UNIVERISTY OF CALIFORNIA
RIVERSIDE

Chemical Syntheses and Cellular Replication Studies of DNA Alkyl Phosphotriester Lesions

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

Doctor of Philosophy

in

Environmental Toxicology

by

Jiabin Wu

December 2019

Dissertation Committee:
Dr. Yinsheng Wang, Chairperson
Dr. Connie Nugent
Dr. Jikui Song

The Dissertation of Jiabin Wu is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

First of all, I would like to give my gratitude to my advisor, Dr. Yinsheng Wang for his guidance and support all along through my Ph. D study. He instructed me, not only by his word but also by his behavior, about how to become a researcher, and his enthusiasm and dedication to the research work greatly impressed me over these years. Without Dr. Wang, it is impossible for me to accomplish these works in the thesis.

Besides, I would like to thank my other committee member, Profs. Connie Nugent and Jikui Song. They are all outstanding researchers, and they provided me many helpful suggestions on my research. Additionally, I also want to thank my former committee member, Dr. Huiwang Ai, who helped me a lot in my first two years.

I am grateful for the support from my labmates and staffs at UCR. I am proud that I was learning and working in such a fantastic research group. Everyone in the group gave me generous assistance during these years. Especially, I would like to thank Dr. Pengcheng Wang, who taught me organic synthesis and *in vivo* replication assay. He put a lot of efforts in assisting me to get into research work in a short time. I also want to thank Jun Yuan and Dr.

Nathan Price for their collaborative work and their efforts contributed greatly to the completion of my thesis.

I wish to extend my acknowledgement to Profs. John M. Essigmann and Graham. C. Walker for generously providing shuttle vector and polymerases- and Ada- deficient *E. coli* strains. I am also grateful that my research was supported by the funding from National Institutes of Health and the fellowships from the Environmental Toxicology Program of the University of California Riverside.

At last, I want to express my gratitude to my family members, who are the first and most important motivation of my studying at UCR. Their love and support without any compensation gave me power and confidence to overcome any barrier. I love you so much.

COPYRIGHT ACKNOWLEDGEMENTS

The text and figures in Chapter 2, in part or in full, are a reprint of the material as it appears in *Nucleic Acids Res.* 2018, 46, 4013-4021. The co-author (Dr. Yinsheng Wang) listed in that publication directed and supervised the research which forms the basis of this Chapter. The co-author (Dr. Pengcheng Wang) listed in that publication conducted experiments in replication assay in Ada-deficient strain.

ABSTRACT OF THE DISSERTATION

Chemical Syntheses and Cellular Replication Studies of DNA Alkyl Phosphotriester Lesions

By

Jiabin Wu

Doctor of Philosophy, Graduate Program in Environmental Toxicology
University of California, Riverside, December 2019
Dr. Yinsheng Wang, Chairperson

Exposure to endogenous and exogenous factors could give rise to DNA damage. DNA alkylation constitutes a major form of DNA damage, and alkyl phosphotriesters (alkyl-PTEs) are particularly important owing to their relatively high frequencies of occurrence and resistance to repair in mammalian cells. However, little is known about how alkyl-PTEs influence fidelity and efficiency of DNA replication in cells and how translesion synthesis (TLS) polymerases modulate the replicative bypass of these lesions.

In Chapter 2. We synthesized and characterized oligodeoxynucleotides (ODNs) containing site-specifically inserted alkyl-PTEs (Me, Et, *n*Pr and *n*Bu) with different alkyl groups and diastereomeric configurations (S_p and R_p) and assessed how these lesions impact DNA replication in *E. coli* cells. Our results revealed that S_p-alkyl-PTEs

could be efficiently bypassed, whereas their R_p counterparts exhibited moderate blockage effects on DNA replication. Additionally, S_p-Me-PTE was found to induce TT→GT and TT→GC mutations, which necessitated Ada proteins.

In Chapter 3. We prepared Me- or *n*Bu-PTE-containing ODNs in two diastereomeric forms with different flanking dinucleotides sequences (XT and TX, X=A, C, G) and examined how these lesions were recognized by DNA replication machinery in *E. coli* cells. We found that S_p-XT-Me-PTEs were efficiently bypassed with similar distributions in replication products (with approximately 85-90% AT and 5-10% TG at XT site). where Ada protein is essential for the generation of the mutagenic products. Additionally, R_p-XT-Me-PTEs and both diastereomers of TX-Me-PTEs displayed error-free bypass with significant blockage effects. We also found that Ada binds strongly to S_p-Me-PTE at AT site, but not TA site. Our results suggest that the role of Ada may be beyond the repair of S_p-Me-PTE lesion and transcriptional modulation of genes involved in repairing alkylated DNA lesions.

In Chapter 4, we analyzed the cytotoxic and mutagenic properties of pyridyloxobutyl (POB)-PTEs in *E. coli* cells. We demonstrated that POB-PTEs did not display strong impediments to DNA replication, and replicative bypass of POB-PTEs does not require TLS polymerases. In Addition, POB-PTEs were not mutagenic.

Together, the research described in this dissertation constitutes the first chemical syntheses of ODNs harboring stereochemically defined alkyl-PTE lesions, and provides a much better understanding about the biological consequences of alkyl-PTE lesions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

CHAPTER 1

General Introduction

This chapter will cover the general background of the research projects discussed in this dissertation.

DNA Alkylation

DNA carries the genetic information, but it has limited chemical stability (1); hence the integrity of DNA is constantly challenged by endogenous and exogenous factors, which lead to different kinds of DNA damage, which may result in mutagenesis and ultimately carcinogenesis (1-3).

DNA alkylation is a common type of DNA damage (4), and it also constitutes a major mechanism for many prescribed chemotherapeutic agents (2). According to Anatomical Therapeutic Chemical classification system code L01A, which is developed by World Health Organization Collaborating Centre for Drug Statistics Methodology, alkylating agents can be classified into 6 groups: nitrogen mustard analogues, alkyl sulfonates, ethylene imines, nitrosoureas, epoxides and other alkylating agents. Alkylating agents can

attack different nucleophiles on DNA to form various alkylation products, and depending on the positions of nucleophiles being attacked, alkylated DNA lesions can be divided into two groups: base alkylation and backbone alkylation. So far, base alkylation constitutes the major concern for DNA alkylation studies (5). However, backbone alkylation products, also known as alkylphosphotriester (alkyl-PTE), which is formed from alkylating agents attacking the non-carbon-bonded oxygen atoms on the phosphate backbone (6), came to our sight in recent years, and it constitutes the focus of this thesis.

Formation of Alkylphosphotriester

In 1948, it has been shown that mustard gas could react with phosphate backbone to form "secondary phosphoryl" adduct (7). Later, detailed studies indicated that many alkylating agents, such as *N*- methy-*N*-lnitrosourea (MNU), *N*-ethyl-*N*-nitrosourea (ENU), diazomethane and methyl methanesulfonate (MMS) could induce the formation of backbone alkylation products (8-10). Recent research also revealed that the methyl diazohydroxide, a bioactivation product of tobacco-specific nitrosamines, *N*-nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and its metabolite 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL), could also react with DNA to form backbone methylation product (11), establishing the relationship between the formation of alkyl-PTEs and tobacco smoking. Since each phosphodiester bond has two non-carbon-bonded oxygen atoms, and alkylating agents can attack either of them, every

type of alkyl-PTE is present in two diastereomeric configurations (Figure 1-1). The diastereomeric difference of alkyl-PTEs has biological significance, which would be illustrated in detail in this dissertation.

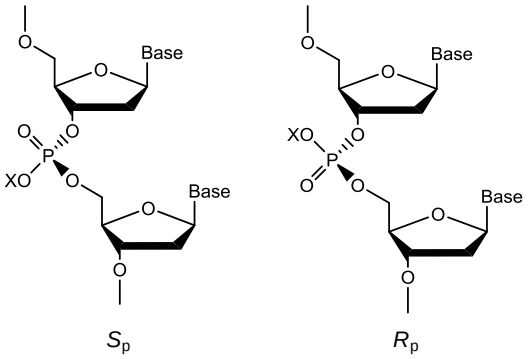


Figure 1-1. The structure of phosphotriester, 'X' represents different alkyl groups.

Chemically, alkyl-PTEs are labile under basic condition and stable under neutral and acidic conditions (8,12), and they also show high stability at high temperature (8). Alkyl-PTEs are relatively stable under physiological conditions. Shooter *et al.* (13) injected C57BL mice with MNU or ENU and they found that the half-life of Me-PTE induced by MNU could be 7 days in kidney, lung and liver, and this number would increase to 10-15 weeks for Et-PTE induced by ENU. Using an HPLC-based method, Engelse *et al.* (14) compared the half-life of alkyl-PTEs with other base alkylation products in liver tissues of rats treated with ENU or *N,N*-dimethylnitrosamine (DMN), and found that alkyl-PTEs persist for more than 56 days.

The formation of alkyl-PTEs is also sequence-dependent. Guichard *et al.* (15), using ^{32}P postlabeling, showed that tissues of *N*-nitrosodiethylamine (NDEA)-treated mice contains a higher level of Et-PTEs with the 5′-nucleobase being thymine or guanine. Mass spectrometry analysis also revealed similar non-random distribution of alkyl-PTEs in different flanking sequences contexts in lung tissue DNA of mice treated with tobacco-specific nitrosoamine NNK and its metabolite NNAL (11).

Formation of POB- and PHB-phosphotriester

Many tobacco-specific nitrosamines, including *N*′-nitrosonornicotine (NNN), NNK as well as its metabolic product NNAL, were characterized as strong carcinogens in various laboratory animal studies (16-18), and both NNN and NNK have been classified as group 1 carcinogens by the International Agency for Research on Cancer (19). Since 2012, several studies revealed that phosphate adducts, including methyl (11), pyridyloxobutyl (POB) and pyridylhydroxybutyl (PHB) (11,20,21), could be detected in lung tissue of mice treated with the aforementioned nitrosamines (22) as well as lung cancer patients (23). The pathways for metabolic activation of NNN and NNK are shown in Figure 1-2 (11,20). Briefly speaking, after several steps of metabolic reactions, highly-reactive diazonium ions (compounds 1 and 2) could pyridyloxobutylate or pyridylhydroxybutylate phosphate backbone to form the corresponding POB- or PHB-PTEs. Additionally, one of the byproducts, methane diazohydroxide (compound 3) can also react with DNA to form Me-

PTEs. Similar as alkyl-PTEs, POB- and PHB-PTEs are also highly persistent *in vivo*, some of the PTEs could persist in tissues in over 70 weeks (24), suggesting their potential in serving as robust biomarkers for tobacco-induced carcinogenesis.

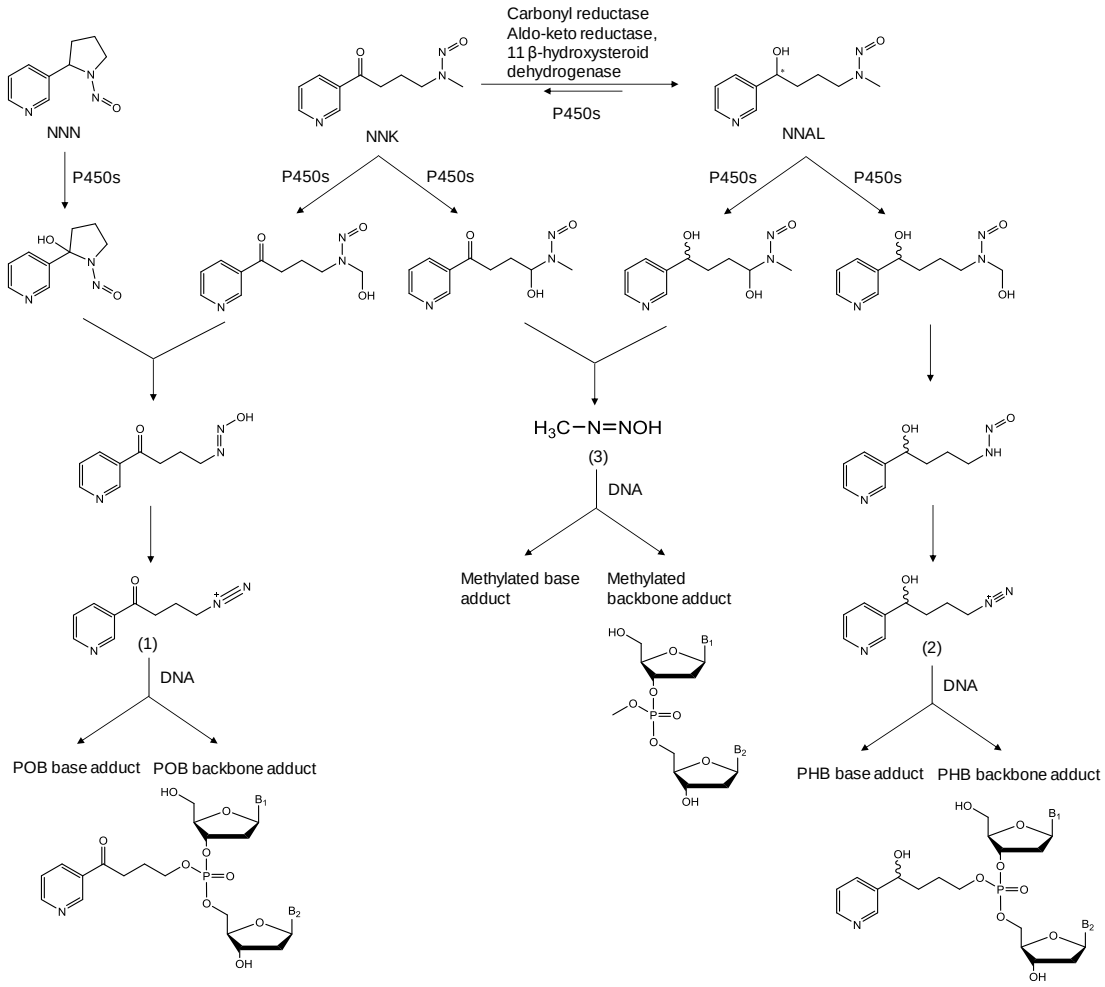


Figure 1-2. Formation of phosphate backbone adducts from metabolic activation of NNN, NNK and NNAL.

Repair of Alkylphosphotriester Lesions

The repair mechanism of alkyl-PTEs *in vivo* is under-investigated. In *E. coli* cells, Me-PTEs are removed by *E. coli* O^6-methylguanine DNA methyltransferase (Ada protein) (25). Ada is a 39-kDa protein having a wide range of substrates, including O^6-methylguanine, O^4-methylthymine and Me-PTEs. Meanwhile, Ada protein could also repair bulky alkylated DNA lesions (*n*-propyl, *i*-propyl and *n*-butyl) formed on the O^6 position of guanine at a lower rate (26).

The Ada-mediated repair of Me-PTEs proceeds through a zinc- and methylation-dependent electrostatic switch mechanism, where the methyl group on the phosphate backbone would be transferred into Cys-38 (27). In N-ada domain, four cysteine residues, Cys-24, -38, -69 and -72, coordinate with a zinc ion to achieve its methyltransferase activity. The sulfur atoms situated on the side chains of three cysteines, Cys-24, -69 and -72 form hydrogen bond to main chain of Ada protein, eventually stabilizing the structure but lessening their reactivity, while Cys-38 could be free from hydrogen bonding interaction, resulting in maintenance of its nucleophilicity. Lindahl *et al.* (28) observed the loss of Ada activity accompanied by the formation of methylated cysteine, indicating that Ada works as a suicide protein rather than an enzyme, and the methylated protein could be recovered upon addition of added external sacrificial reagent, e.g. methanethiol (29). Additionally, the repair of Me-PTEs mediated by Ada protein is specific to the S_p diastereomer. This can

be explained from the facts that, in B-form duplex DNA, the methyl group on S_p-Me-PTE projects into the solution from the phosphodiester backbone, whereas the methyl group on R_p diastereomer points inwards into DNA with greater steric crowding. As a result, the S_p diastereomer reacts more readily with Cys-38 of Ada protein (30).

Ada is a bifunctional protein: Apart from being a repair protein, it also works as a transcriptional regulator (31). The fact that Ada works as a suicidal protein would suggest that its removal of Me-PTEs is expensive and inefficient since one protein molecule could only remove one methyl group from the phosphate backbone. It turned out that methylated Ada protein could increase the *E. coli* capability in tolerating alkylating agents (32-34) via its interaction with the promoter sequence of *ada* gene, also known as "ada box" (27), followed by increased expression of Ada protein. The ada box is on upstream of RNA polymerase-binding site and works as a positive regulator of *ada* gene (35,36). Aside from *ada* gene, ada box also exists in other genes including *alkA*, *alkB* and *aidB* (37-39), which can be upregulated by adaptive response.

In mammalian cells, so far no repair pathway has been reported to remove alkyl-PTEs (6). Human O^6-methylguanine methyltransferase (MGMT) serves same function as *E. coli* Ada in repairing O^6-MedG but it cannot repair alkyl-PTEs (40,41). However, Le Pla *et al.* (42) raise the possibility of repair pathway in mammalian cells. In that study, the half-life of total Et-PTEs in liver DNA of mice treated with NDEA was measured by HPLC,

and an initial sharp decline of Et-PTEs (with $t_{1/2}$<24h) was found, which is probably attributed to active repair pathway.

Biological Consequences of Alkylphosphotriester Lesions

Even though the PTEs have come to our sight recently, research about their biological consequences still remains overlooked. Since PTE formation could neutralize the negative charge on the phosphate backbone, it could interfere with the interaction of proteins to DNA substrates. There are many proteins found to be sensitive to PTEs: MutS (43), MutY (44), RNA polymerase (45,46) and APE protein, a class-II apurinic/apyrimidinic endonuclease (47), etc.

Even though lacking research in how PTEs affect DNA replication *in vivo*, several *in vitro* studies infer that existence of PTEs could inhibit replication mediated by many polymerases. Miller *et al.* (48) revealed that an Et-PTE in a 12 mer oligodeoxyribonucleotide (ODN) could result in a decreased elongation rate (25 % for isomer I and 50 % for isomer II) mediated by *E. coli* polymerase I. Additionally, subsequent *in vitro* study also indicated that a mixture of 50:50 S_p and R_p of Et-PTE in ODN could impede primer extension by T4 DNA polymerase, which could be tolerated by increasing dNTP level (49). In addition to influencing polymerase elongation, *i*Pr-PTEs could also affect the unwinding process mediated by SF1 helicases *in vitro* (50).

LC-MS/MS-coupled Competitive Replication and Adduct Bypass (CRAB)Assay

One critical reason about why little attention has been paid on PTEs is the lack of suitable and precise analytical methods. PTEs are resistant to many enzymes commonly used in biochemistry and molecular studies for analyzing biological sequences, such as T4 PNK (51), DNase I, nuclease P1 and alkaline phosphatase (52-54), etc. Additionally, another major hurdle on PTE research is the lack of the methods for isolating and identifying independent S_p and R_p diastereomer-containing ODNs. In this dissertation, a HPLC-NMR coupled analysis has been employed to separate and characterize two diastereomers of PTE-carrying ODNs, and CRAB assay was utilized for analyzing the biological consequences of PTEs.

Competitive replication and adduct bypass (CRAB) assay was introduced by Dr. Essigmann and coworkers (55-57), which allows for the assessment about how a particular lesion interferes DNA replication and repair *in vivo*. In this assay, the lesion-containing ODNs (22 mer) or competitor ODN (25 mer, internal standard) are incorporated into M13mp7 bacteriophage, and the lesion/competitor ratio can be accurately normalized by ^{32}P-labeled PAGE gel. After normalization, the lesion-containing or lesion-free control vectors are mixed with competitor vector in a specific ratio and transfected into *E. coli* strains proficient or deficient in polymerases or repair protein. After several cycles' proliferation, the replication products of the aforementioned vectors are extracted and

purified, followed by PCR amplification and restriction endonuclease digestion to obtain short ODNs containing the replication products containing the initial lesion-containing site, which can be subjected to PAGE gel analysis to quantify bypass efficiency and mutation frequency by the following equation: Bypass efficiency (%) = (lesion signal/competitor signal)/(control signal/competitor signal)×100%. A brief workflow is shown on Figure 1-3. In CRAB assay, what being analyzed is the replication products of lesion-containing vectors, as a result, the resistance of PTEs to enzymes can be disregarded.

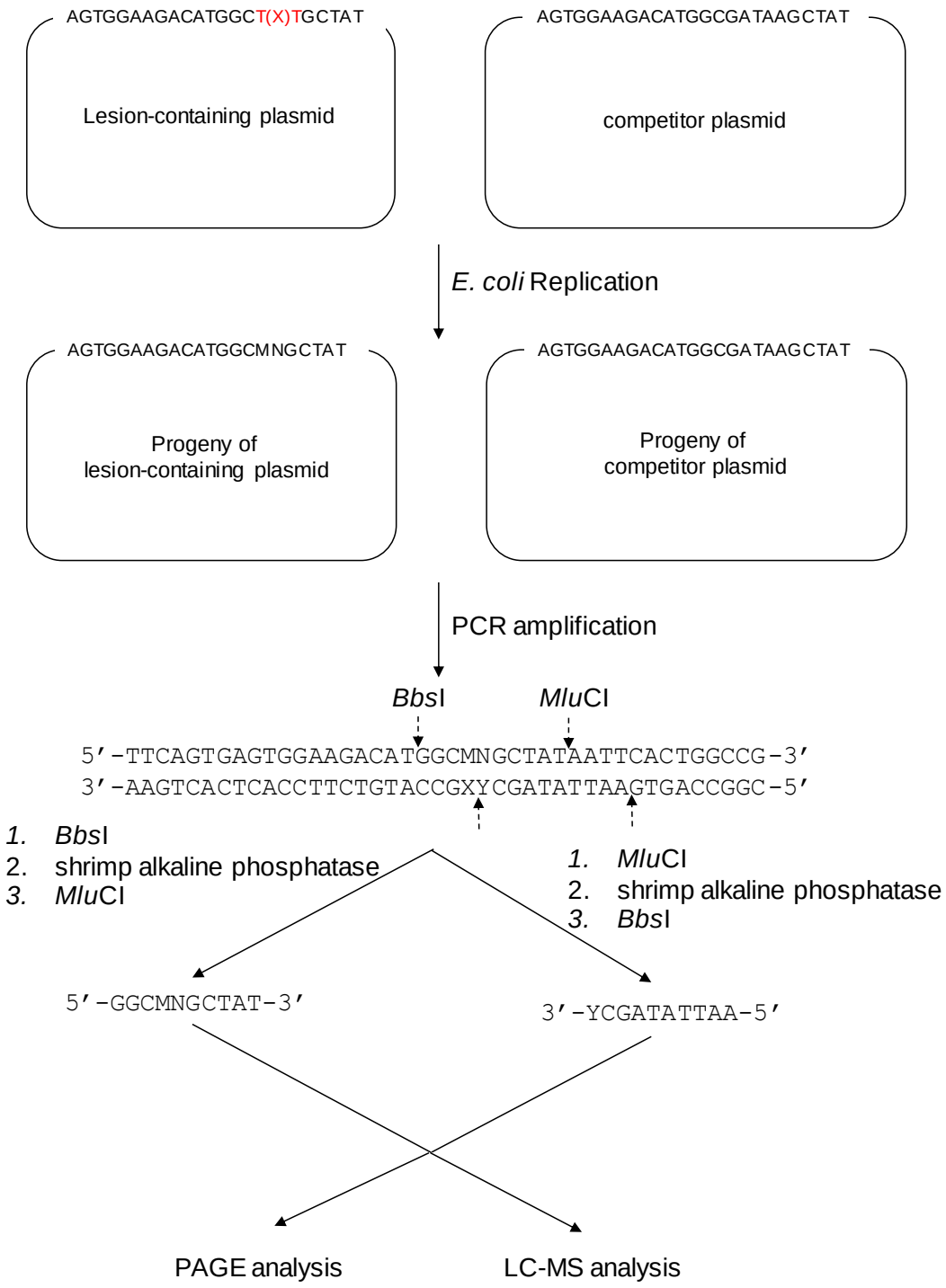


Figure 1-3. The workflow of CRAB assay.

Besides CRAB assay, mass spectrometry also becomes a powerful tool for PTE analysis due to its high sensitivity and specificity. LC-MS/MS can be coupled with CRAB assay to provide unambiguous insight about replication products. First, previous study has elucidated how to determine the sequence of ODNs by mass spectrometry (Figure 1-4)(60): In MS/MS mode, two types of fragmentation ions, w_n and

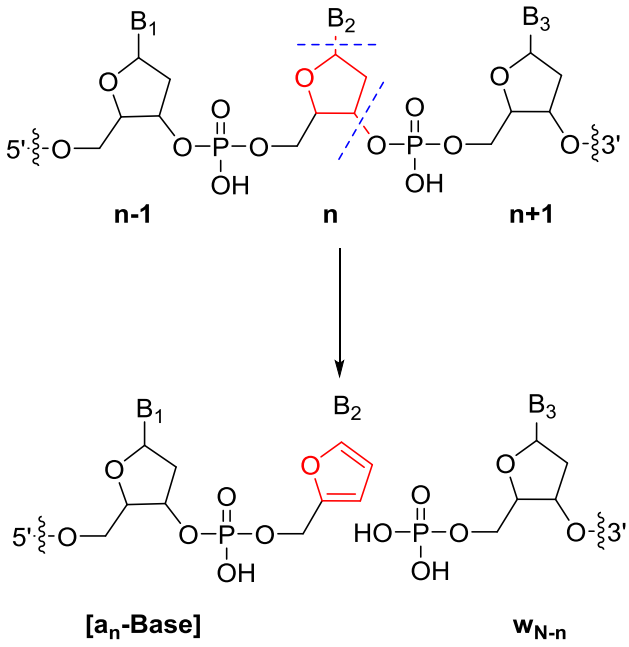


Figure 1-4. MS/MS fragmentation scheme of ODN.

[a_n-base] ions, could be generated, and (w_n-w_{n-1}) or ([a_{n+1}-base]-[a_n-base]) represents the mass-to-charge ratio (*m/z*) of nucleoside n, and each canonical nucleoside has a unique mass. With this method, the sequence of the target ODN can be identified. Second, LC-MS/MS analysis also made it possible to quantify the relative levels of different replication products (61).

Electrophoretic Mobility Shift Assay (EMSA)

Electrophoretic mobility shift assay (EMSA) is a powerful tool for analyzing the nucleic acid-protein interactions (62-64). The method is based on the electrophoretic mobility difference between free nucleic acids and protein-nucleic acid complex (65). In this assay, the nucleic acids are incubated with protein at different concentrations. After electrophoresis, free nucleic acid can be separated with protein-nucleic acid complex, and the dissociation constant (K_d) can be quantified using the following equation: [Unbound nucleic acid]/[Bound nucleic acid]= $K_d \times 1/$ [Protein] (66).

EMSA is a powerful tool for quantifying binding capacity of a single protein to nucleic acid. First, by employing ^{32}P-labeling, the assay can obtain high sensitivity (to ~0.1 nM of nucleic acid). Apart from radioisotope-labeling, EMSA is also compatible with other labeling strategies, including fluorescence (67) or chemiluminescence (68,69). Second, EMSA has less requirement to size or structure of nucleic acids and protein. It has been reported that the method is compatible with a broad range of nucleic acids structure [single strand, duplex, triplex (70) or circular DNA (71)) or size (from 5 nt (72) to several hundred bp (73)], protein size (74) and complexity of protein components (75).

EMSA, however, also has some disadvantages: The electrophoresis can influence the chemical equilibrium of protein-nucleic acid interaction, which can be compensated by low temperature and short time for electrophoresis. Moreover, except for dissociation constant,

EMSA can provide little information about protein-nucleic acids binding, like the sequences of nucleic acids involved in direct interaction with proteins, and other binding assays, e.g. quantitative DNase footprint titration (76), are necessary in order for obtaining detailed binding information.

Scope of This Dissertation

As mentioned above, multiple studies have shown the relationship between formation of PTEs and tobacco smoking, suggesting their potential in serving as a persistent biomarker for tobacco-related diseases, like lung cancer. In contrast, little is known about the biological consequences of PTEs, especially how PTEs influence DNA replication, and how they are repaired *in vivo*. Unlike nucleobase modifications, the structures of PTEs are complicated, and many factors, including properties of backbone modifications, stereochemistry and sequence contexts, could affect the biological consequences of PTEs.

In this dissertation, we attempt to attain the following objectives: 1) To synthesize and characterize PTEs with different groups (including methyl, ethyl, *n*-propyl, *n*-butyl, POB and PHB), stereochemistry (S_p and R_p) or flanking base sequences. 2) To reveal how size, stereochemistry, flanking base sequences of alkyl-PTEs, and the SOS-induced translesion synthesis polymerases (Pol II, Pol IV and Pol V in *E. coli*) bypass of these lesions in *E. coli*. 3) To define the role of Ada protein in modulating the repair and translesion synthesis of alkyl-PTE lesion in *E. coli* cells.

References

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.

2. Fu, D., Calvo, J.A. and Samson, L.D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer*, **12**, 104-120.

3. Liu, S. and Wang, Y. (2015) Mass spectrometry for the assessment of the occurrence and biological consequences of DNA adducts. *Chem. Soc. Rev.*, **44**, 7829-7854.

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

5. Drablos, F., Feyzi, E., Aas, P.A., Vaagbo, C.B., Kavli, B., Bratlie, M.S., Pena-Diaz, J., Otterlei, M., Slupphaug, G. and Krokan, H.E. (2004) Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair*, **3**, 1389-1407.

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

7. Elmore, D.T., Gulland, J.M., Jordan, D.O. and Taylor, H.F. (1948) The reaction of nucleic acids with mustard gas. *Biochem. J.*, **42**, 308-316.

8. Bannon, P. and Verly, W. (1972) Alkylation of phosphates and stability of phosphate triesters in DNA. *Eur. J. Biochem.*, **31**, 103-111.

9. Brimacombe, R., Griffin, B.E., Haines, J., Haslam, W.J. and Reese, C. (1965) An approach to the methylation of polynucleotides. *Biochemistry*, **4**, 2452-2458.

10. Lawley, P.D. (1973) Reaction of *N*-methyl-*N*-nitrosourea (MNUA) with ^{32}P-labelled DNA: evidence for formation of phosphotriesters. *Chem. Biol. Interact.*, **7**, 127-130.

11. Ma, B., Zarth, A.T., Carlson, E.S., Villalta, P.W., Upadhyaya, P., Stepanov, I. and Hecht, S.S. (2018) 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*, **31**, 48-57.

12. Sun, L. and Singer, B. (1975) Specificity of different classes of ethylating agents toward various sites of HeLa cell DNA *in vitro* and *in vivo*. *Biochemistry*, **14**, 1795-1802.

13. Shooter, K.V. and Slade, T.A. (1977) The stability of methyl and ethyl phosphotriesters in DNA *in vivo*. *Chem. Biol Interact.*, **19**, 353-361.

14. Den Engelse, L., Menkveld, G.J., De Brij, R.J. and Tates, A.D. (1986) 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*, **7**, 393-403.

15. Guichard, Y., Jones, G.D. and Farmer, P.B. (2000) Detection of DNA alkylphosphotriesters by ^{32}P postlabeling: evidence for the nonrandom manifestation of phosphotriester lesions *in vivo*. *Cancer Res.*, **60**, 1276-1282.

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

17. Hecht, S.S. (2003) Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat. Rev. Cancer*, **3**, 733.

18. Hecht, S.S., Stepanov, I. and Carmella, S.G. (2016) Exposure and metabolic activation biomarkers of carcinogenic tobacco-specific nitrosamines. *Acc. Chem. Res.*, **49**, 106-114.

19. Cancer, I.A.f.R.o. (2018).

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

21. Ma, B., Zarth, A.T., Carlson, E.S., Villalta, P.W., Upadhyaya, P., Stepanov, I. and Hecht, S.S. (2018) 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*, **39**, 232-241.

22. Ma, B., Stepanov, I. and Hecht, S.S. (2019) Recent studies on DNA adducts resulting from human exposure to tobacco smoke. *Toxics.* , **7**. 16

23. Ma, B., Villalta, P.W., Hochalter, J.B., Stepanov, I. and Hecht, S.S. (2019) Methyl DNA phosphate adduct formation in lung tumor tissue and adjacent normal tissue of lung cancer patients. *Carcinogenesis*.

24. Ma, B., Villalta, P.W., Zarth, A.T., Kotandeniya, D., Upadhyaya, P., Stepanov, I. and Hecht, S.S. (2015) 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.*, **28**, 2151-2159.

25. Weinfeld, M., Drake, A.F., Saunders, J.K. and Paterson, M.C. (1985) Stereospecific removal of methyl phosphotriesters from DNA by an *Escherichia coli ada*$^{+}$ extract. *Nucleic Acids Res.*, **13**, 7067-7077.

26. Morimoto, K., Dolan, M.E., Scicchitano, D. and Pegg, A.E. (1985) Repair of O^6-propylguanine and O^6-butylguanine in DNA by O^6-alkylguanine-DNA alkyltransferases from rat liver and *E. coli*. *Carcinogenesis*, **6**, 1027-1031.

27. He, C., Hus, J.C., Sun, L.J., Zhou, P., Norman, D.P., Dotsch, V., Wei, H., Gross, J.D., Lane, W.S., Wagner, G. *et al.* (2005) A methylation-dependent electrostatic switch controls DNA repair and transcriptional activation by *E. coli ada*. *Mol. Cell*, **20**, 117-129.

28. Lindahl, T., Demple, B. and Robins, P. (1982) Suicide inactivation of the E. coli O^6-methylguanine-DNA methyltransferase. *EMBO J.,* **1**, 1359-1363.

29. He, C., Wei, H. and Verdine, G.L. (2003) Converting the sacrificial DNA repair protein N-Ada into a catalytic methyl phosphotriester repair enzyme. *J. Am. Chem. Soc.*, **125**, 1450-1451.

30. Samson, L. and Cairns, J. (1977) A new pathway for DNA repair in *Escherichia coli*. *Nature*, **267**, 281-283.

31. Teo, I., Sedgwick, B., Demple, B., Li, B. and Lindahl, T. (1984) Induction of resistance to alkylating agents in *E. coli*: the *ada*$^{+}$ gene product serves both as a regulatory protein and as an enzyme for repair of mutagenic damage. *EMBO J.*, **3**, 2151-2157.

32. Mccarthy, J.G., Edington, B.V. and Schendel, P.F. (1983) Inducible repair of phosphotriesters in *Escherichia coli*. *P. Natl. Acad. Sci.-Biol.*, **80**, 7380-7384.

33. Nakamura, T., Tokumoto, Y., Sakumi, K., Koike, G., Nakabeppu, Y. and Sekiguchi, M. (1988) Expression of the *ada* gene of *Escherichia coli* in response to alkylating agents. Identification of transcriptional regulatory elements. *J. Mol. Biol.*, **202**, 483-494.

34. Samson, L. (1992) The suicidal DNA-repair methyltransferases of microbes. *Mol. Microbiol.*, **6**, 825-831.

35. Sedgwick, B. (1983) Molecular-cloning of a gene which regulates the adaptive response to alkylating-agents in *Escherichia Coli*. *Mol. Gen. Genet.*, **191**, 466-472.

36. Teo, I., Sedgwick, B., Kilpatrick, M.W., McCarthy, T.V. and Lindahl, T. (1986) The intracellular signal for induction of resistance to alkylating agents in *E. coli*. *Cell*, **45**, 315-324.

37. Falnes, P.O., Johansen, R.F. and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature*, **419**, 178-182.

38. Trewick, S.C., Henshaw, T.F., Hausinger, R.P., Lindahl, T. and Sedgwick, B. (2002) Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature*, **419**, 174-178.

39. Kataoka, H. and Sekiguchi, M. (1985) Molecular cloning and characterization of the alkB gene of *Escherichia coli*. *Mo.l Gen. Genet.*, **198**, 263-269.

40. Yarosh, D.B., Fornace, A.J. and Day III, R.S. (1985) 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*, **6**, 949-953.

41. Yarosh, D.B., Rice, M., Day, R.S., 3rd, Foote, R.S. and Mitra, S. (1984) O^6-Methylguanine-DNA methyltransferase in human cells. *Mutat. Res.*, **131**, 27-36.

42. Le Pla, R.C., Guichard, Y., Bowman, K.J., Gaskell, M., Farmer, P.B. and Jones, G.D.D. (2004) Further development of P-32-postlabeling for the detection of alkylphosphotriesters: Evidence for the long-term nonrandom persistence of ethyl-phosphotriester adducts *in vivo*. *Chem. Res. Toxicol.*, **17**, 1491-1500.

43. Biswas, I. and Hsieh, P. (1997) Interaction of MutS protein with the major and minor grooves of a heteroduplex DNA. *J. Biol. Chem.*, **272**, 13355-13364.

44. Lu, A.L., Tsai-Wu, J.J. and Cillo, J. (1995) DNA determinants and substrate specificities of *Escherichia coli* MutY. *J. Biol. Chem.*, **270**, 23582-23588.

45. Marushige, K. and Marushige, Y. (1983) Template properties of DNA alkylated with *N*-methyl-*N*-nitrosourea and *N*-ethyl-*N*-nitrosourea. *Chem. Biol. Interact.*, **46**, 179-188.

46. Siebenlist, U. and Gilbert, W. (1980) Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. *P. Natl. Acad. Sci. Biol.*, **77**, 122-126.

47. Wilson, D.M., 3rd, Takeshita, M. and Demple, B. (1997) Abasic site binding by the human apurinic endonuclease, Ape, and determination of the DNA contact sites. *Nucleic Acids Res.*, **25**, 933-939.

48. Miller, P.S., Chandrasegaran, S., Dow, D.L., Pulford, S.M. and Kan, L.S. (1982) Synthesis and template properties of an ethyl phosphotriester modified decadeoxyribonucleotide. *Biochemistry*, **21**, 5468-5474.

49. Tsujikawa, L., Weinfield, M. and Reha-Krantz, L.J. (2003) Differences in replication of a DNA template containing an ethyl phosphotriester by T4 DNA polymerase and *Escherichia coli* DNA polymerase I. *Nucleic Acids Res.*, **31**, 4965-4972.

50. Suhasini, A.N., Sommers, J.A., Yu, S., Wu, Y., Xu, T., Kelman, Z., Kaplan, D.L. and Brosh, R.M., Jr. (2012) DNA repair and replication fork helicases are differentially affected by alkyl phosphotriester lesion. *J. Biol. Chem.*, **287**, 19188-19198.

51. Weinfeld, M. and Livingston, D.C. (1986) Synthesis and properties of oligodeoxyribonucleotides containing an ethylated internucleotide phosphate. *Biochemistry*, **25**, 5083-5091.

52. Bowman, K.J., Pla, R.L., Guichard, Y., Farmer, P.B. and Jones, G.D. (2001) 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.*, **29**, E101.

53. Le Pla, R.C., Guichard, Y., Bowman, K.J., Gaskell, M., Farmer, P.B. and Jones, G.D. (2004) 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.*, **17**, 1491-1500.

54. Saris, C.P., Damman, S.J., van den Ende, A.M., Westra, J.G. and den Engelse, L. (1995) A ^{32}P-postlabelling assay for the detection of alkylphosphotriesters in DNA. *Carcinogenesis*, **16**, 1543-1548.

55. Delaney, J.C. and Essigmann, J.M. (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions *in vivo*. *Methods Enzymol.*, **408**, 1-15.

56. Delaney, J.C. and Essigmann, J.M. (1999) Context-dependent mutagenesis by DNA lesions. *Chem. Biol.*, **6**, 743-753.

57. Delaney, J.C. and Essigmann, J.M. (2004) Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 14051-14056.

58. Haglund, J., Van Dongen, W., Lemiere, F. and Esmans, E.L. (2004) Analysis of DNA-phosphate adducts *in vitro* using miniaturized LC-ESI-MS/MS and column switching: Phosphotriesters and alkyl cobalamins. *J. Am. Soc. Mass. Spectr*, **15**, 593-606.

59. Zhang, F.G., Bartels, M.J., Pottenger, L.H., Gollapudi, B.B. and Schisler, M.R. (2007) Quantitation of lower levels of the DNA adduct of thymidylyl(3'-5 ')thymidine methyl phosphotriester by liquid chromatography/negative atmospheric pressure chemical ionization tandem mass spectrometry. *Rapid Commun. Mass Spectrom.*, **21**, 1043-1048.

60. McLuckey, S.A., Van Berkel, G.J. and Glish, G.L. (1992) Tandem mass spectrometry of small, multiply charged oligonucleotides. *J. Am. Soc. Mass Spectrom.*, **3**, 60-70.

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

62. Lane, D., Prentki, P. and Chandler, M. (1992) Use of gel retardation to analyze protein-nucleic acid interactions. *Microbiol. Rev.*, **56**, 509-528.

63. Garner, M.M. and Revzin, A. (1986) The use of gel-electrophoresis to detect and study nucleic-acid protein interactions. *Trends Biochem. Sci.*, **11**, 395-396.

64. Fried, M.G. (1989) Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. *Electrophoresis*, **10**, 366-376.

65. Hellman, L.M. and Fried, M.G. (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.*, **2**, 1849-1861.

66. Dai, X.X., Wang, T.L., Gonzalez, G. and Wang, Y.S. (2018) Identification of YTH domain-containing proteins as the readers for *N*1-Methyladenosine in RNA. *Anal. Chem.*, **90**, 6380-6384.

67. Rye, H.S., Drees, B.L., Nelson, H.C. and Glazer, A.N. (1993) Stable fluorescent dye-DNA complexes in high sensitivity detection of protein-DNA interactions. Application to heat shock transcription factor. *J. Biol. Chem.*, **268**, 25229-25238.

68. Berger, R., Duncan, M.R. and Berman, B. (1993) Nonradioactive gel mobility shift assay using chemiluminescent detection. *Biotechniques*, **15**, 650-&.

69. Rodgers, J.T., Patel, P., Hennes, J.L., Bolognia, S.L. and Mascotti, D.P. (2000) Use of biotin-labeled nucleic acids for protein purification and agarose-based chemiluminescent electromobility shift assays. *Anal. Biochem.*, **277**, 254-259.

70. Musso, M., Bianchi-Scarra, G. and Van Dyke, M.W. (2000) The yeast CDP1 gene encodes a triple-helical DNA-binding protein. *Nucleic Acids Res.*, **28**, 4090-4096.

71. Nordheim, A. and Meese, K. (1988) Topoisomer Gel Retardation - Detection of anti-Z-DNA antibodies bound to Z-DNA within supercoiled DNA minicircles. *Nucleic Acids Res.*, **16**, 21-37.

72. Rasimas, J.J., Kar, S.R., Pegg, A.E. and Fried, M.G. (2007) Interactions of human O^6-alkylguanine-DNA alkyltransferase (AGT) with short single-stranded DNAs. *J. Biol. Chem.*, **282**, 3357-3366.

73. Fried, M.G. and Daugherty, M.A. (1998) Electrophoretic analysis of multiple protein-DNA interactions. *Electrophoresis*, **19**, 1247-1253.

74. Talanian, R.V., McKnight, C.J. and Kim, P.S. (1990) Sequence-specific DNA binding by a short peptide dimer. *Science*, **249**, 769-771.

75. Varshavsky, A. (1987) Electrophoretic assay for DNA-binding proteins. *Methods Enzymol.*, **151**, 551-565.

76. Brenowitz, M., Senear, D.F., Shea, M.A. and Ackers, G.K. (1986) Quantitative DNase footprint titration: a method for studying protein-DNA interactions. *Methods Enzymol.*, **130**, 132-181.

CHAPTER 2

Cytotoxic and Mutagenic Properties of Alkyl Phosphotriester Lesions in *Escherichia coli* Cells

INTRODUCTION

DNA carries the genetic information, but it has limited chemical stability, where spontaneous hydrolysis can result in deamination of nucleobases and cleavage of *N*-glycosidic bonds in DNA (1). The integrity of the human genome can also be compromised upon exposure to many endogenous and exogenous agents, which may result in mutagenesis and carcinogenesis (2,3).

Alkylation is a common type of DNA damage (4), and DNA alkylation constitutes the major mechanism of action for many commonly prescribed cancer chemotherapeutic agents (5). So far, the majority of studies about DNA alkylation have focused on nucleobase modifications, which is attributed to the potential mutagenic effects of these lesions (4). Aside from forming nucleobase adducts, alkylating agents can also attack one of the non-carbon-bonded oxygen atoms of internucleotide phosphate group to yield backbone alkylation products, i.e. the alkyl phosphotriester (alkyl-PTE) lesions (6). In this vein, Me-

and Et-PTE (Figure 2-1) were found to account for 12-17% and 55-57% of the total alkylation in duplex DNA treated with *N*-nitroso-*N*-methylurea and *N*-nitroso-*N*-ethylurea, respectively (7). In addition, a very recent study showed that pyridylhydroxybutyl-PTE adducts constitute 38-55% and 34-40% of all the measured pyridine-containing DNA adducts in lung and liver tissues, respectively, of rats treated through drinking water with a tobacco-specific nitrosamine, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (8). Depending on which of the two non-carbon-bound oxygen atoms on backbone phosphate is alkylated, alkyl-PTEs can form in the S_p or R_p configuration (Figure 2-1) (6).

Alkyl-PTEs are known to persist in mammalian tissues. A previous study showed that, in liver tissues of mice, Et-PTE displayed longer half-life ($t_{1/2}$, up to 32 days) than any nucleobase ethylation products (9). Another study by Shooter *et al*. (10) revealed that the $t_{1/2}$ for Me-PTE was 7 days in kidney and lung tissues of C57BL mice, which is much shorter than the $t_{1/2}$ values for Et-PTE in the same tissues (10-15 weeks). The poorer repair of the alkyl-PTE lesions in mammalian tissues suggest that they are more likely to be encountered by DNA replication machinery than nucleobase alkylation products.

The repair and biological consequences of alkyl-PTE lesions are under-investigated (6). In this respect, the formation of alkyl-PTEs results in neutralization of the negative charge of the phosphate backbone, which interferes with the binding of proteins to DNA, thereby resulting in an inhibition of the normal functions of many DNA-binding proteins. For

instance, a single isopropyl-PTE could inhibit SF2 family of DNA helicases, including RecQ and WRN (11). It was also observed that the S_p, but not the R_p diastereomer of Me-PTE could be removed by Ada protein in *Escherichia coli* (12,13). In addition, *in vitro* biochemical studies revealed that a 50:50 mixture of S_p- and R_p-Et-PTE could inhibit primer extension mediated by T4 DNA polymerase (14). However, it remains unclear how the alkyl-PTE lesions affect DNA replication in cells.

In the present study, we employed a robust shuttle vector-based method (15,16), together with mass spectrometry, to assess how alkyl-PTE lesions with different sizes of the alkyl group (from methyl to *n*-butyl) and stereochemical configurations (S_p and R_p) affect the fidelity and efficiency of DNA replication in *E. coli* cells, and how replication across these lesions is influenced by translesion synthesis DNA polymerases and the Ada protein.

MATERIALS AND METHODS

Materials

All chemicals, unless specified, were from Sigma-Aldrich (St Louis, MO, USA) or Thermo Fisher Scientific (Pittsburg, PA, USA). Common reagents for solid-phase DNA synthesis were from Glen Research (Sterling, VA, USA) and all the unmodified oligodeoxyribonucleotides (ODNs) were from Integrated DNA Technologies (Coralville,

IA, USA). 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) was from Oakwood Products Inc. (West Columbia, SC, USA), and [γ-^{32}P]ATP was obtained from Perkin Elmer (Piscataway, NJ, USA). All enzymes were purchased from New England Biolabs (Ipswich, MA, USA).

M13mp7(L2) plasmid, wild-type AB1157 *E. coli* strains, C215 *E. coli* strains and Ada-deficient *E. coli* strains were kindly provided by Prof. John M. Essigmann (17,18). Polymerase-deficient AB1157 strains [*ΔpolB1*::spec (Pol II deficient), *ΔdinB* (Pol IV-deficient), *ΔumuC*::kan (Pol V deficient) and *ΔpolB1*::spec *ΔdinB ΔumuC*::kan (Pol II, Pol IV, Pol V-triple knockout)] were generously provided by Prof. Graham C. Walker (17).

Chemical Synthesis

The thymidine alkylphosphoramidites were synthesized following previously published procedures (Scheme 2-1) (19,20). Compound **1** (200 mg, 0.75 mmol) was dissolved in anhydrous diethyl ether (7 ml) in an ice bath under argon atmosphere, to which solution were added the corresponding alcohol (140 μl) and triethylamine (4.0 eq., 306 mg). The solution was stirred at room temperature overnight, mixed with 8 ml petroleum ether, and filtered. The resulting precipitate was washed with ether. The filtrate and the ether solution were pooled and evaporated under reduced pressure to yield **2a-b**, which were employed directly for the synthesis in the next step.

For the preparation of compounds **2c-d**, compound **1** (300 mg, 1.13 mmol) in 2 ml tetrahydrofuran, the corresponding alcohol (140 μl) and triethylamine (230 μl, 1.63

mmol) were added to a round bottom flask, which was under an argon atmosphere and placed in an ice water bath. The solution was stirred at room temperature for 1 h and filtered. The filtrate was concentrated *in vacuo*. Pentane (10 ml) was added to the resulting solid, and the mixture was filtered. The solvent in the filtrate was removed under vacuum to yield compounds **2c-d**.

For dimethoxytrityl (DMTr) protection of the 5′-hydroxyl group, compound **4** was prepared following the previously published procedures (21). Briefly, compound **3** (100 mg, 0.41 mmol) was dissolved in anhydrous pyridine (10 ml), and to the mixture were added dimethoxytrityl chloride (1.5 eq.) and 4-dimethylaminopyridine (0.5% mol). After stirring at room temperature for 10 h, the solution was concentrated *in vacuo* and the residues were purified with silica gel column chromatography by using ethyl acetate as the mobile phase to yield compound **4**.

For phosphoramidite building block synthesis, to a round bottom flask, which contained a solution of compound **4** (70 mg, 0.13 mmol) in dry dichloromethane (1.5 ml), were added compound **2a-b** (1 eq.) and diisopropylamine hydrotetrazolide (1 eq.), and the mixture was stirred at room temperature under argon. After 1 h, TLC analysis indicated complete conversion of the starting material. The solvent was removed *in vacuo* and isolated by silica gel column chromatography with a mixture of hexane, ethyl acetate, and triethylamine (49:49:2, v/v) to yield the desired products **5a-b**.

For the syntheses of compounds **5c-d**, compound **2c-d** (1 eq.) and tetrazole in acetonitrile (0.3 ml, 0.135 mmol) were added to a round bottom flask containing a solution of compound **4** (70 mg, 0.13 mmol) in dry dichloromethane (1.5 ml), and the mixture was stirred at room temperature under argon atmosphere. The reaction was continued for 5 h, and the desired products **5c-d** were purified following the aforementioned procedures for compounds **5a-b**.

The reaction yields and spectroscopic characterizations of the above-synthesized products are provided in the online Supplementary Materials. The nuclear magnetic resonance (NMR) spectra for these compounds are shown in Figures 2-2 and 2-3.

ODN synthesis

The 12-mer lesion-containing ODNs, 5′-ATGGCT(X)TGCTAT-3′ ('X' designates the location of the alkyl-PTE), and Me-PTE-containing dimer 5′-T(Me)T-3′ were synthesized on a Beckman Oligo 1000M DNA synthesizer (Fullerton, CA) at 1 μmol scale. The corresponding phosphoramidite building block was dissolved in anhydrous acetonitrile at a concentration of 0.067 mM. Commercially available ultramild phosphoramidite building blocks were used for the incorporation of unmodified nucleotides (Glen Research Inc., Sterling, VA, USA) following the standard ODN assembly protocol. The ODNs were cleaved from the controlled pore glass support and deprotected with concentrated ammonium hydroxide at room temperature for 55 min. The solvent was

removed by using a Speed-vac, and the resulting solid residues were redissolved in water and purified by high-performance liquid chromatography (HPLC).

HPLC

The HPLC separation was performed on an Agilent 1100 system with a Kinetex XB-C18 column (4.6 × 150 mm, 5 μm in particle size and 100 Å in pore size; Phenomenex Inc., Torrance, CA, USA). For the purification of the synthesized 12-mer ODNs, the mobile phases consisted of 50 mM triethylammonium acetate (pH 7.0, Solution A) and a mixture of solution A and acetonitrile (70/30, v/v, Solution B). A gradient of 5-30% B in 5 min and 30-60% B in 65 min was employed, and the flow rate was 0.7 ml/min. The HPLC traces for the purification of the 12-mer lesion-containing ODNs are shown in Figure 2-6 and the electrospray ionization-mass spectra (ESI-MS) and tandem MS (MS/MS) of these ODNs are displayed in Figures 2-7 to 2-10.

Determination of stereochemical configurations of Me-PTE-containing ODNs

The Me-PTE-containing dimer 5′-T(Me)T-3′ was first purified by HPLC using a Synergi Fusion-RP column (10 × 150 mm, 4 μm in particle size and 80 Å in pore size; Phenomenex Inc., Torrance, CA, USA). The mobile phases were water (Solution A) and methanol (Solution B). A gradient of 5-25% B in 5 min and 25-50% B in 75 min was used, and the flow rate was 1.1 ml/min. The HPLC traces are shown in Figure 2-4a and

the products were identified by MS and ^{31}P-NMR. For NMR analyses, the two purified diastereomers were dissolved in 300 µl CD_3OD, and triphenylphosphine was added as the reference for chemical shift calibration. The ^{31}P-NMR spectra were recorded on a Bruker Avance Neo 400 spectrometer at 80 MHz for ^{31}P (Figure 2-4b, c).

To determine the stereochemical configurations of the Me-PTEs in 12 mer ODNs, we digested the lesion-containing ODNs with a cocktail of four enzymes and analyzed the digestion mixture by HPLC to examine whether the two diastereomers in 12 mer ODNs exhibited the same elution order as those in the dimers. The digestion was conducted following previously published procedures (22). Briefly, nuclease P1 (0.1 U/µg DNA), phosphodiesterase 2 (0.000125 U/µg DNA), and 1 nmol of 12 mer ODNs with S_p- or R_p-Me-PTE were incubated in a 100-µl solution containing 300 mM sodium acetate (pH 5.6) and 10 mM $ZnCl_2$. After incubation at 37°C for 24 h, alkaline phosphatase (0.1 U/µg DNA), phosphodiesterase 1 (0.00025 U/µg DNA), and 50 µl of 0.5 M Tris-HCl (pH 8.9) were added to the digestion mixture, and the mixture was incubated at 37°C for 2 h. The enzymes were subsequently removed by chloroform extraction and the solution was dried by lyophilization. The resulting solid residue was reconstituted in doubly distilled water (100 µl) and analyzed by HPLC. A Thermo Hypersil Gold C18 column (4.6×250 mm, 3 µl in particle size, 300 Å in pore size, Thermo Scientific Inc., Waltham, MA) was employed for the separation of the digestion

mixture. A solution of 10 mM ammonium formate and methanol were used as mobile phases A and B, respectively. The gradient profile was 0% B in 42 min, 0-2% B in 1 min, 2% B for 17 min, 2-5% B in 1 min, 5-25% B in 5 min, 25-35% B in 75 min. The HPLC traces are shown in Figure 2-5.

Construction of single-stranded lesion-containing and lesion-free competitor M13 genomes

The 12-mer alkyl-PTE-containing ODNs were 5′-phosphorylated and ligated with a 10-mer ODN (5′-ACTGGAAGAC-3′) in the presence of ligation buffer with T4 DNA ligase and ATP at 16°C for 8 h. The resulting 22-mer ODNs were purified by denaturing polyacrylamide gel.

The lesion-containing and lesion-free M13mp7 (L2) genomes were prepared following previously reported procedures (21,23). First, 20 pmol of M13 plasmid was digested with 40 U *Eco*RI-HF at room temperature for 8 h to linearize the vector. The linearized vector was mixed with 2 scaffolds, 5′-CTTCCACTCACTGAATCATGGTCATAGCTTTC-3′ and 5′-AAAACGACGGCCAGTGAATTATAGC-3′ (25 pmol), each spanning one end of the linear vector. To the mixture was subsequently added 30 pmol of 5′-phosphorylated 22-mer lesion-containing ODN or 25-mer competitor ODN (5′-GCAGGATGTCATGGCGATAAGCTAT-3′). The resulting mixture was treated with T4

DNA ligase at 16°C for 8 h, followed by incubation with T4 DNA polymerase (22.5 U) at 16°C for 2.5 h to degrade the unligated vector and excess scaffolds. The lesion-containing and lesion-free plasmids were purified using Cycle Pure Kit (Omega) and subsequently normalized against the competitor plasmid, following published procedures (21,23).

Transfection of lesion-free, lesion-containing and competitor plasmids into *E. coli* cells

The lesion-free or lesion-containing M13 genome was mixed with the competitor genome at a 1:1 molar ratio. The mixtures were transfected into SOS-induced, electrocompetent AB1157 *E. coli* strain as well as the isogenic cells deficient in Pol II, Pol IV, Pol V, or all three in combination. The lesion-free or lesion-containing M13 genome was also mixed with the competitor genome at a 4:1 molar ratio and transfected into electrocompetent Ada-proficient (C215) and Ada-deficient (C217) *E. coli* cells, as described elsewhere (21). The SOS induction was conducted by irradiating *E. coli* cells with 254 nm ultraviolet light at a dose of 45 J/m^2 (17). The AB1157, C215 and C217 *E. coli* cells were then grown in lysogeny broth (LB) medium at 37°C for 5.5 h. The phage was recovered from the supernatant by centrifugation at 13200 r.p.m. for 5 min and further transfected into SCS110 cells to increase the progeny/lesion-genome ratio. The amplified phage was purified by Qiaprep Spin M13 Kit (Qiagen) to obtain the ssM13 DNA template for PCR amplification.

Quantification of bypass efficiencies and mutation frequencies

A modified version of the competitive replication and adduct bypass (CRAB) assay (15,16,23,24) was employed to assess the bypass efficiencies and mutation frequencies of the alkyl-PTE lesions in *E. coli* cells. The sequence region of interest in the M13 template was amplified by PCR with the use of Phusion high-fidelity DNA polymerase. The primers were 5′-YCAGCTATGACCATGATTCAGTGAGTGGA-3′ and 5′-YTCGGTGCGGGCCTCTTCGCTATTAC-3′ ('Y' denotes the $H_2N(CH_2)_6$- group conjugated to the 5′ phosphate group of the ODNs). The amplification started from annealing at 98°C for 30 s, followed by 35 cycles of amplification, each of which consisted of 10 s at 98°C, 30 s at 65°C and 15 s at 72°C, and then with a final 5-min extension at 72°C and ending at 4°C. The PCR products were purified by using Cycle Pure Kit (Omega). For determining the bypass efficiencies and mutation frequencies, 100 ng of PCR products were digested with 10 U *Bbs*I-HF restriction endonucleases together with 10 U recombinant shrimp alkaline phosphatase (rSAP) in 10 μl CutSmart buffer (New England Biolabs) at 37°C for 25 min, followed by heating at 80°C for 10 min to deactivate the phosphatase. To the mixture were subsequently added 5 mM DTT, 1.66 pmol [γ-^{32}P] ATP, 10 U T4 polynucleotide kinase (T4 PNK), T4 PNK buffer and water to give a total volume of 15 μl. After a 30-min incubation at 37°C, the T4 PNK was deactivated by heating the solution at 80°C for 10 min. To the resulting solution was

added 10 U *Mlu*CI restriction endonuclease, and the mixture was then incubated at 37°C for 25 min. The reaction was subsequently quenched by adding 15 μl formamide gel loading buffer which contained xylene cyanol FF and bromophenol blue dyes. The mixture was resolved by using 30% native polyacrylamide gel (acrylamide:bis-acrylamide = 19:1), and the gel band intensities were determined by phosphorimager analysis on a Typhoon 9410 variable-mode imager.

The two-step restriction endonuclease digestion gave rise to a 10-mer duplex d(p*GGCMNGCTAT)/d(AATTATAGCY), where 'M' and 'N' designate the nucleobases at the dinucleotide sites initially flanking the alkyl-PTE lesions, 'Y' is the complementary base of 'N' in the opposite strand, and p* denotes the [5′-^{32}P]-labeled phosphate. The bypass efficiency was calculated by using following equation: Bypass efficiency (%) = (lesion signal/competitor signal)/(control signal/competitor signal)×100%.

Identification of mutagenic products by LC-MS and MS/MS

The PCR products were digested with 50 U *Bbs*I-HF restriction endonuclease and 20 U rSAP in 250 μl CutSmart buffer at 37°C for 2 h, and the phosphatase was subsequently deactivated by heating at 80°C for 20 min. To the mixture was added 50 U *Mlu*CI and the solution was incubated at 37°C for 1 h. The solution was then extracted once with phenol/chloroform/isoamyl alcohol (25:24:1, v/v). The aqueous layer was dried, desalted by HPLC and redissolved in 20 μl water. A 10-μl aliquot was injected for LC-MS/MS

analysis on an LTQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA, USA). 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 35 min of 20–50% methanol in 400 mM HFIP (pH was adjusted to 7.0 with triethylamine). The mass spectrometer was set up for monitoring the fragmentation of the $[M-3H]^{3-}$ ions of 10-mer d(GGCMNGCTAT) and d(AATTATAGCY), with 'M', 'N', and 'Y' being 'A', 'T', 'C' or 'G'. The fragment ions detected in MS/MS were manually assigned.

RESULTS

The objectives of this study were to assess how alkyl-PTE lesions with different stereochemical configurations and various sizes of alkyl group compromise DNA replication, and to define the respective roles of the three SOS-induced polymerases and Ada protein in bypassing and repairing these lesions in *E. coli* cells.

We first synthesized 12-mer ODNs harboring a site-specifically inserted alkyl-PTE lesion (Scheme 2-1) and characterized these ODNs by ESI-MS and MS/MS analyses (Figures 2-7 to 2-10). The ESI-MS and MS/MS results support the site of the alkyl-PTE lesion and the identity of the alkyl group being conjugated with backbone phosphate. In addition, we established the configurations of the two diastereomers of Me-PTEs based on ^{31}P-NMR and HPLC analyses (Figures 2-4 to 2-5). Previously, Weinfeld *et al.* (13,25)

showed that, under the same solvent conditions, the ^{31}P signal for the S_p diastereomer of Me-PTE appeared at a lower field than that for the R_p counterpart. We first determined the stereochemistry of the Me-PTE-containing T(Me)T based on ^{31}P-NMR analyses, where the earlier and later eluting T(Me)T-1 and T(Me)T-2 were found to contain the S_p and R_p diastereomers of the Me-PTE, respectively. A chemical shift difference of 0.16 ppm for the two diastereomers was nearly identical to what was previously reported (26). We next digested the HPLC-purified Me-PTE-containing 12-mer ODNs with a cocktail of four enzymes, which release Me-PTE as T(Me)T. By comparing the HPLC elution properties of the T(Me)T liberated from the 12-mer ODNs with those of the standards, we demonstrated that the earlier- and later-eluting 12-mer ODNs contained the S_p- and R_p-Me-PTEs, respectively. Based on the observations that the S_p- and R_p-Me-PTE lesions in a dimer and a 12-mer ODN exhibit the same elution orders on reversed-phase HPLC columns, we assign the stereochemistry of the purified Et-, *n*Pr- and *n*Bu-PTEs based on their relative orders of elution from the HPLC column, i.e., the S_p diastereomer elutes earlier than the R_p diastereomer.

We next ligated these lesion-containing ODNs into single-stranded M13mp7 plasmid and measured the bypass efficiencies and mutation frequencies of the alkyl-PTE lesions using a modified version of the competitive replication and adduct bypass assay (Figures 2-11 to 2-16) (15,16). After PCR amplification and restriction digestion, the released ODNs

were analyzed by LC-MS/MS and native PAGE to identify the replication products, as described elsewhere (24). As shown in Figure 2-11, we were able to radiolabel the 5′ terminus of either the original lesion-situated strand, i.e. d(p*GGCMNGCTAT), or its complementary strand, i.e. d(p*AATTATAGCY), by switching the orders of *Bbs*I and *Mlu*CI digestions. Results from native PAGE analyses of the resulting radiolabeled fragments revealed that the S_p diastereomer of Me-PTE could give rise to TT→GT and TT→GC mutations at the flanking TT dinucleotide sites, whereas no mutagenic products were detectable for any other alkyl-PTE lesions. In this regard, when *Bbs*I was added first, we could resolve [5′-^{32}P]-labeled d(p*GGCTTGCTAT) (non-mutagenic product, 10 mer TT) from the products with TT→GT or TT→GC mutation, i.e. d(p*GGCGTGCTAT) (10 mer GT) and d(p*GGCGCGCTAT) (10 mer GC). Nonetheless, 10 mer GT and 10 mer GC could not be distinguished by native PAGE analysis (Figure 2-11b). To resolve these two mutagenic products, we first added *Mlu*CI to radiolabel the opposite strand, which enabled us to separate the product with TT→GT mutation, i.e. d(p*AATTATAGCA) (10 mer A), from that with TT→GC mutation, namely d(p*AATTATAGCG) (10 mer G), as shown in Figure 2-11c and Figures 2-12 to 2-14.

We further confirmed the identities of the aforementioned restriction digestion products by LC-MS/MS analysis. We monitored the $[M-3H]^{3-}$ ions of d(GGCMNGCTAT) and d(AATTATAGCY), where ‘M’ and ‘N’ designate the nucleotides inserted at the initial

TT dinucleotide flanking the alkyl-PTE site, and 'Y' represents the opposing nucleotide of 'N' in the complementary strand, respectively (Figures 2-15, 2-16). The LC-MS/MS data showed that only S_p-Me-PTE could induce TT→GT and TT→GC mutations, which are in keeping with the results obtained from the aforementioned PAGE analyses.

Having identified the replication products, we determined the bypass efficiencies of the alkyl-PTE lesions by measuring the ratio of combined intensities of all 10-mer bands arising from the lesion-containing genome over that of the 13-mer band emanating from the lesion-free competitor genome (see Materials and Methods). The results showed that the S_p diastereomers of all four alkyl-PTEs do not appreciably impede DNA replication in *E. coli* cells, whereas the R_p diastereomers elicit moderate blockage effects on DNA replication (Figure 2-17a).

We also examined the roles of the three SOS-induced polymerases (i.e. Pol II, Pol IV and Pol V) in bypassing the alkyl-PTE lesions by conducting the replication experiments using the isogenic *E. coli* strains that are deficient in these polymerases. The results showed that the bypass of the alkyl-PTE lesions does not require any of the three TLS polymerases (Figure 2-17a). Moreover, our results showed that the bypass efficiencies of these lesions were not modulated by SOS induction, either in wild-type AB1157 cells or the isogenic cells with all three SOS-induced DNA polymerases being knocked out (Figure 2-17a).

We further determined the mutation frequencies of the Me-PTE in wild-type and polymerase-deficient AB1157 *E. coli* strains (Figure 2-17b). It turned out that the S_p-Me-PTE is moderately mutagenic, resulting in TT→GT and TT→GC mutations. Additionally, deletion of any of the three SOS-induced polymerases alone did not confer any significant alterations in mutation frequencies, though simultaneous removal of the three polymerases in both SOS-induced and uninduced cells led to a slight, albeit statistically significant decrease in TT→GT mutation (Figure 2-17b).

The unique mutagenic property of S_p-Me-PTE parallels the previous finding that Ada protein removes the methyl group exclusively from the S_p-Me-PTE at an efficiency that is much higher than the corresponding removal of larger alkyl groups from alkyl-PTE lesions (13). Thus, we next asked whether Ada protein assumes any role in modulating the cytotoxic and mutagenic properties of the two diastereomers of Me- and Et-PTEs by conducting replication experiments with the use of isogenic *E. coli* cells that are proficient or deficient in Ada protein (Figure 2-18). Our results showed that the deletion of Ada led to a significant decline in bypass efficiency for S_p-Me-PTE, which is accompanied with the abrogation of both TT→GT and TT→GC mutations for the lesion, suggesting the involvement of Ada protein in repairing the S_p-Me-PTE and inducing mutations for this lesion. Similar as what we found for AB1157 cells, replication across S_p-Me-PTE and the two diastereomers of Et-PTE lesions are error-free in Ada-deficient C217 cells, and

deletion of Ada protein did not appreciably alter the bypass efficiencies of these three lesions (Figure 2-18).

DISCUSSION

We systematically investigated the cytotoxic and mutagenic properties of alkyl-PTE lesions in *E. coli* cells and our results led to several important conclusions. First, we found that the S_p and R_p diastereomers of the alkyl-PTE lesions exerted distinct effects on DNA replication in *E. coli* cells. The lesions in the S_p configuration displayed higher replication bypass efficiencies than the corresponding unmodified DNA, whereas those in the R_p configuration suppress DNA replication in AB1157 cells. The replication bypass efficiencies for the S_p-alkyl-PTEs, except for the methyl adduct which is removed by Ada protein, decrease with the size of alkyl group, which might be attributed to the elevated steric hindrance imposed by larger alkyl groups. In duplex DNA, the alkyl group in the S_p diastereomer is known to point perpendicularly out from the DNA double helix, whereas that of the R_p diastereomer projects into the major groove (6). Hence, the alkyl groups in the S_p configuration may assume some favorable interaction with DNA polymerases to promote translesion synthesis, where steric hindrance imposed by larger alkyl groups may diminish this enhancement effect. However, the R_p-alkyl-PTE lesions display an opposite trend, with *n*Pr- and *n*Bu-PTEs exhibiting higher bypass efficiencies than Me- and Et-PTEs. This result supports that the difference in stereochemical configurations of the alkyl group

conjugated with the phosphate backbone can modulate the replicative bypass of these lesions. By employing a DNA template containing an equimolar mixture of S_p and R_p diastereomers of Et-PTE at a defined site, Tsujikawa *et al.*(14) found that the lesions could inhibit partially primer extension catalyzed by T4 DNA polymerase, but could be bypassed by *E. coli* DNA polymerase I. The results from that study also suggested that one of the two diastereomers was a major block to DNA replication (14), though the identity of the diastereomer could not be ascertained owing to the lack of stereochemically pure DNA substrates. It is worth noting that, due to the differential accessibilities of the two non-carbon-bound oxygen atoms on backbone phosphate to DNA alkylating agents, the S_p diastereomer is produced at a higher yield than the R_p counterpart (27). Hence, the assessment about overall biological consequences of alkyl-PTEs should also take into account the relative frequencies of occurrence of these lesions.

Second, our results demonstrated that the bypass efficiencies of the alkyl-PTE lesions were not affected by deletion of any of the three SOS-induced TLS polymerases, alone or all three in combination. Additionally, the mutation frequencies induced by S_p-Me-PTE were not influenced by the deletion of any of the three TLS polymerases alone, though the absence of all three polymerases resulted in a slight, yet statistically significant decrease in the frequency of TT→GT mutation.

Third, Ada protein, which could repair the S_p-Me-PTE, is required for the TT→GT and TT→GC mutations induced by this lesion. The removal of Me-PTE by Ada protein takes time. In this vein, we also measured the bypass efficiencies and mutation frequencies of the two diastereomers of Me-PTEs at 1, 3 and 6 h following transfection, and we observed that at 1 h, the bypass efficiency of S_p-Me-PTE at 1 h is lower than that at 3 and 6 h (Figure 2-19). Ada protein plays a crucial role in bacterial tolerance toward exposure to alkylating agents in a process known as adaptive response, where Ada acts both as a DNA repair protein and a transcription activator (4,28,29). In that process, the 39 kDa Ada protein undergoes proteolysis at the sole Lys-Gln linkage to yield a 19 kDa C-terminal fragment (C-Ada) and a 20 kDa N-terminal fragment (N-Ada); while Cys321 in C-Ada is a methyltransferase for the repair of O^6-methylguanine and O^4-methylthymine (30), Cys38 in N-Ada could remove the methyl group specifically from the S_p-Me-PTE (31). Moreover, He et al. (31) illustrated, from detailed solution (NMR) and crystal (X-ray) structure studies, that this latter methyl group transfer stimulated the binding of N-Ada with DNA through an electrostatic switch mechanism. The heightened DNA affinity led to promoter binding and transcriptional activation of several genes that are important in repairing alkylated DNA lesions including Ada, Aid, AlkA, and AlkB (6). Thus, Ada protein is instrumental in protecting the genomic integrity of bacterial cells by promoting the repair of alkylated DNA lesions. Here we made an unexpected finding that, while the Ada-stimulated repair of the

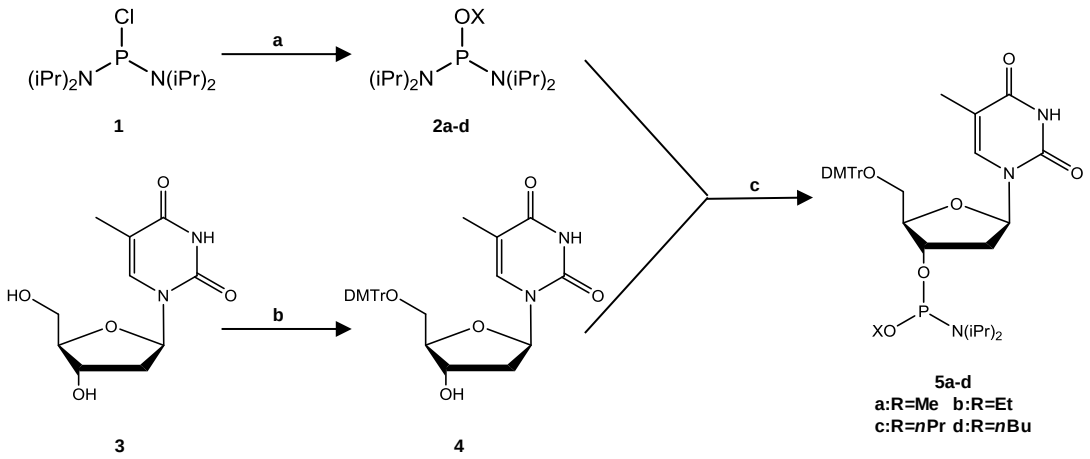
S_p-Me-PTE results in enhanced replicative bypass, it comes at an expense by inducing TT→GT and TT→GC mutations at the flanking nucleoside sites. Further studies are warranted for understanding how the selective interaction between Ada and S_p-Me-PTE would interfere with Watson-Crick base pairing during DNA replication and trigger specific mutations at the lesion site.

A recent genome-wide mutation analysis of 32 *E. coli* strains treated with *N*-methyl-*N*'-nitrosoguanidine showed that, among the 4099 identified mutations, 96.6%, 2.17% and 0.46% were G→A, T→C and T→G substitutions, respectively (32). While G→A and T→C mutations are thought to arise from methylation at the O^6 of guanine and O^4 of thymine (21,33), respectively, our results suggest that T→G mutation may be attributed, in part, to the Me-PTE formed on the 3' side of thymidine. In this context, it is worth noting that the formation frequency of the Me-PTE lesions was found to be affected by flanking sequences (34), and flanking sequences may also modulate the mutagenic properties of alkyl-PTE lesions. Thus, it will be important to investigate how the cytotoxic and mutagenic properties of the alkyl-PTE lesions are influenced by flanking sequences in the future.

In summary, our study demonstrated that the two diastereomers of alkyl-PTEs exert distinct effects on DNA replication; the S_p-alkyl-PTEs did not appreciably impede DNA replication, whereas the R_p counterparts could moderately inhibit this process. We also revealed that the mutagenic properties of the S_p-Me-PTE could be substantially modulated

by Ada protein, but not by any of the three SOS-induced polymerases. Hence, our systematic shuttle-vector study on alkyl-PTE lesions offered new insights about how this group of DNA alkylation products are recognized by the *E. coli* DNA replication machinery. It will be important to examine how the alkyl-PTE lesions compromise DNA replication in mammalian cells in the future.

FIGURES AND SCHEMES



Scheme 2-1. Syntheses of alkylphosphoramidite building blocks. Reagents and conditions: (a) for methyl and ethyl: ROH, triethylamine, diethyl ether, ice bath, 14 h; for *n*-propyl and *n*-butyl: ROH, triethylamine, THF, 0°C,1 h; (b) DMTr-Cl, DMAP, pyridine, room temperature, 10 h; (c) for methyl and ethyl: diisopropylamine hydrotetrazolide, CH_2Cl_2, room temperature, 1 h; for *n*-propyl and *n*-butyl: 0.45 M tetrazole in acetonitrile, CH_2Cl_2, room temperature, 5 h.

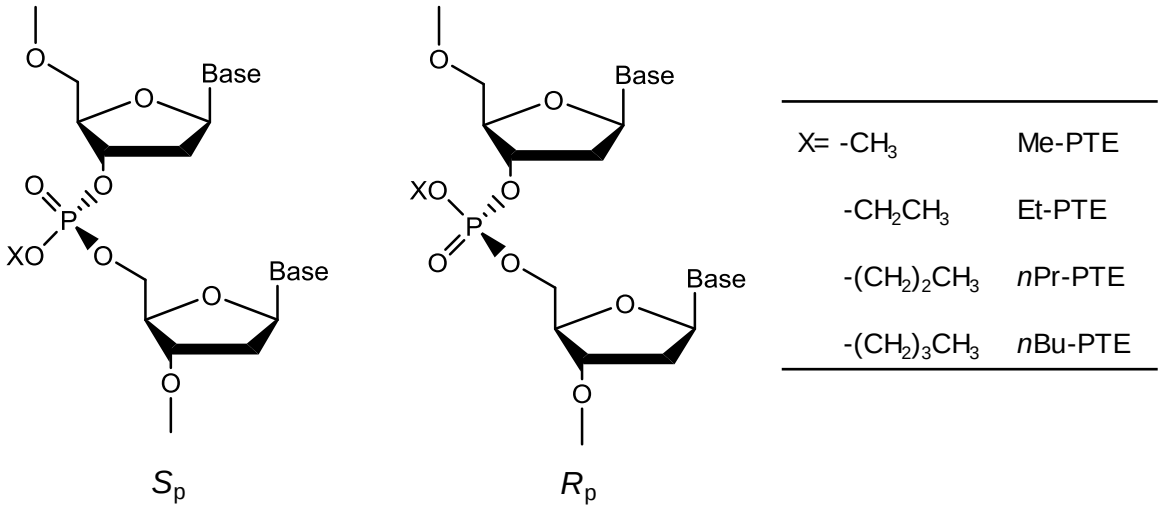


Figure 2-1. S_p and R_p diastereomers of alkyl phosphotriester residues in DNA.

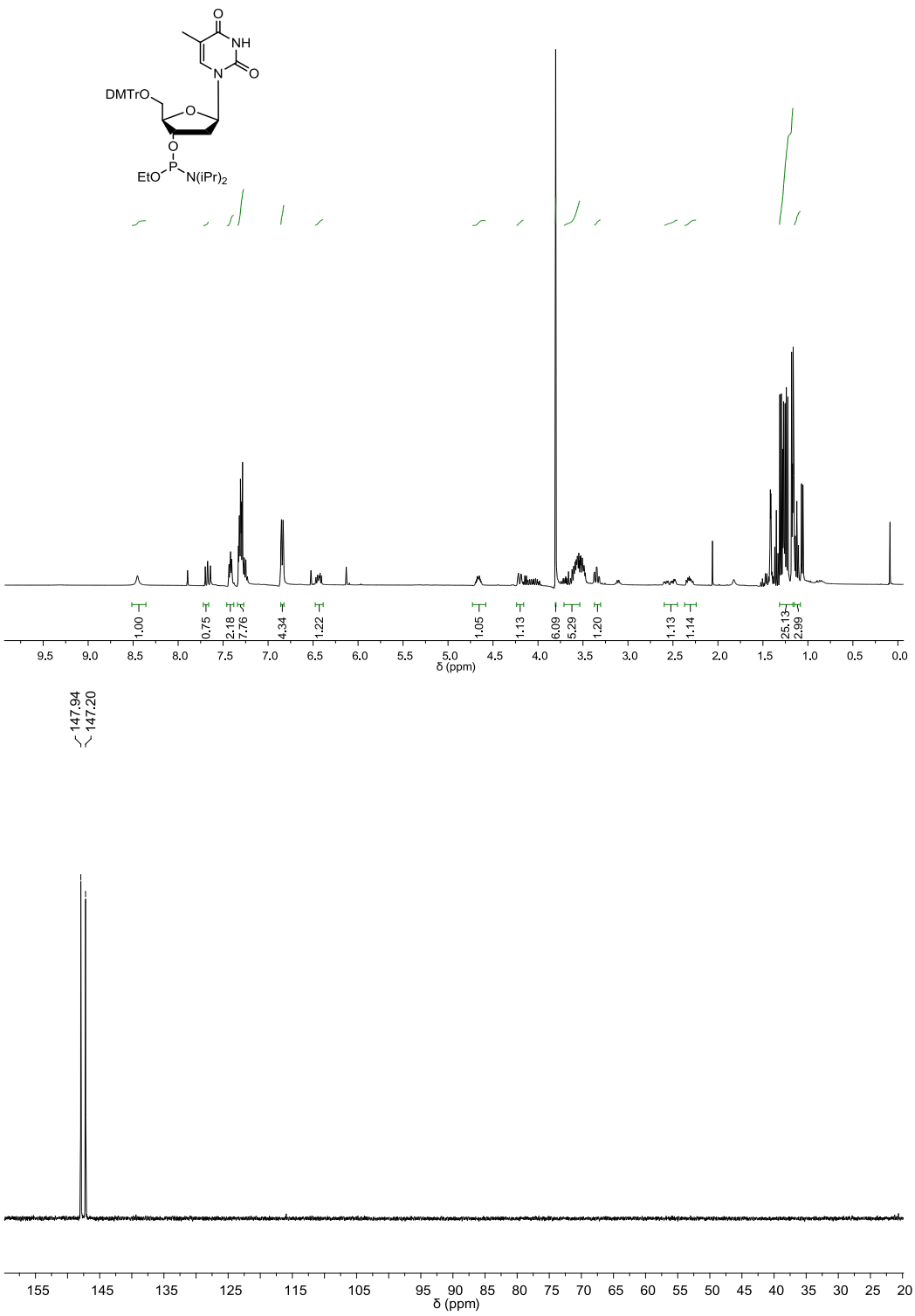


Figure 2-2. The ^{1}H NMR (400 MHz, $CDCl_3$, 25°C) and ^{31}P NMR (80 MHz, $CDCl_3$, 25°C) spectra of dT-ethylphosphoramidite.

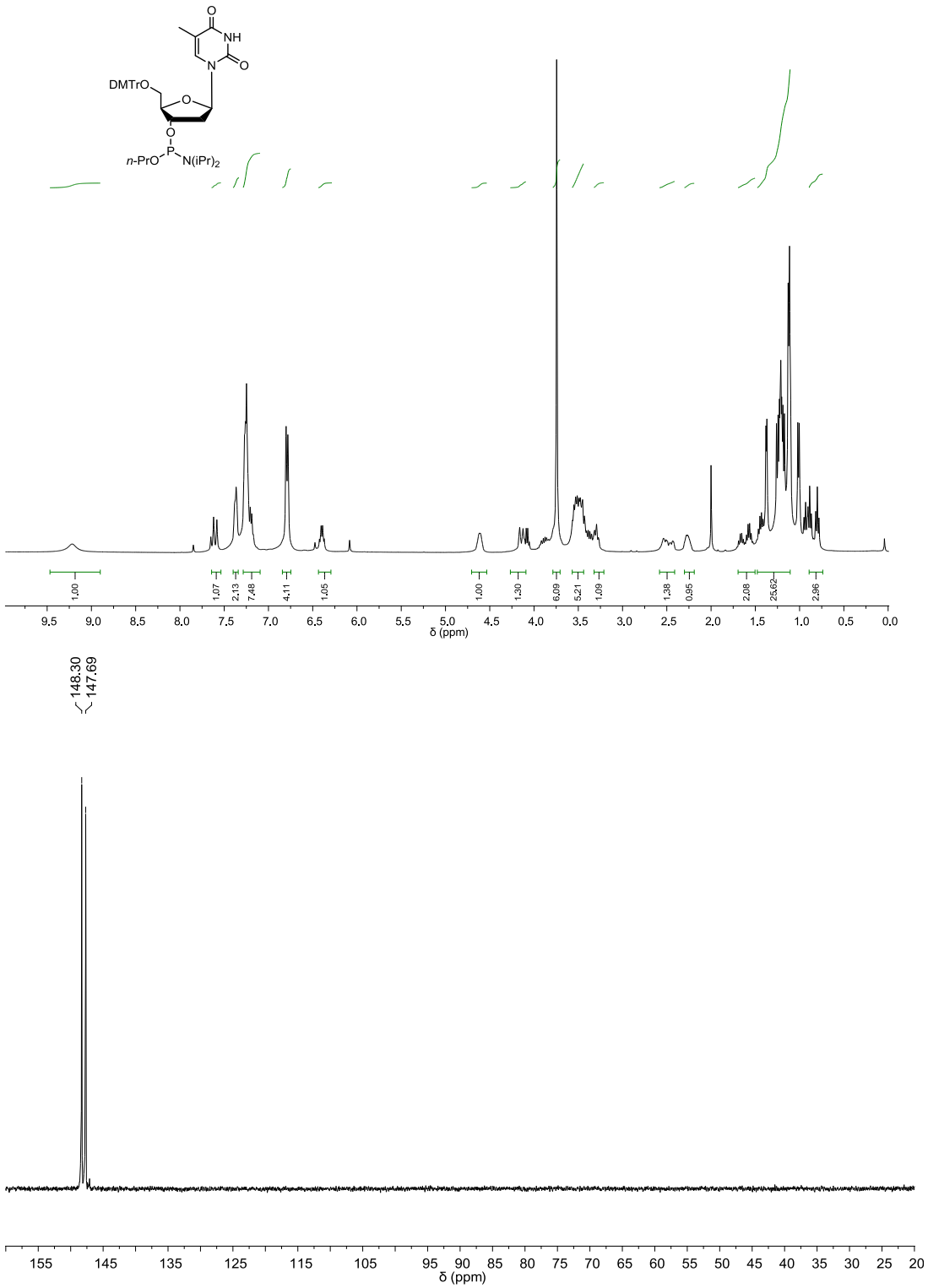


Figure 2-3. The ^{1}H NMR (400 MHz, $CDCl_3$, 25°C) and ^{31}P NMR (80 MHz, $CDCl_3$, 25°C) spectra of dT-*n*-prophosphoramidite

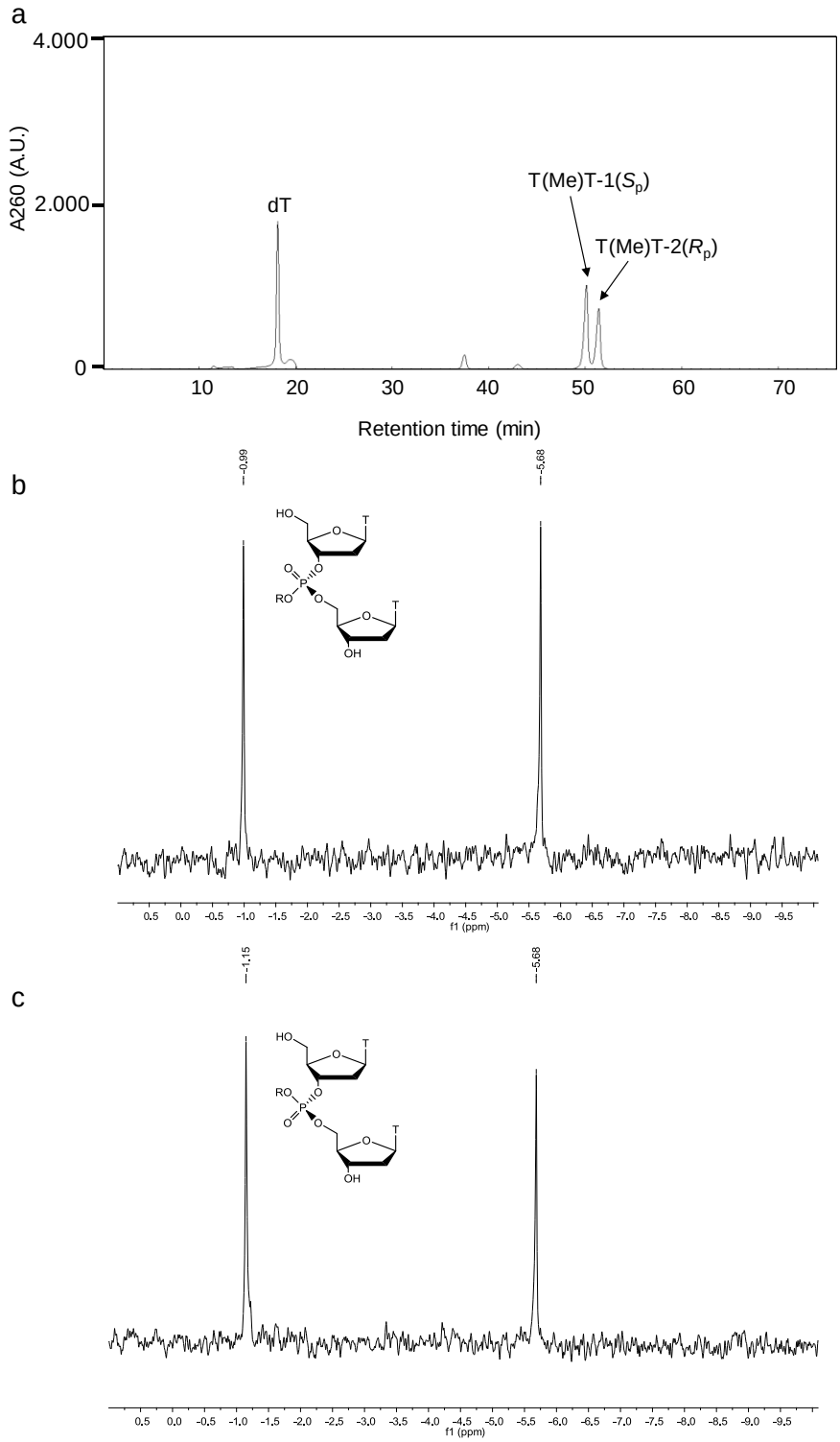


Figure 2-4. (a) The HPLC trace for the purification of T(Me)T-dimer and ^{31}P NMR spectrum of the S_p-T(Me)T-1 (b) and R_p-T(Me)T-2 (c) (80 MHz, MeOD, 25ºC).

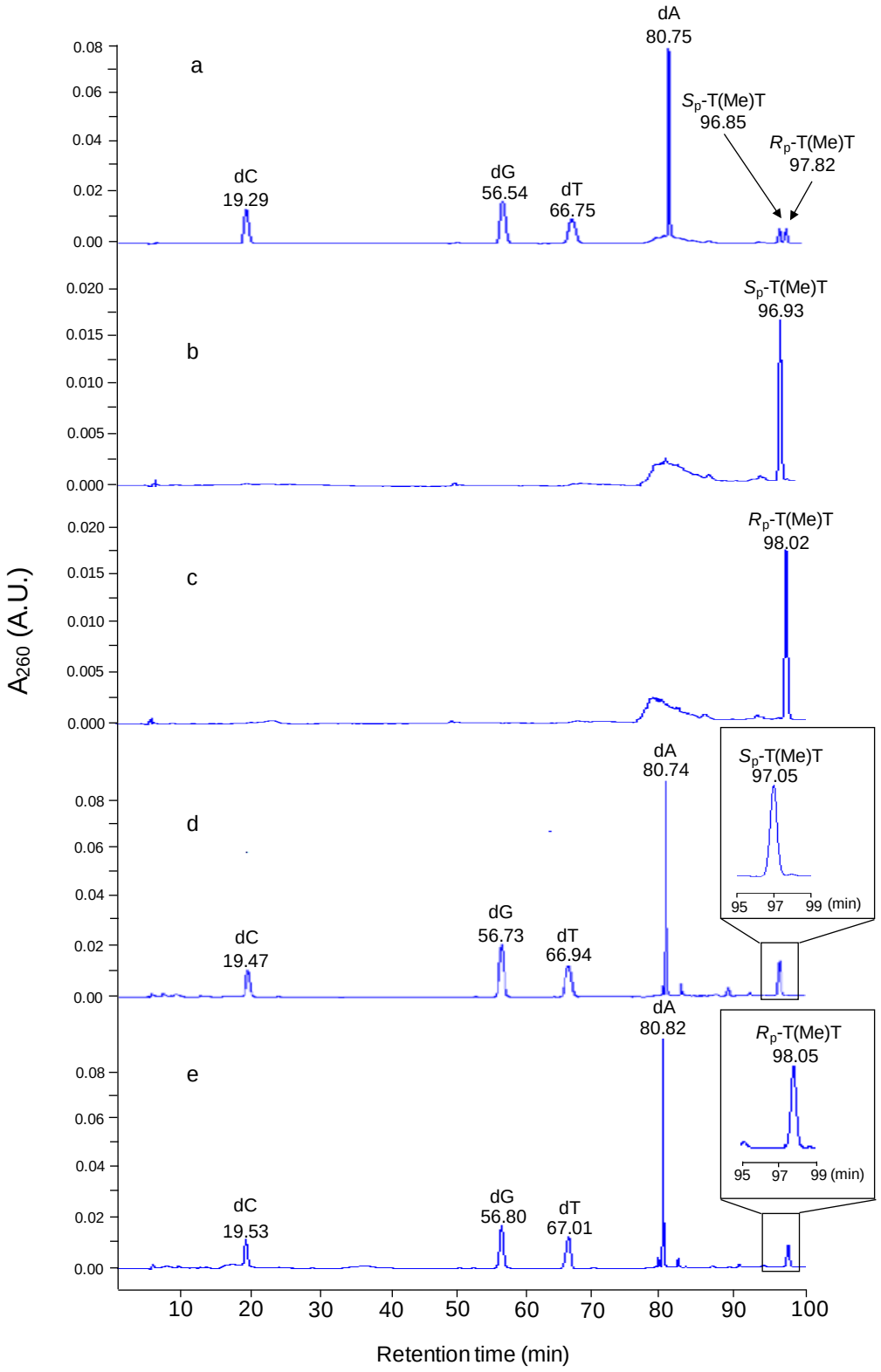

a
dA
80.75
S_p-T(Me)T
96.85
R_p-T(Me)T
97.82
dC
19.29
dG
56.54
dT
66.75
b
S_p-T(Me)T
96.93
c
R_p-T(Me)T
98.02
d
dA
80.74
S_p-T(Me)T
97.05
95 97 99 (min)
dC
19.47
dG
56.73
dT
66.94
e
dA
80.82
R_p-T(Me)T
98.05
95 97 99 (min)
dC
19.53
dG
56.80
dT
67.01
A_{260} (A.U.)
Retention time (min)

Figure 2-5. The HPLC traces for the separation of S_p-T(Me)T and R_p-T(Me)T in the digestion mixtures of synthesized 12mer site-specific lesion-containing ODN (5′-ATGGCT(Me)TGCTAT-3′). ‘dC’, ‘dG’, ‘dT’, ‘dA’ represent 2'-deoxycytidine, 2'-deoxyguanosine, thymidine, and 2'-deoxyadenosine, respectively. (a) The HPLC trace for the separation of the synthesized T(Me)T dimer together with canonical 2′-deoxynucleosides. (b) The HPLC trace for separation of the S_p-T(Me)T standard. (c) The HPLC trace for the separation of the R_p-T(Me)T standard. (d) The HPLC trace for the separation of the digestion mixture of 5′-ATGGCT(Me$_{Sp}$)TGCTAT-3′. (e) The HPLC trace for the separation of the digestion mixture of 5′-ATGGCT(Me$_{Rp}$)TGCTAT-3′.

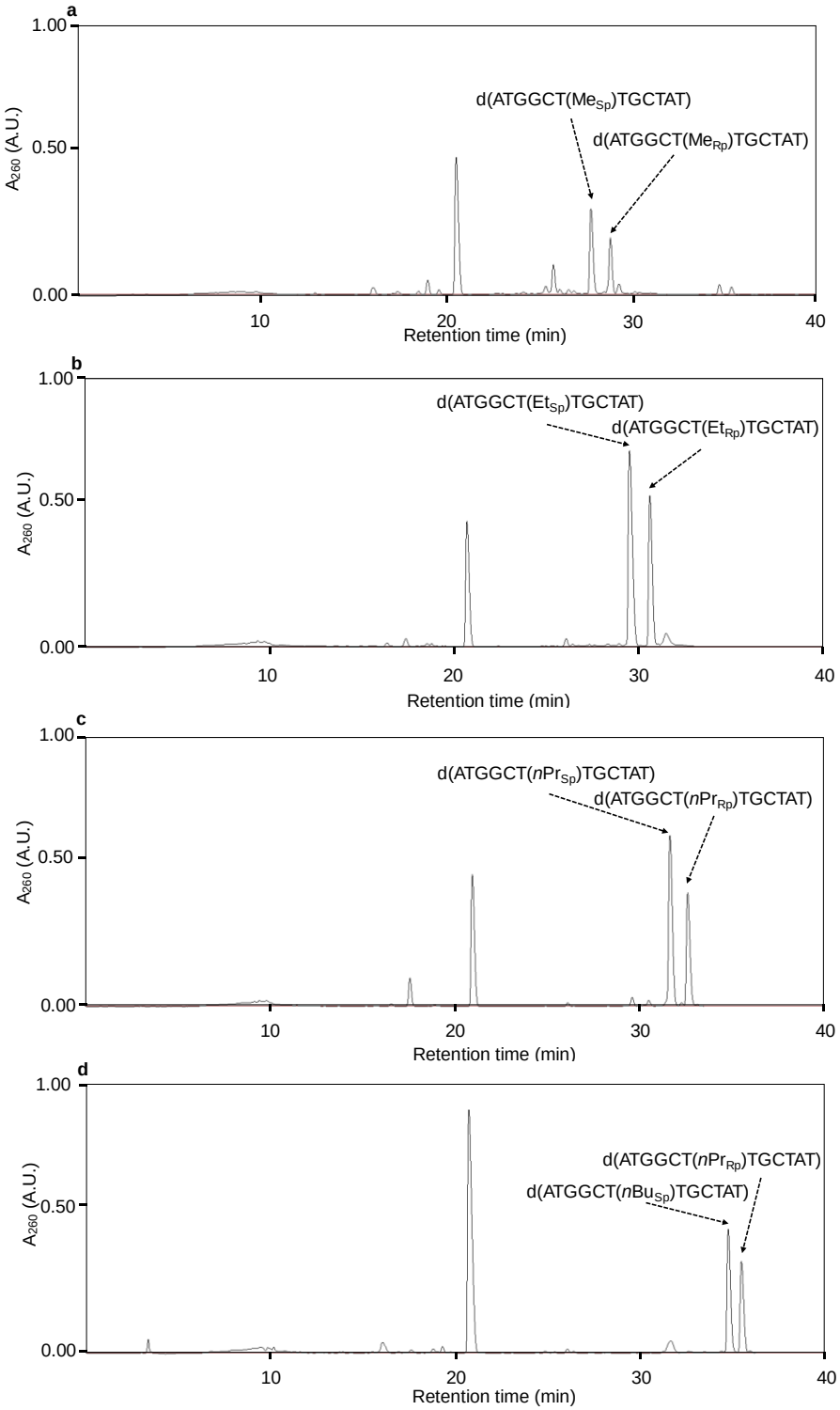


Figure 2-6. HPLC traces for the separations of the synthesized 12mer alkyl phosphotriester-bearing ODNs: (a) methyl; (b) ethyl; (c) *n*-propyl; (d) *n*-butyl.

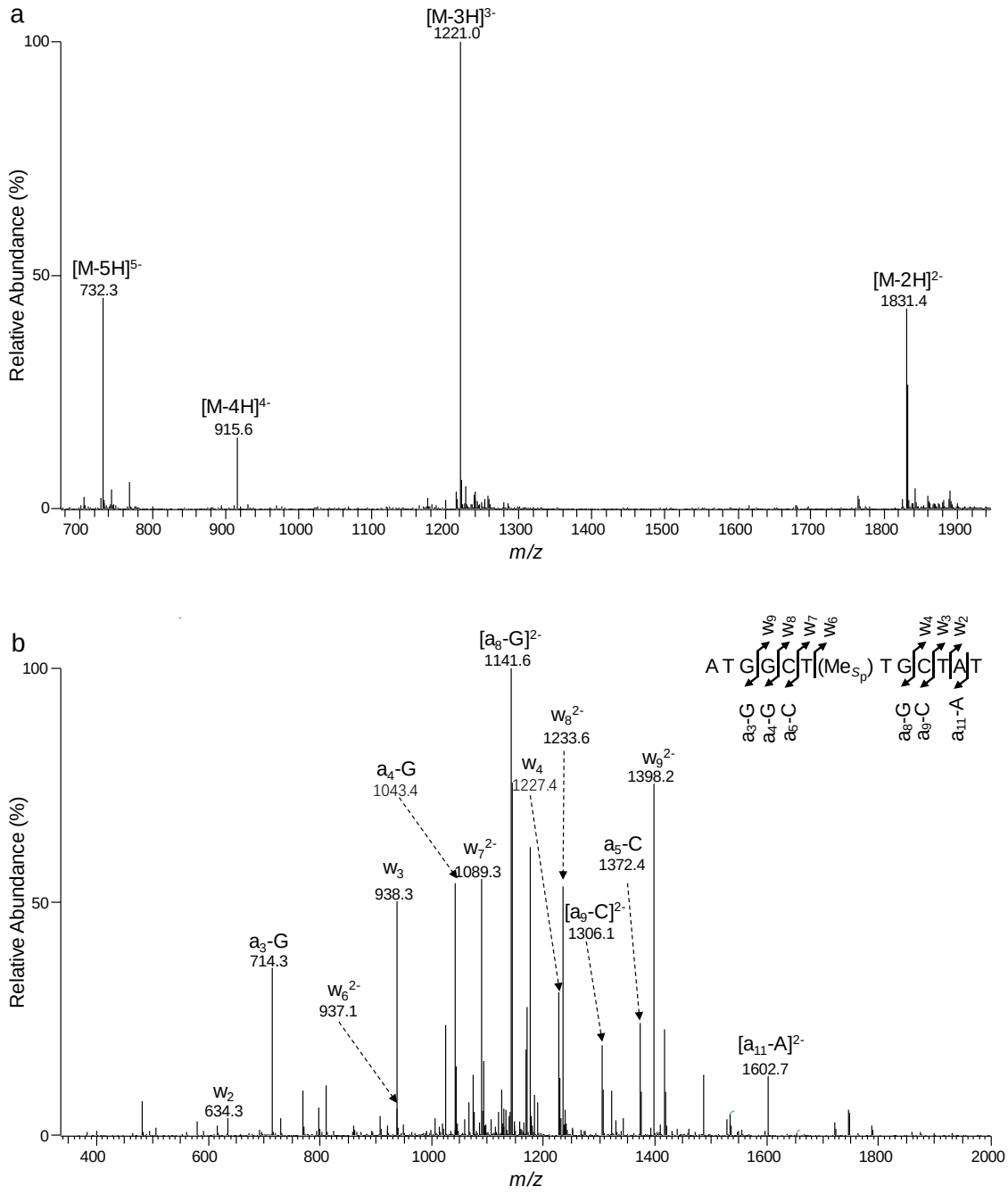


Figure 2-7. ESI-MS & MS/MS characterizations of d(ATGGCT(Me*Sp*)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m/z* 1221.0)

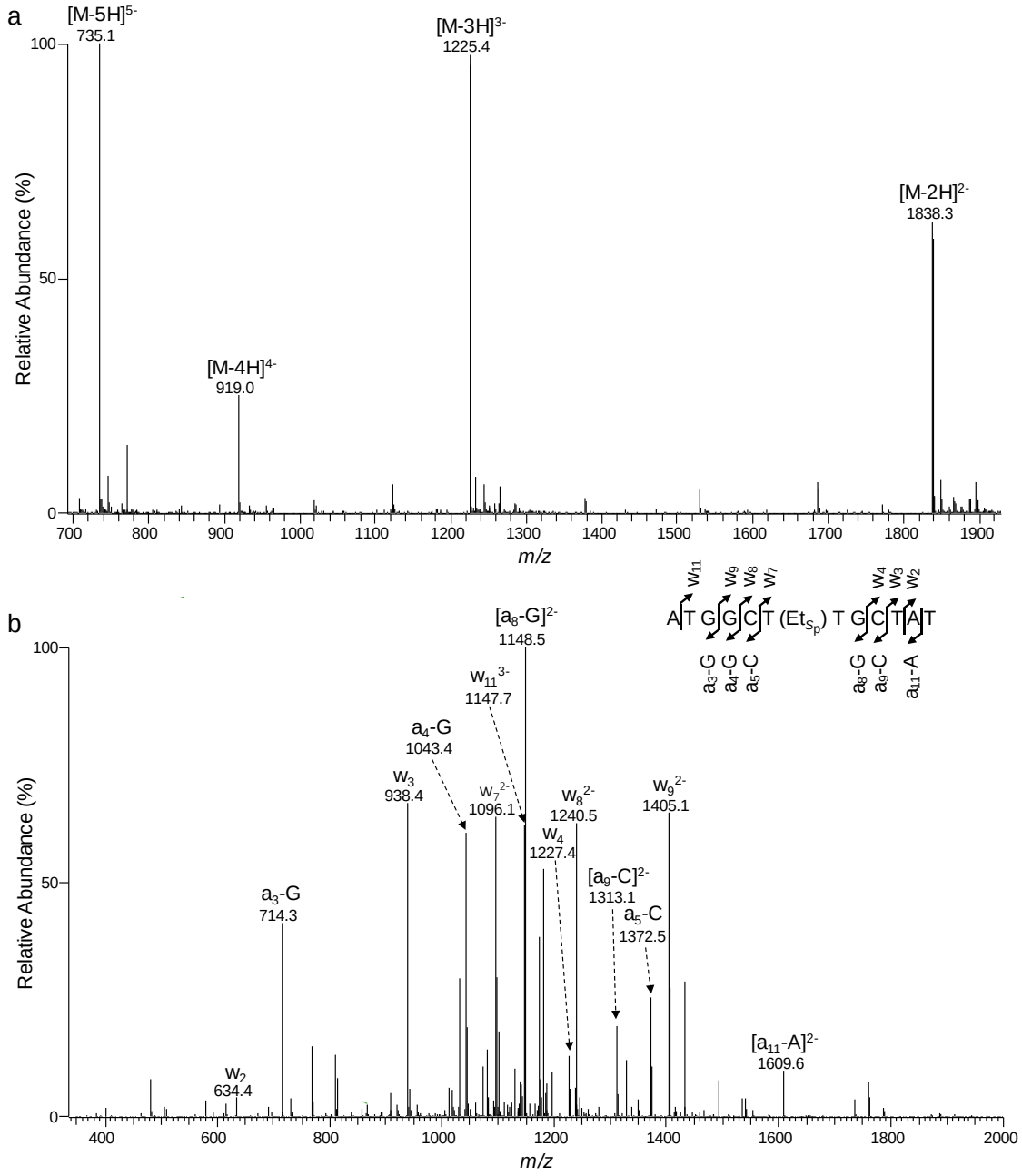


Figure 2-8. ESI-MS & MS/MS characterizations of d(ATGGCT(Et$_{Sp}$)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1225.4)

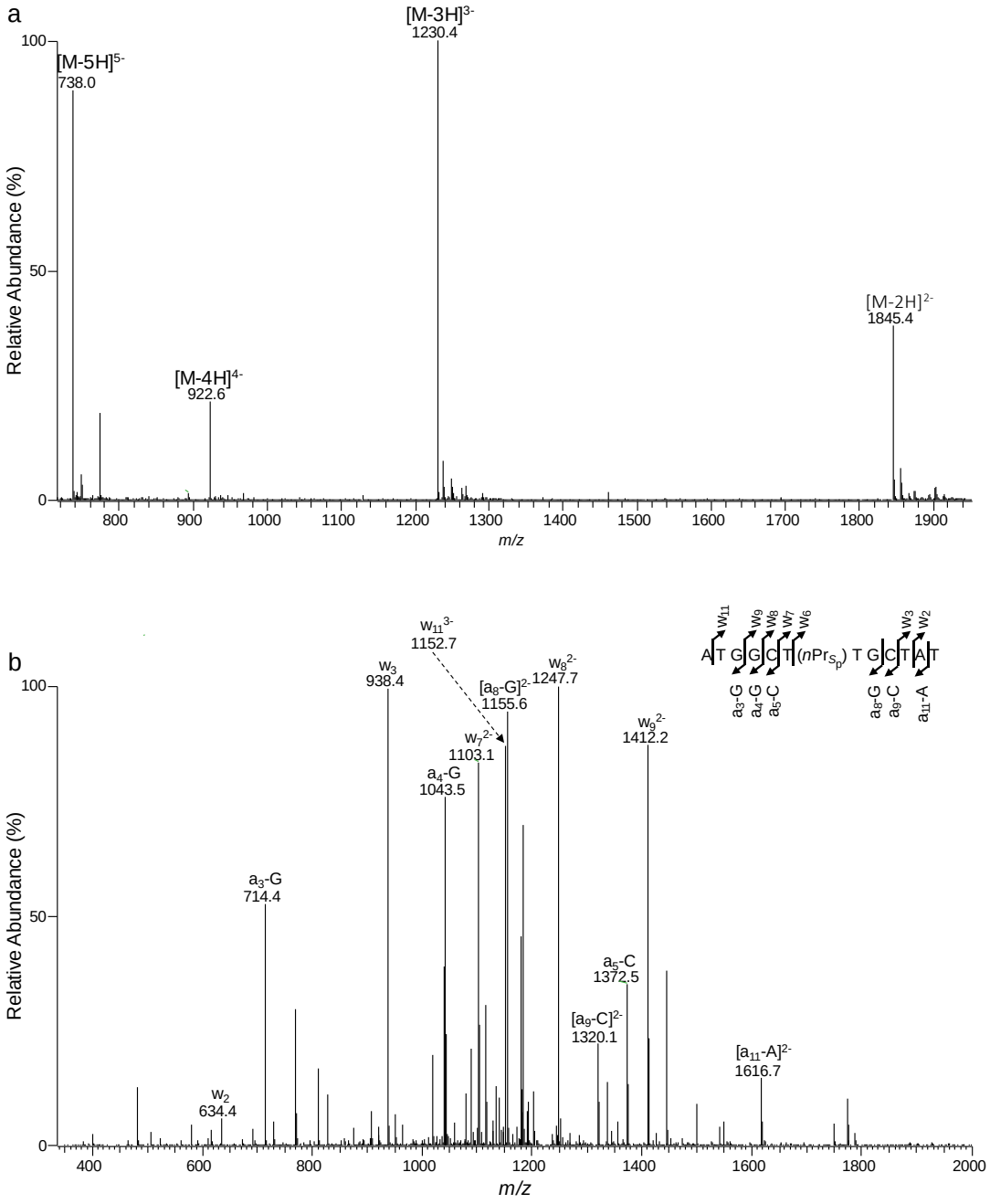


Figure 2-9. ESI-MS & MS/MS characterizations of d(ATGGCT(*n*Pr$_{Sp}$)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1230.4)

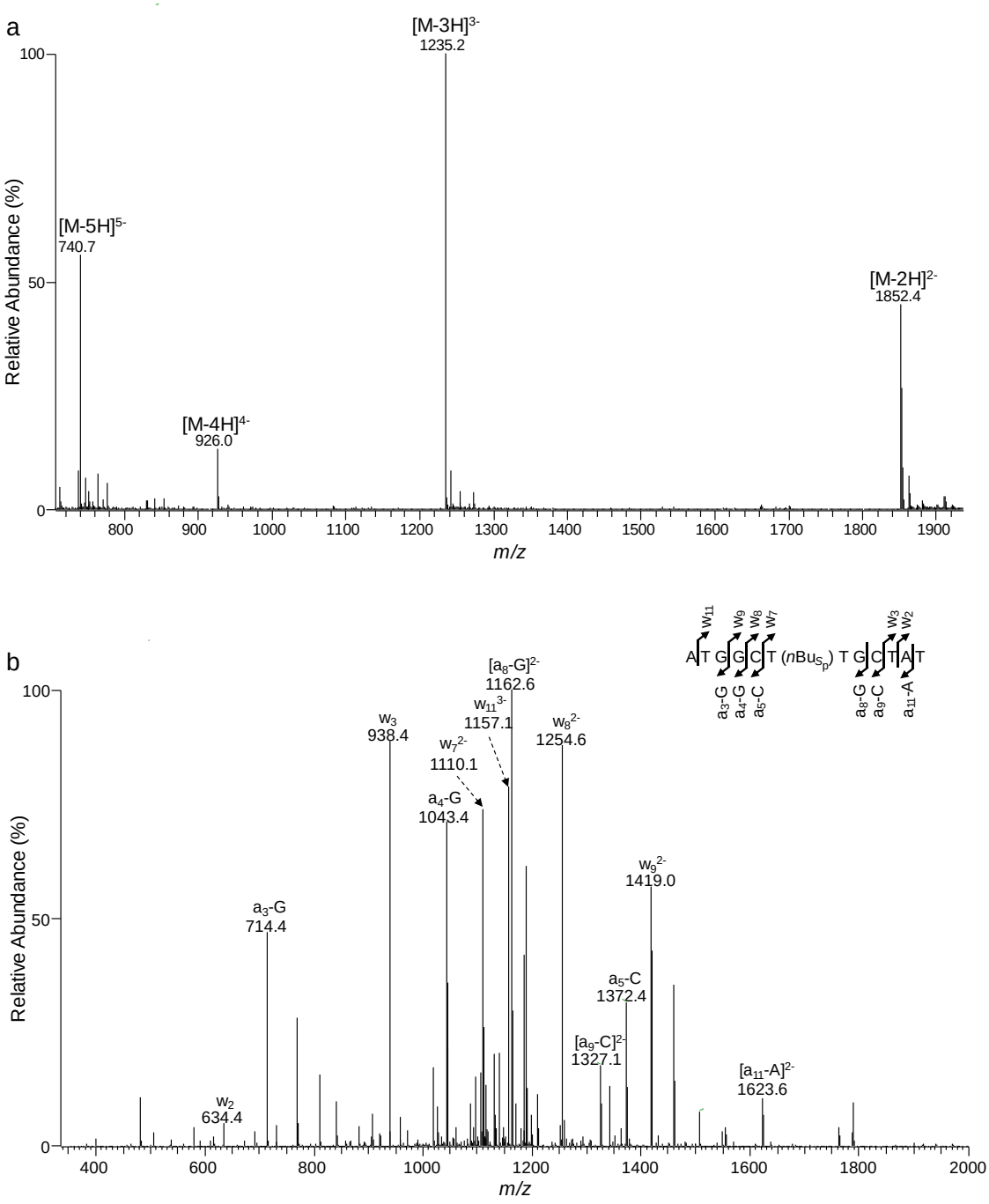


Figure 2-10. ESI-MS & MS/MS characterizations of d(ATGGCT(*n*Bu*S*p)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1235.2)

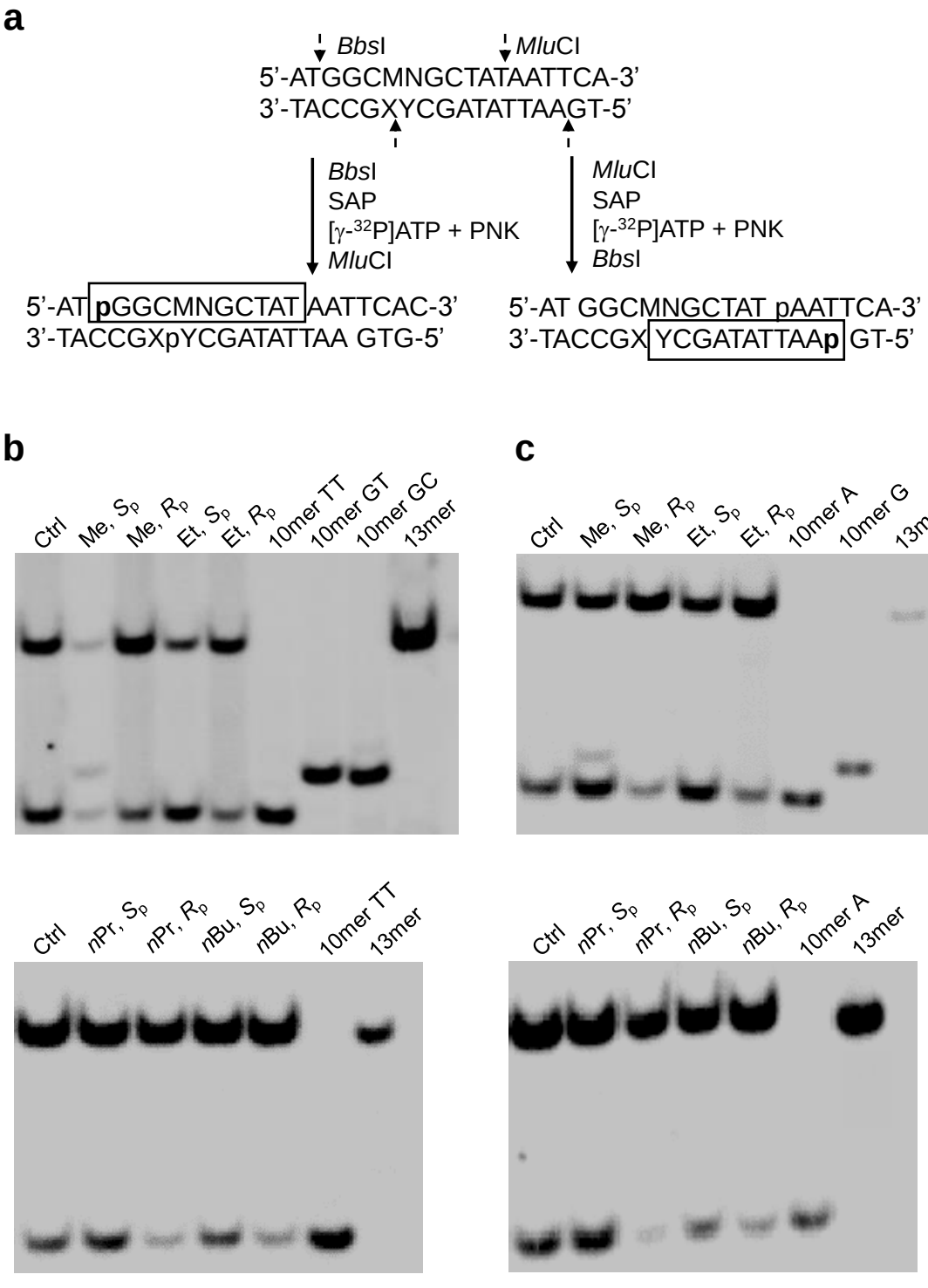

a
BbsI
MluCI
5'-ATGGCMNGCTATAATTCA-3'
3'-TACCGXYCGATATTAAGT-5'
BbsI
SAP
[γ-32P]ATP + PNK
MluCI
MluCI
SAP
[γ-32P]ATP + PNK
BbsI
5'-AT pGGCMNGCTAT AATTCAC-3'
3'-TACCGXpYCGATATTAA GTG-5'
5'-AT GGCMNGCTAT pAATTCA-3'
3'-TACCGX YCGATATTAAp GT-5'
b
Ctrl
Me, Sp
Me, Rp
Et, Sp
Et, Rp
10mer TT
10mer GT
10mer GC
13mer
c
Ctrl
Me, Sp
Me, Rp
Et, Sp
Et, Rp
10mer A
10mer G
13mer
Ctrl
nPr, Sp
nPr, Rp
nBu, Sp
nBu, Rp
10mer TT
13mer
Ctrl
nPr, Sp
nPr, Rp
nBu, Sp
nBu, Rp
10mer A
13mer

Figure 2-11. Restriction enzyme digestion and radiolabeling, followed by native PAGE (30%) analysis for quantifying the bypass efficiencies and mutation frequencies of alkyl phosphotriester lesions in wild-type AB1157 *E. coli* cells. (a) Selective labelling of the original lesion-containing strand and its complementary strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. 'MN' in the sequence denotes the site where the TT dinucleotide flanking the alkyl phosphotriester lesions were initially situated. (b) Gel image showing the 13 mer (competitor genome) and 10 mer (control or lesion-containing genome) digestion products formed from the original strand, where 10 mer TT, 10 mer GC and 10 mer GT designate [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with MN being TT, GC and GT, respectively. (c) Gel image showing 13 mer (competitor genome) and 10 mer (control or lesion-containing genome) digestion products formed from the opposite strand, where 10 mer A and 10 mer G designate the [5′-^{32}P]-labeled standard ODNs 5′-AATTATAGCY-3′, with Y being A and G, respectively.

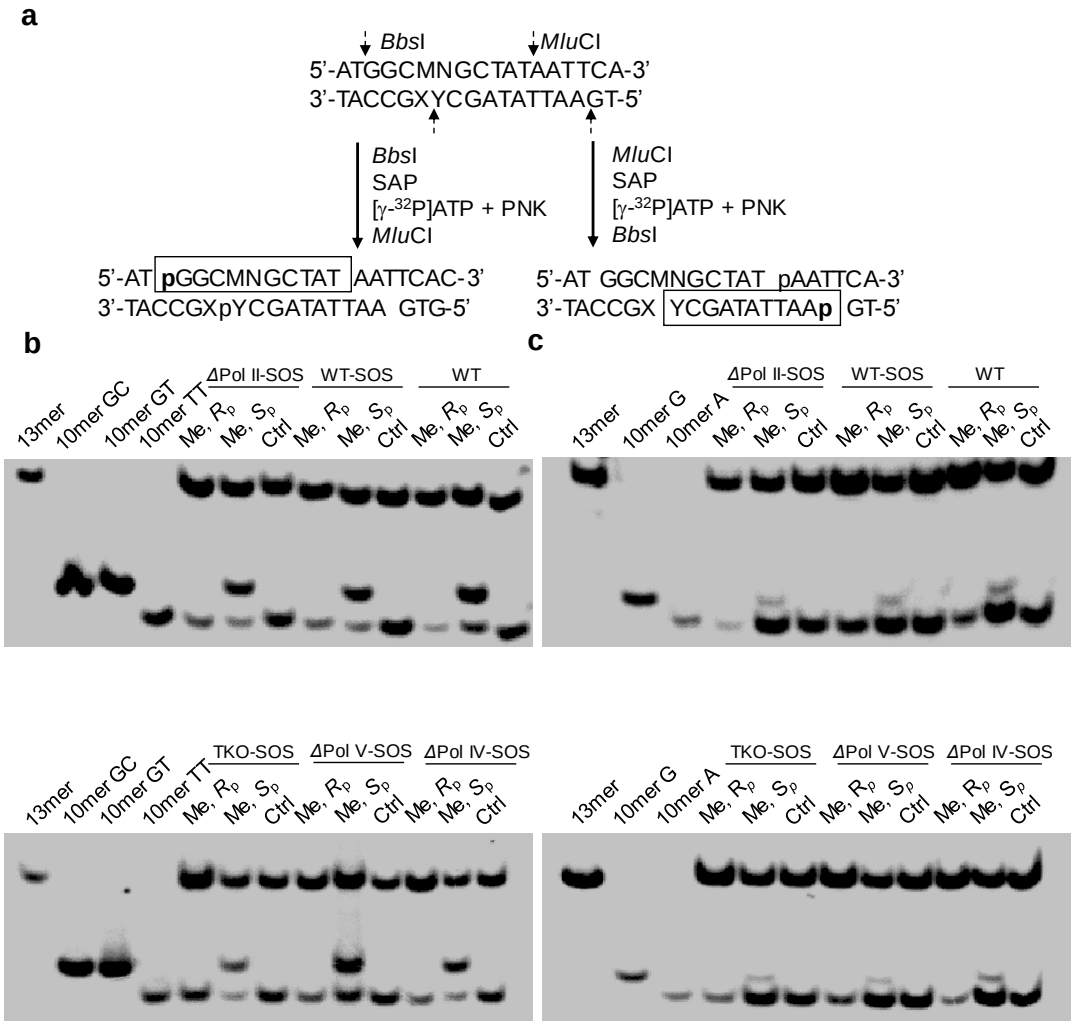


Figure 2-12. Native PAGE (30%) for monitoring the bypass efficiencies and mutation frequencies of S_p- and R_p-methyl phosphotriester in wild-type (WT), SOS-WT, SOS-*Δ*Pol II, IV, V and triple knockout (TKO) AB1157 *E. coli* cells. (a) Selective labelling of original lesion-containing strand and its complementary strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where 10mer TT, 10mer GC, and 10mer GT represent the [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with 'MN' being TT, GC and GT, respectively. (c) Gel images showing the 13-mer and 10-mer products released from the bottom-strand (opposite to the lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where 10mer A and 10mer G represent the [5′-^{32}P]-labeled standard ODNs 5′-AATTATAGCA-3′ and 5′-AATTATAGCG-3′, respectively.

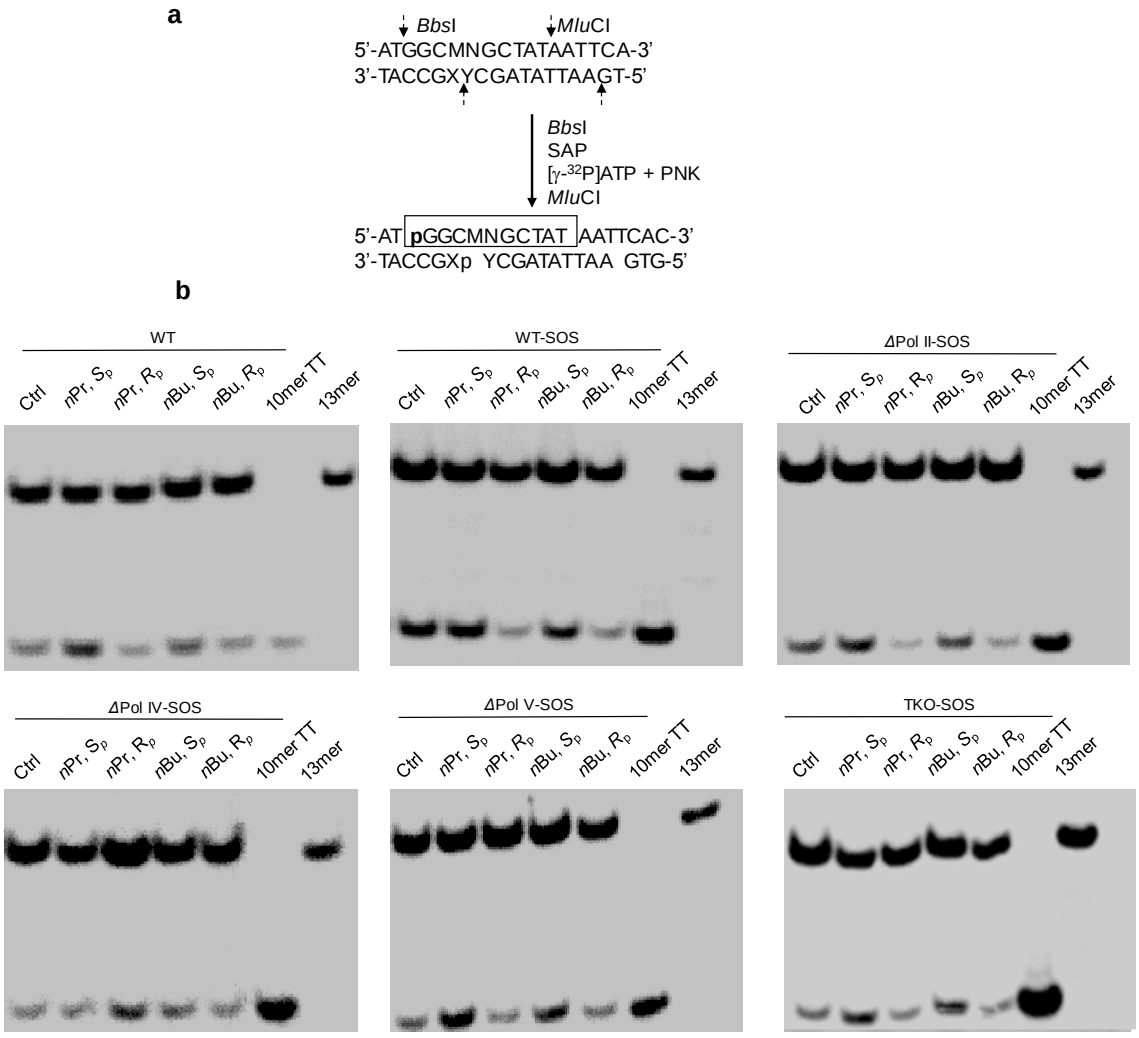


Figure 2-13. Native PAGE (30%) for monitoring the bypass efficiencies and mutation frequencies of *n*-propyl and *n*-butyl phosphotriester in wild-type AB1157 cells without SOS induction (WT), or wild-type AB1157 (SOS-WT) and isogenic cells that are deficient in Pol II, Pol IV, or PolV (*Δ*Pol II-SOS, *Δ*Pol IV-SOS, *Δ*Pol V-SOS) and all three polymerases (TKO-SOS) with SOS induction. (a) Selective labelling of original lesion-containing strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where 10mer TT represents the [5′-^{32}P]-labeled standard ODNs 5′-GGCTTGCTAT-3′.

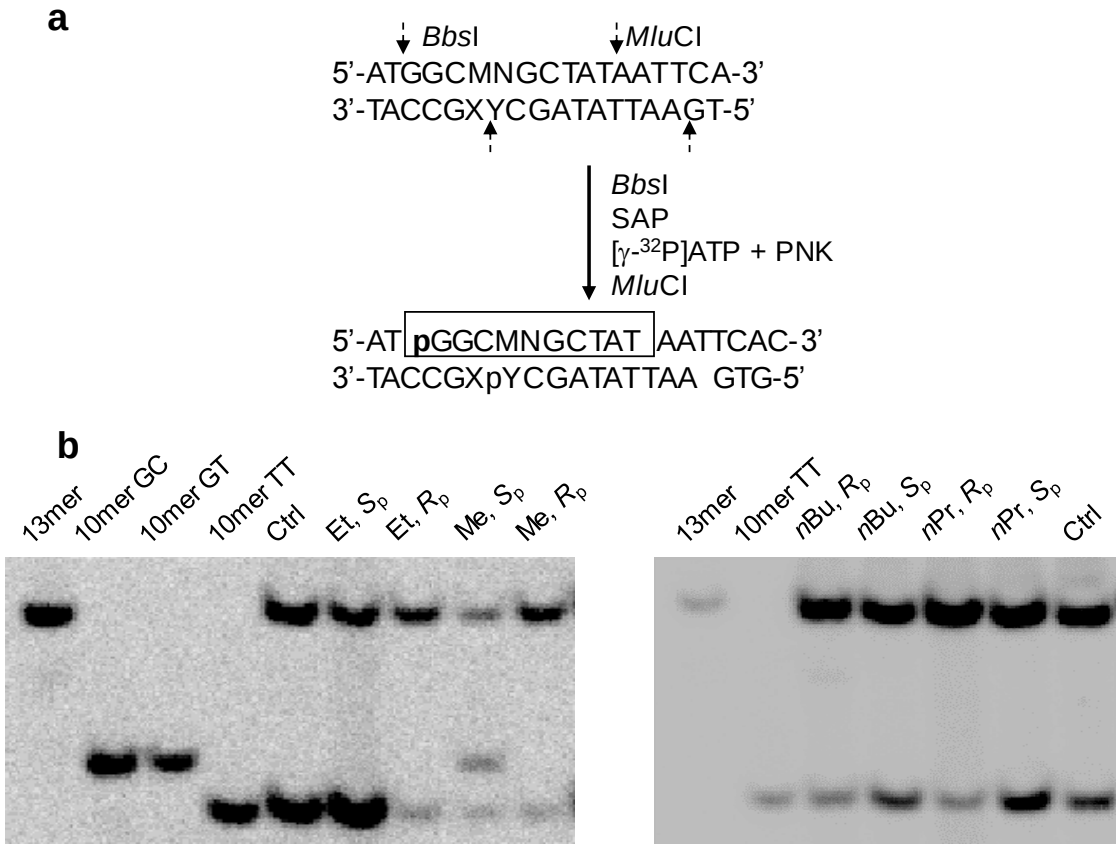


Figure 2-14. Native PAGE (30%) for monitoring the bypass efficiencies and mutation frequencies of alkylphosphotriester in triple knockout (TKO) AB1157 cells without SOS induction (WT). (a) Selective labelling of original lesion-containing strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where where 10mer TT, 10mer GC, and 10mer GT represent the [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with 'MN' being TT, GC and GT, respectively.

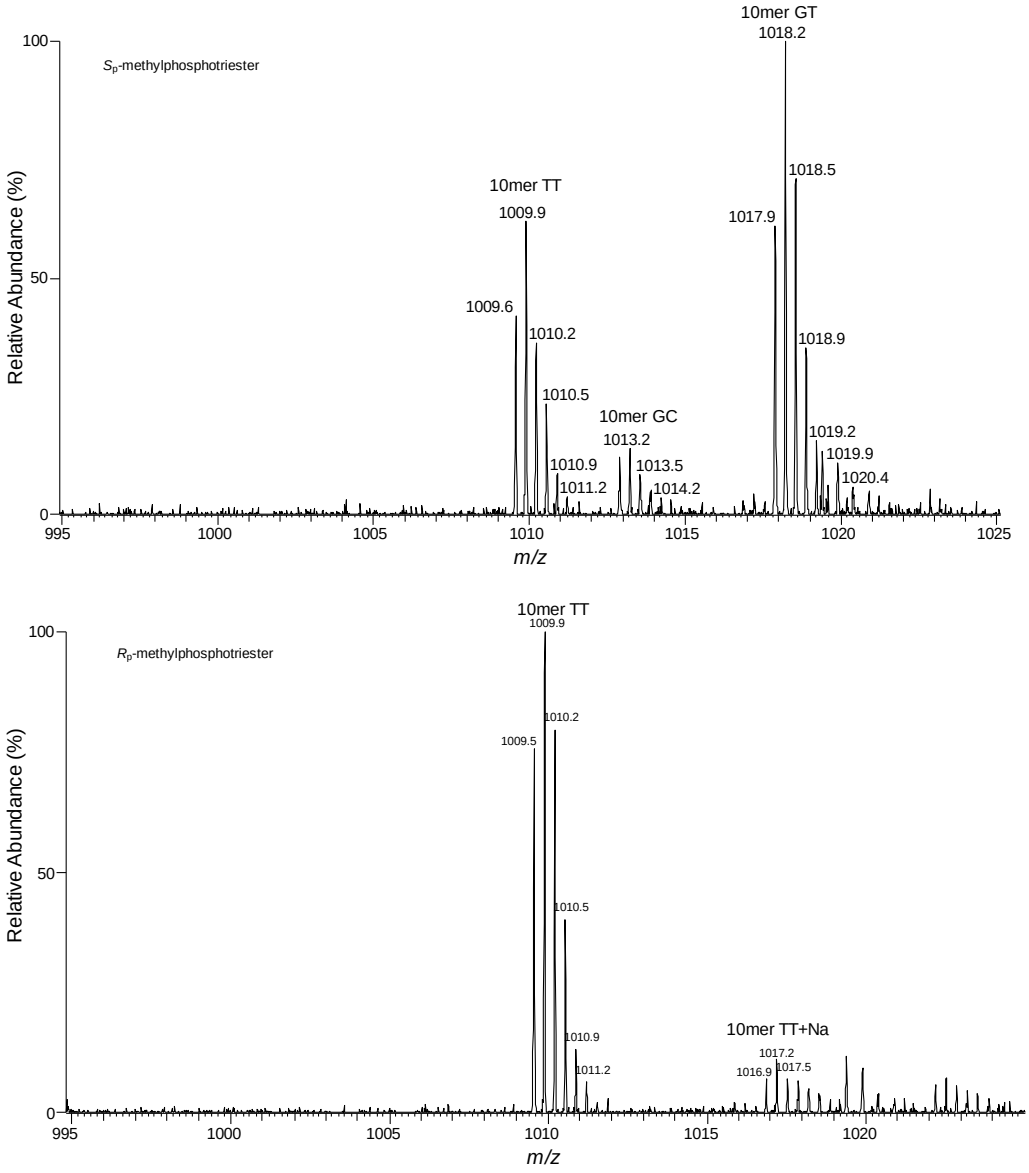


Figure 2-15. Higher-resolution "ultra-zoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of S_p- and R_p-methyl phosphotriester-bearing single-stranded M13 genomes in SOS-induced wild-type AB1157 cells. Displayed are the $[M-3H]^{3-}$ ions for the lesion-containing top strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for S_p-methyl phosphotriester are shown in Figure 2-16.

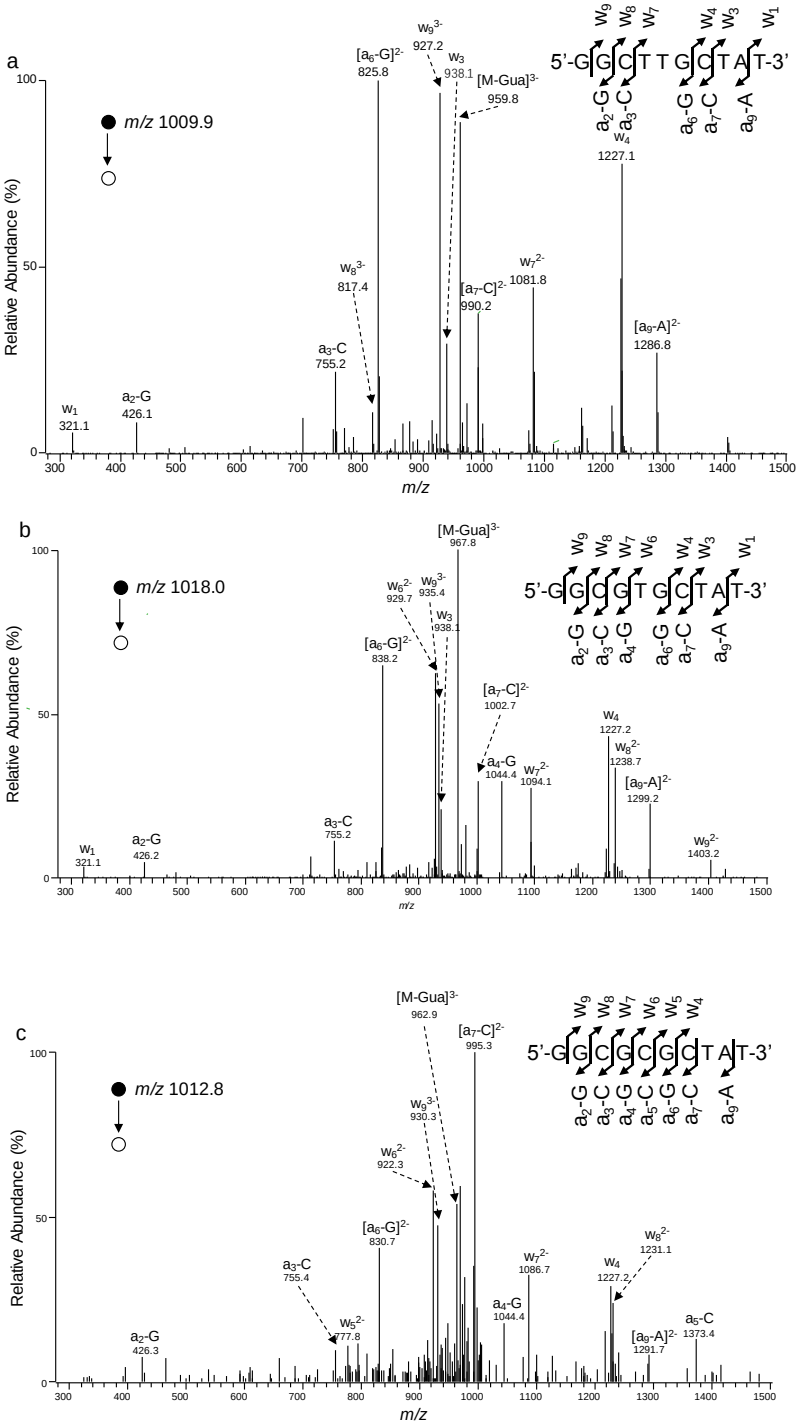
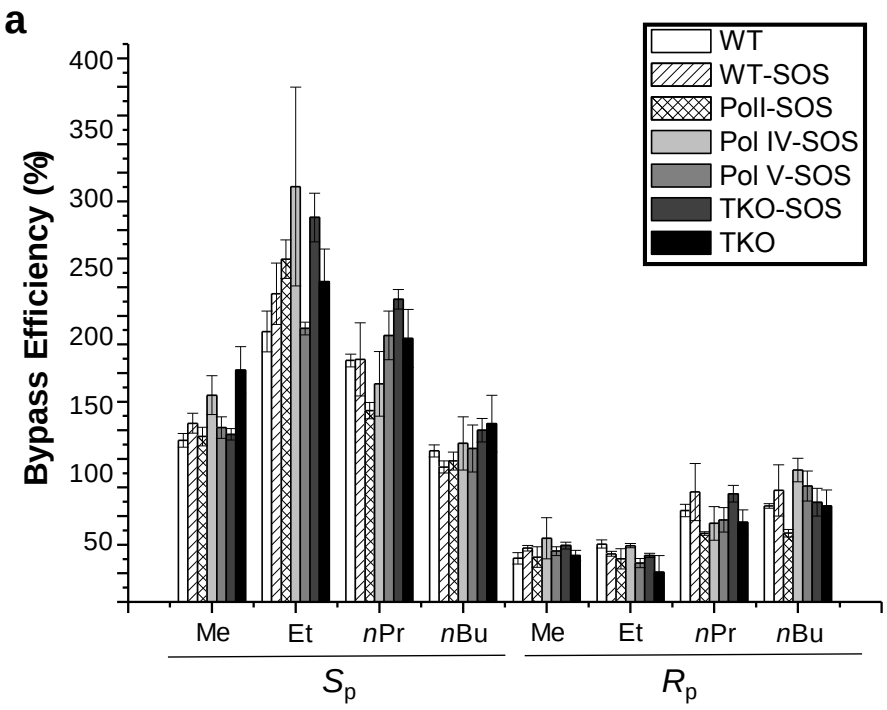


Figure 2-16. LC-MS and MS/MS for the identification of restriction fragments of PCR products. MS/MS for the $[M - 3H]^{3-}$ ions of (a) 10 mer TT (non-mutagenic product), (b) 10 mer GT (TT→GT mutation), and 10 mer GC (TT→GC mutation).


a
Bypass Efficiency (%)
400
350
300
250
200
150
100
50
WT
WT-SOS
PolI-SOS
Pol IV-SOS
Pol V-SOS
TKO-SOS
TKO
Me
Et
nPr
nBu
Me
Et
nPr
nBu
Sp
Rp

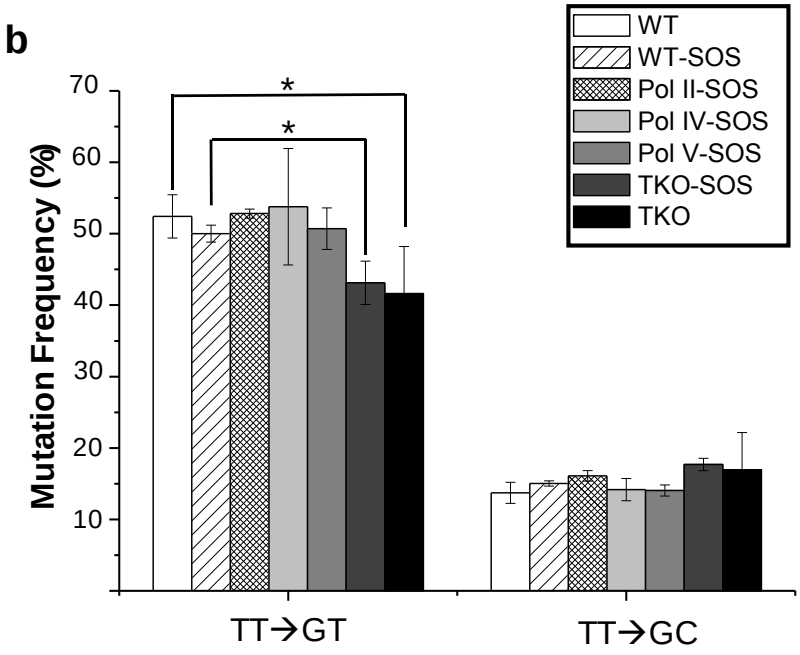

b
Mutation Frequency (%)
70
60
50
40
30
20
10
*
*
WT
WT-SOS
Pol II-SOS
Pol IV-SOS
Pol V-SOS
TKO-SOS
TKO
TT→GT
TT→GC

Figure 2-17. Bypass efficiencies of alkyl phosphotriester lesions (a) and mutation frequencies of the S_p methyl phosphotriester lesion (b) in AB1157 *E. coli* strains that are proficient or deficient in SOS-induced DNA polymerases, Pol II, Pol IV, Pol V, alone or all three in combination. The data represent the means and standard deviations of results from three independent replication experiments. *, $0.01 < p < 0.05$. The p values were calculated by using an unpaired two-tailed t-test and referred to comparisons with the data obtained for control cells.

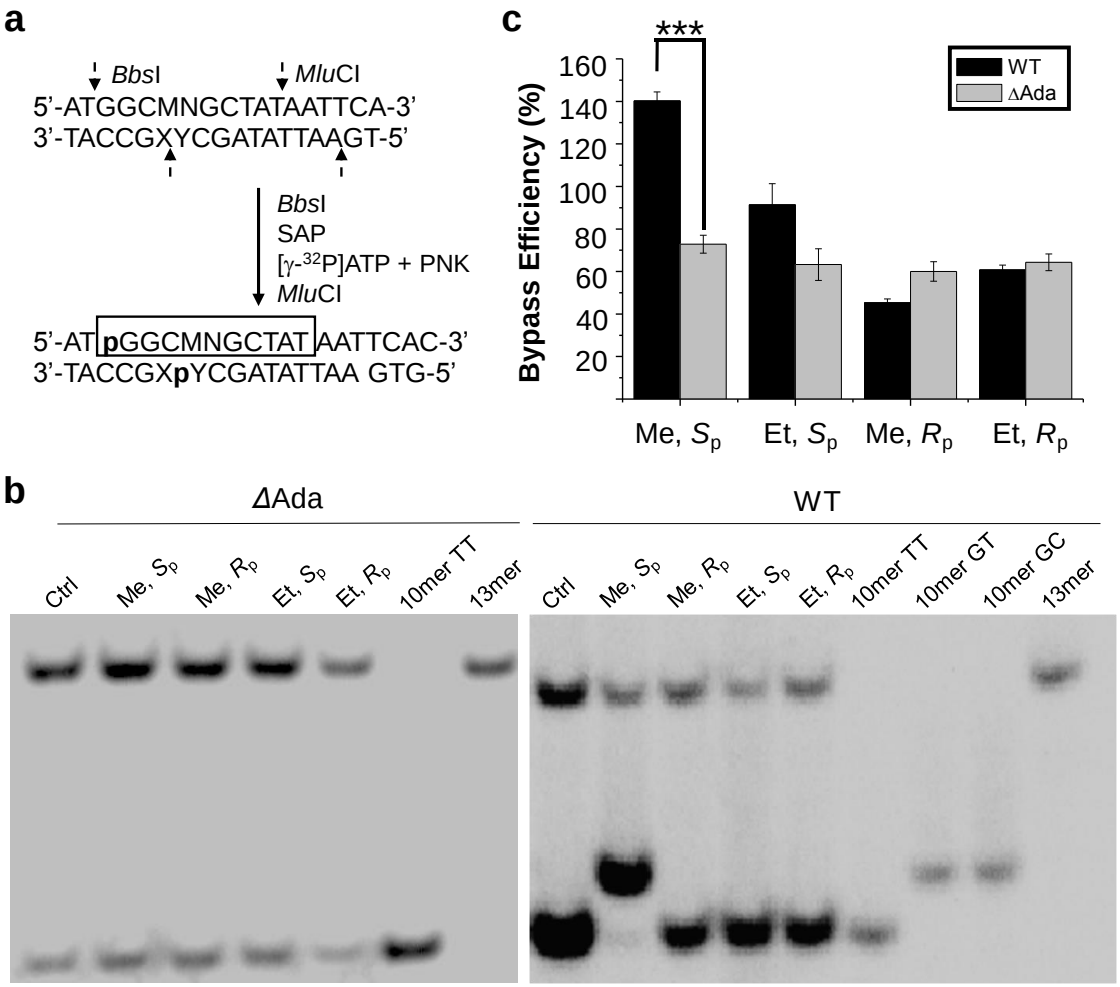


Figure 2-18. Ada protein promotes the replicative bypass of the S_p-Me-PTE, but not the R_p-Me-PTE or the S_p-/R_p-Et-PTE lesion, and Ada is required for the mutations induced by S_p-Me-PTE. (a) Selective labelling of original lesion-containing strand and its complementary strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing 13 mer (from the competitor genome) and 10 mer (from the control or lesion-containing genome) digestion products formed from the strand initially containing the lesion, where 10 mer TT, 10 mer GC and 10 mer GT designate [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with MN being TT, GC and GT, respectively. (c) The bypass efficiencies of the two diastereomers of Me- and Et-PTE lesions in *E. coli* cells that are proficient in DNA repair or deficient in both Ada and Ogt (*Δ*Ada, *Δ*Ogt). The data represent the means and standard deviations of results from three independent replication experiments. ***, $p < 0.001$. The p values were calculated by using an unpaired two-tailed t-test and referred to comparisons with the data obtained for control cells.

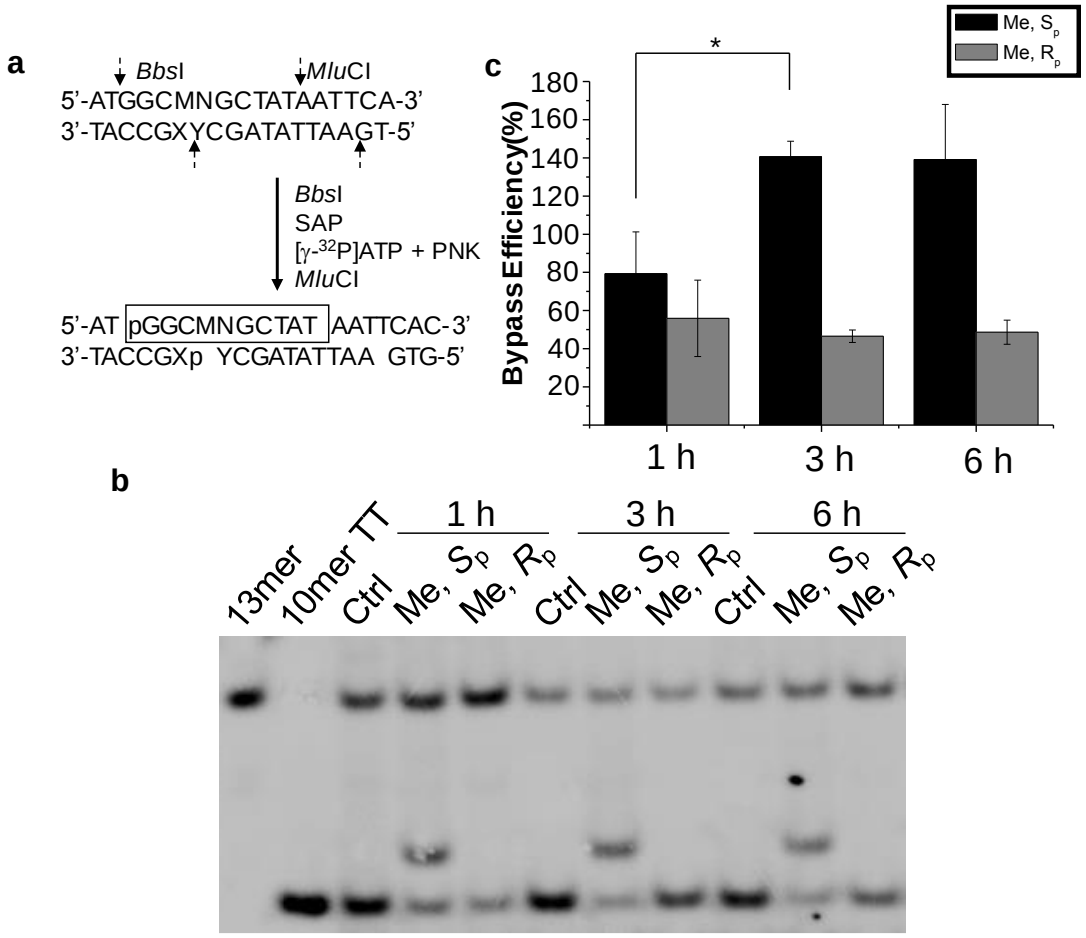


Figure 2-19. Gel image showing the time-dependent CRAB assay of 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome. (a) Selective labelling of original lesion-containing strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where 10mer TT represent the [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with 'MN' being TT, respectively. (c) Bypass efficiencies of alkyl phosphotriester lesions in wild-type AB1157 *Escherichia coli* strains. The data represent the means and standard deviations of results from three independent replication experiments. **, $0.001 < p < 0.01$. The p values were calculated by using an unpaired two-tailed t-test and referred to comparisons with the data obtained for control cells.

REFERENCES

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.

2. Friedberg, E.C., Walker, G.C., Siede, W. and Wood, R.D. (2005) *DNA Repair and Mutagenesis*. American Society for Microbiology Press.

3. Liu, S. and Wang, Y. (2015) Mass spectrometry for the assessment of the occurrence and biological consequences of DNA adducts. *Chem. Soc. Rev.*, **44**, 7829-7854.

4. Shrivastav, N., Li, D. and Essigmann, J.M. (2009) Chemical biology of mutagenesis and DNA repair: cellular responses to DNA alkylation. *Carcinogenesis*, **31**, 59-70.

5. Fu, D., Calvo, J.A. and Samson, L.D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer*, **12**, 104-120.

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

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

8. Ma, B., Zarth, A.T., Carlson, E.S., Villalta, P.W., Upadhyaya, P., Stepanov, I. and Hecht, S.S. (2018) 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*, **39**, 232-241.

9. Den Engelse, L., Menkveld, G.J., De Brij, R.J. and Tates, A.D. (1986) 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*, **7**, 393-403.

10. Shooter, K.V. and Slade, T.A. (1977) The stability of methyl and ethyl phosphotriesters in DNA *in vivo*. *Chem. Biol. Interact.*, **19**, 353-361.

11. Suhasini, A.N., Sommers, J.A., Yu, S., Wu, Y., Xu, T., Kelman, Z., Kaplan, D.L. and Brosh, R.M., Jr. (2012) DNA repair and replication fork helicases are differentially affected by alkyl phosphotriester lesion. *J. Biol. Chem.*, **287**, 19188-19198.

12. McCarthy, T.V. and Lindahl, T. (1985) Methyl phosphotriesters in alkylated DNA are repaired by the Ada regulatory protein of *E. coli*. *Nucleic Acids Res*, **13**, 2683-2698.

13. Weinfeld, M., Drake, A.F., Saunders, J.K. and Paterson, M.C. (1985) Stereospecific removal of methyl phosphotriesters from DNA by an *Escherichia coli ada*$^{+}$ extract. *Nucleic Acids Res.*, **13**, 7067-7077.

14. Tsujikawa, L., Weinfield, M. and Reha-Krantz, L.J. (2003) Differences in replication of a DNA template containing an ethyl phosphotriester by T4 DNA polymerase and *Escherichia coli* DNA polymerase I. *Nucleic Acids Res.*, **31**, 4965-4972.

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

16. Yuan, B., Cao, H., Jiang, Y., Hong, H. and Wang, Y. (2008) Efficient and accurate bypass of N^2-(1-carboxyethyl)-2'-deoxyguanosine by DinB DNA polymerase *in vitro* and *in vivo*. *Proc. Natl. Acad. Sci. USA*, **105**, 8679-8684.

17. Neeley, W.L., Delaney, S., Alekseyev, Y.O., Jarosz, D.F., Delaney, J.C., Walker, G.C. and Essigmann, J.M. (2007) DNA polymerase V allows bypass of toxic guanine oxidation products *in vivo*. *J. Biol. Chem.*, **282**, 12741-12748.

18. Rye, P.T., Delaney, J.C., Netirojjanakul, C., Sun, D.X., Liu, J.Z. and Essigmann, J.M. (2008) Mismatch repair proteins collaborate with methyltransferases in the repair of O^6-methylguanine. *DNA Repair (Amst)*, **7**, 170-176.

19. Dohno, C., Matsuzaki, K., Yamaguchi, H., Shibata, T. and Nakatani, K. (2015) A hybridisation-dependent membrane-insertable amphiphilic DNA. *Org. Biomol. Chem.*, **13**, 10117-10121.

20. Hamamoto, S. and Takaku, H. (1986) New approach to the synthesis of deoxyribonucleoside phosphoramidite derivatives. *Chem. Lett.*, **15**, 1401-1404.

21. Wang, P., Amato, N.J., Zhai, Q. and Wang, Y. (2015) Cytotoxic and mutagenic properties of O^4-alkylthymidine lesions in *Escherichia coli* cells. *Nucleic Acids Res.*, **43**, 10795-10803.

22. Yu, Y., Wang, J., Wang, P. and Wang, Y. (2016) 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.*, **88**, 8036-8042.

23. Delaney, J.C. and Essigmann, J.M. (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions *in vivo*. *Methods Enzymol.*, **408**, 1-15.

24. Zhai, Q., Wang, P., Cai, Q. and Wang, Y. (2014) Syntheses and characterizations of the *in vivo* replicative bypass and mutagenic properties of the minor-groove O^2-alkylthymidine lesions. *Nucleic Acids Res.*, **42**, 10529-10537.

25. Potter, B.V., Connolly, B.A. and Eckstein, F. (1983) Synthesis and configurational analysis of a dinucleoside phosphate isotopically chiral at phosphorus. Stereochemical course of Penicillium citrum nuclease P1 reaction. *Biochemistry*, **22**, 1369-1377.

26. Hayashi, J., Hamada, T., Sasaki, I., Nakagawa, O., Wada, S. and Urata, H. (2015) Synthesis of novel cationic spermine-conjugated phosphotriester oligonucleotide for improvement of cell membrane permeability. *Bioorg. Med. Chem. Lett.*, **25**, 3610-3615.

27. Weinfeld, M., Drake, A.F., Kuroda, R. and Livingston, D.C. (1989) Isolation and characterization of the diastereoisomers of a series of phosphate-ethylated dinucleoside monophosphates. *Anal. Biochem.*, **178**, 93-101.

28. Nakamura, T., Tokumoto, Y., Sakumi, K., Koike, G., Nakabeppu, Y. and Sekiguchi, M. (1988) Expression of the ada gene of *Escherichia coli* in response to alkylating agents. Identification of transcriptional regulatory elements. *J. Mol. Biol.*, **202**, 483-494.

29. Samson, L. and Cairns, J. (1977) A new pathway for DNA repair in *Escherichia coli*. *Nature*, **267**, 281-283.

30. Samson, L. (1992) The suicidal DNA repair methyltransferases of microbes. *Mol. Microbiol.*, **6**, 825-831.

31. He, C., Hus, J.C., Sun, L.J., Zhou, P., Norman, D.P., Dotsch, V., Wei, H., Gross, J.D., Lane, W.S., Wagner, G. *et al.* (2005) A methylation-dependent electrostatic switch controls DNA repair and transcriptional activation by *E. coli ada*. *Mol. Cell*, **20**, 117-129.

32. Harper, M. and Lee, C.J. (2012) Genome-wide analysis of mutagenesis bias and context sensitivity of *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (NTG). *Mutat Res.*, **731**, 64-67.

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

34. Ma, B., Zarth, A.T., Carlson, E.S., Villalta, P.W., Upadhyaya, P., Stepanov, I. and Hecht, S.S. (2018) 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.*, **31**, 48-57.

CHAPTER 3

Ada- and Sequence Context-dependent Mutagenesis of Alkyl Phosphotriester Lesions in *Escherichia coli* cells

INTRODUCTION

The specific sequence of DNA within an organism imparts the genetic code for all domains of life; however, this code is susceptible to alterations due to limited chemical stability of DNA (1). As a result, the genetic integrity of DNA can be compromised by endogenous metabolites and exogenous chemicals, resulting in different types of damage (1,2).

Alkylation is a major type of DNA damage (3), and the cytotoxic effects of DNA alkylation adducts are manifested by the fact that DNA alkylation constitutes the central mechanism of action for many prescribed chemotherapeutic agents (2). Nucleobase modifications have been the major focus of DNA alkylation studies (3), though efficient formation of alkyl phosphotriesters (alkyl-PTEs) have also been demonstrated (4). The latter lesions are induced when alkylating agents react with one of the two non-carbon-bonded oxygen atoms and, based on which of these oxygen atoms is attacked, alkyl-PTEs can form in S_p or R_p configuration (4). It was also shown that chronic exposure of rats to

4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), a major tobacco-specific *N*-nitrosamine, through drinking water induces the formation of Me-PTE lesions in lung tissues (5).

There have been some studies about the repair and biological consequences of alkyl-PTEs (4). Since the addition of an alkyl group to the backbone phosphate neutralizes its negative charge, the presence of alkyl-PTE lesions in DNA may perturb its interactions with proteins. For instance, a mixture of S_p- and R_p-ethyl-PTE inhibits *in vitro* primer extension catalyzed by T4 DNA polymerase and *E. coli* polymerase I (6). Isopropyl-PTE was also found to inhibit the unwinding of duplex DNA mediated by SF2 DNA helicases (6). Our previous study showed that the two diastereomers of alkyl-PTEs exhibited different replication bypass efficiencies in *E. coli* cells, where replication across S_p-Me-PTE at TT dinucleotide site is mutagenic and the mutagenicity of the lesion requires the presence of Ada protein (7).

Previous research indicated that the frequencies for the formation of alkyl-PTEs are influenced by flanking sequences. For instance, Guichard *et al.* (8) employed *N*-nitrosodiethylamine (NDEA) to treat three strains of mice, detected the levels of Et-PTE products by ^{32}P-postlabeling, and observed higher frequencies of Et-PTE lesions with the 5′-flanking nucleoside being a thymidine or 2′-deoxyguanosine compared to a 2′-deoxyadenosine or 2′-deoxycytidine. Likewise, the relative frequencies of 5′-nucleobases

at PTE sites exhibit non-random distribution in calf thymus DNA and liver DNA of BALB/c mice treated with diethyl sulfate (9). LC-MS/MS results also revealed the effects of flanking sequences on the levels of Me-PTE lesions induced in lung tissues of rats treated with NNK and its metabolite, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) (5). Several models were proposed to rationalize the sequence-dependent accumulation of alkyl-PTE lesions: 1) Other than the non-carbon-bound oxygen atoms on backbone phosphate groups, each nucleobase possesses unique nucleophiles, which exhibit different reactivities toward alkylating agents; 2) the alkylating agents' electrophilicity may also influence the non-random distribution of alkyl-PTE lesions, where highly reactive alkylating agents may yield a more random distribution (4); or 3) the repair efficiencies for alkyl-PTEs may vary with flanking base sequences. However, not much is known about how the flanking base sequence context modulates the biological consequences of alkyl-PTEs.

Here, we employed a shuttle vector-based method, in conjunction with LC-MS (10,11), to analyze how the flanking sequences of methyl- and *n*-butyl-PTE lesions influence the fidelity and efficiency of DNA replication in *E. coli* cells, and how replication past alkyl-PTEs is modulated by Ada protein and translesion synthesis DNA polymerases. Considering that we have previously investigated the TT sequence, here we examined the

Me- and *n*Bu-PTE lesions in both S_p and R_p configurations and in six different combinations of flanking sequences (TX and XT, with 'X' being A, C or G) (Figure 3-1).

MATERIALS AND METHODS

Materials

Chemicals were, unless otherwise specified, from Sigma-Aldrich (St Louis, MO, USA) or Thermo Fisher Scientific (Pittsburg, PA, USA). Common reagents for solid-phase DNA synthesis were from Glen Research (Sterling, VA, USA) and all unmodified oligodeoxyribonucleotides (ODNs) were from Integrated DNA Technologies (Coralville, IA, USA). 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) was from Oakwood Products Inc. (West Columbia, SC, USA), and [γ-^{32}P] ATP was obtained from Perkin Elmer (Piscataway, NJ, USA). All enzymes were purchased from New England Biolabs (Ipswich, MA, USA).

M13mp7 (L2) plasmid and wild-type AB1157 *E. coli* strains were kindly provided by Prof. John M. Essigmann (13,14). Polymerase-deficient AB1157 *E. coli* strain [*ΔpolB1*::spec *ΔdinB ΔumuC*::kan (Pol II, Pol IV, Pol V-triple knockout)] was generously provided by Prof. Graham C. Walker (13). Ada-deficient AB1157 *E. coli* strain was generated following published procedures (15).

Chemical syntheses

The 5′-DMTr-protected nucleoside derivatives were synthesized according to the published procedure (Scheme 3-1) (12-15). For methoxyphosphine synthesis (compound **10**, X=Me), compound **9** (200 mg, 0.75 mmol) was dissolved in 10 ml of diethyl ether in an ice bath under argon protection. Methanol (114 μl, 2.8 mmol) and triethylamine (4 ml, 2.8 mmol) were subsequently added to the mixture. After stirring at room temperature for 12 h, 10 ml of petroleum ether was added, and the precipitate was removed by filtration. The supernatant was concentrated to yield methoxyphosphine. For *n*-butoxyphosphine synthesis (compound **10**, X=*n*Bu), *n*-butanol (82 μl, 0.9 mmol) was gradually added to a mixture of compound **9** (200 mg, 0.75 mmol) and triethylamine (130 μl, 0.9 mmol) in tetrahydrofuran (2 ml) under an argon atmosphere in an ice water bath. The mixture was stirred at room temperature for 1 h and filtered. The filtrate was concentrated, redissolved in 5 ml pentane and filtered again. The solvent in the filtrate was removed under vacuum to yield *n*-butoxyphosphine.

For the preparation of compounds **11a1-d1**, compound **10** (1 eq. X=Me) and diisopropylamine hydrotetrazolide (1 eq.) was added to a round bottom flask containing a solution of compound **8a-d** in 1 ml anhydrous dichloromethane. The mixture was stirred at room temperature under argon protection for 1 h. The solvent was removed, and the mixture was separated using flash column chromatography with a mixture of hexane, ethyl

acetate and triethylamine to yield the phosphoramidite building block **11a1-d1**. For the preparation of compounds **11a2-d2**, a solution of compound **8a-d** (60 mg) and tetrazole (1 eq.) in 1 ml acetonitrile was added to a solution of compound **10** (1 eq. X=*n*Bu) in dichloromethane (1 ml) under argon protection. The reaction mixture was stirred for 5 h, concentrated by rotary evaporation, and then purified by flash column chromatography to yield building block **11a2-d2**. The NMR spectra for these compounds are shown in Figures 3-2 to 3-5.

ODN synthesis

A Beckman Oligo 1000M DNA synthesizer (Fullerton, CA) was used to synthesize the 12-mer lesion-containing ODNs, 5′-ATGGCX(Y)TGCTAT-3′ and 5′-ATGGCT(Y)XGCTAT-3′ (‘X’ represents A, C or G; ‘Y’ represents a Me or *n*Bu group) at 1 μmol scale. The synthesized phosphoramidite building block was dissolved in anhydrous acetonitrile at a concentration of 67 mM. Incorporation of unmodified nucleotides was conducted by using commercially available ultramild phosphoramidite building blocks (Glen Research Inc., Sterling, VA, USA) following standard protocols. Synthesized ODNs were cleaved and deprotected from controlled pore glass with concentrated ammonium hydroxide at room temperature for 55 min. After solvent removal using a Speed-Vac, the solid residues were dissolved in water and purified by high-performance liquid chromatography (HPLC).

HPLC

HPLC separation was conducted on an Agilent 1100 system with a Synergi Fusion-RP column (10 × 150 mm, 4 μm in particle size and 80 Å in pore size; Phenomenex Inc., Torrance, CA, USA). Triethylammonium acetate (TEAA) solution (50 mM, pH 6.8), and a mixture of 50 mM TEAA and acetonitrile (70:30, v/v) were employed as mobile phases A and B, respectively. The gradient profile was 5-30% B in 5 min and 30-60% B in 70 min, and the flow rate was 0.8 ml/min. The HPLC traces for the purification of the 12-mer lesion-containing ODNs are shown in Figure 3-6 and their electrospray ionization-mass spectra (ESI-MS) and tandem MS (MS/MS) are provided in Figures 3-7 to 3-18.

Construction of single-stranded lesion-containing and lesion-free competitor M13 genomes

The 12-mer lesion-containing ODNs were 5′-phosphorylated and ligated with a 10-mer ODN (5′-AGTGGAAGAC-3′) to yield 22-mer lesion-containing ODNs, and the ensuing ligation products were purified by denaturing polyacrylamide gel electrophoresis (PAGE).

The single-stranded lesion-containing and lesion-free competitor M13 genomes were prepared following published procedures (12). Briefly, 20 pmol of M13mp7 (L2) plasmid was digested with 40 U *EcoR*I-HF at 25 °C for 8 h to linearize the vector. Two scaffolds, 5′-CTTCCACTCACTGAATCATGGTCATAGCTTTC-3′ and 5′-AAAACGACGGCCAGTGAATTATAGC-3′ (25 pmol each), were subsequently annealed

with the linearized vector. A phosphorylated 22-mer lesion-containing or lesion-free control ODN, or a 25-mer competitor ODNs, was then added to the mixture and incubated with scaffold ODNs and T4 DNA ligase at 16 °C for 8 h. Unligated linear vector and ODNs were removed by the exonuclease activity of T4 DNA polymerase (22.5 U, 16 °C for 2 h). The resulting plasmids were purified using Cycle Pure Kit (Omega), and the purified lesion-containing and lesion-free control plasmids were subsequently normalized against the competitor plasmid (12).

Preparation of Ada-deficient AB1157 *E. coli* strain

P1 transduction was employed to obtain the Ada-deficient *E. coli* strain (*Δada*::kan) in AB1157 background (13). Genotype of the ensuing deficient strain was confirmed by antibiotic resistance and polymerase chain reaction (PCR) followed by sequencing.

Transfection of M13 genomes into *E. coli* cells

The lesion-free control or lesion-containing plasmids were mixed with the competitor plasmid at a 1:1 molar ratio. The mixtures were transfected into electrocompetent wild-type AB1157 *E. coli* strains as well as the isogenic cells deficient in Ada, or all three SOS-induced DNA polymerases (i.e. Pol II, Pol IV and Pol V) (11). The transfected *E. coli* cells were cultured at 37 °C for 5.5 h and the M13 phage was isolated from the supernatant by centrifugation at 13200 r.p.m. for 5 min. The purified M13 phage was transfected into

SCS110 *E. coli* cells for amplification, followed by extraction with QIAprep Spin M13 kit (Qiagen) to obtain M13 ssDNA template for PCR amplification.

Quantification of bypass efficiency by the competitive replication and adduct bypass (CRAB) assay

We utilized a modified version of the competitive replication and adduct bypass (CRAB) assay to assess the bypass efficiency of alkyl-PTE lesions in *E. coli* cells (10-12,14). The interested regions in the M13 ssDNA templates were amplified by PCR with the use of Phusion high-fidelity DNA polymerase. The primers were 5′-YCAGCTATGACCATGATTCAGTGAGTGGA-3′ and 5′-YTCGGTGCGGGCCTCTTCGCTATTAC-3′, where 'Y' represents a 5′-amino modifier conjugated to the 5′-phosphate group of ODNs, i.e., $H_2N(CH_2)_6$-. The PCR amplification started from 98 °C for 30 s, followed by 35 cycles of amplification, with each cycle consisting of 98 °C for 10 s, 65 °C for 30 s and 72 °C for 15 s, and then with a final extension at 72 °C for 5 min, ending at 4 °C. The PCR products were purified by Cycle Pure kit (Omega).

The PCR products (100 ng) were digested with *Bbs*I-HF restriction endonuclease (10 U) and recombinant shrimp alkaline phosphatase (rSAP, 10 U) in 10 μl 1× CutSmart buffer (New England Biolabs) at 37 °C for 25 min, followed by deactivation of rSAP at 80 °C for 10 min. To the above mixture were added 5 mM dithiothreitol (DTT), 1.66 pmol [γ-^{32}P]ATP,

10 U T4 polynucleotide kinase (T4 PNK), CutSmart buffer and water to give a total volume of 15 μl. The mixture was incubated at 37 °C for 30 min, and T4 PNK was then deactivated by heating the solution at 70 °C for 10 min. The resulting mixture was further digested with 10 U *Mlu*CI at 37 °C for 25 min and subsequently quenched by adding 15 μl formamide gel-loading buffer containing xylene cyanol FF and bromophenol blue dyes. The radiolabeled digestion mixtures were resolved using a 30% native polyacrylamide gel (19:1) and the intensities for the radiolabeled gel bands were measured by using a Typhoon 9410 imager.

The aforementioned digestion procedures yield a 10-mer duplex: 5′-p*GGCMNGCTAT-3′/5′-AATTATAGCY-3′ for full-length replication products, with 'M' and 'N' being the nucleobases at the dinucleotides initially flanking the alkyl-PTE lesions, 'Y' being the complementary base of 'N' in the opposite strand, and p* being the radiolabeled phosphate. The bypass efficiency was calculated by: Bypass efficiency (%) = (lesion signal/competitor signal)/ (control signal/competitor signal) × 100%.

Identification and quantification of mutation frequencies by restriction endonuclease and mass spectrometry (REAMS) assay

Approximately 3 μg PCR products in 250 μl of 1× CutSmart buffer was mixed with 50 U *Bbs*I-HF, 20 U rSAP, and incubated at 37 °C for 2 h, followed by deactivation of rSAP at 80 °C for 20 min. *Mlu*CI (20 U) was subsequently added to the mixture and the digestion

was continued at 37 °C for 1 h. The resulting mixture was extracted once with phenol/chloroform/isoamyl alcohol (25:24:1, v/v) and the aqueous phase was evaporated, desalted by Waters Oasis HLB extraction cartridges (Milford, MA), and redissolved in 10 μl water. A 5-μl aliquot was analyzed by LC-MS/MS on an LTQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA, USA) with an Agilent Zorbax SB-C18 column (0.5 × 150 mm, 5 μm in particle size). The gradient was 5 min of 5–20% methanol followed by 35 min of 20–50% methanol in 400 mM HFIP (pH was adjusted to 7.0 with triethylamine). The temperature for the ion-transfer tube was 300 °C, and the mass spectrometer was set up for acquiring the higher-resolution 'ultra-zoom scan' MS and full-scan MS/MS for the $[M-3H]^{3-}$ ions of 10 mer ODNs, d(GGCMNGCTAT), with 'M' and 'N' being A, T, C and G.

Plasmid construction and protein purification.

E. coli ada gene was amplified from AB1157 strains with primers containing *BamH*I and *Xho*I restriction recognition sites and inserted into the pGEX-4T1 vector, which was generously provided by Prof. Chuan He. The sequences of primers are listed in Table 3-1.

Recombinant Ada protein was obtained by inducing transformed Rosetta (DE3) pLysS *E. coli* cells with isopropyl 1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 0.5 mM when DE3 culture reached $OD_{600} \approx 0.55$, followed by culture at 37 °C for 4 h. The recombinant protein was extracted using Pierce Glutathione Agarose (Thermos) and

then eluted with a reduced glutathione solution (3 mg/ml), followed by concentration using Microcon ultracentrifugal filters (Millipore).

Electrophoresis mobility shift assay (EMSA)

Single-stranded ODNs were radiolabeled on the 5′ termini at 37 °C for 30 min in a mixture containing 10 U T4 PNK and 1.66 pmol [γ-^{32}P]ATP in 10 μl PNK buffer, followed by deactivation of the enzyme by heating at 70 °C for 20 min. For duplex DNA, the radiolabeled ODNs were annealed with the corresponding complementary strands at a 1:1 molar ratio.

The binding experiments were conducted using solutions containing 0-1.5 μM Ada protein, 1 nM of labeled 22 mer lesion-containing or lesion-free DNA in a binding buffer [10 mM Tris (pH 7.5), 1 mM EDTA, 0.1 M KCl, 0.1 mM DTT, 5% glycerol and 10 μg/ml BSA]. The mixtures were incubated at room temperature for 1 h, followed by separation on a 10% 29:1 polyacrylamide gel in TAE buffer (40 mM Tris, 2.5 mM EDTA, pH 7.8) at 4 °C for 45 min.

RESULTS

The aim of the present study was to gain a comprehensive understanding about the impact flanking base sequences have on DNA replication past alkyl-PTE lesions in *E. coli*, and to

examine the roles of Ada protein and TLS polymerases in modulating the replicative bypass of these lesions.

We synthesized 12-mer ODNs containing a site-specifically inserted alkyl-PTE lesion in different sequence contexts following our recently published procedures (7), with the modification that the exocyclic amino groups of adenine, cytosine and guanine bases in the phosphoramidite building blocks were protected (see Materials and Methods). The synthesized ODNs were purified by using HPLC (Figure 3-6). Because the S_p and R_p diastereomers of the T(Me)C- and T(*n*Bu)A-containing ODNs cannot be resolved from each other by HPLC, a mixture of the two diastereomers were utilized for these two ODNs in the subsequent experiments. We characterized the ODNs using ESI-MS and MS/MS and the results confirmed the expected site of alkyl-PTE incorporation and the sequences of the modified ODNs (Figures 3-7 to 3-18).

We employed a previously reported shuttle vector method to assess the bypass efficiencies and mutation frequencies of the alkyl-PTE lesions (12). The lesion-containing ODNs were ligated into single-stranded M13 phage. After replication, progeny recovery, PCR amplification and restriction enzyme digestion, the released ODNs were analyzed by native PAGE and LC-MS/MS to identify the replication products (Figures 3-19 to 3-33). As shown in Figure 3-19a, we employed two restriction enzymes, *Bbs*I and *Mlu*CI, to digest the PCR products of the progeny genome, resulting in the liberation of the initial damage-

containing region as 10 mer ODNs for the lesion-containing or control genome, or a 13 mer ODN for the corresponding region in the competitor genome. By switching the order of digestion of the two restriction enzymes, we selectively radiolabel the 5′-terminus of either the original lesion-situated strand (p*GGCMNGCTAT) or the opposite strand (p*AATTATAGCY).

Replication across the alkyl-PTEs may yield up to 16 potential products (i.e. with the four natural nucleotides being incorporated at the two nucleosides flanking the PTE site), which could not be completely resolved from one another by PAGE analysis. Thus, we employed a restriction endonuclease and mass spectrometry (REAMS) assay (10) to identify the mutagenic products and to quantify the mutation frequency. In this vein, the aforementioned digestion products were subjected to LC-MS and MS/MS analyses, where we monitored the fragmentation of the $[M-3H]^{3-}$ ions. The mutation frequencies were quantified by the calibrated ratios of peaks found in the selected-ion chromatograms (SICs) for the mutagenic and non-mutagenic products (Figures 3-22, 3-23 and 3-29b). Intriguingly, we found that the replication products of the S_p-Me-PTEs at the three XT sites ('X' = A, C or G) are largely independent of the neighboring 5′ nucleosides, with ~85-90% and ~5-10% of products carrying AT and TG at bases flanking the initial damage site, respectively, though replication across C(Me)T also yields 7% of non-mutagenic replication products (Figures 3-20, 3-21). Additionally, none of the R_p diastereomers of the Me-PTEs in the TX

sequences, neither diastereomer of Me-PTEs in the TX sequences, and none of the *n*Bu-PTEs in the TX or XT sequence are mutagenic (Figures 3-20, 3-24, 3-27 and 3-28).

The bypass efficiencies of the Me-PTE lesions were quantified by comparing the relative intensity of signal for the 10-mer products from the lesion-containing or control genome, to that of the 13-mer replication product from the competitor genome; this value was adjusted based on the molar ratios of lesion/competitor and control/competitor genomes used for the transfection (Figure 3-29a). For Me-PTEs, none of the S_p-PTEs at the three XT sites were strong impediments to DNA replication, though the corresponding R_p diastereomers significantly blocked DNA replication. Additionally, the blockage effects were more pronounced when the flanking 5′-nucleobase was a purine (adenine and guanine). Meanwhile, all TX-PTEs elicited moderate blockage effects on DNA replication.

We also examined whether SOS-induced DNA polymerases (Pol II, Pol IV and Pol V) promote the replicative bypass of different Me-PTEs. Our results showed that, similar to the results obtained for the alkyl-PTE lesions at TT site (7), genetic depletion of all three SOS-induced DNA polymerases did not exert any apparent effects on the replication bypass efficiencies for the Me-PTE lesions. For *n*Bu-PTEs, we found that none of the S_p-PTEs at XT sites significantly impede DNA replication in *E. coli* cells, whereas the R_p diastereomer in the XT sequence and both diastereomers of *n*Bu-PTEs in the TX sequence confer moderate blockage effects (Figure 3-29c).

Considering our previous observation that Ada protein can influence the mutagenicity and cytotoxicity of S_p-Me-PTE at TT dinucleoside site (7), we next asked how Ada protein affects the replication bypass efficiencies and mutation patterns of S_p-Me-PTEs in different flanking base sequences (Figures 3-30 to 3-33). We found that removal of the *ada* gene resulted in a moderate decline in bypass efficiencies for the S_p-Me-PTEs in all three XT sequences, and a complete abrogation of mutations. However, depletion of *ada* gene did not elicit significant effects on the replication efficiencies or mutations arising from S_p-Me-PTEs in any of the three TX sequences.

The above strong flanking sequence- and Ada-dependent mutagenesis of S_p-Me-PTE lesions prompted us to examine the binding capabilities of purified Ada protein to single- and double-stranded DNA containing a S_p-Me-PTE in two flanking base sequences (AT and TA) (Figure 3-34). Our results from electrophoretic mobility shift assay (EMSA) showed that Ada protein binds more strongly to single-stranded S_p-A(Me)T-containing ODN than its unmethylated counterpart, whereas S_p-T(Me)A-containing ODN did not display significant difference in binding toward Ada relative to the corresponding TA-containing ODN. In this vein, Ada protein also exhibits stronger binding to single-stranded lesion-free ODN with an AT sequence relative to the corresponding ODN with a TA sequence, though the binding affinities of the protein to duplex DNA are independent of whether the DNA carries a Me-PTE.

DISCUSSION

We previously investigated how size and stereochemistry of the alkyl-PTEs at TT site influence DNA replication in *E. coli* (7). Here, we systematically investigated how compositions of the flanking nucleobases of the alkyl-PTE lesions affect DNA replication. We found that, for Me-PTEs in XT sequences (X = A, C, G), some of the characteristics were the same as what we found for the replicative bypass of T(Me)T (7). First, we demonstrated that neither the replication bypass efficiency nor the mutation frequency of Me-PTEs was impacted by genetic depletion of all three SOS-induced DNA polymerases. Second, we revealed that none of S_p-Me-PTEs in the three XT sequences suppress DNA replication in *E. coli* cells, whereas the R_p-Me-PTEs in the three sequence contexts exhibit significant replication blockage effects. Third, we show that S_p-Me-PTEs at XT sites are mutagenic and Ada protein is indispensable for the mutagenic bypass. In this vein, removal of Ada protein also resulted in a decrease in bypass efficiencies for the S_p-Me-PTEs in the three XT sequences.

We also uncover some unique features regarding replication past alkyl-PTEs with different flanking base sequences. Strikingly, we found that the distribution of replication products for S_p-Me-PTEs in the three XT sequences (X = A, C or G) was largely independent of the 5′-neighboring nucleobase being an A, C, or G, where approximately 85-90% and 5-10% of the replication products for the S_p-Me-PTEs at XT sties were AT and

TG, respectively. In this context, it is worth noting that S_p-Me-PTE at TT site induced 50% TT→GT and 15% TT→GC mutations, and Ada protein is also essential for the induction of these mutations (Figure 3-33) (7).

Ada was reported to remove the methyl group from S_p-Me-PTEs, where Cys38 on N-terminal domain of the protein interacts with the methyl group on the DNA backbone (15,16). A previous structural study showed that the methylated N-terminal domain of Ada protein preferentially recognizes A/T in the Ada box (15,17). In this recognition, Arg45 participates in hydrogen bonding interactions with thymine residues in both strands, and the binding with Arg45 is sterically incompatible with a G positioned in the last two bases of the Ada box (15,17), which is consistent with the sequence specificity observed for replication past alkyl-PTEs.

In contrast to what we found for Me-PTEs at XT sites (X = A, C, G, T), replication across the Me-PTEs at TX sites (X=A, C and G) was accurate. In addition, genetic ablation of Ada did not alter the efficiency or fidelity of replication across any of the Me-PTEs at TX site. We also examined the replication past *n*Bu-PTEs and found that the trends in bypass efficiency was similar to what we found for Me-PTEs (Figure 3-29c); however, none of *n*Bu-PTEs were mutagenic.

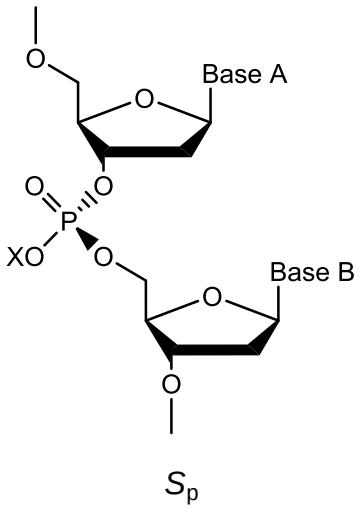
The results from our replication studies are supportive of a model where Ada protein binds to S_p-Me-PTE lesions formed at XT dinucleotide sites, and this binding is maintained

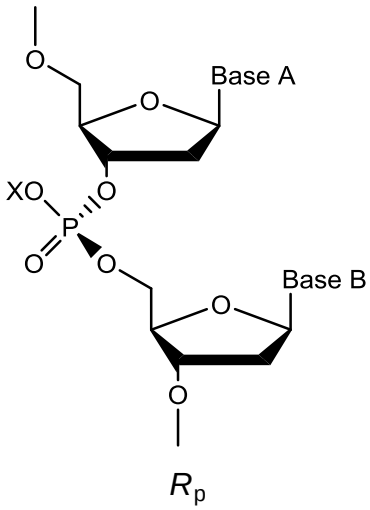
during the replicative bypass of these lesions in single-stranded DNA. Moreover, this interaction with Ada protein also assists the replicative bypass of these lesions in *E. coli* cells. The lack of apparent dependence of profiles of replication products for the S_p-Me-PTE lesion in the three XT sequences (X = A, C or G) on the identities of the 5′-flanking nucleobases strongly suggests their lack of recognition by DNA polymerases during nucleotide incorporation at the site. These results, together with the previous observation that an arginine residue in REV1 could direct the incorporation of dCMP through direct hydrogen bonding interaction with the nucleobase in the incoming nucleotide (20), suggest that some amino acid(s) in the Ada protein may direct nucleotide incorporation when replicating past the 5′ flanking nucleotide of the Me-PTE lesions. Consistent with this notion, we found that single-stranded (ss) DNA with an S_p-A(Me)T displayed augmented binding relative to the corresponding ssDNA with an S_p-T(Me)A, suggesting that the flanking 3′-T enhances the interaction between S_p-Me-PTEs and Ada protein.

Harper *et. al.* (21) analyzed the mutations induced by *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine in sixteen different strains of *E. coli*, and they found that 96.6% of the 4099 detected mutations were G→A transition mutation, which were attributed to O^6-MedG (14,22). Our results, however, suggest that the G→A mutation may also arise, in part, from S_p-Me-PTE formed at GT site.

In summary, the results from our shuttle vector-based replication study showed that flanking base sequences play important roles in DNA replication across Me- and *n*Bu-PTE lesions. S_p-Me- PTEs and S_p-*n*Bu-PTEs at XT sites (X = A, C, G) were not strong impediments to DNA replication, whereas their R_p counterparts exhibited blockage effects. Meanwhile, Me- and *n*Bu-PTEs at TX sites moderately block DNA replication in *E. coli*. Furthermore, the S_p-Me-PTEs at XT sites resulted in 85-90% XT→AT and 5-10% XT→ TG mutations, which require the presence of Ada protein. However, this phenomenon was not observed for R_p-Me-PTEs at TX sites, which is consistent with the notion that Ada protein does not recognize R_p-Me-PTEs. There are two established functions for Ada protein, i.e. removal of the methyl group from S_p-Me-PTE or O^6-MedG, and transcriptional activation of *ada*, *alkA*, *alkB* and *aidB* genes (3,18,19); both functions are mediated by the N-terminal domain of Ada (15). Our work suggests that, aside from these two well-characterized functions, Ada may assume other important functions in cells. Along this line, it will be important to examine whether Me-PTE lesions are endogenous modifications and whether they assume regulatory roles in *E. coli* cells. In addition, future studies about how alkyl-PTE lesions influence DNA replication in mammalian cells will provide additional insights into the biological impacts of this unique class of DNA damage.

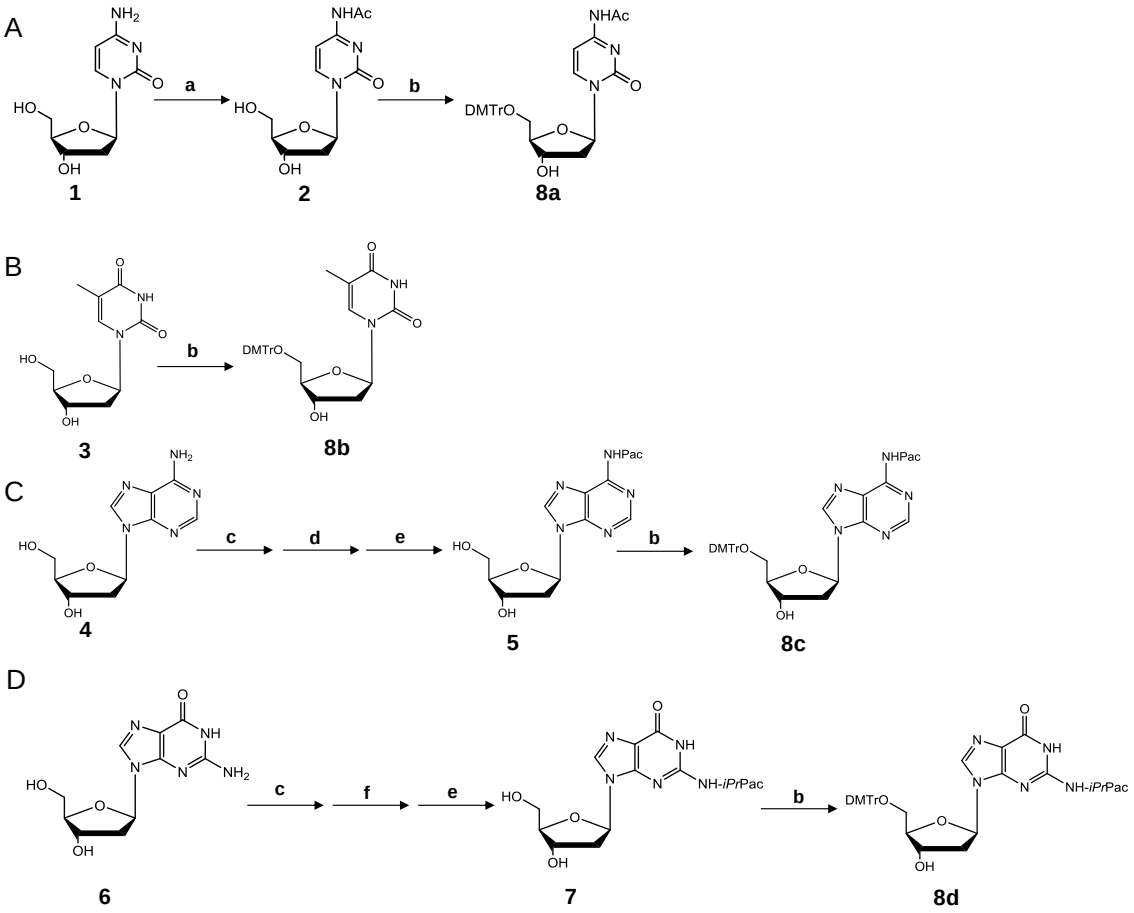
FIGURES AND SCHEMES





Base A	Base B
A	T
C	T
G	T
T	A
T	C
T	G

Figure 3-1. S_p and R_p diastereomers of alkyl phosphotriester residues in DNA (X=Me or *n*Bu).



a1: Base=C^{Ac}, X=Me **a2: Base=C^{Ac}, X=*n*Bu**
b1: Base=T , X=Me **b2: Base=T , X=*n*Bu**
c1: Base=A^{Pac}, X=Me **c2: Base=A^{Pac}, X=*n*Bu**
d1: Base=G^{iPrPac}, X=Me **d2: Base=G^{iPrPac}, X=*n*Bu**

Scheme 3-1. Syntheses of alkyl phosphoramidite building blocks. Reagents and conditions: (a) acetic anhydride, *N,N*-dimethylformamide, R. T. , 24 h; (b) 4,4′-dimethoxytrityl chloride, 4-dimethylaminopyridine, pyridine, R. T. , 10 h; (c) trimethylsilyl chloride, pyridine, R.T., 1 h; (d) phenoxyacetyl chloride, R.T., 2 h; (e) saturated $NaHCO_3$ solution, R.T., 5 h; (f) 4′-isopropylphenoxyacetyl chloride, 4-dimethylaminopyridine, R.T., 2 h; (g) X=Me: Methanol, triethylamine, diethyl ether, ice bath, 14 h; X=*n*Bu: *n*-butanol, triethylamine, tetrahydrofuran, R. T. , 1 h; (h) X=Me: diisopropylamine hydrotetrazolide, dichloromethane, R. T. , 1 h; X=*n*Bu: tetrazole, acetonitrile, dichloromethane, R.T., 5 h.

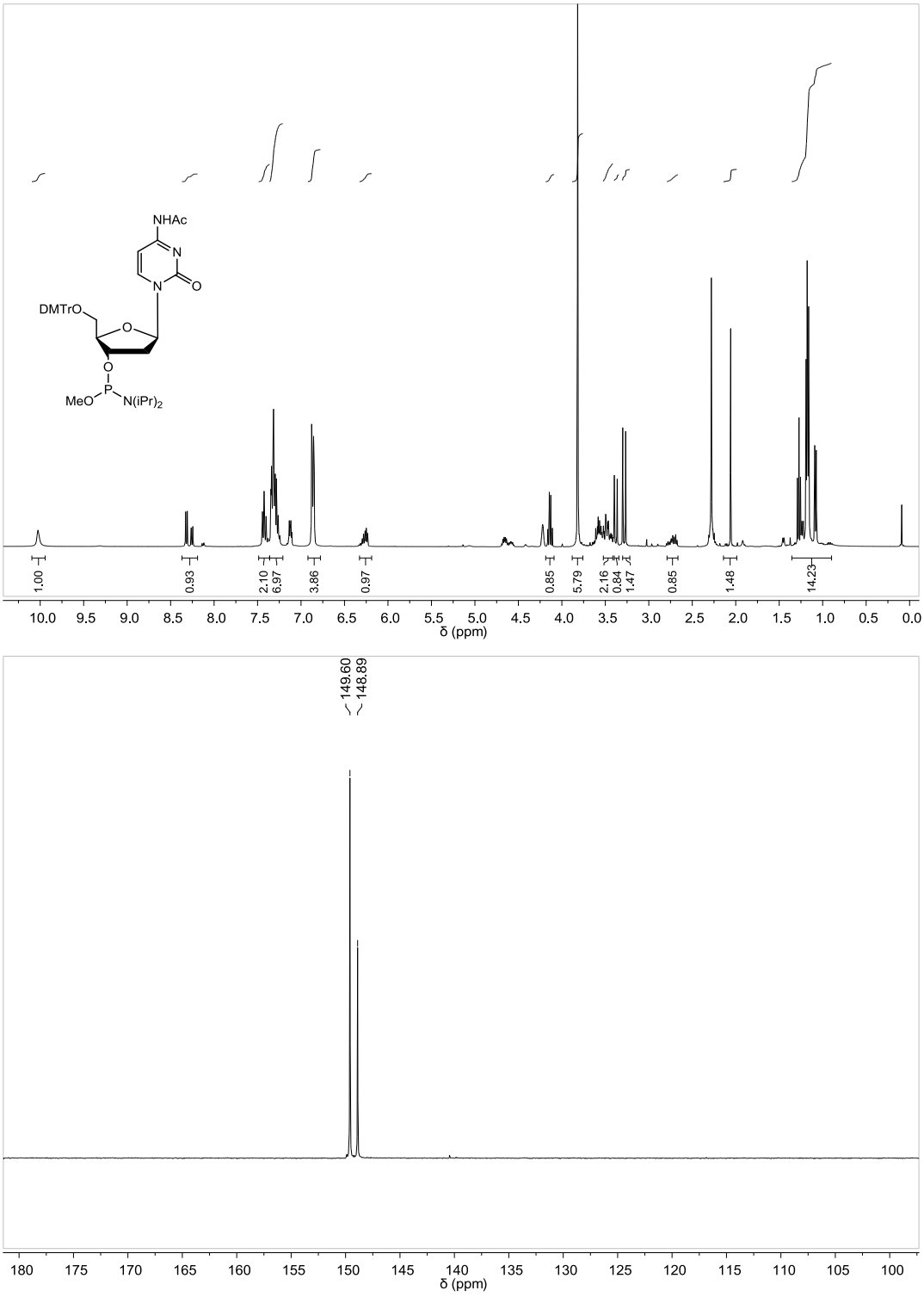


Figure 3-2. The ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, top) and ^{31}P NMR (80 MHz, $CDCl_3$, 25ºC, bottom) spectra of N^4-AcdC-methylphosphoramidite.

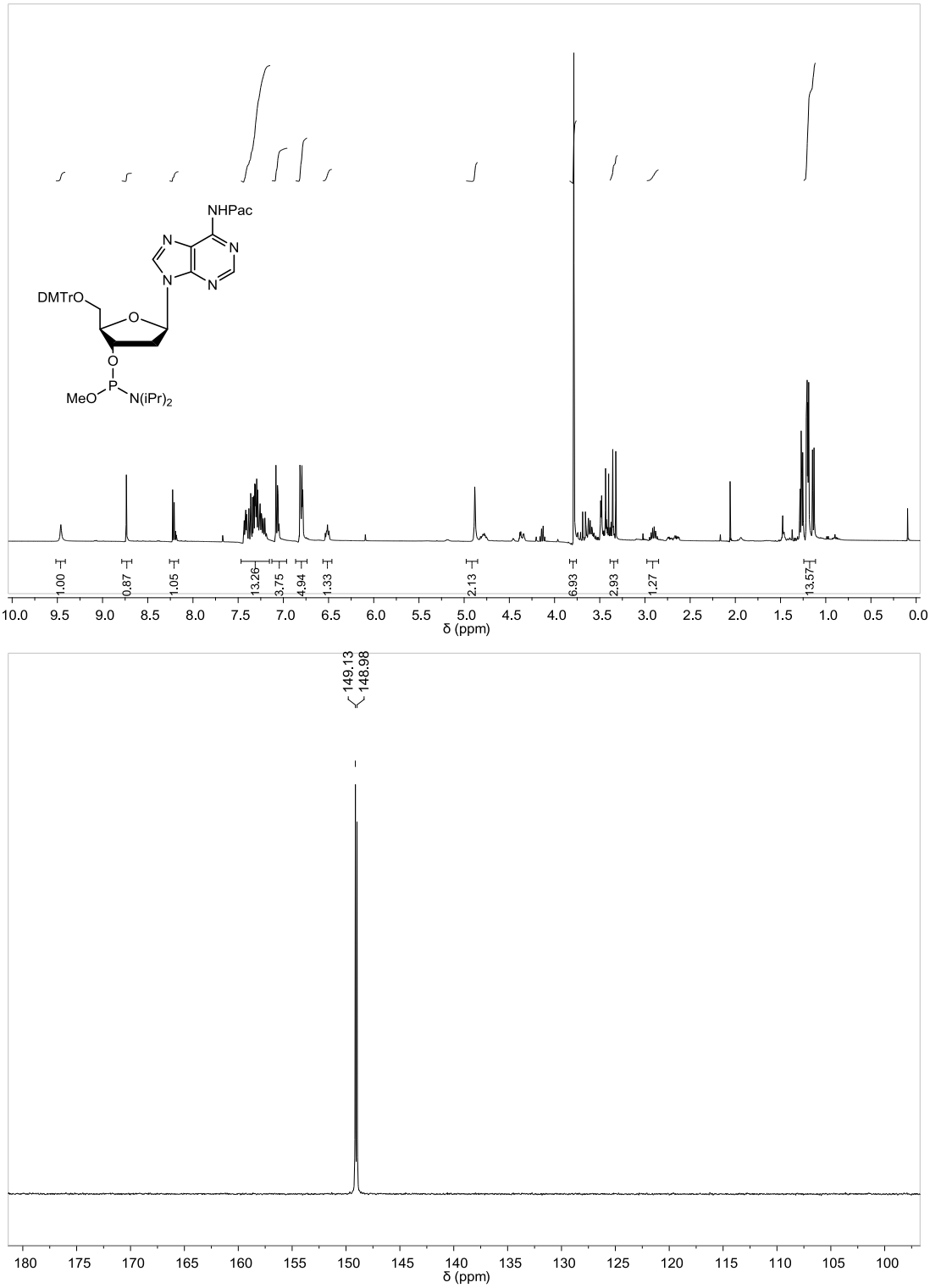


Figure 3-3. The ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, top) and ^{31}P NMR (80 MHz, $CDCl_3$, 25ºC, bottom) spectra of N^6-PacdA-methylphosphoramidite.

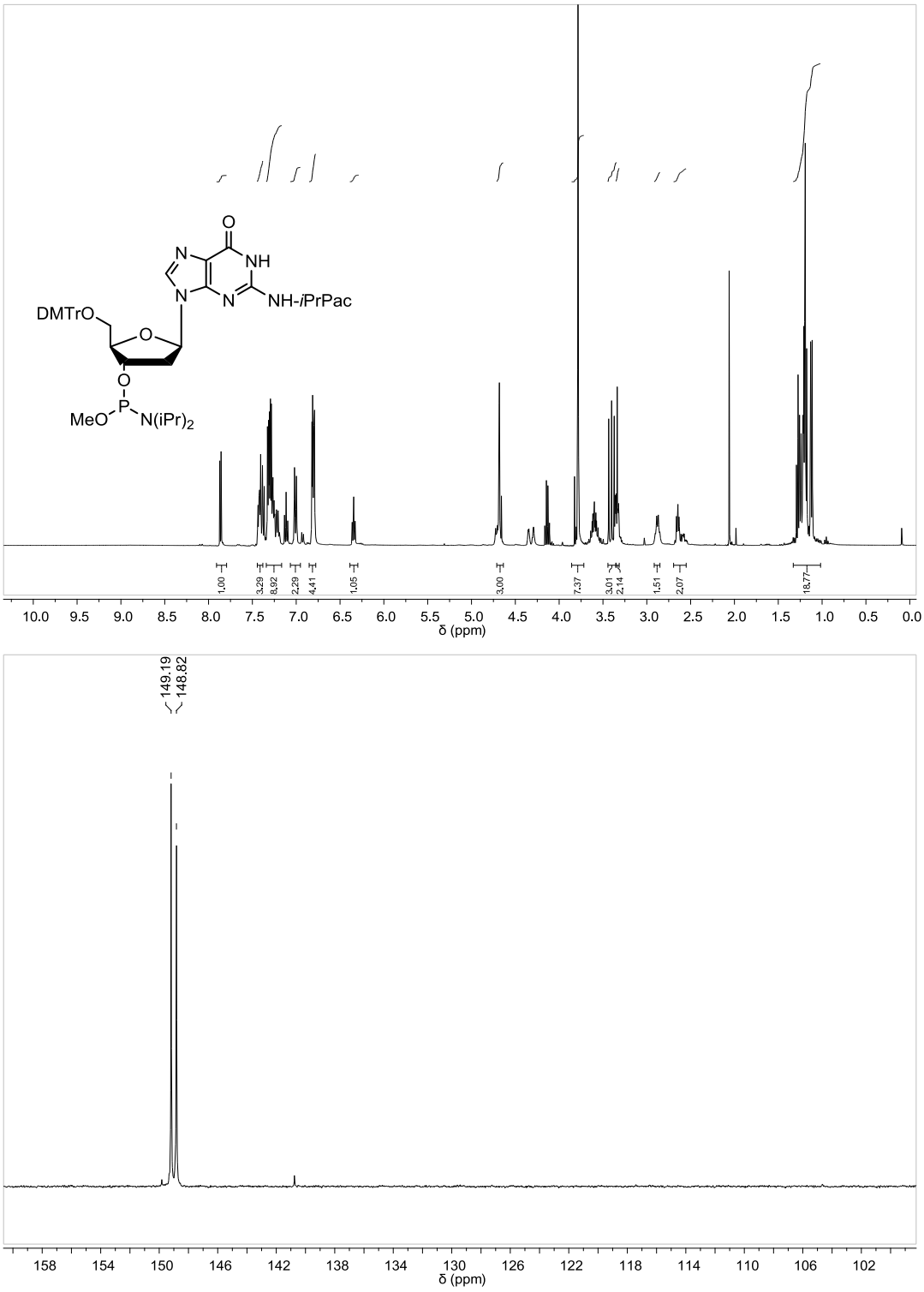


Figure 3-4. The ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, top) and ^{31}P NMR (80 MHz, $CDCl_3$, 25ºC, bottom) spectra of N^2-*i*PrPacdG-methylphosphoramidite.

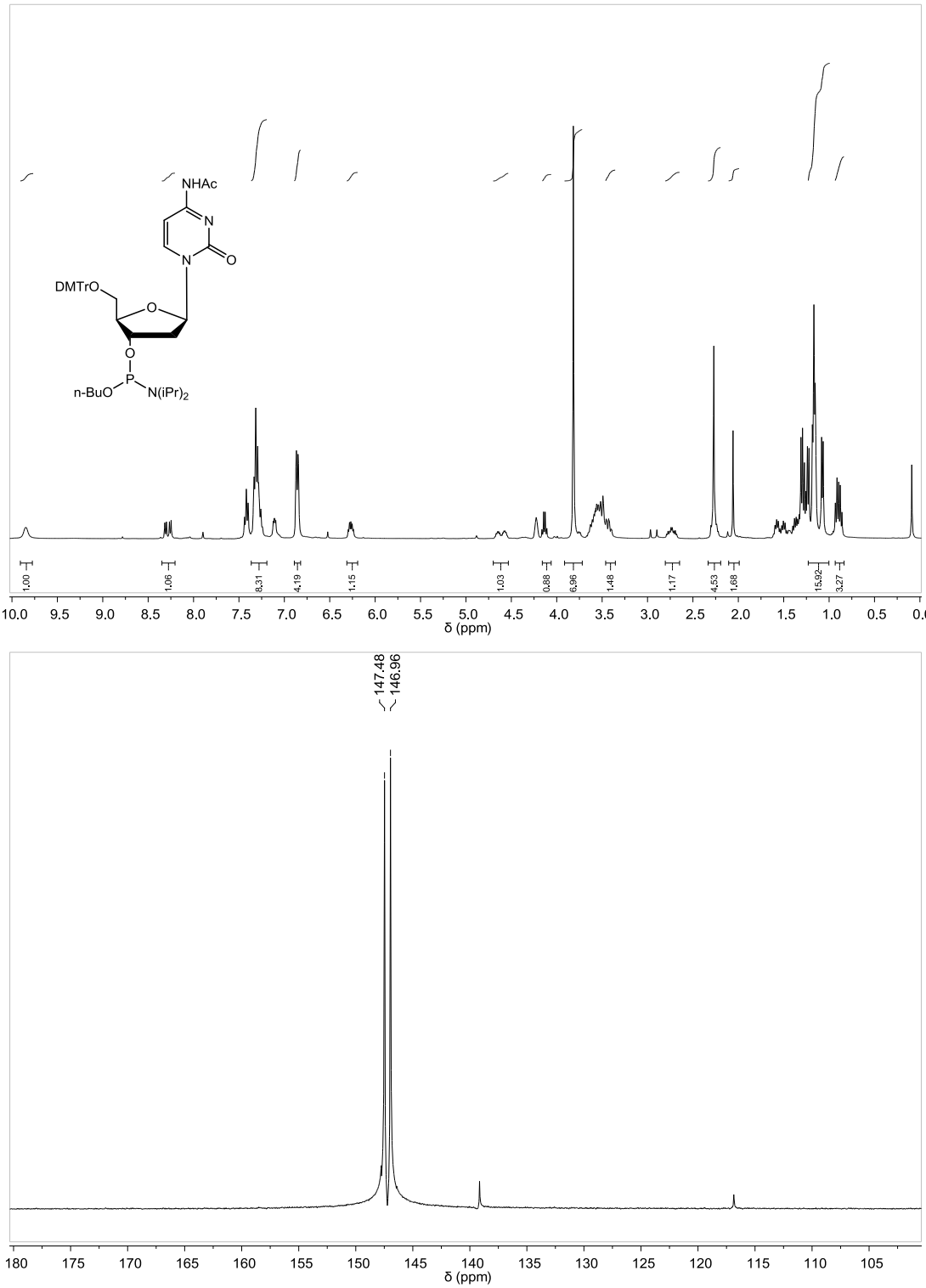
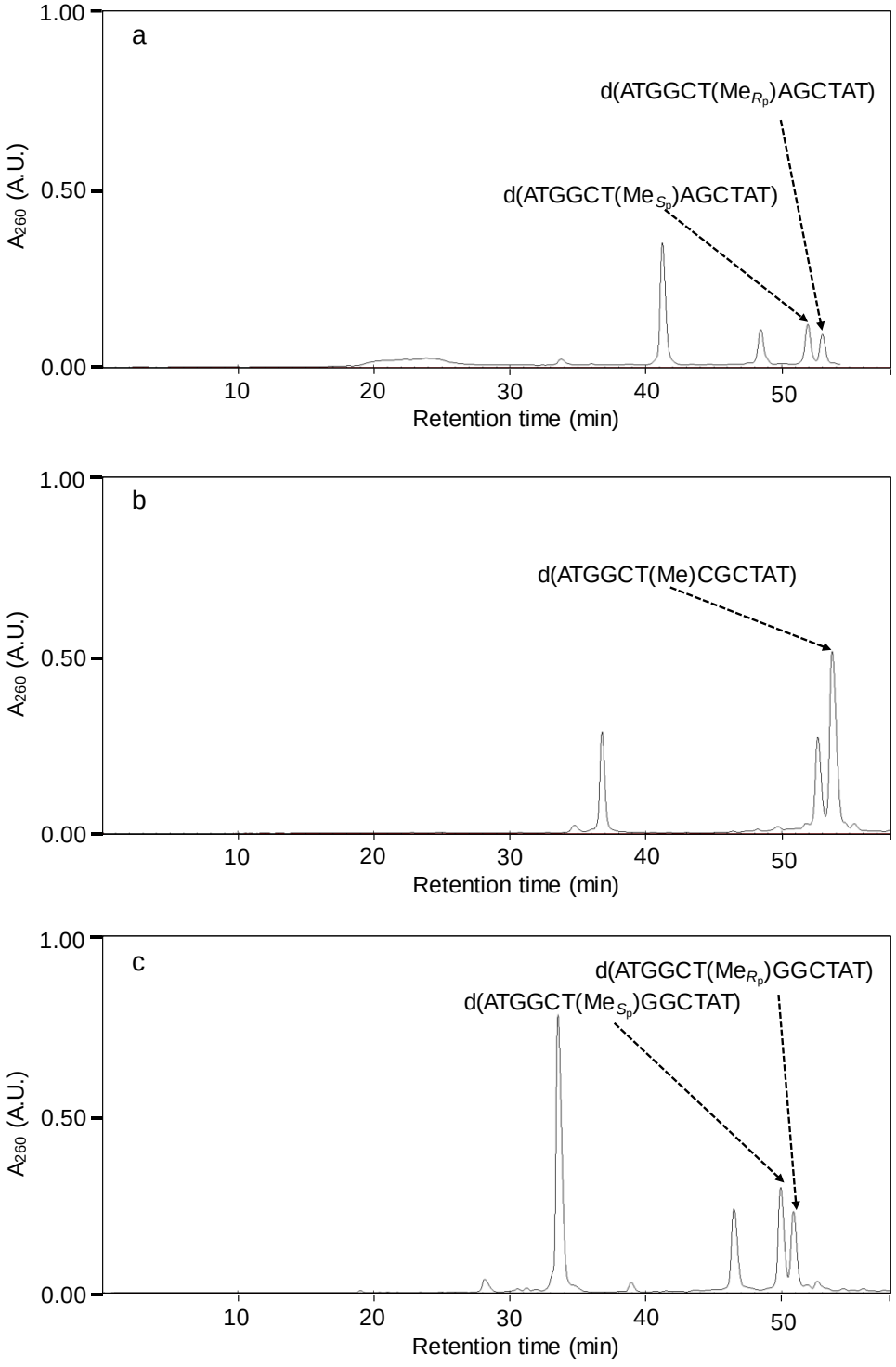
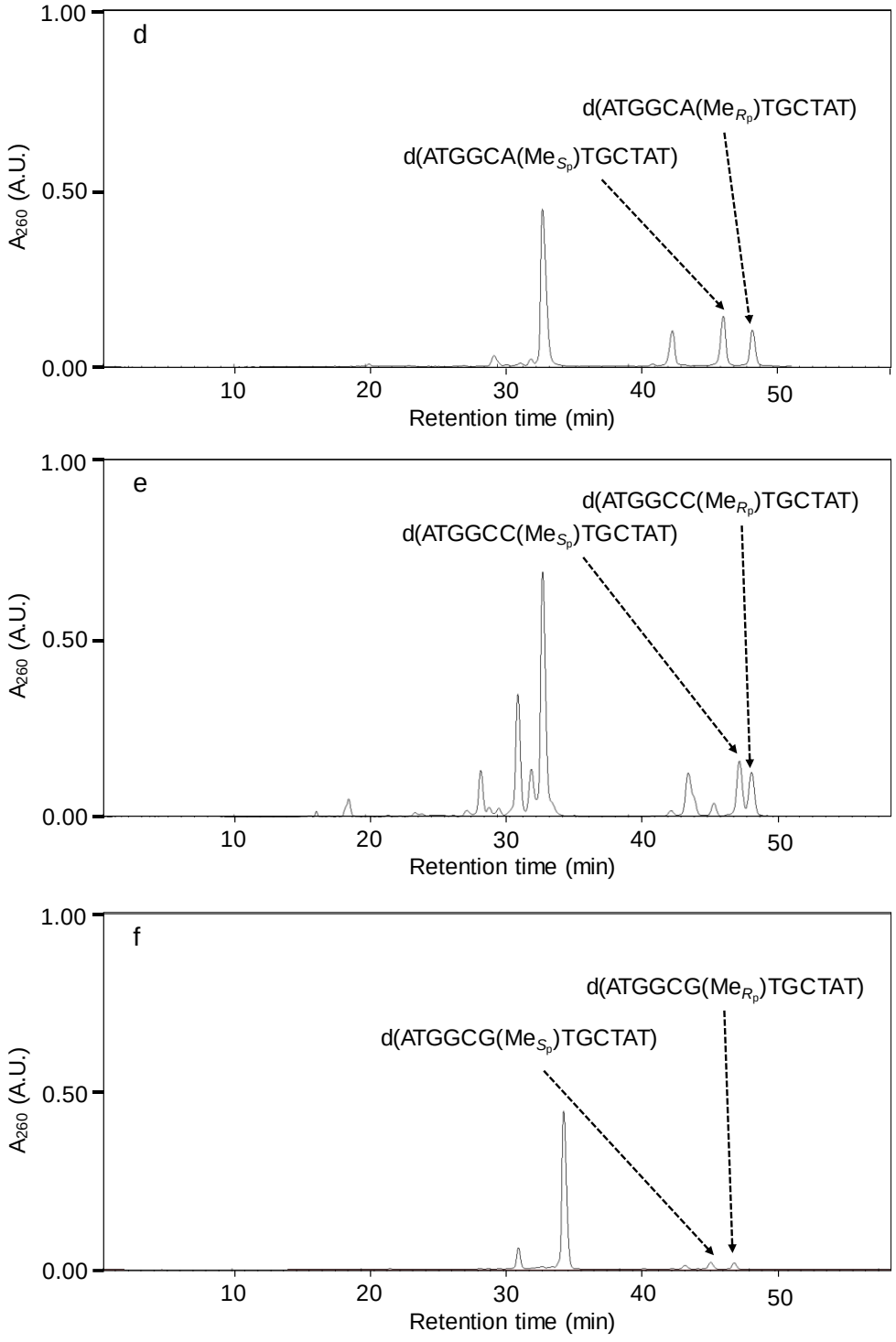
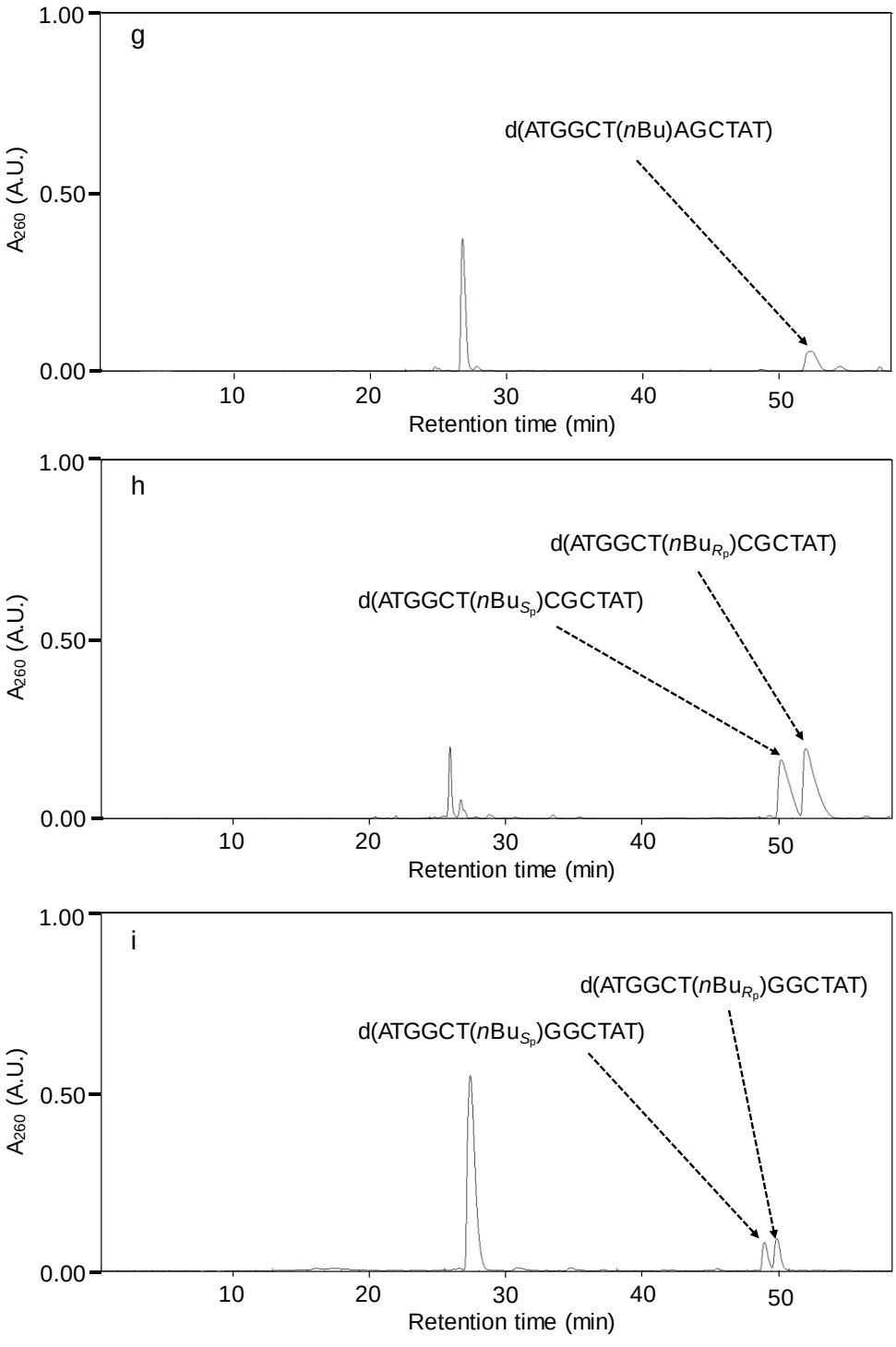


Figure 3-5. The ^{1}H NMR (400 MHz, $CDCl_3$, 25°C, top) and ^{31}P NMR (80 MHz, $CDCl_3$, 25ºC, bottom) spectra of N^4-AcdC-*n*-butylphosphoramidite.


a
d(ATGGCT(Me$_{R_p}$)AGCTAT)
d(ATGGCT(Me$_{S_p}$)AGCTAT)
b
d(ATGGCT(Me)CGCTAT)
c
d(ATGGCT(Me$_{R_p}$)GGCTAT)
d(ATGGCT(Me$_{S_p}$)GGCTAT)
A$_{260}$ (A.U.)
1.00
0.50
0.00
10
20
30
40
50
Retention time (min)


d
d(ATGGCA(Me$_{R_p}$)TGCTAT)
d(ATGGCA(Me$_{S_p}$)TGCTAT)
e
d(ATGGCC(Me$_{R_p}$)TGCTAT)
d(ATGGCC(Me$_{S_p}$)TGCTAT)
f
d(ATGGCG(Me$_{R_p}$)TGCTAT)
d(ATGGCG(Me$_{S_p}$)TGCTAT)
A$_{260}$ (A.U.)
1.00
0.50
0.00
10
20
30
40
50
Retention time (min)


g
d(ATGGCT(nBu)AGCTAT)
h
d(ATGGCT(nBu$_{R_p}$)CGCTAT)
d(ATGGCT(nBu$_{S_p}$)CGCTAT)
i
d(ATGGCT(nBu$_{R_p}$)GGCTAT)
d(ATGGCT(nBu$_{S_p}$)GGCTAT)
A$_{260}$ (A.U.)
1.00
0.50
0.00
10
20
30
40
50
Retention time (min)

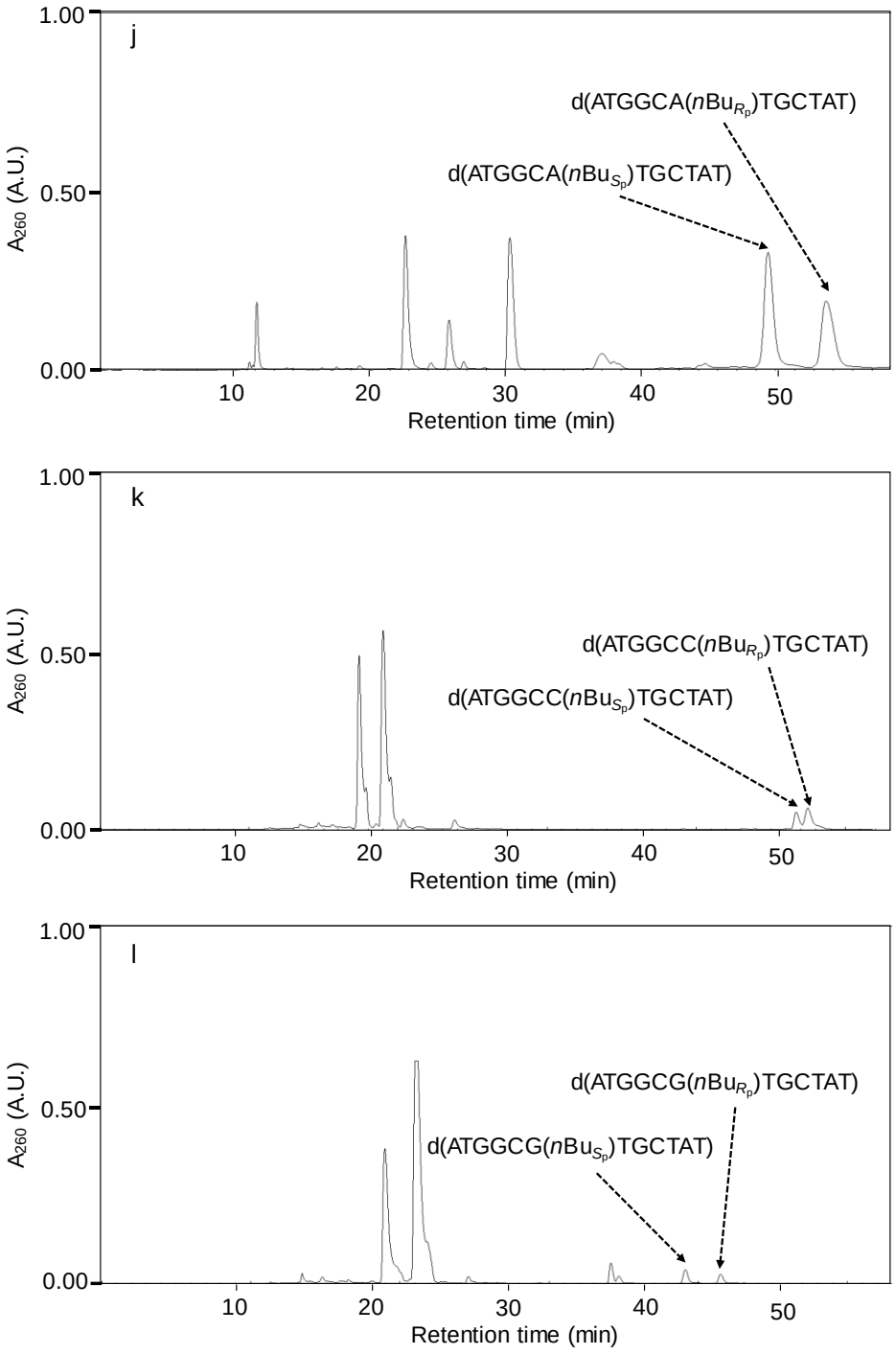


Figure 3-6. HPLC traces for the separations of the synthesized 12mer methylphosphotriester-bearing ODNs: (a) T(Me)A; (b) T(Me)C; (c) T(Me)G; (d) A(Me)T; (e) C(Me)T; (f) G(Me)T; (g) T(*n*Bu)A; (h) T(*n*Bu)C; (i) T(*n*Bu)G; (j) A(*n*Bu)T; (k) C(*n*Bu)T; (l) G(*n*Bu)T.

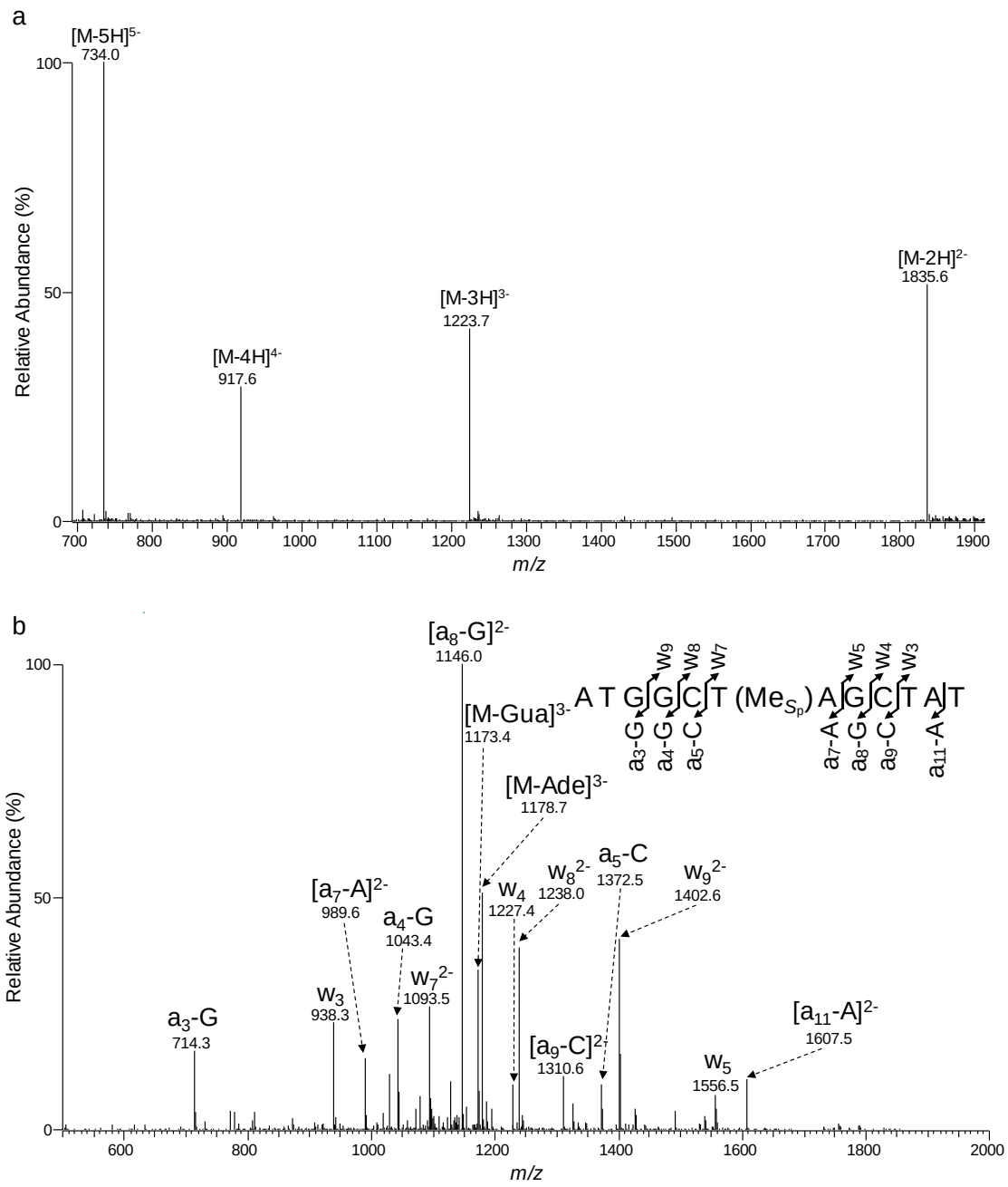


Figure 3-7. ESI-MS & MS/MS characterizations of d(ATGGCT(Me*Sp*)AGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1223.7).

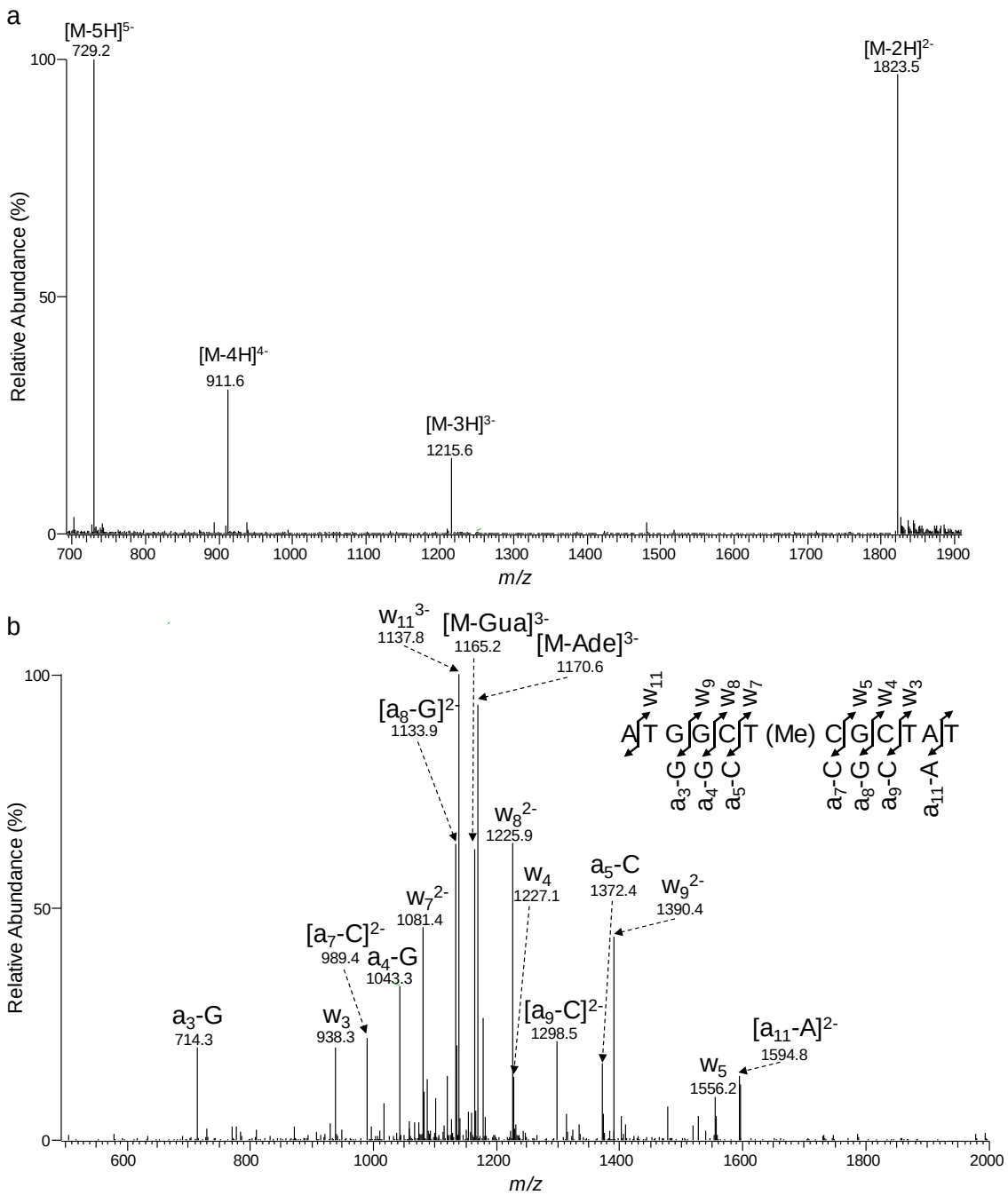


Figure 3-8. ESI-MS & MS/MS characterizations of d(ATGGCT(Me)CGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m/z* 1215.6).

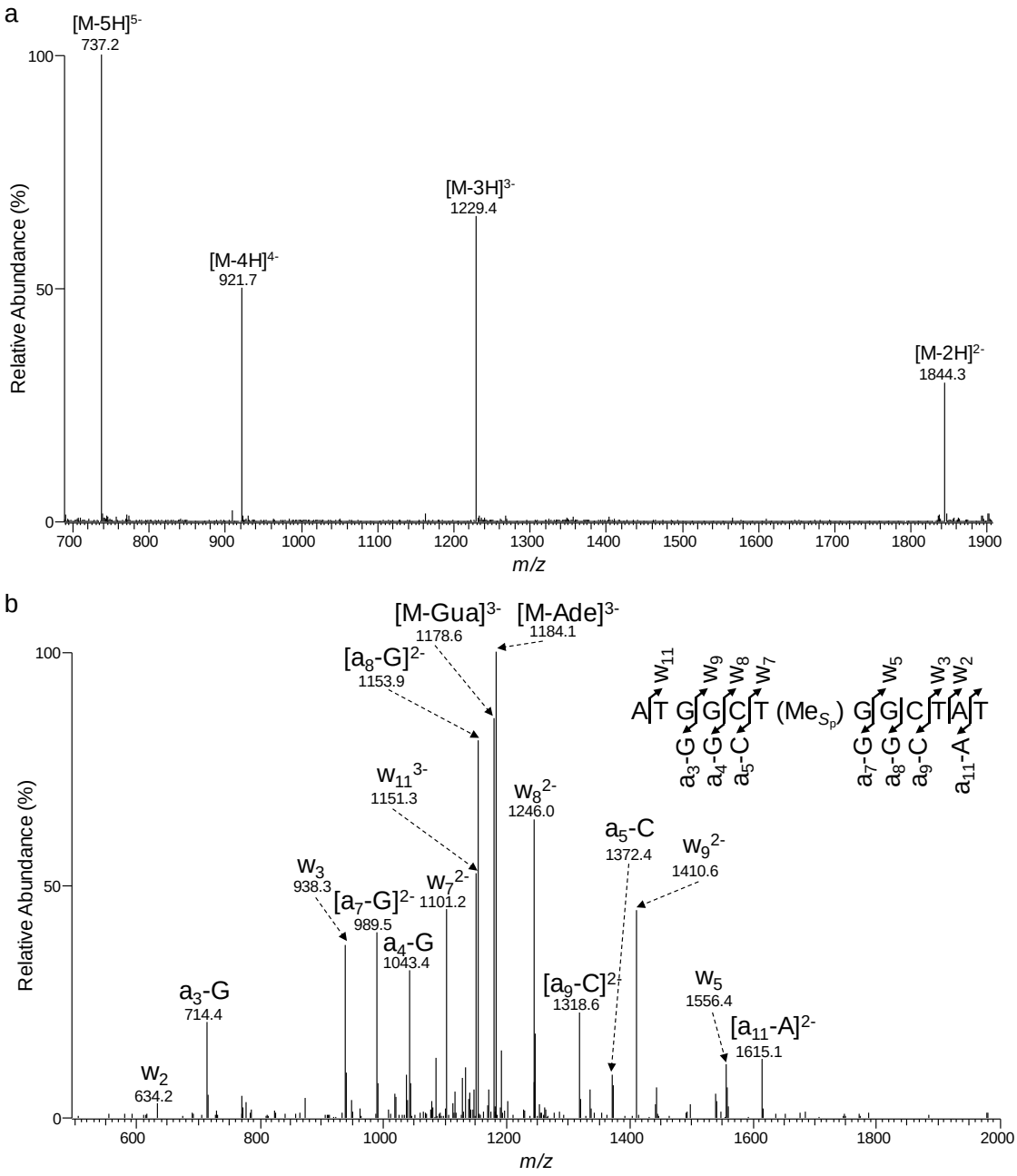


Figure 3-9. ESI-MS & MS/MS characterizations of d(ATGGCT(Me*Sp*)GGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1229.4).

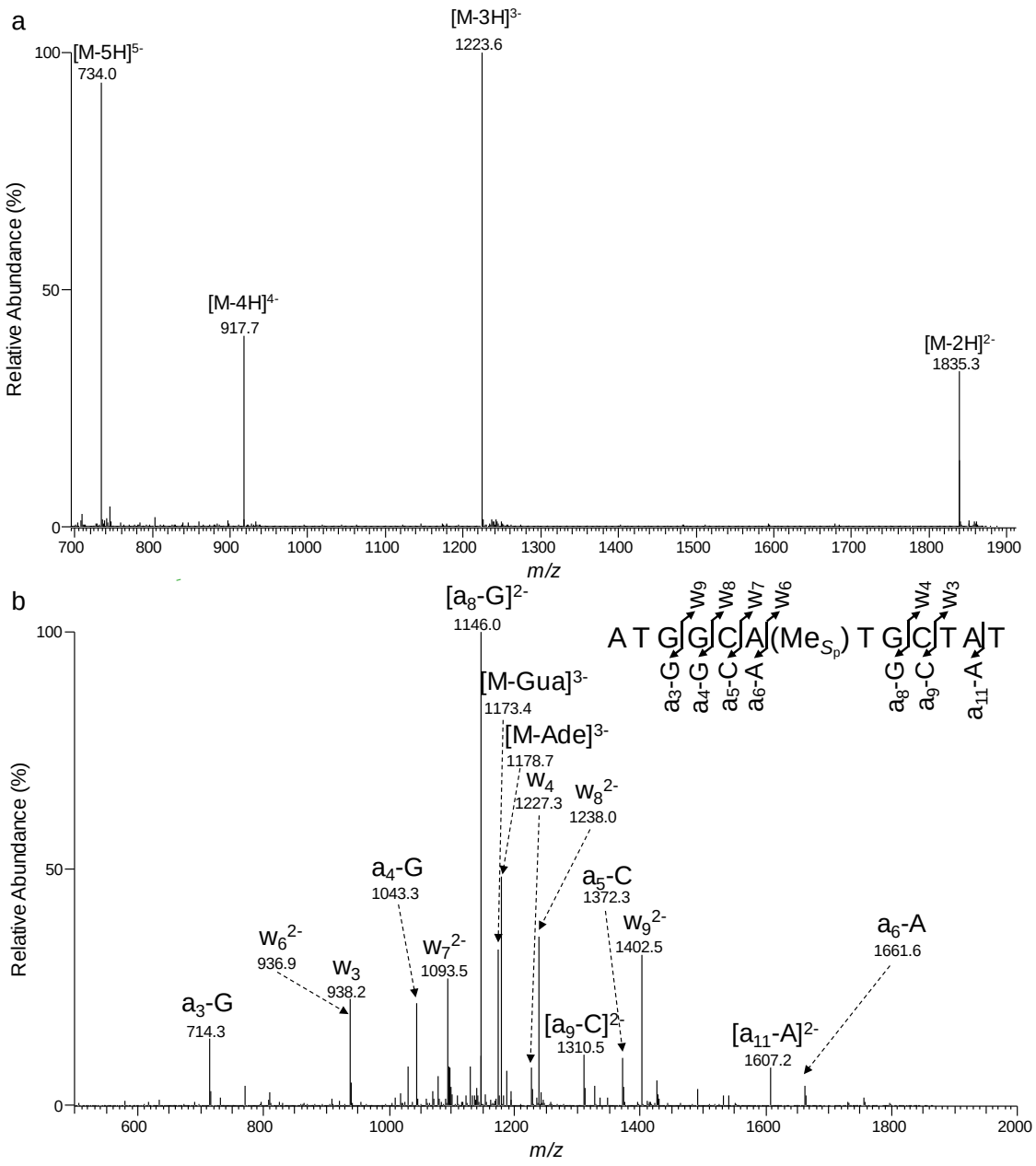


Figure 3-10. ESI-MS & MS/MS characterizations of d(ATGGCA(Me$_{Sp}$)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1223.6).

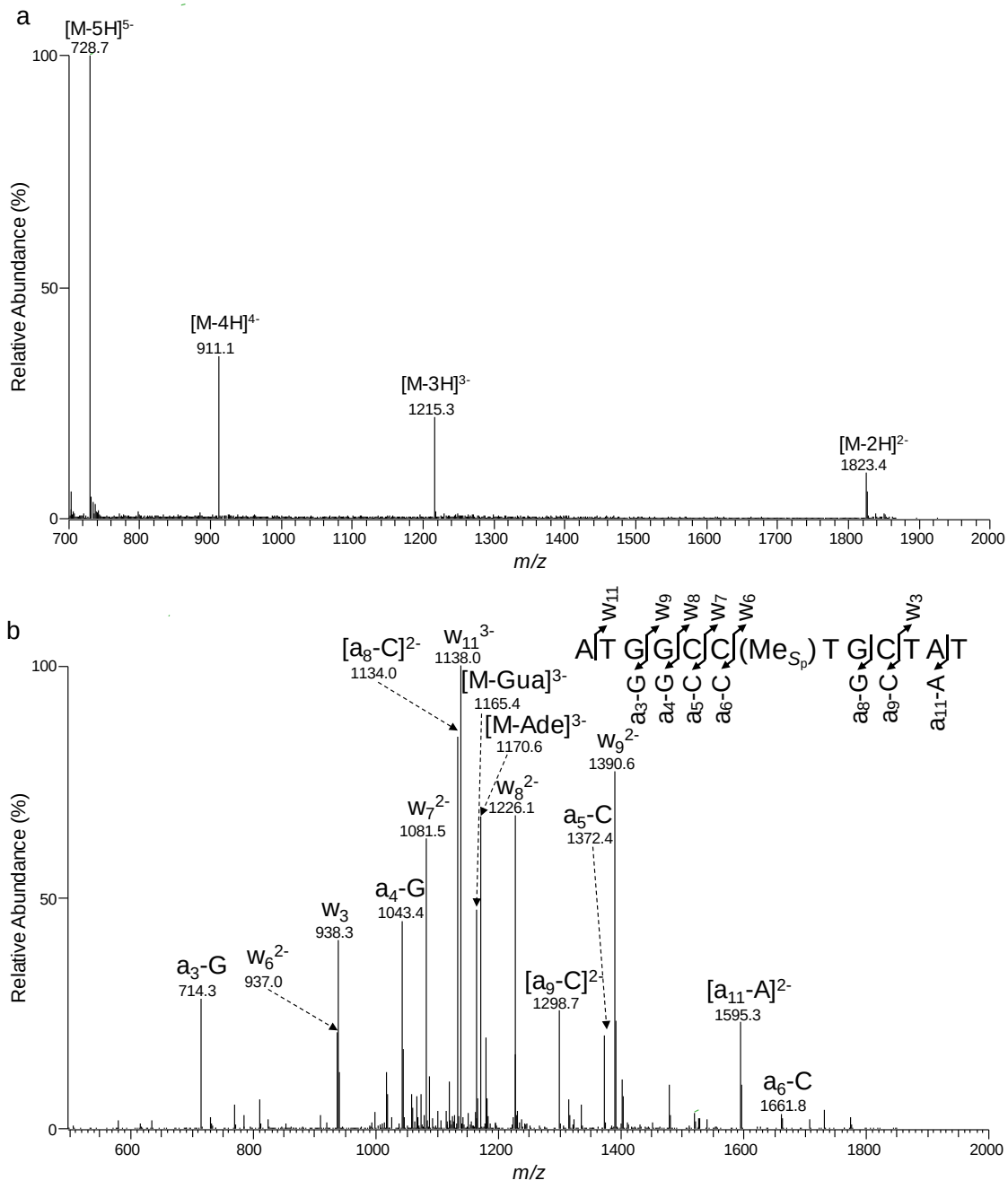


Figure 3-11. ESI-MS & MS/MS characterizations of d(ATGGCC(Me$_{Sp}$)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1215.3).

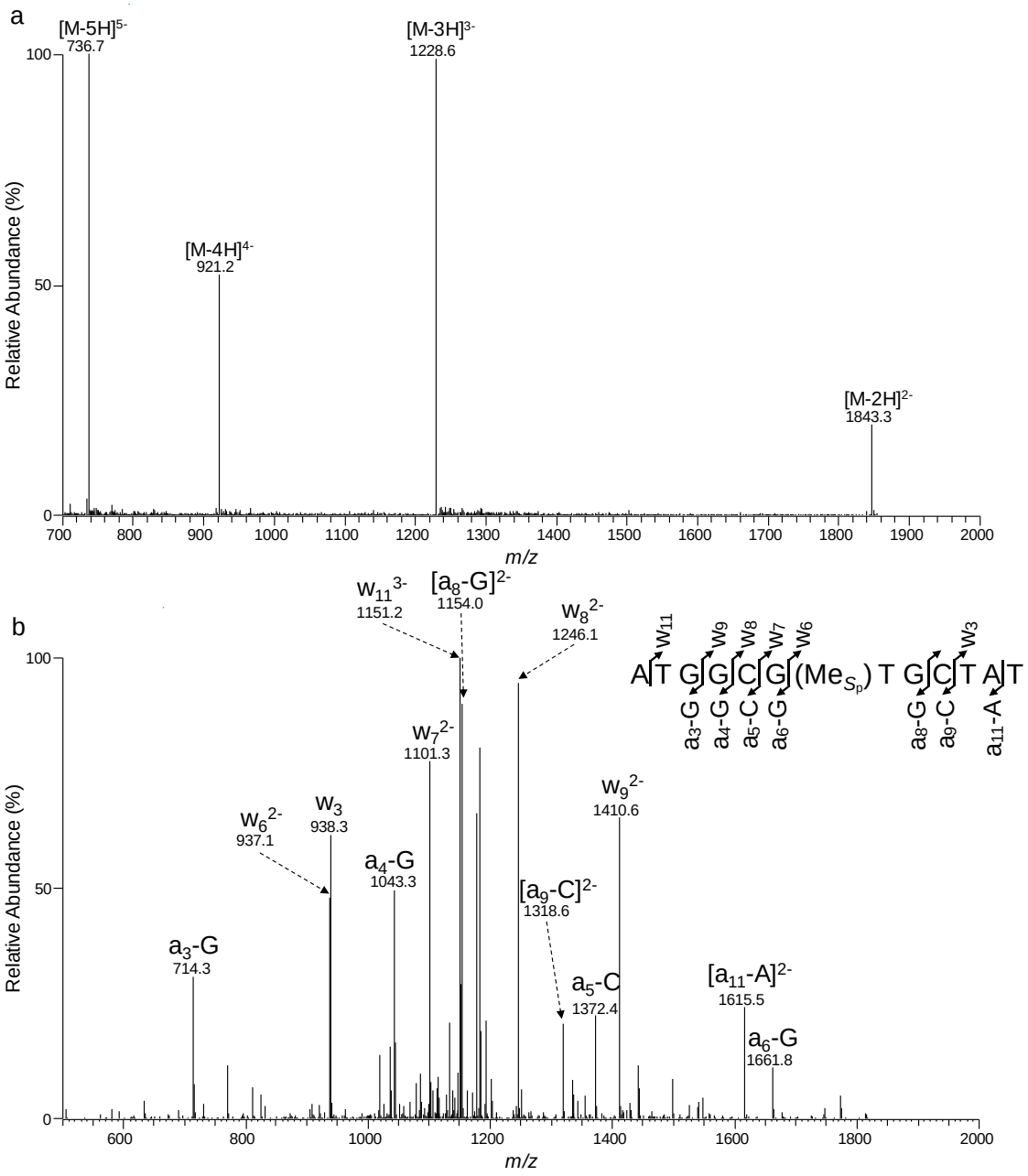


Figure 3-12. ESI-MS & MS/MS characterizations of d(ATGGCG(Me$_{Sp}$)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the [M-3H]$^{3-}$ ion (*m*/*z* 1228.6).

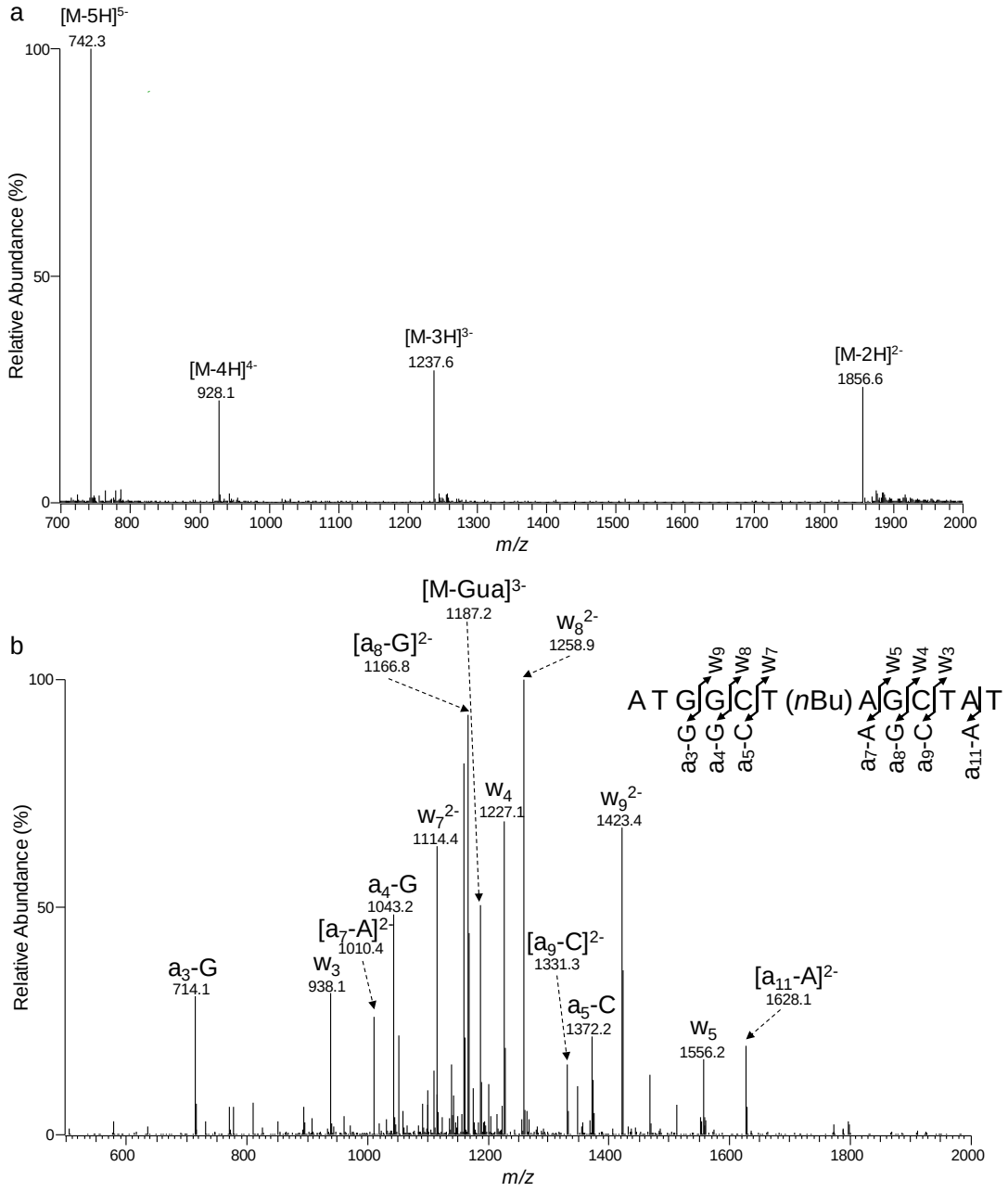


Figure 3-13. ESI-MS & MS/MS characterizations of d(ATGGCT(*n*Bu)AGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1237.6).

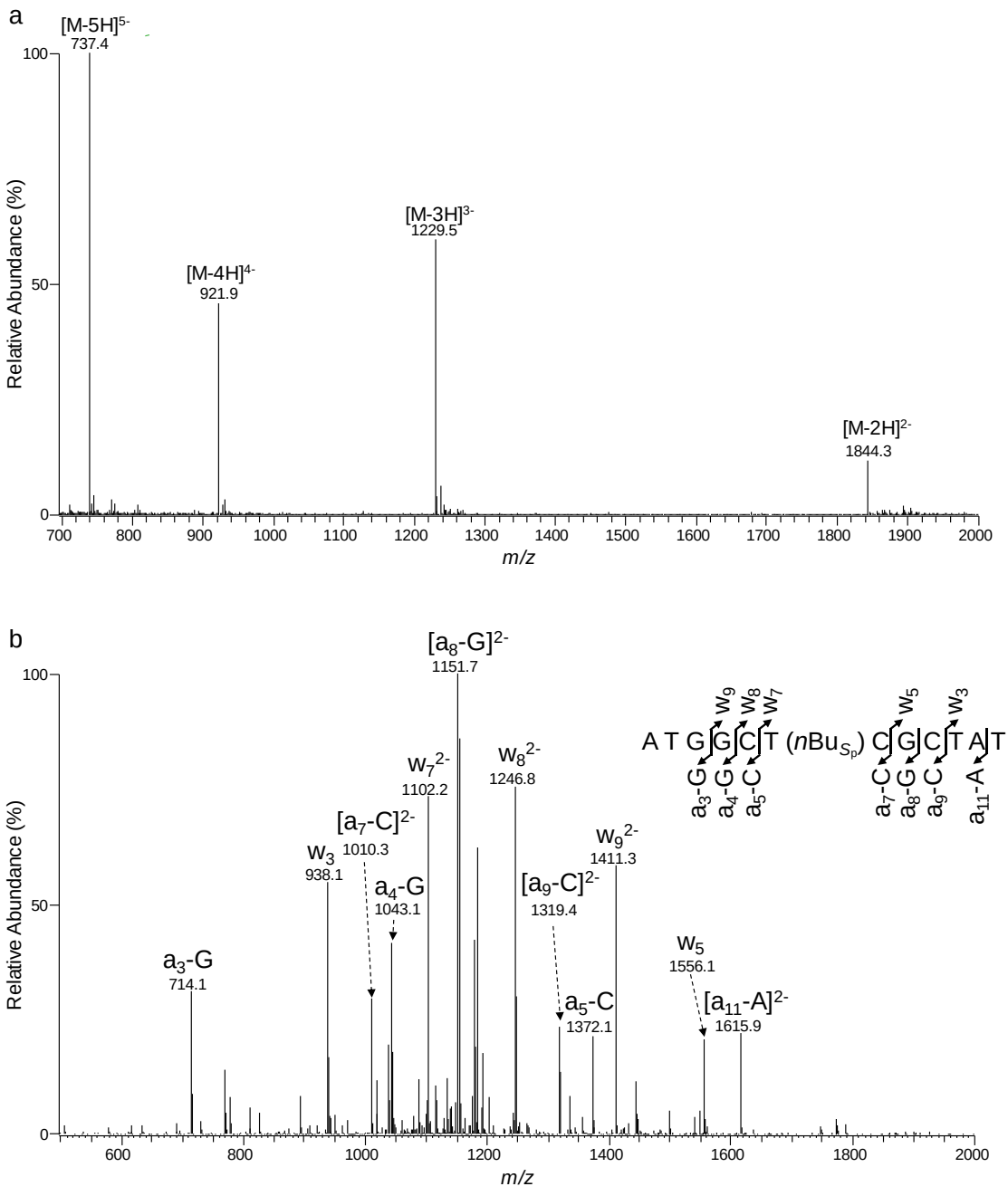


Figure 3-14. ESI-MS & MS/MS characterizations of d(ATGGCT(*n*Bu$_{Sp}$)CGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1229.5).

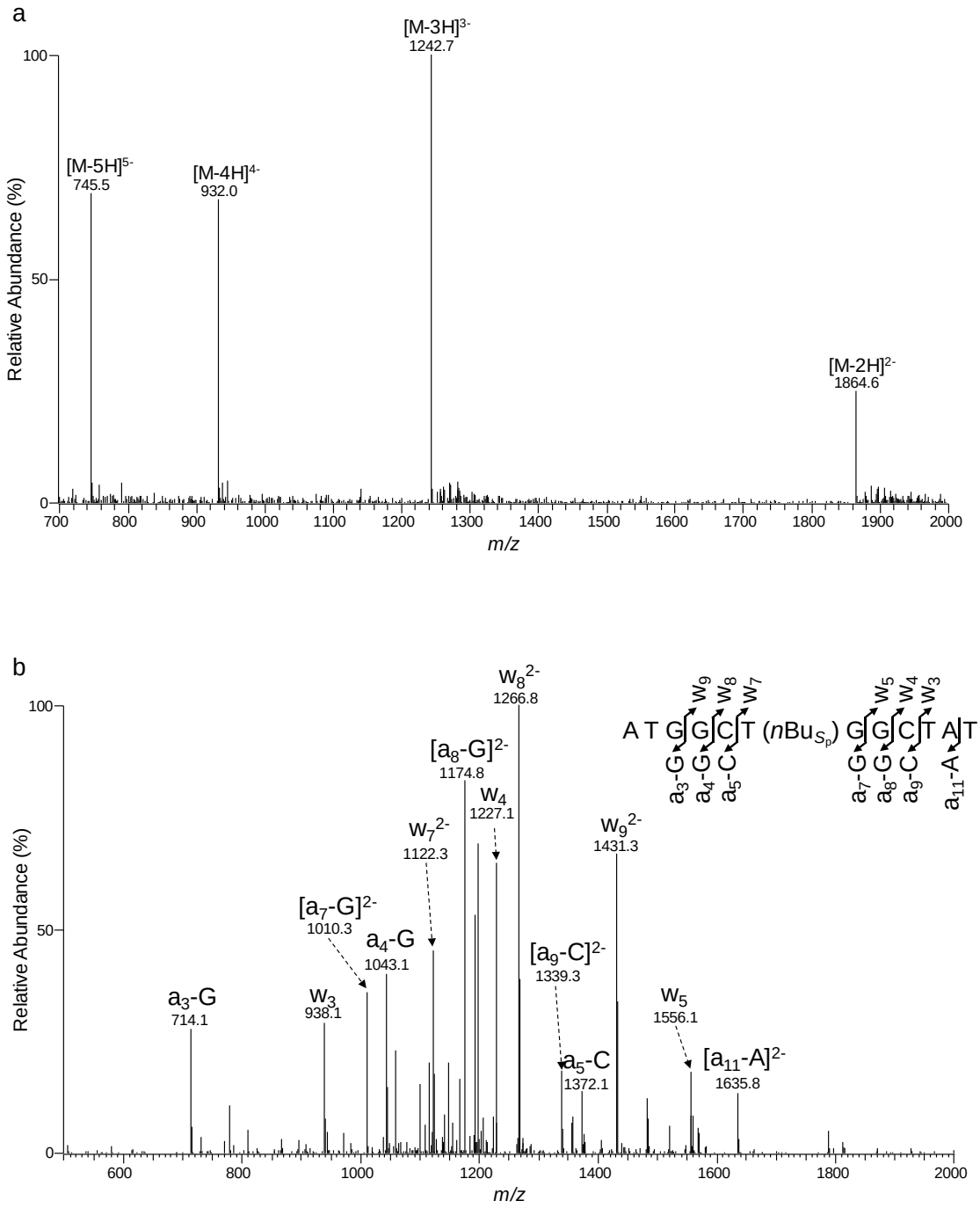


Figure 3-15. ESI-MS & MS/MS characterizations of d(ATGGCT(*n*Bu$_{Sp}$)GGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1242.7).

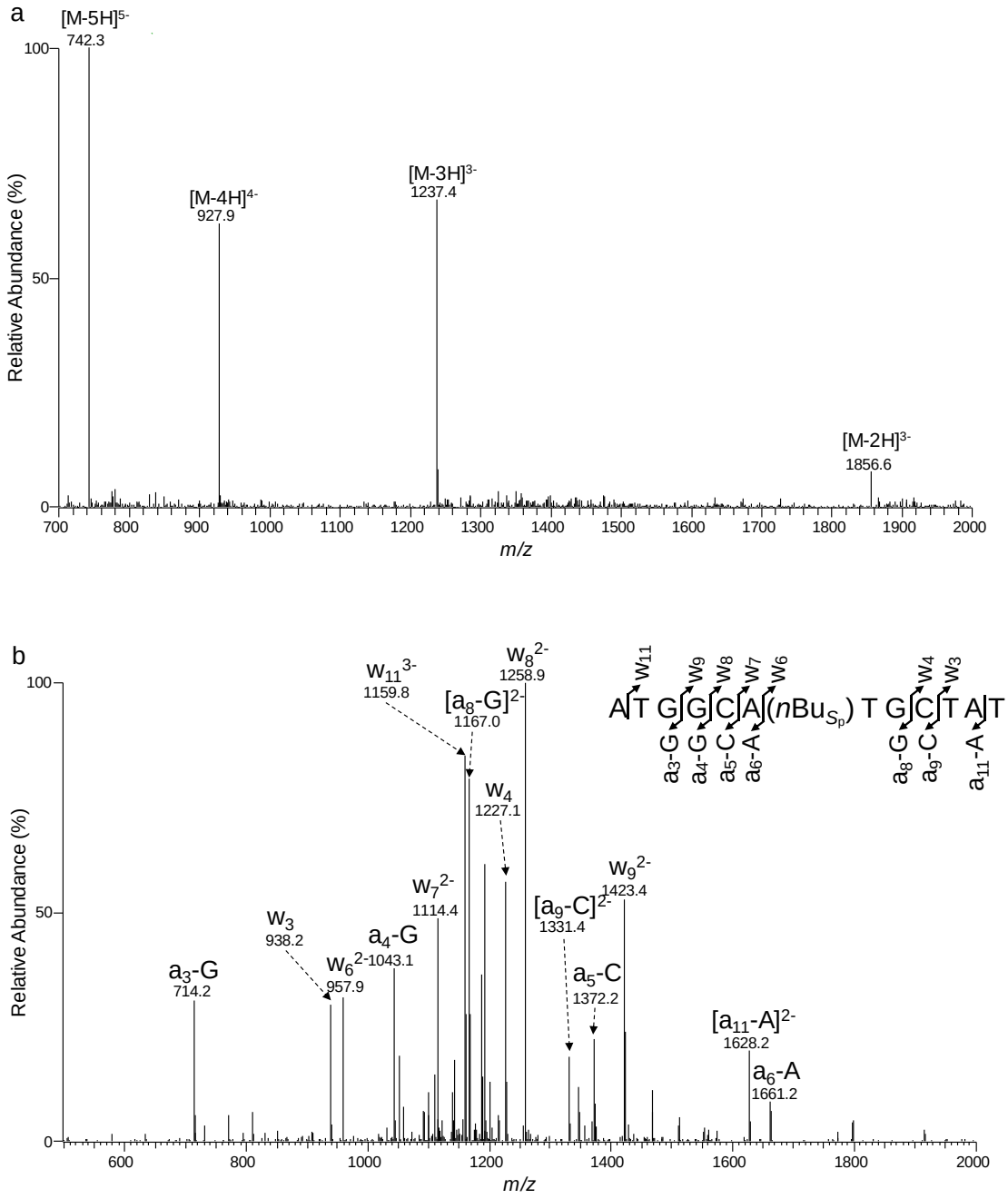


Figure 3-16. ESI-MS & MS/MS characterizations of d(ATGGCA(*n*Bu$_{Sp}$)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1237.4).

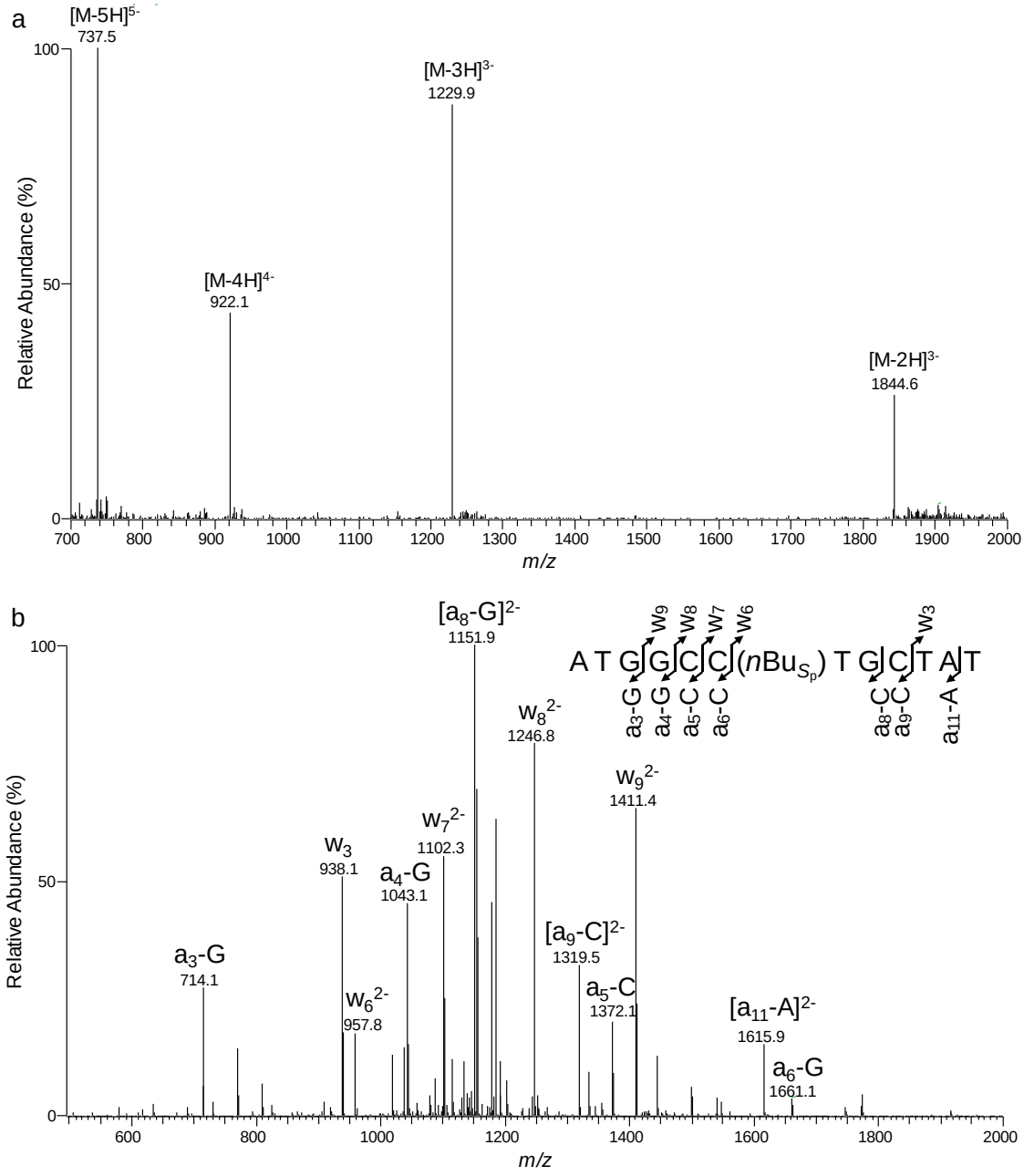


Figure 3-17. ESI-MS & MS/MS characterizations of d(ATGGCC(*n*Bu*Sp*)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z* 1229.9).

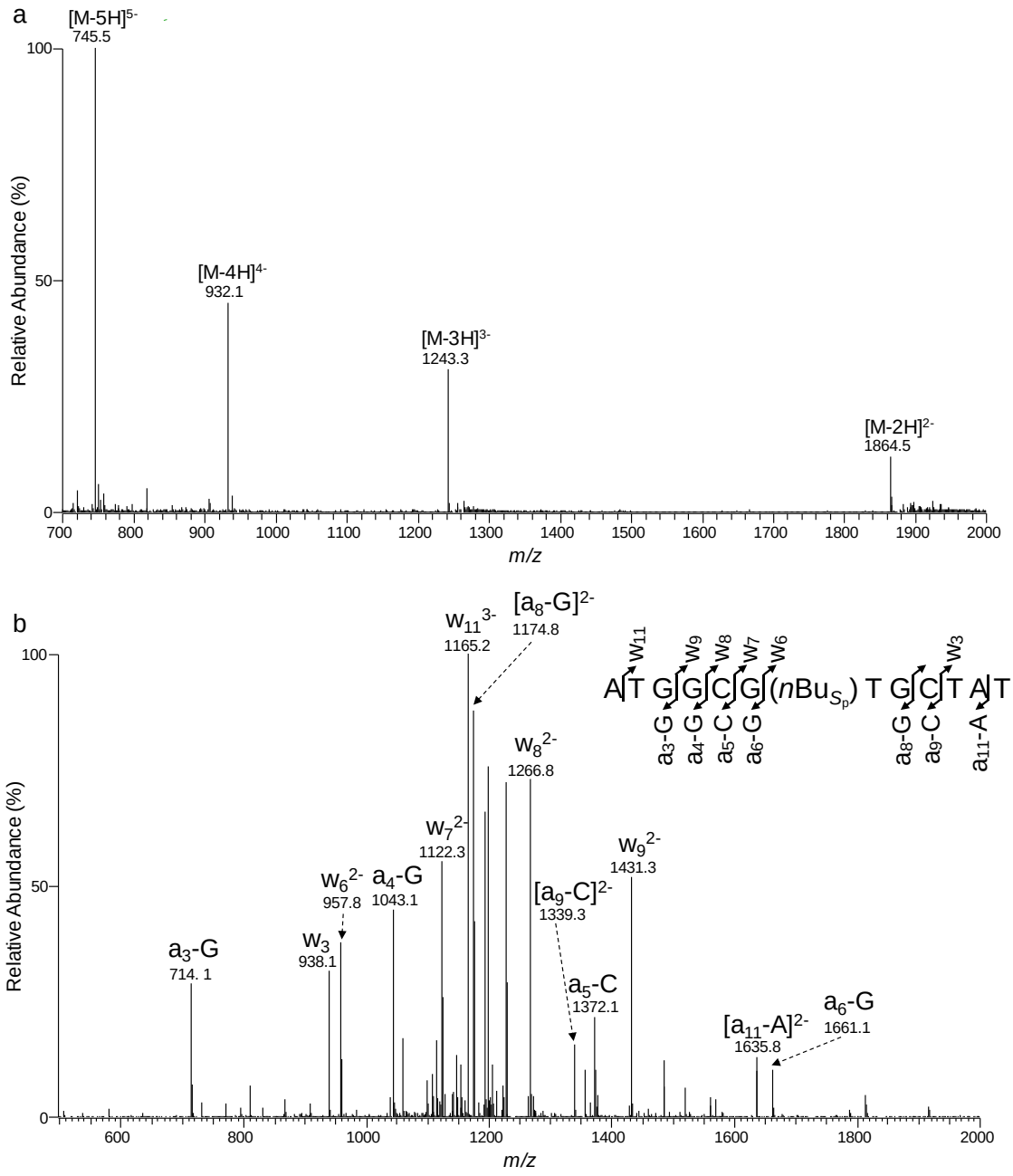


Figure 3-18. ESI-MS & MS/MS characterizations of d(ATGGCG(*n*Bu*Sp*)TGCTAT): (a) Negative-ion ESI-MS; (b) the product-ion spectrum of the $[M-3H]^{3-}$ ion (*m*/*z*1243.3).

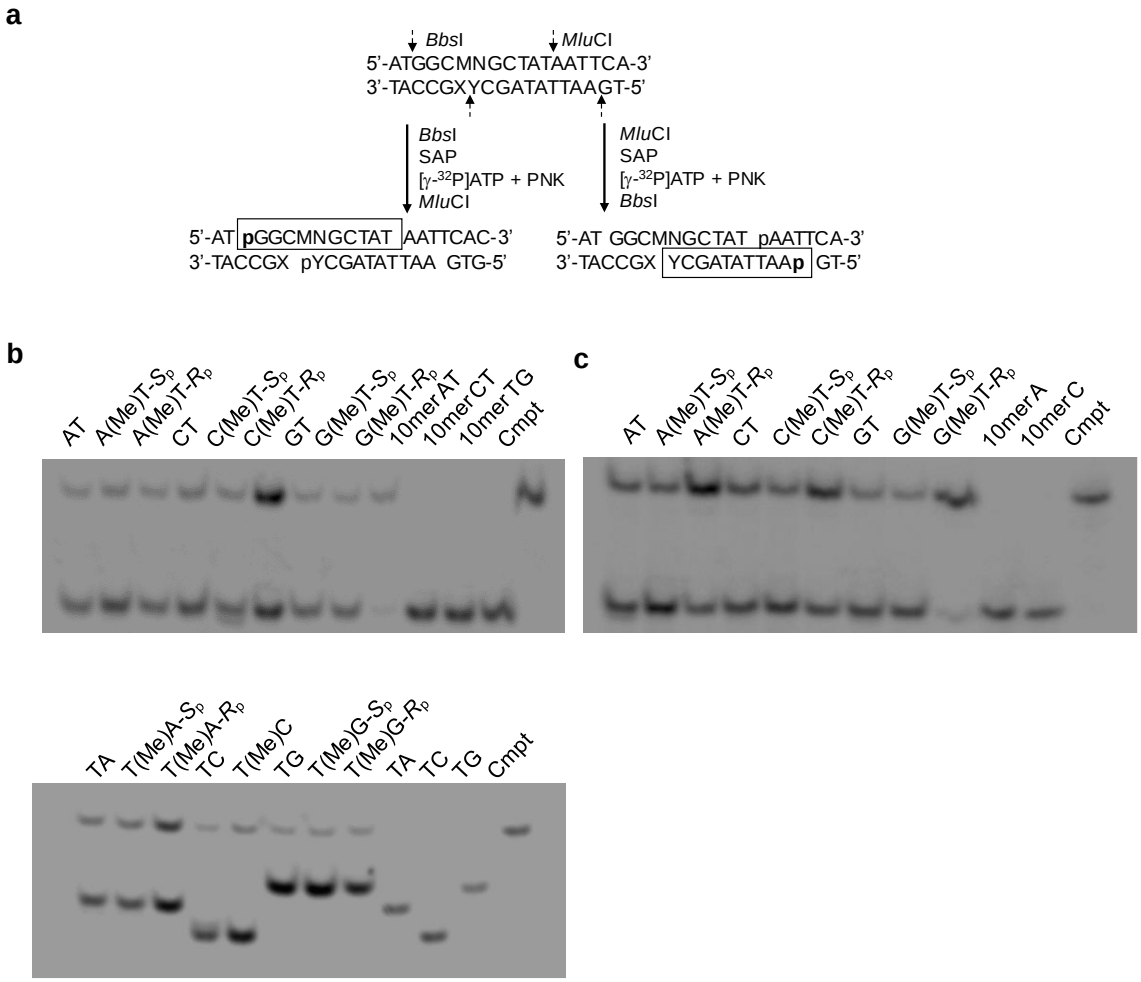


Figure 3-19. Restriction enzyme digestion and radiolabeling, followed by native PAGE (30%) analysis for quantifying the bypass efficiencies of Me-PTE lesions located in different sequence contexts in wild-type AB1157 *E. coli* cells. (a) Selective labelling of the original lesion-containing strand and its complementary strand via sequential restriction digestion. 'SAP' and 'PNK' designate shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. 'MN' in the sequence denotes the site where the TX or XT dinucleotide flanking the Me-PTE lesions were incorporated. (b) Gel image showing the 13 mer (competitor genome) and 10 mer (control or lesion-containing genome) digestion products formed from the original strand, where 10 mer AT, 10 mer CT and 10 mer TG designate [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with MN being AT, CT and TG, respectively. (c) Gel image showing the 13 mer (competitor genome) and 10 mer (control or lesion-containing genome) digestion products formed from the opposite strand, where 10 mer A and 10 mer C indicate the [5′-^{32}P]-labeled standard ODNs 5′-AATTATAGCY-3′, with 'Y' being A and C, respectively.

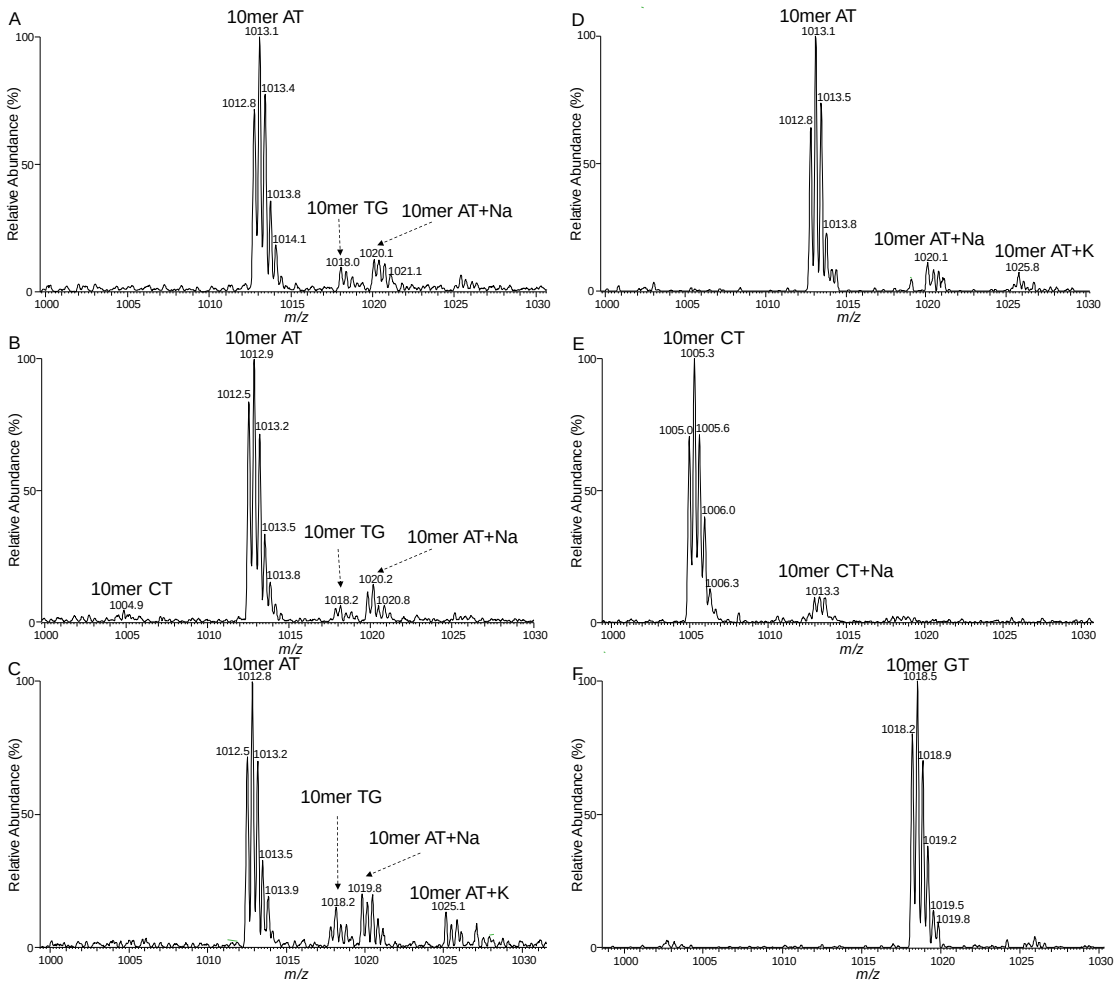
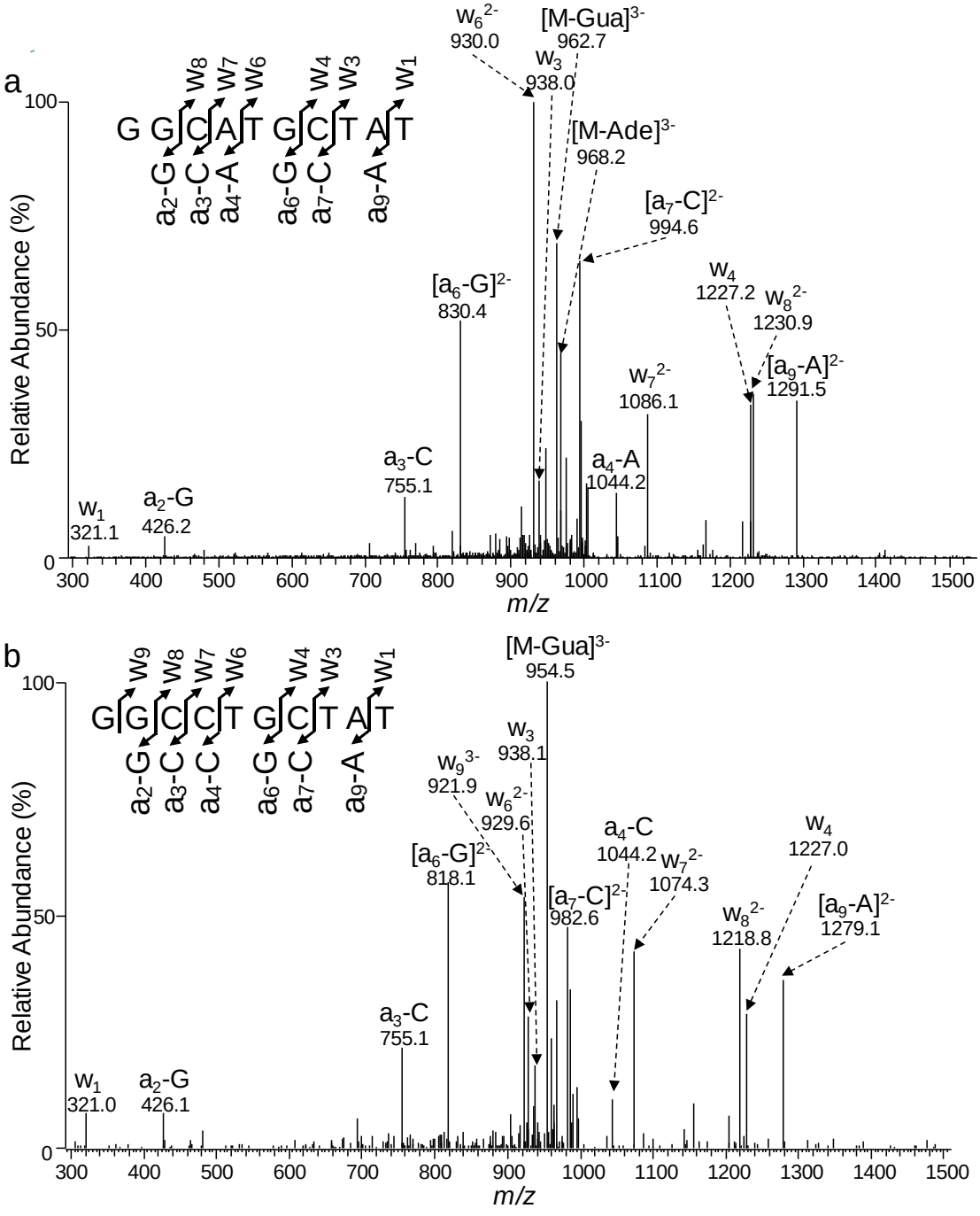


Figure 3-20. Higher-resolution "zoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of single-stranded M13 genomes harboring a site-specifically incorporated (a) S_p-A(Me)T, (b) S_p-C(Me)T, (c) S_p-G(Me)T, (d) R_p-A(Me)T, (e) R_p-C(Me)T, and (f) R_p-G(Me)T in wild-type AB1157 cells. Displayed in (a)-(f) are the $[M-3H]^{3-}$ ions for the lesion-containing strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for Me-PTE lesions are shown in Figures 3-21.


a
G G C A T G C T A T
w8 w7 w6 w4 w3 w1
a2-G a3-C a4-A a6-G a7-C a9-A
w6 2- 930.0
[M-Gua]3- 962.7
w3 938.0
[M-Ade]3- 968.2
[a7-C]2- 994.6
[a6-G]2- 830.4
w4 1227.2
w8 2- 1230.9
[a9-A]2- 1291.5
w7 2- 1086.1
a3-C 755.1
a4-A 1044.2
w1 321.1
a2-G 426.2
Relative Abundance (%)
m/z
b
G G C C T G C T A T
w9 w8 w7 w6 w4 w3 w1
a2-G a3-C a4-C a6-G a7-C a9-A
[M-Gua]3- 954.5
w3 938.1
w9 3- 921.9
w6 2- 929.6
a4-C 1044.2
w4 1227.0
[a6-G]2- 818.1
w7 2- 1074.3
[a7-C]2- 982.6
w8 2- 1218.8
[a9-A]2- 1279.1
a3-C 755.1
w1 321.0
a2-G 426.1
Relative Abundance (%)
m/z

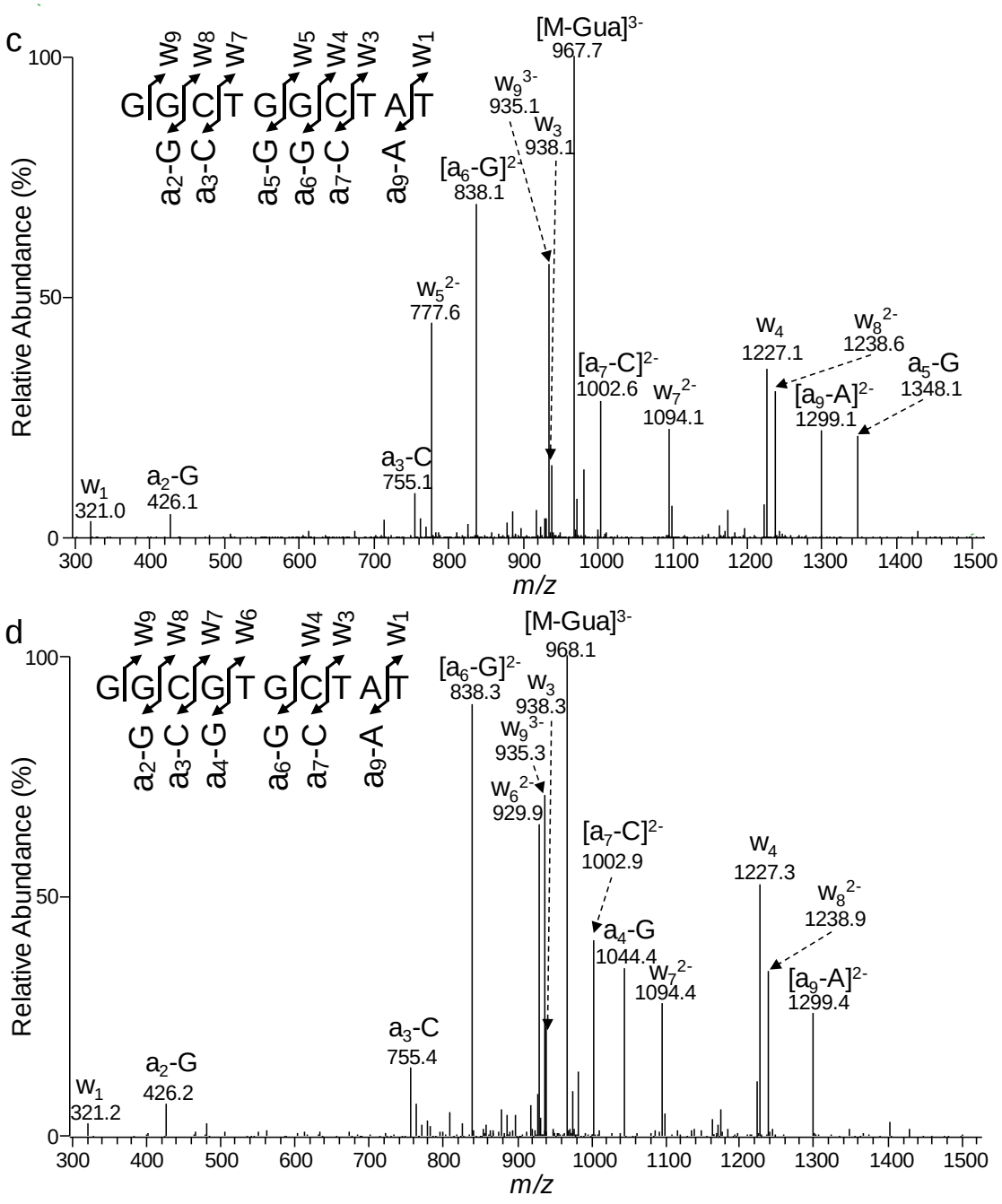


Figure 3-21. LC-MS and MS/MS for the identification of restriction fragments of PCR products. MS/MS for the $[M-3H]^{3-}$ ions of (a) 10 mer AT, (b) 10 mer CT, (c) 10 mer TG and (d) 10 mer GT.

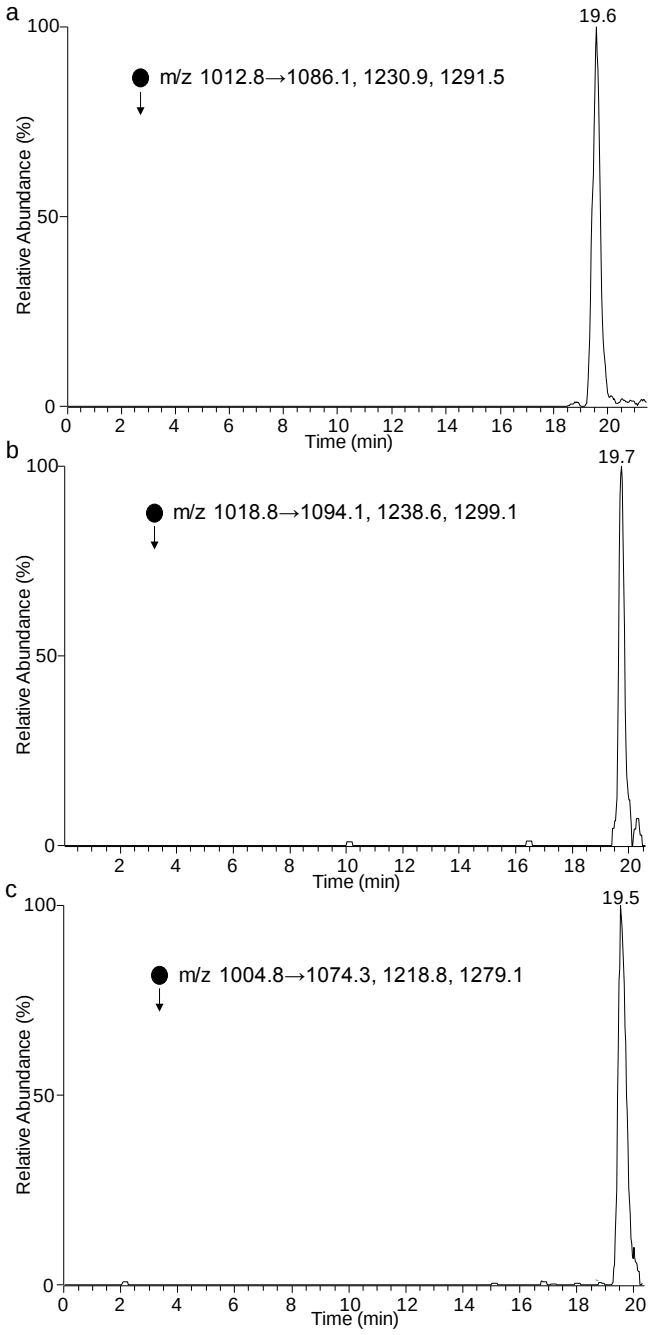
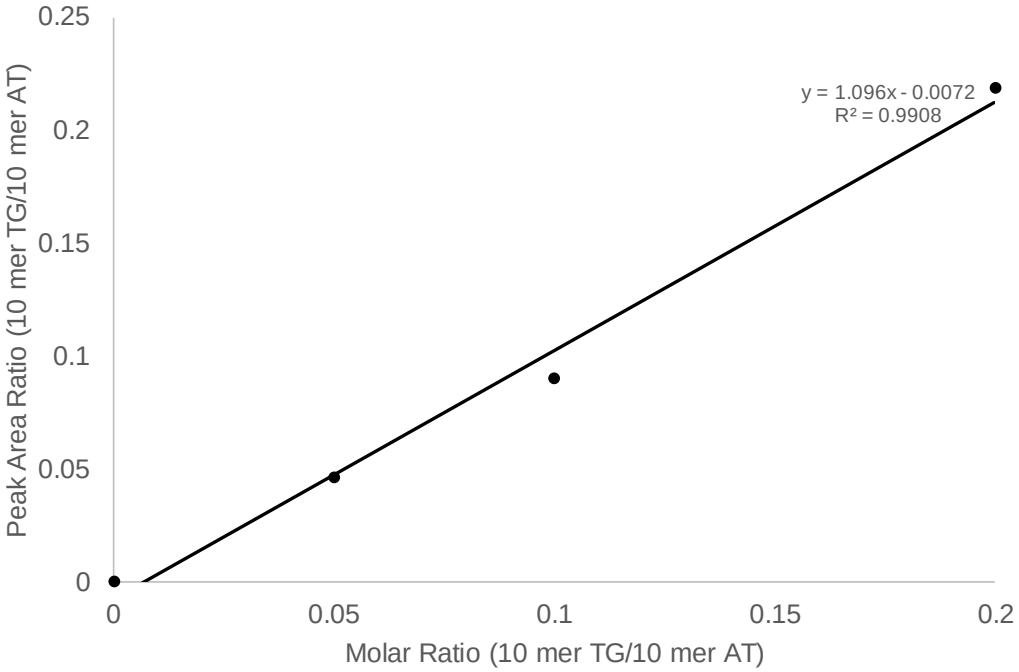


Figure 3-22. LC-MS/MS for monitoring the restriction fragments of interest without mutation or with a XT→AT and XT→TG mutations at the original guanine portion of the lesion [i.e. d(GGCATGCTAT) and d(GGCTGGCTAT)]. Shown in (a) and (b) and (c) are the SICs for the formation of indicated fragment ions (i.e. w_7^{2-}, w_8^{2-} and $[a_9\text{-A}]^{2-}$) of these three ODNs.

a



b

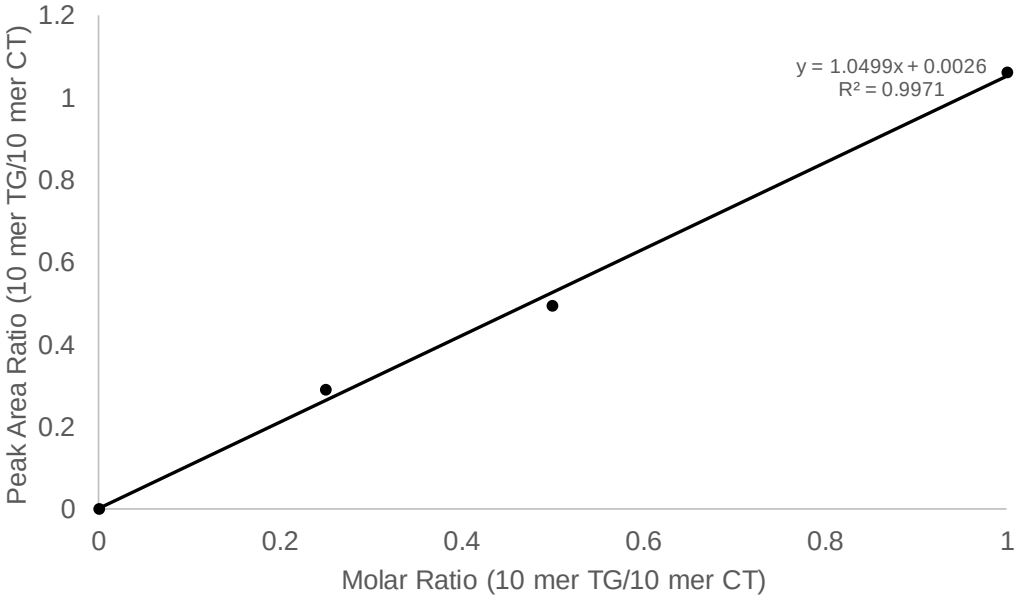


Figure 3-23. Calibration curves for quantification the molar ratio between (a) 10mer TG/10mer AT and (b) 10mer TG/10mer CT.

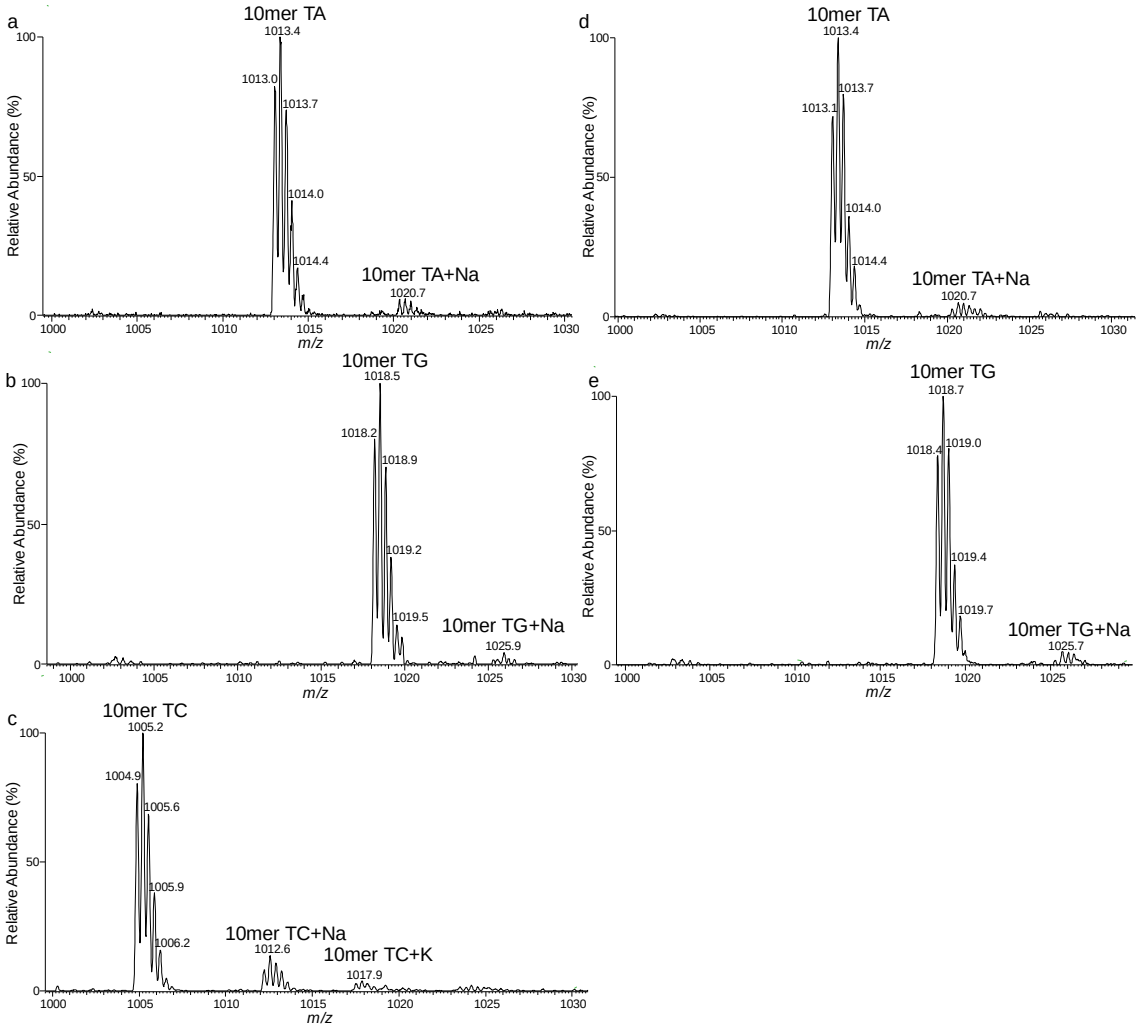


Figure 3-24. Higher-resolution "zoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of single-stranded M13 genomes harboring a site-specifically inserted (a) S_p-T(Me)A, (b) S_p-T(Me)G, (c) T(Me)C, (d) R_p-T(Me)A, and (e) R_p-T(Me)G in wild-type AB1157 cells. Displayed in (a)-(e) are the $[M-3H]^{3-}$ ions for the lesion-containing strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for Me-PTE lesions are shown in Figures 3-25.

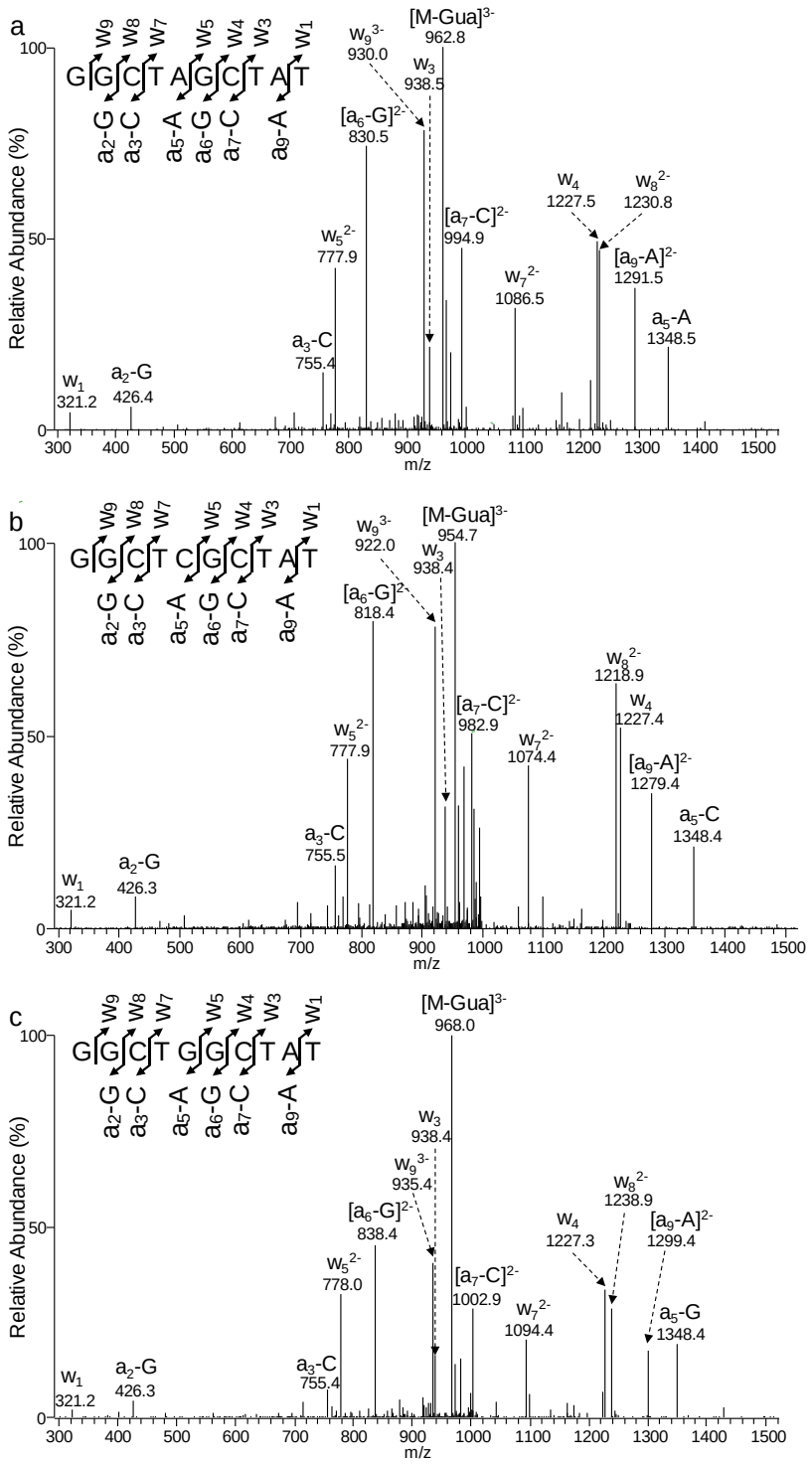


Figure 3-25. LC-MS and MS/MS for the identification of restriction fragments of PCR products. MS/MS for the $[M-3H]^{3-}$ ions of (a) 10 mer TA, (b) 10 mer TC, and (c) 10 mer TG.

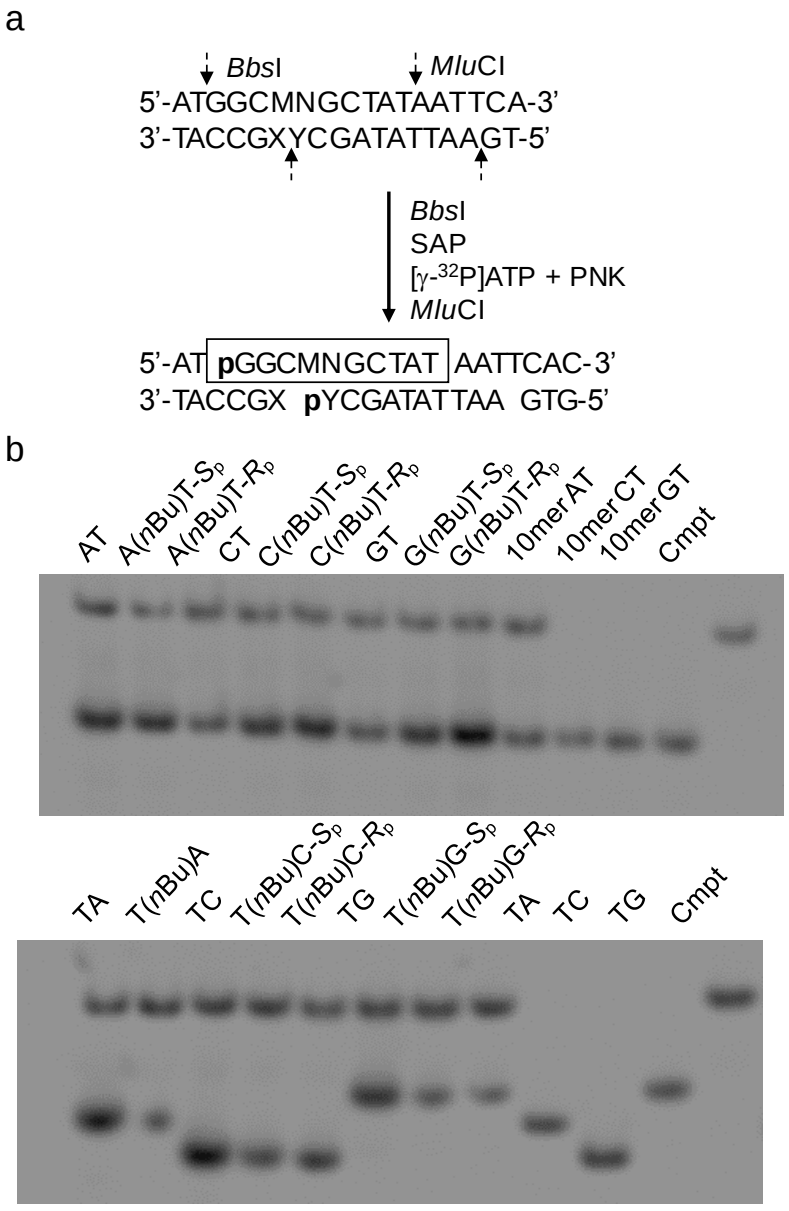


Figure 3-26. Native PAGE (30%) for monitoring the bypass efficiencies of *n*Bu-PTEs in wild-type AB1157 *E. coli* cells. (a) Selective labelling of original lesion-containing strand and its complementary strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where AT, CT, GT, TA, TC and TG represent the [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with 'MN' being AT, CT, GT, TA, TC and TG, respectively.

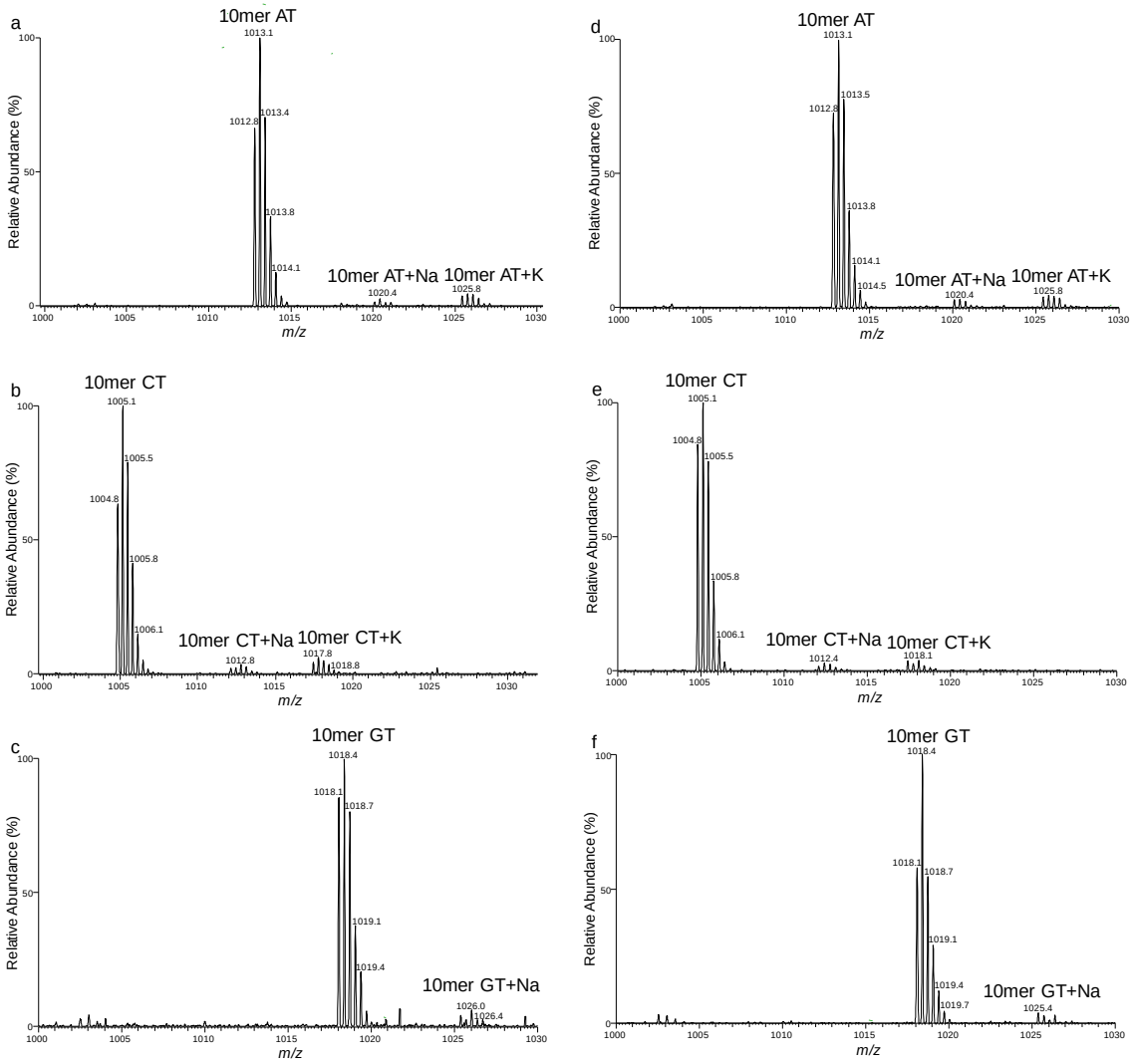


Figure 3-27. Higher-resolution "ultrazoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of single-stranded M13 genome harboring a site-specifically inserted (a) S_p-A(*n*Bu)T; (b) S_p- C(*n*Bu)T; (c) S_p- G(*n*Bu)T; (d) R_p- A(*n*Bu)T; (e) R_p- C(*n*Bu)T and (f) R_p- G(*n*Bu)T in WT cells. Displayed in (a)-(f) are the $[M-3H]^{3-}$ ions for the lesion-containing strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for *n*Bu-PTE lesions are shown in Figures 3-21.

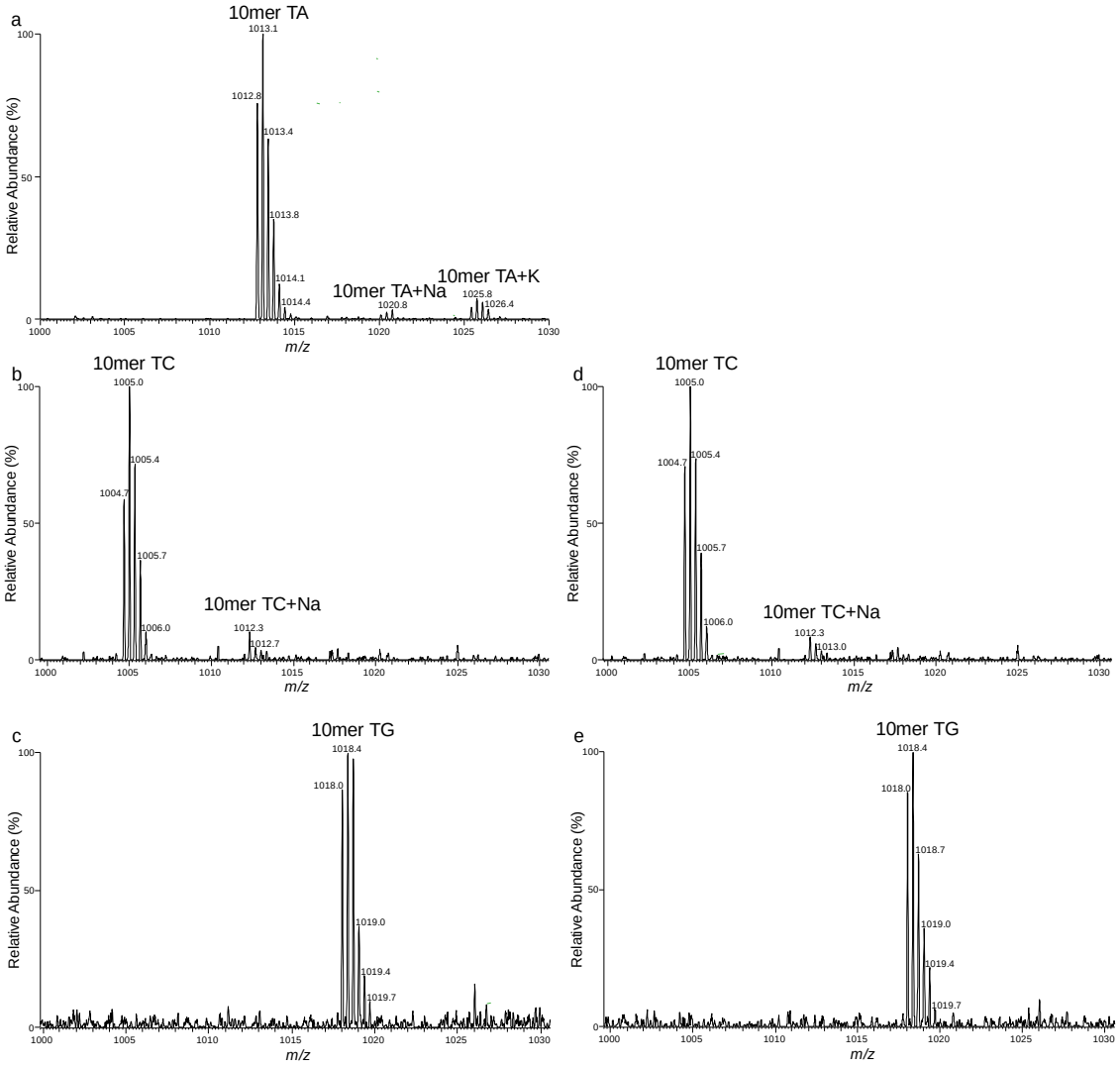


Figure 3-28. Higher-resolution "ultrazoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of single-stranded M13 genomes harboring a site-specifically inserted (a) T(*n*Bu)A; (b) S_p- T(*n*Bu)C; (c) S_p- T(*n*Bu)G; (d) R_p- T(*n*Bu)C and (e) R_p- T(*n*Bu)G in WT cells. Displayed in (a)-(e) are the $[M-3H]^{3-}$ ions for the lesion-containing strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for *n*Bu-PTE lesions are shown in Figures 3-25.

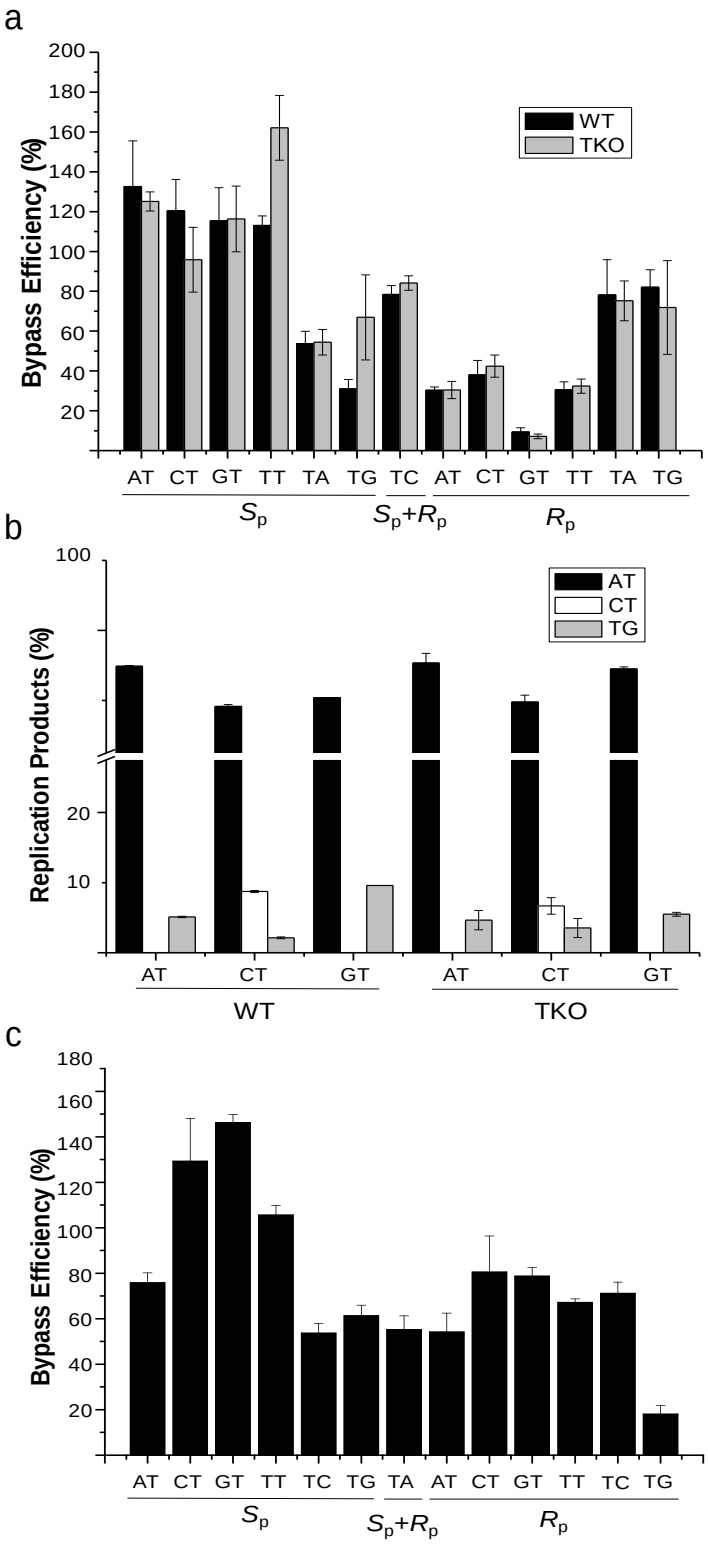
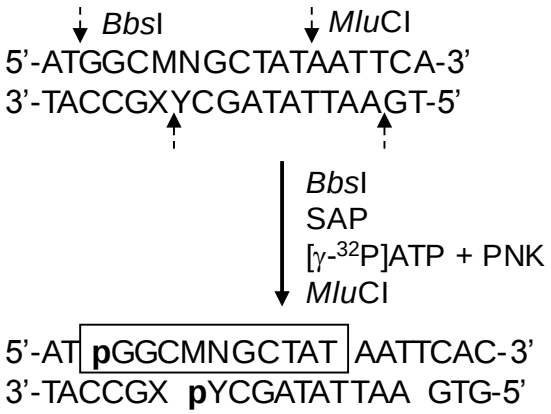
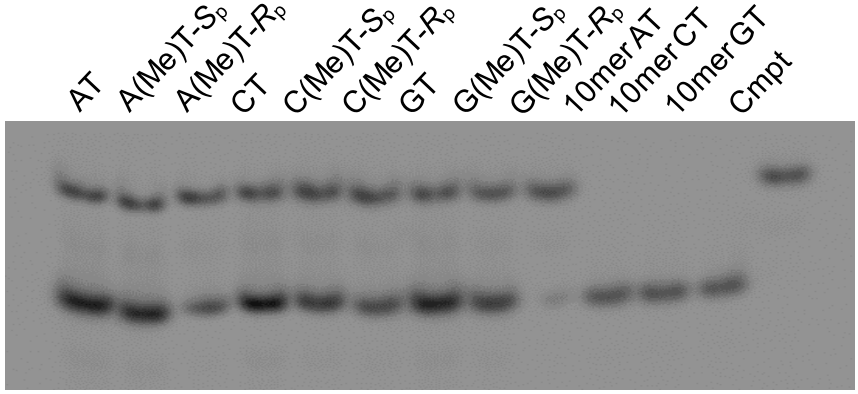

a
Bypass Efficiency (%)
200
180
160
140
120
100
80
60
40
20
WT
TKO
AT CT GT TT TA TG TC AT CT GT TT TA TG
S_p
S_p+R_p
R_p
b
Replication Products (%)
100
20
10
AT
CT
TG
AT CT GT AT CT GT
WT
TKO
c
Bypass Efficiency (%)
180
160
140
120
100
80
60
40
20
AT CT GT TT TC TG TA AT CT GT TT TC TG
S_p
S_p+R_p
R_p

Figure 3-29. Bypass efficiencies of Me-PTEs (a) and *n*Bu-PTEs (c), and mutation frequencies of the S_p-Me-PTE lesions (b) in AB1157 *E. coli* strains that are proficient in translesion synthesis or with all three SOS-induced DNA polymerases (Pol II, Pol IV and Pol V) being genetically depleted. The data represent the mean and standard deviation of results from three independent replication experiments.

a


BbsI
MluCI
5'-ATGGCMNGCTATAATTCA-3'
3'-TACCGXYCGATATTAAGT-5'
BbsI
SAP
[γ-32P]ATP + PNK
MluCI
5'-AT pGGCMNGCTAT AATTCAC-3'
3'-TACCGX pYCGATATTAA GTG-5'

b


AT
A(Me)T-Sp
A(Me)T-Rp
CT
C(Me)T-Sp
C(Me)T-Rp
GT
G(Me)T-Sp
G(Me)T-Rp
10merAT
10merCT
10merGT
Cmpt

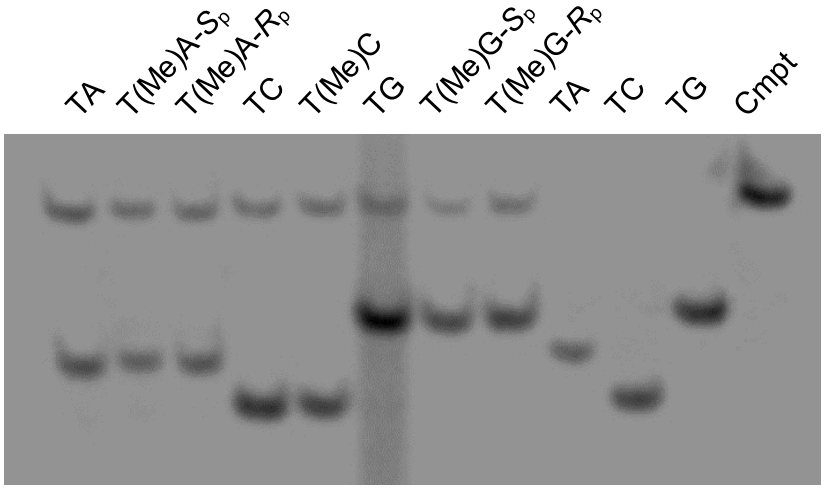

TA
T(Me)A-Sp
T(Me)A-Rp
TC
T(Me)C
TG
T(Me)G-Sp
T(Me)G-Rp
TA
TC
TG
Cmpt

Figure 3-30. Native PAGE (30%) for monitoring the bypass efficiencies of Me-PTEs in *ada*-deficient AB1157 *E. coli* cells. (a) Selective labelling of the original lesion-containing strand and its complementary strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) Gel image showing the 13-mer and 10-mer products released from the top-strand (lesion-containing strand) of the PCR products of the progeny of the competitor genome and the control or lesion-carrying genome, where AT, CT, GT, TA, TC and TG represent the [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with 'MN' being AT, CT, GT, TA, TC and TG, respectively.

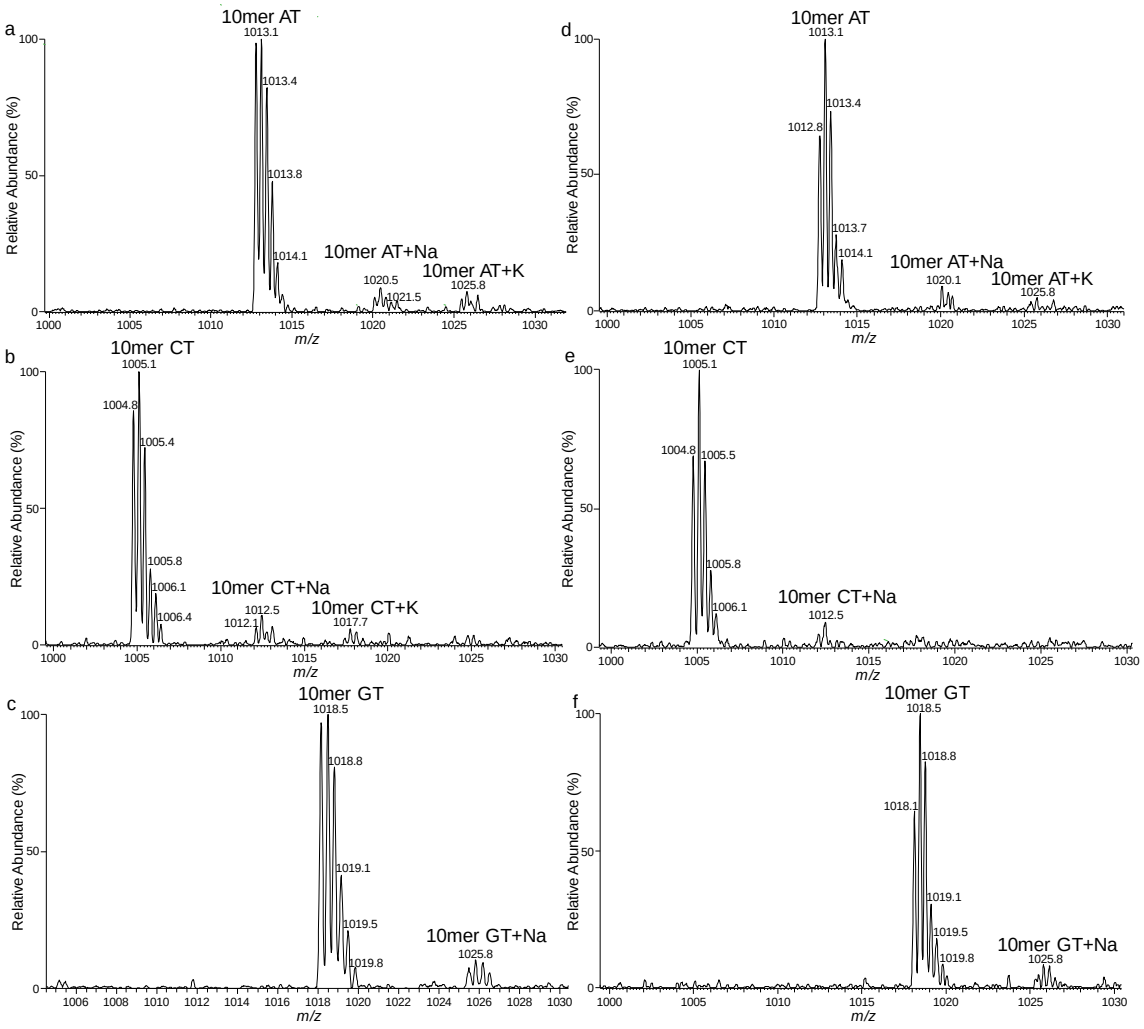


Figure 3-31. Higher-resolution "zoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of single-stranded M13 genomes harboring a site-specifically inserted (a) S_p-A(Me)T, (b) S_p-C(Me)T, (c) S_p-G(Me)T, (d) R_p-A(Me)T, (e) R_p-C(Me)T, and (f) R_p-G(Me)T in *ada*-deficient cells. Displayed in (a)-(f) are the $[M-3H]^{3-}$ ions for the lesion-containing strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for Me-PTE lesions are shown in Figures 3-21.

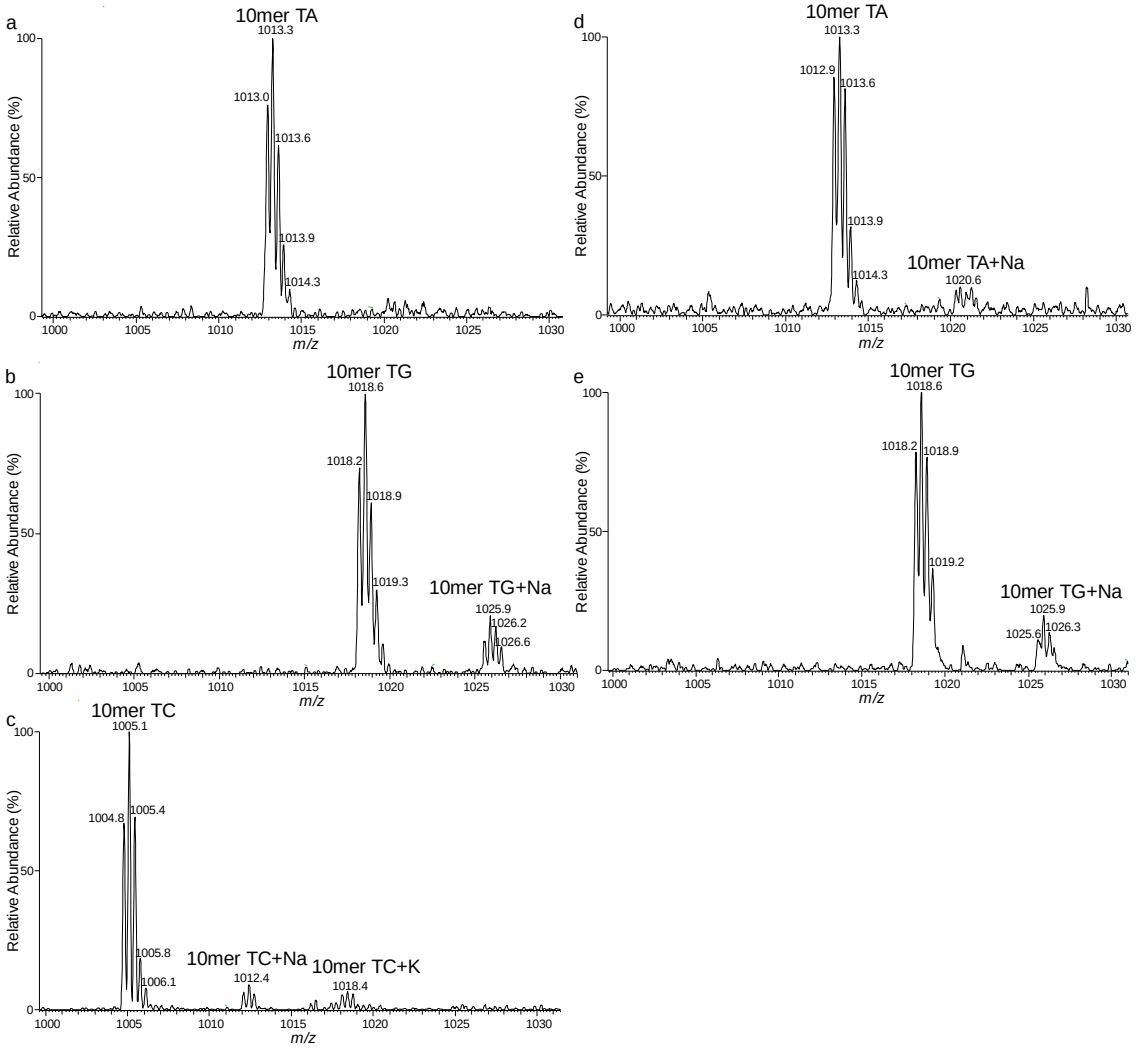


Figure 3-32. Higher-resolution "zoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of single-stranded M13 genomes harboring a site-specifically inserted (a) S_p-T(Me)A, (b) S_p-T(Me)G, (c) T(Me)C, (d) R_p-T(Me)A, and (e) R_p-T(Me)G in *ada*-deficient cells. Displayed in (a)-(e) are the $[M-3H]^{3-}$ ions for the lesion-containing strand products. All the mutagenic products were further confirmed by MS/MS analyses, and representative MS/MS results for the restriction fragments corresponding to replication products for Me-PTE lesions are shown in Figures 3-25.

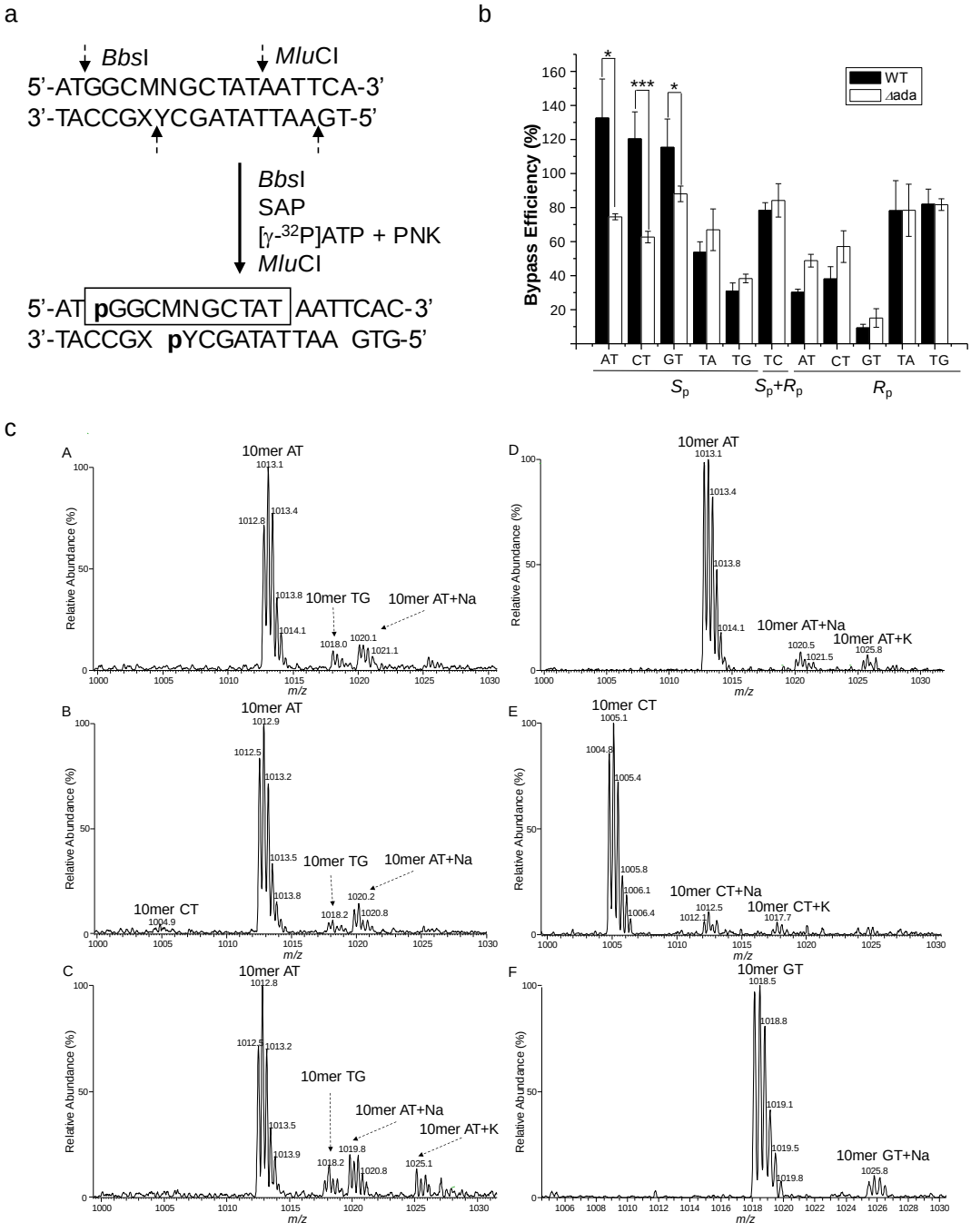
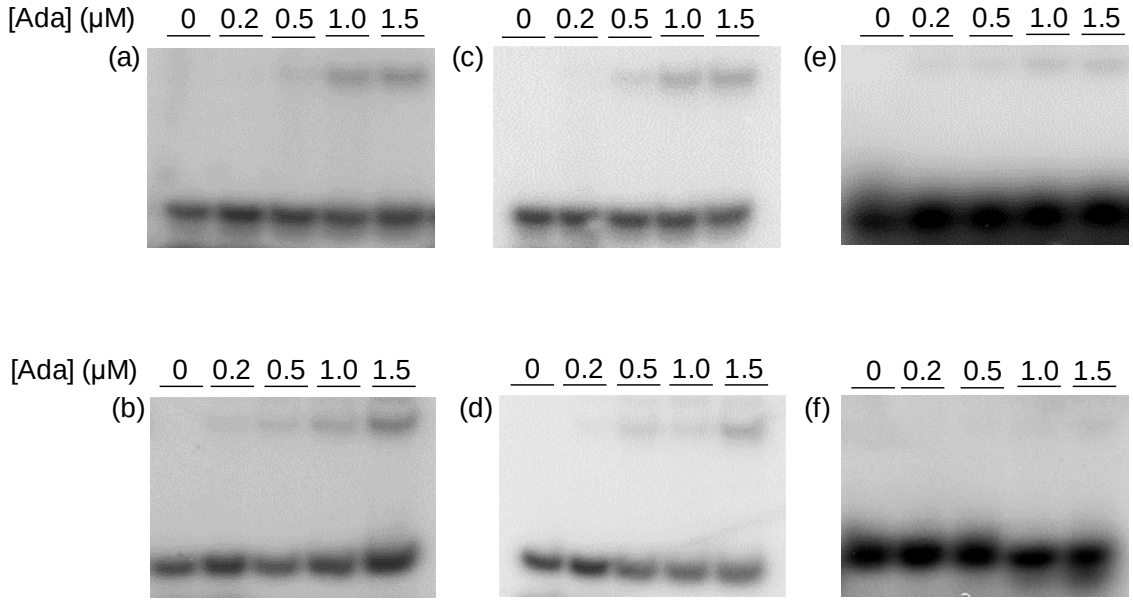

a
BbsI
MluCI
5'-ATGGCMNGCTATAATTCA-3'
3'-TACCGXYCGATATTAAGT-5'
BbsI
SAP
[γ-32P]ATP + PNK
MluCI
5'-AT pGGCMNGCTAT AATTCAC-3'
3'-TACCGX pYCGATATTAA GTG-5'
b
Bypass Efficiency (%)
WT
Δada
AT CT GT TA TG TC AT CT GT TA TG
Sp
Sp+Rp
Rp
c
A
10mer AT
10mer TG
10mer AT+Na
B
10mer AT
10mer CT
10mer TG
10mer AT+Na
C
10mer AT
10mer TG
10mer AT+Na
10mer AT+K
D
10mer AT
10mer AT+Na
10mer AT+K
E
10mer CT
10mer CT+Na
10mer CT+K
F
10mer GT
10mer GT+Na
Relative Abundance (%)
m/z

Figure 3-33. Ada protein promotes the replicative bypass of the S_p-XT-Me-PTEs and is required for the mutations induced by the lesions. (a) Selective labelling of original lesion-containing strand and its complementary strand via sequential restriction digestion. 'SAP' and 'PNK' represent shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. (b) The bypass efficiencies of the Me-PTE lesions in *E. coli* cells that are proficient in or deficient in Ada (*Δ*Ada). The data represent the mean ± S.D. of results from three independent replication experiments. *, $0.01 \leq p < 0.05$; **, $0.001 \leq p < 0.01$. The p values were calculated by using an unpaired two-tailed t-test and referred to comparisons with the data obtained for parental AB1157 cells. (c) Higher-resolution "zoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of S_p-XT-PTEs bearing single-stranded M13 genomes in SOS-induced wild-type and the isogenic Ada-deficient cells. Displayed are the $[M-3H]^{3-}$ ions of the restriction fragments for the bottom strand of the replication products of: (A) A(Me)T, (B) C(Me)T, and (C) G(Me)T in AB1157 cells, respectively, and (D) A(Me)T, (E) C(Me)T and (F) G(Me)T in the isogenic Ada-deficient cells, respectively. All the mutagenic products were confirmed by MS/MS analyses.

Primer name	Primer Sequence
ada-Forward Primer	5′- CGGCTCGAGTTACCTCTCCTCATTTTCAGCT -3′
ada-Reverse Primer	5′- CGCGGATCCATGAAAAAAGCCACATGCTT -3′

Table 3-1. The list of primers used in the present study.



	Description	Sequence	K_d (μM)	*p* value
a	ssAT	AGTGGAAGACATGGCATGCTAT	4.28±0.21	2.03×10^{-4}
b	ssA(Me)T-S_p	AGTGGAAGACATGGCA(Me)TGCTAT	1.05±0.04	
c	dsAT	AGTGGAAGACATGGCATGCTAT TTTTTCACCTTCTGTACCGTACGATATTTT	2.32±0.09	0.86
d	dsA(Me)T-S_p	AGTGGAAGACATGGCA(Me)TGCTAT TTTTTCACCTTCTGTACCGT ACGATATTTT	2.35±0.09	
e	ssTA	AGTGGAAGACATGGCTAGCTAT	15.93±1.25	0.26
f	ssT(Me)A-S_p	AGTGGAAGACATGGCT(Me)AGCTAT	24.68±0.88	

Figure 3-34. Electrophoresis mobility shift assay (EMSA) for quantifying the binding capacity of Ada protein with (a) ssAT, (b) ssA(Me)T-S_p, (c) dsAT, (d) dsA(Me)T-S_p, (e) ssTA and (f) ssT(Me)A-S_p. The DNA sequences, K_d values and *p* values are listed in the table.

REFERENCES

1. Lindahl, T. (1993) Instability and decay of the primary structure of DNA. *Nature*, **362**, 709-715.

2. Fu, D., Calvo, J.A. and Samson, L.D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer*, **12**, 104-120.

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

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

5. Ma, B., Zarth, A.T., Carlson, E.S., Villalta, P.W., Upadhyaya, P., Stepanov, I. and Hecht, S.S. (2018) 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.*, **31**, 48-57.

6. Tsujikawa, L., Weinfield, M. and Reha-Krantz, L.J. (2003) Differences in replication of a DNA template containing an ethyl phosphotriester by T4 DNA polymerase and *Escherichia coli* DNA polymerase I. *Nucleic Acids Res.*, **31**, 4965-4972.

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

8. Guichard, Y., Jones, G.D. and Farmer, P.B. (2000) Detection of DNA alkylphosphotriesters by ^{32}P postlabeling: evidence for the nonrandom manifestation of phosphotriester lesions *in vivo*. *Cancer Res.*, **60**, 1276-1282.

9. Le Pla, R.C., Guichard, Y., Bowman, K.J., Gaskell, M., Farmer, P.B. and Jones, G.D. (2004) 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.*, **17**, 1491-1500.

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

11. Wang, P., Amato, N.J., Zhai, Q. and Wang, Y. (2015) Cytotoxic and mutagenic properties of O^4-alkylthymidine lesions in *Escherichia coli* cells. *Nucleic Acids Res.*, **43**, 10795-10803.

12. Delaney, J.C. and Essigmann, J.M. (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions *in vivo*. *DNA Repair*, **408**, 1-15.

13. Thomason, L.C., Costantino, N. and Court, D.L. (2007) *E. coli* genome manipulation by P1 transduction. *Curr. Protoc. Mol. Biol.*, **Chapter 1**, Unit 1 17.

14. Zhai, Q., Wang, P., Cai, Q. and Wang, Y. (2014) Syntheses and characterizations of the in vivo replicative bypass and mutagenic properties of the minor-groove O^2-alkylthymidine lesions. *Nucleic Acids Res.*, **42**, 10529-10537.

15. He, C., Hus, J.C., Sun, L.J., Zhou, P., Norman, D.P., Dotsch, V., Wei, H., Gross, J.D., Lane, W.S., Wagner, G. *et al.* (2005) A methylation-dependent electrostatic switch controls DNA repair and transcriptional activation by *E. coli* Ada. *Mol. Cell*, **20**, 117-129.

16. McCarthy, T.V. and Lindahl, T. (1985) Methyl phosphotriesters in alkylated DNA are repaired by the Ada regulatory protein of *E. coli*. *Nucleic Acids Res.*, **13**, 2683-2698.

17. Myers, L.C., Terranova, M.P., Ferentz, A.E., Wagner, G. and Verdine, G.L. (1993) Repair of DNA methylphosphotriesters through a metalloactivated cysteine nucleophile. *Science*, **261**, 1164-1167.

18. Nakamura, T., Tokumoto, Y., Sakumi, K., Koike, G., Nakabeppu, Y. and Sekiguchi, M. (1988) Expression of the *ada* gene of *Escherichia coli* in response to alkylating agents. Identification of transcriptional regulatory elements. *J. Mol. Biol.*, **202**, 483-494.

19. Samson, L. and Cairns, J. (1977) A new pathway for DNA repair in *Escherichia coli*. *Nature*, **267**, 281-283.

20. Nair, D.T., Johnson, R.E., Prakash, L., Prakash, S. and Aggarwal, A.K. (2005) Rev1 employs a novel mechanism of DNA synthesis using a protein template. *Science*, **309**, 2219-2222.

21. Harper, M. and Lee, C.J. (2012) Genome-wide analysis of mutagenesis bias and context sensitivity of *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (NTG). *Mutat. Res.*, **731**, 64-67.

22. Wang, P., Leng, J. and Wang, Y. (2019) DNA replication studies of *N*-nitroso compound-induced O^6-alkyl-2'-deoxyguanosine lesions in *Escherichia coli*. *J. Biol. Chem.*, **294**, 3899-3908.

CHAPTER 4

Chemical Syntheses and Replication Studies of Pyridyloxobutylphosphotriester Lesions in *Escherichia coli* Cells

INTRODUCTION

Due to its chemical instability, DNA is susceptible to damage by endogenous and exogenous factors, resulting in the formation of various DNA adducts (1,2). If not efficiently repaired, these DNA adducts may induce mutations and ultimately result in carcinogenesis (3). Tobacco and its combustion products contain more than 8,000 compounds (4), over 70 of which are classified as carcinogens by the International Agency for Research on Cancer (IARC) (5). Among them, two tobacco-specific nitrosamines, *N*′-nitrosonornicotine (NNN) and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), are considered group-I carcinogens (carcinogenic to humans) by the IARC (6). Additionally, a reduced metabolite of NNK, 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) was also revealed to induce cancer in rodents (7-9).

NNN and NNK can be metabolically activated by cytochrome P450s to form reactive intermediates, which could pyridyloxobutylate (POB) DNA to form numerous lesions (9-11). Among them, O^2-POB-thymidine (O^2-POBdT) and O^6-POB-2′-deoxyguanosine (O^6-POBdG) were detected at significant levels in lung tissues of NNK-

treated mice, and owing to the lack of *N*-glycosidic bond stability, O^2-POB-cytosine (O^2-POBCyt) and *N*7-POB-guanine (*N*7-POBGua) were also detected as nucleobase adducts (12). Additionally, some of the POB-modified nucleobase derivatives were found to be mutagenic in cells. O^6-POBdG could induce G→A mutation in *E. coli* cells and both G→A and G→T mutations in HEK293T cells (13). Du *et al.* also revealed that, in HEK293T cells, O^2-POBdT and O^4-POBdT could induce T→A and T→C mutations, respectively (14).

Apart from nucleobase adducts, backbone alkylation products, also known as alkyl phosphotriesters (alkyl-PTEs) came to our sight in recent years (15). Many alkylating agents, e.g. *N*-methyl-*N*-nitrosourea, diazomethane and methyl methanesulfonate, could attack non-carbon-bonded oxygen atoms at phosphate backbone to form PTEs (16-18). Additionally, PTEs constitute approximately 55 % of total alkylated DNA lesions for *N*-ethyl-*N*-nitrosourea (19). Because conjugation with an alkyl group could neutralize the negative charge on backbone phosphate, PTEs were shown to perturb the interaction of DNA with DNA-binding proteins, including MutS (20), MutY (21) and RNA polymerase (22). Results from *in vitro* primer extension assay revealed that Et-PTE-containing ODN could inhibit the replication mediated by purified T4 DNA polymerase and *E. coli* DNA polymerase I (23). Our previous replication study in *Escherichia coli* cells also unveiled that S_p-T(Me)T induced TT→GT and TT→GC mutations, which required the Ada (24).

Aside from simple backbone alkylation products, active metabolites of NNN and NNK can attack phosphate backbone to form pyridyloxobutylphosphotriesters (POB-PTEs) (Figure 4-1A). Both NNN and NNK can be activated to form POB-phosphate adducts *in vitro* and in rats (25,26). Similar to other simple alkyl-PTE products, POB-PTEs are also highly persistent in mammalian tissues. POB-PTEs could remain highly abundant in over 70 weeks in lung tissue of rats treated with drinking water containing 5 ppm NNN (26). Additionally, after chronic exposure to NNK, 55-73% of the total POB DNA adducts in rat lung tissues reside on phosphate backbone, suggesting that POB-PTEs can serve as biomarkers for some tobacco-related diseases, like lung cancer (25). However, no research has elucidated how POB-PTEs affect DNA replication *in vitro* or in cells.

Herein, we synthesized oligodeoxyribonucleotides (ODNs) harboring a site-specifically incorporated and stereochemically defined POB-PTE lesions, we also employed a shuttle vector-based method coupled with LC-MS/MS to assess the impact of POB-PTEs at TT dinucleoside site on the efficiency and fidelity of DNA replication in *E. coli* cells. Moreover, we assessed how replication past POB-PTEs is modulated by translesion synthesis (TLS) DNA polymerases and Ada protein.

EXPERIMENTAL DETAILS

Chemicals and supplies

Unless further specified, all chemicals were from Thermo Fisher Scientific (Pittsburg, PA, USA). Common reagents for solid-phase ODN synthesis were obtained from Glen Research (Sterling, VA, USA). ODNs without modification were purchased from Integrated DNA Technologies (Coralville, IA, 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 Perkin Elmer (Piscataway, NJ, USA). All enzymes were from New England Biolabs (Ipswich, MA, USA).

M13mp7(L2) plasmid and wild-type AB1157 *E. coli* strains were kindly provided from Prof. John. M. Essigmann (27,28). Polymerase-deficient AB1157 *E. coli* strain [*ΔpolB1*::spec (Pol II deficient), *ΔdinB* (Pol IV-deficient), *ΔumuC*::kan (Pol V deficient) and *ΔpolB1*::spec *ΔdinB ΔumuC*::kan (Pol II, Pol IV, Pol V-triple knockout)] were generously provided by Prof. Graham. C. Walker (27). Ada-deficient AB1157 *E. coli* strain was generated following published procedures (28).

Chemical Synthesis

The 5′-DMTr-protected thymidine (29) and 4-(1,3-dithian-2-yl)-4-((3-pyridyl) butan-1-ol (POB-OH) (30) was synthesized according to published procedures. For dT-

phosphoramidite synthesis, compound **2** (100 mg, 0.18 mmol) was dissolved in 3 ml of anhydrous CH_2Cl_2 at room temperature under argon protection. Bis(diisopropylamino)chlorophosphine (146 mg, 0.54 mmol) and *N,N*-diisopropylethylamine (94 μl, 0.54 mmol) were subsequently added and the reaction mixture was stirred for 1 h, followed by vacuum concentration. The residue was isolated using column chromatography with a mixture of hexane, ethyl acetate and triethylamine (49/49/2, v/v) to obtain compound **3**. For the preparation of compound **4**, POB-OH (33 mg, 1.1 eq) was added into a round bottom flask with compound **3** (100 mg, 0.12 mmol), followed by addition of 0.25 M 1H-ethylthiotetrazole in acetonitrile (1 eq, Glen Research). The reaction mixture was stirred under an argon atmosphere at room temperature for 1 h, concentrated by rotary evaporation, and then purified by flash column chromatography to obtain compound **4**. The synthetic procedures are shown in Figure 4-1B.

ODN synthesis

A Beckman Oligo 1000M DNA synthesizer (Fullerton, CA) was employed to synthesize the 12-mer lesion-containing ODN: 5′-ATGGCT(X)TGCTAT-3′ at 1 μM scale, where 'X' designates the dithiane-protected POB group. The phosphoramidite building block (compound **4**) was dissolved in anhydrous acetonitrile at a final concentration of 0.067 M, and commercially available phosphoramidite building blocks were used for the

incorporation of unmodified nucleotides (Glen Research, Sterling, VA, USA) following standard protocols. A 55-min treatment with concentrated ammonium hydroxide was employed to cleave the synthesized ODNs from controlled pore glass support. After solvent removal by Speed-Vac, the solid residue was dissolved in water for HPLC purification.

HPLC

An Agilent 1100 system with a Thermo Hypersil Gold aQ column (250×4.6 mm, 3 μm in particle size and 175 Å in pore size; Thermo Fisher Scientific, Waltham, MA, USA) was used for HPLC separation of the synthesized 12-mer ODNs. The mobile phases consisted of 50 mM triethylammonium acetate (TEAA, pH 6.8, phase A) and 30% acetonitrile in 50 mM TEAA (phase B). The gradient profile was 5%-25% B in 5 min, followed by 25-80% B in 65 min, and the flow rate was 0.5 ml/min. The HPLC trace for the purification of the 12-mer lesion-containing ODNs is displayed in Figure 4-2.

After HPLC purification, the dithiane group on POB was removed by incubation with *N*-chlorosuccinimide (5 eq.) in acetonitrile and water (50:50, v/v, prechilled on ice) at room temperature for 10 min. The reaction mixture was purified by HPLC with the aforementioned mobile phases and gradient. The electrospray ionization-mass spectra (ESI-MS) and tandem MS (MS/MS) for the purified ODNs are provided in Figure 4-3.

Construction of single-stranded lesion-containing and lesion-free control M13 genomes.

The 12-mer lesion-containing ODNs were 5′-phosphorylated and ligated with a 10-mer ODN (5′-AGTGGAAGAC-3′) in the presence of T4 DNA ligase and ATP in the ligase buffer (New England Biolabs) at 10 °C for 12 h. The resulting 22-mer ligation product was purified by denaturing polyacrylamide gel electrophoresis (PAGE).

The lesion-containing and lesion-free single-stranded M13mp7 plasmids were prepared following previously published procedures (27). Briefly, 20 pmol of M13 vector was linearized by digestion using 40 U *EcoR*I-HF at 25 °C for 8 h. Two scaffolds were used for the ligation, 5′-CTTCCACTCACTGAA TCATGGTCATAGCTTTC-3′ and 5′-AAAACGACGGCCAGTGAATTATAGC-3′ (25 pmol), each of which spanned one end of the linearized vector. The above-described 5′-phosphorylated 22-mer lesion-containing or 25-mer competitor ODNs were subsequently ligated to M13 plasmid and incubated with T4 DNA ligase at 16 °C for 8 h. Excess ODNs and unligated vector were cleaved by the exonuclease activity of T4 polymerase (22.5 U) at 16 °C for 2 h. Cycle Pure Kit (Omega) was employed to purify the lesion-containing or lesion-free vectors.

Transfection of M13 plasmids into *E. coli* cells

The lesion-containing and lesion-free control vectors were mixed individually with the competitor vector at a 1:1 molar ratio. The mixtures were then transfected into wild-type (WT) AB1157 *E. coli* strains, or the isogenic strains deficient in Pol II, Pol IV, Pol V, or all three in combination (TKO), or Ada. The transfected *E. coli* cells were cultured at 37 °C for 5.5 h, and the M13 phage was recovered from the supernatant by centrifugation at 13,200 rpm for 5 min and then transfected to SCS110 *E. coli* strain for amplification, and the M13 phages were extracted by using QIAprep Spin M13 Kit (Qiagen).

Quantification of bypass efficiency by competitive replication and adduct bypass (CRAB) assay

A modified version of the competitive replication and adduct bypass (CRAB) assay was employed to quantify the bypass efficiencies of POB-PTE lesions in *E. coli* cells (27,29,31,32). The regions of the replication products harboring the lesion-containing site were amplified by PCR with the use of Phusion high-fidelity DNA polymerase, with the primers being 5′-YCAGCTATGACCATGATTCAGTGAGTGGA-3′ and 5′-YTCGGTGCGGGCCTCTTCGCTATTAC-3′, where 'Y' denotes the $H_2N(CH_2)_6$- group being conjugated to the 5′-phosphate group of the ODNs. The PCR amplification began from 98 °C for 30 s, followed by 35 cycles of amplification, with each cycle consisting

98 °C for 10 s, 65 °C for 30 s and 72 °C for 15 s, and then with a final extension at 72 °C for 5 min, ending at 4 °C. The PCR products were purified by Cycle Pure Kit (Omega).

PCR products (100 ng) were digested with *Bbs*I-HF restriction endonuclease (5 U) and recombinant shrimp alkaline phosphatase (rSAP, 10 U) in a 10 μl CutSmart buffer (New England Biolabs) at 37 °C for 30 min, followed by deactivation at 70 °C for 10 min. To the mixture were subsequently added 10 U T4 polynucleotide kinase (T4 PNK), 5 mM dithiothreitol (DTT), 1.66 pmol [γ-^{32}P]ATP, NEB CutSmart buffer, and water to a total volume of 15 μl. The reaction was continued at 37 °C for 30 min, and T4 PNK was deactivated by heating the solution at 70 °C for 10 min. Next, 10 U *Mlu*CI was added to the above mixture, which was incubated at 37 °C for 25 min. The digestion reaction was terminated by adding 15 μl formamide gel-loading buffer containing xylene cyanol FF and bromophenol blue. The aforementioned digestion yielded 10 mer radiolabeled fragments for lesion-containing or control genome (5′-GGCMNGCTAT-3′, 'M' and 'N' being nucleobases at the dinucleotides initially flanking the lesion site), and a 13 mer fragment for competitor genome (5′-GGCGATAAGCTAT-3′), which were resolved using 30% native polyacrylamide gel (19:1). The intensities for the radiolabeled bands were quantified by using a Typhoon 0410 imager, and the bypass efficiency was calculated by using the following equation: Bypass efficiency (%) = (lesion signal/competitor signal)/(control signal/competitor signal) × 100%.

Identification of replication products by mass spectrometry

Approximately 3 μg of PCR products were digested by 50 U *Bbs*I-HF and 20 U rSAP in 250 μl CutSmart buffer at 37 °C for 2 h, followed by deactivation at 70 °C for 20 min. To the above mixture was subsequently added 20 U *MluC*I, and the digestion was continued at 37 °C for 1 h. The solution was extracted once with phenol/ chloroform/isoamyl alcohol (25:24:1 v/v) and the aqueous layer was collected and concentrated. The resulting solid residue was dissolved in 100 μl water and desalted using Waters Oasis HLB extraction cartridge (Milford, MA). The eluate was evaporated by Speed-Vac, redissolved in 10 μl water and subjected to LC-MS and MS/MS analyses. An LTQ linear ion trap mass spectrometer (Thermo Electron, San Jose, CA, USA) with an Agilent 1200 capillary HPLC pump was used to analyze the digestion products and a Zorbax SB-C18 column (0.5 × 150 mm, 5 μm in particle size) was employed for the separation. Mobile phases A and B contained 400 mM HFIP (pH 7.0) and methanol, respectively. The gradient was 5-20% B in 5 min and 20-50% B in 35 min. The ion-transfer tube temperature was 275 °C and the mass spectrometer was set up for recording the higher-resolution ultra-zoom scan MS and MS/MS for the cleavages of $[M-3H]^{3-}$ ions for 10 mer d(GGCMNGCTAT), with 'N' and 'M' being A, T, C and G.

RESULTS

The objectives of this study were to assess the impact of POB-PTEs on DNA replication, to define the roles of three translesion synthesis (TLS) polymerases in bypassing these lesions, and to examine the function of Ada in modulating the replication across the POB-PTEs in *E. coli* cells.

We synthesized and characterized 12-mer ODNs containing a site-specifically inserted POB-PTE lesion. The synthesized ODNs were purified by HPLC and characterized with ESI-MS and MS/MS. Our previous results revealed that using C18 reverse phase column with TEAA/acetonitrile as mobile phases, the S_p diastereomer of alkyl-PTEs eluted earlier than the corresponding R_p diastereomer (24). Here, we assign the earlier- and later-eluting fractions as the S_p- and R_p-POB-PTE-harboring ODNs, respectively. The MS/MS results confirmed the sequence of the synthesized ODNs and verified the location of POB group conjugated with the phosphate backbone (Figure 4-3).

We employed a shuttle-vector method to assess the bypass efficiencies of the POB-PTEs lesions (27). The mixture of lesion-containing and competitor vectors was transfected into AB1157 *E. coli* strains proficient or deficient with SOS-induced DNA polymerases. After replication, the progeny was extracted, followed by PCR amplification. The PCR products were subsequently digested with two restriction endonucleases, *Bbs*I and *MluC*I, yielding10 mer ODNs, which contain the replication

products of initial damage-containing site, for lesion-containing or lesion-free control genome, or 13 mer for the corresponding region of the competitor genome, of which the 5′-terminius was radiolabeled (Figure 4-4). By switching the order of digestion for the two restriction enzymes, either the lesion-situated strand (p*GGCMNGCTAT) or the corresponding complementary strand (p*AATTATAGCY) can be radiolabeled (see Experimental Procedures). The resulting ODNs were analyzed by native PAGE and LC-MS/MS.

Alkylation of the phosphate backbone can potentially influence the fidelity of replication of both nucleosides flanking the PTE site (24), yielding up to 16 replication products (A, T, C or G being incorporated at the flanking nucleoside sites), which cannot be completely resolved by PAGE. Therefore, we subjected the aforementioned digestion products to LC-MS and MS/MS analyses, where we monitored the $[M-3H]^{3-}$ ions of d(GGCMNGCTAT). The LC-MS/MS data showed that neither S_p- nor R_p-POB-PTEs was mutagenic, which is agreement with the PAGE analysis results (Figure 4-6, 4-7).

After identification of replication products, the bypass efficiencies of POB-PTEs were determined by comparing the intensity ratio of 10-mer band from the lesion-containing genome over the corresponding 13-mer bands from the competitor genome (Figure 4-5). The results showed that neither S_p- nor R_p-POB-PTE was strong impediment to DNA replication. Additionally, we assessed the role of three SOS-induced translesion synthesis

polymerases (Pol II, Pol IV and Pol V) in bypassing the POB-PTEs by conducting the CRAB assay using isogenic *E. coli* strains deficient in Pol II, Pol IV, Pol V, or all three in combination. The result showed that the replicative bypass of POB-PTEs did not require any of the SOS-induced DNA polymerases.

Previous research has found that *E. coli* Ada protein could remove S_p-Me-PTE, and is required for the mutagenic bypass of the S_p-Me-PTE at TT site (24,33). We next asked if Ada protein can influence the bypass efficiency or mutation patterns of the bulky POB-PTEs, and we observed that depletion of Ada protein could not affect the replication efficiencies or induce mutations of the POB-PTEs (Figure 4-5).

DISCUSSION

In the present study, we assessed how POB-PTEs flanked by TT influence DNA replication. First, we revealed that neither S_p- nor R_p-POB-PTEs showed significant blockage effect, which is, to some extent, similar to what we found for *n*Bu-PTEs in same sequence context (24). S_p-POB-PTE did not display significant difference in bypass efficiency compared with *n*Bu-PTE, whereas R_p-POB-PTEs exhibited an elevated bypass efficiency relative to R_p-*n*Bu-PTEs. Tsujikawa *et al.* (23) indicated that R_p-Et-PTE at TT sequence could cause bending of the template strand, which may introduce dynamic change to DNA-polymerase interactions, resulting in stronger steric hinderance and lower bypass efficiency of R_p-Et-PTE than the S_p counterpart (24). However, our results

indicated that this distinction in bypass efficiency for the two diastereomers of PTEs were narrowed with the increasing size of alkyl group conjugated with the backbone phosphate, suggesting that some other factors may modulate the aforementioned distortion-induced steric hindrance.

Second, we observed that the bypass efficiencies and mutagenic properties of POB-PTEs were not impacted by genetic depletion of the three SOS-induced DNA polymerases (Pol II, Pol IV and Pol V), alone or all three together, which is consistent with what we observed for simple alkyl-PTEs with different size of alkyl group and stereochemistry, as well as flanking base sequences around PTE site (24), supporting replication across these lesions does not require these TLS polymerases.

Previous research has revealed that *E. coli* Ada protein can remove S_p-Me-PTE, albeit at the expense of eliciting mutations (24,33). Here we found no apparent change in the efficiencies or fidelity of replication across POB-PTEs introduced by depletion of Ada, suggesting the lack of interaction between Ada and POB-PTEs.

In summary, the results from our shuttle vector-based replication study showed that POB-PTEs did not impede DNA replication or introduce mutations during DNA replication. Meanwhile, genetic depletion of TLS polymerases does not impact the replication bypass efficiencies of POB-PTEs, and Ada is incapable of removing POB phosphate adducts. Our study provided, for the first time, the knowledge about how bulky

POB group on phosphate backbone influences DNA replication *in vivo.* Future studies about how POB-PTEs affect DNA replication and transcription in mammalian cells will offer additional insights into the biological impacts of this unique class of DNA damage.

FIGURES AND SCHEMES

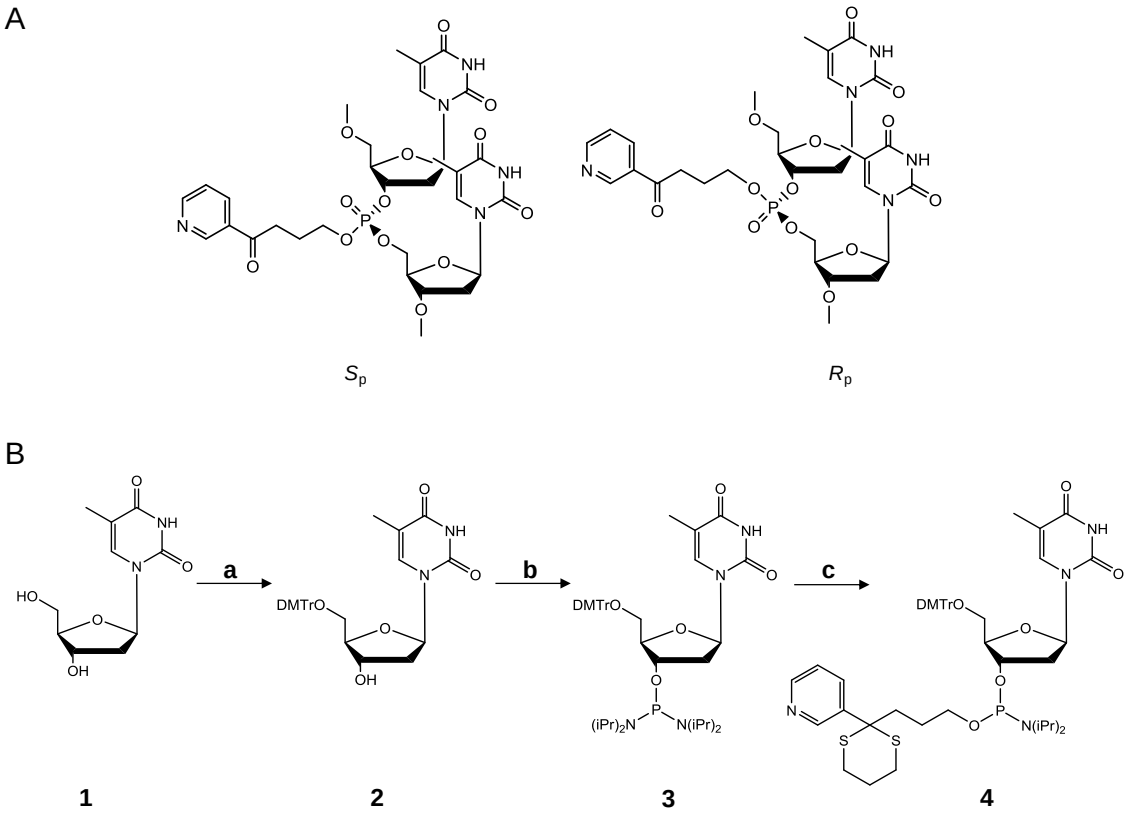


Figure 4-1. (A) The structures of the S_p and R_p diastereomers of POB-PTEs at TT site. (b) Chemical Synthesis of the POB the phosphoramidite building block. Reagents and conditions: (B) 4,4′-dimethoxytrityl chloride, 4-dimethylaminopyridine, pyridine, R. T., 10 h; (b) bis(diisopropylamino)chlorophosphine, *N,N*-diisopropylethylamine, dichloromethane, R. T. 1 h, (c) 4-(1,3-dithian-2-yl)-4-((3-pyridyl) butanol, ethylthiotetrazole, acetonitrile, dichloromethane, R. T. 1 h.

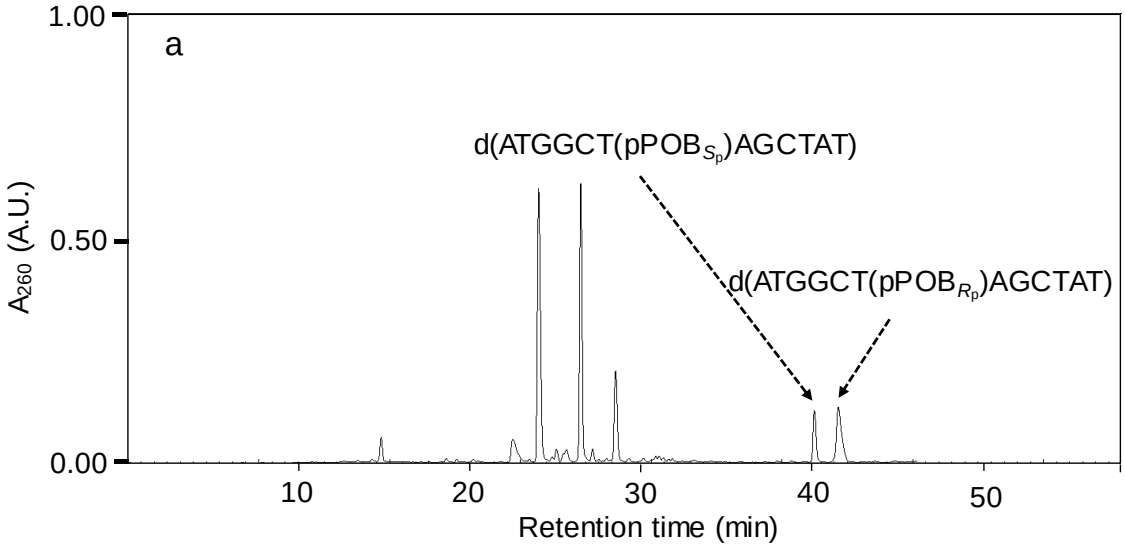


Figure 4-2. HPLC trace for the separation of the synthesized 12 mer dithiane-protected POB-PTE-bearing ODNs.

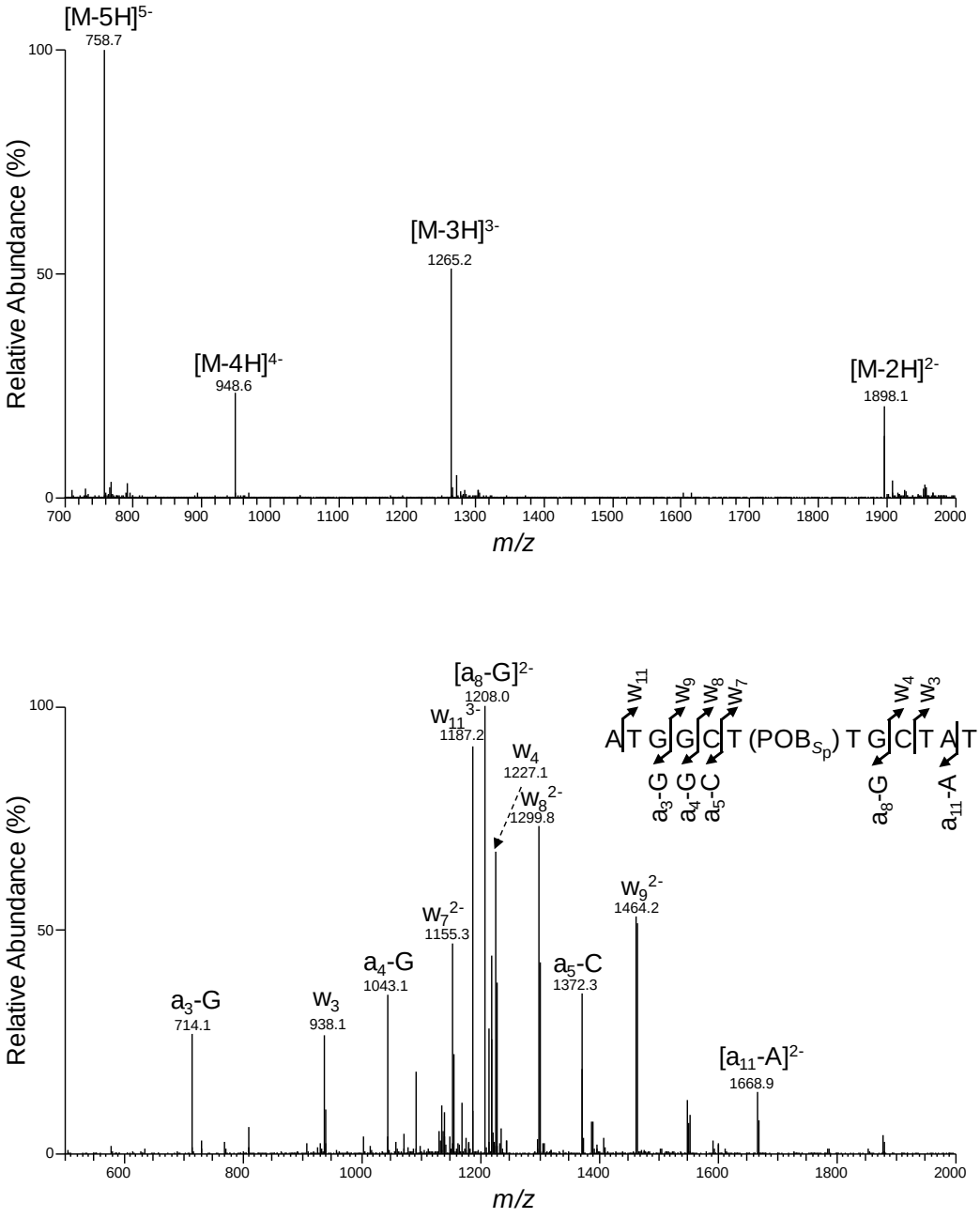


Figure 4-3. ESI-MS & MS/MS characterizations of d(ATGGCT(POB$_{Sp}$)TGCTAT): Negative-ion ESI-MS (top) and the product-ion spectrum of the [M-3H]$^{3-}$ ion (*m/z* 1265.2)

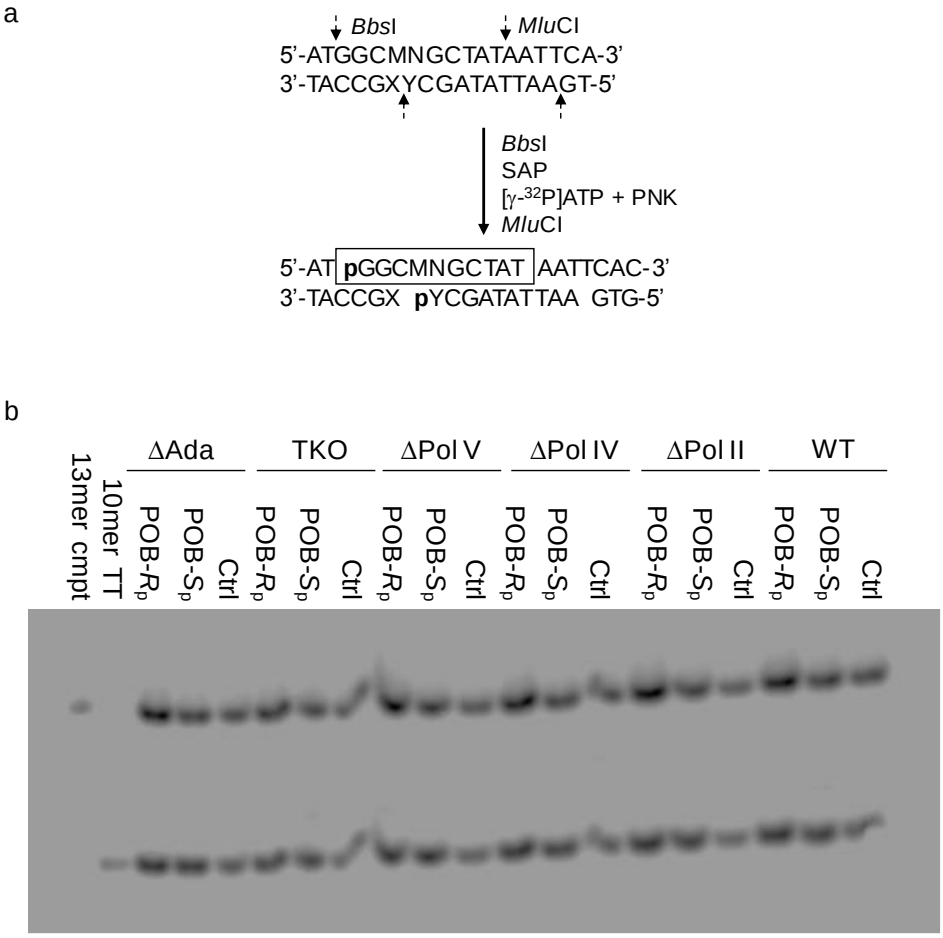


Figure 4-4. Restriction enzyme digestion and radiolabeling, followed by native PAGE (30%) analysis for quantifying the bypass efficiencies of POB-PTE lesions located at TT dinucleotide site in AB1157 *E. coli* cells. (a) Selective labelling of the original lesion-containing strand and its complementary strand via sequential restriction digestion. 'SAP' and 'PNK' designate shrimp alkaline phosphatase and T4 polynucleotide kinase, respectively. 'MN' indicate the flanking nucleobases of the site where the POB-PTE lesions was initially situated. (b) Gel image showing the 13 mer (competitor genome) and 10 mer (control or lesion-containing genome) digestion products formed from the original strand, where 10 mer TT designates [5′-^{32}P]-labeled standard ODNs 5′-GGCMNGCTAT-3′, with MN being TT, respectively.

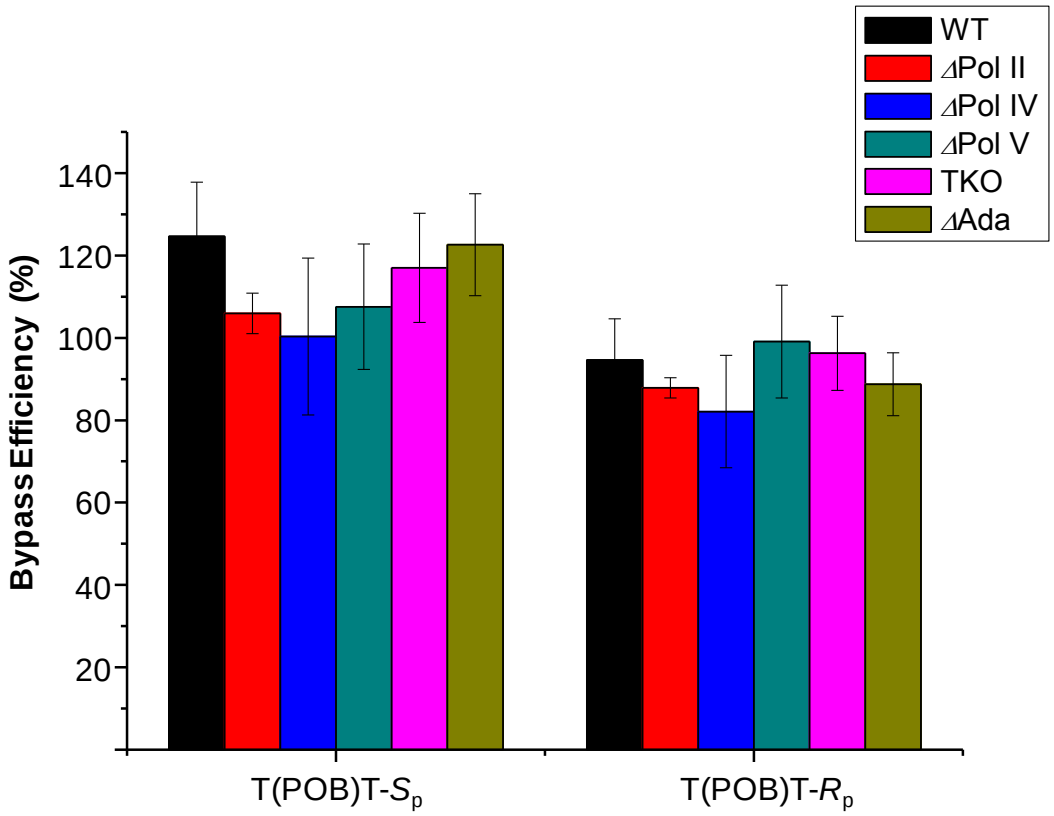


Figure 4-5. Bypass efficiencies of POB-PTEs in AB1157 *E. coli* strains that are proficient in translesion synthesis and Ada, or with Ada protein or three SOS-induced DNA polymerases (Pol II, Pol IV and Pol V) being genetically depleted alone or in combination. The data represent the mean and standard deviation of results from three independent replication experiments.

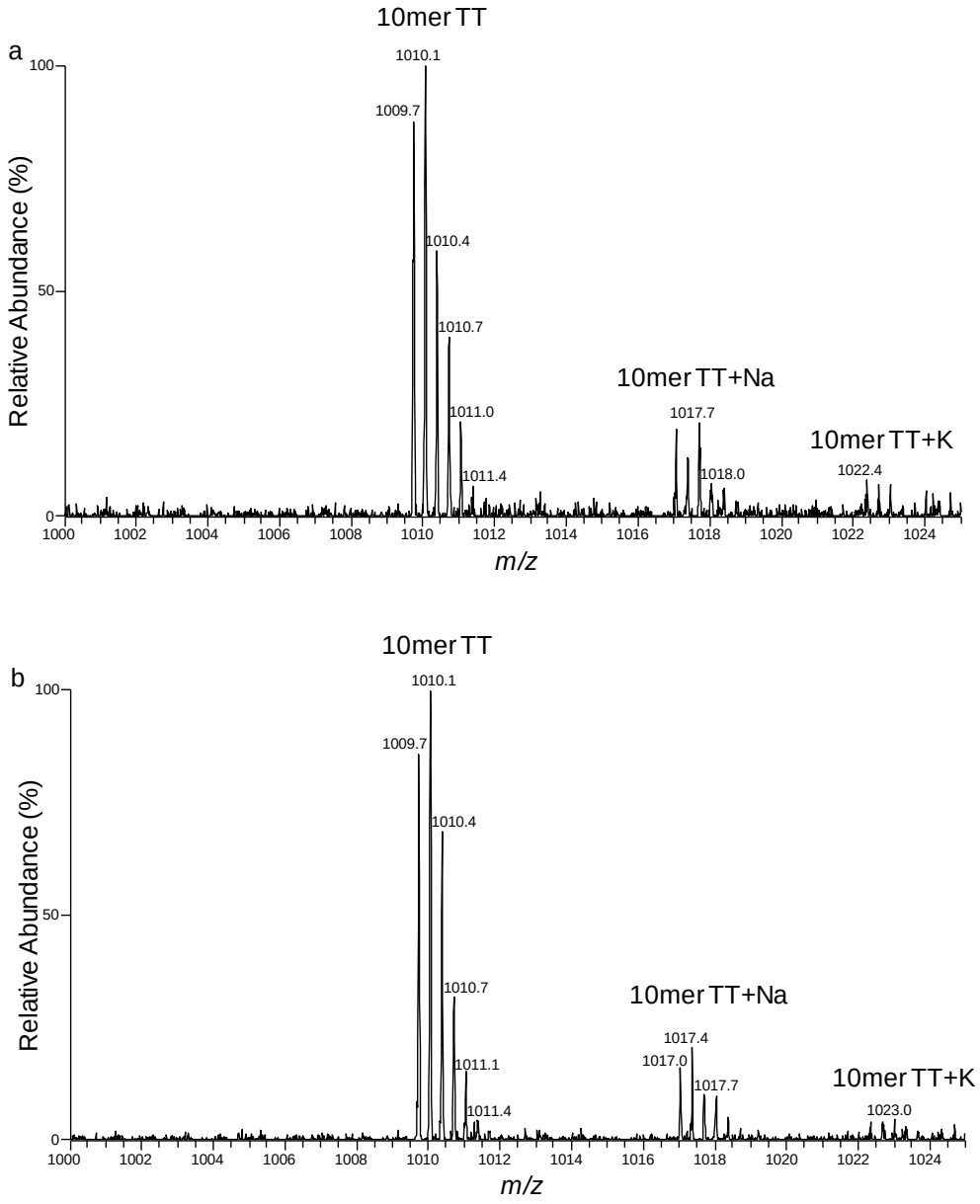


Figure 4-6. Higher-resolution "ultrazoom scan" ESI-MS of the restriction fragments for the PCR products from the replication of single-stranded M13 genomes harboring a site-specifically inserted (a) S_p-T(POB)T and (b) R_p-T(POB)T in AB1157 cells. Displayed in (a) and (b) are the $[M-3H]^{3-}$ ions of the strand initially contained the lesion.

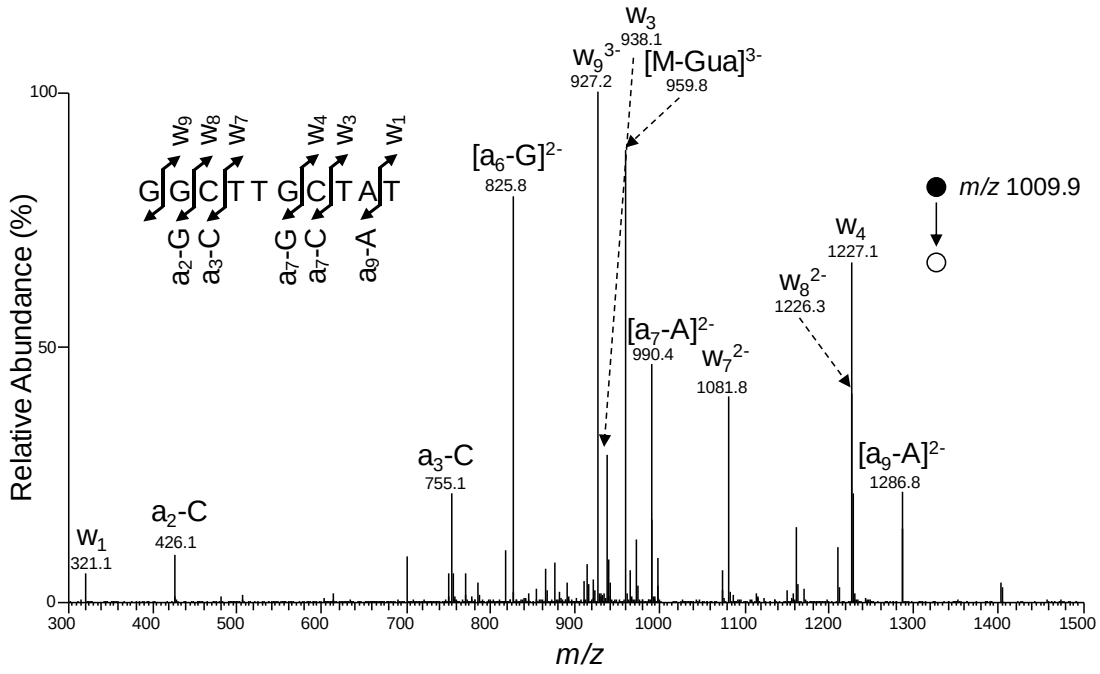


Figure 4-7. MS/MS for the identification of restriction fragments of PCR products. MS/MS for the $[M-3H]^{3-}$ ions of 10mer TT sequence from replication product of S_p-POB-PTE-containing genome.

REFERENCES

1. Fu, D., Calvo, J.A. and Samson, L.D. (2012) Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer.*, **12**, 104-120.

2. Lindahl, T. and Wood, R.D. (1999) Quality control by DNA repair. *Science*, **286**, 1897-1905.

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

4. Rodgman, A. and Perfetti, T.A. (2016) *The chemical components of tobacco and tobacco smoke*. CRC press.

5. Hecht, S.S. (2012) Research opportunities related to establishing standards for tobacco products under the Family Smoking Prevention and Tobacco Control Act. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco*, **14**, 18-28.

6. Ma, B., Zarth, A.T., Carlson, E.S., Villalta, P.W., Upadhyaya, P., Stepanov, I. and Hecht, S.S. (2018) 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.*, **31**, 48-57.

7. Hoffmann, D. and Hecht, S.S. (1985) Nicotine-derived *N*-nitrosamines and tobacco-related cancer - current status and future-directions. *Cancer Res.*, **45**, 935-944.

8. Hecht, S.S. and Hoffmann, D. (1988) Tobacco-specific nitrosamines, an important group of carcinogens in tobacco and tobacco-smoke. *Carcinogenesis*, **9**, 875-884.

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

10. Hecht, S.S. (2003) Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat. Rev. Cancer*, **3**, 733.

11. Hecht, S.S., Stepanov, I. and Carmella, S.G. (2016) Exposure and metabolic activation biomarkers of carcinogenic tobacco-specific nitrosamines. *Acc. Chem. Res.*, **49**, 106-114.

12. Hecht, S.S. (1999) DNA adduct formation from tobacco-specific *N*-nitrosamines. *Mutat. Res.*, **424**, 127-142.

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

14. Du, H., Leng, J., Wang, P., Li, L. and Wang, Y. (2018) Impact of tobacco-specific nitrosamine-derived DNA adducts on the efficiency and fidelity of DNA replication in human cells. *J. Biol. Chem.*, **293**, 11100-11108.

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

16. Brimacombe, R., Griffin, B.E., Haines, J., Haslam, W.J. and Reese, C. (1965) An approach to the methylation of polynucleotides. *Biochemistry*, **4**, 2452-2458.

17. Bannon, P. and Verly, W. (1972) Alkylation of phosphates and stability of phosphate triesters in DNA. *Eur. J. Biochem.*, **31**, 103-111.

18. Lawley, P.D. (1973) Reaction of *N*-methyl-*N*-nitrosourea (MNUA) with ^{32}P-labelled DNA: evidence for formation of phosphotriesters. *Chem. Biol. Interact.*, **7**, 127-130.

19. Beranek, D.T., Weis, C.C. and Swenson, D.H. (1980) A comprehensive quantitative-analysis of methylated and ethylated DNA using high-pressure liquid-chromatography. *Carcinogenesis*, **1**, 595-606.

20. Biswas, I. and Hsieh, P. (1997) Interaction of MutS protein with the major and minor grooves of a heteroduplex DNA. *J. Biol. Chem.*, **272**, 13355-13364.

21. Lu, A.L., Tsai-Wu, J.J. and Cillo, J. (1995) DNA determinants and substrate specificities of *Escherichia coli* MutY. *J. Biol. Chem*, **270**, 23582-23588.

22. Marushige, K. and Marushige, Y. (1983) Template properties of DNA alkylated with *N*-methyl-*N*-nitrosourea and *N*-ethyl-*N*-nitrosourea. *Chem. Biol. Interact.*, **46**, 179-188.

23. Tsujikawa, L., Weinfield, M. and Reha-Krantz, L.J. (2003) Differences in replication of a DNA template containing an ethyl phosphotriester by T4 DNA polymerase and *Escherichia coli* DNA polymerase I. *Nucleic Acids Res.*, **31**, 4965-4972.

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

25. Li, Y., Ma, B., Cao, Q., Balbo, S., Zhao, L., Upadhyaya, P. and Hecht, S.S. (2019) Mass spectrometric quantitation of pyridyloxobutyl DNA phosphate adducts in rats chronically treated with *N*-nitrosonornicotine. *Chem. Res. Toxicol*.

26. Ma, B., Villalta, P.W., Zarth, A.T., Kotandeniya, D., Upadhyaya, P., Stepanov, I. and Hecht, S.S. (2015) 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.*, **28**, 2151-2159.

27. Delaney, J.C. and Essigmann, J.M. (2006) Assays for determining lesion bypass efficiency and mutagenicity of site-specific DNA lesions *in vivo*. *Methods Enzymol.*, **408**, 1-15.

28. Thomason, L.C., Costantino, N. and Court, D.L. (2007) *E. coli* genome manipulation by P1 transduction. *Curr. Protoc. Mol. Biol.*, **Chapter 1**, Unit 1 17.

29. Wang, P., Amato, N.J., Zhai, Q. and Wang, Y. (2015) Cytotoxic and mutagenic properties of O^4-alkylthymidine lesions in *Escherichia coli* cells. *Nucleic Acids Res.*, **43**, 10795-10803.

30. Spratt, T.E., Peterson, L.A., Confer, W.L. and Hecht, S.S. (1990) Solvolysis of model compounds for alpha-hydroxylation of *N'*-nitrosonornicotine and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone: evidence for a cyclic oxonium ion intermediate in the alkylation of nucleophiles. *Chem. Res. Toxicol.*, **3**, 350-356.

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

32. Zhai, Q., Wang, P., Cai, Q. and Wang, Y. (2014) Syntheses and characterizations of the *in vivo* replicative bypass and mutagenic properties of the minor-groove O^2-alkylthymidine lesions. *Nucleic Acids Res.*, **42**, 10529-10537.

33. Weinfeld, M., Drake, A.F., Saunders, J.K. and Paterson, M.C. (1985) Stereospecific removal of methyl phosphotriesters from DNA by an *Escherichia coli ada*$^+$ extract. *Nucleic Acids Res.*, **13**, 7067-7077.

CHAPTER 5

Concluding Remarks and Future Directions

In this dissertation, we investigated how differences in sizes, structures, stereochemical configuration and sequences contexts of alkyl phosphotriesters (alkyl-PTEs) influence the efficiency and fidelity of DNA replication by using the competitive replicative adduct bypass (CRAB) assay. We also examined how translesion synthesis (TLS) polymerases modulate the replicative bypass of alkyl-PTEs in *E. coli* cells. Additionally, since *E. coli* Ada protein has been found to remove S_p-Me-PTEs, we also characterized the cytotoxicity and mutagenicity of S_p-Me-PTEs induced by Ada protein and analyzed the sequence-dependent interaction of Ada protein with S_p-Me-PTEs.

In Chapter 2, we synthesized oligodeoxyribonucleotides (ODNs) containing alkyl-PTEs (with the four alkyl groups being Me, Et, *n*Pr or *n*Bu) at TT site, followed by separation and identification of the two diastereomers by using HPLC and NMR analysis. We also conducted replication studies in *E. coli* cells, and our results showed that the S_p diastereomers of alkyl-PTEs could be efficiently bypassed, whereas their R_p counterparts exhibited moderate blockage effects. Additionally, S_p-Me-PTEs induced TT→GT and TT→GC mutations, which depend on the Ada protein. Moreover, genetic depletion of

polymerase II, IV and V, alone or all three in conjunction, conferred subtle changes on bypass efficiencies and mutation frequencies of the alkyl-PTEs in *E. coli*. Taken together, these results provided a comprehensive understanding about the recognition of alkyl-PTE lesions by DNA replication machinery in *E. coli* cells and revealed for the first time the Ada-dependent induction of mutations at S_p-Me-PTE site.

In Chapter 3, we investigated the impact of flanking base sequences on the replication across alkyl-PTE lesions. We prepared ODNs with Me- and *n*Bu-PTEs in different flanking base sequences (XT and TX, with X being A, C and G) and assessed how these lesions perturb DNA replication in *E. coli* cells. Our results indicated that S_p-XT-Me-PTEs were not strong impediments to DNA replication. Moreover, their replication products were highly sequence-independent, which comprise 85-90% AT and 5-10% TG. Ada protein was again critical for the formation of these mutagenic products. The R_p diastereomers, however, exhibited error-free bypass with strong blockage effect. Unlike what we found for XT-Me-PTEs, TX-Me-PTEs were not mutagenic, and elicited moderate blockage effects on DNA replication. Additionally, none of S_p-XT-*n*Bu-PTEs significantly impeded DNA replication, whereas R_p- XT-*n*Bu-PTEs and all TX-*n*Bu-PTEs moderately blocked replication, with none of the *n*Bu-PTEs being mutagenic. Consistent with what we found for T(alkyl)T, depletion of all three TLS polymerases did not exert apparent effects on replicative bypass of Me-PTEs in different sequence contexts.

The differences in Ada-induced mutagenicity between XT- and TX-S_p-Me-PTEs suggested that Ada's interaction with PTE site was also sequence-dependent. We prepared 22-mer S_p-Me-PTE with flanking base sequences being AT or TA and employed electrophoresis mobility shift assay (EMSA) to analyze their binding capacities to Ada protein, and we found that S_p-A(Me)T bound strongly to Ada protein relative to that of S_p-T(Me)A. In summary, our study provides further understanding about how alkyl-PTE lesions are recognized by DNA replication machinery in prokaryotic cells, and our results suggest that the role of Ada may be beyond the repair of S_p-Me-PTE lesion and transcriptional modulation of genes involved in repairing alkylated DNA lesions.

In chapter 4, we assessed how POB-PTEs modulated the efficiency and fidelity of DNA replication and how their replicative bypass is mediated by TLS polymerases and Ada protein. Our replication studies showed that neither S_p- nor R_p-POB-PTEs elicited significant impediment to DNA replication in *E. coli*. We also examined the roles of translesion synthesis polymerases in bypassing POB-PTEs by conducting the replication assay in isogenic *E. coli* strains with SOS-induced DNA polymerases being depleted individually and together, and the results indicated that the replicative bypass of POB-PTEs did not require TLS polymerases. LC-MS/MS results also showed that POB-PTEs were not mutagenic. Therefore, our data provided a comprehensive understanding about how

tobacco-specific nitrosamine-induced POB-PTE adducts are recognized by the entire *E. coli* DNA replication machinery.

Further studies can be conducted, and they can focus on the following aspects: 1) It remains elusive how alkyl-PTEs influence replication and transcription in mammalian cells. Ma *et al.* (1) detected alkyl-PTEs in lung tissue of lung cancer patients, and an increased level of Me-PTEs was observed in lung tissues of smokers compared with that of non-smokers. In this dissertation, we mainly focus on the cytotoxic and mutagenic effects of alkyl-PTEs in *E. coli* cells, whereas the replication, repair and transcription machinery in mammalian cells are more complicated and distinct. For instance, *E. coli* Ada protein do not have mammalian homologues. Thus, it will be important to further extend our replication study to mammalian cells. 2) Phosphate adducts were found not only being lesions, but also being endogenous modifications. Many studies have detected and sequenced endogenous phosphorothioation in bacterial genomes and suggested its function in a restriction-modification system (2). It will be of interest to investigate whether endogenous alkyl-PTEs can be detected, if so, how backbone alkylation products contribute to gene regulation in *E. coli* cells.

REFERENCES

1. Ma, B., Villalta, P.W., Hochalter, J.B., Stepanov, I. and Hecht, S.S. (2019) Methyl DNA phosphate adduct formation in Lung tumor tissue and adjacent normal tissue of lung cancer patients. *Carcinogenesis*.

2. Wang, L.R., Chen, S., Vergin, K.L., Giovannoni, S.J., Chan, S.W., DeMott, M.S., Taghizadeh, K., Cordero, O.X., Cutler, M., Timberlake, S. *et al.* (2011) DNA phosphorothioation is widespread and quantized in bacterial genomes. *Proc. Nat. Acad. Sci. U. S. A.*, **108**, 2963-2968.