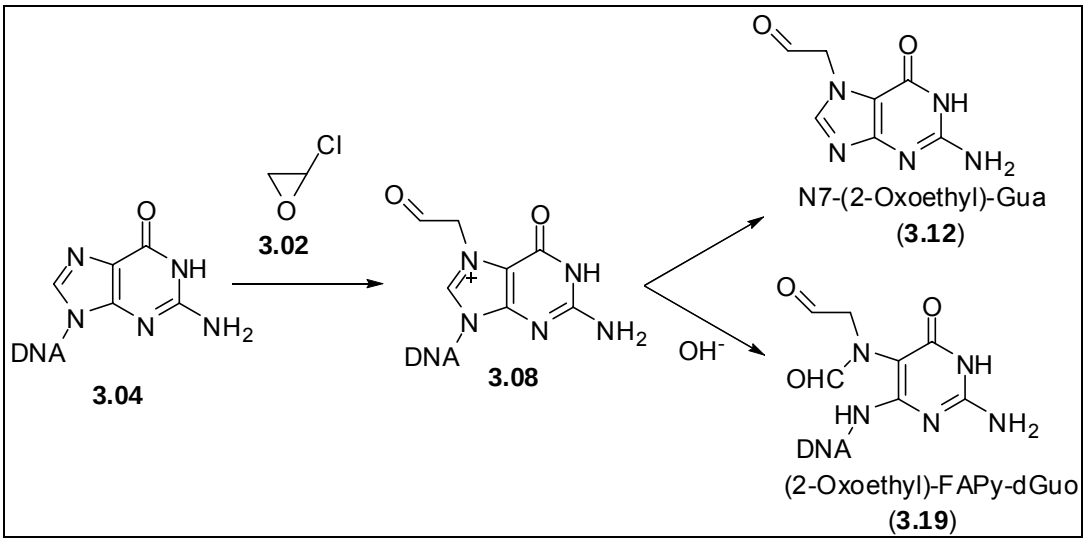



**Figure 3-04**. Possible formation of (2-oxoethyl)-FAPy-dGuo adduct (**3.19**) via basic hydrolysis of the imidazole ring of the N7-(2-oxoethyl)-dGuo cationic species **3.08**.

The failure of investigators to observe the (2-oxoethyl)-FAPy-dGuo lesion (**3.19**) might reflect the complex chemistry of FAPy-dGuo adducts. As discussed in Chapter II, an opening and reclosure of the deoxyribose ring of a FAPy-dGuo nucleoside can generate a mixture of α- and β-furanose and pyranose nucleosides. $N^5$-Alkylated FAPy-dGuo adducts have also a hindered rotation around the C5-$N^5$ bond leading to atropisomers. Slow reorientation of the planar formamide creates geometrical isomers. The various isomeric forms are frequently separable but still able to interconvert leading to the nucleoside having poor chromatographic behavior. Even at the FAPy-Gua base level atropisomers and geometrical isomers can exist. The N7 adducts of dGuo that have been investigated so far have little or no mutagenic activity, but several of the corresponding FAPy-dGuo lesions have been found to be strongly promutagenic (Figure **3-05**). The unsubstituted FAPy-dGuo species (**3.20**) formed by oxidation of guanine is mutagenic in mammalian systems (*14*). The dominant form of the $AFB_1$-FAPy-dGuo lesion (**3.21**) is potently mutagenic in *E. coli* and more mutagenic than the corresponding N7-cationic-dGuo adduct *(15)*.

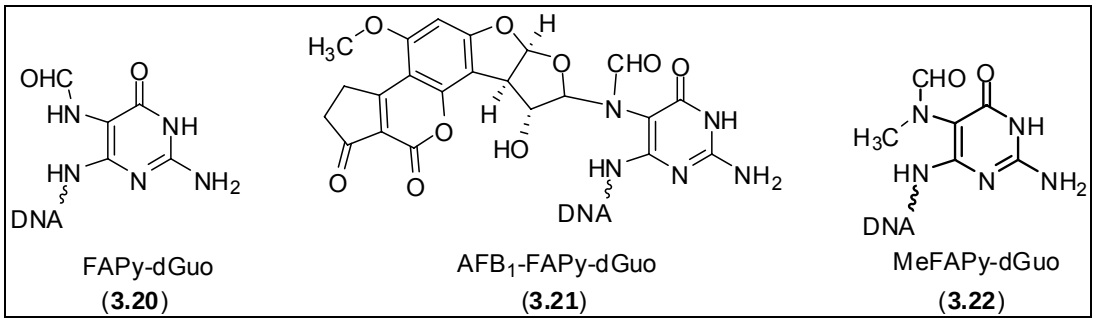



**Figure 3-05**. Structures of the FAPy-dGuo (**3.20**), $AFB_1$-FAPy-dGuo (**3.21**) and MeFAPy-dGuo (**3.22**) lesions.

However, in contrast to these results, the MeFAPy-dGuo lesion (**3.22**) derived from the N7-methyl derivative of dGuo has been reported to be a strong block to replication by the Klenow fragment exo$^-$ but bypass, to the extent that it occurs, is with high fidelity (*16*). Our results, described in Chapter II, with prokaryotic DNA polymerases also showed that MeFAPy-dGuo lesion (**3.22**) is highly mis-coding, but this does not necessary reflect its mutagenic potential since the initial incorporation of dCTP opposite the adduct is preferentially extended. In contrast the *in vitro* bypass and extension past the MeFAPy-dGuo lesion (**3.22**), described also in Chapter II, with eukaryotic DNA polymerases showed that the MeFAPy-dGuo lesion (**3.22**) is bypassed and extended in both error-free and error-prone manner. The question still remains of whether the MeFAPy-dGuo lesion (**3.22**) is mutagenic in living organisms.

## Results and Discussion

### *Preliminary Results from the Reaction of 2-Acetoxyoxirane (3.27) and dGuo (3.30)*

2-Chlorooxirane (**3.02**) is usually prepared by photochemical chlorination of ethylene oxide (**3.23**) with t-butyl hypochlorite (**3.24**) (*17, 18*) (Figure **3-06**). It is difficult to obtain 2-chlorooxirane (**3.02**) in pure form, because trace amounts of HCl, which may be present as a contaminant, can catalyze rearrangement to 2-chloroacetaldehyde (**3.03**). For this reason, we chose to use 2-acetoxyoxirane (**3.27**) as a surrogate, although it likely to be less reactive with dGuo (**3.30**) and is also prone to hydrolysis. However, the acetic acid, which may contaminate the 2-acetoxyoxirane (**3.27**) is a less potent catalyst of rearrangement than HCl. Substantial excesses of the epoxide (**3.27**) would be required to obtain reasonable yields of the N7 adduct.

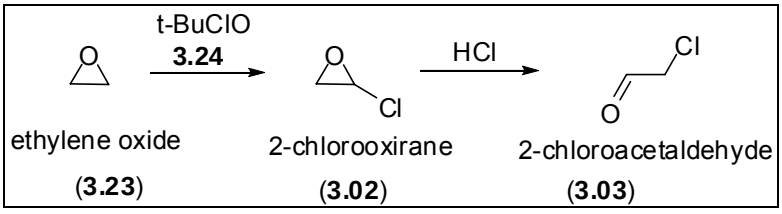



**Figure 3-06**. Synthesis of 2-chlorooxirane (**3.02**) from ethylene oxide (**3.23**).

We found that 2-acetoxyoxirane (**3.27**) could be conveniently prepared by epoxidation of vinyl acetate (**3.25**) with dimethyl dioxirane (**3.26**) (Figure **3-07**).

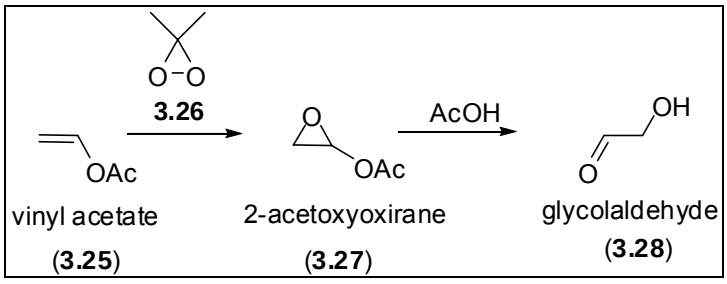



**Figure 3-07**. Synthesis of 2-acetoxyoxirane (**3.27**) from vinyl acetate (**3.25**).

Several solvent mixtures were investigated for the reaction of dGuo (**3.30**) with 2-acetoxyoxirane (**3.27**). Dimethyl sulfoxide (DMSO) was examined first because an aprotic polar solvent favors the N7-alkylation of dGuo by stabilizing the transition state in the formation of the cationic adduct. HPLC analysis of the DMSO reaction mixture showed the major product was N7-(2-oxoethyl)-Gua (**3.12**). In addition, glyoxal-dGuo (**3.31**), 1,$N^2$-ε-dGuo (**3.09**) and the $N^2$,3-ε-Gua (**3.11a**) were observed (Figure **3-08** and **3-09**); their identity was confirmed by co-injection with authentic standards. We were unable to detect the (2-oxoethyl)-FAPy-dGuo species (**3.19**). The cationic adduct (**3.08**) may have deglycosylated due to the presence of acetic acid arising from hydrolysis of the 2-acetoxyoxirane (**3.27**) by adventitious moisture. The (2-oxoethyl)-FAPy-dGuo (**3.19**) was not detected, which is attributed to the absence of hydroxide; an equivalent of

hydroxide is required to achieve the imidazole ring-opening reaction.

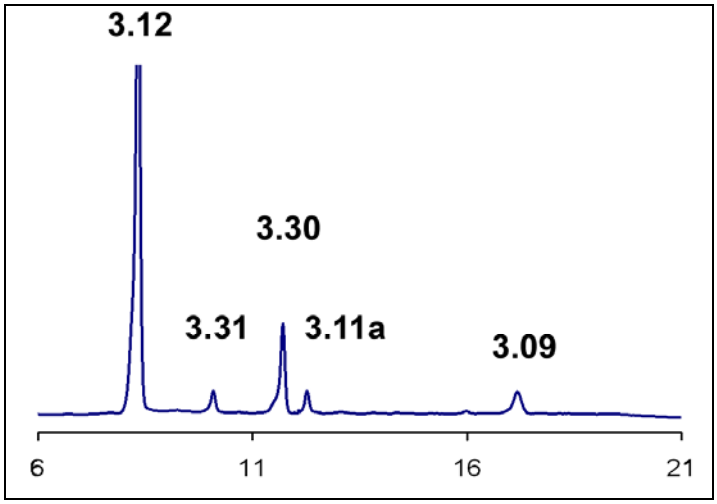



**Figure 3-08**. HPLC traces of the reaction mixture of 2-acetoxyoxirane (**3.27**) with dGuo (**3.30**)

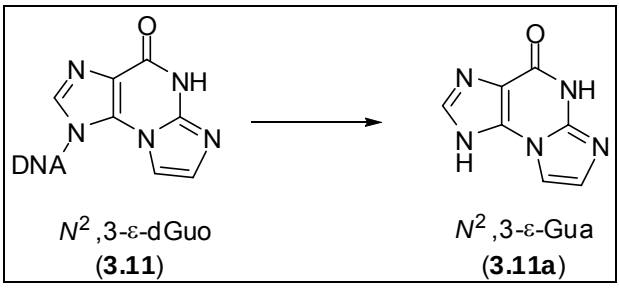



**Figure 3-09**. Formation of the $N^2$,3-ε-Gua (**3.11a**).

Glyoxal (**3.29**) was a by-product of the adduction reaction, leading to formation of the glyoxal adduct **3.31** of dGuo (Figure **3-10**). Glyoxal (**3.29**) arises by oxidation of 2-acetoxyoxiarane (**3.27**) and/or glycolaldehyde (**3.28)**, which is

the decomposition product of the 2-acetoxyoxirane (**3.27**). We have proposed that the latter oxidation is mediated by DMSO (*19, 20*). Both reactions require molecular oxygen. In subsequent reactions, the formation of glyoxal (**3.29**) and its dGuo adduct **3.31** was suppressed by carrying out the reactions under an inert atmosphere and using degassed solvents.

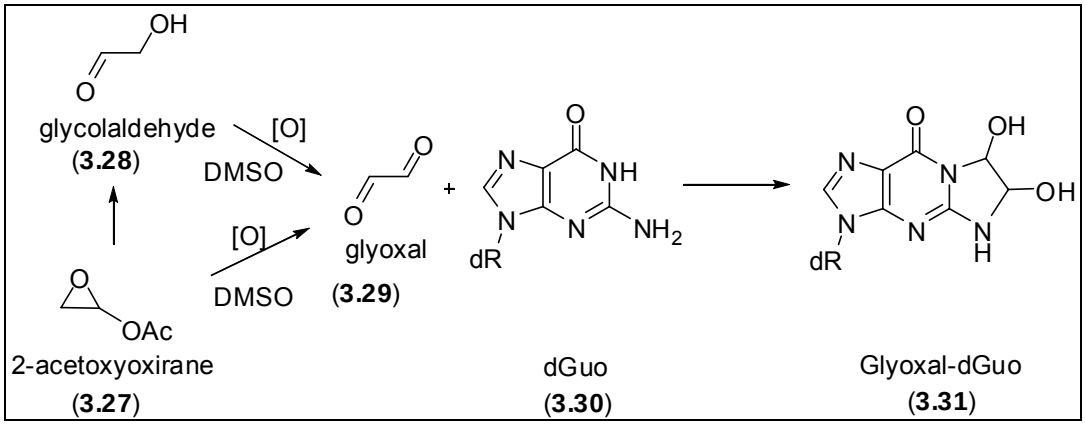



**Figure 3-10**. Oxidation of glycolaldehyde (**3.28**) and 2-acetoxyoxirane (**3.27**) to glyoxal (**3.29**); the later reacts with dGuo (**3.30**) to form glyoxal-dGuo (**3.31**) adduct.

In the next experiment, dGuo (**3.30**) and 2-acetoxyoxirane (**3.27**) were incubated in DMSO for 20 min, after which the DMSO was removed in vacuo, pH 7.0 phosphate buffer was added in an effort to convert the cationic adduct to the (2-oxoethyl)-FAPy-dGuo derivative (**3.19**), and further reaction was monitored by reverse phase HPLC. N7-(2-oxoethyl)-Gua (**3.12**) was identified as the major product. When the reaction of dGuo (**3.30**) and 2-acetoxyoxirane (**3.27**) was done in phosphate buffers at pH's ranging from 7.0 to 9.0, once again only N7-(2-

oxoethyl)-Gua (**3.12**) was detected.

We recognized the possibility that (2-oxoethyl)-FAPy-dGuo nucleoside (**3.19**) might have complex chromatographic behavior such that the product could be overlooked. Prior investigations of MeFAPy-dGuo nucleosides have shown that they exist as mixtures of isomeric forms. In the nucleosides, the carbohydrate moiety undergoes ring-opening and reclosure leading to mixtures of $\alpha$ and $\beta$-anomers of furanosides and pyranosides. Furthermore, hindered rotation around the C5-$N^5$ bond can lead to slowly equilibrating atropisomers. Slow reorientation of the formamide linkage can create chromatographically separable geometrical isomers. This potential mixture of isomers can be simplified somewhat by analyzing the FAPy-Gua rather than the nucleoside although the potential remains for the product to be a mixture of geometrical isomers. Thus, the focus of the study shifted to the use of 0.5 M NaOH as the base with the work-up procedure for subsequent reactions involving deglycosylation under acidic conditions (2 h at pH 4-5 and room temperature). LC-ESI/MS/MS analysis of the acid hydrolysates gave a chromatographic peak having a molecular ion at *m/z* 212, which is consistent with $[M + H]^+$ of the (2-oxoethyl)-FAPy-Gua (**3.35**). The N7-(2-oxoethyl)-Gua (**3.12**) as a hydrated form also gives a molecular ion at *m/z* 212. An independent synthesis of the N7-(2-oxoethyl)-Gua (**3.12**) was performed (*21, 22*) and the species was analyzed by LC-ESI/MS/MS. The MS/MS analysis of the N7-(2-oxoethyl)-Gua (**3.12**) showed a different pattern of fragmentation than the observed fragmentation of the tentatively assigned (2-oxoethyl)-FAPy-Gua (**3.35**).

### *Synthesis and Spectral Properties of (2-Oxoethyl)-FAPy-Gua (3.35)*

To establish the structure of the base, tentatively assigned as the (2-oxoethyl)-FAPy-Gua (**3.35**), an independent synthesis was developed as shown in Figure **3-11** to provide an authentic sample for the MS analysis of ring opening of N7 adduct **3.19**. The synthetic approach that was taken involved preparation of $N^5$-allyl-FAPy-dGuo (**3.32**), which promised to be relatively straightforward, followed by transformation of it to (2-oxoethyl)-FAPy-Gua (**3.35)**. N7-Alkylation of dGuo (**3.30**) with allyl bromide was achieved in good yield. The cationic adduct was treated with sodium hydroxide to give $N^5$-allyl-FAPy-dGuo (**3.32**) in an overall 41% yield after purification. Purification was complicated by the fact that **3.32** was a complex product mixture for the reasons discussed above. To simplify the mixture, acid-catalyzed deglycosylation was performed to give the $N^5$-allyl-FAPy-Gua (**3.33**) in high yield. Dihydroxylation of **3.33** using $OsO_4$ gave diol **3.34** in 85% yield. Oxidative cleavage of the diol **3.34** with $NaIO_4$ at neutral pH was achieved in 79% yield to give the desired product.

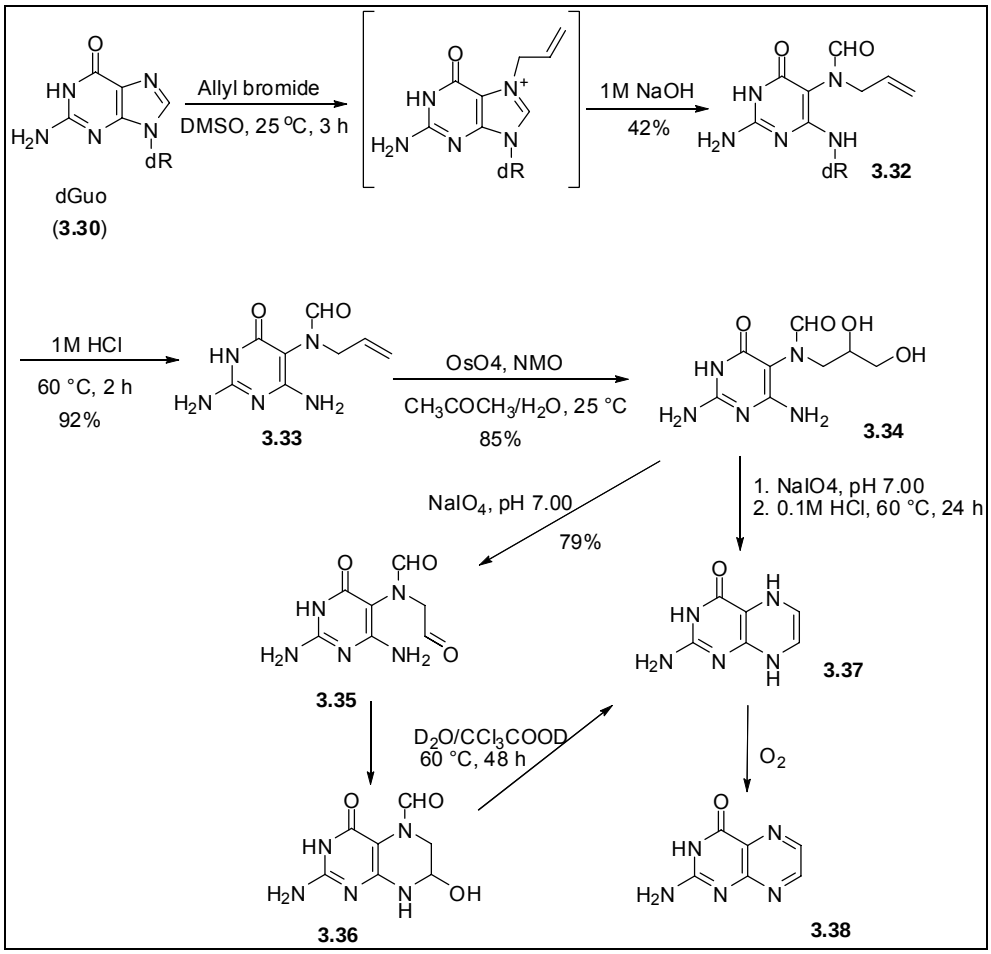



**Figure 3-11**. Synthesis of the cyclized (2-oxoethyl)-FAPy-Gua (**3.36**).

HPLC analysis immediately after periodate oxidation gave a broad peak centered at about 5 min which was tentatively assigned as (2-oxoethyl)-FAPy-Gua (**3.35**) (Figure **3-12**, Panel A). When the sample was allowed to stand, the chromatographic behavior changed to give two sharp peaks of similar intensity, eluting at $t_R$ 2.89 min and 3.57 min, along with a broad peak between them (Figure **3-12**, Panel B); however, when normal phase chromatographic system was used for **3.35** the peak shape was satisfactory (Figure **3-12**, Panel C).

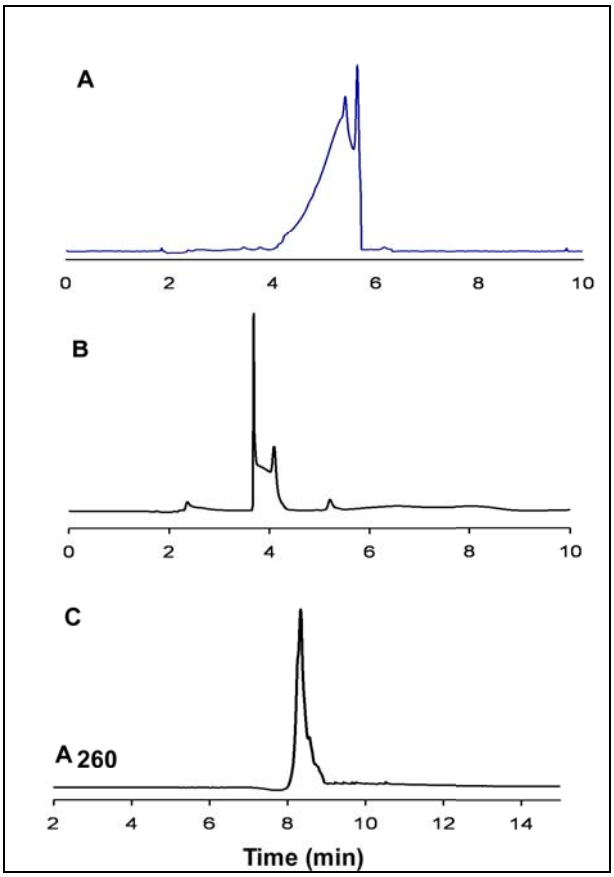



**Figure 3-12**. HPLC traces of compound(s) **3.35**. **A**) Reverse phase HPLC traces of compound **3.35** after periodate cleavage of **3.34**; **B**) Reverse phase HPLC traces of compound **3.35** after 1h in water; **C**) Normal phase HPLC traces of compound **3.35** after 1h in water.

This change occurred over a period of 1 h but was rapid if a trace of acid was added to the sample. The sharp peaks were collected individually at –78 °C and subsequently re-analyzed by HPLC. Each peak gave a similar chromatogram (Figure **3-12**, Panel B), which suggested that the two species were interconverting on the time scale of the separation. Subsequently, the two components were collected as a single fraction.

Solubility and kinetic constraints precluded obtaining NMR spectra on the initially formed product of the peroxide oxidation. NMR studies on the equilibrated products were carried out in $D_2O$ containing trichloroacetic acid which was required to obtain adequate solubility. The product was assigned as carbinolamine **3.36** arising by reaction of the aldehyde of **3.35** and the C4-amino group of the pyrimidione on the basis of $^{1}H$, $^{13}C$, COSY, NOESY, and HSQC spectra described below (the NMR spectra are available in the Appendix). The two interconverting chromatographic components which are designated **3.36a** and **3.36b** are assigned as involving slow reorientation of the formyl moiety of the formamide (Figure **3-13**). This assignment is consistent with earlier assignments for other $N^5$-alkylated FAPy-Guo adducts. There was no evidence for uncyclized aldehyde species **3.35** in the $^{1}H$ spectrum, which would have given an aldehyde signal in the range 9.0 - 10.0 ppm. The formyl protons for **3.36a** and **3.36b** were observed at 8.56 and 7.89 ppm, respectively; in both cases the methine proton appeared as a broad triplet at 5.17 ppm (Figure **3-13**). The methylene signals of **3.36b** appeared at 3.11 and 3.76 ppm, whereas those of **3.36a** appeared at 2.73 and 4.42 ppm. It is noteworthy that the difference in chemical shifts for the methylene protons of **3.36b** is much smaller that the difference **3.36a**.

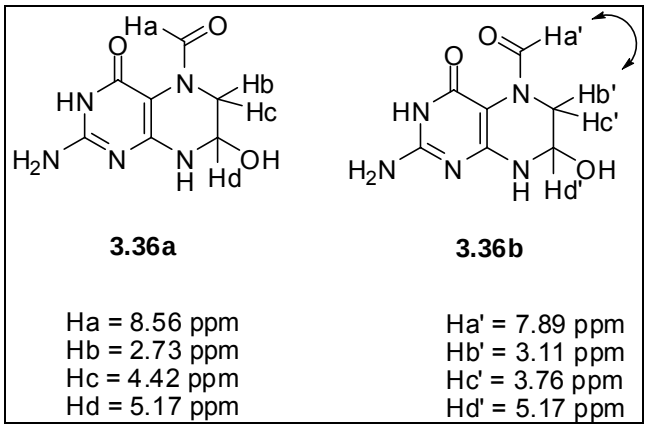



**Figure 3-13**. Structures of the two interconverting chromatographic components designated as the **3.36a** and **3.36b** of the carbinolamine **3.36**.

Strong geminal coupling was observed between the methylene protons (13.4 Hz in **3.36b** and 18 Hz in **3.36a**) and weak vicinal coupling (1.8 Hz) to the methine protons. The COSY spectrum showed correlations among the three signals for each streoisomer. The fact that the vicinal coupling constants are small and equal for both isomers indicates that the preferred conformation of each is a half chair having the hydroxyl group in the axial position leading to torsional angle between the methine proton and the each of the methylene protons of approximately 60°. Had the product actually been the hydrated form for of aldehyde **3.35,** the preferred conformation would place the formamide across from one of the hydroxyl groups and adjacent to the other. Thus, the torsional angle of the methine proton with one of the methylene protons would have been 180° and with other 60° leading to one vicinal coupling being large and the other small.

The NOESY spectrum of **3.36b** showed correlation of the formamide

proton with one of the methylene protons (3.11 ppm) allowing this isomer to be assigned as having Z conformation with the formyl proton proximal to the methylene group. No correlation was observed to the other methylene proton of **3.36a**. For **3.36a**, the half-chair conformation of the ring places the formyl group much close to the equatorial proton of the methylene group to the axial proton. Consequently, we can assign the signal at 3.11 ppm to the equatorial proton and the one at 3.76 ppm to the axial. No correlation was observed of the formyl proton of **3.36a** with either the methylene protons, which is consistent with an E configuration for the formamide. The large difference between the chemical shifts of the methylene protons of **3.36a** versus the much smaller difference between the shifts of **3.36b** is consistent with the carbonyl group of **3.36a** to shielding the equatorial proton of **3.36a**.

The $^{13}C$ chemical shifts were correlated with the proton spectrum via an HSQC spectrum. The spectra of **3.36a** and **3.36b** were very similar with the exception that the methylene carbon of **3.36b** was 5.68 ppm farther downfield than the methylene carbon of **3.36a** (42.32 versus 48.10 ppm). This is consistent with the methylene carbon of **3.36a** lying in the shielding cone of the carbonyl group.

The NOESY spectrum clearly showed a weak cross-peak between the formyl proton signals of phase opposite to those of the NOE signals. It can be assigned to chemical exchange involving reorientation of the formyl group on the time scale of the NMR experiment and is consistent with the observation that **3.36a** and **3.36b** are in equilibrium on the time scale of the HPLC separation.

A detailed analysis was made of the mass spectrum of **3.36**, acquired by LC-ESI/MS/MS, in an attempt to develop a sensitive assay for the compound. It eluted as a reasonably sharp peak from a normal phase column with a molecular ion $[M + H]^+$ at *m/z* 212. Fragmentation of the molecular ion at 15% relative energy (RE) gave a product ion of *m/z* 184, which is assigned as lost of CO. Increasing the RE to 25% gave two product ions of *m/z* 184 and *m/z* 166; the latter ion is assigned as loss of $H_2O$ from *m/z* 184. Further fragmentation of *m/z* 184 at 20% RE gave a product ion of *m/z* 166, and when 30% RE was used, only a product ion of *m/z* 166 was observed (Figure **3-14**).

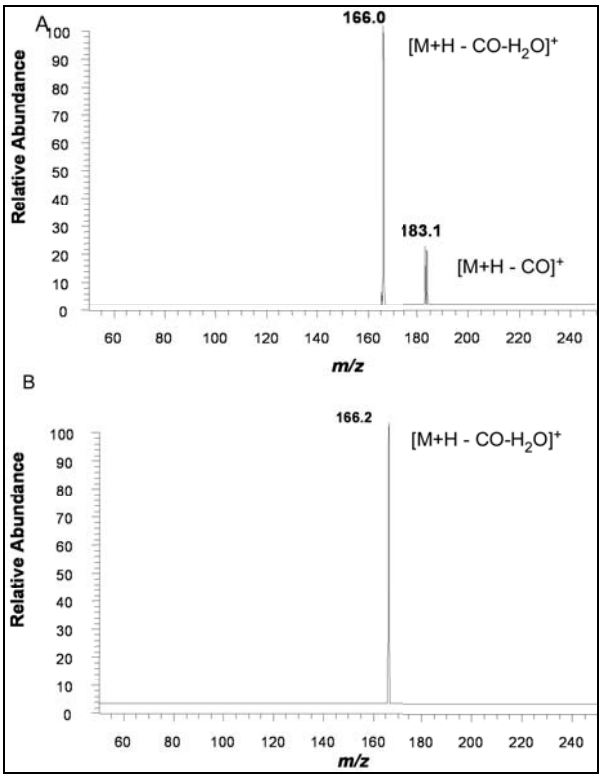



**Figure 3-14**. MS/MS fragmentation of the (2-oxoethyl)-FAPy-Gua **3.36**; **A**) MS spectrum of the major product ion transitions *m/z* 212 → 184 at 20% RE; **B**) $MS^2$ spectrum of the ion transition *m/z* 184 → 166 at 30% RE.

As monitoring of the *m/z* 212 → 184 transition was not diagnostic for our purpose, we chose to use the 30% RE in our $MS^2$ parameters to give only the *m/z* 166 product ion. $MS^3$ fragmentation was carried on the *m/z* 166 product ion at 35% RE. This led to the formation of a major product ion *m/z* 149, due to loss of ammonia, plus several less intense product ions at *m/z* 138 and *m/z* 124, representing losses of CO and NC-$NH_2$ respectively, from the *m/z* 166 ion (Figure **3-15**).

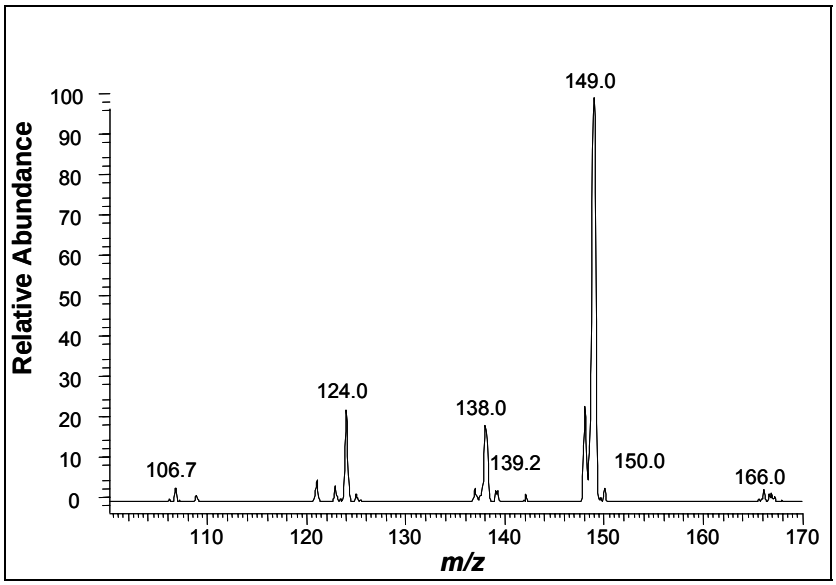



**Figure 3-15**. $MS^3$ profile of molecular ion peak with *m/z* 166, derived from (2-oxoethyl)-FAPy-Gua (**3.36**) standard at 35% RE.

Based on these MS fragmentations, an LC-ESI/MS/MS method was developed for detection of (2-oxoethyl)-FAPy-Gua (**3.36**) consisting of a full scan followed by fragmentation of the *m/z* 212 molecular ion at 25% RE to give

product ions *m/z* 166 and *m/z* 184, then in a third event *m/z* 212 was fragmented at 30% RE to give only *m/z* 166, which was further fragmented at 35% RE. Possible structures of the product ions are shown in Figure **3-16**.

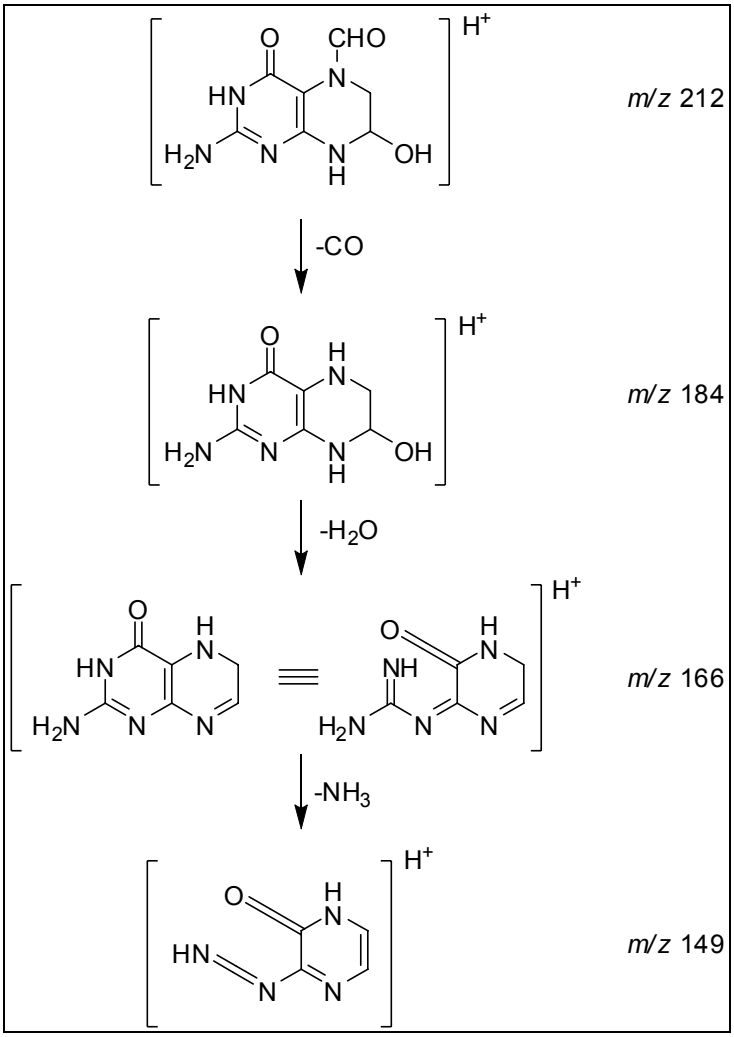



**Figure 3-16**. Proposed assignments of mass spectrometric products ions derived from MS/MS fragmentation of the (2-oxoethyl)-FAPy-Gua (**3.36**).

***Conversion of (2-Oxoethyl)-FAPy-Gua (3.36) into Pyrimidopyrazine (3.38)***

The (2-oxoethyl)-FAPy-Gua (**3.36**) was heated at 60 °C for 2 days in the in the $D_2O/CCl_3COOD$ solution that had been used for NMR measurements. HPLC analysis showed the appearance of a single compound, eluting at $t_R$ 5.18 min (Figure **3-17**).

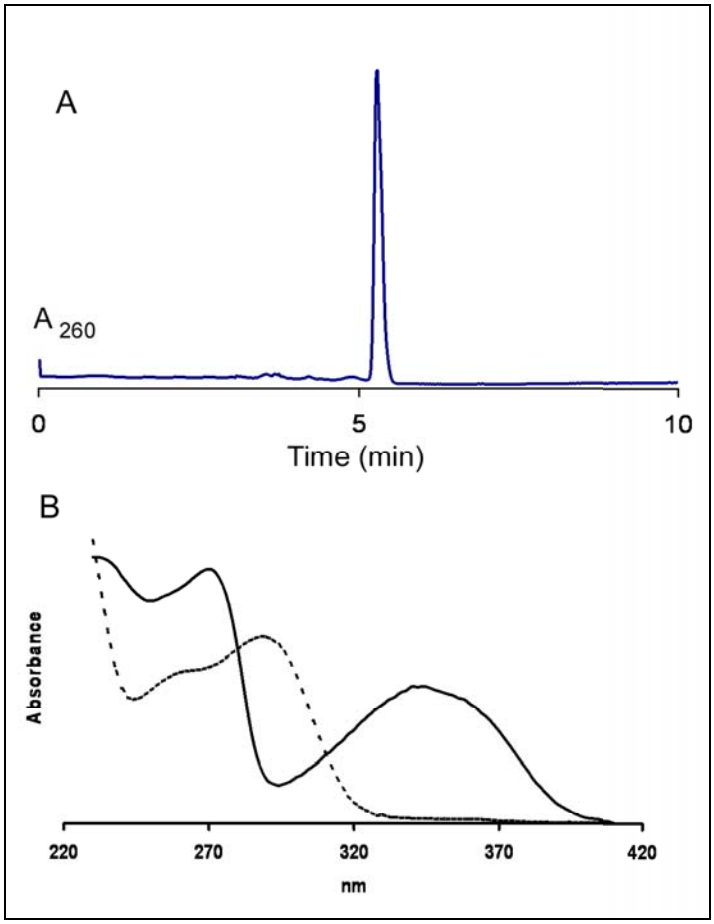



**Figure 3-17**. Characterization of pyramidopyrazine **3.38**. **A**) HPLC traces of pyramidopyrazine **3.38**. **B**) Comparison of the UV spectra of (2-oxoethyl)-FAPy-Gua (**3.36**) (---) and pyramidopyrazine (**3.38**) (—)

Based on the $^{1}H$ and $^{13}C$ NMR data, all of which were acquired under acidic conditions, this compound was identified as dihydropteridinone (**3.37**) (Figure **3-11**). However, **3.37** was only stable under strongly acidic conditions; under neutral conditions **3.37** underwent air oxidation to pyrimidopyrazine **3.38**, which was identified by comparison with an authentic sample prepared by a published procedure (*23*). Compounds **3.37** and **3.38** were readily distinguished by their MS and NMR. However, the oxidation of **3.37** to **3.38** occurred readily (*24-27*); HPLC and LC-ESI/MS/MS of compound **3.37** under neutral conditions showed only the presence of **3.38**.

### *Reaction of 2-Acetoxyoxirane (3.27) with dGuo (3.30) and Oligonucleotide (3.39) Followed by Treatment with 0.5 M NaOH*

The next reactions were carried out with dGuo (**3.30**) and 2-acetoxyoxirane (**3.27**) in dry degassed DMSO. After treatment with 0.5 M NaOH and acid hydrolysis (pH 4 at 25 °C for 2 h) of the glycosyl bond, the samples were analyzed by LC-ESI/MS/MS using the aforementioned MS method. The analyses showed that (2-oxoethyl)-FAPy-Gua (**3.36**) was formed (Figure **3-18**, Panel A). When the acid hydrolysis step was performed for 2 h at 60 °C, LC-ESI/MS/MS analysis showed the presence of **3.38** together with remaining **3.36**.

In order to investigate the formation of **3.36** further, a self-complementary oligomer (5′-AGGCGCCT-3′, 8mer) (**3.39**) having two adjacent dGuo residues was incubated with 2-acetoxyoxirane (**3.27**) in degassed DMSO. This sequence was chosen due to the enhanced nucleophilicity of the N7 position of the dGuo

moieties in duplex DNA. After base treatment and acidic work-up (pH 4 at 25 °C for 2 h), LC-ESI/MS/MS analysis showed the formation of the **3.36** (Figure **3-18**, Panel B).

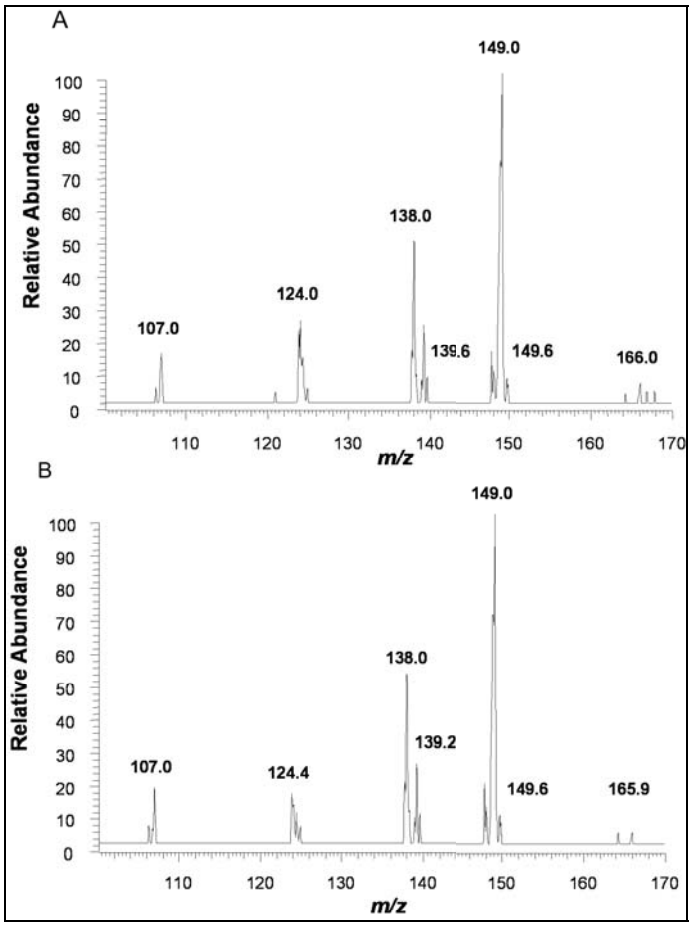



**Figure 3-18**. $MS^3$ profile of molecular ion peak with *m/z* 166, derived from (2-oxoethyl)-FAPy-Gua (**3.36**) at 35% RE; **A**) the reaction product mixture of dGuo (**3.30**) and 2-acetoxyoxirane (**3.27**) in dry degassed DMSO, followed by treatment with 0.5 M NaOH; **B**) the reaction product mixture of oligonucleotide (**3.39**) and 2-acetoxyoxirane (**3.27**) in dry degassed DMSO, followed by treatment with 0.5 M NaOH.

The percentages of formation of the (2-oxoethyl)-FAPy-Gua (**3.36**) from the reaction of dGuo (**3.30**) and ds-DNA (**3.39**) with 2-acetoxyoxirane (**3.27**) were determined by the LC-ESI-MS/MS method based on the calibration curve shown in Figure **3-19**. The $d_3$-(2-oxoethyl)-FAPy-Gua (**3.40**) was synthesized as an internal standard according to Figure **3-10** using $d_5$-allyl bromide (Figure **3-20**).

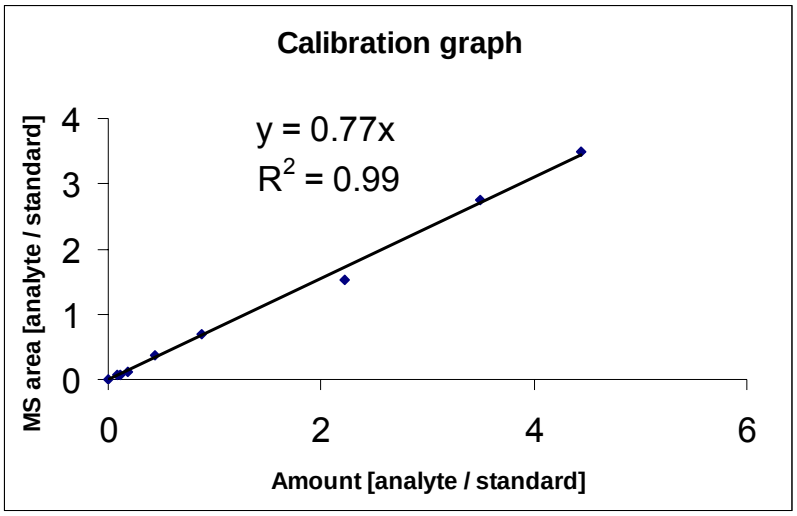



**Figure 3-19**. The calibration curve for the (2-oxoethyl)-FAPy-Gua (**3.36**) and its corresponding deuterated standard $d_3$-(2-oxoethyl)-FAPy-Gua (**3.40**).

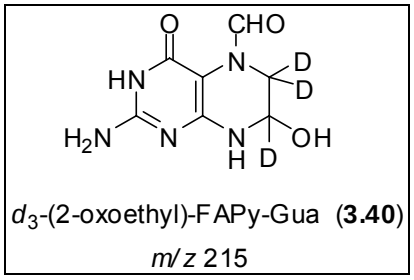



**Figure 3-20**. The structure of the $d_3$-(2-oxoethyl)-FAPy-Gua (**3.40**).

Employing 150-fold excess of the 2-acetoxyoxirane (**3.27**), the reaction dGuo (**3.30**) gave ~ 4% yield of **3.36**, while the yield was less than 1% for the reaction of 330-fold excess of 2-acetoxyoxirane (**3.27**) with the oligonucleotide **3.39** (Table **1**).

| Substrate | Solvent | Time (min) | Yield of 2.36 (%) |
|---|---|---|---|
| dGuo (**3.30**) | DMSO | 15 | 3 |
| | | 30 | 4 |
| | | 60 | 4.5 |
| 5'-AGGCGCCT-3' (**3.39**) | DMSO | 15 | 0.3 |
| | | 30 | 1 |
| | | 60 | 0.7 |

**Table 1**. Yields of formation of (2-oxoethyl)-FAPy-Gua (**3.36**) when dGuo (**3.30**) or 5′-AGGCGCCT-3′ (**3.39**) were treated with 2-acetoxyoxirane (**3.27**).

### *Reaction of 2-Acetoxyoxirane (3.27) and dGuo (3.30) in Phosphate Buffers*

The reaction of dGuo (**3.30**) and 2-acetoxyoxirane (**3.27**) was examined in degassed phosphate buffers at pH ranging from 7.0 to 9.0, followed by acid hydrolysis. HPLC analysis of the acid hydrolates showed that the N7-(2-oxoethyl)-Gua (**3.12**) was the major product. The ESI-LC/MS/MS analysis of the hydrolates showed that (2-oxoethyl)-FAPy-Gua (**3.36**) was not formed.

### *Reaction of 2-Chlorooxirane (3.02) with dGuo (3.30)*

We next examined the reaction of dGuo (**3.30**) and 2-chlorooxirane (*18*) (**3.02**) in dry degassed DMSO or DMF, followed by base treatment and acid hydrolysis. The LC-ESI/MS/MS analysis showed that (2-oxoethyl)-FAPy-Gua (**3.36**) was not formed, presumably due to the reaction of the solvents with 2-chlorooxirane (**3.02**). To avoid this, other solvents, such as THF or $CH_2Cl_2$ were used. Since the solubility of dGuo (**3.30**) in these solvents is very low, 3,5-protected dGuo **3.41** was used instead (Figure **3-21**).

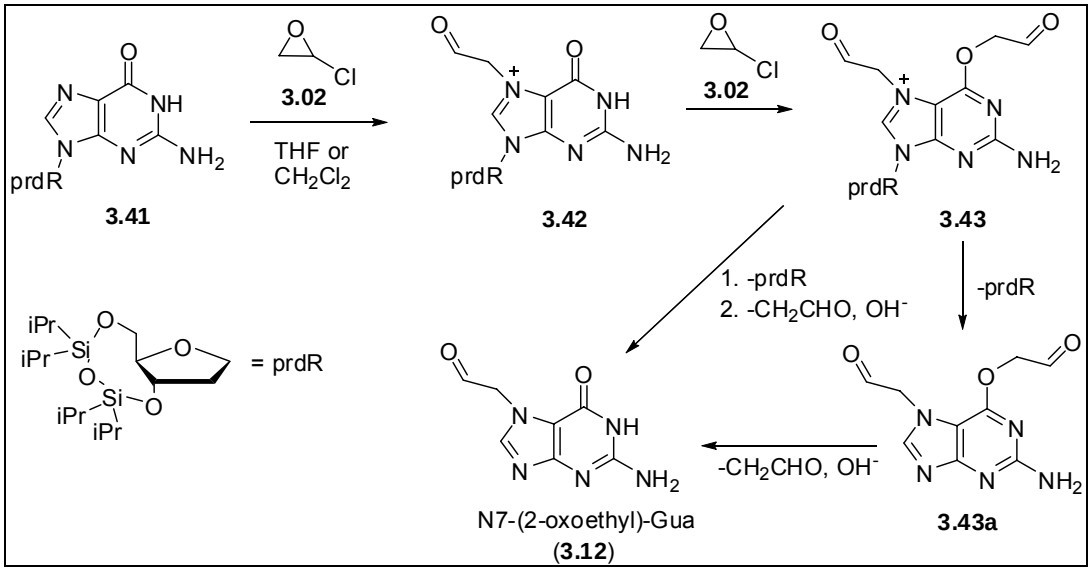



**Figure 3-21**. Reaction of the 2-chlorooxirane (**3.02**) and the 3,5-diprotected dGuo **3.41**.

The LC-ESI/MS/MS analysis showed that (2-oxoethyl)-FAPy-Gua (**3.26**) was not formed, which could be a result of the use of 100-fold excess of 2-

chlorooxirane (**3.02**). The excess of 2-chlorooxirane (**3.02**) may cause alkylation of the $O^6$ position of **3.42** to form the N7,$O^6$-di-(2-oxoethyl)-dGuo species (**3.43**) (Figure **3-21**). The N7,$O^6$-di-(2-oxoethyl)-dGuo species (**3.43**) is much more prone to deglycosylation than to hydroxide ring-opening. Thus, 2-chlorooxirane (**3.02**) was incubated with an excess of the **3.41**, and the reaction mixture was either treated with aqueous sodium hydroxide, followed by acid hydrolysis or only subjected to acid hydrolysis in order to detect the possible formation of the N7,$O^6$-di-(2-oxoethyl)-Gua (**3.43a**), which species if treated with base could lose the $O^6$-2-oxoethyl group to form N7-(2-oxoethyl)-Gua (**3.12**). The LC-ESI/MS/MS analysis of both mixtures showed no evidence for the formation of the (2-oxoethyl)-FAPy-Gua (**3.36**) nor the N7,$O^6$-di-(2-oxoethyl)-Gua species (**3.43a**).

It was suggested that the difference of the reactivity of 2-chlorooxirane (**3.02**) and 1-acetoxyoxirane (**3.27**) could be due to the counter anion of the N7-dGuo cationic species, the chloride versus the acetoxy ion, respectively. The acetoxy group is a stronger base and will better support the formation of the N7-(2-oxoethyl)-dGuo cationic species than the chloride group. In order to test this, the reaction of 2-chlorooxirane (**3.02**) and **3.41** was performed in the presence of tetrabutylammonium acetate, which could provide the necessary acetate base. The LC-ESI/MS/MS analysis of the reaction mixture showed that (2-oxoethyl)-FAPy-Gua (**3.36**) was not formed.

### *Reaction of 2-Chloroacetaldehyde (3.03) with dGuo (3.30)*

dGuo (**3.30**) was incubated with 2-chloroacetaldehyde (**3.03**) (50 % wt. in water) in DMSO and the reaction was monitored by reverse phase HPLC and LC-ESI/MS/MS. The HPLC analysis after 4 h showed the formation of a single species with a long retention time (19.5 min), which was identified by NMR as the carbinolamine **3.44** (Figure **3-22**). The stability of the carbinolamine was investigated in 0.1 M NaOH, DMSO and 0.1 M HCl.

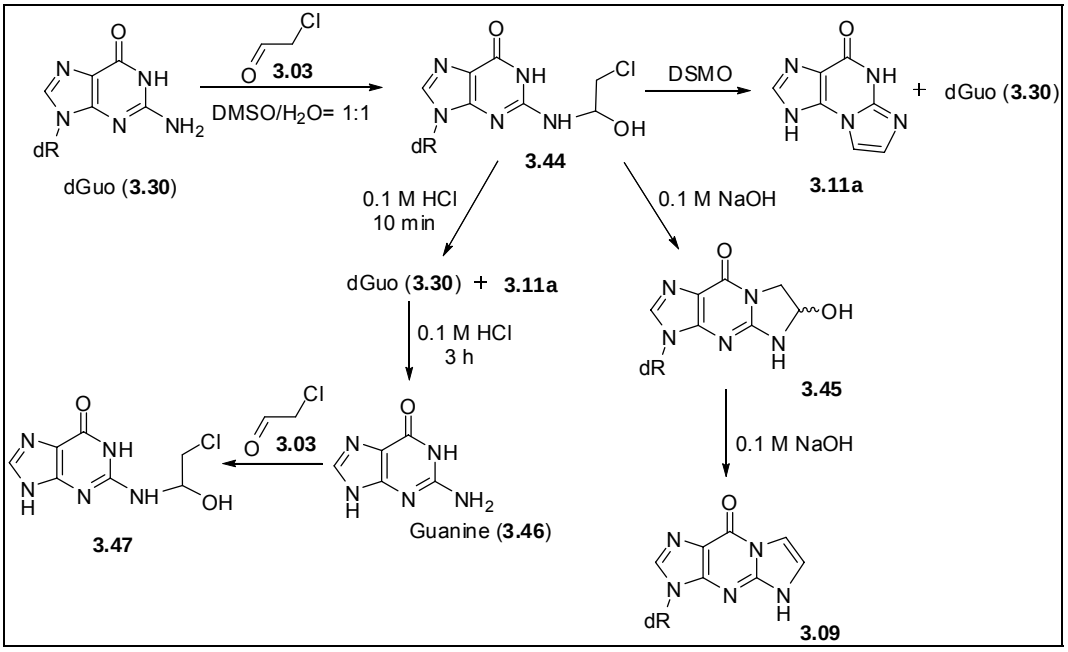



**Figure 3-22**. Reaction of dGuo (**3.30**) and 2-chloroacetaldehyde (**3.02**) in a mixture of DMSO/$H_2O$ = 1:1.

When compound **3.44** was treated with 0.1 M NaOH this led to the formation of **3.45**, which is formed by initial attack of the N1 at C-Cl and displacement of the chlorine (Figure **3-23**, Panel A). Compound **3.45** underwent dehydratation and tautomerization resulting 1,$N^2$-ε-dGuo (**3.09**) (Figure **3-23**, Panel B). The identities of both **3.45** and **3.09** were confirmed by co-injection with authentic standards.

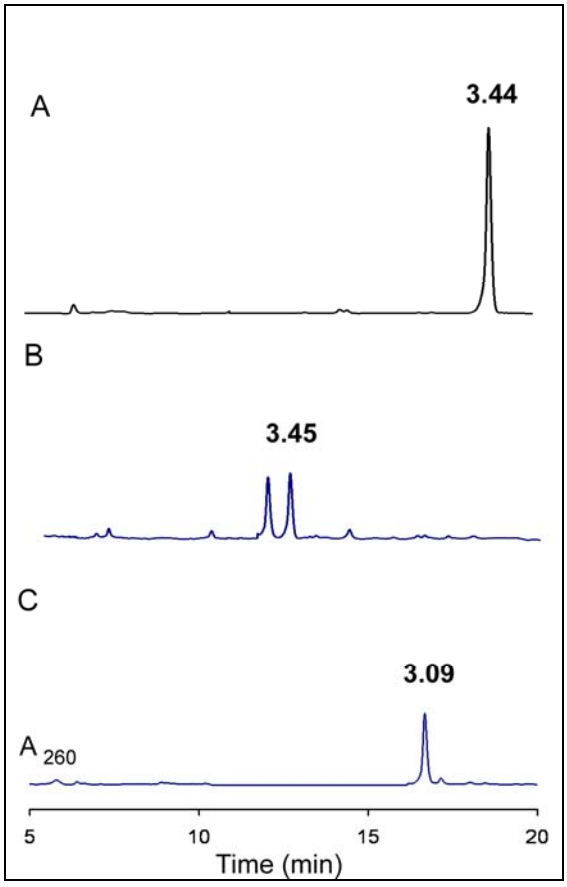



**Figure 3-23**. HPLC traces of the product(s) formed from carbinolamine **3.44** when treated with 0.1 M NaOH; **A**) HPLC traces of pure carbinolamine **3.44**; **B**) HPLC traces of the formation of **3.45**; **C**) HPLC traces of the 1,$N^2$-ε-dGuo (**3.09**) formed via dehydratation of **3.45** and further tautomerization.

The carbinolamine **3.44** was unstable in DMSO and decomposed to $N^2$,3-ε-Guo (**3.11a**) and dGuo (**3.30**) (Figure **3-24**, Panel A). Treatment of carbinolamine **3.44** with acid for a period of 10 min, gave rise to $N^2$,3-ε-Gua (**3.11a**) and dGuo (**3.30**) (Figure **3-24**, Panel B). A mixture of $N^2$,3-ε-Gua (**3.11a**) guanine (**3.46**) and carbinolamine **3.47** was observed after 3 h (Figure **3-24**, Panel C).

If the reaction of 2-chloroacetaldehyde (**3.03**) and dGuo (**3.30**) was allowed to take place for more than 8 hours, the formation of $N^2$,3-ε-Gua (**3.11a**), guanine (**3.46**) and carbinolamine **3.47** (Figure **3-24**, Panel D) were observed. The formation of carbinolamine **3.47** may also be attributed to the hydrolysis of dGuo (**3.30**) to guanine (**3.46**), which reacts with the excess of 2-chloroacetladehyde (**3.03**). The LC-ESI/MS/MS analysis of the reaction mixture of 2-chloroacetaldehyde (**3.03**) and dGuo (**3.30**) did not show that (2-oxoethyl)-FAPy-Gua (**3.36**) was formed.

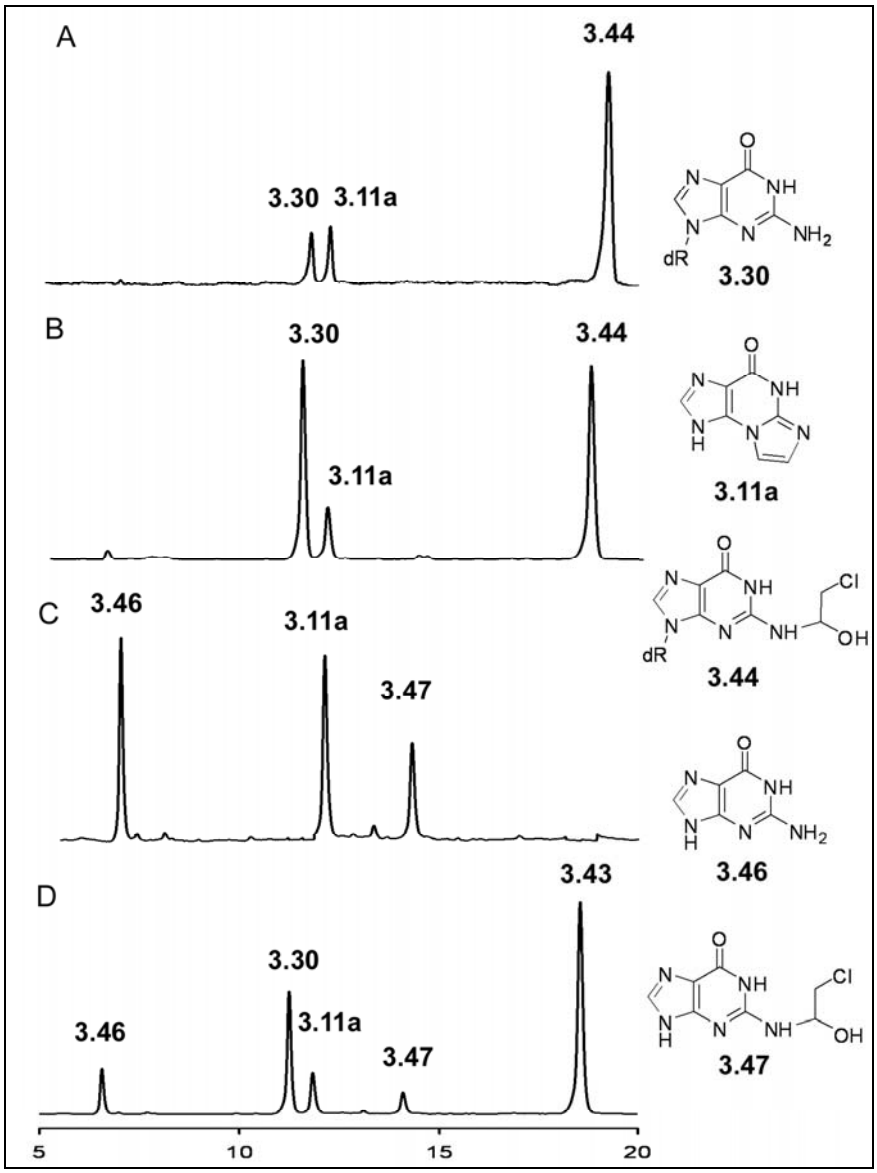



**Figure 3-24**. HPLC traces of the reaction of 2-chloroacetaldehyde (**3.03**) and dGuo (**3.30**); **A**) HPLC traces of the decomposition of carbinolamine **3.44** to dGuo (**3.30**) and $N^2$,3-ε-Gua (**3.11a**) in DMSO; **B**) HPLC traces of the acid hydrolysis of carbinolamine **3.44** for a period of 10 min; **C**) HPLC traces of the acid hydrolysis of carbinolamine **3.44** for a period of 3 h; **D**) HPLC traces of the reaction of 2-chloroacetladehyde (**3.03**) and dGuo (**3.03**) after 8 h.

## *Summary*

We have hypothesized that (2-oxoethyl)-FAPy-dGuo (**3.19**) lesion may be formed via basic hydrolysis of the imidazole ring of the N7-(2-oxoethyl) alkylated dGuo (**3.08**) and could contribute to the mutagenicity of vinyl chloride (**3.01**) and related species (Figure **3-25** and **3-26**). To identify the formation of the (2-oxoethyl)-FAPy-dGuo (**3.19**) lesion, an independent synthesis was developed to provide an authentic sample of the corresponding FAPy-Guo **3.35**. The spectral analysis of **3.35** showed that this species exists as the carbinolamine **3.36**, which is arising by reaction of the aldehyde of **3.35** and the amino group on the C4 of the pyrimidione (Figure **3-25**).

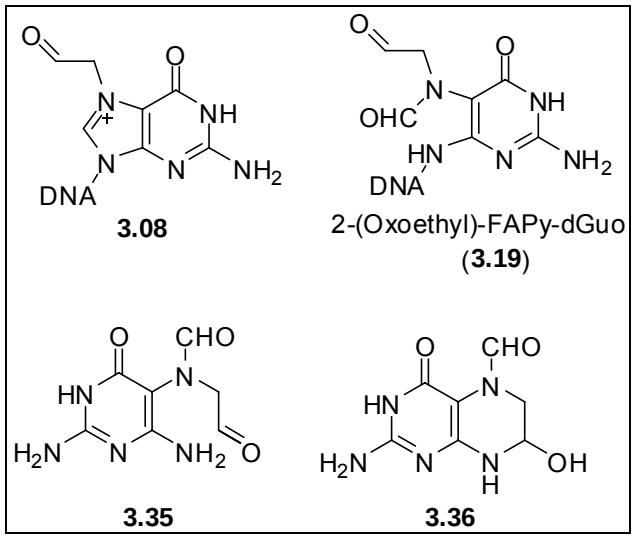



**Figure 3-25**. Structures of the N7-(2-oxoethyl)-dGuo cationic species (**3.08**), (2-oxoethyl)-FAPy-dGuo (**3.19**), (2-oxoethyl)-FAPy-Gua (**3.35**) and carbinolamine **3.36**.

The carbinolamine **3.36** consists of two interconverting species identified as a geometric isomers arising by the rotation of the formyl group. The (2-oxoethyl)-FAPy-Gua (**3.35**), was detected when 2-acetoxyoxirane (**3.27**), a surrogate for 2-chlorooxirane (**3.02**), was used to treat dGuo (**3.30**) and oligonucleotide (**3.39**) followed by treatment with 0.1 M NaOH and subsequent acid hydrolysis (Figure **3-17**). When 2-chlorooxirane (**3.01**) was used as an alkylating agent, the (2-oxoethyl)-FAPy-Gua (**3.35**) was not detected.

We also investigated the reaction of dGuo (**3.30**) with 2-chloroacetaldehyde (**3.03**), a metabolite of 2-chlorooxirane (**3.02**), to form (2-oxoethyl)-FAPy-Gua (**3.35**). The spectral analysis of the product mixture showed that the $N^5$-alkylated FAPy-Gua (**3.35**) was not formed; instead the carbinolamine **3.44** was formed (Figure **3-26**).

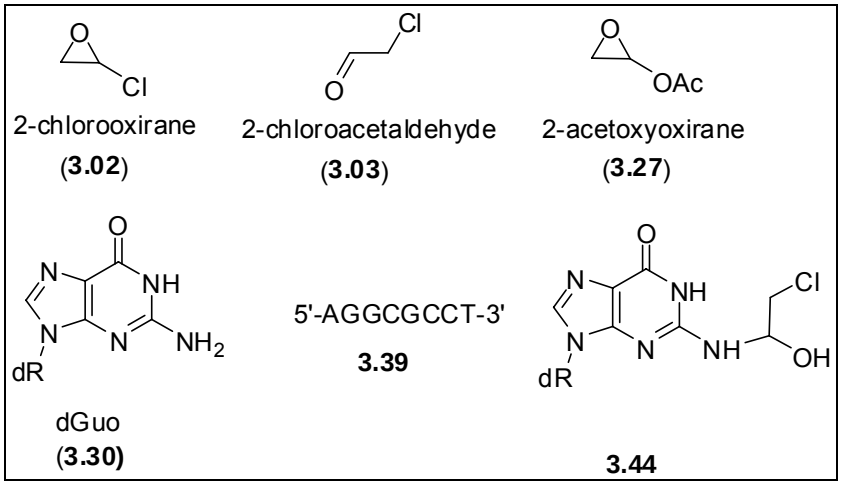



**Figure 3-26**. Structures of the 2-chloroxirane (**3.02**), 2-chloroacetaldehyde (**3.03**), 2-acetoxyoxirane (**3.27**), dGuo (**3.30**), oligonucleotide (**3.39**) and carbinolamine (**3.44**).

## Experimental Procedures

### *Materials and Methods*

All commercial chemicals were of the highest quality available and used without further purification.

### *NMR Spectra*

$^{1}H$ NMR spectra were recorded at 400 or 500 MHz in $D_2O$, acetone-$d_6$ or DMSO-$d_6$; $^{13}C$ NMR spectra were recorded at 125 MHz. The two-dimensional nuclear Overhauser/chemical exchange spectroscopy experiments (NOESY) were performed on a 500 MHz spectrometer with the water peak suppressed by presaturation. A total of 4096 scans were collected using a spectral width of 5000 Hz. The acquisition, pre-acquisition delay, and mixing times were 204 ms, 2 s, and 600 ms, respectively.

### *Chromatography*

HPLC analysis was carried out on a gradient HPLC (Beckman Instruments; System Gold Software) equipped with pump module 125 and photodiode array detector module 168. For monitoring reactions by HPLC, a C-18 reverse phase column (YMC ODS-AQ, 250 x 4.6 mm, flow rate 1.5 mL/min and Phenomenex Gemini-C18, 250 x 4.6 mm, flow rate 1.5 mL/min) was employed with the effluent monitored at 254 nm. Sample purifications were carried out using C-18 reverse phase columns either by HPLC (Phenomenex

Gemini-C18 column, 250 x 10 mm, flow rate 5 mL/min and YMC ODS-AQ, 250 x 10 mm, flow rate 5 mL/min) or by medium pressure chromatographic system (Biotage SP1, Charlottesville, VA) using Biotage C18-HS-(12 + M) with a flow rate 12 mL/min. Effluent composition from both systems was monitored at 254 nm.

Gradient A: 0.1 M aqueous ammonium formate buffer (A) and acetonitrile (B); initially 99% A; 15 min linear gradient to 90% A; 5 min linear gradient to 80% A; isocratic at 80% A for 5 min; 3 min linear gradient to 20% A; isocratic at 20% A for 3 min; 3 min linear gradient to 99% A (initial conditions).

Gradient B: 0.1 M aqueous ammonium formate buffer (A) and methanol (B); initially 99% A; 10 min linear gradient to 60% A; 2 min linear gradient to 20% A; isocratic at 20% A for 2 min; 3 min linear gradient to 99% A (initial conditions).

Gradient C: 0.02 M phosphate buffer (A) and methanol (B); initially 97% A; 10 min linear gradient to 87% A; 2 min linear gradient to 20% A; isocratic at 20% A for 2 min; 3 min linear gradient to 97% A (initial conditions).

Gradient D: 0.1 M aqueous ammonium formate buffer (A) and acetonitrile (B); isocratic at 100% A for 20 min; 40 min linear gradient to 70% A; 5 min linear gradient to 0% A; isocratic at 0% A for 10 min; 15 linear gradient to 100% A (initial conditions).

Gradient E: water (A) and acetonitrile (B); initially 99% A; linear gradient to 90% A for 15 min; 5 min linear gradient to 80% A; isocratic at 80% A for 5 min; 3min linear gradient to 20% A; isocratic at 20% A for 2 min; linear gradient to 99% A (initial conditions).

Normal phase HPLC analysis of (2-oxoethyl)-FAPy-Gua (**3.36**) was performed with Phenomenex Luna 5 µ Silica column (150 mm x 2 mm, flow rate 0.60 mL/min). Effluent composition was monitored at 254 nm. Gradient F: acetonitrile (A) and 38.8% acetonitrile, 60% $H_2O$, 0.2% AcOH, 1% (1M) ammonium acetate (B); isocratic at 99% A for 3 min; 3 min linear gradient to 0% A; isocratic at 0% A for 11 min; 2 min linear gradient to 99% A , isocratic at 99% A for 9 min (initial conditions).

***Mass Spectrometry***

FAB mass spectra (low and high resolution) were obtained at the Mass Spectrometry Facility at the University of Notre Dame (Notre Dame, IN) using a matrix of nitrobenzyl alcohol.

LC-ESI/MS/MS was performed on a DecaXP ion trap instrument (ThermoFinnigan, San Jose, CA) using an Agilent 1100 A pump system (Agilent, Foster City, USA). Detection and quantification of the (2-oxoethyl)-FAPy-Gua species (**3.36**) were carried out on a Phenomenex Luna 5 µ Silica column (150 mm x 2 mm). Buffer A consisted of 96.8% ACN, 2% $H_2O$, 0.2% AcOH, 1% (1M) ammonium acetate; buffer B consisted of 38.8% acetonitrile, 60% $H_2O$, 0.2% AcOH, 1% (1M) ammonium acetate. The following gradient program was used with a flow rate of 0.35 mL/min: isocratic at 99% A for 3 min; 3 min linear gradient to 0% A; isocratic at 0% A for 5 min; 1 min linear gradient to 99% A; isocratic at 99% A for 8 min (initial conditions). The (2-oxoethyl)-FAPy-Gua species (**3.36**) eluted at 7.27 min. Samples were injected using an autosampler. ESI conditions:

source voltage 4 kV, $N_2$ sheath gas setting 64 units, $N_2$ auxiliary sweep gas setting 11 units, capillary voltage 3 V, capillary temperature 300 °C, tube lens offset 0 V. A method consisting of three scan events was used: 1) full scan, 2 microscans, ion accumulation time 200 ms, *m/z* [100.00-500.00]; 2) selected reaction monitoring (SRM): 1 microscan, spectral width 2, ion accumulation time 50 ms, MS *m/z* 212.10 @25 [165.50-166.50, 183.50-184.50]; 3) $MS^3$ 1 microscan, spectral width 2, ion accumulation time 200 ms, *m/z* 212.10 → *m/z* 166 @ 35 [100-170].

***Reaction of dGuo (3.30) and 2-Acetoxyoxirane (3.27) in Phosphate Buffers***

dGuo:$H_2O$ (**3.30**) (0.485 mg, 0.0017 mmol) in degassed phosphate buffers (pH 7, pH 8 or pH 9) (200 µL) was treated with 2-acetoxyoxirane (**3.27**) (0.26 mmol, 24 µL) at room temperature. The progress of the reaction was monitored by HPLC using gradient E and the LC-ESI/MS/MS method for the detection of (2-oxoethyl)-FAPy-Gua (**2.36**).

***Reaction of dGuo (3.30) and 2-Acetoxyoxirane (3.27) in DMSO***

dGuo:$H_20$ (**3.30**) (0.5 mg, 0.00175 mmol) in dry, degassed DMSO (100 µL) was treated with 2-acetoxyoxirane (**3.27**) (0.26 mmol, 24 µL) at room temperature. The progress of the reaction was monitored by HPLC using gradient E and the LC-ESI/MS/MS method for the detection of (2-oxoethyl)-FAPy-Gua (**2.36**).

***Reaction of dGuo (3.03) and 2-Acetoxyoxirane (3.27) in DMSO Followed by Treatment with 0.5 M NaOH***

dGuo:$H_2O$ (**3.30**) (0.462 mg, 0.0016 mmol) in dry, degassed DMSO (100 µL) was treated with 2-acetoxyoxirane (**3.27**) (0.24 mmol, 22 µL) at room temperature. Aliquots (25 µL) were withdrawn at 15, 30, and 60 min and treated with 0.5 M NaOH (200 µL) for 3 min. HCl (10 µL, 6 M) was then added to these solutions to lower the pH to 4.0-5.0; the mixtures were allowed to stand at room temperature for 1.5 h. The mixtures were spiked with measured quantities (0.3019 µg) of $d_3$-**3.40** as an internal standard and analyzed by LC-ESI/MS/MS (10 µL samples). Quantitation involved comparison of peak areas for signals at *m/z* 212 in the full scan spectra with the respective signals derived from $d_3$-**3.40**, which appeared at *m/z* 215.

***Reaction of the Oligonucleotide 5'-AGGCGCCT-3' (3.39) with 2-Acetoxyoxirane (3.27) in DMSO Followed by Treatment with 0.5 M NaOH***

The oligonucleotide (**3.39**) (25 $A_{260}$ units) dissolved in degassed DMSO (300 µL) was treated with 2-acetoxyoxirane (**3.27**) (10 µL) at room temperature. Aliquots (25 µL) were withdrawn at 15, 30 and 60 min and treated with 0.5 M NaOH (200 µL) for 3 min. HCl (<10 µL, 6 M) was added to these solutions to lower the pH to 4.0-5.0; the mixtures were allowed to stand at room temperature for 2.0 h. Quantitation was accomplished as described above using $d_3$-**3.40** as the internal standard. Samples (10 µL) were analyzed by LC-ESI/MS/MS.

***Preparation of $d_3$-(2-Oxoethyl)-FAPy-Gua (3.40) and Construction of the Calibration Curve***

The $d_3$-(2-oxoethyl)-FAPy-Gua (**3.40**) was synthesized using $d_5$-allyl bromide according to the procedure for the synthesis of the unlabeled sample (procedure to follow). Its structure was confirmed by $^1$H NMR, MS and UV. EI analysis showed the anticipated molecular ion peak at *m/z* 215. The standard calibration curve was constructed using nine concentrations of the undeuterated analyte (30 to 3000 ng/mL injected) and a constant amount of the deuterated standard (0.3019 µg). The calibration graph showed good linearity ($R^2$ = 0.9962).

***2-Acetoxyoxirane (3.27)***

Vinyl acetate (1.04 mL, 0.017 mmol) (**3.25**) was added dropwise to an acetone solution of dimethyl dioxirane (*28*) (375 mL, 0.05 M) (**3.26**) at -78 °C. The mixture was allowed to warm up to room temperature. After stirring for 1 h, the mixture was concentrated under vacuum (70 – 80 Torr). The distillation was stopped after the volume had been reduced by ~30%. The remaining solution was dried with anhydrous $K_2CO_3$ for 15 min at 0 °C, filtered and distilled (80°C, 70 – 80 Torr) to give 1.8 g of 2-acetoxyoxirane (**3.27**) (78%). $^1$H NMR (acetone-$d_6$): δ 5.36 (dd, 1H, *J* = 2.4 Hz, *J* = 1.2 Hz), 2.72 (dd, 1H, *J* = 2.4 Hz, *J* = 4.4 Hz), 2.69 (dd, 1H, *J* = 1.2 Hz, *J* = 4.4 Hz), 1.94 (s, 3H).

***N-[2-Amino-6-[(2-deoxy-β-D-erythro-pentopyranosyl)amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-N-(2-propenyl)-formamide (3.32)***

dGuo:$H_2O$ (**3.30**) (300 mg, 1.12 mmol) was dissolved in 10 mL of degassed DMSO. Allyl bromide (1.94 ml, 22.40 mmol) was added dropwise and the reaction mixture was stirred at room temperature for 3 h. Excess allyl bromide was removed *in vacuo* and the residue was treated with aqueous NaOH (30 mL, 1 M) for 1 h. The reaction was monitored with an YMC ODS-AQ column (gradient A) and purified by Biotage SP1 chromatography (gradient D) to afford **3.32** (150 mg, 41% yield). $^1$H NMR (DMSO-$d_6$) mixture of isomers: δ 10.69 (broad s, 1H, amide), 7.76-7.69 (multiple s, 1H, CHO), 7.06 (d, 1H, NH, *J* = 7.6 Hz), 6.64 (broad s, 2H, $NH_2$), 5.77-5.74 (m, 1H, =CH ), 5.42-5.35 (m, 1H, H-1'), 5.14-4.97 (m, 3H, OH-3', $=CH_2$), 4.70-4.43 (m, 1H, OH-5'), 4.10-3.85 (m, 3H, H-3', $CH_2$-N), 3.78-3.65 (m, 1H, H-4'), 3.67-3.24 (m, 2H, H-5'), 1.93-1.71 (m, 2H, H-2'). HRMS (FAB$^+$) *m/z* calcd for $C_{13}H_{20}N_5O_5$ $[M + H]^+$ 326.1464, found 326.1449.

***N-(2,6-Diamino-3,4-dihydro-4-oxo-5-pyrimidinyl)-N-(2-propenyl)-formamide (3.33)***

Compound **3.32** (130 mg, 0.39 mmol) was heated in aqueous HCl (1M, 5mL) for 2 h at 60 °C. The reaction was monitored with an YMC ODS-AQ column (gradient A) and purified with the Biotage SP1 chromatograph (gradient D) to afford **3.33** (75 mg, 92% yield). $^1$H NMR (DMSO-$d_6$) first isomer: δ 11.50 (broad s, 1H, amide) 8.34 (s, 1H, CHO), 6.36-6.30 (broad s, 4H, $NH_2$), 5.81-5.71 (m, 1H, =CH, *J* = 8 Hz), 5.11-5.03 (m, 2H, $=CH_2$), 4.03-3.97 (m, 2H, $CH_2$-N); second

isomer: δ 11.35 (broad s, 1H, amide) 7.74 (s, 1H, CHO), 6.14-6.00 (broad s, 4H, $NH_2$), 5.70-5.68 (m, 1H, =CH, *J* = 8 Hz), 4.99-4.95 (m, 2H, $=CH_2$), 3.90-3.78 (m, 2H, $CH_2$-N). HRMS ($FAB^+$) *m/z* calcd for $C_8H_{12}N_5O_2$ $[M + H]^+$ 210.0991, found 210.0997.

***N-(2,6-Diamino-3,4-dihydro-4-oxo-5-pyrimidinyl)-N-(2,3-dihydroxypropyl)-formamide (3.34)***

A solution of **3.33** (70 mg, 0.33 mmol) in water (0.3 mL) was added to a mixture of water (1 mL), acetone (0.5 mL), N-methylmorpholine-N-oxide (46.39 mg, 0.39 mmol) and of $OsO_4$ (~ 1 mg). The reaction mixture was stirred overnight at room temperature then the solvents were evaporated *in vacuo*. Progress of the reaction was monitored with an YMC ODS-AQ column (gradient B). The product was purified by Biotage SP1 chromatography (gradient D) to afford **3.34** (69 mg, 85% yield). $^1H$ NMR (DMSO-$d_6$) mixture of isomers: δ 12.00 (broad s, 1H, amide), 8.07-7.74 (s, 1H, CHO), 6.70-6.50 (broad s, 2H, $NH_2$), 6.43-6.20 (broad s, 2H, $NH_2$), 4.02-3.99 (m, 1H, $CH_2OH$), 3.77-3.75 (m, 1H, $CH_2OH$), 3.53-3.48 (m, 1H, C<u>H</u>OH), 3.35-3.25 (m, 2H, $CH_2N$). HRMS ($FAB^+$) *m/z* calcd for $C_8H_{14}N_5O_4$ $[M + H]^+$ 244.1046, found 244.1058.

***2-Amino-3,4,7,8-tetrahydro-7-hydroxy-4-oxopteridine-5(6H)-carbaldehyde ((2-oxoethyl)-FAPy-Gua) (3.36)***

Compound **3.34** (60 mg, 0.25 mmol) was dissolved in water (1 mL), then treated with aqueous $NaIO_4$ solution (16 mL, 20 mM in 0.05 M phosphate buffer,

pH 7) for 45 min at room temperature. The reaction was monitored and purified by HPLC using a Phenomenex Gemini-C18 column (gradient C; Solvent system B consisted of 0.1 M aqueous ammonium formate buffer) to afford **3.36** as a mixture of the geometrical isomers. Compound **3.36** was liophilized several times with water to remove the ammonium formate (42 mg, 79% yield). $^{1}$H NMR (500 MHz, $D_2O$, 10% $CCl_3COOD$) isomer **3.36b**: δ 7.89 (s, 1H, CHO), 5.17 (broad tr, 1H, C<u>H</u>OH), 3.76 (dd, 1H, $CH_2$, $J_1$ = 1.8 Hz, $J_2$ = 13.4 Hz), 3.11 (dd, 1H, $CH_2$, $J_1$ = 1.8 Hz, $J_2$ = 13.4 Hz); isomer **3.36a**: δ 8.56 (s, 1H, CHO), 5.17 (broad tr, 1H, C<u>H</u>OH), 4.42 (dd, 1H, $CH_2$, $J_1$ = 1.8 Hz, $J_2$ = 18.0 Hz), 2.73 (dd, 1H, $CH_2$, $J_1$ = 1.8 Hz, $J_2$ = 18.0 Hz). $^{13}$C NMR (125 MHz, $D_2O$, 10% $CCl_3COOD$) isomer **3.36b**: δ 163.55, 158.13, 154.25, 153.45, 91.72, 72.70, 42.32; isomer **3.36a**: δ 165.42, 158.79, 155.25, 153.84, 93.04, 72.70, 48.10. HRMS ($FAB^+$) *m/z* calcd for $C_7H_{10}N_5O_3$ $[M + H]^+$ 212.0784, found 212.0793; UV ($H_2O$) $\lambda_{max}$ 290 nm.

***2-Amino-5,8-dihydro-4(3H)-pteridinone (3.37)***

Compound **3.36** (5 mg, 0.023 mmol) was dissolved in $D_2O$ (0.5 mL) containing $CCl_3COOD$ (0.05 mL). The solution was heated in an NMR tube for 2 d at 60 °C to give **3.37**. The reaction was monitored (gradient C) by HPLC using a Phenomenex Gemini-C18 column. $^{1}$H NMR (500 MHz, $D_2O$, 10% $CCl_3COOD$): δ 8.85 (d, 1H, *J* = 2.32 Hz), 8.78 (d, 1H, *J* = 2.32 Hz); $^{13}$C NMR (125 MHz, $D_2O$, 10% $CCl_3COOD$): δ 165.75, 160.00, 150.09, 147.83, 144.32, 128.22. HRMS ($FAB^+$) *m/z* calcd for $C_6H_8N_5O$ $[M + H]^+$ 166.0729, found 166.0701.

***2-Amino-4-hydroxypyrimido[4,5-b]pyrazine (3.38)***

Compound **3.38** was synthesized as previously described (*23*). The HPLC analysis was performed by HPLC using a Phenomenex Gemini-C18 column (gradient C). $^{1}$H NMR (500 MHz, $D_2O$, 10% $CCl_3COOD$): δ 7.48 (d, 1H, *J* = 2.25 Hz), 7.39 (d, 1H, *J* = 2.25 Hz); $^{13}$C NMR (125 MHz, $D_2O$, 10% $CCl_3COOD$): δ 159.83, 151.85, 150.99, 147.65, 143.90, 127.90. HRMS ($FAB^+$) m/z calcd for $C_6H_6N_5O$ $[M + H]^+$ 164.0572, found 164.0565; UV ($H_2O$), $\lambda_{max}$ 270, 343 nm.

***$N^2$-(2-Chloro-1-hydroxy-ethyl)-(2-deoxy-β-D-erythro-pentopyranosyl)-guanine (3.44)***

dGuo:$H_2O$ (**3.30**) (20 mg, 0.070 mmol) dissolved in DMSO (220 µL) was treated with 2-chloroacetaldehyde (**3.03**) (50 % wt in water, 0.35 mmol, 54.95 µL) for 2 h at room temperature. The product was purified by reverse phase HPLC using YMC ODS-AQ column and gradient E. $^{1}$H NMR (400MHz, DMSO-$d_6$): δ 10.65 (broad s, 1H, amide), 8.01 (s, 1H, H-8), 6.87 (m, 1H, NH), 6.19-6.16 (m, 1H, H1'), 5.64-5.61 (m, 1H, C<u>H</u>OH), 4.35-4.33 (m, 1H, H3'), 3.82-3.73 (m, 1H, H4'), 3.72-3.70 (m, 2H, $CH_2Cl$), 3.54-3.58 (m, 2H, H5'), 2.60-2.55 (m, 1H, H2'), 2.25-2.19 (m, 1H, H2'). HRMS ($FAB^+$) *m/z* calcd for $C_{12}H_{16}ClN_5O_5$ $[M + H]^+$ 345.0840, found 345.0835.

***Stability Studies on the Carbinolamine (3.44)***

Carbinolamine (**3.44**) (0.1 mg) was dissolved in aqueous NaOH (300 μL, 0.1 M), DMSO (300 μL) or HCl (300 μL, 0.1 M) and the reaction mixtures was stirred at room temperature. Aliquots were withdrawn and analyzed by HPLC using YMC ODS-AQ column and gradient E.

## References


(1) Lee, F. I. and Harry, D. S. (1974) Angiosarcoma of the liver in a vinyl-chloride workers. *Lancet 1*, 1316-1318.

(2) Martsteller, H. I., Lelbach, W. K., Muler, R. and Gedigk, P. (1975) Unusual splenomegalic liver disease as evidenced by peritoneoscopy and guided liver biopsy among polyvynyl chloride production workers. *Ann. N. Y. Acad. Sci. 246*, 95-134.

(3) Nicholson, W., Hammond, C., Seidmon, H. and Selikoff, J. (1975) Mortality experience of a cohort of a vinyl chloride-polyvinyl chloride workers. *Ann. N. Y. Acad. Sci. 246*, 225-230.

(4) Creech, J. L. and Johnson, M. N. (1974) Angiosarcoma of liver in the manufacture of polyvinyl chloride. *J. Occup. Med. 16*.

(5) Guengerich, F. P., Crawford, W. M. and Watanabe, P. G. (1979) Activation of vinyl chloride to covalently bound metabolites: roles of 2-chloroethylene and 2-chloroacetaldehyde. *Biochemistry 18*, 5177-5182.

(6) Barbin, A., Laib, R. J. and Bartsch, H. (1985) Lack of miscoding properties of 7-(2-oxoethyl)guanine, the major vinyl chloride-DNA adduct. *Cancer Res. 45*, 2440-2444.

(7) Kusmierek, J. T. and Singer, B. (1982) Chloroacetaldehyde-treated ribo- and deoxyribopoly-nucleotides. 2. Errors in transcription by different

polymerases resulting from ethenocytosine and its hydrated intermediate. *Biochemistry 21*, 5723-5728.

(8) Singer, B., Abbott, L. and Spenger, S. (1984) Assessment of mutagenic efficiency of the two carcinogen modified nucleosides, 1,$N^6$-ethenodeoxyadenosine and $O^4$-methyldeoxythymidine, using polymerases of varying fidelity. *Carcinogenesis 5*, 1165-1171.

(9) Singer, B., Spenger, S., Chavez, F. and Kusmierek, J. T. (1987) The vinyl chloride-derived nucleoside $N^2$,3-ethenoguanosine, is a highly efficient mutagen in transcription. *Carcinogenesis 8*, 745-747.

(10) Singer, B., Kusmierek, J. T., Folkman, W., Chavez, F. and Dosanjh, M. K. (1991) Evidence for the mutagenic potential of the vinyl chloride induced $N^2$,3-etheno-deoxyguanosine, using a site-directed kinetic assay. *Carcinogenesis 12*, 745-747.

(11) Fedtke, N., Boucheron, J. A., Turner, M. J., Jr. and Swenberg, J. A. (1990) Vinyl chloride-induced DNA adducts. I: Quantitative determination of $N^2$,3-ethenoguanine based on electrophore labeling. *Carcinogenesis 11*, 1279-1285.

(12) Nair, J., Barbin, A., Guichard, Y. and Bartsch, H. (1995) 1,$N^6$-Ethenodeoxyadenosine and 3,$N^4$-ethenodeoxycytine in liver DNA from humans and untreated rodents detected by immunoaffinity/$^{32}$P-postlabeling. *Carcingenesis 16*, 613-617.

(13) El Ghissassi, F., Barbin, A., Naor, J. and Bartsch, H. (1995) Formation of 1,$N^6$-ethenoadenine and 3,$N^4$-ethenocytosine by lipid peroxidation products and nucleic acid bases. *Chem. Res. Toxicol. 8*, 278-283.

(14) Kalam, M. A., Haraguchi, K., Chandani, S., Loechler, E. L., Moriya, M., Greenberg, M. M. and Basu, A. K. (2006) Genetic effects of oxidative DNA damages: comparative mutagenesis of the imidazole ring-opened formamidopyrimidines (Fapy lesions) and 8-oxo-purines in simian kidney cells. *Nucleic Acids Res. 34*, 2305-2315.

(15) Smela, M. E., Hamm, M. L., Henderson, P. T., Harris, C. M., Harris, T. M. and Essigmann, J. M. (2002) The aflatoxin $B_1$ formamidopyrimidine adduct plays a major role in causing the types of mutations observed in human hepatocellular carcinoma. *Proc. Nat. Acad. Sci. USA 99*, 6655-6660.

(16) Asagoshi, K., Terato, H., Ohyama, Y. and Ide, H. (2002) Effects of a guanine-derived formamidopyrimidine lesion on DNA replication. *J. Biol. Chem. 277*, 14589–14597.

(17) Walling, C. and Fredricks, P. S. (1962) Positive halogen compound. IV. Radical reactions of chlorine and t-butyl hypochlorite with some small ring compounds. *J. Am. Chem. Soc. 84*, 3326-3331.

(18) Kesselmayer, M. A. and Sheridan, R. S. (1986) Methoxychlorocarbene. Matrix spectroscopy and photochemistry. *J. Am. Chem. Soc. 108*, 99-107.

(19) Awada, M. and Dedon, P. C. (2001) Formation of the 1,$N^2$-glyoxal adduct of deoxyguanosine by phosphoglycolaldehyde, a product of 3'-deoxyribose oxidation in DNA. *Chem. Res. Toxicol. 14*, 1247-1253.

(20) Thornalley, P., Wolff, S., Ceabbe, J. and Stem, A. (1984) The autoxidation of glyceraldehyde and other simple monosaccharides under physiological conditions catalyzed by buffer ions. *Bioch. Biophys. Acta 797*, 276-287.

(21) Piper, J. R., Laseter, A. G. and Montgomery, J. A. (1980) Synthesis of potential inhibitors of hypoxanthine-guanine phosphoribosyltransferase for testing as antiprotozal agents. 1. 7-Substitutes 6-oxopurines. *J. Med. Chem. 23*, 357-364.

(22) Roe, J., Jr., Paul, J. S. and Montgpmery, J. A. (1973) Synthesis and PMR spectra of 7-hydroxyalkylguanosinium acetates. *J. Het. Chem. 10*, 849-857.

(23) Cain, D. K., Mallette, M. F. and Tayler, E. C. (1946) Pyrimido[4,5-b]pyrazines. I. Synthesis of 6,7-symetrically substituted derivatives. *J. Am. Chem. Soc. 68*, 1996-1999.

(24) Kwee, S. and Lund, H. (1973) Electrochemistry of some substituted pteridines. *Bioch. Biophys. Acta 297*, 285-296.

(25) Kwee, S. and Lund, H. (1974) Indirect electrolysis of macromolecules by means of pteridone mediators. *Bioelectrochemistry and Bioenergetics 1*, 87-95.

(26) Farahani, M., Surdhar, P. S., Allen, S., Armstrong, D. A., Schoneich, C., Mao, Y. and Asmus, K.-D. (1991) Reactions of $CO_2^{\cdot-}$ radicals with pterin and pterin-6-carboxylate ions. *J. Chem. Soc. Perkin Trans 2*, 1687-1693.

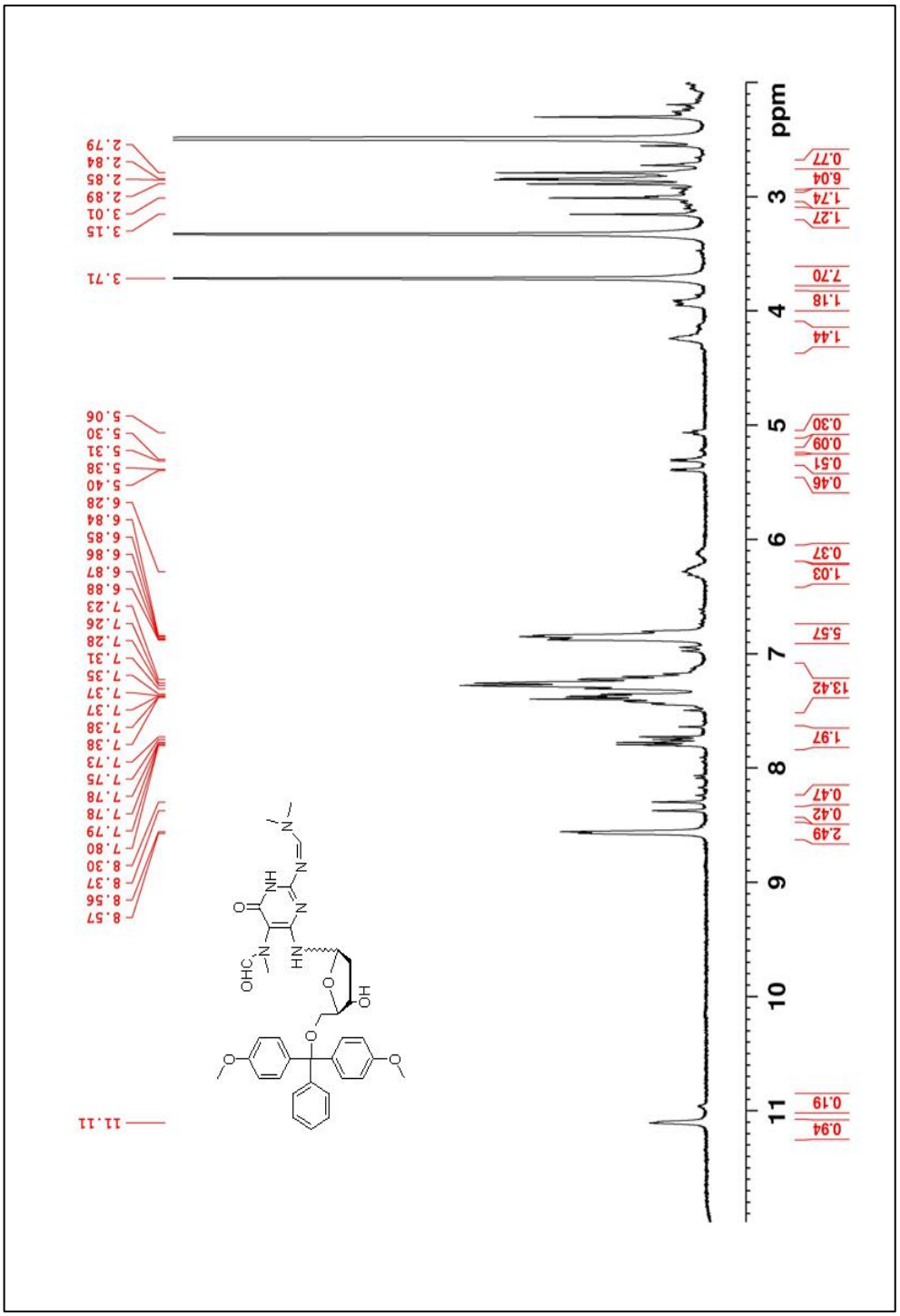
(27) Forrest, H. S., Baalen, V. C., Viscontini, M. and Piraux, M. (1960) Reaktion von $CN^-$ mit hydriertem 2-amino-6-hydroxy-pteridin. *Helv. Chim. Acta 128*, 1006-1010.

(28) Adam, W., Bialas, J. and Hadjiarapoglou, L. (1991) A convenient preparation of acetone solutions of dimethyl dioxirane. *Chem. Berich. 124*, 2377.

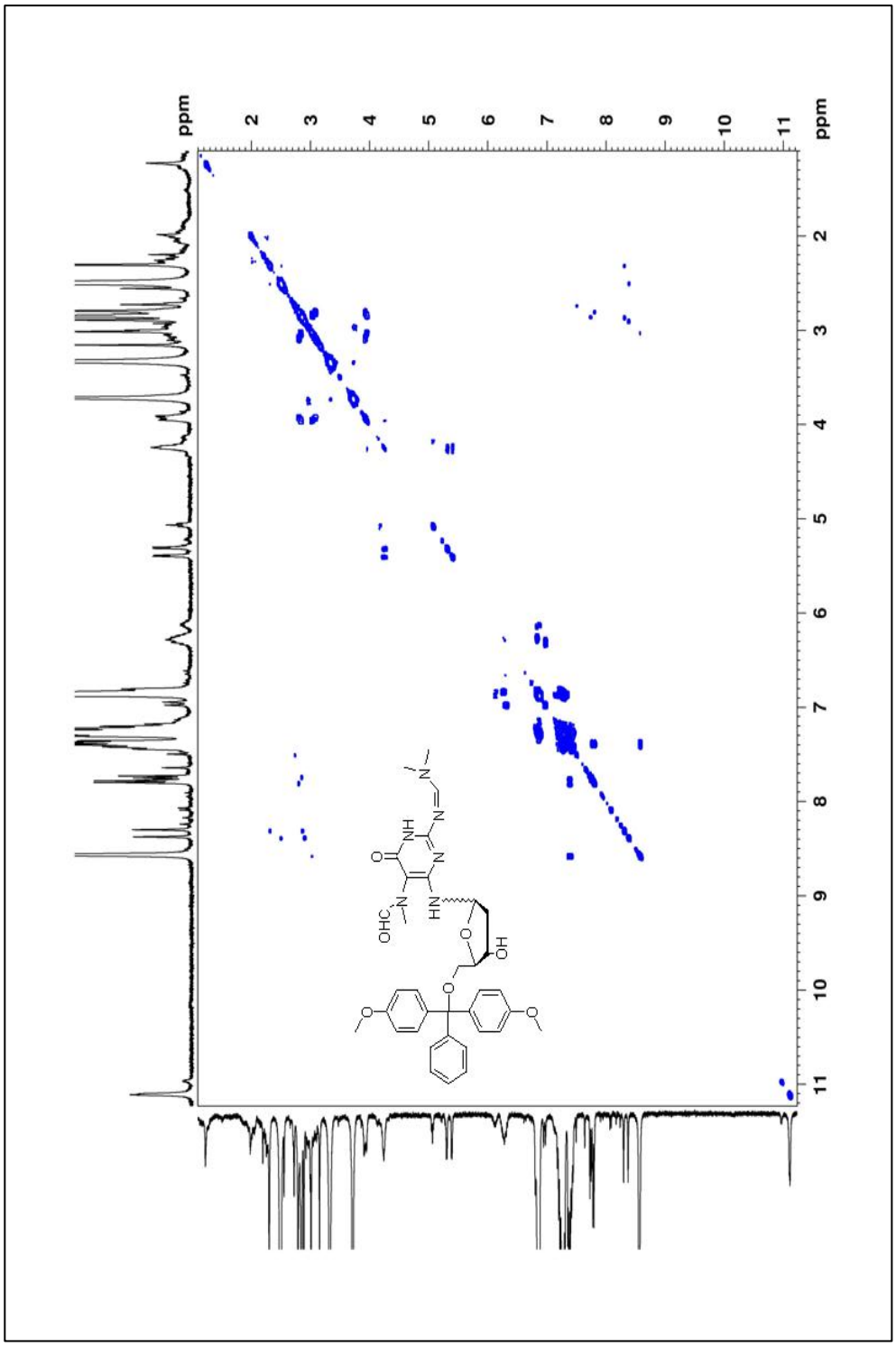
# Appendix I:

# $^{1}H$, $^{13}C$, COSY and $^{31}P$ NMR Spectra

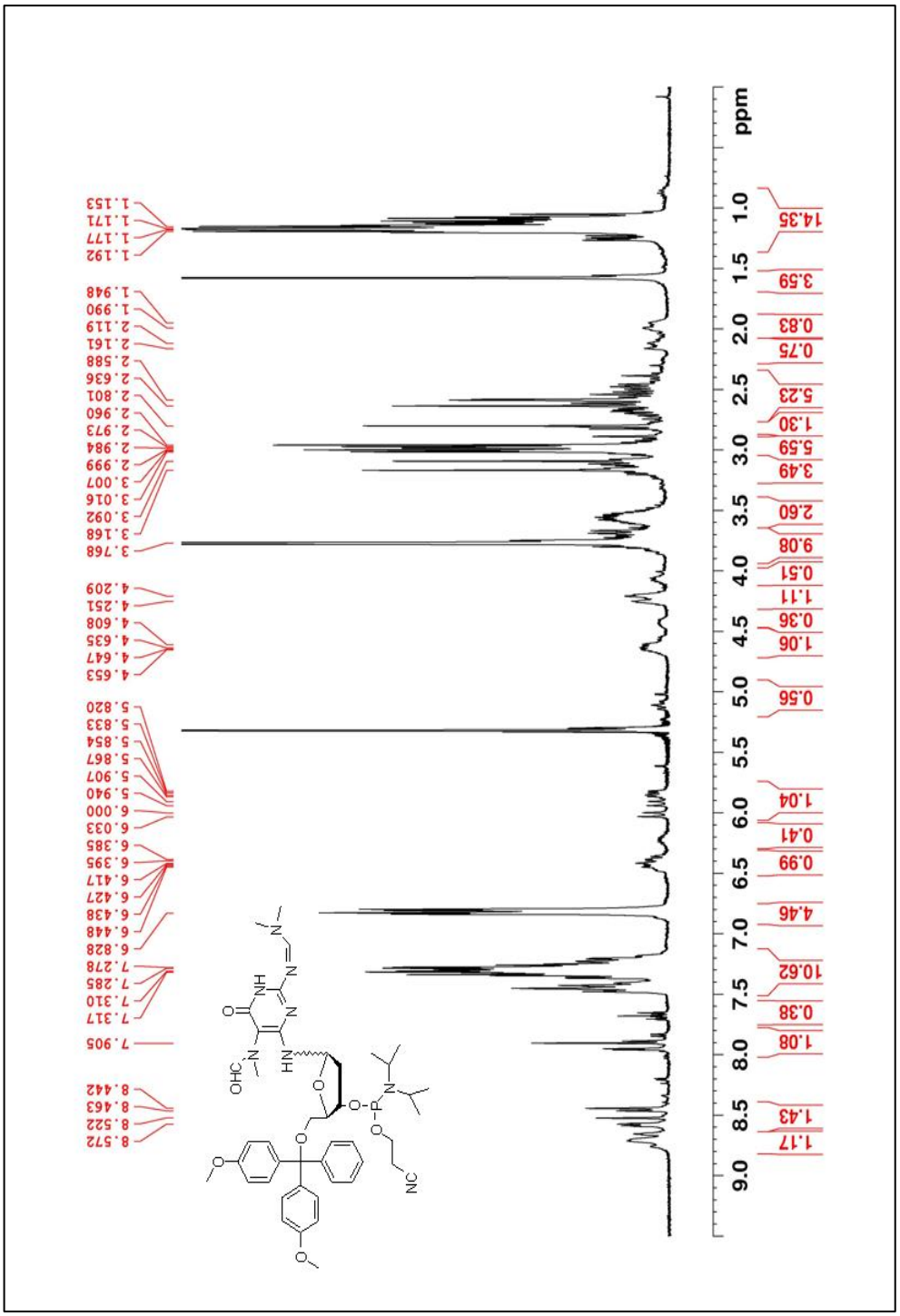
# for Chapter II



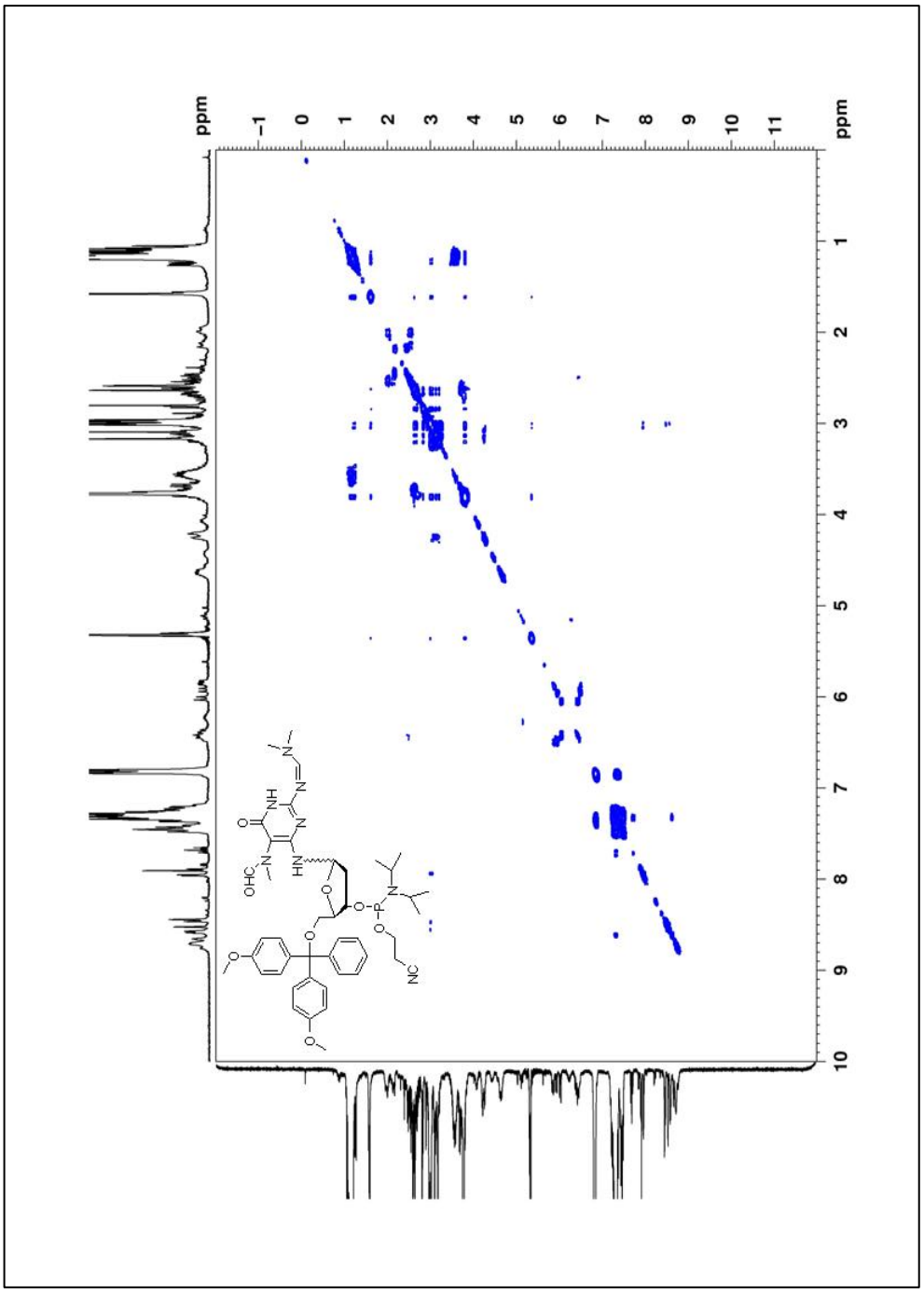
**$^{1}$H NMR of N-3-[(dimethylamino)methylene]-(5'-O-[bis(4-methoxy phenyl)phenylmethyl-2'-deoxy-β-D-*erythro*-pentofyranosyl)amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-N-(methyl)-formamide (2.35)**



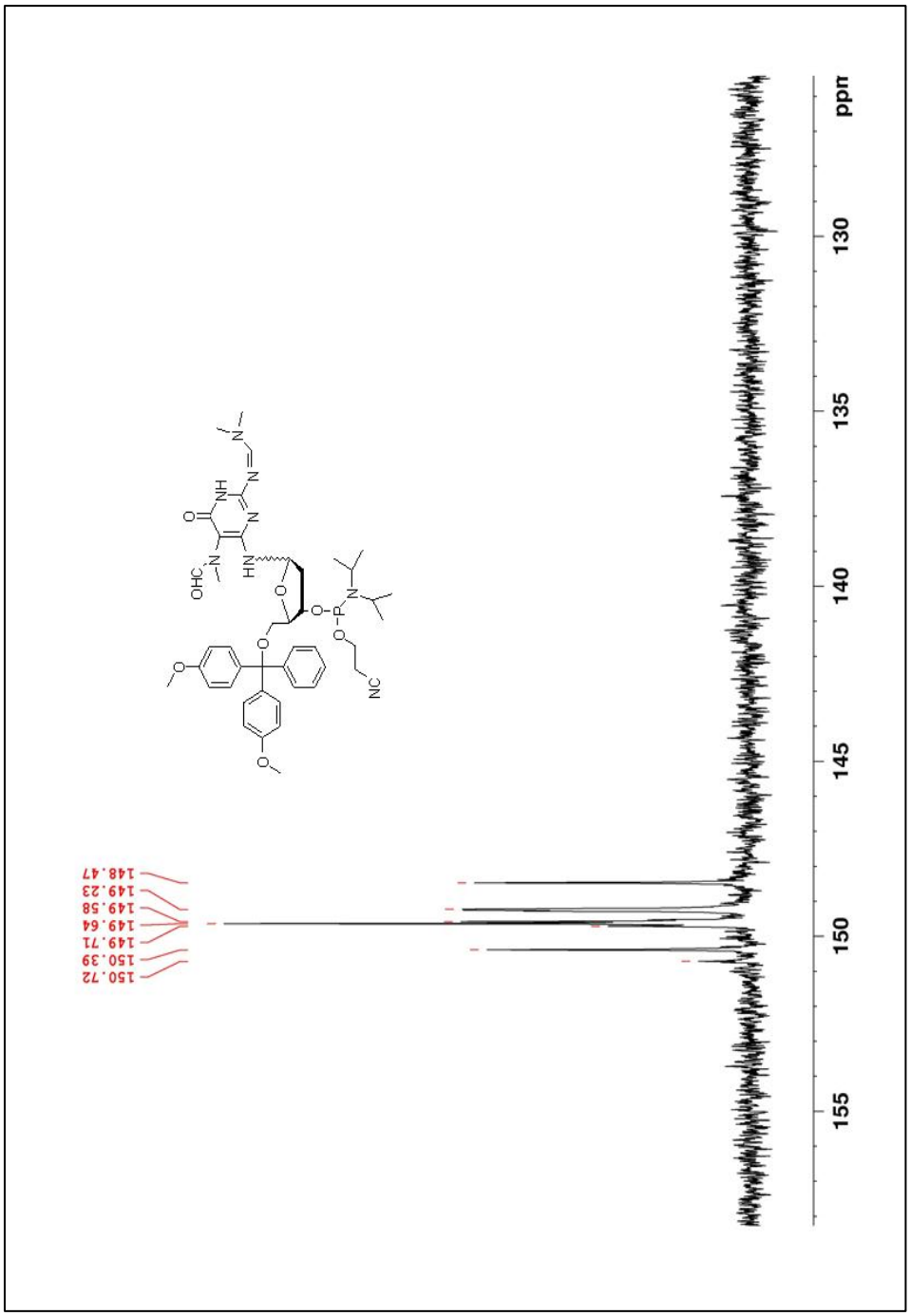

**COSY NMR of N-3-[(dimethylamino)methylene]-(5'-O-[bis(4-methoxyphenyl)-phenylmethyl-2'-deoxy-β-D-*erythro*-pento fyranosyl)-amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-N-(methyl)-formamide (2.35)**



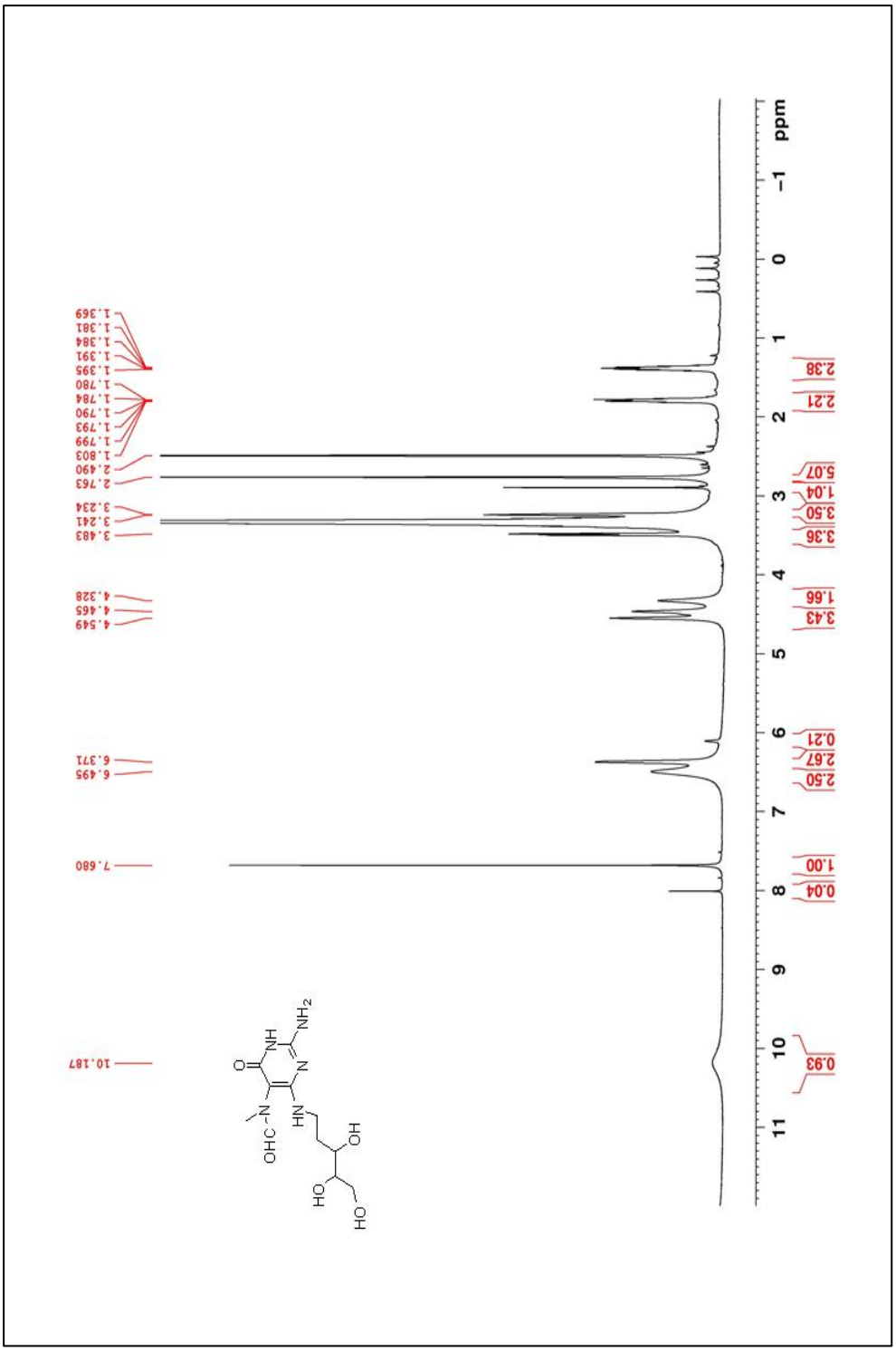
**$^{1}$H NMR of N-3-[(dimethylamino)methylene]-3'O-[(2-cyanoethyl)-(N,N-diisopropyl)phosphoroamidite-(5'-O-[bis(4-methoxyphenyl)-phenylmethyl-2'-deoxy-β-D-*erythro*-pentofyranosyl)amino]-3,4-dihydro-4-oxo-5-pyrimidi-nyl]-N-(methyl)-formamide (2.36)**



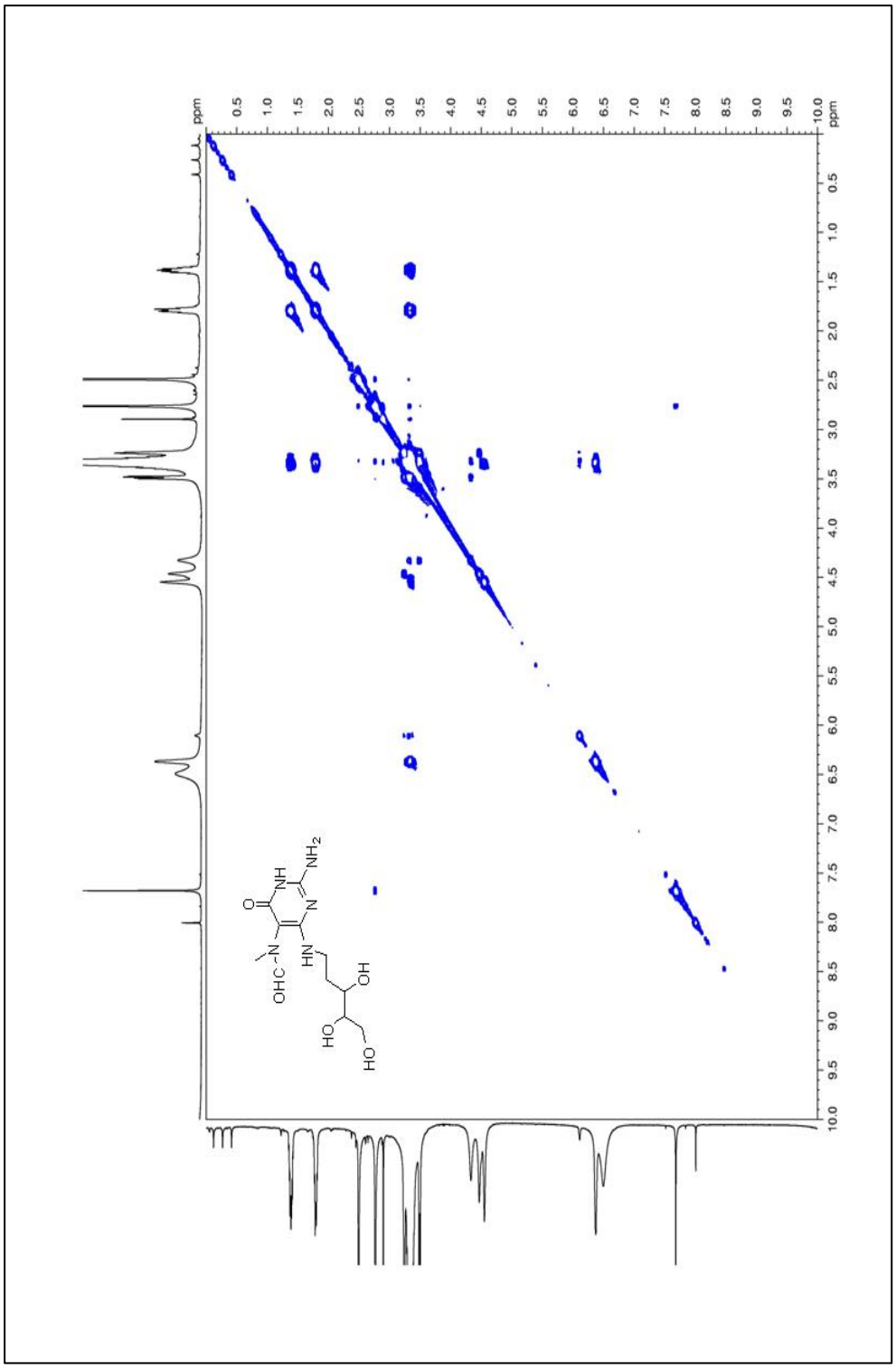

**COSY NMR of N-3-[(dimethylamino)methylene]-3'O-[(2-cyanoethyl)-(N,N-diisopropyl)phosphoroamidite-(5'-O-[bis(4-methoxyphenyl)phenylmethyl-2'-deoxy-β-D-*erythro*-pentofyranosyl)amino]-3,4-dihydro-4-oxo-5-pyrimidi-nyl]-N-(methyl)-formamide (2.36)**



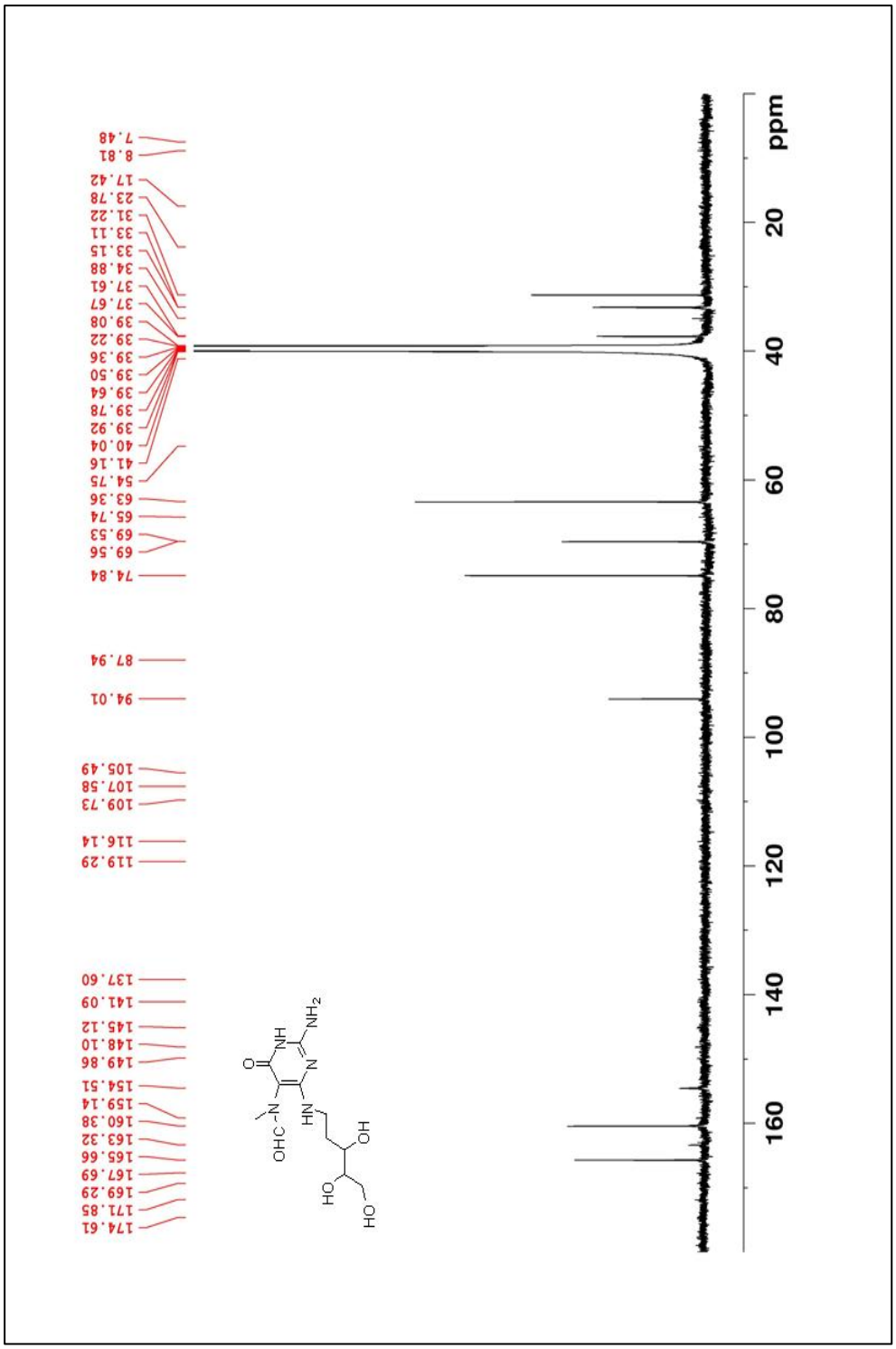

**$^{31}$P NMR of N-3-[(dimethylamino)methylene]-3'O-[(2-cyanoethyl)-(N,N-diisopropyl)phosphoroamidite-(5'-O-[bis(4-methoxy phenyl)phenylmethyl-2'-deoxy-β-D-*erythro*-pentofyranosyl)amino]-3,4-dihydro-4-oxo-5-pyrimidi-nyl]-N-(methyl)-formamide (2.36**




**$^{1}$H NMR of N-[(1,2-dideoxy-D-*erythro*-pentitol)2-amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-6-amino-N-(methyl)-formamide (2.50)**




**COSY NMR of N-[(1,2-dideoxy-D-*erythro*-pentitol)2-amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-6-amino-N-(methyl)-formamide (2.50)**

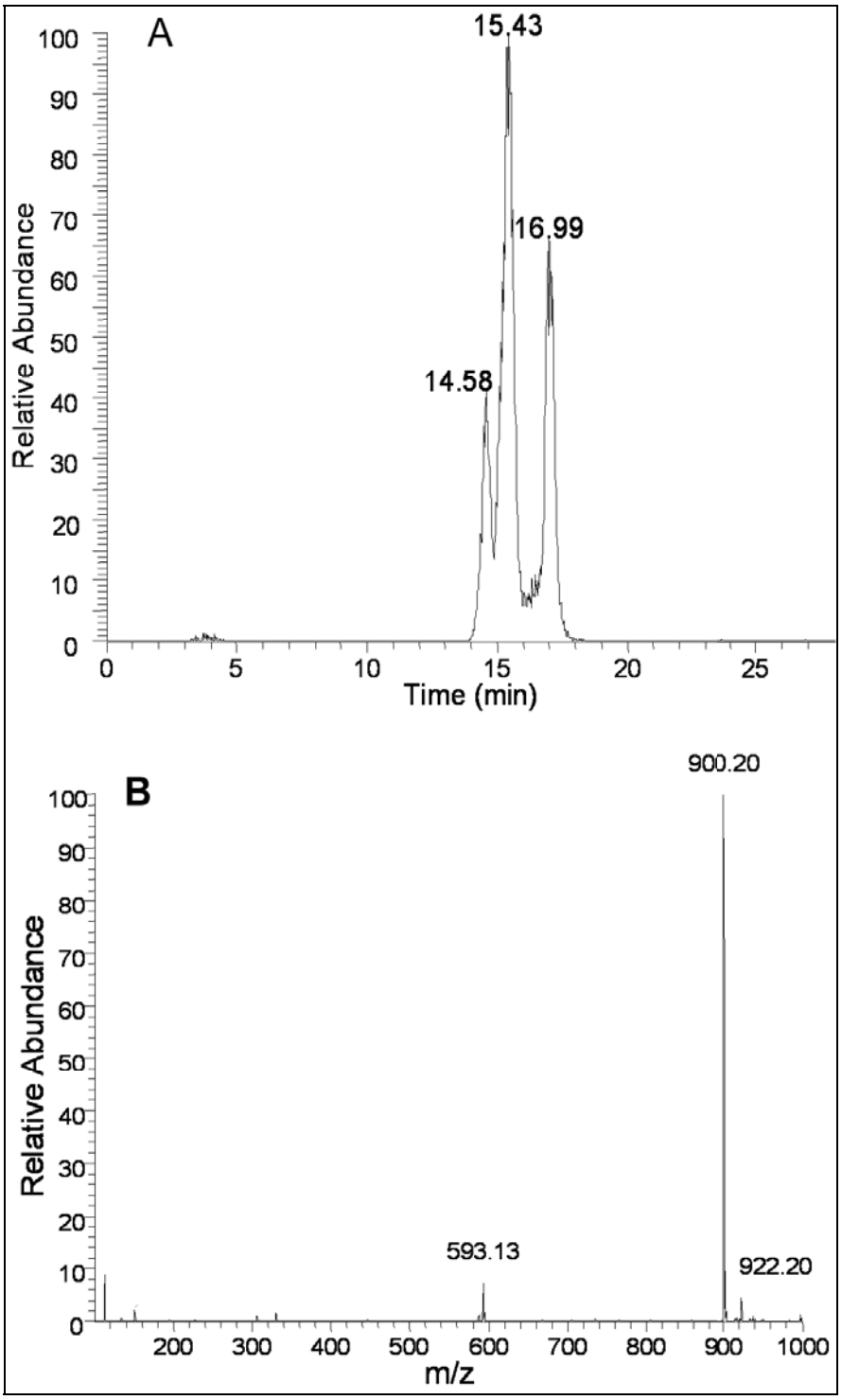



**$^{13}$C NMR of N-[(1,2-dideoxy-D-*erythro*-pentitol)2-amino]-3,4-dihydro-4-oxo-5-pyrimidinyl]-6-amino-N-(methyl)-formamide (2.50)**

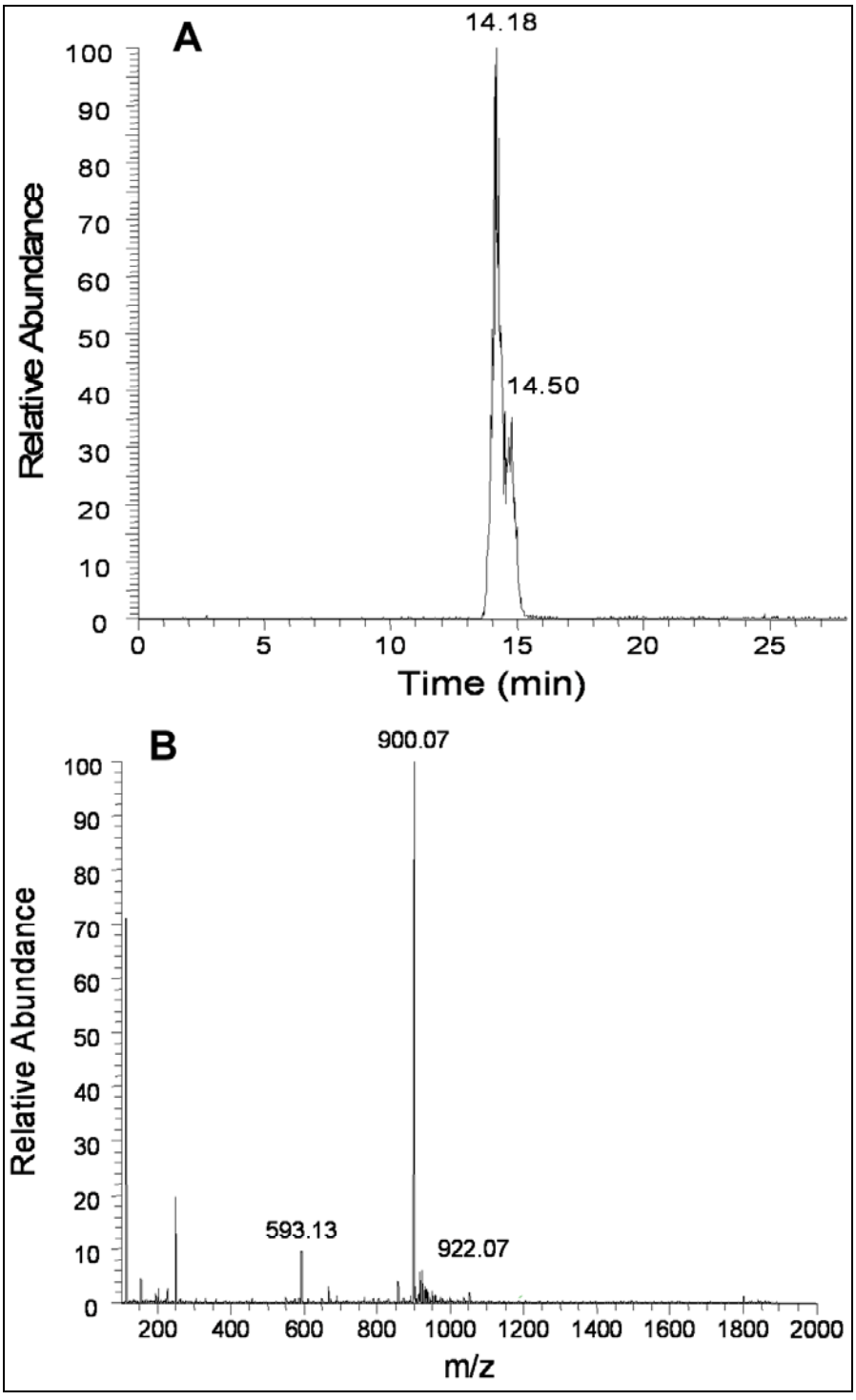
# Appendix II:

# MALDI-TOF Mass Spectra

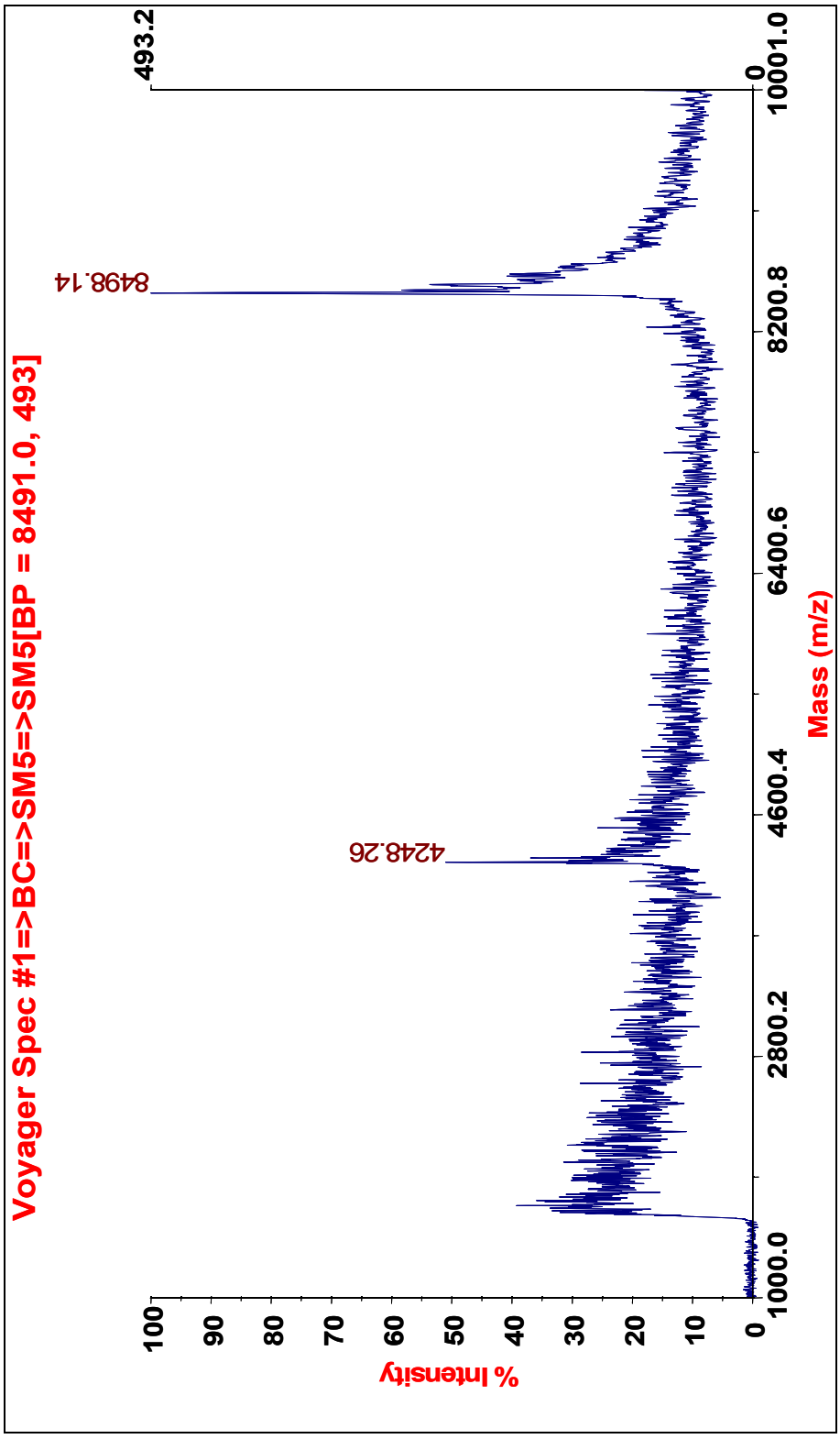
# for Chapter II



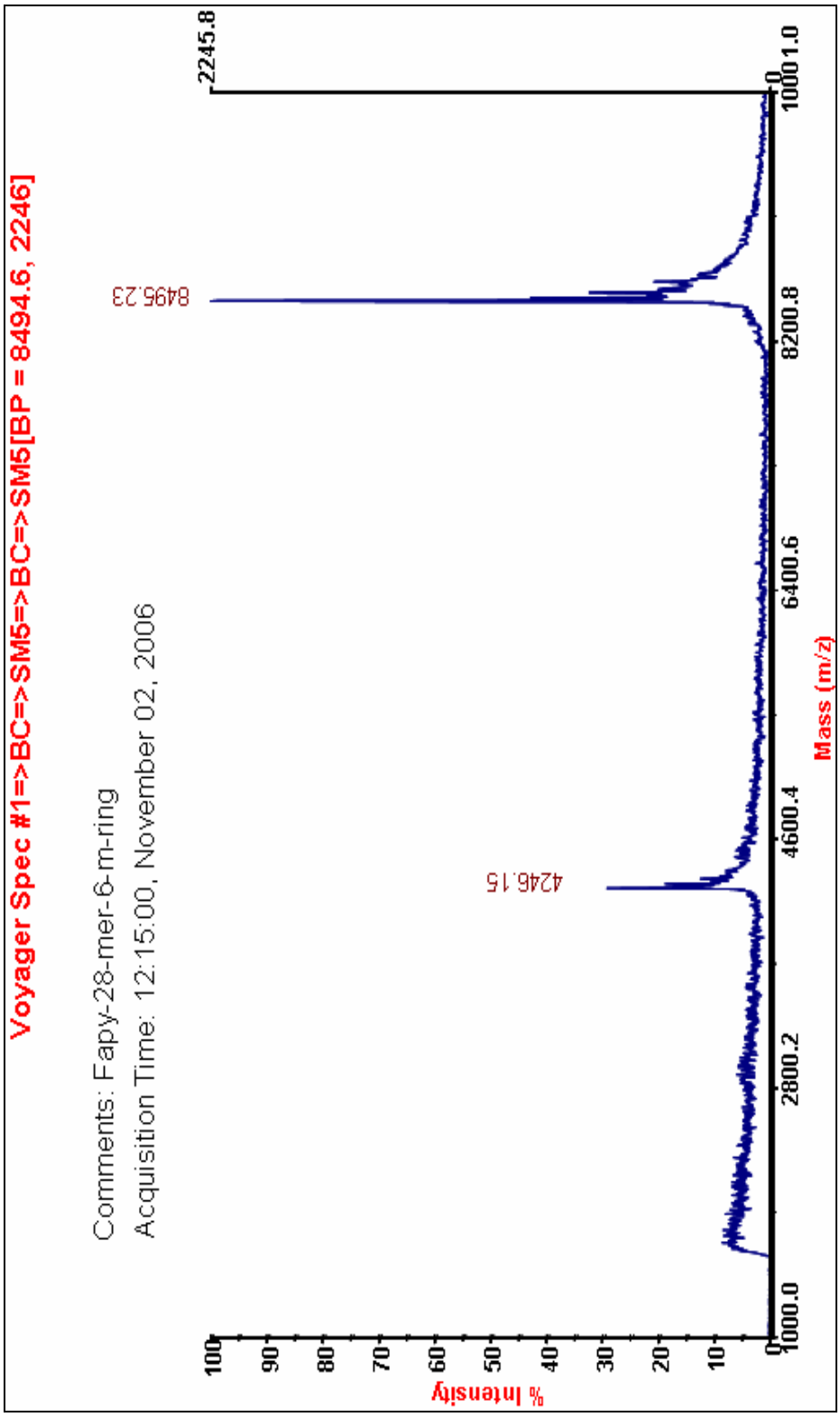

**LC-ESI/MS analysis of trinucleotide of 5'-A-(MeFAPy-dGuo)-C-3' (2.38a); A) Select ion profile of 2.38a with *m/z* 900.20; B) The molecular mass of the 2.38a.**



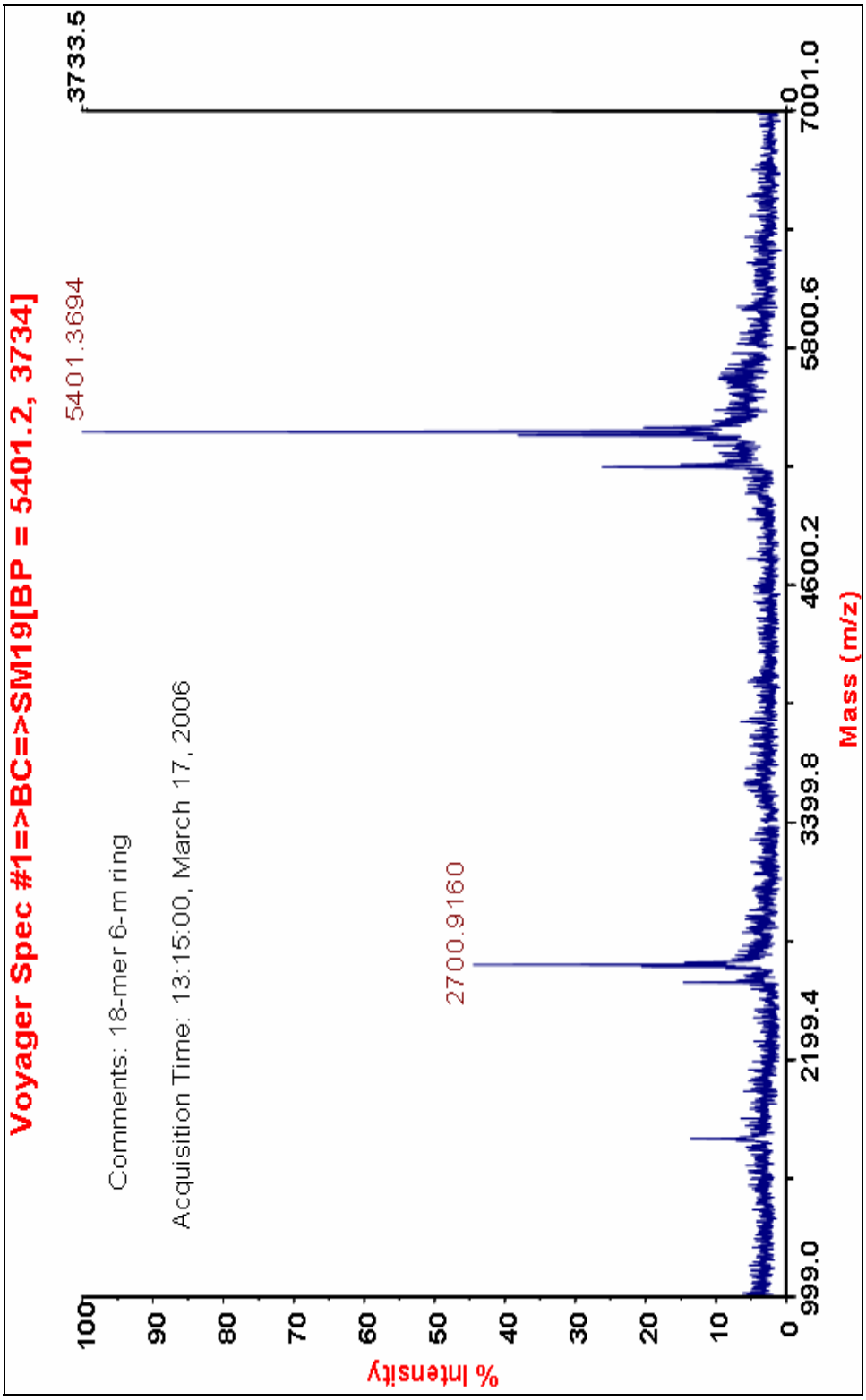

**LC-ESI/MS analysis of trinucleotide of 5'-A-(MeFAPy-dGuo)-C-3' (2.38b); A) Select ion profile of 2.388 with *m/z* 900.07; B) The molecular mass of the 2.38b.**



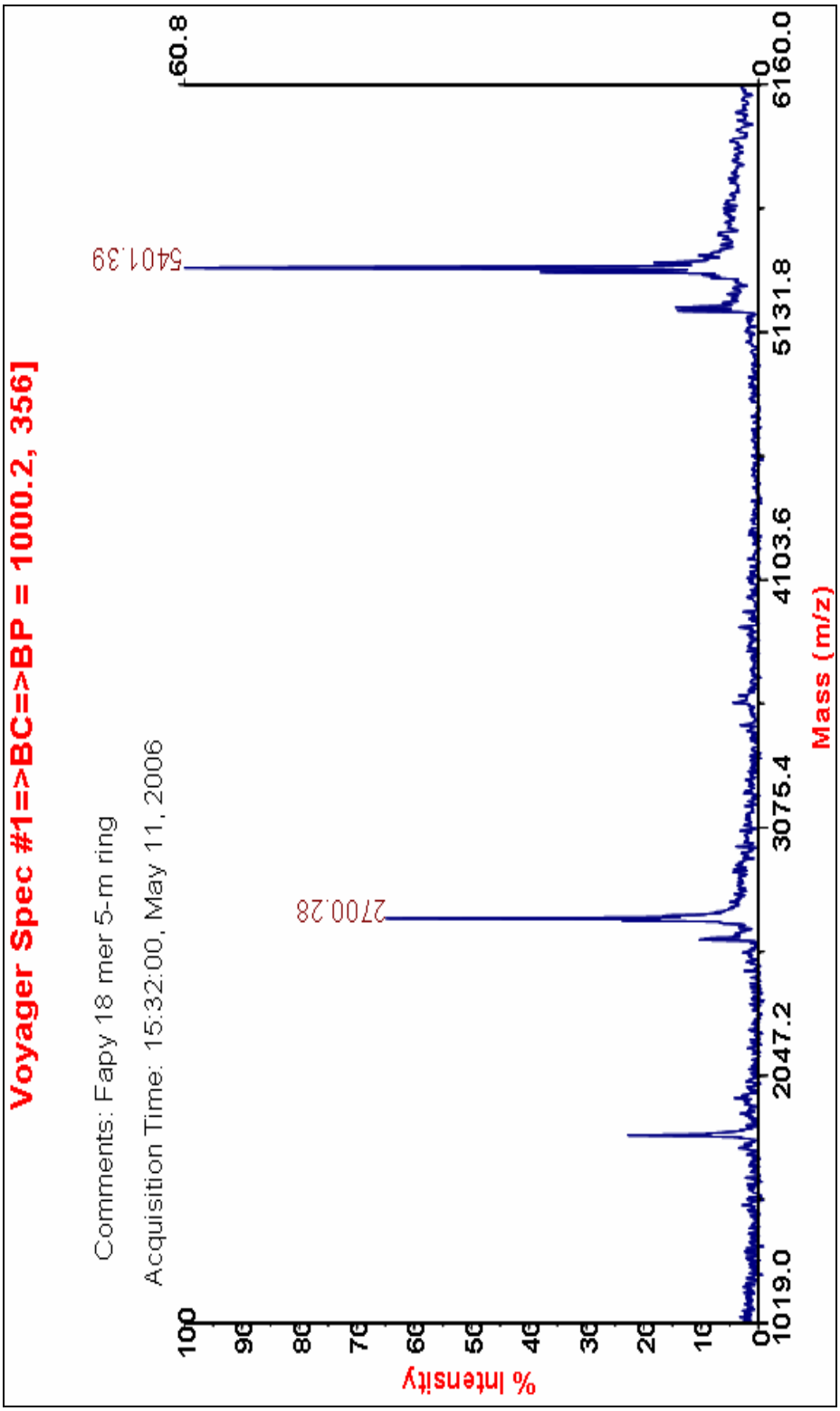

**5'-TCAT-(MeFAPy-dGuo)-GAATCCTTACGAGCATCGCCCCC-3'**
**(2.52a, furanose)**



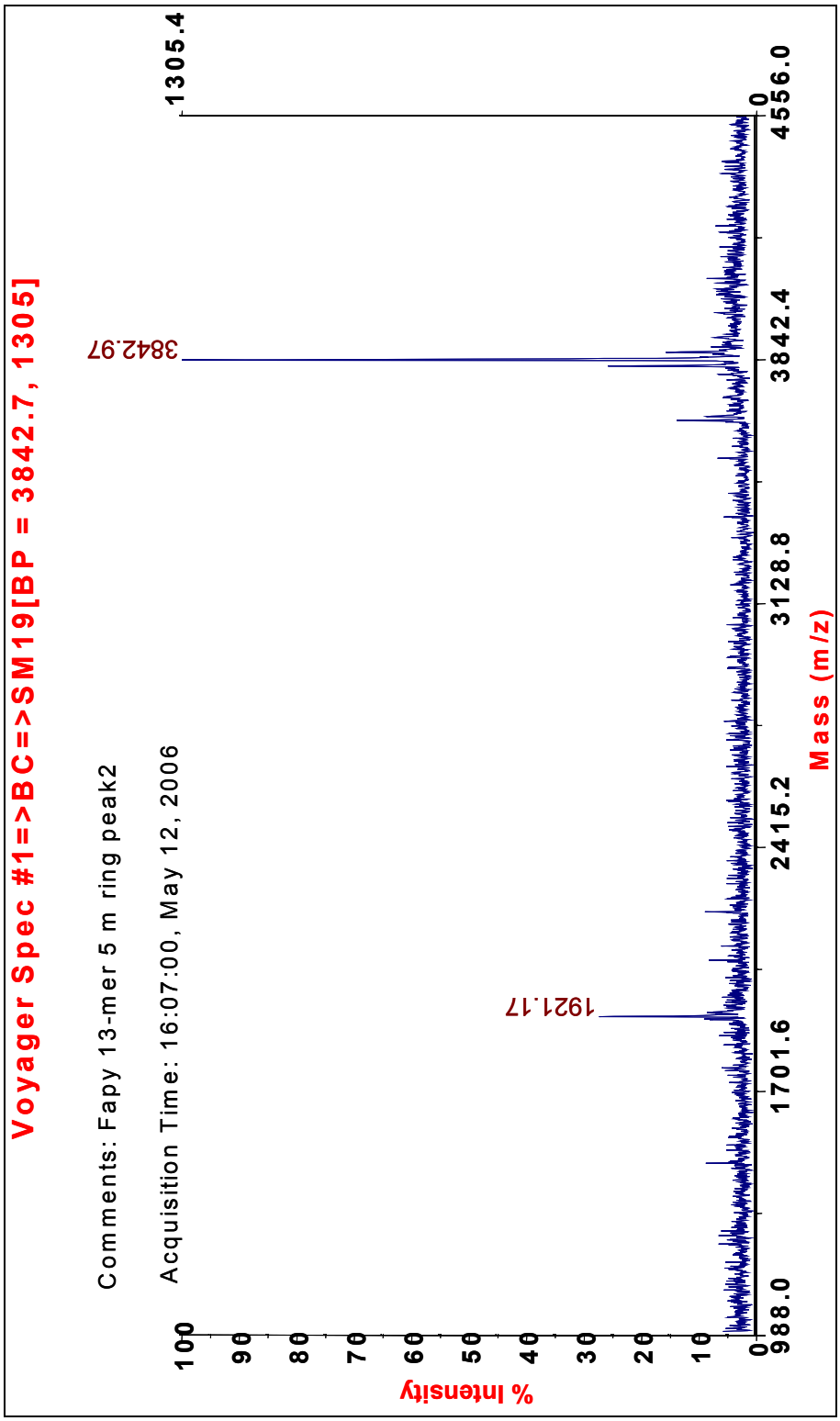

**5'-TCAT-(MeFAPy-dGuo)-GAATCCTTACGAGCATCGCCCCC-3'**
**(2.52b, pyranose)**



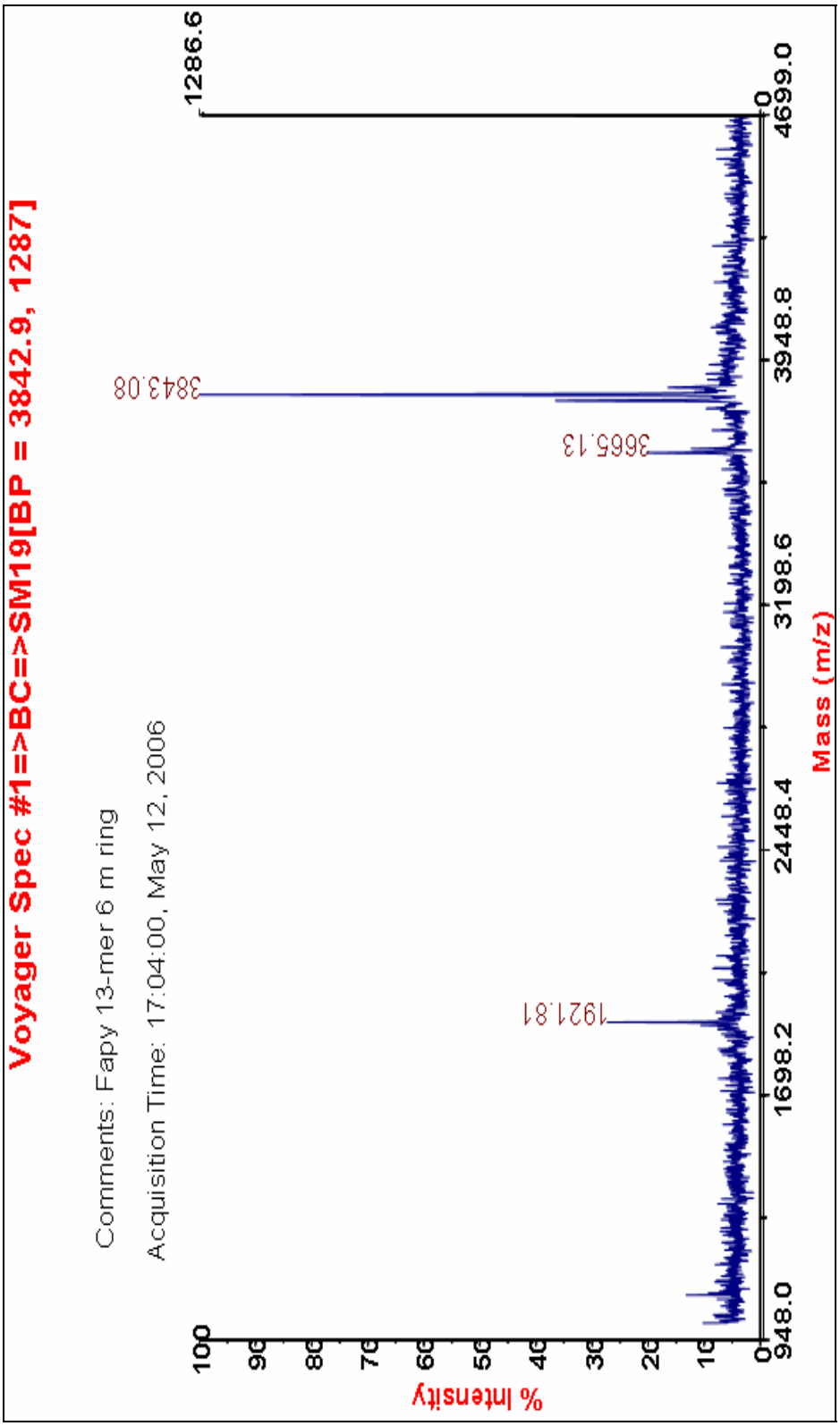

**5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCCCCC-3' pyranose**



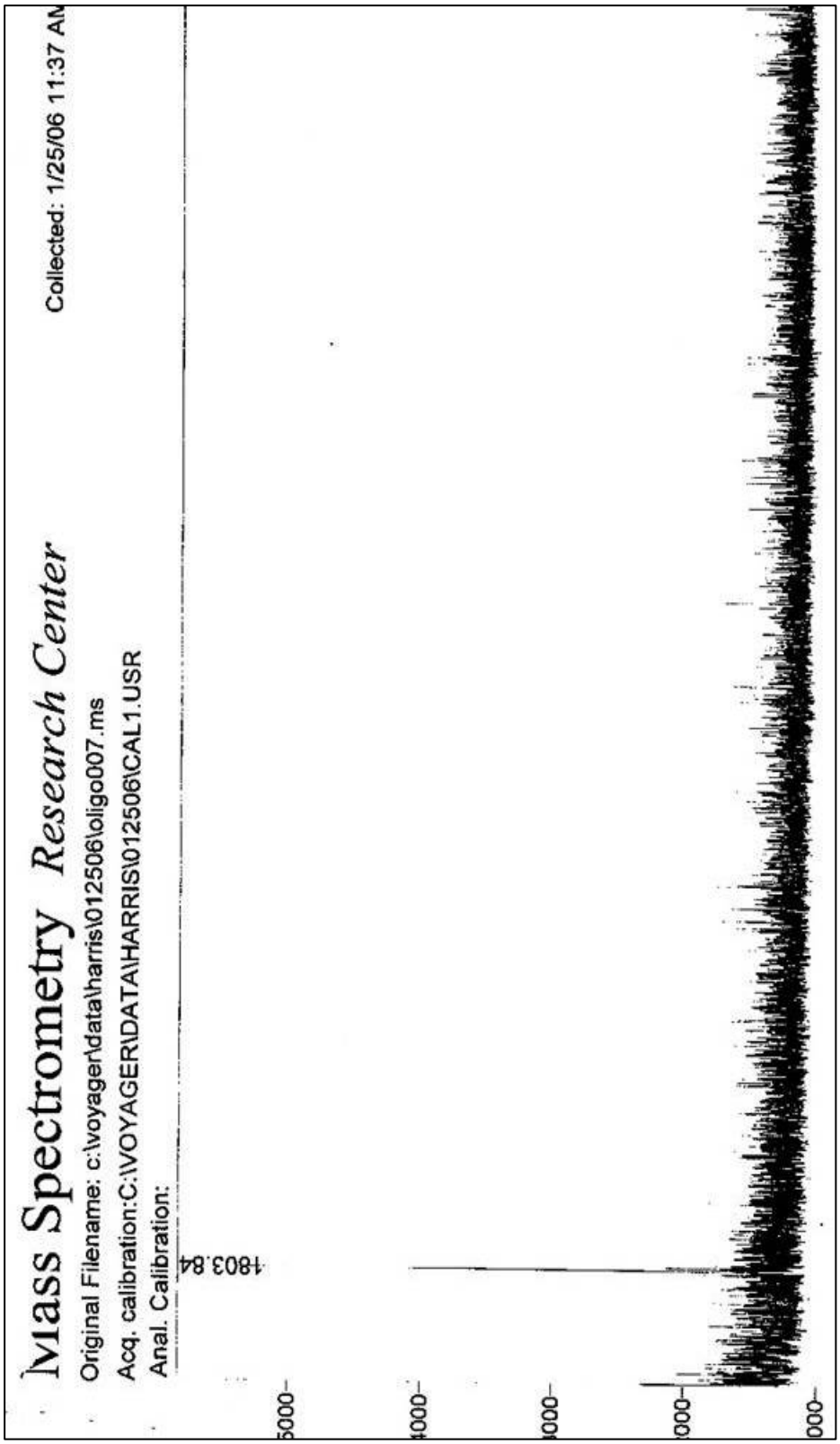

**5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCCCCC-3' furanose**



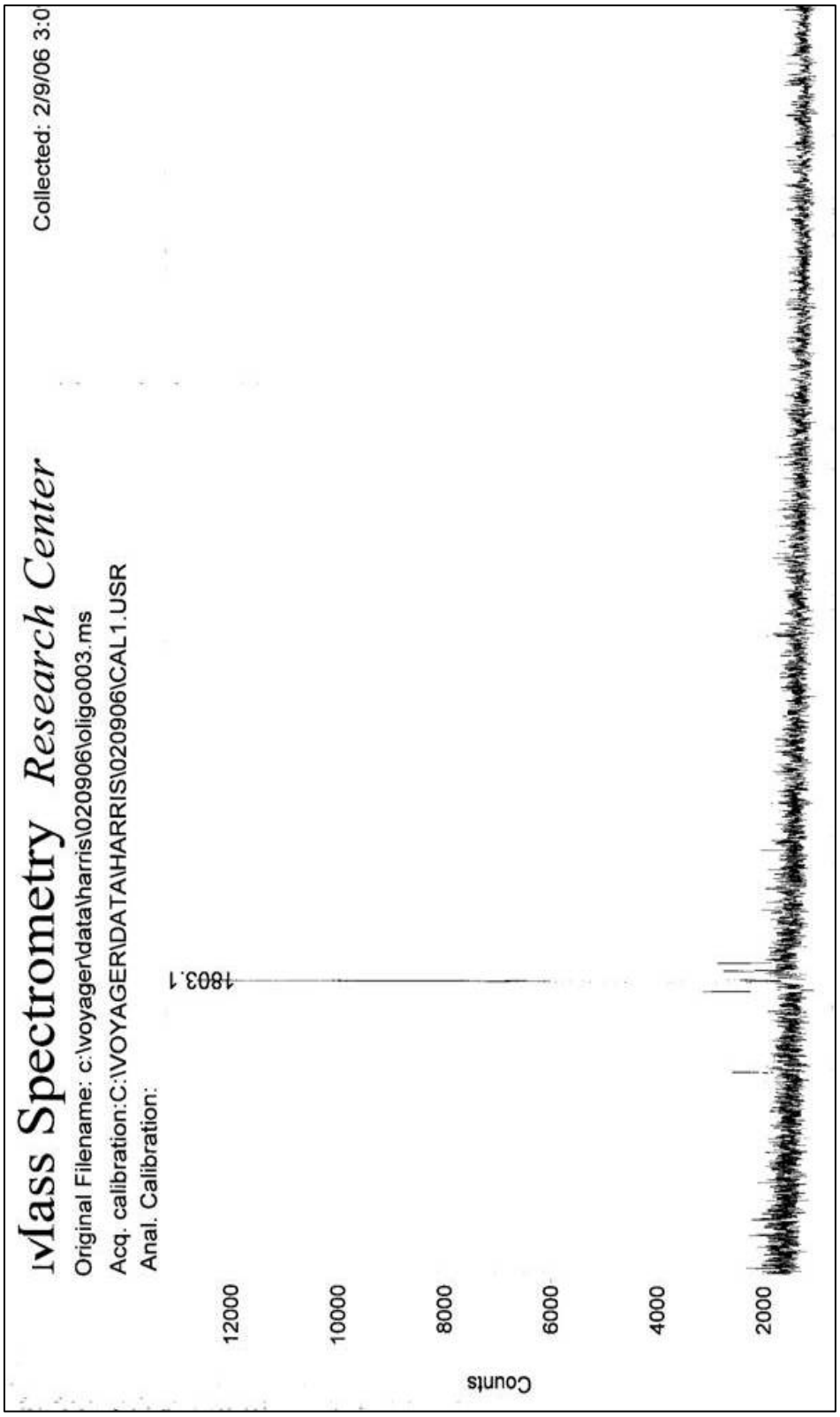

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.43a, furanose)**




**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.43b, pyranose)**




**5’-CTT-(MeFAPy-dGuo)-TT-3’ (2.37a, furanose)**

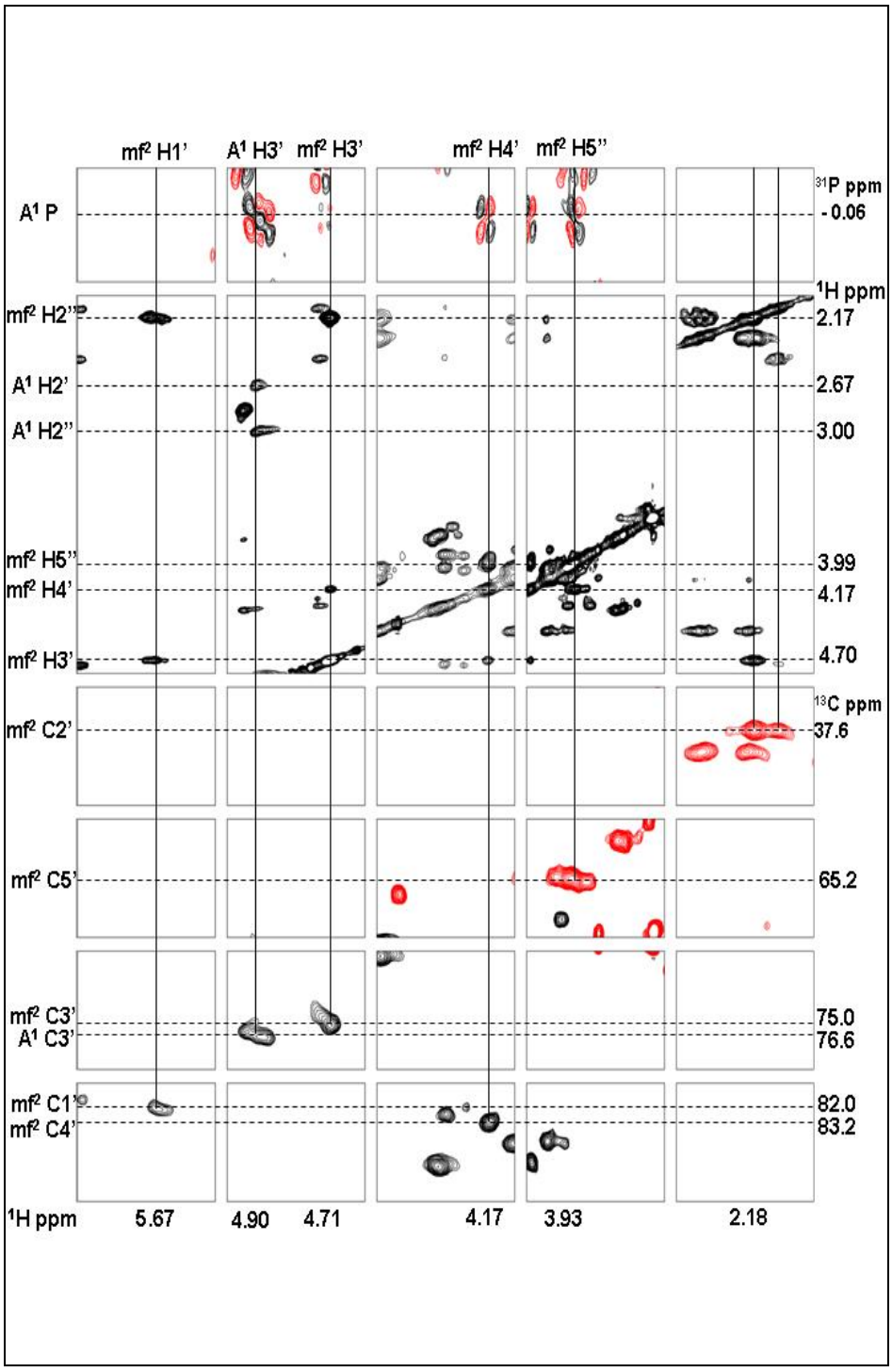



**5’-CTT-(MeFAPy-dGuo)-TT-3’ (2.37b, pyranose)**

# Appendix III

## NMR spectra of 5′-A-(MeFAPy-dGuo)-C-3′

## for Chapter II




**NMR of 5′-A-(MeFAPy-dGuo)-C-3′ (2.38a, furanose)**

| Group | Atom | Nuc | Shift | SDev | Assignments |
|---|---|---|---|---|---|
| 1A1 | 31P | 31P | -0.065 | 0.005 | 2 |
| 1A1 | C1' | 13C | 86.027 | 0.000 | 1 |
| 1A1 | C2' | 13C | 37.367 | 0.008 | 2 |
| 1A1 | C3' | 13C | 77.114 | 0.000 | 1 |
| 1A1 | C4' | 13C | 86.926 | 0.000 | 1 |
| 1A1 | C5' | 13C | 61.866 | 0.000 | 1 |
| 1A1 | H1' | 1H | 6.350 | 0.004 | 4 |
| 1A1 | H2' | 1H | 2.678 | 0.005 | 3 |
| 1A1 | H2'' | 1H | 3.006 | 0.008 | 3 |
| 1A1 | H3' | 1H | 4.909 | 0.005 | 6 |
| 1A1 | H4' | 1H | 4.324 | 0.003 | 3 |
| 1A1 | H5' | 1H | 3.780 | 0.002 | 2 |
| 1C3 | C1' | 13C | 85.419 | 0.000 | 1 |
| 1C3 | C2' | 13C | 39.350 | 0.003 | 2 |
| 1C3 | C3' | 13C | 70.370 | 0.000 | 1 |
| 1C3 | C4' | 13C | 85.013 | 0.000 | 1 |
| 1C3 | C5' | 13C | 69.464 | 0.000 | 1 |
| 1C3 | H1' | 1H | 6.220 | 0.003 | 4 |
| 1C3 | H2' | 1H | 2.192 | 0.005 | 3 |
| 1C3 | H2'' | 1H | 2.328 | 0.006 | 3 |
| 1C3 | H3' | 1H | 4.490 | 0.002 | 5 |
| 1C3 | H4' | 1H | 4.104 | 0.006 | 3 |
| 1C3 | H5 | 1H | 5.958 | 0.000 | 1 |
| 1C3 | H5' | 1H | 3.888 | 0.003 | 2 |
| 1C3 | H6 | 1H | 7.776 | 0.000 | 1 |
| 1mf2 | 31P | 31P | -0.187 | 0.000 | 1 |
| 1mf2 | C1' | 13C | 81.375 | 0.000 | 1 |
| 1mf2 | C2' | 13C | 37.543 | 0.012 | 2 |
| 1mf2 | C3' | 13C | 75.368 | 0.000 | 1 |
| 1mf2 | C4' | 13C | 81.998 | 0.000 | 1 |
| 1mf2 | H1' | 1H | 5.893 | 0.004 | 4 |
| 1mf2 | H2' | 1H | 2.106 | 0.003 | 3 |
| 1mf2 | H2'' | 1H | 2.471 | 0.008 | 3 |
| 1mf2 | H3' | 1H | 4.735 | 0.007 | 6 |
| 1mf2 | H4' | 1H | 4.247 | 0.001 | 4 |
| 1mf2 | H5' | 1H | 3.932 | 0.000 | 1 |
| 1mf2 | H5'' | 1H | 4.041 | 0.000 | 1 |
| 2A1 | 31P | 31P | -0.220 | 0.001 | 2 |
| 2A1 | C1' | 13C | 85.353 | 0.000 | 1 |
| 2A1 | C2' | 13C | 37.423 | 0.000 | 1 |
| 2A1 | C3' | 13C | 76.704 | 0.000 | 1 |
| 2A1 | C5' | 13C | 61.415 | 0.000 | 1 |
| 2A1 | H1' | 1H | 6.404 | 0.000 | 3 |
| 2A1 | H2'' | 1H | 2.857 | 0.002 | 3 |
| 2A1 | H3' | 1H | 4.948 | 0.001 | 6 |
| 2A1 | H4' | 1H | 4.334 | 0.001 | 2 |
| 2A1 | H5' | 1H | 3.810 | 0.000 | 3 |
| 2C3 | C1' | 13C | 85.305 | 0.000 | 1 |

| | | | | | |
|---|---|---|---|---|---|
| 2C3 | H1' | 1H | 6.143 | 0.002 | 4 |
| 2C3 | H2' | 1H | 2.186 | 0.000 | 1 |
| 2C3 | H2'' | 1H | 2.312 | 0.000 | 1 |
| 2C3 | H3' | 1H | 4.480 | 0.000 | 1 |
| 2mf2 | 31P | 31P | -0.646 | 0.009 | 2 |
| 2mf2 | C1' | 13C | 82.098 | 0.000 | 1 |
| 2mf2 | C2' | 13C | 37.547 | 0.000 | 1 |
| 2mf2 | C3' | 13C | 76.096 | 0.000 | 1 |
| 2mf2 | C4' | 13C | 83.276 | 0.000 | 1 |
| 2mf2 | C5' | 13C | 65.181 | 0.000 | 1 |
| 2mf2 | H1' | 1H | 5.677 | 0.003 | 3 |
| 2mf2 | H2'' | 1H | 2.175 | 0.001 | 3 |
| 2mf2 | H3' | 1H | 4.706 | 0.003 | 5 |
| 2mf2 | H4' | 1H | 4.179 | 0.001 | 4 |
| 2mf2 | H5' | 1H | 3.937 | 0.002 | 2 |
| 2mf2 | H5'' | 1H | 3.988 | 0.001 | 2 |
| 3C3 | H1' | 1H | 6.079 | 0.002 | 3 |
| 3C3 | H2' | 1H | 2.135 | 0.000 | 1 |
| 3C3 | H2'' | 1H | 2.345 | 0.000 | 1 |
| 3C3 | H3' | 1H | 4.480 | 0.000 | 2 |
| 3C3 | H4' | 1H | 4.022 | 0.000 | 1 |
| 3mf2 | C4' | 13C | 82.587 | 0.000 | 1 |
| 3mf2 | C5' | 13C | 64.902 | 0.000 | 1 |
| 3mf2 | H1' | 1H | 5.935 | 0.000 | 1 |
| 3mf2 | H3' | 1H | 4.724 | 0.009 | 2 |
| 3mf2 | H4' | 1H | 4.302 | 0.001 | 4 |
| 3mf2 | H5' | 1H | 3.923 | 0.000 | 1 |
| 3mf2 | H5'' | 1H | 4.012 | 0.005 | 3 |

**$^{13}$C NMR of 5′-A-(MeFAPy-dGuo)-C-3′ (2.38a, furanose)**

| Group | Atom | Nuc | Shift | SDev | Assignments |
|---|---|---|---|---|---|
| 1A1 | C1' | 13C | 84.863 | 0.000 | 1 |
| 1A1 | C2' | 13C | 38.136 | 0.004 | 2 |
| 1A1 | C3' | 13C | 76.258 | 0.000 | 1 |
| 1A1 | C4' | 13C | 86.703 | 0.000 | 1 |
| 1A1 | C5' | 13C | 61.633 | 0.000 | 1 |
| 1A1 | H1' | 1H | 6.431 | 0.002 | 4 |
| 1A1 | H2' | 1H | 2.742 | 0.007 | 3 |
| 1A1 | H2" | 1H | 2.830 | 0.005 | 3 |
| 1A1 | H3' | 1H | 4.941 | 0.007 | 7 |
| 1A1 | H4' | 1H | 4.385 | 0.000 | 1 |
| 1A1 | H5' | 1H | 3.777 | 0.001 | 2 |
| 2mf1 | C1' | 13C | 85.384 | 0.000 | 1 |
| 2mf1 | C2' | 13C | 39.652 | 0.039 | 2 |
| 2mf1 | C3' | 13C | 70.779 | 0.000 | 1 |
| 2mf1 | C4' | 13C | 68.230 | 0.000 | 1 |
| 2mf1 | C5' | 13C | 66.119 | 0.002 | 2 |
| 2mf1 | H1' | 1H | 5.914 | 0.003 | 5 |
| 2mf1 | H2' | 1H | 2.038 | 0.001 | 3 |
| 2mf1 | H2" | 1H | 2.278 | 0.006 | 4 |
| 2mf1 | H3' | 1H | 4.343 | 0.003 | 6 |
| 2mf1 | H4' | 1H | 4.031 | 0.008 | 5 |
| 2mf1 | H5' | 1H | 3.679 | 0.000 | 2 |
| 2mf1 | H5" | 1H | 4.079 | 0.002 | 2 |
| 2mf2 | C1' | 13C | 74.637 | 0.000 | 1 |
| 2mf2 | C2' | 13C | 34.679 | 0.007 | 2 |
| 2mf2 | C3' | 13C | 70.975 | 0.000 | 1 |
| 2mf2 | C4' | 13C | 69.687 | 0.000 | 1 |
| 2mf2 | C5' | 13C | 69.429 | 0.000 | 1 |
| 2mf2 | H1' | 1H | 5.535 | 0.002 | 4 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 3 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 4 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 5 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 6 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 7 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 8 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 9 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 10 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 11 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 12 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 13 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 14 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 15 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 16 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 17 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 18 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 19 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 20 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 21 |

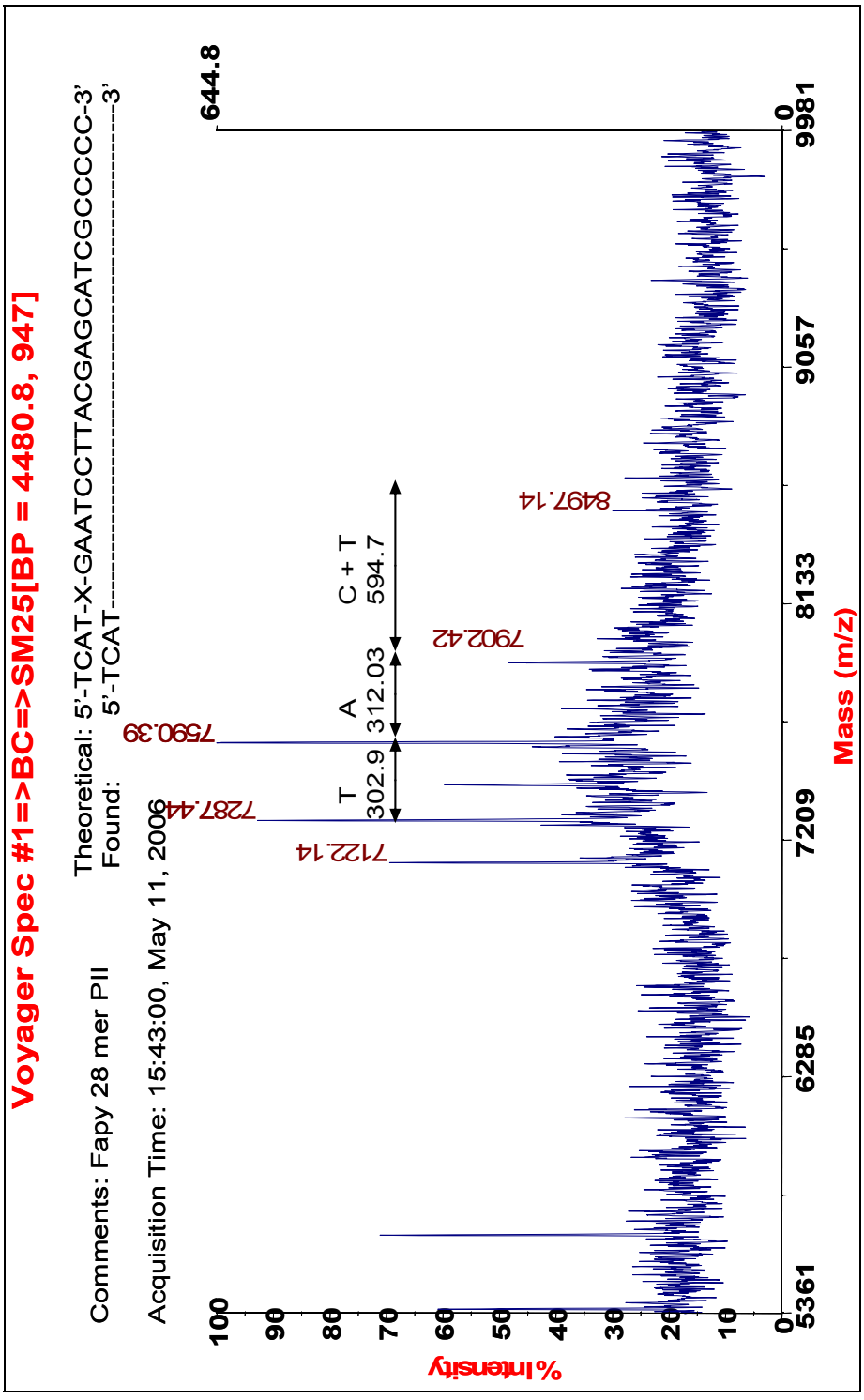
| | | | | | |
|---|---|---|---|---|---|
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 22 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 23 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 24 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 25 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 26 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 27 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 28 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 29 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 30 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 31 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 32 |
| 2mf2 | H2' | 1H | 2.027 | 0.003 | 33 |

**$^{13}$C NMR of 5′-A-(MeFAPy-dGuo)-C-3′ (2.38b, furanose)**

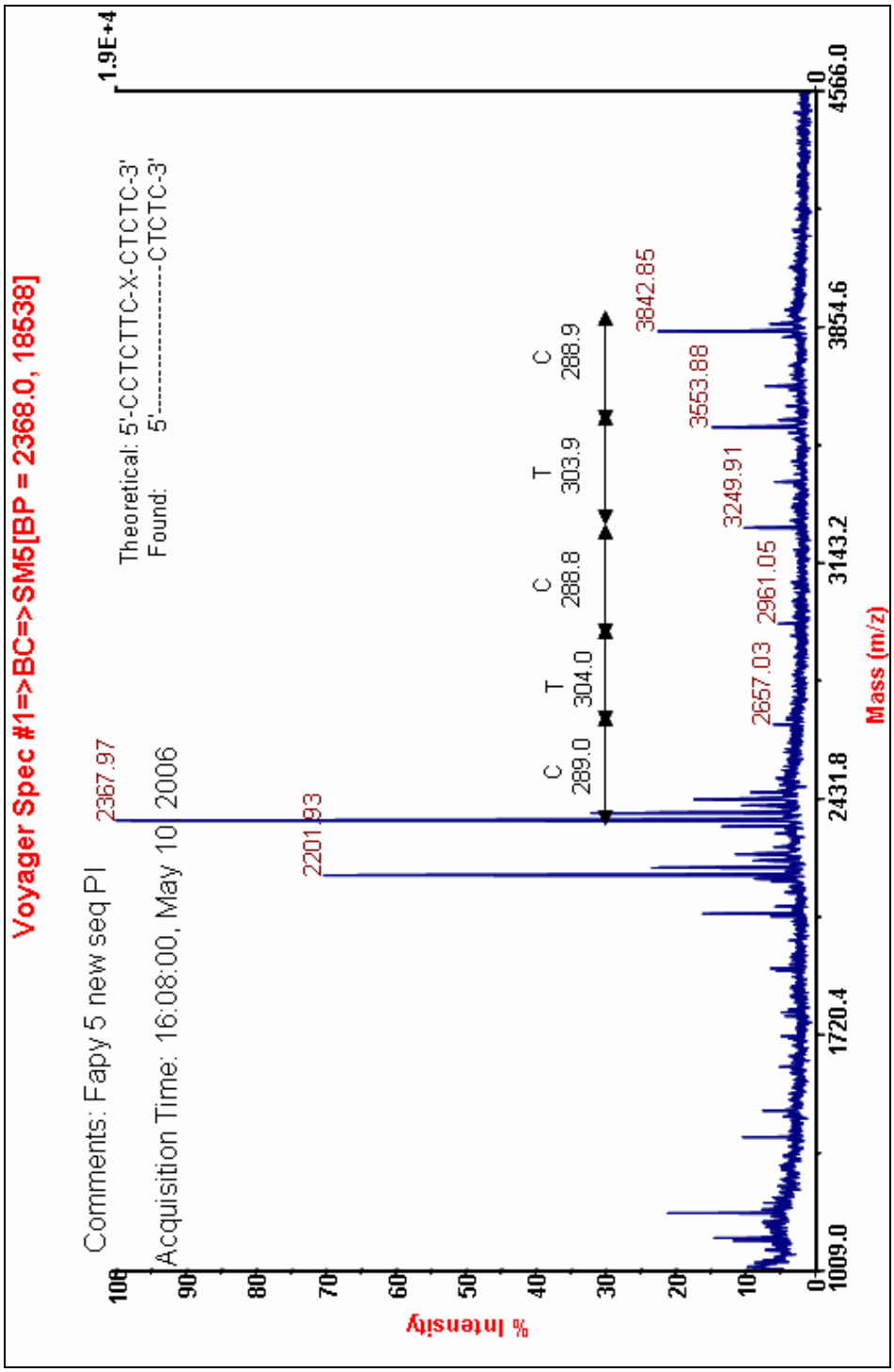
# Appendix IV

# MALDI-TOF Sequencing

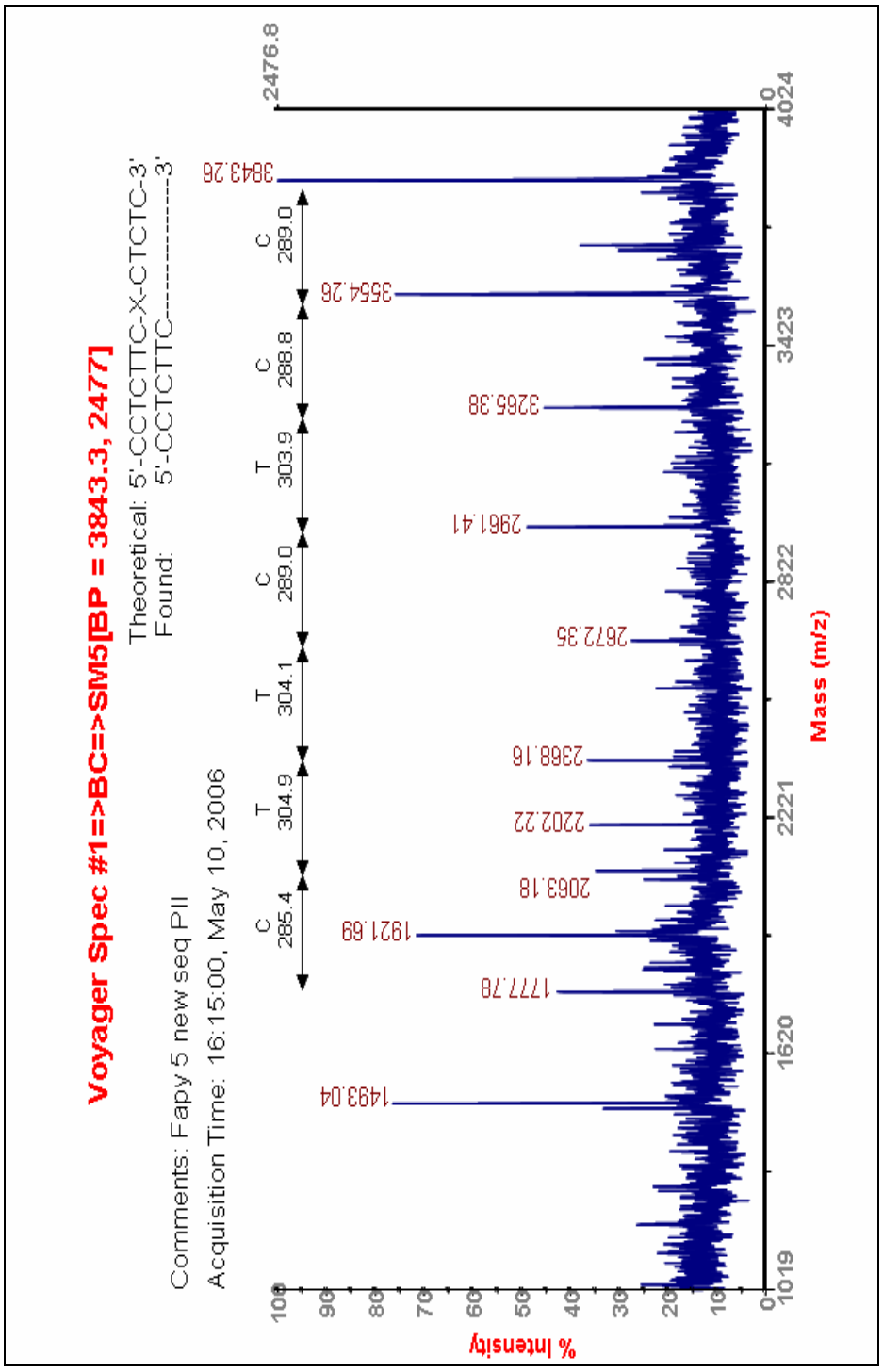
# for Chapter II



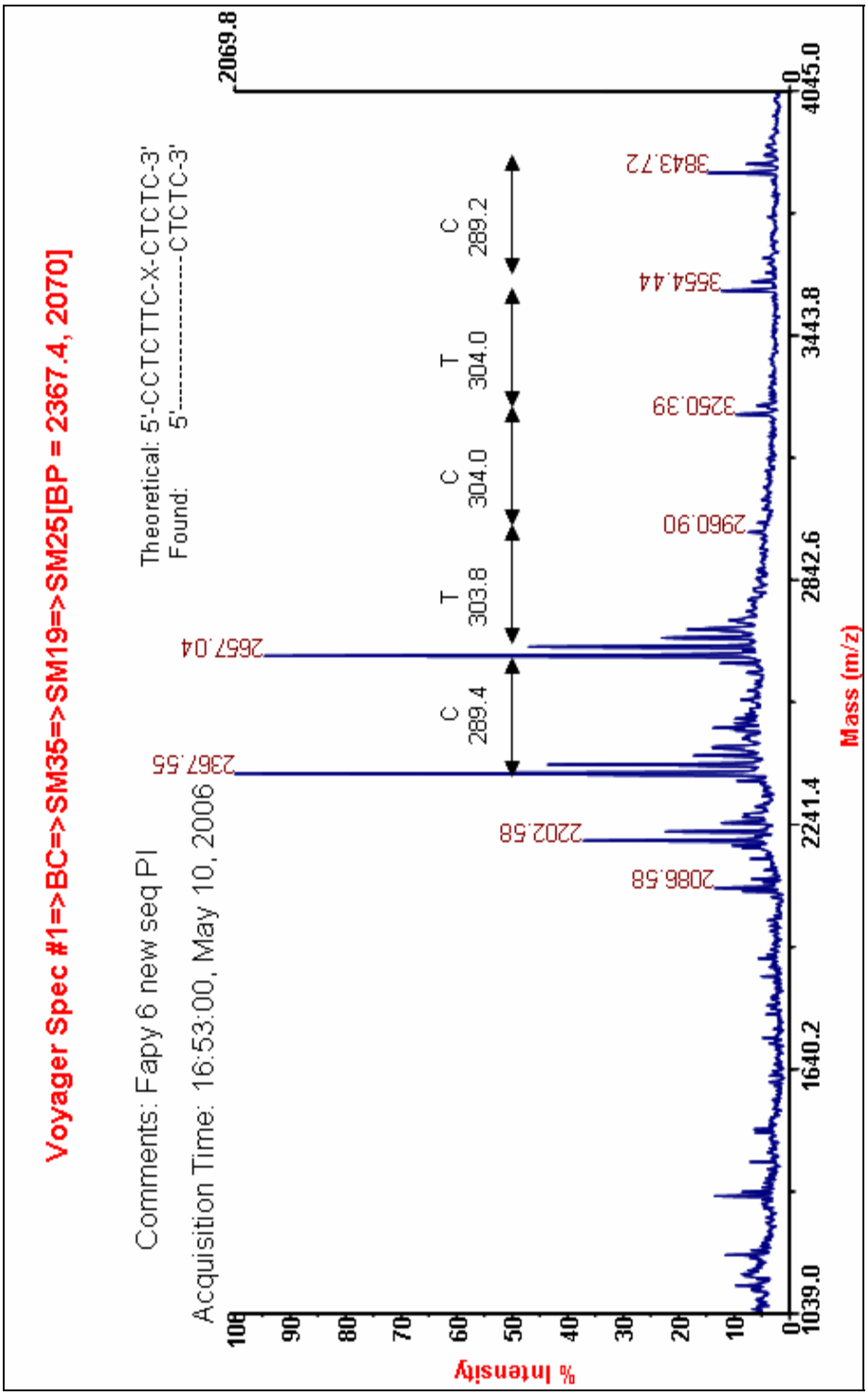

**5'-TCAT-(MeFAPy-dGuo)-GAATCCTTACGAGCATCGCCCCC-3'**
**(2.52a, furanose) with Phosphodiesterase II**



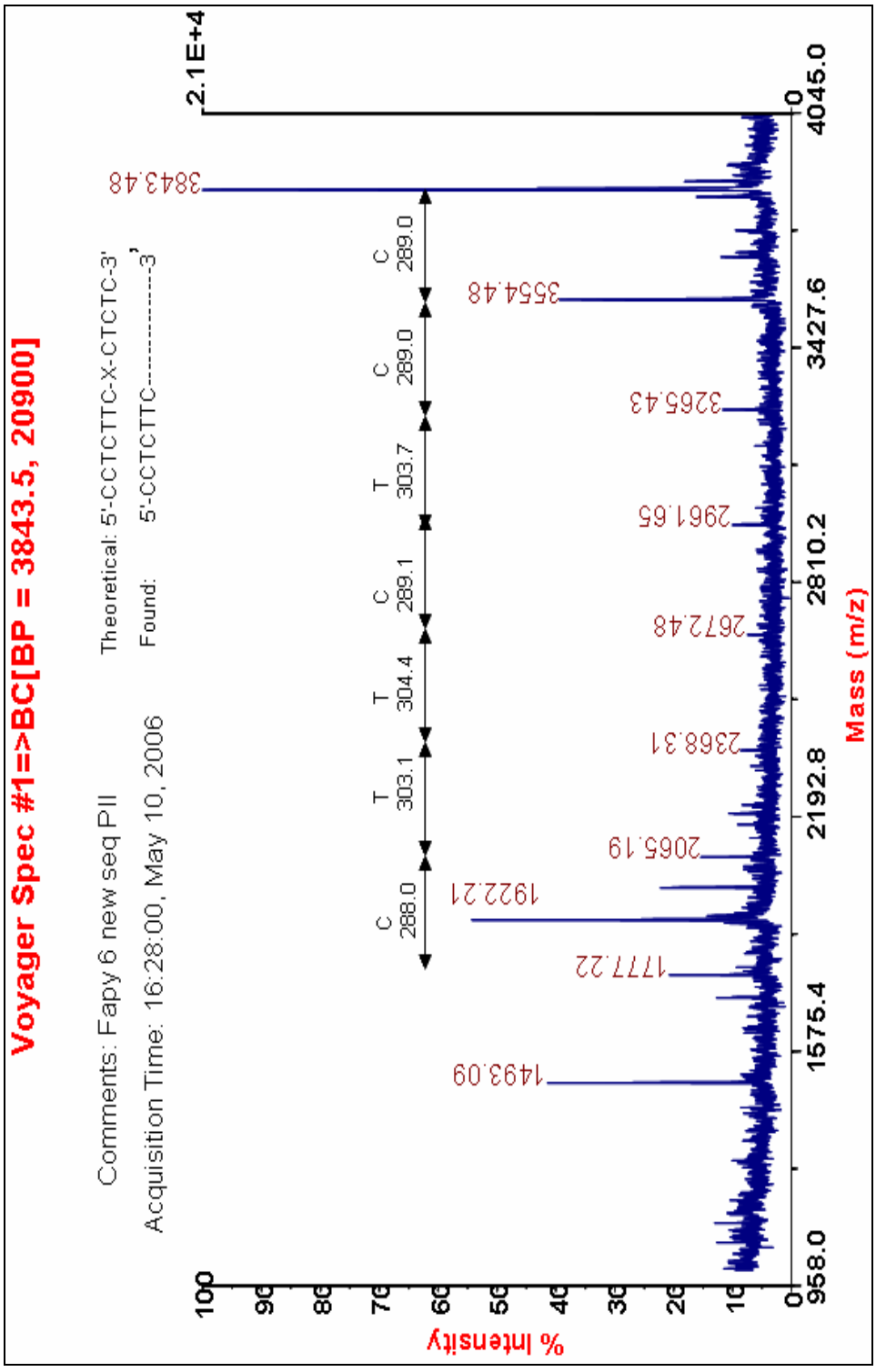

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' furanose with Phosphodiesterase I**



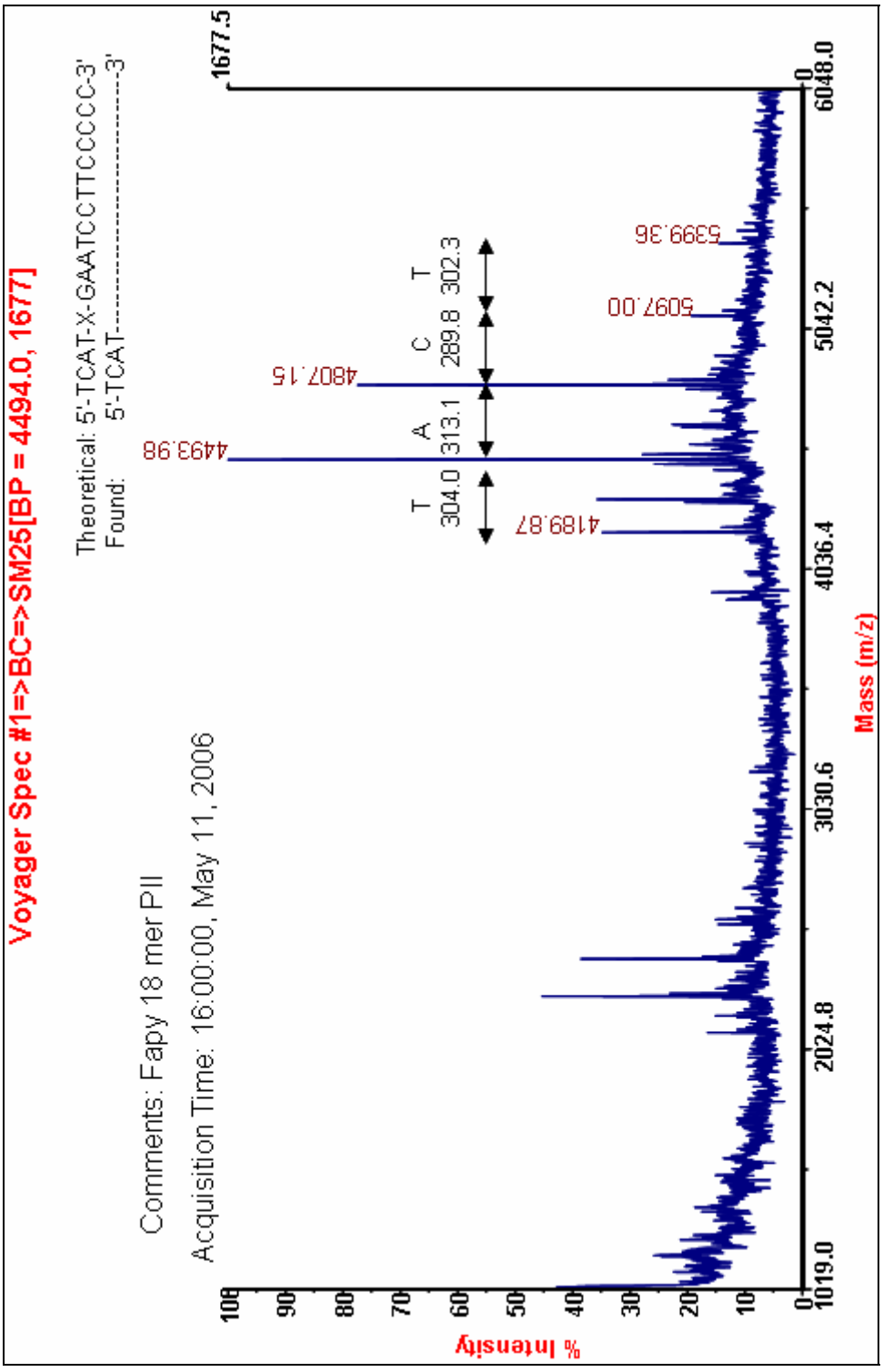

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' furanose with Phosphodiesterase II**



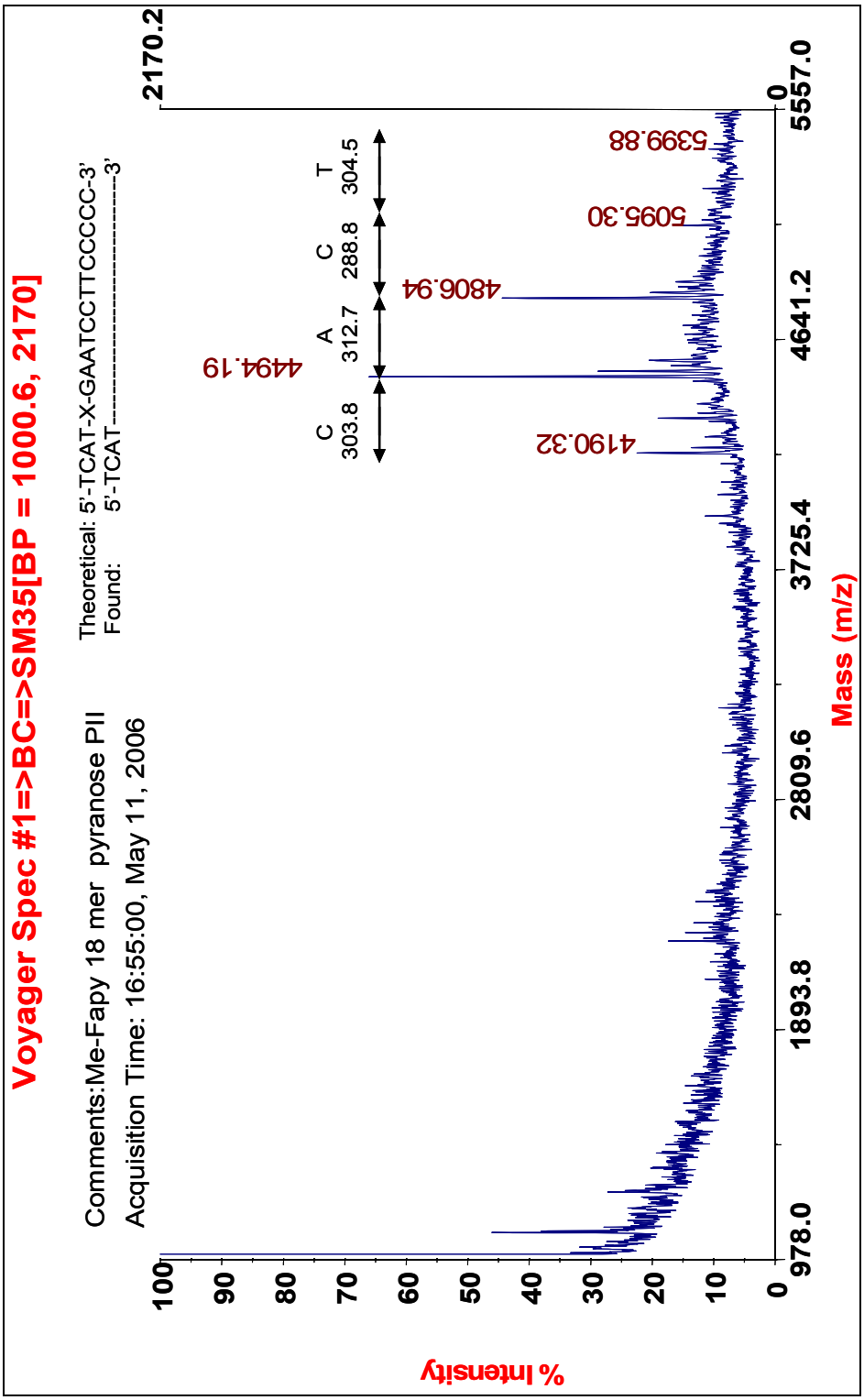

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) with Phosphodiesterase I**




**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.43b ,pyranose) with Phosphodiesterase II**



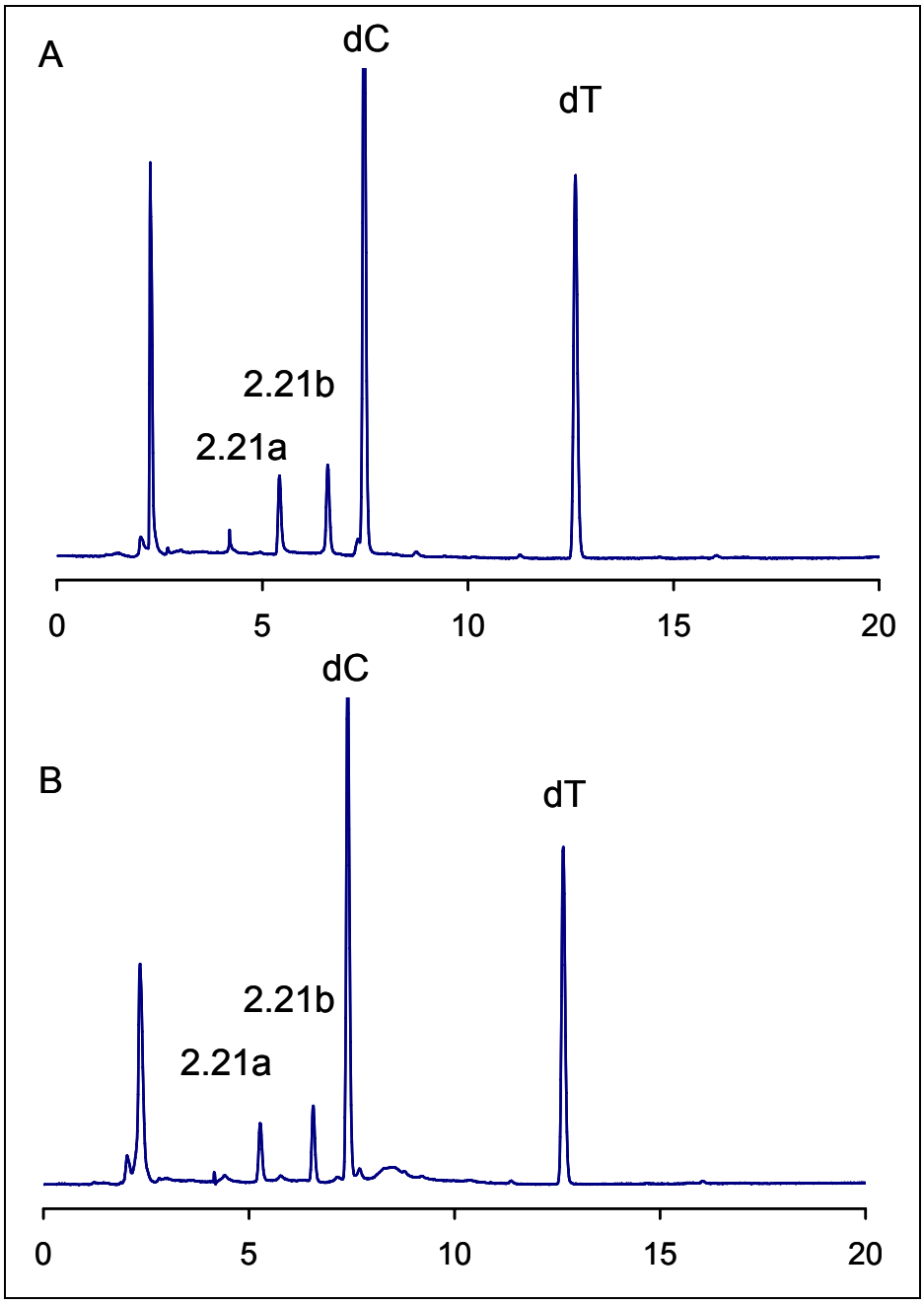

**5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCCCCC-3' (furanose) with Phosphodiesterase II**



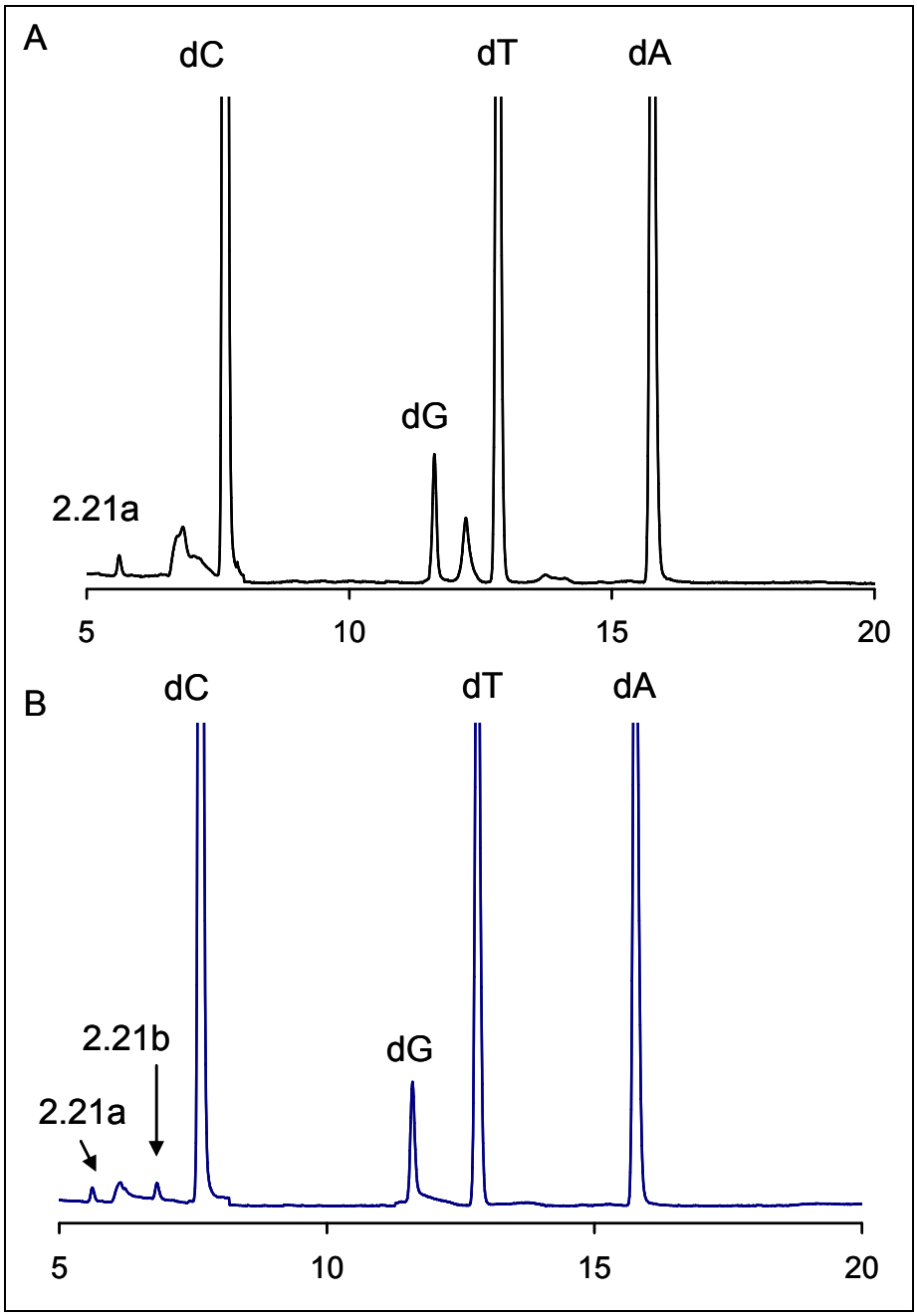

**5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCCCCC-3' (pyranose) with Phosphodiesterase II**

# Appendix V

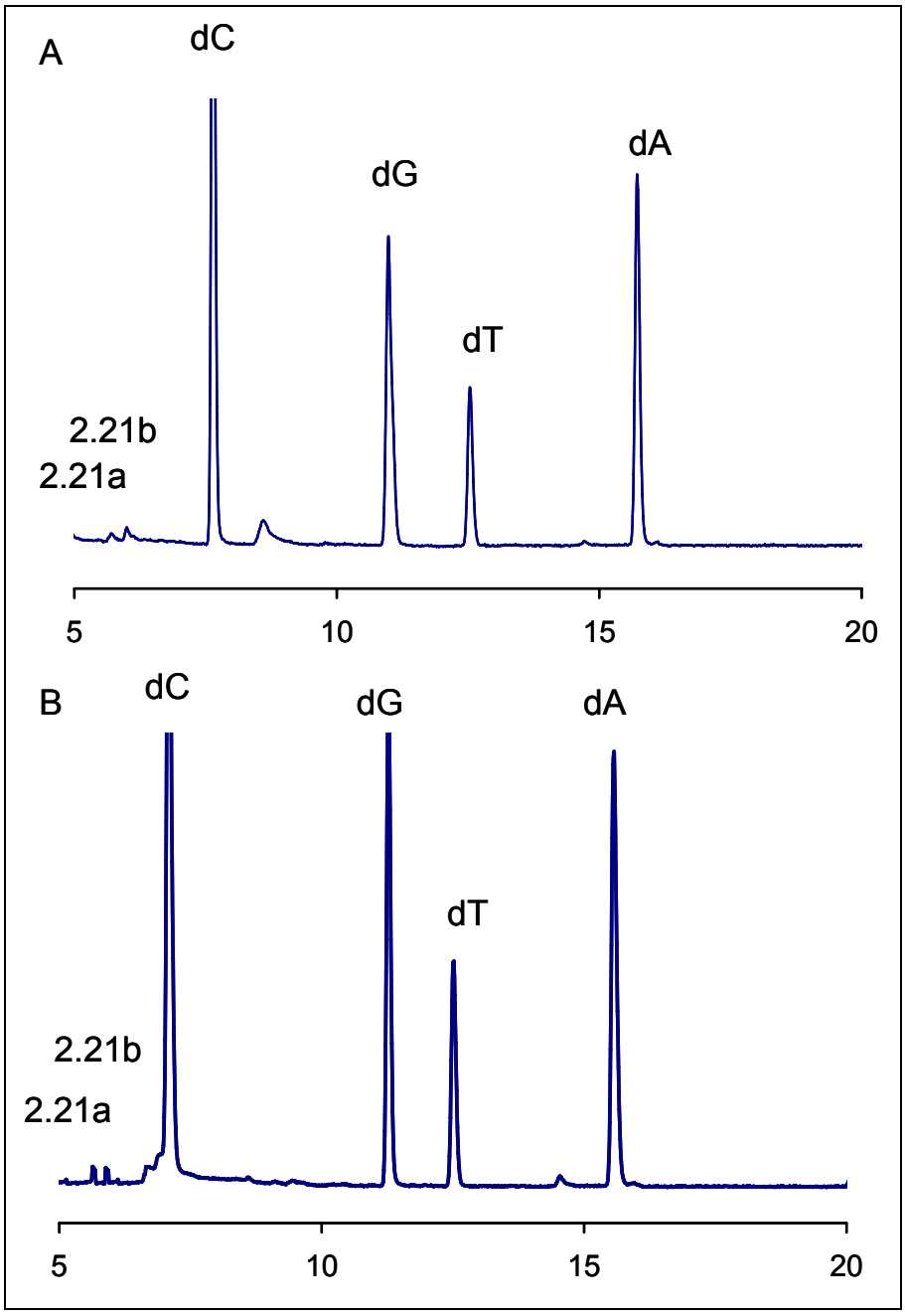
# Enzyme Digestion of MeFAPy-dGuo containing oligonucleotides for Chapter II



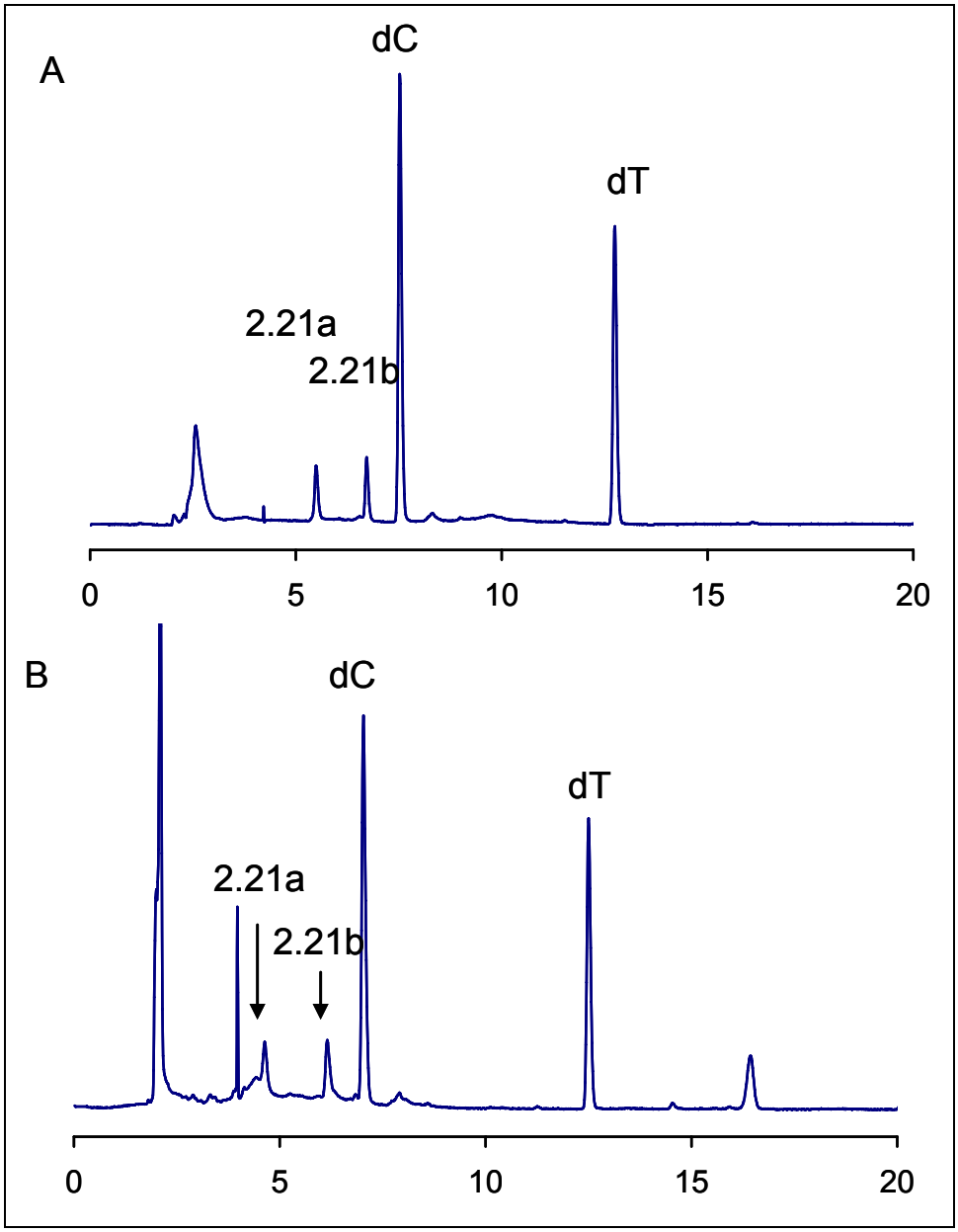

**Enzyme digestion of A: 5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.43a, furanose); B: 5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.43b, pyranose)**




**Enzyme digestion of A: 5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCCCCC-3' (furanose); B: 5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCCCCC-3' (pyranose)**



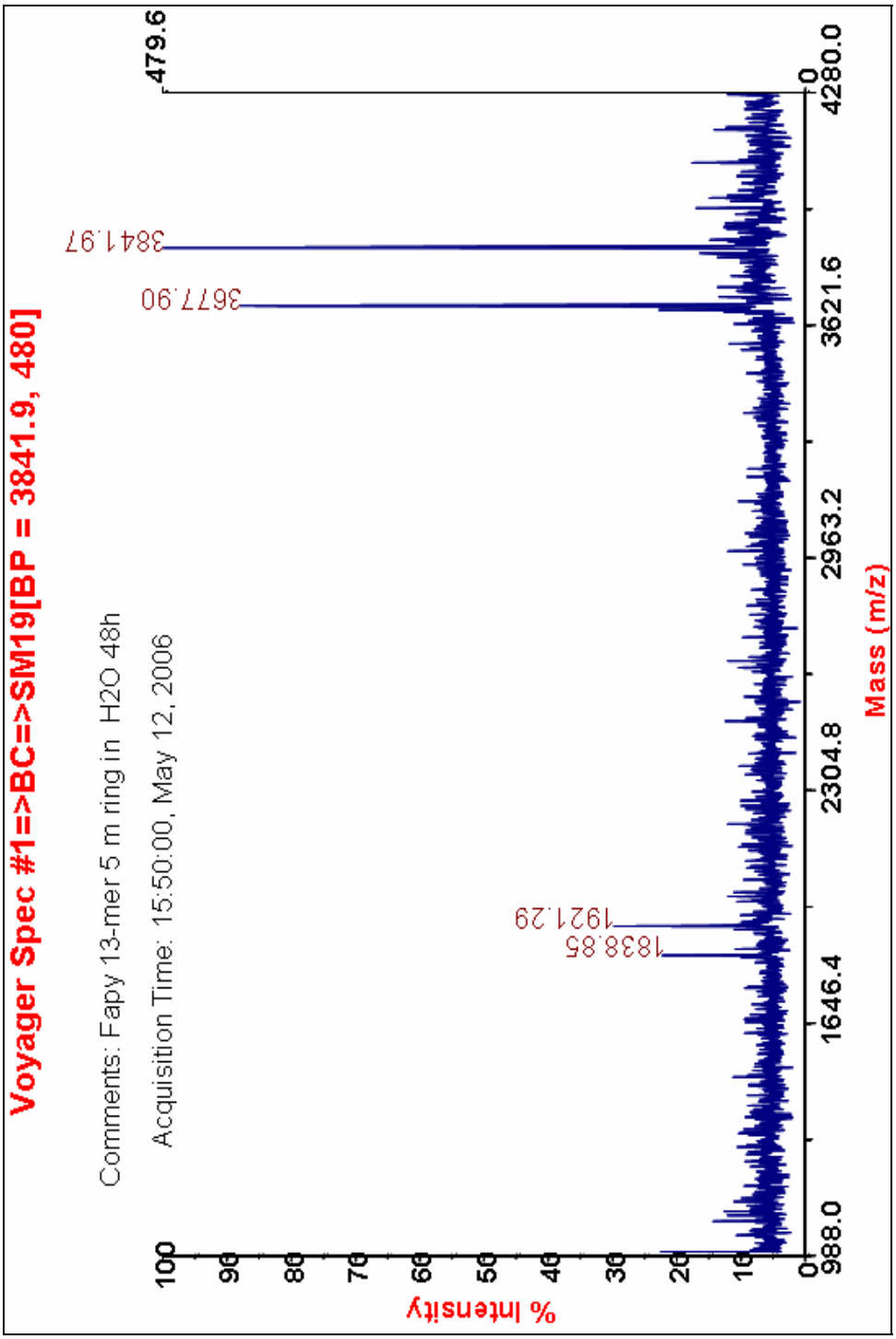

**Enzyme digestion of A: 5'-TCAT-(MeFAPy-dGuo)-GAATCCTTACG AGCATCGCCCCC-3' (2.52a, furanose); B: : 5'-TCAT-(MeFAPy-dGuo)-GAATCCTTACGAGCATCGCCCCC-3' (2.52b, pyranose)**



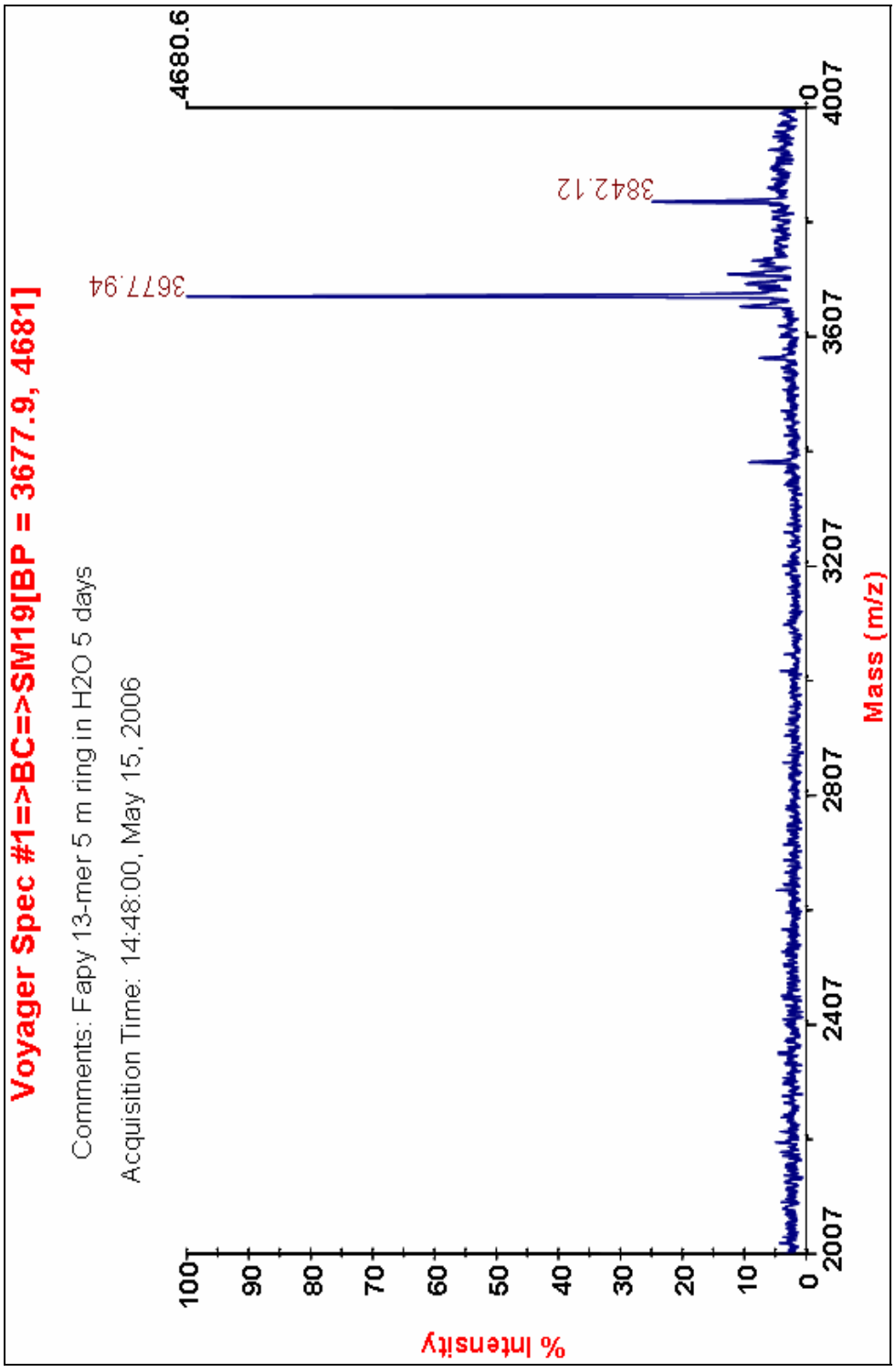

**Enzyme digestion of A: 5'-CTT-(MeFAPy-dGuo)-TT-3' (2.37a, furanose); B: 5'-CTT-(MeFAPy-dGuo)-TT-3' (2.37b, pyranose)**

# Appendix VI

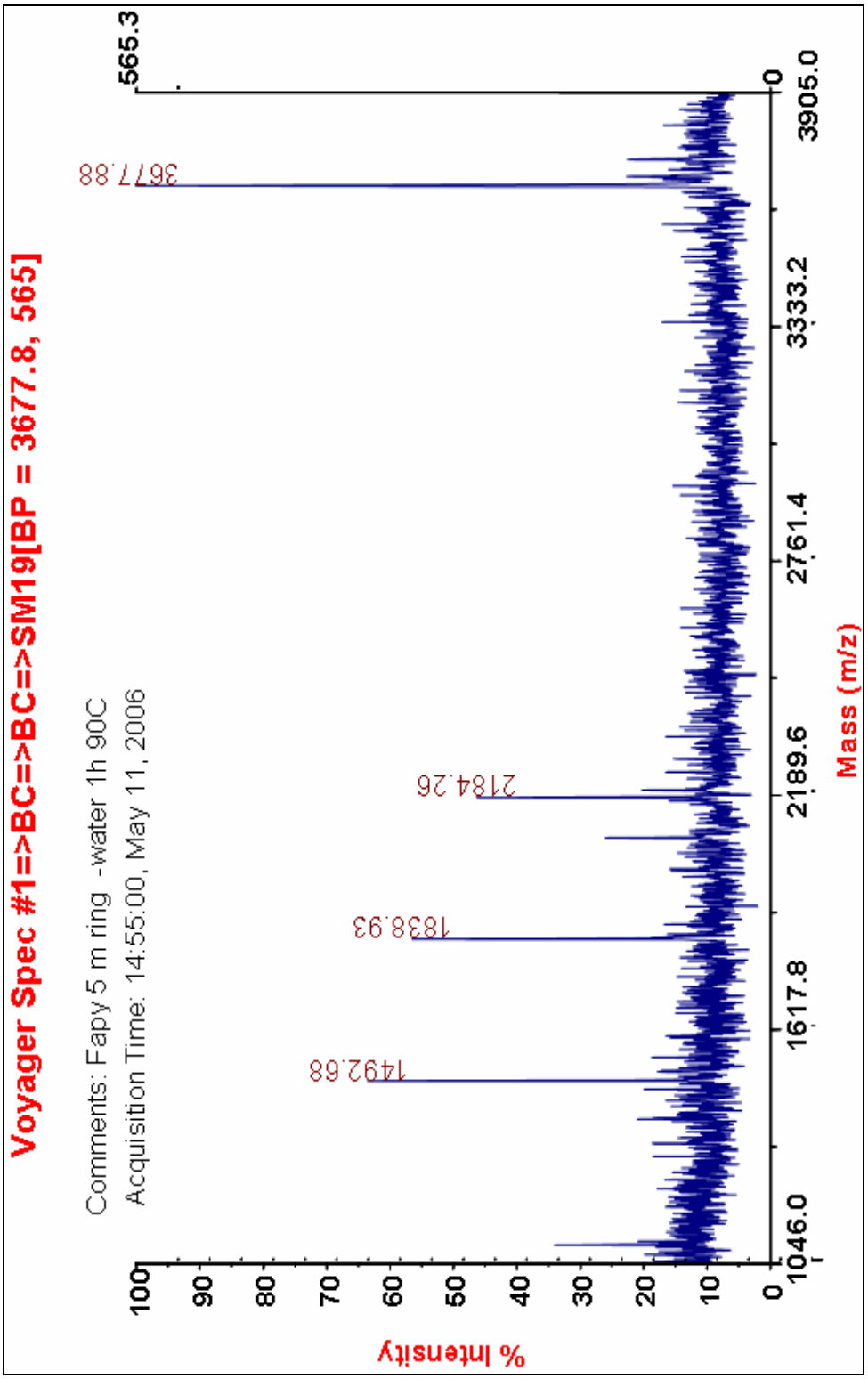
# Stability of MeFAPy-dGuo containing oligonucleotides (2.42a and 2.42b) for Chapter II



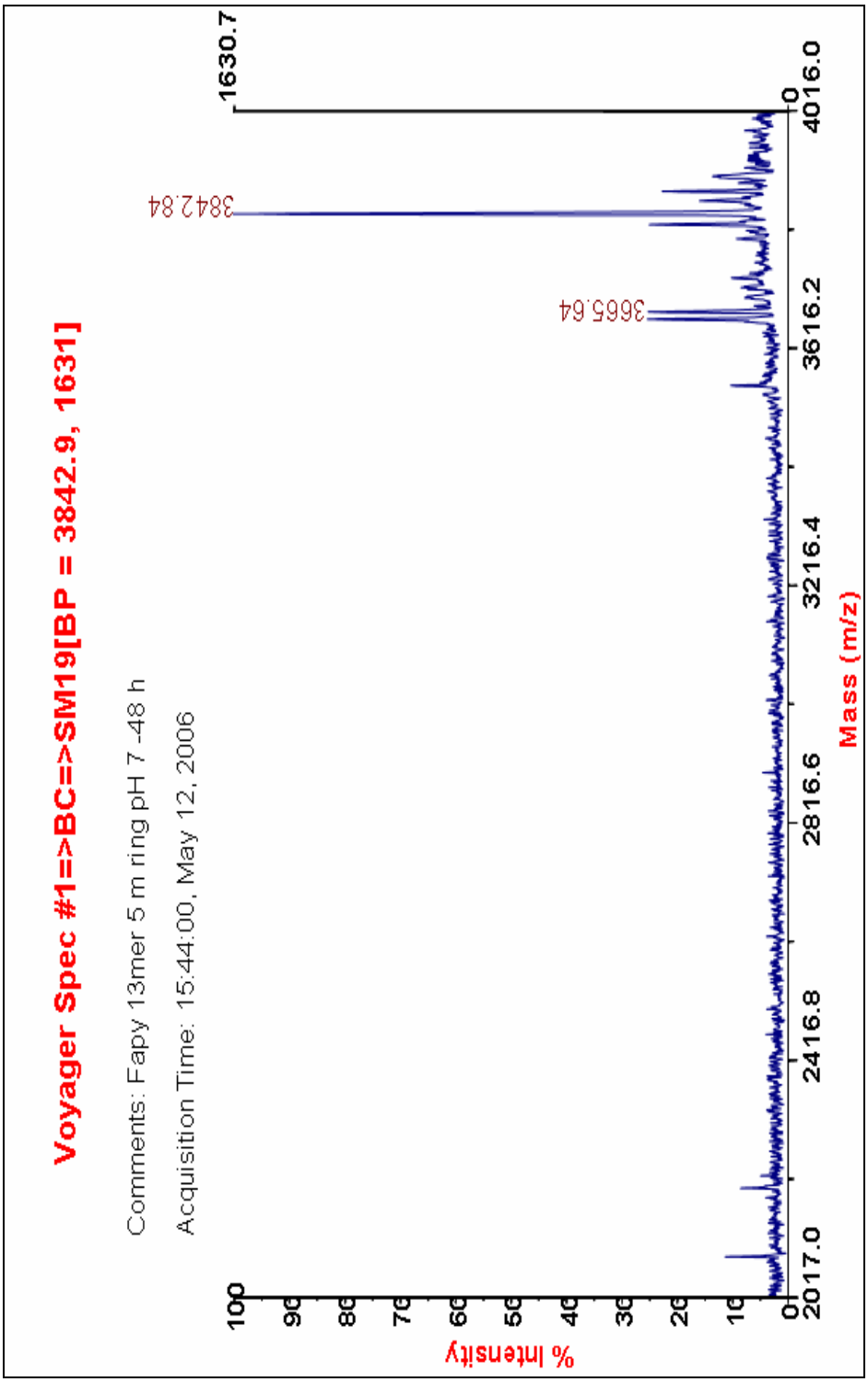

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in water for 2 days**



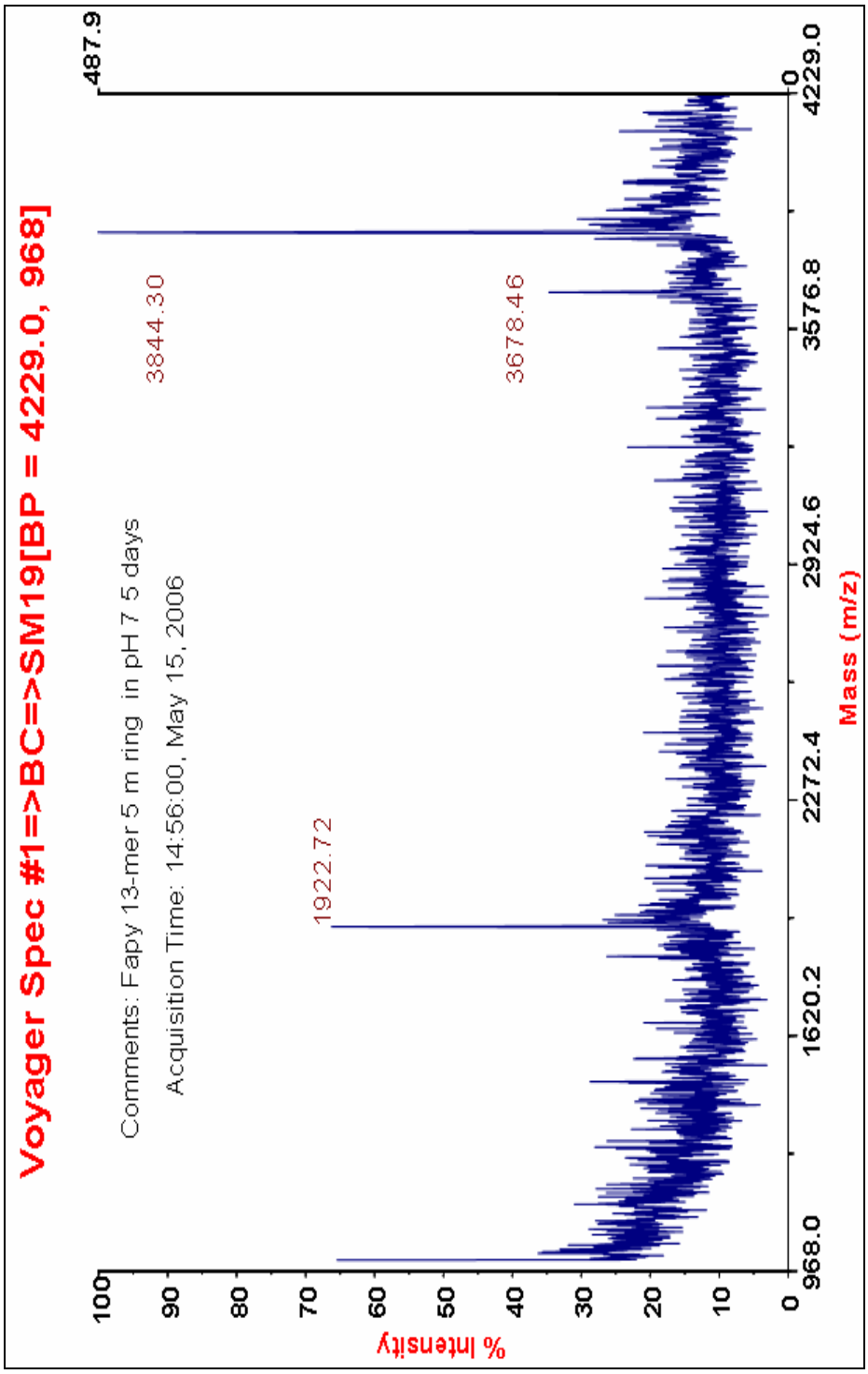

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in water for 5 days**



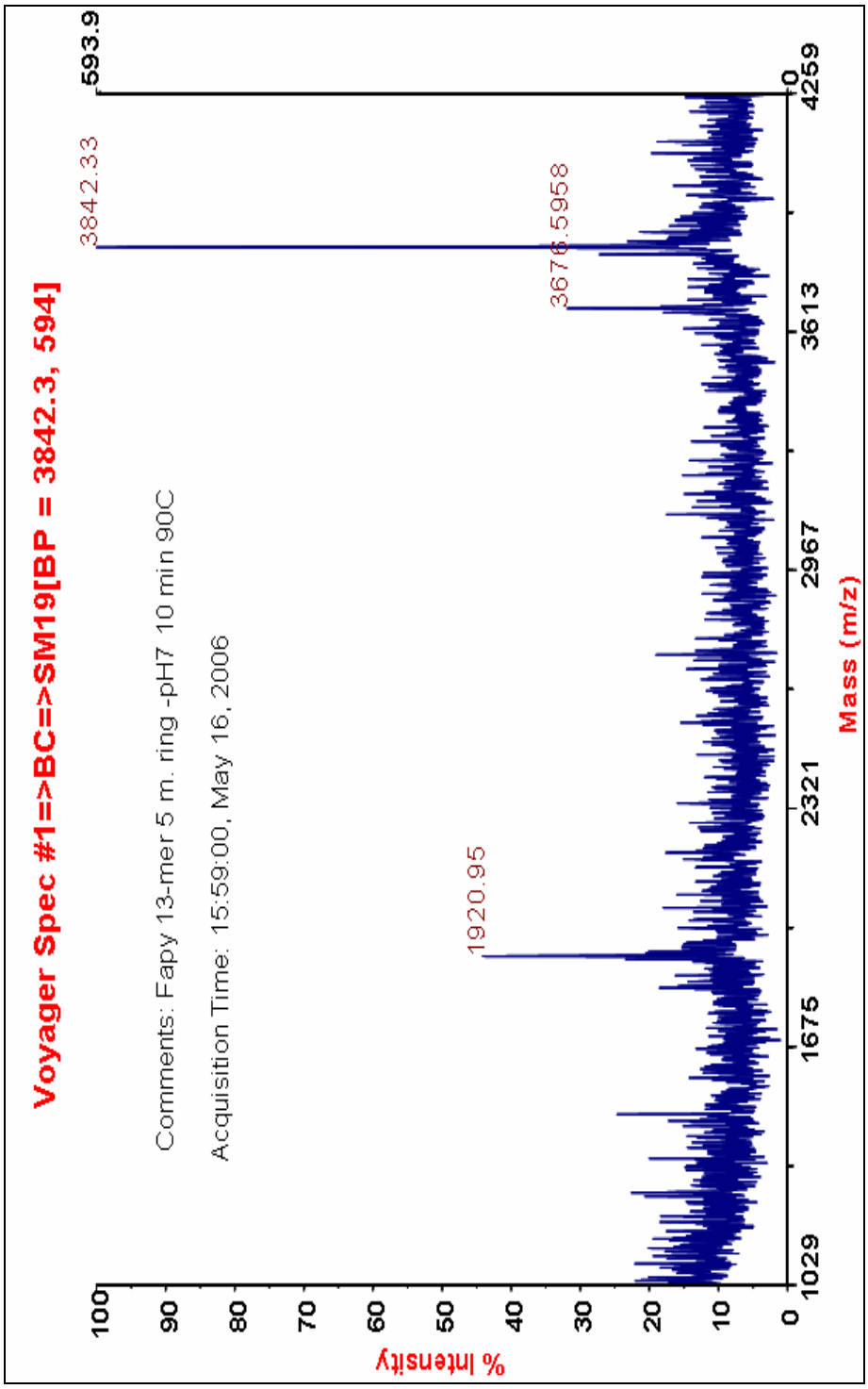

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in water for 1h at 90 °C**



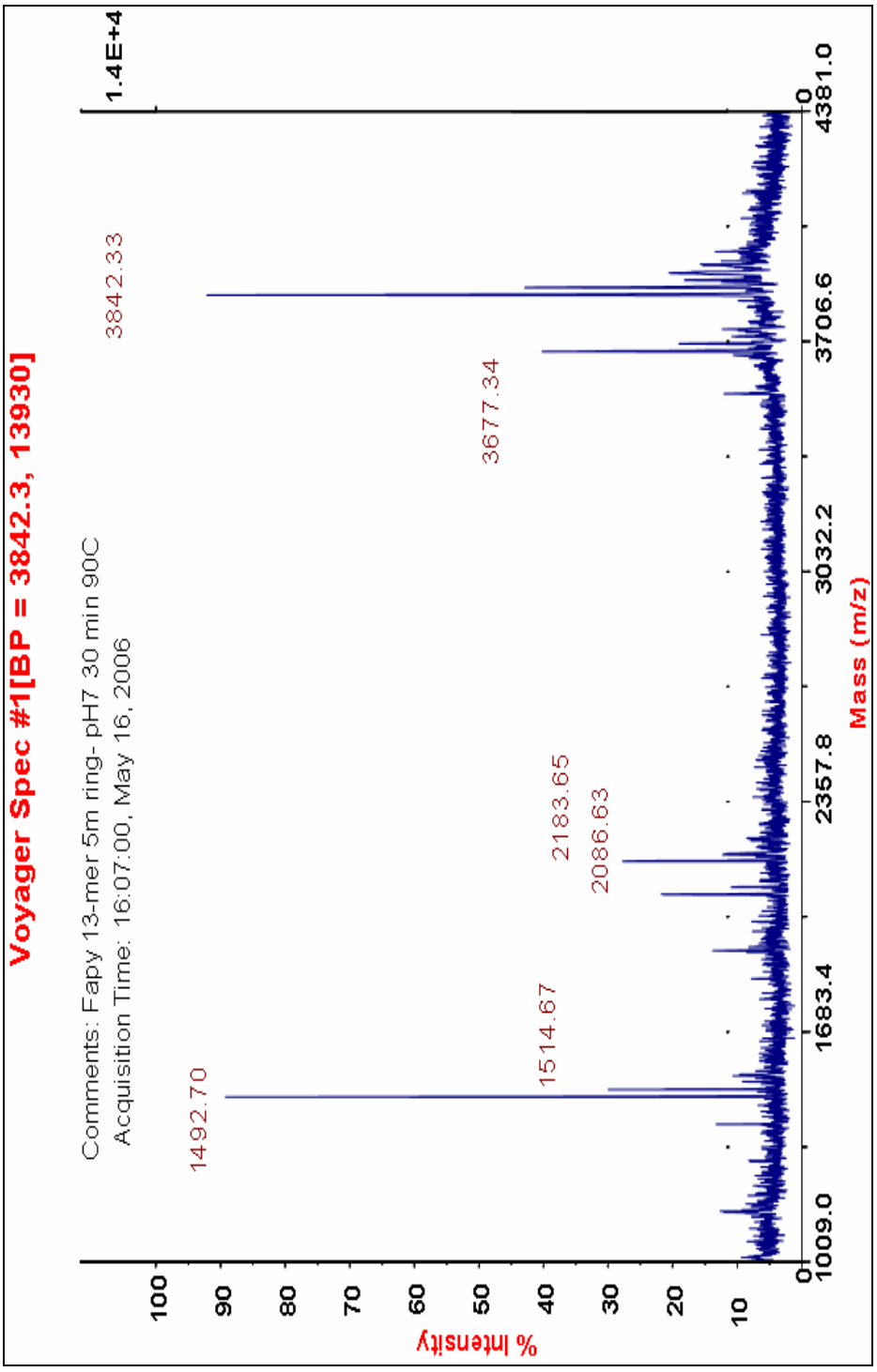

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in buffer for 2 days**



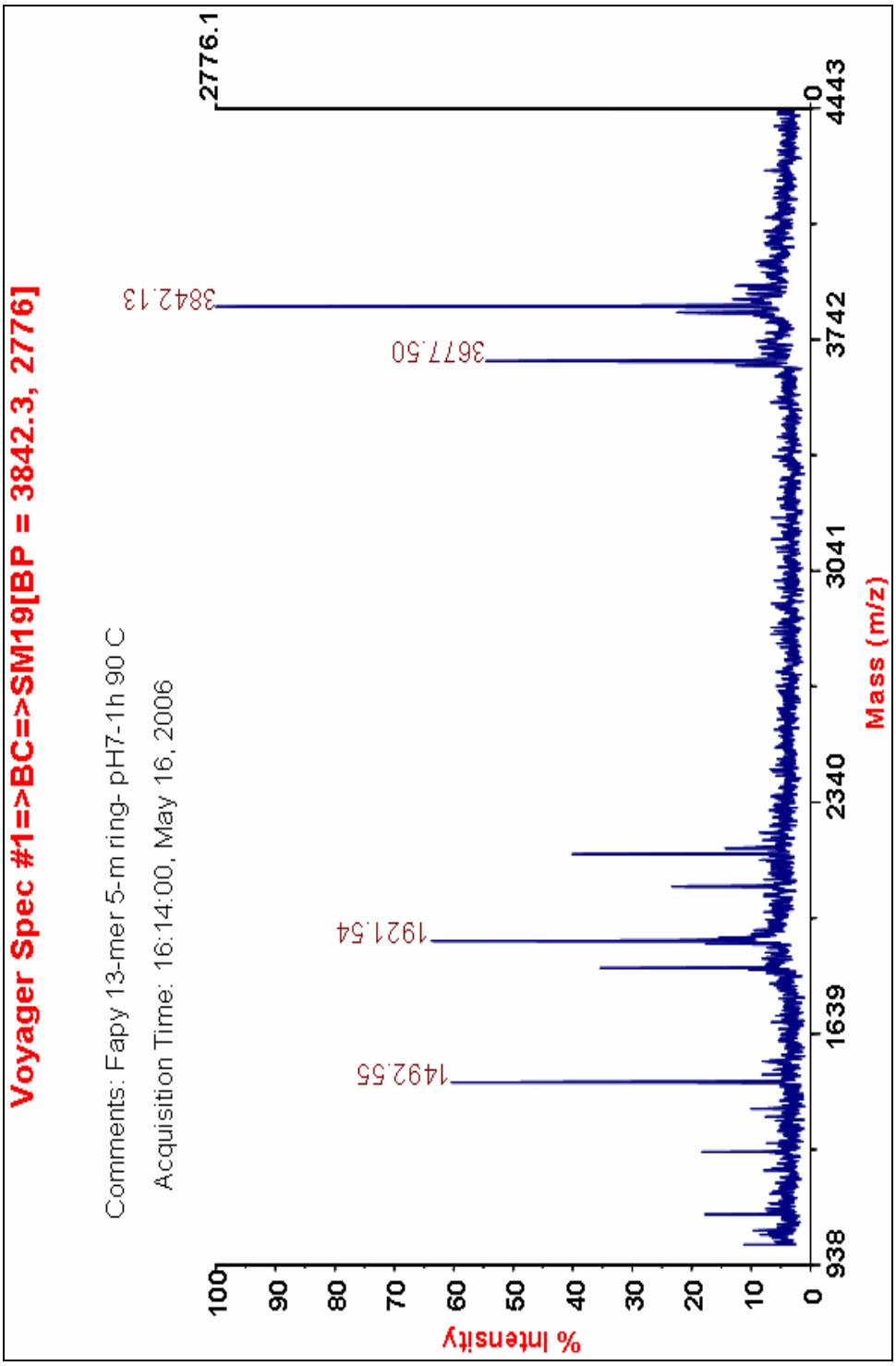

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in buffer for 5 days**



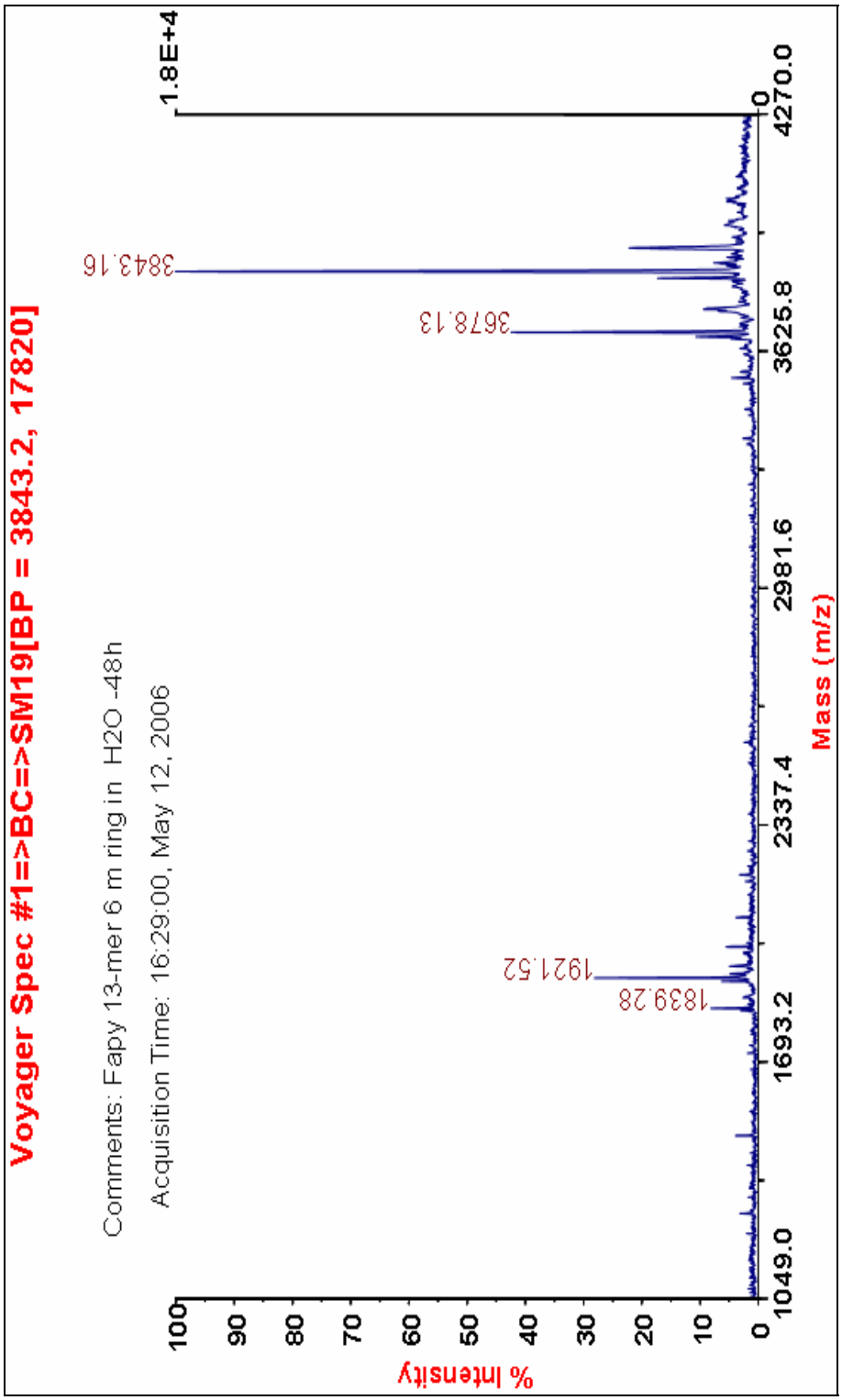

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in buffer for 10 min at 90 °C**



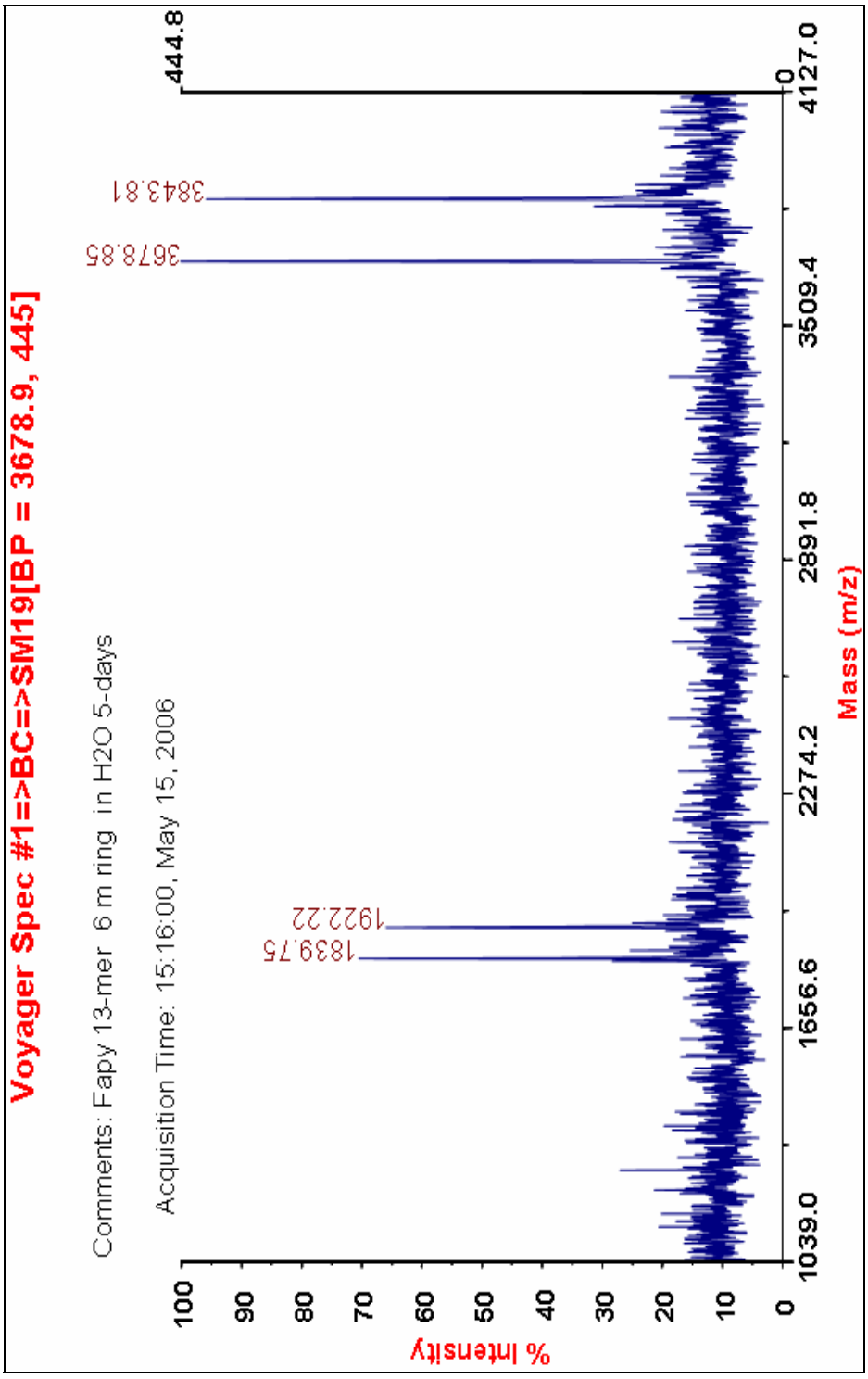

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in buffer for 30 min at 90 °C**



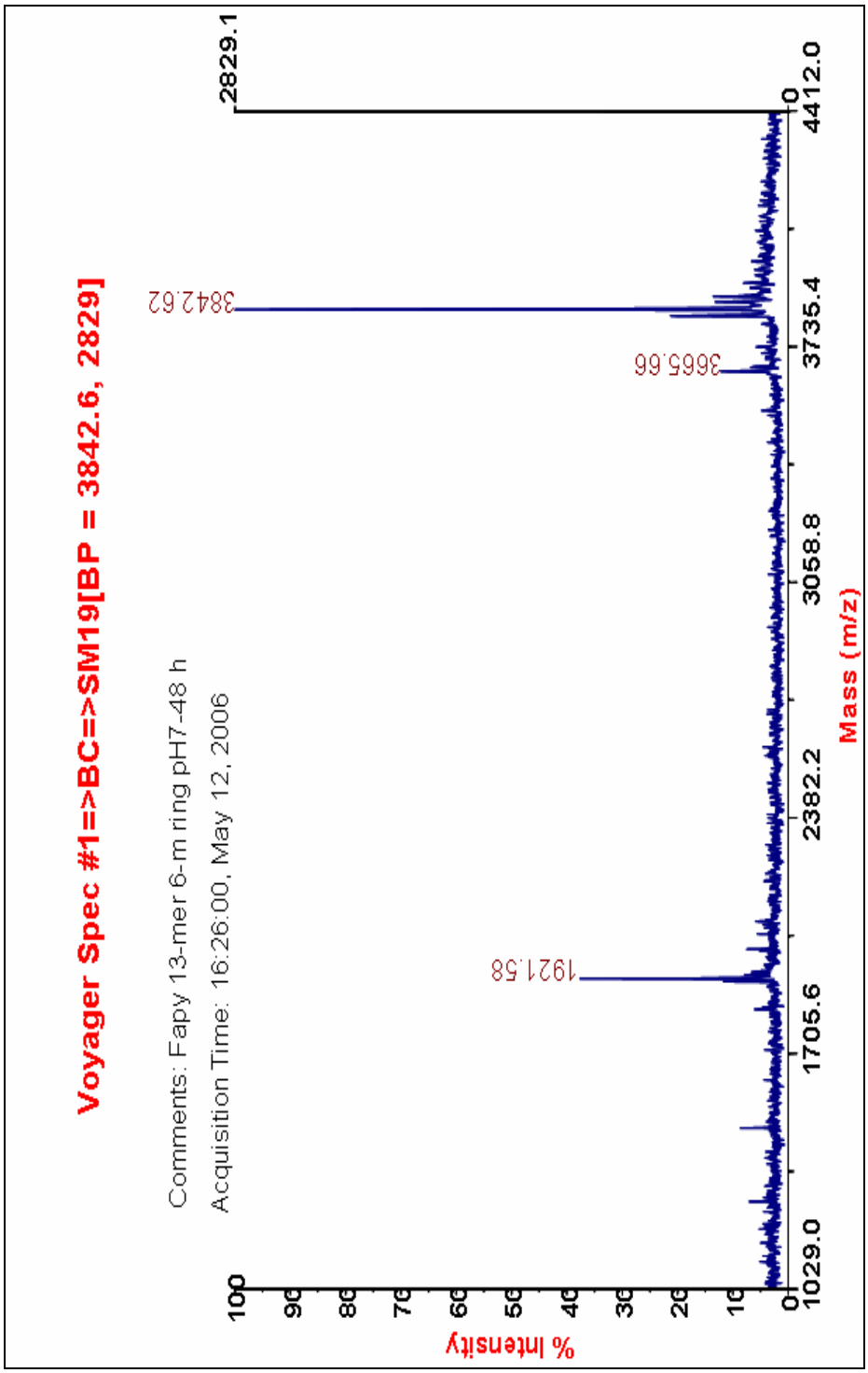

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in buffer for 1h at 90 °C**



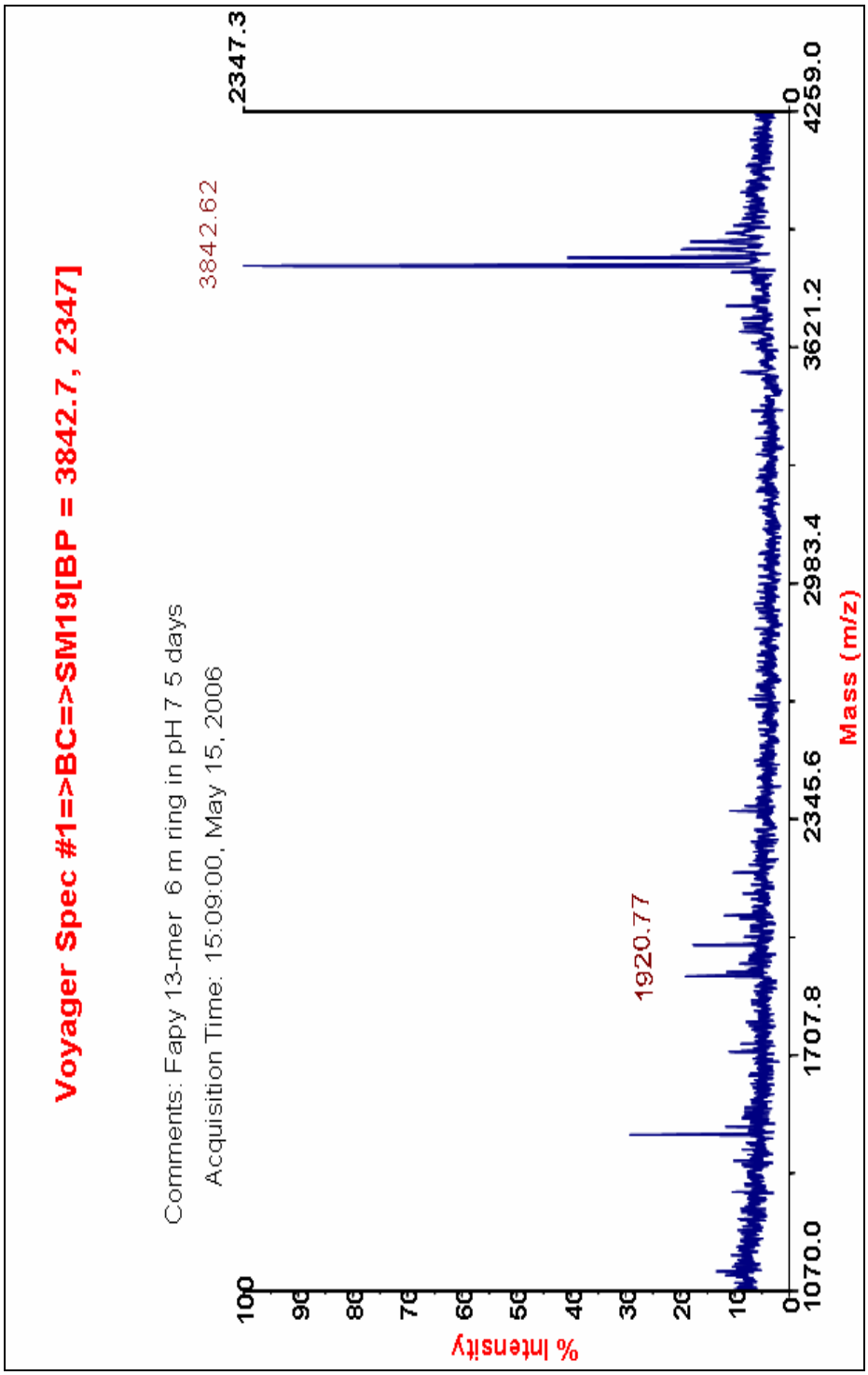

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42b, pyranose) in water for 2 days**



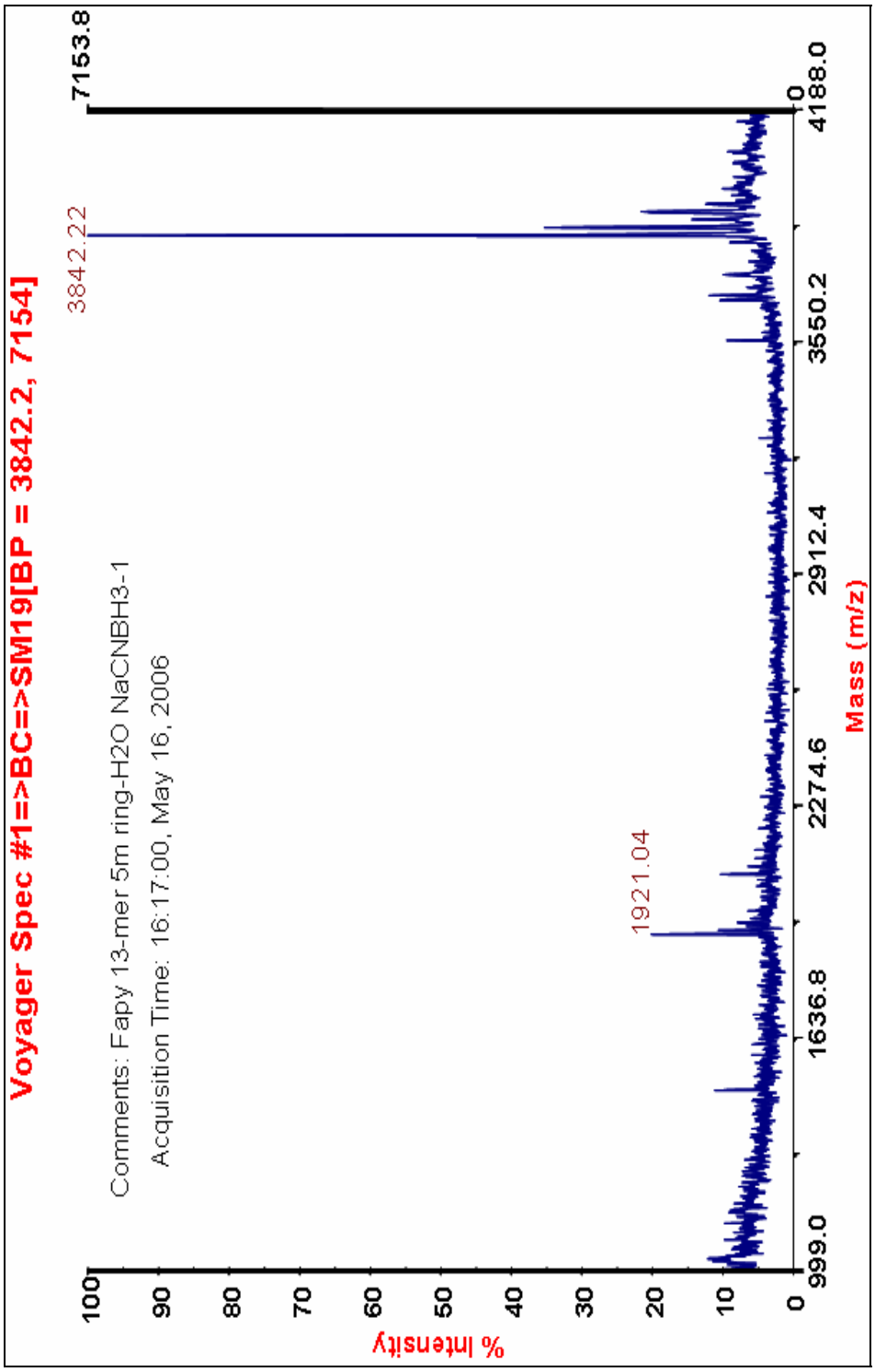

**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42b, pyranose) in water for 5 days**




**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42b, pyranose) in buffer for 2 days**




**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42b, pyranose) in buffer for 5 days**

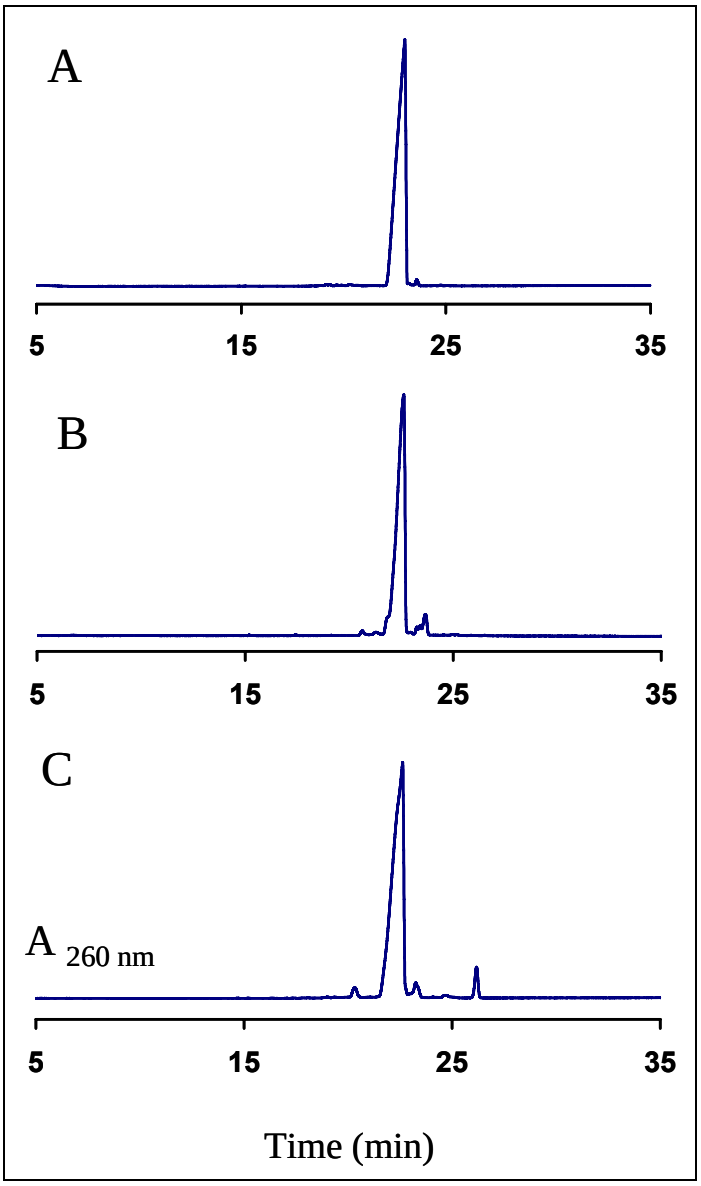



**5'-CCTCTTC-(MeFAPy-dGuo)-CTCTC-3' (2.42a, furanose) in water with $Na(CN)BH_3$ for 2 days**

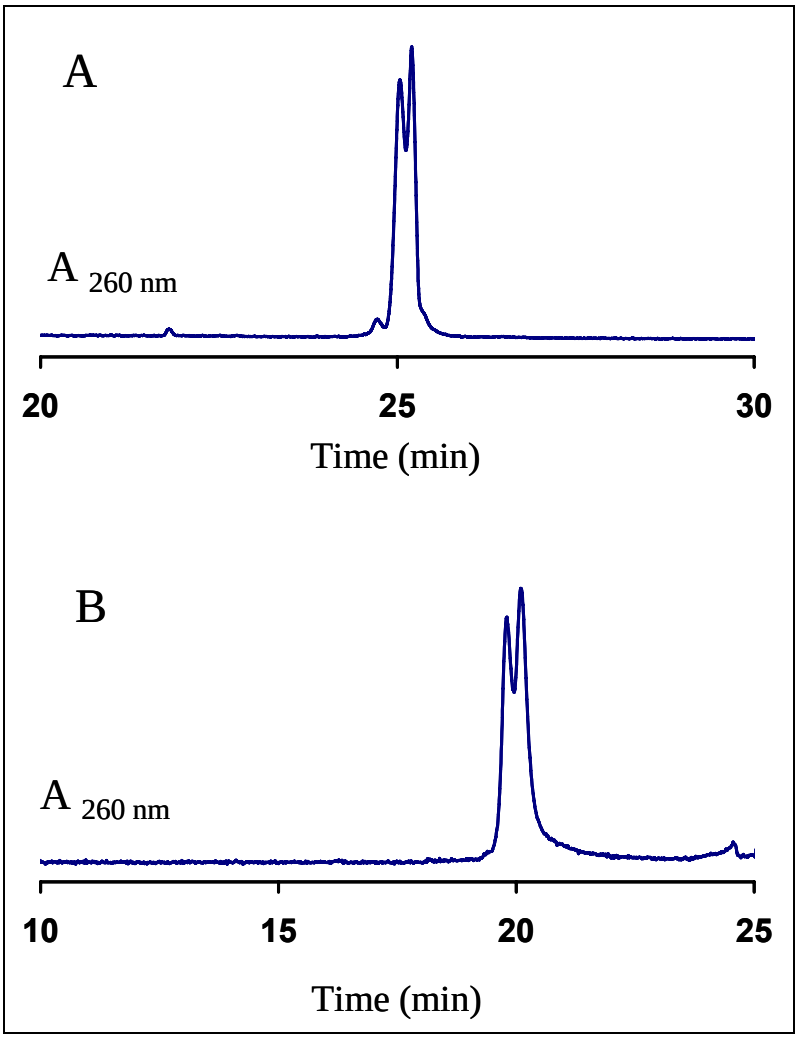
# Appendix VII

# HPLC and CZE Traces

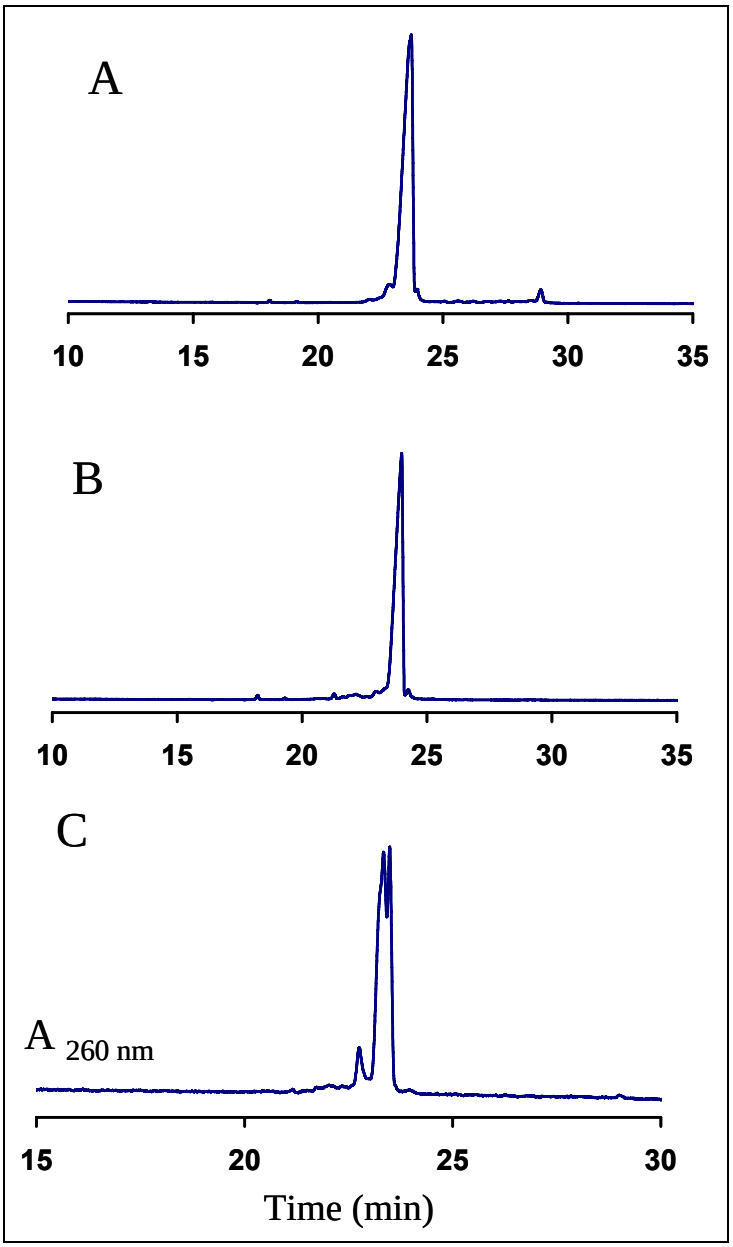
# for Chapter II



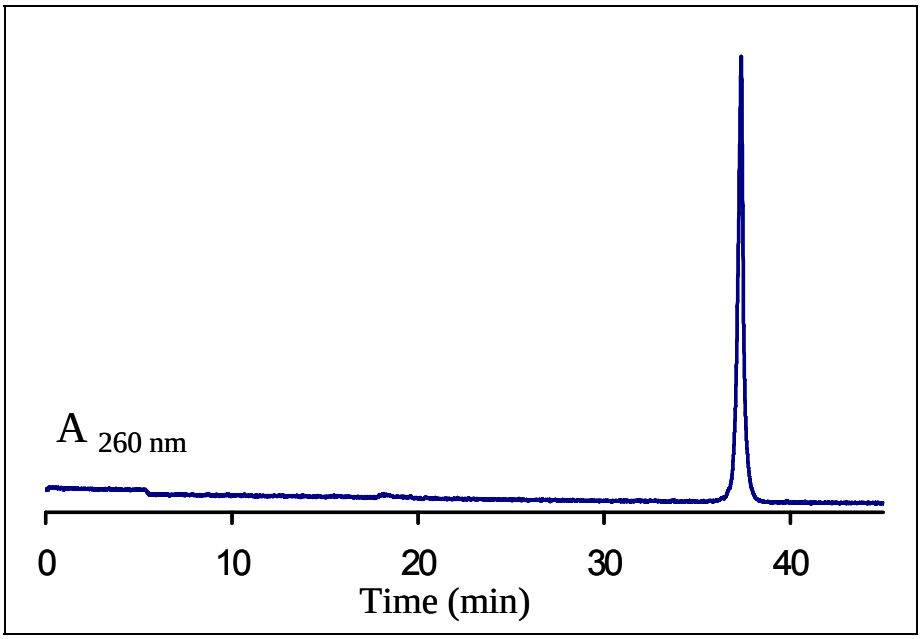

**CZE analysis of 5'-CTT-(MeFAPy-dGuo)-TT-3' (2.27); A) CZE traces of 2.27a (furanose form); B CZE traces of 2.27b (pyranose form); C) CZE traces of the co-injection of 2.27a and 2.27b.**




**CZE (A) and HPLC (B) traces of 5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCCCCC-3' (furanose form).**




**CZE analysis of 5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCCCCC-3'; A) CZE traces of the furanose form; B) CZE traces of the pyranose form; C) CZE traces of the co-injection of the furanose and pyranose forms.**

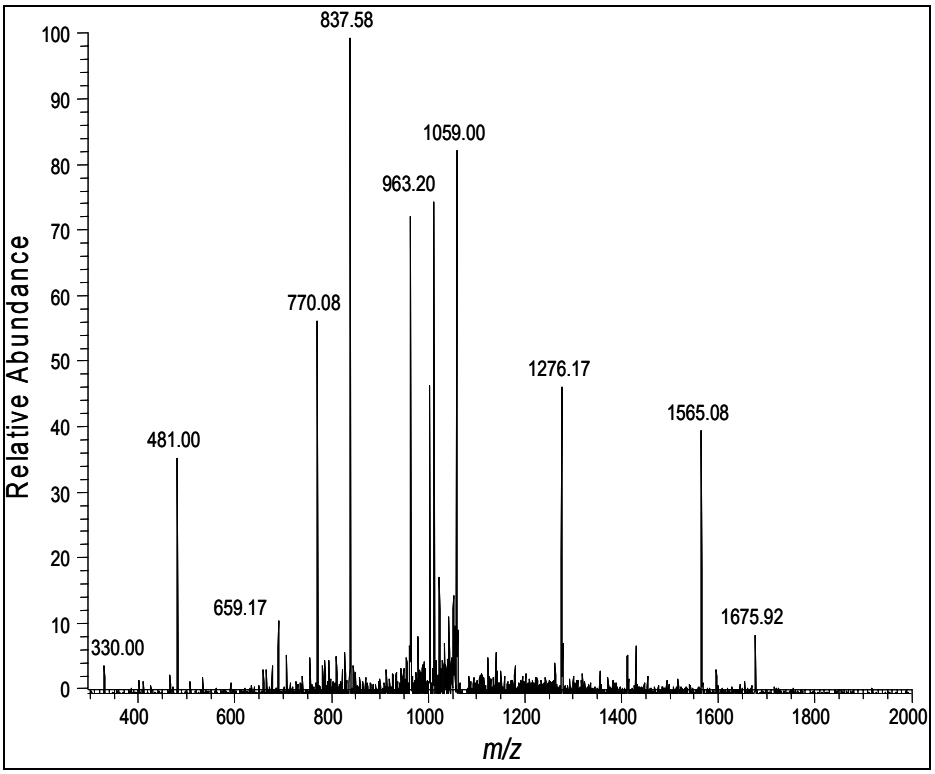



**CZE analysis of 5'-TCAT-(MeFAPy-dGuo)-GAATCCTTCGAGCATC GCCCCC-3' 92.52A).**

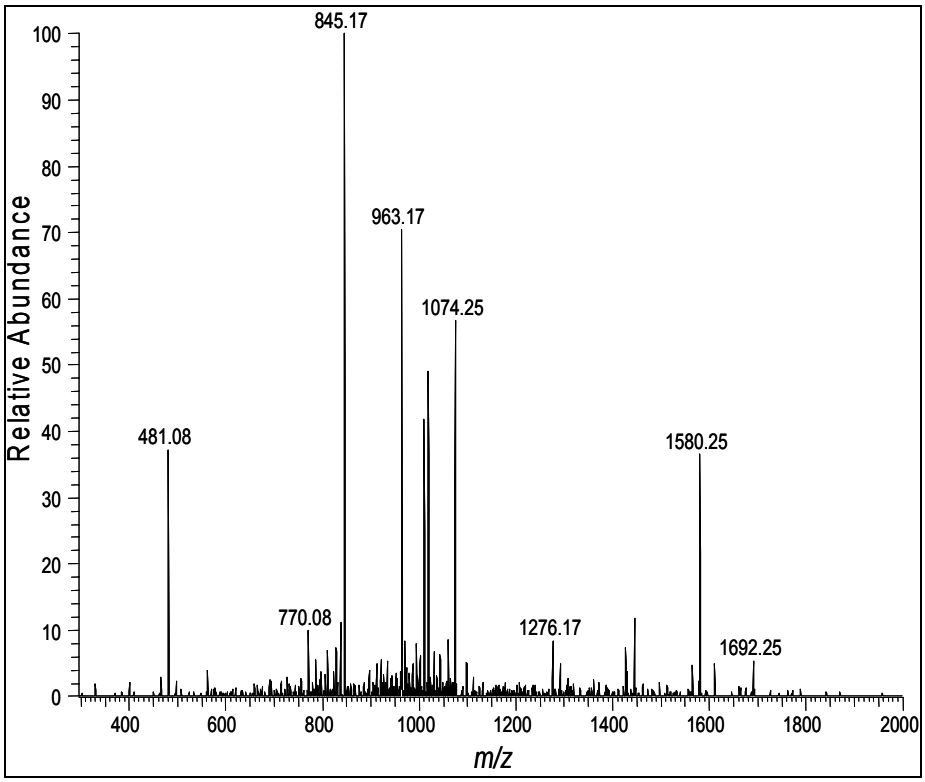
# Appendix VIII

# LC-ESI/MS/MS Sequencing

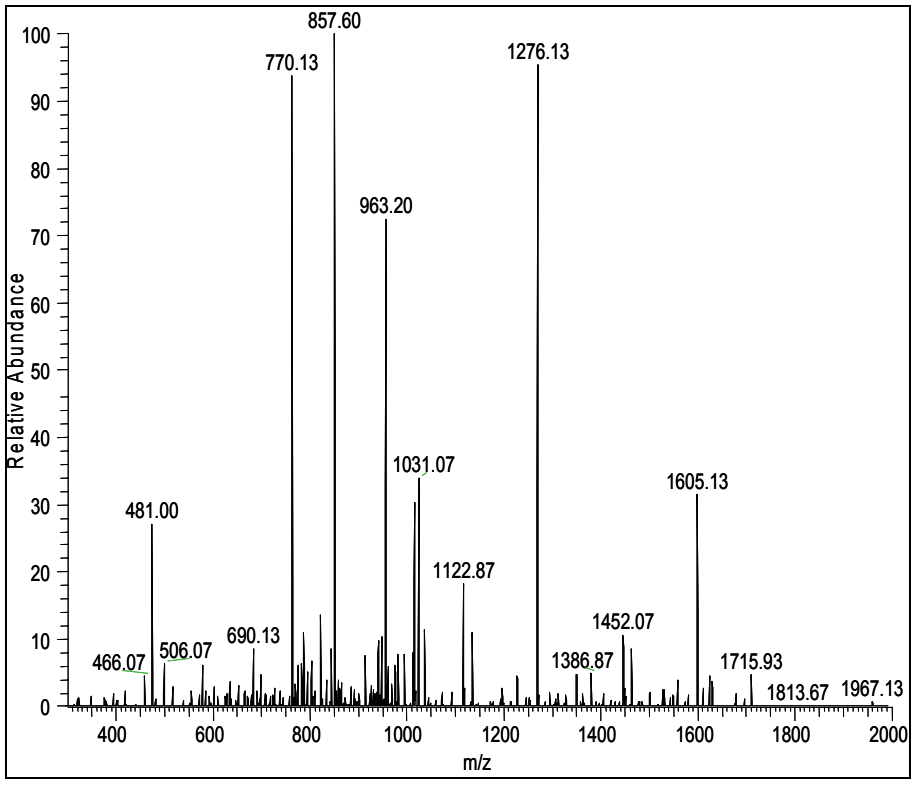
# for Chapter II



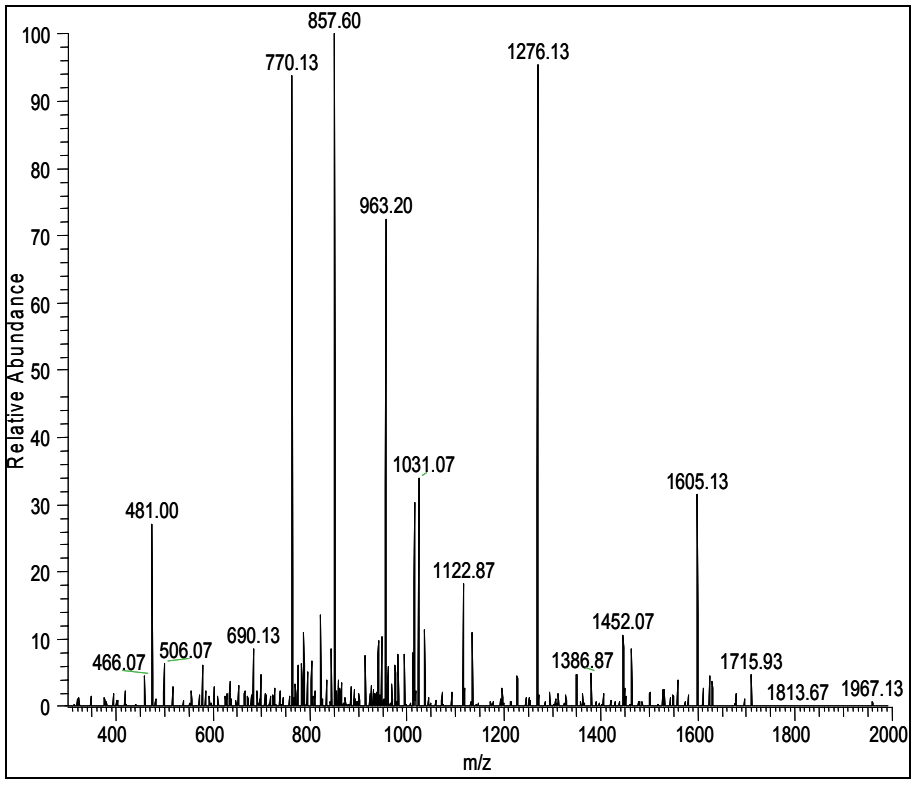

**CID fragmentation of commercially obtained sequence 5′-pTCCATGA-3′**



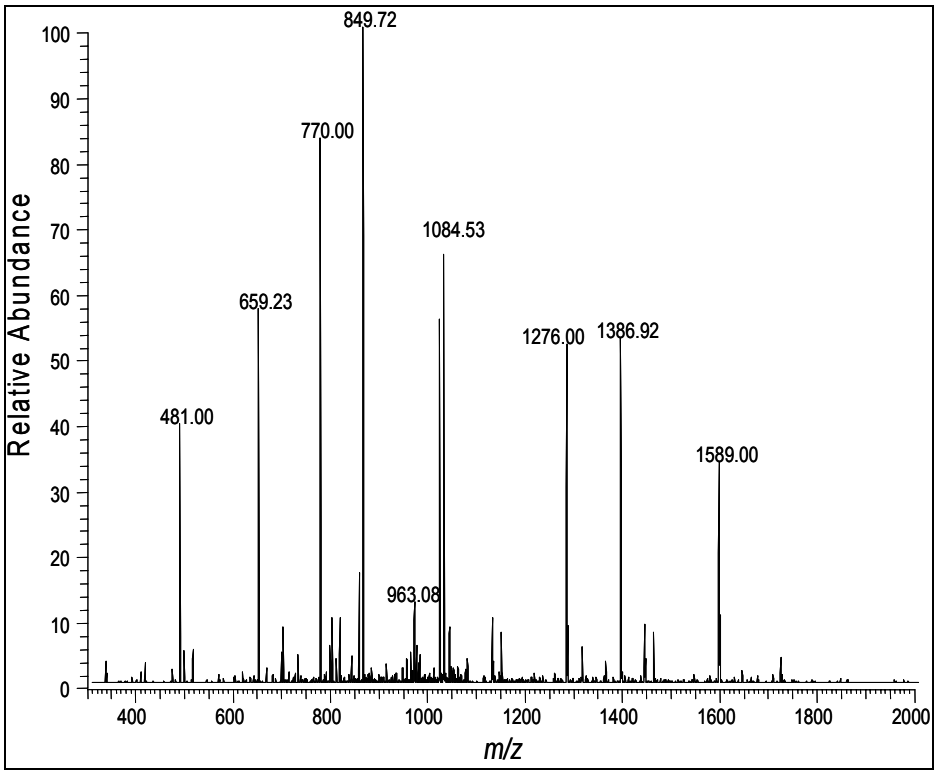

**CID fragmentation of commercially obtained sequence 5′-pTCTATGA-3′**



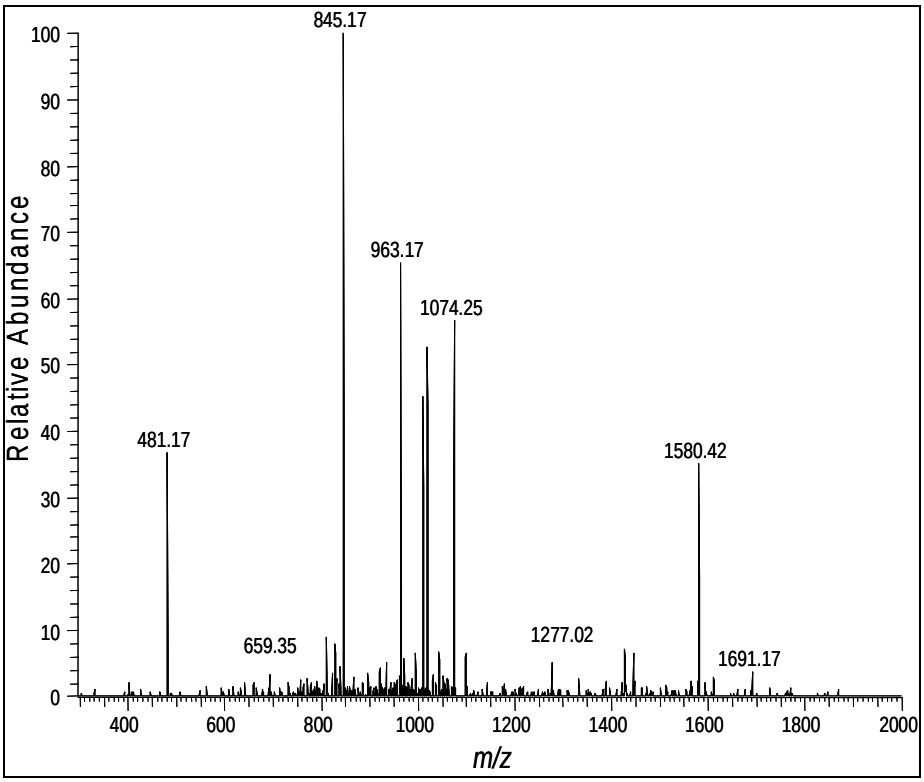

**CID fragmentation of commercially obtained sequence 5′-pTCGATGA-3′**




**CID spectrum of the molecular ion peak *m/z* 1099.25 of the DPO4 full-length extension product, identified as 5′-pTCGATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide 2.52a using primer 5′GGGGGCGATGCTCGTAAGGATTCG-3’ (2.58).**




**CID spectrum of the molecular ion peak *m/z* 1090.68 of the DPO4 full-length extension product, identified as 5′-pTCAATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide 2.52a using primer 5′GGGGGCGATGCTCGTAAGGATTCA-3' (2.56).**




**CID spectrum of the molecular ion peak *m/z* 1086.83 of the DPO4 full-length extension product, identified as 5′-pTCTATGA-3′, past the MeFAPy-dGuo lesion in oligonucleotide 2.52a using primer 5′GGGGGCGATGCTCGTAAGGATTCT-3' (2.57).**

**Sequence :**
**TCC ATGA**
5'p - DNA[7mer] - 3'OH
C:2 T:2 A:2 G:1
monoisotopic mass, negative mode

**CID FRAGMENTS**

| n | ch | a-B | w | y | d-H2O |
|---|---|---|---|---|---|
| 1 | -1 | | 330.059 | 250.093 | 303.037 |
| | -2 | | | | 151.015 |
| 2 | -1 | 481.040 | 659.111 | 579.145 | 592.083 |
| | -2 | 240.016 | 329.051 | 289.068 | 295.538 |
| | -3 | | | | 196.689 |
| 3 | -1 | 770.086 | 963.157 | 883.191 | 881.129 |
| | -2 | 384.539 | 481.074 | 441.091 | 440.061 |
| | -3 | 256.023 | 320.380 | 293.725 | 293.038 |
| | -4 | | | | 219.526 |
| 4 | -1 | 1059.132 | 1276.214 | 1196.248 | 1194.187 |
| | -2 | 529.062 | 637.603 | 597.620 | 596.589 |
| | -3 | 352.372 | 424.733 | 398.077 | 397.390 |
| | -4 | 264.027 | 318.297 | 298.306 | 297.790 |
| | -5 | | | | 238.031 |
| 5 | -1 | 1372.189 | 1565.260 | 1485.294 | 1498.232 |
| | -2 | 685.590 | 782.126 | 742.143 | 748.612 |
| | -3 | 456.724 | 521.081 | 494.426 | 498.739 |
| | -4 | 342.291 | 390.559 | 370.567 | 373.802 |
| | -5 | 273.631 | 312.245 | 296.252 | 298.840 |
| | -6 | | | | 248.865 |
| 6 | -1 | 1676.235 | 1854.306 | 1774.340 | 1827.285 |
| | -2 | 837.613 | 926.649 | 886.666 | 913.138 |
| | -3 | 558.073 | 617.430 | 590.775 | 608.423 |
| | -4 | 418.302 | 462.820 | 442.829 | 456.065 |
| | -5 | 334.440 | 370.055 | 354.061 | 364.650 |
| | -6 | 278.532 | 308.211 | 294.883 | 303.707 |
| | -7 | | | | 260.176 |

**Sequence :**
**TCT ATG A**
5'p - DNA[7mer] - 3'OH
C:1 T:3 A:2 G:1
monoisotopic mass, negative mode

**CID FRAGMENTS**

| n | ch | a-B | w | y | d-H2O |
|---|---|---|---|---|---|
| 1 | -1 | | 330.059 | 250.093 | 303.037 |
| | -2 | | | | 151.015 |
| 2 | -1 | 481.040 | 659.111 | 579.145 | 592.083 |
| | -2 | 240.016 | 329.051 | 289.068 | 295.538 |
| | -3 | | | | 196.689 |
| 3 | -1 | 770.086 | 963.157 | 883.191 | 896.129 |
| | -2 | 384.539 | 481.074 | 441.091 | 447.560 |
| | -3 | 256.023 | 320.380 | 293.725 | 298.037 |
| | -4 | | | | 223.276 |
| 4 | -1 | 1074.131 | 1276.214 | 1196.248 | 1209.186 |
| | -2 | 536.561 | 637.603 | 597.620 | 604.089 |
| | -3 | 357.372 | 424.733 | 398.077 | 402.390 |
| | -4 | 267.777 | 318.297 | 298.306 | 301.540 |
| | -5 | | | | 241.031 |
| 5 | -1 | 1387.189 | 1580.260 | 1500.294 | 1513.232 |
| | -2 | 693.090 | 789.626 | 749.643 | 756.112 |
| | -3 | 461.724 | 526.081 | 499.426 | 503.738 |
| | -4 | 346.041 | 394.309 | 374.317 | 377.552 |
| | -5 | 276.631 | 315.245 | 299.252 | 301.840 |
| | -6 | | | | 251.365 |
| 6 | -1 | 1691.234 | 1869.306 | 1789.340 | 1842.284 |
| | -2 | 845.113 | 934.149 | 894.166 | 920.638 |
| | -3 | 563.073 | 622.430 | 595.774 | 613.423 |
| | -4 | 422.052 | 466.570 | 446.579 | 459.815 |
| | -5 | 337.440 | 373.055 | 357.061 | 367.650 |
| | -6 | 281.032 | 310.711 | 297.383 | 306.207 |
| | -7 | | | | 262.319 |

**Sequence :**
**TCA ATG A**
5'p - DNA[7mer] - 3'OH
C:1 T:2 A:3 G:1
monoisotopic mass, negative mode

**CID FRAGMENTS**

| n | ch | a-B | w | y | d-H2O |
|---|---|---|---|---|---|
| 1 | -1 | | 330.059 | 250.093 | 303.037 |
| | -2 | | | | 151.015 |
| 2 | -1 | 481.040 | 659.111 | 579.145 | 592.083 |
| | -2 | 240.016 | 329.051 | 289.068 | 295.538 |
| | -3 | | | | 196.689 |
| 3 | -1 | 770.086 | 963.157 | 883.191 | 905.141 |
| | -2 | 384.539 | 481.074 | 441.091 | 452.066 |
| | -3 | 256.023 | 320.380 | 293.725 | 301.041 |
| | -4 | | | | 225.529 |
| 4 | -1 | 1083.143 | 1276.214 | 1196.248 | 1218.198 |
| | -2 | 541.067 | 637.603 | 597.620 | 608.595 |
| | -3 | 360.375 | 424.733 | 398.077 | 405.394 |
| | -4 | 270.029 | 318.297 | 298.306 | 303.793 |
| | -5 | | | | 242.833 |
| 5 | -1 | 1396.200 | 1589.271 | 1509.305 | 1522.244 |
| | -2 | 697.596 | 794.132 | 754.149 | 760.618 |
| | -3 | 464.728 | 529.085 | 502.430 | 506.742 |
| | -4 | 348.294 | 396.562 | 376.570 | 379.805 |
| | -5 | 278.433 | 317.048 | 301.054 | 303.642 |
| | -6 | | | | 252.867 |
| 6 | -1 | 1700.246 | 1878.317 | 1798.351 | 1851.296 |
| | -2 | 849.619 | 938.655 | 898.672 | 925.144 |
| | -3 | 566.076 | 625.434 | 598.778 | 616.426 |
| | -4 | 424.305 | 468.823 | 448.832 | 462.068 |
| | -5 | 339.243 | 374.857 | 358.864 | 369.453 |
| | -6 | 282.534 | 312.213 | 298.885 | 307.709 |
| | -7 | | | | 263.607 |

**Sequence :**
**TCG ATG A**
5'p - DNA[7mer] - 3'OH
C:1 T:2 A:2 G:2
monoisotopic mass, negative mode

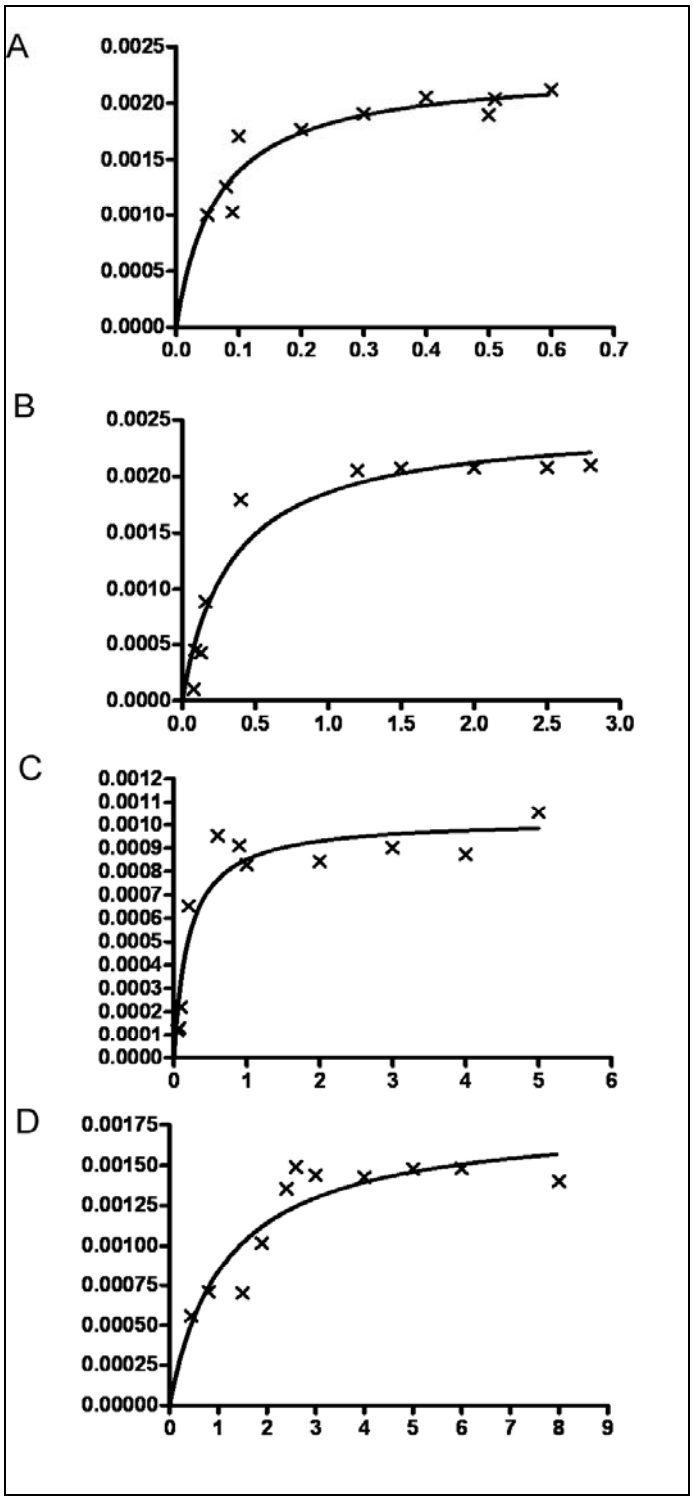
**CID FRAGMENTS**

| n | ch | a-B | w | y | d-H2O |
|---|---|---|---|---|---|
| 1 | -1 | | 330.059 | 250.093 | 303.037 |
| | -2 | | | | 151.015 |
| 2 | -1 | 481.040 | 659.111 | 579.145 | 592.083 |
| | -2 | 240.016 | 329.051 | 289.068 | 295.538 |
| | -3 | | | | 196.689 |
| 3 | -1 | 770.086 | 963.157 | 883.191 | 921.136 |
| | -2 | 384.539 | 481.074 | 441.091 | 460.064 |
| | -3 | 256.023 | 320.380 | 293.725 | 306.373 |
| | -4 | | | 229.528 | |
| 4 | -1 | 1099.138 | 1276.214 | 1196.248 | 1234.193 |
| | -2 | 549.065 | 637.603 | 597.620 | 616.592 |
| | -3 | 365.707 | 424.733 | 398.077 | 410.725 |
| | -4 | 274.028 | 318.297 | 298.306 | 307.792 |
| | -5 | | | | 246.032 |
| 5 | -1 | 1412.195 | 1605.266 | 1525.300 | 1538.239 |
| | -2 | 705.593 | 802.129 | 762.146 | 768.615 |
| | -3 | 470.059 | 534.417 | 507.761 | 512.074 |
| | -4 | 352.293 | 400.560 | 380.569 | 383.803 |
| | -5 | 281.632 | 320.247 | 304.253 | 306.841 |
| | -6 | | | 255.533 | |
| 6 | -1 | 1716.241 | 1894.312 | 1814.346 | 1867.291 |
| | -2 | 857.616 | 946.652 | 906.669 | 933.141 |
| | -3 | 571.408 | 630.765 | 604.110 | 621.758 |
| | -4 | 428.304 | 472.822 | 452.830 | 466.066 |
| | -5 | 342.442 | 378.056 | 362.063 | 372.652 |
| | -6 | 285.200 | 314.878 | 301.551 | 310.375 |
| | -7 | | | | 265.892 |

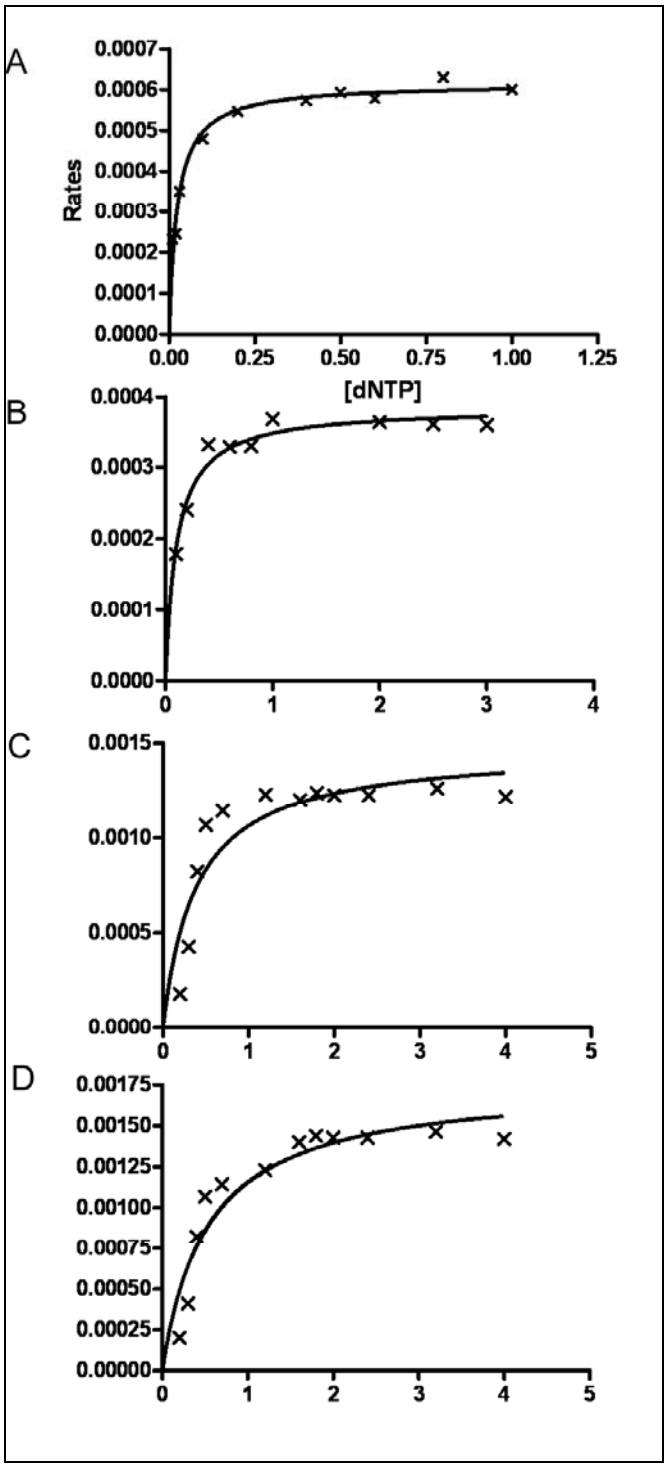
# Appendix IX

# Steady-state Kinetics

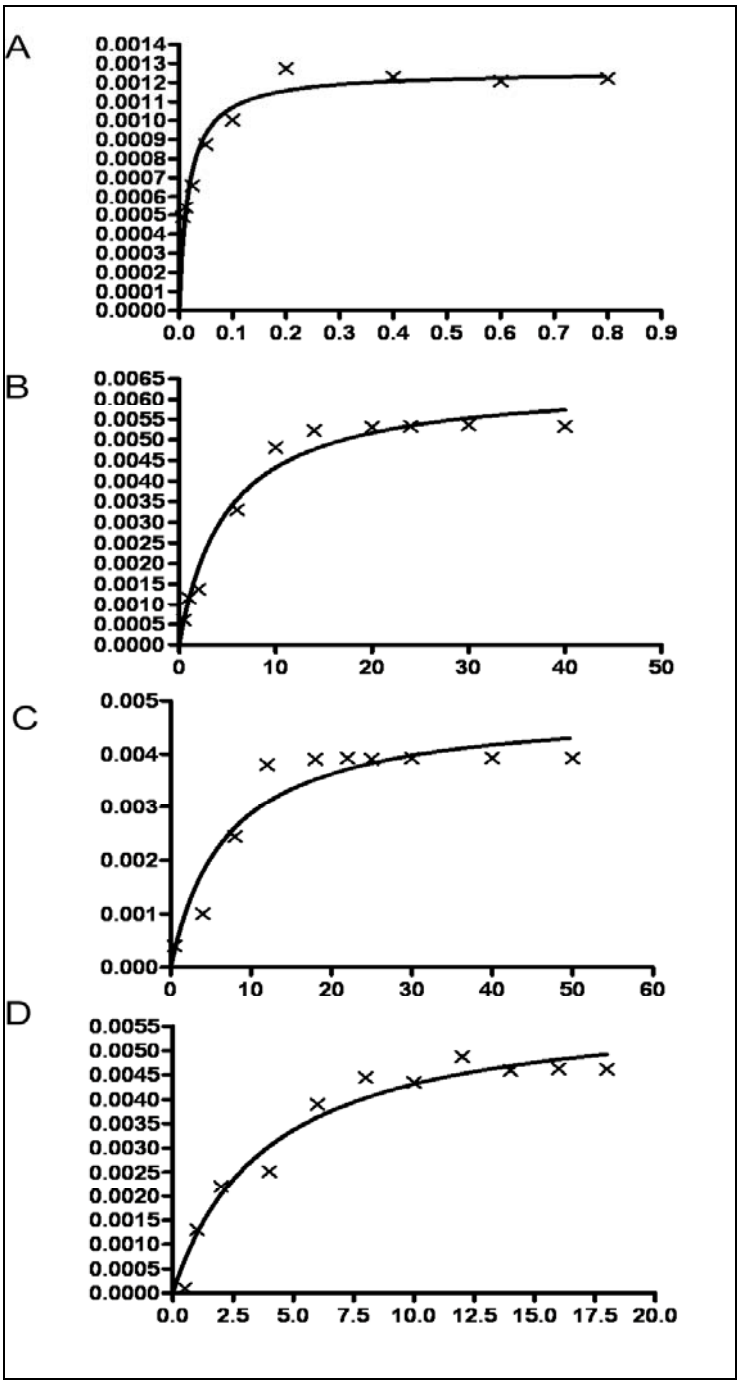
# for Chapter II



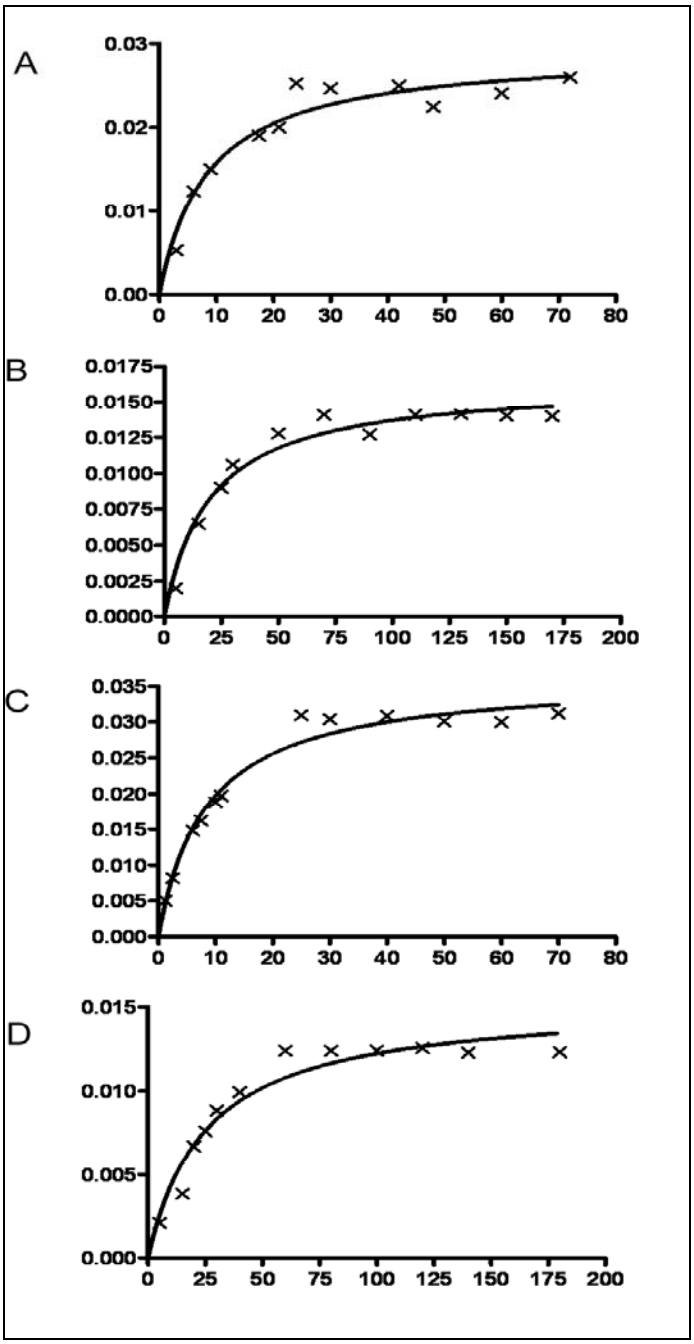

**Steady-State kinetics for $Kf^{-}$ with dCTP(A), dTTP ( B), dGTP(C) and dATP (D)**



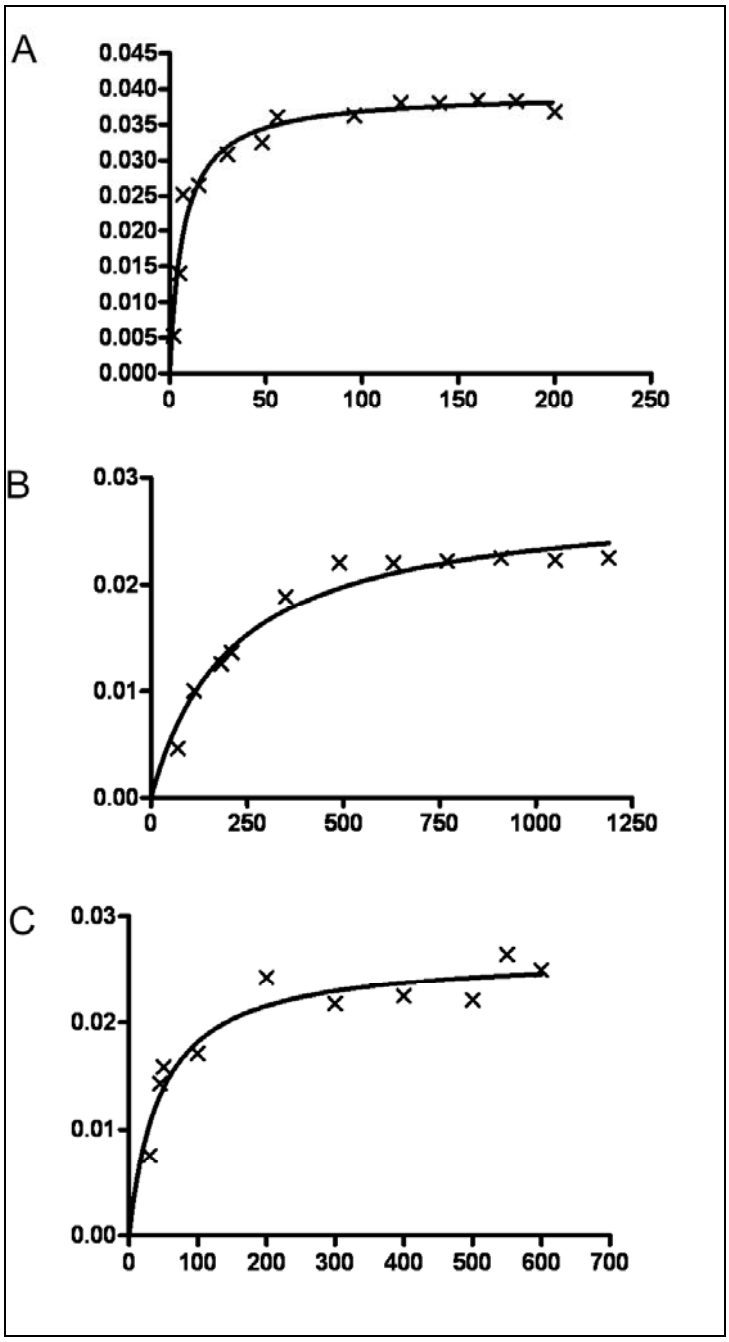

**Steady-State kinetics for Dpo4- with dCTP(A), dTTP ( B), dGTP(C) and dATP (D)**




**Steady-State kinetics for Pol II$^{-}$ with dCTP(A), dTTP ( B), dGTP(C) and dATP (D)**



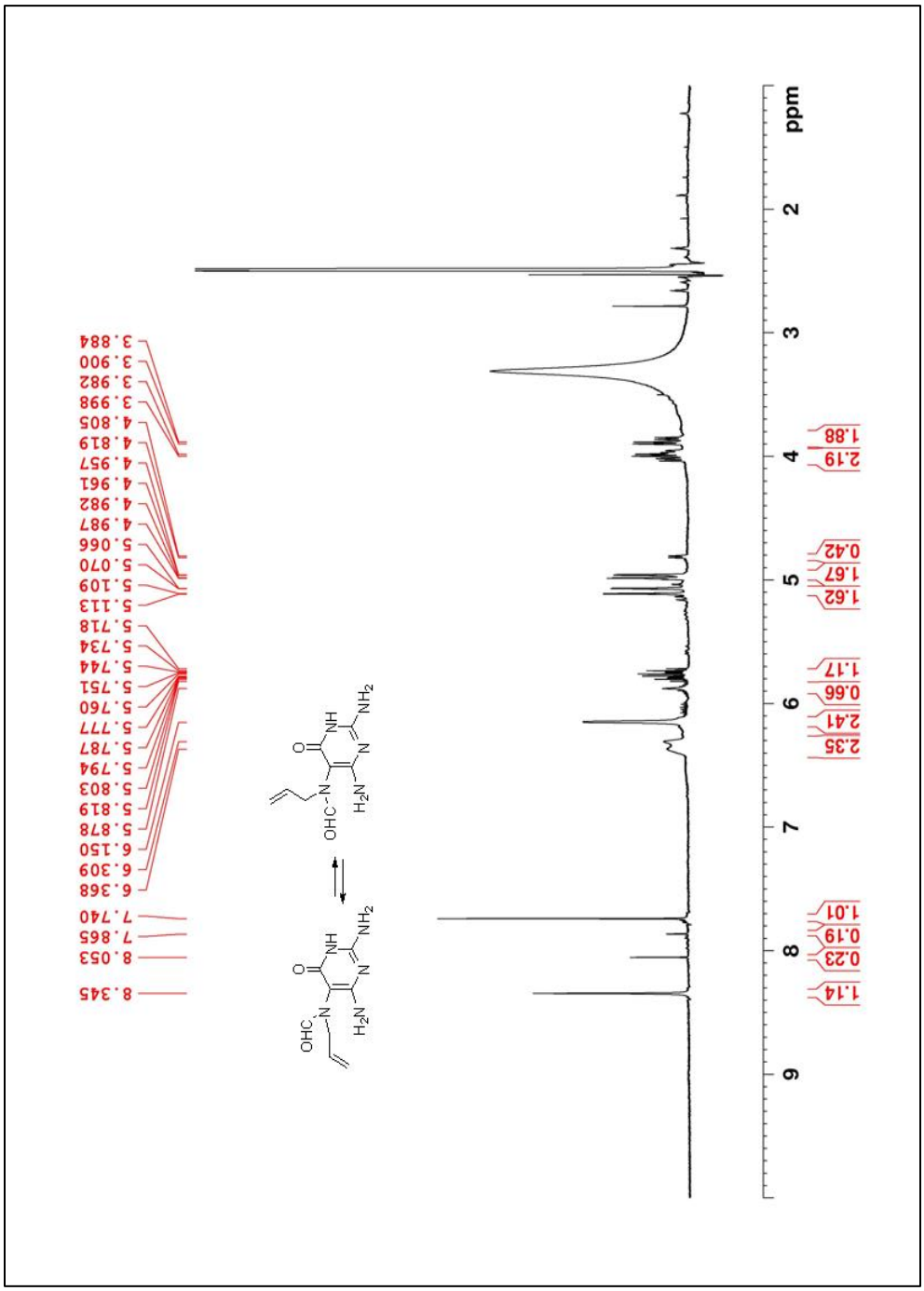

**Steady-State kinetics for pol η with dCTP(A), dTTP ( B), dGTP(C) and dATP (D)**



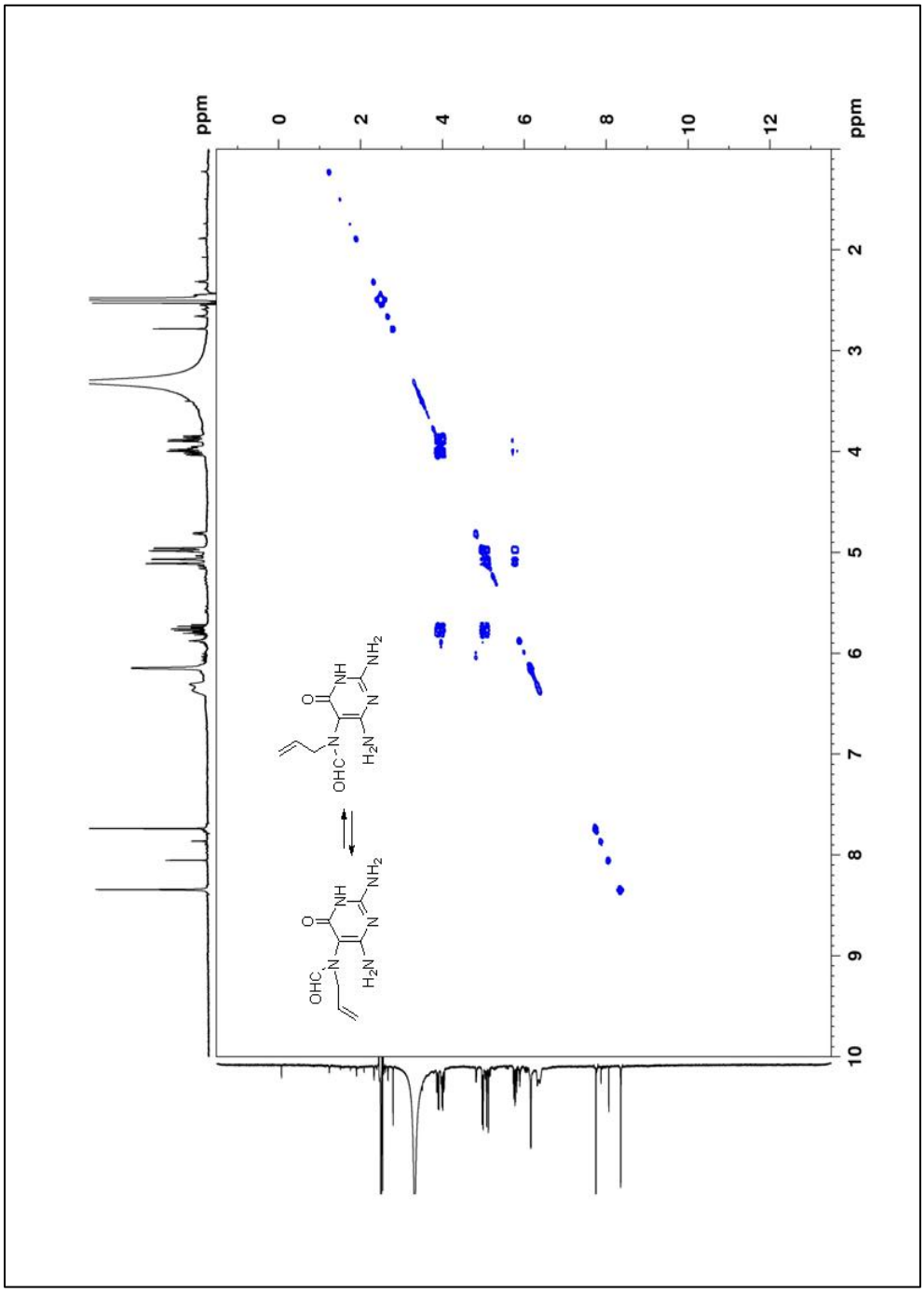

**Steady-State kinetics for pol κ with dCTP(A), dGTP ( B) and dTTP(C)**

# Appendix X

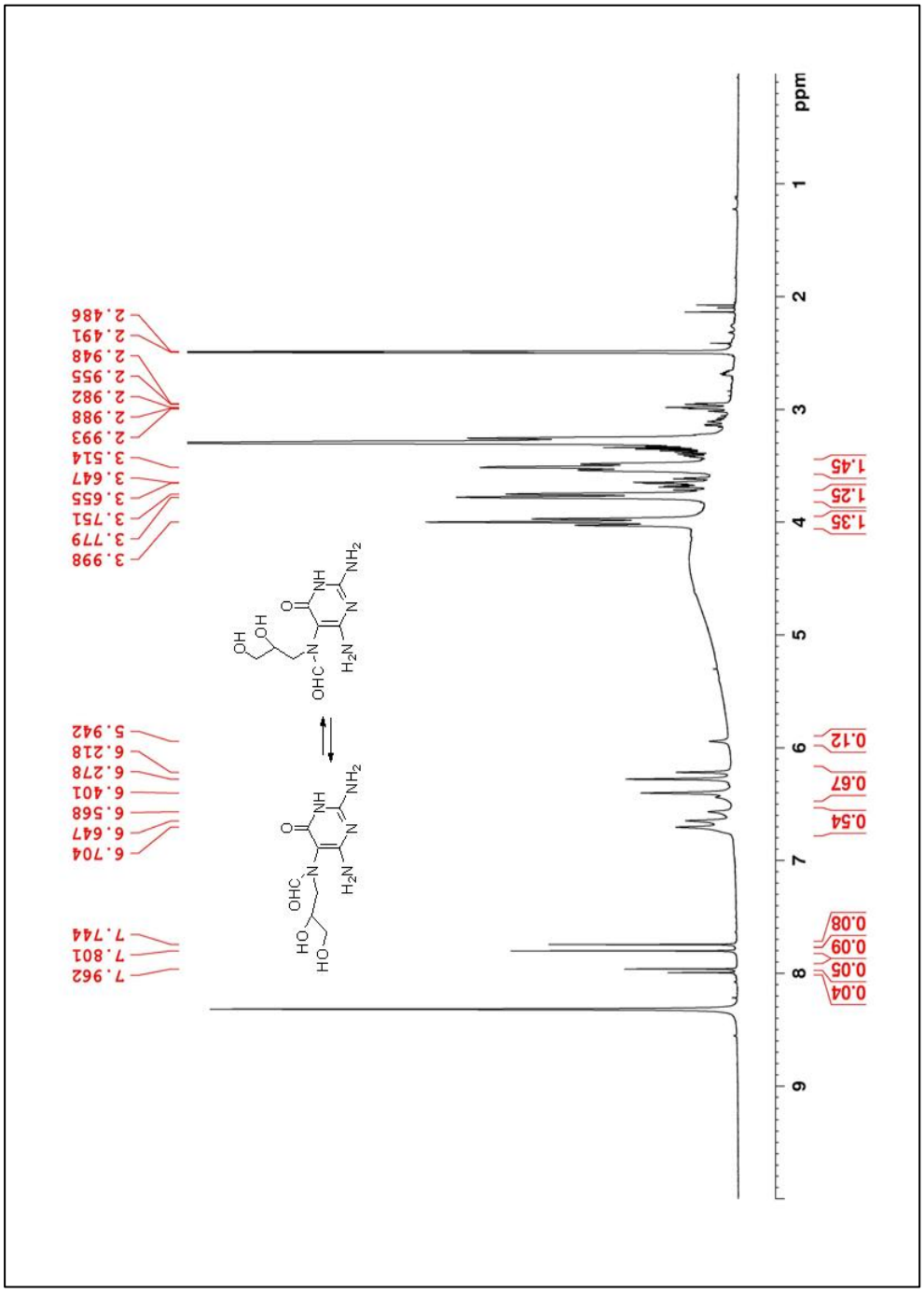
# $^{1}H$, $^{13}C$, COSY, NOESY, HMBC and HSQC NMR Spectra for Chapter III



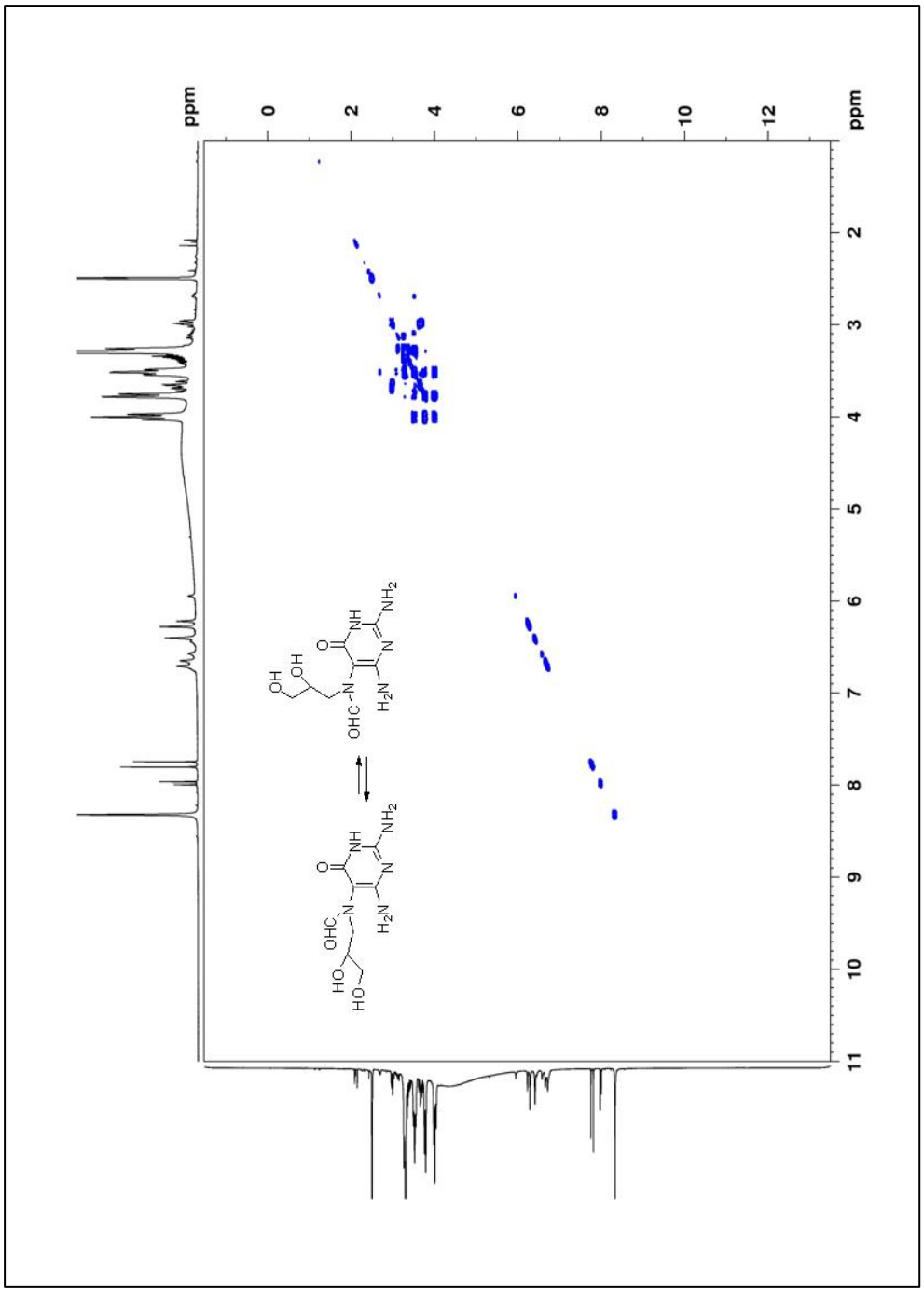

**$^{1}$H NMR of N-(2,6-Diamino-3,4-dihydro-4-oxo-5-pyrimidinyl)-N-(2-propenyl)-formamide (3.33)**



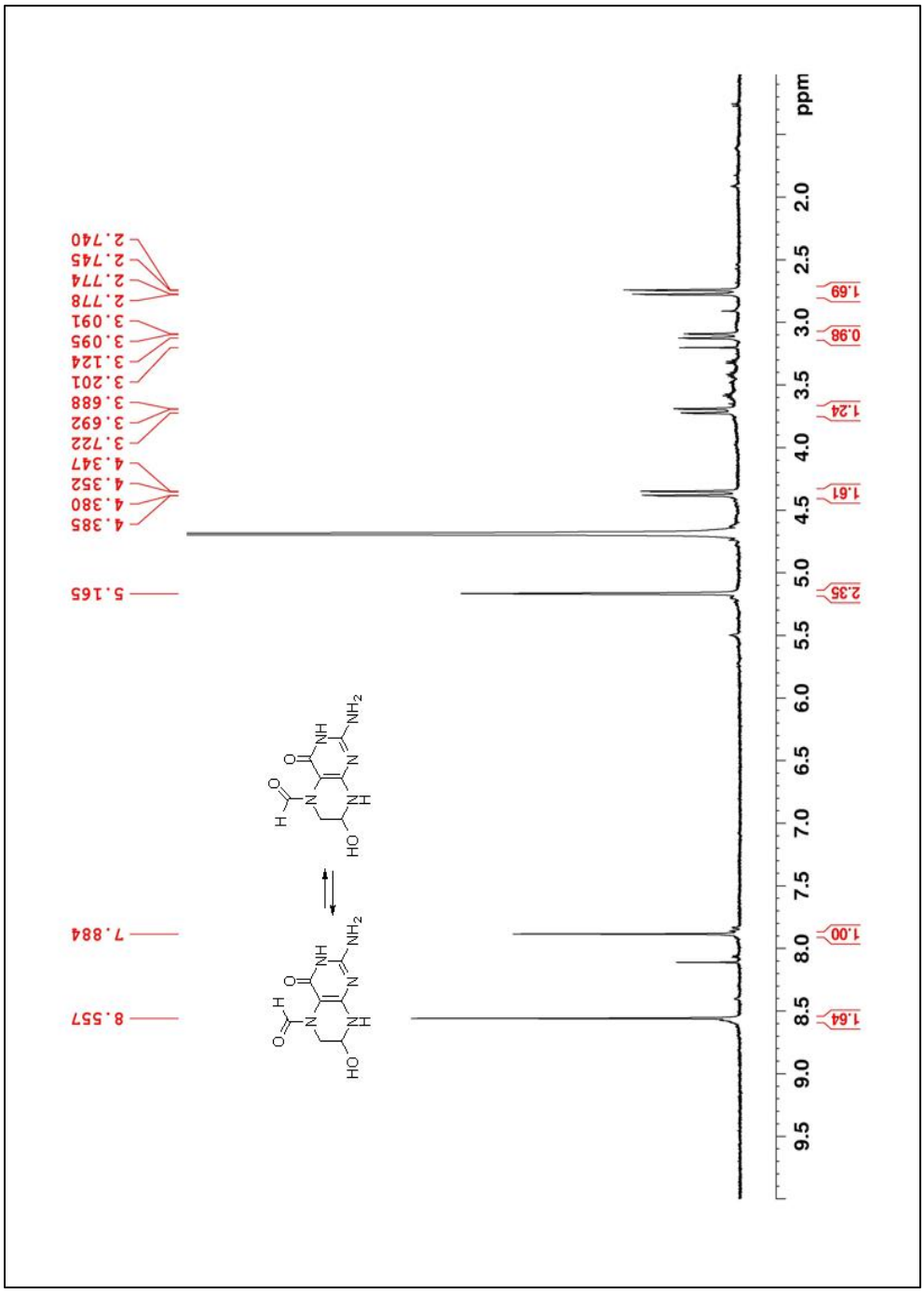

**COSY NMR of N-(2,6-Diamino-3,4-dihydro-4-oxo-5-pyrimidinyl)-N-(2-propenyl)-formamide (3.33)**



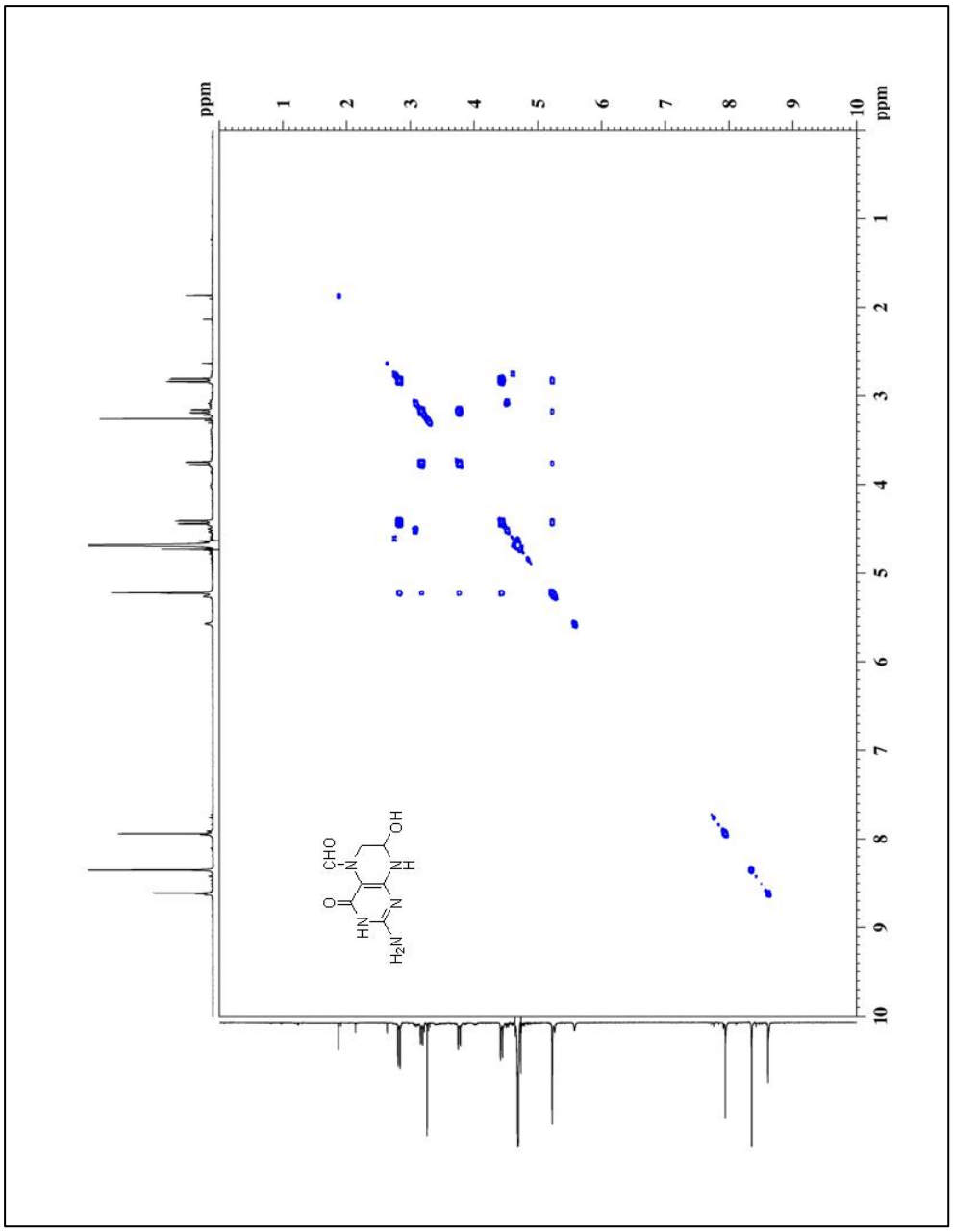

**$^{1}$H NMR of N-(2,6-Diamino-3,4-dihydro-4-oxo-5-pyrimidinyl)-N-(2,3-dihydroxypropyl)-formamide (3.34)**



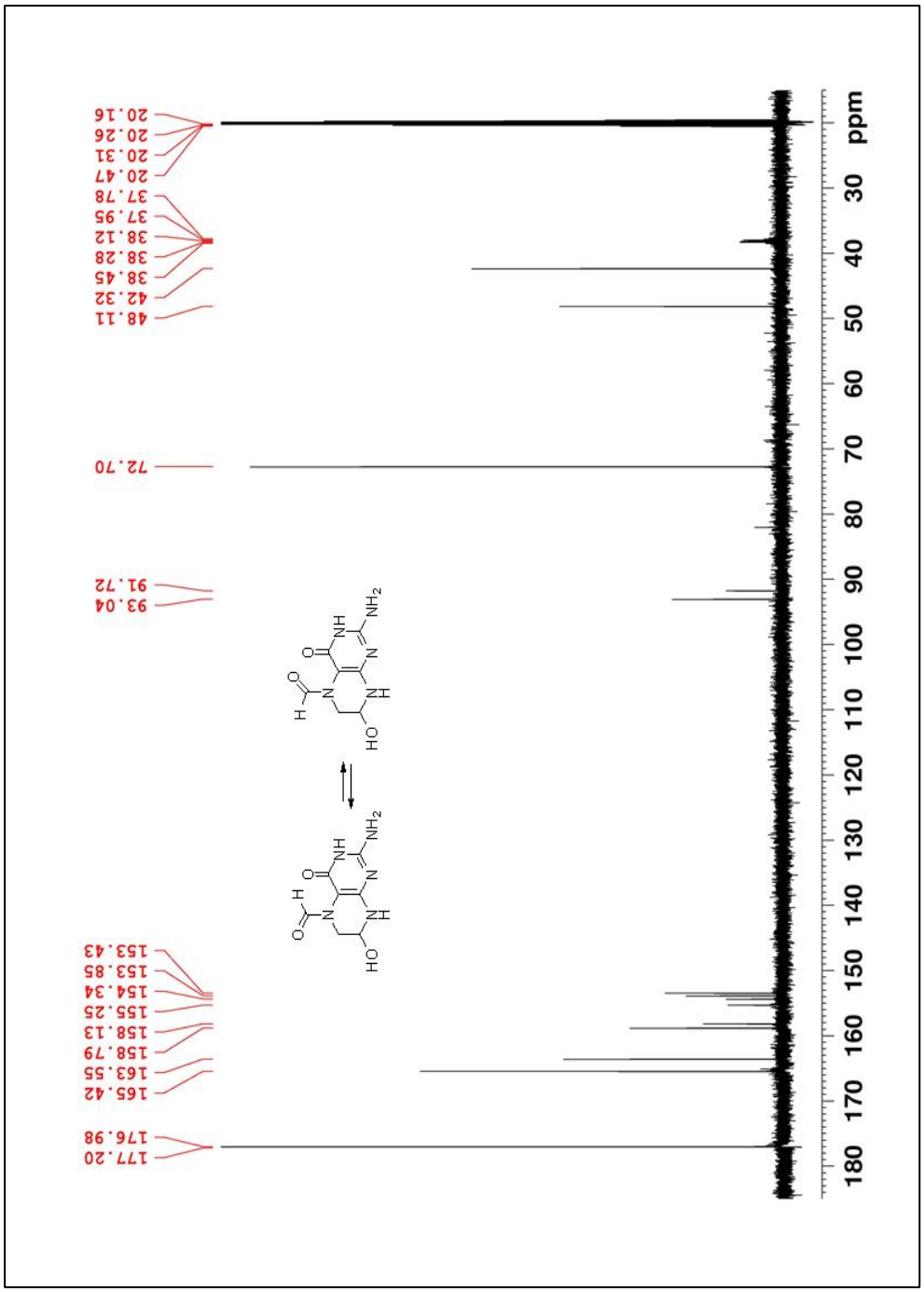

**COSY NMR of N-(2,6-Diamino-3,4-dihydro-4-oxo-5-pyrimidinyl)-N-(2,3-dihydroxypropyl)-formamide (3.34)**



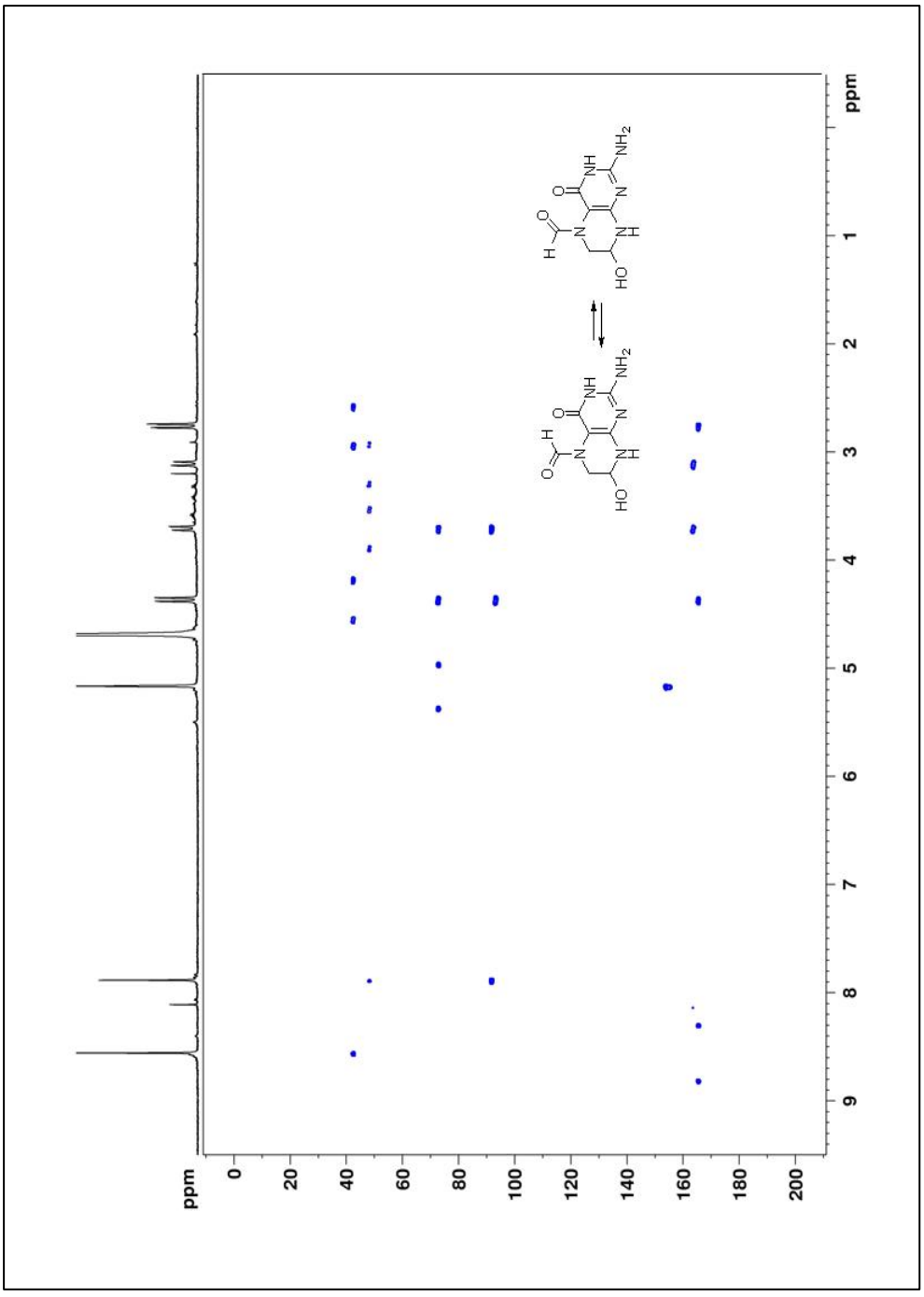

**[1]H NMR of 2-Amino-3,4,7,8-tetrahydro-7-hydroxy-4-oxopteridine-5(6H)-carbaldehyde (3.36)**



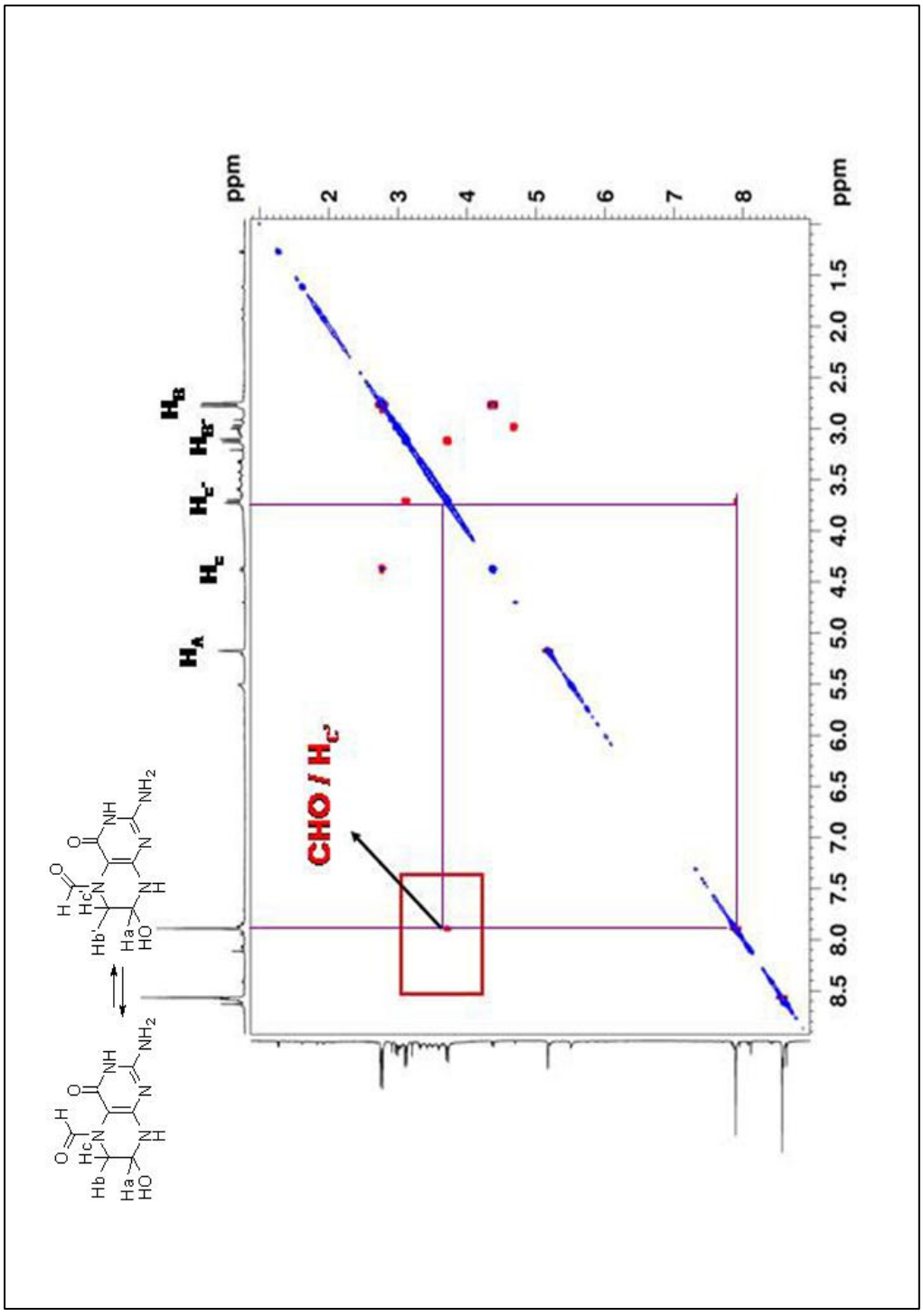

**COSY NMR of N-3-[(dimethylamino)methylene]-3'O-[(2-cyanoethyl)-(N,N-diisopropyl)phosphoroamidite-(5'-O-[bis(4-methoxyphenyl)-phenylmethyl-2'-deoxy-β-D-*erythro*-pentofyranosyl)amino]-3,4-dihydro-4-oxo-5-pyrimidi-nyl]-N-(methyl)-formamide (2.36)**



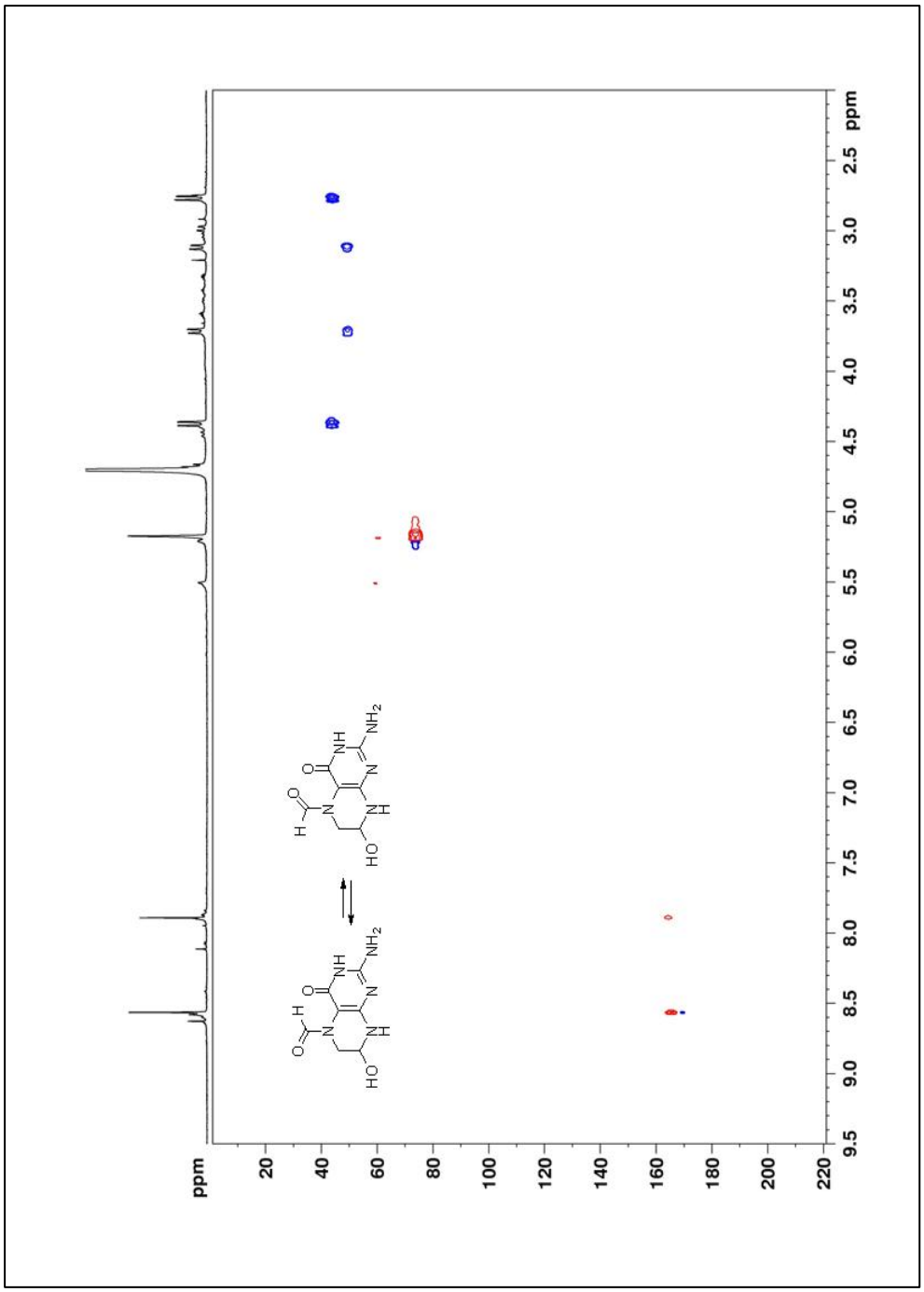

**$^{13}$C NMR of 2-Amino-3,4,7,8-tetrahydro-7-hydroxy-4-oxopteridine-5(6H)-carbaldehyde(3.36)**



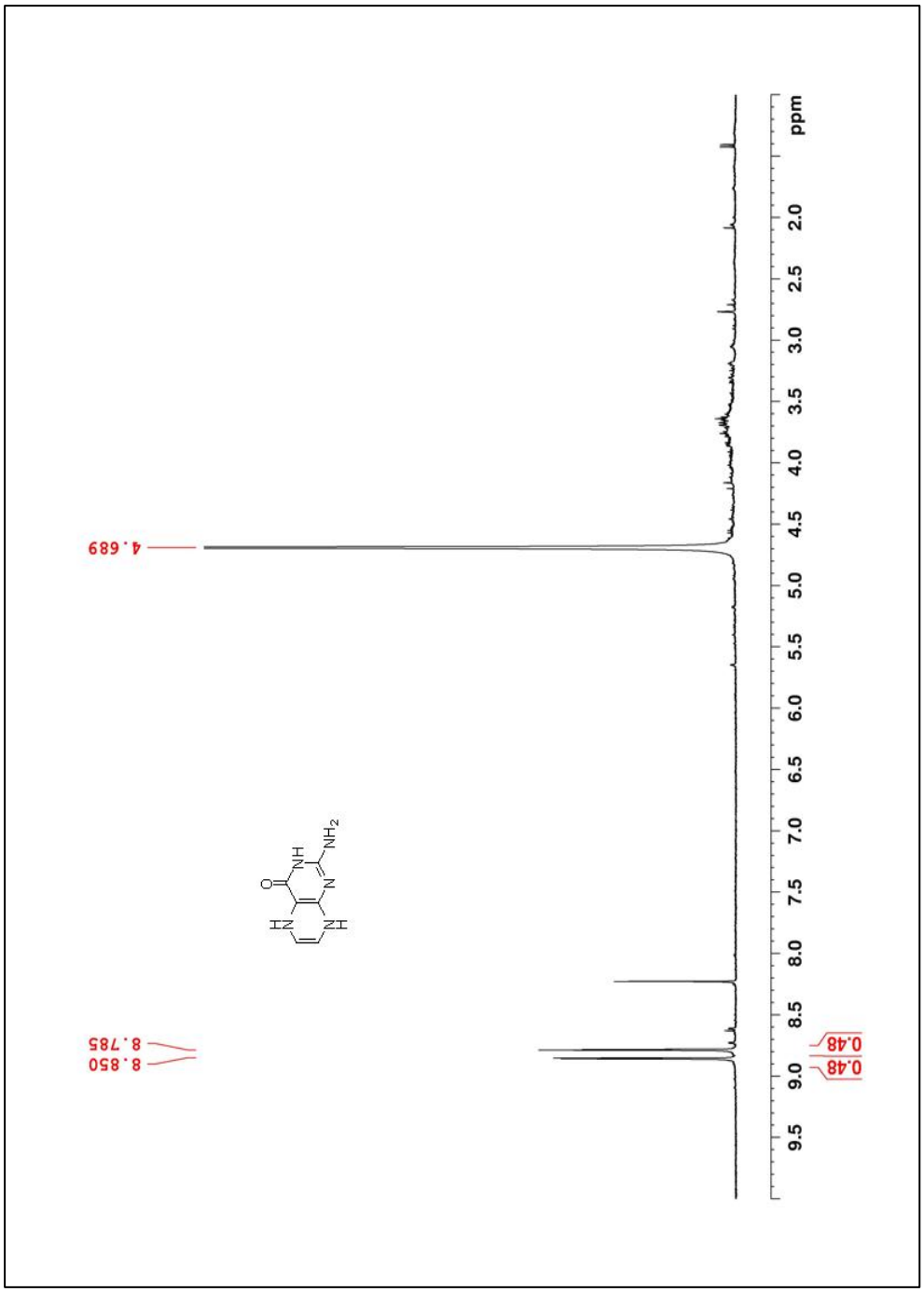

**HMBC NMR of 2-Amino-3,4,7,8-tetrahydro-7-hydroxy-4-oxopteridine-5(6H)-carbaldehyde (3.36)**



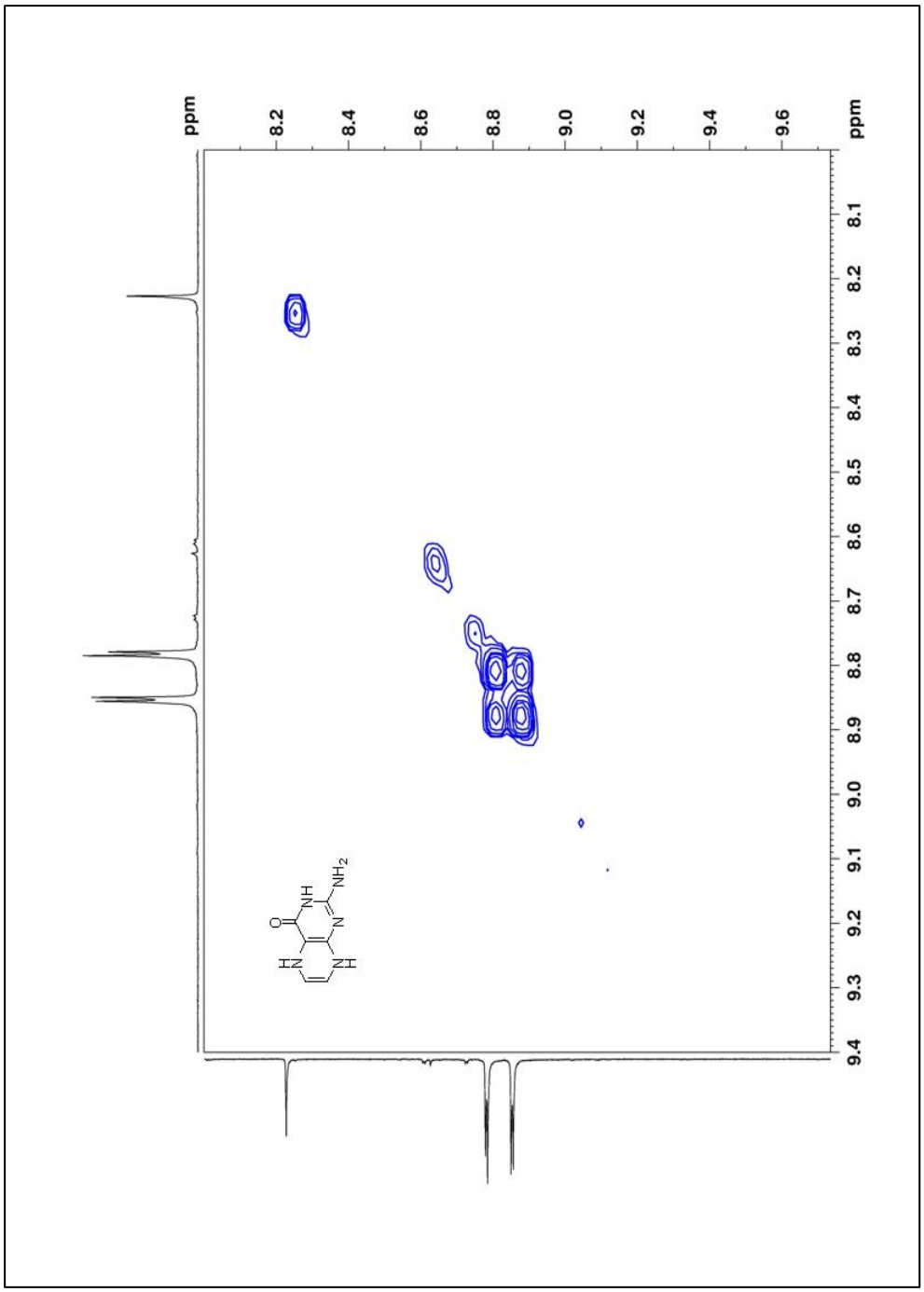

**NOESY NMR of 2-Amino-3,4,7,8-tetrahydro-7-hydroxy-4-oxopteridine-5(6H)-carbaldehyde(3.36)**



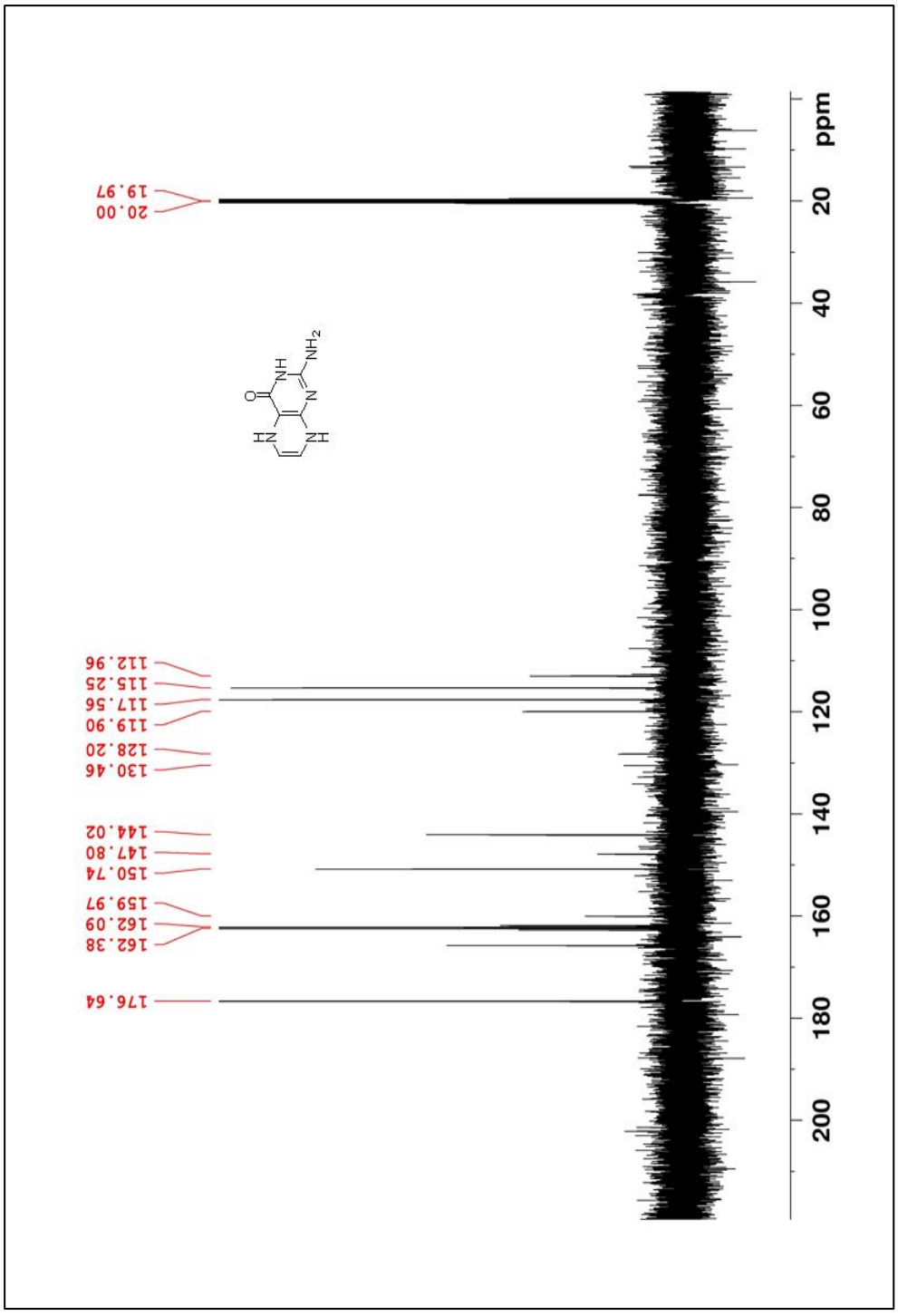

**HSQC NMR of 2-Amino-3,4,7,8-tetrahydro-7-hydroxy-4-oxopteridine-5(6H)-carbaldehyde(3.36)**



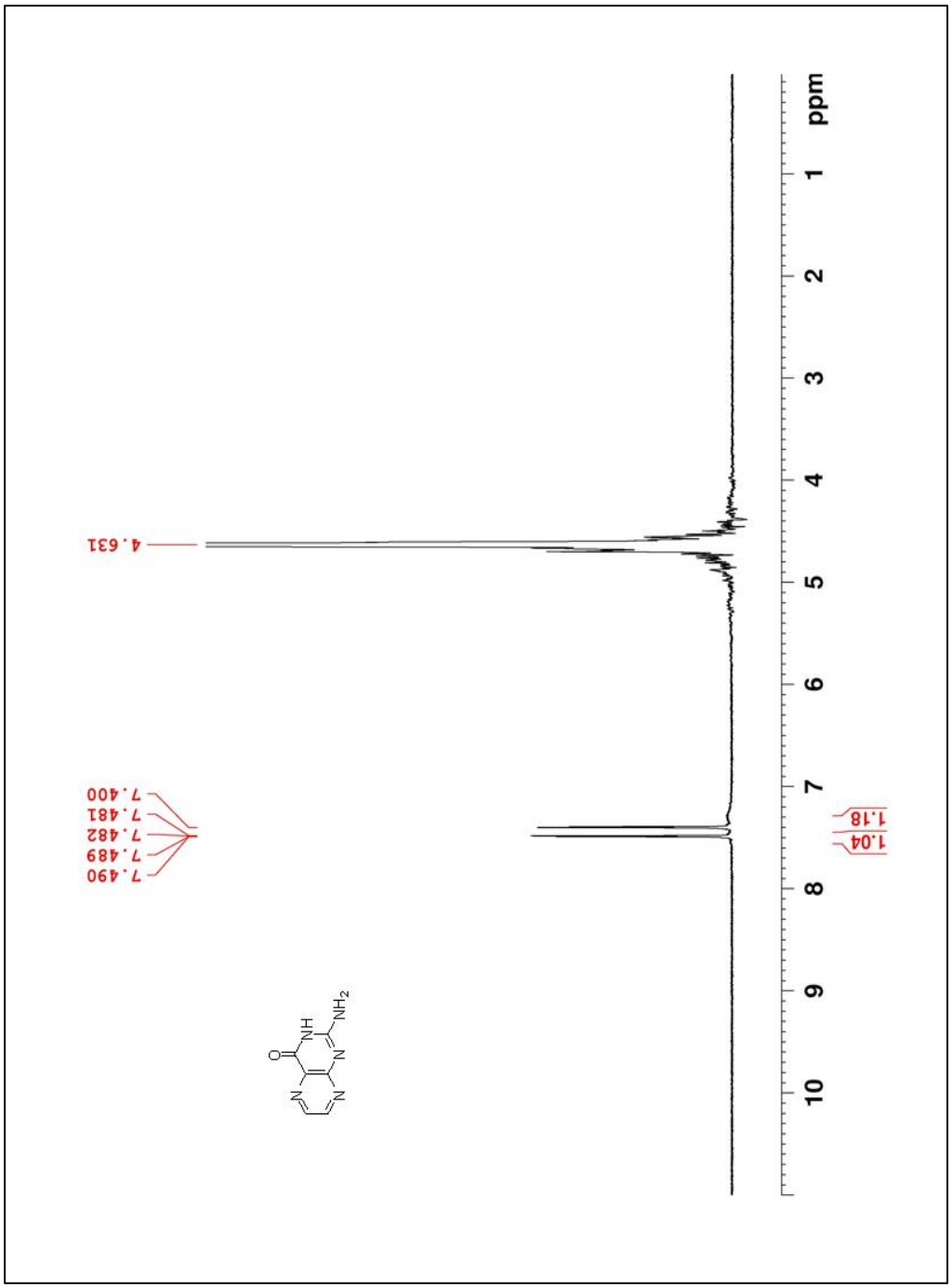

**$^{1}$H NMR of 2-Amino-5,8-dihydro-4(3H)-pteridinone (3.37)**



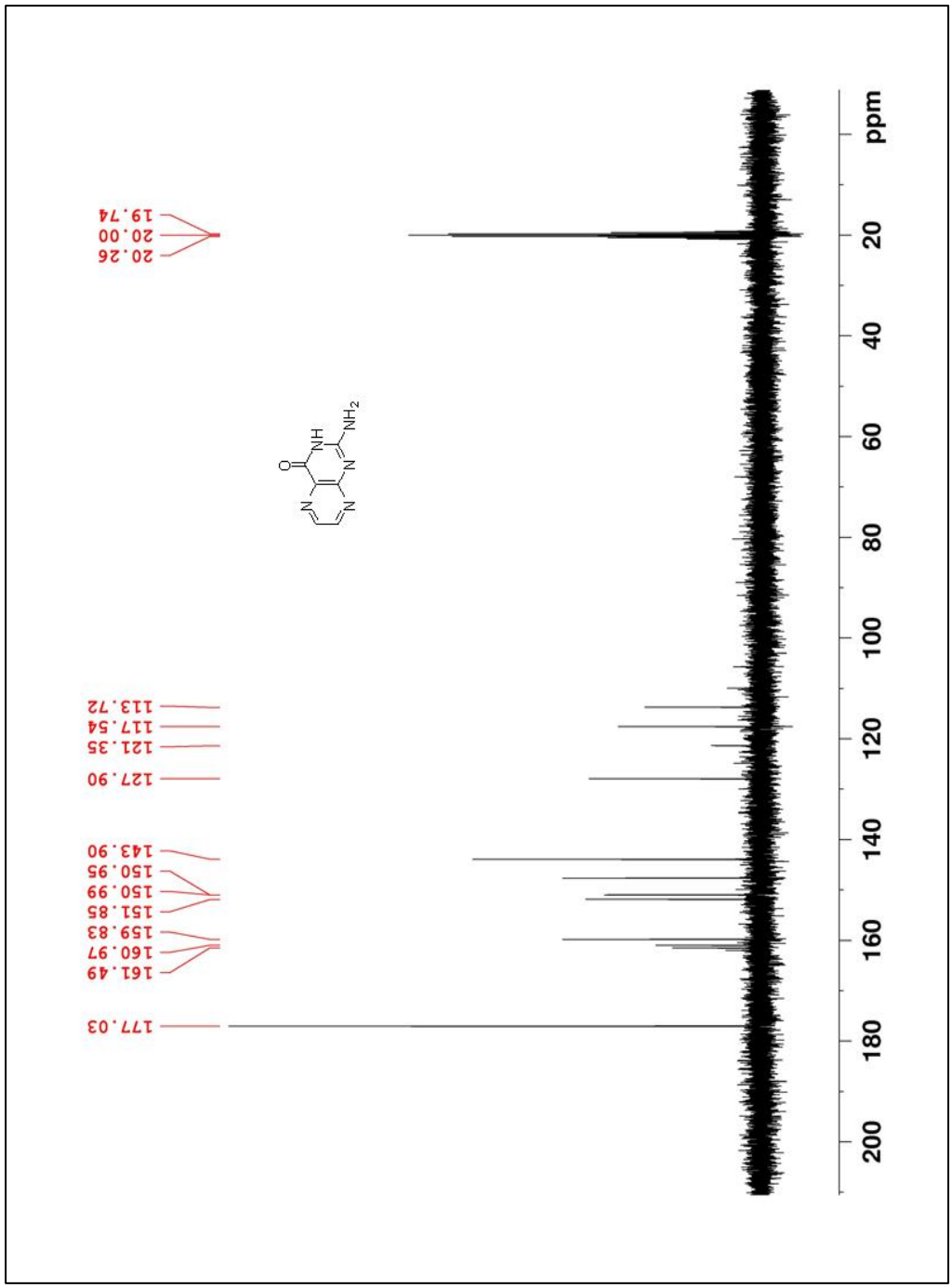

**COSY NMR of 2-Amino-5,8-dihydro-4(3H)-pteridinone (3.37)**



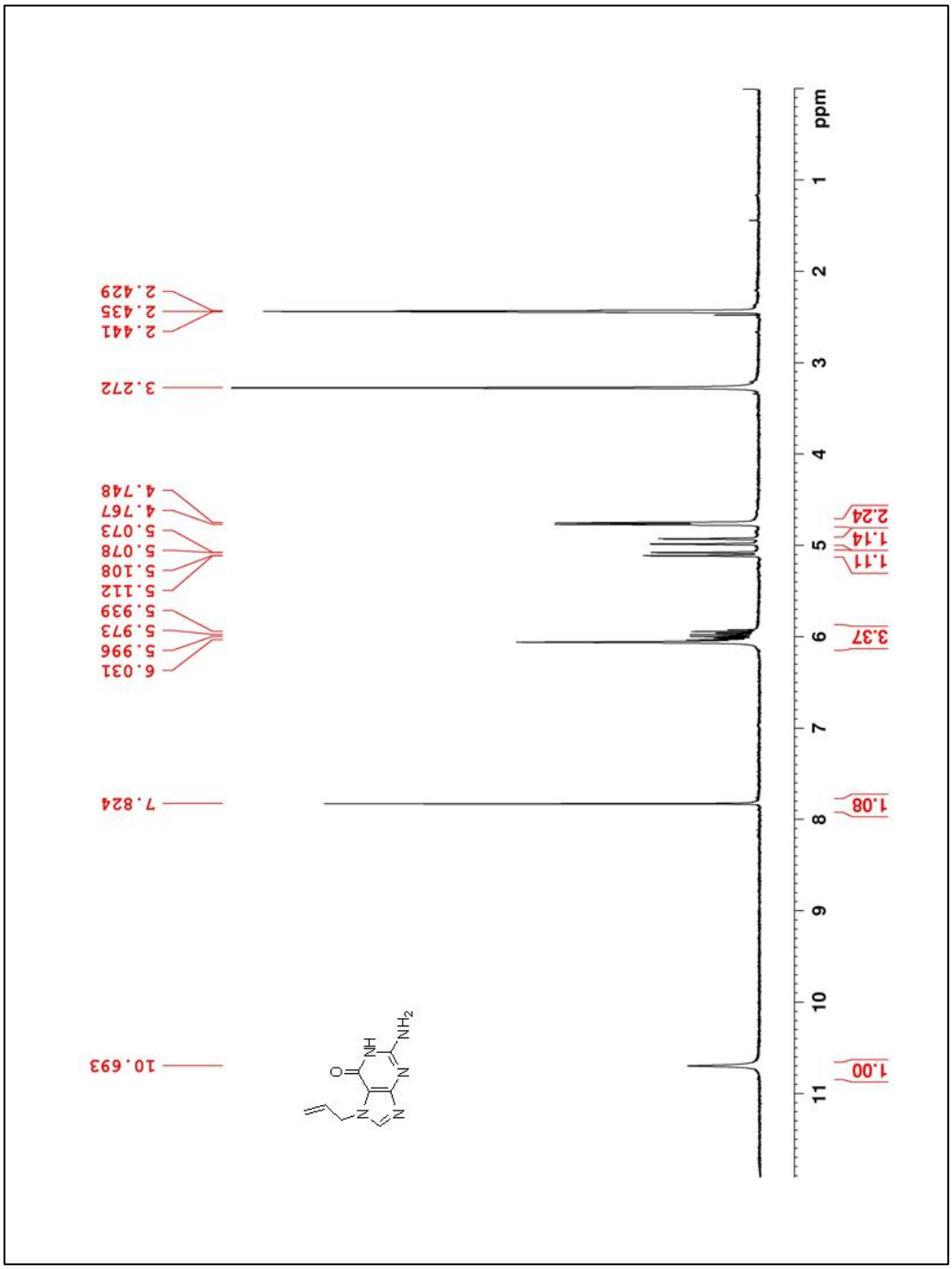

**$^{13}$C NMR of 2-Amino-5,8-dihydro-4(3H)-pteridinone (3.37)**



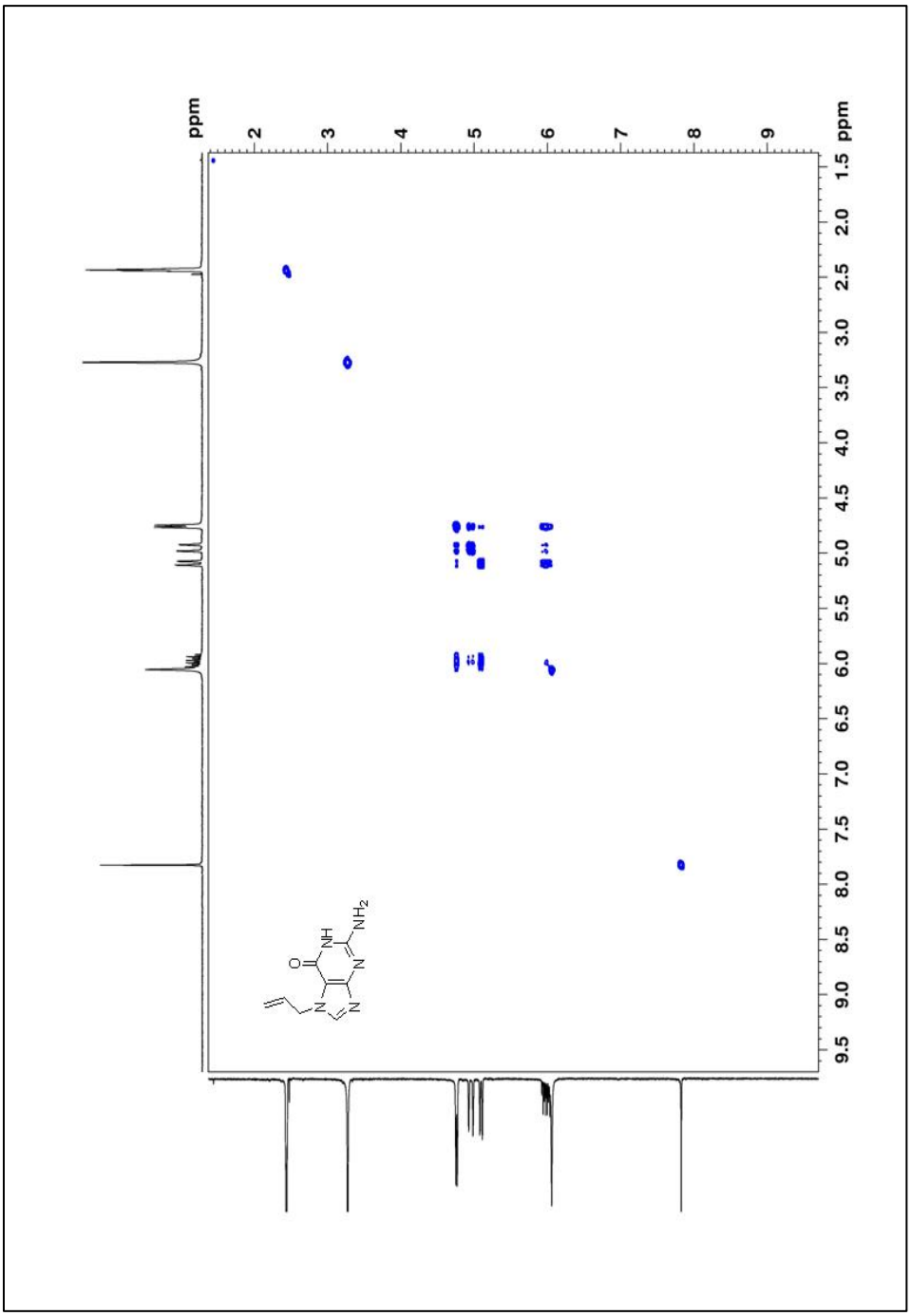

**$^{1}$H NMR of 2-Amino-4-hydroxypyrimido[4,5-b]pyrazine (3.38)**



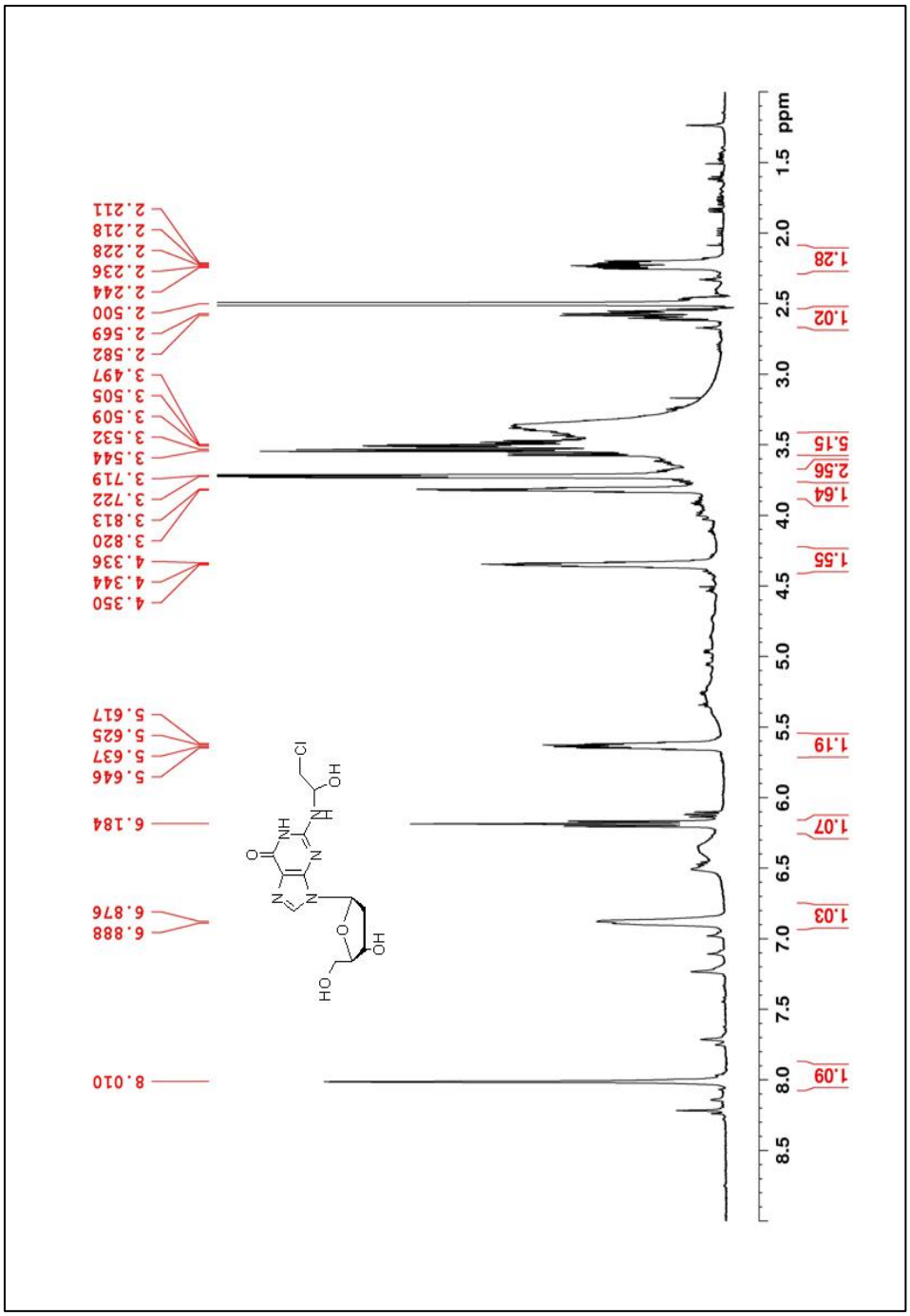

**$^{13}$C NMR of 2-Amino-4-hydroxypyrimido[4,5-b]pyrazine (3.38)**



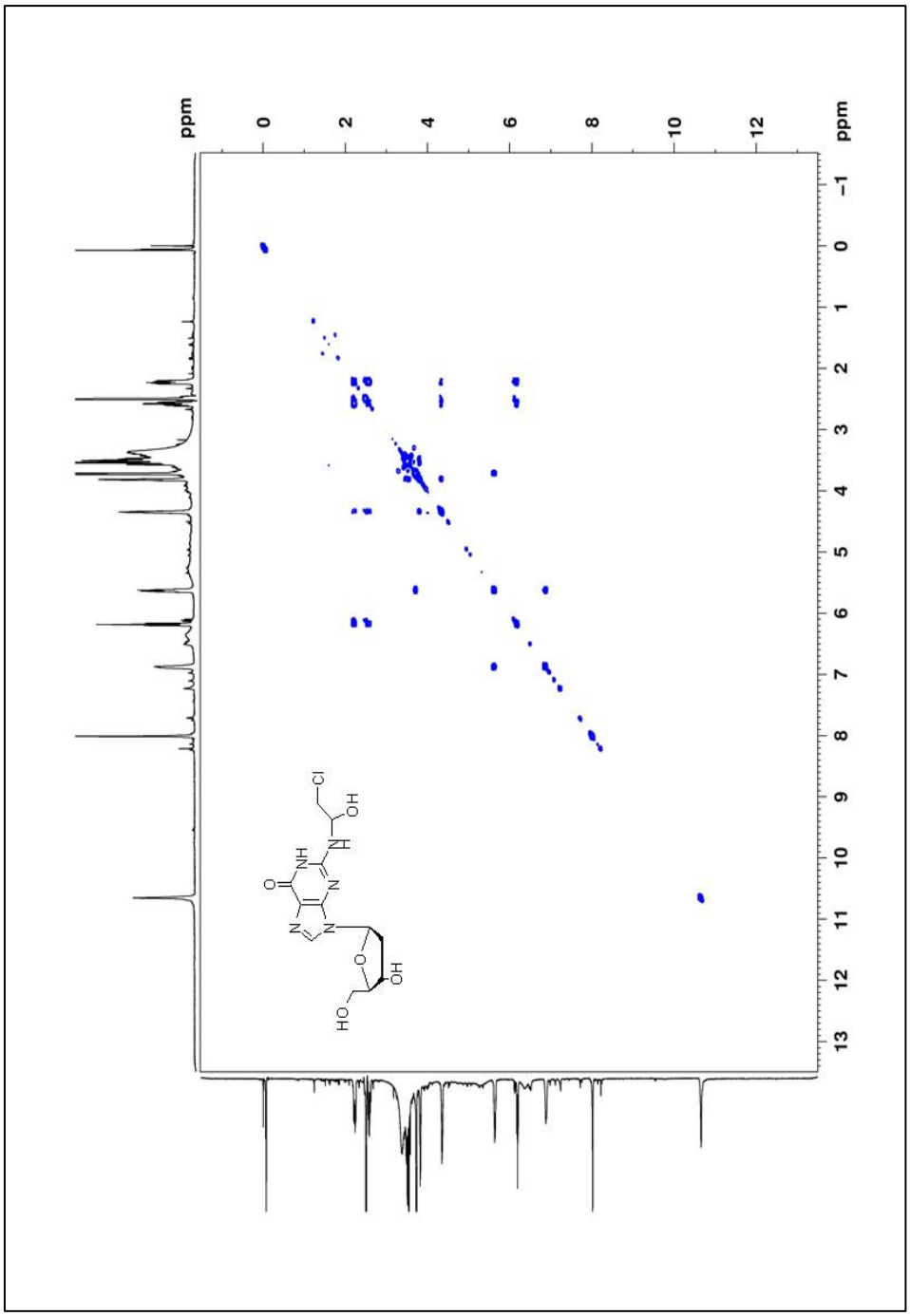

**$^{1}$H NMR of N7-(2-Propenyl)-guanine**




**COSY NMR of N7-(2-Propenyl)-guanine**




$^{1}$H NMR of N2-(2-Chloro-1-hydroxy-ethyl)-(2-deoxy-β-D-erythro-pentopyranosyl)-guanine (3.43)




**COSY NMR of N2-(2-Chloro-1-hydroxy-ethyl)-(2-deoxy-β-D-erythro-pentopyranosyl)-guanine (3.43)**