# Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen

**ABSTRACT**

**SYNTHESIS AND DETECTION OF DNA DAMAGE**

**by**
**Xun Gao**

The biological effects of DNA alkyl adducts are difficult to evaluate at the cellular level due to their instability. Synthesis of oligonucleotides that contain a single N7-alkylguanine has become a vital tool to achieve the above goal. However, the instability of N7-alkyguanines is not compatible with the phosphoramidite chemistry used by solid-phase oligonucleotide synthesis either. Development of chemically stable analogues of unstable DNA lesions enables accurate study of polymerase bypass. The design and successful synthesis of N7-hydroxyethyl-9-deaza-2′-deoxyguanosine and N7-oxoethyl-9-deaza-2′-deoxyguanosine as the stable analogues of N7-hydroxyethyl- 2′-deoxyguanosine and N7-oxoethyl-2′-deoxyguanosine, respectively, are reported. The synthesis of these two nucleosides whose N7 side chains are protected by TBS for the convenience of conversion to phosphoramidites are also developed. The C-glycosidic bonds of these compounds are demonstrated to be stable under strong acidic and basic conditions. These analogues will become versatile tools to study the replication and repair of DNA alkylation damages.

DNA oxidation product 8-oxoGua has been suggested as a biomarker for early cancer diagnosis. An artificial receptor for the free base of 8-oxoGua on a triplex DNA backbone was previously developed. However, accurate detection of 8-oxoGua in urine samples was affected by the presence of a large excess of guanine. Herein, a unique strategy to convert such a receptor to a colorimetric biosensor by conjugating DNA strands to gold nanoparticles (GNP)

is developed. Binding of 8-oxoGua to the receptor caused the conjugation of GNP, resulting in diagnostic red-to-purple color changes. The presence of multiple binding cavities enhances the binding-induced stabilization effect and widened the temperature window used for detection. By simply incubating our sensor with a sample, 8-oxoGua can be detected at submicromolar concentrations with a UV–vis spectrometer or even by naked eye. The detection limit in a urine matrix is determined as 126 nM and the response range covers a major portion of the biologically relevant concentration range.

# SYNTHESIS AND DETECTION OF DNA DAMAGE

**by**
**Xun Gao**

**A Dissertation**
**Submitted to the Faculty of**
**New Jersey Institute of Technology**
**in Partial Fulfillment of the Requirements for the Degree of**
**Doctor of Philosophy in Chemistry**

**Department of Chemistry and Environmental Science**

**May 2017**

**APPROVAL PAGE**

**SYNTHESIS AND DETECTION OF DNA DAMAGE**

**Xun Gao**

---

Dr. Edgardo T. Farinas, Dissertation Advisor — Date
Associate Professor of Chemistry and Environmental Science, NJIT

---

Dr. Haidong Huang, Dissertation Co-advisor — Date
Senior Scientist of Quantum-Si Inc.

---

Dr. Xiaoyang Xu, Committee Member — Date
Assistant Professor of Chemical Biological and Pharmaceutical Engineering, NJIT

---

Dr. Tamara M. Gund, Committee Member — Date
Professor of Chemistry and Environmental Science, NJIT

---

Dr. Yong-Ick Kim, Committee Member — Date
Assistant Professor of Chemistry and Environmental Science, NJIT

---

Dr. Pradyot Patnaik, Committee Member — Date
Consultant, Radiance, New Jersey

## BIOGRAPHICAL SKETCH

**Author:** **Xun Gao**

**Degree:** **Doctor of Philosophy**

**Date:** **May 2017**

**Undergraduate and Graduate Education:**

- Doctor of Philosophy in Chemistry,
  New Jersey Institute of Technology, Newark, NJ, 2017

- Master of Science in Chemistry,
  Xiamen University, Xiamen, Fujian, P. R. China, 2011

- Bachelor of Science in Applied Chemistry,
  Anhui University, Hefei, Anhui, P. R. China, 2008

**Major:** Chemistry

**Presentations and Publications:**

Xun Gao and Haidong Huang, (2017) Synthesis of N7-alkyl-9-deazaguanosines containing polar N7 chains. Examples of chemically stable analogues of N7-hydroxyethyl and N7-oxoethyl adducts of 2'-deoxyguanosine, Journal of Organic Chemistry, 81, 11697-11705.

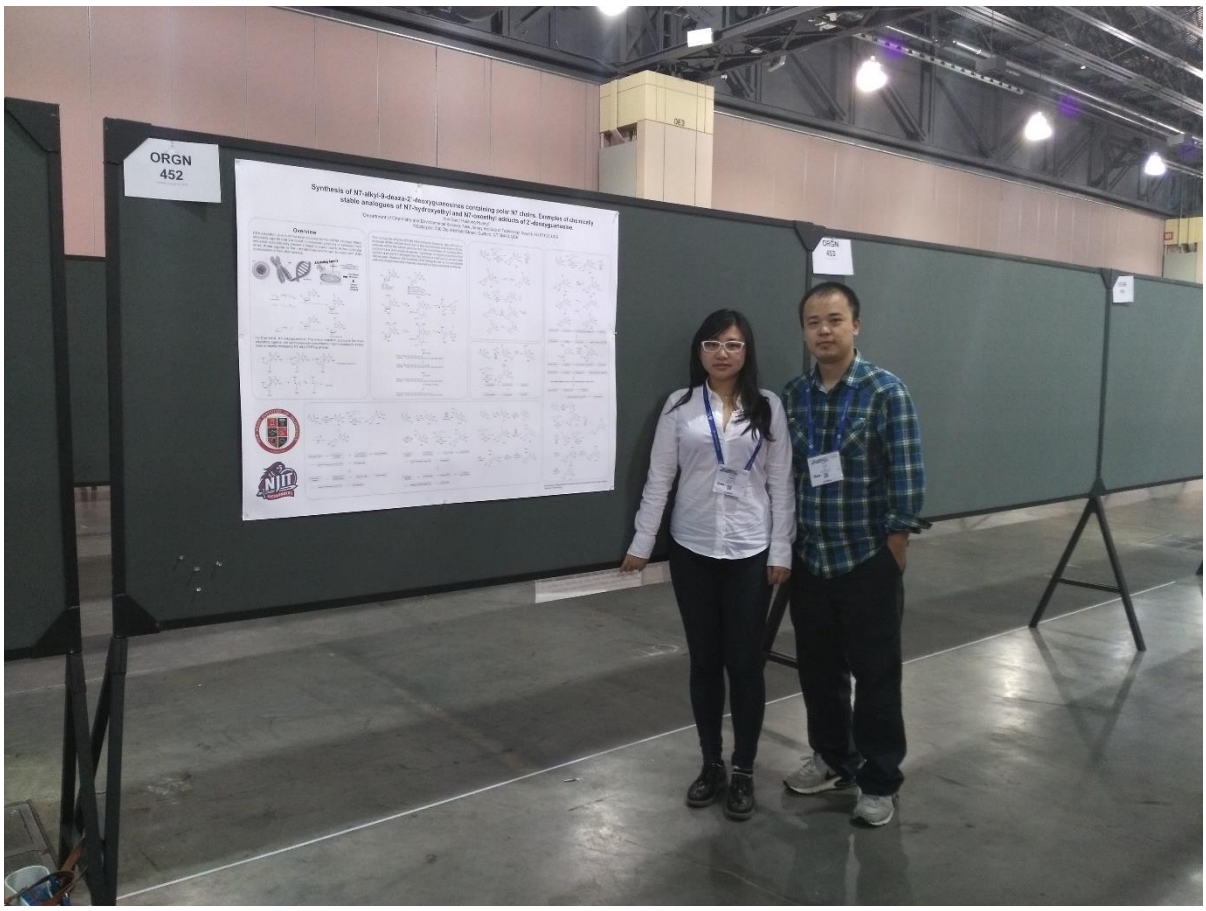
Xun Gao, Yung-Hao Tsou, Marina Garis, Haidong Huang, Xiaoyang Xu, (2016) Highly specific colorimetric detection of DNA oxidation biomarker using gold nanoparticles/triplex DNA conjugates, Nanomedicines: Nanotechnology, Biology and Medicine, 12, 2101-2105.

Jagruti Rana, Xun Gao and Haidong Huang, (2013) Polymerase bypass of Chemically stable analogues of N7-alkylguanine, the most abundant DNA alkylation damage products, 9th Annual Graduate Student Research Day, New Jersey Institute of Technology, Newark, NJ.

Jagruti Rana, Xun Gao and Haidong Huang, (2014) Polymerase bypass of Chemically stable analogues of N7-alkylguanine, the most abundant DNA alkylation damage products, Organic Chemistry Conference in Rutgers, New Brunswick, NJ

Xun Gao and Haidong Huang, (2015) Synthesis of N7-alkyl-9-deazaguanosines containing polar N7 chains. Examples of chemically stable analogues of N7-hydroxyethyl and N7-oxoethyl adducts of 2'-Deoxyguanosine, 11th Annual Graduate Student Research Day, New Jersey Institute of Technology, Newark, NJ.

Xun Gao and Haidong Huang, (2016) Synthesis of N7-alkyl-9-deazaguanosines containing polar N7 chains. Examples of chemically stable analogues of N7-hydroxyethyl and N7-oxoethyl adducts of 2'-Deoxyguanosine, 252nd American Chemical Society National Meeting & Exposition, Philadelphia Convention Center, Poster Presentation of Organic Section, Philadelphia, Pennsylvania.




感谢老婆陈媛一直以来的付出和陪伴。此生有你无憾。

感谢家人的支持和鼓励。作为爸妈的儿子，我很骄傲。

感谢黄海东老师的关心和指导。

人生之路，皆是修行，如是而已。

## ACKNOWLEDGEMENT

It is my great fortune to join the group of Dr. Haidong Huang during my PhD study. You made me so proud as one of your students. The research project is not easy, but with you standing by my side and guiding me with your wisdom and learnedness, I can finally come this far today. Besides, it is your great passion to science that encourages me to keep learning, your deliberate thinking that enlightens me to be creative and visionary, and your humble personality that educates me to be what kind of person. When I encountered some difficulties during the synthesis of the target molecule, I could stay focus only because I know you always trust me. Even when you were busy working you still spent the weekend for the tedious proof reading for me. There are so many other things I would like to bring out here, and in a word I am so thankful that I could join your lab in NJIT, that would be one of my most precious experiences in my life.

I joined Prof. Edgardo Farinas's lab at my fifth academic year. I have to give my special thanks to Prof. Farinas for providing me great opportunity to study another different but exciting topic. His enthusiasm to science as being a professor along with great management skills as being a chairman make a great example of success to me.

In my fourth year, I really appreciated to have a seat in Prof. Xiaoyang Xu's lab and learn from him. Your perception and determination in research as your profession motivated me to pursue my dream no matter what it takes.

I would like to thank Prof. Mengyan Li here, for your kind help when I was lack of basic knowledge of the research project. You and your students generously shared a lot of guidance, thank you very much.

I also want to say thank you to Prof. Tamura Gund, Prof. Yong-Ick Kim and Dr. Pradyot Patnaik, for being my committee member, you make me proud of a student in NJIT.

I would like to thank Yogesh Ghandi and Pin Gu, for your generous help; Sylvana Brito and Genti Price, for your great work to take care of us as graduate students; Frank Ellis and Miriam Gulotta, for your consideration when I was your TA, it was so nice working with you.

I would like say thank you to all of my fellows and friends here: Jagruti Rana, Peter Tlatelpa, Zixing Zhuang, Yung-Hao Tsou, Silu Sheng, Joydeep Chakraborty, Han Jia, Daiyong Deng, Fei Li. I want to say it is so wonderful to work with you, it means a lot to me.

I want to say that during this five years studying abroad, my parents were very proud of me as a son, of being whom I am so proud as well. In the end, I want to say to my wife, Yuan Chen, that you are the very reason I am trying to be better. Since I met you, I finally found the meaning of my life. Love you forever.

## TABLE OF CONTENTS

**TABLE OF CONTENTS**
**(Continued)**

**Chapter** **Page**

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF FIGURES
## (Continued)

## LIST OF ABBREVIATIONS

| | |
|---|---|
| DNA | Deoxyribonucleic acid |
| ROS | Reactive oxygen species |
| 8-OxodG | 8-Oxo-2'-deoxyguanosine |
| BER | Base excision repair |
| NER | Nucleotide excision repair |
| TLS | Translesion synthesis |
| 8-OxoGua | 8-Oxoguanine |
| AP | Apurinic sites |
| N7-HEG | N7-hydroxyethyl guanine |
| N7-OEG | N7-oxoethyl Guanine |
| $^{7HE9C}$dG | N7-hydroxyethyl-9-deaza- 2'-deoxyguanosine |
| $^{7OE9C}$dG | N7-oxoethyl-9-deaza -2'-deoxyguanosine |
| DMF | N,N-dimethylformamide (solvent) |
| dmf | N,N-dimethylformamide (protecting group) |
| DBU | 1,8-Diazabicyclo[5.4.0]undec-7-ene |
| DIPEA | N,N-diisopropylethylamine |
| 4-DMAP | 4-Dimethylaminopyridine |
| TBHP | tert-Butyl hydroperoxide |
| TBAF | Tetra-n-butylammonium fluoride |

## LIST OF ABBREVIATION
## (Continued)

| | |
|---|---|
| TBS | tert-Butyldimethylsilyl ether |
| Bn | Benzyl Group |
| THF | Tetrahydrofuran |
| OGG1 | Oxoguanine DNA glycosylase 1 |
| SPR | Surface plasmon resonance |
| AuNP | Gold Nanoparticle |

# CHAPTER 1

# GENERAL INTRODUCTION

## 1.1 DNA Damage and Mutation

Deoxyribonucleic acid (DNA) is the hereditary material in humans and almost all other organisms. The DNA code contains instructions needed to make the proteins and molecules essential for our growth, development and health. DNA is made up of molecules called nucleotides. Each nucleotide contains a phosphate group, a sugar group and a nucleobase. The four types of nucleobases are adenine (A), thymine (T), guanine (G) and cytosine (C). Nucleotides are attached together to form two long strands that spiral to create a structure called a double helix. The bases on one strand pair with the bases on another strand: adenine pairs with thymine (two hydrogen bonds in the center of DNA strand), and guanine pairs with cytosine (three hydrogen bonds in the center of DNA strand). Two factors are mainly responsible for the stability of the DNA double helix: base pairing between complementary strands and stacking between adjacent bases. The sequence of the nucleobases, A, C, G and T, in DNA determines our unique genetic code and provides the instructions for producing molecules in the body, stress the importance of the integrity and stability of the DNA. Unfortunately, DNA molecules are susceptible to the attack, as research showed that among ~$10^{13}$ cells in the human body there were tens of thousands of DNA lesions per day.[1] The vast majority of DNA damage in human tissues is certainly of endogenous origin. Endogenous damage arises from endogenous cellular processes, including hydrolysis and oxidative damage caused by reactions with reactive oxygen species (ROS). Hydrolytic

DNA damage involves deamination or the total removal of individual bases. Loss of DNA bases, known as AP (apurinic/apyrimidinic) sites, can be particularly mutagenic and if left unrepaired they can inhibit transcription. Oxidative DNA damage refers to the oxidation of specific bases. 8-Hydroxydeoxyguanosine (8-oxodG) is the most common marker for oxidative DNA damage and can be measured in virtually any species. It is formed most often by chemical carcinogens.

Faced with DNA damage, cells always elicit DNA damage response (DDR) consist of DNA damage recognition, signal transduction, transcriptional regulation, cell cycle control and DNA repair to attenuate DNA damage. Modified bases, abasic site, and the DNA single-strand breaks are repaired by base excision repair (BER) pathway. DNA double-strand breaks are dealt with homologous recombination repair (HRR) and non-homologous end-joining (NHEJ) pathways. DNA mismatch repair is a system for recognizing and repairing erroneous insertion, deletion, and misincorporations of bases. The nucleotide excision repair (NER) pathway deals with modified nucleotides that distort the structure of the double helix. In addition to these major pathways, nucleotide damage in the form of adducts that can block replication can be bypassed by a mechanism known as translesion synthesis.

The bulk of DNA damage is repaired promptly and accurately by the cellular repair machinery. This machinery, however, is not perfect (becoming even less so with ageing), and it may occasionally miss sites of damage or make mistakes - for example, an oxidation damage such as 8-oxodG. 8-OxodG are usually recognized and repaired by a base excision repair system. But if the DNA template containing 8-oxodG is replicated despite the presence of damage, the resulting daughter strands will carry different 'meaning' in their DNA (G to T

transversion). During the replication cycle, the altered DNA molecule will produce daughter molecules carrying the altered sequence, transmitting it on to their progeny. Eventually, a mutant clone is created. Mutations are changes in the nucleotide sequence, involving deletions, insertions, substitutions or rearrangements of base pairs, and can lead to dysfunctional proteins. As a permanent alteration in the DNA sequence that makes up a gene, either hereditary mutations inherited from a parent or acquired mutations caused by environmental factors, are common enough to be considered as a normal variation in DNA. Mutations are essential to evolution. Some of the mutations are beneficial, which keeps populations healthy, some have no effect at all and some we hear about most often are ones cause diseases such as cystic fibrosis, sickle cell anaemia, Tay-Sachs disease, phenylketonuria and color-blindness. Genomic stability depends on an efficient DNA damage repair system to keep the chromosomes intact. Unrepaired DNA damage not only causes cell cycle arrest, apoptosis but also accumulates genome mutations. It has been postulated that the two main causes of cellular ageing are the accumulation of mutations in DNA resulting from replication errors and the oxidative stress resulting from the genotoxic action of reactive oxygen species (ROS) on mitochondrial DNA.

### 1.2 DNA Alkylation and Oxidation Formation

Almost 20,000 lesions happened one day per cell, including mainly spontaneous depurination, reactive oxygen species damage and deamination. Even though the contribution of alkylation damage is unclear, the mere existence of specific repair system for this damage justifies the importance of the understanding interaction of exogenous agents and influences

with endogenous processes in the induction of cancer and other diseases. Alkylating agents are ubiquitous. Humans are exposed to alkylating compounds produced endogenously and in the environment, such as in the air, foods, waters and drugs.

Alkylating agents can form adducts at mainly O- and N-atoms in nucleobases through the electrophilic attack. Reaction was shown to mainly occur via $S_N1$ followed by first order kinetics or $S_N2$ which is heavily dependent on steric effect. In DNA molecules, the ring nitrogens and the exocyclic oxygens are the preferred sites for alkylation, such as the N7 of guanine, the N3 of adenine and the O6 of guanine. S-adenosylmethionine (SAM), a methyl donor in many biochemical reactions as well as a weak chemical methylating agent that has been shown to induce mutations in DNA, generate 7-methylguanine and 3-methyladenine as the major product when incubated with DNA.[2] Trace amounts of O-alkylated bases such as O6-methylguanine are also formed, although these latter adducts are generated more efficiently by alkylating agents which react by an $S_N1$ mechanism, for example, N-methyl-N'-nitro-N-nitrosoguanidine. The comparison of the persistence of N7-guanine adducts to O6-guanine adducts in brain and liver of mice treated by N-methyl-N-nitrosourea (MNU) suggested that O6-guanine are most likely mutagenesis and carcinogenesis other than N7-guanine.[3] The steric effect can dominate the binding site selectivity, generally the larger the alkyl group is, the more efficient for reactions at the O6 position of guanine, for example, N-ethyl-N-nitrosourea (ENU) has more preference to the O6 position than N-methyl-N-nitrosourea (MNU).[4] N-Ethyl-N-nitrosourea (ENU) is a potent carcinogen to the nervous system as MNU and it generates N7-ethyl-guanine (N7-Et-Gua), O6-ethyl-guanine (O6-Et-Gua), O2-ethyl-thymidine (O2-Et-Thy) and O4-ethyl-thymidine (O4-Et-Thy)

as the main adducts. The repair capacity of O6-Et-Gua, O4-Et-Thy and O2-Et-Thy is tissue-specific, but N7-Et-Gua is not. O6-Me-Gua and O6-Et-Gua adducts can be repaired in liver other than the brain. In contrast, the less stable N7-Et-Gua and N3-Et-adenine adducts have similar repairing rates in both liver and brain, suggesting that these adducts may be eliminated due to spontaneous depurination.[5]

N-nitrosamines (such as N, N-Dimethylnitrosamine (DMN) and 4-(methylnitrosamine) -1-(3-pyridyl)-1-butanone (NNK)), and Hydrazine compounds (such as 2-dimethylhydrazine (SDMH)) require metabolic activation, particularly functioned through P450 enzyme to form highly reactive diazonium ions and aldehyde, to exhibit their mutagenic and carcinogenic effects. This kind of compounds is a potent rodent carcinogen, responsible for the formation of N7-Me-Gua and O6-Me-Gua.[6] Three kinds of olefins: ethylene (ET), propylene, and 1, 3-butadiene (BD) have been well studied in respect of DNA alkylation agents. Been metabolised to the corresponding epoxides by the P450 enzyme, olefins can easily form DNA adducts. Since they are relatively more stable than the other metabolites such as diazonium and hydrazine, they will be spread similarly in the target and non-target tissues.

Aflatoxin B1 ($AFB_1$) is the most prevalent and carcinogenic of the aflatoxins, and the International Agency for Research on Cancer reported that there is sufficient evidence to classify $AFB_1$ as a Group I carcinogen. $AFB_1$ is bioactivated by epoxidation of the terminal furan ring double bond, generating the electrophilic intermediate $AFB_1$-8,9-epoxide, which attack guanine to form N7-$AFB_1$-Gua as the the primary $AFB_1$-DNA adduct.[7] The N7-$AFB_1$-Gua converts to secondary lesions including AP sites and AFB1-FAPy. Compared to N7-AFB1-Gua, AFB1-FAPy is highly persistent in rat liver DNA, reaching maximum amounts two

weeks after exposure.[8] The increased chemical stability and altered secondary DNA structure are responsible for the much greater mutagenicity of $AFB_1$-FAPy compared to N7-$AFB_1$-Gua.

It has been well established that cellular metabolism is the source of reactive oxygen species. These nonpathogenic cellular processes inevitably affect the background level of endogenous DNA oxidation damage. Leaking electrons in the line of electron transportation can flow to oxygen and give birth to the formation of superoxide. Other than that, superoxide can also be generated by the phagocytic cell aimed to destroy the infected cells. ROS can also be produced by ionizing or ultraviolet radiation and generate superoxide or reactive hydroxyl radicals, leading to potential damage to the DNA in the cell.

The highly reactive hydroxyl radicals (OH•) can add to the C5-C6 double bond of pyrimidines to generate the C5-OH and C6-OH adduct radicals. It can also abstract the H atom on the methyl group of thymine leading to the allyl radical. In terms of their redox properties, C5-OH adduct radicals are reducing and C6-OH adduct radicals are oxidizing. The 5-hydroxymethyluracil and 5-formyluracil are generated by addition of oxygen to the allyl radicals of thymine. Thymine peroxyl radicals can also be reduced and then protonated to give hydroxy hydroperoxides, which will spontaneously decompose and produce thymine glycol, 5-hydroxymethyluracil, 5-formyluracil, and 5-hydroxy-5-methylhydantoin.

In purines, hydroxyl radical can add to the C4, C5 and C8 position. C4-OH and C8-OH adduct radicals of adenine have been detected. C4-OH and C5-OH adduct radicals of purines can be dehydrated and generate an oxidizing purine radical, followed by reduction and protonation to form purine C4-OH adduct radicals. These adduct radicals are oxidizing,

whereas C5-OH and C8-OH adduct radicals are mainly reducing. The redox properties may depend on the mesomeric structures (redox ambivalence). C4-OH and C5-OH adduct radicals of purines can be dehydrated to form an oxidizing purine (-H•) radical, followed by reduction to reconstitute the purine. C4-OH adduct radical of guanine may form guanine radical cation (guanine$^{+}$•) and may deprotonate at desired pH condition to give rise to guanine radical (-H•). The free guanine radical cation molecule is not the source of C8-OH adduct radical which lead to 8-hydroxyguanine (8-oxoguanine, 8-OH-Gua), otherwise, it reacts with 2-deoxyribose in DNA by hydrogen abstraction causing DNA strand breaks. In double-stranded DNA, the hydration of guanine radical cation may form the C8-OH adduct radical and form 8-OH-Gua by oxidation. C8-OH adduct radicals of purines may be oxidized by oxidants including oxygen. 8-Hydroxypurines (7,8-dihydro-8-oxopurines) in DNA can be formed by oxidation reaction or in the absence of oxygen with less efficiency. On the other hand, the C8-OH adduct radicals can also go through the unimolecular opening of the imidazole ring by bond breaking between C8 and N9, forming 2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyGua) out of guanine and 4,6-diamino-5-formamidopyrimidine (FapyAde) from adenine. Further oxidation of 8-OH-Gua by $^{1}O_2$ lead to 4-HO-8-oxodGuo and oxazolone as the major products.[9] The oxidative pathway of 8-OH-Gua via one-electron oxidant has also been investigated. Two stable products including guanidinohydantoin and spiroiminodihydantoin were characterized and the later was the predominant product in the nucleoside at pH 7.[10]

### 1.3 Repair and Translesion Synthesis of DNA Damage

Organisms have developed several DNA-repair pathways as well as DNA-damage checkpoints to cope with the frequent challenge of endogenous and exogenous DNA insults. In the absence or impairment of such repair or checkpoint mechanisms, the genomic integrity of the organism is often compromised. On the other hand, when the damage may not be repairable in some cases, translesion synthesis (TLS) which is a DNA damage tolerance process allow the DNA replication machinery to replicate past DNA lesions.

### 1.3.1 DNA Damage Repair

The removal of DNA lesions is certainly important for the limitation of mutagenesis, cytostasis, and cytotoxicity. Most of the DNA lesions are subject to multiple repair pathways. The redundancy ensures that the attenuation or malfunction of one repair process does not completely cease the repair pathway of a particular lesion. Which mechanisms are most important after low level endogenous or environmental exposure remains undetermined, but would presumably be direct repair, BER, NER, HRR and ICL repair (Figure 1.1).

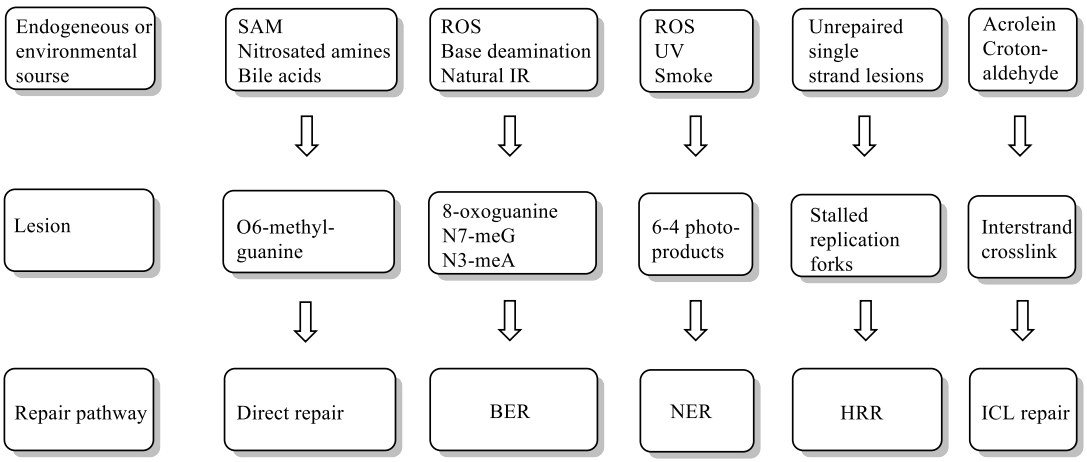



**Figure 1.1** Sources of DNA damage and their repair pathway.

### 1.3.2 Translesion Synthesis

Although DNA lesions can be removed by NER and BER, many lesions escape repair and present a block to continued transcriptional elongation by RNA polymerases and to replication by DNA polymerases. In eukaryotes, Translesion Synthesis (TLS) DNA polymerases, which belong to the Y family, and DNA polymerase ζ, which is a member of the B family, promote replication through DNA lesions. Structural and Biochemical studies have revealed striking differences among them in the roles they play in lesion bypass. In contrast to replicative DNA polymerases which are blocked by lesions that significantly distort the geometry of DNA, TLS DNA polymerases synthesize DNA with much higher error rates and are able to bypass lesions that block replicative polymerases.

DNA Polymerase η was first identified in yeast and shown proficient ability to replicate DNA containing a cis-syn thymine-thymine (T-T) dimer by inserting two adenines across from the dimer.[11] This proficiency derives from its unique structural feature - its active site is open enough to accommodate both residues of a dimer. Other than CPD lesion, Pol η plays a prominent role in efficient and accurate replication through the 8-oxoguanine (8-oxoG) lesion.[12] 8-OxoG is highly mutagenic and causes G:C to T:A transversions. In general, most DNA polymerases tested show either a preference for incorrect insertion of dATP or are indifferent to dCTP or dATP. Pol η can bypass the lesion much more efficiently than misincorporation of dATP.[13] Notably, the fidelity of 8-oxodG bypass by human Pol η is greatly increased by the inclusion of the accessory factors PCNA and RPA.

Pol ι is unable to replicate through a cis-syn T-T dimer like Pol η. However, Pol ι can incorporate nucleotides opposite an abasic site, but it is not capable of performing extension step. In addition, Pol ι can efficiently incorporate nucleotides opposite an N2-adducted guanine.

For instance, acrolein reacts with the highly reactive N2 of G followed by ring closure at N1 to generate γ-HOPdG, which blocks the Watson-Crick pairing in DNA. Pol ι incorporates a C or a T residue opposite γ-HOPdG, followed by efficiently extension performed by Pol κ from the C: γ-HOPdG primer terminus but not from the T:γ-HOPdG terminus. Thus, the accuracy and efficiency of the translesion synthesis of γ-HOPdG is ensured by the cooperation of Pol ι and Pol κ.

Pol κ is a member of Y-family DNA polymerases. It is unable to insert nucleotides opposite most of the lesion sites such as an abasic site or T-T dimer; but it can extend from nucleotides inserted by another DNA polymerase opposite certain DNA lesions. Highly inefficient in inserting nucleotides opposite an O6-methylguanine lesion, it efficiently extends from the T or C nucleotide incorporated opposite this lesion by Pol δ. Opposite an 8-oxoG lesion, Pol κ inserts an A quite efficiently and then proficiently extend from it.

Thymine glycol (Tg) is one of the principal DNA lesions that can be induced by oxidation and ionizing radiation. Pol ζ is exclusively efficient in extending Tg since no other polymerase has such ability. Other than extending beyond Tg and abasic sites, Pol ζ is highly inefficient at replicating through other lesions.

Overall, replication through certain DNA lesions can be accomplished by one polymerase, but through many DNA lesions the replication requires the action of two different polymerases, one for insertion reaction opposite the lesion and the other for the subsequent extension reaction (Table 1.1).

**Table 1.1** Lesion Bypass by One or Two DNA Polymerases

| | Lesion bypass by | | |
|---|---|---|---|
| | | Two Pol | |
| DNA lesion | One Pol | Inserter | Extender |
| 8-oxoG | Pol η | | |
| CPDs | Pol η | | |
| Abasic sites | | Pol η or Pol ι | Pol ζ |
| γ-HOPdG | | Pol ι | Pol κ |
| Tg | | Pol δ | Pol ζ |

### 1.4 Cellular Impact of Alkylation and Oxidation DNA Damage

There are a significant number of DNA lesions whose effects on replication and transcription have been well established. Some lesions can influence the replication and transcription fidelities, and many are mutagenic regardless whether they are formed in situ or arise by misincorporation from the deoxynucleotide pool. For example, 8-oxodG formed in situ results in G:T substitutions; alternatively, 8-oxo-dGTP may be misincorporated opposite dA, producing an A:C substitution.[14]

Conformational changes and the structural alterations arise from DNA lesions enhance the potential for mutagenicity. In single stranded DNA, 8-oxo-dG may adopt syn conformation during the replication or transcription other than the native anti conformation. This results in mispairing with dA or dT. The most abundant alkylation damage: guanine-N7 methylation, according to recent research results, alters hydrogen-bonding patterns of the guanine and affects

the stability of duplex DNA. The formation of a stable Watson-Crick-like N7-mdG:dT base pair in duplex DNA implicate that N7-mdG may induce G to A transition mutations if unrepaired.[15] A further factor about the mutagenicity or otherwise of a DNA damage is the level of difficulty in the repair pathway of a particular lesion. The preferences of the repair enzymes dominate the ease with which the lesion is repaired. In mammalian cells, the 8-oxoGua:C pair can be effectively repaired by OGG1, but the 8-oxoGua:A pair is poorly repaired, and replication is occurring before completion of DNA repair which explains the high mutagenic potency of the 8-OH-Gua:A mispair.[16]

The cell line or polymerase may also need to be considered regarding mutagenicity. The former point is well illustrated by O6-Alkylguanine-DNA alkyltransferase (alkyltransferase). In human, rat and mouse tissues, liver containing the highest level of activity, and low levels of alkyltransferase activity relative to protein were seen in human brain, rat brain and small intestine, and mouse kidney. In general, the rank of alkyltransferase activity relative to DNA for each tissue was human > rat > mouse, which suggests that the mouse is more susceptible to alkylating reagent such as nitrosoureas. The latter point can be demonstrated by 8-oxodG, the most abundant oxidation product of guanine. Pol η can stabilize the normal *anti* conformation of 8-oxodG to allow the Watson-Crick face orient toward the incoming dNTP. In contrast to Pol η, Pol κ favors dATP opposite the 8-oxodG. The structural studies have revealed the different molecular features between Pol κ and Pol η. In the complex of Pol η and 8-oxodG, the Lys-498 forms a hydrogen bond with O8 atom, but in Pol κ the residue is Leu-508 which apparently cannot stabilize the *anti* form of 8-oxodG. On the other hand, Pol κ has the N-terminal extension called “N-clasp” which encircles the DNA and serves to stabilize the binding

site. In the presence of N-clasp, the conformational fluctuation of the ligand is limited, thus the thermodynamically stable *syn* orientation of 8-oxoG is favored.

Sequence context is another factor that contributes to the DNA damage sites and the polymerase performance. The formamidopyrimidines FapyGua and FapyAde are efficiently generated by hydroxyl radicals in DNA. FapyAde, when present in DNA, is a moderately toxic lesion. Formation or the ability of FapyAde to inhibit DNA synthesis depended on the sequence context: almost no inhibition of DNA replication was observed in poly (dA) run and the strongest inhibition was observed in some adenine and guanine-rich sequences. FapyGua might possess similar properties.[17] N3-methyladenine (3-MeA) adducts which are expected to be cytotoxic. When p53 cDNA was treated *in vitro* with Me-lex which is a neutral N-methylpyrrolecarboxamide-based dipeptide used to generate 3-MeA, most of the 3-MeA formed at $A^{601}$ and $A^{602}$ in the presence of a mutational hotspot at the $A^{597}AATTT^{602}$ lex-binding site.[18]

While the mutagenic and carcinogenic effects of DNA damage are largely well recognized, a rising number of epigenetic studies about DNA damage are emerging in the recent work. It is clear that the presence of lesions in the transcribed regions of genes can lead to mutation, but what is the performance of lesions in a nontranscribed region? It has been demonstrated that the presence of 8-oxo-dG in promoter elements can affect transcription factor binding. It has been confirmed that a single oxidative 8-oxodG moiety can completely inhibit transcription factor binding to AP-1 and Sp1.[19] These results are consistent with previous reports showing that promoter regions containing alkylation damage inhibit transcription factor binding.[20] Further, sequence text dependent low binding of Sp1 was observed when guanine

residue of the 5′-GGGGCGGGG-3′ (GC box) was substituted by 8-oxo-dG.[21] This effect may have pathological consequences, as the interference of Sp1 binding in the kidney and liver of diabetic rats was proposed to result from ROS-mediated damage to DNA. In addition, oxidation or alkylation of bases can interfere with the ability of mammalian cell DNA to serve as a substrate for DNA-methyltransferases (DNA-MTase), leading to altered patterns in the distribution of 5-methylcytosine (5-MeC) residues at CpG sites. As an epigenetic change, methylation of DNA can result in changes in chromatin structure, render chromatin more compacted, and this is often accompanied by modified patterns of gene expression. Alterations in the DNA methylation may play a central role in tissue-specific gene expression, as well as in physiological processes, such as aging, carcinogenesis, etc.[22]

## 1.5 Role of DNA Damage in Cancer, Disease and Aging

DNA damages are a major problem for life. Put simply, by-products of normal metabolism cause damage to DNA, protein, and lipid. We argue that this damage (the same as that produced by radiation) is a major contributor to aging and to degenerative diseases of aging such as cancer, cardiovascular disease, immune-system decline, brain dysfunction, and cataracts.

### 1.5.1 Cancer

The well-known relationships between changes in DNA damage, gene expression, and disease have been described for cancer. In tumors and cultured cell lines, significant changes in genome-wide DNA damage have been observed. This includes a general genome-wide

demethylation that affects repeated sequences, a more gene-specific hypermethylation, and accumulation of DNA damage such as 8-oxo-dG. Deficient DNA methylation might contribute to the genomic instability of some colorectal tumor cell lines. DNA hypomethylation has also been proved to cause chromosomal instability in cells treated with the demethylating agent 5-azadeoxycytidine.[23] As a consequence, genomic instability can cause mutations in genes, thereby representing an indirect way in which changes in methylation patterns can affect gene expression. 8-oxo-dG is established biomarker of oxidative stress. In addition, 8-oxo-dG with potential mutagenicity in mammalian cells can also be seen as intermediate markers of a disease endpoint such as cancer. Evidence are the findings that GC to TA transversions derived from 8-oxo-dG has been detected *in vivo* in the ras oncogene and the p53 tumor suppressor gene in lung and liver cancer.

Numerous studies have attempted to establish a relationship between levels of DNA damage and cancer. Elevated levels of damage arise because of an environment in the tumor low in repair enzymes. One of the evidence is the correlations between DNA methylation inducer, the level of methylation in different CGIs and gastric cancer development. The research results clarify that accumulated levels of aberrant DNA methylation are associated with a risk of gastric cancer development. [24] Normal cells are protected by antioxidant enzymes from the toxic effects of high concentrations of reactive oxygen species generated during cellular metabolism. Even though cancer cells generate reactive oxygen species, it has been demonstrated biochemically that antioxidant enzyme levels are low in most animal and human cancers. Tumor cells are always lack of catalase, which responsible for the high level of $H_2O_2$ in tumor cells without exogenous stimulation, potentially explain the high levels of

oxidative damage seen.[25] However, it is still too early to directly link the elevated level of DNA damage to cancer until we could confirm that whether the former is the epiphenomenon to the later, or whether the coding region is damaged.

### 1.5.2 Diseases

Alzheimer's disease (AD) is a complex progressive, degenerative brain disorder characterized by an insidious loss of memory and other cognitive functions that usually results in death 5-10 years after onset. The absence of sufficiently efficient DNA repair mechanisms, involved in the removal of small base damages in the brain, could result in the accumulation of misrepaired or nonrepaired DNA damage and might ultimately lead to the neuronal degeneration as observed in AD patientsAlzheimer's disease cells, unlike normal cells, fail to repair methyl-methane sulfonate-induced DNA damage.[26] Besides, AD lymphocytes and age-matched normal lymphocytes respond similarly to methylation damage, as assessed by multiple techniques.[27] The fact that oxidative stress also plays a crucial role in AD pathogenesis seems clear. Markers of oxidative damage include the increase of 8-oxoG in AD brain and the characteristic pattern of altered hallmark structures of AD including AGE-modification, protein crosslinking and carbonyl- and acyl-modification.[28]

Inflammation can result from a range of sources including microbial infections, exposure to allergens and toxic chemicals and obesity. The chronic inflammatory response is generally harmful as the balance of the immune response is broken and linkage between chronic inflammation and cancer is now well accepted. Aberrant DNA methylation has been observed in many chronic diseases such as chronic biliary tract inflammation, Barrett's esophagus,

Helicobacter pylori infection and inflammatory bowel disease (IBD). In chronic inflammation during IBD, the level of pro-inflammatory cytokine IL-6 and TNF-α is elevated upon the increase of NF-κB, which in turn increases NF-κB and STAT3 in the epithelial cells, resulting in inhibition of apoptosis and increased proliferation of epithelial cells. The increased IL-6 can also increase DNMT levels in the epithelial cells, which can alter gene expression such as that of tumor suppressor genes. Overall, this provides an environment conducive to malignant transformation.[29] Another Evidence of DNA damage related inflammation shows that 8-oxodG content in livers with chronic hepatitis was significantly higher than the 8-oxodG content in normal livers. There was also a significant correlation between the 8-oxodG content in noncancerous liver tissues with individual serum alanine aminotransferase concentration which is a sensitive indicator of liver cell injury.[30] Rheumatoid arthritis (RA) and Systemic lupus erythematosus (SLE) are both chronic inflammatory diseases. It has been suggested that ROS may be able to modify both IgG and DNA in RA and SLE to make them become more susceptible to bind to the antibodies; this result implicates the importance of ROS in their pathogenesis. In addition, SLE cells show a considerable decrease in this rate of 8-oxodG removal, suggesting an impaired ability to excise 8-oxodG from the DNA which would contribute to the increased number of cells dying.[31]

The consequences of DNA damage in cardiovascular diseases have also drawn a lot of concerns in the latest decades. Lipid peroxidation is a major oxidative effect in which lipids after combination with oxygen through peroxyl radical formation, leads to lipid peroxidation, which undergoes homolytic decomposition to the cytotoxic downstream products such as the a,b-unsaturated aldehyde genotoxins, 4-oxo-2-nonenal, 4,5-epoxy-

2(E)-decenal, and 4- hydroxy-2-nonenal.[32] Lipid hydroperoxides can also be derived from the action of lipoxygenases and cyclooxygenases on polyunsaturated fatty acids. Lipid peroxidation along with its cytotoxic products are proved to be responsible for membrane disruption, leading to cardiovascular diseases such as angina pectoris. The reduction of nitric oxide (NO) which directly inhibits cell proliferation or LDL oxidation is expected to cause vasoconstriction followed by atherosclerosis and hypertension. ROS that generated by NADPH oxidase family of enzymes weakly couples cardiac mitochondria under oxidative stress conditions and leads to the development of fatal ventricular arrhythmia and when ROS oxidize Low-density lipoprotein to O-LDL (Oxidized LDL) it leads to hyperlipidemia.[33]

### 1.5.3 Aging

The well-accepted DNA damage theory of aging propose that the accumulation of DNA damage, especially nuclear DNA damage (nDNA damage) is the main inducement of aging. Nuclear DNA damage can contribute to aging indirectly by increasing apoptosis or directly by increasing cell dysfunction. Mammalian lifespans correlate with the effectiveness of nDNA repair. The most severe forms of accelerated aging disease in humans are due to nDNA repair defects, and many of these diseases do not exhibit increased cancer incidence. High rates of cellular senescence and apoptosis due to high rates of nDNA damage are apparently the main cause of the elderly phenotype in these diseases.

Biologic molecules are susceptible to spontaneous chemical reactions, mostly hydrolysis, oxidation and alkylation. The variable consequences of DNA damage along with the defect in DNA repair cause cellular dysfunction that manifests as aging independent of

apoptosis or cellular senescence. Mouse cells show significant age-dependent nDNA damage which correlates with pathology. A study target 8-oxodG in multiple tissues of rat suggest that the rate of 8-oxodG formation in old rats is more than tripled as compared to young rats, and the damage resulting from 8-oxodG are almost doubled for old rats in all brain areas.[34] In the human frontal cortex, the evidently reduced gene expression was found after age 40 and most pronounced after age 70.[35] One recent study suggests that the concentration of 8-oxodG in leukocyte DNA or urine is positively related with ages between 20 and 70, indicating that aging is associated with dysdifferentiation.

The persistence of damage and a subsequent increase in replication errors correlating with life span can also be explained in part by the decline of DNA repair capability with aging as there are sufficient evidence justifying that all pathways of DNA repair including MMR, NER, BER and DSB repair become less efficient with age leading to accumulation of mutations.[36] Many studies have shown a decline in NER for human dermal fibroblasts with age due to the reduced repair protein levels and activity.[37] The decline in base excision repair (BER) with age results from the reduction of glycosylase activity. Human fibroblasts and leukocytes from old donors show reduced BER glycosylase activity compared to cells from young donors.[38] The activity of uracil DNA glycosylase (UDG) in the cerebellum of the mouse brain decline nearly 50% and for 8-oxoguanine DNA glycosylase (OGG1) the reduction is almost 90%, and 8-OHdG repair in kidney and liver tissue of young and aged rats showed a significantly lower BER in the older rats. In human, the DNA glycosylase enzyme required for BER of 8-OHdG in lymphocytes of older showed a significant linear decline to half of the newborn values.[39]

## 1.6 DNA Damage Detection

Over a period of few decades, a number of methods have been invented to detect DNA damage in various organisms, such as PCR-based assays, comet assay, mass spectrometry-based methods, electrochemical methods, immunological assays, and so on.

### 1.6.1 PCR-Based Assays

The polymerase chain reaction (PCR) has been adapted for use in just about every branch of molecular biology and beyond. With no surprise, it should find a place in the study of DNA damage repairing. PCR is one of the most reliable techniques for detecting DNA damage as the amplification stops at the site of the damage. Quantitative PCR (QPCR) aims to measure the aggregate damage on both strands in the target gene region. Sub-gene functional regions such as introns, exon and promoters can also be efficiently detected. The detection range is desired to be between 300 and 3000 bp; however the upper limit can reach up to 20-30 kbp if using "long-PCR" reagents, allowing QPCR to be used to study entire genes. Strand-specific QPCR (ss-QPCR) incorporates adaptations that allow damage to be measured in the same region as QPCR but in a strand-specific way. Strand-specificity can be achieved by performing the reverse transcription reaction in the presence of only one primer specifically annealing with a unique region of the viral negative strand before amplifying the obtained cDNA by adding the second primer or a specific primer pair for PCR amplification. Unfortunately, strand-specific RT-PCR is highly susceptible to false-positive results. Single-strand ligation PCR (sslig-PCR) push the detection limit to reach single nucleotide damage on individual strands in a single copy

gene in mammalian cells. If antibodies to the DNA adducts of interest are available, these can be used to capture and isolate adducted DNA for use in sslig-PCR.

### 1.6.2 Comet Assay

Single Cell Gel Electrophoresis assay (SCGE, also known as comet assay), is used to detect DNA damage in individual cell and estimate its distribution in the cell population. Cells are lysed to form nucleoids containing supercoiled loops of DNA linked to the nuclear matrix. After the electrophoresis, comet structure will be the observed by fluorescence microscopy and the intensity of the comet tail compared to the head is related to the number of DNA breaks. The possible theory behind this is that loops containing a break may lose their supercoiling to be free to extend toward the anode. To improve the sensitivity and specificity of the assay, nucleoids are incubated with bacterial repair endonucleases that recognize specific kinds of damage to the DNA and convert lesions to DNA breaks, increasing the amount of DNA in the comet tail. The assay has applications in testing novel chemicals for genotoxicity, monitoring environmental contamination with genotoxins, human biomonitoring and molecular epidemiology, and fundamental research in DNA damage and repair. DNA damage in the patients with Down syndrome has been assessed by using new optimized comet assay. [40]

### 1.6.3 Mass Spectrometry-Based Methods

In the latest decade, mass spectrometry (MS) has become a powerful and reliable tool for the identification of specific DNA adducts up to 10-100 µg. HPLC-ES tandem mass spectrometry (HPLC-ES-MS/MS) can be used to detect DNA adducts such as etheno-DNA adducts, malondialdehyde-derived DNA adducts, tamoxifenDNA adducts, acrylamide-derived

DNA adducts, and this technology is especially useful for detection of oxidative base damage such as 8-OHdG, 8-OHdA.[41] GC-MS is commonly used for analysis of oxidative DNA damage due to its ability to identify a wide range of base oxidation products. In this method, polar bases are converted into thermally stable derivatives which possess mass spectra in a process called derivatization. However, derivatization at high temperature in the presence of air can cause 'artifactual' oxidation of some undamaged bases, leading to an overestimation of their oxidation products, including 8-hydroxyguanine. By optimizing derivatization conditions such as lower the temperature or adding reducing reagents this problem can be minimized.[42] Fapy (formamido-pyrimidines) derivatives are also measured by GC-MS.[43]

### 1.6.4 Electrochemical Methods

Electrochemical methods have revealed their considerable advantages in the fast response detection of DNA damage as a result of the sensitivity, selectivity, low cost and miniaturized devices. As an electroactive and surface-active substance, DNA can generate analytically valuable electrochemical signals for detection. Adenine, cytosine, and guanine are subject to redox processes by the mercury electrodes while guanine and adenine can be oxidized on carbon electrodes. The electrochemical signal response may also reflect the DNA structure. The reason is that the difference of accessibility of specific electroactive target may alter the electrochemical signal.[44] The 8-oxoguanine has been detected by this method via carbon electrodes. Besides, electrochemical measurements at mercury or solid amalgam electrodes offer a much more sensitive detection of DNA strand breaks than any other electrochemical method.[45]

### 1.6.5 Immunological Assays

Immunoassays are used to quantify molecules of biological interest based on the specificity and selectivity of antibody reagents generated. Immunoassays have become very popular in view of their high sensitivity, safety, economy and simple instrument requirements. By using a monoclonal antibody (D1OA1) in an immuno-slot-blot assay, the detection limit of highly mutagenic and carcinogenic malondialdehyde (MDA) adduct DNA can reach as low as 1 μg of DNA sample. This method can detect 2.5 adducts out of 108 bases. (Determination of malondialdehyde-induced DNA damage in human tissues using an immuno-slot-blot assay.) Another simple and efficient blotting method was developed to determine the frequency of thymine dimers in a variety of aquatic organisms such as cyanobacteria, phytoplankton and macroalgae, utilizing a thymine dimer-specific antibodies followed by blotting and chemiluminating.[46]

Immunoassays which employ enzymes are referred to as enzyme-linked immunosorbent assays (ELISAs). By introducing a specific enzyme-linked antibody, the final added substrate will be converted to produce a detectable or mostly colorimetric signal. ELISA technology has been mainly developed to detect the most abundant oxidative DNA biomarker: 8-oxodG.[47] Up to date, the ELISA kit on the market can detect as little as 100 pg/mL 8-oxodG.

# CHAPTER 2

# SYNTHESIS OF N7-ALKYL-9-DEAZA-2'-DEOXYGUANOSINES CONTAINING POLAR CHAINS AND EXAMPLES OF CHEMICALLY STABLE ANALOGUES OF N7-HYDROXYETHYL AND N7-OXOETHYL ADDUCTS OF 2'-DEOXYGUANOSINE

## 2.1 Introduction

DNA alkylation or adduct formation occurs at nucleophilic sites in DNA. N7-position of guanine is the most reactive site among these nucleophilic sites. Although remain debatable, so far the N7-alkyl guanine adducts are generally classified as non-promutagenic. Research about the N7-guanine adduct is mainly focused on the characterization of the lesions and the relationship between the formation of the guanine adduct and the biological consequences. However, due to the technical limitations, it is still not an easy task to justify the mutagenic potential of the DNA adducts and to identify the chemical source of the damage. Due to the higher abundance of the guanine adducts compared to other DNA adducts, they become good biomarkers and their involvement into mutagenesis remain to be fully investigated.

### 2.1.1 Formation of N7-Hydroxyethyl and N7-Oxoethyl Guanine Adducts

It is obvious that carcinogens comprise a variety of chemicals. Some of them were from endogenous sources or natural products, while others arise from synthetic products of modern human life. These chemicals are able to react with nucleophilic sites (electron rich, S, N, and O), in DNA and proteins. *In vitro* and *in vivo* experiment confirm that under physiological conditions (pH 7.4, 37 °C), alkylation of DNA primarily occurred at the N7-position of guanine,

although sometimes the larger group will favor O6-position of guanine via SN1 reaction. Other than directly react with DNA, studies also demonstrate that some carcinogens can be activated by metabolic processes to obtain their abilities to react with DNA and to exhibit their mutagenic and carcinogenic effects. Consequently, carcinogenic compounds were classified as "direct-acting" or "metabolically activated" carcinogens. Various technologies have been applied to animal and human exposure studies for routine analysis of N7-guanine adducts and other DNA adducts. These studies have increased our understanding of formation and persistence of DNA adducts, and their relationship to carcinogenesis.

The brief introduction about the methylation on the N7 position of guanine was discussed above. Here we mainly focus on the formations of N7-hydroxyethyl and N7-oxoethyl guanine adducts. Olefins, also known as unsaturated hydrocarbons, including ethylene (ET), propylene, and 1,3-butadiene (BD) which are the best studied in respect to DNA adduct formation. Ethylene oxide (EO), which can be metabolically activated from BD and the ET or exist as a chemical pollutant, has been classified as human carcinogens mutagenic in various *in vitro* and *in vivo* test systems and carcinogenic in rodents. EO forms different adducts upon reaction with DNA, N7-2-hydroxyethyl guanine (N7-HEG) being the main adduct. Inhalation of ET in mice and rats shows the dose-dependent response of N7-HEG formation. Under chronic exposure to ET, N7-HE-Gua adducts accumulation reach steady-state after the first week, suggesting the metabolic saturation of ET.

Chloroethylene oxide (CEO), the primary oxidative metabolite of vinyl chloride (VC) by the cytochrome P-450, can rearrange into chloroacetaldehyde (CAA). It is well believed that CEO is the ultimate mutagenic/carcinogenic intermediate of VC. The DNA adducts produced

from chloroethylene oxide include 7-(2-oxoethyl)-guanine, $N^2$,3- etheno-guanine, 3, $N^4$-etheno-cytosine, among which the 7-(2-oxoethyl)-guanine is quantitatively predominant. Only a small portion of CEO metabolically generated from VC is available for DNA alkylation due to the isomerization to CAA or the suicidal inactivation of P450 which cease the metabolization of VC.[48]

### 2.1.2 Biological Significance of N7-Hydroxyethyl and N7-Oxoethyl Guanine Adducts

Humans are subjected to the environmental or dietary exposure of ET which could be converted to EO by P450 in the liver, and physiological background level of N7-HEG between 2.1 and 5.8 pmol/mg DNA has been detected in human blood.[49] Another study using Drosophila as the model shows that under the relatively higher dose of EO (62.5-1000ppm), up to 20 fold increase in the mutation rate was observed in the nucleotide repair-deficient flies, suggesting the importance of NER in the N7-HEG repair and the involvement of the DNA adducts other than the free nucleotide.[50]

The hydrogen-bonding pattern of N7-alkylguanine has been investigated by using a novel polβ-host-guest complex system followed by analysis of crystal structure. This system reveals that the N7 methylation moderately alters the base pair geometry of dG: dC. In addition, the methylation may also stabilize the enol tautomer of guanine which may alter the hydrogen bonding pattern. The stable Watson-Click-like base pairing observed in 7-MEdG: dT, for example, suggest the potential G to A transition mutations. The additionally substituted group in the N7 along with the formal charge resulting from the alkylation may contribute to this tautomerization alteration, as we can also boldly prospect that the hydrogen bonding patterns

between 7-HEG or 7-OEG and the other four nucleobases may also provide some interesting information.

As the product of spontaneous depurination, apurinic sites (AP) are the most common form of endogenous DNA damage. The positively charged N7-alkyl-dG has a half-life of several hours to days in duplex DNA and can undergo spontaneous depurination to produce apurinic sites. AP sites are also intermediates of damage repairing system such as BER. Abasic sites dramatically affect the thermodynamic stability of duplex DNA due to the loss of hydrogen bonding and other noncovalent interactions. Other than that, AP can also induce G to T transversion mutations and interstrand cross-links. In monkey kidney cells, the preferential incorporation opposite the AP sites is dA (48%) > dC (39%) > dG (13%) >> dT (none), suggesting the potential GC-AT transversions induced by the AP site. The relationship among the N7-HEG, AP sites and mutations has been investigated by Rusyn et al. The repeated exposure to EO or ET does not cause the increase of AP sites, but are responsible for the increase of 7-HEG in the tissue, suggesting that the accumulation AP sites due to the DNA alkylation are not a primary mechanism for the mutagenicity and carcinogenicity of EO. However, the decrease of the proapoptotic genes in the spleen may contribute to the carcinogenicity of EO.[51] The presence of AP sites makes the interpretation of cellular mutagenic studies of 7-HEG and 7-OEG complicated.

N7-guanine adducts are susceptible to nucleophilic attack by hydroxide on the C8 carbon, followed by spontaneous ring opening reaction to form 5-N-alkyl-2,6, -diamino-4-hydroxyformamidopyrimidine (Alkyl-FAPy). Since the negative charge on the N9-position of the N7 adducts is delocalized to the formyl group, the glycosidic bond is stabilized and can

no longer spontaneous depurinate anymore. This effect has important biological implications, as studies show that FAPy adducts become highly persist in the tissues. The replication behavior of DNA Polymerases bypass the lesion MeFAPy-dGuo demonstrated that this FAPy lesion was miscoding with all four dNTPs at various efficiencies depending on the DNA polymerase, besides, deletion product and the misincorporation of dA:Fapy-dG lesion product arise with 5'-T-(MeFapydGuo)-T-3' local sequence.[52]

### 2.1.3 Need for Stable Analogues of N7-Hydroxyethyl and N7-Oxoethyl Guanine Adducts

To quantitatively analyze the mutagenic potential of a DNA lesion, one of the most reliable methods is to study the relative rate of dNTP incorporation by each polymerase. However, the low stability of N7-alkylguanine becomes an obstacle to synthesize oligonucleotides containing a single N7-alkylguanine for the analysis of DNA polymerase bypass. To construct single-strand DNA template with a single lesion, spontaneous depurination and the hydrolysis ring-open reaction are two main concerns. It has been confirmed that spontaneous depurination is generally faster in single-stranded DNA than in double-stranded DNA, which may indicate that the persistence of N7-alkylguanine lesions in duplex DNA is not only controlled by the inherent chemical stability of the N-glycosidic bond but also affected by the three-dimensional structure of the double helix in the region surrounding the lesion.[53] As a result, it is possible that the DNA adduct which distorts the helix structure may undergo depurination more rapidly.[54] Spontaneous hydrolysis which generates FAPy structure may proceed under neutral pH, although in low efficiency.

One study has been reported to use single guanine adduct containing single-strand plasmid DNA as a tool to investigate the fidelity of replication and transformation efficiency *in vivo* of N7-HEG in *E. Coli*. By carefully treating the plasmid with the enzymes cntaining AP endonuclease activity of exonuclease III, and the 5'- 3' exonuclease activity of T7 DNA polymerase just prior to transformation into *E. Coli* cells, this condition, according to the author, minimize the presence of the depurinated molecules. However, this strategy inevitably encounters the challenge of depurination and FAPY hydrolysis after transformation into *E. Coli*, which somehow weakens the specificity of the conclusion. Furthermore, the requirement for a single G site greatly limited the possibility of investigating different local sequences. Despite the large number of samples they used (300), the detection limit is still relatively high (0.3%), thus a high-throughput screening method is desired in this strategy.

N7-oxoethyl Guanine (N7-OEG) has been recognized as almost devoid of promutagenic effects since the lack of miscoding property according to early research back to 1980s.[55] However, These studies were mostly carried out at the cellular lever and the molecular behaviors were not clearly understood.

Hence, how and which polymerases bypass N7-HEG and N7-OEG has never been investigated. In the past, we have used N7-methyl-9-deazaguanine as the model of N7-methylguanine to study its polymerase replication. To expand the DNA lesions to the adducts formed between guanine and epoxides, particularly those generated *in vivo* from various olefins, N7-hydroxyethyl-2'-deoxyguanine and N7-oxoethyl-2'-deoxyguanosine became the molecules of interest.

### 2.1.4 Design of the Analogue

A stable analogue of an unstable DNA lesion can provide a good model to study its various biophysical and biochemical properties. Nucleobase analogues and their corresponding nucleosides and nucleotides are extensively used in enzymology as inhibitors, drugs, and probes. Deazapurines in particular are used as analogues of natural purine substrates to make stable enzyme-substrate complexes. 9-Dezapurines in which the N9 is substituted by a carbon atom are good substrate mimics. This substitution alters the electron distribution of the ring which consequently stablizes the glycosidic bond and renders the N7 hydrogen less acidic. Purine analogues including 9-deaza purines have been employed as inhibitors of well-known enzymes such as purine nucleoside phosphorylase (PNP), hypoxanthine-guanine phosphoribosyl-transferase (HGPRT), etc.

The 7-deazapurines and their derivatives in which the N7 was replaced by a carbon atom have shown useful pharmaceutical properties. As the N7 guanine adducts are labile to the alkaline treatment due to the electron-withdrawing $N^+$, 7-deaza-7-nitro-dATP (7-$NO_2$-dATP) and 7-deaza-7-nitro-dGTP (7-$NO_2$-dGTP) were designed to provide similar sensitivity to alkaline treatment as the N7 guanine adduct while gaining thermal stability to survive PCR cycles. The potent electron-withdrawing nitro group facilitates the cleavage of glycosidic bonds under alkaline conditions.[56]

The idea that the lack of formal charge on the N7 may affect the competence of 9-deaza or 7-deaza guanine as good purine mimics is still remained debatable. The modern synthetic chemistry allows the researcher to find out the difference between the analogue with and without positive charge on it. Chen uses the decreased affinity of 9-deaza analogue toward the

enzyme to confirm the cation-π interactions they predict, but they also found that the permeability will be improved if the candidate is a neutral compound.[57] To investigate DNA interstrand crosslink (ICL) damage induced by bifunctional carcinogens such as cisplatin or nitrogen mustard using PCR technology, the stable guanine analog 7-Deaza-7-(2,3-diacetoxypropyl)-dG CEP can be used for the site-specific introduction of ICL precursor nucleoside.[58]

Similar to 7-deazaguanine but maintain slightly structural difference, 9-deazaguanine retain the unmodified Watson-Click edge and the stability of the glycosidic bond which ensure the survival in the PCR. However, the 7-deazaguanines such as Me-7-deaza-G which has the N-C glycosidic bond are still substrate of BER enzyme. It was also discovered that 3-Me-3-deaza-A can be removed by AlkA. As a result, the 9-deazaguanine analogues will have these advantages: 1) structural stability and availability via phosphoramidite chemistry; 2) stability against hydrolysis and BER; 3) The synthetic method can be modular for the preparation of other N7-alkylG adducts for future studies. Finally, the synthesis of N7-hydroxyethyl-9-deaza-2'-deoxyguanosine ($^{7HE9C}$dG) and N7-oxoethyl-9- deaza-2'-deoxyguanosine ($^{7OE9C}$dG) was selected as the two N7-alkyl-dG analogues that are resistant to glycosidic bond cleavage. The synthesis of side-chain-protected nucleosides that can be readily converted to phosphoramidites for oligonucleotide synthesis was also studied.

## 2.2 Synthesis of the Analogues

The N2-dmf-protected 9-deazaguanine is obtained from 2-amino-4-hydroxy-6-methyl pyrimidine in five steps following Rana's protocol (Figure 2.1).[59] The glycosyl donor 1-(a,b)-

O-Methyl-3,5-di-(O-*p*-toluoyl)-2′-deoxyribose was prepared according to a reported procedure (Figure 2.2).[60]

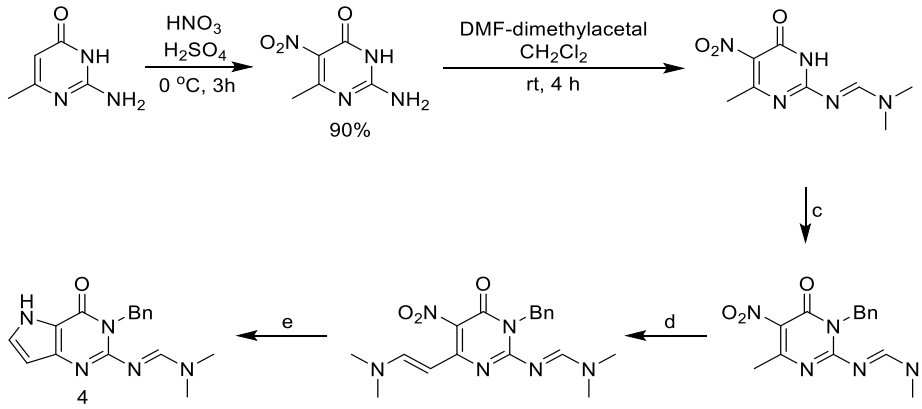



**Figure 2.1** Synthesis of starting material **4**. (a) $H_2SO_4$, $HNO_3$, 0°C, 3 h,90%; (b) DMF-dimethylacetal, $CH_2Cl_2$, rt, 4 h, 80%; (c) dry DMF, DIPEA, Benzyl bromide, rt, 20 h, 80%; (d) dry DMF, DMF-dimethylacetal, 65ºC, 4 h, 70% (e) DMF, $Na_2S_2O_4$ , 65ºC, 3 h, 80%.

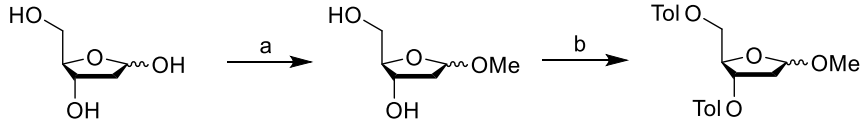



**Figure 2.2** Synthesis of 1-(α, β)-*O*-methyl-3,5-di-(*O*-p-toluoyl)-2-deoxy-D-ribose. (a) $CH_3COCl$, MeOH, 0°C to rt, 3 h (b) DMAP, p-Toluoyl chloride, dry pyridine, rt, 16 h.

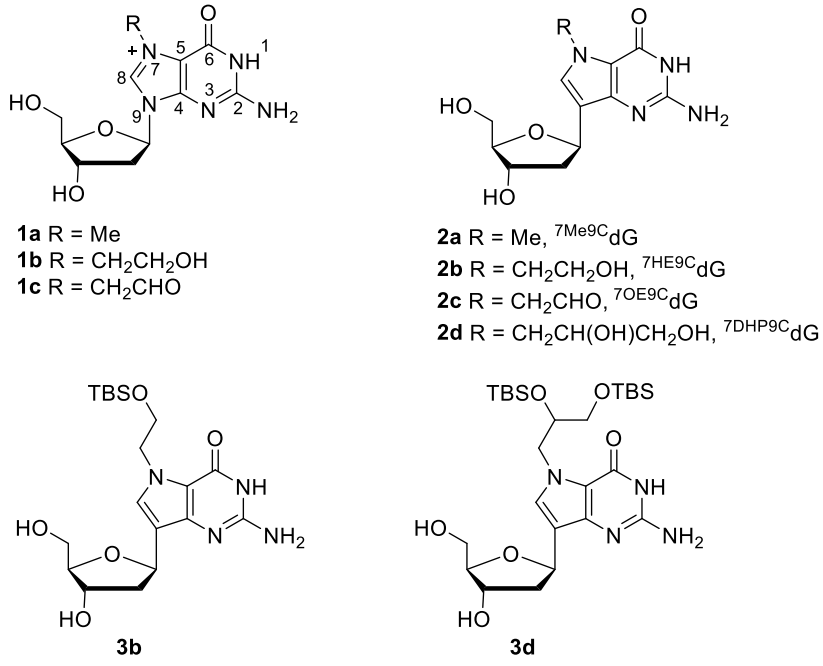



**Figure 2.3** Compounds of interest.

The N2-dmf protected 9-deazaguanine undergoes Lewis-acid-mediated Friedel-Crafts reaction with a 1-O-methyl-3,5-O-ditoluated deoxyribose derivative to yield C-nucleoside. The β-anomer is afforded by this reaction in 26% yield (Figure 2.4A). Although comparable to the reaction of N2- and N7- unprotected 9-deazaguanine, the yield of this reaction is dramatically lower than that of N2-dmf-N7-methyl-9-deazaguanine (44%) (Figure 2.3). The hydroxyethyl group on the N7 position is introduced by using ethylene carbonate as the alkylating agent, followed by the removal of toluoyl and dmf groups under strong alkaline condition, compound **7** is afforded. To remove the Bn group on N1 position, Pd/C-ammonium carbonate is proved to be very efficient reagent under mild heating.[59] As a result, this reaction leads to the formation of $^{7HE9C}$dG (**2b**).

By following the similar strategy, $^{7OE9C}$dG can also be generated (Figure 2.4B). Allyl group is introduced to N7 position to generate N7-allyl nucleoside **8** by using allyl bromide as the alkylating agent. Then the dihydroxylation reaction is performed on the olefin to form a vicinal diol (**9**) using the popular oxidant - osmium tetroxide. Then the deprotections of dmf, toluoyl, and benzyl groups are following the same procedures discussed above, and the $^{7DHP9C}$dG (**2d**) is obtained. By using potassium periodate to break apart vicinal diols to form aldehyde, $^{7OE9C}$dG (**2c**) **i**s generated (Figure 2.4B). Since $^{7OE9C}$dG is not very stable under room temperature due to the highly reactive aldehyde, the characterization of this compound is performed immediately after the purification.

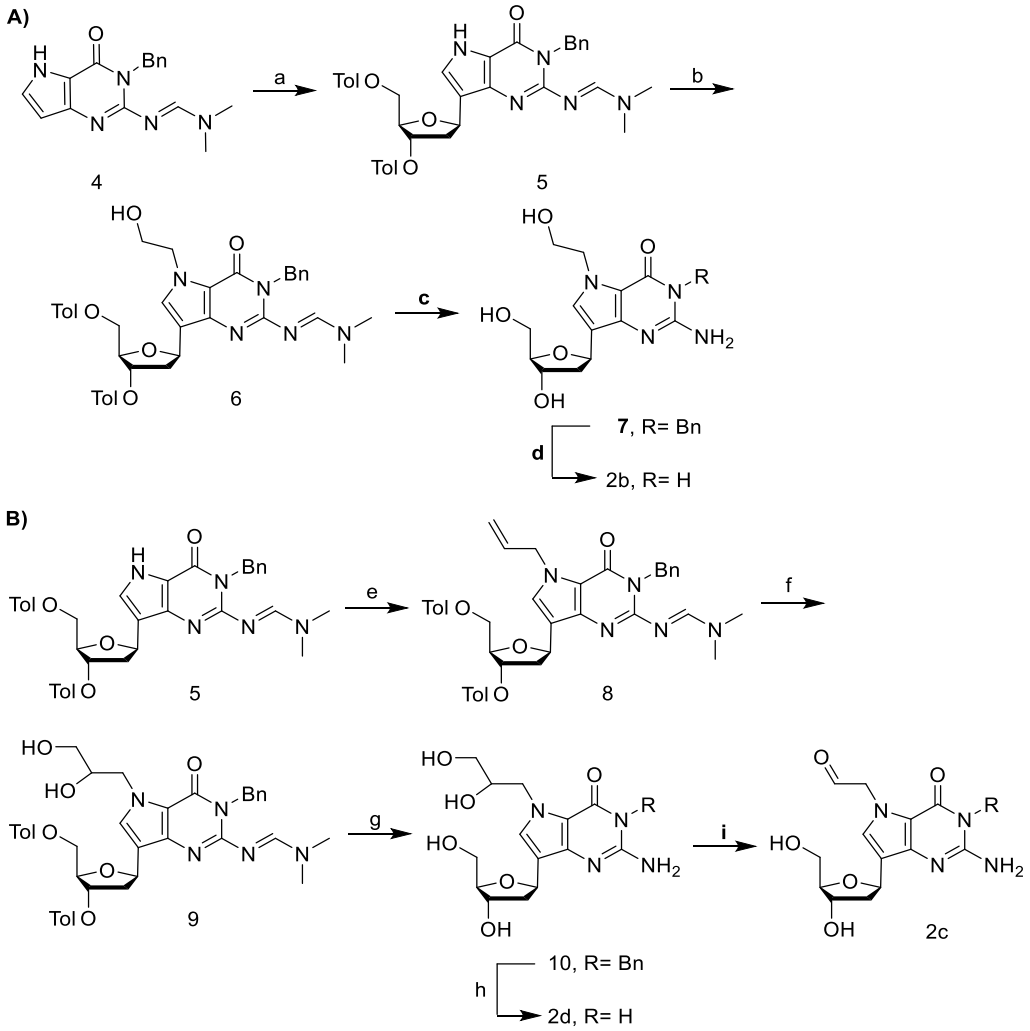



**Figure 2.4** Synthesis of target molecules **2b** and **2d**. (a) 1-(α,β)-*O*-methyl-3,5-di- (*O*-p-toluoyl) -2-deoxy-D-ribose, $SnCl_4$, 1:1 (v/v) dry $CH_3CN$-$CH_2Cl_2$, 65 °C, 16 h, 26%; (b) ethylene carbonate, DBU, dry DMF, 90 °C, 4 h, 49%; (c) 1 M NaOH, 2:1 (v/v) $CH_3OH$-$H_2O$, 70 °C, 16 h, 72%; (d) Ammonium formate, Pd/C, $CH_3OH$, 75 °C, 16 h, 48%; (e) allyl bromide, NaH, dry THF, rt, 16 h, 48%; (f) Osmium tetraoxide, TBHP, TBAF, 4:1 (v/v) acetone-$H_2O$, rt, 76%; (g) 1 M NaOH, 6:1 (v/v) $CH_3OH$-$H_2O$, 70 °C, 16 h. 54%; (h) Ammonium formate, Pd/C, $CH_3OH$, 75 °C, 16 h, 74%; i) $KIO_4$, 1:1 (v/v) $CH_3OH$-$H_2O$, rt, 30 min, 85%.

To examine the stabilities of the glycosidic bonds of $^{7HE9C}$dG and $^{7DHP9C}$dG, these two compounds are incubated under the following three conditions for 8 h: i) HCl (pH 2.5), r.t., ii) NaOH (pH 11.7), r.t., and iii) phosphate buffer (pH 7.2), 70 °C. The glycosidic bond is confirmed to be very stable by using the NMR and epimerization was not observed, and these results demonstrate that the stable analogues of N7-alkyl-dG are good candidates of

phosphoramidite chemistry. Considering the naturally instability of the aldehyde group, $^{7OE9C}$dG is excluded from the stability study.

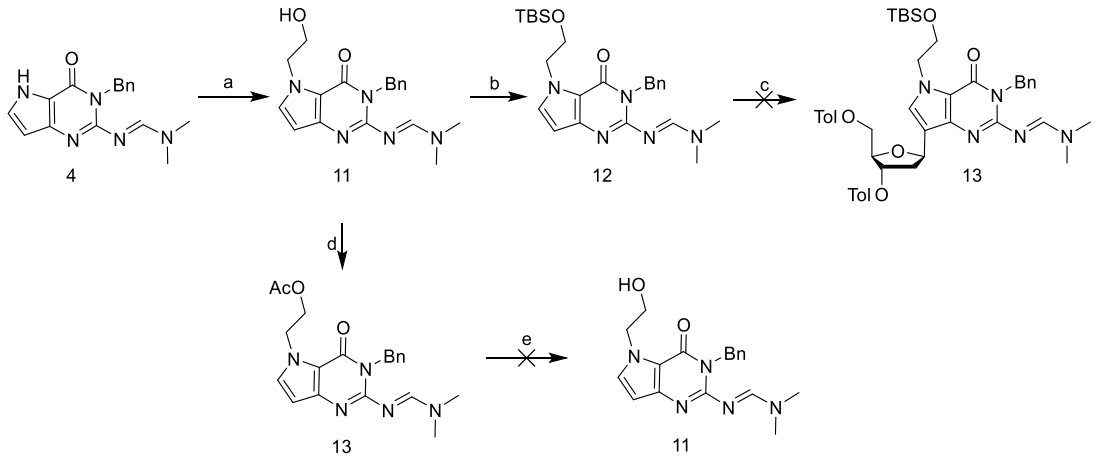



**Figure 2.5** Incompatibility between $SnCl_4$ and TBS group, unsuccessful deprotection of acetate group by lipase B. (a) Ethylene carbonate, DBU, dry DMF, 90 °C, 4 h, 49% (b) TBSCl, Imidazole, 4-DMAP, DMF, 20 h, 62% (c) 1-(α,β)-*O*-methyl-3,5-di- (*O*-p-toluoyl)-2-deoxy-D-ribose, $SnCl_4$ (1M) in $CH_2Cl_2$ , 1:1 (v/v)dry $CH_3CN$- $CH_2Cl_2$, reflux, 20h. (d) AcOAc, $NEt_3$, rt, 16h, 78% (e) Lipase B, 1:1 (v/v) $H_2O$-DMSO, 37°C, pH = 7.2, 2d.

TBS is the typically used in the solid phase synthesis of DNA/RNA to protect hydroxy group. It can be easily removed under mild condition by TBAF. Having completed the synthesis of $^{7HE9C}$dG and $^{7OE9C}$dG, to construct the molecules which can be readily convert to phosphoramidites and used in the solid phase synthesis, the side chain on N7 position needs to be protected by TBS to obtain target molecules **3b** and **3d**. Initially, according to the previous results that the N7 capped nucleobase may give better yield in the coupling reaction, the hydroxyethyl group of compound **11** was protected by TBS. However, the following glycosylation reaction fails due to the speculated reason that the TBS may not be compatible with the lewis acid using in the coupling reaction - $SnCl_4$ (Figure 2.5). To circumvent this problem, the alternative route is first using the acetate to protect of the hydroxy group and then

selectively deprotect it and replace with TBS after the glycosylation reaction. As reported, lipase B could efficiently and selectively deprotect acetate group under mild condition, however this enzymatic deprotection by lipase B is futile on compound **13** (Figure 2.5). Thus, the strategies in figure **4** are maintained to perform the coupling reaction before the N7 substitution.

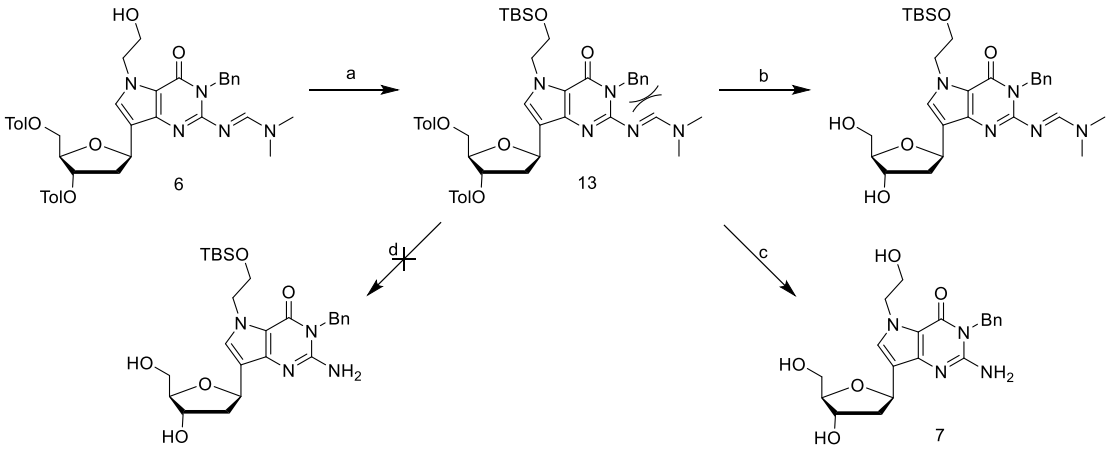



**Figure 2.6** Unsuccessful deprotection on compound **13**. (a) TBSCl, Imidazole, 4-DMAP, DMF, rt, 16 h, 75%; (b) 5:1 (v/v) $CH_3OH$-$H_2O$, 0.05 M NaOH, rt, 1 h, 75%; (c) 1 M NaOH, 2:1 (v/v) $CH_3OH$-$H_2O$, 70 °C, 16 h, 72%; (d) different bases (ammonia, KOH, or NaOH), temperatures (0-75 °C), reaction times (1-48 h), and solvent ratios (water-ethanol-THF mixture).

After the hydroxyethyl group of compound **6** is protected by TBS, multiple conditions are examined to selectively deprotect the toluoyl and dmf groups while maintaining the TBS group. Unfortunately, great difficulties are encountered in this step. Although it was well documented that the dmf group was typically used in the solid phase synthesis for the protection of the exocyclic amine of nucleobases and it could be removed under a mild condition, during the previous course of the study of $^{7Me9C}$dG (**2a**), it is found that strong basic conditions and high temperature are required to remove the N2-dmf if the N1 of 9-deazaguanine is protected by a benzyl (Bn) group. It should also be noted that TBS cannot survive strong basic alkaline condition, and the glycosidic bond is very sensitive to acidic conditions (Figure 2.6). To

perform the selective deprotection, different bases (concentrated ammonia, diluted or 1 M KOH/NaOH), temperatures (0-75 °C), and solvent ratios (water-ethanol-THF mixture) are tested. In general, mild conditions only enable the removal of the toluoyl groups while keeping the dmf unprotected, and harsh conditions lead to the deprotection of toluoyl, dmf, and TBS at the same time.

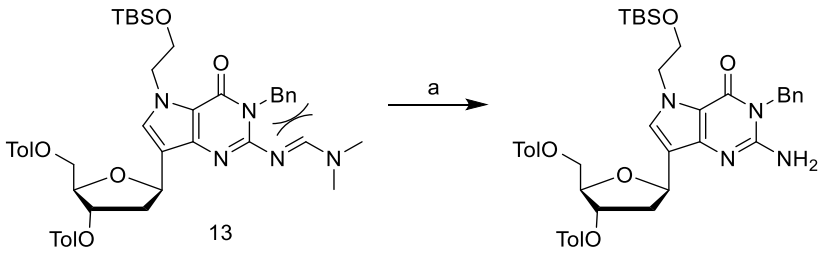



**Figure 2.7** Hydrogenolysis on compound **13**. (a) Ammonium formate, Pd/C, DMSO, reflux, 48 h, 10 %.

Anticipating that the removal of the N1-Bn group can release the steric hinderance and facilitates the deprotection of dmf under milder conditions while keeping TBS intact, compound **13** is treated with Pd/C-ammonium formate in refluxing methanol overnight. Unfortunately, the same reaction that works efficiently in the syntheses of **2b** and **2d** is completely futile on compound **13**. Change of the conditions to refluxing DMF leads to partial removal of the N2-dmf group and maintain the N1-Bn, but the low yield limits the use of this condition for the whole synthetic plan (48 hrs, less than 10% conversion) (Figure 2.7). As a result, it appears that N1-Bn and N2-dmf of compound **13** have a remarkable steric effect on each other, which blocks facile deprotection of either group.

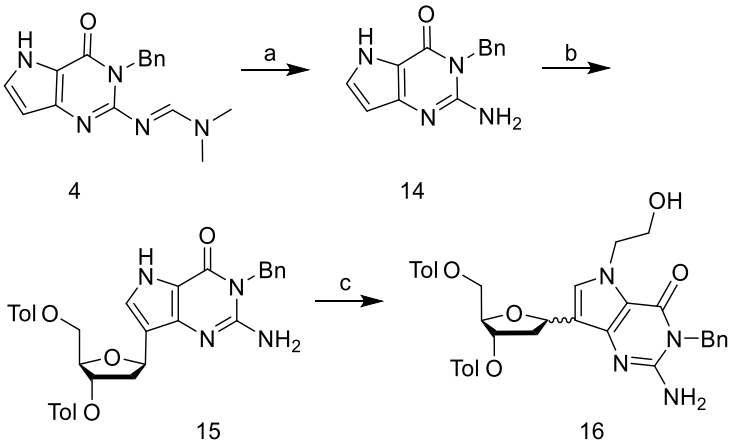



**Figure 2.8** Synthesis of compound **16**. a) 1 M NaOH, 1:4 (v/v) $H_2O$-$CH_3OH$, reflux, 20 h, 75%; , b) 1-(α, β)-*O*-methyl-3,5-di-(*O*-p-toluoyl)-2-deoxy-D-ribose, $SnCl_4$, 1:1 (v/v) dry $CH_3CN$-$CH_2Cl_2$, 65 °C, 16 h, 26%; c) ethylene carbonate, DBU, dry DMF, 110 °C, 4 h, 46%, α:β = 1:3.

Obviously the removal of the dmf has to be ensured in the later stage of the synthesis. Once the carbohydrate is introduced which carries two toluoyl groups, selective deprotection would become very difficult to achieve. It seems that the removal of dmf group should be done prior to glycosylation step, on condition that the unprotected N2 amino group would not compete with the N7 alkylation under basic conditions. The dmf-free nucleobase **14** still reacts with 9-deaza-dG through coupling reaction in moderate yield (26%) However, during the N7-alkylation, the reaction does not proceed under the previously adopted temperature (90 °C) using DBU and ethylene carbonate. This observation suggests that the removal of N2-dmf group renders the N7-H more difficult to dissociate. Until the temperature is elevated to 110 °C, the desired hydroxyethylated product is generated, however, as a mixture of inseparable α/β isomers (1:3, $^1$H NMR integration) (Figure 2.8). Therefore, an alternative solution is required to achieve the synthesis of **3b**.

Epimerization of C-nucleosides under basic conditions had also been found by other chemists. Hamm et al first found epimerization of unalkylated 9-deaza-dG in concentrated ammonia at 55 °C for 15 h,[61] while the same observation could not be found at room temperature, suggesting that the high temperature was essential for the epimerization. In contrast, the N7-alkylated N7-methyl-9-deaza-dG (**2a**) was not prone to epimerization under strong alkaline conditions even at high temperature.[59] Therefore it seems that the epimerization of 9-deazaguanine nucleosides requires the deprotonating of N7-H through either general base-catalysis or specific base-catalysis. Furthermore, the electronic effect also contributes to the epimerization in this case. With the N2-dmf protecting group, no epimerization is observed with compound **5**, suggesting that less electron-donating effect exerted by N2-dmf can attenuate epimerization. This result is consistent with the literature record of the C-glycoside of 6-aminopyridone.[62] In short, the availability of N7-H, alkaline pH, high temperature, and electron donating groups on the nucleobase in combination create favorable conditions for the epimerization of 9-deaza-dG.

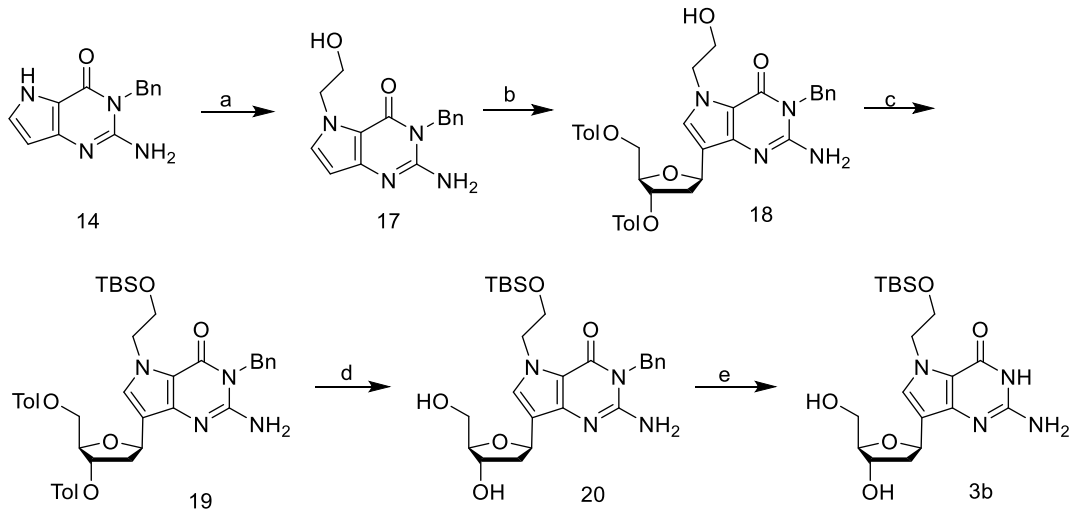



**Figure 2.9** Synthesis of compound **3b** (a) ethylene carbonate, DBU, dry DMF, 90 °C, 4 h, 47%; (b) 1-(α, β)-*O*-methyl-3,5-di-(*O*-p-toluoyl)-2-deoxy-D-ribose, $SnCl_4$, 1:1 (v/v) dry $CH_3CN$-$CH_2Cl_2$, reflux, 16 h, 28%; (c) TBSCl, imidazole, 4-DMAP, DMF, rt, 16 h, 93% (d) 0.05 M NaOH, 10:1 (v:v) $CH_3OH$-$H_2O$, rt, 1 h, 74%; (e) ammonium formate, Pd/C, $CH_3OH$, 75 °C, 16 h, 43%.

Non-reacting hydroxyl groups are nucleophilic and in many cases they need to be rationally protected in a Lewis acid-promoted glycosylation reaction. However, if the intermediate is an O-glycoside, attacking of the alkoxyl group to the oxocarbenium intermediate is reversible. As such, the more stable C-glycosidic bond will be favored over an O-glycosydic bond. This analysis is consistent with the experimental result. The unprotected N7-hydroxyethyl-9-deazaguanine **17** successfully reacts with the sugar donor to form the desired product with no significantly lower yield (Figure 2.9). Then after TBS protection of the hydroxy group followed by removal of the two toluoyl groups and the N1-Bn group, compound **3b** is finally obtained.

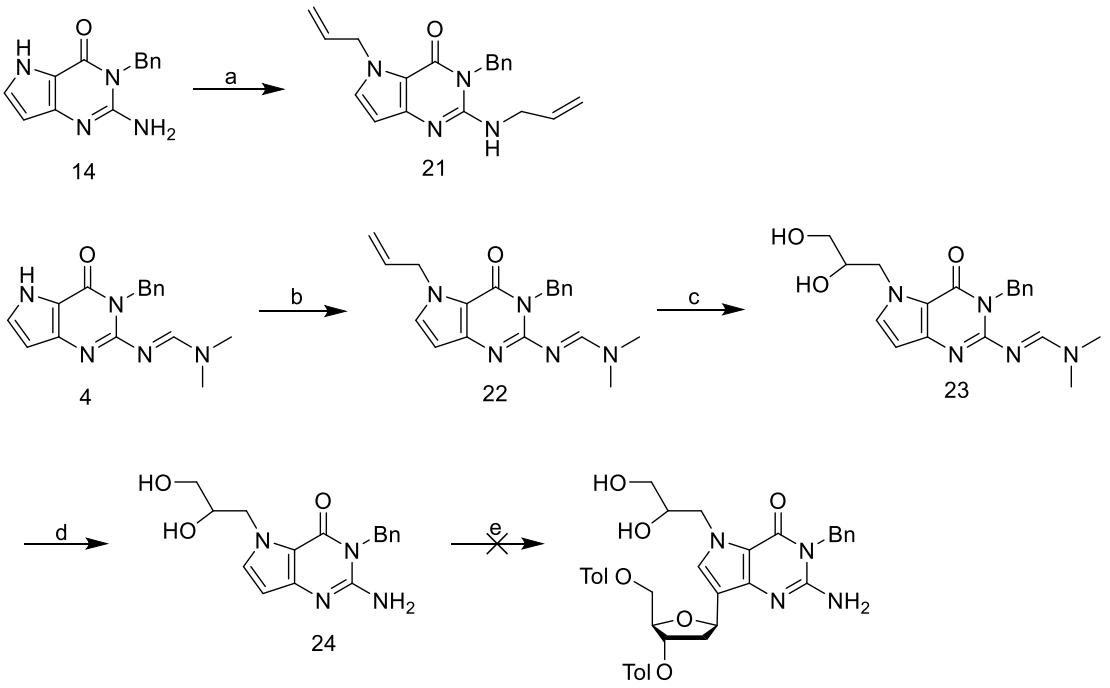



**Figure 2.10** Unsuccessful reactions towards target molecule **3d** (a) allyl bromide, NaH, dry THF, rt, 4 h. (b) allyl bromide, NaH, dry THF, rt, 4 h, 64%. (c) osmium tetraoxide, TBHP, TBAF, 4:1 (v/v) acetone-$H_2O$, rt, 16 h, 81%. (d) 1 M NaOH, 1:4 (v/v) $H_2O$-$CH_3OH$, 70 °C, 16 h, 53%. (e) 1-(α,β)-O-methyl-3,5-di-(O-p-toluoyl)-2-deoxy-D-ribose, $SnCl_4$, 1:1 (v/v) dry $CH_3CN$-$CH_2Cl_2$, reflux, 16 h.

The synthesis of **3d** is carried out follow a similar strategy as that of **3b** (Figure 6B). In the beginning, starting from the compound **14**, the allylation reaction using allyl bromide as the alkylating agent produces a mixture of N2- and N7-allyl products, suggesting the alkylating reaction of **14** is dominated by kinetic control (Figure 2.10). Instead, the N2-dmf-protected 9-deazaguaine (**4**) is then used as the starting material, which allows the allyl group to be specifically introduced to N7. After the removal of N2-dmf group under strong alkaline conditions, the olefin is dihydroxylated to generate compound **23**. However, the glycosylation reaction between the vicinal diol **24** and the sugar donor is completely futile. It appears that the diol may chelate and thus deactivate the $SnCl_4$ catalyst. Therefore, the dmf-free N7-allyl-9-deazaguanine is employed to carry on the coupling reaction. (Figure 2.11) The N7 allyl group

successfully survives in this reaction to produce compound **27** in 23 %. Then dihydroxylation is performed to obtain a vicinal diol **28**, followed by TBS protection of both hydroxyl groups (**29**). After the removal of the two toluoyl groups and the N1-Bn, **3d** is successively generated.

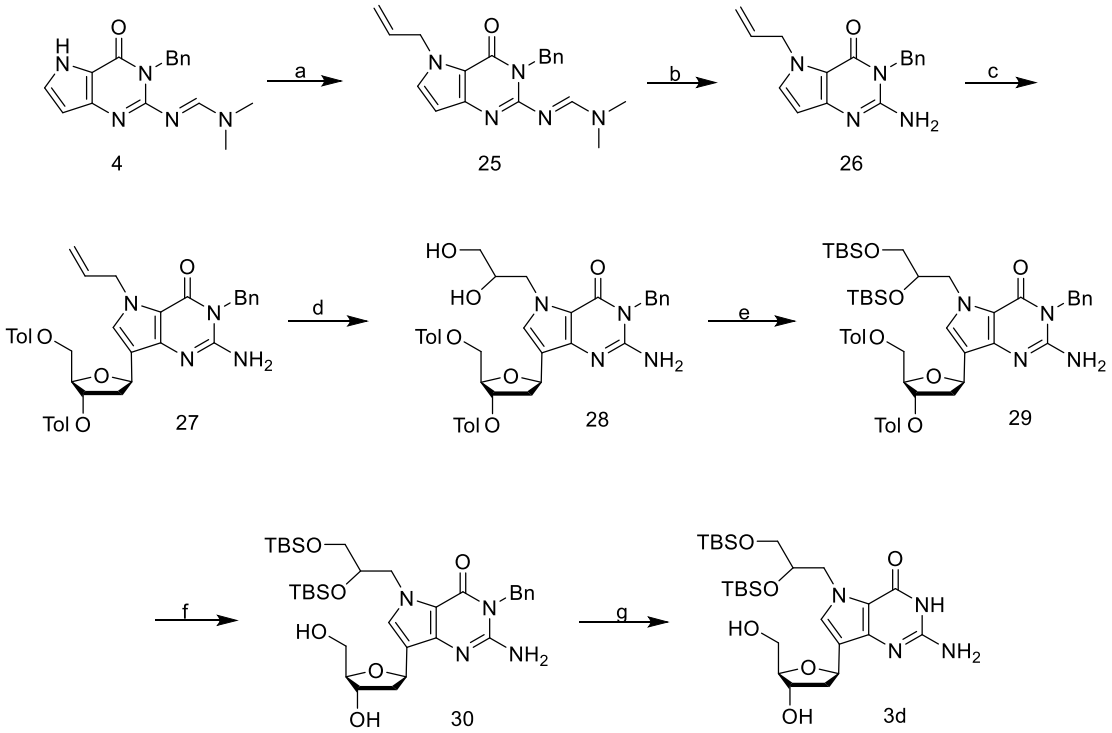



**Figure 2.11** Synthesis of compound **3d**. (a) allyl bromide, NaH, dry THF, rt, 4 h, 64%; (b) 1 M NaOH, 1:4 (v/v) $H_2O$-$CH_3OH$, 70 °C, 16 h, 70%; (c) 1-(α,β)-O-methyl-3,5-di-(O-p-toluoyl) -2-deoxy-D-ribose, $SnCl_4$, 1:1 (v/v) dry $CH_3CN$-$CH_2Cl_2$, 65 °C, 16 h, 23%; (d) osmium tetraoxide, TBHP, TBAF, 4:1 (v/v) acetone-$H_2O$, rt, 16 h, 40%; (e) TBSCl, imidazole, 4-DMAP, DMF, rt, 16 h, 76%; (f) 0.1 M NaOH, 10:1 (v:v) $CH_3OH$-$H_2O$, rt, 1 h, 68%; (g) ammonium formate, Pd/C, $CH_3OH$, 75 °C, 16 h, 56%.

In summary, two polar-chain-containing N7-alkyl-9-deaza-dGs were successfully synthesized. This work has expanded the series of chemically stable analogues of N7-alkyl-dGs. The stability test confirmed that the C-glycosidic bonds of the two new compounds were stable under strong acidic and basic conditions and at high temperatures. An efficient synthetic route to side-chain-protected 9-deazaguanosine in spite of the difficulties in stereoselectivity and regioselectivity when the exocyclic amine of 9-deazaguanine was not

protected. These side-chain-protected nucleosides can be readily converted to phosphoramidites for solid-phase oligonucleotide synthesis to allow the study of polymerase actions on N7-hydroxyethyl and N7-oxoethyl adducts of 2'-deoxyguanosine.

### 2.2.1 Methods and Materials

**General Information.** $^{1}H$ and $^{13}C$ NMR spectra were recorded on a NMR spectrometer operating at 500 or 600 MHz for $^{1}H$ and 125 or 150 MHz for $^{13}C$ using the solvent as an internal reference. The coupling constants (*J*) for $^{1}H$ NMR are recorded in hertz. High resolution mass spectra (HRMS) of compound **9**, **13**, **19**, and **29** were obtained with MALDI-TOF and all others with ICR (ESI). Melting points were recorded on a microscopic instrument.

THF was distilled freshly from $LiAlH_4$. Acetonitrile, dichloromethane, and DMF were distilled freshly from $CaH_2$. 1-Benzyl-9-deazaguanine (**4)** was prepared following Jagruti's protocol and its characterization was described previously by Gibson, *et al*. [63].1-(α, β)-*O*-methyl-3,5-di-(*O*-*p*-toluoyl)-2-deoxy-D-ribose was prepared according to a reported procedure.[60]

**Stability Test.** $^{7HE9C}$dG and $^{7DHP9C}$dG were incubated under the following three conditions for 8 h: i) HCl (pH 2.5), r.t., ii) NaOH (pH 11.7), r.t., and iii) phosphate buffer (pH 7.2), 70 °C. After that, the reaction mixture was neutralized, concentrated and purified by flash column chromatograp using relatively more polar eluent (methanol/ethyl acetate 1:2, compared to the 1:5 for **2b** and **2d**) to retain possible products. The NMR spectra are compared with the spectra of the pure compound to determine any possible bond cleavage and epimerization.

(2R,3S,5R)-5-(3-Benzyl-2-(((dimethylamino)methylene)amino)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl) tetrahydrofuran-3-yl-4-methyl benzoate (**5**)**.** To a suspension of N'-(3-benzyl-4-oxo- 4,5-dihydro-3H-pyrrolo-[3,2-d]pyrimidin-2-yl)-N,N-dimethylformimidamide **4** (3.52 g, 12.0 mmol) and 1-(α,β)-*O*-methyl-3,5-di- (*O*-*p*-toluoyl)-2-deoxy-D-ribose (6.6 g, 17.2 mmol) in a mixture of methylene chloride (20.0 mL) and acetonitrile (20.0 mL) was added a solution of $SnCl_4$ (23.2 mL, 23.2 mmol, 1 M in $CH_2Cl_2$). The reaction mixture was heated at 65 °C for 16 h. The reaction mixture was diluted with methylene chloride and washed successively with sat. $NaHCO_3$ and brine. The organic layer was separated, dried over $MgSO_4$. The solution was concentrated and purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with hexane/ethyl acetate 2:1 to 1:2) to give compound **5** as a light yellow solid (1.96 g, 26%). mp 92-93°C. $^1$H NMR (500 MHz, $CDCl_3$) δ ppm 2.36 (s, 3H), 2.42 (s, 3H), 2.49-2.53 (m, 1H), 2.91-2.95 (m, 1H), 3.05 (s, 3H), 3.12 (s, 3H), 4.50 (s, 1H), 4.58 (dd, *J* = 10.5, 5.0 Hz, 1H), 4.66 (dd, *J* = 11.0, 4.5Hz, 1H), 5.49-5.56 (m, 3H), 5.74 (s, 1H), 7.15-7.36 (m, 10H), 7.90 (d, *J* = 7.0 Hz, 2H), 8.01 (s, 2H), 8.58 (s, 1H), 10.36 (s, 1H). $^{13}$C NMR (125 MHz, $CDCl_3$) 21.6, 21.7, 35.0, 37.9, 40.8, 45.6, 60.4, 64.8, 73.8, 81.8, 114.4, 116.2, 126.8, 127.0, 127.2, 127.23, 128.0, 128.1, 129.1, 129.14, 129.7, 129.8, 138.9, 143.6, 144.0, 153.7, 156.2, 156.8, 166.2, 166.5. ESI-MS $(M+H)^+$ for $C_{37}H_{37}N_5O_6$: expected 648.2822, found 648.2852.

(2R,3S,5R)-5-(3-Benzyl-2-(((dimethylamino)methylene)amino)-5-(2-hydroxyethyl)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl) tetrahydrofuran-3-yl-4-methylbenzoate (**6**). To a solution of **5** (0.52 g, 0.8 mmol) in dry DMF (5.0 mL), DBU (280.0 µL, 1.8 mmol) and ethylene carbonate

(0.21 g, 2.4 mmol) was added. The reaction was heated under 90 °C for 4 h. After removal of DMF, the reaction mixture was diluted with ethyl acetate and washed successively with water and brine. The solution was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with hexane/ethyl acetate 1:4 to1:8) to give compound **6** as a white solid (0.27 g, 49%). mp 75-76°C. $^1$H NMR (500 MHz, $CDCl_3$) δ ppm 2.40 (s, 3H), 2.44(s, 3H), 2.52-2.57 (m, 1H), 2.85-2.89 (m, 1H), 3.05 (s, 3H), 3.15 (s, 3H), 3.90 (s, 2H), 4.46 (s, 2H), 4.51 (s, 1H), 4.59 (dd, *J* = 11.0, 4.0 Hz, 1H), 4.70 (dd, *J* = 11.3, 4.5 Hz, 1H), 5.53 (s, 3H), 5.76 (s, 1H), 7.09 (s, 1H), 7.19-7.21 (m, 3H), 7.26-7.29 (m, 4H), 7.35 (d, *J* = 5.0 Hz, 2H), 7.94 (d, *J* = 7.5 Hz, 2H), 8.01 (d, *J* = 7.5 Hz, 2H), 8.60 (s, 1H). $^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm 21.5, 21.6, 34.9, 37.8, 40.7, 45.2, 50.8, 63.0, 64.5, 73.5, 81.7, 113.4, 115.2, 126.6, 127.0, 127.04, 127.6, 128.0, 129.0, 129.4, 129.6, 130.7, 138.6, 143.6, 143.8, 153.7, 156.3, 156.6, 166.0, 166.3. ESI-MS $(M+H)^+$ for $C_{39}H_{41}N_5O_7$: expected 692.3084, found 692.3122.

2-Amino-3-benzyl-7-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-5-(2-hydroxyethyl)-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**7**). A suspension of **6** (0.15 g, 0.22 mmol) in 1 M sodium hydroxide in a mixture of methanol (2.0 mL) and water (1.0 mL) was heated under 70 °C for 16 h and allowed to cool to room temperature. The reaction mixture was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with hexane/ethyl acetate 1:5 to 1:10) to give **7** as a white solid (0.063 g, 72 %). mp 99-100°C. $^1$H NMR (500 MHz, $CD_3OD$) δ ppm 2.08 (dd, *J* = 13.0, 5.0 Hz, 1H), 2.46-2.52 (m, 1H), 3.69 (d, *J* = 12.0 Hz, 1H), 3.80–3.85 (m, 3H), 4.02 (s, 1H), 4.36-4.41 (m, 3H), 4.45 (d, *J* = 5.0 Hz, 1H), 5.28-5.32 (m, 3H), 7.22 (d, *J* = 7.5 Hz, 2H). 7.25 (d, *J* = 7.5 Hz, 2H), 7.31 (t, *J* = 8.0 Hz, 2H). $^{13}$C NMR (125 MHz, $CD_3OD$) δ ppm 42.1, 43.7, 50.5, 61.7, 63.4, 74.4, 74.7, 87.9,

112.4, 112.8, 126.17, 127.2, 128.4, 131.7, 135.7, 142.6, 151.0, 154.7. ESI-MS $(M+H)^+$ for $C_{20}H_{24}N_4O_5$: expected 401.1825, found 401.1839.

2-Amino-7-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-5-(2-hydroxyethyl)-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**2b**). To a mixture of **7** (0.063 g, 0.16 mmol) and 10% palladium on carbon (0.03 g) in methanol (2.0 mL) was added ammonium formate (0.10 g, 1.5 mmol) under Ar. The reaction mixture was heated under 75 °C for 16 h and then filtered through Celite. The filtrate was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with methanol/ethyl acetate 1:50 to 1:5) to give **2b** as a white solid (0.024 g, 48%). mp 124-125°C. $^1$H NMR (500 MHz, $CD_3OD$) δ 2.04 (dd, *J* = 11.0, 4.5 Hz, 1H), 2.43 (ddd, *J* = 11.0, 9.5, 4.5 Hz, 1H), 3.70 (dd, *J* = 10.0, 2.0 Hz, 1H), 3.79-3.83 (m, 3H), 4.00 (s, 1H) 4.34-4.37 (m, 2H), 4.44 (d, *J* = 4.5 Hz, 1H), 5.26 (dd, *J* = 9.5, 4.5 Hz, 1H). 7.20 (s, 1H). $^{13}$C NMR was not available due to the poor solubility. ESI-MS $(M+H)^+$ for $C_{13}H_{18}N_4O_5$: expected 311.1355, found 311.1377.

(2R,3S,5R)-5-(5-Allyl-3-benzyl-2-(((dimethylamino)methylene)amino)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3- yl 4-methylbenzoate (**8**). To a suspension of **5** (1.61 g, 2.4 mmol) in dry THF (25.0 mL), sodium hydride (0.1 g, 2.3 mmol, 60% in mineral oil) was added. The mixture was stirred for 15 min before allyl bromide (150 µL, 3.5 mmol) was added. The mixture was stirred at room temperature for 16 h. After removing the solvent, the reaction was diluted with methylene chloride and washed successively with water and brine. The organic layer was separated, dried over $MgSO_4$, and concentrated. Purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with hexane/ethyl acetate 5:1 to 3:1) to give compound **8** as a white solid (0.82 g, 48%).

mp 94-95°C. $^{1}$H NMR (500 MHz, $CDCl_3$) δ ppm 2.41 (s, 3H), 2.45 (s, 3H), 2.53-2.57 (m, 1H), 2.89-2.92 (m, 1H), 3.04 (s, 3H), 3.12 (s, 3H), 4.52-4.55 (m, 1H), 4.60 (dd, *J* = 11.5, 4.0 Hz, 1H), 4.71 (dd, *J* = 11.5, 5.0 Hz, 1H), 5.01-5.09 (m, 2H), 5.13-5.23 (m, 2H), 5.55 (s, 3H), 5.77 (t, *J* = 5.0 Hz, 1H), 6.00-6.05 (m, 1H), 7.06 (s, 1H), 7.19-7.23 (m, 3H), 7.26-7.29 (m, 4H), 7.38-7.39 (m, 2H), 7.95 (d, *J* = 7.5 Hz, 2H), 8.01 (d, *J* = 7.5 Hz, 2H), 8.59 (s, 1H).$^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm 21.8, 21.83, 35.2, 38.4, 41.0, 45.3, 50.6, 64.8, 68.6, 73.8, 82.1, 114.0, 117.6, 126.8, 127.3, 127.9, 128.3, 129.0, 129.2, 129.23, 129.26, 129.3, 129.85, 129.9, 130.1, 134.5, 143.8, 144.0, 154.2, 166.5. ESI-MS $(M+H)^+$ for $C_{40}H_{41}N_5O_6$: expected 688.3135, found 688.3202.

(2R,3S,5R)-5-(3-Benzyl-5-(2,3-dihydroxypropyl)-2-(((dimethylamino)methylene) amino)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl 4-methylbenzoate (**9**). To the suspension of **8** (0.82 g, 1.2 mmol) in a mixture of acetone (4.0 mL) and water (1.0 mL) was added TBHP (220 µL, 1.1-1.3 mmol, 5.0-6.0 M in decane), TBAF (0.028 g, 0.11 mmol) and $O_SO_4$ (trace). The reaction was stirred at room temperature for overnight. The reaction mixture was diluted with methylene chloride and washed successively with water and brine. The organic layer was separated, dried over $MgSO_4$, and concentrated. Purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with hexane/ethyl acetate 1:1 to 1:3) to give compound **9** as a light yellow solid (0.66 g, 76%). mp 80-81°C. $^{1}$H NMR (500 MHz, $CDCl_3$) δ ppm 2.40 (s, 3H), 2.44 (s, 3H), 2.52-2.60 (m, 1H), 2.84-2.93 (m, 1H), 3.06 (s, 3H), 3.15 (s, 3H), 3.51 (s, 2H), 3.88-3.94 (m, 2H), 4.30-4.34 (m, 1H), 4.51-4.60 (m, 3H), 4.69-4.72 (m, 1H), 5.54 (s, 3H), 7.11 (s, 1H), 7.19–7.35 (m, 9H), 7.93 (d, *J* = 8.0 Hz, 2H), 8.01 (s, 2H), 8.58 (s, 1H). $^{13}$C NMR for both

diastereomers (125 MHz, $CDCl_3$) δ ppm 21.0, 21.6, 35.0, 37.8, 40.8, 45.4, 50.0, 50.1, 60.4, 62.9, 63.0, 64.6, 71.8, 73.5, 73.6, 81.8, 113.8, 126.8, 127.08, 127.1, 127.7, 128.1, 129.0, 129.09, 129.1, 129.2, 129.68, 129.7, 129.72, 131.7, 138.7, 143.7, 143.9, 153.8, 156.8, 166.4. ESI-MS $(M+H)^+$ for $C_{40}H_{43}N_5O_8$: expected 722.3190, found 722.3193.

2-Amino-3-benzyl-5-(2,3-dihydroxypropyl)-7-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**10**). A suspension of **9** (0.66 g, 0.91 mmol) in 1 M sodium hydroxide in a mixture of methanol (3 mL) and water (0.5 mL) was heated under 70 °C for 16 h and allowed to cool to room temperature. The reaction mixture was concentrated and purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with methanol/ethyl acetate 1:20 to 1:10) to give **10** as a white solid (0.21 g, 54%). mp 93-94°C. $^1$H NMR (500 MHz, $CD_3OD$) δ ppm 2.07 (dd, *J* = 13.0, 5.5 Hz, 1H), 2.51 (ddd, *J* = 13.0, 11.5, 5.5 Hz, 1H), 3.48 (ddd, *J* = 11.5, 5.5, 3.0 Hz, 1H), 3.54 (ddd, *J* = 11.5, 5.0, 2.0 Hz, 1H), 3.69 (dd, *J* = 9.5, 2.5 Hz, 1H), 3.82 (dd, *J* = 9.5, 2.5 Hz, 1H), 3.93–3.97 (m, 1H), 4.00–4.02 (m, 1H), 4.26 (dt, *J* = 14.0, 5.0 Hz, 1H), 4.45–4.46 (d, *J* = 5.0 Hz, 1H), 4.51-4.56 (m, 1H), 5.29-5.35 (m, 3H), 7.23-7.29 (m, 4H), 7.32-7.35 (m, 2H). $^{13}$C NMR for both diastereomers (125 MHz, $CD_3OD$) δ ppm 44.18, 44.2, 45.9, 52.6, 62.2, 65.4, 65.5, 73.9, 76.5, 76.8, 90.1, 114.9, 115.1, 128.3, 129.3, 130.5, 134.3, 137.9, 145.3, 153.1, 157.2. ESI-MS $(M+H)^+$ for $C_{21}H_{26}N_4O_6$: expected 431.1931, found 431.1966.

2-Amino-5-(2,3-dihydroxypropyl)-7-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-2-yl)-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**2d**). To a mixture of **10** (0.21 g, 0.49 mmol) and 10% palladium on carbon (0.1 g) in methanol (5 mL) was added ammonium formate (0.29 g, 4.5 mmol) under Ar. The reaction mixture was heated under 75 °C

for 16 h and then filtered through Celite. The filtrate was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with methanol/ethyl acetate 1:20 to 1:5) to give **2d** as a white solid (0.12 g, 74%). mp 135-136 °C. $^1$H NMR (500 MHz, $d_6$-DMSO) δ ppm 1.90 (dd, *J* = 13.8, 4.8 Hz, 1H), 2.11-2.17 (m, 1H), 3.27 (t, *J* = 5.0 Hz, 2H), 3.45 (s, 2H), 3.71 (s, 2H), 4.03-4.09 (m, 1H), 4.18 (s, 1H), 4.32-4.35 (m, 1H), 4.66 (s, 1H), 4.91 (s, 2H), 5.07 (dd, *J* = 11.2, 6.0 Hz, 1H), 5.20 (br, 1 H), 5.72 (s, 2H), 7.15 (s, 1H), 10.53 (s, 1H). $^{13}$C NMR for both of the diastereomers (150 MHz, $d_6$-DMSO) δ ppm 42.7, 51.6, 59.9, 63.9, 64.4, 72.3, 73.5, 74.0, 88.4, 113.3, 114.4, 131.3, 132.3, 151.6, 156.3. ESI-MS $(M+H)^+$ for $C_{14}H_{20}N_4O_6$: expected 341.1461, found 341.1486.

2-(2-Amino-7-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-4-oxo-3,4-dihydro-5H-pyrrolo[3,2-d]pyrimidin-5-yl)acetaldehyde (**2c**). To compound **2d** (0.050 g, 0.15 mmol) in a mixture of methanol (1 mL) and water (1 mL) was added potassium periodate (0.035 g, 0.15 mmol). The reaction mixture was stirred under room temperature for 30 min, then concentrate and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with methanol/ethyl acetate 1:20 to 1:5) to give **2c** as a white solid (0.039 g, 85%). mp 117-118 °C. $^1$H NMR (600 MHz, $d_6$-DMSO) δ ppm 1.91-1.99 (m, 1H), 2.10-2.13 (m, 1H), 3.44 (s, 2H), 3.72 (s, 1H), 4.20-4.22 (m, 1H), 4.93-4.95 (m, 1H), 5.07 (s, 2H), 5.84-6.01 (m, 2H), 7.16 (s, 1H), 9.60 (s, 1H), 10.66-10.72 (s, 1H). $^{13}$C NMR was not available due to the poor solubility. ESI-MS $(M+H)^+$ for $C_{13}H_{16}N_4O_5$: expected 309.1199, found 309.1220.

(2R,3S,5R)-5-(3-Benzyl-5-(2-((tert-butyldimethylsilyl)oxy)ethyl)-2-(((dimethylamino)methylene)amino)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl 4-methylbenzoate (**13**). To a

solution of **6** (0.45 g, 0.65 mmol) in DMF was added imidazole (0.13 g, 1.9 mmol), 4-DMAP (0.002 g, 0.018 mmol), TBSCl (0.29 g, 1.9 mmol). The reaction was stirred under room temperature for 16 h. After removal of the solvent, the mixture was diluted by methylene chloride and washed with water and brine. The solution was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with hexane/ethyl acetate 5:1 to 3:1) to give compound **13** as a light yellow solid (0.43 g, 82 %). mp 50-51°C. $^1$H NMR (500 MHz, $CDCl_3$) δ ppm -0.12 (s, 6H), 0.83 (s, 9H), 2.40 (s, 3H), 2.44 (s, 3H), 2.49-2.52 (m, 1H), 2.87-2.89 (m, 1H), 3.03 (s, 3H), 3.10 (s, 3H), 3.93 (s, 2H), 4.44 (s, 2H), 4.52 (s, 1H), 4.60 (dd, *J* = 11.3, 4.5 Hz, 1H), 4.66 (dd, *J* = 11.5, 5.0 Hz, 1H), 5.56 (s, 3H), 5.78 (s, 1H), 7.12 (s, 1H), 7.19–7.21 (m, 3H), 7.25-7.29 (m, 4H), 7.35-7.37 (m, 2H), 7.96 (d, *J* = 7.5 Hz, 2H), 8.01 (d, *J* = 6.5 Hz, 2H), 8.59 (s, 1H). $^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm -5.4, 18.4, 21.9, 21.94, 26.1, 31.8, 35.2, 41.0, 45.3, 51.3, 63.5, 65.0, 73.6, 82.1, 113.0, 114.7, 126.8, 127.46, 127.5, 127.9, 128.3, 129.3, 129.4, 129.96, 130.0, 139.3, 143.1, 143.9, 144.1, 154.2, 156.0, 156.8, 166.4, 166.6. ESI-MS $(M+H)^+$ for $C_{45}H_{55}N_5O_7Si$: expected 806.3943, found 806.3944.

2-Amino-3-benzyl-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**14**). A suspension of **6** (3 g, 10.2 mmol) in 1 M sodium hydroxide in a mixture of methanol (12 mL) and water (6 mL) was heated under 70 °C overnight and allowed to cool to room temperature. The reaction mixture was concentrated and purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with ethyl acetate) to give **13** as a white solid (1.86 g, 75%). The characterization of this compound was previously described by Gibson *et al*.

(2R,3S,5R)-5-(2-Amino-3-benzyl-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl-4-methylbenzoate (**15**). To

compound **14** (0.31 g, 1.3 mmol) and 1-(α,β)-*O*-methyl-3,5-di-(*O*-*p*-toluoyl)- 2-deoxy-D-ribose (0.7 g, 1.9 mmol) in a mixture of methylene chloride (3.0 mL) and acetonitrile (3.0 mL) was added a solution of $SnCl_4$ (2.5 mL, 2.5 mmol, 1 M in $CH_2Cl_2$). The reaction mixture was heated at 65 °C for 16 h. The reaction mixture was diluted with methylene chloride and washed successively with satured $NaHCO_3$ and brine. The organic layer was separated and dried over $MgSO_4$. The solution was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with hexane/ethyl acetate 1:1 to 1:2) to give compound **15** as a light yellow solid (0.20 g, 26%). mp 84-85°C. $^1$H NMR (500 MHz, $CDCl_3$) δ ppm 2.37 (s, 3H), 2.41 (s, 3H), 2.60-2.63 (m, 2H), 4.53 (s, 1H), 4.64 (dd, *J* = 12.0, 4.0 Hz, 1H), 4.78 (dd, *J* = 11.5, 4.5 Hz, 1H), 5.34-5.42 (m, 2H), 5.47-5.50 (m, 1H), 5.65-5.66 (m, 1H), 7.18-7.33 (m, 10H), 7.91 (d, *J* = 7.5 Hz, 2H), 7.98 (d, *J* = 8.0 Hz, 2H), 10.54 (s, 1H). $^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm 21.9, 22.0, 30.0, 39.3, 45.1, 65.1, 73.6, 82.9, 113.0, 113.7, 127.0, 127.2, 127.3, 128.5, 129.3, 129.39, 129.4, 129.5, 130.2, 134.8, 143.7, 144.1, 144.2, 151.3, 166.5, 166.8. ESI-MS $(M+H)^+$ for $C_{34}H_{32}N_4O_6$: expected 593.2401, found 593.2422.

(2R,3S,5R,5S)-5-(2-amino-3-benzyl-5-(2-hydroxyethyl)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl-methylbenzoate (**16**). To compound **15** (0.100 g, 0.17 mmol) in dry DMF (2 mL), DBU (28 µL, 0.18 mmol) and ethylene carbonate (0.04 g, 0.45 mmol) was added. The reaction was heated under 110 °C for 4 h. After removal of the solvent, the reaction mixture was diluted with ethyl acetate and washed with water and brine. The residue was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with hexane/ethyl acetate 1:5 to 1:8) to give compound **16** as a white solid (0.050g, 46%). The ratio of the epimers (α:β = 1:3) was

determined by $^{1}$H NMR. $^{1}$H NMR for both epimers (600 MHz, $CDCl_3$) δ ppm 2.38 (s, 4.78H), 2.41 (s, 3H), 2.50 (dd, *J*= 11.2, 5.6 Hz, 1.35H), 2.62–2.65 (m, 1.38H), 3.35 (s, 0.66H), 3.67 (s, 2H), 3.89 (t, *J*= 5.6 Hz, 2.88H), 4.42 (t, *J*= 5.6 Hz, 2H), 4.46-4.47 (m, 1.59H), 4.51-4.53 (m, 0.53H), 4.55 (dd, *J*= 11.2, 5.6 Hz, 1H), 4.64-4.65 (m, 0.39H), 4.75 (dd, *J*= 11.2, 5.6 Hz, 1H), 5.11 (s, 0.61H), 5.20-5.29 (m, 3.21H), 5.42 (dd, *J*= 9.6, 4.8 Hz, 1H), 5.48 (t, *J*= 6.6 Hz, 0.38H), 5.52–5.58 (m, 0.34H), 5.64 (d, *J*= 5.4 Hz, 1H), 7.08 (s, 1H), 7.16-7.36 (m, 13.72H), 7.78 (d, *J* = 7.2 Hz, 0.62H), 7.93 (d, *J* = 7.8 Hz, 2.47H), 7.97 (d, *J* = 8.2 Hz, 2H). $^{13}$C NMR for both epimers (150 MHz, $CDCl_3$) δ ppm 21.80, 21.82, 42.2, 44.6, 51.3, 62.2, 64.7, 73.2, 83.2, 110.6, 110.9, 112.4, 126.9, 127.1, 127.3, 128.6, 129.2, 129.3, 129.8, 129.9, 130.1, 130.2, 131.2, 133.4, 144.0, 144.1, 144.2, 151.36, 151.4, 152.5, 166.5, 166.6, 166.7.

2-Amino-3-benzyl-5-(2-hydroxyethyl)-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**17**). To compound **14** (1.3g, 4.4 mmol) in dry DMF (5 mL), DBU (280 µL, 1.83 mmol) and ethylene carbonate (1.0 g, 11.7 mmol) was added. The reaction was heated under 90 °C for 4 h. After removal of the solvent, the reaction mixture was diluted with ethyl acetate and washed with water and brine. The residue was chromatographed by ethyl acetate to give compound **17** as a white solid (0.6 g, 47%). mp 168-169 °C. $^{1}$H NMR (500 MHz, $d_6$-DMSO) δ ppm 3.65 (dd, *J* = 9.5, 5.0 Hz, 2H), 4.30 (t, *J* = 4.5 Hz, 2H), 5.23 (s, 2H), 5.90 (s, 1H), 6.25 (s, 1H), 7.21–7.26 (m, 4H), 7.32 (t, *J* = 6.0 Hz, 2H). $^{13}$C NMR (150 MHz, $CD_3OD$) δ ppm 45.1, 51.9, 63.2, 100.3, 112.7, 127.5, 128.5, 129.8, 134.6, 137.3, 146.9, 152.8, 156.2. ESI-MS $(M+H)^+$ for $C_{15}H_{16}N_4O_2$: expected 285.1352, found 285.1337.

(2R,3S,5S)-5-(2-Amino-3-benzyl-5-(2-hydroxyethyl)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl-4-

methyl benzoate (**18**). To compound **17** (0.6 g, 2.2 mmol) and 1-(α, β)-O-methyl-3,5-di-(O-*p*-toluoyl)- 2-deoxy-D-ribose (2.4 g, 3.3 mmol) in a mixture of methylene chloride (5.0 mL) and acetonitrile (5.0 mL) was added a solution of $SnCl_4$ (4.5 mL, 4.5 mmol, 1 M in $CH_2Cl_2$). The reaction mixture was heated at 65 °C for 16 h. The reaction mixture was diluted with methylene chloride and washed successively with sat. $NaHCO_3$ and brine. The organic layer was separated, dried over $MgSO_4$. The solution was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with hexane/ethyl acetate 1:1 to 1:2) to give compound **18** as a white solid (0.39 g, 28%). mp 89-90 °C. $^1$H NMR (500 MHz, $CDCl_3$) δ ppm 2.40 (s, 3H), 2.43 (s, 3H), 2.50 (dd, *J* = 13.0, 5.0 Hz, 1H), 2.67 (ddd, *J* = 10.5, 10.5, 4.5 Hz, 1H), 3.60 (s, 1H), 3.90 (s, 2H), 4.42 (s, 3H), 4.55 (dd, *J* = 10.5, 4.5 Hz, 1H), 4.73 (dd, *J* = 10.5, 4.5 Hz, 1H), 5.02 (s, 2H), 5.19-5.26 (m, 2H), 5.45 (dd, *J* = 10.5, 5.0 Hz, 1H), 5.65 (d, *J* = 5.0 Hz, 1H), 7.09 (s, 1H), 7.20-7.32 (m, 9H), 7.95 (d, *J* = 8.0 Hz, 2H), 7.99 (d, *J* = 8.0 Hz, 2H). $^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm 21.87, 21.9, 38.6, 44.8, 51.3, 63.0, 65.0, 73.5, 82.6, 112.8, 113.1, 126.7, 127.3, 127.4, 128.2, 129.38, 129.4, 129.9, 130.0, 131.0, 135.3, 144.1, 144.3, 150.9, 155.2, 166.4, 166.7. ESI-MS $(M+H)^+$ for $C_{36}H_{36}N_4O_7$: expected 637.2663, found 637.2697.

(2R,3S,5S)-5-(2-Amino-3-benzyl-5-(2-((tert-butyldimethylsilyl)oxy)ethyl)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl) tetrahydrofuran-3-yl 4-methylbenzoate (**19**). To Compound **18** (0.39 g, 0.6 mmol) in DMF was added imidazole (0.13 g, 1.9 mmol), 4-DMAP (0.002 g, 0.018 mmol), TBSCl (0.3 g, 2.0 mmol). The reaction was stirred under room temperature for 16 h. After removal of the solvent, the mixture was diluted by methylene chloride and washed with water and brine. The solution was concentrated and purified by column ($SiO_2$, 0.06-0.20mm, eluting with hexane/ethyl acetate 5:1

to 3:1) to give compound **19** as a gummy solid (0.42 g, 93%). mp 55-56 °C. $^{1}$H NMR (500 MHz, $CDCl_3$) δ ppm -0.11 (d, *J* = 5.0 Hz, 6H), 0.82 (s, 9H), 2.39 (s, 3H), 2.41 (s, 3H), 2.44 (dd, *J* = 11.5, 4.5 Hz, 1H), 2.62 (ddd, *J* = 11.5, 11.5, 4.5 Hz, 1H), 3.90 (t, *J* = 3.5 Hz, 2H), 4.31-4.37 (m, 2H), 4.44 (dt, *J* = 11.5, 3.5 Hz, 1H), 4.51 (dd, *J* = 9.5, 3.5 Hz, 1H), 4.67 (dd, *J* = 9.5, 4.0 Hz, 1H), 5.14-5.28 (m, 3H), 5.45 (dd, *J* = 9.5, 4.5 Hz, 1H), 5.65 (d, *J* = 4.5 Hz, 1H), 7.08 (s, 1H), 7.19-7.30 (m, 9H), 7.94 (d, *J* = 7.0 Hz, 2H), 7.97 (d, *J* = 6.5 Hz, 2H). $^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm -5.6, 18.1, 21.6, 21.65, 25.8, 37.8, 44.3, 51.1, 53.5, 63.1, 64.7, 72.6, 82.3, 111.7, 112.2, 126.1, 127.1, 127.3, 127.8, 129.0, 129.1, 129.7, 129.74, 130.9, 135.3, 143.7, 144.0, 151.4, 154.7, 166.0, 166.2. ESI-MS $(M+H)^+$ for $C_{42}H_{50}N_4O_7Si$: expected 751.3528, found 751. 3505.

2-Amino-3-benzyl-5-(2-((tert-butyldimethylsilyl)oxy)ethyl)-7-((2S,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**20**). To Compound **19** (0.42 g, 0.56 mmol) in a mixture of methanol (10.0 mL) and water (1.0 mL) was added sodium hydroxide (0.05 M). The reaction mixture was stirred under room temperature for 1 h. The reaction mixture was concentrated and purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with methanol/ethyl acetate 1:50 to 1:20) to give **20** as a white solid ( 0.21 g, 74%). mp 57-58 °C. $^{1}$H NMR (500 MHz, $CD_3OD$) δ ppm -0.09 (d, *J*= 10.0 Hz, 6H), 0.82 (s, 9H), 2.02 (dd, *J* = 11.0, 5.0 Hz, 1H), 2.51–2.56 (m, 1H), 3.68 (d, *J* = 12.0 Hz, 1H), 3.87–3.90 (m, 2H), 4.08 (s, 1H), 4.31 (dt, *J* = 14.0, 5.0 Hz, 1H), 4.53–4.58 (m, 2H), 5.13 (d, *J* = 16.0 Hz, 1H), 5.26 (dd, *J* = 11.5, 5.5 Hz, 1H), 5.40 (d, *J* = 16.0 Hz, 1H), 7.06 (s, 1H), 7.22–7.35 (m, 5H). $^{13}$C NMR (125 MHz, $CD_3OD$) δ ppm -4.5, 19.7, 27.2, 44.4,

45.7, 52.6, 64.9, 65.5, 76.7, 89.9, 114.3, 114.7, 128.2, 129.2, 130.4, 134.0, 137.7, 144.9, 152.9, 156.6. ESI-MS $(M+H)^+$ for $C_{26}H_{38}N_4O_5Si$: expected 515.2690, found 515.2706.

2-Amino-5-(2-((tert-butyldimethylsilyl)oxy)ethyl)-7-((2S,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**3b**). To compound **20** (0.21 g, 0.41 mmol) and 10% palladium on carbon (0.11 g) in methanol (5.0 mL) was added ammonium formate (0.28 g, 4.5 mmol) under Ar. The reaction mixture was heated under 75 °C for 16 h and then filtered through celite. The filtrate was concentrate and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with methanol/ethyl acetate 1:20 to 1:5) to give 3b as a white gummy solid (0.075 g, 43%). mp 49-50 °C. $^1$H NMR (600 MHz, $d_6$-DMSO) δ ppm -0.13 (s, 6H), 0.79 (s, 9H), 1.88 (dd, *J* = 12.2, 5.4 Hz, 1H), 2.06-2.11 (m, 1H), 3.42 (s, 2H), 3.72 (s, 1H), 3.81 (t, *J* = 5.4 Hz, 2H), 4.17 (s, 1H), 4.25 (t, *J* = 4.8 Hz, 2H), 4.90 (s, 1H), 5.07 (dd, *J* = 12.0, 5.4 Hz, 1H), 5.71 (s, 1H), 7.12 (s, 1H), 10.78 (s, 1H). $^{13}$C NMR (150 MHz, $d_6$-DMSO) δ ppm -4.7, 18.8, 26.7, 42.8, 51.1, 63.9, 73.4, 74.0, 88.4, 112.8, 114.5, 130.9, 145.8, 151.6, 155.6. ESI-MS $(M+H)^+$ for $C_{19}H_{32}N_4O_5Si$: expected 425.2221, found 425.2260.

N'-(5-Allyl-3-benzyl-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-2-yl)-N,N-dimethylformimidamide (**25**). To a suspension of compound 4 (2.0 g, 6.8 mmol) in dry THF (25.0 mL), sodium hydride (0.28 g, 6.7 mmol, 60% in mineral oil) was added. The mixture was stirred for 15min before allyl bromide (330.0 μL, 8.0 mmol) was added. The mixture was stirred at room temperature for 16 h. After removing the solvent, the reaction was diluted with methylene chloride and washed successively with water and brine. The organic layer was separated, dried over $MgSO_4$ and concentrated. Purified by column chromatography ($SiO_2$,

0.06–0.20mm, eluting with hexane/ethyl acetate 5:1 to 3:1) to give compound **25** as a light yellow solid (1.4 g, 64%). mp 93-94 °C. $^{1}$H NMR (500 MHz, $CDCl_3$) δ ppm 3.06 (s, 3H), 3.15 (s, 3H), 5.06-5.09 (m, 3H), 5.19 (d, *J* = 10.5 Hz, 1H), 5.52 (s, 2H), 6.05 (ddt, *J* = 16.5, 10.5, 5.0 Hz, 1H), 6.31 (s, 1H), 7.03 (s, 1H), 7.19-7.21 (m, 1H), 7.27 (t, *J* = 7.5 Hz, 2H), 7.37 (d, *J* = 7.5 Hz, 2H), 8.61 (s, 1H). $^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm 35.0, 40.8, 45.1, 50.4, 101.3, 114.3, 116.9, 126.7, 127.9, 128.1, 130.4, 134.8, 139.0, 144.4, 154.3, 155.8, 156.2. ESI-MS $(M+H)^+$ for $C_{19}H_{21}N_5O$: expected 336.1824, found 336.1832.

5-Allyl-2-amino-3-benzyl-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**26**). To Compound **25** (1.0 g, 3 mmol) in a mixture of methanol (20.0 mL) and water (5.0 mL) was added sodium hydroxide (1.0 M). The reaction mixture was heated under 70 °C for 16 h. The reaction mixture was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with ethyl acetate) to give **26** as a white solid (0.58 g, 70%). mp 121-122 °C. $^{1}$H NMR (500 MHz, $CD_3OD$) δ ppm 4.99-5.05 (m, 3H), 5.16 (dd, *J* = 10.5, 1.5 Hz, 1H), 5.33 (s, 2H), 6.04-6.12 (m, 2H), 7.21-7.30 (m, 4H), 7.35 (t, *J* = 7.5 Hz, 2H). $^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm 44.5, 50.5, 101.3, 112.3, 117.0, 126.5, 127.9, 129.1, 131.1, 134.6, 135.7, 144.9, 151.1, 154.8. ESI-MS $(M+H)^+$ for $C_{16}H_{16}N_4O$: expected 281.1403, found 281.1387.

(2R,3S,5R)-5-(5-Allyl-2-amino-3-benzyl-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl-4-methylbenzoate (**27**). To compound **26** (0.58 g, 2.1 mmol) and 1-(α,β)-*O*-methyl-3,5-di-(*O*-*p*-toluoyl)-2-deoxy-D-ribose (2.3 g, 3.1 mmol) in a mixture of methylene chloride (5.0 mL) and acetonitrile (5.0 mL) was added a solution of $SnCl_4$ (4.2 mL, 4.2 mmol, 1 M in $CH_2Cl_2$). The reaction mixture was heated at 65 °C for 16 h. The reaction mixture was diluted with

methylene chloride and washed successively with sat. $NaHCO_3$ and brine. The organic layer was separated, dried over $MgSO_4$, and concentrated. Purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with hexane/ethyl acetate 2:1 to 1:2) to give compound **27** as a light yellow solid (0.30 g, 23%). mp 59-60 °C. $^1$H NMR (600 MHz, $CDCl_3$) δ ppm 2.40 (s, 3H), 2.41 (s, 3H), 2.56-2.59 (m, 2H), 4.47 (dt, *J* = 4.2, 1.8 Hz, 1H), 4.57 (dd, *J* = 12.0, 4.2 Hz, 1H), 4.76 (dd, *J* = 12.0, 4.2 Hz, 1H), 4.94–4.96 (m, 2H), 5.09 (d, *J* = 10.4 Hz, 1H), 5.17 (d, *J* = 10.4 Hz, 1H), 5.29 (s, 2H), 5.48 (t, *J* = 8.4 Hz, 1H), 5.64-5.66 (m, 1H), 5.95–6.02 (m, 1H), 7.04 (s, 1H), 7.23-7.29 (m, 7H), 7.31-7.34 (m, 2H), 7.94 (d, *J* = 8.4 Hz, 2H), 7.97 (d, *J* = 8.4 Hz, 2H). $^{13}$C NMR (125 MHz, $CDCl_3$) δ ppm 21.75, 21.8, 38.6, 44.5, 50.6, 64.8, 73.2, 77.4, 82.5, 112.8, 113.1, 117.6, 126.4, 127.2, 127.3, 128.0, 129.0, 129.2, 129.23, 129.8, 134.2, 135.4, 143.8, 144.1, 151.3, 154.7, 166.2, 166.4. ESI-MS $(M+H)^+$ for $C_{37}H_{36}N_4O_6$: expected 633.2714, found 633.2745.

(2R,3S,5R)-5-(2-Amino-3-benzyl-5-(2,3-dihydroxypropyl)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl)tetrahydrofuran-3-yl 4-methylbenzoate (**28**). To the suspension of compound **27** (0.30 g, 0.48 mmol) in a mixture of acetone (4.0 mL) and water (1.0 mL) was added TBHP (100 µL, 0.5-0.6 mmol, 5.0-6.0 M in decane), TBAF (0.012 g, 0.048 mmol) and $O_SO_4$ (trace). The reaction was stirred at room temperature for 16 h. The reaction mixture was diluted with methylene chloride and washed successively with water and brine. The organic layer was separated, dried over $MgSO_4$, and concentrated. The solution was purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with hexane/ethyl acetate 1:1 to 1:3) to give compound **28** as a light yellow solid (0.13 g, 40%). mp 61-62 °C. $^1$H NMR (500 MHz, $CDCl_3$) δ ppm 2.36 (s, 3H), 2.39 (s, 3H), 2.50-2.58

(m, 2H), 3.46 (dd, *J* = 12.0, 5.5 Hz, 1H), 3.52 (ddd, *J* = 11.5, 4.5, 1.5 Hz, 1H), 3.88-3.93 (m, 1H), 4.18 (dt, *J* = 14.0, 7.0 Hz, 1H), 4.40 (dt, *J* = 5.0, 2.0 Hz, 1H), 4.44 (dd, *J* = 14.0, 3.5Hz, 1H ), 4.54 (ddd, *J* = 12.0, 5.0, 2.0 Hz, 1H), 4.67 (dt, *J* = 12.0, 5.4 Hz, 1H), 5.26 (s, 2H), 5.43 (dd, *J* = 10.0, 6.0 Hz, 1H), 5.59-5.60 (m, 1H), 7.19-7.28 (m, 10H), 7.89 (d, *J* = 8.5 Hz, 2H), 7.94 (d, *J* = 8.0 Hz, 2H). $^{13}$C NMR for both diastereomers (125 MHz, $CDCl_3$) δ ppm 22.5, 22.56, 46.0, 52.79, 52.8, 65.4, 65.4, 66.7, 73.8, 73.85, 75.5, 79.51, 84.6, 114.1, 114.3, 128.3, 129.0, 129.1, 129.40, 130.6, 131.1, 131.5, 131.53, 133.7, 137.7, 146.2, 146.3, 153.4, 156.9, 168.4, 168.6. ESI-MS $(M+H)^+$ for $C_{37}H_{38}N_4O_8$: expected 667.2769, found 667.2786.

(2R,3S,5R)-5-(2-Amino-3-benzyl-5-(2,3-bis((tert-butyldimethylsilyl)oxy)propyl)-4-oxo-4,5-dihydro-3H-pyrrolo[3,2-d]pyrimidin-7-yl)-2-(((4-methylbenzoyl)oxy)methyl) tetrahydrofuran-3-yl-4-methylbenzoate (**29**). To Compound **28** (0.13 g, 0.19 mmol) in DMF was added imidazole (0.08 g, 1.1 mmol), 4-DMAP (0.002 g, 0.018 mmol), TBSCl (0.17 g, 1.1 mmol). The reaction was stirred under room temperature for 16 h. After removal of the solvent, the mixture was diluted by methylene chloride and washed with water and brine. The solution was concentrated and purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with hexane/ethyl acetate 5:1 to 3:1) to give compound **29** as a white gummy solid (0.13 g, 76%). mp 44-45 °C. $^1$H NMR (600 MHz, $CDCl_3$) δ ppm -0.31 (s, 3H), -0.07 (d, *J* = 10.2 Hz, 3H), 0.07 (s, 6H), 0.78 (d, *J* = 2.4 Hz, 9H), 0.91 (s, 9H), 2.39 (s, 3H), 2.40 (s, 4H), 2.42-2.44 (m, 1H), 3.52-3.55 (m, 1H), 3.61 (dd, *J* = 10.2, 3.6 Hz, 1H), 3.99-4.14 (m, 2H), 4.50-4.52 (m, 1H), 4.60-4.64 (m, 2H), 4.77 (dd, *J* = 11.4, 4.8 Hz, 1H), 5.39-5.43 (m, 3H), 5.61-5.63 (m, 1H), 7.11 (s, 1H), 7.18-7.20 (m, 2H), 7.24-7.27 (m, 4H), 7.32-7.35 (m, 5H), 7.91 (d, *J* = 7.8 Hz, 2H), 8.02 (d, *J* = 7.8 Hz, 2H). $^{13}$C NMR for both diastereomers (600 MHz, $CDCl_3$)

δ ppm -5.1, -5.0, -4.5, 18.2, 18.6, 21.9, 22.0, 26.1, 26.2, 29.9, 38.5, 44.6, 52.3, 52.6, 65.2, 65.8, 72.9, 73.0, 73.6, 82.6, 112.2, 112.3, 126.7, 126.8, 127.4, 127.5, 128.3, 129.4, 129.5, 130.0, 130.1, 131.4, 131.6, 135.5, 143.9, 144.2, 150.9, 151.2, 166.4, 166.6. ESI-MS $(M+H)^+$ for $C_{49}H_{66}N_4O_8Si_2$: expected 895.4498, found 895.4525.

2-Amino-5-(2,3-bis((tert-butyldimethylsilyl)oxy)propyl)-7-((2R,4S,5R)-4-hydroxy-5-(hydroxymethyl)tetrahydrofuran-2-yl)-3,5-dihydro-4H-pyrrolo[3,2-d]pyrimidin-4-one (**3d**). To Compound **29** (0.13 g, 0.14 mmol) in a mixture of methanol (5 mL) and water (0.5 mL) was added sodium hydroxide (0.1 M). The reaction mixture was stirred under room temperature for 1 h. The reaction mixture was concentrated and purified by column chromatography ($SiO_2$, 0.06-0.20mm, eluting with methanol/ethyl acetate 1:50 to 1:20) to give **30** (0.065 g, 68%). To compound **30** (0.065 g, 0.10 mmol) and 10% palladium on carbon (0.030 g) in methanol (3.0 mL) was added ammonium formate (0.063 g, 1.0 mmol) under Ar. The reaction mixture was heated under 75 °C for 16 h and then filtered through celite. The filtrate was concentrated and purified by column chromatography ($SiO_2$, 0.06–0.20mm, eluting with methanol/ethyl acetate 1:20 to 1:5) to give **3d** as a white gummy solid (0.031 g, 56%). mp 53-54 °C. $^1$H NMR (600 MHz, $d_6$-DMSO) δ ppm -0.28 (d, *J* = 6.0 Hz, 3H), -0.08 (s, 3H), 0.05 (s, 6H), 0.79 (s, 9H), 0.89 (s, 9H), 1.88-1.91 (m, 1H), 2.02-2.11 (m, 1H), 3.46-3.50 (m, 2H), 3.70 (d, *J* = 15.6 Hz, 1H), 3.97-4.03 (m, 2H), 4.18 (d, *J* = 12.0 Hz, 1H), 4.34 (d, *J* = 10.2 Hz, 1H), 4.90 (s, 1H), 5.06-5.10 (m, 1H), 5.69 (d, *J* = 20.0 Hz, 1H), 7.02 (s, 1H), 10.50 (d, *J* = 15.2 Hz, 1H). $^{13}$C NMR (150 MHz, $d_6$-DMSO) δ ppm -4.5, -3.9, 18.6, 19.0, 26.7, 49.6, 52.1, 52.2, 63.9, 64.0, 66.3, 73.6, 73.9, 74.2, 88.3, 113.2, 114.8, 131.2, 145.8, 151.4, 155.3. ESI-MS $(M+H)^+$ for $C_{26}H_{48}N_4O_6Si_2$: expected 569.3191, found 569.3234.

# CHAPTER 3

# HIGHLY SPECIFIC COLORIMETRIC DETECTION OF DNA OXIDATION BIOMARKER USING GOLD NANOPARTICLE/TRILEX DNA CONJUGATES

## 3.1 Introduction

DNA is inevitably subjected to oxidative damage during cellular metabolism, and the oxidative damage can cause a variety of diseases including cancer. More than 20 oxidatively damaged DNA base lesions were identified in 2003 by Cooke et al. Among these lesions, 8-oxo-2'-deoxyguanosine (8-oxo-dG) and 8-oxo-guanosine (8-oxoGuo) have become the focus of intense research interest due to their relative higher abundance compared to other lesions. DNA repair is considered the major factor contributing to the urinary 8-oxoGua levels, although the precise pathways have not been fully understood. To get a better understanding of the role of such DNA lesion in the aetiology of many diseases such as cancer and ageing, methods of the 8-oxoGua analysis have been developed rapidly. Chomatographic methods and immunoassay are major detection methods for urinary 8-oxoGua. It has been reported that some of the detection methods may actually generate oxidative nucleobase during the sample preparation, although efforts have been made to minimize this problem.

### 3.1.1 Formation and Repair Mechanism of 8-OxodG

Normal cellular metabolism continuously generates reactive oxygen species (ROS) as byproducts, such as peroxides, superoxide, hydroxyl radical, and singlet oxygen. These molecules, produced as byproducts during the mitochondrial electron transport of aerobic respiration or by oxidoreductase enzymes and metal-catalyzed oxidation, have the potential to

cause a number of deleterious events. In addition, daily activities such as fatigue and smoking, environmental factors such as smoking, exposure to radiation and polluted chemicals, may contribute to the increase of the oxidative stress. Cigarette smoke contains reactive oxygen-generating substances such as hydroquinone and catechol and induces oxidative DNA-damage *in vitro*. According to several studies, the urinary 8-OHGua levels of the smokers studied here were certainly higher than those of nonsmokers. [64]

Bruskov et al have suggested that 8-oxodG can be formed due to heat-mediated generation of ROS and their reactions with DNA.[65] The OH• radical is shown to attack the C4, C5, and C8 positions of guanine among which the C8 position is the most favored reaction site. In addition, $H_2O_2$ that is void of reactivity with any of the nucleotides can produce the highly reactive OH• radical and an OH- anion through the Fenton and Haber-Weiss mechanisms that involve metal cations.[66, 67] According to a computational study of the 8-oxoGua formation mechanism, the addition reaction of an OH• radical at the C8 position of guanine is found to form 8-oxoGua, which a second OH• radical at the C8 position of guanine leading to the formation of 8-oxoGua complexed with a water molecule by two different pathways. According to one mechanism, the 8-OHGua complexed with a water molecule followed by tautomerization to 8-oxoGua, while in the other mechanism the five-membered ring open and close to form 8-oxoGua complexed with a water molecule.

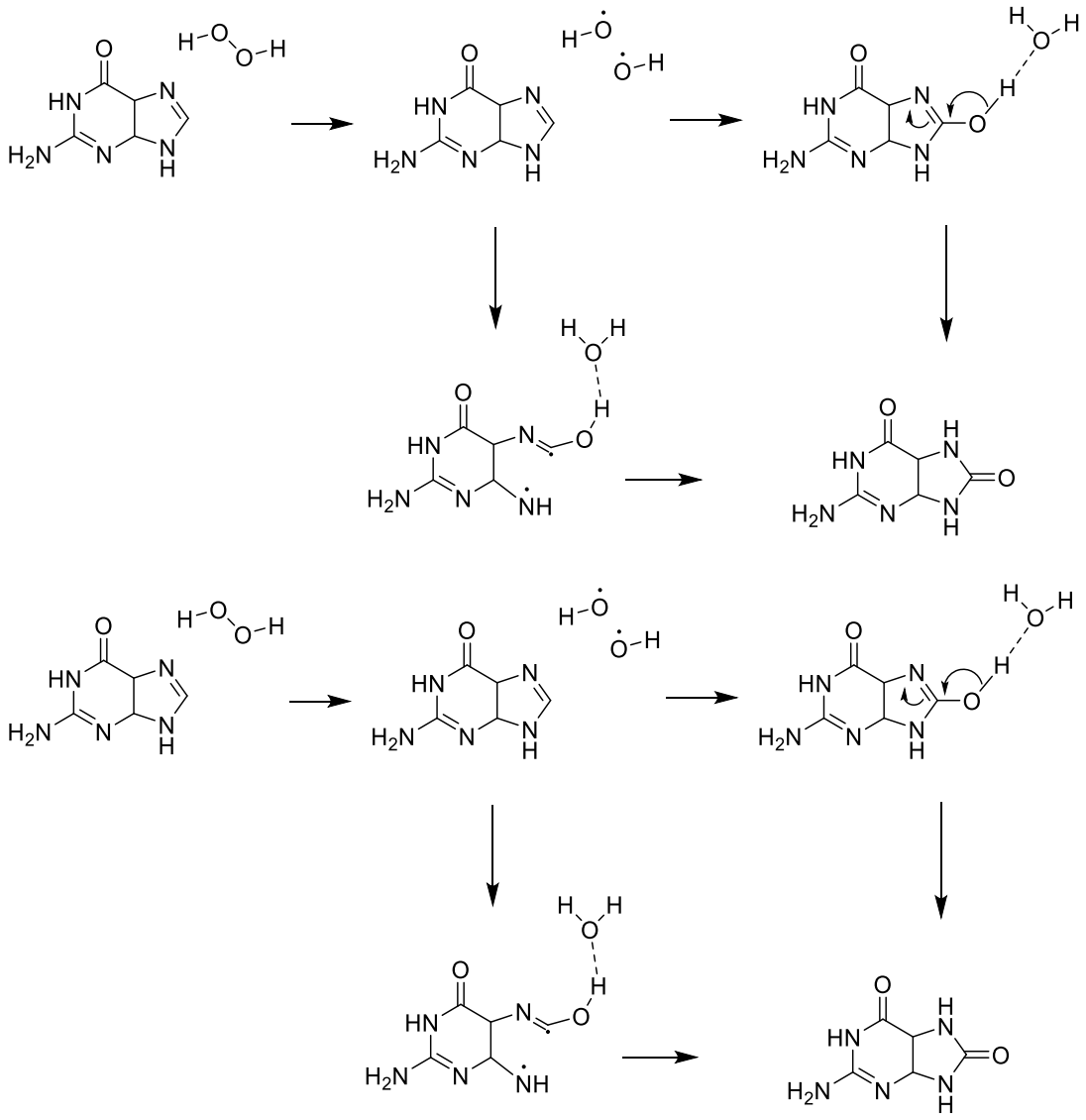


**Figure 3.1** Mechanisms of 8-oxoG formation.

Direct oxidation of guanine in DNA can generate 8-oxodG, meanwhile, the oxidation of dGTP in the nucleotide pool can form 8-oxodGTP, which can be incorporated into DNA during replication. 8-oxodG is mainly repaired by oxoguanine DNA glycosylase 1 (OGG1), releasing 8-oxoGua for urinary excretion. Nei endonuclease VIII-like 1 (NEIL1) in *E. Coli*, DNA translesion synthesis enzymes, and nucleotide excision repair along with the putative nucleotide incision repair all act as backup systems assisting OGG1. The mutY homolog (MUTYH) in *E. Coli* remove the adenine mispaired with 8-oxoGua to reduce the risk of

mutations. Nudix (nucleoside diphosphate linked moiety X) - type motif 1 (NUDT1) hydrolyze the 8-oxodGTP to di or mono- nucleotide to prevent the incorporation during the replication. Obtułowicz et al found positive correlations between 8-oxoGua excretion and OGG1 mRNA expression according to the fact that mRNA levels of OGG1 increased in colorectal cancer (CRC) patients, which could explain the increased 8-oxoGua excision rate in CRC patients and the upregulated OGG1 mRNA expression by oxidative stress.[68]

### 3.1.2 Aging and Diseases Related to 8-OxoGua Excretion

The association between aging and 8-oxoGua excretion is of interest since it is well believed that one of the major contributors of aging is DNA oxidative damage. Upon the studies about effects of aging on the level of oxidatively damaged DNA in various organs, most investigated organs such as brain, heart, kidney, lung, and liver are more susceptible to aging-associated oxidative damage to DNA, while the organs like intestine, spleen, and testis seem to be less susceptible. Although in one study the positive association between age and 8-oxoGua excretion is reported,[69] in general it seems that aging has little effect on the 8-oxoGua excretion based on results of other large studies.[70, 71]

The level of urinary 8-oxoGua is addressed as a possible factor of cancer or pre-cancerous diseases. Loft *et al* using nested-in-cohort case-control model to investigate the association between urinary 8-oxoGua and lung cancer, and it is of importance that they found association between risk of lung cancer and 8-oxoGua excretion among men, former smokers and nonsmokers.[72] According to reported results, urinary and leukocytic 8-oxoGua levels in patients with colorectal tumor were significantly higher than healthy controls.[68] Similarly,

patients with head and neck cancer excrete more 8-oxoGua compared with healthy controls. Elevated levels of 8-oxoGua both in the cerebrospinal fluid and the urine were also found in patients with mixed Alzheimer disease/vascular dementia.[73] In contrast, a study also shows that no significant difference in urinary 8-oxoGua was found among patients with lung cancers and healthy controls.[74] Under certain circumstances, elevated 8-oxoGua excretion is the consequence of the disease other than the reflection of oxidative stress. This is supported by the results that only the untreated patients with celiac cancer have elevated urinary 8-oxoGua, while the treated patients with limited diseases have a similar level of urinary 8-oxoGua compared with healthy controls.[75]

### 3.1.3 Urinary Excretion of 8-OxoGua as a Biomarker

Validating 8-oxoGua as a biomarker has been advanced for a long time. Even though high pH and heating used to compromise the poor solubility of 8-oxoGua can somehow sabotage the long-term storage of 8-oxoGua, it is also reported that the similar levels of 8-oxoGua in urine samples are found when stored under -130°C for 9-13 years.

To clarify the cellular resource of urinary 8-oxoGua, Rozalski et al found a 26% reduction in the urinary level of 8-oxo-7,8-dihydroguanine in OGG1 deficient mice in comparison with the wild-type strain.[76] This small difference could be because that the 8-oxodG in DNA is higher in OGG1 deficient mice due to the accumulation of the oxidative damage, and assume first order kinetics, less backup repair systems are required for this excretion. Depending on this result, it should be emphasized that assuming the steady state kinetics, urinary 8-oxoGua excretion is related to the rate of ongoing oxidative damage

regardless of repair capacity. Thus, correlations between urinary excretion and cellular level of 8-oxoGua would be of interest in steady state and the association between these two has been confirmed regarding different diseases such as colorectal cancer, lung cancer, celiac cancer and so on. In contrast, research on mixed dementia found elevated urinary 8-oxoGua level without increasing leukocytic 8-oxoGua level.

It has also been confirmed that both the urinary 8-oxoGua and 8-oxodG can reflect the oxidative stress by a number of studies. [69, 71, 77] Although large variation was found between these two biomarkers in different species, the augment of these two from the same samples will be taken into accounted. Frequently, spot urine samples corrected for creatinine have been used for the analysis of 8-oxodG. However, the correlation between the 8-oxodG to creatinine ratio in spot samples and the 24 h excretion of 8-oxodG has been shown to be rather poor.[78] Creatinine correction is commonly used and usually found no problems for repeated measurements, but creatinine excretion varies significantly by muscle mass which is related to age, sex and diseases.

As a well-accepted biomarker, 8-oxoGua excretion is still in its early development. Problems such as analytical validation and dietary sources still need to be addressed. Besides, it is impossible to relate the urinary excretion of 8-oxoGua to DNA damage in a particular organ.

### 3.1.4 General Methods for the Measurement of 8-OxoGua or 8-OxodG

In 1986, a method of detecting 8-oxodG using HPLC-EC was introduced by Floyd *et al*.[79] The excellent sensitivity allows this method to be widely applied in the last decade with various modifications. In general, the DNA was hydrolyzed and separated on C18 columns followed

by EC detection. The inefficient release of 8-oxodG by the hydrolytic enzymes is thought to be the major reason of the underestimation of 8-oxodG in DNA.[80] In addition, small quantities of DNA sample may greatly increase the ratio of 8-oxodG/dG due to the trace metals in the buffers or enzymes, and as a result hydrolysis of >100µg of DNA samples would be preferable to allow a greater margin of error. The detection of urinary 8-oxodG was found to be difficult until a two-step solid-phase extraction method was developed, and the detection limit was as low as 0.9 nM in urinary 8-oxoG measurement.

Gas chromatography with mass spectrometry (GC-MS) is a highly specific, sensitive and versatile technique for the quantitative analysis of individual products of oxidized DNA bases including 8-oxoguanine. However, an artifactal oxidation of guanine during the derivatization reaction including the silylation reaction prior to the GC-MS was demonstrated.[81] One of the solutions to this problem is to purify 8-oxoGua by immunoaffinity chromatography prior to the silylation reaction or derivatization by HPLC.

$^{32}$P-postlabeling assays based on TLC have been developed to selectively and sensitively detect the oxidation DNA damage including 8-oxodG.[82] The deoxyribonucleoside-3'-monophosphates was prepared by enzymetic digestion, followed by 5′ phosphorylation using λ-$^{32}$P ATP. Radiolabelled digests was separated by 2-directional PEI- cellulose TLC. This method has high sensitivity (<1 8-oxodG per $10^{5-6}$ nucleotides) when using nanogram quantity of DNA.[83] To reduce the risk of artifactual oxidation of dG to 8-oxodG caused by the ionizing radiation, an improvement in which the unmodified nucleotides was removed prior to the post-labelling was developed. The enrichment of 8-oxodG allow using the larger quantities of DNA

and higher specific activity [$\gamma$-$^{32}$P]ATP, and as a result, the sensitivity was increased to <1 8-oxodG per $10^7$ nucleotides using microgram quantity of DNA.

Compared to HPLC which is relatively complicated and inconvenient to use in the clinical laboratories, enzyme-linked immunosorbent assay (ELISA) are easier to perform. To compare these two methods to determine whether the ELISA method is applicable to detect 8-oxodG, Kayoko et al found a good correlation between them, but ELISA estimates were about 2-fold higher than the HPLC.[84] This was explained by a possible nonspecific binding of materials prior to the assay, or crossreactivity with other modified bases present in the samples, or the lack of selectivity between urinary 8-oxodG and 8-oxoGua.

8-OxodG and 8-oxoGua in human DNA are released by DNA repair pathways and the modified nucleotides or nucleobases are excreted into urine. Unlike determination of 8-oxodG or 8-oxoGua in leukocyte DNA, the measurement of urinary 8-oxodG has some advantages: (1) it is a non-invasive method (2) it can minimize the artifactual oxidation product during sample procedure or derivatization, (3) 8-oxodG shows high stability in urine (4) the urinary 8-oxodG is thought to present the whole body oxidative damage. Urinary 8-oxodG is preferentially analyzed by HPLC-EC and GC-MS. To improve the accuracy of the measurement of urinary 8-oxodG, separation of the analyte from the urinary matrix containing complex materials has become a major challenge. Herein different procedure has been described for HPLC techniques such as solid phase extraction and coupled-column HPLC. In addition, tandem mass spectrometry method (MS/MS) has been used with HPLC. Without any sample pre-treatment, HPLC-MS/MS is able to analyse all types of 8-hydroxylated guanine modifications.[85]

### 3.1.5 Aptamers as Small Molecules Recognition Elements and Their Applications

Nucleic acids are well-known for their abilities to store genetic information. Although not as complex as proteins, the nucleic acid can also fold into tertiary structures which are able to perform multiple functions such as catalytic activities and ligand binding. Systematic evolution of ligands by exponential enrichment (SELEX) was first introduced in 1990 and the nucleic acids using in this process are called aptamers. Aptamers are short nucleic acid oligomers which have high affinity and specificity toward their targets through a combination of molecular-shape complimentary, hydrogen bonding and stacking interactions. As a growing need for effective sensing methods, the focus on the molecular recognition probes dramatically promotes the development of aptamers. Compare to the antibodies which are also broadly used as recognition units, aptamers offer several key advantages. First, the aptamer sequence can be easily and economically generated by the automated solid-phase synthesis. This also means specific designs are feasible such as adding a fluorescent beacon at a certain position. In addition, the nucleic acid nature of aptamer makes it more robust against the harsh conditions such as high temperature. Especially, aptamers can be reversibly denatured under conditions that irreversibly denature the antibodies. And toxicity and low immunogenicity of particular antigens do not interfere with the aptamer development.

Small molecules play key roles in many biological processes because of their ability to diffuse across cell membranes. They can have a variety of biological functions, such as cell signaling molecules, drugs in medicine, pesticides in farming, and in many other roles.[86] One major challenge to the small molecule-binding aptamers is that most of the aptamers cannot bind targets with high affinity required for most sensing applications. In general, the larger

target molecular weight resulted in lower $K_d$.[87] Efforts have been done to seek the natural aptamer with high affinity to target such as riboswitch or to study the structure-affinity relationship of the aptamer.[88] Considering the effectiveness of the naturally developed aptamers, there is still great potential in this field.

Till now, plenty of research regarding the engineering of aptamer-based biosensors has been reported. These biosensors were well-constructed by a variety of methodologies, including electrochemical biosensors and optical biosensors. One notable approach to small-molecule biosensing has been the "structure-switching" strategy. Upon binding to the target, the duplex structure of aptamer switches to another tertiary structure. The signal generated by the conformational change can be detected by EC method, or it can be converted to other colorimetric signals such as fluorescence. The switch from the duplex to the aptamer-target complex is a general approach that allows the recognition of the target to generate signals for detection. This method has been applied to the detection of a lot of small molecules such as adenosine or ATP, theophylline, cocaine, histidine and arginine.[89-92] Aptamers can also serve as regulatory elements through the ability to combine with other functional moieties without affecting recognition and bind to their cognate target. By replacing one of the stem-loop sections of the widely studied hammerhead ribozyme with the ATP-binding aptamer, ATP binding was required for activation.[93] This chimeric system can be used as regulatory elements for nucleic acid enzymes (either natural ribozymes or synthetic DNAzymes) or other actuator parts, allowing diversifying the signal generation methods.[94]

### 3.1.6 Gold Nanoparticles in Biological Sensing of Small Molecules

The distinct physical and chemical attributes possessed by gold nanoparticles make them excellent scaffolds for the fabrication of novel chemical and biological sensors. First, AuNPs can be synthesized by a straightforward approach. Second, they possess unique optical and electronic properties. Third, excellent biocompatibility can be demonstrated by the high surface-to-volume ratio provided by AuNPs. Fourth, all of these properties of AuNPs are tunable by varying their size, shape and the surrounding chemical environment. Notably, AuNPs provide a novel platform for multi-functionalization with a wide range of organic or biological ligands for the detection of small molecules and biological targets and allow researchers to develop advanced sensing strategies with improved sensitivity, stability and selectivity.

AuNPs feature a surface plasmon resonance (SPR). SPR is the results of collective resonant oscillation of the free electrons of the conduction band of the AuNPs excited by incoming photons. The nature of this surface plasmon resonance (SPR) was elucidated by Mie in 1908.[95] The SPR dynamics are influenced by size, solvent, ligand, and temperature. The spectral shift induced by solvent refractive index changes was confirmed by Murray and this result was consistent with Mie theory.[96] The dependence of optical thickness on the refractive index altered by solvent and ligand can be used for impurities detection due to different refractive indexes of gold oxide and gold chloride.[97] Notably, the SPR frequency is dependent on the proximity of nanoparticles. In another word, the aggregation of nanoparticles results in significant red-shifting (from ~520 nm to ~650 nm), changing the solution color from red to blue. This attractive phenomenon has made AuNPs a promising candidate for colorimetric sensors.

DNA-based AuNP assembly was developed in 1996.[98] Fabrication of AuNPs conjugated with thiolated DNA strand allowed the self-aggregation of nanoparticles to be rational and reversible. By tailoring the properties of the AuNPs, researchers have developed many oligonucleotide-mediated AuNP aggregations for colorimetric detection methods.[99-101] In one of these approaches, unfolded and folded DNA could be discriminated by unmodified DNA. Before binding to the target, the coil-like aptamers do not protect the AuNPs against self-aggregation. Once binding to the target, the aptamer form 3D structure and can no longer stablize the AuNPs against salt-induced aggregation. Since the color change is visible to naked eye, this technology has been used for the colorimetric detection of potassium ion and proteins.[102, 103] In another approach, two ssDNA-modified AuNP probes were designed to be complementary to both ends of the target oligonucleotide. The color change due to the AuNP aggregation as well as the sharp transition of UV-vis spectroscopy demonstrated that nanoparticle probes can be used for the homogeneous colorimetric detection of oligonucleotides as the alteration of expensive fluorescence-based methods.[101]

### 3.2 Need for the Reliable Sensors of 8-OxoGua

Single-stranded DNA aptamer sensor provides a direct approach to quantify the concentration of biologically important ligand, but with limited binding affinity and selectivity. For example, the dissociation constant of adenosine aptamer/adenine complex is 6 μM as reported, and the aptamers do not distinguish adenosine from AMP or ATP.[104] DNA-duplex or DNA-triplex

aptamer with an abasic site has been reported to be a good system for the selective and strong binding where an abasic (AP) site is utilized as an active cavity for binding events. It is well known that AP sites are formed by spontaneous depurination or by the DNA glycosylase during the base excision repair (BER). By intentionally incorporate such lesion sites into the duplex or triplex, a normal DNA is allowed to complementarily hybridize with an AP-site-containing DNA. The high stability of the AP-site-based aptamer is achieved by a combination of hydrogen bonding with target molecules and stacking interaction with nucleobases flanking the AP site. More importantly, the charge repulsion between the phosphate groups of the aptamer in the binding site ensures the selectivity of nucleoside over its monophosphate or triphosphate derivative.[105]

Duplex DNA aptamer of 8-oxoG was previously designed.[106] Although the great fluorescence quenching ability of 8-oxoG resulted in a highly selective method to detect free 8-oxoG against other tested nucleobases, the binding affinity was only moderate, which limited its application in the low concentration range. Since almost every existing duplex DNA aptamer showed only moderate binding affinity to its respective target, it is unlikely that a different design based on a duplex scaffold will significant promote the binding affinity[107, 108] To build a sensor that can provide a sensitive and reliable tool to quantify nucleoside and nucleobase ligands, a triplex DNA aptamer with Hoogsteen hydrogen bonding site was designed. Based on the theory that the stable triplex DNA containing 8-oxodG would be able to convert into a stable triplex aptamer-8-oxoGua complex, a DNA triplex that contains a center AP site was designed to serve as the receptor for free 8-oxoGua. By incorporating Pyrrolo-C as the signal reporter, significant fluorescence quenching was detected by the addition of 8-

oxoGua in the triplex containing A·OG-C triad or a C·OG-A triad. The combined detection range was 3 nM - 1 μM and remarkable selectivity was achieved. However, the fluorescence sensors overestimate 8-oxoGua concentrations by 1.5-2.0 folds in presence of high concentration of guanine which is typical in urine. Moreover, the fluorescent dyes suffered from photobleaching over a long period of exposure. Hence, a more reliable method which is highly resistant from the interference of guanine is desired.

### 3.3 Gold Nanoparticle Conjugated DNA Triplex Sensor

In 2006, a colorimetric assay by using oligonucleotide-directed AuNP assembly is developed for the screening of triplex DNA binders.[109] This approach contained three components: two sets of AuNPs functionalized with non-complementary single-stand DNA and a free single-strand DNA. With proper design of the sequence, triplex structure formation is possible, however, the low stability of the structure prevent the aggregation of the AuNPs under the room temperature. Only under the presence of appropriate triplex DNA binders which could stabilize the triplex structure, the GNP aggregation would be induced and exhibit significant different color. This approach offers a convenient and high-throughput method for screening triplex binders from large combinatorial libraries. This unique design enlightens us to use the similar strategy to develop a colorimetric and high-throughput method for the detection of urinary 8-oxoGua.

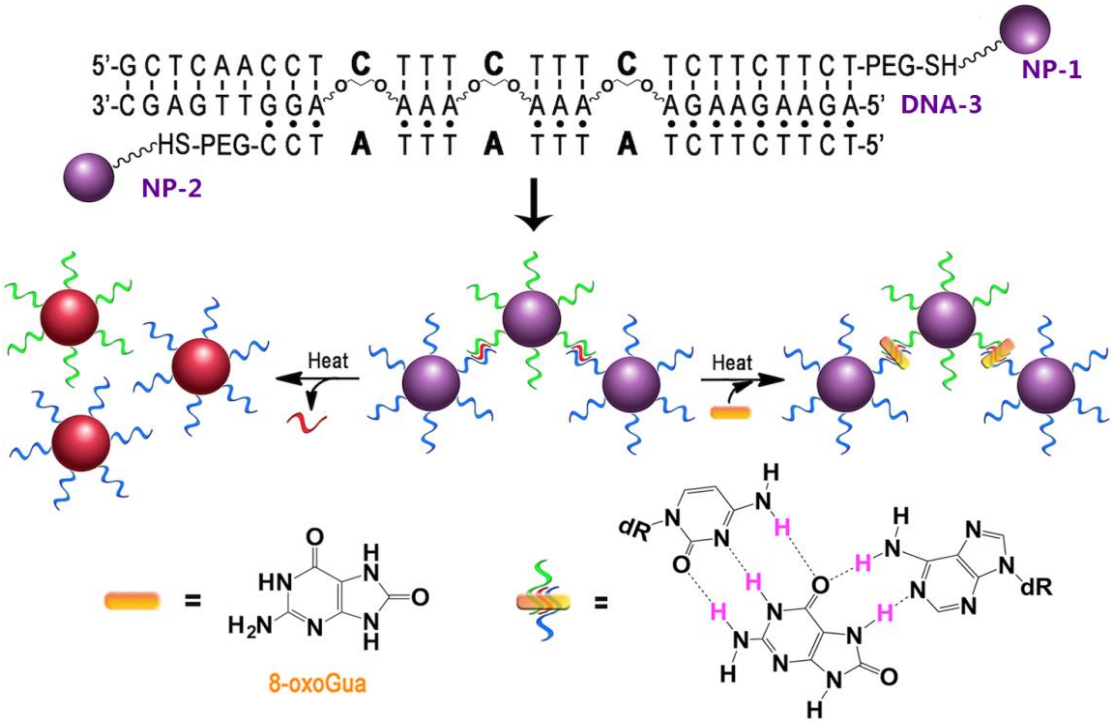



**Figure 3.2** Schematic illustration of DNA-GNP based colorimetric sensor for 8-oxoGua detection.

The assay consists of two sets of gold nanoparticles NP-1 and NP-2 and a free strand of DNA, DNA-3. NP-1 and NP-2 are functionalized with pyrimidine-rich thiol-modified oligonucleotides strand. These two strands are non-complementary and do not interact. DNA-3 which is complementary to NP-1 is rationally designed to amplify the stabilizing effect. According to the research about the structural dependence of the binding effect, first, large binding site would dramatically decrease the stability of the triplex formation, hence single binding sites such as C3 spacers are considered in this assay; second, it is envisioned that three propylene linkers used for binding pockets with three inner purine segments between the binding sites would be able to amplify the stabilizing effect without sacrificing the stability of the triplex. NP-2 DNA has the proper sequence to form a triplex structure once the NP-1/DNA-3 duplex has formed.

The thermal stability of the same triplex DNA in the absence of AuNPs was tested by UV-melting experiment (Figure 3.1). No triplex-to-duplex transition was observed above 20 °C, suggesting that the three C3 spacers separated by two -AAA- segments significantly destabilize the triplex formation. The instability of triplex structure in the absence of the ligand allows the amplification of the stability effect. Unlike the fluorescent sensor which is still sensitive to guanine which is abundant in urine, this well-designed triplex DNA aptamer/AuNPs system will not be stabilized by the binding of guanine. In other words, urinary guanine can not affect the melting points of aggregated AuNPs and the resulting purple-to-red color change in the assay.

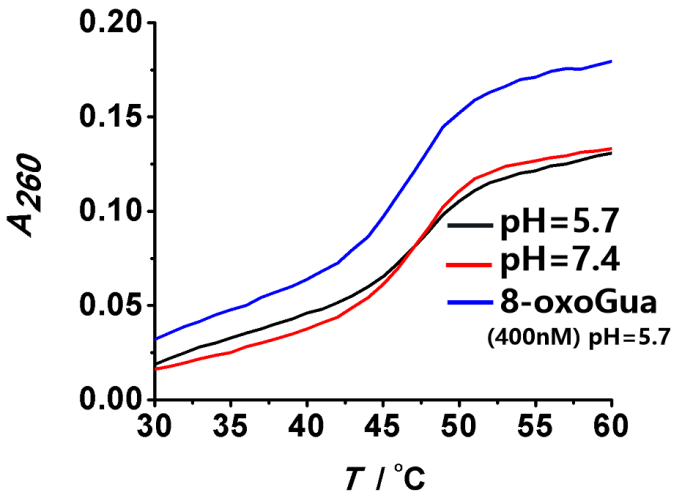



**Figure 3.3** Melting curves of AuNP-free triplex or duplex DNA monitored at 260 nm.

The relationship between the thermal stability of the triplex helix and the concentration of 8-oxoGua was determined through a series of UV-Vis melting experiments monitored at the wavelength of 520 nm. Fixed amount of NP-1, NP-2 and DNA-3 (2 nM) were used in this experiment and different concentration of 8-oxoGua was added to stabilize the triplex. As expected, a sharp melting transition was observed in each experiment due to the cooperative

dissociation of GNP aggregates (Figure 3.2), along with the distinct purple-to-red color change. In the absence of 8-oxoGua, the melting temperature was 40 °C. Within the 8-oxoGua concentration range of 400 nM-1600 nM, each time the 8-oxoGua concentration was doubled, the melting temperature was increased by approximately 2 °C (Table 1). Unlike the common melting curve of the DNA (Figure 1), the signature sharp melting transitions by using AuNPs offer relatively wide temperature windows to distinguish different 8-oxoGua concentration in the samples.

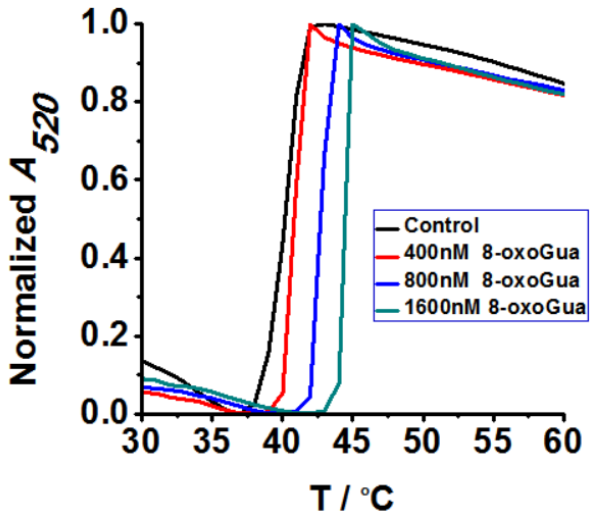



**Figure 3.4** Melting curves of GNP aggregates in the presence of different concentrations of 8-oxoGua.

**Table 3.1** Melting Temperatures in Different Conditions

| Entry | [AuNP-DNA] (nM) | [8-oxoGua] (nM) | Tm (°C) |
|---|---|---|---|
| 1 | 2 | 0 | 40.0 |
| 2 | 2 | 400 | 42.0 |
| 3 | 2 | 800 | 43.9 |
| 4 | 2 | 1600 | 46.0 |
| 5 | 1 | 0 | 40.1 |
| 6 | 1 | 400 | 44.0 |
| 7 | 1 | 0 | 41.0 |
| 8 | 1 | 400 | 43.0 |

Benefit from the two-degree temperature window established from the UV-Vis melting experiment, it is feasible to carry out the in situ colorimetric assays, which could allow the otherwise complicated detection of urinary 8-oxoGua to become much more convenient, economical, and rapid, particularly compared to the traditional chromatographic or enzymatic methods. A set of samples with the same amount of GNP aggregates produced by 2 nM of the AuNP-DNA were incubated with the different compound: 8-oxoGua, adenine, guanine, cytosine, and thymine (400 nM each) at 40 °C for 2 min. As expected, only the sample containing 8-oxoGua survived through the incubation and remain pink color (Figure 3.3), whereas the sample in the absence of 8-oxoGua turned into red. To verify this result, the visible spectrum of the same sample was measured and the significant blue shift was observed when 8-oxoGua was added (Figure 3.4). Notably, the samples in the presence of other potential interfering species which typically exist in urine also dissociated and turned red. This experiment demonstrated that under the certain temperature, this assay could excellently

discriminate between 8-oxoGua and other chosen nucleobases. Hence, this assay can perform higher accuracy in the presence of urinary guanine then the fluorescent sensor which overestimates urinary 8-oxoGua concentration due to the interference of guanine.

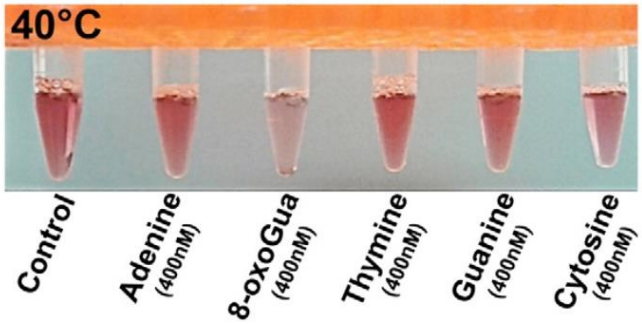



**Figure 3.5** Colorimetric assays of GNP-DNA aggregates after incubation under 40 °C for 2 minutes in presence of different nucleobases.

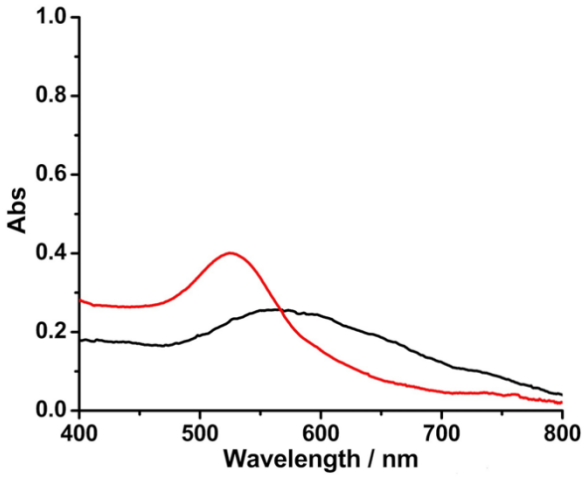



**Figure 3.6** Visible spectra of DNA-modified gold nanoparticle before and after the addition of 8-oxoGua when samples were heated to 40 ºC. Black: control; Red: 400nM 8-oxoGua.Conditions: GNP-DNA conjugates, 2 nM of each, linker, 360 nM, PBS, 10 mM, PIPES, 100 mM, additional NaCl, 100 mM, $MgCl_2$, 10 mM, SDS, 0.1%wt, pH 5.7.

The same colorimetric assay was performed to qualitatively determine the concentration range of 8-oxoGua (Figure 3.5). By two minutes of incubation at 40 °C, the sample in the absence of 8-oxoGua melted and turned red, while all the other samples containing 8-oxoGua

(400-1600 nM) remained pink. When the temperature increased to 42 °C, the sample containing 400 nM 8-oxoGua began to turn into red. At 44 °C, the sample with 800 nM 8-oxoGua also became red, while the only sample that remained pink was in the presence of 1600 nM 8-oxoGua. These colorimetric performances of the GNP aggregates under mild heating were consistent with the above UV – Vis melting results. The sensor may be used for a wide range of different applications since the threshold of detection was temperature-dependent.

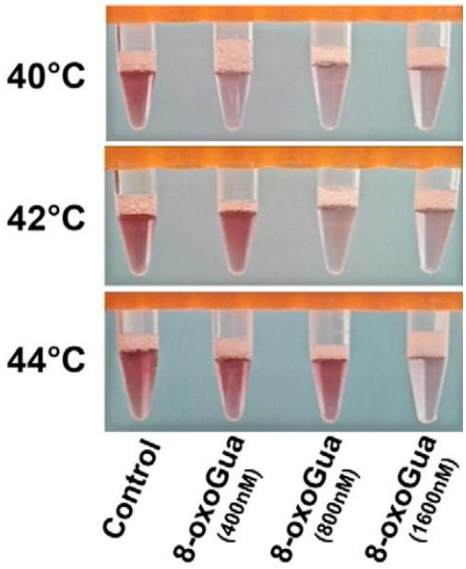



**Figure 3.7** Colorimetric assays of GNP-DNA aggregates after incubation under different temperatures in presence of different concentrations of 8-oxoGua.

The peak of absorbance ratio of two wavelengths ($A_{650}/A_{520}$) recorded on a UV – vis spectrometer indicates the onset of significant GNP aggregation. The double wavelength monitoring greatly increased the sensitivity of the sensor. For instance, although clear aggregations could be observed when a sample containing 200 nM 8-oxoGua was incubated at 40 °C, after agitation, the sample appeared to be no different from the control and thus cannot be directly detected with naked eyes. However, the $A_{650}/A_{520}$ ratio of the 200 nM sample

under stirring conditions was significantly different from the control. Based on the melting curves of GNP in presence of different 8-oxoGua concentrations, at under 41°C, the control sample has been fully melting, while the sample containing 400 nM 8-oxoGua was melting and the sample in presence of 800 nM and 1600 nM 8-oxoGua remain aggregates. By plotting the $A_{650}/A_{520}$ ratio versus 8-oxoGua concentrations at 41°C to generate an 8-oxoGua response curve, the concentration of 8-oxoGua in an unknown sample could be quantitatively determined by this assay (Figure 3.6). The limit of detection using the spectrometer was 128 nM. The detection limit was higher than the fluorescent sensor developed previously, but the detection range still covers the major portion of the biological relevant concentration range.

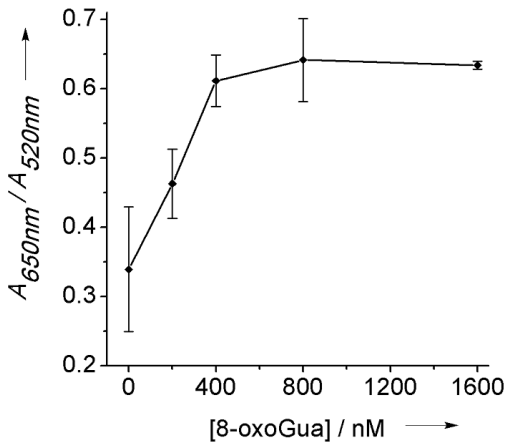



**Figure 3.8** The plot of absorption ratio ($A_{650}/A_{520}$) vs. different concentrations of 8-oxoGua at 41 °C.

A 2 nM suspension of GNP-DNA contained approximately 600 nM binding cavities, but the detection range of this assay can be up to 1600nM. The unique design of the triplex which contains three consecutive cavities allows the binding of 8-oxoGua to be positively cooperative due to the effect of binding-induced preorganization. Therefore, it is possible that 200 nM of 8-oxoGua would only be able to bind one of the three cavities in each receptor with

a weak binding constant, and this speculation is consistent with the high upper limit of the detection range. To further promoting the detection range to a lower limit, the GNP concentration was changed from 2 nM to 1 nM. In this way, the same amount of 8-oxoGua was able to occupy a bigger portion of the cavities, and the triplex helix would be more stabilized. Indeed, the melting temperature difference was changed from 2 °C to 4 °C when the concentration of the GNP was changed from 2 nM to 1 nM (Figure 3.7, table 3.1).

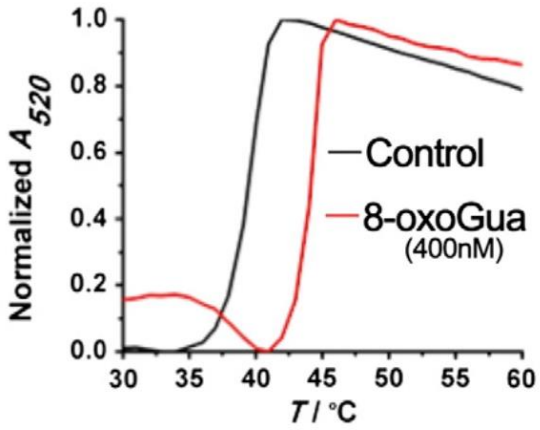



**Figure 3.9** Melting curves of GNP aggregates at using 1 nM GNP concentration.

To demonstrate that this assay can overcome the impact of the highly concentrated urinary nucleobases and thereby can be practically used in the detection of urinary 8-oxoGua, we performed the melting experiment and colorimetric assay with a urine mimic which contains 100 mM urea, 4 μM adenine, 0.4 μM cytosine, 1.2 μM guanine, and 6 μM uracil. In order to increase the sensitivity, 1 nM of the GNP was employed. The sample with the urine mimic and without the 8-oxoGua was melted at 41°C, which is slightly higher than the one without the urine mimic (40°C), suggested that the non-specific binding was caused by the highly concentrated nucleobases. The melting temperature of the sample in the presence of both urine mimic and 400 nM 8-oxoGua increased to 43 °C, indicating that the stability of the triplex-

induced by the 8-oxoGua was not interfered by the urine mimic. The two-degree window also allow the colorimetric assay to show a clear red versus the pink difference between the spiked and unspiked sample, consistent with the results of the melting study (Figure 3.8).

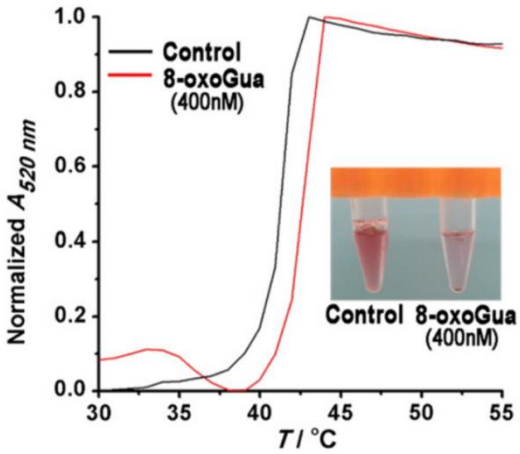



**Figure 3.10** Melting curves of 1 nM GNP-DNA aggregates in urine mimic.

Although a variety of methods have been developed to quantify urinary 8-oxoGua concentrations, the methods that are able to determine the absolute amount of 8-oxoGua in concentration units or nmol/24h are rare. The GNP-DNA aptamer sensor developed by us combine the specificity and tunable detection range by changing the GNP concentrations and incubation time, which allow a protocol for each specific application to be flexible to meet the clinical demand for 8-oxoGua quantification.

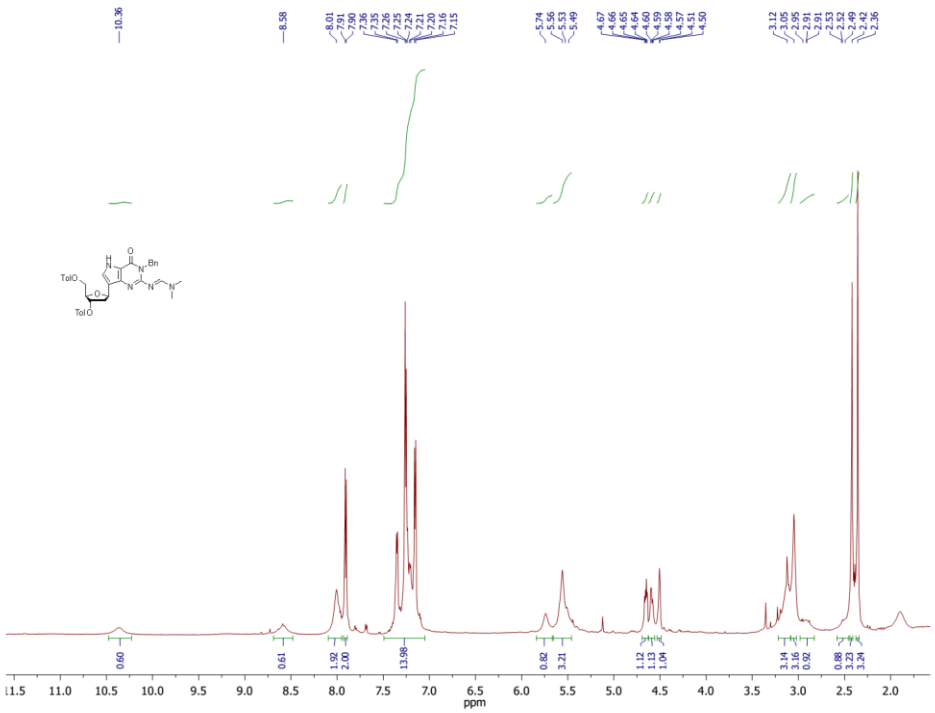
In summary, the conjugation of the two pyrimidine-rich DNA strand to gold nanoparticles promote the detection of the assembling of multi-gapped triplex receptors. The multiple binding sites of the receptor increase the binding-induced stabilization effect and thus widened the temperature window used for the colorimetric detection. The triplex aptamer guarantee the selectivity for 8-oxoGua over guanine, a commonly known interfering species.

For the first time, 8-oxoGua can be directly detected at sub-micromolar concentrations without using a major instrument. This methodology has the potential to be used on the detection of nucleobases and nucleosides in biological fluids.

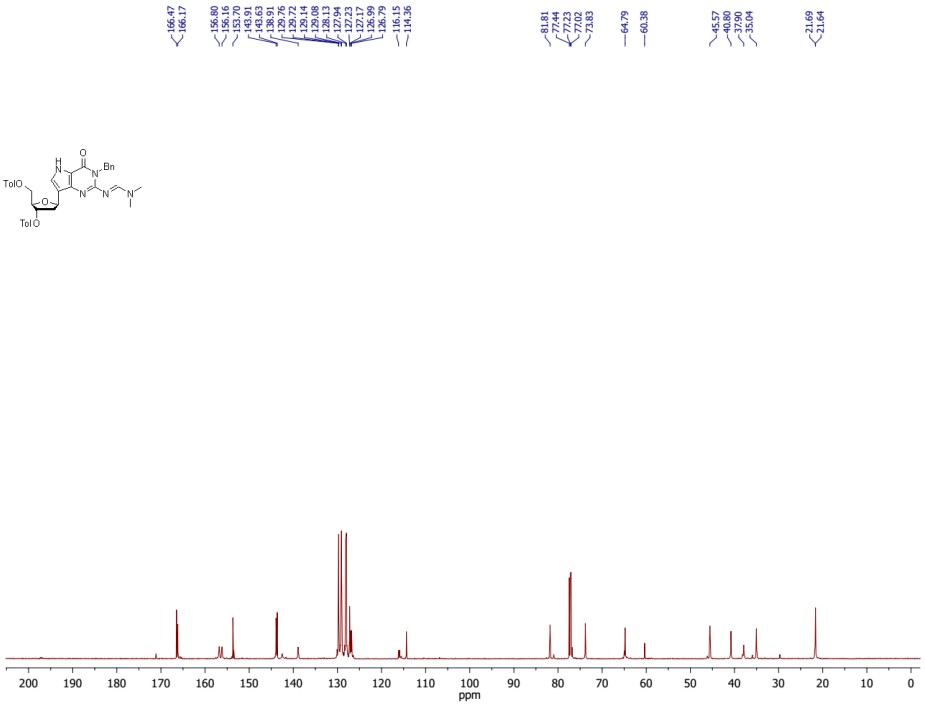
## APPENDIX

## $^{1}$H AND $^{13}$C SPECTRA

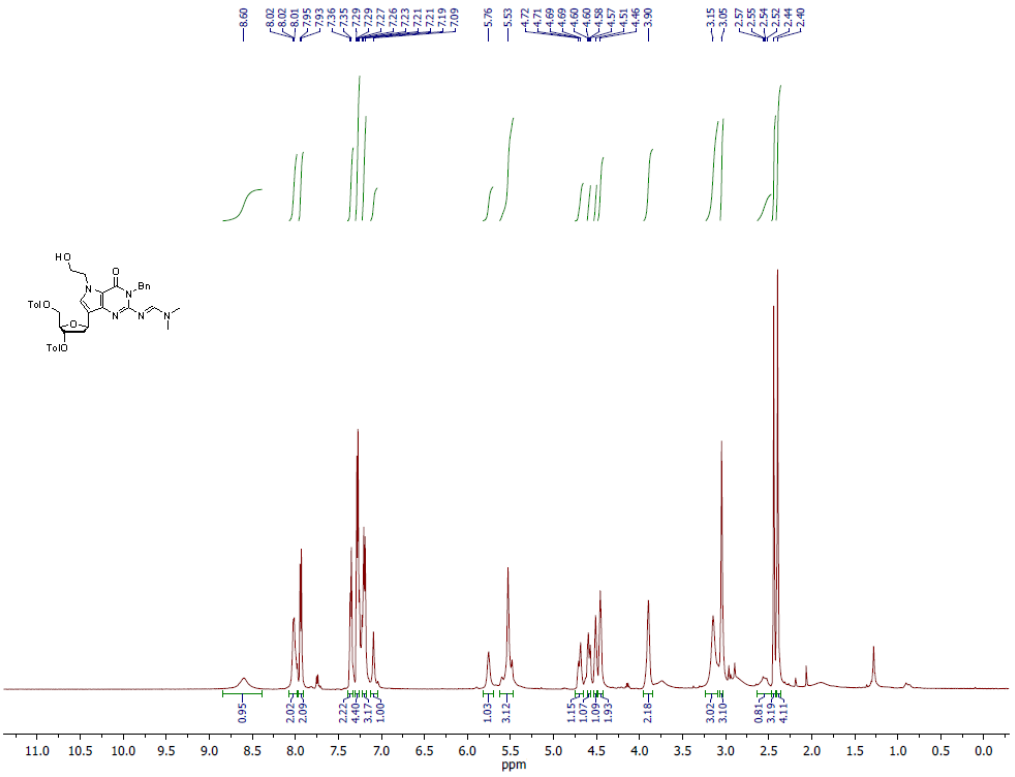
$^{1}$H and $^{13}$C spectra taken from NMR instrument are shown from p86 to p108.



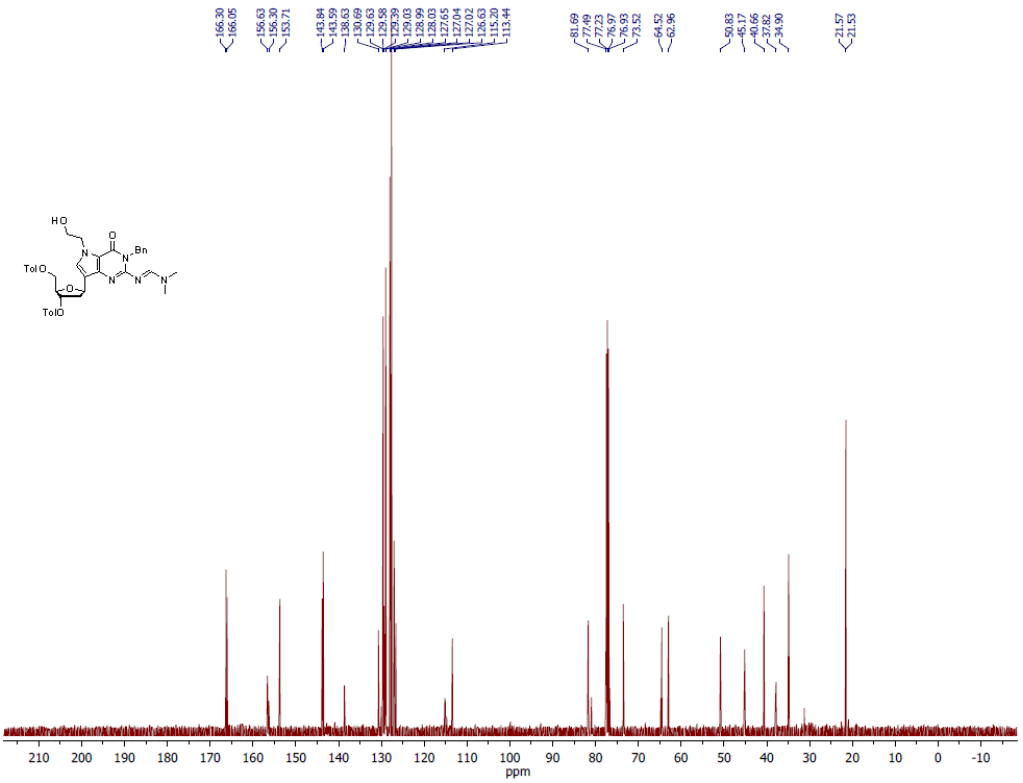
$^{1}$H spectrum of compound **5** in $CDCl_3$.



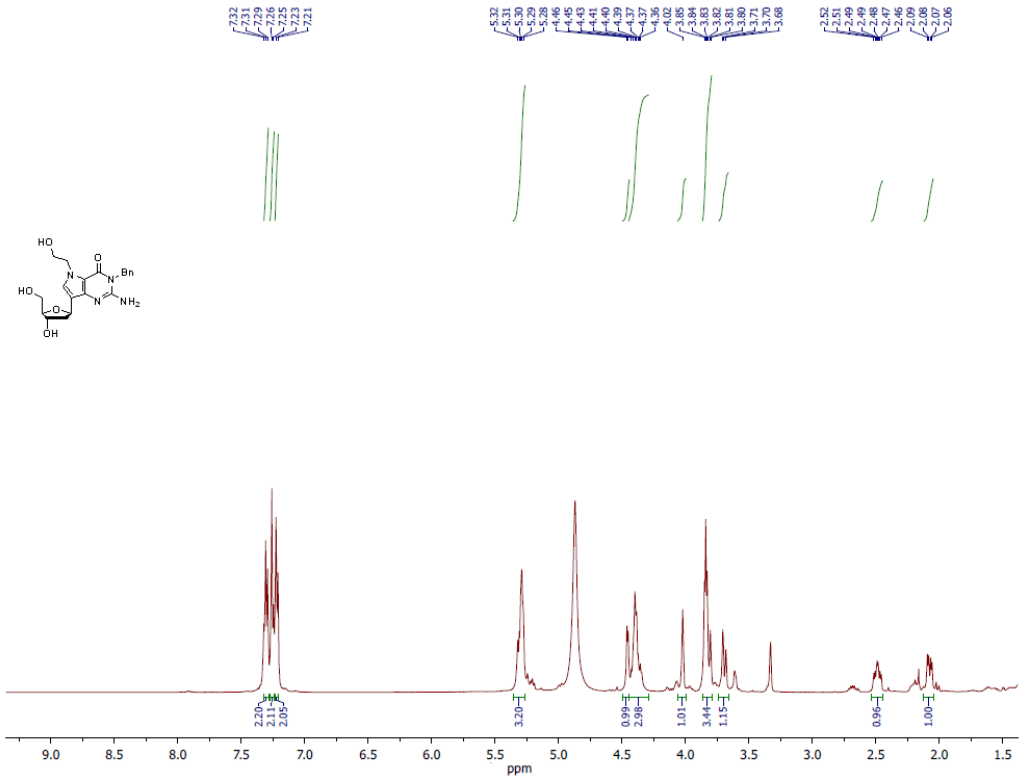
$^{13}$C spectrum of compound **5** in $CDCl_3$.



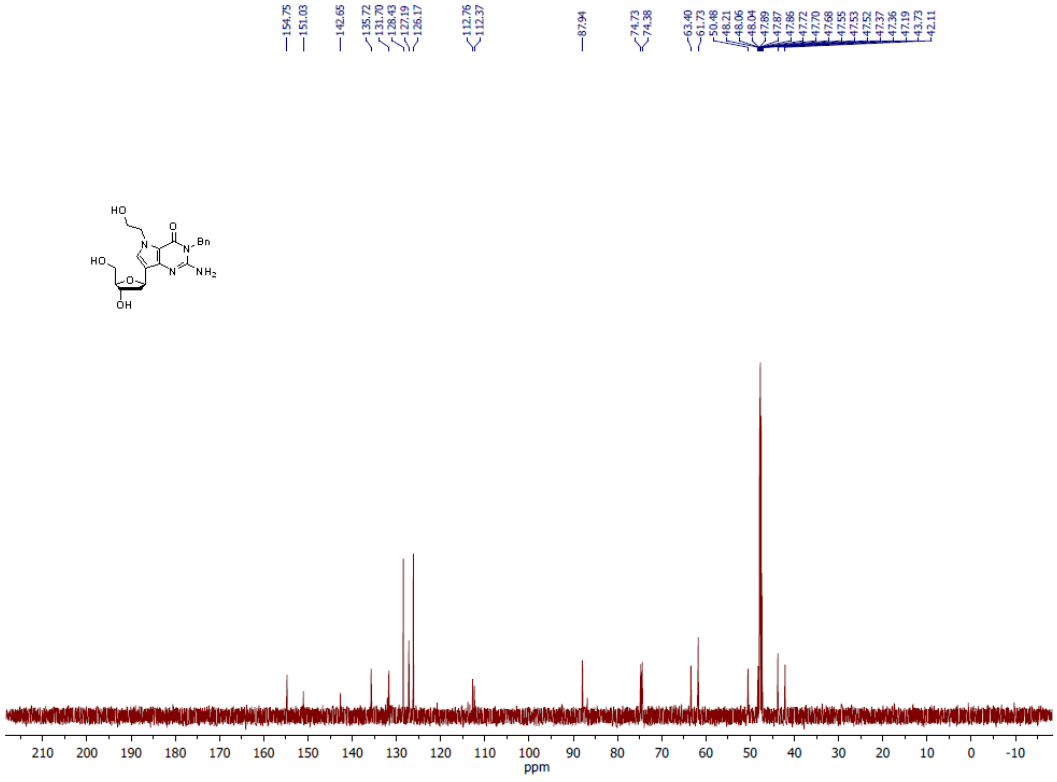
$^{1}$H spectrum of compound **6** in $CDCl_3$.



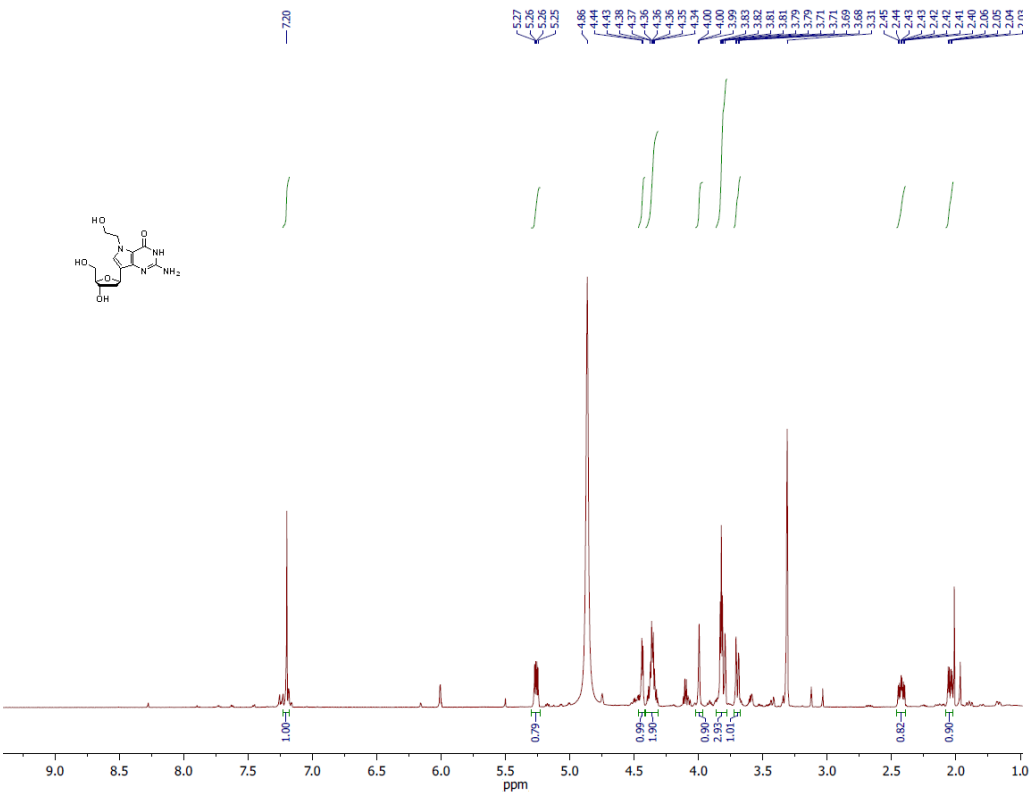
$^{13}$C spectrum of compound **6** in $CDCl_3$.



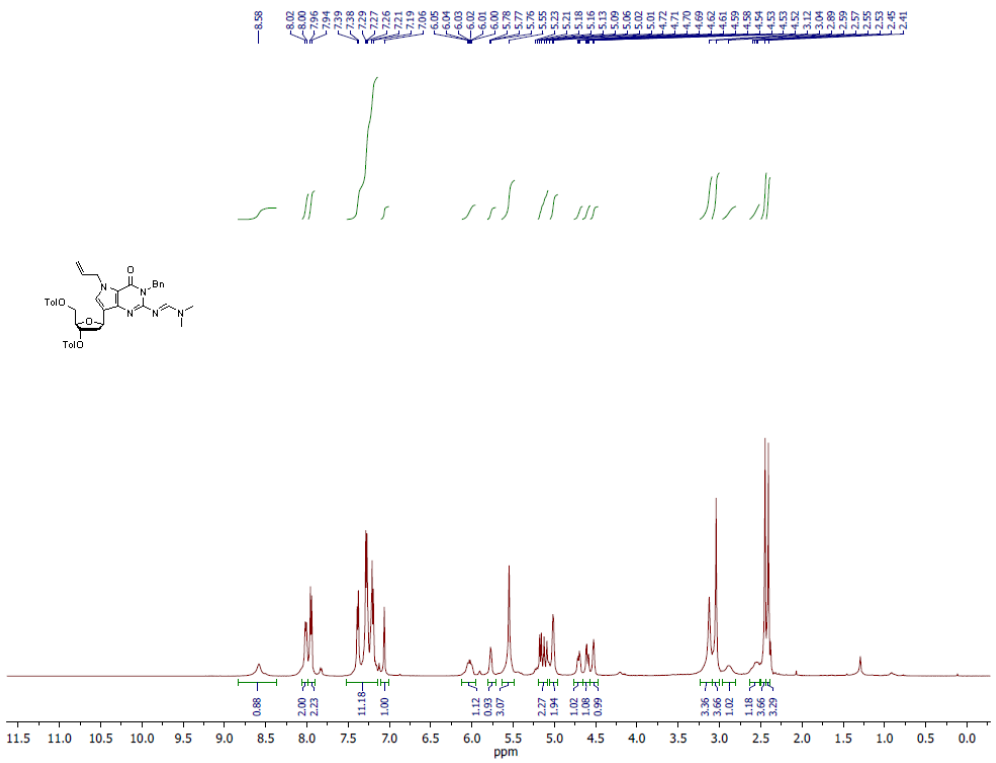
$^{1}$H spectrum of compound **7** in $CD_3OD$.



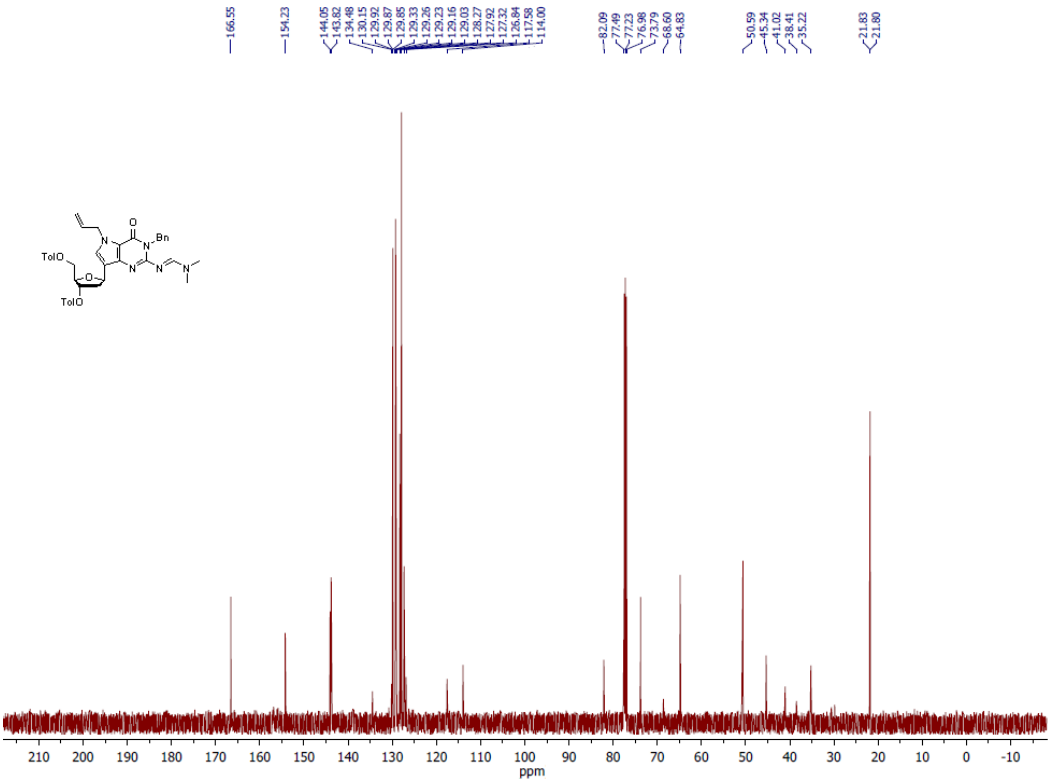
$^{13}$C spectrum of compound **7** in $CD_3OD$.



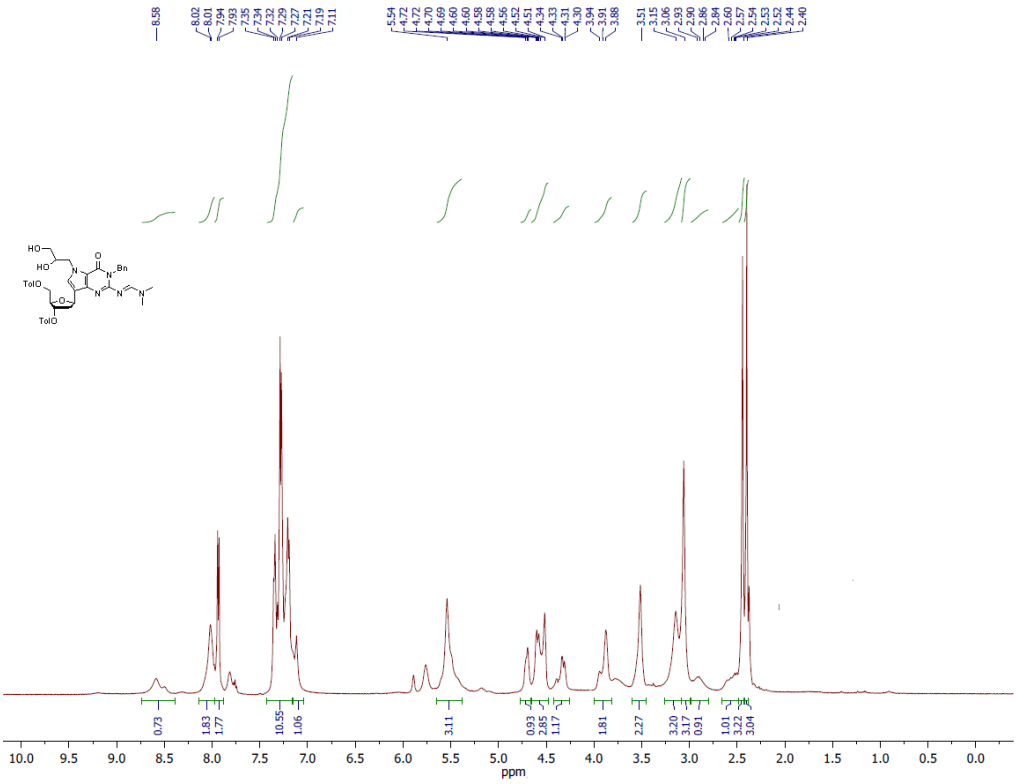

$^{1}$H spectrum of compound **2b** in $CD_3OD$.



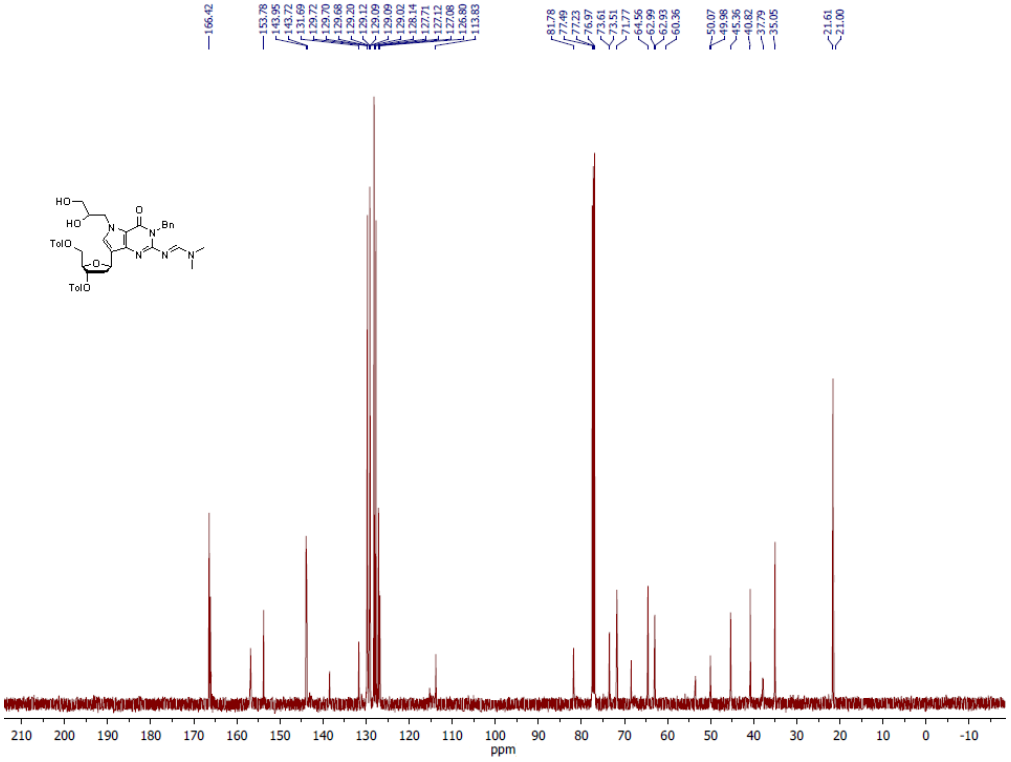
$^{1}$H spectrum of compound **8** in $CDCl_3$.



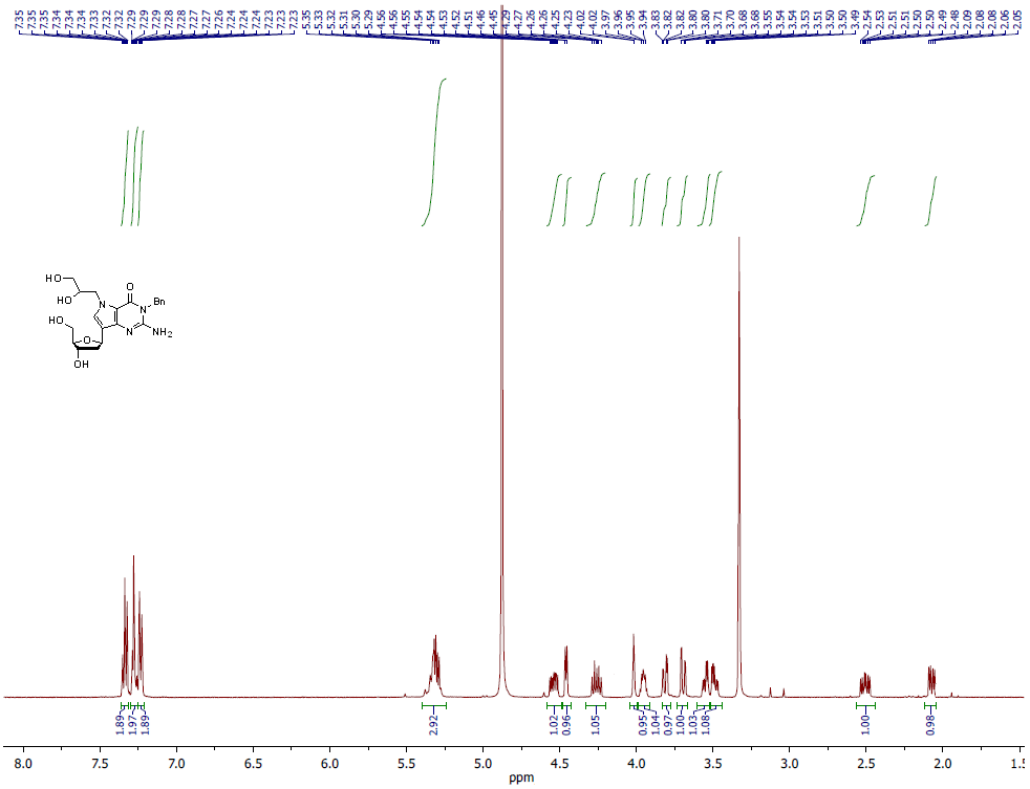
$^{13}$C spectrum of compound **8** in $CDCl_3$.



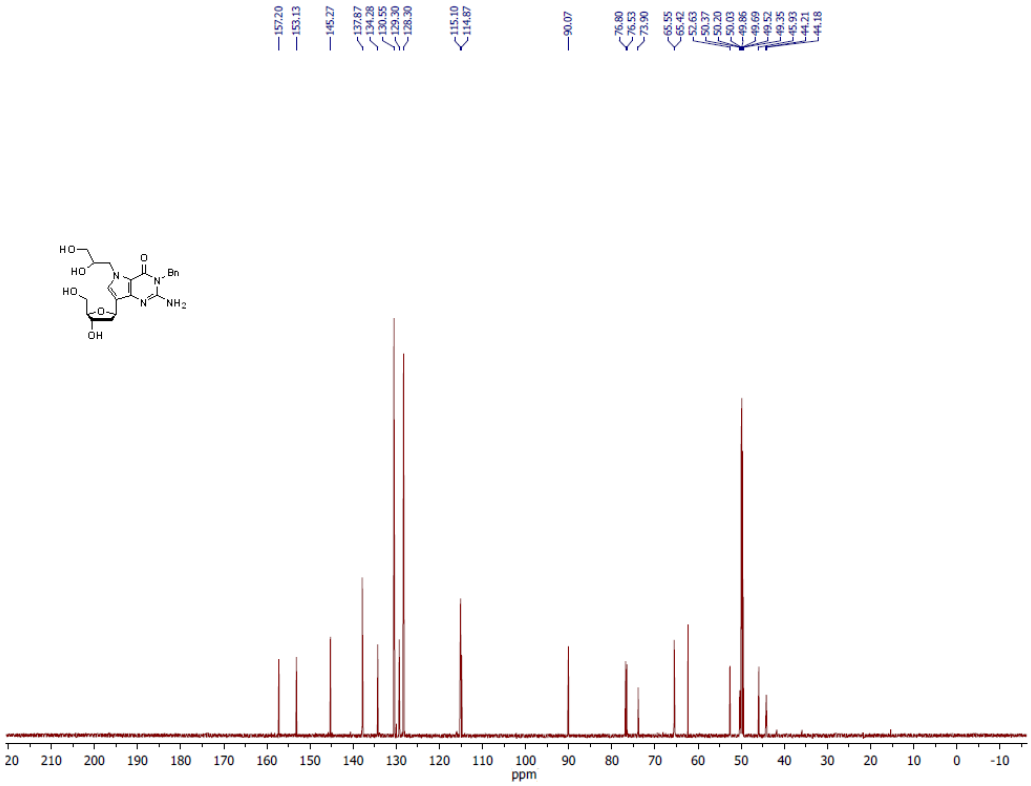
$^{1}$H spectrum of compound **9** in $CDCl_3$.



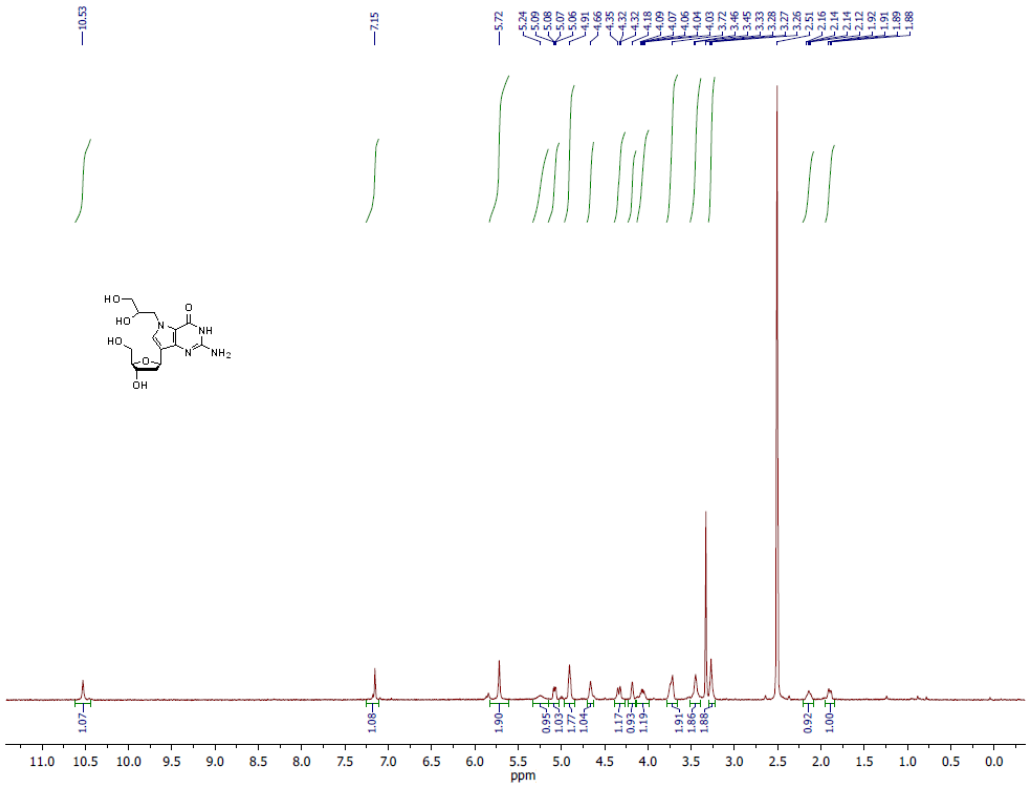
$^{13}$C spectrum of compound **9** in $CDCl_3$.



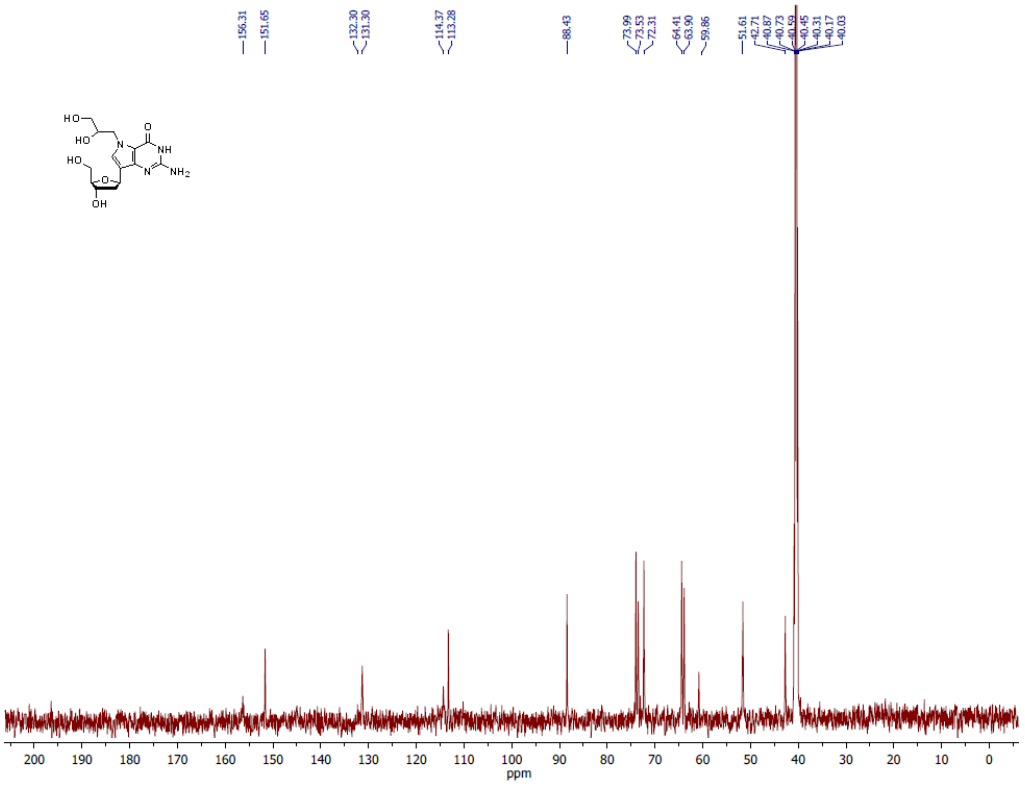
$^{1}$H spectrum of compound **10** in $CD_3OD$.



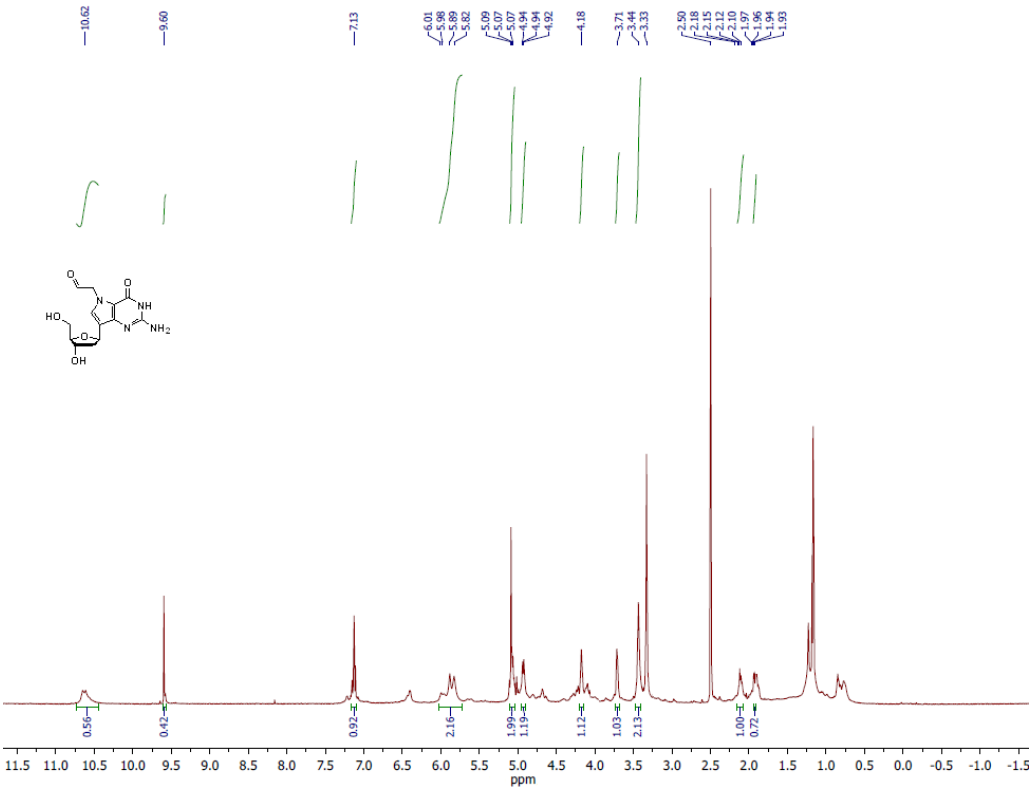
$^{13}$C spectrum of compound **10** in $CD_3OD$.



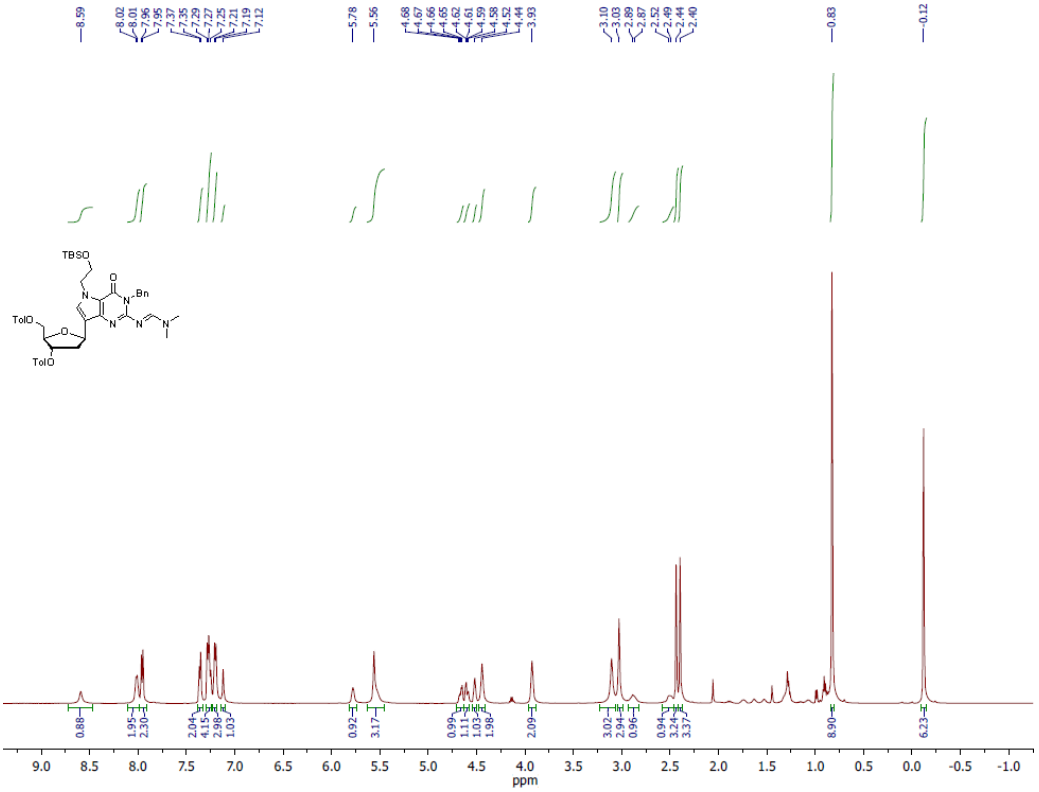
$^{1}$H spectrum of compound **2d** in $d_6$-DMSO.



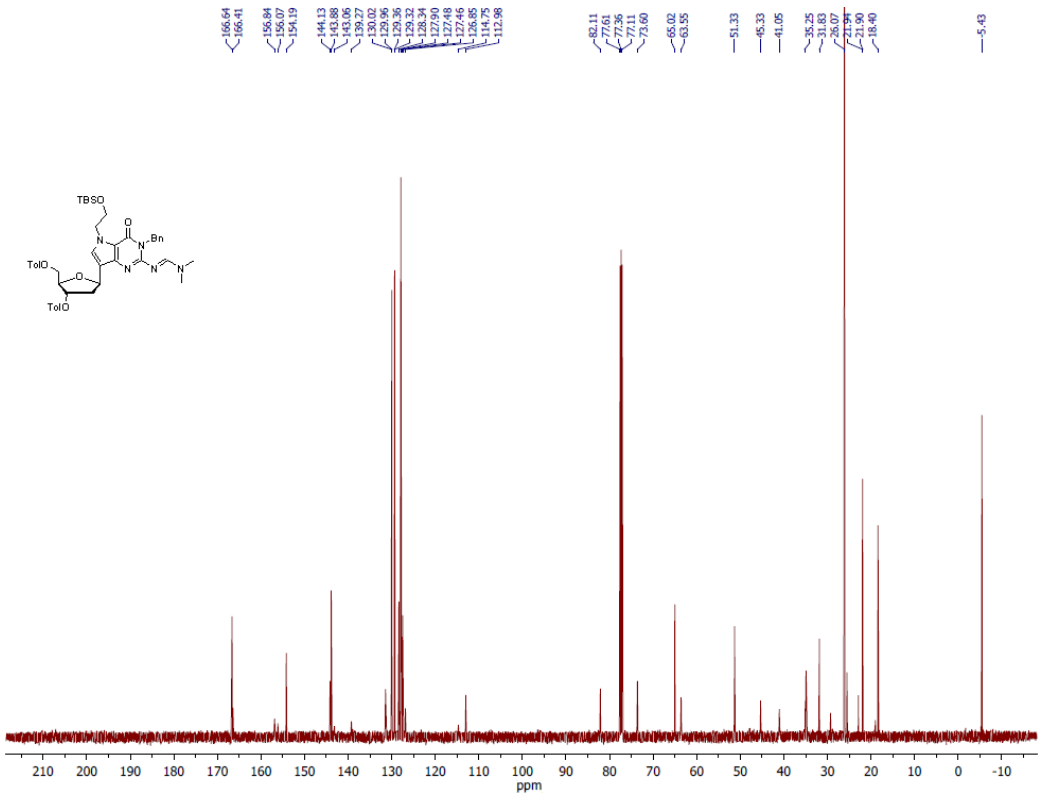
$^{13}$C spectrum of compound **2d** in $d_6$-DMSO.



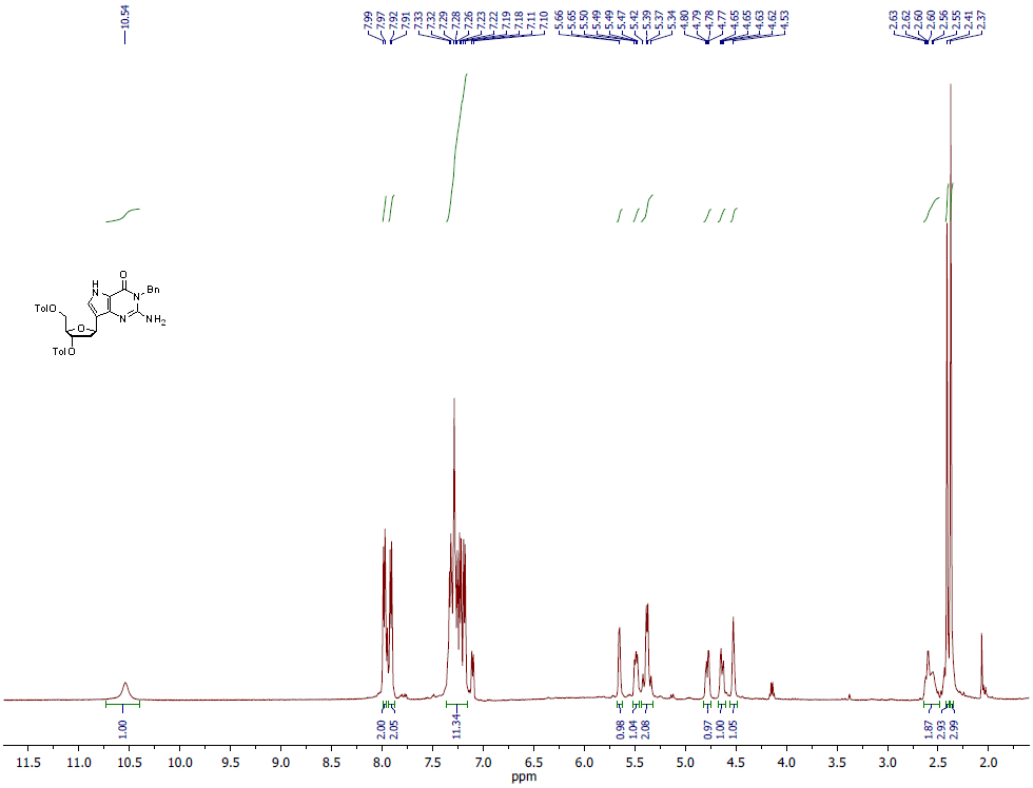

$^{1}$H spectrum of compound **2c** in $d_6$-DMSO.



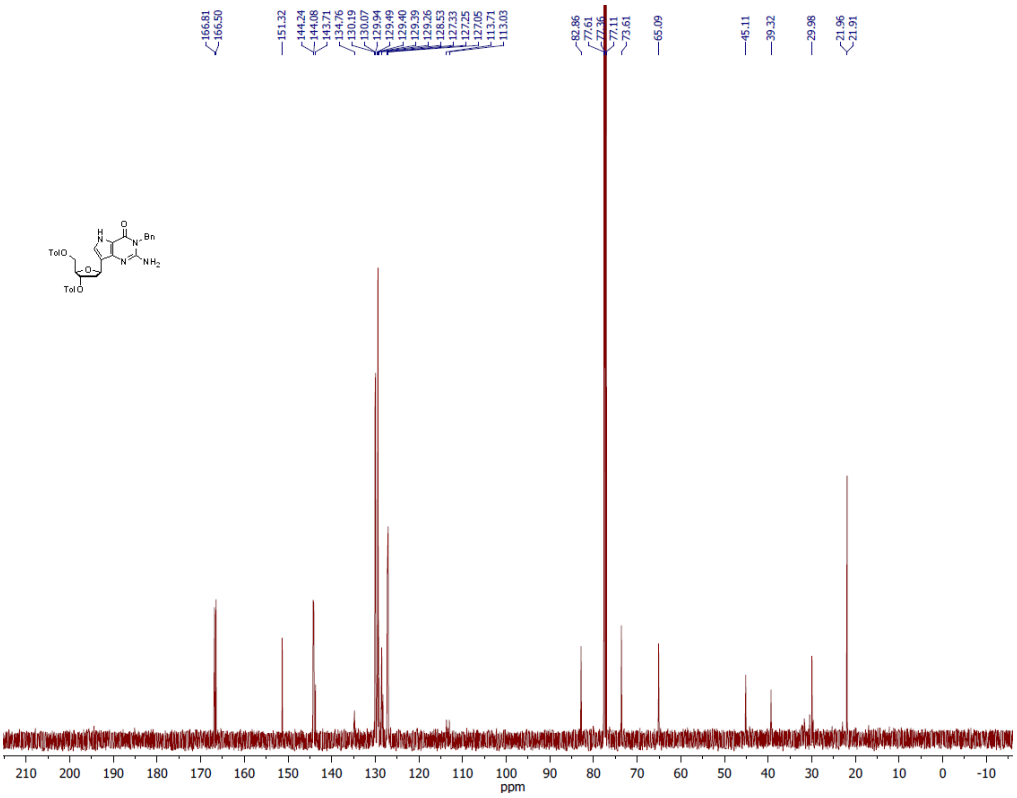

$^{1}$H spectrum of compound **13** in $CDCl_3$.



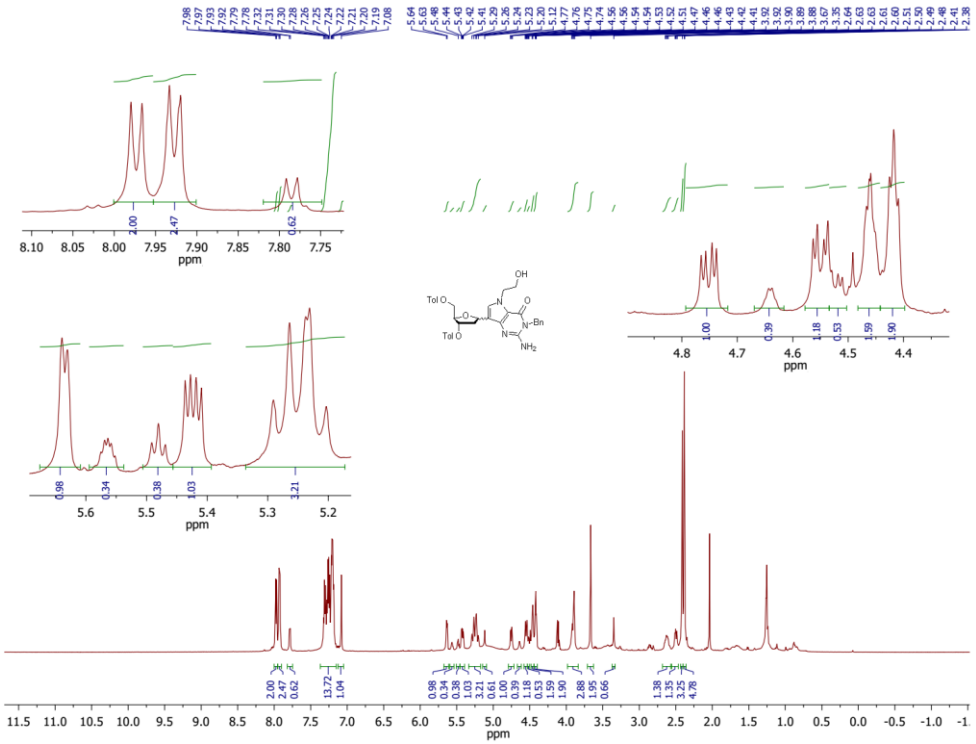

$^{13}$C spectrum of compound **13** in $CDCl_3$.



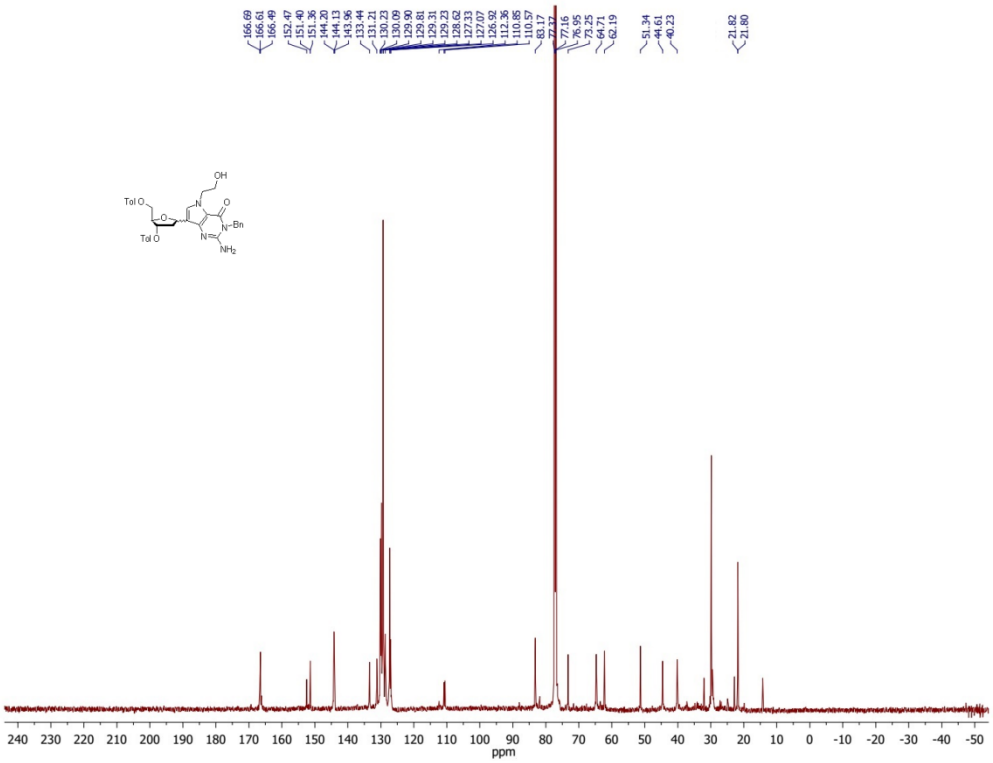
$^{1}$H spectrum of compound **15** in $CDCl_3$.



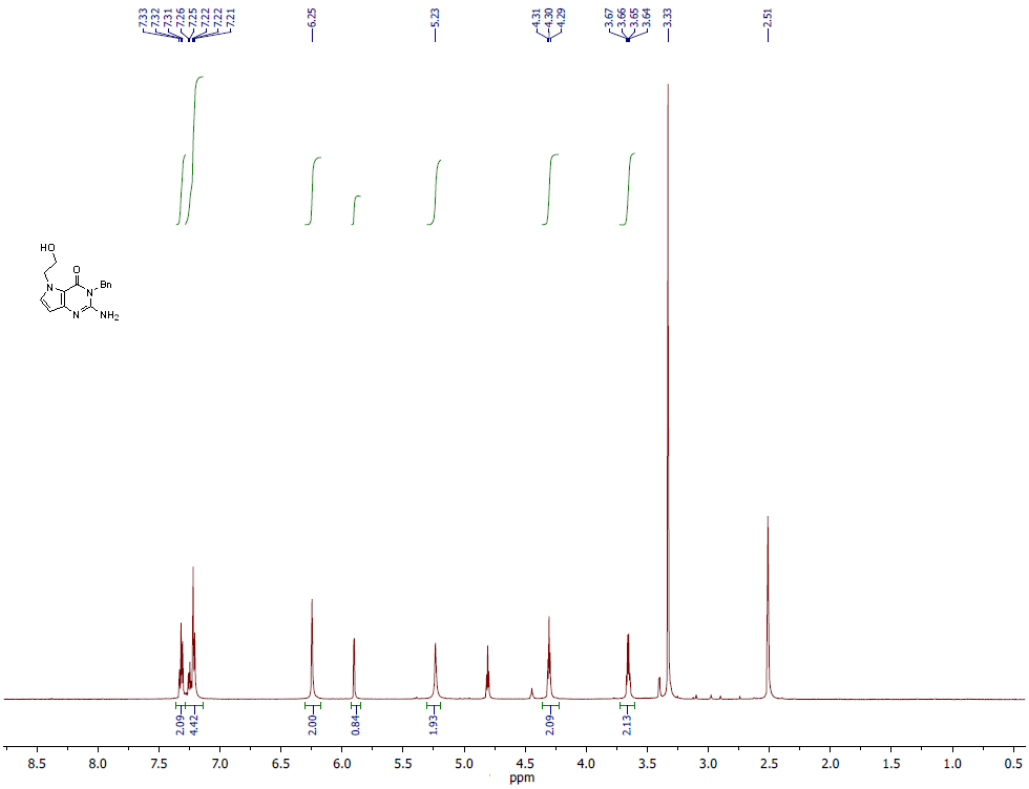
$^{13}$C spectrum of compound **15** in $CDCl_3$.



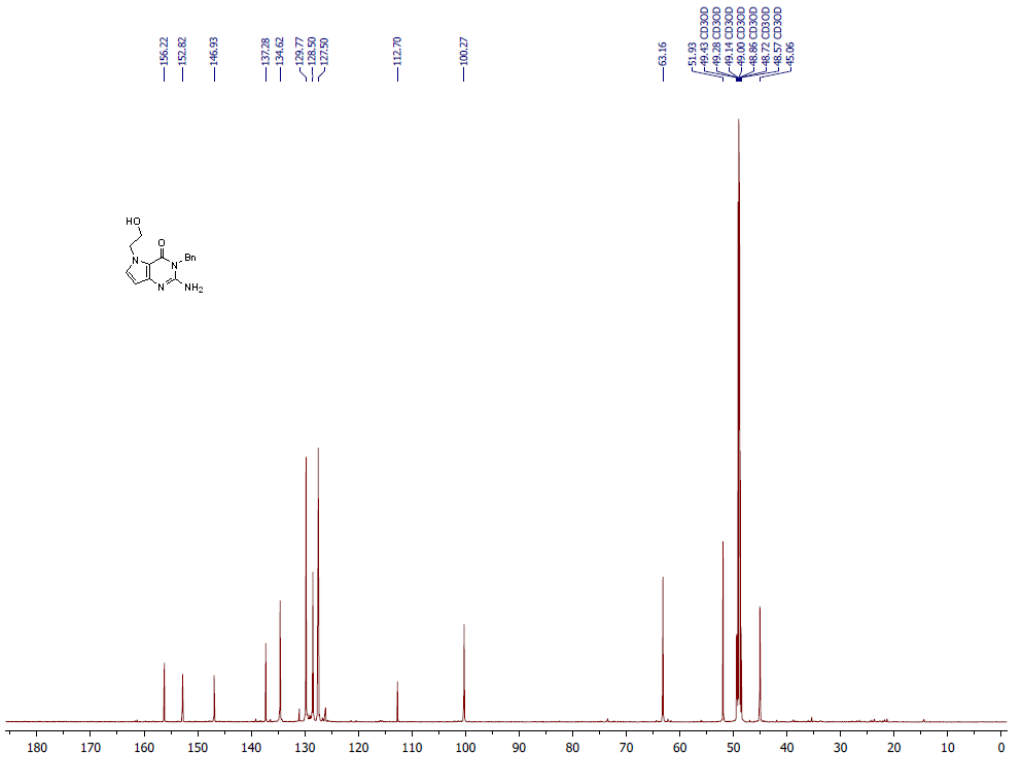
$^{1}$H spectrum of compound **16** in $CDCl_3$.



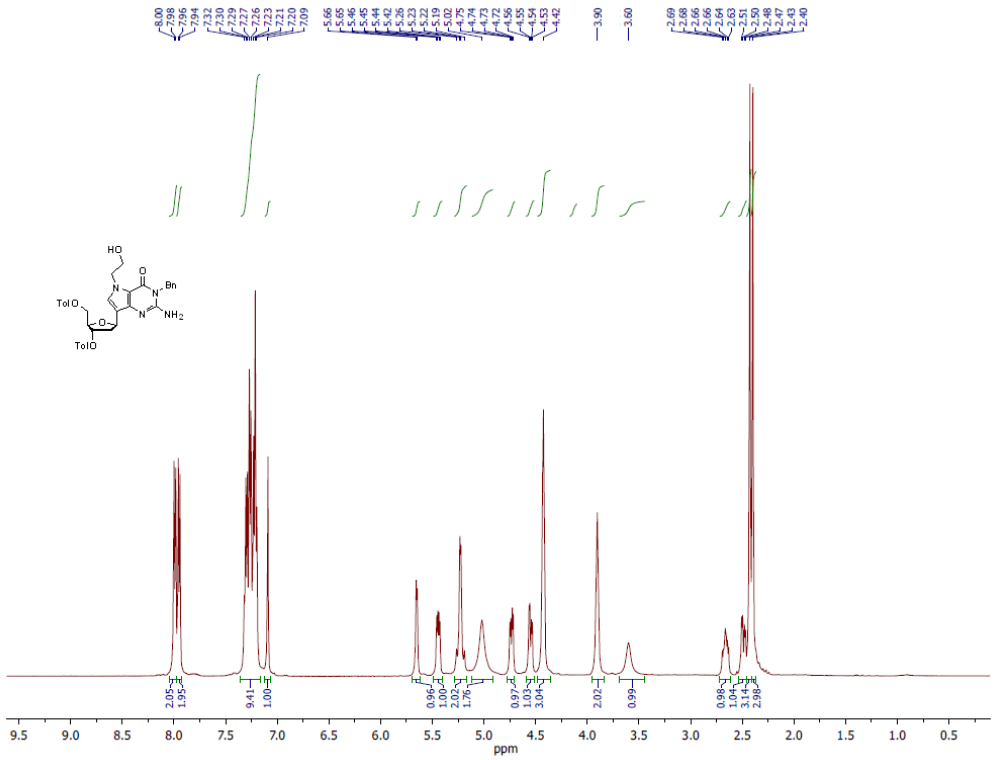
$^{13}$C spectrum of compound **16** in $CDCl_3$.



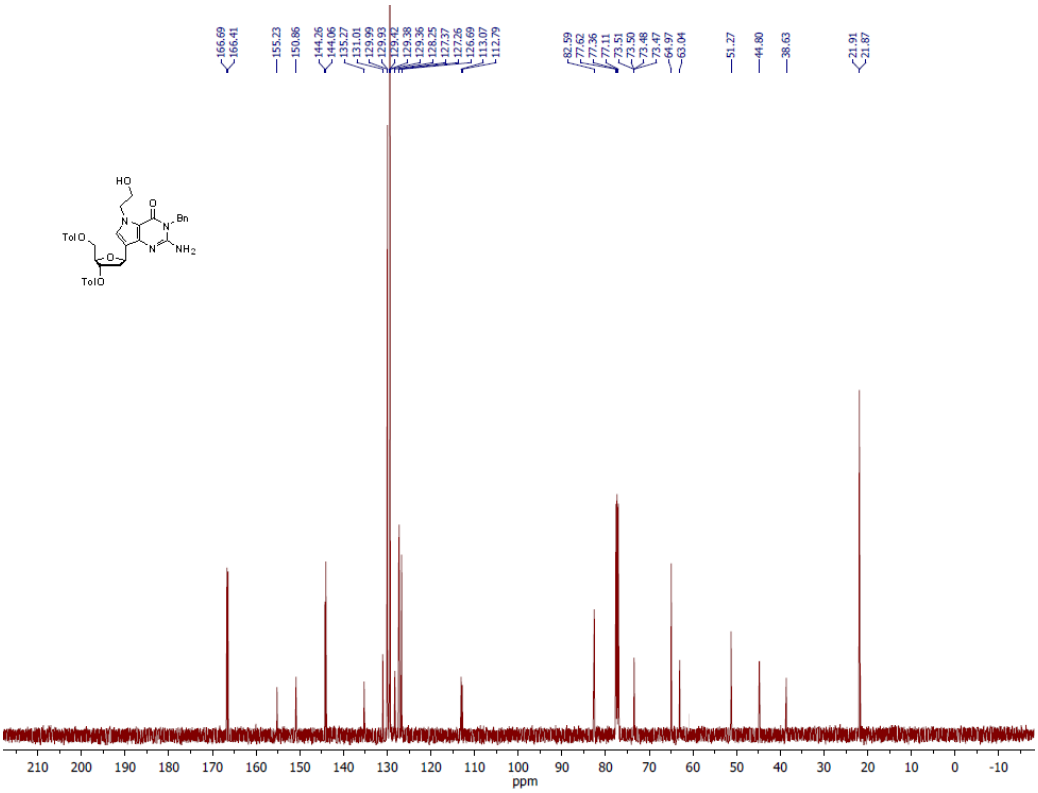

$^{1}$H spectrum of compound **17** in $d_6$-DMSO.



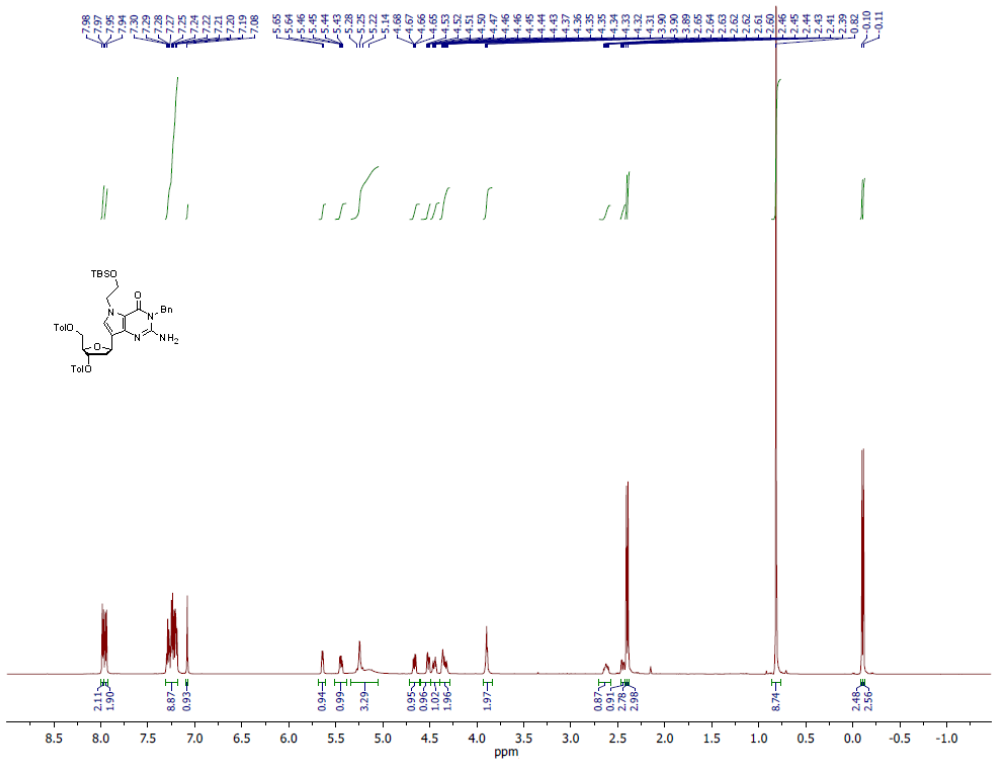

$^{13}$C spectrum of compound **17** in $CD_3OD$.



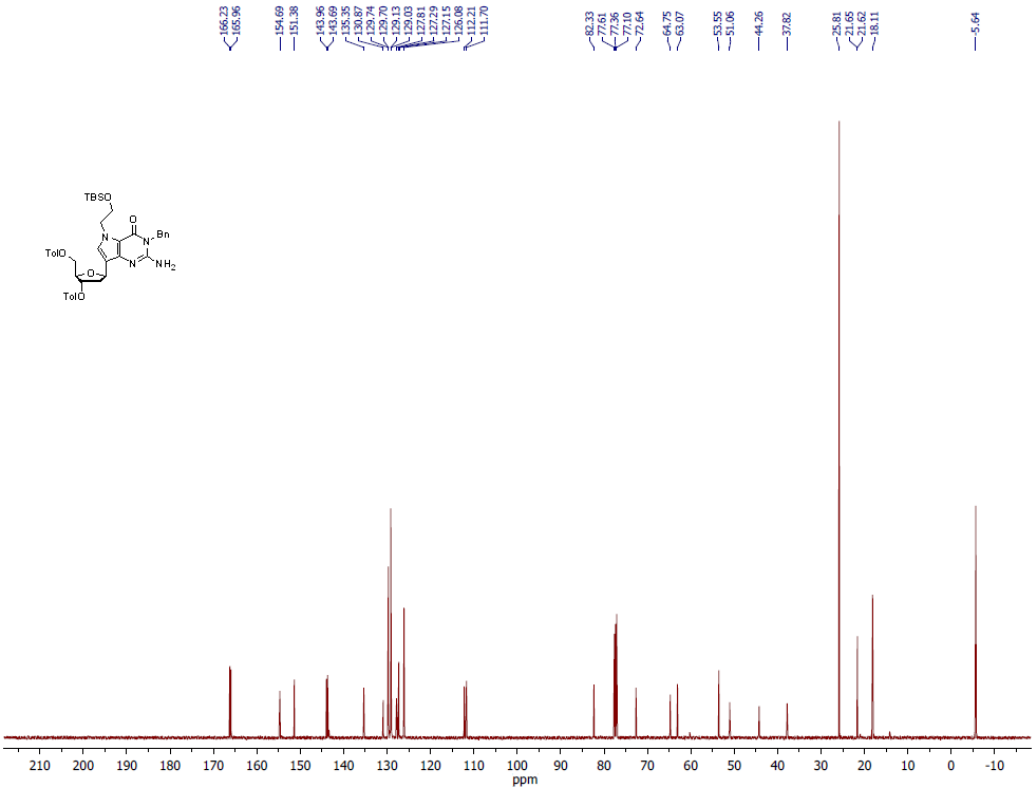
$^{1}$H spectrum of compound **18** in $CDCl_3$.



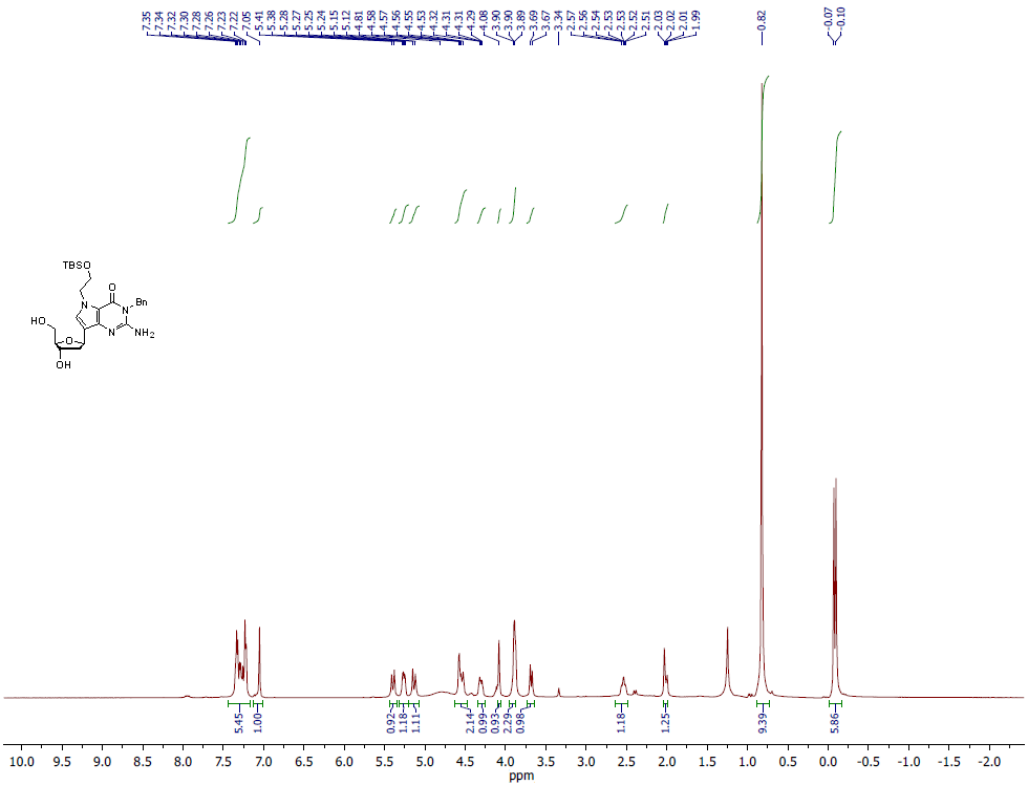
$^{13}$C spectrum of compound **18** in $CDCl_3$.



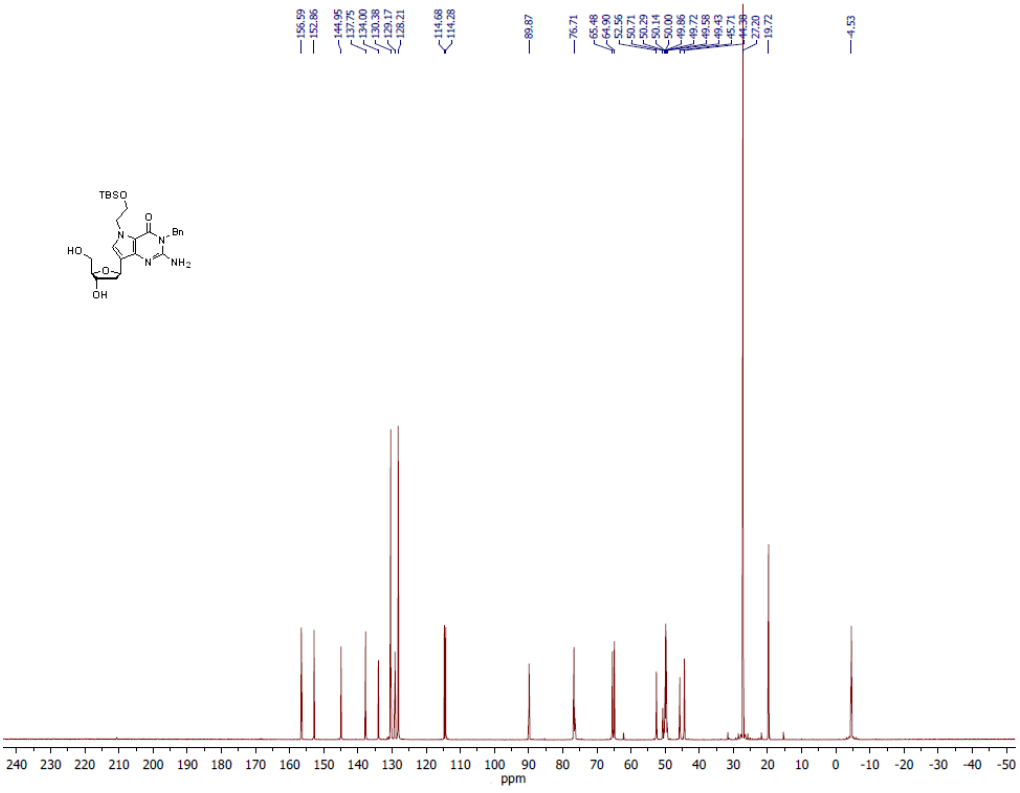

$^{1}$H spectrum of compound **19** in $CDCl_3$.



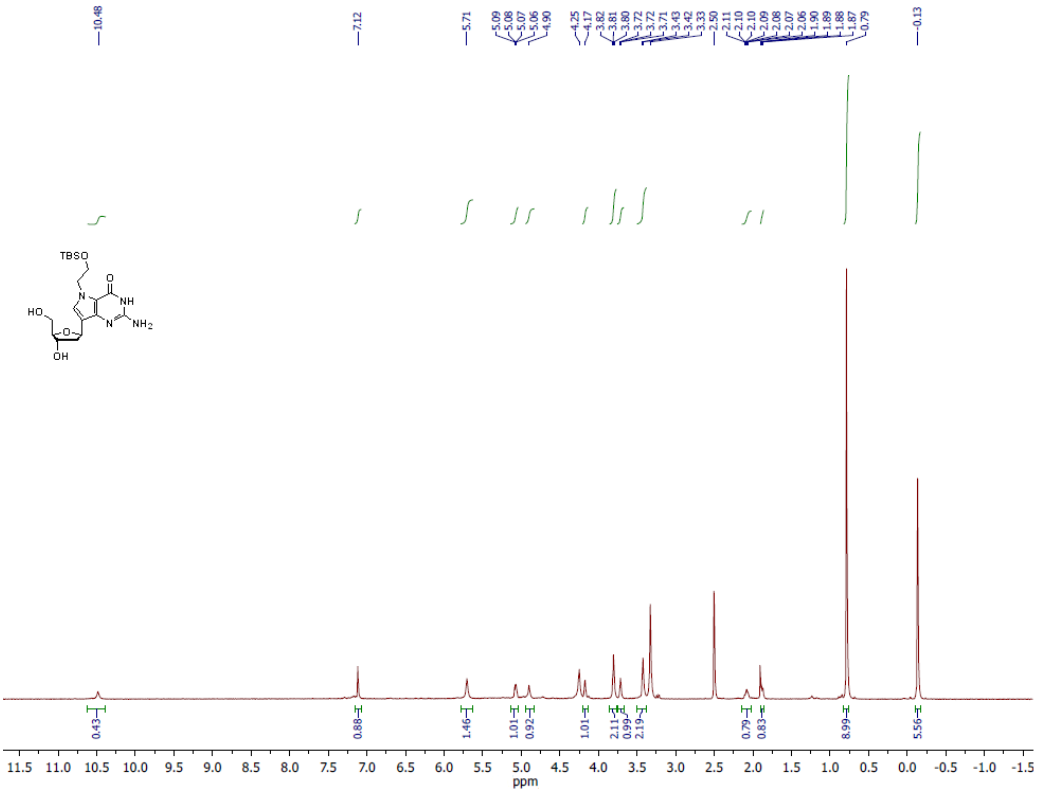

$^{13}$C spectrum of compound **19** in $CDCl_3$.



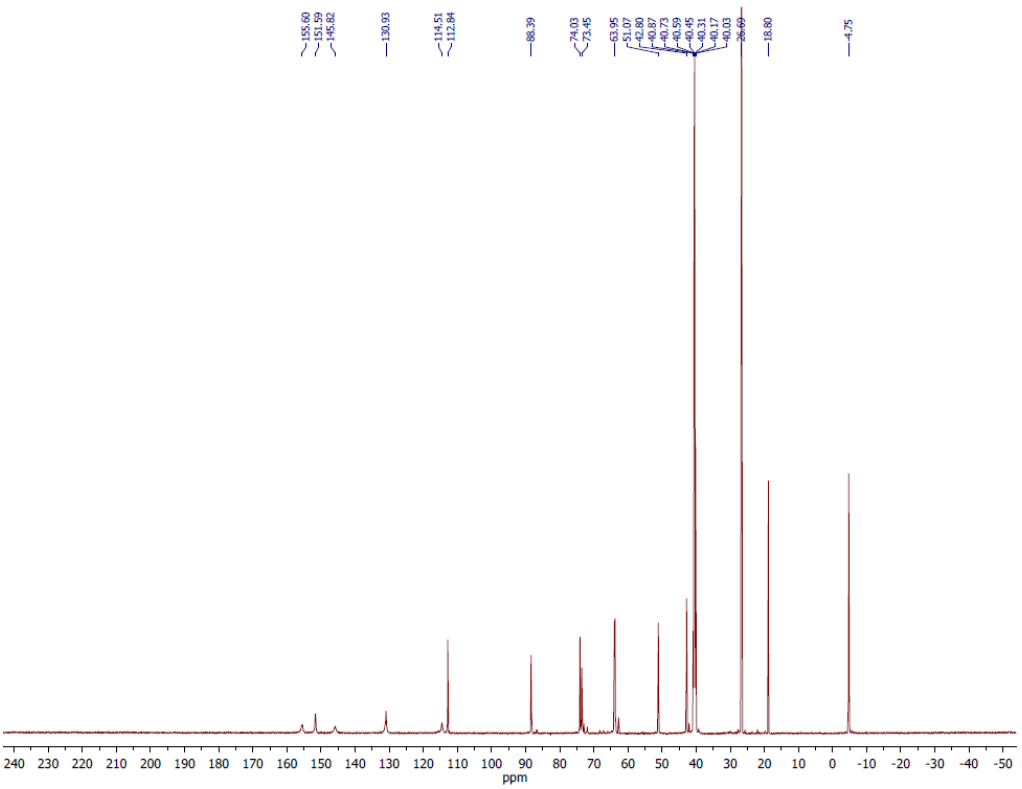

$^{1}$H spectrum of compound **20** in $CD_3OD$.



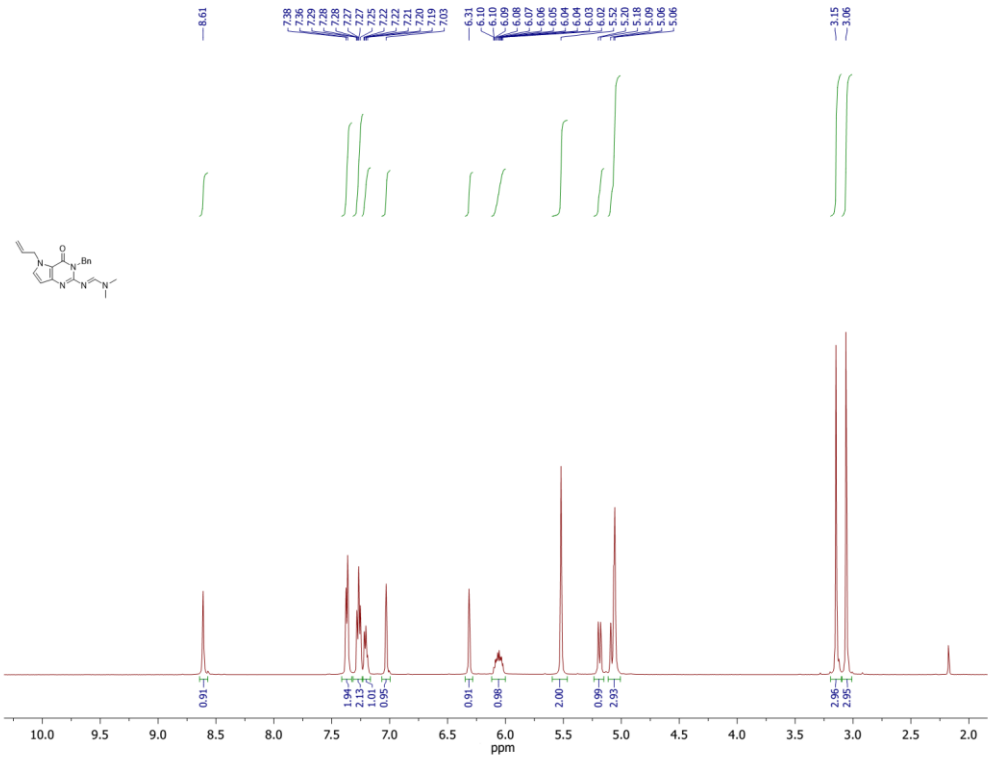

$^{13}$C spectrum of compound **20** in $CD_3OD$



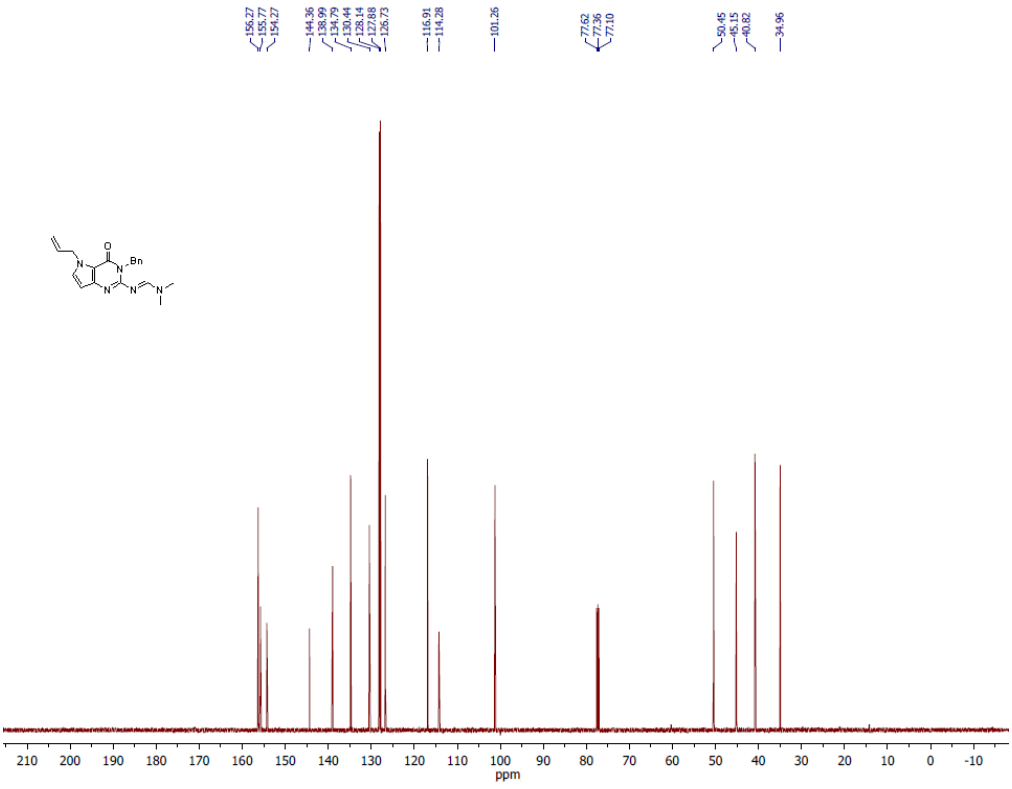

$^{1}$H spectrum of compound **3b** in $d_6$-DMSO.



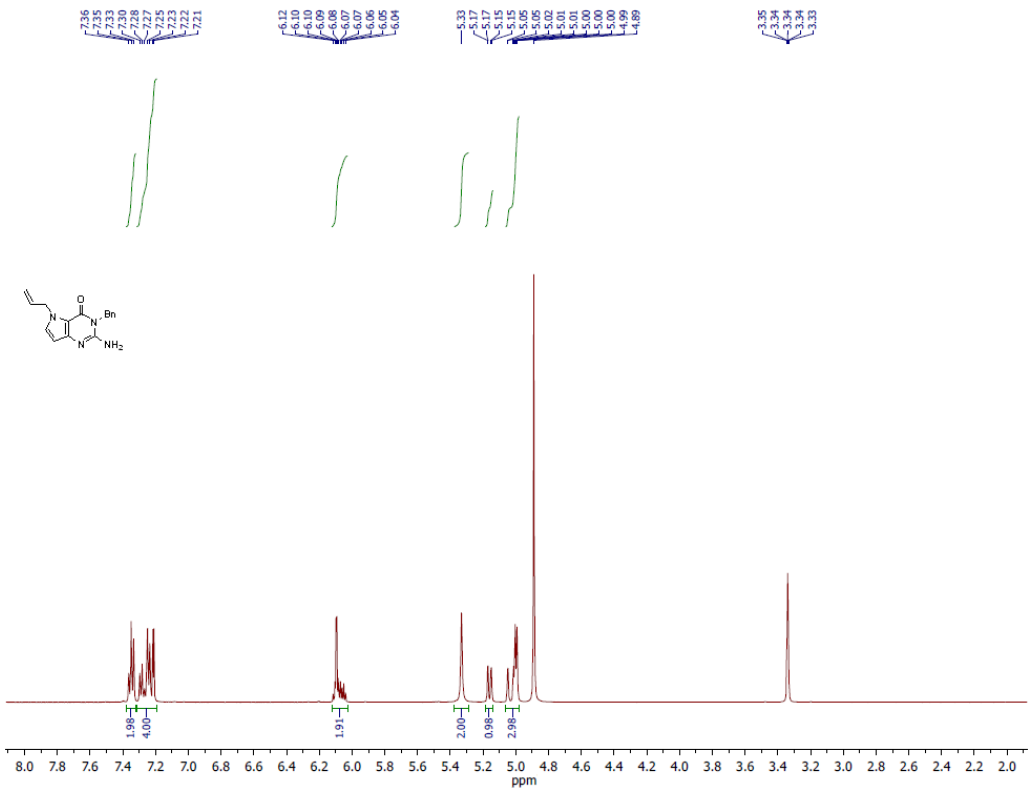

$^{13}$C spectrum of compound **3b** in $d_6$-DMSO.



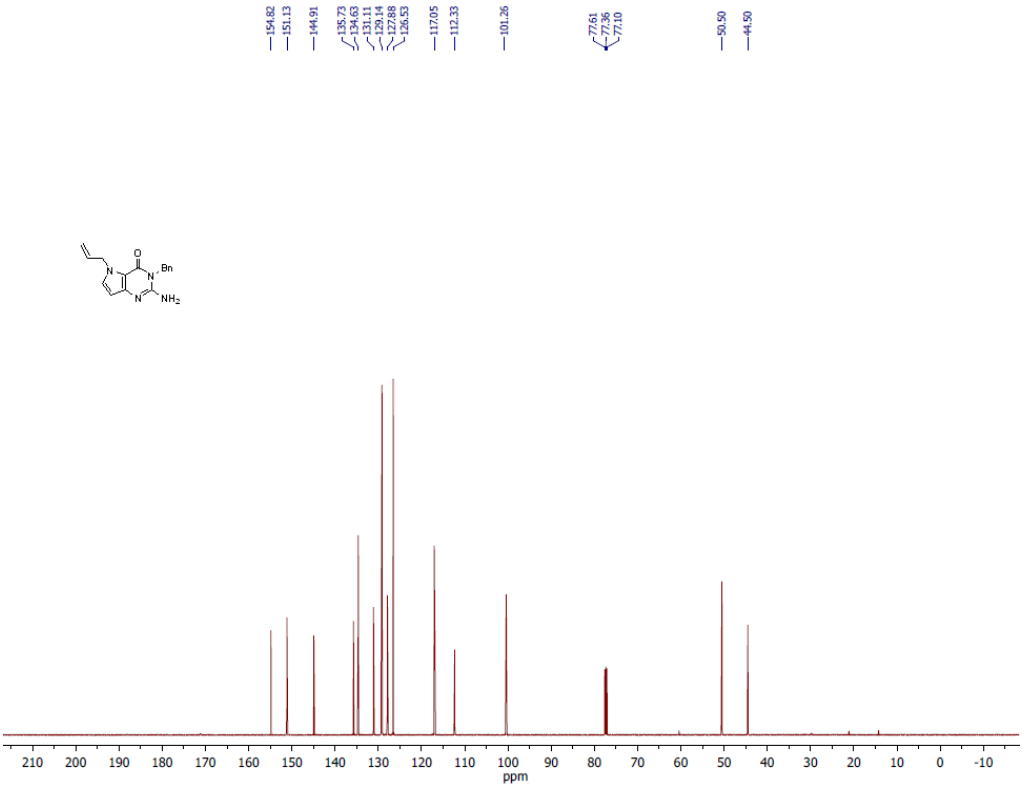
$^{1}H$ spectrum of compound **25** in $CDCl_3$.



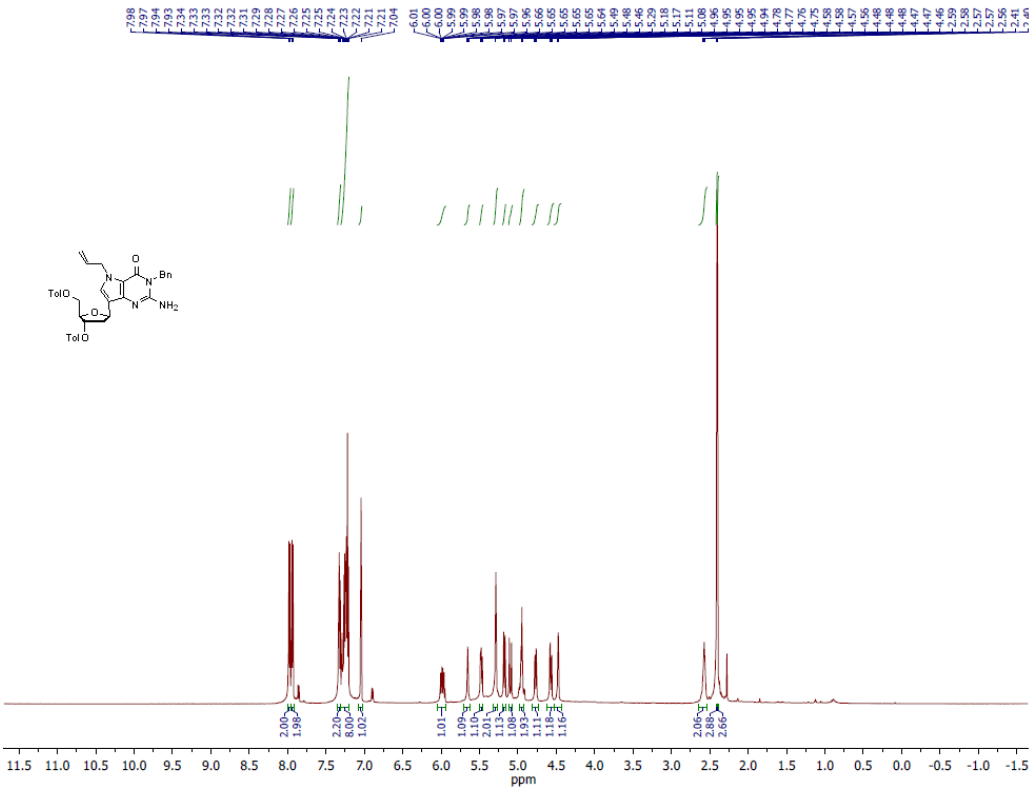
$^{13}C$ spectrum of compound **25** in $CDCl_3$.



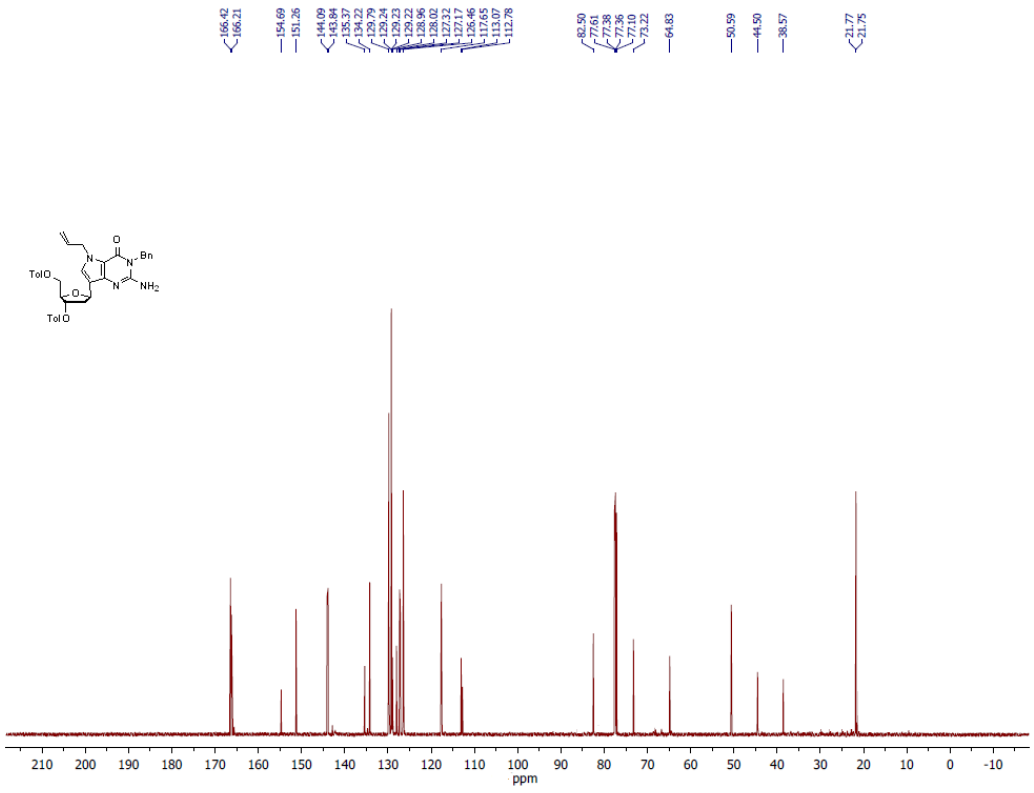
$^{1}$H spectrum of compound **26** in $CD_3OD$.



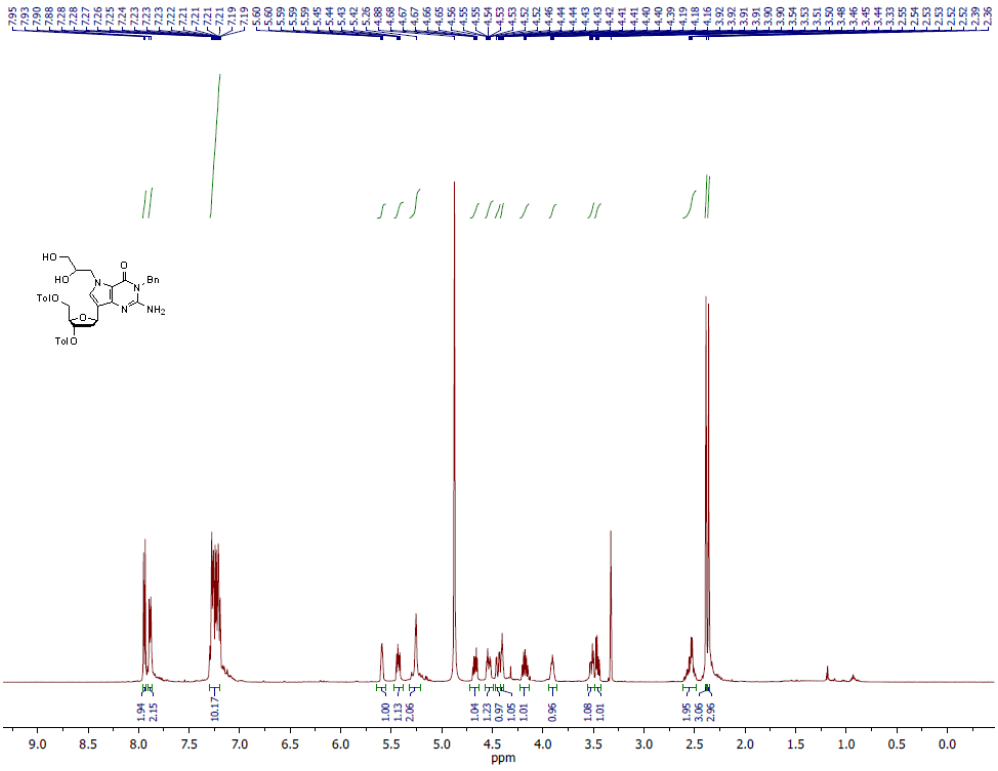
$^{13}$C spectrum of compound **26** in $CDCl_3$.



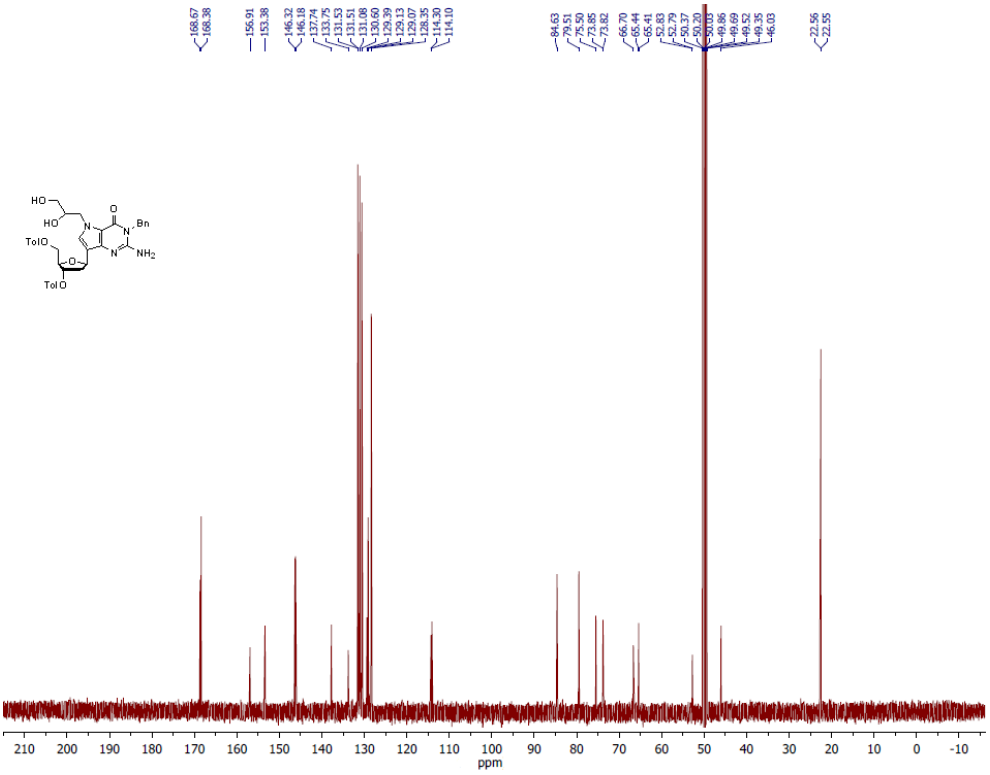
$^{1}$H spectrum of compound **27** in $CDCl_3$.



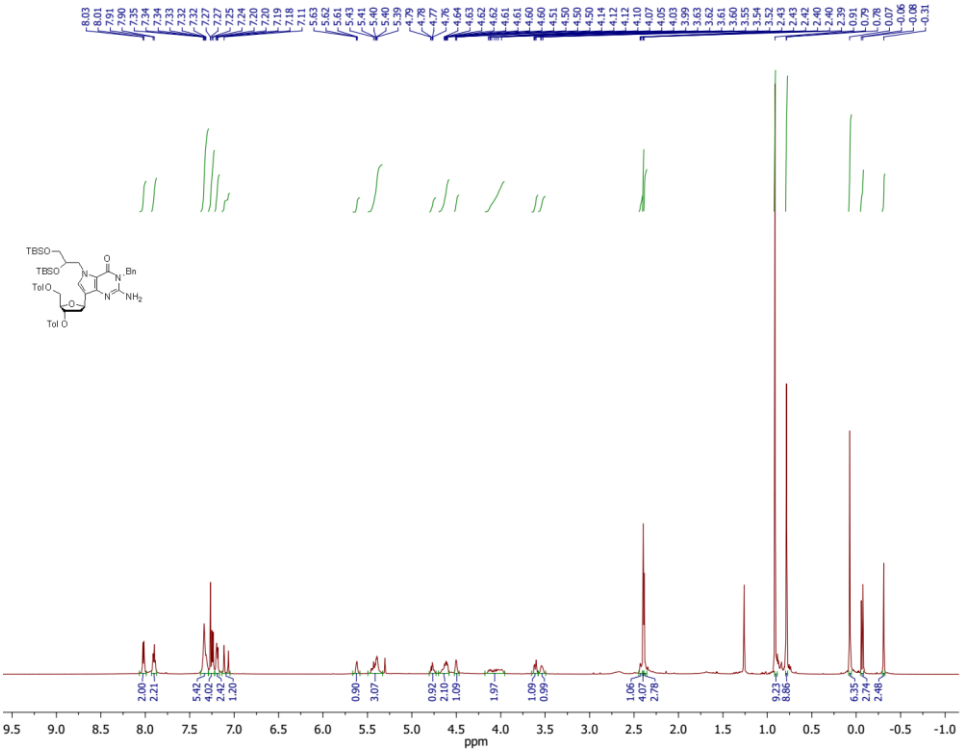
$^{13}$C spectrum of compound **27** in $CDCl_3$.



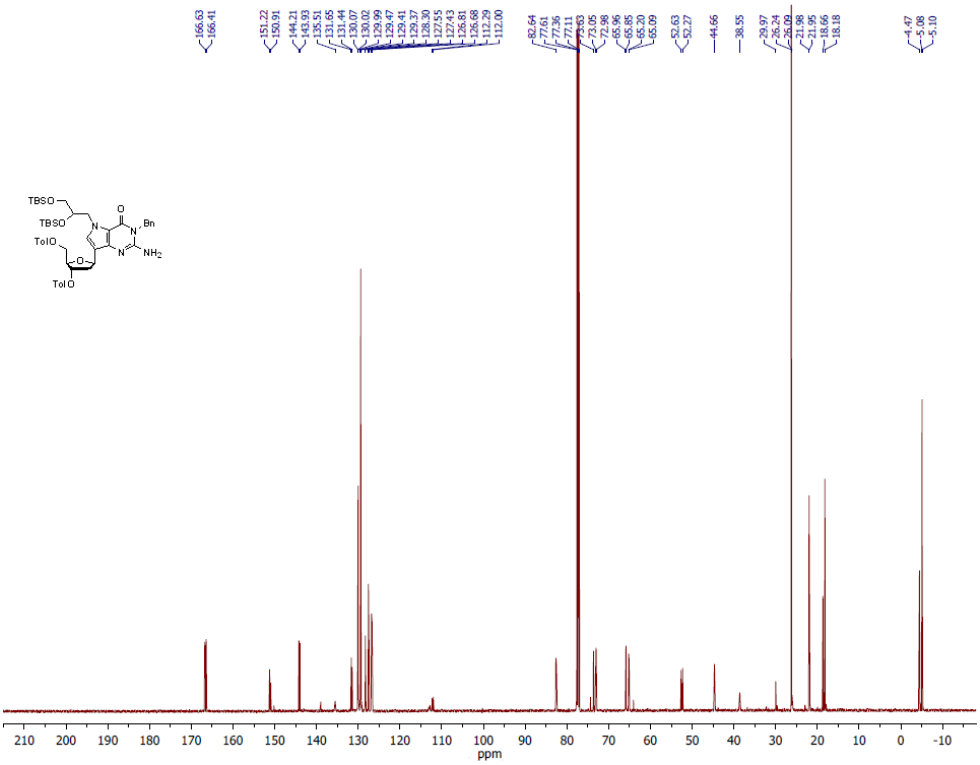
$^{1}$H spectrum of compound **28** in $CDCl_3$.



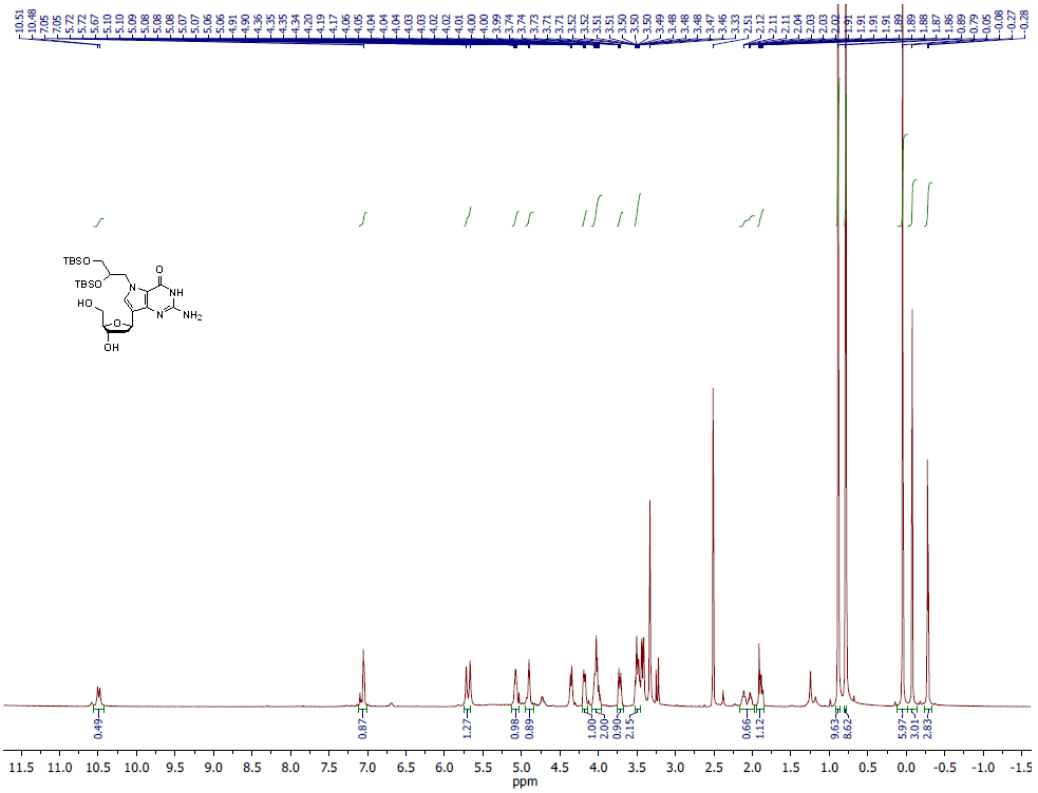
$^{13}$C spectrum of compound **28** in $CDCl_3$.



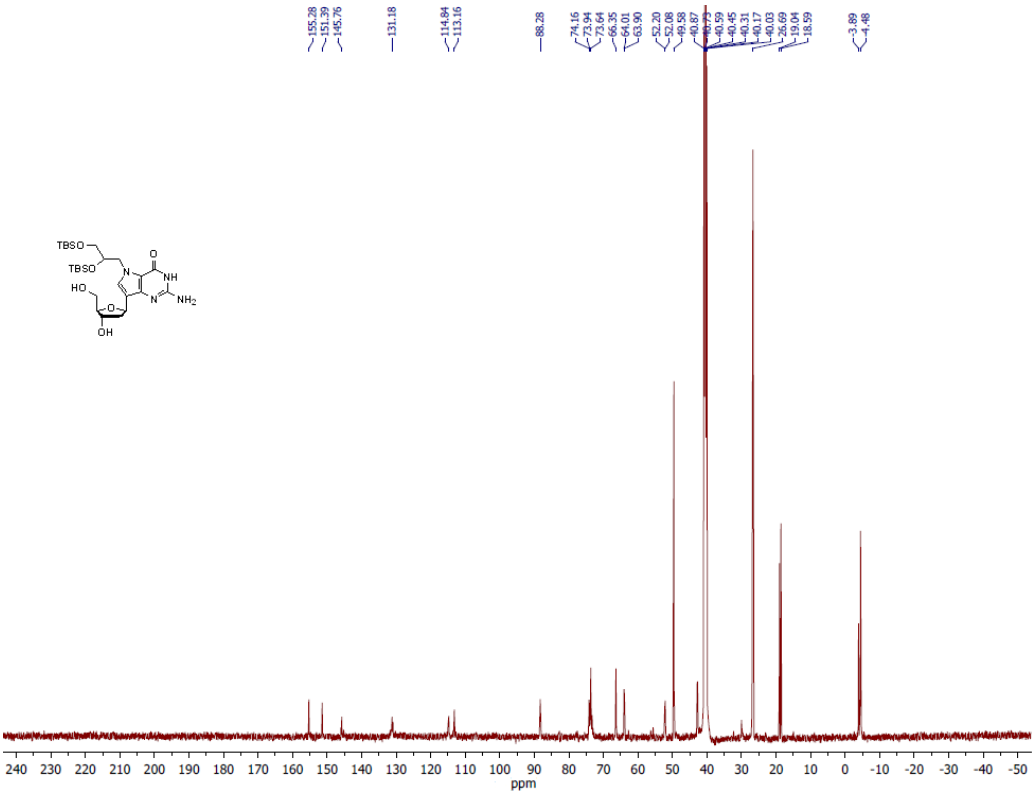
$^{1}$H spectrum of compound **29** in $CDCl_3$.



$^{13}$C spectrum of compound **29** in $CDCl_3$.




$^{1}$H spectrum of compound **3d** in $d_6$-DMSO.




$^{13}$C spectrum of compound **3d** in $d_6$-DMSO.

## REFERENCES


1. Jackson SP, Bartek J. The DNA-damage response in human biology and disease. Nature. 2009;461(7267):1071-8.

2. Rydberg B, Lindahl T. Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. The EMBO Journal. 1982;1(2):211.

3. Gaubatz JW, Tan BH. Introduction, distribution, and removal of 7-methylguanine in different liver chromatin fractions of young and old mice. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 1997;375(1):25-35.

4. Beranek DT. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 1990;231(1):11-30.

5. Kleihues P, Margison P. Brief Communication: Carcinogenicity of N-Methyl-N-nitrosourea: Possible Role of Excision Repair of O6-Methylguanine From DNA. Journal of the National Cancer Institute. 1974;53(6):1839-41.

6. Pegg AE, Hui G. Formation and subsequent removal of O6-methylguanine from deoxyribonucleic acid in rat liver and kidney after small doses of dimethylnitrosamine. Biochemical Journal. 1978;173(3):739-48.

7. Bedard LL, Massey TE. Aflatoxin B 1-induced DNA damage and its repair. Cancer Letters. 2006;241(2):174-83.

8. Croy R, Essigmann J, Reinhold V, Wogan G. Identification of the principal aflatoxin B1-DNA adduct formed in vivo in rat liver. Proceedings of the National Academy of Sciences. 1978;75(4):1745-9.

9. Adam W, Saha-Möller CR, Schönberger A. Photooxidation of 8-oxo-7, 8-dihydro-2 '-deoxyguanosine by thermally generated triplet-excited ketones from 3-(hydroxymethyl)-3, 4, 4-trimethyl-1, 2-dioxetane and comparison with type I and type II photosensitizers. Journal of the American Chemical Society. 1996;118(39):9233-8.

10. Luo W, Muller JG, Rachlin EM, Burrows CJ. Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-oxo-7, 8-dihydroguanosine. Organic Letters. 2000;2(5):613-6.

11. McCulloch SD, Kokoska RJ, Masutani C, Iwai S, Hanaoka F, Kunkel TA. Preferential cis–syn thymine dimer bypass by DNA polymerase η occurs with biased fidelity. Nature. 2004;428(6978):97-100.

12. Haracska L, Yu S-L, Johnson RE, Prakash L, Prakash S. Efficient and accurate replication in the presence of 7, 8-dihydro-8-oxoguanine by DNA polymerase η. Nature Genetics. 2000;25(4):458-61.

13. Eoff RL, Irimia A, Angel KC, Egli M, Guengerich FP. Hydrogen bonding of 7, 8-dihydro-8-oxodeoxyguanosine with a charged residue in the little finger domain determines miscoding events in Sulfolobus solfataricus DNA polymerase Dpo4. Journal of Biological Chemistry. 2007;282(27):19831-43.

14. Cheng KC, Cahill DS, Kasai H, Nishimura S, Loeb LA. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. Journal of Biological Chemistry. 1992;267(1):166-72.

15. Kou Y, Koag M-C, Lee S. N7 Methylation Alters Hydrogen-Bonding Patterns of Guanine in Duplex DNA. Journal of the American Chemical Society. 2015;137(44):14067-70.

16. Le Page F, Sarasin A, Gentil A, Guy A, Cadet J. Repair and mutagenic potency of 8-oxoG: A and 8-oxoG: C base pairs in mammalian cells. Nucleic Acids Research. 1998;26(5):1276-81.

17. Grąziewicz MA, Zastawny TH, Oliński R, Speina E, Siedlecki J, Tudek B. Fapyadenine is a moderately efficient chain terminator for prokaryotic DNA polymerases. Free Radical Biology and Medicine. 2000;28(1):75-83.

18. Kelly JD, Inga A, Chen F-X, Dande P, Shah D, Monti P, et al. Relationship between DNA methylation and mutational patterns induced by a sequence selective minor groove methylating agent. Journal of Biological Chemistry. 1999;274(26):18327-34.

19. Ghosh R, Mitchell DL. Effect of oxidative DNA damage in promoter elements on transcription factor binding. Nucleic Acids Research. 1999;27(15):3213-8.

20. Gray PJ. Sulphur mustards inhibit binding of transcription factor AP2 in vitro. Nucleic Acids Research. 1995;23(21):4378-82.

21. Ramon O, Sauvaigo S, Gasparutto D, Faure P, Favier A, Cadet J. Effects of 8-oxo-7, 8-dihydro-2′-deoxyguanosine on the binding of the transcription factor Sp1 to its cognate target DNA sequence (GC box). Free Radical Research. 1999;31(3):217-29.

22. Wachsman JT. DNA methylation and the association between genetic and epigenetic changes: relation to carcinogenesis. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 1997;375(1):1-8.

23. Chen RZ, Pettersson U, Beard C, Jackson-Grusby L, Jaenisch R. DNA hypomethylation leads to elevated mutation rates. Nature. 1998;395(6697):89-93.

24. Maekita T, Nakazawa K, Mihara M, Nakajima T, Yanaoka K, Iguchi M, et al. High levels of aberrant DNA methylation in Helicobacter pylori–infected gastric mucosae and its possible association with gastric cancer risk. Clinical Cancer Research. 2006;12(3):989-95.

25. Oxidative D. damage: mechanisms, mutation, and disease Cooke, Marcus S.; Evans, Mark D.; Dizdaroglu, Miral; Lunec, Joseph. FASEB Journal. 2003;17(10):1195-214.

26. Robison SH, Munzer JS, Mrcp RT, Bradley WG. Alzheimer's disease cells exhibit defective repair of alkylating agent—induced DNA damage. Annals of Aeurology. 1987;21(3):250-8.

27. Tobi SE, Neary D, Itzhaki RF. Alkylation damage and repair in Alzheimer's disease lymphocytes. Gerontology. 1993;39(5):241-51.

28. Smith MA, Rottkamp CA, Nunomura A, Raina AK, Perry G. Oxidative stress in Alzheimer's disease. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease. 2000;1502(1):139-44.

29. Hartnett L, Egan LJ. Inflammation, DNA methylation and colitis-associated cancer. Carcinogenesis. 2012:bgs006.

30. Shimoda R, Nagashima M, Sakamoto M, Yamaguchi N, Hirohashi S, Yokota J, et al. Increased formation of oxidative DNA damage, 8-hydroxydeoxyguanosine, in human livers with chronic hepatitis. Cancer Research. 1994;54(12):3171-2.

31. Lunec J, Herbert K, Blount S, Griffiths HR, Emery P. 8-Hydroxydeoxyguanosine: a marker of oxidative DNA damage in systemic lupus erythematosus. Febs Letters. 1994;348(2):131-8.

32. Lee SH, Blair IA. Oxidative DNA damage and cardiovascular disease. Trends in Cardiovascular medicine. 2001;11(3):148-55.

33. Dhalla NS, Temsah RM, Netticadan T. Role of oxidative stress in cardiovascular diseases. Journal of Hypertension. 2000;18(6):655-73.

34. Haripriya D, Sangeetha P, Kanchana A, Balu M, Panneerselvam C. Modulation of age-associated oxidative DNA damage in rat brain cerebral cortex, striatum and hippocampus by L-carnitine. Experimental Gerontology. 2005;40(3):129-35.

35. Lu T, Pan Y, Kao S-Y, Li C, Kohane I, Chan J, et al. Gene regulation and DNA damage in the ageing human brain. Nature. 2004;429(6994):883-91.

36. Gorbunova V, Seluanov A, Mao Z, Hine C. Changes in DNA repair during aging. Nucleic Acids Research. 2007;35(22):7466-74.

37. Goukassian D, Gad F, Yaar M, Eller MS, Nehal US, Gilchrest BA. Mechanisms and implications of the age-associated decrease in DNA repair capacity. The FASEB Journal. 2000;14(10):1325-34.

38. Atamna H, Cheung I, Ames BN. A method for detecting abasic sites in living cells: age-dependent changes in base excision repair. Proceedings of the National Academy of Sciences. 2000;97(2):686-91.

39. Chen S-K, Hsieh WA, Tsai M-H, Chen C-C, Hong AI, Wei Y-H, et al. Age-associated decrease of oxidative repair enzymes, human 8-oxoguanine DNA glycosylases (hOgg1), in human aging. Journal of Radiation Research. 2003;44(1):31-5.

40. Tiano L, Littarru GP, Principi F, Orlandi M, Santoro L, Carnevali P, et al. Assessment of DNA damage in Down Syndrome patients by means of a new, optimised single cell gel electrophoresis technique. Biofactors. 2005;25(1-4):187-95.

41. Beland FA, Churchwell MI, Von Tungeln LS, Chen S, Fu PP, Culp SJ, et al. High-performance liquid chromatography electrospray ionization tandem mass spectrometry for the detection and quantitation of benzo [a] pyrene-DNA adducts. Chemical Research in Toxicology. 2005;18(8):1306-15.

42. Jenner A, Timothy G, ARUOMA OI, HALLIWELL B. Measurement of oxidative DNA damage by gas chromatography–mass spectrometry: ethanethiol prevents artifactual generation of oxidized DNA bases. Biochemical Journal. 1998;331(2):365-9.

43. Douki T, Spinelli S, Ravanat J-L, Cadet J. Hydroxyl radical-induced degradation of 2′-deoxyguanosine under reducing conditions. Journal of the Chemical Society, Perkin Transactions 2. 1999;(9):1875-80.

44. Fojta M. Electrochemical sensors for DNA interactions and damage. Electroanalysis. 2002;14(21):1449-63.

45. Cahová-Kuchaříková K, Fojta M, Mozga T, Paleček E. Use of DNA repair enzymes in electrochemical detection of damage to DNA bases in vitro and in cells. Analytical Chemistry. 2005;77(9):2920-7.

46. Sinha RP, Dautz M, Hader D. A simple and efficient method for the quantitative analysis of thymine dimers in cyanobacteria, phytoplankton and macroalgae. Acta Protozoologica. 2001;40(3):187-96.

47. Yin B, Whyatt RM, Perera FP, Randall MC, Cooper TB, Santella RM. Determination of 8-hydroxydeoxyguanosine by an immunoaffinity chromatography-monoclonal antibody-based ELISA. Free Radical Biology and Medicine. 1995;18(6):1023-32.

48. Barbin A, Bartsch H. Nucleophilic selectivity as a determinant of carcinogenic potency (TD 50) in rodents: a comparison of mono-and bi-functional alkylating agents and vinyl chloride metabolites. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 1989;215(1):95-106.

49. Bolt HM, Leutbecher M, Golka K. A note on the physiological background of the ethylene oxide adduct 7-(2-hydroxyethyl) guanine in DNA from human blood. Archives of Toxicology. 1997;71(11):719-21.

50. Nivard MJ, Czene K, Segerbäck D, Vogel EW. Mutagenic activity of ethylene oxide and propylene oxide under XPG proficient and deficient conditions in relation to N-7-(2-hydroxyalkyl) guanine levels in Drosophila. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 2003;529(1):95-107.

51. Rusyn I, Asakura S, Li Y, Kosyk O, Koc H, Nakamura J, et al. Effects of ethylene oxide and ethylene inhalation on DNA adducts, apurinic/apyrimidinic sites and expression of base excision DNA repair genes in rat brain, spleen, and liver. DNA Repair. 2005;4(10):1099-110.

52. Christov PP, Angel KC, Guengerich FP, Rizzo CJ. Replication past the N5-methyl-formamidopyrimidine lesion of deoxyguanosine by DNA polymerases and an improved procedure for sequence analysis of in vitro bypass products by mass spectrometry. Chemical Research in Toxicology. 2009;22(6):1086.

53. Vodička P, Hemminki K. Depurination and imidazole ring-opening in nucleosides and DNA alkylated by styrene oxide. Chemico-biological Interactions. 1988;68(1-2):117-26.

54. Philippin G, Cadet J, Gasparutto D, Mazon G, Fuchs RP. Ethylene oxide and propylene oxide derived N7-alkylguanine adducts are bypassed accurately in vivo. DNA Repair. 2014;22:133-6.

55. Barbin A, Laib RJ, Bartsch H. Lack of miscoding properties of 7-(2-oxoethyl) guanine, the major vinyl chloride-DNA adduct. Cancer Research. 1985;45(6):2440-4.

56. Kawate T, Allerson CR, Wolfe JL. Regioselective syntheses of 7-nitro-7-deazapurine nucleosides and nucleotides for efficient PCR incorporation. Organic Letters. 2005;7(18):3865-8.

57. Chen X, Kopecky DJ, Mihalic J, Jeffries S, Min X, Heath J, et al. Structure-guided design, synthesis, and evaluation of guanine-derived inhibitors of the eIF4E mRNA–cap interaction. Journal of Medicinal Chemistry. 2012;55(8):3837-51.

58. Angelov T, Guainazzi A, Schärer OD. Generation of DNA Interstrand Crosslinks by Post-Synthetic Reductive Amination. Organic Letters. 2009;11(3):661.

59. Rana J, Huang H. Corrigendum to ‘‘Actions of the Klenow fragment of DNA polymerase I and some DNA glycosylases on chemically stable analogues of N7-methyl-20-deoxyguanosine’’[Bioorg. Med. Chem. Lett. 21. Bioorganic & Medicinal Chemistry. 2014;22(9):2825.

60. Hoffer M. α-Thymidin. Chemische Berichte. 1960;93(12):2777-81.

61. Hamm ML, Parker AJ, Steele TW, Carman JL, Parish CA. Oligonucleotide incorporation and base pair stability of 9-deaza-2′-deoxyguanosine, an analogue of 8-oxo-2′-deoxyguanosine. The Journal of Organic Chemistry. 2010;75(16):5661-9.

62. Wellington KW, Benner SA. A review: Synthesis of aryl C-glycosides via the Heck coupling reaction. Nucleosides, Nucleotides, and Nucleic Acids. 2006;25(12):1309-33.

63. Gibson ES, Lesiak K, Watanabe KA, Gudas LJ, Pankiewicz KW. Synthesis of a Novel C-Nucleoside, 2-Amino-7-(2-deoxy-β-D-erythro-pentofuranosyl)-3H, 5H-pyrrolo-[3, 2-d] pyrimidin-4-one (2′-Deoxy-9-deazaguanosine). Nucleosides, Nucleotides & Nucleic Acids. 1999;18(3):363-76.

64. Suzuki J, Inoue Y, Suzuki S. Changes in the urinary excretion level of 8-hydroxyguanine by exposure to reactive oxygen-generating substances. Free Radical Biology and Medicine. 1995;18(3):431-6.

65. Bruskov VI, Malakhova LV, Masalimov ZK, Chernikov AV. Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA. Nucleic Acids Research. 2002;30(6):1354-63.

66. Kehrer JP. The Haber–Weiss reaction and mechanisms of toxicity. Toxicology. 2000;149(1):43-50.

67. Aruoma OI, Halliwell B, Gajewski E, Dizdaroglu M. Copper-ion-dependent damage to the bases in DNA in the presence of hydrogen peroxide. Biochemical Journal. 1991;273(3):601-4.

68. Obtułowicz T, Swoboda M, Speina E, Gackowski D, Rozalski R, Siomek A, et al. Oxidative stress and 8-oxoguanine repair are enhanced in colon adenoma and carcinoma patients. Mutagenesis. 2010:geq028.

69. Andreoli R, Mutti A, Goldoni M, Manini P, Apostoli P, De Palma G. Reference ranges of urinary biomarkers of oxidized guanine in (2′-deoxy) ribonucleotides and nucleic acids. Free Radical Biology and Medicine. 2011;50(2):254-61.

70. Siomek A, Tujakowski J, Gackowski D, Rozalski R, Foksinski M, Dziaman T, et al. Severe oxidatively damaged DNA after cisplatin treatment of cancer patients. International Journal of Cancer. 2006;119(9):2228-30.

71. Manini P, De Palma G, Andreoli R, Marczynski B, Hanova M, Mozzoni P, et al. Biomarkers of nucleic acid oxidation, polymorphism in, and expression of, hOGG1 gene in styrene-exposed workers. Toxicology Letters. 2009;190(1):41-7.

72. Loft S, Danielsen P, Løhr M, Jantzen K, Hemmingsen JG, Roursgaard M, et al. Urinary excretion of 8-oxo-7, 8-dihydroguanine as biomarker of oxidative damage to DNA. Archives of biochemistry and biophysics. 2012;518(2):142-50.

73. Gackowski D, Rozalski R, Siomek A, Dziaman T, Nicpon K, Klimarczyk M, et al. Oxidative stress and oxidative DNA damage is characteristic for mixed Alzheimer disease/vascular dementia. Journal of the Neurological Sciences. 2008;266(1):57-62.

74. Crohns M, Saarelainen S, Erhola M, Alho H, Kellokumpu-Lehtinen P. Impact of radiotherapy and chemotherapy on biomarkers of oxidative DNA damage in lung cancer patients. Clinical Biochemistry. 2009;42(10):1082-90.

75. Szaflarska-Popławska A, Siomek A, Czerwionka-Szaflarska M, Gackowski D, Różalski R, Guz J, et al. Oxidatively damaged DNA/oxidative stress in children with celiac disease. Cancer Epidemiology and Prevention Biomarkers. 2010;19(8):1960-5.

76. Rozalski R, Siomek A, Gackowski D, Foksinski M, Gran C, Klungland A, et al. Substantial decrease of urinary 8-oxo-7, 8-dihydroguanine, a product of the base excision repair

pathway, in DNA glycosylase defective mice. The International Journal of Biochemistry & Cell Biology. 2005;37(6):1331-6.

77. Foksinski M, Gackowski D, Rozalski R, Siomek A, Guz J, Szpila A, et al. Effects of basal level of antioxidants on oxidative DNA damage in humans. European Journal of Nutrition. 2007;46(3):174-80.

78. LoftlH S, Poulsen HE. Estimation of oxidative DNA damage in man from urinary ex-cretion of repair products. 1998.

79. Floyd RA, Watson JJ, Wong PK, Altmiller DH, Rickard RC. Hydroxyl free radical adduct of deoxyguanosine: sensitive detection and mechanisms of formation. Free radical research communications. 1986;1(3):163-72.

80. Helbock HJ, Beckman KB, Shigenaga MK, Walter PB, Woodall AA, Yeo HC, et al. DNA oxidation matters: the HPLC–electrochemical detection assay of 8-oxo-deoxyguanosine and 8-oxo-guanine. Proceedings of the National Academy of Sciences. 1998;95(1):288-93.

81. Stadlert RH. Determination of 8-oxoguanine in DNA by gas chromatography-mass spectrometry and HPLC-electrochemical detection: overestimation of the background level of the oxidized base by the gas chromatography-mass spectrometry assay. Chem Res Toxicol. 1995;8(8):1039-45.

82. Podmore K, Farmer P, Herbert K, Jones G, Martin E. 32 P-Postlabelling approaches for the detection of 8-oxo-2′-deoxyguanosine-3′-monophosphate in DNA. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis. 1997;378(1):139-49.

83. Devanaboyina U-s, Gupta RC. Sensitive detection of 8-hydroxy-2′-deoxyguanosine in DNA by 32P-postlabeling assay and the basal levels in rat tissues. Carcinogenesis. 1996;17(5):917-24.

84. Shimoi K, Kasai H, Yokota N, Toyokuni S, Kinae N. Comparison between high-performance liquid chromatography and enzyme-linked immunosorbent assay for the determination of 8-hydroxy-2′-deoxyguanosine in human urine. Cancer Epidemiology and Prevention Biomarkers. 2002;11(8):767-70.

85. Weimann A, Belling D, Poulsen HE. Quantification of 8-oxo-guanine and guanine as the nucleobase, nucleoside and deoxynucleoside forms in human urine by high-performance liquid chromatography–electrospray tandem mass spectrometry. Nucleic Acids Research. 2002;30(2):e7-e.

86. Jayasena SD. Aptamers: an emerging class of molecules that rival antibodies in diagnostics. Clinical Chemistry. 1999;45(9):1628-50.

87. Carothers JM, Goler JA, Kapoor Y, Lara L, Keasling JD. Selecting RNA aptamers for synthetic biology: investigating magnesium dependence and predicting binding affinity. Nucleic Acids Research. 2010:gkq082.

88. Mandal M, Boese B, Barrick JE, Winkler WC, Breaker RR. Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. Cell. 2003;113(5):577-86.

89. Lau PS, Coombes BK, Li Y. A General Approach to the Construction of Structure-Switching Reporters from RNA Aptamers. Angewandte Chemie International Edition. 2010;49(43):7938-42.

90. Zheng D, Zou R, Lou X. Label-free fluorescent detection of ions, proteins, and small molecules using structure-switching aptamers, SYBR gold, and exonuclease I. Analytical chemistry. 2012;84(8):3554-60.

91. Liang J, Chen Z, Guo L, Li L. Electrochemical sensing of L-histidine based on structure-switching DNAzymes and gold nanoparticle–graphene nanosheet composites. Chemical Communications. 2011;47(19):5476-8.

92. Null EL, Lu Y. Rapid determination of enantiomeric ratio using fluorescent DNA or RNA aptamers. Analyst. 2010;135(2):419-22.

93. Tang J, Breaker RR. Rational design of allosteric ribozymes. Chemistry & Biology. 1997;4(6):453-9.

94. Win MN, Liang JC, Smolke CD. Frameworks for programming biological function through RNA parts and devices. Chemistry & Biology. 2009;16(3):298-310.

95. Gustav M. Optics of turbid Media. Ann Physik. 1908;25(3):377-445.

96. Templeton AC, Pietron JJ, Murray RW, Mulvaney P. Solvent refractive index and core charge influences on the surface plasmon absorbance of alkanethiolate monolayer-protected gold clusters. The Journal of Physical Chemistry B. 2000;104(3):564-70.

97. Swanson N, Billard B. Optimization of extinction from surface plasmon resonances of gold nanoparticles. Nanotechnology. 2003;14(3):353.

98. Mirkin CA, Letsinger RL, Mucic RC, Storhoff JJ. A DNA-based method for rationally assembling nanoparticles into macroscopic materials. Nature. 1996;382(6592):607.

99. Chakrabarti R, Klibanov AM. Nanocrystals modified with peptide nucleic acids (PNAs) for selective self-assembly and DNA detection. Journal of the American Chemical Society. 2003;125(41):12531-40.

100. Li H, Rothberg LJ. DNA sequence detection using selective fluorescence quenching of tagged oligonucleotide probes by gold nanoparticles. Analytical chemistry. 2004;76(18):5414-7.

101. Reynolds RA, Mirkin CA, Letsinger RL. Homogeneous, nanoparticle-based quantitative colorimetric detection of oligonucleotides. Journal of the American Chemical Society. 2000;122(15):3795-6.

102. Wang L, Liu X, Hu X, Song S, Fan C. Unmodified gold nanoparticles as a colorimetric probe for potassium DNA aptamers. Chemical Communications. 2006;(36):3780-2.

103. Wei H, Li B, Li J, Wang E, Dong S. Simple and sensitive aptamer-based colorimetric sensing of protein using unmodified gold nanoparticle probes. Chemical Communications. 2007;(36):3735-7.

104. Huizenga DE, Szostak JW. A DNA aptamer that binds adenosine and ATP. BIOCHEMISTRY-PENNSYLVANIA THEN WASHINGTON-. 1995;34:656-.

105. Patel M, Dutta A, Huang H. A selective adenosine sensor derived from a triplex DNA aptamer. Analytical and bioanalytical chemistry. 2011;400(9):3035-40.

106. Roy J, Chirania P, Ganguly S, Huang H. A DNA aptamer sensor for 8-oxo-7, 8-dihydroguanine. Bioorganic & medicinal chemistry letters. 2012;22(2):863-7.

107. Sankaran N, Nishizawa S, Seino T, Yoshimoto K, Teramae N. Abasic-Site-Containing Oligodeoxynucleotides as Aptamers for Riboflavin. Angewandte Chemie International Edition. 2006;45(10):1563-8.

108. Li M, Sato Y, Nishizawa S, Seino T, Nakamura K, Teramae N. 2-Aminopurine-modified abasic-site-containing duplex DNA for highly selective detection of theophylline. Journal of the American Chemical Society. 2009;131(7):2448.

109. Han MS, Lytton-Jean AK, Mirkin CA. A gold nanoparticle based approach for screening triplex DNA binders. Journal of the American Chemical Society. 2006;128(15):4954.