UNVIERSTIY OF CALIFORNIA

RIVERSIDE

Occurrence and Biological Consequences of Alkyl Phosphotriester Lesions in DNA

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Garrit Michael Clabaugh

December 2025

Dissertation Committee:

Dr. Yinsheng Wang, Chairperson
Dr. Ryan Jullian
Dr. Linlin Zhao

The Dissertation of Garrit Michael Clabaugh is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGMENTS

First and foremost, I would like to give thanks to my principal investigator and advisor, Dr. Yinsheng Wang, for his exceptional guidance, unending support, and insightful mentorship throughout my Ph. D studies. His dedication to our field, creativity, and scientific vision has been a constant source of inspiration, and his patience and thoughtful mentorship has left a lasting impression on me as a scientist. I am honored to have served under his guidance, and I can say that the accomplishments presented in this dissertation would not have been possible without his leadership and encouragement.

I would also like to thank my committee members, Dr. Ryan Julian and Dr. Linlin Zhao, as well as my former committee member Dr. Wenwan Zhong for their valuable input and constructive feedback throughout my journey. Their constructive feedback, especially during my second-year research exam and oral qualifying exam, helped sharpen both my scientific approach and communication skills. I am grateful for the time, effort, and patience they generously provided throughout my time at the University of California, Riverside.

I am incredibly thankful for Dr. Drew Kellum and all of my lab mates in the Wang Lab. Your camaraderie, willingness to collaborate, and helpful discussions created an environment that was stimulating and rewarding. Thank you for your friendship and walking along this path with me. I couldn’t have asked for a better team to grow alongside.

To my father, Gary Clabaugh, my best friend, Eric Matthies, and all family members, your unending support, encouragement, and belief in me have been the

cornerstone of my resilience. You all have been a constant source of motivation and strength. I could not have reached this milestone without your support and presence in my life.

Lastly, to my girlfriend, Rachel Forbes, thank you for your unwavering love, patience, and support throughout this journey. Your encouragement and understanding, even during the most stressful and uncertain moments, meant more to me than words can express. You believed in me when I doubted myself and lifted me up when things felt overwhelming. Thank you for being my rock and my biggest supporter. I am grateful to have shared this chapter with you.

ABSTRACT OF DISSERTATION

Occurrence and Biological Consequences of Alkyl Phosphotriester Lesions in DNA

by

Garrit Michael Clabaugh

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, December 2025
Dr. Yinsheng Wang, Chairperson

Owing to its limited chemical stability, the human genome is vulnerable to damage by exogenous and endogenous factors, giving rise to DNA damage. A major form of this damage is DNA alkylation, with alkyl phosphotriesters (alkyl-PTEs) being particularly significant due to their high formation frequency and resistance to repair in mammalian cells. However, the occurrence and biological consequences of these lesions remain poorly unexplored.

In Chapter 2, we describe the synthesis of a dithiane-protected pyridyloxobutyl (POB) alcohol and its coupling to a thymidine phosphoramidite, which was incorporated into oligodeoxynucleotides (ODNs) using solid-phase synthesis to construct ODNs harboring the site-specific POB-PTE lesion in both R_P and S_P configurations. These protected ODNs were subsequently converted into pyridylhydroxybutyl-PTE (PHB-PTE) ODNs and characterized by LC-MS/MS. These substrates will serve as valuable tools for

future studies aimed at investigating the impact of PHB-PTE lesions on DNA replication and transcription, as well as the cellular response to this form of DNA damage.

In Chapter 3, we constructed ODNs bearing site-specific alkyl-PTEs (Me, Et, *n*Pr, *n*Bu) lesions flanked by two thymidine residues in both R_P and S_P diastereomeric configurations to investigate their impact on DNA replication in human cells. While replicative bypass was accurate, these lesions moderately impeded DNA replication. Genetic ablation of Pol η or Pol ζ significantly reduced bypass efficiency for the bulkier *n*Pr- and *n*Bu-PTEs in both configurations, as well as the S_P diastereomer of Et-PTE. Additionally, loss of Pol κ impaired bypass of the R_P diastereomers of both *n*Pr- and *n*Bu-PTEs. Thus, our results showed that most alkyl-PTE lesions moderately hinder DNA replication in human cells, with minimal differences between the diastereomers.

In Chapter 4, we synthesized a methyl ester of carboxymethyl-PTE (CM-PTE) phosphoramidite building block of thymidine and constructed ODNs containing a site-specifically inserted CM-PTE modification. Additionally, we demonstrated that treatment of calf thymus DNA with diazoacetate induced the formation of CM-PTE adducts, where we identified a total of 16 CM-PTE lesions across all possible combinations of flanking nucleobases. These results provide evidence for DNA phosphate backbone carboxymethylation and establish a foundation for future studies aimed at elucidating the *in vivo* formation and biological consequences of CM-PTE lesions.

Together, this work advances our understanding of DNA PTE adducts by developing synthetic tools to explore the formation, replication, and biological consequences of site-specific PHB-, alkyl-, and CM-PTE lesions.

TABLE OF CONTENTS

CHAPTER 4

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

CHAPTER 1

General Introduction

This chapter provides the general background for the research projects presented in this dissertation.

DNA alkylation

Due to its limited chemical stability, genomic integrity of living cells is constantly threatened by endogenous and exogenous agents that induce many types of DNA damage, potentially leading to mutagenesis and ultimately carcinogenesis.[1–3] Among these genotoxicants are *N*-nitroso compounds (NOCs), a class of carcinogenic alkylating agents that are present both exogenously and endogenously. Lifestyle factors such as tobacco use and the consumption of cured or preserved red meat can influence NOC formation and contribute to DNA alkylation.[4,5]

DNA alkylation is a class of damage that results from the transfer of an alkyl group from an electrophilic species to nucleophilic sites on DNA, forming DNA alkylation products.[6] While most studies focus on nucleobase alkylation owing to its miscoding potential during DNA replication and promoting mutagenesis, alkylation can also occur on the non-bridging oxygen atoms of the phosphate backbone, giving rise to alkyl-phosphotriester (alkyl-PTE) lesions.[7,8] Common alkylating agents capable of inducing PTE lesions include *N*-nitroso compounds (NOCs) and diazo compounds, which are capable of generating reactive intermediates that can modify the phosphate backbone of DNA.[9,10]

Although underinvestigated, PTE lesions are recognized for their biological significance and their contribution to genomic instability, forming the central focus of this thesis.

Alkyl phosphotriester lesion formation

Alkyl-PTE lesions were first observed in the late 1940s, when researchers discovered that mustard gas, a potent alkylating agent, could form covalent adducts with nucleic acids, including modifications to the DNA phosphate backbone in the form of “secondary phosphoryl” adducts.[11] Later studies in the 1970s revealed that NOCs could alkylate the DNA phosphate backbone, inducing methyl-PTE (Me-PTE) lesions.[12] Furthermore, recent studies showed that the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and its metabolite 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) can be bioactivated to form potent alkylating agents that react with DNA to form pyridyloxobutyl (POB), pyridylhydroxybutyl (PHB), and methyl phosphate backbone adducts.[13–15] This research led to phosphotriester lesions gaining attention for their chemical stability and potential to disrupt DNA processing.

Alkylation can occur on different non-bridging oxygen atoms on the phosphate backbone. Depending on which oxygen atom is alkylated, S_P or R_P diastereomers can form due to the tetrahedral configuration of the phosphate group (Figure 1-1).[8] Every type of alkyl-PTE can be categorized into these two configurations. The alkyl group of the S_P diastereomer projects perpendicular out from the DNA double helix, whereas the R_P counterpart projects into the major groove of DNA.[8] The structural differences of the diastereomers can influence DNA-protein interactions, replication, and repair pathways.[16–18]

The extent of PTE formation is dependent on the electrophilic strength of alkylating agents and mirrors the induction profiles of other mutagenic oxygen alkylation products, such as O^6-alkyl-2'-deoxyguanosine (O^6-alkyl-dG) and O^4-alkyl-thymidine.[19,20] For example, ethylnitrosourea (ENU), a potent carcinogen, induces ethyl phosphotriesters (Et-PTEs) that account for 55% of total alkylation products in salmon sperm DNA, whereas ethylmethanesulfonate (EMS), a weaker carcinogen, constituted only 12% of products. Similarly, O^6-ethyl-2'-deoxyguanosine (O^6-Et-dG) represents 9% of ENU induced lesions, but only 2% of those were induced by EMS.[9] Thus, PTEs may serve as surrogate markers for known mutagenic lesions in DNA.

The formation of alkyl-PTEs is thought to have a non-random occurrence. Guichard *et al.*[20] developed a sensitive ^{32}P labeling method capable of detecting total methyl and ethyl PTE lesions in DNA at the femtomole level, providing a semiquantitative approximation of their distribution. Upon treating calf thymus DNA (ctDNA) with methylating or ethylating reagents, the residues flanking PTE adducts appeared random. In contrast, analysis of liver DNA from three strains of mice treated with the ethylating agent *N*-nitrosodiethylamine (NDEA) revealed that the frequency of thymidine and dG residues 5' to the ethyl PTE was significantly different from the corresponding normal nucleotide content, suggesting *in vivo* formation of alkyl-PTEs are nonrandom. In a future study, Jones *et al.*[21] found that PTE lesions formed preferentially at the 3' position of pyrimidine residues compared to purines in human B-lymphoblastoid cells, their isolated nuclei, and their purified DNA following treatment with alkylating agents of varying electrophilic character. Notably, PTE formation exhibited a nonrandom distribution in all three samples. Together, these results suggest that the induction of alkyl-PTEs is non-random *in vivo*.

Persistence and repair of alkyl-PTE lesions

Unlike nucleobase adducts that can be efficiently recognized and repaired, PTE lesions are poorly repaired and can accumulate at appreciably high levels. For example, Me- and Et-PTEs induced by *N*-methyl-*N*-nitrosourea (MNU) and ENU represented 12-17% and 55-57% of total DNA alkylation products *in vitro*, respectively[12]. Additionally, Engelse *et al.*[22] discovered that the half-lives ($t_{1/2}$) for Me- and Et-PTEs were 1 and 10-15 weeks in the kidney and lung tissues of mice treated with methylating and ethylating reagents, respectively. Due to their persistence and relatively high frequency, alkyl-PTEs are more likely to be encountered by DNA replication and transcription machineries. Furthermore, owing to their relative chemical stability under physiological conditions and their resistance to DNA repair mechanisms, PTE adducts have the potential to be used as biomarkers for measuring exposure to alkylating agents.[23,24]

As mentioned previously, the presence of a phosphate backbone modification can perturb DNA-protein interactions. For example, an isopropyl-PTE (iPr-PTE) residue positioned on either the helicase-translocating or non-translocating strand was found to inhibit members of superfamily 2 helicases involved in DNA repair pathways, such as recombinational repair and cross-link repair.[16] In contrast, replication fork helicases exhibited full tolerance to the isopropyl modification, regardless of the strand. Moreover, the RECQ1 helicase was selectively trapped by the iPr-PTE lesion when located on the helicase-translocating strand.[16] These findings highlight how the presence of alkyl-PTE can disrupt DNA repair, with important implications for genomic stability.

Alkyl-PTEs are resistant to repair and their biological consequences are poorly understood. In 2018, Wu *et al.*[25] investigated how alkyl-PTEs of varying chain lengths

affect the replicative bypass by SOS-induced translesion synthesis (TLS) polymerases in *E. coli.* The R_P diastereomer of Me-PTEs (*Rp*-Me-PTE) were more blocking to replication than the S_P counterpart (S_P-Me-PTE).[25] Notably, S_P-Me-PTE flanked by two thymidine residues induced TT→GT and TT→GC mutations at the lesion site, and the mutations were a consequence of the repair by the Ada protein.[25] Additionally, the mutation outcome was largely independent of the 5' flanking nucleoside.[17] A related repair mechanism was also found in the eukaryote *Aspergillus nidulans*, but a repair mechanism has not been found in higher eukaryotic systems.[26]

In recent studies, we investigated how alkyl-PTEs with varying chain length and stereochemical configuration affect DNA replication and transcription in human cells.[18,27] In our replication experiments, while no mutations were detected for any of the lesions in either the R_P or S_P configuration, the adducts moderately impeded replication.[27] In our transcription experiments, we found that no mutant transcripts were produced in human cells either proficient or deficient in the TLS polymerases Pol κ, Pol ι, Pol ξ and REV1 for the Me- and *n*Pr-PTE lesions in either configurations; however, both Me- and nPr-PTE lesions in the R_P configuration elicited moderate to strong transcriptional impediment.[18] In contrast, neither adducts in the S_P configuration appreciably perturbed transcriptional bypass.[18]

Pyridylhydroxybutyl phosphotriester formation from exogenous *N*-nitroso compound exposure

Exposure to exogenous NOCs, notably through tobacco use, introduces potent carcinogens such as NNK, a tobacco-specific nitrosamine classified as a group I

carcinogen by the International Agency for Research on Cancer (IARC) due to its carcinogenicity toward humans.[28] Tobacco products and tobacco smoke carry over 80 carcinogens and exclusively contain NNK, which has been linked to lung, oral cavity, and esophageal cancers in tobacco users.[29–31] In the United States, tobacco consumption is the largest source of NOCs, with estimated daily intakes reaching up to 25,000 ng.[32] Notably, the use of tobacco products remains the leading cause of preventable cancer-related deaths.[33]

In humans, NNK requires bioactivation to exert its toxic effects.[34] The primary metabolic pathway in humans involves reduction of its carbonyl group by enzymes such as carbonyl reductases, aldo-keto reductases, and 11β-hydroxysteroid dehydrogenase type 1, converting NNK to its metabolite 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanol (NNAL).[35,36] Both NNK and NNAL are subsequently oxidized primarily by cytochrome P450 2A13, and is highly expressed in respiratory tissues, particularly in the nasal mucosa and lung.[37] Hydroxymethylation at the α-methylene position of both NNK and NNAL produces a methyldiazonium ion, capable of inducing DNA methylation, while α-methyl hydroxylation produces POB and PHB diazonium ions, generating POB and PHB DNA adducts, respectively.[37] Outlined in Figure 1-2 are the metabolic activation pathways of NNK and NNAL. Alongside alkylating nucleobases, these diazonium ions have been shown to alkylate the oxygen atoms of the internucleotide phosphodiester moiety, forming PHB-PTE and methyl-PTE adducts, respectively.[15,38] The generation of these diazonium ions are responsible for the mutagenicity and carcinogenicity of NNK exposure.[39]

Occurrence, persistence, and repair of pyridylhydroxybutyl phosphotriester lesions

Pyridylhydroxybutyl PTE adducts have been quantified and characterized in the lung tissues of rats treated with NNK. In 2017, Ma *et al.*[15] Investigated the formation of PHB-PTE lesions in the lung tissues of rats chronically treated with 5 ppm NNK in their drinking water for up to 70 weeks. By using LC-MS/MS analysis, they characterized 107 structurally unique PHB-PTE lesions in the lungs of rats treated with NNK.[15] They also demonstrated that the PHB-PTE adducts accounted for 38-55 and 34-40% of all measure pyridine-containing DNA adducts in rate lung and liver, respectively.[15] Furthermore, some of these lesions were detectable for over 70 weeks and occurred at levels of 3950-8160 fmol/mg of DNA.[15]

Currently, there are no studies investigating how PHB-PTE adducts are repaired in cells. However, Wu *et al.*[40] investigated how POB-PTE adducts, which can also be formed by NNK exposure, are tolerated in *E. coli* cells proficient or deficient in the SOS-induced TLS polymerases Pol II, Pol IV, Pol V, as well as the ada protein. They found that both R_P- and S_P- diastereomers of POB-PTE flanked by two thymidine residues neither impeded replicative bypass nor induced mutation.[40] Furthermore, the depletion of TLS polymerases or the ada protein did not affect replicative bypass, suggesting these lesions are tolerated in *E. coli*.[40] Despite these findings, PHB-PTEs remain largely underexplored, and further studies are needed to elucidate their repair mechanisms and how they are tolerated in human cells.

Endogenous NOC exposure and carboxymethyl DNA adduct formation

Endogenous toxicants can give rise to DNA damage that threatens genomic stability, induces mutagenesis and drives carcinogenesis.[1] Among these toxicants are the endogenous NOCs, which are estimated to account for 45-75% of total human NOC exposure.[41] These compounds are formed *in vivo* through nitrosation reactions involving dietary precursors such as nitrates, nitrites, and secondary amines.[42] Processed meat is a major dietary source of the aforementioned compounds, all of which are precursors necessary for endogenous formation of NOCs in the stomach.[43,44] Additionally, nitrate may be reduced to nitrite by bacteria in the oral cavity, further facilitating endogenous NOC formation.[45] A widely accepted model for the development of gastric cancer implicates DNA damage resulting from intragastric nitrosation as a major factor for contributing to the early stages of the disease. In support of this claim, a study from the European Prospective investigation into Cancer and Nutrition revealed a link between elevated levels of endogenous NOCs and an increased risk of noncardia gastric cancer.[46]

A subset of endogenous NOCs can be metabolically activated to release diazoacetate, a potent alkylating agent that is capable of inducing carboxymethyl and methyl DNA modifications.[47,48] Notably, some endogenous NOCs, such as *N*-(*N*′-acetyl-L-prolyl)-*N*-nitrosoglycine (APNG) and *N*-nitrosoglycholic acid (NOGC), have been shown to undergo metabolic activation to form diazoacetate, which can decompose into carboxymethyldiazonium and methyldiazonium ions. These ions are potent alkylating agents that are capable of inducing carboxymethyl (CM), and to a lesser extent methyl DNA adducts, such as O^6-carboxymethyl-2′-deoxyguanosine (O^6-CMdG) and O^6-methyl-

2′-deoxyguanosine (O^6-MedG), respectively.[49–51] Furthermore, Gottschalg *et al.*[52] demonstrated that replication of a human p53-containing plasmid treated with diazoacetate in yeast cells produced mutation profiles that closely resembled the mutation profiles observed at non-CpG sites in gastrointestinal (GI) tract tumors. This similarity suggests that DNA carboxymethylation, but not methylation, plays a role in elevating cancer risk in GI tissues.

Recent work from our laboratory showed that upon diazoacetate treatment, carboxymethyl PTE lesions were induced in calf thymus DNA across all possible combinations of flanking nucleobases.[10] However, it remains unexplored whether carboxymethylation of the phosphate backbone contributes to adverse health effects associated with diazoacetate exposure. In contrast, the impact of carboxymethylation of nucleobases has been more thoroughly investigated. The known carboxymethylated nucleobase lesions induced by diazoacetate include *N*3-carboxymethylthymidine (*N*3-CMdT), O^4-carboxymethylthymidine (O^4-CMdT), N^4-carboxymethyl-2′-deoxycytidine (N^4-CMdC), N7-carboxymethylguanine (N7-CMGua), N^6-carboxymethyl-2′ deoxyadenosine (N^6-CMdA), and *O*6-CMdG.[51,53,54] Replication assays in HEK293T cells performed by Wu *et al.*[54] revealed that O^6-CMdG, O^4-CMdT, and N3-CMdT induce G→A, T→C, and T→A mutations, respectively, and all moderately impede DNA replication. Additionally, neither N^6-CMdA nor N^4-CMdC induced mutations or hindered DNA replication. Furthermore, genetic depletion of TLS polymerases provided information on how these lesions are tolerated. Specifically, depletion of Pol η led to decreased replicative bypass of both N^4-CMdC and N3-CMdT, Pol κ depletion reduced bypass of N^6-CMdA and O^6-CMdG, and loss of Pol ζ decreased the bypass of O^6-CMdG.

Altogether, the biological consequences of phosphotriester lesions are underexplored compared to nucleobase modifications in DNA. Considering their occurrence, persistence, and resistance to DNA repair, there is much to be learnt about how these adducts impact genomic stability. The rest of this chapter will discuss various experimental approaches that are employed in this dissertation to investigate the biological consequences of PTE lesions.

Oligodeoxynucleotides

Oligodeoxynucleotides (ODNs) are short, synthetic DNA sequences that can be designed to match any desired DNA sequence and contain site-specific modifications. Due to their structural similarity to native DNA, and their ability to be modified with DNA lesions, they can be used as tools in chemical biology. Although they are commonly used as primers or probes, ODNs can be synthesized to contain various DNA lesions, such as abasic sites, oxidative base damage and alkylated nucleosides.[55] Lesion-bearing ODNs enable researchers to study how DNA modifications impact cellular processes, such as DNA replication and transcription, in a controlled and reproducible manner, as well as to assess the frequency of their formation *in vivo* and *in vitro*.[18,40]

Automated synthesis of ODNs was a significant advancement in the field of molecular biology, biotechnology, and the development of nucleic-acid-based therapeutics. The concept of synthesizing short DNA sequences first occurred in the early 1950s, shortly after the discovery of the DNA double helix by Watson and Crick.[56] These earlier efforts relied on manual solution-phase methods involving the stepwise coupling of nucleotide

monomers, which were tedious, time-consuming, and characterized by low coupling efficiency and extensive purification steps.[56]

In the 1970s, an innovation by Letsinger and colleagues introduced the phosphite coupling method, which has been applied to the construction of ODNs, oligoribonucleotides, and nucleic acid derivatives by demonstrating that ODNs can be constructed by linking the nucleosides through the formation of phosphotriester linkages under mild conditions.[57] This involves reacting a protected nucleoside, a bifunctional phosphitylating reagent, and a second protected nucleoside, improving coupling efficiency, yield, and the overall reliability of the synthesis process. Their work reduced the time required to synthesize ODNs and the complexity of the synthetic process, laying the foundation for further developments in the ODN synthesis process; however, the stability of the reaction intermediates are susceptible to air oxidation and hydrolysis, leaving room for improvement to the method.[57]

In the early 1980s, Caruthers and Beaucage revolutionized the field of nucleic acid chemistry by introducing a new class of nucleoside phosphoramidites.[58] These phosphoramidite building blocks were superior to the previous compounds and are designed for fast and efficient nucleoside coupling under mild conditions. The phosphoramidites can be prepared following standard organic chemistry procedures and are stable under normal laboratory conditions, including both hydrolysis and air oxidation. Moreover, their chemical stability enables them to be stored as dry, stable powders, laying the foundation for the rapid and widespread adoption for their use in automated solid-phase ODN synthesis.

In December 1982, Caruthers and Beaucage's contribution culminated in the installation of the first automated ODN synthesizer, Applied Biosystems 380A.[59] This automated process allowed the rapid, efficient, and reproducible construction of custom ODNs that revolutionized the field of nucleic acids and setting the stage for modern chemical biology. The mechanism of automated ODN synthesis is based on solid-phase phosphoramidite chemistry and proceeds through an efficient and repetitive four-step cycle: deprotection, coupling, capping, and oxidation.[60] By first synthesizing a phosphoramidite bearing a PTE modification, it can be site-specifically incorporated into ODNs during solid-phase synthesis. The lesion-bearing ODNs will serve as tools for molecular studies aimed at investigating the biological consequences of PTE lesions. The remainder of this chapter will explore their applications.

Strand-specific PCR-Competitive Replication and Adduct Bypass (SSPCR-CRAB) Assay coupled with LC-MS/MS

Understanding the biological consequences of PTE lesions is important, as they can impede DNA replication and transcription in human cells.[27] However, the lack of suitable analytical methods complicates their investigation, as these lesions are resistant to many enzymes commonly used in studies due to the neutralization of the negative charge of the phosphate backbone, disrupting enzyme activity. Some examples of these enzymes are T4 PNK, nuclease P1, and alkaline phosphatase.[61–63] In this dissertation, we employed a method using HPLC and LC-MS/MS to separate and characterize PTE-containing ODNs and subsequently applied a modified CRAB assay coupled with LC-MS/MS to investigate the replicative bypass and mutagenesis of alkyl-PTE lesions in human cells.

The Competitive Replication and Adduct Bypass Assay (CRAB) was originally developed by Dr. John Essigmann and colleagues to evaluate the degree of replication impediment and mutagenicity of DNA lesions.[64] This method involves constructing M13mp7 bacterial vectors containing a synthesized lesion-bearing ODN, which is co-transfected into *E. coli* alongside a competitor bacterial vector containing an unmodified ODN serving as an internal standard. Following several rounds of replication, the progeny DNA is isolated, PCR-amplified, restriction digested and subjected to gel electrophoresis to quantify bypass efficiency and mutation frequency.

Building upon the CRAB assay, a strand-specific PCR-competitive replication and adduct bypass assay (SSPCR-CRAB) is used to assess the biological consequences of alkyl-PTE lesions in human cells.[27,65] In this assay, double-stranded shuttle vectors containing either a lesion-bearing ODN or an unmodified competitor ODN (internal standard) were constructed. The lesion/competitor ratio can be accurately normalized by ^{32}P-labeled PAGE gel. To correct for reduced replication efficiency caused by the presence of the alkyl-PTE, a C/C mismatch was incorporated two nucleotides away from the lesion site, which will also distinguish the replication products formed from the two strands.[66] After normalization, the lesion-bearing or unmodified control vectors are mixed individually with an unmodified competitor vector at specific molar ratios and co-transfected into HEK293T cells proficient or deficient in polymerases (Pol η, Pol ι, Pol κ, and Pol ζ). Relative to the control vector, the competitor vector harbored three more nucleotides between two restriction sites. After cellular replication, the products are extracted, purified, PCR-amplified, and subjected to restriction endonuclease digestion to release short DNA fragments containing the original lesion site. During amplification, the primer (P1) carried a terminal 3'-G complementary to the C/C mismatch site to selectively amplify the progeny

genomes arising from the lesion-bearing strand. Furthermore, P1 was designed to harbor a C/A mismatch three bases from its 3'-end to increase the specificity of strand-specific PCR. After amplification and restriction digestion, the replication products are resolved by PAGE to quantify bypass efficiency and mutation frequency. Bypass efficiency is calculated using the formula: Bypass efficiency (%) = (lesion signal/competitor signal)/(control signal/competitor signal) x 100%. This assay avoids the use of enzymatic steps that are inhibited by PTE modifications and instead enables direct analysis of replication products across the lesion site, making it well-suited for evaluating PTE adducts. The general workflow of this assay is shown in Figure 1-3.

Liquid chromatography-tandem mass spectrometry (LC-MS/MS) has been widely used for the detection and characterization of many DNA adducts, including PTE modifications, owing to its high sensitivity and selectivity. When coupled with the SSPCR-CRAB assay, LC-MS/MS enables accurate identification of replication products, providing insights to the mutation profiles induced by PTE adducts.[17,27,67,68] Previous studies have shown how DNA undergoes fragmentation inside a mass spectrometer, primarily through cleavage at the 3' C-O bond, generating two types of fragmentation ions: w_n and [a_n-base] (Figure 1-4).[69] Each canonical nucleobase produces a unique mass and can be determined based on the difference between consecutive fragment ions (e.g., $w_n - w_{n-1}$ or [a_{n+1}-base] - [a_n-base]. By analyzing the fragmentation pattern of the replication products, we can determine progeny DNA sequences at the lesion site. Furthermore, LC-MS/MS allows for the quantification of relative abundances of each replication product.

Synthesis of carboxymethyl-PTE ODN standards

The synthesis of lesion-bearing ODNs can be used as a tool for detecting DNA lesions induced by alkylating agents. Oligodeoxynucleotides with site-specific modifications can be used as synthetic standards to enable researchers to investigate the occurrence of DNA damage, setting the stage for assessing the *in vivo* formation and biological consequences of these adducts using LC-MS/MS.[10,15] In particular, PTE adducts are understudied and may hinder DNA-protein interactions required for cellular processes due to their steric hinderance and phosphate backbone charge neutralization. The ability to introduce these lesions at specific positions within an ODN unlocks the ability to first confirm the occurrence of PTE adducts, paving the way for future studies utilizing LC-MS/MS to investigate their impact on DNA replication and transcription under controlled conditions.

Scope of this Dissertation

As discussed previously, many studies have investigated the formation of PTEs arising from both exogenous and endogenous NOCs, suggesting their potential in serving as persistent biomarkers of exposure. However, the biological consequences of alkyl-PTEs, such as generic alkyl-, PHB-, and CM-PTEs remain poorly understood. In particular, the various functional groups and stereochemistry of these adducts can hinder cellular processes, such as DNA replication. Understanding how PTE lesions impact cellular properties will elucidate the mutagenic and cytotoxic potential of PTEs, as well as define their role in genome instability and contribution to diseases.

In this dissertation, we attempt to accomplish the following objectives: 1) To synthesize and characterize ODNs harboring site-specific PHB-PTE modifications in both R_P and S_P diastereomeric configurations. 2) To elucidate how PTE chain length and stereochemistry influence bypass by translesion synthesis (TLS) polymerases (Pol η, Pol ι, Pol κ, or Pol ζ) across methyl, ethyl, *n*-propyl, and *n*-butyl PTE lesions in both R_P and S_P configurations. 3) To synthesize and characterize CM-PTE ODNs in both R_P and S_P configurations, and to confirm the presence of these lesions in all possible flanking nucleobase sequences in calf thymus DNA.

FIGURES AND SCHEMES

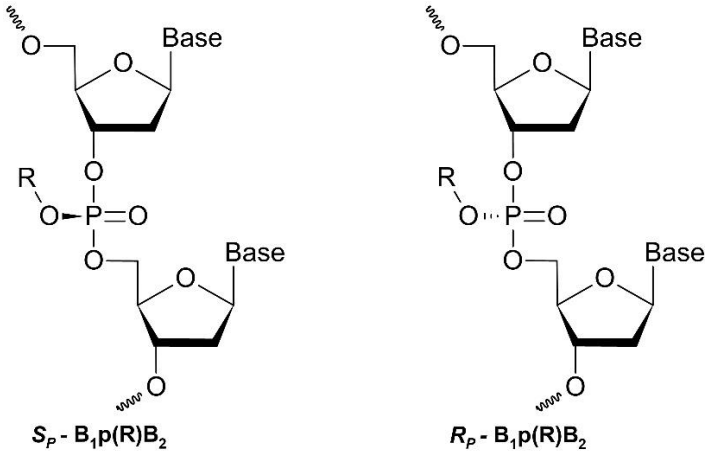


Figure 1-1. The structure of phosphotriester adducts, where 'R' designates the different alkylmodifications.

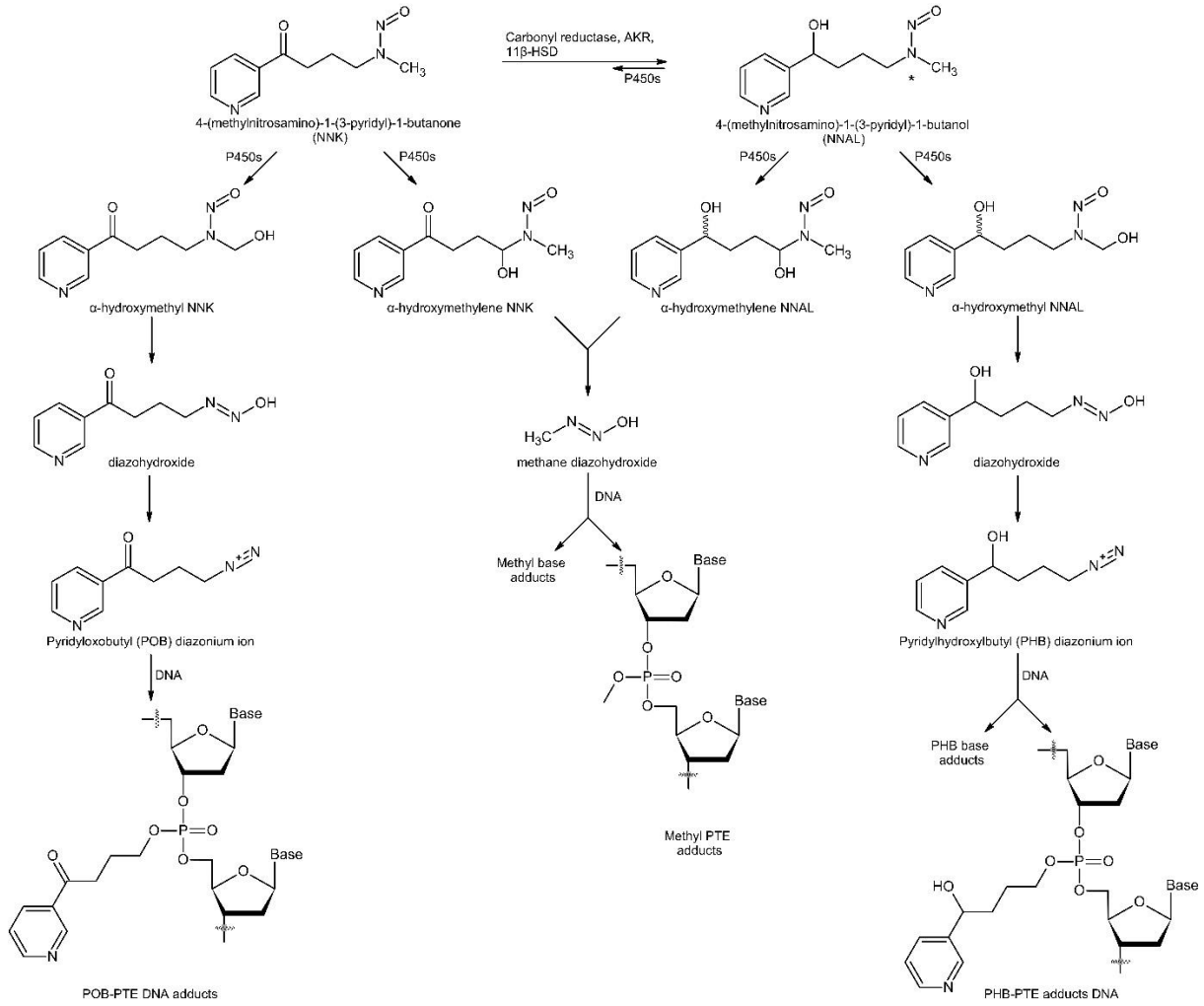


Figure 1-2. Phosphate backbone adduct formation through metabolic activation of NNK and NNAL.

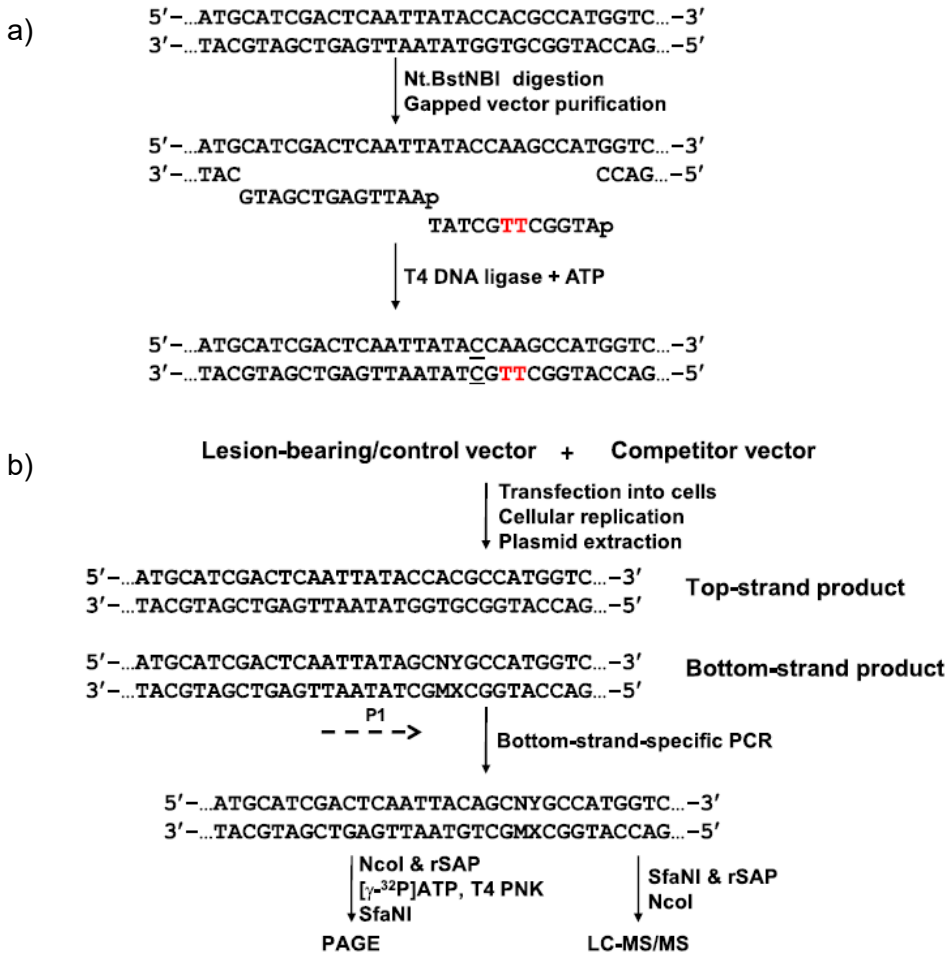


Figure 1-3. Schematic diagram showing the construction of the lesion-bearing double stranded plasmid (a) and the workflow of the SSPCR-CRAB assay (B).

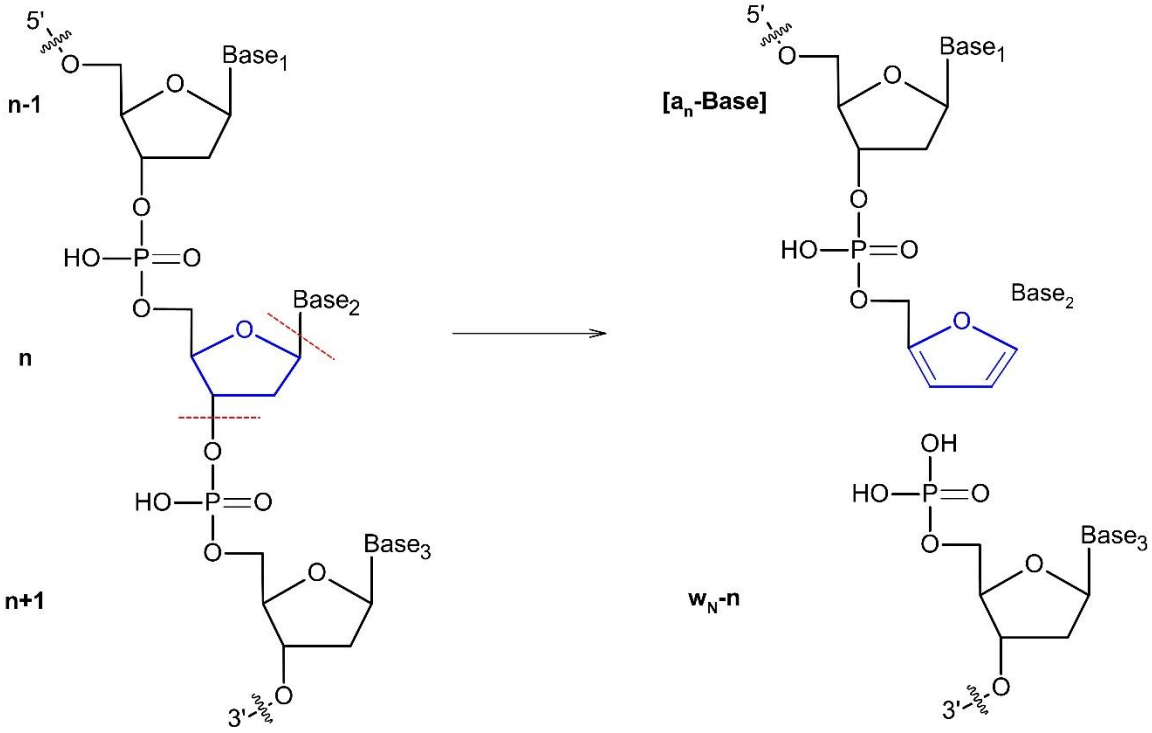


Figure 1-4. Scheme showing MS/MS fragmentation of DNA

REFERENCES

(1) Lindahl, T. Instability and Decay of the Primary Structure of DNA. *Nature* **1993**, *362* (6422), 709–715.

(2) Pegg, A. E. Methylation of the O^6 Position of Guanine in DNA Is the Most Likely Initiating Event in Carcinogenesis by Methylating Agents. *Cancer Invest.* **1984**, *2* (3), 223–231.

(3) Hoeijmakers, J. H. J. DNA Damage, Aging, and Cancer. *N. Engl. J. Med.* **2009**, *361* (15), 1475–1485.

(4) Environ and Mol Mutagen - 2004 - Cross - Meat-related Mutagens Carcinogens in the Etiology of Colorectal Cancer.Pdf.

(5) Hecht, S. S. Tobacco Carcinogens, Their Biomarkers and Tobacco-Induced Cancer. *Nat. Rev. Cancer* **2003**, *3* (10), 733–744.

(6) Shrivastav, N.; Li, D.; Essigmann, J. M. Chemical Biology of Mutagenesis and DNA Repair: Cellular Responses to DNA Alkylation. *Carcinogenesis* **2010**, *31* (1), 59–70.

(7) Drabløs, F.; Feyzi, E.; Aas, P. A.; Vaagbø, C. B.; Kavli, B.; Bratlie, M. S.; Peña-Diaz, J.; Otterlei, M.; Slupphaug, G.; Krokan, H. E. Alkylation Damage in DNA and RNA—Repair Mechanisms and Medical Significance. *DNA Repair* **2004**, *3* (11), 1389–1407.

(8) Jones, G. D. D.; Le Pla, R. C.; Farmer, P. B. Phosphotriester Adducts (PTEs): DNA's Overlooked Lesion. *Mutagenesis* **2010**, *25* (1), 3–16.

(9) Ma, B.; Villalta, P. W.; Zarth, A. T.; Kotandeniya, D.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S. Comprehensive High-Resolution Mass Spectrometric Analysis of DNA Phosphate Adducts Formed by the Tobacco-Specific Lung Carcinogen 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone. *Chem. Res. Toxicol.* **2015**, *28* (11), 2151–2159.

(10) Clabaugh, G.; Wang, Y. Formation of Carboxymethyl-Phosphotriester Adducts in DNA. *Chem. Res. Toxicol.* **2025**, *38* (5), 892–899.

(11) Elmore, D. T.; Gulland, J. M.; Jordan, D. O.; Taylor, H. F. W. The Reaction of Nucleic Acids with Mustard Gas. *Biochem. J.* **1948**, *42* (2), 308–316.

(12) Beranek, D. T. Distribution of Methyl and Ethyl Adducts Following Alkylation with Monofunctional Alkylating Agents. *Mutat. Res. Mol. Mech. Mutagen.* **1990**, *231* (1), 11–30.

(13) Haglund, J.; Henderson, A. P.; Golding, B. T.; Törnqvist, M. Evidence for Phosphate Adducts in DNA from Mice Treated with 4-(*N* -Methyl- *N* -Nitrosamino)-1-(3-Pyridyl)-1-Butanone (NNK). *Chem. Res. Toxicol.* **2002**, *15* (6), 773–779.

(14) Ma, B.; Zarth, A. T.; Carlson, E. S.; Villalta, P. W.; Stepanov, I.; Hecht, S. S. Pyridylhydroxybutyl and Pyridyloxobutyl DNA Phosphate Adduct Formation in Rats Treated Chronically with Enantiomers of the Tobacco-Specific Nitrosamine Metabolite 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanol. *Mutagenesis* **2017**, *32* (6), 561–570.

(15) Ma, B.; Zarth, A. T.; Carlson, E. S.; Villalta, P. W.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S. Identification of More than 100 Structurally Unique DNA-Phosphate Adducts Formed during Rat Lung Carcinogenesis by the Tobacco-Specific Nitrosamine 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone. *Carcinogenesis* **2018**, *39* (2), 232–241.

(16) Suhasini, A. N.; Sommers, J. A.; Yu, S.; Wu, Y.; Xu, T.; Kelman, Z.; Kaplan, D. L.; Brosh, R. M. DNA Repair and Replication Fork Helicases Are Differentially Affected by Alkyl Phosphotriester Lesion. *J. Biol. Chem.* **2012**, *287* (23), 19188–19198.

(17) Wu, J.; Yuan, J.; Price, N. E.; Wang, Y. Ada Protein– and Sequence Context–Dependent Mutagenesis of Alkyl Phosphotriester Lesions in Escherichia Coli Cells. *J. Biol. Chem.* **2020**, *295* (26), 8775–8783.

(18) Tan, Y.; Wu, J.; Clabaugh, G.; Li, L.; Du, H.; Wang, Y. Size- and Stereochemistry-Dependent Transcriptional Bypass of DNA Alkyl Phosphotriester Adducts in Mammalian Cells. *DNA* **2022**, *2* (4), 221–230.

(19) Singer, B. In Vivo Formation and Persistence of Modified Nucleosides Resulting From Alkylating Agents.

(20) Guichard, Y.; Jones, G. D. D.; Farmer, P. B. Detection of DNA Alkylphosphotriesters by 32P Postlabeling: Evidence for the Nonrandom Manifestation of Phosphotriester Lesions in Vivo.

(21) Le Pla, R. C.; Bowman, K. J.; Farmer, P. B.; Jones, G. D. D. Phosphate Alkylation in Different DNA Substrates: The Role of Local DNA Sequence and Electrophile Character in Determining the Nonrandom Nature of Phosphotriester Adduct Formation. *Chem. Res. Toxicol.* **2006**, *19* (3), 407–413.

(22) Engelse, L. D.; Menkveld, G. J.; De Brij, R.-J.; Tates, A. D. Formation and Stability of Alkylated Pyrimidines and Purines (Including Imidazole Ring-Opened 7-Alkylguanine) and Alkylphosphotriesters in Liver DNA of Adult Rats Treated with Ethylnitrosourea or Dimethylnitrosamine. *Carcinogenesis* **1986**, *7* (3), 393–403.

(23) Bannon, P.; Verly, W. Alkylation of Phosphates and Stability of Phosphate Triesters in DNA. *Eur. J. Biochem.* **1972**, *31* (1), 103–111.

(24) Shooter, K. DNA Phosphotriesters as Indicators of Cumulative Carcinogen-Induced Damage. *Nature* **1978**, *274*, 612–614.

(25) Wu, J.; Wang, P.; Wang, Y. Cytotoxic and Mutagenic Properties of Alkyl Phosphotriester Lesions in Escherichia Coli Cells. *Nucleic Acids Res.* **2018**, *46* (8), 4013–4021.

(26) Baker, S. M.; Margison', G. P.; Strike, P. Inducible Alkyltransferase DNA Repair Proteins in the Filamentous Fungus Aspergfflus Nidulans.

(27) Wu, J.; Wu, J.; Clabaugh, G.; Wang, Y. Replication Studies of Alkyl Phosphotriester Lesions in Human Cells. *Chem. Res. Toxicol.* **2024**, *37* (3), 451–454.

(28) Hecht, S. S. Research Opportunities Related to Establishing Standards for Tobacco Products Under the Family Smoking Prevention and Tobacco Control Act. *Nicotine Tob. Res.* **2012**, *14* (1), 18–28.

(29) IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. *Smokeless Tobacco ; and, Some Tobacco-Specific N-Nitrosamines*; IARC monographs on the evaluation of carcinogenic risks to humans; World Health Organization, International Agency for Research on Cancer: Lyon, France Geneva, Switzerland, 2007.

(30) Hecht, S. S. It Is Time to Regulate Carcinogenic Tobacco-Specific Nitrosamines in Cigarette Tobacco. *Cancer Prev. Res. (Phila. Pa.)* **2014**, *7* (7), 639–647.

(31) Hecht, S. S. Biochemistry, Biology, and Carcinogenicity of Tobacco-Specific *N* -Nitrosamines. *Chem. Res. Toxicol.* **1998**, *11* (6), 559–603.

(32) Gushgari, A. J.; Halden, R. U. Critical Review of Major Sources of Human Exposure to N-Nitrosamines. *Chemosphere* **2018**, *210*, 1124–1136.

(33) Islami, F.; Goding Sauer, A.; Miller, K. D.; Siegel, R. L.; Fedewa, S. A.; Jacobs, E. J.; McCullough, M. L.; Patel, A. V.; Ma, J.; Soerjomataram, I.; Flanders, W. D.; Brawley, O. W.; Gapstur, S. M.; Jemal, A. Proportion and Number of Cancer Cases and Deaths Attributable to Potentially Modifiable Risk Factors in the United States. *CA. Cancer J. Clin.* **2018**, *68* (1), 31–54.

(34) Sun, Y.; Wang, H.; Chen, H.; Zhang, S.; Li, J.; Zhang, J.; Tian, J.; Zhang, Y.; Hou, H.; Hu, Q. Nicotine Inhibits the Cytotoxicity and Genotoxicity of NNK Mediated by CYP2A13 in BEAS-2B Cells. *Molecules* **2022**, *27* (15), 4851.

(35) Richter, E.; Engl, J.; Friesenegger, S.; Tricker, A. R. Biotransformation of 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone in Lung Tissue from Mouse, Rat, Hamster, and Man. *Chem. Res. Toxicol.* **2009**, *22* (6), 1008–1017.

(36) Breyer-Pfaff, U.; Martin, H.-J.; Ernst, M.; Maser, E. ENANTIOSELECTIVITY OF CARBONYL REDUCTION OF 4-METHYLNITROSAMINO-1-(3-PYRIDYL)-1-BUTANONE BY TISSUE FRACTIONS FROM HUMAN AND RAT AND BY

ENZYMES ISOLATED FROM HUMAN LIVER. *Drug Metab. Dispos.* **2004**, *32* (9), 915–922.

(37) Jalas, J. R.; Hecht, S. S.; Murphy, S. E. Cytochrome P450 Enzymes as Catalysts of Metabolism of 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone, a Tobacco Specific Carcinogen. *Chem. Res. Toxicol.* **2005**, *18* (2), 95–110.

(38) Ma, B.; Zarth, A. T.; Carlson, E. S.; Villalta, P. W.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S. Methyl DNA Phosphate Adduct Formation in Rats Treated Chronically with 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone and Enantiomers of Its Metabolite 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanol. *Chem. Res. Toxicol.* **2018**, *31* (1), 48–57.

(39) Peterson, L. A.; Urban, A. M.; Vu, C. C.; Cummings, M. E.; Brown, L. C.; Warmka, J. K.; Li, L.; Wattenberg, E. V.; Patel, Y.; Stram, D. O.; Pegg, A. E. Role of Aldehydes in the Toxic and Mutagenic Effects of Nitrosamines. *Chem. Res. Toxicol.* **2013**, *26* (10), 1464–1473.

(40) Wu, J.; Wang, Y. Replication of Pyridyloxobutyl Phosphotriester Lesions in Cells. *Chem. Res. Toxicol.* **2020**, *33* (2), 308–311.

(41) Tricker, A. R. N-Nitroso Compounds and Man: Sources of Exposure, Endogenous Formation and Occurrence in Body Fluids. *Eur. J. Cancer Prev. Off. J. Eur. Cancer Prev. Organ. ECP* **1997**, *6* (3), 226–268.

(42) Mirvish, S. S. Formation of N-Nitroso Compounds: Chemistry, Kinetics, and in Vivo Occurrence. *Toxicol. Appl. Pharmacol.* **1975**, *31* (3), 325–351.

(43) Mirvish, S. S. Role of N-Nitroso Compounds (NOC) and N-Nitrosation in Etiology of Gastric, Esophageal, Nasopharyngeal and Bladder Cancer and Contribution to Cancer of Known Exposures to N 0 C. *Cancer Lett.* **1995**.

(44) *Ingested Nitrate and Nitrite, and Cyanobacterial Peptide Toxins*; IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, International Agency for Research on Cancer, World Health Organization, Eds.; IARC monographs on the evaluation of carcinogenic risks to humans; International Agency for Research on Cancer ; Distributed by WHO Press: Lyon, France : Geneva, 2010.

(45) Dietrich, M.; Block, G.; Pogoda, J. M.; Buffler, P.; Hecht, S.; -Martin, S. P. A Review: Dietary and Endogenously Formed N-Nitroso Compounds and Risk of Childhood Brain Tumors. *Cancer Causes Control* **2005**, *16* (6), 619–635.

(46) Jakszyn, P.; Bingham, S.; Pera, G.; Agudo, A.; Luben, R.; Welch, A.; Boeing, H.; Del Giudice, G.; Palli, D.; Saieva, C.; Krogh, V.; Sacerdote, C.; Tumino, R.; Panico, S.; Berglund, G.; Simán, H.; Hallmans, G.; Sanchez, M. J.; Larrañaga, N.; Barricarte, A.; Chirlaque, M. D.; Quirós, J. R.; Key, T. J.; Allen, N.; Lund, E.; Carneiro, F.; Linseisen, J.; Nagel, G.; Overvad, K.; Tjonneland, A.; Olsen, A.; Bueno-de-Mesquita, H. B.; Ocké, M. O.; Peeters, P. Hm.; Numans, M. E.; Clavel-

Chapelon, F.; Trichopoulou, A.; Fenger, C.; Stenling, R.; Ferrari, P.; Jenab, M.; Norat, T.; Riboli, E.; Gonzalez, C. A. Endogenous versus Exogenous Exposure to N -Nitroso Compounds and Gastric Cancer Risk in the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST) Study. *Carcinogenesis* **2006**, *27* (7), 1497–1501.

(47) Cheng, S.; Qiu, H.; Ding, W.; Kong, C.; Ma, J.; Hou, R.; Liu, C.; Ji, L. Insight into the Molecular Initiating Event of Mutagenic *N*-Nitrosamines: A Computational Study on DNA Alkylation by Their Diazonium Ions. *Environ. Toxicol. Chem.* **2025**, *44* (3), 777–785.

(48) Fahrer, J.; Christmann, M. DNA Alkylation Damage by Nitrosamines and Relevant DNA Repair Pathways. *Int. J. Mol. Sci.* **2023**, *24* (5), 4684.

(49) Harrison, K. L.; Jukes, R.; Cooper, D. P.; Shuker, D. E. G. Detection of Concomitant Formation of O^6 -Carboxymethyl- and O^6 -Methyl-2‘-Deoxyguanosine in DNA Exposed to Nitrosated Glycine Derivatives Using a Combined Immunoaffinity/HPLC Method. *Chem. Res. Toxicol.* **1999**, *12* (1), 106–111.

(50) Carboxymethylation of DNA Induced by N-Nitroso Compounds and Its Biological Implications. In *Advances in Molecular Toxicology*; Elsevier, 2011; pp 219–243.

(51) Yu, Y.; Wang, J.; Wang, P.; Wang, Y. Quantification of Azaserine-Induced Carboxymethylated and Methylated DNA Lesions in Cells by Nanoflow Liquid Chromatography-Nanoelectrospray Ionization Tandem Mass Spectrometry Coupled with the Stable Isotope-Dilution Method. *Anal. Chem.* **2016**, *88* (16), 8036–8042.

(52) Gottschalg, E.; Scott, G. B.; Burns, P. A.; Shuker, D. E. G. Potassium Diazoacetate-Induced P53 Mutations in Vitro in Relation to Formation of O6-Carboxymethyl- and O6-Methyl-2’-Deoxyguanosine DNA Adducts: Relevance for Gastrointestinal Cancer. *Carcinogenesis* **2006**, *28* (2), 356–362.

(53) Swanson, A. L.; Wang, J.; Wang, Y. In Vitro Replication Studies of Carboxymethylated DNA Lesions with *Saccharomyces Cerevisiae* Polymerase η. *Biochemistry* **2011**, *50* (35), 7666–7673.

(54) Wu, J.; Wang, P.; Li, L.; Williams, N. L.; Ji, D.; Zahurancik, W. J.; You, C.; Wang, J.; Suo, Z.; Wang, Y. Replication Studies of Carboxymethylated DNA Lesions in Human Cells. *Nucleic Acids Res.* **2017**, *45* (12), 7276–7284.

(55) Dahlmann, H. A.; Vaidyanathan, V. G.; Sturla, S. J. Investigating the Biochemical Impact of DNA Damage with Structure-Based Probes: Abasic Sites, Photodimers, Alkylation Adducts, and Oxidative Lesions. *Biochemistry* **2009**, *48* (40), 9347–9359.

(56) Michelson, A. M.; Todd, S. R. Nuckotides Part X X X I I .* Xynthesis of a Dithymidine Dinuleotide Containing a 3’ :5’-Internucleotidic Linkuge.

(57) Letsinger, R. L.; Finnan, J. L.; Heavner, G. A.; Lunsford, W. B. Nucleotide Chemistry. XX. Phosphite Coupling Procedure for Generating Internucleotide Links.

J. Am. Chem. Soc. **1975**, *97* (11), 3278–3279.

(58) Beaucage, S. L.; Caruthers, M. H. Deoxynucleoside Phosphoramidites—A New Class of Key Intermediates for Deoxypolynucleotide Synthesis. *Tetrahedron Lett.* **1981**, *22* (20), 1859–1862.

(59) Caruthers, M. H. The Chemical Synthesis of DNA/RNA: Our Gift to Science. *J. Biol. Chem.* **2013**, *288* (2), 1420–1427.

(60) Bartosik, K.; Debiec, K.; Czarnecka, A.; Sochacka, E.; Leszczynska, G. Synthesis of Nucleobase-Modified RNA Oligonucleotides by Post-Synthetic Approach. *Molecules* **2020**, *25* (15), 3344.

(61) Bowman, K. J. Evaluation of Phosphodiesterase I-Based Protocols for the Detection of Multiply Damaged Sites in DNA: The Detection of Abasic, Oxidative and Alkylative Tandem Damage in DNA Oligonucleotides. *Nucleic Acids Res.* **2001**, *29* (20), 101e–1101.

(62) Le Pla, R. C.; Guichard, Y.; Bowman, K. J.; Gaskell, M.; Farmer, P. B.; Jones, G. D. D. Further Development of^{32} P-Postlabeling for the Detection of Alkylphosphotriesters: Evidence for the Long-Term Nonrandom Persistence of Ethyl-Phosphotriester Adducts in Vivo. *Chem. Res. Toxicol.* **2004**, *17* (11), 1491–1500.

(63) Saris, C. P.; Damman, S. J.; Van Ende, A. M. C. den; Westra, J. G.; Den Engelse, L. A ^{32}P-Postlabelling Assay for the Detection of Alkylphosphotriesters in DNA. *Carcinogenesis* **1995**, *16* (7), 1543–1548.

(64) Delaney, J. C.; Essigmann, J. M. Assays for Determining Lesion Bypass Efficiency and Mutagenicity of Site-Specific DNA Lesions In Vivo. In *Methods in Enzymology*; Elsevier, 2006; Vol. 408, pp 1–15.

(65) You, C.; Swanson, A. L.; Dai, X.; Yuan, B.; Wang, J.; Wang, Y. Translesion Synthesis of 8,5′-Cyclopurine-2′-Deoxynucleosides by DNA Polymerases η, ι, and ζ. *J. Biol. Chem.* **2013**, *288* (40), 28548–28556.

(66) Levine, R. L.; Yang, I.-Y.; Hossain, M.; Pandya, G. A.; Grollman, A. P.; Moriya, M. Mutagenesis Induced by a Single 1,N6-Ethenodeoxyadenosine Adduct in Human Cells.

(67) Mcluckey, S. A.; Van Berkel, G. J.; Glish, G. L. Tandem Mass Spectrometry of Small, Multiply Charged Oligonucleotides. *J. Am. Soc. Mass Spectrom.* **1992**, *3* (1), 60–70.

(68) Wang, P.; Amato, N. J.; Zhai, Q.; Wang, Y. Cytotoxic and Mutagenic Properties of O^4-Alkylthymidine Lesions in *Escherichia Coli* Cells. *Nucleic Acids Res.* **2015**, *43* (22), 10795–10803.

(69) Habibi-Goudarzi, S.; McLuckey, S. A. Ion Trap Collisional Activation of the Deprotonated Deoxymononucleoside and Deoxydinucleoside Monophosphates. *J. Am. Soc. Mass Spectrom.* **1995**, *6* (2), 102–113.

CHAPTER 2

Syntheses and Characterizations of Pyridylhydroxybutyl Phosphotriester-containing Oligodeoxynucleotides

INTRODUCTION

Due to its limited chemical stability, DNA is vulnerable to damage by endogenous and exogenous genotoxins, resulting in the induction of various DNA adducts, potentially inducing mutagenesis and ultimately carcinogenesis if they are not efficiently repaired.[1] Tobacco products and tobacco smoke contain over 80 compounds categorized as carcinogens by the International Agency for Research on Cancer (IARC), and consumption of tobacco products remains the leading cause of preventable cancer deaths.[2,3] Among these carcinogens, the tobacco specific nitrosamines 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK) and *N*'-nitrosonornicotine (NNN) are categorized as group 1 carcinogens by the IARC and have been linked to lung, oral cavity, and esophageal cancers in tobacco users.[4]

All tobacco products contain the tobacco-specific nitrosamine NNK, a potent pulmonary carcinogen that is capable of inducing lung tumors in multiple mammalian species, regardless of the route of administration.[4,5,6] For example, male F-344 rats exposed to 5 ppm NNK in their drinking water over a 90-week period developed lung tumors, with metastases to the pancreas also being observed.[7] Similarly, A/J mice injected

intraperitoneally with 7.5 μmol NNK induced lung adenomas, which was notably exacerbated by concurrent nasal exposure to acetaldehyde, formaldehyde, and carbon dioxide.[8]

NNK is metabolized into 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanol (NNAL) in both rodents and humans, and is also a potent pulmonary carcinogen, detectable in the urine of humans who use tobacco products.[9] During this process, NNK is reduced to produce two enantiomers of NNAL: (*R*)-NNAL and (*S*)-NNAL, with proportions of each depending on the stereoselectivity of reducing enzymes.[10] In 2023, Xing *et al.* [11] found that intraperitoneal exposure to (*S*)-NNAL led to robust tumor development in A/J mice, with all rodents developing lung tumors of more than four tumors per mouse, which mirrored the results found from NNK exposure. In contrast, exposure to (*R*)-NNAL resulted in lower tumor frequencies, with fewer than half of the mice developing tumors at an average of only one tumor per rodent. These findings suggest that the carcinogenicity of the two enantiomers is different due to distinct adductomic profiles.

The carcinogenicity of NNK is believed to result from its ability to form DNA adducts and promote mutagenesis.[12] As mentioned previously, NNK is enzymatically reduced to produce (*R*)-NNAL and (*S*)-NNAL, which can be further oxidized by cytochrome P450 enzymes to ultimately produce potent alkylating agents that form pyridylhydroxybutyl (PHB) DNA adducts (Scheme 1).[9] In an earlier study by Hecht *et al.,*[13] F-344 mice were treated with up to 10 ppm of NNK, (*R*)- and (*S*)-NNAL in drinking for up to 20 weeks. They observed that (*S*)-NNAL and NNK induced similar levels of O^2-pyridylhydroxybutyl-thymidine (O^2-PHB-dT), O^6-PHB-2'-deoxyguanosine (O^6-PHB-dG), and N7-PHB-guanine (N7-PHB-Gua) in lung tissues, with O^2-PHB-dT being the most abundant through 16 weeks of treatment.[13] The similarity in adductomic profiles between (*S*)-NNAL and NNK is

likely due to their ability to be interconverted.[14] In contrast, (*R*)-NNAL produced significantly greater adduct levels than both NNK and (*S*)-NNAL, suggesting cytochrome P450 enzymes efficiently catalyze the metabolic activation of (*R*)-NNAL. However, it was less effective at inducing lung tumors, likely due to its inability to be converted back to NNK.[13]

Alkylating agents can react with the non-carbon-bound oxygen atoms of the DNA internucleotide phosphodiester moiety, forming phosphotriester (PTE) adducts.[15] Due to the tetrahedral coordination around the phosphate center, PTE lesions can occur in the R_P or S_P configuration. Alkyl-PTEs in the S_p configuration projects outwards, perpendicular to the DNA double helix, while the R_P counterpart projects into the major groove of DNA.[15] Upon alkylation, the negative charge on the phosphate group is neutralized, perturbing DNA-protein interactions.[16,17,18] Additionally, an isopropyl-PTE modification can impede superfamily 2 helicases from unwinding DNA substrates; however, replicative helicases were fully tolerant of the modification.[19]

Alkyl-PTEs can occur at high frequencies and can persist in DNA. Methyl- and ethyl-PTEs (Me- and Et-PTEs) comprise of 12-17% and 55-57% of total DNA alkylation products following treatment with methylating and ethylating reagents, respectively.[20] Den Engelse *et al.*[21] reported that the half-lives ($t_{1/2}$) of methyl- and ethyl-phosphotriester (Me- and Et-PTE) lesions in the kidney and lung tissues of mice exposed to methylating and ethylating agents were approximately 1 week and 10-15 weeks, respectively. Moreover, Ma *et al.*[22] uncovered the distribution of pyridylhydroxybutyl (PHB) adducts in rats treated with NNK and discovered that PHB-PTEs accounted for 38-55% and 34-40% of total detectable pyridine-containing DNA adducts in rat lung and liver, respectively. Similarly, pyridyloxobutyl (POB)-PTEs represented 32-48% and 40-44% of total pyridine DNA

adducts in the lung and liver of rats, respectively.[22] Additionally, POB- and PHB-PTE adducts were detected in rat lung and liver tissues 70 weeks following chronic exposure to NNK,[22] indicating inefficient repair pathways, suggesting that they are more likely to be encountered by DNA replication and transcription machineries. In this vein, studies investigating the biological consequences of PHB-PTEs are limited and there is much to learn about how they impact these processes.

The objectives of the research described in this chapter are to synthesize a dithiane-protected POB alcohol and to couple it to a thymidine phosphoramidite for use in solid-phase synthesis to construct oligodeoxynucleotides (ODNs) harboring site-specifically inserted PHB-PTE modifications. These modified ODNs will be utilized for investigating in the future, the effects of PHB lesions on DNA replication and transcription, and to elucidate cellular responses to the formation of these DNA lesions.

MATERIALS AND METHODS

All chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) unless otherwise specified. Common reagents and phosphoramidite building blocks for solid-phase DNA synthesis were purchased from Glen Research (Sterling, VA, USA). Oakwood Products, Inc. (West Columbia, SC, USA) supplied 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP).

Synthesis of 2-(3-pyridyl)-1,3-dithiane (1)

A solution of 3-pyridinecarboxaldehyde (53.3 mmol) in 25 mL anhydrous tetrahydrofuran (THF) was added to a round bottom flask under argon atmosphere. To this solution was added 1,3-propanedithiol (1.3 eq.) and boron trifluoride etherate (0.33 eq.) sequentially. The reaction was refluxed for 3 days and quenched with 20 mL of water. The organic layer was transferred to a separatory funnel and the aqueous phase was adjusted to pH 4.0 using 1- M HCl, which was extracted three times with ethyl acetate. The organic layers were pooled and washed with brine, dried over anhydrous sodium sulfate, and concentrated under reduced pressure. The crude product was purified by flash column chromatography using a mobile phase consisting of 2:1 *n*-hexane:ethylacetate to yield compound **1** (56% yield).

^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ 8.71 – 8.64 (m, 1H), 8.53 (dd, *J* = 4.8, 1.6 Hz, 1H), 7.82 (ddd, *J* = 7.9, 2.7, 1.4 Hz, 1H), 7.32 – 7.24 (m, 1H), 5.17 (s, 1H), 3.14 – 3.00 (m, 2H), 2.92 (ddd, *J* = 14.6, 4.4, 3.0 Hz, 2H), 2.18 (dtt, *J* = 14.2, 4.7, 2.5 Hz, 1H), 2.05 – 1.83 (m, 1H). ESI-MS: $[M + H]^+$ calc *m/z* 198.3, found 198.3.

Synthesis of 3-[2-[2-(1,3-dioxolan-2-yl)ethyl]-1,3-dithian-2-yl]pyidine (2)

Compound **1** (8.68 mmol) was dissolved in 30 mL anhydrous THF under an argon atmosphere in an oven-dried round bottom flask. The solution was cooled to -78 °C and *n*-butyllithium (2.5 M in hexanes, 1.3 eq) was slowly added dropwise to the reaction mixture and was stirred at this temperature for 1 hr. In a separate oven-dried vessel, 2-(2-bromoethyl)-1,3-dioxolane (1.3 eq) was added to 4 mL anhydrous THF and the mixture was added dropwise. The reaction mixture was gradually warmed to room

temperature and stirred overnight. The reaction was quenched with 15 mL of water and the aqueous layer was extracted with ethyl acetate three times. The organic layers were pooled, dried over anhydrous sodium sulfate, filtered, and concentrated under reduced pressure. The crude product was purified by flash column chromatography with a mobile phase of 1:1 *n*-hexane/ethylacetate to yield compound **2** (57% yield).

^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ 9.12 (d, *J* = 2.4 Hz, 1H), 8.52 – 8.48 (m, 1H), 8.21 (dt, *J* = 8.1, 1.7 Hz, 1H), 7.31 (dd, *J* = 8.2, 4.8 Hz, 1H), 4.77 (t, *J* = 4.4 Hz, 1H), 3.90 – 3.75 (m, 4H), 2.67 (ddt, *J* = 24.1, 14.4, 4.3 Hz, 4H), 2.17 – 2.11 (m, 2H), 1.94 (tt, *J* = 9.8, 4.2 Hz, 2H), 1.64 (dt, *J* = 8.6, 4.3 Hz, 2H). ESI-MS: $[M + H]^+$ calc *m/z* 298.4, found 298.4.

Synthesis of 3-(2-pyridin-3-yl-1,3-dithiane-2-yl)propanal (3)

Compound **2** (5.04 mmol) was dissolved in 5 mL dichloromethane (DCM) in a round bottom flask. Fifty mL of a 0.4 M oxalic acid solution was added and the reaction was stirred at room temperature for 3 days. The pH of the aqueous layer was adjusted to basic, and was extracted with ethyl acetate 3 times. The organic layers were pooled, dried over anhydrous sodium sulfate, and concentrated under reduced pressure to yield compouund **3** (55% yield).

^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ 9.68 (s, 1H), 9.30 (d, *J* = 2.1 Hz, 1H), 8.98 (dd, *J* = 8.4, 2.0 Hz, 1H), 8.75 (d, *J* = 5.5 Hz, 1H), 7.98 (dd, *J* = 8.3, 5.5 Hz, 1H), 2.78 (dt, *J* = 14.8, 3.9 Hz, 3H), 2.70 (t, *J* = 7.3 Hz, 2H), 2.52 (ddd, *J* = 14.6, 11.8, 2.8 Hz, 3H), 2.32 (t, *J* = 7.3 Hz, 2H). ESI-MS: $[M + H]^+$ calc *m/z* 254.4, found 254.2.

Synthesis of 3-(2-pyridin-3-yl-1,3-dithiane-2-yl)propanol (4)

Crude compound **3** (4.28 mmol) was transferred to an oven-dried round bottom flask and dissolved in 10 mL THF under an argon atmosphere. The temperature was reduced to 0°C using a ice bath and lithium aluminum hydride (1.5 eq) was added dropwise to the reaction vessel, which was stirred at this temperature for 1 hr. The reaction was then quenched with sequential addition of 0.6 mL water, 0.6 mL 15% NaOH in water, and 1.6 mL water. The resulting precipitate was filtered, extracted with ethyl acetate, and dried over anhydrous sodium sulfate. The pooled organic layers were concentrated under reduced pressure and purified using flash column chromatography using ethyl acetate as the mobile phase to yield compound **4** (65% yield).

^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ 9.10 (dd, *J* = 4.7, 1.6 Hz, 1H), 8.46 (dd, *J* = 4.7, 1.6 Hz, 1H), 8.21 (ddd, *J* = 8.1, 2.5, 1.6 Hz, 1H), 7.30 (ddd, *J* = 8.1, 4.7, 0.8 Hz, 1H), 3.52 (t, *J* = 6.3 Hz, 2H), 2.76 – 2.55 (m, 5H), 2.16 (s, 1H), 2.14 – 2.05 (m, 2H), 2.01 – 1.88 (m, 2H), 1.63 – 1.43 (m, 2H). ESI-MS: $[M + H]^+$ calc *m/z* 256.4, found 256.4.

The NMR and ESI mass spectra of the above-mentioned products are shown in figures 2.1-2.8.

Synthesis of dithiane-protected pyridyloxobutyl phosphotriester of dT phosphoramidite building block (6)

Compound **4** (0.0980 mmol) was dissolved in 5 mL anhydrous DCM in an oven-dried round bottom flask at room temperature under an argon atmosphere. A separate solution of **5** (1.3 eq) and tetrazole in acetonitrile (1.3 eq) was prepared under argon and added dropwise to the reaction vessel and stirred at room temperature for 3 hr. The reaction

mixture was concentrated under reduced pressure to obtain the crude product and was purified by flash column chromatography with a mobile phase of *n*-hexane:ethylacetate:triethylamine 49:49:2 to yield compound **6**.

ESI-MS: $[M + H]^+$ calc *m/z* 930.1, found 930.3

The ESI mass spectrum of the the lesion-bearing phosphoramidite is shown in Figure 2.9.

ODN synthesis

The dithiane-protected 12-mer ODN **(7)** was synthesized on a Beckman Oligo 1000M DNA synthesizer (Fullerton, CA, USA) at a 1-μmol scale, where "X" represents the dithiane-protected POB group. The dithiane-protected pyridyloxobutyl phosphotriester phosphoramidite (**5**) was dissolved in anhydrous acetonitrile at a concentration of 0.1 mM. Commercially available ultramild phosphoramidites for unmodified nucleosides were purchased from Glen Research (Sterling, VA, USA) and used for ODN synthesis following standard protocols. The ODNs were cleaved from controlled pore glass (CPG) support and deprotected following standard ultramild deprotection protocol. The aqueous solvent was removed using a Speed-vac and subsequently dissolved in Milli-Q water for HPLC purification.

HPLC purification of ODNs

HPLC purification of ODNs was carried out using an Agilent 1100 system equipped with a BDS Hypersil C18 column (4.6 x 250 mm, 5 μm in particle size and 100 Å in pore size, Thermo Fisher Scientific, San Jose, CA). The mobile phases included 50 mM

triethylammonium acetate (pH 6.8, solution A) and a 70:30 (v/v) mixture of solution A and acetonitrile (solution B). A gradient of 5-20% B in 5 min, 20-35% B in 15 min, 35-45% B in 30 min, 45-75% B in 25 min was employed at a flow rate of 0.5 mL/min. The HPLC traces for the purification of the 12-mer lesion-bearing ODNs are shown in Figure 2.10.

After HPLC purification, dithiane deprotection of **7** was achieved by treatment with *N*-chlorosuccinimide (5 eq.) in a 1:1 mixture of acetonitrile and water (v/v) at room temperature for 30 mins to yield compound **8**. The POB-containing ODN was desalted by a NAP-5 column, dried by SpeedVac, reconstituted in Milli-Q water and subsequently purified using the previously described HPLC mobile phases and gradient. Reduction to PHB-PTE ODN **(9)** was achieved by treatment with sodium borohydride (10 eq.) in water at 0 °C for 1 hr. The reaction mixture was again desalted, dried, and purified by HPLC using the aforementioned mobile phases and gradient. Representative electrospray ionization-mass spectra (ESI-MS) and tandem MS (MS/MS) for the lesion-bearing ODNs are shown in Figure 2.11-2.13.

Mass spectrometry (MS) and NMR

Electrospray ionization (ESI) MS analyses of the synthesized products were performed on an LCQ Deca XP ion-trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). A 1:1 (v/v) mixture of acetonitrile and water was used as the electrospray solvent. The heated capillary temperature was maintained at 275°C, and the spray voltage was 3.0 kV. ^{1}H spectra were recorded on a Bruker Avance 600 spectrometer.

LC-MS/MS

All ODNs were characterized by LC-MS/MS using an Agilent 1200 capillary HPLC pump (Agilent Technologies) coupled to an LTQ linear ion-trap mass spectrometer (Thermo Fisher Scientific). Separation was performed on a Zorbax SB C18 column (0.5 × 150 mm, 5 μm particle size, 80 Å pore size, Agilent Technologies). The mobile phases consisted of HFIP (Solution A, pH 7.0) and methanol (solution B). A linear gradient of 5% B over 5 min, 5-70% B over 25 min, 70-90% B over 1 min at a flow rate of 8 μL/min was used. The ESI spray voltage was 3.0 kV and the ion transfer tube was held at a temperature of 275°C.

RESULTS

The biological consequences of PHB-PTE adducts remain largely unexplored. The objective of this study was to synthesize ODNs bearing a site-specific PHB-PTE modification to enable future investigations into the biological consequences of the lesion. We reason that the successful construction of the PHB-PTE-containing ODN will serve as a molecular tool for elucidating the impact of this adduct on cellular processes and for understanding how cells respond to the damage.

We synthesized the dithiane-protected POB alcohol, 3-(2-pyridin-3-yl-1,3-dithiane-2-yl)propanol **(4)**, based on the Corey-Seeback reaction strategy (scheme 2).[23] The first step involves a condensation reaction between 3-pyridinecarboxaldehyde and 1,3-propanedithiol, in the presence of the Lewis acid catalyst boron trifluoride diethyl etherate ($BF_3 \cdot Et_2O$). In this reaction, 3-pyridinecarboxaldehyde is an electrophilic carbonyl substrate, while 1,3-propanedithiol condenses with the aldehyde group to form a six-membered 1,3-dithiane ring, serving as a protecting group for the carbonyl functionality in subsequent synthetic steps. The Lewis acid catalyst activates the carbonyl group,

facilitating nucleophilic attack by 1,3-propanedithiol. Upon the formation of compound **1**, the acidity of the methine proton adjacent to the dithiane ring is greatly increased, enabling its deprotonation and subsequent carbon–carbon bond formation via a nucleophilic substitution reaction.

The second step proceeds through a lithiation reaction, where the methine proton of compound **1** (2-(3-pyridyl)-1,3-dithiane) is deprotonated by n-butyllithium (*n*-BuLi). The low temperature allows *n*-BuLi to act as a strong base, selectively abstracting the acidic proton adjacent to the dithiane ring to generate and stabilize a nucleophilic carbanion, which undergoes nucleophilic substitution with 2-(2-bromoethyl)-1,3-dioxlane, forming a new carbon-carbon bond between the lithiated dithiane and the bromoethyl-dioxolane moiety. This step yields a key intermediate bearing both dithiane and acetal protection groups. The bifunctional product is designed to allow for subsequent deprotection and functionalization steps toward the synthesis of the target POB alcohol to construct PHB-PTE ODNs.

The final two steps involve deprotection of the acetal group to yield an aldehyde, which is followed by reduction to the dithiane-protected POB alcohol for the synthesis of a phosphoramidite for solid-phase ODN synthesis (**6**). Acetal deprotection of compound **3** was achieved using 0.4 M aqueous solution of oxalic acid to yield the corresponding aldehyde intermediate. Oxalic acid is a water-soluble organic acid that is known to efficiently catalyze the hydrolysis of acetal protection groups under aqueous conditions, which minimizes the formation of side products and protecting the acid-sensitive dithiane protection group.[24] The oxygen atom of the acetal group is protonated, causing the cleavage of the dioxolane ring to generate the aldehyde functional group. Subsequent treatment with sodium borohydride reduces the aldehyde to a primary alcohol owing to its

chemoselectivity for aldehyde and ketone groups under mild conditions.[25] This two-step process enables the production of the dithiane-protected POB alcohol in high yield to be used to couple to a dT phosphoramidite for ODN synthesis. The structures of compounds **1-5** were supported by ESI-MS and ^{1}H-NMR analyses (Figures 2.1-2.8).

We next synthesized the dithiane-protected POB phosphotriester of a dT phosphoramidite by first preparing a dT phosphoramidite (**5**, Scheme 3). After purification by flash column chromatography, Compound **4** was coupled to compound **5** using the activator 1-H tetrazole, activating the phosphorus and facilitating nucleophilic attack by the dithiane-protected POB alcohol to yield compound **6**, which was subsequently purified and incorporated in a 12-mer ODN via solid-phase synthesis.

Construction of PHB-PTE ODN and its characterizations

To investigate the biological consequences of PHB-PTE modifications in future studies, we constructed ODNs containing a site-specifically inserted PHB-PTE modification (Scheme 4). The lesion-bearing phosphoramidite building block was synthesized by using the aforementioned procedures and incorporated into a 12-mer sequence 5'-ATGGCT(X)TGCTAT-3', where X denotes the dithiane-protected POB group. The measured *m/z* values for the doubly, triply, and quadruply deprotonated ions were *m/z* 1943.7, 1295.4, and 970.9, respectively, and are consistent with their calculated ones (Figure 2.11). Furthermore, MS^{2} of the $[M - 3H]^{3-}$ ion (*m/z* 1295.4) showed the formation of [a_n-Base] and their complementary w_n ions, resulting from the cleavages of the *N*-glycosidic and 3′-C-O bonds of the same nucleoside and bear the 5′ and 3′ termini of the ODN, respectively.[26] For example, [a_3-G], [a_4-G], [a_5-C], [a_8-G], [a_9-C], [a_{11}-A], w_2, w_3, w_4, w_8, w_9, and w_{11} ions were detected. Fragment ions bearing the modification, i.e., [a_8-G],

[a_9-C], [a_{11}-A], w_8, w_9, and w_{11}, exhibited a 239 Da higher in mass than the corresponding fragment ions from the unmodified ODN, which is consistent with the site of a dithiane-protected POB modification in the ODN (Figure 2.11). The S_P- and R_P- ODN diastereomers were purified and their corresponding fractions were collected by HPLC for removal of the dithiane protection group (Figure 2.10).

The collected fractions of the S_P- and R_P- dithiane-protected ODNs were collected for the removal of the dithiane protection groups to yield the corresponding POB-PTE ODN 5'-ATGGCT(POB)TGCTAT-3' (Figure 2.10). Each fraction was treated with 5 molar equivalents of NCS in a 1:1 mixture of acetonitrile and water at room temperature for 30 minutes, resulting in the removal of the dithiane protection group, yielding the POB-PTE ODNs, which is supported by LC-MS/MS analysis. The ultrazoom scan MS showed a triply charged ion with a monoisotopic *m/z* of 1264.5, exhibiting a mass decrease of 90 Da, consistent with the removal of the dithiane protection group to produce a POB ketone (Figure 2.12). Additionally, the fragment ions detected in the ms^2 of the $[M - 3H]^{3-}$ ion bearing the modification also showed a 90 Da decrease in mass relative to the fragment ions from the protected ODN, further confirming successful deprotection. The R_{P-} and S_{P-} POB-PTE ODNs were subsequently desalted and purified by HPLC to prepare for the final reduction step.

Treatment with 10 molar equivalents of sodium borohydride in water at 0 °C for 1 hour reduced POB to PHB, yielding the PHB-PTE ODN 5'-ATGGCT(PHB)TGCTAT-3' (Scheme 4). After HPLC purification (Figure 2.10), ultrazoom scan MS showed a triply charged ion with a monoisotopic *m/z* of 1265.3, exhibiting of mass increase of 2 Da with respect to the POB-PTE ODN (Figure 2.13). Additionally, fragment ions from the cleavage of $[M - 3H]^{3-}$ ion (*m/z* 1265.8) bearing the PHB modification also exhibit a mass increase

of 2 Da, e.g. [a_8-G], [a_9-C], [a_{11}-A], w_7, w_8, w_9, and w_{11}, confirming the identity of the PHB-PTE ODN.

DISCUSSION

Constructing lesion-bearing ODNs unlocks the capacity to understand how DNA lesions affect cellular processes such as replication, transcription, and how they are repaired in cells. Furthermore, they can be used in untargeted proteomics to understand the cellular response upon damage formation. Outlined in the chapter were the steps to synthesize an ODN harboring a site-specifically inserted PHB-PTE lesion to use as a tool for future experiments.

The synthesis of ODNs bearing alkyl-PTE lesions have previously been used as a tool to elucidate how these adducts impact DNA replication. For example, Wu *et al.*[27] constructed ODNs harboring alkyl-PTE modifications with varying alkyl chain lengths (Me, Et, *n*Pr, and *n*Bu), which were separately ligated into replicative plasmids and transfected into *Escherichia coli* (*E. coli*) strains either proficient or deficient in translesion synthesis (TLS) polymerases Pol II, Pol IV, and Pol V, either alone or all three in combination. Using the Competitive Replication and Adduct Bypass (CRAB) assay,[28] it was observed that Me-PTEs in the S_P configuration (S_P-Me-PTE) showed lower levels of replication impediment in *Escherichia coli* than the R_P diastereomer. Notably, the S_P-Me-PTE adduct flanked by two thymidine residues induced TT→GT and TT→GC mutations at the dinucleotide site, and the replication products was largely independent of the flanking 5' nucleoside.[29] Overall, alkyl-PTEs in the S_P configuration were less blocking than those in the R_P configuration, and no other lesions induced detectable mutations. In addition, Wu *et al.*[30] synthesized ODNs containing a POB-PTE modification flanked by two thymidine residues

and found that while the R_P-POB-PTE adducts slightly perturbed replication bypass in *E. coli.*, neither diastereomer led to mutations.

We recently investigated how the alkyl-PTEs (Me, Et, nPr, nBu) affect DNA replication in human cells.[31] Oligodeoxynucleotides bearing PTE modifications were synthesized and separately ligated into double-stranded replicative plasmids, which were subsequently transfected into human cells proficient or deficient in TLS polymerases Pol η, ι, κ, or ζ.[31] While no mutations were observed for any of the lesions in either configuration, the adducts moderately inhibited replication.[31] In a separate study, we synthesized alkyl-PTE ODNs to investigate how alkyl group size and stereochemical configurations impact transcription in human cells using the Competitive Transcription and Adduct Bypass assay (CTAB) assay.[32] Alkyl-PTE (Me, nPr) ODNs were synthesized and separately ligated into nonreplicative plasmids and subsequently ligated into human cells either proficient or deficient in the TLS polymerases Pol κ, Pol ι, Pol ξ and REV1. We found that no mutant transcripts were produced for any of the adducts and both Me- and nPr-PTE lesions in the R_P configuration elicited moderate to strong transcriptional impediment.[32] In contrast, neither adducts in the S_P configuration appreciably perturbed transcriptional bypass.[32]

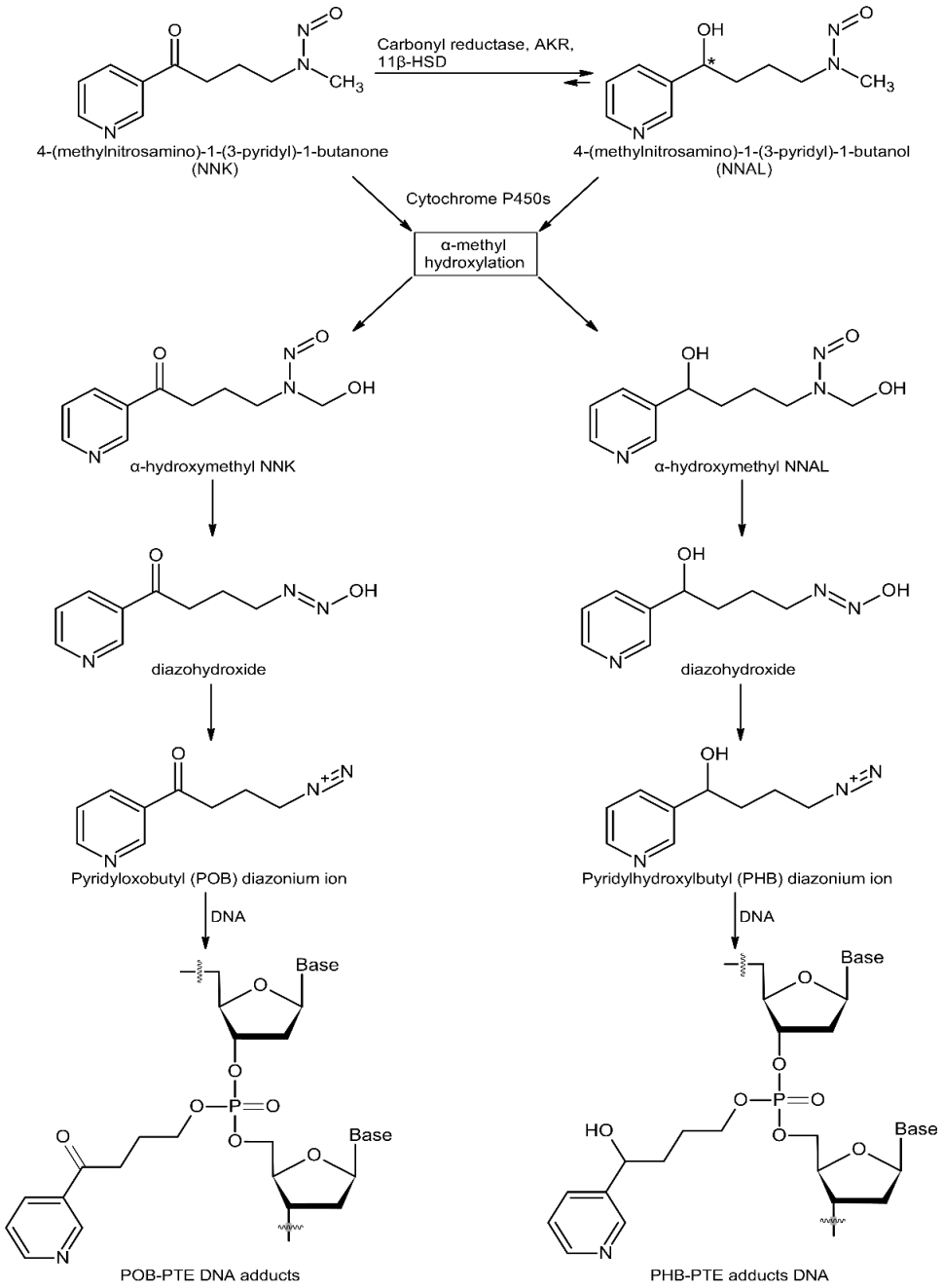
The construction of lesion-bearing ODNs also enables untargeted proteomic approaches to provide insights into the cellular responses to DNA damage and the mechanisms involved in their repair. Zhao *et al.*[33] recently constructed ODNs harboring the N^2-*n*Bu-2'-deoxyguanosine (N^2-*n*Bu-dG) lesion and used them as a biotinylated duplex DNA photoaffinity probe to identify nuclear proteins that bind to the modification and assisting in its repair. They found that high-mobility group protein B3 (HMGB3) and SUB1 bind to N^2-*n*Bu-dG, promoting its repair.[33] Additionally, they also discovered that HMGB3

and SUB1 protects cells against N^2-benzo[*a*]pyrene-7,8-diol-9,10-epoxide (BPDE), which elicits the formation of N^2-BPDE-dG, the minor grove lesion.[33]

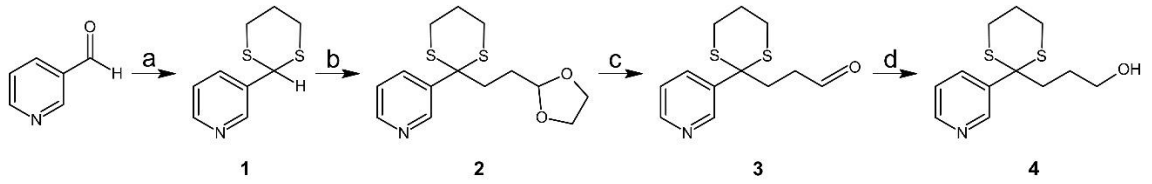
The construction of the dithiane-protected POB alcohol can also provide opportunities to investigate the mutations induced by PHB nucleobase lesions. Currently, the mutagenicity of these lesions are under-investigated. One of the few PHB-modified nucleobase lesions studied is O^6-PHB-2'-deoxyguanosine (O^6-PHB-dG), which can induce G→A and G→T mutations in both *E. coli* and human cells.[34,35] While specific mutation data are lacking for other PHB nucleobase adducts, such as O^2-PHB-thymidine (O^2-PHB-dT), studies on structurally similar POB adducts have shown that O^2-POB-dT is mutagenic, inducing T→A and T→G in *E. Coli* cells and T→A in human cells.[36,37] These findings may suggest that O^2-PHB-dT may exert similar effects on replication fidelity and warrants further investigation.

Together, we successfully synthesized a dithiane-protected POB-PTE phosphoramidite of dT and incorporated it into ODNs, which were subsequently converted to PHB-PTE ODNs, and the site-specific incorporation of the lesion was confirmed by LC-MS/MS analysis. It remains largely unexplored how cells repair the PHB-PTE lesion and respond to its formation. The synthesis of a dithiane-protected POB alcohol and the construction of ODNs bearing the PHB-PTE modification provided valuable substrates for future studies aimed at elucidating the mechanisms by which cells recognize and repair the lesion, as well as the biological consequences of their formation.

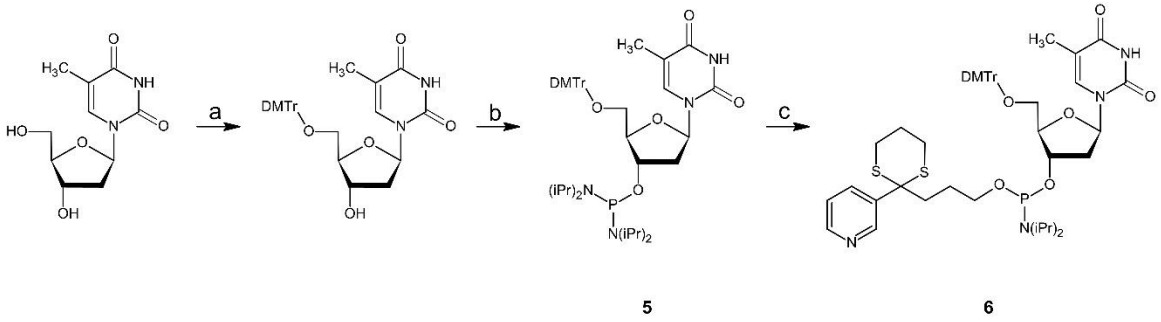
FIGURES AND SCHEMES



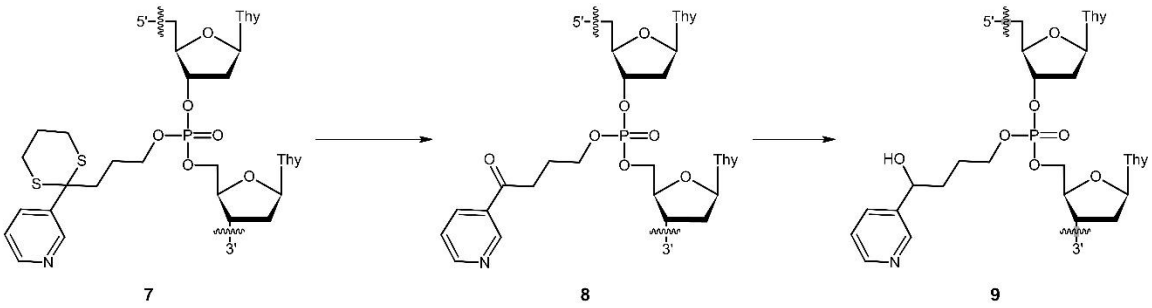
Scheme 2-1. Metabolic activation pathways for 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanone (NNK) and 4-(methylnitrosoamino)-1-(3-pyridyl)-1-butanol (NNAL), leading to pyridyloxobutyl- and pyridylhydroxybutyl phosphotriester (POB-PTE, PHB-PTE) DNA adducts.



Scheme 2-2. Synthesis of dithiane-protected POB alcohol. *A-D,* reagents and conditions. a, 1,3-propanedithiol, boron trifluoride etherate, anhydrous THF, reflux, 72 h. b, 2-(2-bromoethyl)-1,3-dioxolane, *n*-butyllithium, anhydrous THF, -78 °C, 12 h. c, oxalic acid, DCM, water, rt, 5 days. d, lithium aluminum hydride, anhydrous THF, 0 °C, 1 h.



Scheme 2-3. Synthesis of dithiane-protected POB-PTE dT phosphoramidite building block. *A-C*, reagents and conditions. a, 4,4'-dimethoxytrityl chloride, 4-dimethylaminopyidine, anhydrous pyridine, rt, 12h; b, bis(diisopropylamino)chlorophosphine, *N,N*-diisopropylethylamine, anhydrous DCM, rt, 1 h. c, 3-(2-pyridin-3-yl-1,3-dithiane-2-yl)propanol (**4**), 1-H tetrazole, acetontrile, DCM, rt, 5 h.



7 8 9

Scheme 2-4. Deprotection of dithiane-protected POB-PTE ODN followed by reduction to the PHB-PTE ODN. *A*-*B,* reagents and conditions. a, *N*-chlorosuccinimide, water, acetonitrile, rt, 30 min. b, sodium borohydride, water, 0 °C, 1 hr.

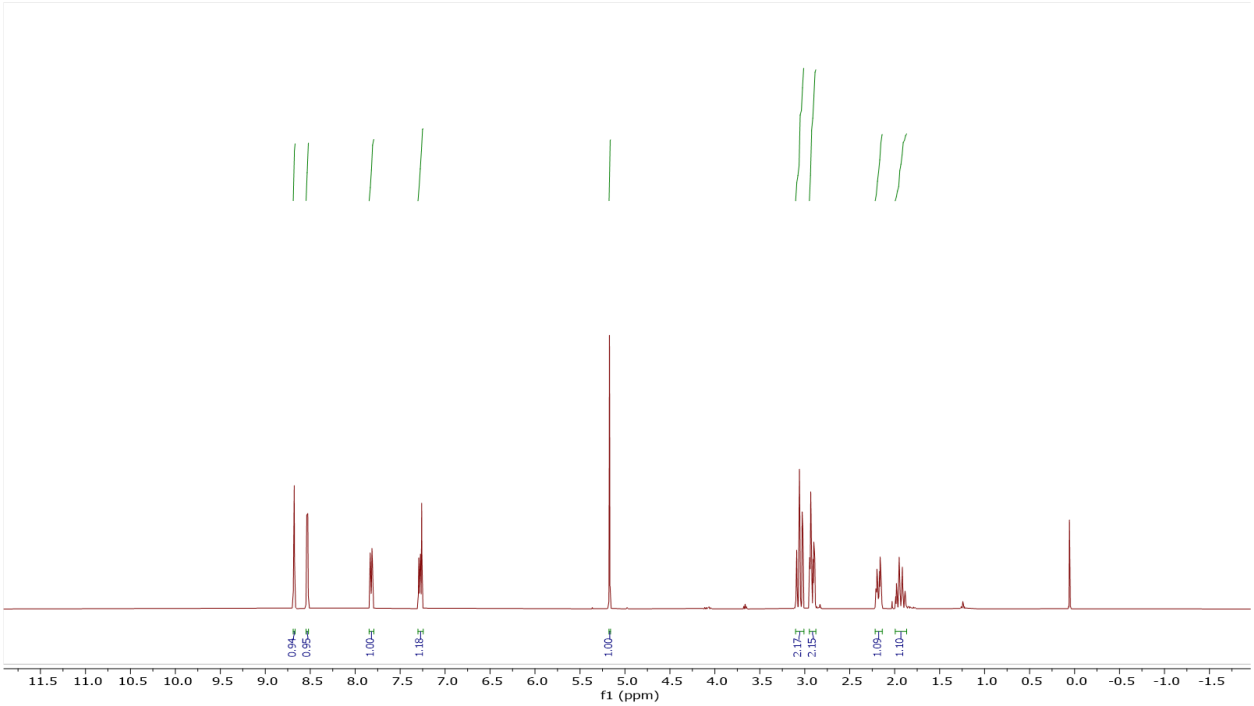


Figure 2.1. `^{1}H NMR spectrum of 2-(3-pyridyl)-1,3-dithiane **(1)**. ^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ 8.71 – 8.64 (m, 1H), 8.53 (dd, *J* = 4.8, 1.6 Hz, 1H), 7.82 (ddd, *J* = 7.9, 2.7, 1.4 Hz, 1H), 7.32 – 7.24 (m, 1H), 5.17 (s, 1H), 3.14 – 3.00 (m, 2H), 2.92 (ddd, *J* = 14.6, 4.4, 3.0 Hz, 2H), 2.18 (dtt, *J* = 14.2, 4.7, 2.5 Hz, 1H), 2.05 – 1.83 (m, 1H).

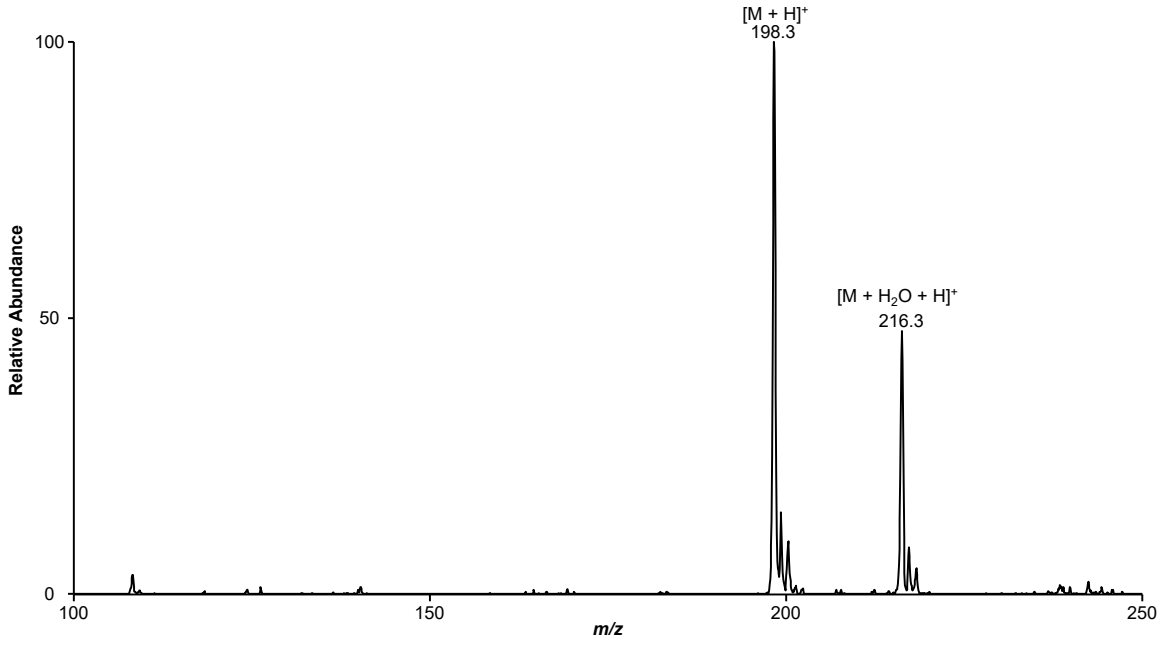


Figure 2-2. Positive-ion ESI-MS of 2-(3-pyridyl)-1,3-dithiane **(1)**. The ion of *m/z* 198.3 corresponds to the $[M + H]^+$ ion (theoretical *m/z* 198.3) of compound **1**.

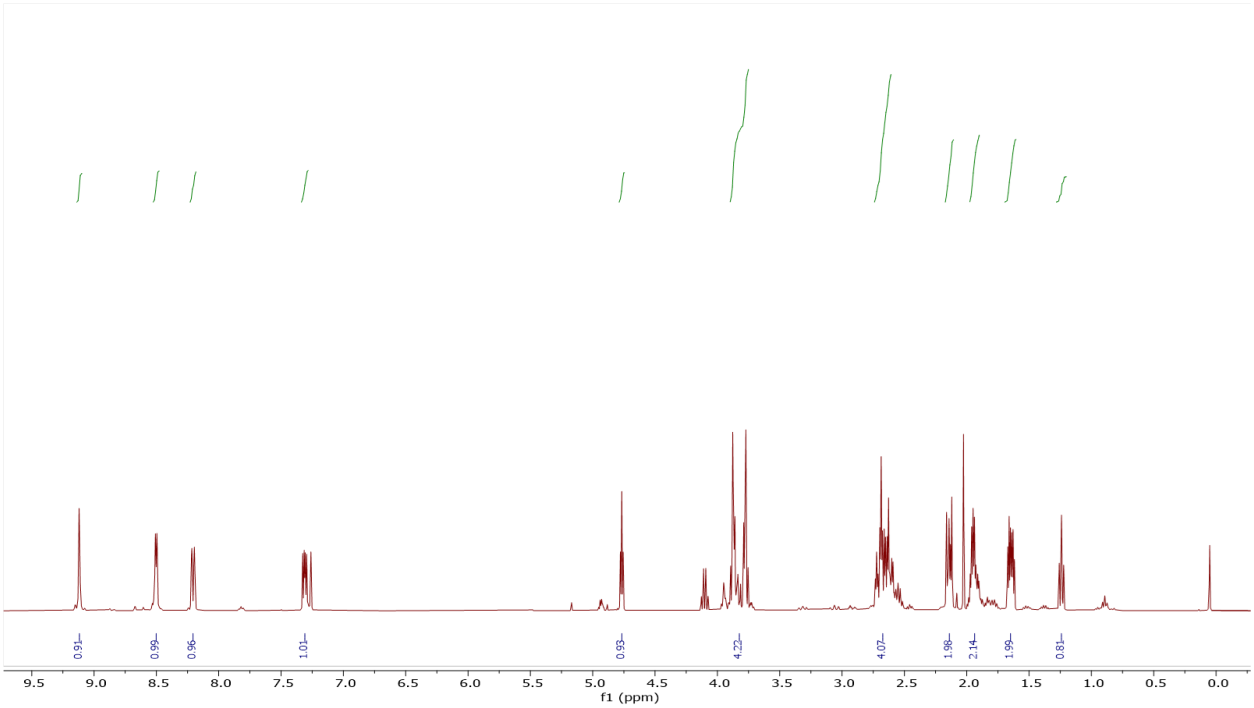


Figure 2-3. ^{1}H NMR spectrum of 3-[2-[2-(1,3-dioxolan-2-yl)ethyl]-1,3-dithian-2-yl]pyridine **(2).** ^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ 9.12 (d, *J* = 2.4 Hz, 1H), 8.52 – 8.48 (m, 1H), 8.21 (dt, *J* = 8.1, 1.7 Hz, 1H), 7.31 (dd, *J* = 8.2, 4.8 Hz, 1H), 4.77 (t, *J* = 4.4 Hz, 1H), 3.90 – 3.75 (m, 4H), 2.67 (ddt, *J* = 24.1, 14.4, 4.3 Hz, 4H), 2.17 – 2.11 (m, 2H), 1.94 (tt, *J* = 9.8, 4.2 Hz, 2H), 1.64 (dt, *J* = 8.6, 4.3 Hz, 2H).

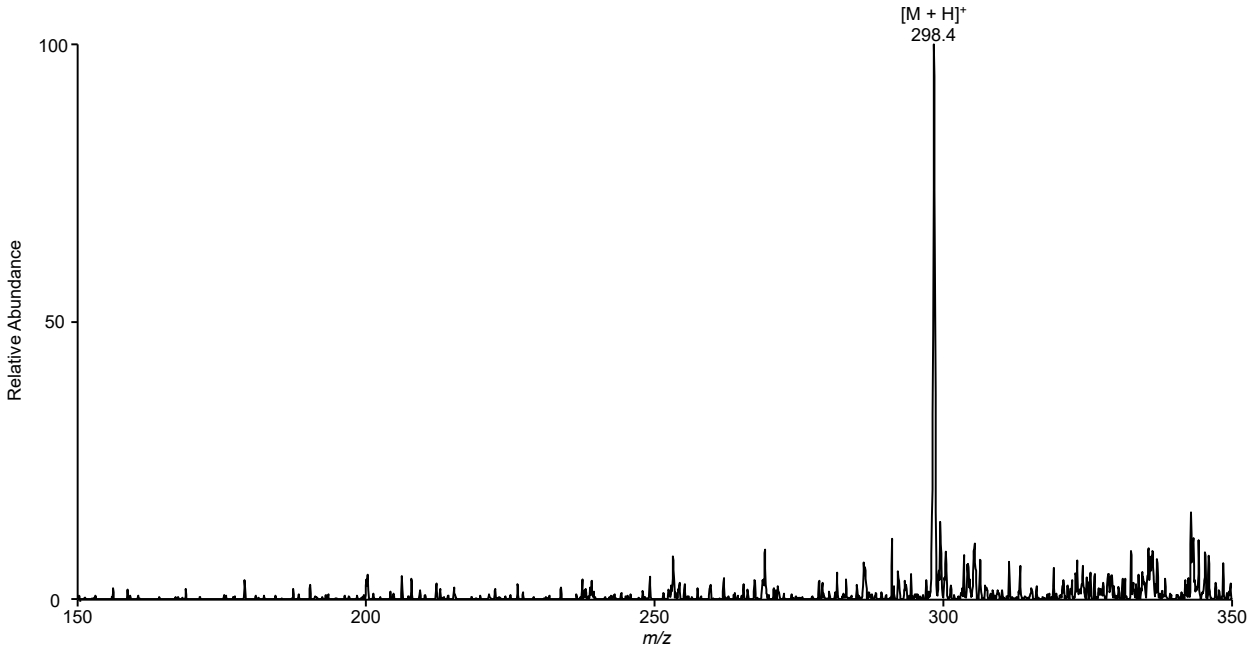


Figure 2-4. Positive-ion ESI-MS of 3-[2-[2-(1,3-dioxolan-2-yl)ethyl]-1,3-dithian-2-yl]pyridine **(2)**. The ion of *m/z* 298.4 corresponds to the $[M + H]^+$ ion (theoretical *m/z* 298.4) of compound **2**.

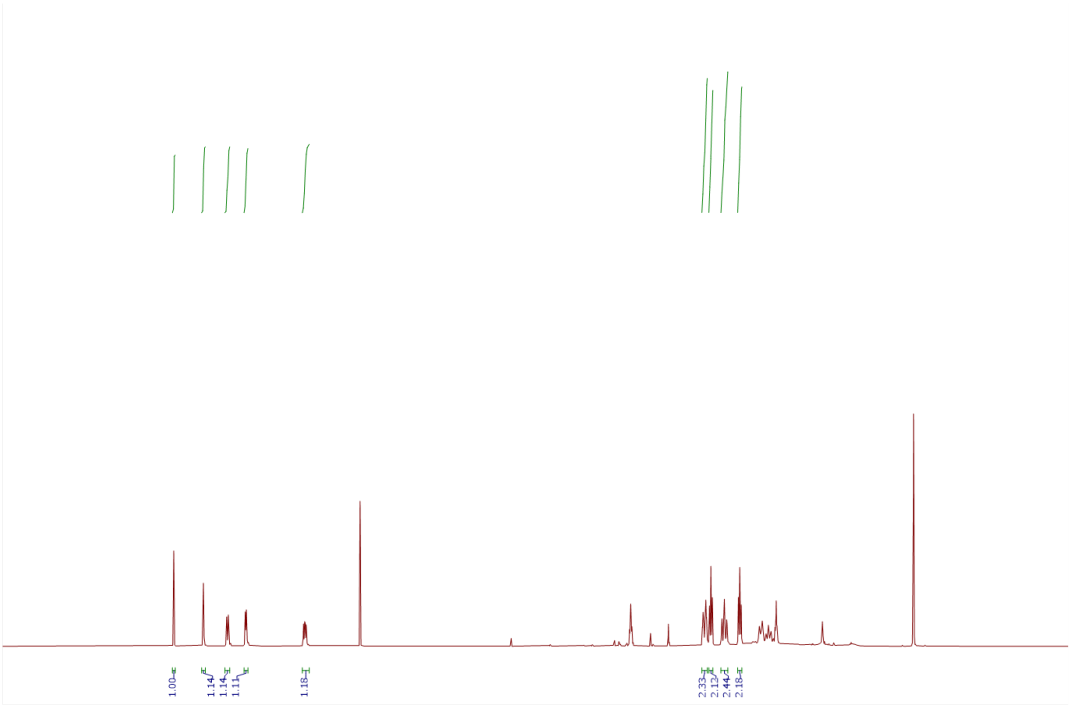


Figure 2-5. ^{1}H NMR spectrum of 3-(2-pyridin-3-yl-1,3-dithiane-2-yl)propanal **(3)**. ^{1}H NMR (400 MHz, $CDCl_3$, 25°C) δ 9.68 (s, 1H), 9.30 (d, *J* = 2.1 Hz, 1H), 8.98 (dd, *J* = 8.4, 2.0 Hz, 1H), 8.75 (d, *J* = 5.5 Hz, 1H), 7.98 (dd, *J* = 8.3, 5.5 Hz, 1H), 2.78 (dt, *J* = 14.8, 3.9 Hz, 3H), 2.70 (t, *J* = 7.3 Hz, 2H), 2.52 (ddd, *J* = 14.6, 11.8, 2.8 Hz, 3H), 2.32 (t, *J* = 7.3 Hz, 2H).

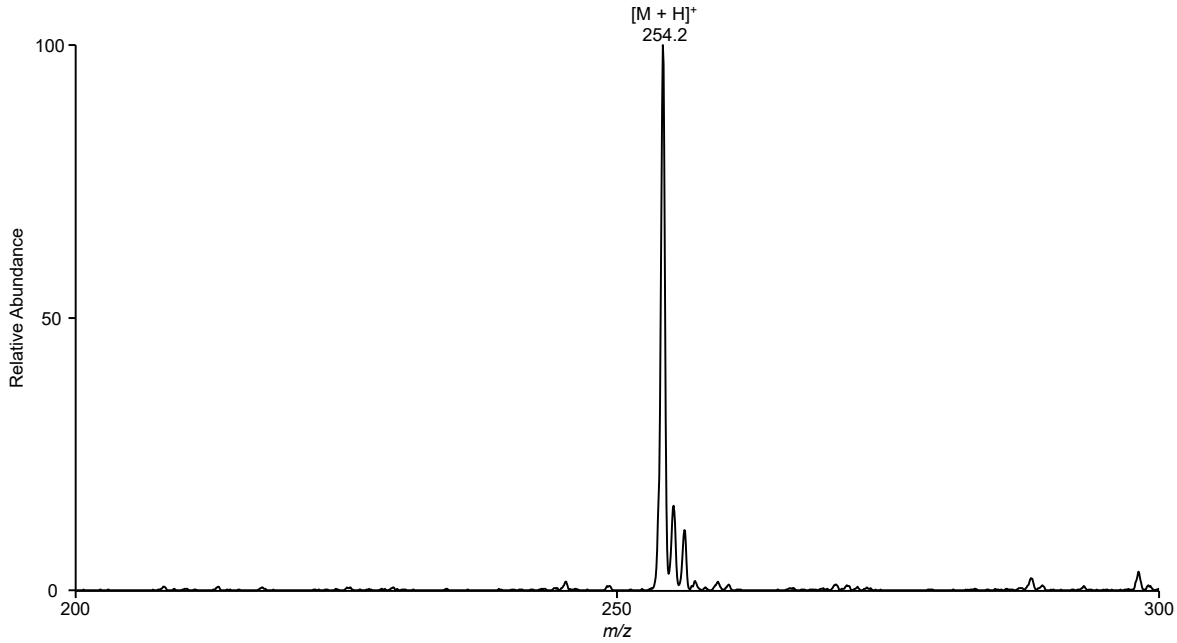


Figure 2-6. Positive-ion ESI-MS of 3-(2-pyridin-3-yl-1,3-dithiane-2-yl)propanal **(3)**. The ion of *m/z* 254.2 corresponds to the $[M + H]^+$ ion (theoretical *m/z* 254.4) of compound **3**.

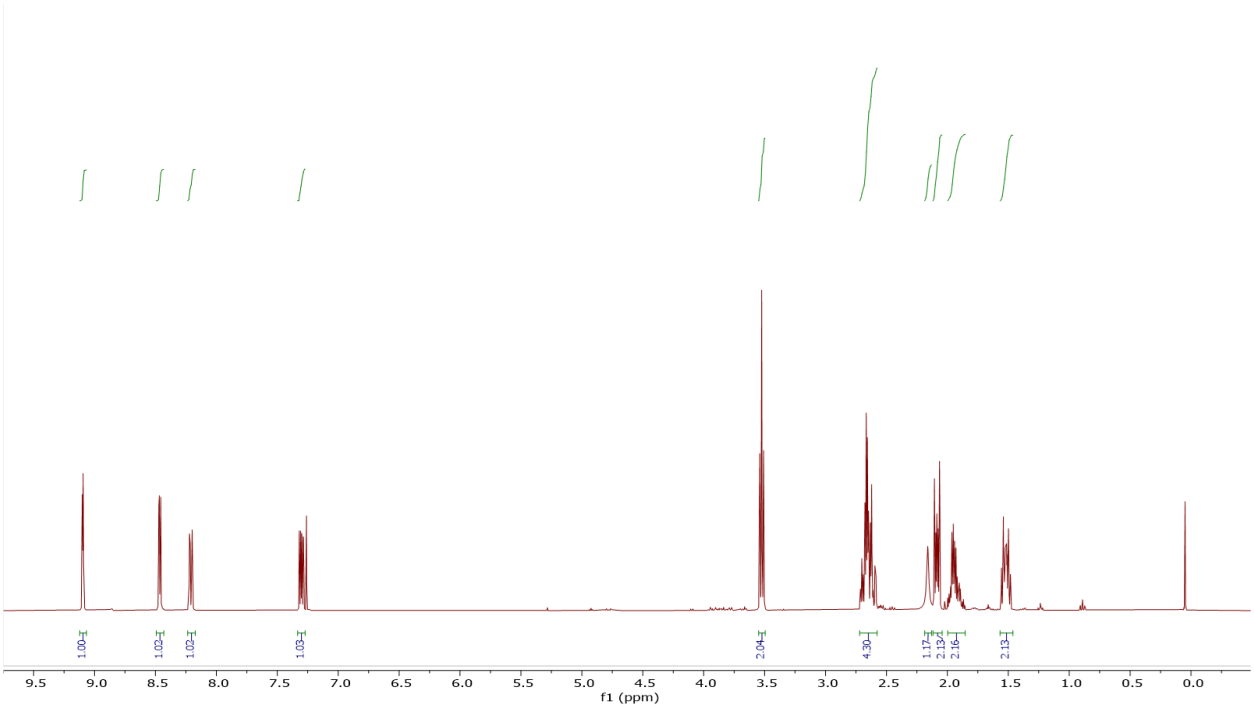


Figure 2-7. ^{1}H NMR spectrum of 3-(2-pyridin-3-yl-1,3-dithiane-2-yl)propanol **(4)**. ^{1}H NMR (400 MHz, $CDCl_3$, 25°C): δ 9.10 (dd, *J* = 4.7, 1.6 Hz, 1H), 8.46 (dd, *J* = 4.7, 1.6 Hz, 1H), 8.21 (ddd, *J* = 8.1, 2.5, 1.6 Hz, 1H), 7.30 (ddd, *J* = 8.1, 4.7, 0.8 Hz, 1H), 3.52 (t, *J* = 6.3 Hz, 2H), 2.76 – 2.55 (m, 5H), 2.16 (s, 1H), 2.14 – 2.05 (m, 2H), 2.01 – 1.88 (m, 2H), 1.63 – 1.43 (m, 2H).

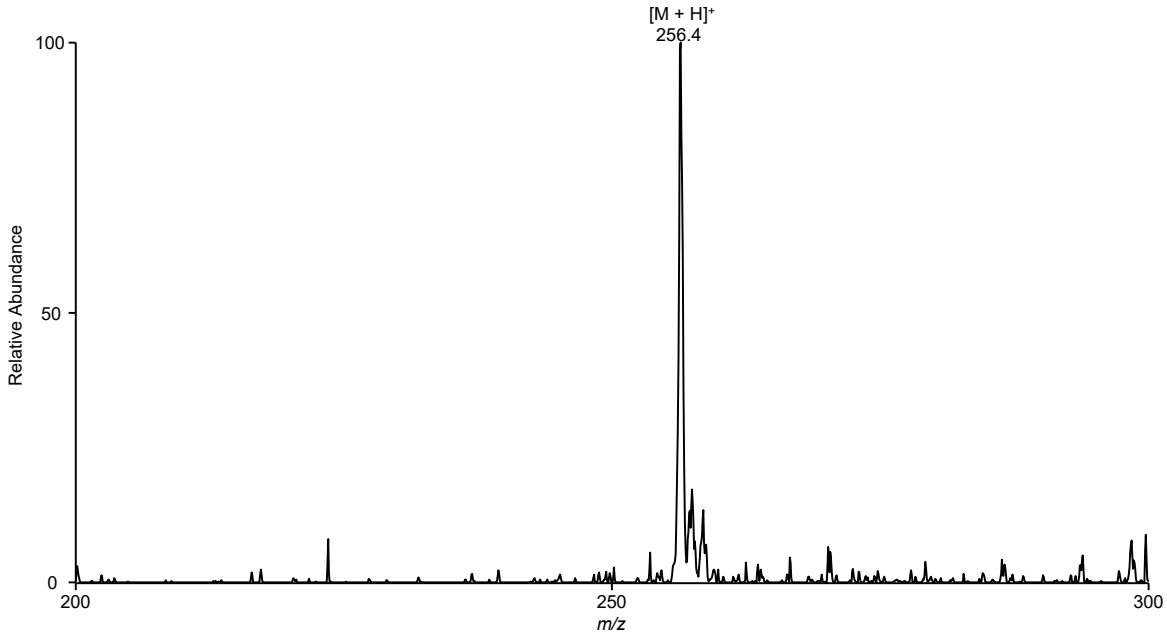


Figure 2-8. Positive-ion ESI-MS of 3-(2-pyridin-3-yl-1,3-dithiane-2-yl)propanol **(4)**. The ion of *m/z* 256.4 corresponds to the $[M + H]^+$ ion (theoretical *m/z* 256.4) of compound **4**.

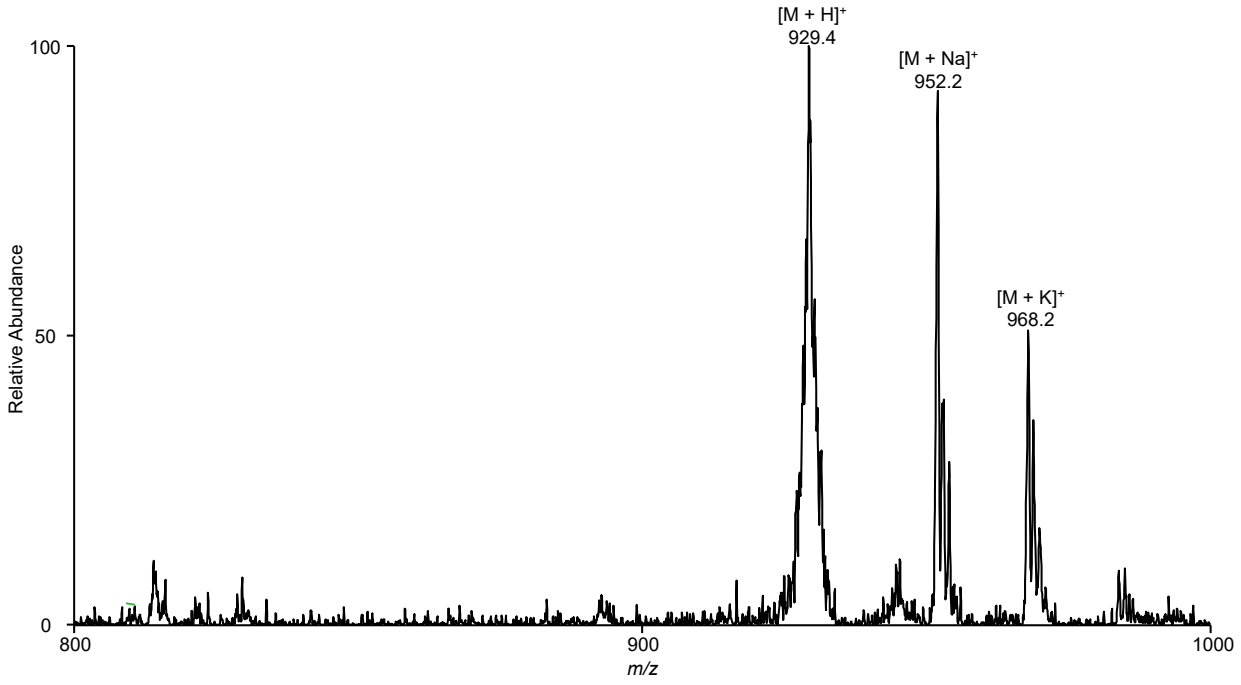


Figure 2-9. Positive-ion ESI-MS of dithiane-protected pyridyloxobutyl phosphotriester of dT phosphoramidite building block **(6)**. The ion of *m/z* 929.4 corresponds to the $[M + H]^+$ ion (theoretical *m/z* 930.3) of compound **6**.

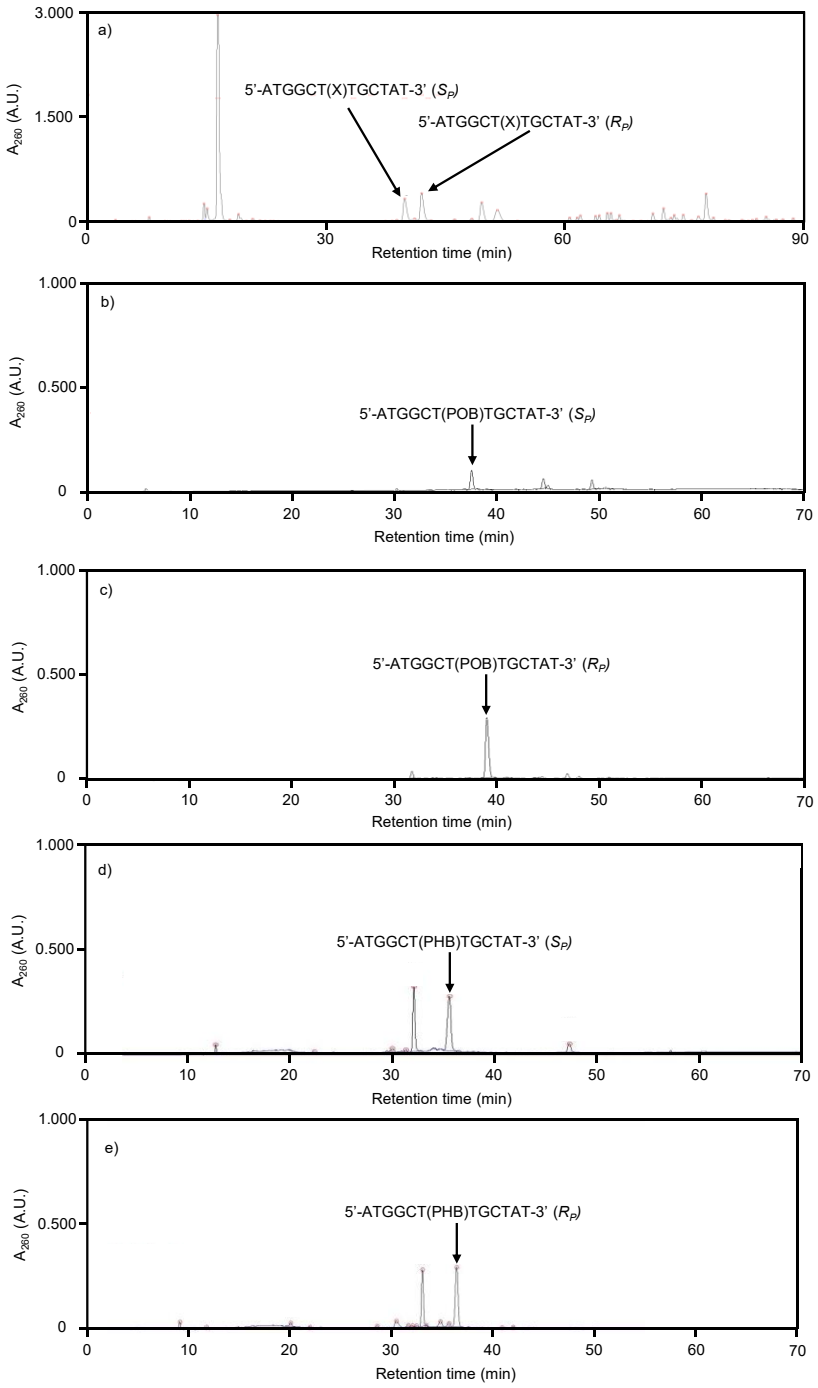


Figure 2-10. HPLC traces of (a) the synthesized dithiane-protected POB-PTE 12-mer; (b and c) The S_P and R_P diastereomers of the POB-PTE 12-mer ODN, respectively; (d and e) S_P and R_P diastereomers of the PHB-PTE 12-mer ODN, respectively.

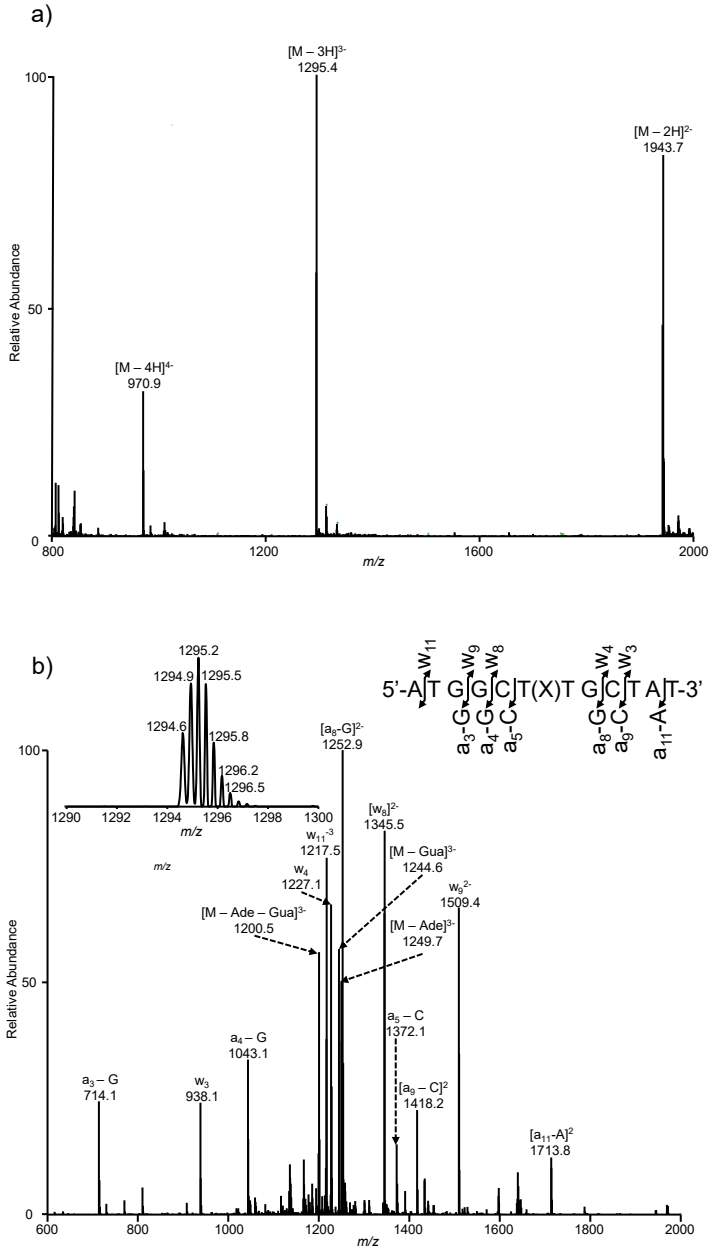


Figure 2-11. ESI-MS and MS/MS characterizations of the dithiane-protected POB-PTE 12-mer ODN **(7),** where 'X' denotes the site of modification. Shown are the negative-ion ESI-MS (a) and the product-ion spectrum of the $[M - 3H]^{3-}$ ion (b). Displayed in the inset is the ultrazoom scan MS of the $[M - 3H]^{3-}$ ion.

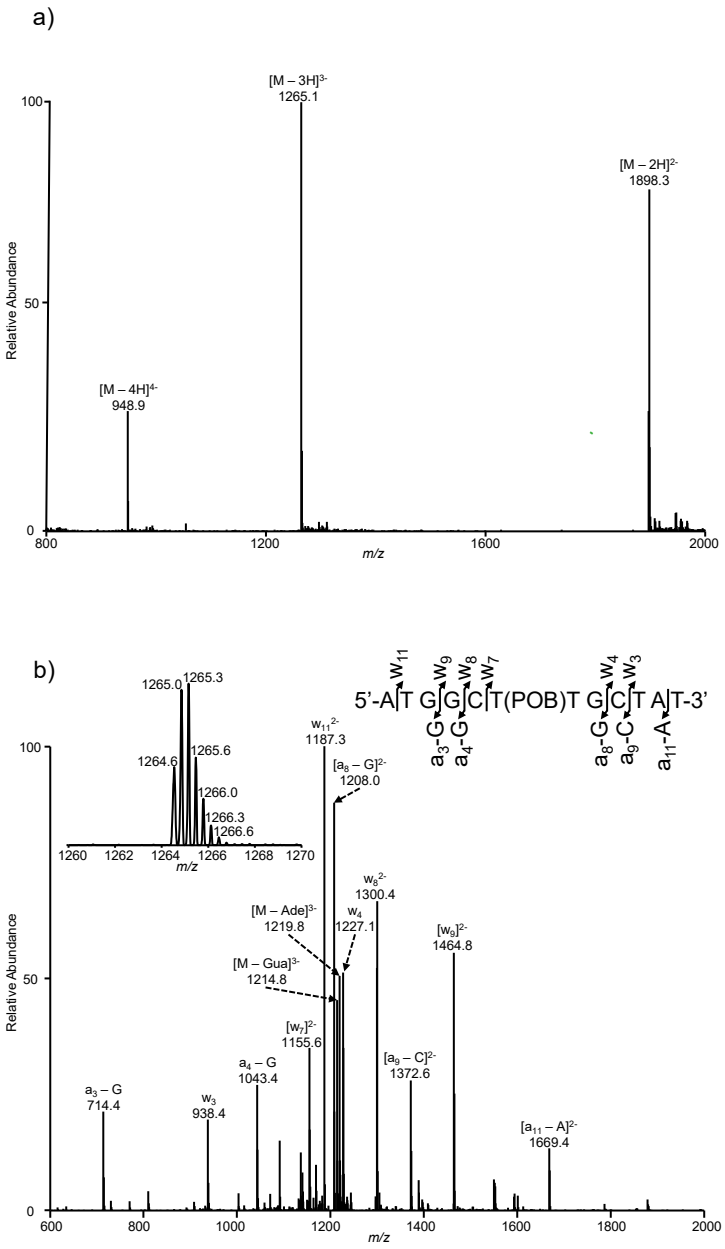


Figure 2-12. ESI-MS and MS/MS characterizations of the POB-PTE 12-mer ODN **(8).** Shown are the negative-ion ESI-MS (a) and the product-ion spectrum of the $[M - 3H]^{3-}$ ion (b). Displayed in the inset is the ultrazoom scan MS of the $[M - 3H]^{3-}$ ion.

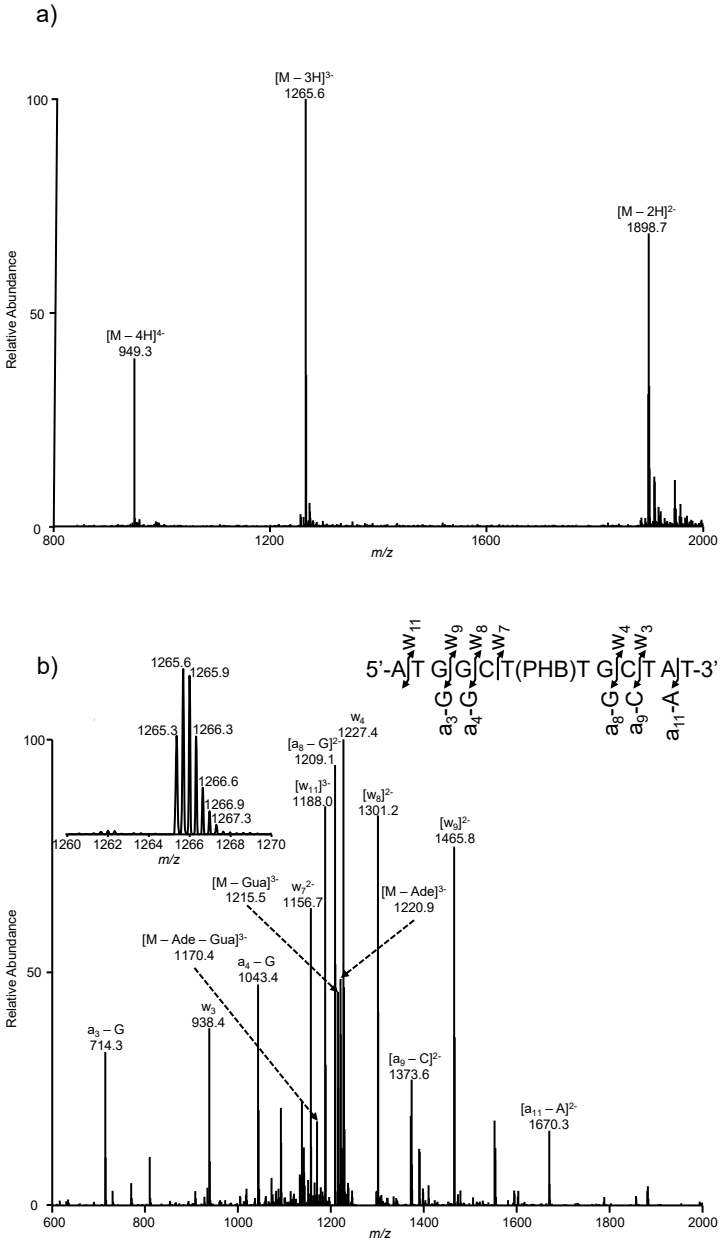


Figure 2-13. ESI-MS and MS/MS characterizations of the PHB-PTE 12-mer ODN **(9).** Shown is the negative-ion ESI-MS (a) and the product-ion spectrum of the $[M - 3H]^{3-}$ ion (b). Displayed in the inset is the ultrazoom scan MS of the $[M - 3H]^{3-}$ ion.

REFERECES

1. Lindahl, T. Instability and Decay of the Primary Structure of DNA. *Nature* **1993**, *362* (6422), 709–715.

2. Li, Y.; Hecht, S. S. Carcinogenic Components of Tobacco and Tobacco Smoke: A 2022 Update. *Food Chem. Toxicol.* **2022**, *165*, 113179.

3. Islami, F.; Goding Sauer, A.; Miller, K. D.; Siegel, R. L.; Fedewa, S. A.; Jacobs, E. J.; McCullough, M. L.; Patel, A. V.; Ma, J.; Soerjomataram, I.; Flanders, W. D.; Brawley, O. W.; Gapstur, S. M.; Jemal, A. Proportion and Number of Cancer Cases and Deaths Attributable to Potentially Modifiable Risk Factors in the United States. *CA. Cancer J. Clin.* **2018**, *68* (1), 31–54.

4. Hecht, S. S. It Is Time to Regulate Carcinogenic Tobacco-Specific Nitrosamines in Cigarette Tobacco. *Cancer Prev. Res. (Phila. Pa.)* **2014**, *7* (7), 639–647.

5. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. *Smokeless Tobacco ; and, Some Tobacco-Specific N-Nitrosamines*; IARC monographs on the evaluation of carcinogenic risks to humans; World Health Organization, International Agency for Research on Cancer: Lyon, France Geneva, Switzerland, 2007.

6. Hecht, S. S. Biochemistry, Biology, and Carcinogenicity of Tobacco-Specific *N* -Nitrosamines. *Chem. Res. Toxicol.* **1998**, *11* (6), 559–603.

7. Balbo, S.; Johnson, C. S.; Kovi, R. C.; James-Yi, S. A.; O'Sullivan, M. G.; Wang, M.; Le, C. T.; Khariwala, S. S.; Upadhyaya, P.; Hecht, S. S. Carcinogenicity and DNA Adduct Formation of 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone and Enantiomers of Its Metabolite 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanol in F-344 Rats. *Carcinogenesis* **2014**, *35* (12), 2798–2806.

8. Peterson, L. A.; Oram, M. K.; Flavin, M.; Seabloom, D.; Smith, W. E.; O'Sullivan, M. G.; Vevang, K. R.; Upadhyaya, P.; Stornetta, A.; Floeder, A. C.; Ho, Y.-Y.; Zhang, L.; Hecht, S. S.; Balbo, S.; Wiedmann, T. S. Coexposure to Inhaled Aldehydes or Carbon Dioxide Enhances the Carcinogenic Properties of the Tobacco-Specific Nitrosamine 4-Methylnitrosamino-1-(3-Pyridyl)-1-Butanone in the A/J Mouse Lung. *Chem. Res. Toxicol.* **2021**, *34* (3), 723–732.

9. A Tobacco-Specific Lung Carcinogen in the Urine of Men Exposed to Cigarette Smoke.

10. Upadhyaya, P.; Carmella, S. G.; Guengerich, F. P.; Hecht, S. S. Formation and Metabolism of 4-(Methylnitrosamino)-1-(3-Pyridyl)- 1-Butanol Enantiomers in Vitro in Mouse, Rat and Human Tissues.

11. Hu, Q.; Upadhyaya, P.; Hecht, S. S.; Aly, F. Z.; Huo, Z.; Xing, C. Characterization of Adductomic Totality of NNK, (*R*)-NNAL and (*S*)-NNAL in A/J Mice, and Their Correlations with Distinct Lung Carcinogenicity. *Carcinogenesis* **2022**, *43* (2), 170–181.

12. Xue, J.; Yang, S.; Seng, S. Mechanisms of Cancer Induction by Tobacco-Specific NNK and NNN. *Cancers* **2014**, *6* (2), 1138–1156.

13. Upadhyaya, P.; Kalscheuer, S.; Hochalter, J. B.; Villalta, P. W.; Hecht, S. S. Quantitation of Pyridylhydroxybutyl-DNA Adducts in Liver and Lung of F-344 Rats Treated with 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone and Enantiomers of Its Metabolite 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanol. *Chem. Res. Toxicol.* **2008**, *21* (7), 1468–1476.

14. Zhang, S.; Wang, M.; Villalta, P. W.; Lindgren, B. R.; Upadhyaya, P.; Lao, Y.; Hecht, S. S. Analysis of Pyridyloxobutyl and Pyridylhydroxybutyl DNA Adducts in Extrahepatic Tissues of F344 Rats Treated Chronically with 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone and Enantiomers of 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanol. *Chem. Res. Toxicol.* **2009**, *22* (5), 926–936.

15. Jones, G. D. D.; Le Pla, R. C.; Farmer, P. B. Phosphotriester Adducts (PTEs): DNA's Overlooked Lesion. *Mutagenesis* **2010**, *25* (1), 3–16.

16. Biswas, I.; Hsieh, P. Interaction of MutS Protein with the Major and Minor Grooves of a Heteroduplex DNA. *J. Biol. Chem.* **1997**, *272* (20), 13355–13364.

17. Lu, A.-L.; Tsai-Wu, J.-J.; Cillo, J. DNA Determinants and Substrate Specificities of Escherichia Coli MutY. *J. Biol. Chem.* **1995**, *270* (40), 23582–23588.

18. Marushige, K. TEMPLATE PROPERTIES OF IbNA ALKYLATED WI N-METHYL-N.NITROSOUREA AND N-ETHYL-N-NITROSO.

19. Suhasini, A. N.; Sommers, J. A.; Yu, S.; Wu, Y.; Xu, T.; Kelman, Z.; Kaplan, D. L.; Brosh, R. M. DNA Repair and Replication Fork Helicases Are Differentially Affected by Alkyl Phosphotriester Lesion. *J. Biol. Chem.* **2012**, *287* (23), 19188–19198.

20. Beranek, D. T. Distribution of Methyl and Ethyl Adducts Following Alkylation with Monofunctional Alkylating Agents. *Mutat. Res. Mol. Mech. Mutagen.* **1990**, *231* (1), 11–30.

21. Engelse, L. D.; Menkveld, G. J.; De Brij, R.-J.; Tates, A. D. Formation and Stability of Alkylated Pyrimidines and Purines (Including Imidazole Ring-Opened 7-Alkylguanine) and Alkylphosphotriesters in Liver DNA of Adult Rats Treated with Ethylnitrosourea or Dimethylnitrosamine. *Carcinogenesis* **1986**, *7* (3), 393–403.

22. Ma, B.; Zarth, A. T.; Carlson, E. S.; Villalta, P. W.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S. Identification of More than 100 Structurally Unique DNA-Phosphate Adducts Formed during Rat Lung Carcinogenesis by the Tobacco-Specific Nitrosamine 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone. *Carcinogenesis* **2018**, *39* (2), 232–241.

23. Corey, E. J.; Seebach, D. Synthesis of 1,n-Dicarbonyl Derivates Using Carbanions from 1,3-Dithianes. *Angew. Chem. Int. Ed. Engl.* **1965**, *4* (12), 1077–1078.

24. Gopalaiah, K. Oxalic Acid: A Very Useful Brønsted Acid in Organic Synthesis. *Synlett* **2004**, *2004* (15), 2838–2839.

25. Chaikin, S. W.; Brown, W. G. Reduction of Aldehydes, Ketones and Acid Chlorides by Sodium Borohydride. *J. Am. Chem. Soc.* **1949**, *71* (1), 122–125.

26. Habibi-Goudarzi, S.; McLuckey, S. A. Ion Trap Collisional Activation of the Deprotonated Deoxymononucleoside and Deoxydinucleoside Monophosphates. *J. Am. Soc. Mass Spectrom.* **1995**, *6* (2), 102–113.

27. Wu, J.; Wang, P.; Wang, Y. Cytotoxic and Mutagenic Properties of Alkyl Phosphotriester Lesions in Escherichia Coli Cells. *Nucleic Acids Res.* **2018**, *46* (8), 4013–4021.

28. Delaney, J. C.; Essigmann, J. M. Assays for Determining Lesion Bypass Efficiency and Mutagenicity of Site-Specific DNA Lesions In Vivo. In *Methods in Enzymology*; Elsevier, 2006; Vol. 408, pp 1–15.

29. Wu, J.; Yuan, J.; Price, N. E.; Wang, Y. Ada Protein– and Sequence Context–Dependent Mutagenesis of Alkyl Phosphotriester Lesions in Escherichia Coli Cells. *J. Biol. Chem.* **2020**, *295* (26), 8775–8783.

30. Wu, J.; Wang, Y. Replication of Pyridyloxobutyl Phosphotriester Lesions in Cells. *Chem. Res. Toxicol.* **2020**, *33* (2), 308–311.

31. Wu, J.; Wu, J.; Clabaugh, G.; Wang, Y. Replication Studies of Alkyl Phosphotriester Lesions in Human Cells. *Chem. Res. Toxicol.* **2024**, *37* (3), 451–454.

32. Tan, Y.; Wu, J.; Clabaugh, G.; Li, L.; Du, H.; Wang, Y. Size- and Stereochemistry-Dependent Transcriptional Bypass of DNA Alkyl Phosphotriester Adducts in Mammalian Cells. *DNA* **2022**, *2* (4), 221–230.

33. Zhao, T.; He, X.; Liang, X.; Kellum, A. H.; Tang, F.; Yin, J.; Guo, S.; Wang, Y.; Gao, Z.; Wang, Y. HMGB3 and SUB1 Bind to and Facilitate the Repair of N^2 -Alkylguanine Lesions in DNA. *J. Am. Chem. Soc.* **2024**, *146* (32), 22553–22562.

34. Wang, P.; Leng, J.; Wang, Y. DNA Replication Studies of N-Nitroso Compound–Induced O6-Alkyl-2′-Deoxyguanosine Lesions in Escherichia Coli. *J. Biol. Chem.* **2019**, *294* (11), 3899–3908.

35. Pauly, G. T.; Peterson, L. A.; Moschel, R. C. Mutagenesis by O^6 -[4-Oxo-4-(3-Pyridyl)Butyl]Guanine in *Escherichia Coli* and Human Cells. *Chem. Res. Toxicol.* **2002**, *15* (2), 165–169.

36. Jasti, V. P.; Spratt, T. E.; Basu, A. K. Tobacco-Specific Nitrosamine-Derived O^2 -Alkylthymidines Are Potent Mutagenic Lesions in SOS-Induced *Escherichia Coli*. *Chem. Res. Toxicol.* **2011**, *24* (11), 1833–1835.

37. Weerasooriya, S.; Jasti, V. P.; Bose, A.; Spratt, T. E.; Basu, A. K. Roles of Translesion Synthesis DNA Polymerases in the Potent Mutagenicity of Tobacco-Specific Nitrosamine-Derived O2-Alkylthymidines in Human Cells. *DNA Repair* **2015**, *35*, 63–70

Chapter 3

Replication Studies of Alkyl Phosphotriester Lesions in Human Cells

INTRODUCTION

Endogenous and exogenous genotoxic agents constantly challenge the integrity of the genome.[1] Among them, various alkylating agents are known to react with DNA to yield a battery of DNA lesions with alkyl groups being conjugated to different positions of nucleobases.[2-3] The ensuing DNA adducts may inhibit DNA replication and transcription, and induce mutations in these important DNA metabolic processes.[3] Apart from nucleobase modifications, alkylating agents can attack directly one of the two non-carbon-bound oxygen atoms of the internucleotide phosphodiester linkage, leading to the formation of alkyl phosphotriester (alkyl-PTE) lesions in two diastereomeric configurations, i.e., S_P and R_P (Figure 3-1).[4-6]

Several studies have been conducted to examine the occurrence of alkyl-PTE lesions. For instance, it was found that approximately 1% and 15% of total alkylation adducts were alkyl-PTE lesions upon treatment of DNA with methyl methanesulfonate and ethyl methanesulfonate, respectively[7]. In addition, Me- and Et-PTE lesions were shown to be highly persistent in mammalian tissues, suggesting the poor repair of these lesions *in vivo*[8-9]. Recent studies also documented that exposure of rodents to tobacco-specific *N*-nitrosamines gives rise to the accrual of substantial levels of alkyl-PTE lesions in tissue DNA[10-12]. The persistence of the alkyl-PTE lesions in tissues suggests that they may be

appropriate biomarkers for assessing cumulative exposure to genotoxic alkylating agents, and they are likely encountered by cellular DNA replication and transcription machineries.

Some studies have been conducted to examine the repair and biological endpoints of alkyl-PTE lesions. Along this line, the S_P-Me-PTE could be repaired by the Ada methyltransferase in *Escherichia coli* [13-14], though no such methyltransferase activity has yet been detected in human cells [4, 15]. Additionally, the activities of SF2 helicases, some of which are involved in DNA repair, were strongly inhibited by *i*Pr-PTE *in vitro*[16], and Et-PTE was found to inhibit T4 DNA polymerase[17]. Recently, Wu et al.[18-20] conducted comprehensive investigations about the replication of a series of alkyl-PTE lesions (with the alkyl group being Me, Et, *n*Pr, *n*Bu, and pyridyloxobutyl) in R_P or S_P configuration at TT dinucleotide site in *E. coli* cells. Interestingly, it was found that only the replication across the S_P-Me-PTE is mutagenic, and the mutagenic bypass of this lesion entails Ada protein[18-20]. In addition, Tan et al.[21] showed that S_P diastereomer of the Me- and *n*Pr-PTEs lesions did not appreciably perturb transcription efficiency in cultured human cells, whereas the R_P diastereomer of the two adducts moderately and strongly blocked transcription, respectively. No replication study, however, has yet been conducted for alkyl-PTE lesions in human cells.

Here, we generated double-stranded plasmids containing a site-specifically inserted alkyl-PTE lesion (the alkyl group is a Me, Et, *n*Pr or *n*Bu) in two diastereomeric configurations, R_P and S_P (Figure 3-1), and examined how these lesions block DNA replication and elicit mutations in cultured human cells that are proficient in translesion synthesis (TLS), or deficient in TLS polymerases, Pol η, ι, κ or ζ[22]. We employed a previously established strand-specific PCR-competitive replication and adduct bypass (SSPCR-CRAB) assay[23]. To this end, we incorporated a single stereochemically defined

alkyl-PTE lesion at a specific site in a double-stranded shuttle vector carrying an SV40 replication origin, and prepared the corresponding lesion-free control vector (Figure 3-2A). We also included a C/C mismatch two nucleotides away from the lesion site to differentiate the replication products originated from lesion-situated strand and the complementary lesion-free strand[23]. We premixed the lesion-carrying or control plasmids individually with a lesion-free competitor vector at fixed molar ratios and transfected them concurrently into HEK293T cells. Relative to the control vector, the competitor vector harbored three additional nucleotides in the region sandwiched with the recognition sites of the two restriction enzymes used for the cleavage of the PCR products (i.e. NcoI and SfaNI).

The progeny genomes were extracted from human cells at 24 h following the transfection, and the residual unreplicated plasmids were removed by treatment with DpnI and exonuclease III. The progeny plasmids were then PCR-amplified using a pair of primers flanking the site where the lesion was initially installed. In this vein, one of the primers (P1) contained a G at the 3′-terminus, which corresponded to the C/C mismatch locus (Figure 3-2B). This allowed for the selective amplification of the progeny genomes emanating from the replication of the lesion-containing bottom strand. Furthermore, we placed a C/A mismatch in the P1 primer two nucleotides from its 3′ terminus to further improve the specificity of strand-specific PCR, as documented elsewhere[24]. The ensuing PCR products were digested by NcoI and SfaNI (Figure 3-2B), and the restriction fragments were analyzed by native PAGE and LC-MS/MS (Figures 3-3 & 3-4 and Figures 3-5 to 3-7). The quantification data from these analyses were then employed to calculate the bypass efficiencies and mutation frequencies, as described in Materials and Methods.

Our results showed that the Me- and Et-PTE lesions were more readily bypassed than the *n*Pr- and *n*Bu-PTE adducts. The bypass efficiencies for the S_p and R_p diastereomers of Me-PTE were ~43% and 58%, respectively, and the corresponding values were 52% and 38% for Et-PTE lesions, 30% and 32% for *n*Pr-PTE adducts, and 28% and 25% for *n*Bu-PTE lesions (Figure 3-4, Table 3-1).

Replication experiments conducted in HEK293T cells with Pol η, Pol ι, Pol κ and Pol ζ being individually ablated by CRISPR-Cas9[22] unveiled the functions of these polymerases in supporting the replicative bypass of alkyl-PTE lesions. In particular, we found that the bypass efficiencies for S_P-Et-PTE and both diastereomers of the *n*Pr- and *n*Bu-PTE lesions were significantly attenuated in HEK293T cells depleted of Pol η or Pol ζ, supporting the roles of these two polymerases in bypassing these lesions (Figure 3-4, Table 3-1). Furthermore, removal of Pol κ resulted in significantly diminished bypass efficiencies for the R_p diastereomer of *n*Pr- and *n*Bu-PTE lesions. Knockout of Pol ι resulted in moderate diminutions in bypass efficiencies for both diastereomers of the *n*Pr- and *n*Bu-PTE products (Figure 3-4, Table 3-1).

The results from PAGE and LC-MS/MS analyses of restriction fragments of PCR products from progeny genomes also allowed us to assess the mutagenic properties of the alkyl-PTE lesions. It turned out that none of the alkyl-PTE adducts were mutagenic in parental HEK293T cells, or the isogenic cells deficient in any of the four TLS polymerases (Figures 3-5 to 3-7).

Exposure to alkylating agents is known to induce appreciable levels of alkyl-PTE lesions in DNA. Although a number of studies have been conducted for assessing the occurrence and persistence of alkyl-PTE lesions, their biological consequences in mammalian cells are largely ignored[4, 21]. In the present study, we aim to attain a comprehensive understanding about the biological end points of the alkyl-PTE lesions by examining how these lesions impede DNA replication and induce mutations in human cells, and how replication across these lesions is modulated by TLS DNA polymerases. Hence, this is the first study about the effects of any alkyl-PTE lesions on DNA replication in human cells.

RESULTS

Our results showed that most alkyl-PTE lesions studied herein constitute moderate impediments to DNA replication in human cells, with no pronounced differences being observed for the S_P and R_P diastereomers. The latter finding differs from the observations made in *E. coli*, where these lesions in the R_P configuration exert stronger blockage effects on DNA replication than those in the S_P configuration[25]. While the exact reason behind this difference is unclear, the alkyl group in the R_P and S_P diastereomers projects into the major groove and extends out perpendicularly of double-stranded DNA[4], respectively; such differences in structural perturbation to duplex DNA elicited by the two diastereomers may be differentially recognized by bacterial DNA replication machinery, but not its mammalian counterpart. In addition, the replication blockage effects of *n*Pr- and *n*Bu-PTE lesions are exacerbated upon individual depletion of Pol η, Pol ι, Pol κ, or Pol ζ, underscoring the important roles of these polymerases in bypassing these lesions (Figure 3-4, Table 3-1).

Together, our results support that the bypass efficiencies of the alkyl-PTE lesions are influenced by both the length of the alkyl chain in the lesions and TLS DNA polymerases in host cells. The exact mechanisms underlying the involvements of these TLS polymerases in supporting the replicative bypass of the alkyl-PTE lesions remain unclear, and future biochemical and structural biology studies may offer some insights into these mechanisms.

No mutagenic products were detectable for either diastereomer of any of the four alkyl-PTE lesions in HEK293T cells or the isogenic cells depleted of any of the four TLS polymerases (Figures 3-6 and 3-7). This perhaps can be attributed to the lack of alterations in Watson-Crick hydrogen bonding properties of nucleobases imposed by alkylation of the backbone phosphate group. This finding also indicates that the presence of the alkyl-PTE lesions does not perturb the DNA polymerases' recognition of the hydrogen bonding face of the nucleobases flanking the lesions. This finding differs from the observations made in *E. coli* cells in the respect that the S_P-Me-PTE at TT dinucleotide site could direct substantial frequencies of nucleotide misincorporations at the two nucleosides flanking the damage site, where Ada protein is required for the mutagenic bypass of the lesion [18]. The difference is not surprising viewing that Ada recognizes only S_P-Me-PTE [13-14] and no such activity has been detected in human cells [4, 15]. It is of note that, in the present study, we only considered the alkyl-PTE lesions at the TT dinucleotide site, and it will be interesting to examine how these lesions flanked by other nucleotides are recognized by DNA replication machinery in human cells.

In summary, we investigated comprehensively the influence of alkyl-PTE adducts on the efficiency and fidelity of DNA replication in cultured human cells and revealed the

roles of four major TLS polymerases (Pol η, κ, ι, and ζ) in bypassing these lesions, which afford novel insights into the biological consequences of alkyl-PTE lesions.

MATERIALS AND METHODS

All chemicals, if not specifically mentioned, were from Sigma-Aldrich (St. Louis, MO, USA) or EMD Millipore, and all enzymes, unless otherwise noted, were obtained from New England Biolabs (Ipswich, MA, USA). 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) was obtained from Oakwood Products Inc. (West Columbia, SC, USA), and [γ-^{32}P]ATP was purchased from PerkinElmer (Piscataway, NJ, USA). All unmodified oligodeoxyribonucleotides (ODNs) were from Integrated DNA Technologies (Coralville, IA, USA). The 12-mer ODNs harboring a site-specifically incorporated PTE lesion were previously synthesized and characterized.[18]

HEK293T cells with *POLH*, *POLI*, *POLK* and *REV3L* genes, which encode DNA polymerases η, ι, κ and the catalytic component of DNA polymerase ζ, respectively, being individually depleted by CRISPR-Cas9 were reported previously.[22] The complete and selective ablation of individual TLS polymerases was validated by Western blot analyses.[22]

Construction of lesion-containing and lesion-free plasmids

The lesion-containing and the lesion-free control/competitor genomes were prepared following previously published procedures.[23,26-27] First, a parent vector was constructed by modifying the sequence of the original pTGFP-Hha10 plasmid[28] to enable the ligation of the alkyl-PTE-containing ODNs. The parent vector was subsequently digested with

Nt.BstNBI to yield a gapped vector (Figure 3-2A), followed by removal of the ensuing 25-mer single-stranded ODN through annealing with a 25-mer complementary ODN in large excess. The gapped plasmid was then isolated using 100 kDa cutoff ultracentrifugal filter units (Millipore). To ensure complete removal of the 25-mer restriction fragment, the steps of annealing with its complementary strand and centrifugation were repeated. The gapped vector was ligated with a 5′-phosphorylated 13-mer lesion-free ODN (5′-AATTGAGTCGATG-3′) and a 5′-phosphorylated 12-mer lesion-carrying ODN (5′-ATGGCTpTGCTAT-3′), where TpT represents Me-, Et-, *n*Pr-, and *n*Bu-PTE lesions in R_P and S_P configurations flanked by two thymidines, or the corresponding lesion-free ODN by using T4 DNA ligase and ATP in the ligation buffer (Figure 3-2A). The ligation mixture was separated by using agarose gel electrophoresis in the presence of ethidium bromide to purify the successfully ligated, supercoiled plasmid. The amounts of the constructed lesion-containing double-stranded vectors were independently normalized against that of the lesion-free competitor vector following published procedures.[27]

Cellular DNA replication and plasmid isolation

The lesion-bearing and the corresponding non-lesion control plasmids were mixed individually with the competitor plasmid at molar ratios of 3:1 and 1:1, respectively. The HEK293T and the isogenic polymerase-deficient cells were cultured at 37°C in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA, USA) and 100 U/ml penicillin in an incubator containing 5% CO_2. The cells (1×10^5) were seeded in a 24-well plate and cultured for 24 h prior to transfection with a total of 300 ng of the aforementioned control/competitor and lesion/competitor genome

mixtures by using Lipofectamine 2000. At 24 h later, the cells were harvested, and the progeny genomes were extracted using Qiagen Spin kit (Qiagen, Valencia, CA, USA).[23] The residual unreplicated plasmids were further digested by DpnI, followed by removal of the resulting linear DNA with exonuclease III digestion, as described elsewhere.[29] In this vein, the parental plasmid carried 25 DpnI recognition sites; hence, cleavage at any of these sites by DpnI would result in the degradation of the entire plasmid by exonuclease III and prohibit the subsequent PCR amplification of the parent vector.

PCR and polyacrylamide gel electrophoresis (PAGE) analyses

The above-isolated progeny genomes were PCR-amplified by employing GoTaq Hot Start DNA polymerase (Promega, Madison, WI, USA). The two primers were 5′-GCTAGCGGATGCATCGACTCAATTACAG-3′ and 5′-GCTGATTATGATCTAGAGTTGCGGCCGC-3′, and the PCR amplifications began at 95°C for 2 min, followed by 35 cycles at 95°C for 30 s, 64°C for 30 s, and 72°C for 1 min, and a final 5-min extension at 72°C. The PCR products were purified with Cycle Pure Kit (Omega, Norcross, GA) and stored at -20°C till use. For PAGE analysis, a fraction of the PCR products was treated with 5 U NcoI and 1 U shrimp alkaline phosphatase (SAP) at 37°C in 10 μl of NEB buffer 3 for 1 h, followed by heating at 80°C for 20 min to deactivate the SAP (Figure 3-3A). The above mixture was then treated with 5 U of T4 polynucleotide kinase (T4 PNK) in 15 μl of NEB buffer 3 containing 5 mM dithiothreitol and ATP (10 pmol cold and 1.66 pmol [γ-^{32}P]ATP). The reaction was continued at 37 °C for 30 min, followed by heating at 65 °C for 20 min to deactivate the polynucleotide kinase. To the reaction mixture was subsequently added 2 U of SfaNI in 5 μl NEB buffer 3 (Figure 3-3A), and the

solution was incubated at 37 °C for 1.5 h, followed by quenching with 20 μl formamide gel-loading buffer containing xylene cyanol FF and bromophenol blue dyes. The mixture was loaded onto a 30% native polyacrylamide gel (acrylamide/bis-acrylamide = 19:1), and the gel band intensities were quantified using phosphorimaging analysis (Figure 3-3B and Supplementary Figure 3-1).

The effects of DNA lesions on replication efficiency are represented by the relative bypass efficiency (RBE). The RBE values were calculated using the following formula, %RBE = (lesion signal/competitor signal)/(non-lesion control signal/competitor signal) ×100%.

Identification of potential mutagenic products by LC-MS/MS

The restriction digestion products were interrogated by LC-MS and MS/MS to identify potential mutagenic products.[30] The PCR products were digested with 30 U SfaNI restriction endonuclease and 15 U SAP in 150 μl NEB buffer 3 at 37°C for 2 h, followed by deactivation of the phosphatase through incubation at 80°C for 20 min. To the mixture was added 50 U NcoI in 50 μl NEB buffer 3, and the solution was incubated at 37°C for another 2 h. The resulting solution was extracted with phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v). To the aqueous layer were subsequently added 2.5 volumes of 100% ethanol and 0.1 volume of 3.0 M sodium acetate, and the solution was incubated at -20°C overnight to precipitate the DNA. The DNA pellet was then dissolved in doubly distilled water for LC-MS/MS analysis (Figure 3-7 and Supplementary Figure 3-2).

An Agilent 1200 capillary HPLC system (Agilent Technologies, Santa Clara, CA, USA) and an LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA,

USA) were used for all the LC-MS and MS/MS experiments. An Agilent Zorbax SB-C18 column (0.5 × 250 mm, 5 μm in particle size) was employed, and the gradient for LC-MS/MS analysis was 5 min of 5-20% methanol followed by 50 min of 20-45% methanol in 400 mM HFIP (pH was adjusted to 7.0 with triethylamine). The ion-transport tube temperature was set at 300°C throughout the LC-MS experiments. The mass spectrometer was set up for monitoring the fragmentations of the $[M-3H]^{3-}$ ions of the 13-mer ODNs, d(AATTACAGCMNGC) and d(CATGGCMNGCTGT), where 'M' and 'N' represents A, T, C or G. The fragment ions detected in MS/MS were manually assigned. Since replication across all the alkyl-PTE lesions in human cells was error-free (*vide infra*), the mutation frequency was essentially zero for all these lesions.

FIGURES AND SCHEMES

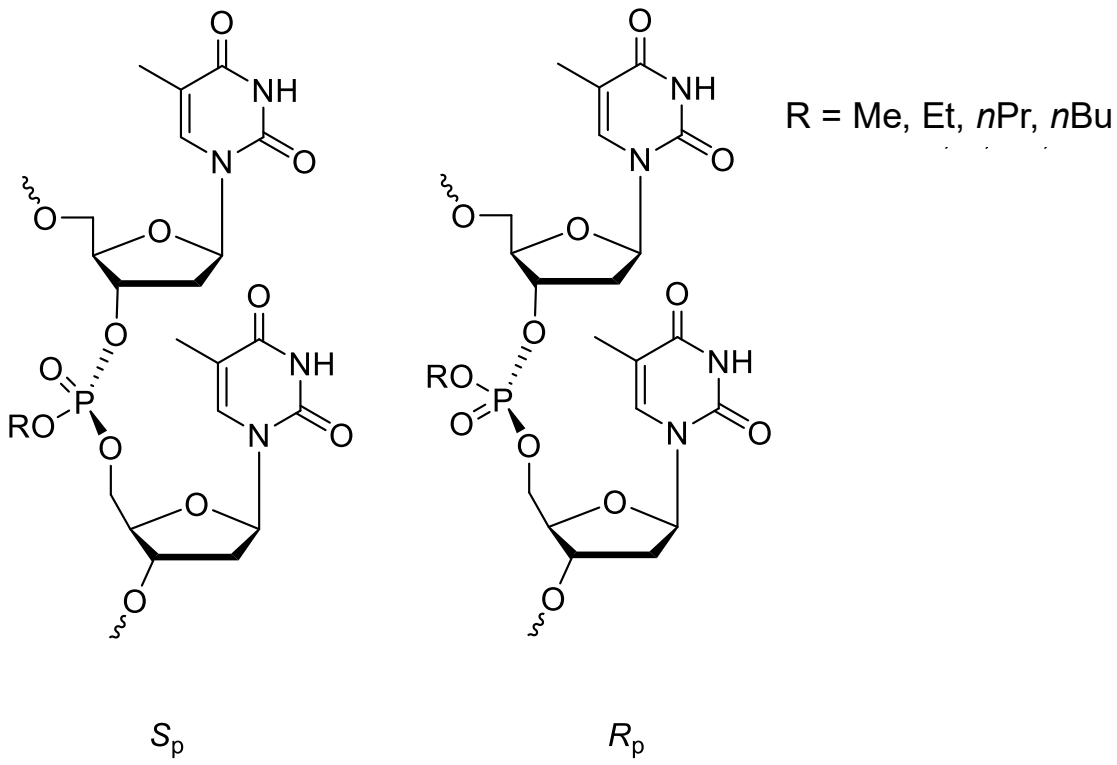


Figure 3-1. The structures of the two diastereomers of alkyl-PTE lesions employed in the present study.

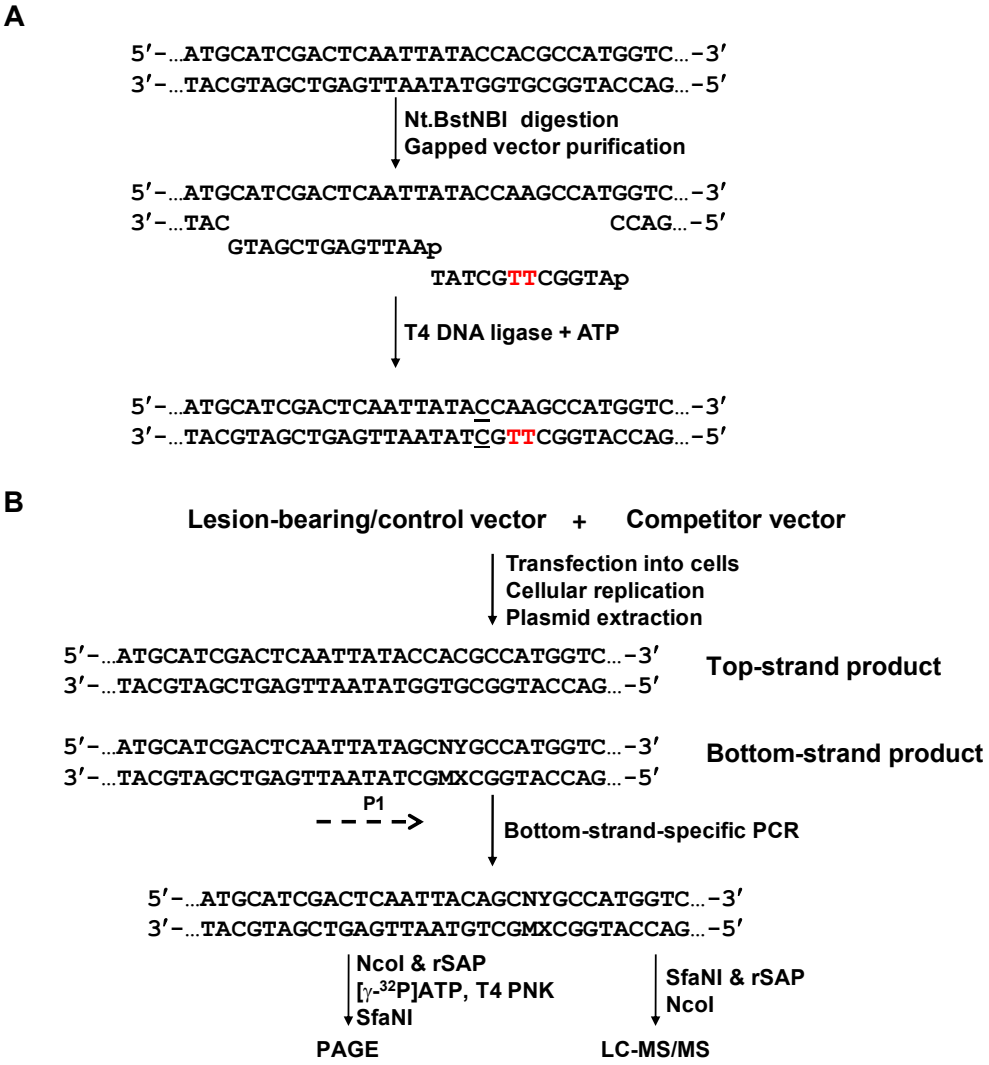


Figure 3-2. Schematic diagrams showing the procedures for the preparation of the lesion-bearing plasmid (A) and the SSPCR-CRAB assay (B). TT in red denotes the site of alkyl-PTE lesions. The C/C mismatch site is underlined. 'P1' designates one of the PCR primers and it contains a G as the terminal 3′-nucleotide, corresponding to the C/C mismatch site of the lesion-bearing genome. It also contains a C/A mismatch two nucleotides away from its 3′-terminus for improving the specificity of PCR. 'M' and 'X' denote the nucleobase generated at the two nucleosides flanking the alkyl-PTE lesion after replication, and 'N' and 'Y' represent the paired nucleobase of M and X, respectively, in the complementary strand.

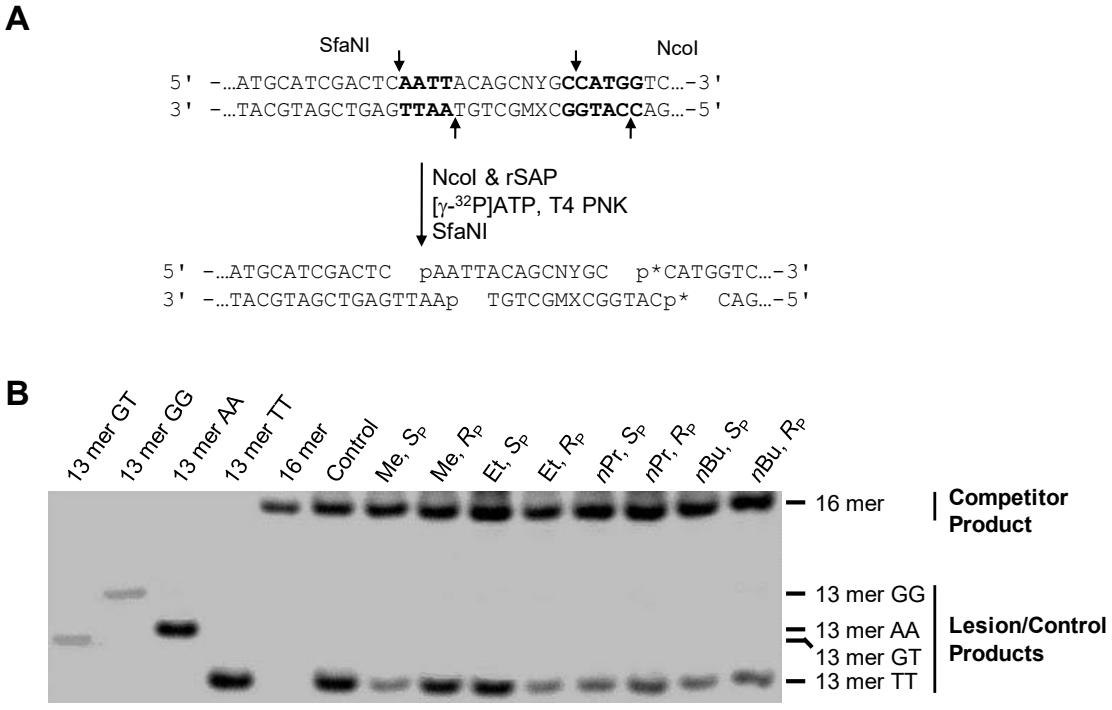


Figure 3-3. Restriction digestion and post-labeling method for determining the bypass efficiencies and mutation frequencies of the PTE lesions in HEK293T cells. (A) Sample processing for restriction digestion using NcoI and SfaNI and post-labeling assay (p* indicates a ^{32}P-labeled phosphate group). The recognition sequences for restriction enzymes are indicated with arrows. (B) Representative gel images showing the NcoI/SfaNI-produced restriction fragments of interest. Synthetic ODN with the same sequence as the restriction fragment arising from the competitor vector, i.e., d(CATGGCGATATGCTGT), is designated as '16-mer Comp'; '13-mer GT', '13-mer GG', '13-mer AA' and '13-mer TT' denote the standard synthetic ODNs d(CATGGCXMGCTGT), where 'XM' is GT, GG, AA and TT, respectively.

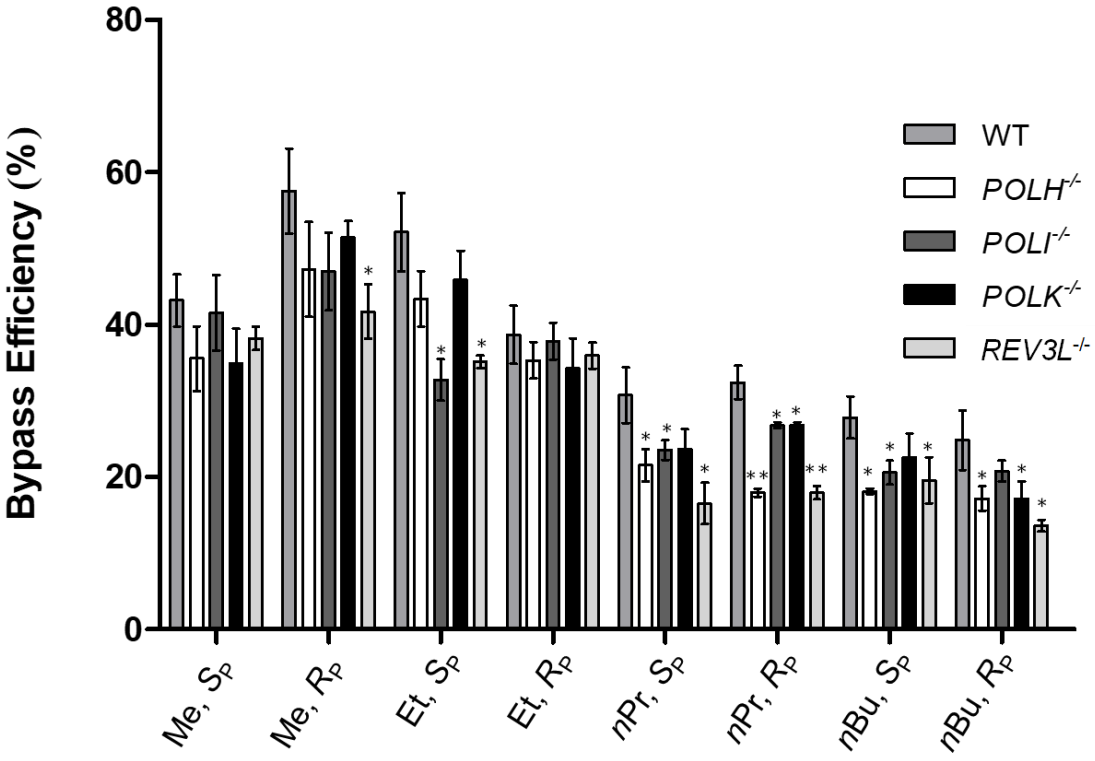


Figure 3-4. The bypass efficiencies of alkyl-PTE lesions. The data represent the mean and S.D. of results from three independent cellular replication experiments. The *p* values referred to the differences between the wild-type (WT) HEK293T cells and the isogenic polymerase-deficient cells, and they were calculated using two-tailed, unpaired t-test. *, $0.01 < p < 0.05$; **, $0.001 < p < 0.01$.

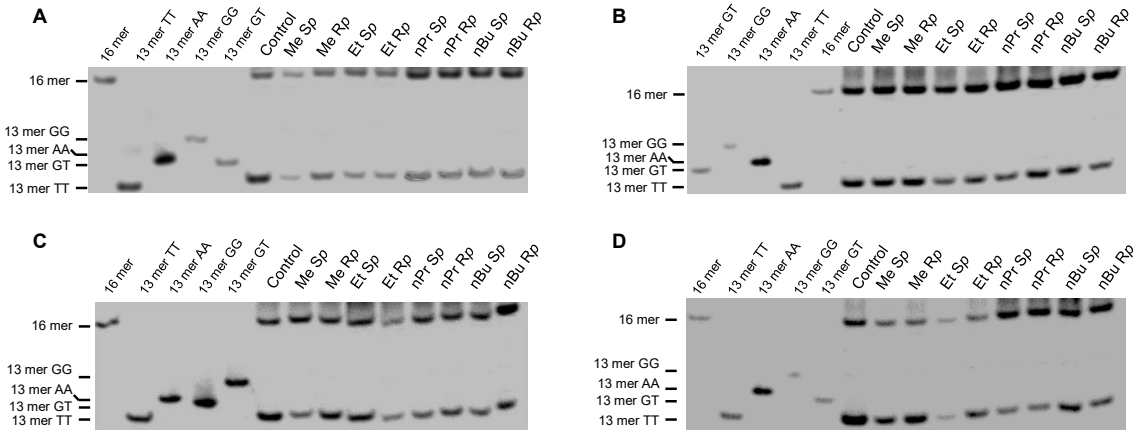


Figure 3-5. Representative gel images showing the NcoI/SfaNI-produced restriction fragments of the PTE DNA lesions in (A) Pol η-, (B) Pol ι-, (C) Pol κ-, and (D) Pol ζ-deficient cells. The restriction fragment arising from the competitor vector, i.e., d(CATGGCGATATGCTGT), is designated as '16-mer'; '13-mer GT', '13-mer GG', '13-mer AA' and '13-mer TT' represent the standard synthetic ODNs d(CATGGCMXGCTGT), where 'MX' is GT, GG, AA and TT, respectively.

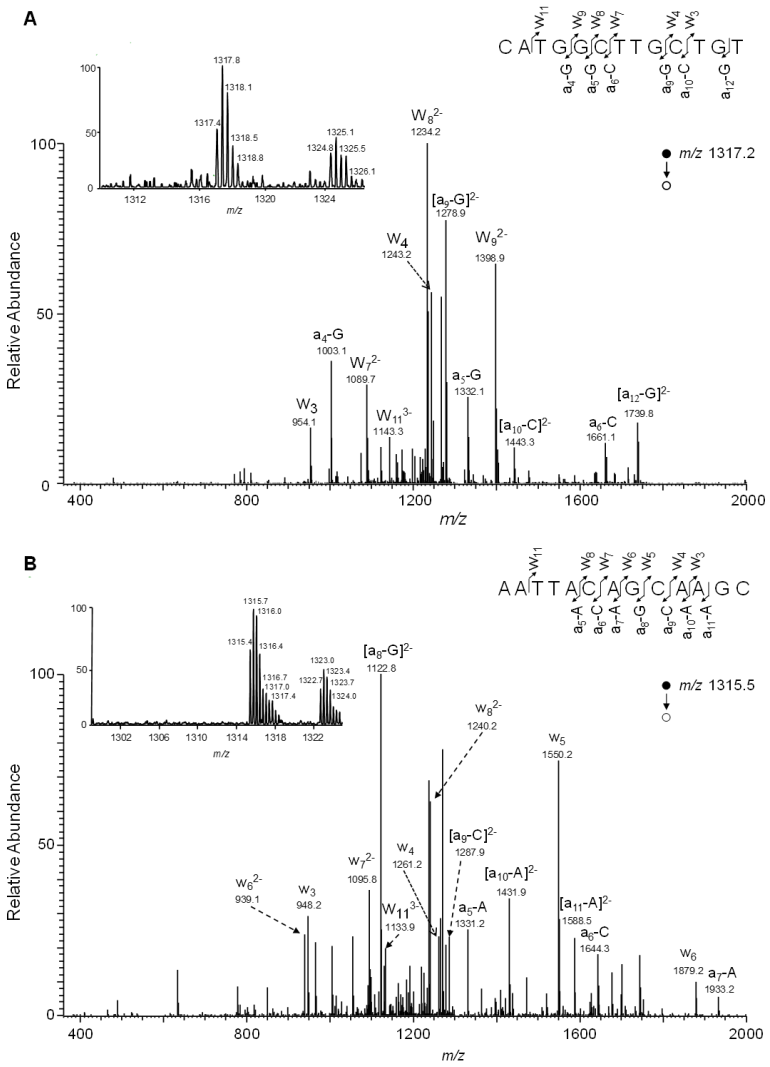
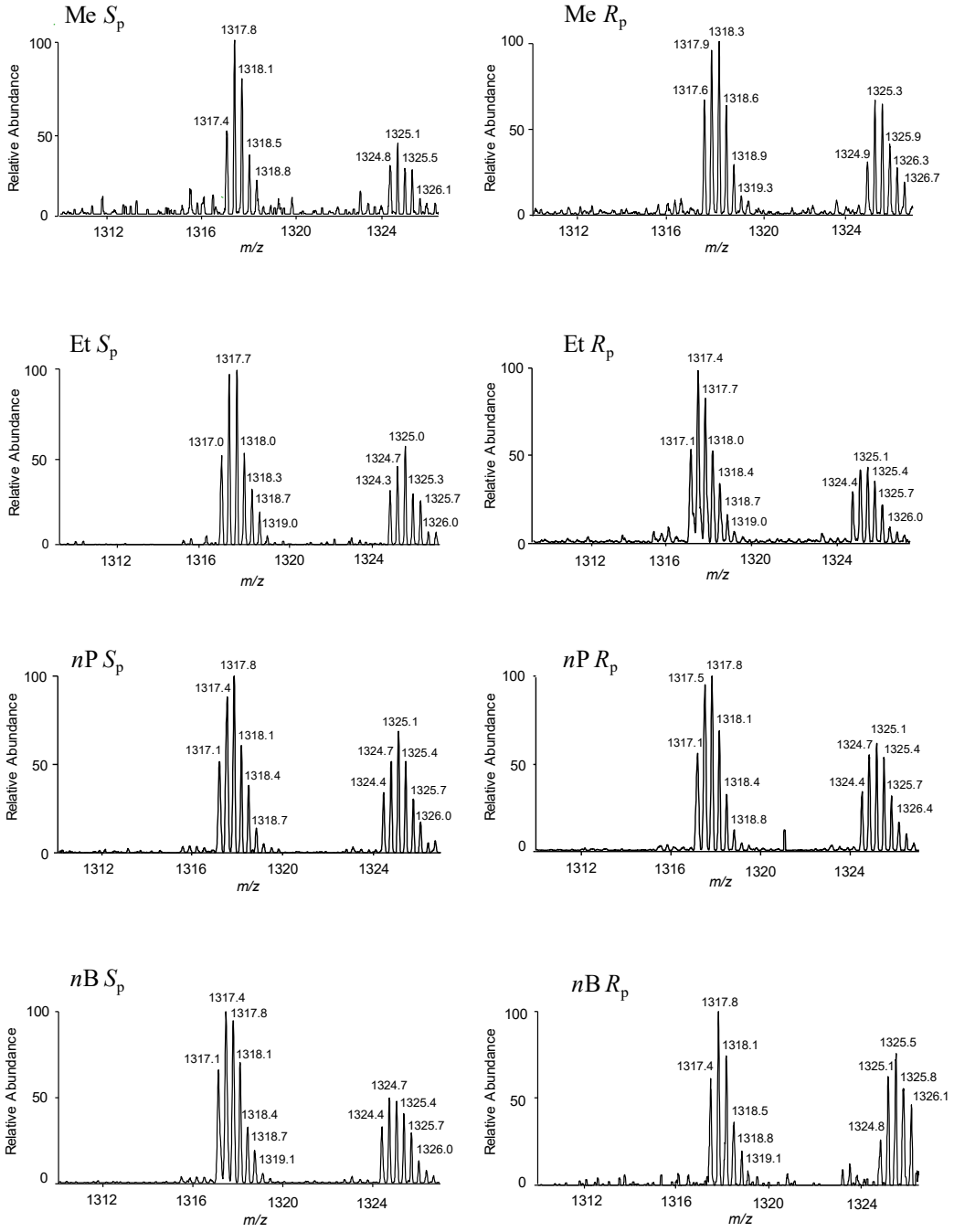


Figure 3-6. Restriction digestion followed by LC-MS/MS for the identification of restriction digestion products (with NcoI and SfaNI) of PCR amplicons of progeny genome arising from the replication of Me-PTE lesions in HEK293T cells (A) and the isogenic Pol η-deficient cells (B). Shown are the MS/MS for monitoring the fragmentations of the $[M–3H]^{3–}$ ions of 5′-CATGGCTTGCTGT-3′ and 5′-AATTACAGCAAGC-3′, respectively. The restriction digestion method is the same as what is described in Figure 3A, except that the shrimp alkaline phosphatase was added after the addition of the two restriction enzymes and that the [^{32}P]-post-labeling step was omitted. Shown in the insets are schemes summarizing the observed fragment ions and a higher-resolution 'ultra-zoom scan' ESI-MS for monitoring the products of $[M–3H]^{3-}$ ions.

A


Me S_p
Me R_p
Et S_p
Et R_p
nP S_p
nP R_p
nB S_p
nB R_p
Relative Abundance
m/z

B

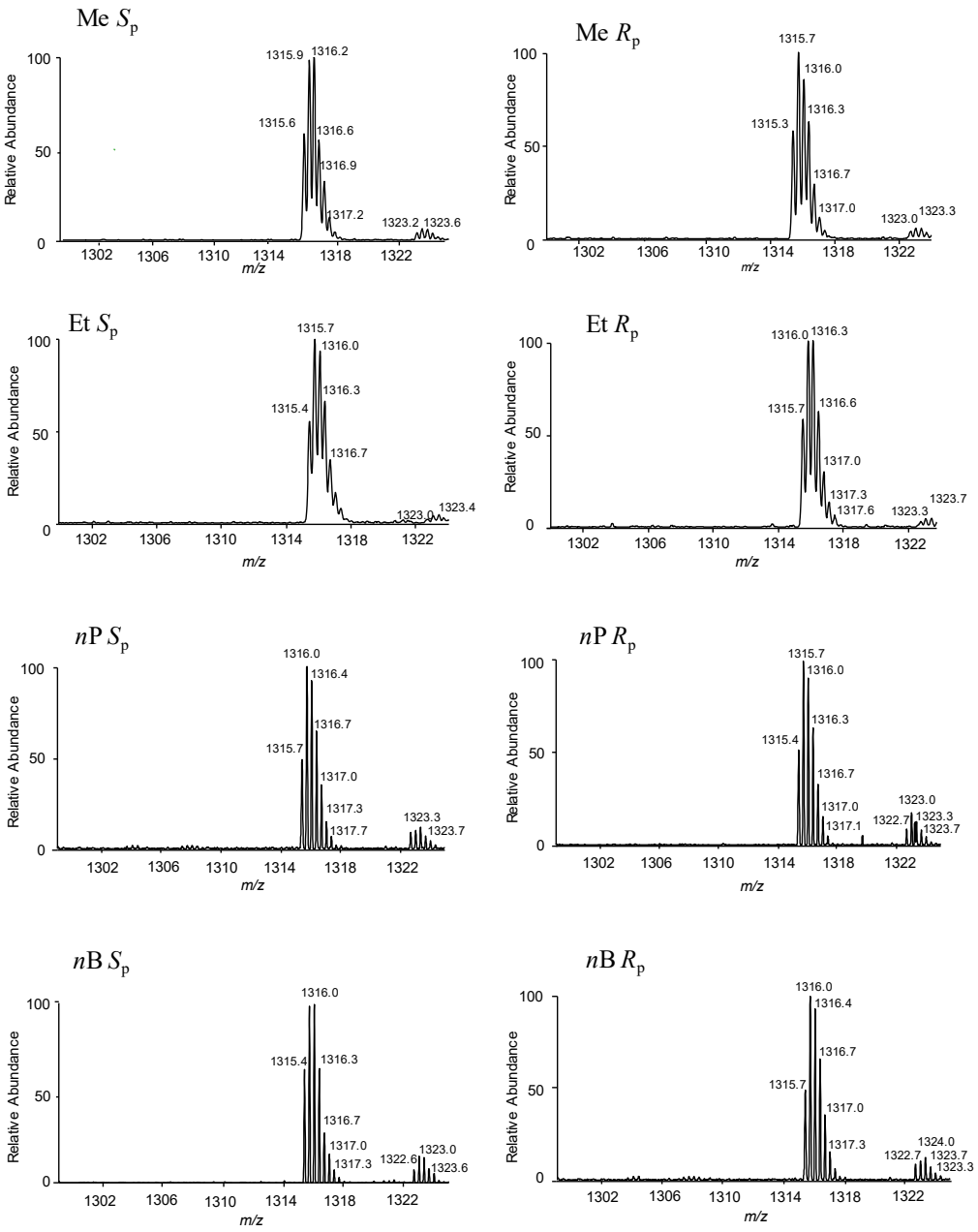

Me S_p
Me R_p
Et S_p
Et R_p
nP S_p
nP R_p
nB S_p
nB R_p
Relative Abundance
m/z

Figure 3-7. Higher-resolution 'ultra-zoom scan' ESI-MS for monitoring the $[M–3H]^{3-}$ ions of the bottom-strand (i.e., the initial lesion-containing strand) restriction fragments of PCR products arising from the replication of alkyl-PTE-containing double-stranded plasmids in HEK293T (A) and the isogenic Pol η-deficient cells (B).

Lesions	WT	*POLH*$^{-/-}$	*POLI*$^{-/-}$	*POLK*$^{-/-}$	*REV3L*$^{-/-}$
S_P, Me-PTE	43.2 ± 3.4	35.5 ± 4.3	41.5 ± 5.0	34.9 ± 4.6	38.2 ± 1.5
R_P, Me-PTE	57.5 ± 5.6	47.2 ± 6.2	47.0 ± 5.1	51.4 ± 2.2	41.7 ± 3.6
S_P, Et-PTE	52.1 ± 5.2	43.4 ± 3.6	32.7 ± 2.7	45.9 ± 3.8	35.1 ± 0.8
R_P, Et-PTE	38.7 ± 3.8	35.3 ± 2.4	37.8 ± 2.4	34.2 ± 4.0	35.9 ± 1.7
S_P, *n*Pr-PTE	30.7 ± 3.6	21.5 ± 2.2	23.5 ± 1.3	23.6 ± 2.6	16.5 ± 2.7
R_P, *n*Pr-PTE	32.4 ± 2.2	17.9 ± 0.6	26.8 ± 0.4	26.8 ± 0.4	17.9 ± 0.9
S_P, *n*Bu-PTE	27.8 ± 2.8	18.1 ± 0.4	20.6 ± 1.6	22.5 ± 3.2	19.6 ± 3.0
R_P, *n*Bu-PTE	24.8 ± 3.9	17.2 ± 1.6	20.8 ± 1.4	17.1 ± 2.2	13.6 ± 0.8

Table 3-1. Relative bypass efficiencies (%, mean ± S.D., n = 3) for the alkyl-PTE lesions in HEK293T and the isogenic TLS polymerase knockout cells.

REFERENCES

1. Friedberg, E. C.; Walker, G. C.; Siede, W.; Wood, R. D.; Schultz, R. A.; Ellenberger, T., *DNA Repair and Mutagenesis*. ASM Press: Washington, D.C., 2006.

2. Liu, S.; Wang, Y., Mass spectrometry for the assessment of the occurrence and biological consequences of DNA adducts. *Chem. Soc. Rev.* **2015,** *44*, 7829-54.

3. Shrivastav, N.; Li, D.; Essigmann, J. M., Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis* **2010,** *31*, 59-70.

4. Jones, G. D.; Le Pla, R. C.; Farmer, P. B., Phosphotriester adducts (PTEs): DNA's overlooked lesion. *Mutagenesis* **2010,** *25*, 3-16.

5. Maccubbin, A. E.; Caballes, L.; Riordan, J. M.; Huang, D. H.; Gurtoo, H. L., A cyclophosphamide/DNA phosphoester adduct formed *in vitro* and *in vivo*. *Cancer Res.* **1991,** *51*, 886-892.

6. Singer, B., In vivo formation and persistence of modified nucleosides resulting from alkylating agents. *Environ. Health Perspect.* **1985,** *62*, 41.

7. Bannon, P.; Verly, W., Alkylation of phosphates and stability of phosphate triesters in DNA. *Eur. J. Biochem.* **1972,** *31*, 103-11.

8. Shooter, K. V.; Slade, T. A., The stability of methyl and ethyl phosphotriesters in DNA *in vivo*. *Chem. Biol. Interact.* **1977,** *19*, 353-61.

9. Den Engelse, L.; De Graaf, A.; De Brij, R. J.; Menkveld, G. J., O^2- and O^4-ethylthymine and the ethylphosphotriester dTp(Et)dT are highly persistent DNA modifications in slowly dividing tissues of the ethylnitrosourea-treated rat. *Carcinogenesis* **1987,** *8*, 751-7.

10. Ma, B.; Villalta, P. W.; Zarth, A. T.; Kotandeniya, D.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S., Comprehensive high-resolution mass spectrometric analysis of DNA phosphate adducts formed by the tobacco-specific lung carcinogen 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Chem. Res. Toxicol.* **2015,** *28*, 2151-9.

11. Ma, B.; Zarth, A. T.; Carlson, E. S.; Villalta, P. W.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S., Identification of more than one hundred structurally unique DNA-phosphate adducts formed during rat lung carcinogenesis by the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Carcinogenesis* **2017,** *39*, 232-241.

12. Ma, B.; Zarth, A. T.; Carlson, E. S.; Villalta, P. W.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S., Methyl DNA phosphate adduct formation in rats treated chronically with 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone and enantiomers of its metabolite 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol. *Chem. Res. Toxicol.* **2018,** *31*, 48-57.

13. McCarthy, T. V.; Lindahl, T., Methyl phosphotriesters in alkylated DNA are repaired by the Ada regulatory protein of *E. coli*. *Nucleic Acids Res.* **1985,** *13*, 2683-98.

14. Weinfeld, M.; Drake, A. F.; Saunders, J. K.; Paterson, M. C., Stereospecific removal of methyl phosphotriesters from DNA by an *Escherichia coli* ada+ extract. *Nucleic Acids Res.* **1985,** *13*, 7067-77.

15. Yarosh, D. B.; Fornace, A. J.; Day, R. S., 3rd, Human O^6-alkylguanine-DNA alkyltransferase fails to repair O^4-methylthymine and methyl phosphotriesters in DNA as efficiently as does the alkyltransferase from *Escherichia coli*. *Carcinogenesis* **1985,** *6*, 949-53.

16. Suhasini, A. N.; Sommers, J. A.; Yu, S.; Wu, Y.; Xu, T.; Kelman, Z.; Kaplan, D. L.; Brosh, R. M., Jr., DNA repair and replication fork helicases are differentially affected by alkyl phosphotriester lesion. *J. Biol. Chem.* **2012,** *287*, 19188-98.

17. Tsujikawa, L.; Weinfield, M.; Reha-Krantz, L. J., Differences in replication of a DNA template containing an ethyl phosphotriester by T4 DNA polymerase and *Escherichia coli* DNA polymerase I. *Nucleic Acids Res.* **2003,** *31*, 4965-72.

18. Wu, J.; Wang, P.; Wang, Y., Cytotoxic and mutagenic properties of alkyl phosphotriester lesions in *Escherichia coli* cells. *Nucleic Acids Res* **2018,** *46*, 4013-4021.

19. Wu, J.; Yuan, J.; Price, N. E.; Wang, Y., Ada protein- and sequence context-dependent mutagenesis of alkyl phosphotriester lesions in *Escherichia coli* cells. *J. Biol. Chem.* **2020,** *295*, 8775-8783.

20. Wu, J.; Wang, Y., Replication of pyridyloxobutyl phosphotriester lesions in cells. *Chem. Res. Toxicol.* **2020,** *33*, 308-311.

21. Tan, Y.; Wu, J.; Clabaugh, G.; Li, L.; Du, H.; Wang, Y., Size- and stereochemistry-dependent transcriptional bypass of DNA alkyl phosphotriester adducts in mammalian cells. *DNA* **2022,** *2*, 221-230.

22. Wu, J.; Li, L.; Wang, P.; You, C.; Williams, N. L.; Wang, Y., Translesion synthesis of O^4-alkylthymidine lesions in human cells. *Nucleic Acids Res.* **2016,** *44*, 9256-9265.

23. You, C.; Swanson, A. L.; Dai, X.; Yuan, B.; Wang, J.; Wang, Y., Translesion synthesis of 8,5'-cyclopurine-2'-deoxynucleosides by DNA polymerases h, i, and z. *J. Biol. Chem.* **2013,** *288*, 28548-56.

24. Newton, C. R.; Graham, A.; Heptinstall, L. E.; Powell, S. J.; Summers, C.; Kalsheker, N.; Smith, J. C.; Markham, A. F., Analysis of any point mutation in DNA. The amplification refractory mutation system (ARMS). *Nucleic Acids Res.* **1989,** *17*, 2503-16.

25. Wu, J.; Wang, P.; Wang, Y., Cytotoxic and mutagenic properties of alkyl phosphotriester lesions in *Escherichia coli* cells. *Nucleic Acids Res.* **2018,** *46*, 4013-4021.

26. Yuan, B.; O'Connor, T. R.; Wang, Y., 6-Thioguanine and S^6-methylthioguanine are mutagenic in human cells. *ACS Chem. Biol.* **2010,** *5*, 1021-1027.

27. You, C.; Wang, Y., Quantitative measurement of transcriptional inhibition and mutagenesis induced by site-specifically incorporated DNA lesions in vitro and in vivo. *Nat. Protoc.* **2015,** *10* (9), 1389-406.

28. Baker, D. J.; Wuenschell, G.; Xia, L.; Termini, J.; Bates, S. E.; Riggs, A. D.; O'Connor, T. R., Nucleotide excision repair eliminates unique DNA-protein cross-links from mammalian cells. *J. Biol. Chem.* **2007,** *282*, 22592-604.

29. Ziegler, K.; Bui, T.; Frisque, R. J.; Grandinetti, A.; Nerurkar, V. R., A rapid in vitro polyomavirus DNA replication assay. *J. Virol. Methods* **2004,** *122* (1), 123-7.

30. Hong, H.; Cao, H.; Wang, Y., Formation and genotoxicity of a guanine cytosine intrastrand cross-link lesion in vivo. *Nucleic Acids Res.* **2007,** *35*, 7118-7127.

Chapter 4

Formation of Carboxymethyl Phosphotriester Adducts in DNA

INTRODUCTION

DNA carries the genetic blueprint of living organisms; however, the limited chemical stability of DNA renders it susceptible to damage by genotoxic agents, potentially compromising genomic integrity.[1] *N*-nitroso compounds (NOCs), a class of human carcinogens present in food, tobacco, and environmental sources, are capable of inducing cancer in many animal species, including higher primates.[2-4] Intragastric nitrosation is a widely accepted mechanism for cancer development from NOC exposure.[5] In this vein, it was estimated that endogenous NOCs constitute 45-75% of total NOC exposure in humans, and augmented NOC exposure is associated with an increased risk of developing noncardia gastric cancer.[2,6]

Metabolic activation of various NOCs, e.g., *N*-(*N*′-acetyl-L-prolyl)-*N*-nitrosoglycine (APNG), azaserine, and *N*-nitrosoglycholic acid (NOGC), can give rise to diazoacetate, which can form carboxymethyldiazonium and methyldiazonium ions; these ions can react with DNA to yield carboxymethyl and methyl adducts, respectively (Scheme 4-1).[4, 7-9] In addition, other agents, e.g., *N*-nitrosoglycine, *N*-nitrososarcosine, and mesyloxyacetic acid, can elicit carboxymethylation of DNA directly.[4, 7] Earlier research by Wogan *et al.*[10] showed that oral treatment of male Fischer rats with nitrosated bile acid conjugates *N*-nitrosoglycholic acid (NOGC) and *N*-nitrosotaurocholic acid (NOTC) can result in liver and gastric cancers in 54-70% and 12-13% of animals, respectively. Harrison *et al.*[11] later detected O^6-carboxymethyl-2′-deoxyguanosine (O^6-CMdG) and, to a lesser extent, O^6-

methyl-2′-deoxyguanosine (O^6-MedG), in calf thymus DNA treated with APNG. Adenine, cytosine, and thymine can also be carboxymethylated, where Wang *et al.*[12-13] detected, by using LC-MS/MS, N^6-carboxymethyl-2′-deoxyadenosine (N^6-CMdA), N^4-carboxymethyl-2′-deoxycytidine (N^4-CMdC), *N*3- and O^4-carboxymethylthymidine (*N*3- and O^4-CMdT) in calf thymus DNA post diazoacetate incubation. Moreover, Gottschalg *et al.*[14] found that the mutation profiles emanating from the replication of a diazoacetate-treated human *p53* gene-containing plasmid in yeast cells are strikingly similar to *p53* mutations at non-CpG sites in gastrointestinal (GI) tract tumors, suggesting that DNA carboxymethylation contributes to increased GI cancer risk. In this vein, a very recent study also documented the presence of gene clusters for azaserine biosynthesis in some human pathogens, including *Salmonella enterica*.[15]

Most studies of NOC-induced DNA damage focus on nucleobase modifications owing to their miscoding potentials during DNA replication and transcription. However, alkylating agents can also attack one of the non-bridging oxygen atoms of the internucleotide phosphate group, forming DNA backbone adducts known as alkyl phosphotriester (alkyl-PTE) lesions.[16] For instance, methyl- and ethyl-PTEs (Me- and Et-PTEs) represent 12-17% and 55-57% of total DNA alkylation products elicited by *N*-methyl-*N*-nitrosourea and *N*-ethyl-*N*-nitrosourea, respectively.[17] Furthermore, Ma *et al.*[18] uncovered the distribution of pyridylhydroxybutyl (PHB) adducts in rats treated with the tobacco carcinogen nicotine-derived nitrosamine ketone (NNK) and found that PHB-PTEs accounted for 34-40% and 38-55% of total detectable pyridine-containing DNA adducts in rat liver and lung, respectively. The distribution frequencies of pyridyloxobutyl (POB)-PTEs were also quantified, and they constituted 40-44% and 32-48% of total pyridine DNA adducts in the liver and lung of rats, respectively.[18]

While the formation of 2-hydroxyethyl-PTE adducts renders DNA susceptible to form strand breaks,[19] most alkyl-PTEs were shown to persist in mammalian tissues.[16] Den Engelse *et al.*[20] observed that the half-lives ($t_{1/2}$) for Me- and Et-PTEs were 1 and 10-15 weeks in the kidney and lung tissues of mice treated with methylating and ethylating agents, respectively. Furthermore, POB-PTE and PHB-PTE adducts were detected in rat tissues 70 weeks after chronic NNK exposure.[18, 21] The persistence of these adducts in mammalian tissues suggests that they are poorly repaired and are likely encountered by DNA replication and transcription machineries. Thus, it is important to investigate the formation and biological consequences of alkyl-PTEs.

Recently, there has been growing interest in understanding the biological consequences of alkyl-PTE lesions in DNA. Depending on which oxygen atom is alkylated, S_p or R_p diastereomers of PTEs can form due to the tetrahedral configuration of the phosphate group. The alkyl group of the S_P diastereomer projects perpendicularly out from the DNA double helix, whereas that of the R_P counterpart projects into DNA's major groove.[16] Wu et al.[22] found that Me-PTE in the S_P configuration (S_P-Me-PTE) was less blocking to *Escherichia coli* DNA replication machinery than the R_P counterpart. Notably, the S_P-Me-PTE adduct flanked by two thymidine residues induces TT→GT and TT→GC mutations at the dinucleotide site in *E. coli,*[22] and the replication outcome for the Me-PTE lesions was highly independent of the 5′ flanking nucleoside.[23] In addition, replication across POB-PTE adducts was shown not to induce any mutations, though R_p adducts slightly impeded replication bypass in *E. coli*.[24] In human cells, the alkyl-PTE lesions (Me, Et, *n*Pr, *n*Bu) flanked by two thymidine residues did not elicit mutations, but moderately hindered DNA replication.[25] Moreover, the S_P diastereomers of Me- and *n*Pr-PTEs did not appreciably perturb transcriptional bypass in human cells, whereas those in the R_P

configuration conferred moderate and strong blockages, respectively;[26] none of the lesions elicited mutant transcripts in human cells.[26]

Together, recent studies suggest that exposure to some endogenous and exogenous NOCs may give rise to DNA carboxymethylation; however, it has not been examined whether this modification can occur on the phosphate backbone. Here, we synthesized oligodeoxyribonucleotides (ODNs) bearing a site-specifically inserted CM-PTE lesion, and we revealed, by using LC-MS/MS analysis, the formation of CM-PTE adducts in calf thymus DNA upon diazoacetate exposure. Our work sets the stage for assessing the *in vivo* formation and biological consequences of these adducts.

MATERIALS AND METHODS

Materials

Chemicals and reagents were purchased from Sigma-Aldrich (St. Louis, MO, USA) or Fisher Scientific (Pittsburgh, PA, USA) unless otherwise specified. Common reagents and phosphoramidite building blocks for solid-phase DNA synthesis were purchased from Glen Research (Sterling, VA, USA). 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) was from Oakwood Products, Inc. (West Columbia, SC, USA). Calf thymus DNA, phosphodiesterases I and II were obtained from Sigma-Aldrich. Nuclease P1 and quick calf intestinal phosphatase (CIP) were purchased from US Biologicals (Salem, MA) and New England Biolabs (Ipswich, MA), respectively.

Synthesis of the methyl ester of bis(diisopropylamino)carboxymethylphosphine (2). Bis(diisopropylamino)chlorophosphine (**1,** 300 mg, 1.12 mmol) was suspended in anhydrous diethyl ether (12 mL) in an ice bath under argon atmosphere. A solution of methyl glycolate (1.2 eq) and triethylamine (2 eq) in 3 mL anhydrous diethyl ether was added dropwise to the reaction vessel. The mixture was stirred overnight and then filtered. The resulting precipitate was washed with ether, and the supernatant was pooled and evaporated under reduced pressure to yield the crude product, which was purified by silica gel chromatography using ethyl acetate/hexane/triethylamine (49:49:2) as the mobile phase to give **2** (yield 65%).

^{1}H-NMR (600 MHz, $CDCl_3$, 25°C): δ 4.17 (d, J = 9.4 Hz, 2H), 3.72 (s, 3H), 3.54 (ddd, J = 13.6, 11.2, 6.6 Hz, 4H), 1.17 (dd, J = 6.7 Hz, 24H). ^{31}P-NMR (242 MHz, $CDCl_3$, 25°C): δ 128 (s, 1P). ESI-MS: $[M + H]^+$ *calc m/z 321.2, found 321.4.*

Synthesis of the methyl ester of carboxymethyl phosphotriester of dT phosphoramidite building block (5)

5′-*O*-(4,4′-dimethoxytrityl)thymidine (**4)** was synthesized based on previously published procedures.[27] A solution of compound **2** (300 mg, 0.936 mmol) and 1H-tetrazole (0.33 eq) in dichloromethane in a round-bottom flask was stirred under argon atmosphere for 30 min. Compound **4** (0.309 mmol, 0.33 eq) was subsequently added dropwise to the flask. The mixture was stirred at room temperature for 4 h and concentrated under reduced pressure for flash column chromatography purification using ethyl acetate/hexane/triethylamine (49/49/2) as the mobile phase to obtain the final product (yield: 77%).

^{1}H-NMR (600 MHz, $CDCl_3$): δ 8.33 (s, 1H), 7.60 (d, *J* = 8.9 Hz, 1H), 7.40 (d, *J* = 7.8 Hz, 2H), 7.31 – 7.26 (m, 6H), 7.22 (t, *J* = 7.3 Hz, 1H), 6.82 (d, *J* = 8.4 Hz, 4H), 6.40 (dt, *J* = 7.9, 5.6 Hz, 1H), 4.79 (dtd, *J* = 13.5, 6.1, 2.9 Hz, 1H), 4.64 (dd, *J* = 16.3, 9.8 Hz, 1H), 4.47 (dd, *J* = 16.3, 13.1 Hz, 1H), 4.11 (q, *J* = 7.2 Hz, 3H), 4.07 (t, *J* = 11.4 Hz, 1H), 3.78 (s, 6H), 3.77 (s, 3H), 3.58 – 3.51 (m, 2H), 2.54 (m, 1H), 2.39 – 2.31 (m, 1H), 2.04 (s, 3H), 1.26 (dd, *J* = 7.0, 4.6 Hz, 12H)*;* ^{31}P-NMR (242 MHz, DMSO): δ 149.74, 149.15; *ESI-MS: calc* $[M+H]^+$ *m/z 764.3, found m/z 764.1.*

The NMR and ESI mass spectra of the above-mentioned products are shown in Figures 4-1 to 4-6.

ODN synthesis

The 12-mers 5′-ATGGCT(CM)TGCTAT-3′ & 5′-ATGGCT(CM)AGCTAT-3′, and dimers, 5′-T(CM)T-3′ & 5′-T(CM)A-3′, were synthesized on a Beckman Oligo 1000M DNA synthesizer (Fullerton, CA) at a 1-μmol scale. The lesion-bearing phosphoramidite (**5**) was dissolved in anhydrous acetonitrile at a concentration of 0.1 mM. Commercially available ultramild phosphoramidites for unmodified nucleosides were purchased from Glen Research Inc. (Sterling, VA, USA) and used for ODN synthesis following standard protocols. The ODNs were cleaved from controlled pore glass (CPG) support, deprotected, and the methyl ester was converted to the corresponding carboxylic acid by treating with 0.1 M NaOH at room temperature for 24 h. The aqueous solvent was removed using a Speed-vac and subsequently dissolved in Milli-Q water for HPLC purification.

HPLC purification of ODNs

HPLC purifications of ODNs were performed on an Agilent 1100 system with a BDS Hypersil C18 column (4.6 x 250 mm, 5 μm in particle size and 100 Å in pore size, Thermo Fisher Scientific, San Jose, CA). The mobile phases consisted of 50 mM triethylammonium acetate (pH 6.8, solution A) and a mixture of solution A and acetonitrile (70/30, v/v, solution B). A gradient of 5-20% B in 5 min, 20-30% B in 15 min, 30-40% B in 20 min, 40-75% B in 25 min was employed at a flow rate of 0.5 mL/min. The HPLC traces for the purification of the dimer and 12-mer lesion-bearing ODNs are shown in Figure 4-7.

Potassium diazoacetate (KDA) synthesis and treatment with calf thymus DNA

KDA was synthesized following previously published procedures.[28] Briefly, 200 μL of ethyl diazoacetate (2 mmol) was dissolved in a 2-mL solution of 2.0 M KOH and stirred at room temperature for 4.5 h until the solution became homogenous. The concentration of the KDA solution was estimated by assuming that the reaction was quantitative. Calf thymus DNA was treated with freshly synthesized KDA to a final concentration of 20 mM in PBS buffer (pH 7.3), and the mixture was incubated at 37*°C* for 8 h. The resulting DNA was desalted using ethanol precipitation, dried by Speed-vac, and stored at -20*°C until use.*

Enzymatic digestion of KDA-treated calf thymus DNA

The digestion consisted of a 2-step procedure. One unit of nuclease P1, 0.00125 U of phosphodiesterase II, and 4.5 μL of a 10x buffer solution containing 300 mM sodium acetate (pH 5.0) and 10 mM zinc acetate, were added to 30 μg of KDA-treated calf thymus DNA. Milli-Q water was added to obtain a final volume of 45 μL. The digestion proceeded at 37°C for 24 h. To the solution were added 1 U of quick CIP, 0.0025 U of

phosphodiesterase I, and 6 μL solution of 0.5 M Tris-HCl (pH 8.9). Milli-Q water was added to reach a final volume of 60 μL. The digestion continued at 37°C for 2 h, and the enzymes were removed by chloroform extraction. The aqueous layer was desalted, dried by a Speed-vac, and the dried residues were reconstituted in 100 μL of water for LC-MS/MS analysis.

Mass spectrometry (MS) and NMR

Electrospray ionization (ESI) MS analyses of synthetic products were carried out on an LCQ Deca XP ion-trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA). The solvent used for electrospray was a mixture of acetonitrile and water (50:50, v/v). The temperature of the heated capillary was 275°C, and the spray voltage was 3.0 kV. ^{1}H- and ^{31}P-NMR spectra were recorded on a Bruker Avance 600 spectrometer.

LC-MS/MS

LC-MS/MS experiments for the characterizations and identification of T(CM)T and T(CM)A synthetic standards, and T(CM)T and T(CM)A in the digestion mixture of KDA-treated calf thymus DNA were performed on an Agilent 1200 capillary HPLC pump (Agilent Technologies) coupled with an LTQ linear ion-trap mass spectrometer (Thermo Fisher Scientific). A Zorbax SB C18 column (0.5 x 150 mm, 5 μm in particle size and 80 Å in pore size, Agilent Technologies) was used for separation. The mobile phases were 10 mM ammonium formate (Solution A, pH 6.5) and methanol (solution B). A linear gradient of 5% B in 5 min, 5-70% B in 25 min, 70-90% B in 1 min at a flow rate of 0.800 μL/min was used. The ESI spray voltage was 3.0 kV and the metal capillary was held at a temperature of 275°C.

All other DNA adducts and the spiked digestion mixture were identified and analyzed using an Thermo Scientific Easy-nLC 1000 (Thermo Fisher Scientific) coupled to a LTQ XL linear ion-trap mass spectrometer (Thermo Fisher Scientific). Online nLC separation was conducted by employing a homemade trapping (150 μm × 40 mm) and analytical (75 μm × 200 mm) columns packed with Zorbax SB-C18 stationary phase material (5 μm in particle size and 200 Å in pore size, Michrome BioResource, Auburn, CA). Mobile phases A and B were 0.1% formic acid in doubly distilled H_2O and methanol, respectively. Initially, the sample was loaded onto the trapping column with mobile phase A at a flow rate of 2.50 μL/min for 6 min. A linear gradient of 5-20% B in 5 min, followed by 20-60% B in 30 min, was used for separation, where the flow rate was 300 nL/min. The LTQ XL mass spectrometer was operated in the positive-ion mode with an electrospray voltage of 3.0 kV and a heated capillary temperature of 275°C.

RESULTS

Syntheses of CM-PTE-containing ODNs

Although previous literature documented the formation of carboxymethylated nucleobase adducts and their impact on cellular DNA replication and transcription,[22-26] it remains unexplored whether this modification occurs on DNA phosphate backbone. The objectives of this study were to synthesize CM-PTE-containing ODNs and to examine the occurrence of CM-PTE adducts induced by diazoacetate in isolated DNA. We reason that the outcome of these experiments will provide a basis for future investigations about the formation of CM-PTE lesions *in vivo* and their contributions to adverse human health effects associated with endogenous NOC exposure.

We first developed a strategy to synthesize the methyl ester of CM-PTE phosphoramidite of thymidine by first preparing a phosphitylating species (**2**). To this end, the chlorine atom in bis(diisopropylamino)chlorophosphine (**1)** was substituted with methyl glycolate. After flash column chromatography purification, compound **2** was coupled to the 3′-hydroxyl group of 5′-*O*-(4,4′-dimethoxytrityl)thymidine (5′-*O*-DMT-dT, **4**) using 1H-tetrazole to yield the methyl ester of CM-PTE phosphoramidite of dT (**5**, Scheme 4-2). Compounds **2** and **5** were purified by flash column chromatography, and their structures were supported by ESI-MS, ^{1}H- and ^{31}P-NMR analyses (Figures 4-1 to 4-6).

Incorporation of CM-PTE into ODNs and their characterizations

To identify CM-PTE adducts in DNA and to later understand their biological implications, it is necessary to synthesize ODNs harboring a site-specifically inserted CM-PTE modification. The lesion-bearing phosphoramidite building block was synthesized by employing the aforementioned procedures and incorporated into the dimers T(CM)T, T(CM)A, and 12-mers, 5′-ATGGCT(CM)TGCTAT-3′ and 5′-ATGGCT(CM)AGCTAT-3′, by following standard solid-phase synthesis procedures, where ‘CM’ represents the inserted carboxymethyl functionality on the phosphate group. After solid-phase syntheses, the ODNs were cleaved from solid support and deprotected, where the methyl ester was converted to the corresponding carboxylic acid derivative by incubating with 0.1 M NaOH.[12]

ESI-MS and MS/MS analysis of the ODNs confirmed the sequence and site-specific insertion of the CM-PTE modification. In this vein, the MS/MS of the $[M + H]^+$ ions of T(CM)T and T(CM)A, (at *m/z* 605.2 and 614.2, respectively), showed cleavage at the 3′ C-O bond (a_1 and w_1 ions) at *m/z* 225.1 & 380.9 and *m/z* 225.1 & 390.1, (Figures 4-8 and

3-9). *N*-glycosidic bond cleavage was also observed, resulting in ions corresponding to the neutral loss of their respective 5′ and 3′ nucleobases. For instance, collisional activation of the $[M + H]^+$ ion of T(CM)A produced ions of *m/z* 488.2 and 479.2, which emanate from the neutral losses of a thymine and an adenine, respectively (Figures 4-8 and 3-9). Fragmentation of synthesized dimers also produced the c_1 ion of *m/z* 363.0, stemming from cleavage of the 3′ P-O bond of the internucleotide phosphate group (Figures 4-8 and 4-9). The assignment of this ion to c_1 is supported by MS^3 analysis, which revealed the formation of the dominant $a_1 – H_2O$ ion (Figures 4-8 and 4-9). Moreover, we detected a fragment ion arising from the further loss of a CH_2COO from the c_1 ion in the MS/MS of the $[M + H]^+$ ions of the two dimers, which aligns with the presence of a carboxymethyl moiety in the two CM-PTE lesions (Figures 4-8 and 4-9).

The measured *m/z* values for the multiply deprotonated ions of the two 12-mer ODNs, 5′-ATGGCT(CM)AGCTAT-3′ and 5′-ATGGCT(CM)TGCTAT-3′, are consistent with their calculated ones. The MS^2 of $[M - 3H]^{3-}$ ions (*m/z* 1238.5 and 1235.7, Figure 4-10) showed the formation of the $[a_n\text{-Base}]$ and their corresponding w_n ions, which emanate from the cleavages of the *N*-glycosidic and 3′-C-O bonds of the same nucleoside and carry the 5′ and 3′ termini of the ODN, respectively.[29] For example, $[a_3\text{-G}]$, $[a_4\text{-G}]$, $[a_5\text{-C}]$, $[a_7\text{-A}]$, $[a_8\text{-G}]$, $[a_9\text{-G}]$, $[a_{11}\text{-A}]$, w_2, w_3, w_4, w_5, w_7, w_8, w_9, and w_{11} ions were observed for 5′-ATGGCT(CM)AGCTAT-3′. Fragment ions bearing the modification, i.e., $[a_7\text{-A}]$, $[a_8\text{-G}]$, $[a_9\text{-G}]$, $[a_{11}\text{-A}]$, w_7, w_8, w_9, and w_{11}, exhibited a 58 Da higher in mass than the corresponding fragment ions from the unmodified ODN, which is consistent with the site of CM-PTE in the ODN. Similar results were obtained for 5′-ATGGCT(CM)TGCTAT-3′ (Figure 4-10).

KDA induced CM-PTE adducts in calf thymus DNA

To identify CM-PTE adducts, we treated calf thymus DNA with 20 mM KDA, and subsequently precipitated the DNA with ethanol, and digested it with a cocktail of four enzymes (see Experimental Section). The carboxymethyl group on the phosphate backbone blocks the abilities of the phosphodiesterases and endonuclease (i.e., nuclease P1) to cleave the phosphodiester bonds flanking the lesion site, leading to a mixture containing mononucleosides and carboxymethylated dinucleoside monophosphates. After the enzymes were removed, the digestion mixtures were desalted and subjected to LC-MS/MS analysis, where we monitored the fragmentations of the $[M+H]^+$ ions of CM-PTE in all 16 possible dinucleoside monophosphates with the flanking nucleobases being an A, C, G, or T. The identifications of T(CM)T and T(CM)A were based on similar elution times and MS/MS as observed for the $[M + H]^+$ ions of the synthetic standards (Figures 4-11 and 4-12). Moreover, similar LC-MS/MS analyses of the digestion mixture of KDA-treated calf thymus DNA spiked with standard T(CM)T and T(CM)A revealed augmented signal intensities for the two lesions, further supporting their formation in KDA-treated isolated DNA (Figures 4-13 and 4-14). In this vein, it is of note that, while we were able to resolve the S_P and R_P diastereomers of synthetic standards by off-line HPLC with a BDS Hypersil C18 column (Figure 4-7), the two diastereomers co-elute under on-line LC-MS/MS conditions on a Zorbax SB C18 column (Figure 4-12).

The identifications of CM-PTE lesions in other dinucleoside monophosphates were based on their characteristic MS/MS. In this vein, collisional activation of the $[M + H]^+$ ions of the majority of the CM-PTE-bearing dinucleoside monophosphates gives rise to the formation of the w_1 ion, which defines the identity of the 3′ nucleoside (Figures 4-15 to 4-23). Furthermore, MS/MS/MS monitoring the further fragmentation of the w_1 ion further

confirmed the identities of the 3' nucleoside in these adducts, which aided in distinguishing sequence isomers (Figure 4-24 and 4-25). This information, together with the *m/z* value for the precursor ion, enables the determination of the 5' nucleoside flanking the CM-PTE. In addition, the nucleobase compositions of the CM-PTE lesions of the dinucleoside monophosphates are substantiated by the fragment ions stemming from the neutral losses of nucleobases.

DISCUSSION

Exposure to endogenous and exogenous NOCs is implicated to confer increased risk of GI track cancers.[6, 30-31] Some endogenous NOCs can be metabolically activated to form diazoacetate, which has been shown to alkylate nucleobases in DNA.[7, 10-11] In addition, the mutation profiles observed at non-CpG sites in human *p53* gene within human GI cancer are strikingly similar to those induced from the replication of a diazoacetate-treated human *p53* gene-containing plasmid in yeast cells,[14] suggesting that carboxymethyl DNA adducts induced by endogenous NOCs are significant contributors to *p53* mutations observed in GI cancers. This is corroborated by recent identifications of the putative gene clusters for azaserine biosynthesis in human pathogens, including *Salmonella enterica* and *Serratia marcescens*.[15] However, it was not clear if carboxymethylation occurs on the phosphate backbone of DNA or if it contributes to mutagenesis or carcinogenesis.

Pyridyloxobutyl and pyridylhydroxybutyl nucleobase and phosphate adducts can result from NNK and NNN exposure.[32] NNK can be enzymatically converted to 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), where both NNK and NNAL are converted by cytochrome P450 enzymes into their corresponding alkylating species, POB and PHB diazonium ions, respectively.[32] These ions ultimately induce POB and PHB

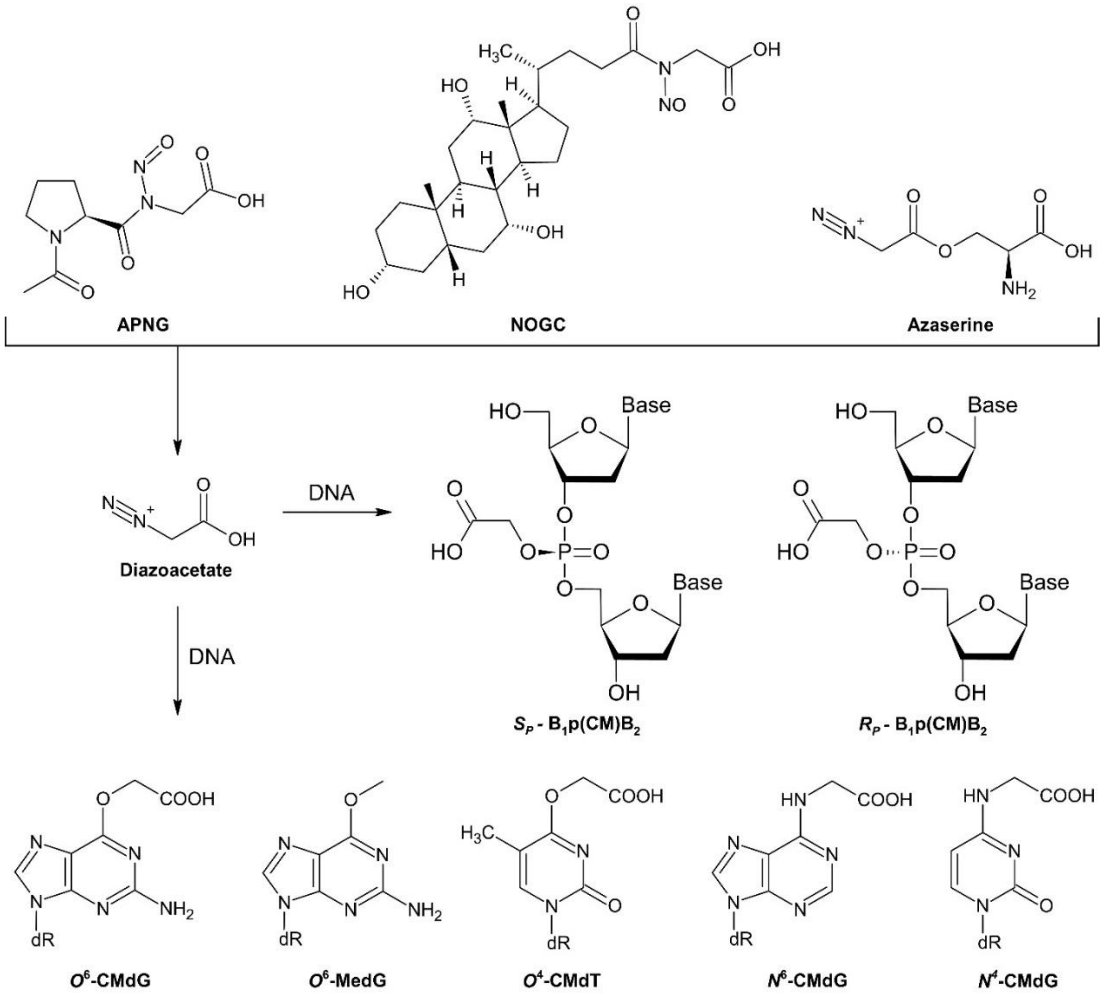
phosphate adducts. Additionally, the extent of PTE formation is thought to parallel that of nucleobase *O*-alkylation products.[16-18] In support of this notion, carboxymethylation of oxygen atoms in nucleobases was shown to occur in calf thymus DNA and genomic DNA of cultured human cells.[7, 9, 11-14] Altogether, this suggests that diazoacetate can induce carboxymethylation of the backbone phosphate in DNA, which we have demonstrated in this study.

The construction of ODNs bearing a CM-PTE modification will unlock the ability to investigate the biological consequences of the adduct. Our laboratory previously established synthetic pathways for ODNs containing N^6-CMdA, N^4-CMdC, *N*3-/O^4-CMdT, and confirmed their formation in isolated DNA.[12-13] These ODNs were also deployed to examine how these adducts influence the efficiency and fidelity of DNA replication by polymerase η using an *in vitro* primer extension assay. The polymerase was able to bypass N^6-CMdA and, to a lesser extent, N^4-CMdC, producing full-length products.[33] However, thymidine lesions severely hindered bypass efficiency, producing extension products up to three nucleotides past the adduct.[33] This observation was supported by a later study investigating how translesion synthesis polymerases pol η, pol ι, pol κ, and pol ξ contribute to the replicative bypass of these lesions in cultured human cells. N^6-CMdA and N^4-CMdC did not block DNA replication, nor was the process mutagenic, though *N*3-CMdT, O^4-CMdT, and O^6-CMdG moderately impeded DNA replication and produced significant frequencies of T→A (81%), T→C (68%), and G→A (6.4%) mutations, respectively.[34]

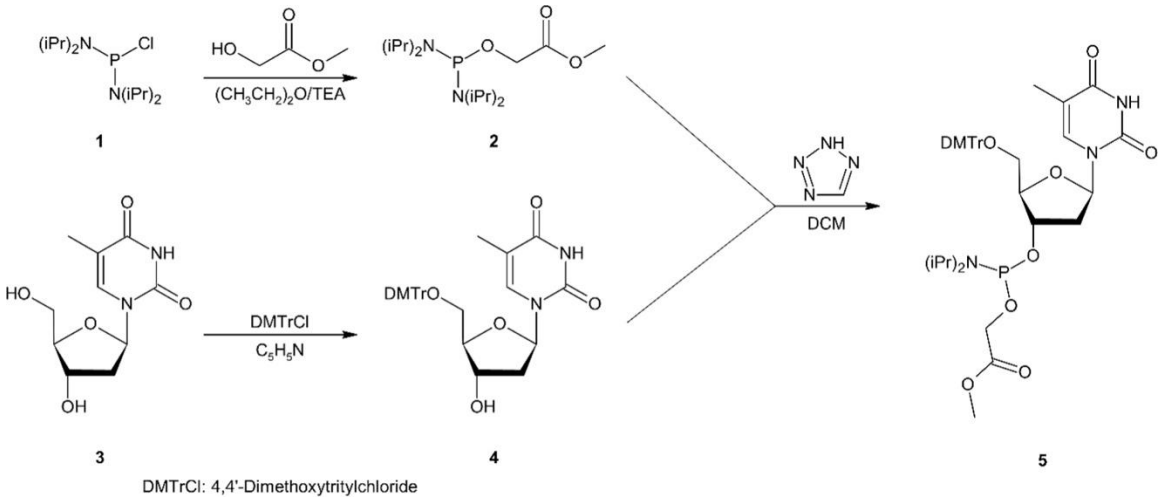
Together, we synthesized authentic CM-PTE ODNs and demonstrated, by using LC-MS/MS analysis, that carboxymethylation of the phosphate backbone can occur across all 16 possible nucleobase combinations in calf thymus DNA upon diazoacetate exposure.

The method for constructing CM-PTE oligomers can assist in synthesizing stable isotope-labeled derivatives to quantify the lesions in human cells, as described previously for O^6-CMdG and N^6-CMdA.[9] Hence, this study laid the foundation for assessing the occurrence and consequences of CM-PTE adducts.

FIGURES AND SCHEMES



Scheme 4-1. Induction of carboxymethyl and methyl adducts in DNA upon diazoacetate exposure from *N*-(*N*′-acetyl-L-prolyl)-*N*-nitrosoglycine (APNG), *N*-nitrosoglycocholic acid (NOGC), and azaserine.



Scheme 4-2. Synthesis of a Methyl Ester of CM-PTE dT Phosphoramidite Building Block for the Construction of Modified Oligodeoxyribonucleotides

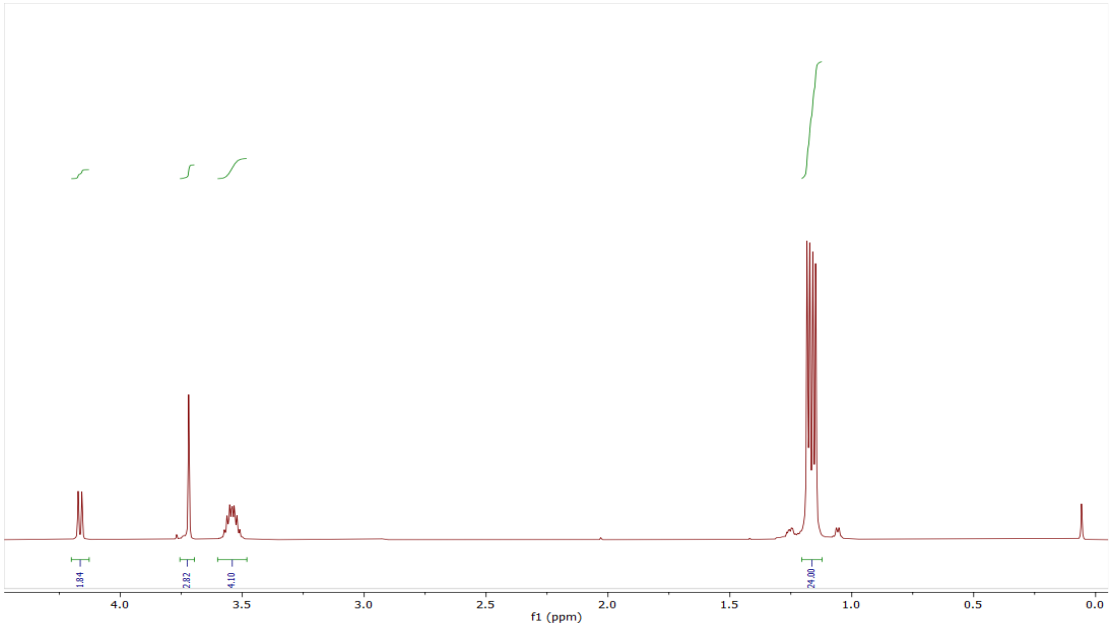


Figure 4-1. ^{1}H NMR spectrum of the methyl ester of bis(diisopropylamino)carboxymethylphosphine (**2**). ^{1}H-NMR (600 MHz, $CDCl_3$, 25°C): δ 4.17 (d, J = 9.4 Hz, 2H), 3.72 (s, 3H), 3.54 (ddd, *J* = 13.6, 11.2, 6.6 Hz, 4H), 1.17 (dd, *J* = 6.7 Hz, 24H).

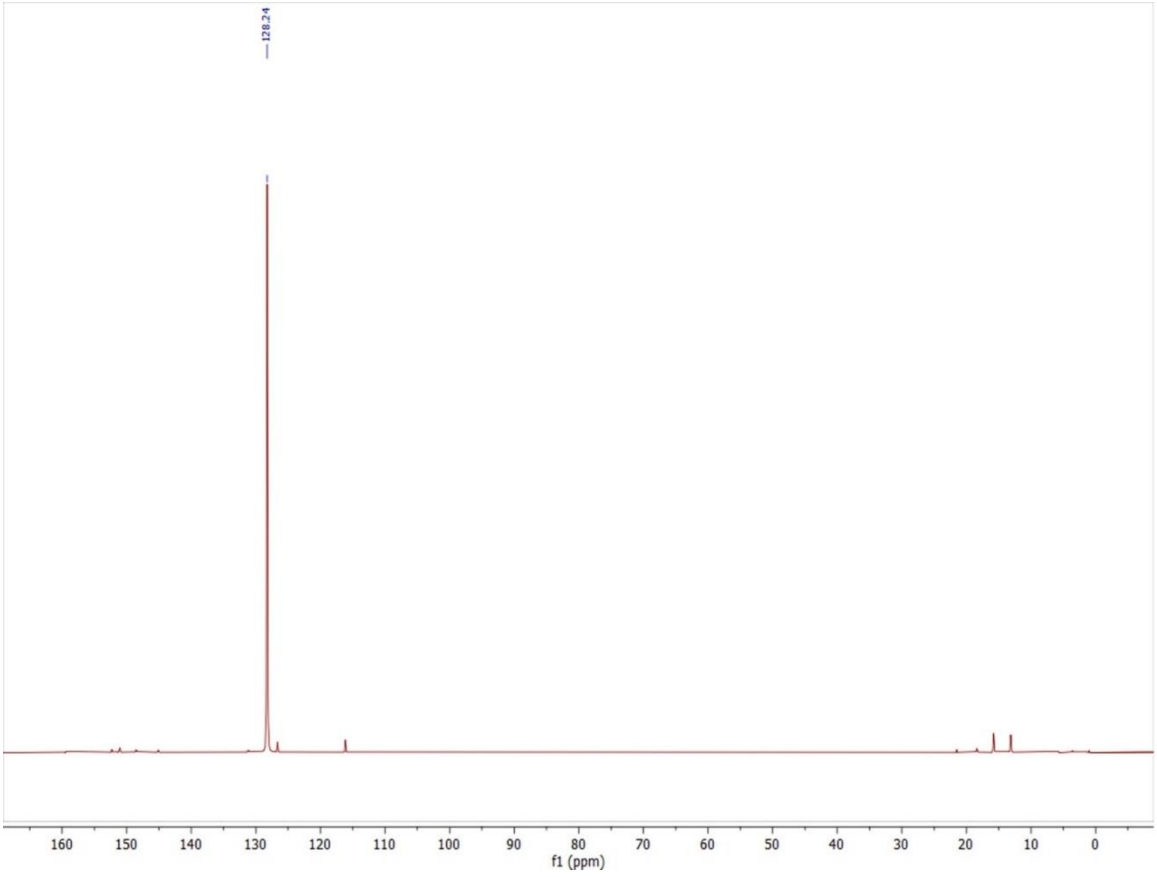


Figure 4-2. ^{31}P NMR spectrum of the methyl ester of bis(diisopropylamino)carboxymethylphosphine (**2**). ^{31}P-NMR (242 MHz, $CDCl_3$, 25°C): δ 128 (s, 1P)

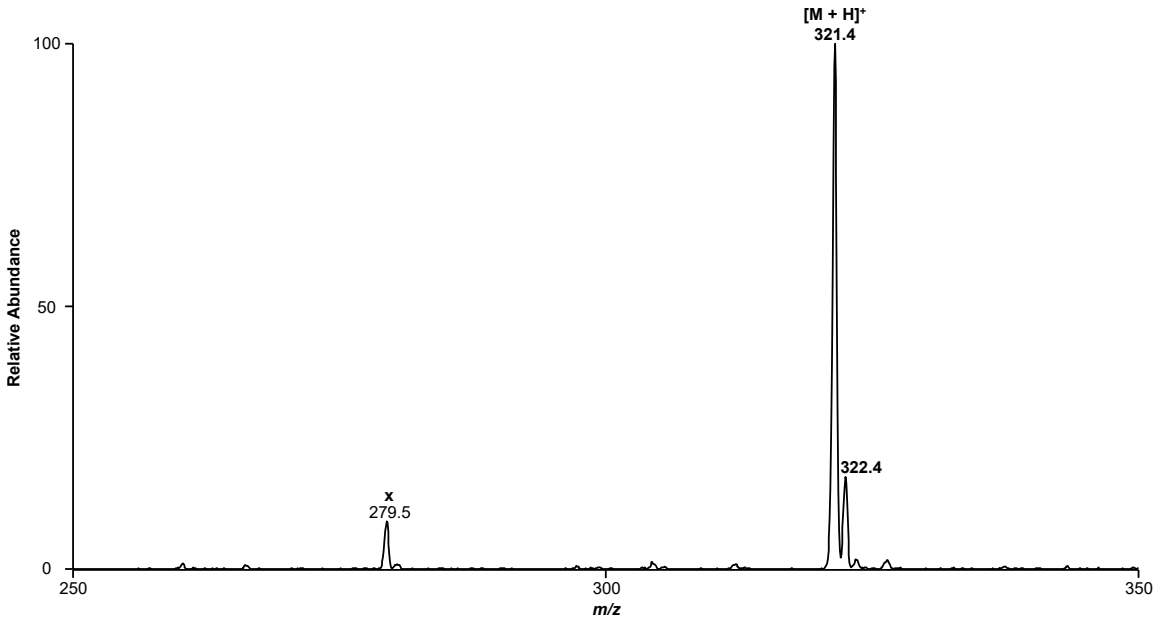


Figure 4-3. Positive-ion ESI-MS of the methyl ester of bis(diisopropylamino)carboxymethylphosphine (**2**). The ion of *m/z* 321.4 corresponds to the $[M + H]^+$ ion (theoretical *m/z* 321.2) of compound **2**.

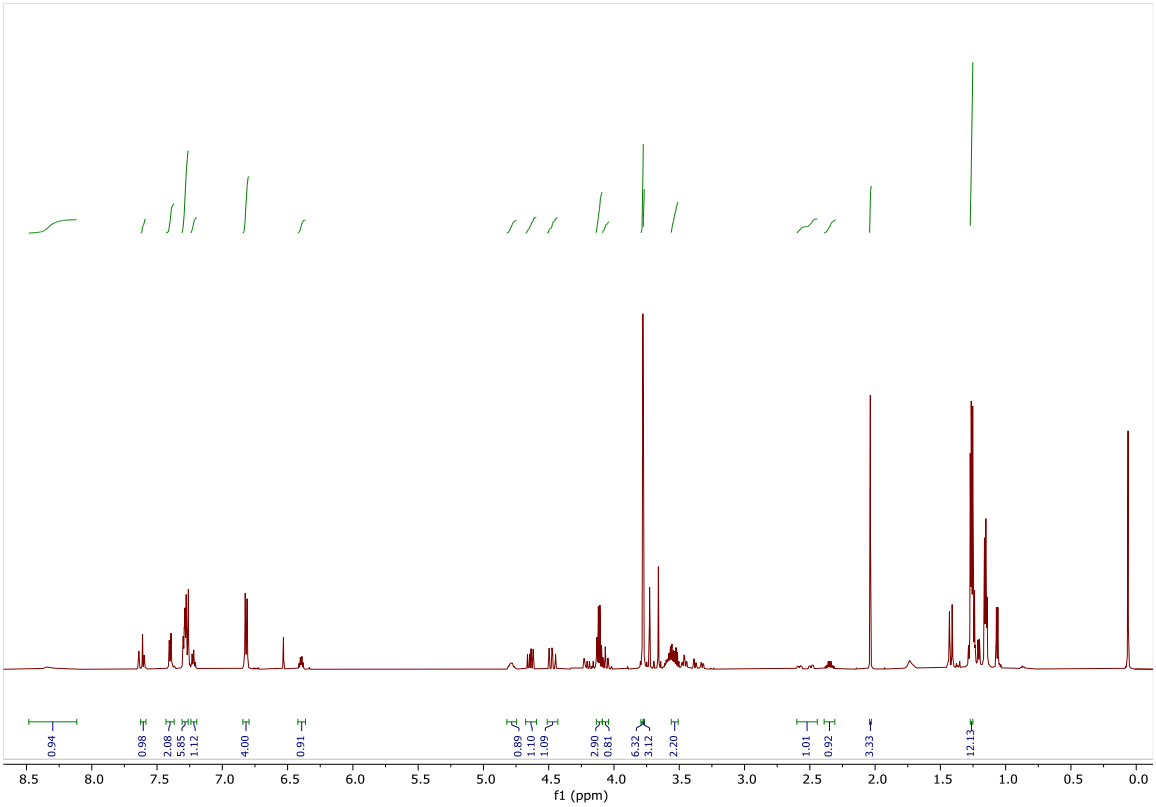


Figure 4-4. ^{1}H NMR spectrum of methyl ester of carboxymethyl phosphotriester of dT phosphoramidite building block (**5**). ^{1}H-NMR (600 MHz, $CDCl_3$): δ 8.33 (s, 1H), 7.60 (d, J = 8.9 Hz, 1H), 7.40 (d, J = 7.8 Hz, 2H), 7.31 – 7.26 (m, 6H), 7.22 (t, J = 7.3 Hz, 1H), 6.82 (d, J = 8.4 Hz, 4H), 6.40 (dt, J = 7.9, 5.6 Hz, 1H), 4.79 (dtd, J = 13.5, 6.1, 2.9 Hz, 1H), 4.64 (dd, J = 16.3, 9.8 Hz, 1H), 4.47 (dd, J = 16.3, 13.1 Hz, 1H), 4.11 (q, J = 7.2 Hz, 3H), 4.07 (t, J = 11.4 Hz, 1H), 3.78 (s, 6H), 3.77 (s, 3H), 3.58 – 3.51 (m, 2H), 2.54 (m, 1H), 2.39 – 2.31 (m, 1H), 2.04 (s, 3H), 1.26 (dd, J = 7.0, 4.6 Hz, 12H).

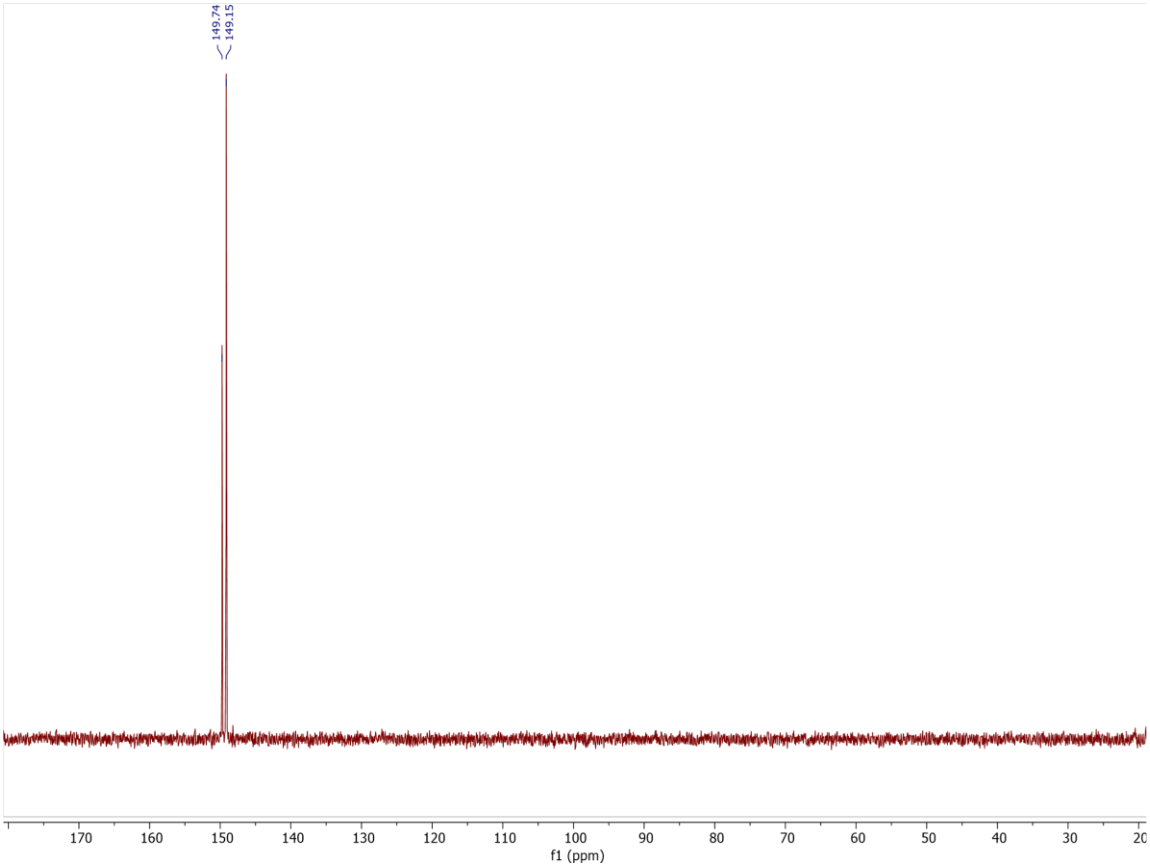


Figure 4-5. ^{31}P NMR spectrum of methyl ester of carboxymethyl phosphotriester of dT phosphoramidite building block (**5**, $CDCl_3$, 242 MHz, 25°C). The two singlet peaks correspond to the two diastereomers of the phosphoramidite.

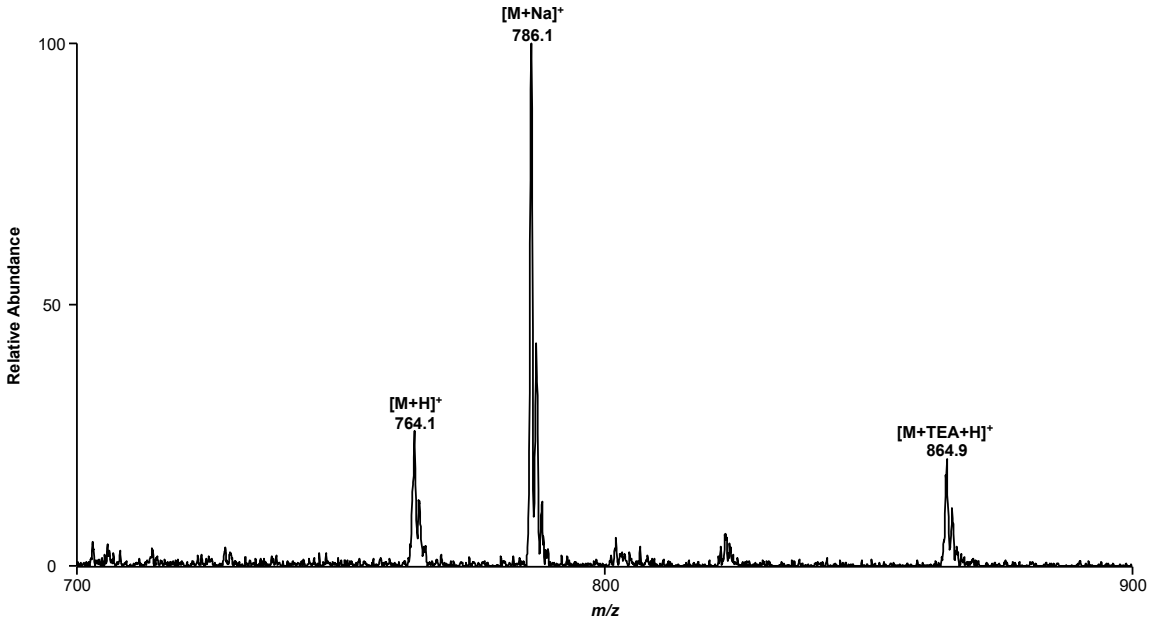


Figure 4-6. ESI-MS of methyl ester of carboxymethyl phosphotriester of dT phosphoramidite building block (**5**). The ion of *m/z* 764.1 corresponds to the $[M + H]^+$ ion of compound **5**.

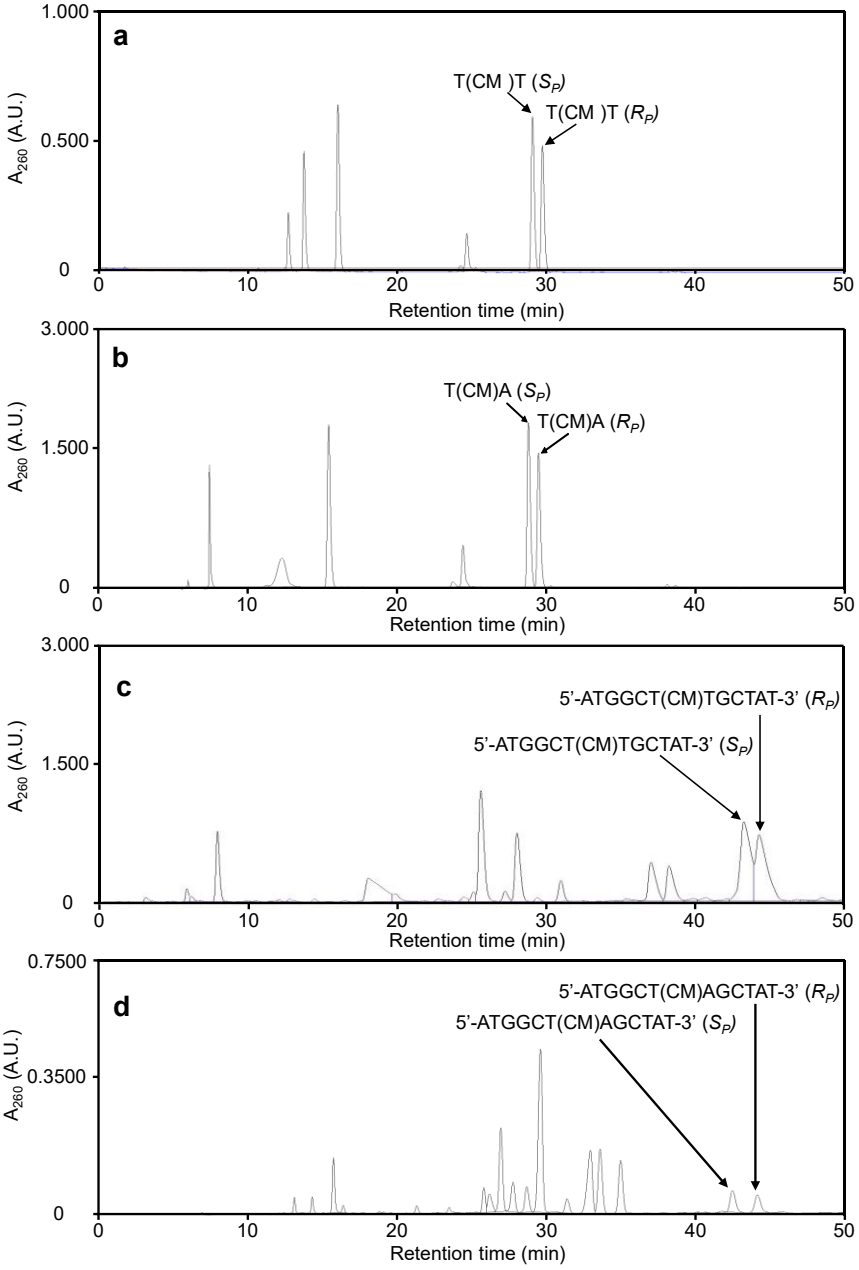


Figure 4-7. HPLC traces of the synthesized dimer (a-b) and 12-mer (c-d) carboxymethyl-phosphotriester ODNs.

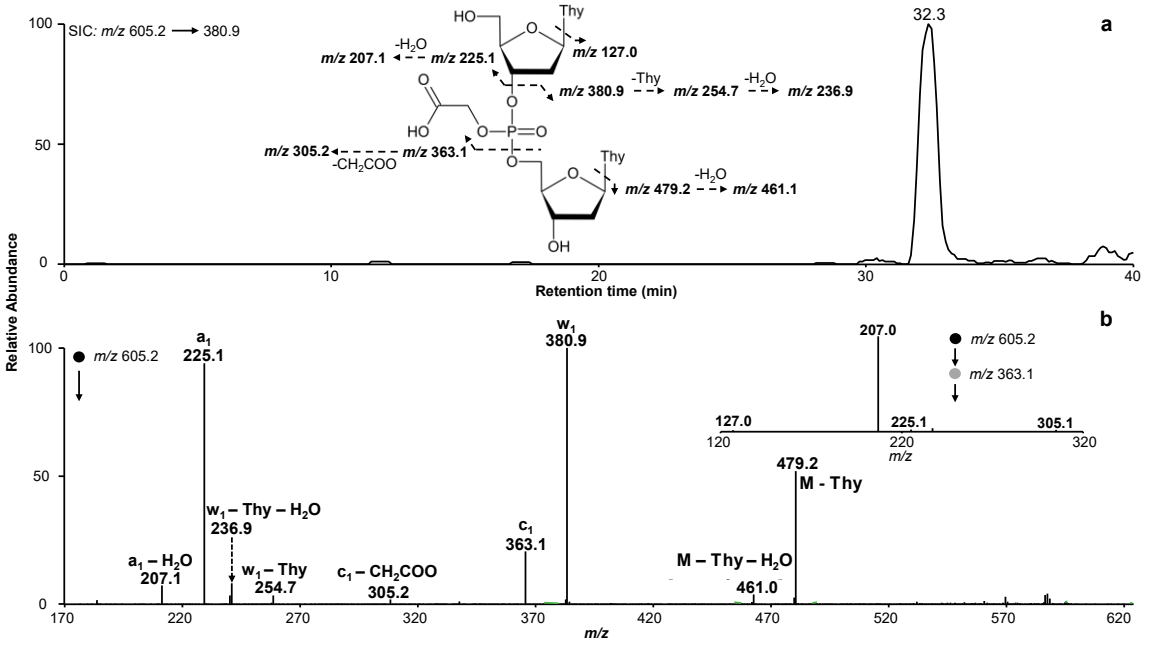


Figure 4-8. LC-MS/MS characterizations of the synthetic T(CM)T standard. (a) SIC for monitoring the *m/z* 605.2 → 380.9 transition, corresponding to the formation of the w_1 ion; (b) MS^2 for the $[M + H]^+$ ion of T(CM)T. Shown in the inset of (a) is the proposed fragmentation pathways for the $[M + H]^+$ ion of T(CM)T, and illustrated in the inset of (b) is the MS^3 of the *m/z* 363.1 ion detected in MS^2, supporting its identity as the c_1 ion.

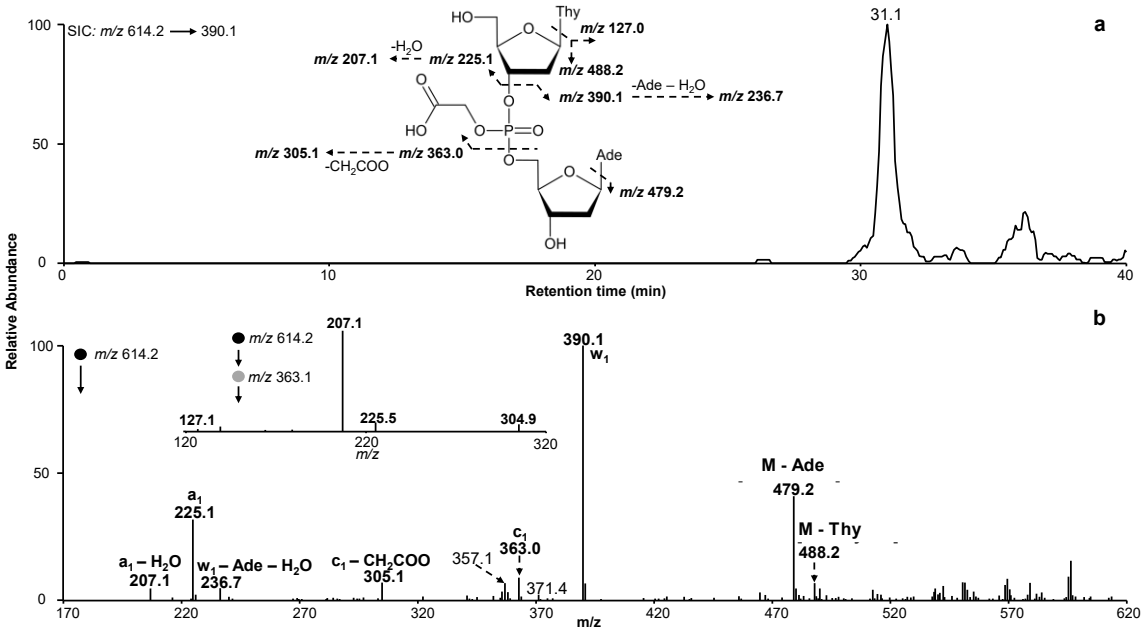


Figure 4-9. LC-MS/MS characterizations of the synthetic T(CM)A standard. (a) SIC for monitoring the *m/z* 614.2 → 390.1 transition, corresponding to the formation of the w_1 ion; (b) MS^2 for the $[M + H]^+$ ion of T(CM)A. Shown in the inset of (a) is the proposed fragmentation pathways for the $[M + H]^+$ ion of T(CM)A, and displayed in the inset of (b) is the MS^3 of the *m/z* 363.0 ion detected in MS^2, supporting its identity as the c_1 ion.

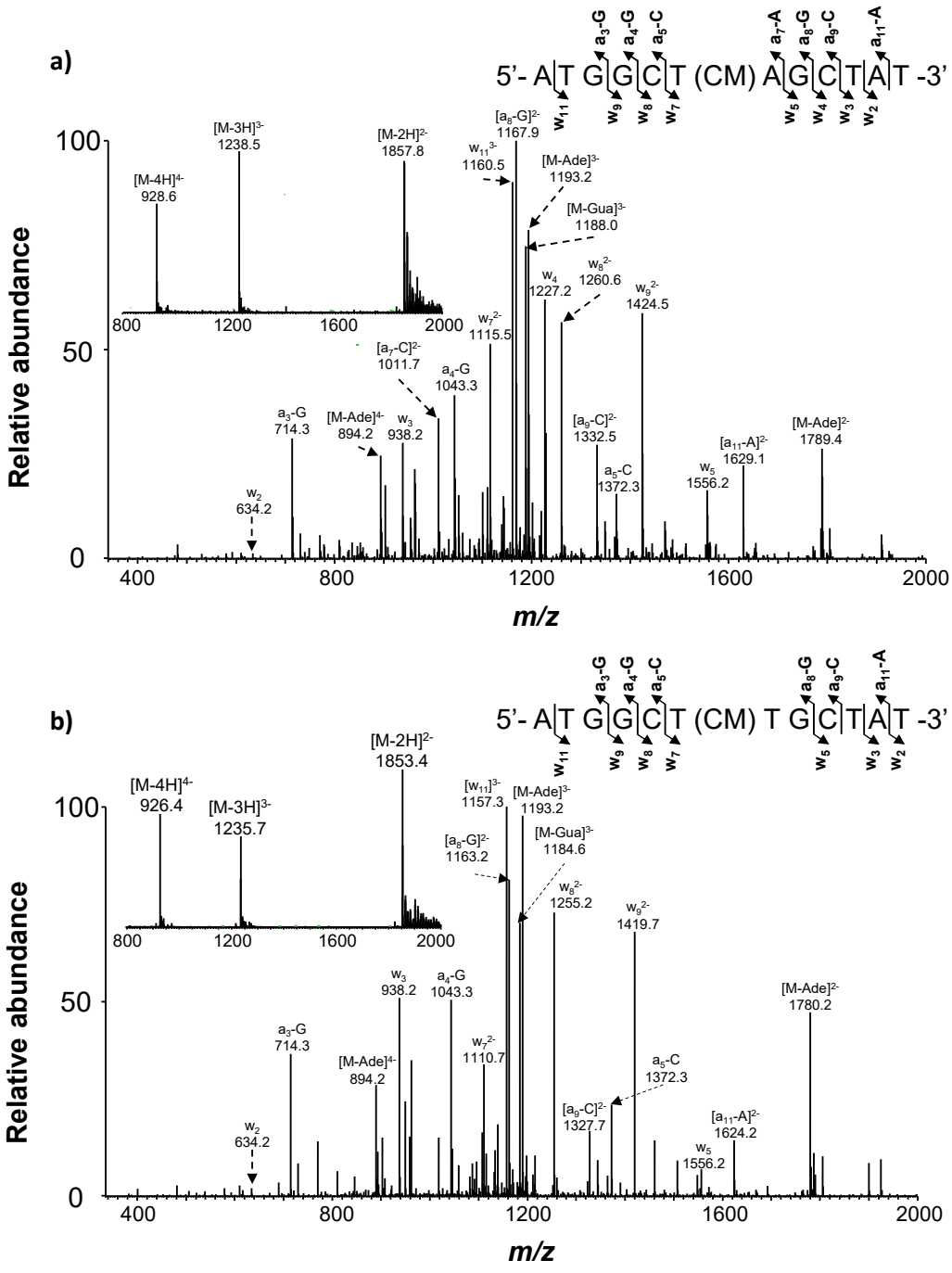


Figure 4-10. The product-ion spectra of the ESI-produced $[M-3H]^{3-}$ ions of the lesion-bearing 12-mers a) ATGGCT(CM)AGCTAT and b) ATGGCT(CM)TGCTAT. Displayed in the insets are the negative-ion ESI-MS of the modified ODNs.

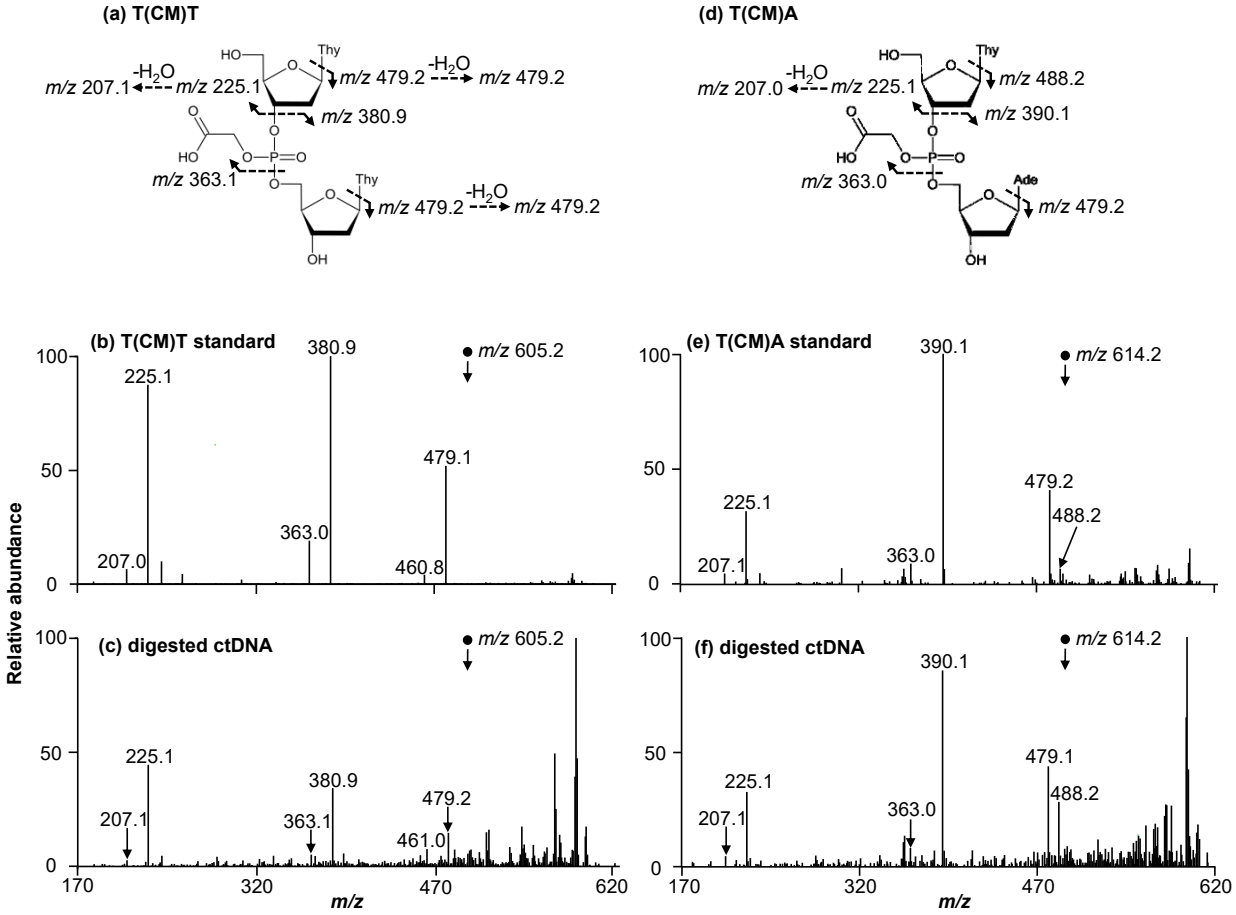


Figure 4-11. MS/MS demonstrating the induction of T(CM)T and T(CM)A in calf thymus DNA. Displayed are the MS/MS results for the $[M + H]^+$ ions of T(CM)T (*m/z* 605.2, **a**) and T(CM)A (*m/z* 614.2, **d**) for: **b**) T(CM)T standard; **c)** the 32.3-min fraction from Figure 2b; **e)** T(CM)A standard; **f)** the 31.1-min fraction from Figure 2d.

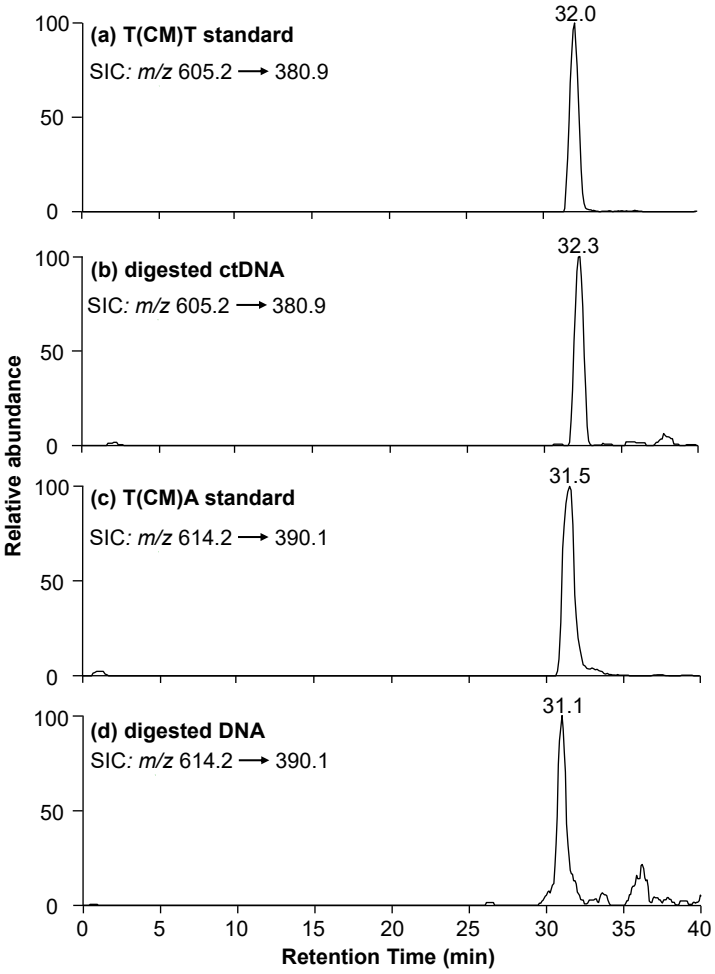


Figure 4-12. LC-MS/MS for monitoring the induction of T(CM)T and T(CM)A in KDA-treated calf thymus DNA. Displayed are the SICs for monitoring *m/z* 605 → 381 transition for: (a) the synthetic T(CM)T standard; (b) the enzymatic digestion mixture of KDA-treated calf thymus DNA; and the *m/z* 614 → 390 transitions for: (c) synthetic standard of T(CM)A; and (d) the same sample as in (b).

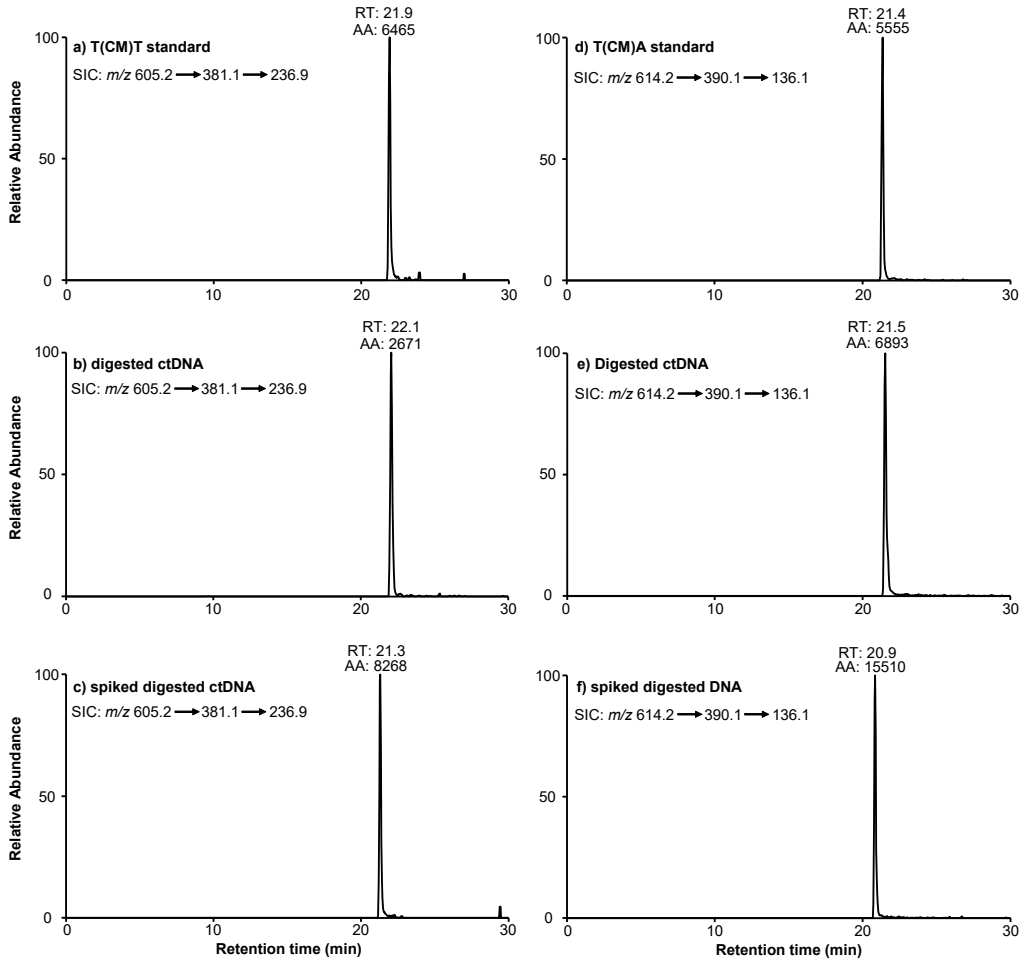


Figure 4-13. LC-MS/MS/MS for monitoring synthetic standards of T(CM)T and T(CM)A (a, d), T(CM)T and T(CM)A in the enzymatic digestion mixture of diazoacetate-treated calf thymus DNA (b, e), and those in enzymatic digestion mixture of diazoacetate-treated calf thymus DNA spiked with synthetic standards of T(CM)T and T(CM)A (c, f). Displayed are the SICs for monitoring *m/z* 605.2 → 381.1 → 236.9 for (a) the synthetic T(CM)T standard; (b) enzymatic digestion mixture of KDA-treated calf thymus DNA; (c) enzymatic digestion mixture spiked with the synthetic T(CM)T standard; SIC for monitoring *m/z* 614.2 → 390.1 → 136.1 for (d) T(CM)A synthetic standard, (e) enzymatic digestion mixture of KDA-treated calf thymus DNA; (f) enzymatic digestion mixture spiked with the synthetic T(CM)A synthetic standard.

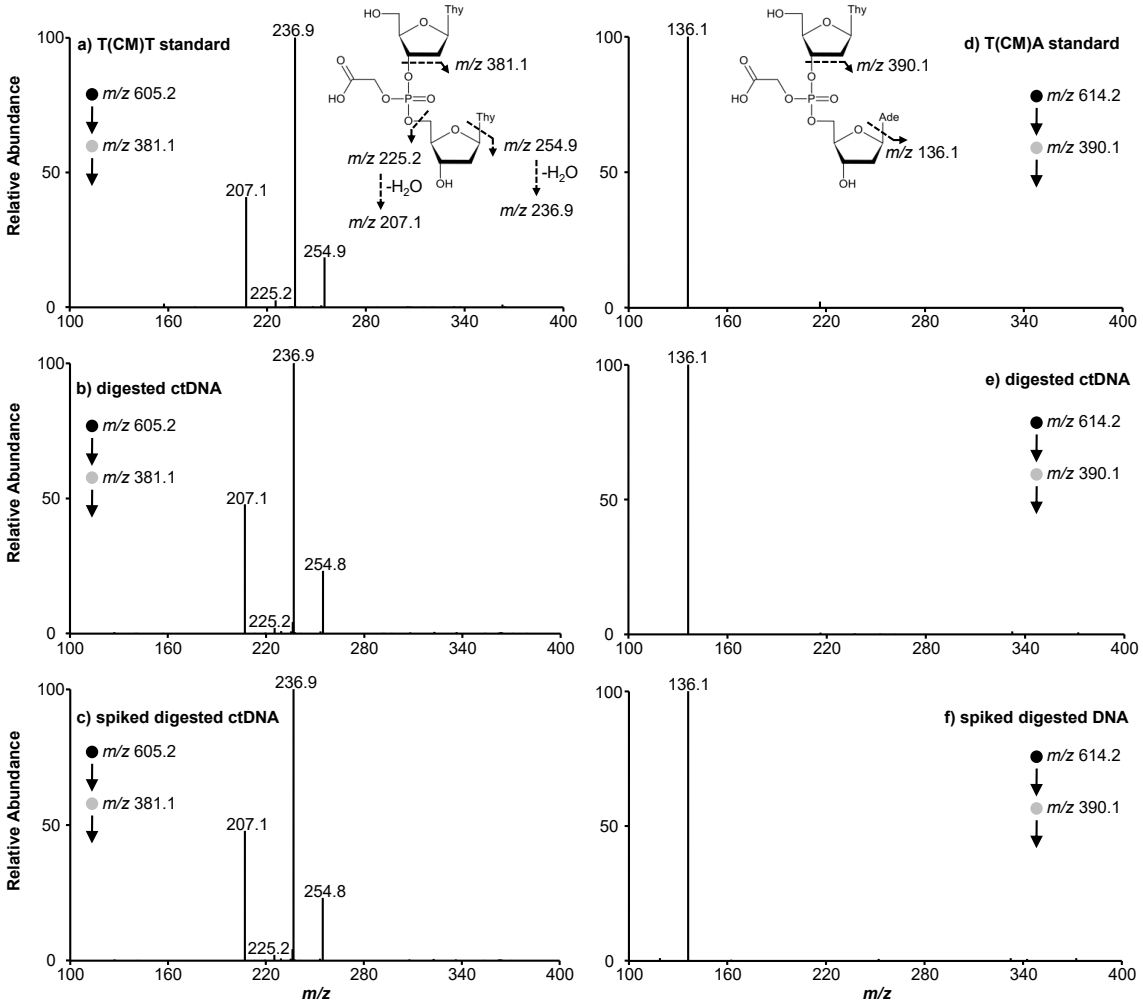


Figure 4-14. LC-MS/MS/MS for monitoring synthetic standards of T(CM)T and T(CM)A (a, d), T(CM)T and T(CM)A in the enzymatic digestion mixture of diazoacetate-treated calf thymus DNA (b, e), and those in enzymatic digestion mixture of diazoacetate-treated calf thymus DNA spiked with synthetic standards of T(CM)T and T(CM)A (c, f). Displayed are the MS^3 for (a) the synthetic T(CM)T standard; (b) T(CM)T in the enzymatic digestion mixture of KDA-treated calf thymus DNA; (c) T(CM)T in enzymatic digestion mixture spiked with the synthetic T(CM)T standard; (d) T(CM)A synthetic standard, (e) T(CM)A in the enzymatic digestion mixture of KDA-treated calf thymus DNA; (f) T(CM)A in enzymatic digestion mixture spiked with the synthetic standard of T(CM)A.

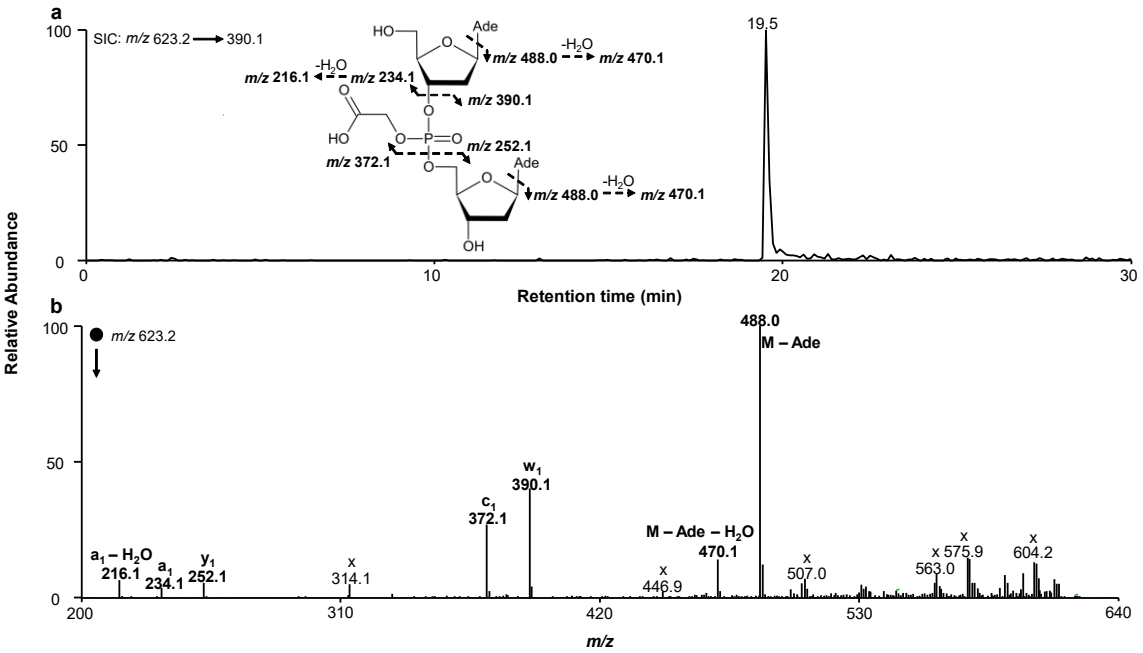


Figure 4-15. LC-MS/MS for monitoring the formation of A(CM)A in calf thymus DNA upon treatment with diazoacetate. (a) SIC for monitoring the *m/z* 623.2 → 390.1 transition, corresponding to the formation of the w_1 ion; (b) MS^2 for the $[M + H]^+$ ion of A(CM)A. Illustrated in the inset of (a) is the proposed fragmentation pathways for the $[M + H]^+$ ion of A(CM)A.

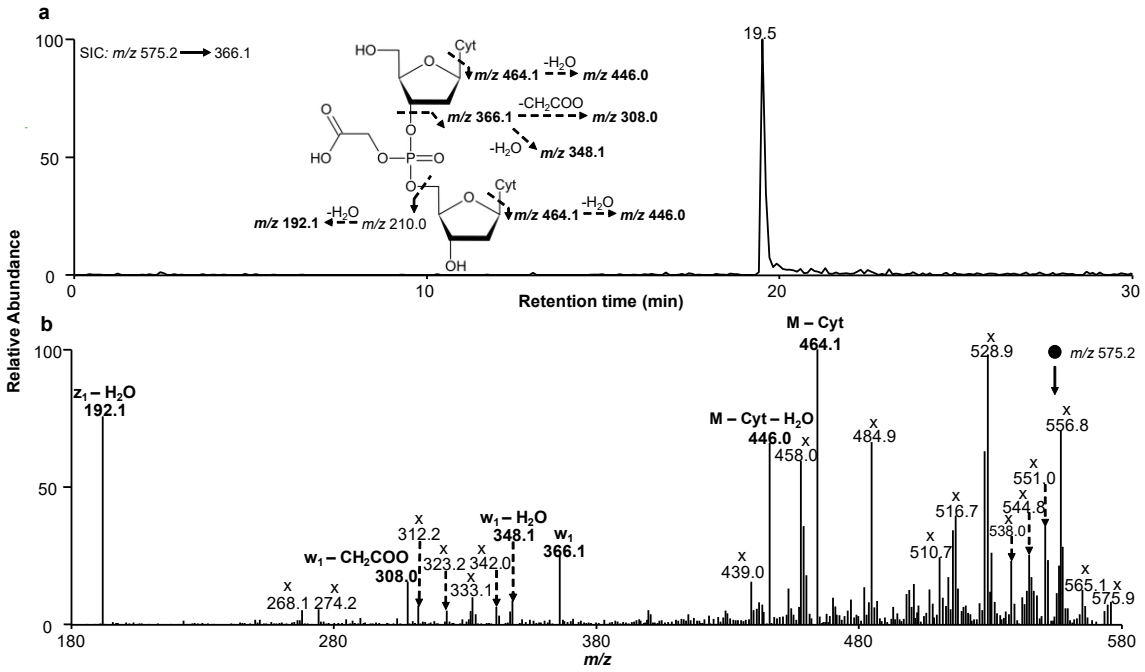


Figure 4-16. LC-MS/MS for monitoring the formation of C(CM)C in calf thymus DNA upon treatment with diazoacetate. (a) SIC for monitoring the *m/z* 575.2 → 366.1 transition, corresponding to the formation of the w_1 ion; (b) The MS^2 for the $[M + H]^+$ ion of C(CM)C. Illustrated in the inset of (a) is the proposed fragmentation pathways for the $[M + H]^+$ ion of C(CM)C.

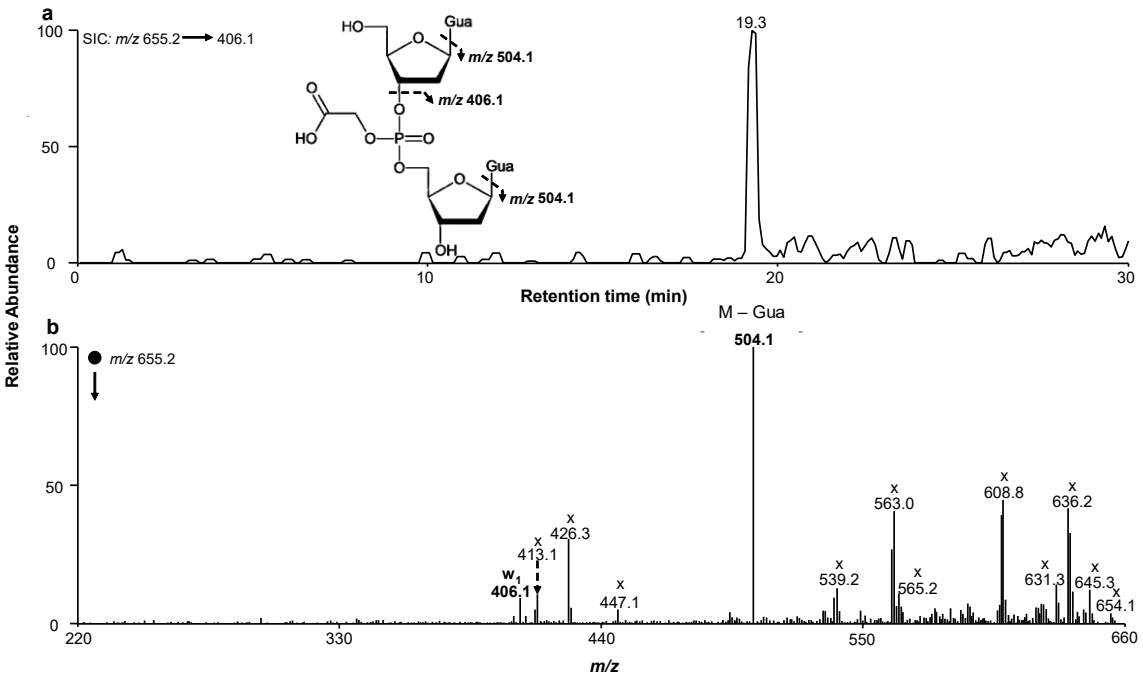


Figure 4-17. LC-MS/MS for monitoring the formation of G(CM)G in calf thymus DNA upon treatment with diazoacetate. (a) SIC for monitoring the *m/z* 655.2 → 406.1 transition, corresponding to the neutral loss of guanine; (b) The MS^2 for the $[M + H]^+$ ion of G(CM)G. Illustrated in the inset of (a) is the proposed fragmentation pathways for the $[M + H]^+$ ion of G(CM)G.

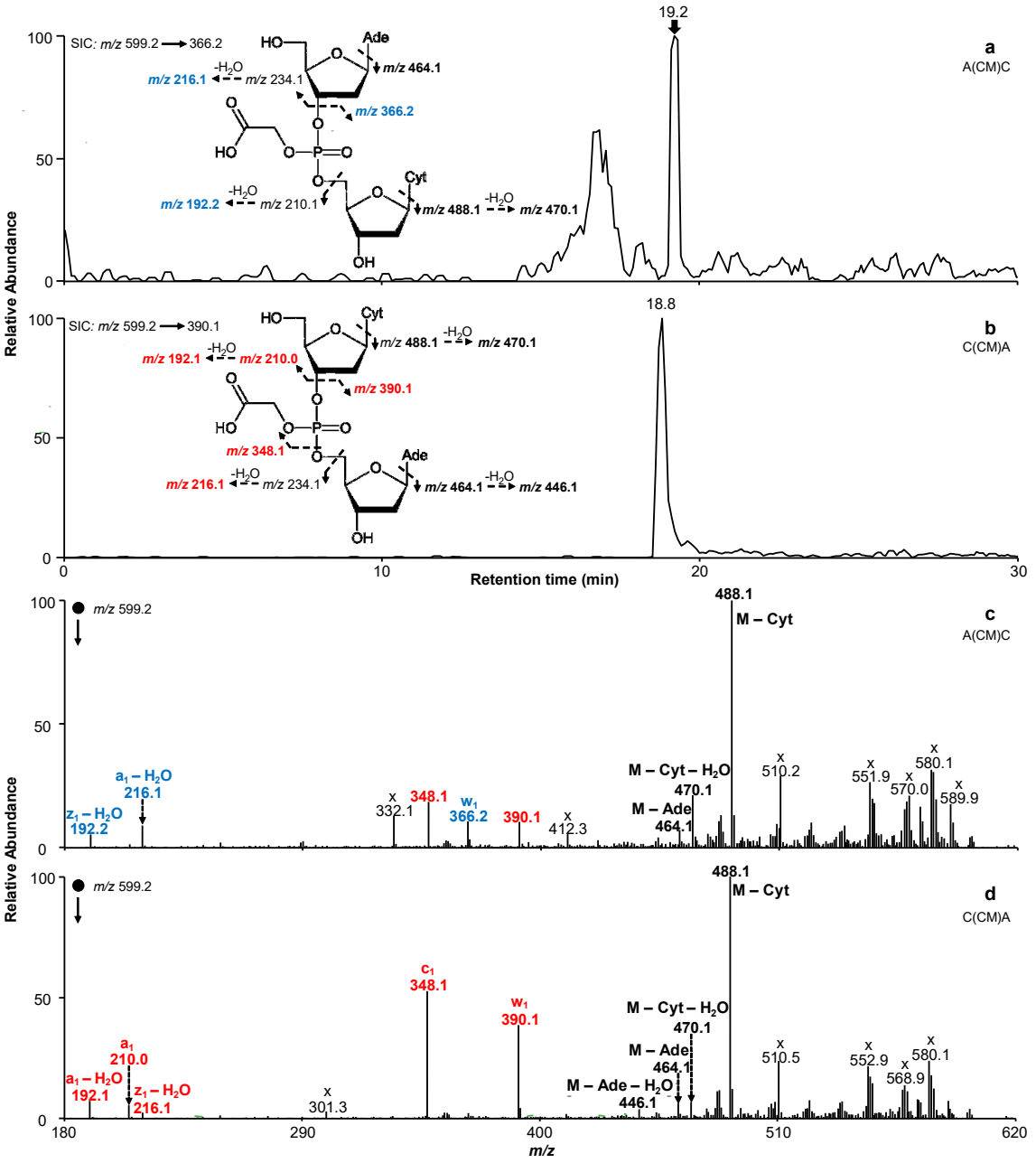


Figure 4-18. LC-MS/MS for monitoring the formation of the sequence isomers of A(CM)C and C(CM)A in calf thymus DNA upon treatment with diazoacetate. (a-b) SICs for monitoring the *m/z* 599.2 → 366.2 and *m/z* 599.2 → 390.1 transitions, corresponding to the formation of the w_1 ion from the $[M + H]^+$ ions of A(CM)C and C(CM)A, respectively; (c-d) the MS^2 for the $[M + H]^+$ ions of A(CM)C and C(CM)A. Illustrated in the insets of (a) and (b) are the proposed fragmentation pathways for the two isomers.

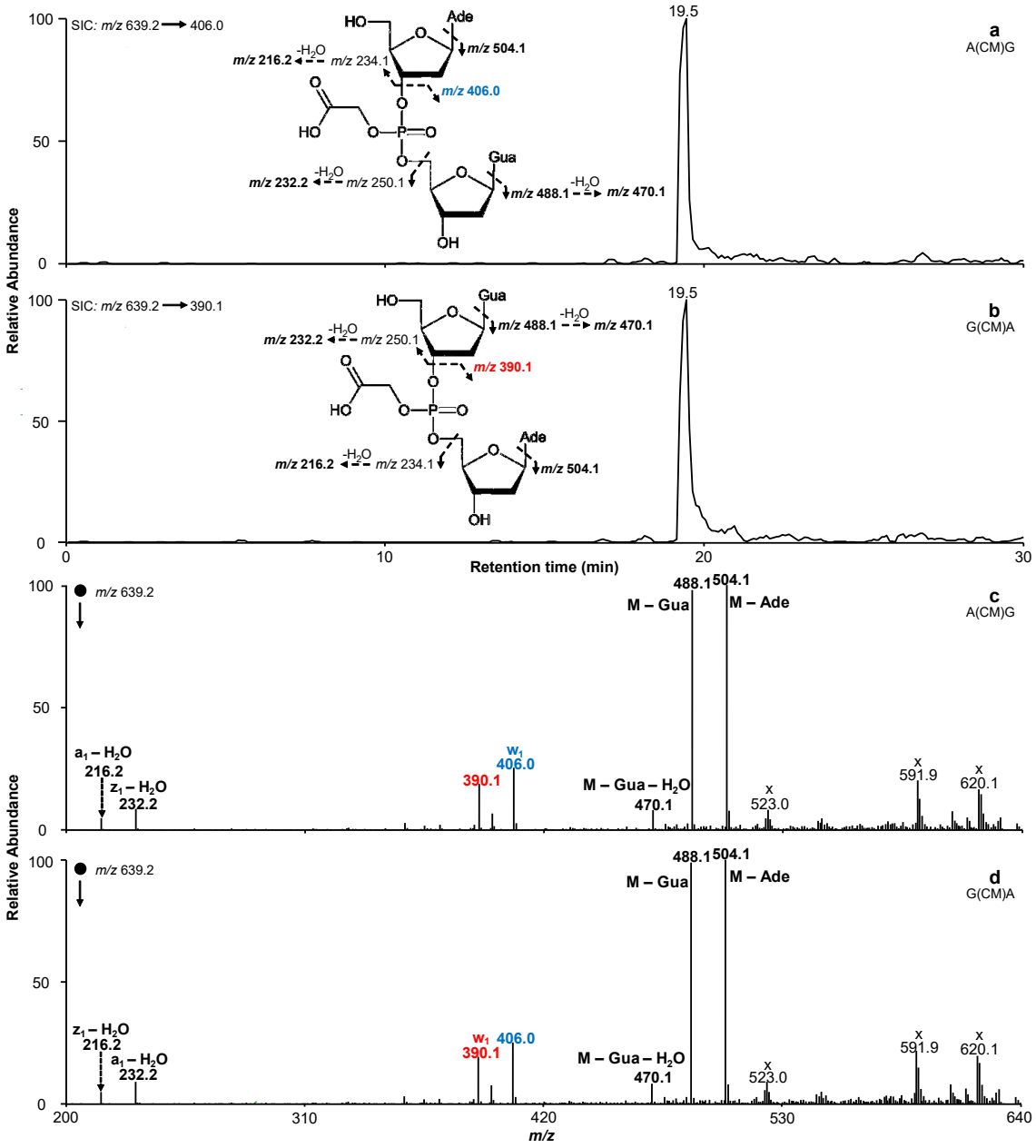


Figure 4-19. LC-MS/MS for monitoring the formation of the sequence isomers of A(CM)G and G(CM)A in calf thymus DNA upon treatment with diazoacetate. (a-b) SICs for monitoring the *m/z* 639.2 → 406.0 and *m/z* 639.2 → 390.1 transitions, corresponding to the formation of the w_1 ion from the $[M + H]^+$ ions of A(CM)G and G(CM)A, respectively; (c-d) the MS^2 for the $[M + H]^+$ ions of A(CM)G and G(CM)A. Illustrated in the insets of (a) and (b) are the proposed fragmentation pathways for the $[M + H]^+$ ions for the two isomers.

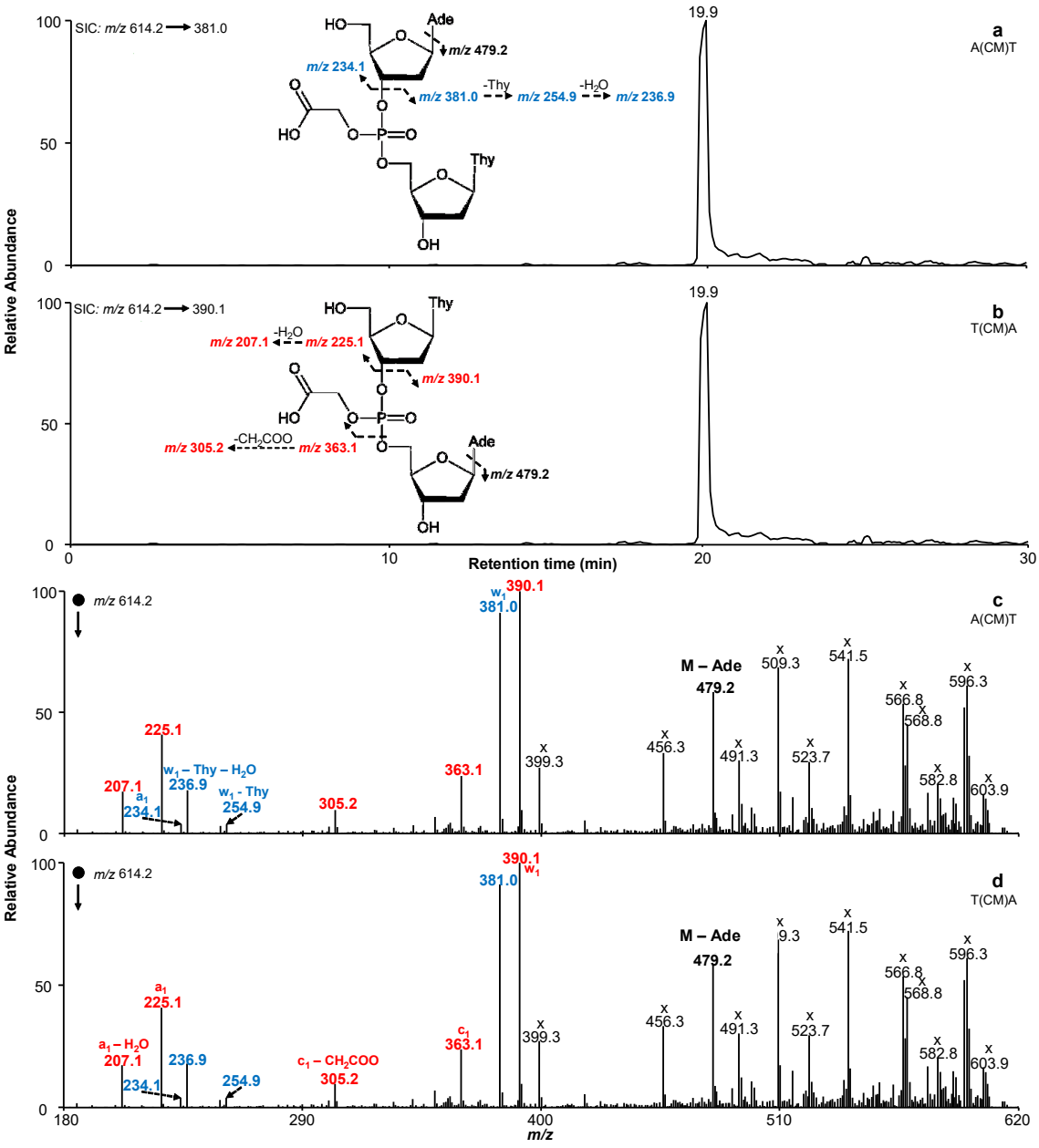


Figure 4-20. LC-MS/MS for monitoring the formation of the sequence isomers of A(CM)T and T(CM)A in calf thymus DNA upon treatment with diazoacetate. (a-b) SICs for monitoring the *m/z* 614.2 → 381.0 and *m/z* 614.2 → 390.1 transitions, corresponding to the formation of the w_1 ions from the $[M + H]^+$ ions of A(CM)T and T(CM)A, respectively; (c-d) MS^2 for the $[M + H]^+$ ions of A(CM)T and T(CM)A. Illustrated in the insets of (a) and (b) are the proposed fragmentation pathways for the $[M + H]^+$ ions for the two isomers.

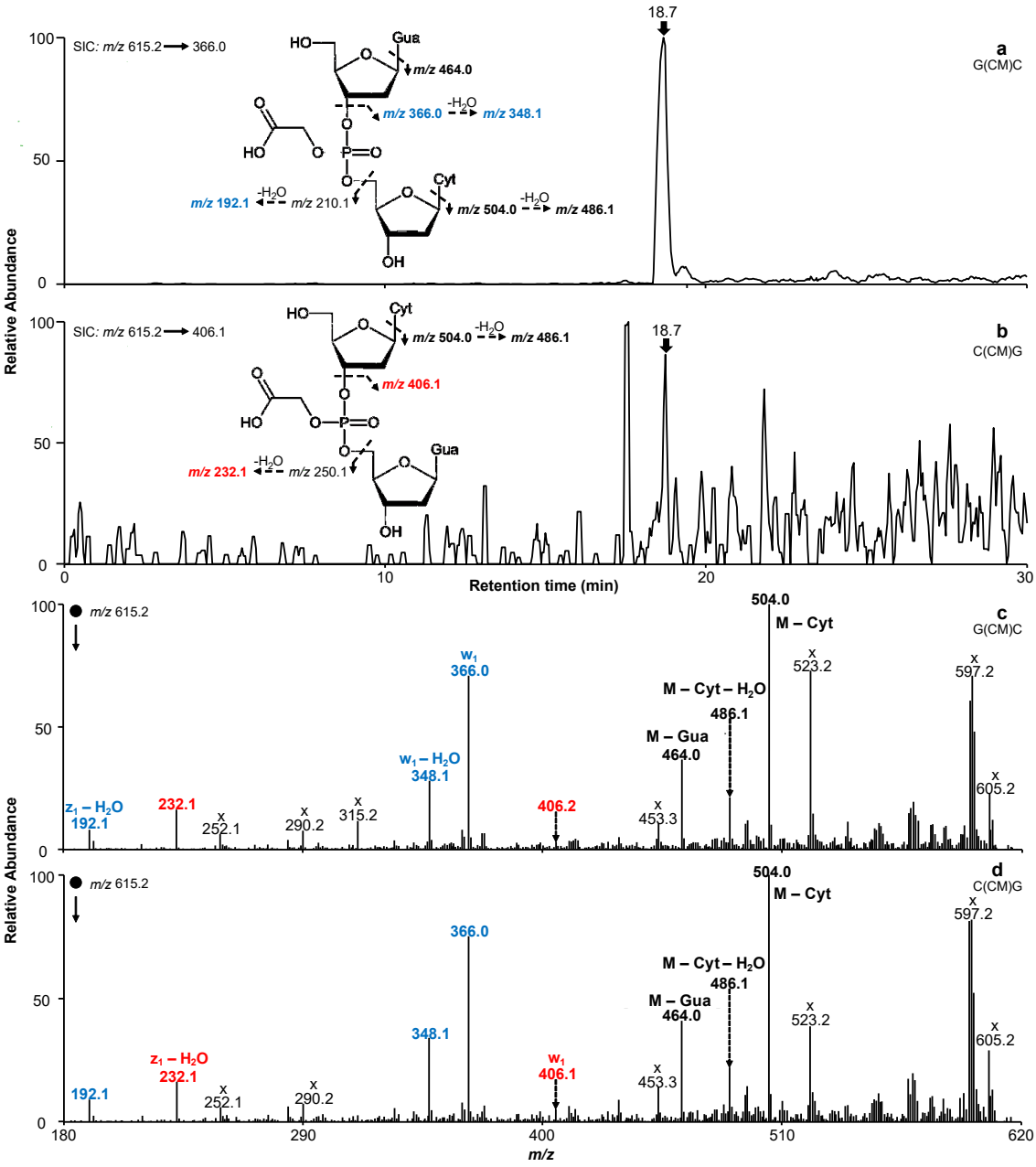


Figure 4-21. LC-MS/MS for monitoring the formation of the sequence isomers of G(CM)C and C(CM)G in calf thymus DNA upon treatment with diazoacetate. (a-b) SICs for monitoring the *m/z* 615.2 → 366.0 and *m/z* 615.2 → 406.1 transition, corresponding to the formation of the w_1 ions from the $[M + H]^+$ ions of G(CM)C and C(CM)G, respectively; (c-d) MS^2 for the $[M + H]^+$ ions of G(CM)C and C(CM)G. Illustrated in the insets of (a) and (b) are the proposed fragmentation pathways for the $[M + H]^+$ ions for the two isomers.

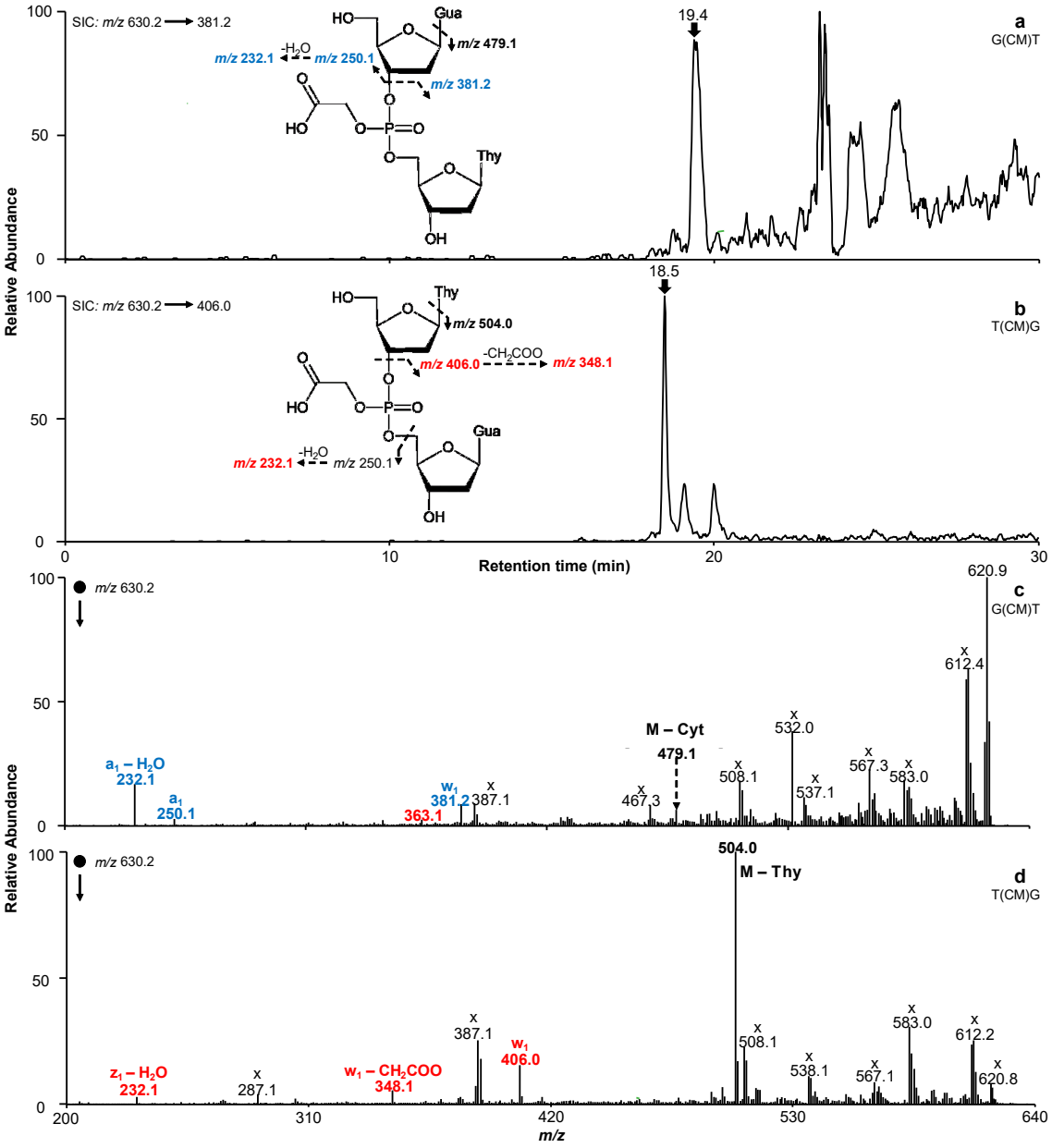


Figure 4-22. LC-MS/MS for monitoring the formation of the sequence isomers of G(CM)T and T(CM)G in calf thymus DNA upon treatment with diazoacetate. (a-b) SICs for monitoring the *m/z* 630.2 → 381.2 and *m/z* 630.2 → 406.0 transitions, corresponding to the formation of the w_1 ions from the $[M + H]^+$ ions of G(CM)T and T(CM)G, respectively; (c-d) MS^2 for the $[M + H]^+$ ions of G(CM)T and T(CM)G. Illustrated in the insets of (a) and (b) are the proposed fragmentation pathways for the $[M + H]^+$ ions of the two isomers.

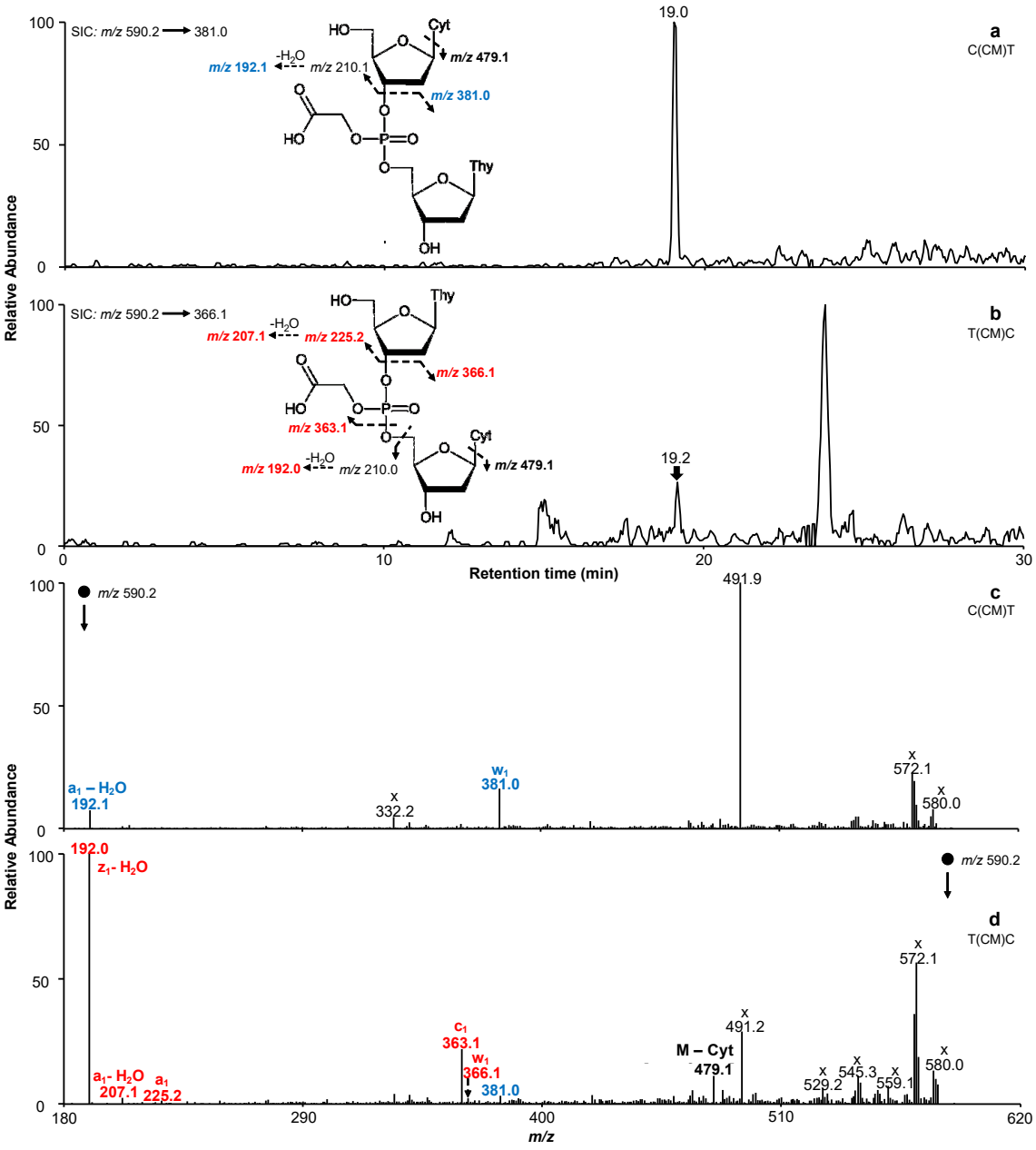


Figure 4-23. LC-MS/MS for monitoring the formation of the sequence isomers of C(CM)T and T(CM)C in calf thymus DNA upon treatment with diazoacetate. (a-b) SICs for monitoring the *m/z* 590.1 → 381.0 and *m/z* 590.1 → 366.1 transition, corresponding to the formation of w_1 ions from the $[M + H]^+$ ions of C(CM)T and T(CM)C, respectively; (c-d) MS^2 for the $[M + H]^+$ ions of C(CM)T and T(CM)C. Illustrated in the insets of (a) and (b) are the proposed fragmentation pathways for the $[M + H]^+$ ions of the two isomers.

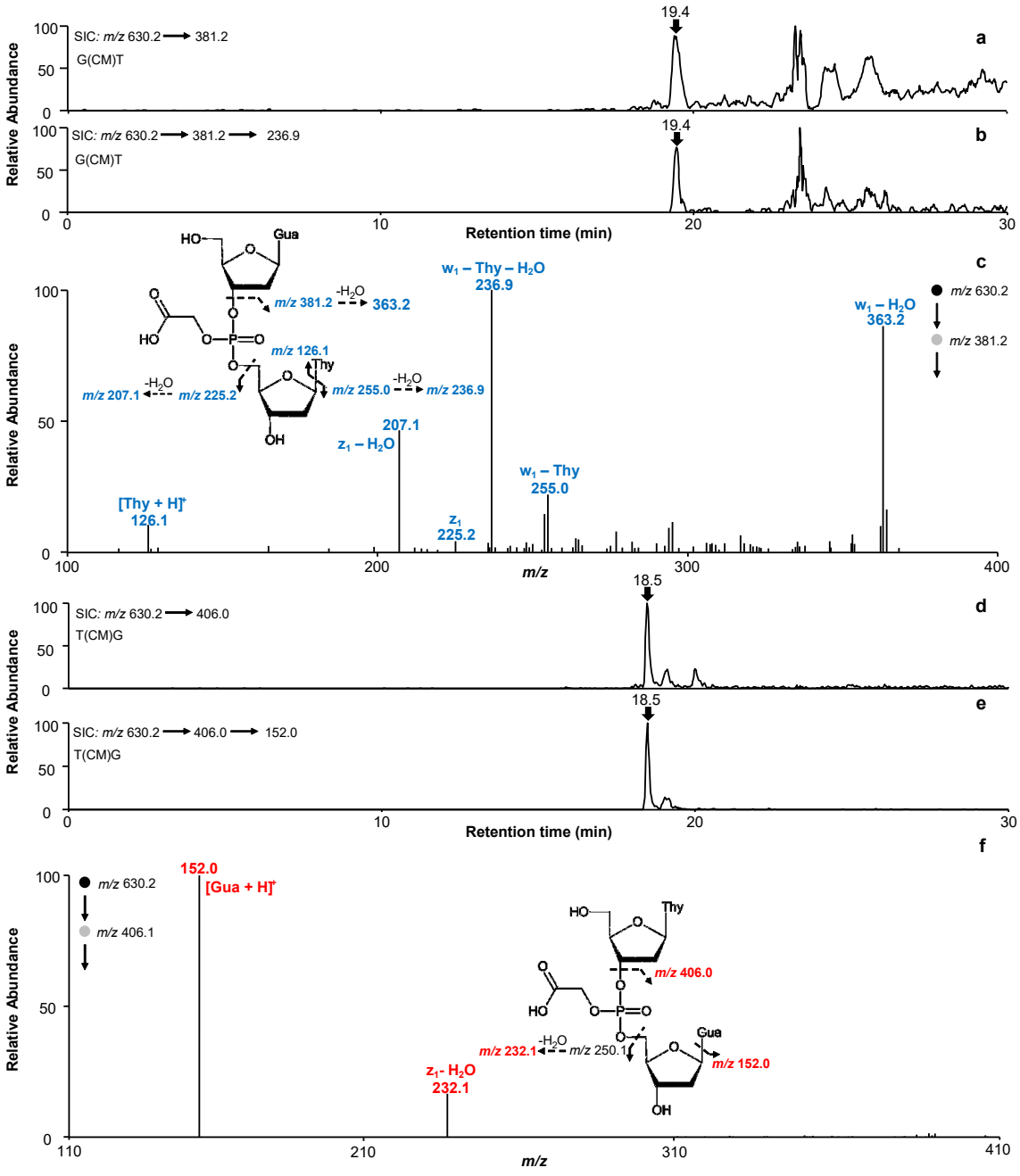

SIC: m/z 630.2 → 381.2
G(CM)T
19.4
a
SIC: m/z 630.2 → 381.2 → 236.9
G(CM)T
19.4
b
Relative Abundance
Retention time (min)
c
m/z 630.2
m/z 381.2
w1 – Thy – H2O
236.9
w1 – H2O
363.2
m/z 381.2 -H2O 363.2
m/z 126.1
m/z 207.1 -H2O m/z 225.2
m/z 255.0 -H2O m/z 236.9
207.1
z1 – H2O
w1 – Thy
255.0
[Thy + H]+
126.1
z1
225.2
m/z
SIC: m/z 630.2 → 406.0
T(CM)G
18.5
d
SIC: m/z 630.2 → 406.0 → 152.0
T(CM)G
18.5
e
Retention time (min)
f
m/z 630.2
m/z 406.1
152.0
[Gua + H]+
m/z 406.0
m/z 232.1 -H2O m/z 250.1
m/z 152.0
z1 - H2O
232.1
m/z

Figure 4-24. LC-MS/MS/MS for characterizing the w_1 ions of the sequence isomers G(CM)T and T(CM)G in calf thymus DNA upon treatment with diazoacetate. (a-b) SICs for monitoring the *m/z* of 630.2 → 381.2 and 630.2 → 381.2 → 236.9 transitions, corresponding to the formation of the w_1 ion from the $[M + H]^+$ ion and $w_1 - Thy - H_2O$ ion from the w_1 ion of G(CM)T, respectively; (c) MS^3 for the w_1 ion of G(CM)T; (d-e) SICs for monitoring the *m/z* transitions of 630.2 → 406.0 and 630.2 → 406.0 → 152.0 transitions, corresponding to the formation of the w_1 ion from the $[M + H]^+$ ion of T(CM)G and $[Gua + H]^+$ ion from the w_1 ion of T(CM)G, respectively; (f) MS^3 for the w_1 ion of T(CM)G. Illustrated in the insets are the proposed fragmentation pathways of the w_1 ions of two isomers.

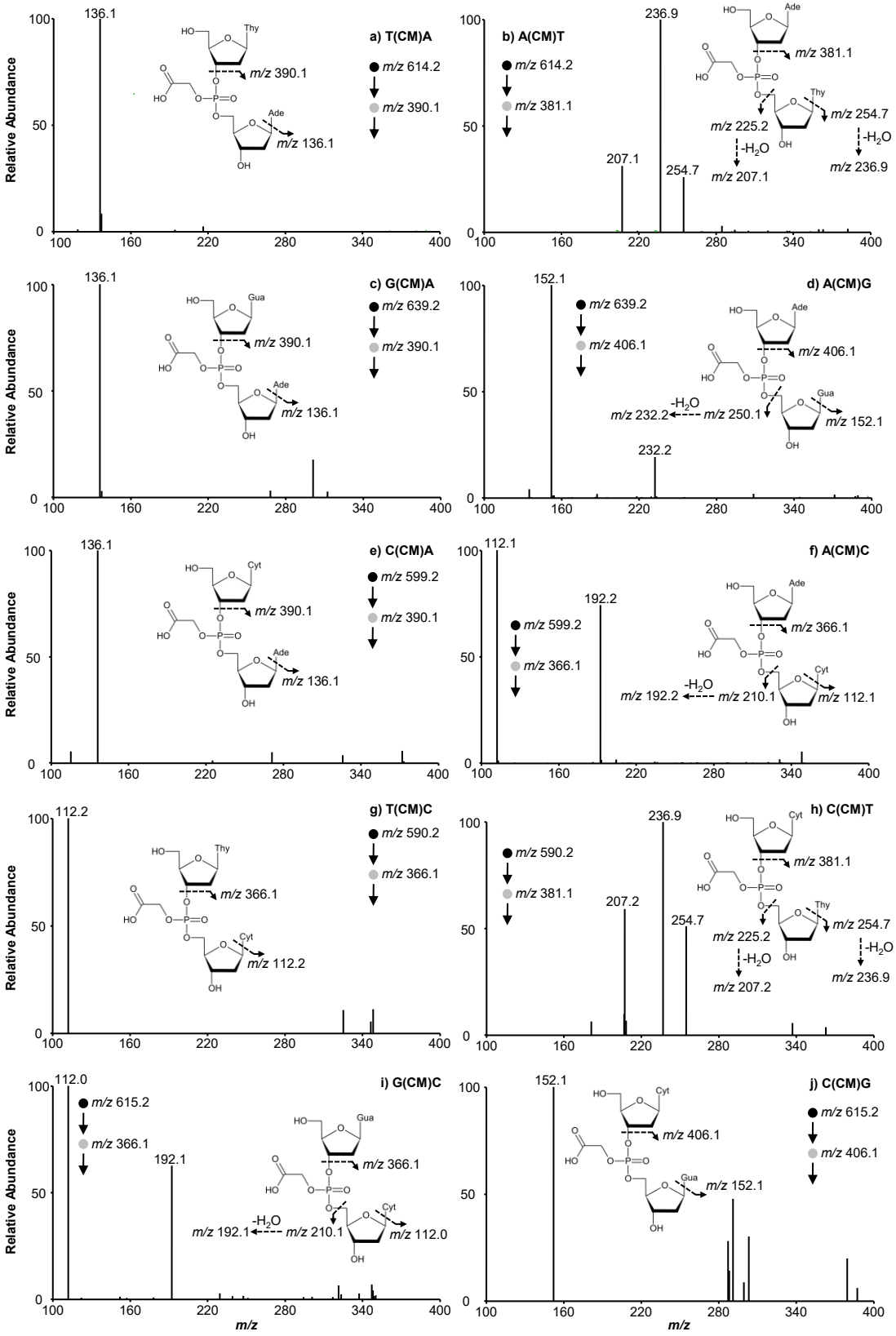

a) T(CM)A
m/z 614.2
m/z 390.1
m/z 390.1
m/z 136.1
136.1
b) A(CM)T
m/z 614.2
m/z 381.1
m/z 381.1
m/z 254.7
m/z 225.2
-H2O
m/z 207.1
-H2O
m/z 236.9
236.9
207.1
254.7
c) G(CM)A
m/z 639.2
m/z 390.1
m/z 390.1
m/z 136.1
136.1
d) A(CM)G
m/z 639.2
m/z 406.1
m/z 406.1
-H2O
m/z 232.2
m/z 250.1
m/z 152.1
152.1
232.2
e) C(CM)A
m/z 599.2
m/z 390.1
m/z 390.1
m/z 136.1
136.1
f) A(CM)C
m/z 599.2
m/z 366.1
m/z 366.1
-H2O
m/z 192.2
m/z 210.1
m/z 112.1
112.1
192.2
g) T(CM)C
m/z 590.2
m/z 366.1
m/z 366.1
m/z 112.2
112.2
h) C(CM)T
m/z 590.2
m/z 381.1
m/z 381.1
m/z 254.7
m/z 225.2
-H2O
m/z 207.2
-H2O
m/z 236.9
236.9
207.2
254.7
i) G(CM)C
m/z 615.2
m/z 366.1
m/z 366.1
-H2O
m/z 192.1
m/z 210.1
m/z 112.0
112.0
192.1
j) C(CM)G
m/z 615.2
m/z 406.1
m/z 406.1
m/z 152.1
152.1
Relative Abundance
m/z

Figure 4-25. MS/MS/MS for monitoring the further fragmentation of the w_1 ions for the sequence isomers of T(CM)A and A(CM)T, G(CM)A and A(CM)G, C(CM)A and A(CM)C, T(CM)C and C(CM)T, G(CM)C and C(CM)G formed in diazoacetate-treated calf thymus DNA. Shown are the MS^3 for the w_1 ion for T(CM)A and A(CM)T (a, b); G(CM)A and A(CM)G (c, d); C(CM)A and A(CM)C (e, f); T(CM)C and C(CM)T (g and h); G(CM)C and C(CM)G (i, j). Illustrated in the insets are the proposed fragmentation pathways of the w_1 ions of the sequence isomers.

REFERENCES

1. Lindahl, T., Instability and decay of the primary structure of DNA. *Nature* **1993,** *362*, 709-15.

2. Tricker, A. R., N-nitroso compounds and man: sources of exposure, endogenous formation and occurrence in body fluids. *Eur. J. Cancer Prev.* **1997,** *6*, 226-68.

3. Bogovski, P.; Bogovski, S., Animal species in which N-nitroso compounds induce cancer. *Int. J. Cancer* **1981,** *27*, 471-4.

4. Li, Y.; Hecht, S. S., Metabolic activation and DNA interactions of carcinogenic N-nitrosamines to which humans are commonly exposed. *Int. J. Mol. Sci.* **2022,** *23*, 4559.

5. Ward, M. H.; Jones, R. R.; Brender, J. D.; de Kok, T. M.; Weyer, P. J.; Nolan, B. T.; Villanueva, C. M.; van Breda, S. G., Drinking water nitrate and human health: An updated review. *Int. J. Environ. Res. Public Health* **2018,** *15*, 1557.

6. Jakszyn, P.; Bingham, S.; Pera, G.; Agudo, A.; Luben, R.; Welch, A.; Boeing, H.; Del Giudice, G.; Palli, D.; Saieva, C.; Krogh, V.; Sacerdote, C.; Tumino, R.; Panico, S.; Berglund, G.; Siman, H.; Hallmans, G.; Sanchez, M. J.; Larranaga, N.; Barricarte, A.; Chirlaque, M. D.; Quiros, J. R.; Key, T. J.; Allen, N.; Lund, E.; Carneiro, F.; Linseisen, J.; Nagel, G.; Overvad, K.; Tjonneland, A.; Olsen, A.; Bueno-de-Mesquita, H. B.; Ocke, M. O.; Peeters, P. H.; Numans, M. E.; Clavel-Chapelon, F.; Trichopoulou, A.; Fenger, C.; Stenling, R.; Ferrari, P.; Jenab, M.; Norat, T.; Riboli, E.; Gonzalez, C. A., Endogenous versus exogenous exposure to *N*-nitroso compounds and gastric cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST) study. *Carcinogenesis* **2006,** *27*, 1497-501.

7. Harrison, K. L.; Jukes, R.; Cooper, D. P.; Shuker, D. E., Detection of concomitant formation of O^6-carboxymethyl- and O^6-methyl-2'-deoxyguanosine in DNA exposed to nitrosated glycine derivatives using a combined immunoaffinity/HPLC method. *Chem. Res. Toxicol.* **1999,** *12*, 106-11.

8. Wang, J.; Wang, Y., Carboxymethylation of DNA induced by *N*-nitroso compounds and its biological implications. *Adv. Mol. Toxicol.* **2011,** *5*, 219-243.

9. Yu, Y.; Wang, J.; Wang, P.; Wang, Y., Quantification of azaserine-induced carboxymethylated and methylated DNA lesions in cells by nanoflow liquid chromatography-nanoelectrospray ionization tandem mass spectrometry coupled with the stable isotope-dilution method. *Anal. Chem.* **2016,** *88*, 8036-42.

10. Busby, W. F., Jr.; Shuker, D. E.; Charnley, G.; Newberne, P. M.; Tannenbaum, S. R.; Wogan, G. N., Carcinogenicity in rats of the nitrosated bile acid conjugates *N*-nitrosoglycocholic acid and *N*-nitrosotaurocholic acid. *Cancer Res.* **1985,** *45*, 1367-71.

11. Harrison, K. L.; Fairhurst, N.; Challis, B. C.; Shuker, D. E., Synthesis, characterization, and immunochemical detection of O^6-(carboxymethyl)-2'-deoxyguanosine: a DNA adduct formed by nitrosated glycine derivatives. *Chem. Res. Toxicol.* **1997,** *10*, 652-9.

12. Wang, J.; Wang, Y., Chemical synthesis of oligodeoxyribonucleotides containing *N*3- and O^4-carboxymethylthymidine and their formation in DNA. *Nucleic Acids Res.* **2009,** *37*, 336-45.

13. Wang, J.; Wang, Y., Synthesis and characterization of oligodeoxyribonucleotides containing a site-specifically incorporated N^6-carboxymethyl-2'-deoxyadenosine or N^4-carboxymethyl-2'-deoxycytidine. *Nucleic Acids Res* **2010,** *38*, 6774-6784.

14. Gottschalg, E.; Scott, G. B.; Burns, P. A.; Shuker, D. E., Potassium diazoacetate-induced *p53* mutations in vitro in relation to formation of O^6-carboxymethyl- and O^6-methyl-2'-deoxyguanosine DNA adducts: relevance for gastrointestinal cancer. *Carcinogenesis* **2007,** *28*, 356-62.

15. Van Cura, D.; Ng, T. L.; Huang, J.; Hager, H.; Hartwig, J. F.; Keasling, J. D.; Balskus, E. P., Discovery of the azaserine biosynthetic pathway uncovers a biological route for a-diazoester production. *Angew. Chem. Int. Ed. Engl.* **2023,** *62*, e202304646.

16. Jones, G. D.; Le Pla, R. C.; Farmer, P. B., Phosphotriester adducts (PTEs): DNA's overlooked lesion. *Mutagenesis* **2010,** *25*, 3-16.

17. Beranek, D. T., Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutat. Res.* **1990,** *231*, 11-30.

18. Ma, B.; Zarth, A. T.; Carlson, E. S.; Villalta, P. W.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S., Identification of more than one hundred structurally unique DNA-phosphate adducts formed during rat lung carcinogenesis by the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone. *Carcinogenesis* **2018,** *39*, 232-241.

19. Conrad, J.; Muller, N.; Eisenbrand, G., Studies on the stability of trialkyl phosphates and di-(2'deoxythymidine) phosphotriesters in alkaline and neutral solution. A model study for hydrolysis of phosphotriesters in DNA and on the influence of a beta hydroxyethyl ester group. *Chem. Biol. Interact.* **1986,** *60*, 57-65.

20. Den Engelse, L.; Menkveld, G. J.; De Brij, R. J.; Tates, A. D., Formation and stability of alkylated pyrimidines and purines (including imidazole ring-opened 7-alkylguanine) and alkylphosphotriesters in liver DNA of adult rats treated with ethylnitrosourea or dimethylnitrosamine. *Carcinogenesis* **1986,** *7*, 393-403.

21. Li, Y.; Ma, B.; Cao, Q.; Balbo, S.; Zhao, L.; Upadhyaya, P.; Hecht, S. S., Mass spectrometric quantitation of pyridyloxobutyl DNA phosphate adducts in rats chronically treated with *N*'-nitrosonornicotine. *Chem. Res. Toxicol.* **2019,** *32*, 773-783.

22. Wu, J.; Wang, P.; Wang, Y., Cytotoxic and mutagenic properties of alkyl phosphotriester lesions in *Escherichia coli* cells. *Nucleic Acids Res* **2018,** *46*, 4013-4021.

23. Wu, J.; Yuan, J.; Price, N. E.; Wang, Y., Ada protein- and sequence context-dependent mutagenesis of alkyl phosphotriester lesions in *Escherichia coli* cells. *J. Biol. Chem.* **2020,** *295*, 8775-8783.

24. Wu, J.; Wang, Y., Replication of pyridyloxobutyl phosphotriester lesions in cells. *Chem. Res. Toxicol.* **2020,** *33*, 308-311.

25. Wu, J.; Wu, J.; Clabaugh, G.; Wang, Y., Replication studies of alkyl phosphotriester lesions in human cells. *Chem. Res. Toxicol.* **2024,** *37*, 451-454.

26. Tan, Y.; Wu, J.; Clabaugh, G.; Li, L.; Du, H.; Wang, Y., Size- and stereochemistry-dependent transcriptional bypass of DNA alkyl phosphotriester adducts in mammalian cells. *DNA* **2022,** *2*, 221-230

27. Gait, M. J., *Oligonucleotide Synthesis: A Practical Approach*. IRL Press Limited: Oxford, England, 1984.

28. Anderson, D.; Hambly, R. J.; Yu, T. W.; Thomasoni, F.; Shuker, D. E., The effect of potassium diazoacetate on human peripheral lymphocytes, human adenocarcinoma colon Caco-2 cells, and rat primary colon cells in the comet assay. *Teratog. Carcinog. Mutagen.* **1999,** *19*, 137-46.

29. Habibi-Goudarzi, S.; McLuckey, S. A., Ion trap collisional activation of the deprotonated deoxymononucleoside and deoxydinucleoside monophosphates. *J. Am. Soc. Mass Spectrom.* **1995,** *6*, 102-13.

30. Lewin, M. H.; Bailey, N.; Bandaletova, T.; Bowman, R.; Cross, A. J.; Pollock, J.; Shuker, D. E.; Bingham, S. A., Red meat enhances the colonic formation of the DNA adduct O^6-carboxymethyl guanine: implications for colorectal cancer risk. *Cancer Res.* **2006,** *66*, 1859-65.

31. Norat, T.; Bingham, S.; Ferrari, P.; Slimani, N.; Jenab, M.; Mazuir, M.; Overvad, K.; Olsen, A.; Tjonneland, A.; Clavel, F.; Boutron-Ruault, M. C.; Kesse, E.; Boeing, H.; Bergmann, M. M.; Nieters, A.; Linseisen, J.; Trichopoulou, A.; Trichopoulos, D.; Tountas, Y.; Berrino, F.; Palli, D.; Panico, S.; Tumino, R.; Vineis, P.; Bueno-de-Mesquita, H. B.; Peeters, P. H.; Engeset, D.; Lund, E.; Skeie, G.; Ardanaz, E.; Gonzalez, C.; Navarro, C.; Quiros, J. R.; Sanchez, M. J.; Berglund, G.; Mattisson, I.; Hallmans, G.; Palmqvist, R.; Day, N. E.; Khaw, K. T.; Key, T. J.; San Joaquin, M.; Hemon, B.; Saracci, R.; Kaaks, R.; Riboli, E., Meat, fish, and colorectal cancer risk: the European Prospective Investigation into cancer and nutrition. *J. Natl. Cancer Inst.* **2005,** *97*, 906-16.

32. Hecht, S. S., Progress and challenges in selected areas of tobacco carcinogenesis. *Chem. Res. Toxicol.* **2008,** *21*, 160-71.

33. Swanson, A. L.; Wang, J.; Wang, Y., *In vitro* replication studies of carboxymethylated DNA lesions with *Saccharomyces cerevisiae* polymerase h. *Biochemistry* **2011,** *50*, 7666-73.

34. Wu, J.; Wang, P.; Li, L.; Williams, N. L.; Ji, D.; Zahurancik, W. J.; You, C.; Wang, J.; Suo, Z.; Wang, Y., Replication studies of carboxymethylated DNA lesions in human cells. *Nucleic Acids Res* **2017,** *45*, 7276-7284.

CHAPTER 5

Concluding Remarks and Future Directions

In this dissertation, we have laid the groundwork for deepening our understanding of the biological consequences of phosphotriester (PTE) lesions. We successfully synthesized pyridylhydroxybutyl-PTE (PHB-PTE) oligodeoxynucleotide (ODN) substrates, which will serve as valuable tools in future studies aimed at uncovering how cells respond, repair, and tolerate these lesions. Additionally, we confirmed the induction of carboxymethyl-PTE lesions in calf thymus DNA upon exposure to diazoacetate, providing the first step toward investigating their impact on genomic instability. Furthermore, we examined how variations in alkyl chain length and stereochemical configuration influence the efficiency and fidelity of DNA replication by using the strand-specific PCR-competitive replication and adduct bypass (SSPCR-CRAB) assay.

In Chapter 2, we synthesized a dithiane-protected pyridylhydroxybutyl (POB) alcohol and coupled it to a thymidine phosphoramidite, which was subsequently incorporated it into ODNs. These modified ODNs were converted into PHB-PTE ODNs, with the lesion positioned between two thymidine residues in both the R_P and S_P configurations. The lesion-containing ODNs were purified by HPLC and characterized by LC-MS/MS, confirming the successful construction of the PHB-PTE modification. These substrates will enable future studies to elucidate the mechanisms by which cells recognize and repair the PHB-PTE lesions, as well as the biological consequences of their formation.

In Chapter 3, we investigated for the first time the effects of alkyl phosphotriester (alkyl-PTE) lesions on DNA replication in HEK239T cells using the SSPCR-CRAB assay. Our results showed that methyl- and ethyl-PTEs (Me-PTE, Et-PTE) were more readily bypassed than the bulkier *n*-propyl- and *n*-butyl-PTE (*n*Pr-PTE, *n*Bu-PTE) lesions. Genetic depletion of Pol η or Pol ζ led to significantly reduced bypass efficiency across *n*Pr- and *n*Bu-PTEs in both R_P and S_P configurations, as well as the S_P-Et-PTE, indicating their role in bypassing these adducts. Diminished replicative bypass was also observed for the R_P diastereomer of *n*Pr- and *n*Bu-PTEs upon the genetic depletion of Pol κ. The CRISPR-Cas9-mediated knockout of Pol ι also resulted in moderate bypass efficiency impediment for both diastereomers of the *n*Pr- and *n*Bu-PTE modifications. Importantly, none of the lesions induced mutations in either the wild-type or isogenic polymerase knockout cell lines. In summary, most alkyl-PTE lesions studied pose moderate impediments to DNA replication in human cells, with minimal observed differences between the S_P or R_P diastereomers. Bypass efficiency was influenced by both alkyl chain length of the lesions and the presence of TLS polymerases in host cells.

In Chapter 4, we report the first identification of carboxymethyl-PTE (CM-PTE) lesions in calf thymus DNA (ctDNA) following treatment with diazoacetate, and we constructed 12-mer ODNs harboring site-specific CM-PTE modifications as substrates for future studies. We first synthesized a methyl ester of CM-PTE phosphoramidite building block of thymidine, which was subsequently incorporated into ODNs in different flanking base sequences (TT and TA). We confirmed the construction of CM-PTE ODNs by HPLC separation and LC-MS/MS analysis, enabling the characterization of both R_P and S_P diastereomers.

To validate the formation of CM-PTE lesions in DNA, we synthesized the carboxymethylated dinucleoside monophosphates T(CM)T and T(CM)A to serve as standards. By comparing the retention times and MS/MS fragmentation patterns between the synthesized standards and the CM-PTE adducts detected in ctDNA treated with diazoacetate, we identified a total of 16 CM-PTE lesions across all possible combinations of flanking nucleobases. Taken together, our findings provide evidence for phosphate backbone carboxymethylation and lay a foundation for future investigations into the *in vivo* formation and biological consequences of CM-PTE lesions.

Further studies should address several key gaps in our understanding of PHB-PTE lesions. First, there are no cellular studies investigating how PHB-PTE lesions affect DNA replication or transcription in human cells. Ma *et al*.[1] identified and characterized 107 distinct PHB-PTE lesions in the lungs of mice chronically treated with NNK, with some lesions detectable for over 70 weeks. Exposure to NNK induces a diverse range of PTE lesions and can persist in DNA, suggesting they are likely to be encountered by and potentially interfere with replication and transcription machineries. However, the only study to date investigating how NNK-induced pyridyl-containing lesions impact replication was conducted by Wu *et al.,*[2] where they examined how POB-PTE lesions influence DNA replication in *E. coli* cells.

Secondly, there is currently no information regarding how cells detect and respond to alkyl-PTE lesions. Ma *et al.*[3] identified alkyl-PTE adducts in lung tissue samples from lung cancer patients and observed elevated levels of Me-PTEs in the lungs of smokers compared to non-smokers, highlighting the potential biological relevance of these lesions in human disease. In this dissertation, we explored how alkyl-PTEs influence DNA replication in human cells; however, the cellular response and mechanisms responsible

for repairing these lesions remain unexplored. Zhao *et al*.[4] recently constructed an ODN photoaffinity probe bearing a N^2-*n*Bu-2'-deoxyguanosine (N^2-*n*Bu-dG) modification to employ in an untargeted proteomics approach for the identification of proteins involved in its recognition and repair. They discovered that the high-mobility group protein (HMGB3) and SUB1 bind to N^2-*n*Bu-dG to promote its repair. In the future, we can construct similar photoaffinity probes harboring alkyl-PTE lesions with varying alkyl chain lengths and stereochemical configurations, which will enable the identification of proteins that interact with distinct PTE lesions and may uncover lesion-specific cellular responses or repair pathways.

Lastly, we identified CM-PTE lesions induced in ctDNA following treatment with diazoacetate. However, it remains unknown whether these lesions are formed in human cells, and at what levels they accumulate. Previous work in our laboratory by Yu *et al.*[5] successfully quantified O^6-carboxymethyl-2′-deoxyguanosine (O^6-CMdG), O^6-methyl-2′-deoxyguanosine (O^6-MedG), and N^6-carboxymethyl2′-deoxyadenosine (N^6-CMdA) in human skin fibroblasts and colorectal carcinoma cells following azaserine treatment using a highly sensitive nanoflow liquid chromatography-nanoelectrospray ionization-multistage tandem mass spectrometry (nLC-nESI-MS^3) method. Building upon this approach, it will be interesting to determine whether CM-PTE lesions can be detected in human cells. This will provide the foundation for future studies aimed at investigating their influence on DNA replication and transcription in human cells.

REFERENCES

(1) Ma, B.; Zarth, A. T.; Carlson, E. S.; Villalta, P. W.; Upadhyaya, P.; Stepanov, I.; Hecht, S. S. Identification of More than 100 Structurally Unique DNA-Phosphate Adducts Formed during Rat Lung Carcinogenesis by the Tobacco-Specific Nitrosamine 4-(Methylnitrosamino)-1-(3-Pyridyl)-1-Butanone. *Carcinogenesis* **2018**, *39* (2), 232–241.

(2) Wu, J.; Wang, Y. Replication of Pyridyloxobutyl Phosphotriester Lesions in Cells. *Chem. Res. Toxicol.* **2020**, *33* (2), 308–311.

(3) Ma, B.; Villalta, P. W.; Hochalter, J. B.; Stepanov, I.; Hecht, S. S. Methyl DNA Phosphate Adduct Formation in Lung Tumor Tissue and Adjacent Normal Tissue of Lung Cancer Patients. *Carcinogenesis* **2019**, *40* (11), 1387–1394.

(4) Zhao, T.; He, X.; Liang, X.; Kellum, A. H.; Tang, F.; Yin, J.; Guo, S.; Wang, Y.; Gao, Z.; Wang, Y. HMGB3 and SUB1 Bind to and Facilitate the Repair of N^2 -Alkylguanine Lesions in DNA. *J. Am. Chem. Soc.* **2024**, *146* (32), 22553–22562.

(5) Yu, Y.; Wang, J.; Wang, P.; Wang, Y. Quantification of Azaserine-Induced Carboxymethylated and Methylated DNA Lesions in Cells by Nanoflow Liquid Chromatography-Nanoelectrospray Ionization Tandem Mass Spectrometry Coupled with the Stable Isotope-Dilution Method. *Anal. Chem.* **2016**, *88* (16), 8036–8042.