1-1-2010

Discovery and Identification of New DNA Photoproducts and Investigation of the Redox Mechanism of Human Apurinic/ Apyrimidinic Endonuclease 1 (hApe1) Protein

Dian Su
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Chemistry

Dissertation Examination Committee:
Dr. Michael L. Gross, Co-chair
Dr. John-Stephen A. Taylor, Co-chair
Dr. Robert E. Blankenship
Dr. Peter M. Burgers
Dr. Ron Bose
Dr. Jacob Schaefer
Dr. Joshua A. Maurer

DISCOVERY AND IDENTIFICATION OF NEW DNA PHOTOPRODUCTS AND INVESTIGATION OF THE REDOX MECHANISM OF HUMAN APURINIC/APYRIMIDINIC ENDONUCLEASE 1 (HAPE1) PROTEIN

by

Dian Su

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2010

Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Discovery and Identification of New DNA Photoproducts and Investigation of the Redox Mechanism of Human Apurinic/Apyrimidinic Endonuclease 1 (hApe1) Protein

by

Dian Su

Doctor of Philosophy in Chemistry

Washington University in St. Louis, 2010

Prof. Michael L. Gross, Co-chair

Prof. John-Stephen A. Taylor, Co-chair

Exposure of cells to ultraviolet light results in the formation of DNA photoproducts, many of which have been linked to skin cancer. The majority of these photoproducts arise from a photoreaction between two pyrimidines. Irradiation of double stranded B form DNA under physiological conditions mainly produces adjacent dimers. Nonadjacent photodimers of both intrastrand and interstrand types are much rarer. Intrastrand-type nonadjacent dimers form when one or more nucleotides between two pyrimidines become extrahelical owing to the formation of single strand DNA or a slipped structure. This allows the two pyrimidines to photodimerize in a colinear arrangement, as if the intervening sequence was not present.

Enzyme digestion, chemical hydrolysis, high performance liquid chromatography (HPLC) coupled with mass spectrometry (MS) have been been powerful tools for

identification and characterization of DNA photoproducts. Using these tools we discovered an unusual nonadjacent thymine dimer product formed between two T's with 5 nucleotides apart in a single-standed 14-mer ODN sequence. This novel and unexpected photoproduct arises from an interstrand-type reaction due to an unusual higher order folded DNA structure.

This finding led us to investigate the photochemistry of biologically relevant G-quadruplexes, which have high biological relevance in telomers. We studied the consequences of UV irradiation of G-quadruplex structures of a 22-mer telomeric sequence by utilizing site-specifically modified DNA, HPLC, CD, and nuclease P1 (NP1) coupled with MS. A large number of specific anti thymine dimers were formed between two T's located in different loops of the telomeric G-quadruplexes. If these unique nonadjacent anti thymine dimer photoproducts also form in vivo, they may constitute a previously unrecognized type of DNA photodamage that interferes with telomere replication and present a unique challenge to DNA repair.

We also describe the characterization of a new photoproduct ^{m}CA*, whose deamination product has the same structure as the well-known TA* photoproduct. The structure proof was done with a combination of NP1 digestion, HPLC correlation, and MS^n techniques.

A second part of the thesis deals with a DNA-binding protein. Human apurinic/apyrimidinic endonuclease 1 (hApe1) contributes to DNA repair by cleaving apurinic/apyrimidinic sites of damaged DNA. Additionally, hApe1 is a bifunctional protein functioning as a redox factor that simulates DNA binding activity of several transcription factors that are involved in cancer promotion and progression including AP-

1, P53 and others. The underlying mechanism of its redox activity, however, is poorly understood. One proposed mechanism is the disulfide exchange between hApe1 and the transcription factors. It's also hypothesized that hApe1 undergoes conformational changes that allows it to interact with the transcription factors. A quinone derivative E3330, (2*E*)-3-[5-(2,3-dimethoxy-*o*-methyl-1,4-benzoquinoyl)]-2-nonyl-2-propenoic acid, is a specific redox-activity inhibitor for hApe1. We used MS-based methods to probe the structure, dynamics, and interactions of HApe1 with ligands. We used amide hydrogen/deuterium exchange (HDX) and *N*-ethylmaleimide (NEM) labeling coupled MS analysis to probe the interaction of hApe1 and E3330. Our study suggests there is a minor amount of a locally unfolded form existing in equilibrium with the major folded form of hAp1. This locally unfolded form interacts with E3330 and is stabilized, allowing formation of disulfide bonds (C65-C93, C65-C99, C93-C99) in the redox active domain. A previous cell-based assay suggests that C65 is essential and thatC93 and C99 are also involved in the redox activity of hApe1. Taken together, we propose that the locally unfolded form is responsible for the redox function of hApe1 via a thiol/disulfide mechanism.

ACKNOWLEDGMENTS

It is my pleasure to thank those who made this thesis possible.

I would like to express my sincere thanks to my two wonderful advisors Prof. Michael L. Gross and Prof. John-Stephen A. Taylor. It is such an honor for me to work with both of them over the past five years at Washington University. I owe my deepest gratitude to them with their enthusiastic guidance, sound advice, perceptive criticisms, and constant encouragement. I thank my advisors for providing me the unique opportunity to work on challenging projects and giving me ample room to explore research possibilities and pursue ideas on my own. I particularly thank them for their critical push and inspiration during the tough time of my research projects, without which I would not have gone this far.

I owe my deepest gratitude to Andre d'Avignon, Henry W. Rohrs, Bich Vu, for their fantastic friendship and unconditional help. I thank Ilan Geerlof-Vidaysky and Maida Geerlof-Vidaysky for adopting me as their family member, which allowed me to feel like home in Saint Louis. Thanks to Drs. Ed Hiss and Alfred Hortmann for treating me as a friend.

I would like to acknowledge gratefully Bich Vu, Richard Huang, Hao Zhang, Weidong Cui, Don Rempel, Huafeng Fang, Ajay Kshetry and Jeff Kao for their contribution and help during my thesis research. I also thank all my colleagues in the Gross group especially, Justin Sperry, Adam Brustkern, Jiawei Chen, Lisa Jones, Joyce Neff, Daryl Gibln, James Walters and those of the Taylor group especially, Xuan (Adele) Yue, Dan Gu, Jianfeng Cai, Gang Sheng, Zhenghui Wang, Ziyan Zhang and Vincent

Cannistraro. I also thank Jianzhong Wen from the Blankenship lab. Their friendship and professional collaboration means a great deal to me.

My special thanks go to our collaborator Dr. Millie M. Georgiadis with Sarah Delaplane from department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana.

I wish to thank my Dissertation Advisory and Examination Committee members, Dr. Robert E. Blankenship, Dr. Peter M. Burgers, Dr. Ron Bose, Dr. Jacob Schaefer, and Dr. Joshua A. Maurer for their reading my thesis and proving valuable advice.

This thesis is dedicated to my parents, sister and my boyfriend Hao Zhang. I can never overstate my gratitude to them for their endless love, encouragement and support. They made me who I am today. My parents are my true heroes always and forever.

This research was supported by NIH Grants CA40463, CA114571 and by the by the Washington University NIH Mass Spectrometry Resource (Grant No. P41 RR000954) and Washington University NMR facility (Grant No. RR1571501).

TABLE OF CONTENTS

Chapter 3

Photocrosslinking of Human Telomeric G-quadruplex Loops by *Anti* Cyclobutane Thymine Dimer Formation

Chapter 4

mCA*: A New Photoproduct of 5' -Methylcytosine and Adenine Characterized by HPLC and Mass Spectrometry

Chapter 5

A Mass Spectrometric Investigation of the Interaction of hApe1 with the Redox Activity Inhibitor, E3330

LIST OF TABLES

Chapter 2

Structure Determination of an Interstrand-Type *Cis-Anti* Cyclobutane Thymine Dimer Produced in High Yield by UVB Light in an Oligodeoxynucleotide at Acidic pH

Chapter 3

Photocrosslinking of Human Telomeric G-quadruplex Loops by *Anti* Cyclobutane Thymine Dimer Formation

Chapter 4

mCA*: A New Photoproduct of 5' -Methylcytosine and Adenine Characterized by HPLC and Mass Spectrometry

LIST OF FIGURES

Chapter 1

Introduction

Chapter 2

Structure Determination of an Interstrand-Type *Cis-Anti* Cyclobutane Thymine Dimer Produced in High Yield by UVB Light in an Oligodeoxynucleotide at Acidic pH

Chapter 3

Photocrosslinking of Human Telomeric G-quadruplex Loops by *Anti* Cyclobutane Thymine Dimer Formation

Chapter 4

mCA*: A New Photoproduct of 5' -Methylcytosine and Adenine Characterized by HPLC and Mass Spectrometry

Chapter 5

A Mass Spectrometric Investigation of the Interaction of hApe1 with the Redox Activity Inhibitor, E3330

LIST OF SCHEMES

Chapter 1

Introduction

Chapter 2

Structure Determination of an Interstrand-Type *Cis-Anti* Cyclobutane Thymine Dimer Produced in High Yield by UVB Light in an Oligodeoxynucleotide at Acidic pH

Chapter 4

mCA*: A New Photoproduct of 5' -Methylcytosine and Adenine Characterized by HPLC and Mass Spectrometry

Chapter 5

A Mass Spectrometric Investigation of the Interaction of hApe1 with the Redox Activity Inhibitor, E3330

Chapter 1

Introduction

This thesis has two parts: discovery and identification of New DNA photoproducts (part 1) and investigation of the redox mechanism of human apurinic/apyrimidinic endonuclease 1 (hApe1) protein (part 2). In this introductory chapter, we will describe the types of DNA photodamage and summarize three new classes of photoproducts that we discovered by a combination of several analytical methods. In the second part, we investigated an important interaction of hApe1 protein with an inhibitor E3330, (2E)-3-[5-(2,3-dimethoxy-o-methyl-1,4-benzoquinoyl)]-2-nonyl-2-propenoic acid, by mass spectrometry (MS)-based methods.

Part 1: Discovery and Identification of New DNA Photoproducts

UVB overlaps the upper end of the DNA absorption spectra and is the range mainly responsible for skin cancer through direct photochemical damage to DNA. Up to now, the most prevalent UV photoproducts induced by UVB or UVC irradiation are found at adjacent pyrimidine sites, giving cyclobutane dimers, (6-4), and their Dewar valence isomers (Figure 1.1) (*1-5*). Other rare UV photoproducts, including the spore photoproduct, TA*, A=A, and AA* are also well known (Figure 1.1) (*6-16*). If not repaired, the photoproducts may interfere with DNA replication or transcription. If bypassed by a DNA polymerase, they may lead to mutations by inserting an incorrect complementary base in the next replication cycle.

Nonadjacent photodimers of both intrastrand and interstrand types are much rarer. Intrastrand-type nonadjacent dimers form when one or more nucleotides between two pyrimidines become extrahelical, owing to the formation of single strand DNA or a slipped structure, thereby allowing the two pyrimidines to photodimerize in a colinear arrangement, as if the intervening sequence were not present (Scheme 1.1). Because

nonadjacent dimer formation effectively shortens the DNA template, this type of photoproduct has been implicated in the formation of UV-induced deletion mutations (*17*). The formation of this interstrand photoproduct also suggests that such photoproducts might also form in biologically important, higher order DNA structures including between the loops of G-quartet structures in telomeres and promoters and between strands in i-motif structures (*18, 19*).

In part I of the thesis, we describe the discovery of new nonadjacent and adjacent DNA photoproducts by using a combined approach of enzymatic and/or chemical degradation, HPLC, NMR, and mass spectrometry.

In chapter 2, we report on an unusual nonadjacent photoproduct that is formed in high yield when irradiating the single-stranded ODN d(GTATCATGAGGTGC) with UVB light at acidic pH(*20*). We obtained evidence from HPLC comparisons, exo- and endo-nuclease coupled MS, and from NMR that this photoproduct is an interstrand-type (see scheme 1.1 for definition of "interstrand-type") of nonadjacent *cis-anti* cyclobutane thymine dimer formed between T2 and T7. This discovery suggests that some folded structure of the short single-stranded oligomer must exist and allow the photocrosslink to occur. This hypothesis led to an extensive study of the photochemistry of human telomeric G-quadruplex structures (Chapter 3).

In chapter 3, we describe the discovery of inter-loop photocrosslinking of a 22-mer human telomeric sequence d[AGGG(TTAGGG)$_3$](*21*). The product forms via a *cis-anti* thymine nonadjacent dimer. Owing to the presence of multiple reactant thymines, it is difficult to locate the sites of nonadjacent thymine dimerization. We implemented an efficient solution for accurately locating the thymine dimers by replacing specific

thymines with the photoequivalent uracils, followed by enzyme-coupled HPLC and tandem MS analysis.

In chapter 4, we report a new photoproduct ^{m}CA* between ^{m}C and A(*22*). We show how comparison of HPLC properties with those of authentic materials and of tandem MS can be used to determine the deamination product of the ^{m}C<>A, T<>A, has the same stereochemistry as the known photoproduct of TA*.

Part 2: Investigation of the Redox Mechanism of HApe1 Protein

HApe1 is a multi-functional protein (Figure 1.2) (*23-29*). It repairs DNA by cleaving apurinic/apyrimidinic sites of damaged DNA. Additionally, hApe1 functions as a redox factor to simulate DNA binding activity of several transcription factors (e.g., AP-1, P53) that are involved in cancer promotion and progression. However, the underlying mechanism of its redox activity is poorly understood. A quinone derivative E3330 inhibits the redox ability of hApe1 (*30-33*).

In chapter 5, we report our probe of the interaction between ApeΔ40, a truncated form lacking the N-terminal 40 residues of hApe1, and its redox-activity inhibitor E3330 with amide hydrogen/deuterium (H/D) exchange and *N*-ethylmaleimide labeling. Most of the analysis is with mass spectrometry, providing an insight on the redox mechanism of hApe1. On the basis of our results, we proposed that a locally unfolded form exists, which interacts with E3330 and exists in equilibrium with the folded form. Interaction with E3330 allows disulfide bonds to form in the redox domain of hApe1. The locally unfolded form is thereby proposed to be the potential redox-active form.

References

(1) Douki, T. and Cadet, J. (1994) Formation of cyclobutane dimers and (6-4) photoproducts upon far-UV photolysis of 5-methylcytosine-containing dinucleotide monophosphates. *Biochemistry 33*, 11942-11950.

(2) Pfeifer, G. P. (1997) Formation and processing of UV photoproducts: effects of DNA sequence and chromatin environment. *Photochem. Photobiol. 65*, 270-283.

(3) Shetlar, M. D., Basus, V. J., Falick, A. M. and Mujeeb, A. (2004) The cyclobutane dimers of 5-methylcytosine and their deamination products. *Photochem. Photobiol. Sci. 3*, 968-979.

(4) Celewicz, L., Mayer, M. and Shetlar, M. D. (2005) The photochemistry of thymidylyl-(3'-5')-5-methyl-2'-deoxycytidine in aqueous solution. *Photochem. Photobiol. 81*, 404-418.

(5) Taylor, J. S. (2006) In *DNA Damage Recognition* (Siede, W., Kow, Y. W. and Doetsch, P. W., Eds.) pp 67-94, Taylor and Francis Group, New York.

(6) Douki, T., Court, M. and Cadet, J. (2000) Electrospray-mass spectrometry characterization and measurement of far-UV-induced thymine photoproducts. *J. Photochem. Photobiol. B: Biol. 54*, 145-154.

(7) Douki, T. and Cadet, J. (2003) Formation of the spore photoproduct and other dimeric lesions between adjacent pyrimidines in UVC-irradiated dry DNA. *Photochem. Photobiol. Sci. 2*, 433-436.

(8) Douki, T., Laporte, G. and Cadet, J. (2003) Inter-strand photoproducts are produced in high yield within A-DNA exposed to UVC radiation. *Nucleic Acids Res. 31*, 3134-3142.

(9) Bose, S. N., Davies, R. J., Sethi, S. K. and McCloskey, J. A. (1983) Formation of an adenine-thymine photoadduct in the deoxydinucleoside monophosphate d(TpA) and in DNA. *Science 220*, 723-725.

(10) Bose, S. N., Kumar, S., Davies, R. J., Sethi, S. K. and McCloskey, J. A. (1984) The photochemistry of d(T-A) in aqueous solution and in ice. *Nucleic Acids Res. 12*, 7929-7947.

(11) Kumar, S., Sharma, N. D., Davies, R. J., Phillipson, D. W. and McCloskey, J. A. (1987) The isolation and characterisation of a new type of dimeric adenine photoproduct in UV-irradiated deoxyadenylates. *Nucleic Acids Res. 15*, 1199-1216.

(12) Sharma, N. D. and Davies, R. J. (1989) Extent of formation of a dimeric adenine photoproduct in polynucleotides and DNA. *J. Photochem. Photobiol. B 3*, 247-258.

(13) Kumar, S., Joshi, P. C., Sharma, N. D., Bose, S. N., Jeremy, R., Davies, H., Takeda, N. and McCloskey, J. A. (1991) Adenine photodimerization in deoxyadenylate sequences: elucidation of the mechanism through structural studies of a major d(ApA) photoproduct. *Nucleic Acids Res. 19*, 2841-2847.

(14) Zhao, X., Nadji, S., Kao, J. L. and Taylor, J. S. (1996) The structure of d(TpA), the major photoproduct of thymidylyl-(3'-5')-deoxyadenosine. *Nucleic Acids Res. 24*, 1554-1560.

(15) Wang, Y., Taylor, J. S. and Gross, M. L. (2001) Isolation and mass spectrometric characterization of dimeric adenine photoproducts in oligodeoxynucleotides. *Chem. Res. Toxicol. 14*, 738-745.

(16) Davies, R. J., Malone, J. F., Gan, Y., Cardin, C. J., Lee, M. P. and Neidle, S. (2007) High-resolution crystal structure of the intramolecular d(TpA) thymine-adenine photoadduct and its mechanistic implications. *Nucleic Acids Res. 35*, 1048-1053.

(17) Nguyen, H. T. and Minton, K. W. (1988) Ultraviolet-induced dimerization of non-adjacent pyrimidines. A potential mechanism for the targeted -1 frameshift mutation. *J. Mol. Biol. 200*, 681-693.

(18) Dai, J., Carver, M. and Yang, D. (2008) Polymorphism of human telomeric quadruplex structures. *Biochimie 90*, 1172-1183.

(19) Qin, Y. and Hurley, L. H. (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie 90*, 1149-1171.

(20) Su, D. G., Kao, J. L., Gross, M. L. and Taylor, J. S. (2008) Structure determination of an interstrand-type cis-anti cyclobutane thymine dimer produced in high yield by UVB light in an oligodeoxynucleotide at acidic pH. *J. Am. Chem. Soc. 130*, 11328-11337.

(21) Su, D. G., Fang, H., Gross, M. L. and Taylor, J. S. (2009) Photocrosslinking of human telomeric G-quadruplex loops by anti cyclobutane thymine dimer formation. *Proc. Natl. Acad. Sci. USA 106*, 12861-12866.

(22) Su, D. G., Taylor, J. S. and Gross, M. L. (2010) A new photoproduct of 5-methylcytosine and adenine characterized by high-performance liquid chromatography and mass spectrometry. *Chem. Res. Toxicol. 23*, 474-479.

(23) Evans, A. R., Limp-Foster, M. and Kelley, M. R. (2000) Going APE over ref-1. *Mutat. Res. 461*, 83-108.

(24) Fritz, G. (2000) Human APE/Ref-1 protein. *Int J Biochem Cell Biol 32*, 925-929.

(25) Dyrkheeva, N. S., Khodyreva, S. N. and Lavrik, O. I. (2007) Multifunctional human apurinic/apyrimidinic endonuclease 1: the role of additional functions. *Mol. Biol. 41*, 450-466.

(26) Bhakat, K. K., Mantha, A. K. and Mitra, S. (2009) Transcriptional regulatory functions of mammalian AP-endonuclease (APE1/Ref-1), an essential multifunctional protein. *Antioxid. Redox Signal 11*, 621-638.

(27) Jeon, B. H. and Irani, K. (2009) APE1/Ref-1: versatility in progress. *Antioxid. Redox Signal 11*, 571-574.

(28) Tell, G., Quadrifoglio, F., Tiribelli, C. and Kelley, M. R. (2009) The many functions of APE1/Ref-1: not only a DNA repair enzyme. *Antioxid. Redox Signal 11*, 601-620.

(29) Luo, M., He, H., Kelley, M. R. and Georgiadis, M. (2009) Redox Regulation of DNA Repair: Implications for Human Health and Cancer Therapeutic Development. *Antioxid. Redox Signal*.

(30) Shimizu, N., Sugimoto, K., Tang, J., Nishi, T., Sato, I., Hiramoto, M., Aizawa, S., Hatakeyama, M., Ohba, R., Hatori, H., Yoshikawa, T., Suzuki, F., Oomori, A., Tanaka, H., Kawaguchi, H., Watanabe, H. and Handa, H. (2000) High-performance affinity beads for identifying drug receptors. *Nat. Biotechnol. 18*, 877-881.

(31) Zou, G. M., Luo, M. H., Reed, A., Kelley, M. R. and Yoder, M. C. (2007) Ape1 regulates hematopoietic differentiation of embryonic stem cells through its redox functional domain. *Blood 109*, 1917-1922.

(32) Zou, G. M. and Maitra, A. (2008) Small-molecule inhibitor of the AP endonuclease 1/REF-1 E3330 inhibits pancreatic cancer cell growth and migration. *Mol. Cancer Ther. 7*, 2012-2021.

(33) Nyland, R. L., Luo, M., Kelley, M. R. and Borch, R. F. Design and synthesis of novel quinone inhibitors targeted to the redox function of apurinic/apyrimidinic endonuclease 1/redox enhancing factor-1 (Ape1/ref-1). *J. Med. Chem. 53*, 1200-1210.

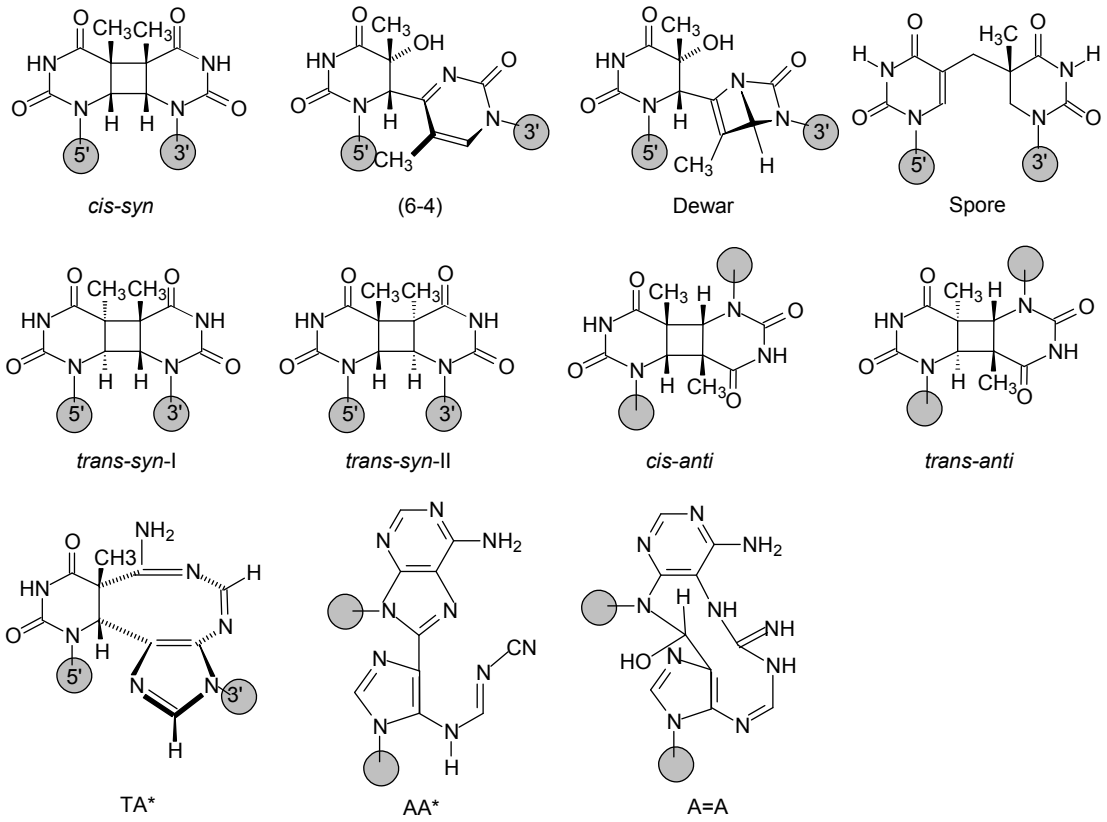


Figure 1.1. Structures of the DNA photoproducts.

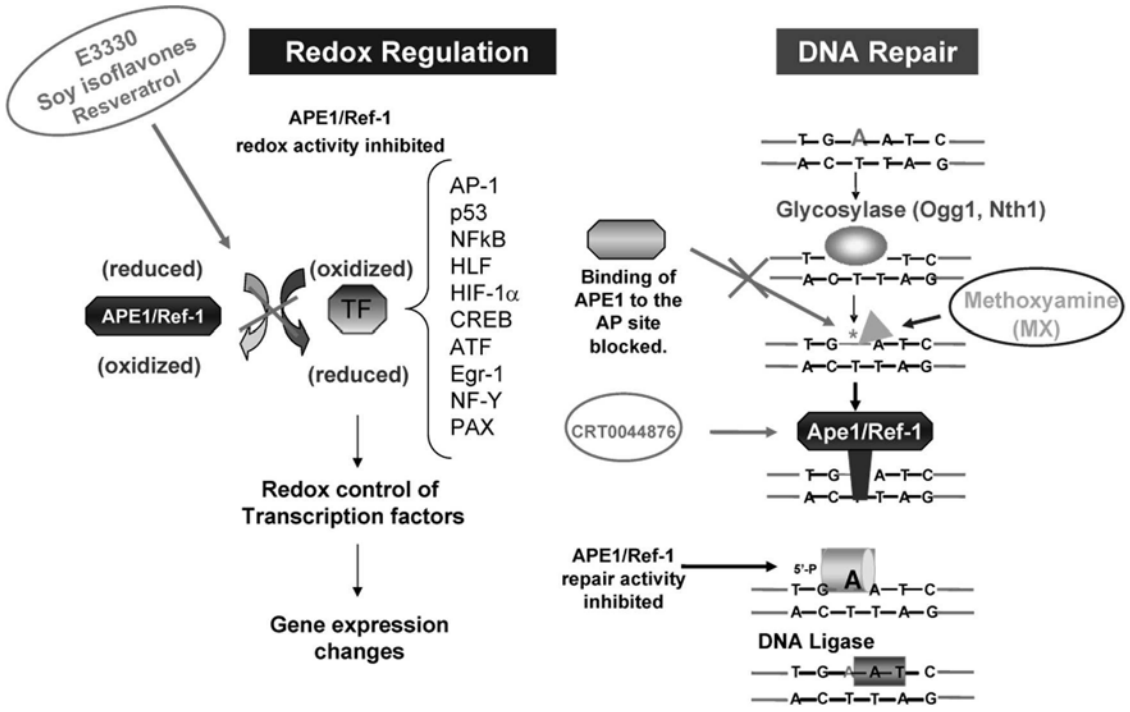
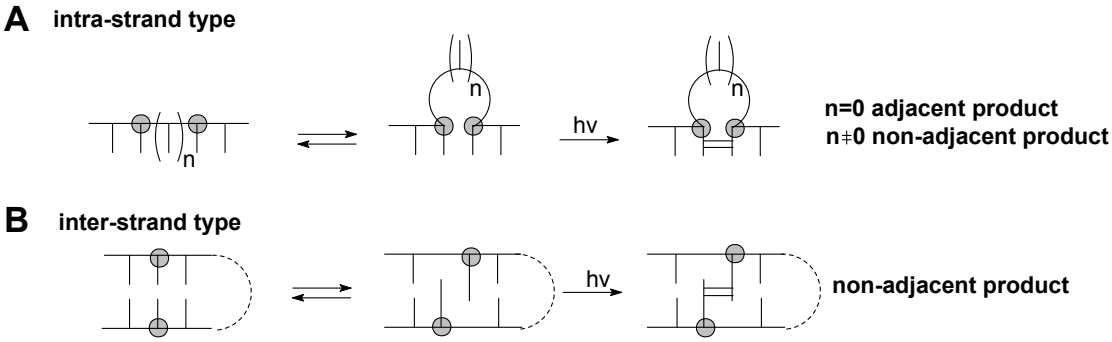


Figure 1.2. The two major functions of hApe1: Redox regulatory/signaling and DNA repair. Through its redox function, Ape1 regulates gene expression by modifying the redox status of some transcription factors which are involved in variety of cancer processes. Small molecules that block Ape1 redox function are shown in ovals. In addition to its redox function, Ape1 plays a critical role in the BER DNA repair pathway as an AP endonuclease, which processes the AP sites. Blocking AP sites using methoxyamine (MX) and/or Ape1 directly using Ape specific inhibitors such as CRT0044876 may decrease DNA repair and lead to tumor cell death.

(Reproduced from Luo et al. Antioxid. Redox Signal, 2009, in press.)



Scheme 1.1. Intra- (A) and inter- (B) strand-type photoreactions, both of which lead to nonadjacent photoproducts, except when n = 0 for A, which results in an adjacent photoproduct.

Chapter 2

Structure Determination of an Interstrand-Type *Cis-Anti* Cyclobutane Thymine Dimer Produced in High Yield by UVB Light in an Oligodeoxynucleotide at Acidic pH

Abstract

The unusual structural forms of telomere DNA, which protect the ends of UVB irradiation of DNA produces photodimers in adjacent DNA bases and on rare occasions in nonadjacent bases. UVB irradiation (312 nm) of d(GTATCATGAGGTGC) gave rise to an unknown DNA photoproduct in approximately 40% yield at acidic pH of about 5. This product has a much shorter retention time in reverse phase HPLC compared to known dipyrimidine photoproducts of this sequence. A large upfield shift of two thymine H6 NMR signals and photoreversion to the parent ODN upon irradiation with 254 nm light indicates that the photoproduct is a cyclobutane thymine dimer. Exonuclease-coupled MS assay establishes that the photodimer forms between T2 and T7, which was confirmed by tandem mass spectrometric MS/MS identification of the endonuclease P1 digestion product pd(T2[A3])=pd(T7[G8]). Acidic hydrolysis of the photoproduct gave a product with the same retention time on reverse phase HPLC and the same MS/MS fragmentation pattern as authentic Thy[*c,a*]Thy. 2D NOE NMR data are consistent with a *cis-anti* cyclobutane dimer between the 3'-sides of T2 and T7 in *anti* glycosyl conformations that had to have arisen from an interstand type reaction. In addition to pH dependency, the photoproduct yield is highly sequence specific and concentration dependent, indicating that it results from a higher order folded structure. The efficient formation of this interstrand-type photoproduct suggests the existence of a new type of folding motif and the possibility that this type of photoproduct might also form in other folded structures, such as G-quadruplexes and i-motif structures which can be now studied by the methods described.

Introduction

Exposure of cells to ultraviolet light results in the formation of DNA photoproducts, many of which have been linked to skin cancer (*1-4*). The majority of these photoproducts arise from a photoreaction between two pyrimidines, the structure and stereochemistry of which depend on the conformation of the two bases at the time of irradiation. Irradiation of double stranded B form DNA under physiological conditions mainly produces *cis-syn* cyclobutane pyridine dimers (CPDs), pyrimidine (6-4) pyrimidone photoproducts, and their Dewar valence isomers (Figure 2.1A) (*5-9*). Minor amounts of the *trans-syn*-I cyclobutane dimer have been detected in irradiated samples of native double stranded DNA and presumably form from single-stranded sections of DNA in which the thymine base has rotated into a *syn*-glycosyl conformation prior to photodimerization. A form DNA, on the other hand, appears to suppress these photoproducts and instead favors formation of the spore photoproduct (*10, 11*).

Nonadjacent photodimers of both intrastrand and interstrand types are much rarer. Intrastrand-type nonadjacent dimers form when one or more nucleotides between two pyrimidines become extrahelical, due to the formation of single strand DNA or a slipped structure, thereby allowing the two pyrimidines to photodimerize in a colinear arrangement, as if the intervening sequence was not present (Figure 2.1B). Because nonadjacent dimer formation effectively shortens the DNA template, this type of photoproduct has been implicated in the formation of UV-induced deletion mutations (*12*). Nonadjacent thymine dimers having a predominantly *cis-syn* configuration form in $d(GT)_n$ upon germicidal lamp irradiation (λ_{max} 254 nm) (*12*) and near-ultraviolet irradiation (λ_{max} 310 nm) (*13*). A nonadjacent *cis-syn* cyclobutane dimer between the two

Ts of TCT was subsequently prepared in good yield by triplet-sensitized irradiation of a bulged loop duplex DNA structure formed between the 13-mer d(GAGTATCTATGAG) and the 12-mer d(CTCATAATACTC) (*14*).

Interstrand-type DNA photoproducts arise from reactions between bases on opposing strands of duplex DNA in a parallel or antiparallel orientation (Figure 2.1C) and have not been detected under native, aqueous conditions. An early study showed that d(GT)n•(CA)n could be photo-cross-linked in ethanolic solutions or in aqueous solutions containing a high concentration of manganese (*15*). More recently, UVC irradiation of freeze-dried and alcoholic solutions of calf thymus DNA have been reported to give significant amounts of interstrand photoproducts (*11*). Enzymatic degradation of DNA irradiated under these conditions revealed the formation of nonadjacent spore photoproduct, as well as nonadjacent *cis-syn*, *cis-anti*, and *trans-anti* cyclobutane pyrimidine dimers. Nonadjacent photoproducts were distinguished from adjacent photoproducts by the absence of an internucleotide phosphate linkage following enzymatic digestion. The formation of the nonadjacent photoproducts was attributed to interstrand-type reactions between the bases (Figure 2.1B, bottom row), but some of these products could have arisen from intrastrand-type reactions (Figure 2.1B, top row). The method used, however, does not allow distinction between inter- and intrastrand nonadjacent photodimers.

In the course of irradiating d(GTATCATGAGGTGC) with UVB light to prepare a site-specific *cis-syn* TC cyclobutane dimer photoproduct for study, we discovered a new DNA photoproduct that was produced in unusually high yield at acidic pH. In this chapter, we describe physical, enzymatic, MS, and NMR evidence that this photoproduct is an interstrand-type nonadjacent *cis-anti* cyclobutane thymine dimer product formed

between T2 and T7. This novel and unexpected photoproduct can be explained to arise from an interstrand-type reaction due to an unusual higher order folded DNA structure. This is the first interstrand-type nonadjacent photoproduct to be characterized in DNA and the first to be produced efficiently in aqueous solution and points to the existence of a new type of DNA folding motif. The facile formation of this interstrand photoproduct also suggests that such a photoproduct may form in other folded structures, such as G-quadruplexes in telomeres and promoters, and i-motif structures (*16, 17*), which could now be studied by the methods described herein.

Material and Methods

Materials

Oligodeoxynucleotides were purchased from Integrated DNA Technologies, Inc. (IDT) (Coralville, Iowa). Snake venom phosphodiesterase (SVP) from *Crotalus adamanteus* was provided by Worthington Biochemical Corp. (Lakewood, NJ). Bovine spleen phosphodiesterase (BSP), nuclease P1 (NP1) from *Penicillium citrinum*, thymine, and 70% hydrogen fluoride in pyridine were all from Sigma (St. Louis, MO). Ammonium citrate and 3-hydroxypicolinic acid (3-HPA) for use as MALDI matrices were purchased from Fluka (Milwaukee, WI). 100% D_2O (99.96% atom %D) for NMR was from Aldrich (St. Louis, MO). Milli-Q (18.2 mΩ/cm) water was obtained from a Milli-Q water purification system (Millipore Corporation, Billerica, MA). HPLC solvents were from Fisher (Fair Lawn, NJ).

Instrumentation

HPLC separation and analysis were carried out on System Gold HPLC BioEssential with a binary gradient 125 pump and a diode array 168 detector (Beckman

Coulter, Inc., Fullerton, CA). An X-Bridge column (C18, 4.6 × 75 mm, 2.5 μm, 135 Å) from Waters Corporation (Milford, MA) was used for reverse-phase HPLC. UVB (280-320 nm) irradiation was carried out with two Spectroline XX-15B UV 15W tubes (312 nm) with peak UV intensity of 1150 $\mu W/cm^2$ at 25 cm and Longlife filter glass from Spectronics Corporation (Westbury, NY). UVC irradiation was carried out with a model UVG-254 mineralight lamp (254 nm, Ultra-Violet Products, Inc., San Gabriel, CA). MALDI mass spectra were collected on an Applied Biosystems 4700 tandem time-of-flight mass spectrometer (Applied Biosystems, Foster City, CA), and ESI mass spectra were collected on an Thermo Finnigan LCQ classic mass spectrometer (Thermo Fisher Scientific, Inc., Waltham, MA). NMR spectra were recorded with Varian Inova-600 (Varian Assoc., Palo Alto, CA) spectrometers, and the data were processed off-line with Varian VNMR software.

UV Irradiation of Oligodeoxynucleotides (ODNs)

ODNs from IDT were used without further purification. ODN d(GTATTATGAGGTGC) in Milli-Q water and d(GTATCATGAGGTGC) in 10 mM acetate buffer, pH 4.8, (50 μM) were purged with nitrogen for 5 min and placed in separate polyethylene Ziplock bags filled with nitrogen. The bags were placed on ice and irradiated for 2-2.5 h at a distance of about 1 cm from a UVB lamp. For the pH study the following buffers were used: 10 mM sodium citrate buffer (pH 3.6), sodium acetate buffer (pH 4.6, 4.8, 5.0, 5.2, 5.4, 5.6), 10 mM sodium 2-(N-morpholino)ethanesulfonate (MES) buffer (pH 5.7), 10 mM potassium phosphate buffer (pH 7.0), and 10 mM tris(hydroxymethyl)aminomethane-HCl (pH 7.8).

Reverse Phase HPLC Analysis and Purification of Products

Four gradients were used, all at a flow rate of 1 mL/min. Method A: 50 min 0-20% solvent B in solvent A (solvent A: 50 mM ammonium acetate, pH 6.8; Solvent B: 50% acetonitrile in 50 mM ammonium acetate, pH 6.8); method B: 50 min 0-15% solvent B in solvent A (solvent A: 5% acetonitrile in 50 mM triethylammonium acetate, pH 7.0; solvent B: 50% acetonitrile in 50 mM triethylammonium acetate, pH 7.0); method C: 50 min 0-20% solvent B in solvent A (solvent A: 50 mM triethylammonium acetate, pH 7.0; solvent B: 50% acetonitrile in 50 mM triethylammonium acetate, pH 7.0); method D: 100% Milli-Q water. The effluent was monitored at 260 nm for methods A-C, and 205 nm for method D.

Thermal and Hydrolytic Stability

Assay The thermal and hydrolytic stability of the photoproducts were determined with 100 μL of 50 μM UVB irradiated ODN in 10 mM sodium acetate buffer (pH = 4.8) in an Eppendorf microcentrifuge tube in a 37 ºC water bath.

Exonuclease-Coupled MS Sequencing

The ODN (150 pmol) was incubated with 0.01 units of SVP in 8 μL of 10 mM ammonium citrate (pH at 9.4) at 37 ºC or with 0.01 units of BSP in 8 μL of Milli-Q water at room temperature. Aliquots (0.5 μL) were removed at various time points and placed in dry ice for 1 min after which they were mixed with 0.5 μL of MALDI matrix solution. The premixed solution (0.5 μL) was spotted on an ABI-192-AB stainless steel plate and allowed to dry at ambient temperature. The matrix solution consisted of 70 mg of 3-hydroxypicolinic acid and 10 mg of ammonium citrate in 700 μL of 50% aqueous acetonitrile. MS spectra were acquired in the reflector positive ion mode with an accelerating voltage of 20 kV. Each spectrum was averaged by summing eight subspectra, each of which consisted of 125 shorts/subspectrum.

Exonuclease-Coupled MS Sequencing of Photoproduct I and Minor Photoproduct II with Bovine Intestinal Mucosa Phosphodiesterase (BIMP)

BIMP was used for 3'→5' sequencing to locate the 3'-thymine of the photodimer in both photoproduct I and minor II. In a typical sequencing experiment, the control or photodamaged ODN (150 pmol) in 7 μL Milli-Q water was mixed with 1 μL BIMP in 10 mM Tris-HCl buffer with 50% Glycerol (pH 7.5) and digested at 37 ºC. Aliquots were removed, spotted and finally analyzed by MALDI-TOF-MS as described for SVP sequencing.

NP1-Coupled MALDI-TOF-MS Assay

NP1 digestion was carried out by adding 1 μL of 1 unit/ μL aqueous NP1 to 1.8 nmol of ODN in 10 μL of Milli-Q water. After 5 h at room temperature, the solution was submitted to reverse phase HPLC using method C. The product eluting after the mononucleotides was isolated, dried, and redissolved in MilliQ water to give a 50 μL solution and mixed with an equal amount of the MALDI matrix. MS experiments were carried out in the reflector positive ion mode as described above. MS/MS experiments were carried out at medium pressure in the positive ion mode. The accelerating voltage was 8 kV and 15 kV for source 1 and source 2, respectively; air was the collision gas. Each spectrum was averaged by summing eight subspectra, each of which consisted of 125 shots/spectrum. Data were collected with the metastable ion suppressor on, and three subspectra used for the precursor whose intensity was optimized. The precursor ion was selected with a relative mass window that was 50-fold the specified resolution (full-width at half-maximum peak height) of the selected precursor.

Glycosidic Bond Hydrolysis

Acid-catalyzed glycosidic bond hydrolysis was carried out by adding 40 μL of 70% hydrogen fluoride in pyridine to 65 μg of ODN in a 1.5 mL polyethylene microcentrifuge tube. The solution was incubated at 37 ºC for 3 h, diluted to 1 mL with Milli-Q water, neutralized with 120 mg of calcium carbonate by vortexing for 5 min, and filtered through an Xpertek 13 mm 0.45 μm nylon syringe filter (P.J. Cobert Associates, Inc., St Louis, MO). The filtrate was evaporated in a SpeedVac (Thermo Savant, Holbrook, NY) and redissolved in 100 μL of Milli-Q water for HPLC analysis with method D. For the hydrolysis of authentic T[*c,s*]T and T[*t,s*]T dinucleotide intermediates, 23 μL of 70% hydrogen fluoride in pyridine was added to 45 μg of the dinucleotide intermediate in a microcentrifuge tube and incubated in 37 ºC water bath for 3 h. The hydrolysis solution was diluted to 600 μL of aqueous solution with Milli-Q water and neutralized by adding 69 mg of calcium carbonate. After filtering and drying, the residue was redissolved in 300 μL of Milli-Q water and analyzed by HPLC with method D. A mixture of authentic thymine cyclobutane dimers for comparison was prepared by irradiating 2 mM thymine in 5% acetone aqueous solution with UVB for 4 h as described above for the ODNs.

ESI-MS and MS/MS Assay

ESI-MS and MS/MS experiments were carried out in the positive ion mode with the ion-trap mass spectrometer. A solution of 80/20 (v/v) methanol/water with 0.1% trifluoroacetic acid (TFA) was used as the spray solvent. The spray voltage was 4.6 kV. The capillary voltage and temperature were 46 V and 200 ºC, respectively. MS/MS experiments were done by using collision-activated dissociation (CAD) with helium as collision gas. The mass window for precursor-ion selection was 3 m/z units. The collision energy was 30% of the maximum value, corresponding to a peak-to-peak excitation

voltage of 5 V. Approximately 20 scans were averaged for each spectrum. Each scan consisted of 3 microscans with a maximum injection time of 300 ms.

NMR Analysis

Approximately 0.26 μmol of photoproduct I purified by HPLC method B was dissolved in 100% D_2O for 1D and 2D NMR analysis. Proton chemical shifts were measured in parts per million (ppm) downfield from the H resonance of an internal 3-(trimethylsilyl)propionic acid (TSP) standard, and ^{31}P chemical shifts were measured using trimethyl phosphate as an external reference. The total correlation (TOCSY) spectra were recorded by using an MLEV-17 mixing sequence of 120 ms flanked by two 2 ms trim pulses. Phase-sensitive 2D spectra were obtained by employing the hypercomplex method. A 2 ×256 ×2048 data matrix with 16 scans per t1 increment was collected. Gaussian and sine-bell apodization functions were used in weighting the t2 and t1 dimensions, respectively. After two-dimensional Fourier transformation, the 2048 ×2048 frequency domain representation was phase and baseline corrected in both dimensions. The NOESY spectrum resulted in a 2 × 256 × 2048 data matrix with 32 scans per t1 increment. Spectra were recorded with 80, 150, 250, and 450 ms mixing times. The hypercomplex method was used to yield phase-sensitive spectra. The time domain data were zero-filled to yield a 2048 × 2048 data matrix and were processed with digital filtering to minimize the water signal. The proton-detected heteronuclear ^{1}H-^{31}P HSQC spectrum was recorded by using a 20 ms delay period in the sequence. The 90º ^{1}H pulse width was 6.5 s, and the 90º ^{31}P pulse width was 18 s. The proton spectral width was set to 1500 Hz, and the ^{31}P spectral width was set to 700 Hz. Phase-sensitive 2D spectra were obtained by employing the hypercomplex method. A 2 × 128 × 2048 data matrix with 64 scans per t1 value was collected. Gaussian line broadening was used in

weighting both the t2 and the t1 dimension. After two-dimensional Fourier transform, the spectra (256 × 2048 data points) were phase and baseline corrected in both dimensions.

Results and Discussion

Our interest in the photochemistry of the 14-mer d(GTATCATGAGGTGC) sequence was to prepare a T[*c,s*]C cyclobutane pyrimidine dimer for mechanistic studies of C→T mutations. The sequence has a single adjoining dipyrimidine site and is devoid of any stable secondary structure based on Watson Crick base pairing. Irradiation of the ODN with UVB light led to the formation of two major photoproducts, I and II, with retention times of 32.5 and 38.5 min, respectively, compared to 40.9 min for the starting ODN (Figure 2.2A). Photoproduct I was produced with an unexpectedly high yield and had an unusually short retention time compared to other known adjacent T=T photoproducts. No such unusual products were previously observed when the identical sequence containing a 5-methyl-CT site in place of the TC site had been irradiated (*18*). In trying to understand why this TC sequence but not the previous ^{m}CT sequence led to this unusual product, we discovered that the pH of the solution that was irradiated was much lower than 7. A systematic study of pH (Figure 2.2B) showed that photoproduct I was produced most efficiently at a pH of approximately 5, whereas photoproduct II was favored at higher pH. MALDI-TOF analysis of the two photoproducts showed that they both have the same mass (within 3 ppm) as the parent ODN, indicating that they are isomeric photoproducts.

Photoreversal of Photoproducts I and II

Irradiation of photoproducts with 254 nm can be used to distinguish various classes of DNA photoproducts. Cyclobutane dimers are reverted to the parent

nucleotides, whereas Dewar photoproducts are reverted back to the parent (6-4) photoproduct. When a pure sample of photoproduct II was irradiated with 254 nm light, it was converted to the (6-4) photoproduct which could be identified by its characteristic absorption at 325 nm (data not shown) (*19, 20*). This indicates that photoproduct II is likely to be a Dewar valence isomer of the (6-4) product of the TC site. This was confirmed by independently preparing the T(6-4)C photoproduct by 254 nm irradiation of the parent ODN and converting the (6-4) photoproduct to its Dewar valence isomer by irradiation with Pyrex-filtered medium pressure mercury arc light. When photoproduct I was irradiated with 254 nm light, however, it was converted to the parent ODN (Figure 2.2C), suggesting that it was a cyclobutane dimer, or some other [2 + 2] cycloadduct, but not of adjacent nucleotides, which would have a longer retention time.

Thermal and Hydrolytic Stability of Photoproduct I

C/^{m}C-containing *cis-syn* cyclobutane pyrimidine dimer photoproducts readily deaminate in neutral or acidic aqueous solution to afford U/T containing *cis-syn* dimers (*18, 21, 22*). This deamination reaction is accompanied by an increase in mass of 1 u. MS analysis of photoproduct I recovered after 1, 5, and 20 h of incubation at pH 4.8 at 37 ºC revealed that the molecular weight was unchanged, excluding the possibility that photoproduct I was a T[*c,s*]C product.

^{1}H NMR Analysis of Photoproduct I

Comparison of the aromatic region of the one-dimensional ^{1}H NMR spectra of photoproduct I and the parent ODN showed that the characteristic H6 protons of the two Cs were present in both cases (confirmed by 2D NMR). On the other hand, two aromatic TH6 protons were missing, consistent with cyclobutane dimer formation between two nonadjacent Ts which would shift the H6 signal upfield (Figure 2.3). The two upfield

shifted H6 protons could be identified through correlations to two methyl groups by 2D NMR spectroscopy (vide infra) and had shifts of 3.59 and 3.62 ppm (Table 2.1).

Assignment of proton chemical shifts of the parent ODN and photoproduct I by TOCSY and NOESY spectra

Cytosines were identified first by strong H5-H6 cross peaks (F, G, H, I of Figure 2.4 and Figure 2.5A). The unique AH8 (i) to CH1' (i-1) NOE cross peaks of A6-C5 (E and J, Figure 2.4A and Figure 2.5A) were then used to start the sequential analysis. The base and sugar moieties are connected through intra-residue H6/H8 (i)-H1'/H2'2" (i) cross-relaxation peaks while inter-residue H6/H8 (i)-H1'/H2'H2" (i)-H6/H8 (i+1) pathways yield sequential assignment along the strand (Figure 2.4 and Figure 2.5). Pyrimidine 5-CH3's of the parent ODN were assigned by the intra-residue H8-5CH3 cross peaks and were also confirmed by the unique A3-T4 and A6-T7 inter-residue AH8 (i)-T5CH_3 (i+1) NOE (Figure 2.4B). The assignment of H3' and H4' sugar protons were made from the combined H3' (i)-p-H4' (i+1) backbone ^{1}H-^{31}P correlation (Figure 2.6A and Figure 2.7A) with the sugar H1' to H4' spin propagation in TOCSY (Figure 2.6B and Figure 2.7B). For instance, by tracing TOCSY T7-H3' cross peak into the ^{1}H-^{31}P correlated spectrum (73' of Figure 2.6) can lead to the observation of the G8-H4' (84' of Figure 2.6). The resonance of G8-H3' (83' of Figure 2.6B) can then be assigned by further tracing the G8-H4' back into the TOCSY spectrum. Comparing Figure 2.6A and Figure 2.7A revealed that unusually downfield shifted 31P resonances at -2.7 and -2.15 ppm were observed in G1-p-T2 and A6-p-T7, respectively. Upfield shifted of T2, T7-H3' (at 4.56 and 4.62 ppm) and downfield shifted of G8-H1' (at 5.64 ppm) were also observed in Figure 2.7B suggesting that the formation of T2=T7 dimer is indeed the cause of the chemical shift changes in the neighboring residues.

DNA Ladder Sequencing

To determine which two of the four nonadjacent Ts had photodimerized we used a previously developed exonuclease-coupled MS method to map DNA damage sites. This method employs phosphodiesterase type I and II enzymes to degrade ODNs from the 3'- and 5'-directions, respectively, and MALDI to assay the fragments formed (*23-26*). This method can map abasic sites and adjacent photodimers, both of which terminate the progressive enzymatic cleavage (*27-29*). For 5'-exonucleolytic cleavage we used BSP, which removes 3'-nucleotide monophosphates from DNA with a free 5'-OH. The activity depends on contact with both the 5'- and 3'- bases flanking a phosphodiester bond, although contact with a 3'-base is not absolutely required (*30, 31*). Digestion of the parent ODN gave the expected fragments corresponding to the sequential loss of 3'-nucleotide monophosphates (+ H_2O - dNp) up to G11 in 5.5 min from which the majority of the sequence could be confirmed (Figure 2.8A). When the same procedure was carried out with BSP on photoproduct I (Figure 2.9A), excision of G1 was much slower than that of normal nucleotides and degradation terminated immediately after loss of only G1, indicating that one of the Ts involved in the unknown photodimer was T2. The reduced activity of BSP for cleaving G1 of the photoproduct containing DNA is consistent with the presence of a damaged base 3' to the phosphodiester bond which would interfere with binding to the enzyme. Subsequent cleavage is then inhibited by the presence of a damaged base that is 5' to the phosphodiester bond.

For 3'-exonucleolytic cleavage, we used SVP, which removes 5'-monophosphates from the 3'-end by binding to the base to the 3'-side of the phosphodiester bond to be cleaved (*30*). With the undamaged parental DNA, sequential excision of 5'-nucleotide monophosphates (+ H_2O - pdN) from the 3'-end was observed up to T4 in 5.5 min (Figure

2.8B). With photoproduct I, sequential loss of seven nucleotides up to and including G8 was observed within the same time frame, indicating that T7 was the other base involved in the photodimer which prevented the enzyme from binding T7 and cleaving to its 5'-side (Figure 2.9B). Upon extended incubation for more than 1 h, however, a loss of 295 u was observed, followed by successive losses corresponding to removal of pdT and pdC. The loss of 295 u can be attributed to the sum of masses resulting from rate-limiting endonucleotytic cleavage of the loop between the dimer at the A3T4 linkage (+ H_2O), presumably via binding to T4, followed by rapid 3'-exonucleolytic excision of pdA3 (+ H_2O - pdA) (Scheme 2.1). Endonuclease activity is intrinsic to venom phosphodiesterase (*32*) but is usually only observed in reactions with excess SVP or when long incubation times are used (*28, 33*). Subsequent loss of pT4 and pdC5 could be explained by the endonucleolytic cleavage steps after which no further enzymatic degradation occurred since there were no more undamaged bases 3' to a phosphodiester bond.

Support for the proposed cleavage pathway and mechanism comes from SVP degradation of the same photoproduct in a different sequence context in which A6 was replaced by a G (Table 2.2, mut6). As with the original sequence, degradation occurred rapidly until reaching T7, after which the same mass loss of 295 u corresponding to rate-limiting endonucleolytic cleavage of the loop between A3T4 followed by exonucleolytic cleavage of pdA was observed, followed by loss of pT4 and pdC5 (Scheme 2.1). Exonucleolytic cleavage of the photoproduct I in the original sequence context by BIMP, another phosphodiesterase I enzyme used for DNA sequencing (*34, 35*), also resulted in cleavage past T7. Unlike SVP, endonucleolytic cleavage of the loop took place at T4C5 and was followed by combined loss of pT by 3'-exonuclease activity and the 5'-phosphate of C5 by a contaminating phosphatase. This was then followed by loss of pdA by 3'-

exonuclease activity and then dC by endonuclease activity followed by rapid loss of the 5'-phosphate by the contaminating phosphatase (Figure 2.10).

NP1-Coupled MALDI-TOF MS Assay of Photoproduct I

To confirm that T2 and T7 are actually involved in the photoproduct, we used an endonuclease-coupled MS assay that we previously developed to identify and quantify adjacent dipyrimidine photoproducts (*18, 36, 37*). NP1 has both endonuclease activity and 3'-phosphatase activity (*38*). It has a binding pocket that is specific for the bases in DNA and, upon binding to the 5'-base, catalyzes hydrolysis of the phosphodiester bond to yield deoxynucleotides with a 3'-hydroxyl group (*30, 31, 39, 40*). Because the bases in dipyrimidine photoproducts are covalently linked together, NP1 cannot bind to either base of the photoproduct and hence cannot cut to the 3'-side of either nucleotide of the photoproduct. As a result, NP1 degradation of ODNs containing a dipyrimidine photodimer of adjacent bases gives a photodimer-containing trinucleotide of the form pd(Py=PyN). NP1 digestion of the parent ODN afforded, as expected, major products corresponding to monomers eluting at 3.0 (pdC), 10.0 (pdT), 10.6 (dG), 11.0 (pdG), and 15.6 (pdA) min when using HPLC method C. NP1 digestion of photoproduct I, however, gave an additional product eluting at 33.3 min. This component, after isolation, gave an $[M + H]^+$ by MALDI of m/z 1287.2347, which is within 3 ppm of the calculated value of 1287.2312 for pd(T[A])=pd(T[G]). Such a tetranucleotide can only result from NP1 digestion of a thymidine dimer formed between T2 and T7 or between T2 and T12, both of which have Ts followed by an A and a G. Involvement of T4 is ruled out because it is followed by C not G. Dimer formation between T2 and T12 is ruled out by SVP digestion which showed no inhibition at T12 but did show an unusual product when degrading past T7.To confirm the assignment of the NP1 digestion product to pd(T2[A3])=pd(T7[G8]),

we carried out MALDI-TOF-MS/MS (Figure 2.11A). Key product ions were as follows: $[Ade + H]^+$ (*m/z* 136.1), $[Gua + H]^+$ (*m/z* 152.1), $[pdA + H]^+$ (*m/z* 332.1), $[M - Ade + H]^+$ (*m/z* 1152.4), $[M - Gua + H]^+$ (*m/z* 1136.4), and $[M - pdG + H]^+$ (*m/z* 940.3). The other product ions are also consistent with the structure as shown in the more detailed fragmentation scheme in Figure 2.11B.

Stereochemistry of the Cyclobutane Dimer in Photoproduct I

To determine the relative stereochemistry of the nonadjacent dimer, the thymine dimer was released from photoproduct I by acid-catalyzed glycosidic bond hydrolysis and was compared with authentic samples of the four possible thymine dimers: *cis-syn*, *d,l-trans-syn*, *d,l-cis-anti*, and *d,l-trans-anti*. The authentic products were produced by sensitized UVB irradiation of thymine in aqueous acetone and separated by reverse phase HPLC according to a recently described procedure (*41*). Thy[*c,s*]Thy, Thy[*t,s*]Thy, Thy[*c,a*]Thy, and Thy[*t,a*]Thy were produced in nearly quantitative yield and eluted at 4.2, 5.2, 2.6, and 4.8 min, respectively, on reverse-phase HPLC with water (Figure 2.12B left). ESI-MS confirmed them to be thymine dimers as all gave $[M + H]^+$ ions of *m/z* 253.1. Irradiation of the mixture with 254 nm light prior to HPLC caused nearly quantitative photoreversal of all the dimers back to thymine (Figure 2.12B, right).

To release the thymine dimer from photoproduct I, we used hydrogen fluoride in pyridine which cleaves both the N-glycosidic and phosphodiester bonds of both undamaged and photodimer-containing DNA without degrading the bases or the photodimers (*42-44*). To test the method, we used authentic samples of d(T[*c,s*]T) and d(T[*t,s*]T) dinucleotide intermediates used in the synthesis of photodimer building blocks for DNA synthesis (*45-47*). The hydrolysis product of the *cis-syn* dimer eluted with the component corresponding to peak 2 and while the hydrolysis product of the *trans-syn*

dimer eluted with the component corresponding to peak 4, confirming our original assignment of the thymine dimers based on their relative retention time (data not shown) (*41*). Hydrolysis of an ODN 14-mer containing authentic *cis-syn* dimer d(GTAT[*c,s*]TATGAGGTGC), gave peak 5 (Figure 2.12C, left) which eluted with the same retention time as Thy[*c,s*]Thy. Photoreversal with 254 nm light resulted in disappearance of the *cis-syn* dimer peak 5 and an increase in the thymine peak (Figure 2.12C, right). Hydrolysis of photoproduct I, however, released a product peak 6 (Figure 2.12A, left) which had the same retention time as the authentic Thy[*c,a*]Thy, establishing it as a cis-anti thymine dimer. Photoreversal with 254 nm light, likewise, resulted in disappearance of the dimer peak 6, and an increase in the thymine peak (Figure 2.12A, right).To confirm further that photoproduct I has the *cis-anti* and not the *cis-syn* stereochemistry, we compared the tandem MS of the hydrolysate of photoproduct I with that of the *cis-syn* thymine dimer-containing ODN 14-mer, and the authentic Thy[*c,a*]Thy (peak 1) and Thy[*c,s*]Thy (peak 2). The thymine dimer molecular ion $[M + H]^+$ of *m/z* 253.1 is known to undergo different fragmentation for the *cis-syn* and *cis-anti* stereoisomers (*48*). MS/MS of the *cis-syn* stereoisomer produced a most abundant *m/z* 210.0 ion by loss of CHNO (43 u) whereas the *cis-anti* stereoisomer produced a most abundant ion of *m/z* 127.1, suggesting that reversion to thymine occurs under the collision activation conditions (Figure 2.13). Reversion of the *cis-anti,* but not the *cis-syn* isomer, during MS/MS is consistent with the thermal instability of the *anti* stereoisomers compared to the *syn* stereoisomers (*41, 49*). MS/MS of the precursor ion of *m/z* 253.1 produced by photoproduct I hydrolysate afforded the ion of *m/z* 127.1 characteristic of the *cis-anti* stereochemistry.

Absolute Stereochemistry of Photoproduct I by 2D NMR Analysis

We assigned proton chemical shifts of the parent ODN and photoproduct I by a standard sequential assignment strategy based on TOCSY and NOESY spectra (Figure 2.14 and Figures 2.4,2.6 and 2.7). The assignments confirm that the photodimer formed between T2 and T7. Owing to a loss of aromaticity of the T2 and T7 bases upon the formation of the cyclobutane dimer, the chemical shift of the H6 protons were significantly shifted upfield. As a result, intraresidue aromatic H6-CH3 cross-peaks were not observed for T2 and T7 but were observed for T4 and T12 (Figure 2.5B). Likewise, aromatic H6/H8 (i)-H1' (i-1) cross-peaks used in the sequential assignment of DNA were not observed for T2-G1 and T7-A6 (Figure 2.5A).

The methyl groups of the thymine dimer were assigned based on the observable G1 H1'/H3'-T2CH_3 and A6 H1'/H3'-T7CH_3 cross-peaks located in the most upfield shifted methyl region (Figure 2.14B,C). The H6s of the thymine dimer were then assigned by tracing two intraresidue H6-5CH3 cross-peaks between 3.5-3.7 ppm and 1.3-1.4 ppm (Figure 2.14D). The corresponding intraresidue T2/T7 H6-H2'2" NOEs were used to reestablish the sequential A3-T2 and G8-T7 pathway through the joint vertical line in the 1.8-2.0 ppm region (Figure 2.14A,D). In this way, T2H6 could be assigned to 3.59 ppm and T7H6 to 3.62 ppm. The observation of four strong T2/T7 H6-CH_3 cross-peaks (Figure 2.14D, dotted box a) is consistent with a *cis-anti* stereochemistry for the dimer for which each H6 proton is within NOE distance of each of the methyl groups (Figure 2.15A). A *cis-syn* dimer, on the other hand, generally shows only three cross-peaks because of the puckered ring conformation which results in one long, NOE inactive H-CH_3 distance. Likewise, a *trans-syn* or *trans-anti* dimer generally shows only two cross-peaks. The presence of T2H1'-T7CH_3 and T7H1'-T2CH_3 (Figure 2.14B, dotted box c) and T2H1'-T2H6 and T7H1'-T7H6 (Figure 2.14B, dotted box d) cross-peaks and the

absence of T2H1'-T7H6 and T7H1'-T2H6 cross-peaks is also consistent with an *anti* orientation of C5-C6 bonds of the dimer as illustrated for Structure I in Figure 2.15B.

Given the *cis-anti* stereochemistry of the cyclobutane dimer, there are still four possible structures of the T2[*c,a*]T7 photoproduct to consider (Figure 2.15B). Structures I and II would result from photodimerization of the two Ts in anti glycosyl conformations with either a 5'- or 3'-staggered stacking arrangement. Structures III and IV would result from photodimerization of the two Ts in a *syn* glycosyl conformation with either a 5'- or 3'-staggered stacking arrangement. The presence of a cross-peak between A3H8 and T7CH_3 (Figure 2.14A, dotted box e) is consistent with either structure I or III which places T7 near A3 (Figure 2.15B, Structure I). The presence of a strong intraresidue cross-peak between H6-H2'2" for both T2 and T7 (Figure 2.14D, dotted box b) at various mixing times ranging from 80 to 450 ms is consistent with an *anti* glycosyl conformation of structure I but not structure III (Figure 2.15B), strongly suggesting that the photoproduct has structure I.

Photoreversion of Minor I and Minor II

The UVB irradiation mixture of d(GTATCATGAGGTGC) produced at pH 3.6 was irradiated with 254 nm light for 30 min. HPLC assay showed that both minor I and minor II disappeared and were presumably converted to the parent ODN (Figure 2.16), suggesting that the two minor photoproducts are cyclobutane dimers.

NP1-Coupled MALDI-TOF MS Assay of Minor I and Minor II

NP1 digestion was carried out on the minor I and minor II photoproducts followed by MALDI-TOF-MS with default calibration. For the minor I photoproduct a molecular ion $[M+H]^+$ was detected at *m/z* 1247.4 corresponding to pd(T[A])=pd(T[C]) and for the minor II photoproduct a molecular ion was detected at *m/z* 1287.4

corresponding to pd(T[A])=pd(T[G]). Further MALDI-TOF-MS/MS experiments confirmed the assignment of the minor I digestion product by the presence of $[Cyt + H]^+$ at *m/z* 112.1 and $[Ade + H]^+$ at *m/z* 136.1 (Figure 2.17A). Likewise, assignment of the minor II digestion product was confirmed by the presence of $[Ade + H]^+$ at *m/z* 136.1 and $[Gua + H]^+$ at *m/z* 152.1 (Figure 2.17B). The results suggest that the thymine dimer of minor I is formed between T2 and T4 and that of minor II is between T2 and T7 or T2 and T12.

DNA Ladder Sequencing of Photoproduct I and Minor Photoproduct II with BIMP

Photoproduct minor II was seen having exactly the same digestion ladder pattern (Figure 2.18) as photoproduct I (Figure 2.10A) and could be explained in the same way (Figure 2.10B). After sequential removal of C14 to G8, exonuclease degradation was inhibited at T7. A mass loss of 366.1 u was observed immediately after loss of G8, which corresponds to addition of water by hydrolysis of the T4C5 phosphodiester bond by its minor endonuclease activity and then loss of T4 by its exonuclease activity along with 5'-HPO3 of C5 by the 5'-phosphatase activity that must contaminate the commercial preparation of BIMP. The enzyme then excised A3 by its 3'-exonuclease activity. Finally the degradation terminated at *m/z* 1127.2 after a loss of and dC5 by its endonuclease activity, and 5'-HPO3 of A6, indicating that the dimer of minor II is formed between T2 and T7.

Cyclobutane Dimer Stereochemistry of Minor I and Minor II

Two minor photoproducts, minor I and minor II (Figure 2.2A), were also analyzed by NP1-coupled MS/MS and exonuclease-coupled MS sequencing and by comparison of acid hydrolysates to the authentic thymine dimers (Figure 2.16-2.20). Hydrolysis of the minor I and minor II photoproducts with hydrogen fluoride in pyridine

released products that had the same retention time as the authentic *cis-syn* thymine dimer and as the authentic *trans-anti* thymine dimer, respectively (Figure 2.19). The molecular ions $[M + H]^+$ of the hydrolysates of both the minor I and minor II photoproducts were detected at *m/z* 253.1 by ESI-MS. ESI-MS/MS of the hydrolysate of the minor I photoproduct gave the same base peak of *m/z* 210.0 as with that of authentic Thy[*c,s*]Thy (Figure 2.20A). ESI-MS/MS of the hydrolysate of the minor II photoproduct, however, gave a base peak of *m/z* 208.0 and another intense peak at *m/z* 127.1 that was also observed for authentic Thy[*t,a*]Thy (Figure 2.20B). On the basis of these data, minor I was assigned to a T2[*c,s*]T4 nonadjacent dimer, and minor II to a T2[*t,a*]T7 nonadjacent dimer.

Photodimerization in Different Sequence Contexts and Concentrations

To determine the sequence specificity of the photodimerization reaction resulting in the T2[*c,a*]T7 product, we irradiated an additional ten ODNs corresponding to site-specific mutations within the loop region, and deletions of the sequence to the 3'-side of the photoproduct (Table 2.2) at pH 4.8. Only irradiation of mut6, in which the A preceding T7 was mutated to G produced the T2[*c,a*]T7 nonadjacent dimer in a similar high yield as the parent sequence (Figure 2.21). The requirements for having A3, T4, C5, G8 and the 3'-end of the sequence to produce the photoproduct in high yield indicates that a set of highly specific base pairing interactions are required to form the necessary folded structure.To determine whether the folded structure leading to the photoproduct was intramolecular or involved the assembly of two or more strands, we determined the photoproduct yield at two different concentrations. To ensure that the DNA in the two samples absorbed the same amount of light, we used samples that differed in concentration by a factor of 10 and placed them in quartz cuvettes with 0.1 and 1 cm

pathlengths, so that they each had an absorbance of 0.3 at 260 nm corresponding to absorption of 50% of the incident light. Thus 100 μL of a 20 μM solution at pH 4.8 in a 0.1 cm path length quartz cuvette and 1000 μL of a 2 μM solution in a 1 cm cuvette were irradiated side by side for 1 h with ice cooling to give the product in 32% and 14% yields, respectively, indicating that the folded structure involves two or more strands.

Conclusions

On the basis of the photoreactivity, spectroscopy, and spectrometry data, we determined the structure of the new photoproduct to be a nonadjacent *cis-anti* cyclobutane dimer formed between T2 and T7 with the absolute stereochemistry depicted by structure I (Figure 2.15). This is the first interstrand type of cyclobutane dimer to be isolated and characterized in DNA. There is no predictable stable Watson-Crick folded structure of d(GTATCATGAGGTGC) that can explain the high yield formation of this product. On the basis of our mutagenesis and concentration studies, we speculate that the sequence adopts a higher order folded structure at pH 5 by inter- and intramolecular base pairing between strands, which remains to be elucidated. The formation of this interstrand photoproduct also suggests that such type of photoproducts might also form in biologically important higher order DNA structures such as between loops of G-quartet structures in telomeres and promoters and between strands in i-motif structures (*16, 17*). The methods described herein could be used to identify and map such photoproducts which could be used as a probe to distinguish between various folding structures in solution and *in vivo*. Once the structure-activity relationships of interstrand photoproduct formation are better understood, it is conceivable that it could also be used for the

assembly and/or characterization of higher order DNA structures for nanotechnological applications.

Acknowledgment

We thank Jeff Lung-Fa Kao for his help with NMR experiments and data analysis. We also thank Ajay Kshetry for providing samples of authentic *cis-syn* and *trans-syn* dinucleotide intermediates. This research was supported by NIH Grant CA40463 and by the Washington University NIH Mass Spectrometry Resource (Grant No. P41 RR000954) and Washington University NMR facility (Grant No. RR1571501).

References

(1) Grossman, D. and Leffell, D. J. (1997) The molecular basis of nonmelanoma skin cancer: new understanding. *Arch. Dermatol. 133*, 1263-1270.

(2) Sarasin, A. (1999) The molecular pathways of ultraviolet-induced carcinogenesis. *Mutat. Res. 428*, 5-10.

(3) Cleaver, J. E. and Crowley, E. (2002) UV damage, DNA repair and skin carcinogenesis. *Front. Biosci. 7*, d1024-1043.

(4) Pfeifer, G. P., You, Y. H. and Besaratinia, A. (2005) Mutations induced by ultraviolet light. *Mutat. Res. 571*, 19-31.

(5) Cadet J.; Vigny, P. (1990) Photochemistry of nucleic acids, In *Bioorganic Photochemistry* (Morrison, H., Ed.) pp 1-272, Wiley, New York.

(6) Taylor, J.-S. (1995) DNA, sunlight and skin cancer. *Pure and Applied Chemistry 67*, 183-190.

(7) Begley, T. P. (1999) In *Comprehensive Natural Products Chemistry* (Poulter, C. D., Ed.) pp 371-399, Elsevier, New York.

(8) Cadet, J., Sage, E. and Douki, T. (2005) Ultraviolet radiation-mediated damage to cellular DNA. *Mutat. Res. 571*, 3-17.

(9) Taylor, J. S. (2006) In *DNA Damage Recognition* (Siede, W., Kow, Y. W. and Doetsch, P. W., Eds.) pp 67-94, Taylor and Francis Group, New York.

(10) Douki, T. and Cadet, J. (2003) Formation of the spore photoproduct and other dimeric lesions between adjacent pyrimidines in UVC-irradiated dry DNA. *Photochem. Photobiol. Sci. 2*, 433-436.

(11) Douki, T., Laporte, G. and Cadet, J. (2003) Inter-strand photoproducts are produced in high yield within A-DNA exposed to UVC radiation. *Nucleic Acids Res. 31*, 3134-3142.

(12) Nguyen, H. T. and Minton, K. W. (1988) Ultraviolet-induced dimerization of non-adjacent pyrimidines. A potential mechanism for the targeted -1 frameshift mutation. *J. Mol. Biol. 200*, 681-693.

(13) Nguyen, H. T. and Minton, K. W. (1989) Extensive photodimerization of non-adjacent pyrimidines. *J. Mol. Biol. 210*, 869-874.

(14) Lingbeck, J. M. and Taylor, J. S. (1999) Preparation and characterization of DNA containing a site-specific nonadjacent cyclobutane thymine dimer of the type implicated in UV-induced -1 frameshift mutagenesis. *Biochemistry 38*, 13717-13724.

(15) Love, J. D., Nguyen, H. T., Or, A., Attri, A. K. and Minton, K. W. (1986) UV-induced interstrand cross-linking of d(GT)n.d(CA)n is facilitated by a structural transition. *J. Biol. Chem. 261*, 10051-10057.

(16) Dai, J., Carver, M. and Yang, D. (2008) Polymorphism of human telomeric quadruplex structures. *Biochimie 90*, 1172-1183.

(17) Qin, Y. and Hurley, L. H. (2008) Structures, folding patterns, and functions of intramolecular DNA G-quadruplexes found in eukaryotic promoter regions. *Biochimie 90*, 1149-1171.

(18) Vu, B., Cannistraro, V. J., Sun, L. and Taylor, J. S. (2006) DNA synthesis past a 5-methylC-containing cis-syn-cyclobutane pyrimidine dimer by yeast pol eta is highly nonmutagenic. *Biochemistry 45*, 9327-9335.

(19) Hauswirth, W. and Wang, S. Y. (1973) Pyrimidine adduct fluorescence in UV irradiated nucleic acids. *Biochem. Biophys. Res. Commun. 51*, 819-826.

(20) Taylor, J. S., Lu, H. F. and Kotyk, J. J. (1990) Quantitative conversion of the (6-4) photoproduct of TpdC to its Dewar valence isomer upon exposure to simulated sunlight. *Photochem. Photobiol. 51*, 161-167.

(21) Lemaire, D. G. and Ruzsicska, B. P. (1993) Kinetic analysis of the deamination reactions of cyclobutane dimers of thymidylyl-3',5'-2'-deoxycytidine and 2'-deoxycytidylyl-3',5'-thymidine. *Biochemistry 32*, 2525-2533.

(22) Celewicz, L., Mayer, M. and Shetlar, M. D. (2005) The photochemistry of thymidylyl-(3'-5')-5-methyl-2'-deoxycytidine in aqueous solution. *Photochem. Photobiol. 81*, 404-418.

(23) Schuette, J. M., Pieles, U., Maleknia, S. D., Srivatsa, G. S., Cole, D. L., Moser, H. E. and Afeyan, N. B. (1995) Sequence analysis of phosphorothioate oligonucleotides via matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *J. Pharm. Biomed. Anal. 13*, 1195-1203.

(24) Bentzley, C. M., Johnston, M. V., Larsen, B. S. and Gutteridge, S. (1996) Oligonucleotide sequence and composition determined by matrix-assisted laser desorption/ionization. *Anal. Chem. 68*, 2141-2146.

(25) Smirnov, I. P., Roskey, M. T., Juhasz, P., Takach, E. J., Martin, S. A. and Haff, L. A. (1996) Sequencing oligonucleotides by exonuclease digestion and delayed extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Anal. Biochem. 238*, 19-25.

(26) Owens, D. R., Bothner, B., Phung, Q., Harris, K. and Siuzdak, G. (1998) Aspects of oligonucleotide and peptide sequencing with MALDI and electrospray mass spectrometry. *Bioorg. Med. Chem. 6*, 1547-1554.

(27) Zhang, L. K., Rempel, D. and Gross, M. L. (2001) Matrix-assisted laser desorption/ionization mass spectrometry for locating abasic sites and determining the rates of enzymatic hydrolysis of model oligodeoxynucleotides. *Anal. Chem. 73*, 3263-3273.

(28) Zhang, L. K., Ren, Y., Rempel, D., Taylor, J. S. and Gross, M. L. (2001) Determination of photomodified oligodeoxynucleotides by exonuclease digestion, matrix-assisted laser desorption/ionization and post-source decay mass spectrometry. *J. Am. Soc. Mass Spectrom 12*, 1127-1135.

(29) Zhang, L. K. and Gross, M. L. (2002) Location of abasic sites in oligodeoxynucleotides by tandem mass spectrometry and by a chemical cleavage initiated by an unusual reaction of the ODN with MALDI matrix. *J. Am. Soc. Mass Spectrom 13*, 1418-1426.

(30) Weinfeld, M., Liuzzi, M. and Paterson, M. C. (1989) Selective hydrolysis by exo- and endonucleases of phosphodiester bonds adjacent to an apurinic site. *Nucleic Acids Res. 17*, 3735-3745.

(31) Weinfeld, M., Soderlind, K. J. and Buchko, G. W. (1993) Influence of nucleic acid base aromaticity on substrate reactivity with enzymes acting on single-stranded DNA. *Nucleic Acids Res. 21*, 621-626.

(32) Pritchard, A. E., Kowalski, D. and Laskowski, M., Sr. (1977) An endonuclease activity of venom phosphodiesterase specific for single-stranded and superhelical DNA. *J. Biol. Chem. 252*, 8652-8659.

(33) Bourdat, A. G., Gasparutto, D. and Cadet, J. (1999) Synthesis and enzymatic processing of oligodeoxynucleotides containing tandem base damage. *Nucleic Acids Res. 27*, 1015-1024.

(34) Landt, M., Everard, R. A. and Butler, L. G. (1980) 5'-Nucleotide phosphodiesterase: features of the substrate binding site as deduced from specificity and kinetics of some novel substrates. *Biochemistry 19*, 138-143.

(35) Garcia-Diaz, M., Avalos, M. and Cameselle, J. C. (1991) Methanol esterification reactions catalyzed by snake venom and bovine intestinal 5'-nucleotide phosphodiesterases. Formation of nucleoside 5'-monophosphate methyl esters from guanosine 5'-triphosphate and other nucleoside 5'-polyphosphates. *Eur. J. Biochem. 196*, 451-457.

(36) Wang, Y., Taylor, J. S. and Gross, M. L. (1999) Nuclease P1 digestion combined with tandem mass spectrometry for the structure determination of DNA photoproducts. *Chem. Res. Toxicol. 12*, 1077-1082.

(37) Wang, Y., Gross, M. L. and Taylor, J. S. (2001) Use of a combined enzymatic digestion/ESI mass spectrometry assay to study the effect of TATA-binding protein on photoproduct formation in a TATA box. *Biochemistry 40*, 11785-11793.

(38) Shishido, K. and Ando, T. (1982) In *Nucleases* (Linn, S. M., Roberts, R. J., Ed.) pp 155-209, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

(39) Volbeda, A., Lahm, A., Sakiyama, F. and Suck, D. (1991) Crystal structure of Penicillium citrinum P1 nuclease at 2.8 A resolution. *EMBO J. 10*, 1607-1618.

(40) Romier, C., Dominguez, R., Lahm, A., Dahl, O. and Suck, D. (1998) Recognition of single-stranded DNA by nuclease P1: high resolution crystal structures of complexes with substrate analogs. *Proteins 32*, 414-424.

(41) Shetlar, M. D., Basus, V. J., Falick, A. M. and Mujeeb, A. (2004) The cyclobutane dimers of 5-methylcytosine and their deamination products. *Photochem. Photobiol. Sci. 3*, 968-979.

(42) Lipkin, D., Howard, F. B., Nowotny, D. and Sano, M. (1963) The Iodination of Nucleosides and Nucleotides. *J. Biol. Chem.*, 2249-2251.

(43) Lipkin, D., Phillips, B. E. and Abrell, J. W. (1969) The action ofhydrogen fluoride on nucleotides and other esters of phosphorus (V) acids. *J. Org. Chem.*, 1539-1547.

(44) Douki, T., Voituriez, L. and Cadet, J. (1995) Measurement of pyrimidine (6-4) photoproducts in DNA by a mild acidic hydrolysis-HPLC fluorescence detection assay. *Chem. Res. Toxicol. 8*, 244-253.

(45) Taylor, J. S., Brockie, I. R. and O'Day, C. L. (1987) A Building Block for the Sequence Specific Introduction of cis - syn Thymine Dimers into Oligonucleotides. Solid Phase Synthesis of TpT[c,s]pTpT. *J. Am. Chem. Soc.*, 6735-6742.

(46) Taylor, J. S. and Brockie, I. R. (1988) Synthesis of a trans-syn thymine dimer building block. Solid phase synthesis of CGTAT[t,s]TATGC. *Nucleic Acids Res. 16*, 5123-5136.

(47) Ordoukhanian, P. and Taylor, J.-S. (1997) Solid phase-supported thymine dimers for the construction of dimer-containing DNA by combined chemical and

enzymatic synthesis: a potentially general method for the efficient incorporation of modified nucleotides into DNA. *Nucleic Acids Res. 25*, 3783–3786.

(48) Douki, T., Court, M. and Cadet, J. (2000) Electrospray-mass spectrometry characterization and measurement of far-UV-induced thymine photoproducts. *J. Photochem. Photobiol. B: Biol. 54*, 145-154.

(49) Wang, M. N. K. a. S. Y. (1972) The photochemical mechanism of pyrimidine cyclobutyl dimerization. *Tetrahedron 28*, 945-957.

Table 2.1. Proton and ^{31}P chemical shifts of the parent ODN and photoproduct I in D_2O at 25 ^{o}C.

		G1	T2	A3	T4	C5	A6	T7	G8	A9	G10	G11	T12	G13	C14
H6/H8	parent	7.85	7.37	8.34	7.36	7.53	8.34	7.22	7.78	8.08	7.75	7.86	7.29	7.94	7.76
	cis-anti	7.87	3.59	8.18	7.69	7.52	8.16	3.62	7.84	8.12	7.74	7.86	7.31	7.94	7.75
H5/CH_3	parent		1.67		1.66	5.91		1.64					1.63		5.87
	cis-anti		1.29		1.85	5.95		1.38					1.65		5.87
H1’	parent	6.08	6.02	6.28	6.04	5.97	6.27	5.86	5.45	5.95	5.64	5.94	5.96	6.07	6.21
	cis-anti	6.22	5.94	6.33	6.34	6	6.22	6.01	5.64	5.97	5.67	5.97	6.04	6.11	6.24
H2’2’	parent	2.63	1.95	2.77	2.11	1.79	2.74	1.73	2.4	2.52	2.42	2.6	1.91	2.61	2.25
			2.27		2.34	2.26		2.11				2.66	2.29	2.72	2.39
	cis-anti	2.66	1.91	2.79	2.31	1.45	2.65	1.81	2.45	2.55	2.41	2.58	1.93	2.63	2.25
		2.81	1.96		2.51	2.13	2.79	1.9				2.67	2.29	2.73	2.37
H3’	parent	4.85	4.81	4.97	4.8	4.75	4.98	4.72	4.82	4.92	4.89	4.94	4.79	4.95	4.55
	cis-anti	4.96	4.56	5.04	4.87	4.68	5.03	4.62	4.87	4.93	4.89	4.93	4.77	4.93	4.54
H4’	parent	4.23	4.18	4.4	4.23	4.15	4.39	4.1	4.17	4.17	4.27	4.34	4.16	4.33	4.13
	cis-anti	4.27	4.12	4.44	4.39	4.15	4.36	4.13	4.23	4.32	4.27	4.32	4.15	4.32	4.12
^{31}P		$\delta_{1\text{-}2}$	$\delta_{2\text{-}3}$	$\delta_{3\text{-}4}$	$\delta_{4\text{-}5}$	$\delta_{5\text{-}6}$	$\delta_{6\text{-}7}$	$\delta_{7\text{-}8}$	$\delta_{8\text{-}9}$	$\delta_{9\text{-}10}$	$\delta_{10\text{-}11}$	$\delta_{11\text{-}12}$	$\delta_{12\text{-}13}$	$\delta_{13\text{-}14}$	
	parent	-3.08	-3.06	-3.24	-2.99	-3.13	-3.24	-3.31	-3.11	-3.22	-3.05	-3.16	-3.18	-2.82	
	cis-anti	-2.7	-2.83	-3.04	-2.78	-3.04	-2.14	-3.21	-3.22	-3.23	-3.03	-3.14	-3.18	-2.81	

Table 2.2. Photoproduct yields for mutant sequences of d(GTATCATGAGGTGC). ODNs (50 μM) were irradiated at pH 4.8 for 2.5 h at 0 ºC[a,b].

name	sequence	% T2[*c,a*]T7	% T2[*c,s*]T4	% T4[Dewar]C5
WT	GTATCATGAGGTGC	42	5	10
Mut1	GT<u>G</u>TCATGAGGTGC	5	4	6
Mut2	GTA<u>A</u>CATGAGGTGC			
Mut3	GTA<u>G</u>CATGAGGTGC			
Mut4	GTAT<u>A</u>ATGAGGTGC			
Mut5	GTAT<u>G</u>ATGAGGTGC			
Mut6	GTATC<u>G</u>TGAGGTGC	30		16
Mut7	GTATCAT<u>A</u>AGGTGC	4	6	6
Mut8	GTATCATGAGGT - -		7	17
Mut9	GTATCATGAG - - - -		9	30
Mut10	GTATCATG - - - - - -		13	7

[a] Mutations are underlined.

[b] Only photoproducts detected in greater than 3% yield are reported.

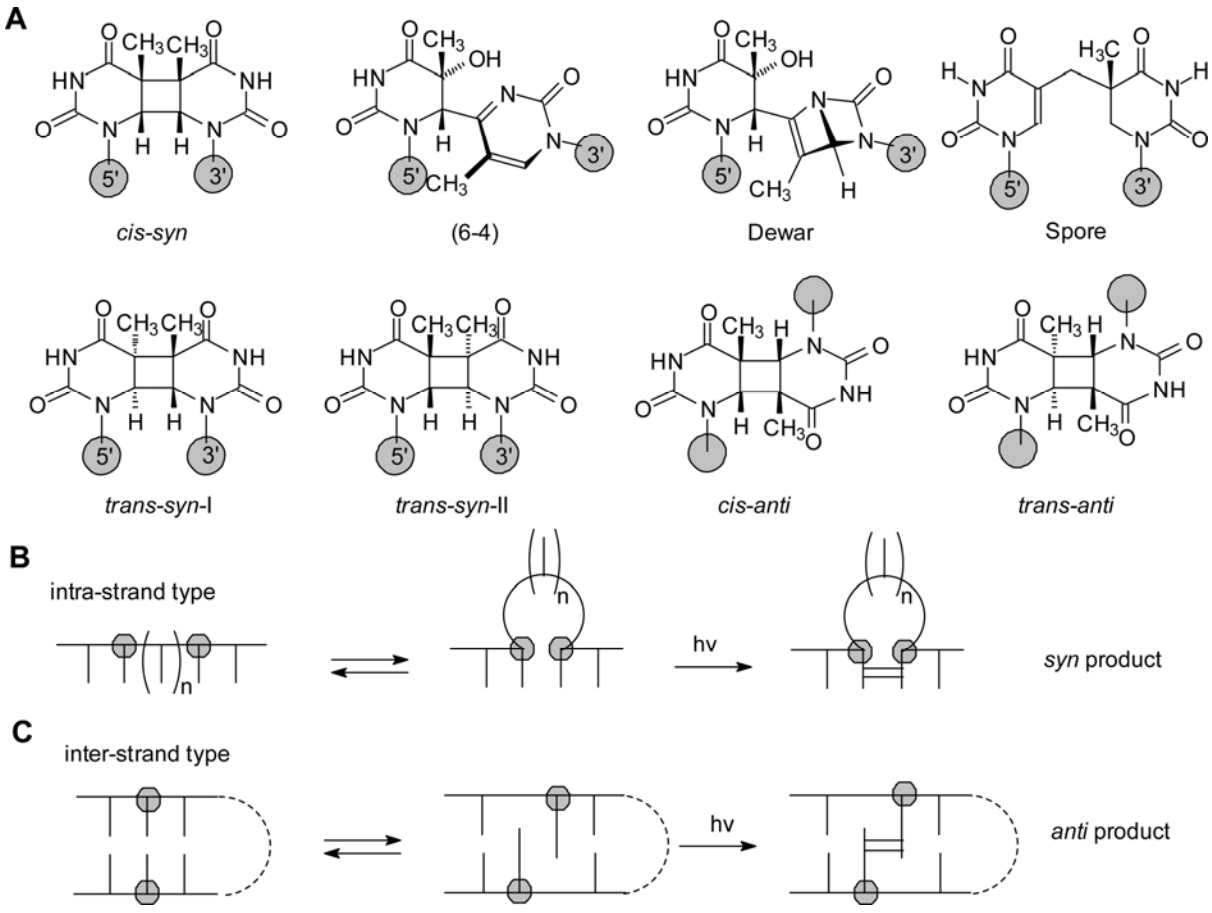


Figure 2.1. Photochemistry of DNA. (A) Structures of the major type of photoproducts. (B) Intra- and (C) interstrand-type photoreactions, both of which lead to nonadjacent photoproducts, except when n = 0 for B, which results in an adjacent photoproduct.

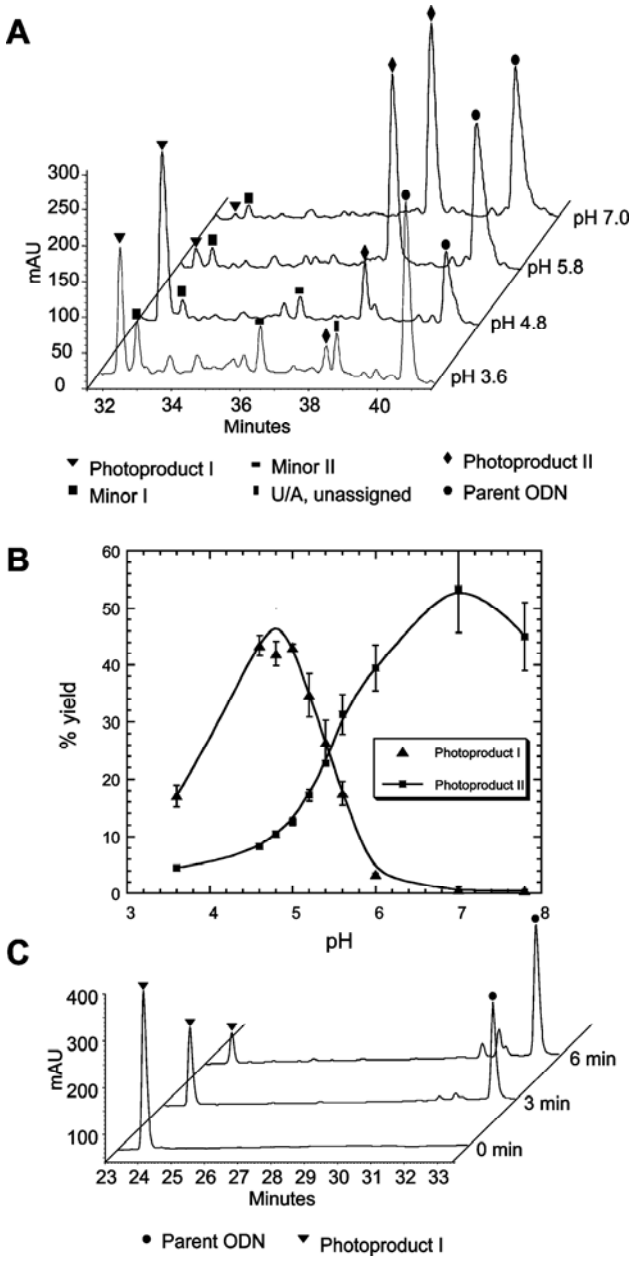


Figure 2.2. HPLC analysis of photoproduct formation and reversal. (A) Reverse phase HPLC analysis with method A of a UVB irradiation mixture of d(GTATCATGAGGTGC) at different pHs. Accurate m/z of the molecular ion $[M + H]^+$ of the unknown photoproduct I is 4317.7778 with mass error of +3.1 ppm compared with a calculated value of 4317.7644. (B) Plot of the percent yield (calculated by peak area) of photoproduct I and the photoproduct II-Dewar photoproduct as a function of pH. The

error bars represent the standard deviation of three independent experiments. (C) Reverse phase HPLC analysis with method B of the products of 254 nm irradiation of photoproduct I for 0, 3, and 6 min. The major reversion product was confirmed to be the parent ODN by MALDI analysis of the PDE I and PDE II digestion products. The smaller peaks eluting around 32 min are attributable 254 nm irradiation products of the reversal product, as confirmed by separate irradiation of the parent ODN.

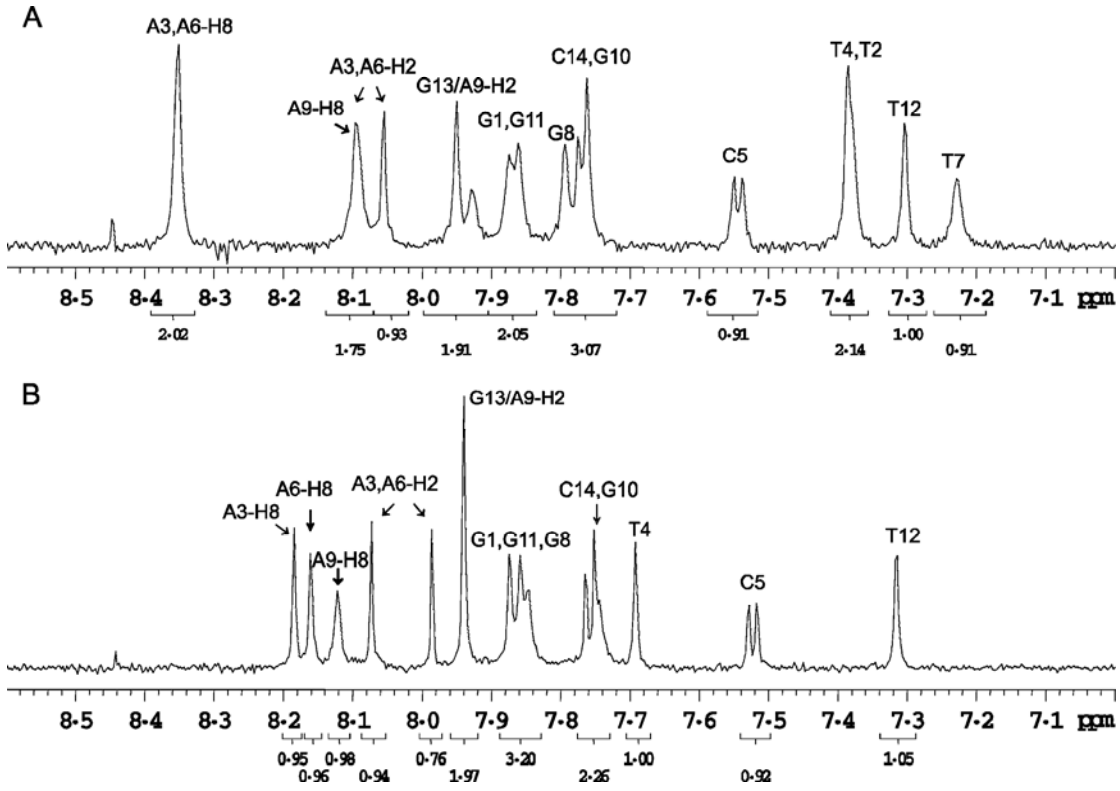


Figure 2.3. NMR spectra of photoproduct I and the parent ODN. 1D ^{1}H NMR spectra of the aromatic region of (A) the parent ODN d(GTATCATGAGGTGC) and (B) photoproduct I at room temperature in D_2O. The spectra show that two of the four aromatic TH6 proton signals are missing in photoproduct I, indicating formation of a T=T photoproduct.

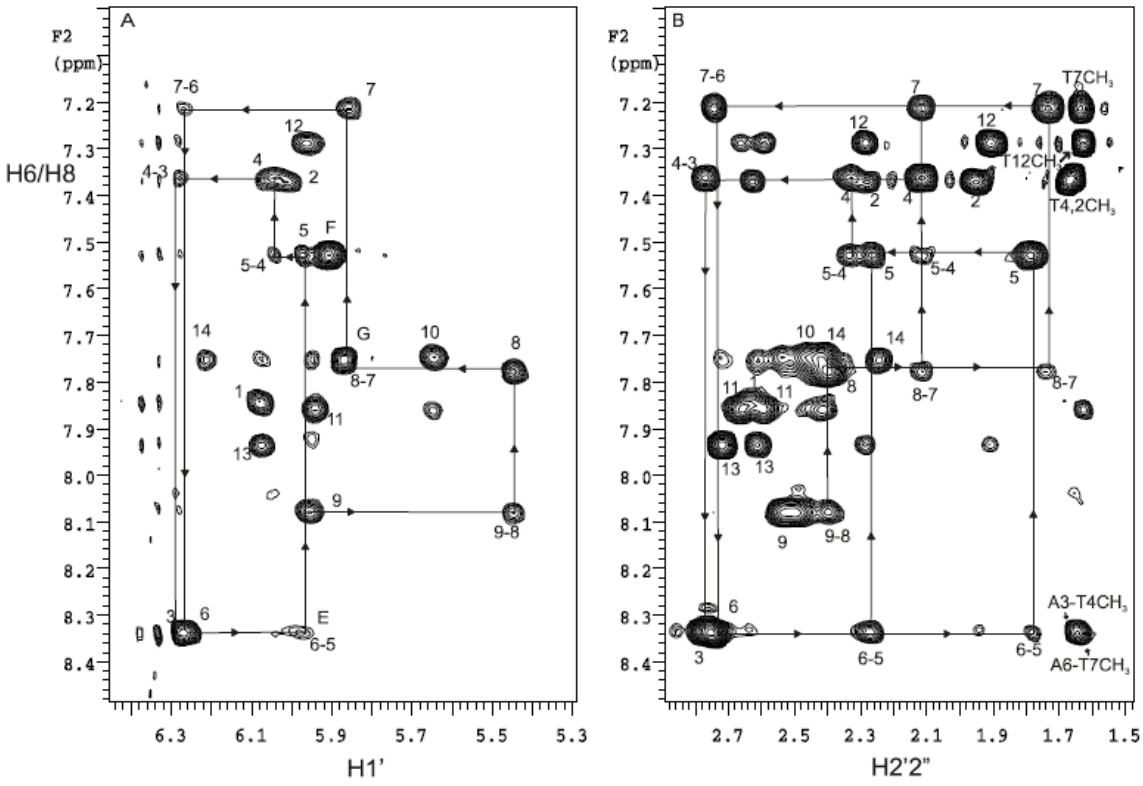


Figure 2.4. NOESY subspectra of the parent ODN. Expansion from the aromatic H6/H8 region to (A) H1'/H5 and (B) H2'2"/TCH3 region of 600 MHz NOESY spectrum of the parent ODN 14-mer in "100%" D2O at 25 °C. Sequence of A3 through A9 is connected by line with number that represents (A) H6/8 (i)-H1'(i), and (B) H6/8 (i)-H2'2"/TCH3 (i) crosspeaks.

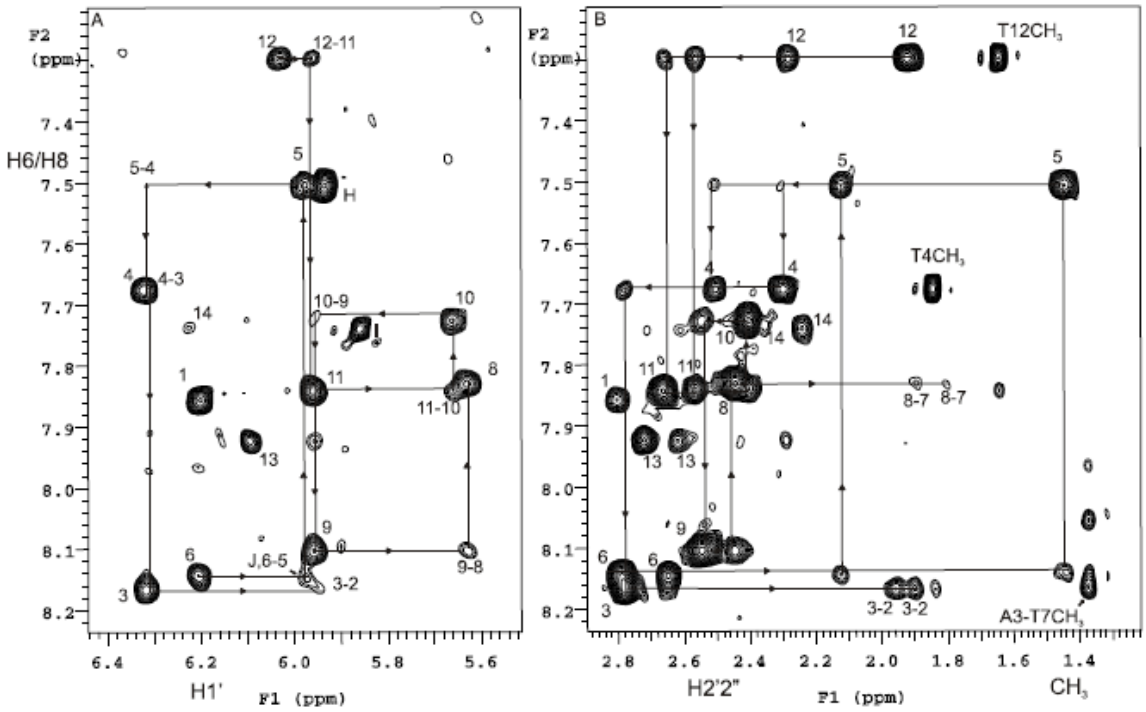


Figure 2.5. NOESY subspectra for photoproduct I. Expansion from the aromatic H6/H8 region to H1'/H5 and H2'2"/T-5CH3 region of 600 MHz NOESY spectrum of photoproduct I 14-mer at 25 °C. Sequence is connected by line with number represents. A) H6/H 8 (i)-H1' (i), and B) H6/H8 (i)-H2'2"/T-5CH3 (i) cross peaks.

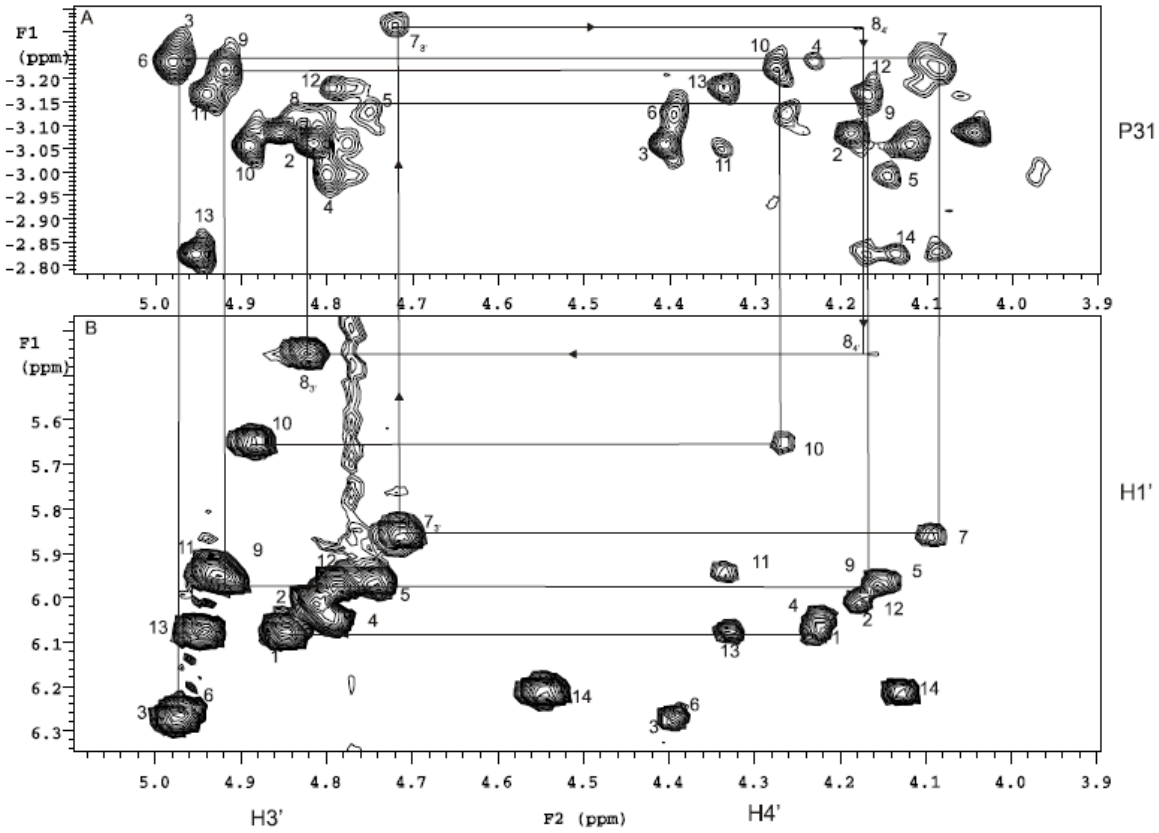


Figure 2.6. ^{1}H-^{31}P HSQC and TOCSY subspectra of the parent ODN. A) ^{1}H-^{31}P HSQC spectrum of the parent ODN 14-mer in "100%" D2O. Horizontal line connections represent H3' (i)-p-H4' (i+1) correlation. B) Expanded H3'/H4' region of TOCSY spectrum. Horizontal line connections represent H3'-H4' spin propagation.

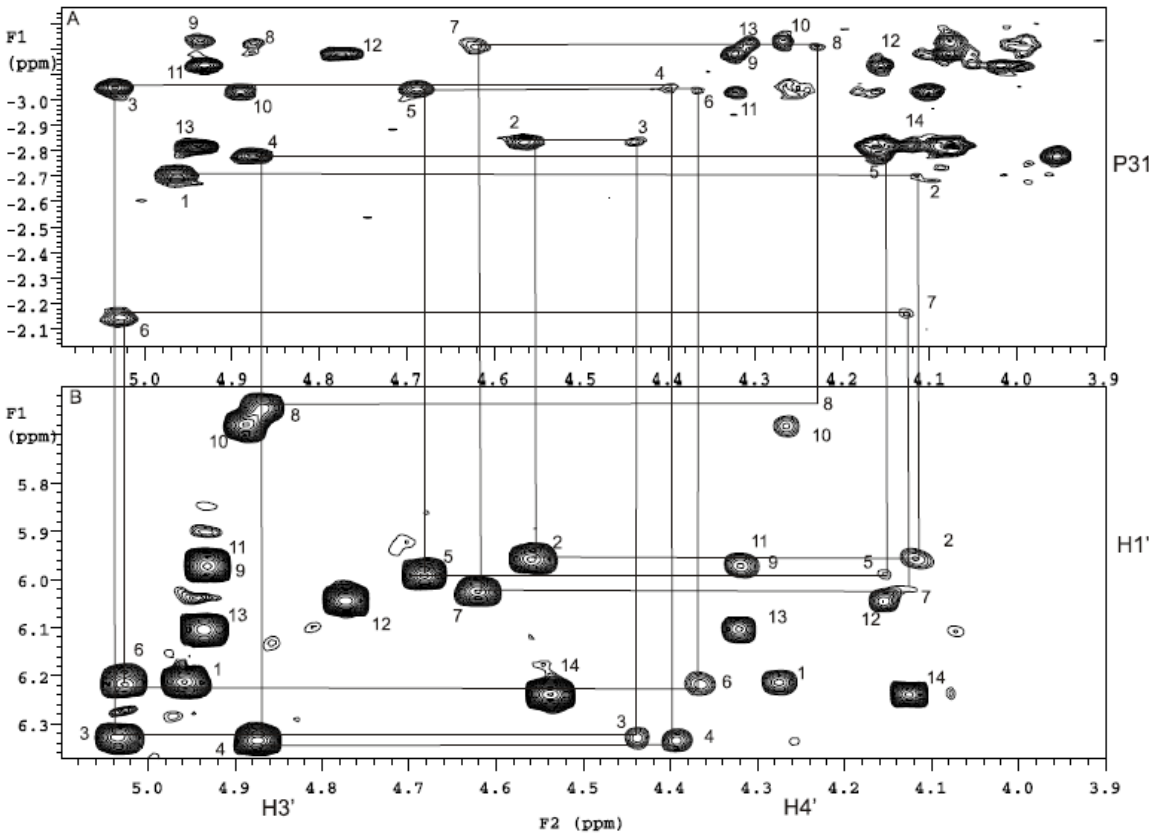


Figure 2.7. ^{1}H-^{31}P HSQC and TOCSY subspectra of photoproduct I. A) ^{1}H-^{31}P HSQC spectrum of photoproduct I in "100%" D2O. Horizontal line connections represent H3'(i)-p-H4'(i+1) correlation, B) Expanded H3'/H4' region of TOCSY spectrum. Horizontal line connections represent H3'-H4' spin propagation.

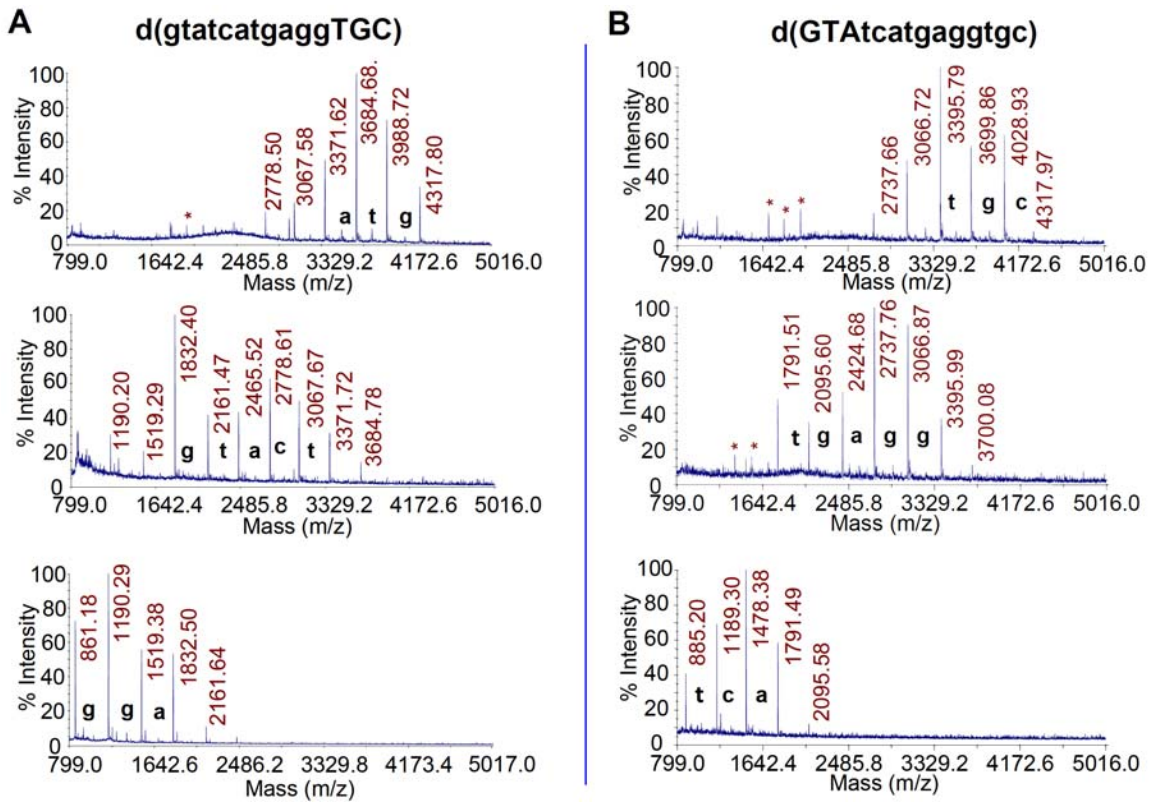


Figure 2.8. Exonuclease-coupled MS ladder sequencing of the parent ODN. A) BSP digestion (5'→3') coupled MALDI sequencing at room temperature for 0.5, 1.5, and 5.5 min (top to bottom). B) SVP digestion (3'→5') coupled MALDI sequencing at 37 °C for 1, 2.5, and 5.5 min (top to bottom). An asterisk refers to a doubly charged ion.

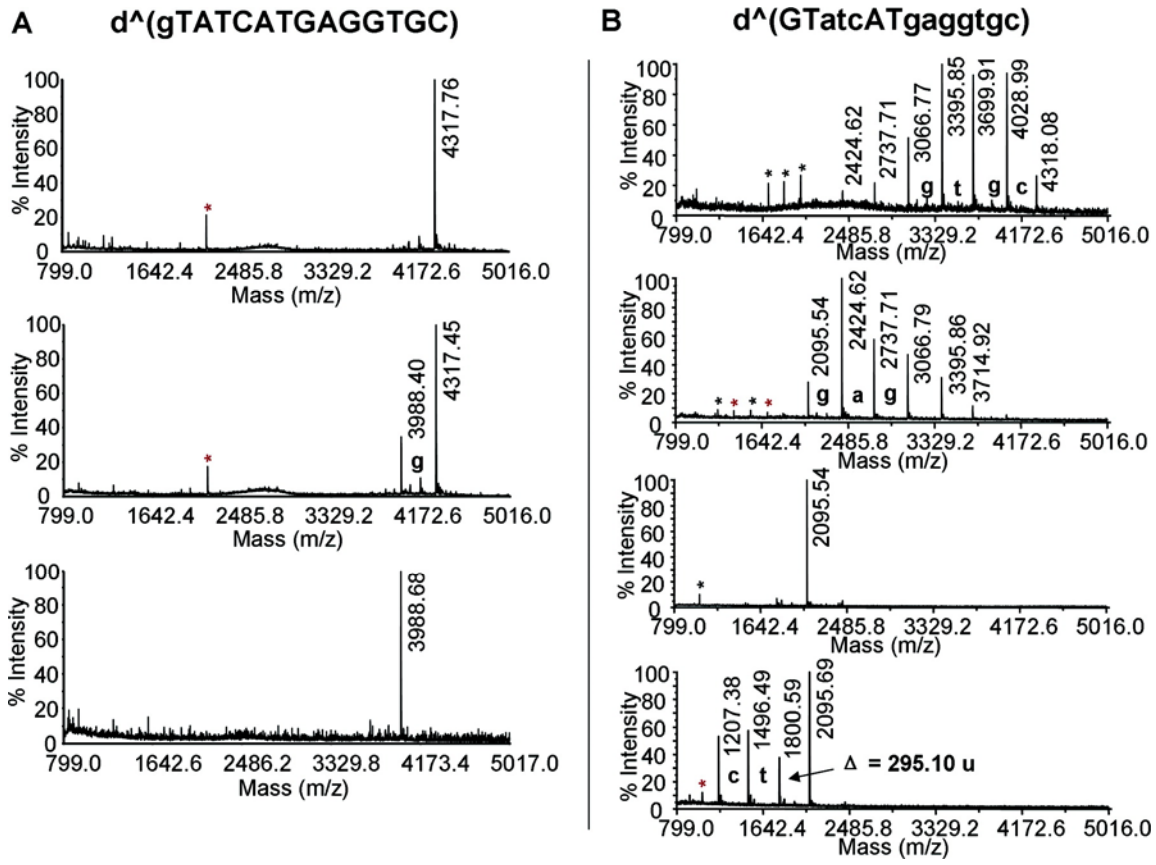


Figure 2.9. Exonuclease-coupled MS ladder sequencing of photoproduct I. (A) BSP digestion coupled MALDI sequencing at room temperature for 0 min, 110 min, and at 37 ºC for 40 min (top to bottom). (B) SVP digestion coupled MALDI sequencing at 37 ºC for 1, 2.5, 5.5, and 85 min (top to bottom). The mass losses of 289.05, 304.05, 313.06, and 329.05 u correspond to exonucleotytic cleavage of pdC, pdT, pdA, and pdG, respectively (+ H2O - pdN). Nucleotides that were enzymatically removed are represented by small-case letters. Peaks labeled with * are doubly charged digestion fragments. indicates photodamaged ODNs. Proton signals are missing in photoproduct I, indicating formation of a T=T photoproduct.

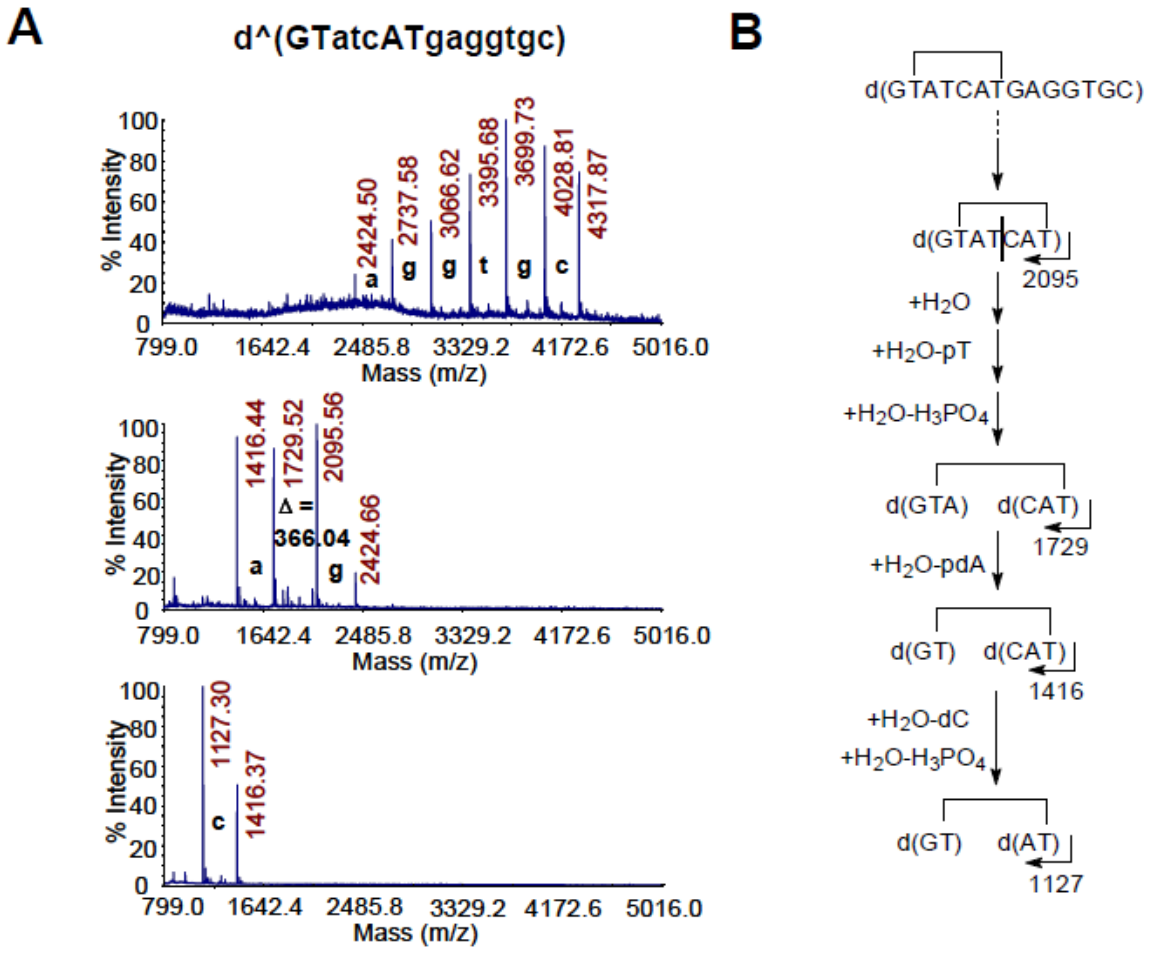


Figure 2.10. BIMP sequencing of photoproduct I. A) BIMP digestion (3'→5') coupled MALDI sequencing of photoproduct I at 37 °C for 0.5 min, 7 min, and 110 min (top to bottom). The mass loss of 289.05, 304.05, 313.06, and 329.05 u corresponding to the excision of pdC, pT, pdA, and pdG, respectively (+ H2O - pdN). Nucleotides that were enzymatically removed are represented by small-case letters. Right: proposed enzymatic cleavage pathway by BIMP for photoproduct I.

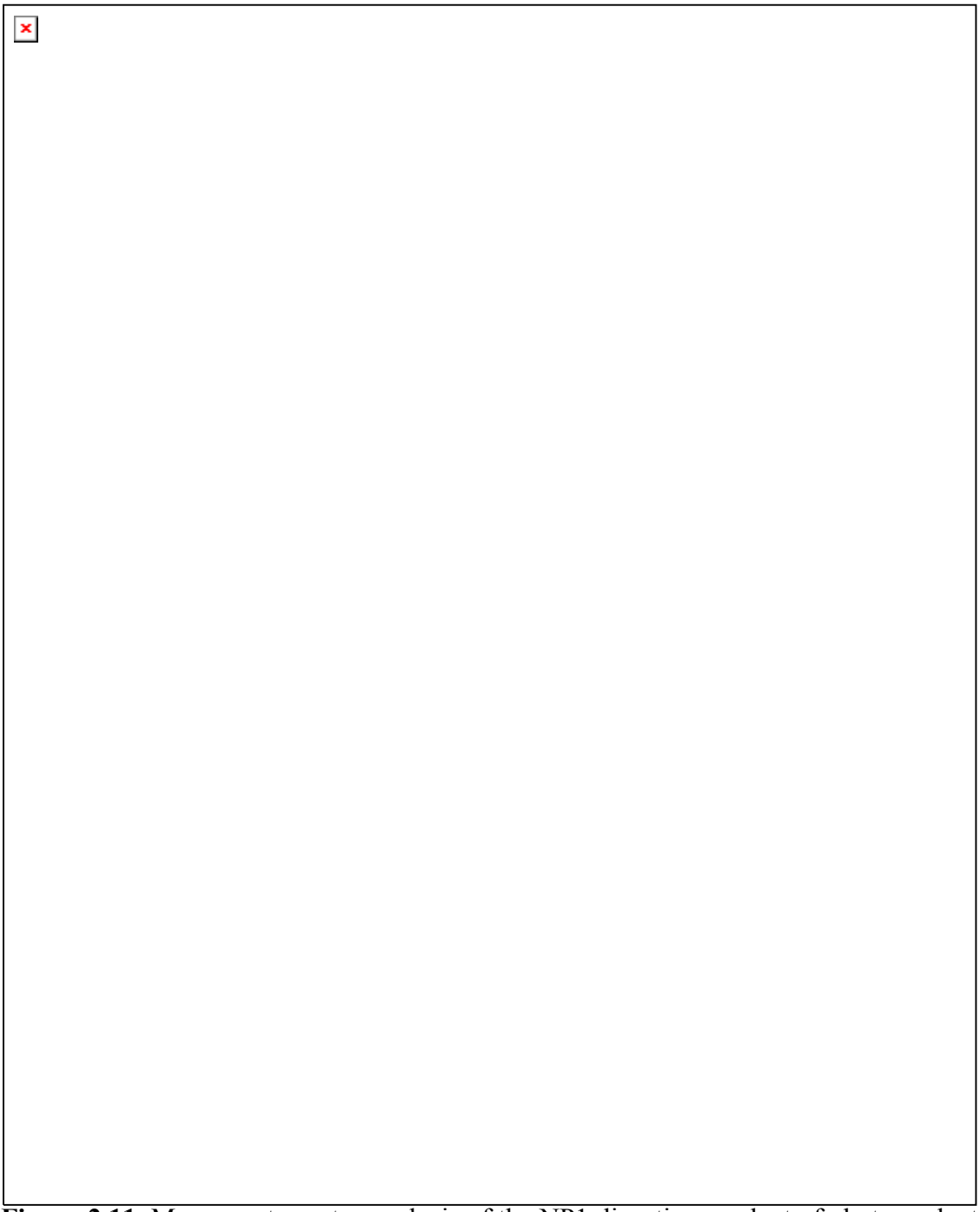


Figure 2.11. Mass spectrometry analysis of the NP1 digestion product of photoproduct I. (A) MALDI-TOF-MS/MS spectrum of parent ion $[M+H]^+$ (*m/z* 1287.3) of HPLC purified ultimate NP1 digestion product of photoproduct I. (B) Proposed fragmentation pathway.

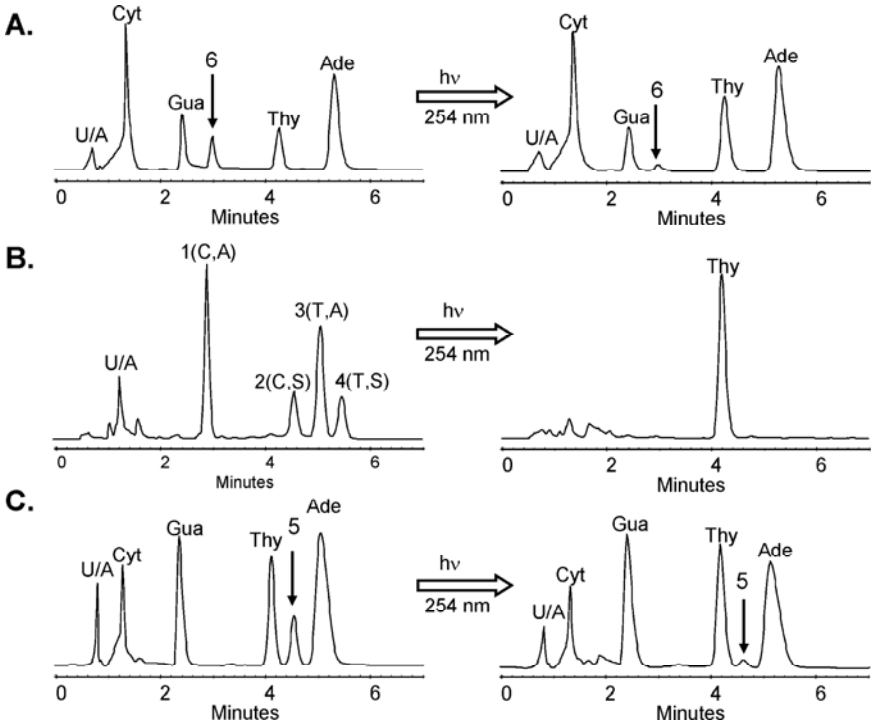


Figure 2.12. Structure correlation of the cyclobutane thymine dimer. The acid hydrolysis products of photoproduct I were compared to an authentic mixture of cyclobutane thymine dimers by HPLC using method D. HPLC traces of (A) the acid hydrolysis products of photoproduct I, (B) the product mixture from the irradiation of thymine, and (C) the acid hydrolysis products of the *cis-syn* TT dimer containing photoproduct d(GTAT[*c,s*]TATGAGGTGC). To the left are the HPLC traces before irradiation with 254 nm light, and to the right, after irradiation. The peaks labeled with U/A are unassigned peaks.

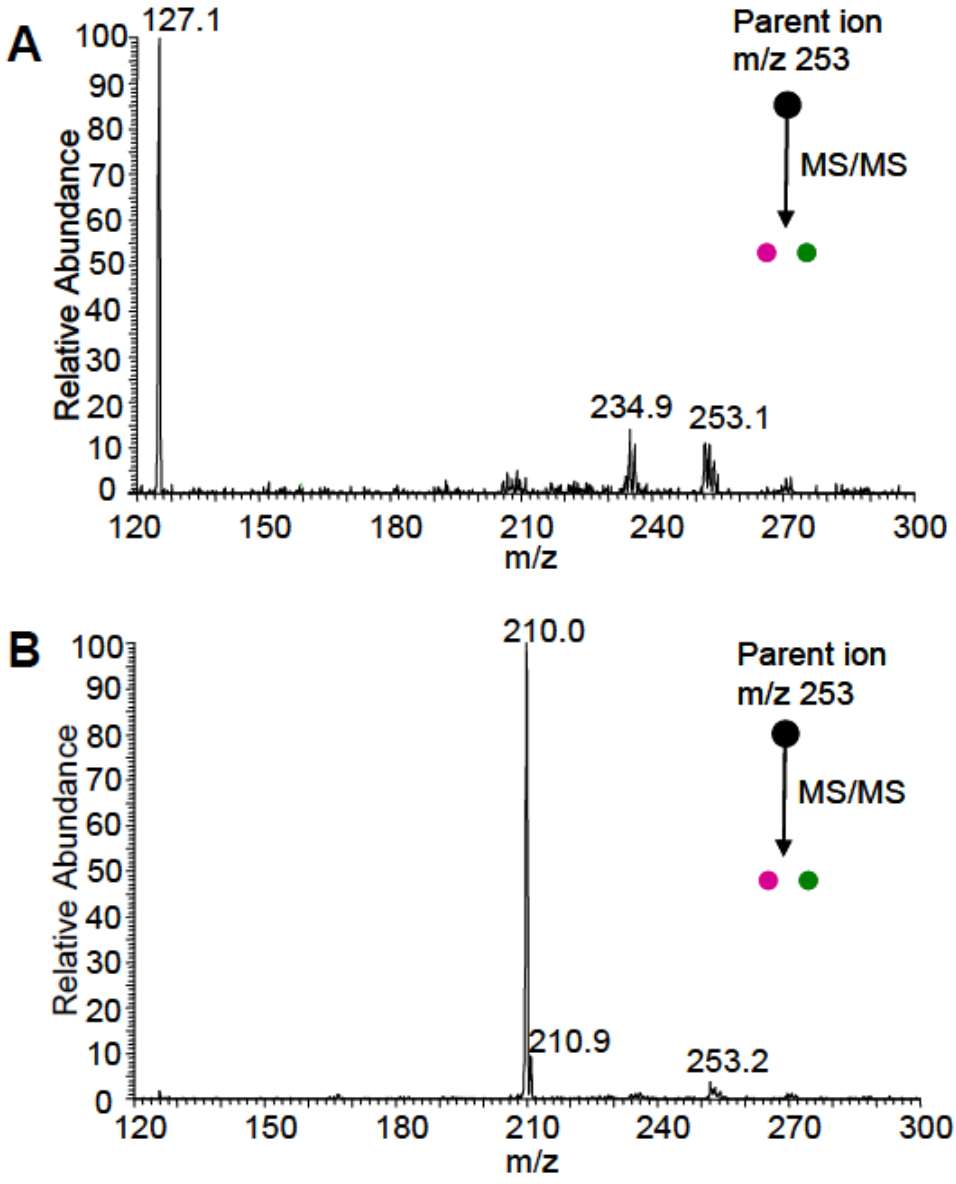


Figure 2.13. ESI-MS/MS characterization of the cyclobutane thymine dimers. A) Thy[*c,a*]Thy dimer from photoproduct I and B) Thy[*c,s*]Thy dimer from d(GTAT[*c,s*]TATGAGGTGC).

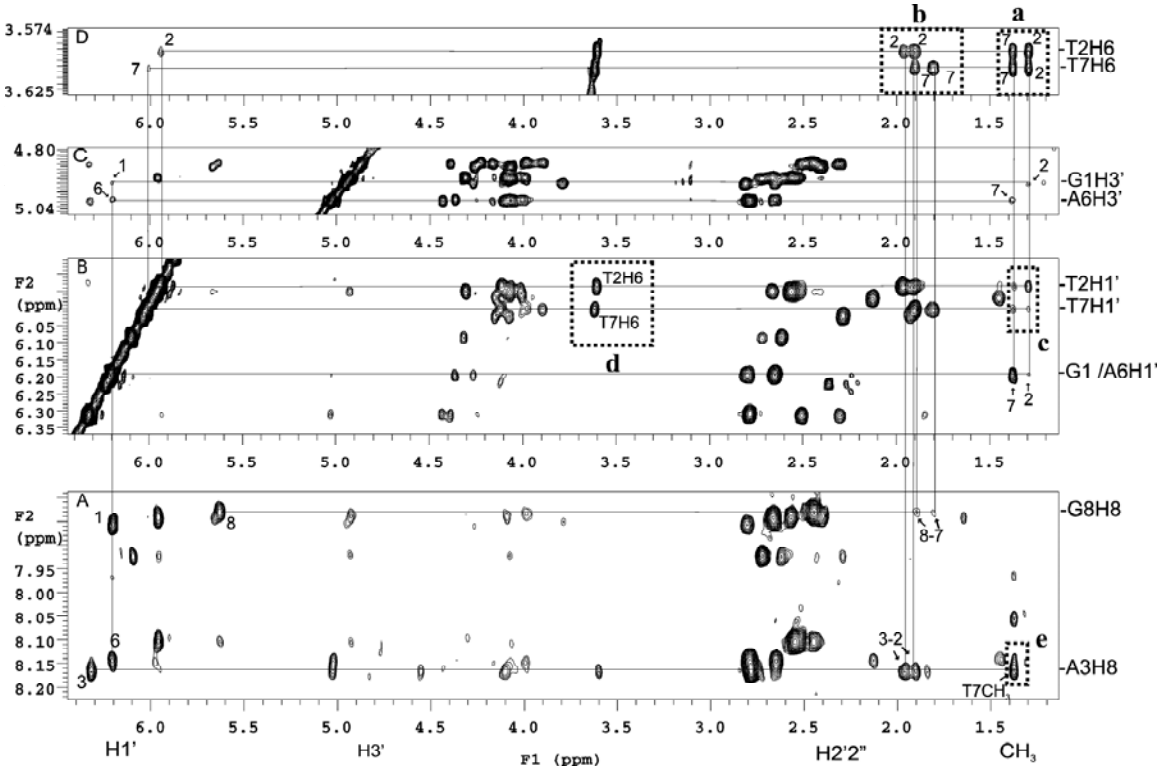


Figure 2.14. Section of NOESY spectra of photoproduct I at 25 ºC. (A) Cross-correlation between the H6, sugar, and methyl protons of T2 and T7. (B) Connection of the preceding G1, A6-H1', T2, T7-H1' to T2/T7 methyl protons. (C) Connection of the preceding G1, A6-H3' to T2/T7 methyl protons. (D) Connection of the T2/T7 H6H2'2"/CH_3 protons. See text for explanation of the cross-peaks enclosed in the dotted boxes.

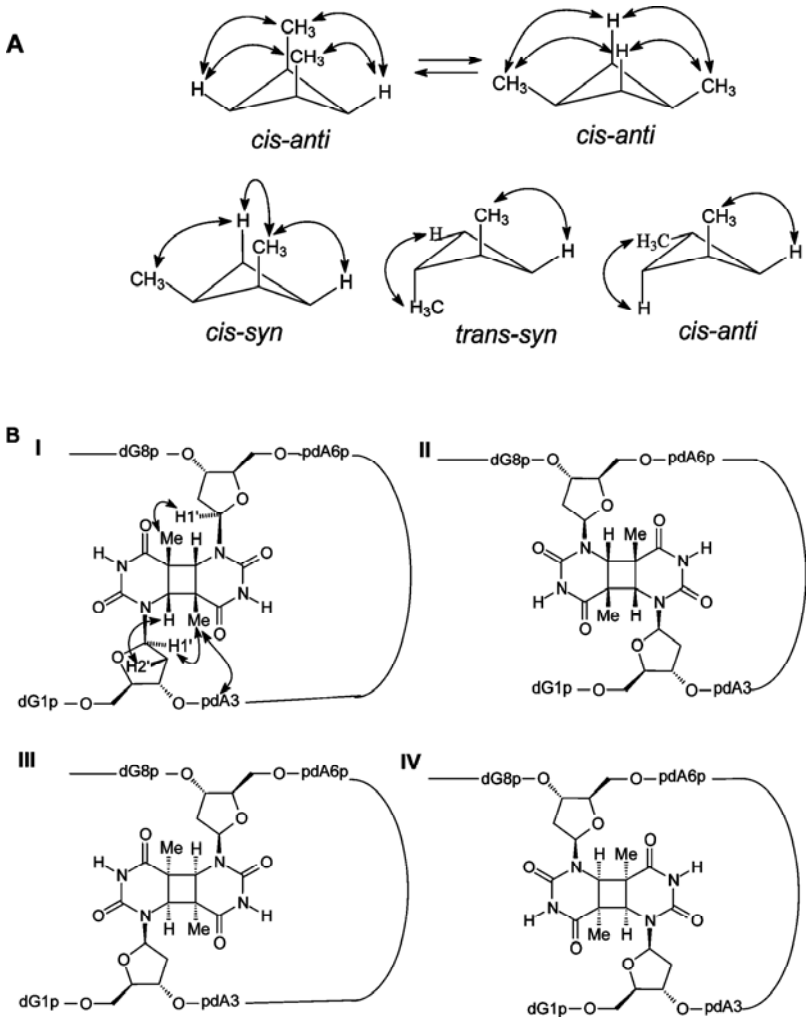


Figure 2.15. Structure and stereochemistry of the *cis-anti* stereoisomers of the T2[*c,a*]T7 product. (A) Conformations of the cyclobutane ring of *cis-anti*, *cis-syn*, *trans-syn*, *trans-anti* thymine dimers and the NOE active H6-CH_3 pairs. (B) The four possible *cis-anti* T2-T7 cyclobutane dimers showing how the observed NOEs are consistent with structure I.

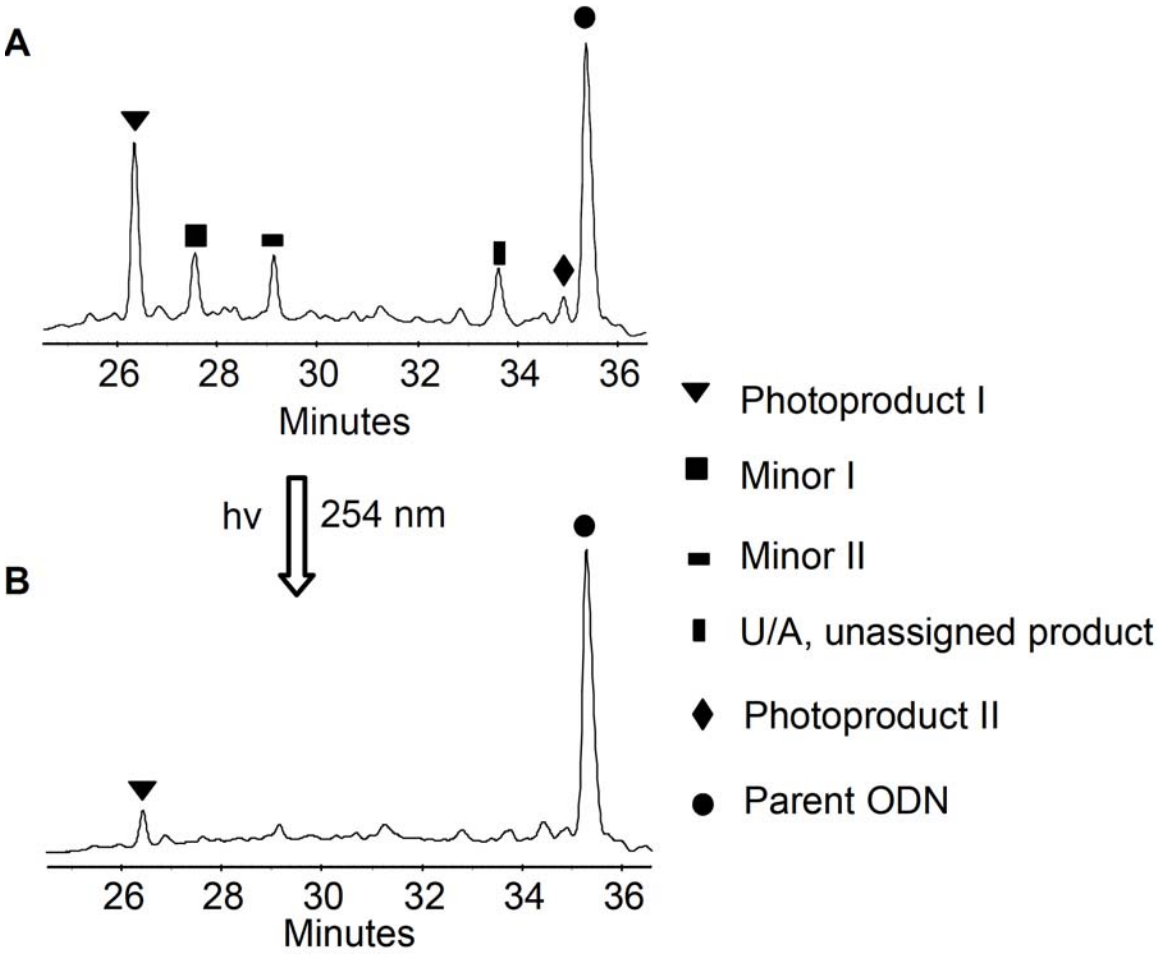


Figure 2.16. Photoreversal of UVB irradiation products to the parent ODN. Reverse phase HPLC analysis (method B) of the photoproducts of d(GTATCATGAGGTGC) produced at pH 3.6 to the parent ODN with 254 nm light: A) UVB irradiation mixture of d(GTATCATGAGGTGC) produced at pH 3.6, and B) UVB irradiation mixture of d(GTATCATGAGGTGC) produced at pH 3.6 after 30 min irradiation with 254 nm light.

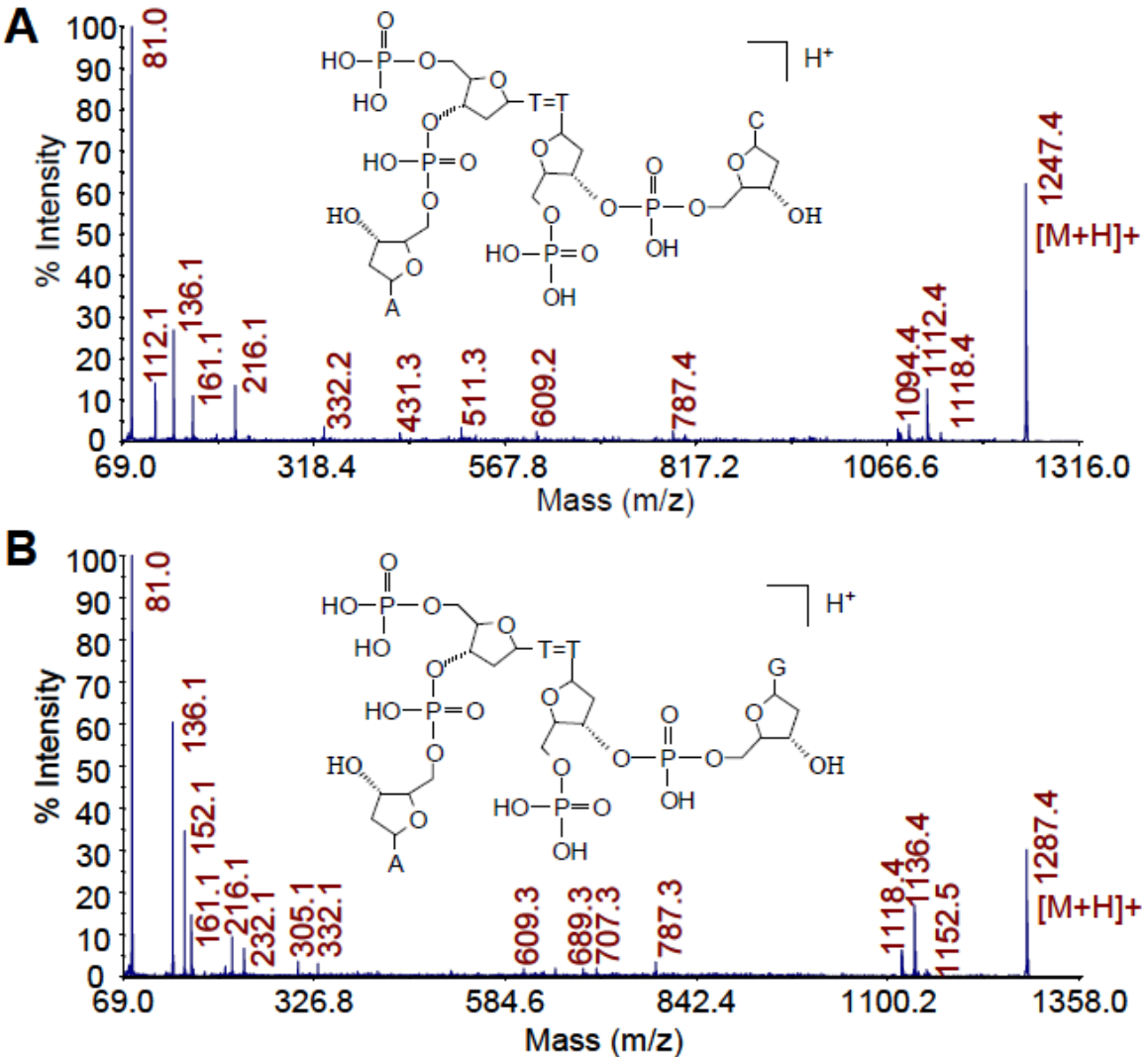


Figure 2.17. Mass spectrometry analysis of NP1 digestion product of minor I and minor II. MALDI-TOF-MS/MS spectra of: A) the parent ion $[M+H]^+$ (*m/z* 1247.4) of HPLC purified ultimate NP1 digestion product of minor I, *c,s*-pd(T[A])=pd(T[C]), showing loss of $[Ade+H]^+$ at 136 and $[Cyt+H]^+$ at 112, and B) the parent ion $[M+H]^+$ (*m/z* 1287.4) of HPLC purified ultimate NP1 digestion product of minor II, *t,a*-pd(T[A])=pd(T[G]) showing the same fragments as for photoproduct I (Figure 2.11).

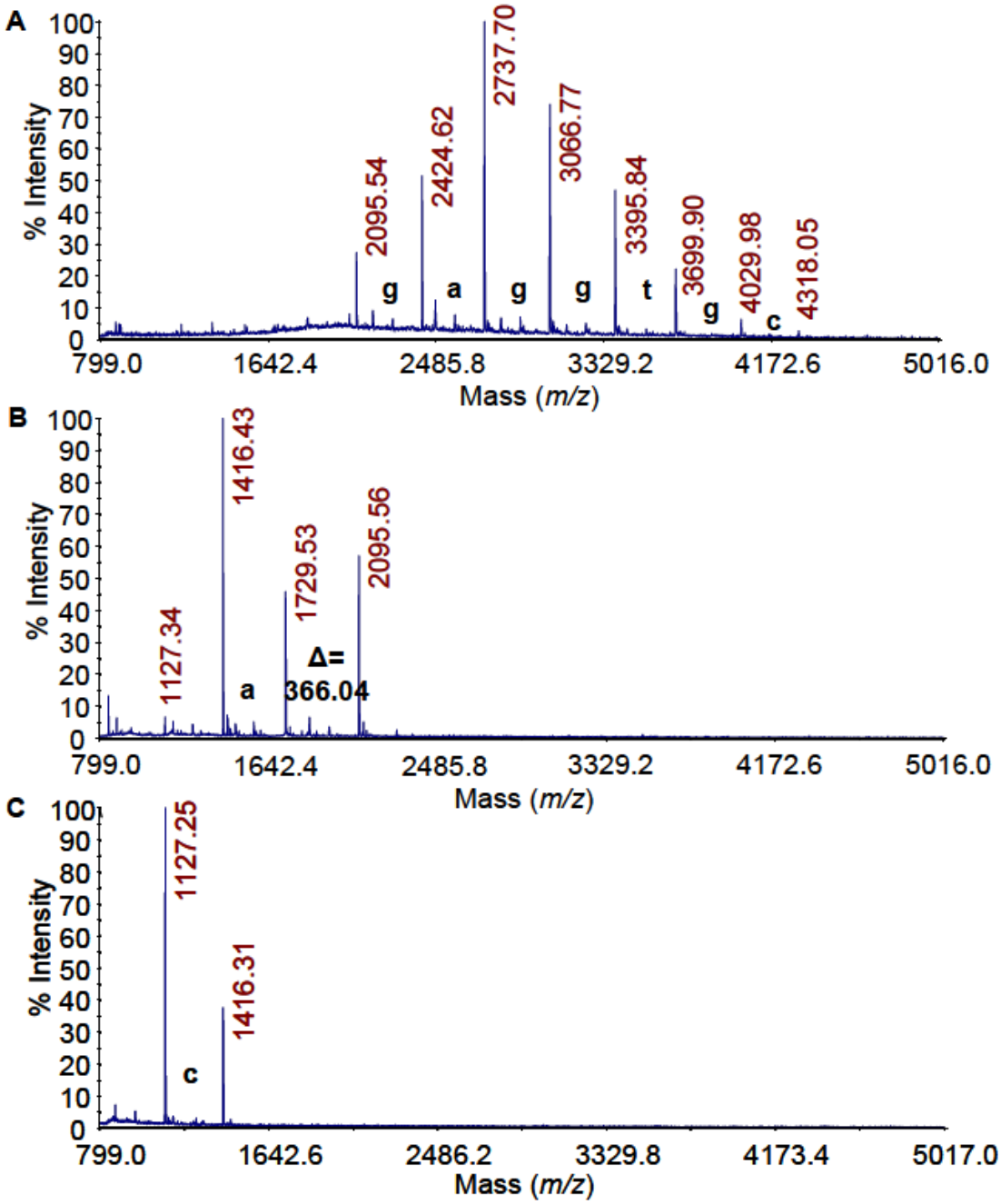


Figure 2.18. BIMP digestion (3'→5') coupled MALDI sequencing of minor photoproduct II. Minor photoproduct II was treated with BIMP at 37 °C for A) 1 min, B) 7 min, and C) 110 min. The mass loss of 289.05, 304.05, 313.06, and 329.05 u corresponds to the excision of pdC, pT, pdA, and pdG, respectively (+H2O-pdN). Nucleotides that were enzymatically removed are represented by small-case letters.

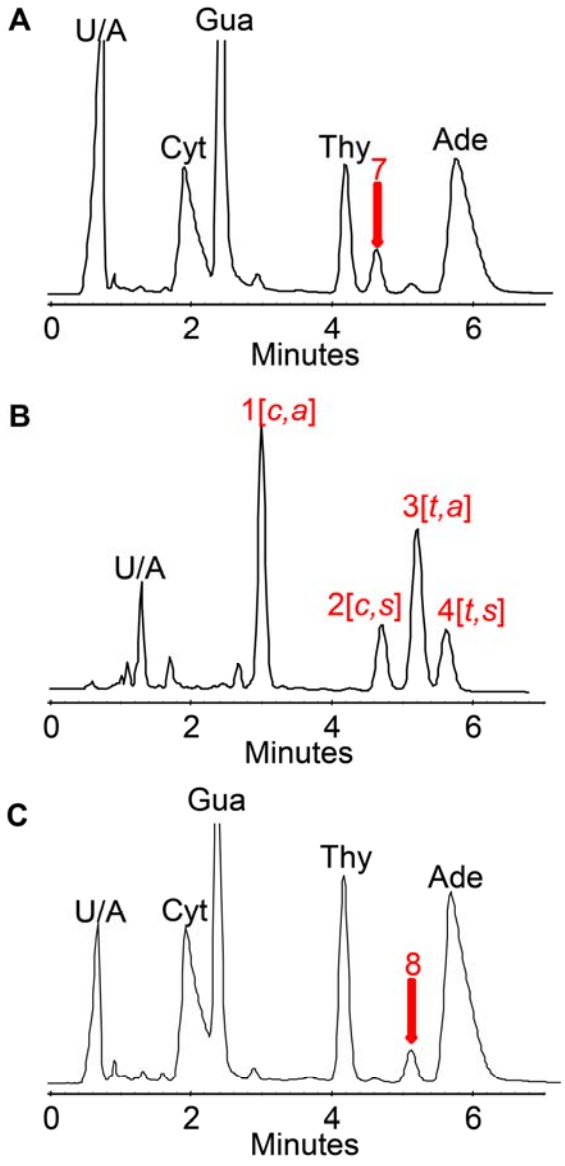


Figure 2.19. Thymine dimer stereochemistry of minor photoproducts I and II by chemical correlation. Comparison of cyclobutane dimers of thymine with the acid hydrolysis products of minor I and minor II photoproducts by reverse phase HPLC (method D). HPLC traces of: A) the hydrolysates of minor I, B) the product mixture from the irradiation of thymine, C) the hydrolysates of minor II.

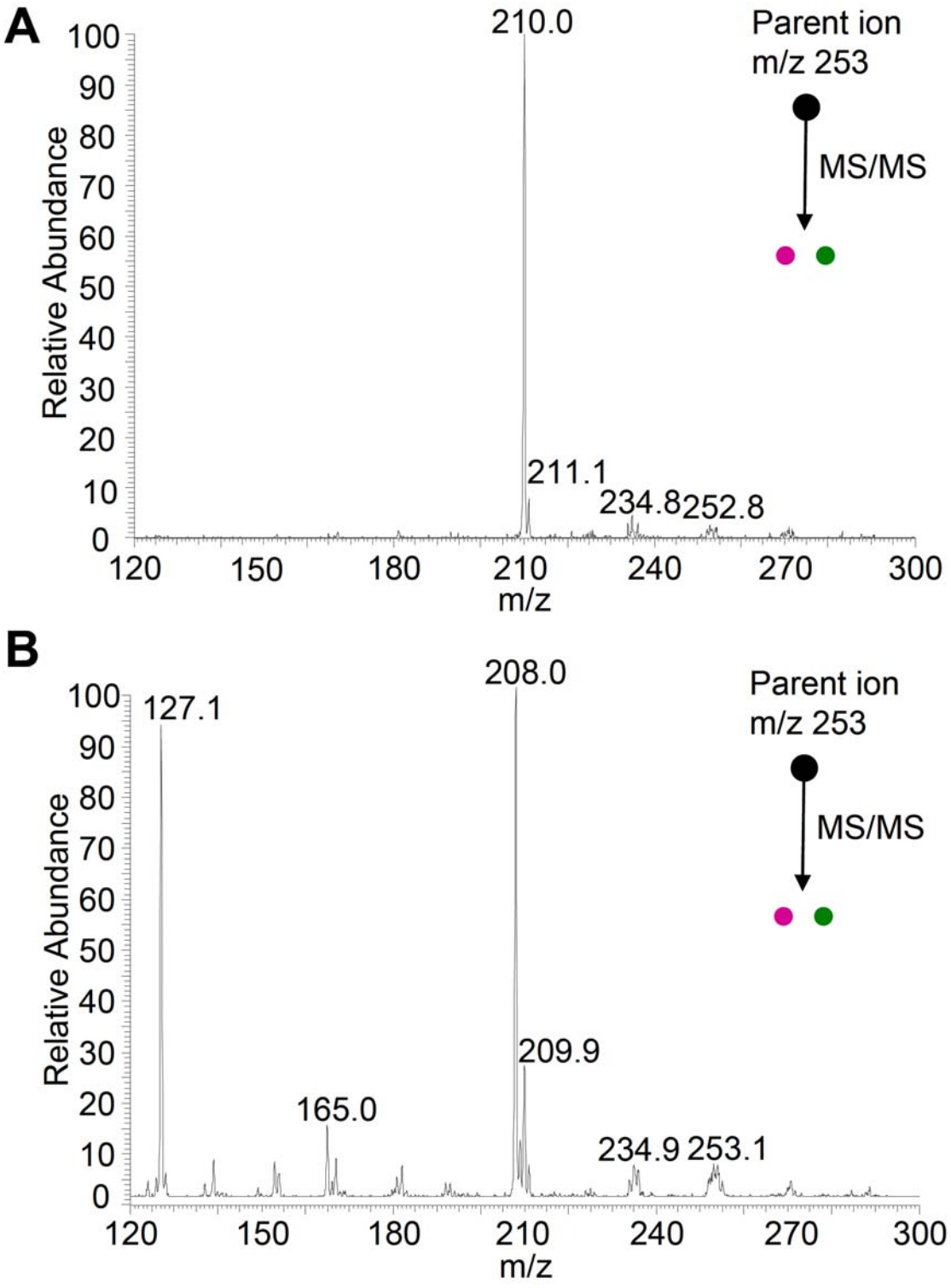


Figure 2.20. ESI-MS/MS characterization of the cyclobutane thymine dimers. A) Thy[*c,s*]Thy dimer from minor I. B) Thy[*t,a*]Thy dimer from minor II.

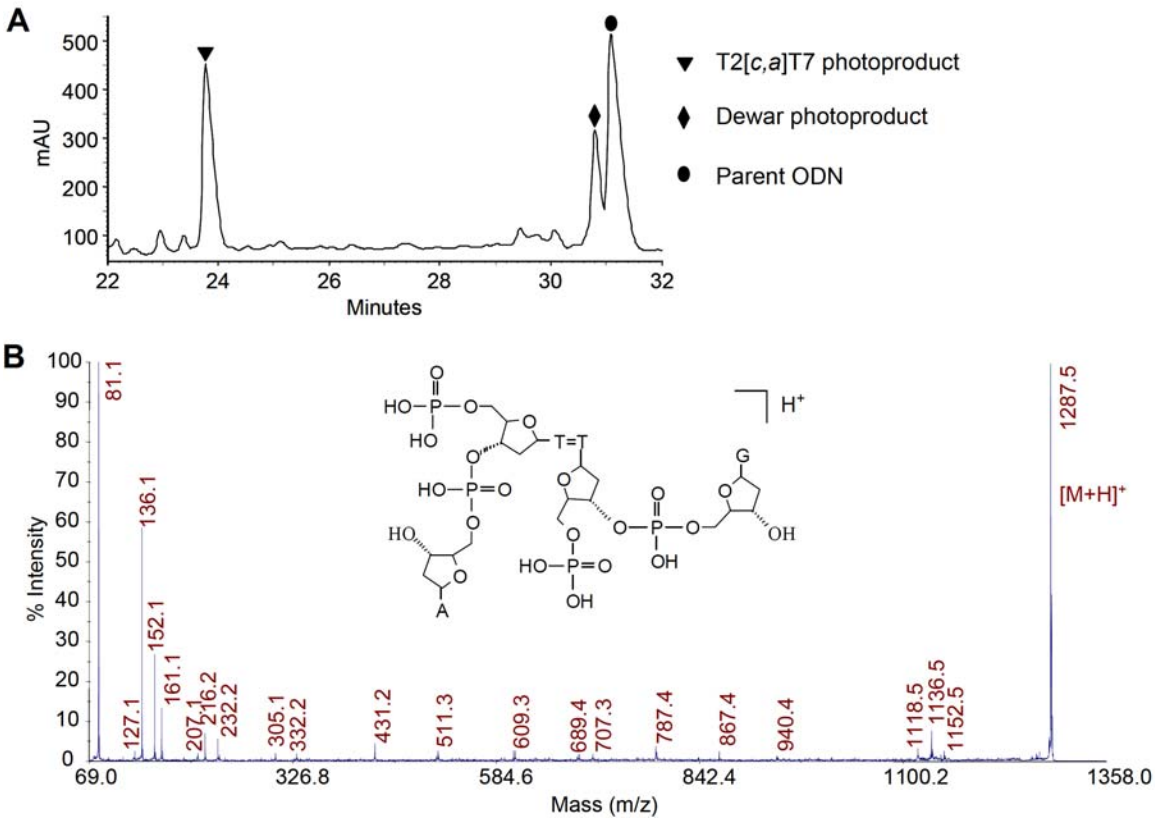
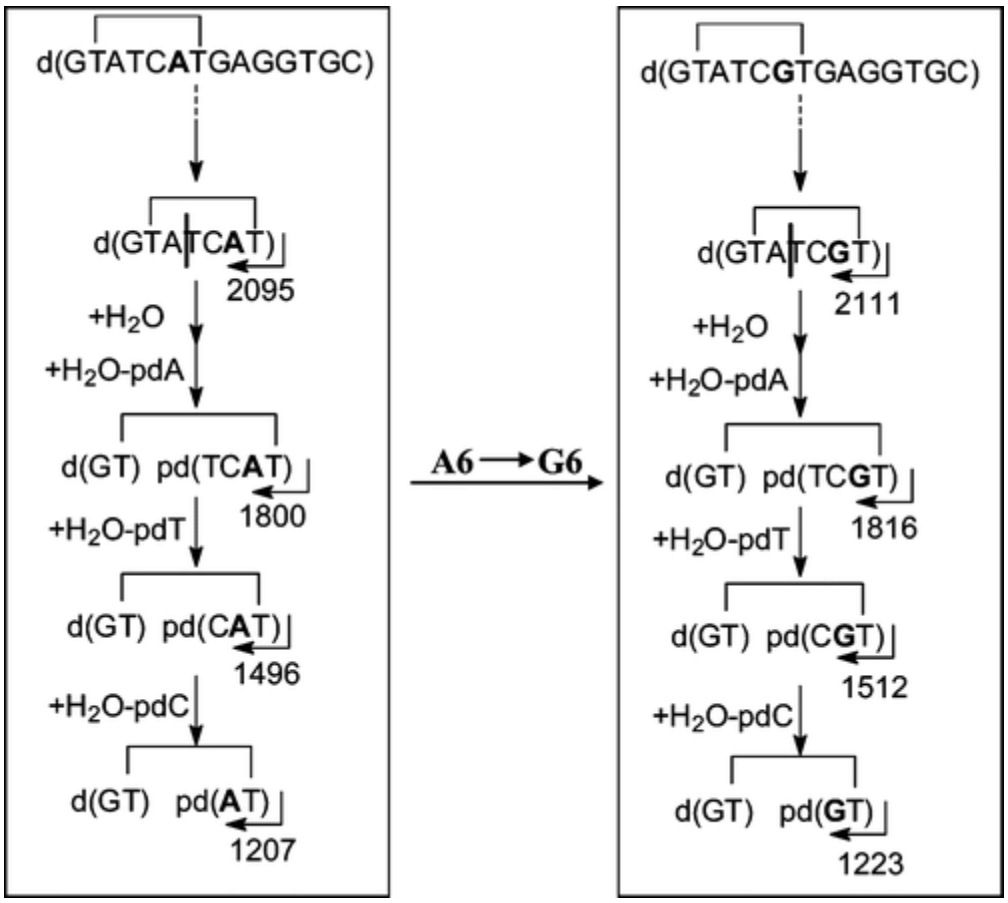


Figure 2.21. Formation and identification of the T2[*c,a*]T7 product in the mut6 sequence. Reverse phase HPLC analysis with method B of the UVB irradiation mixture of A) d(GTATCGTGAGGTGC) at pH 4.8, and B) MALDI-TOF-MS/MS assay of HPLC purified ultimate NP1 digestion product of T2[*c,a*]T7 non-adjacent photoproduct of d(GTATCGTGAGGTGC).



Scheme 2.1. Proposed enzymatic cleavage pathway by SVP for the T2[*c,a*]T7 photoproducts of the original sequence and mut6 (A6G).

Chapter 3

Photocrosslinking of Human Telomeric G-quadruplex Loops by *Anti* Cyclobutane Thymine Dimer Formation

Abstract

The unusual structural forms of telomere DNA, which protect the ends of chromosomes during replication, may render it vulnerable to unprecedented photodamage, possibly involving nonadjacent bases that are made proximate by folding. The G-quadruplex for the human telomere sequence consisting of a repeating d(TTAGGG) is one unusual form. Tel22, d[AGGG(TTAGGG)$_3$], forms a basket structure in the presence of Na^+ and may form multiple equilibrating structures in the presence of K^+ with hybrid-type structures predominating. UVB irradiation of d[AGGG(TTAGGG)$_3$] in the presence of Na^+ results in a *cis,syn* thymine dimer between two adjacent Ts in a TTA loop and a mixture of nonadjacent *anti* thymine dimers between various loops. Irradiation in the presence of K^+, however, produces, in addition to these same products, a large amount of specific *anti* thymine dimers formed between either T in loop 1 and the central T in loop 3. These latter species were not observed in the presence of Na^+. Interloop-specific *anti* thymine dimers are incompatible with hybrid-type structures, but could arise from a chair or basket-type structure or from triplex intermediates involved in interconverting these structures. If these unique nonadjacent *anti* thymine dimer photoproducts also form *in vivo*, they would constitute a previously unrecognized type of DNA photodamage that may interfere with telomere replication and present a unique challenge to DNA repair. Furthermore, these unusual *anti* photoproducts may be used to establish the presence of G-quadruplex or quadruplex-like structures *in vivo*.

Introduction

Telomeres are repetitive sequences at the ends of chromosomes that function in concert with telomerase and a number of other proteins to protect the ends of chromosomes from shortening during successive rounds of replication (*1-3*). The human telomere sequence consists of repeating d(TTAGGG) that may fold into various repeating G-quadruplex structures. Short fragments of the human telomere sequence can exist in a myriad of structures that depend on their sequence and on the cations present (*4-7*). To date, basket (*8, 9*), parallel (*10*), and (3 + 1) structures including hybrid-1 and hybrid-2 (*11-14*) type G-quadruplex forms have been determined by NMR and crystallography. The oligodeoxynucleotide, d[AGGG(TTAGGG)$_3$] (Tel22), adopts a basket structure in solutions containing Na^+ (*8*), whereas in the presence of K^+, the predominant intracellular cation, the hybrid-type structures are favored over the basket (*11, 13*) and/or chair (*15*) structures (Figure 3.1).

In chapter 2, we describe a highly unusual interstrand-type nonadjacent, *anti* thymine photodimer (Figure 3.2B), previously only found in dried or ethanolic solutions of UVC irradiated DNA (*16*), can form in aqueous solution. Irradiation of d(GTATCATGAGGTGC) with UVB light produced a high yield of *cis,anti* thymine photodimer between T2 and T7 under slightly acidic conditions (*17*). This anti photodimer must originate from some folded structure, still unknown, that brings T2 in a head to tail arrangement with T7 followed by photoinduced [2 + 2] cycloaddition reaction between the two 5, 6 double bonds. An important implication of this observation is that *anti* thymine photodimers might also form in G-quadruplex forms of human telomeric DNA. Among the various G-quadruplex structures that could lead to *anti*

thymine photodimers are the basket and chair structures, which have two lateral TTA loops in close proximity. Other structures, such as the parallel and hybrid structures do not have proximate loops, suggesting that photochemical reactions such as this may be developed as structural probes even *in vivo*.

To our knowledge, the photochemistry of native human telomeric DNA has never been investigated, although the photoinduced hydrogen abstraction reactions of iodouracil substituted telomeric DNA has been reported (*18*). In 1989, however, Cech and coworkers did report the formation of crosslinks upon 254 nm irradiation of the folded form of the closely related Oxytricha, $d(T_4G_4)_4$, and Tetrahymena $d(T_2G_4)_4$ telomeric sequences as evidence for a G-quartet structure (*19*). The photocrosslinks formed in the presence of Na^+ were found to occur primarily between T11 and T27 in $d(T_4G_4)_4$ and between T1 or G3 and T13 in $d(T_2G_4)_4$ by gel electrophoresis of chemical cleavage products, and were ascribed to G-quartet chair structures. The structures of the photoproducts involved in the crosslinking, however, were never determined. In this chapter, we describe that irradiation of G-quadruplex forms of the human telomeric sequence $d[AGGG(TTAGGG)_3]$ (Tel22) with biologically relevant UVB light (280-320 nm) and in the presence of the biologically relevant cation K^+ results in a significant amount of *anti* thymine dimers between either T in loop 1 with the central T in loop 3. We also discuss the mechanistic and biological implications of these findings.

Material and Methods

Preparation of G-Quadruplexes

Typically, 50 μM ODN (IDT) in 10 mM Tris-HCl, pH 7.5, with 150 mM KCl or NaCl was heated at 95 ºC for 5-10 min and then either rapidly cooled down in ice or

slowly cooled to room temperature over a couple of hours and then cooled in ice. The formation of G-quadruplexes was confirmed by Circular Dichroism (CD) (Figure.3.3).

CD Experiments

Formation of G-quadruplexes was monitored by CD. Samples were measured immediately after the preparation or after storage for hours to days at 4 ºC. All of the CD spectra were recorded on a J-810 spectropolarimeter (Jasco). The measurements were carried out with 100 µL of 50 µM ODN samples in a 1-mm path quartz cell at 4 ºC under nitrogen to prevent water condensation. Spectra shown (Figure. 3.3) were the average of 3 accumulations in a range from 220 to 300 nm with a band width of 1 nm, response time of 1 s, data pitch of 0.2 nm, and scan speed of 50 nm/min. A blank sample of 10 mM Tris-HCl buffer, pH 7.5, with 150 mM KCl or NaCl was used for baseline correction.

Polyacrylamide Gel Electrophoresis (PAGE) Experiments

The 5' termini of the oligodeoxynucleotides were labeled by a forward reaction with 1 × reaction buffer (500 mM Tris□HCl, pH 7.6, 100 mM $MgCl_2$, 50 mM DTT, 1 mM spermidine, 1 mM EDTA; Fermentas Life Science) with 10UT4 polynucleotide kinase and 20 pmol [γ-^{32}P]-ATP at 37 ºC for 30 min. The reaction was quenched by adding to an equal volume of 2 × loading buffer (98% formamide, 10 mM EDTA), boiling for 2 min, and cooling on ice for another 1 min. The radiolabeled samples were then analyzed on a 15% denaturing polyacrylamide gel and scanned with a PhosphorImager (Bio-Rad Laboratories).

UVB Irradiation

UVB irradiation (280-400 nm with peak intensity at 312 nm) was carried out with two Spectroline XX-15B UV 15-W tubes (rated UV intensity of 1.15 mW/cm2 at 25 cm) filtered by a LONGLIFE filter (Spectronics Corporation) immediately after sample

preparation or after storage at 4 ºC for overnight to days with similar results. G-quadruplex samples were enclosed in a polyethylene Ziplock bag and irradiated on a bed of ice for 22.5 h at a distance of 1 cm from the UVB lamp.

NP1 Digestion

Typically, 1 µL 1 U/µL aqueous NP1 from Penicillium citrinum (Sigma) and 1 µL 10 mM $ZnCl_2$ were added to 100 µL 50 µM UVB irradiation samples and digested at 37 ºC for >36 h. Digested samples were then converted to the ammonium form for mass spectrometry by mixing with the ammonium form of Dowex resin (Sigma).

HF/Pyridine Hydrolysis

Typically, 8 µL 70% HF/pyridine (Sigma) were added to 15 µg dry NP1 digestion product in a polyethylene microcentrifuge tube and incubated at 37 ºC. After 2.5 h, the sample was diluted to 300 µL with Milli-Q water, neutralized with 30 mg calcium carbonate and filtered through an Xpertek 13-mm 0.45-µm nylon syringe filter (P.J. Cobert Associates). The filtrate was evaporated and redissolved in Milli-Q water before HPLC analysis.

HPLC Analysis

Reverse-phase HPLC was carried out with an X-Bridge column (C18, 4.6 × 75 mm, 2.5 µm, 135 Å Waters Corporation) on a System Gold BioEssential HPLC with a Model 125 binary gradient pump and a Model 168 diode array detector (Beckman Coulter). NP1 digestion products were analyzed with a 1 mL/min gradient of 100% A (50 mM triethylammonium acetate, pH 7.5) for 3 min and 0%-20% B for 3-53 min (50% acetonitrile in 50 mM triethylammonium acetate, pH 7.5) and detected at 260 nm. HF/pyridine hydrolysis products were analyzed with an isocratic gradient of 100% Milli-Q water and detected at 205 nm.

ESI-MS Experiments

Intact/whole-length ODN samples were analyzed in the positive-ion mode on a Waters Micromass Q-Tof Ultima spectrometer. The Z-Spray source was operated at 2.8 kV, the cone voltage was 150 V, and RF lens was 50. The source temperature and desolvation temperatures were 80 ºC and 180 ºC, respectively. The collision energy was 10 eV, and the MCP detector was 2,200 V. Spray solvent was 30 mM ammonium acetate in 20% methanol. ODNs from NP1-coupled HPLC separation were analyzed in the negative-ion mode with a Thermo Finnigan LTQ-FT mass spectrometer (Thermo Fisher Scientific). A solution of 50/50 (vol/vol) methanol/water was used as the spray solvent. The spray voltage was 3.5 kV. The capillary voltage and temperature were 46 V and 250 ºC, respectively. MS/MS experiments were done by using CAD with helium as the collision gas. The mass window for precursor-ion selection was 2.5 *m/z* units. The normalized collision energy (15% of the maximum) was adjusted to obtain product ions of good signal-to-noise ratio. At this setting, nearly all of the precursor ions had fragmented.

Results

To determine whether nonadjacent photoproducts can be produced in the G-quadruplex forms of the human telomere, we analyzed the irradiation products of the 22-mer fragment, d[AGGG(TTAGGG)$_3$] (Tel22) and its derivatives by an enzyme-coupled mass spectrometry assay (*17, 20*). This 22-mer fragment was initially chosen because it adopts the basket form in Na^+ solution (*8*), which was expected to lead to efficient photocrosslinking between loops 1 and 3. To determine which nucleotides were involved

in photoproduct formation, we substituted various Ts with Us, which have the same photochemical and base pairing properties but differ in mass.

Formation and Irradiation of Telomeric G-Quadruplex

To prepare the G-quadruplexes, ODNs were first denatured and then either cooled quickly or slowly to 0 ºC with similar results, which is consistent with a rapid rate of folding (*21*). The circular dichroism (CD) spectra of Tel22 in Na^+ and K^+ closely matched previously reported spectra (Figure 3.3). The CD spectra for Tel22 in Na^+ (Figure 3.3A) most closely matched those reported by Li et al., Xu et al., and Gray et al. (*15, 22, 23*) but differed from those reported by Vorlickova et al. and Ambrus et al. (*11, 24*) in the relative magnitude of the 263 nm peak to the 297 nm peak. Likewise, the spectra for Tel22 in K^+ solution (Figure. 3.3B) most closely matched those reported by Ambrus et al., Vorlickova et al., and Gray et al. (*11, 23, 24*) but differed somewhat from that reported by Li et al. (*22*) and more so from that reported by Xu et al. (*15*). These differences, their origin, or their significance, have not, to our knowledge, been addressed in the literature. The CD spectrum for Tel26 in K^+ solution (Figure. 3.3C) matched closely to that reported by Ambrus et al. (*11*). ODNs in which a U was substituted for a T (Figure 3.2A) gave similar CD spectra to those of Tel22. Samples were then irradiated on ice with UVB light for 22.5 h.

Polyacrylamide Gel Electrophoresis (PAGE) of Irradiation Products

Analysis of ^{32}P-end-labeled Tel22 irradiation samples by denaturing PAGE revealed the presence of two major bands I and II (Figure 3.4A). Band I has the same mobility as Tel22, whereas band II had increased mobility, suggesting that altered structures had formed. Most significantly, the rate of formation of band II was greater

with K+ than with Na^+ and is opposite to what had been previously observed for $d(T_4G_4)_4$ and $d(T_2G_4)_4$ (*19*).

Electrospray Ionization-Mass Spectrometry (ESI-MS) of Irradiation Products

For mass spectrometry studies, we submitted nonradiolabeled Tel22 to the same procedure as described above. The individual bands were visualized by brief exposure to UVC light and excised from the gel. The products from irradiation in K^+ solution for 2 h were extracted from bands corresponding to I and II, and from the intervening band. ESI-MS analysis of these products on a quadrupole time-of-flight (Q-Tof) mass spectrometer gave a molecular weight of 6,966, as did those from the unirradiated parent ODN, indicating that they are intramolecular photoproducts.

Nuclease P1 (NP1)-Coupled HPLC Assay of Irradiation Products

NP1 is an endonuclease that degrades photoproduct-containing DNA to smaller photoproduct-containing fragments that are characteristic of the type of photodamage (*20, 25*). NP1 functions by binding to the base portion of a nucleotide after which it hydrolyzes the phosphodiester bond on the 3'-side to yield a deoxynucleotide with a 3'-hydroxyl group (*26, 27*). Because NP1 does not bind to photodimerized bases, it cannot cut on the 3'-side of either nucleotide involved in photodimer formation. As a result, NP1 digestion produces trinucleotides, pd(T=TN), from adjacent thymine dimers, and tetranucleotides, pd[T(N)]=pd[T(N)], from nonadjacent dimers, and mononucleotides, pdN, from undamaged nucleotides (Figure 3.2B) (*17, 20*). With our HPLC method, mononucleotides elute before 20 min, trinucleotides from 20 to 35 min, and tetranucleotides from 35 to 45 min, as confirmed by ESI-MS analysis of the eluents. HPLC analysis of the NP1 digestion products of Tel22 irradiated in Na^+ solution showed a sharp peak corresponding to an adjacent dimer and a small broad peak corresponding to

nonadjacent thymine dimers (Figure 3.5A). Analysis of the irradiation products in the presence of K^+, however, showed numerous additional sharp peaks corresponding to nonadjacent thymine dimers (Figure 3.5B).

DNA sequences with a folded secondary structure have increased mobility in gel electrophoresis compared with the unfolded structure (*19, 28-30*). NP1 digestion and mass spectral analysis of the faster moving band II corresponding to a species formed in K^+ solution showed tetranucleotides (peaks 7, 8, 9, 10, and 11) as the major products (Figure 3.4B). The higher mobility also indicates formation of photodimers that crosslink the DNA into loops, thereby increasing the mobility relative to that of the unmodified ODN. The NP1 digestion products of the intermediate band contained a mixture of trinucleotides (peaks 2 and 3) and tetranucleotides (peaks 7, 8, 9, 10, and 11). NP1 digestion and analysis of band I (Figure 3.4B), however, revealed the presence of only trinucleotides (peaks 2 and 3), indicative of adjacent thymine dimers, which would not be expected to change the mobility of the oligodeoxynucleotides, as was observed.

Nucleotide Composition of the Trimer and Tetramer NP1 Degradation Products

We separated the NP1 degradation products by HPLC and determined their nucleotide compositions by ESI-MS and MS/MS analysis. The nucleotides were identified first by their molecular weight deduced from the MS and second by tandem mass spectrometry (MS/MS) experiments. For example, those nucleobases that were not involved in dimer formation were cleaved from the parent ion upon collisional activation. Thus, HPLC peak 3 was assigned to pd(T=TA) because it gives a molecular ion $[M - H]^-$ at *m/z* 938.2 and a characteristic product ion $[M - Ade - H]^-$ at *m/z* 803.1 (Figure 3.6A). We also detected a small amount of the TA* photoproduct (peak 2) (Figure 3.7B) (*31, 32*). HPLC peaks 7 and 8 were both assigned as isomeric pd[T(A)]=pd[T(A)] on the basis

of a molecular ion $[M - 2H]^{2-}$ at *m/z* 634.3 and characteristic product ions $[pdA - H]^-$ at *m/z* 330.1 and $[M - Ade - 2H]^{2-}$ at *m/z* 566.6 (Figure 3.4C). Using similar reasoning, we assigned HPLC peak 11 to pd[T(T)]=pd[T(A)] (Figure 3.8A). Some peaks appeared to be contaminated with products from flanking peaks and were assigned to the major component, whereas others appeared to contain two products (Table 3.1).

Stereochemistry of the Thymine Photodimers

The stereochemistry of the thymine photodimers was assigned by comparing the base portion of the photoproducts with authentic thymine cyclobutane dimers (*17, 33, 34*) (Figure 3.6B). The base portion of the photodimers was obtained by treating the trinucleotides and tetranucleotides with 70% hydrogen fluoride in pyridine (HF/pyridine) (Figure 3.2B). HPLC peak 3 was thereby determined to correspond to the *cis,syn* thymine photodimer (Figure 3.6 B and C), and peaks 7, 8, 9, 10, and 11 to thymine photodimers of either *cis,anti* or *trans,anti* stereochemistry or both (Figure 3.4D, Figure 3.9B, Figure 3.8B, and Table 3.1). The *syn* stereochemistry is consistent with photodimer formation between adjacent thymines, whereas the anti stereochemistry is expected for photodimer formation between thymines in opposing loops (interstrand-type product) (Figure 3.2B) (*17*). More *cis,anti* and *trans,anti* products were detected than expected based on their symmetry properties (one *t,a* and two *c,a* for pd[T(A)]=pd[T(A)] suggesting that there may be additional isomers because of restricted rotation.

Mapping of Thymine Photodimer Sites

Because the ODN sequence is repetitive, it is not possible to identify directly the specific thymines involved in adjacent and nonadjacent thymine photodimer formation by analysis of the trinucleotide and tetranucleotide degradation products. To solve this problem, we labeled the various loops by constructing three telomere sequences Tel22-U^n

(n = 1, 2, and 3), in which the second T in the TTA loop n was replaced with U, which undergoes the same photodimerization as T but lacks a C5 methyl group (Figure 3.2). Owing to the substrate specificity of NP1, the second pyrimidine is always retained in the trimeric digestion products of an adjacent pyrimidine dimer formed within a single loop, and in the tetrameric digestion products of a photodimer formed between two different loops. Thus, if a pyrimidine photodimerization takes place within or between a loop containing a U, the tri- and tetra-nucleotide digestion products will have different HPLC retention times, and the molecular ions will be 14 u lower than those only involving loops containing only Ts. Furthermore, analysis of the characteristic product ions in the ESI-product-ion spectra can establish whether the U is involved in the photodimerization.

Irradiation of Tel22-U^2 in K^+ followed by NP1 digestion resulted in a nearly identical set of photoproducts as that from Tel22, with the exception of a photoproduct (HPLC peak 2-1, Figure 3.5C), which was identified as pd(T=UA) (Figure 3.7A). This sole U-containing product indicates that the Ts in loop 2 do not photodimerize with Ts in the other 2 loops. In contrast, irradiation and digestion of both Tel22-U^1 and Tel22-U^3 led to HPLC peaks, n-1, n-4, n-6, and n-12 (n=1 and 3, respectively, where n refers to the loop containing the U) (Figure 3.5 D and E). These peaks corresponded to the same photoproducts found in Tel22 except that a T was replaced by a U. The results clearly establish the involvement of loops 1 and 3 in thymine photodimerization, and furthermore, that both Ts of loop 1 photodimerize preferentially with only the second T of loop 3. The basis for the latter conclusion is that Tel22-U^1 yielded primarily pd[T^1(U)]=pd[T^3(A)] and pd[U^1(A)]=pd[T^3(A)] (Figure 3.10), whereas Tel22-U^3 gave primarily pd[T^1(T)]=pd[U^3(A)] and pd[T^1(A)]=pd[U^3(A)] (Figure 3.11). The detailed assignment of photoproducts of Tel-U^n is given in Table 3.2.

It is noteworthy that we observed a broad HPLC peak in the digestion products of irradiated Tel22 (peak 5), and the sequences substituted with U (peaks 15, 25, and 35), whether irradiated in Na^+ or K^+ solution (Figure 3.5). Analysis of this peak indicates that it corresponds to a mixture of thymine photodimers between different pairs of loops (Table 3.2). We speculate that some or all of these photoproducts are not produced from specific G-quadruplex structures but from a small fraction of nonspecifically folded isomers that exist in equilibrium with these structures.

Mechanistic Implications.The inefficient formation of anti thymine photodimers in the Na^+ form of Tel22, which has a basket structure with two proximate lateral loops (*8*) (Figure 3.1), was unexpected. Even more unexpected was the efficient formation of anti thymine photodimers between loops 1 and 3 in the K^+ form of Tel22, which is thought to exist primarily in hybrid-1 and/or hybrid-2 conformations that lack proximate loops (*11, 13*). These hybrid structures, however, are also thought to be in dynamic equilibrium with either or both basket and chair conformations (*11, 13, 15*) that have proximate lateral loops that could photocrosslink (Figures 3.1 and 3.12). Given that the basket form which Tel22 adopts in Na^+ did not facilitate photocrosslinking between loops 1 and 3, it is unlikely that the basket form in K^+ would be any better. It is possible, however, that interloop photodimer formation would be more favorable in the chair form, which has an antiparallel arrangement of loops 1 and 3 compared with the parallel arrangement in the basket form. The higher preference for isomeric pd[T(A)]=pd[T(A)] products (peaks 7 and 8) (16% yield) over isomeric pd[T(T)]=pd[T(A)] products (6.5% yield), is in agreement with a report that the Ts and As involved in the chair conformation can form a T-A-T-A quartet (*22*).

Another explanation for the enhanced photocrosslinking in K^+ comes from the recent discovery that the closely related 22-mer human telomeric sequence d[$(GGGTTA)_3GGGT$] adopts a different basket structure (Form III) with K^+ (*9*) than the one found with Na^+ (*8*) (Figure 3.12). The Form III basket structure corresponds to the Na^+ basket structure in which the first GGG sequence misaligns to disrupt the first G quartet and to increase the size of the first loop by one nucleotide. The increased size of loop 1 in the Form III structure in K^+ is expected to better facilitate photodimer formation between loop 1 and loop 3 than the basket structure formed in Na^+. In support of this, the Oxytricha telomeric sequence $d(T_4G_4)_4$, which adopts a basket structure with two 4-nucleotide loops in Na^+ (*35*), photocrosslinks efficiently between the third T in loop 1 and the third T in loop 3 (*19*).

A third possibility is that the triplex intermediates involved in interconverting the various quadruplex structures (*11, 23, 36*) can themselves facilitate anti thymine dimer formation, and are more accessible and/or stable in K^+ than in Na^+. To test whether triplexes might be involved, we prepared a truncated form of Tel22, Tel19 (Figure 3.2A), that lacks the first three Gs necessary to complete the fourth strand of the quadruplex and, thus, limits the structure to intermediate triplex forms. Deleting the 3 Gs or replacing them with As in Tel22-A_3 (Figure 3.2A) resulted in comparable amounts of the same anti thymine dimers as formed in Tel22 (Figure 3.13 A and B, respectively), except that the relative yields of isomeric *cis,anti* and *trans,anti* pd[T(A)]=pd[T(A)] (peaks 7 and 8) were greatly diminished. These results show that although formation of anti thymine photodimers between loops 1 and 3 can presumably proceed through intermediate G-quadruplex-like triplex structures, they are not produced in the same relative yields as observed for the G-quadruplex form of Tel22.

To determine whether or not photocrosslinking between loops 1 and 3 could be suppressed by choosing a sequence known to more greatly favor a hybrid structure, we investigated the photochemistry of Tel26, d[AAA(GGGTTA)$_3$GGGAA]. This sequence was found by NMR to initially adopt a mixture of hybrid-1 and -2 structures in K^+ that quantitatively converts to the hybrid-1 structure on incubation (*11, 13*). Even though the hybrid-1 structure lacks proximate loops, the overall yield of anti thymine photodimers in Tel26 was nearly the same as was found for Tel22, except that there was an even greater preference for *trans,anti* pd[T(A)]=pd[T(A)] between loops 1 and 3 (peak 8) (Figure 3.13C). These results suggest that the photocrosslinking of G quadruplexes in K^+ is taking place through a highly photoreactive structure other than a triplex, such as the chair or Form III basket, that is in rapid equilibrium with the hybrid structures, and that favors the *trans,anti* dimer between the central Ts of loops 1 and 3.

Conclusion

We have shown that the G-quadruplex form of the human telomere sequence Tel22 is susceptible to both adjacent and interloop-specific thymine photodimer formation in the presence of biologically relevant K^+. The adjacent thymine photodimer has the same *cis,syn* stereochemistry as do the major cyclobutane photodimers induced by UVB light in duplex DNA. The interloop-specific thymine photodimers, however, are all of the *anti* stereochemistry, which is not observed in duplex DNA irradiated under native conditions (*16*). Thus, *anti* thymine dimers, and in particular their pattern of formation, may represent a unique photochemical signature for the potassium quadruplex form of the human telomeric sequences studied herein, and possibly for other folded DNA structures. Extrapolating from these studies of model systems, we suggest that

human telomeric DNA may also be susceptible to UVB-induced *anti* thymine photodimer formation *in vivo*, in addition to *cis,syn* dimer formation, causing hitherto unrecognized effects on the replication and stability of telomere DNA. Formation of a covalently linked loop of 13-14 nt is also expected to pose a unique challenge to repair systems. It remains to be seen, however, whether *anti* thymine photodimers are produced in human telomere DNA *in vivo* and if so, how they affect telomere biochemistry. Nuclease P1 coupled LC-MS has been used to detect psoralen photocrosslinks in human genomic DNA (*37*), but more sensitive ^{32}P-postlabeling assays, or ^{14}C-postlabeling coupled with accelerator mass spectrometry (*38*), may be required to detect *anti* thymine photodimers in telomeres.

Acknowledgment

I thank Huafeng Fang for his contribution to this work by doing PAGE experiments. We thank Professor Robert Blankenship for use of his CD instrument. This work supported by National Institutes of Health (NIH) Grant CA40463 and Washington University National Institutes of Health Mass Spectrometry Resource Grant 2P41RR000954.

References

(1) Gilson, E. and Geli, V. (2007) How telomeres are replicated. *Nat. Rev. Mol. Cell Biol. 8*, 825-838.

(2) Verdun, R. E. and Karlseder, J. (2007) Replication and protection of telomeres. *Nature 447*, 924-931.

(3) Riethman, H. (2008) Human telomere structure and biology. *Annu. Rev. Genomics Hum. Genet. 9*, 1-19.

(4) Patel, D. J., Phan, A. T. and Kuryavyi, V. (2007) Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res. 35*, 7429-7455.

(5) Dai, J., Carver, M. and Yang, D. (2008) Polymorphism of human telomeric quadruplex structures. *Biochimie 90*, 1172-1183.

(6) Gaynutdinov, T. I., Neumann, R. D. and Panyutin, I. G. (2008) Structural polymorphism of intramolecular quadruplex of human telomeric DNA: effect of cations, quadruplex-binding drugs and flanking sequences. *Nucleic Acids Res. 36*, 4079-4087.

(7) Lane, A. N., Chaires, J. B., Gray, R. D. and Trent, J. O. (2008) Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res. 36*, 5482-5515.

(8) Wang, Y. and Patel, D. J. (1993) Solution structure of the human telomeric repeat d[$AG_3(T_2AG_3)_3$] G-tetraplex. *Structure 1*, 263-282.

(9) Lim, K. W., Amrane, S., Bouaziz, S., Xu, W., Mu, Y., Patel, D. J., Luu, K. N. and Phan, A. T. (2009) Structure of the human telomere in K^+ solution: a stable

basket-type G-quadruplex with only two G-tetrad layers. *J. Am. Chem. Soc. 131*, 4301-4309.

(10) Parkinson, G. N., Lee, M. P. and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature 417*, 876-880.

(11) Ambrus, A., Chen, D., Dai, J., Bialis, T., Jones, R. A. and Yang, D. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res. 34*, 2723-2735.

(12) Luu, K. N., Phan, A. T., Kuryavyi, V., Lacroix, L. and Patel, D. J. (2006) Structure of the human telomere in K^+ solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc. 128*, 9963-9970.

(13) Dai, J., Carver, M., Punchihewa, C., Jones, R. A. and Yang, D. (2007) Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K^+ solution: insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res. 35*, 4927-4940.

(14) Phan, A. T., Kuryavyi, V., Luu, K. N. and Patel, D. J. (2007) Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution. *Nucleic Acids Res. 35*, 6517-6525.

(15) Xu, Y., Noguchi, Y. and Sugiyama, H. (2006) The new models of the human telomere d[AGGG(TTAGGG)$_3$] in K^+ solution. *Bioorg. Med. Chem. 14*, 5584-5591.

(16) Douki, T., Laporte, G. and Cadet, J. (2003) Inter-strand photoproducts are produced in high yield within A-DNA exposed to UVC radiation. *Nucleic Acids Res. 31*, 3134-3142.

(17) Su, D. G., Kao, J. L., Gross, M. L. and Taylor, J. S. (2008) Structure determination of an interstrand-type cis-anti cyclobutane thymine dimer produced in high yield by UVB light in an oligodeoxynucleotide at acidic pH. *J. Am. Chem. Soc. 130*, 11328-11337.

(18) Xu, Y., Tashiro, R. and Sugiyama, H. (2007) Photochemical determination of different DNA structures. *Nat. Protoc. 2*, 78-87.

(19) Williamson, J. R., Raghuraman, M. K. and Cech, T. R. (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell 59*, 871-880.

(20) Wang, Y., Taylor, J. S. and Gross, M. L. (1999) Nuclease P1 digestion combined with tandem mass spectrometry for the structure determination of DNA photoproducts. *Chem. Res. Toxicol. 12*, 1077-1082.

(21) Gray, R. D. and Chaires, J. B. (2008) Kinetics and mechanism of K+- and Na+-induced folding of models of human telomeric DNA into G-quadruplex structures. *Nucleic Acids Res. 36*, 4191-4203.

(22) Li, J., Correia, J. J., Wang, L., Trent, J. O. and Chaires, J. B. (2005) Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic Acids Res. 33*, 4649-4659.

(23) Gray, R. D., Li, J. and Chaires, J. B. (2009) Energetics and kinetics of a conformational switch in G-quadruplex DNA. *J. Phys. Chem. B 113*, 2676-2683.

(24) Vorlickova, M., Chladkova, J., Kejnovska, I., Fialova, M. and Kypr, J. (2005) Guanine tetraplex topology of human telomere DNA is governed by the number of (TTAGGG) repeats. *Nucleic Acids Res. 33*, 5851-5860.

(25) Wang, Y., Gross, M. L. and Taylor, J. S. (2001) Use of a combined enzymatic digestion/ESI mass spectrometry assay to study the effect of TATA-binding

protein on photoproduct formation in a TATA box. *Biochemistry 40*, 11785-11793.

(26) Volbeda, A., Lahm, A., Sakiyama, F. and Suck, D. (1991) Crystal structure of Penicillium citrinum P1 nuclease at 2.8 A resolution. *EMBO J. 10*, 1607-1618.

(27) Romier, C., Dominguez, R., Lahm, A., Dahl, O. and Suck, D. (1998) Recognition of single-stranded DNA by nuclease P1: high resolution crystal structures of complexes with substrate analogs. *Proteins 32*, 414-424.

(28) Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I., Jr. and Blackburn, E. H. (1987) Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Cell 51*, 899-908.

(29) Murchie, A. I. and Lilley, D. M. (1994) Tetraplex folding of telomere sequences and the inclusion of adenine bases. *EMBO J. 13*, 993-1001.

(30) Xue, Y., Kan, Z. Y., Wang, Q., Yao, Y., Liu, J., Hao, Y. H. and Tan, Z. (2007) Human telomeric DNA forms parallel-stranded intramolecular G-quadruplex in K+ solution under molecular crowding condition. *J. Am. Chem. Soc. 129*, 11185-11191.

(31) Bose, S. N., Davies, R. J., Sethi, S. K. and McCloskey, J. A. (1983) Formation of an adenine-thymine photoadduct in the deoxydinucleoside monophosphate d(TpA) and in DNA. *Science 220*, 723-725.

(32) Zhao, X., Nadji, S., Kao, J. L. and Taylor, J. S. (1996) The structure of d(TpA), the major photoproduct of thymidylyl-(3'-5')-deoxyadenosine. *Nucleic Acids Res. 24*, 1554-1560.

(33) Cadet, J., Voituriez, L., Hruska, F., KAN, L.-S., LEEUW, F. A. M. D. and A., C. (1985) Characterization of thymine ultraviolet photoproducts. *Can. J. Chem. 63*, 2861-2868.

(34) Shetlar, M. D., Basus, V. J., Falick, A. M. and Mujeeb, A. (2004) The cyclobutane dimers of 5-methylcytosine and their deamination products. *Photochem. Photobiol. Sci. 3*, 968-979.

(35) Wang, Y. and Patel, D. J. (1995) Solution structure of the Oxytricha telomeric repeat d[$G_4(T_4G_4)_3$] G-tetraplex. *J. Mol. Biol. 251*, 76-94.

(36) Mashimo, T., Sannohe, Y., Yagi, H. and Sugiyama, H. (2008) Folding pathways of hybrid-1 and hybrid-2 G-quadruplex structures. *Nucleic Acids Symp. Ser. (Oxf)*, 409-410.

(37) Cao, H., Hearst, J. E., Corash, L. and Wang, Y. (2008) LC-MS/MS for the detection of DNA interstrand cross-links formed by 8-methoxypsoralen and UVA irradiation in human cells. *Anal. Chem. 80*, 2932-2938.

(38) Farmer, P. B. and Singh, R. (2008) Use of DNA adducts to identify human health risk from exposure to hazardous environmental pollutants: the increasing role of mass spectrometry in assessing biologically effective doses of genotoxic carcinogens. *Mutat. Res. 659*, 68-76.

Table 3.1. Stereochemistry, base composition, and loop assignment of the NP1 digestion products of UV-irradiated Tel22 shown in Figure 3.5.

HPLC Peak	Stereochemistry	Base Composition
2	----	T=AG*
3	*c,s*†	T=TA
5	*c,a*	T(T)=T(A) T(A)=T(A)
7	*c,a*‡	$T^1(A)=T^3(A)$
8	*t,a*‡	$T^1(A)=T^3(A)$
9	*c,a*¶	$T^1(A)=T^3(A)$
10	*c,a* & *t,a*‖	$T^1(T)=T^3(A)$ $T^1(A)=T^3(A)$
11	*t,a*§	$T^1(T)=T^3(A)$

Table 3.2. Stereochemistry, composition, and loop assignment of the NP1 digestion products of Tel22-U^n (n=1, 2, 3) in Figure 3.5 C, D, and E. Loop assignment indicated by the superscripted number.

HPLC Peak	Stereochemistry	Base Composition
1-1, 2-1, 3-1	*c,s*	T=UA
1-2, 2-2, 3-2*	----	T=AG
1-3, 2-3, 3-3*	*c,s*	T=TA
1-4	----	$U^1(A)=T^3(A)$ $U^1(A)=T^3(T)$†
1-5	----	$T^1(U)=T(A)$ $U^1(A)=T(A)$
1-6	----	$T^1(U)=T^3(A)$ $U^1(A)=T^3(A)$
1-8	----	$U^1(A)=T^3(A)$ $T^1(U)=T^3(A)$†
1-12	----	$U^1(A)=T^3(A)$
2-5	----	$T(A)=T^2(U)$ $T(A)=U^2(A)$ $T^1(T)=T^3(A)$ $T^1(A)=T^3(A)$
2-7*	*c,a*	$T^1(A)=T^3(A)$
2-8*	*t,a*	$T^1(A)=T^3(A)$
2-9*	*c,a*	$T^1(A)=T^3(A)$
2-10*	*c,a; t,a*‡	$T^1(T)=T^3(A)$ $T^1(A)=T^3(A)$
2-11*	*t,a*	$T^1(T)=T^3(A)$
3-4	----	$T^1(T)=U^3(A)$ $T^1(A)=U^3(A)$
3-5	----	$T(A)=U^3(A)$ $T(A)=T^3(U)$§
3-6	----	$T^1(T)=U^3(A)$ $T^1(A)=U^3(A)$
3-8	----	$T^1(A)=T^2(A)$ $T^1(A)=U^3(A)$¶ $T^1(A)=T^3(U)$¶
3-12	----	$T^1(A)=U^3(A)$

Loop assignment indicated by the superscripted number.

*Same products as assigned for Tel22 in Table 1.

†Less than 20% estimated from MS peak intensity.

‡Could not be assigned to a specific product.

§Less than 10% estimated from MS peak intensity.

¶Less than 15% estimated from MS peak intensity.

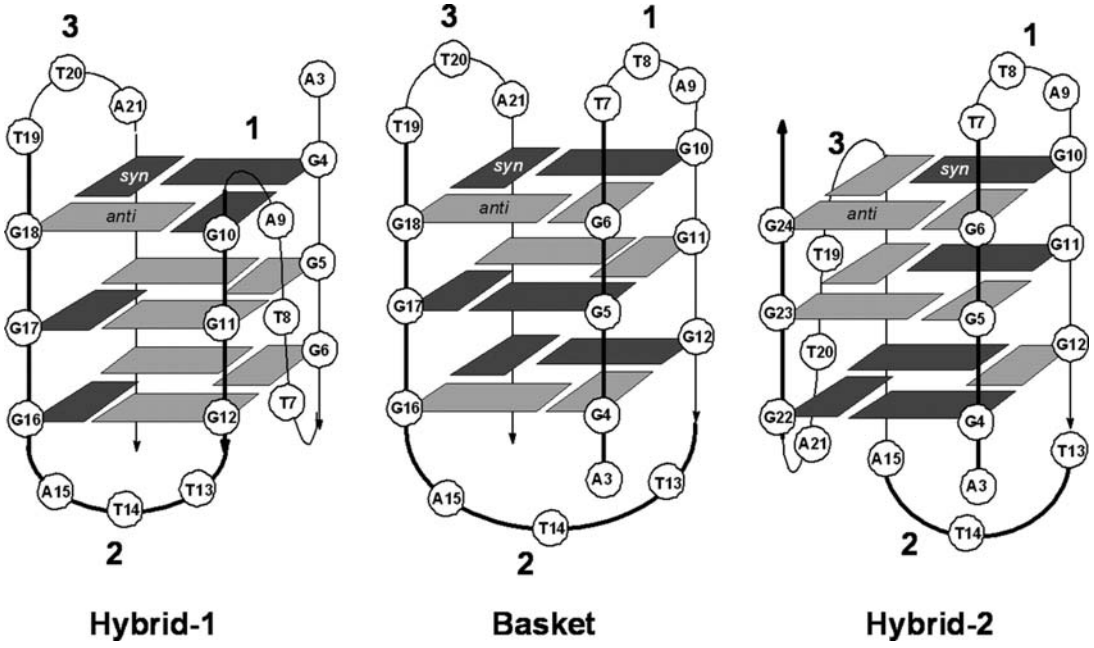


Figure 3.1. Known basket, hybrid-1, and hybrid-2 triple quartet forms of human telomeric DNA that could be adopted by Tel22. The numbering system for the bases starts from 1 in the human telomere sequence $d(TTAGGG)_n$. The loops are numbered sequentially, and dark and light gray boxes represent syn and anti guanines, respectively.

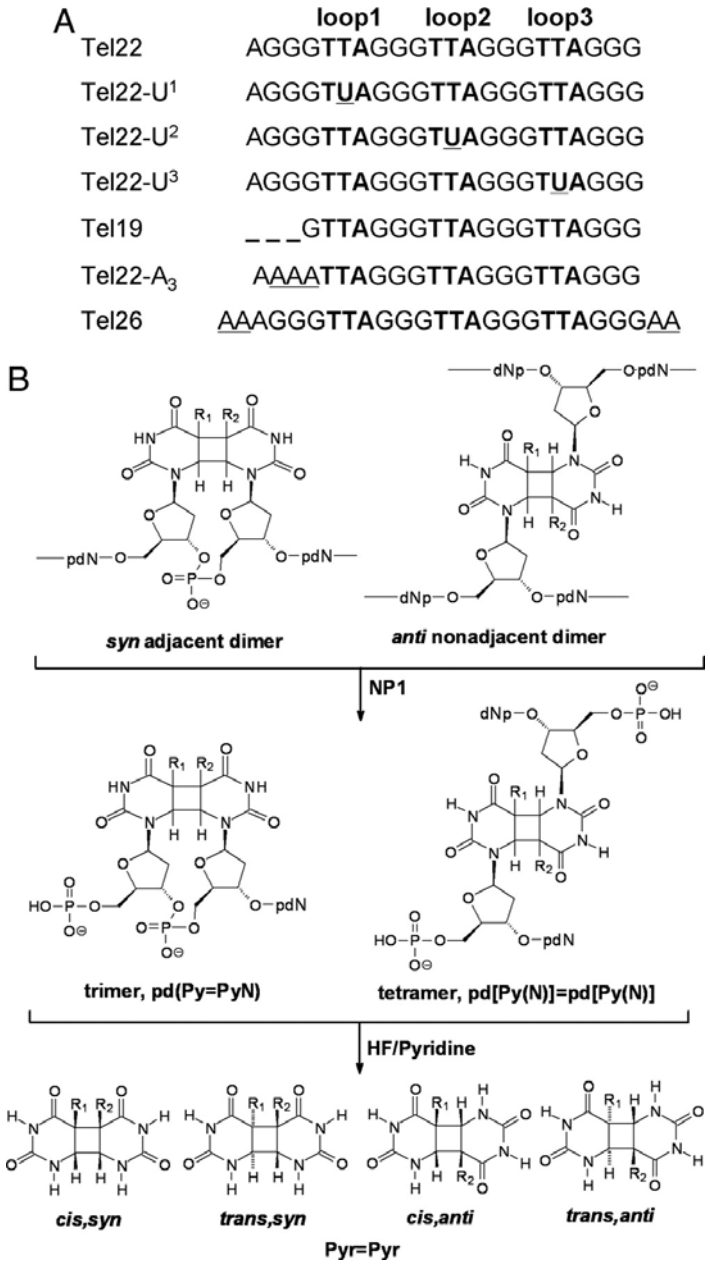


Figure 3.2. Telomeric substrates and cyclobutane pyrimidine dimer photoproducts. (A) Telomeric sequences used in the present work. The superscript n in U^n (n = 1, 2, and 3, respectively) refers to the loop in which the second T was replaced by U. (B) Adjacent and nonadjacent photoproducts and their nuclease P1 (NP1) digestion products and HF/pyridine hydrolysis products. R = CH_3 for thymine, and R = H for uracil. *Trans,syn* and *cis,anti* pyrimidine dimers exist as enantiomeric pairs.

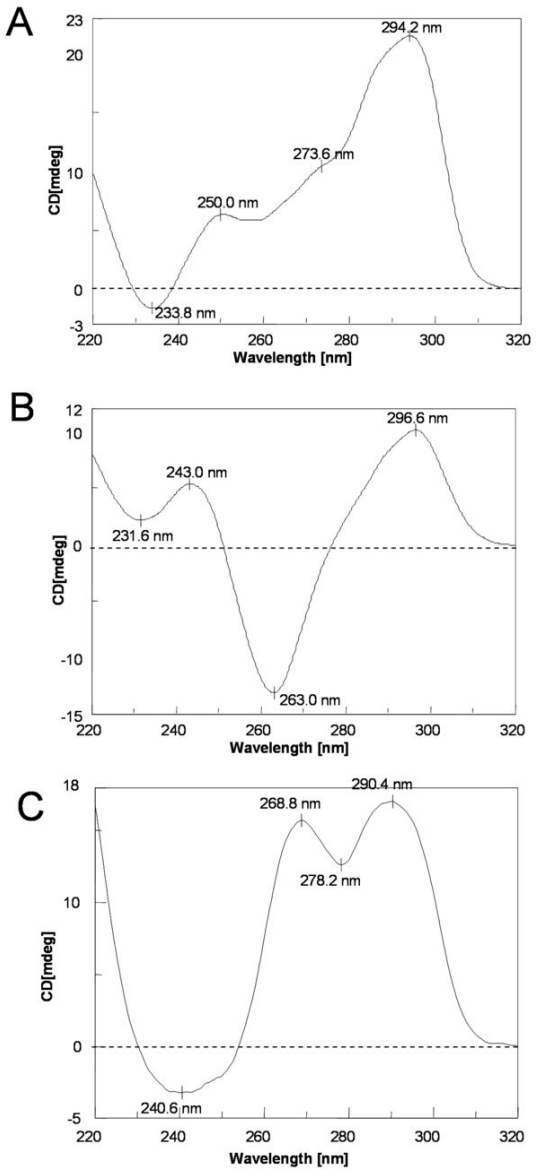


Figure 3.3. CD spectra of 50 μM Tel22 G-quadruplexes in 10mM Tris-HCl (pH 7.5) at 4 ºC. (A) Tel22 in 150 mM K^+ solution after denaturing and renaturing process. (B) Tel22 in 150 mM Na^+ solution after denaturing and renaturing process. (C) CD spectrum of 50 μM Tel26 G-quadruplex at 4 ºC after 24 h.

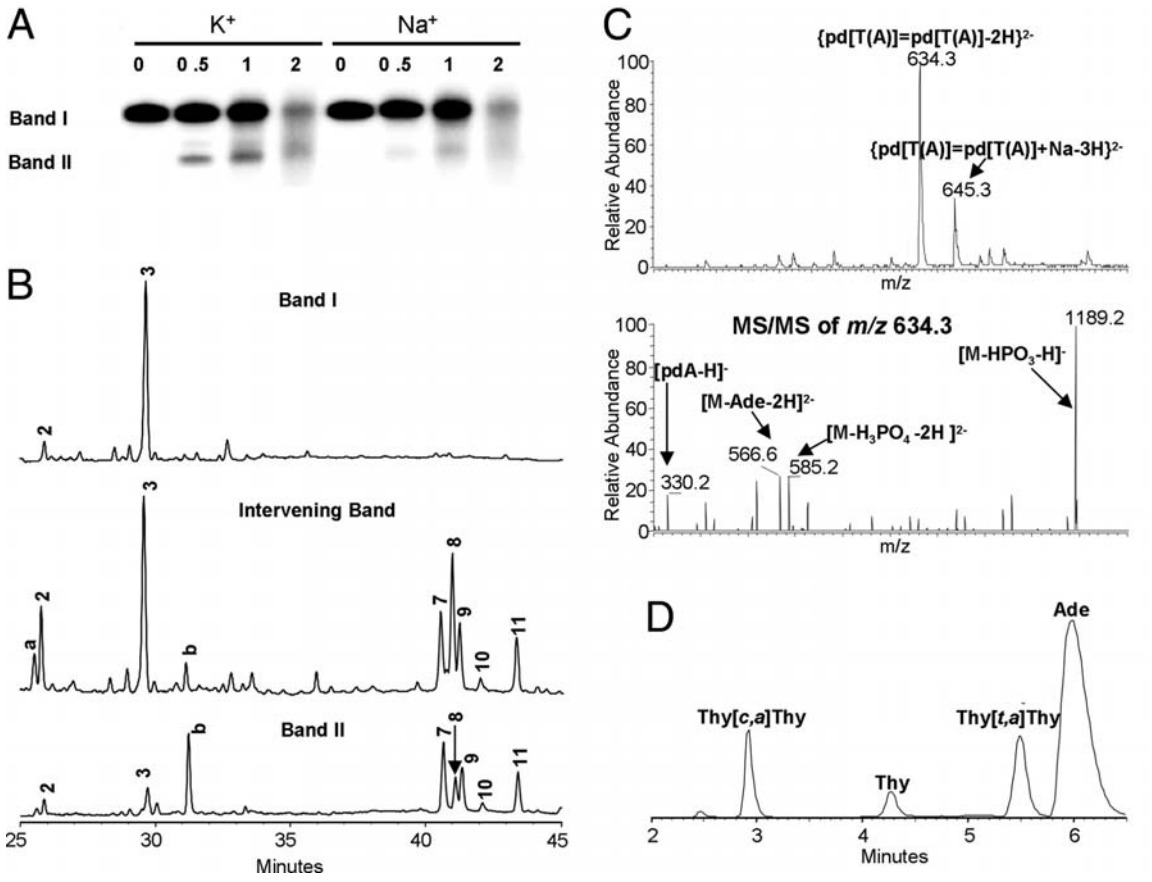


Figure 3.4. Analysis of UVB irradiated telomeric quadruplexes. (A) Tel22 was UVB irradiated in 150 mM Na^+ or K^+ solution at pH 7.5 for the number of hours indicated in the lane headings, ^{32}P-end-labeled, and then subjected to 15% denaturing PAGE. Bands I and II and the intervening band were excised and subjected to NP1 digestion. (B) NP1-coupled HPLC assay of nonradiolabeled samples corresponding to bands I, II, and the intervening band. HPLC peaks a and b were not observed in nonelectrophoresed samples and could not be assigned. Products were detected at 260 nm where the nondimerized bases have their maximum absorption. (C) Mass spectra and product-ion mass spectra of the major nonadjacent photoproduct HPLC peak 8. The observation of $[pdA - H]^-$ and $[M - Ade - 2H]^{2-}$ indicate that adenine is the nonphotodimerized base. (D) HPLC assay of HF/pyridine hydrolysis products of HPLC peak 8 that was contaminated with adjacent

peaks. The hydrolysis products were detected at 205 nm where the photodimers have their maximum absorbance, and their stereochemistry assigned by correlation with authentic thymine photodimers (Figure. 3.6B). The small thymine peak may be caused by the decomposition of thymine anti dimers in the course of hydrolysis (17). Analysis of various HPLC fractions established peak 8 to be the *trans,anti* photodimer.

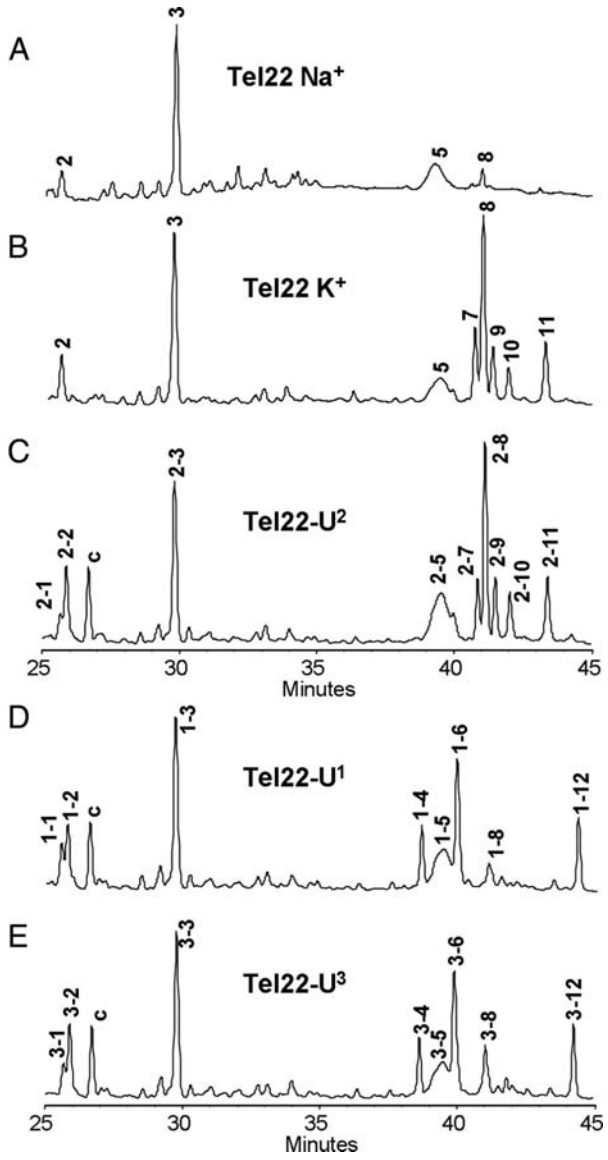


Figure 3.5. NP1-coupled HPLC assay of Tel22 and the U-substituted sequences after 2.5 h UVB irradiation at pH 7.5. (A) Tel22 in 150 mM Na^+ solution. (B) Tel22 in 150 mM K^+ solution. (CE) Tel22-U^n in 150 mM K^+ solution. The HPLC peaks are labeled n-m, in which n refers to the loop that is substituted by U, and m refers to the order of elution. Peak c could not be assigned. Peaks 1-3 are trinucleotides resulting from adjacent photoproducts, and peaks 4-12 are tetranucleotides resulting from nonadjacent photoproducts.

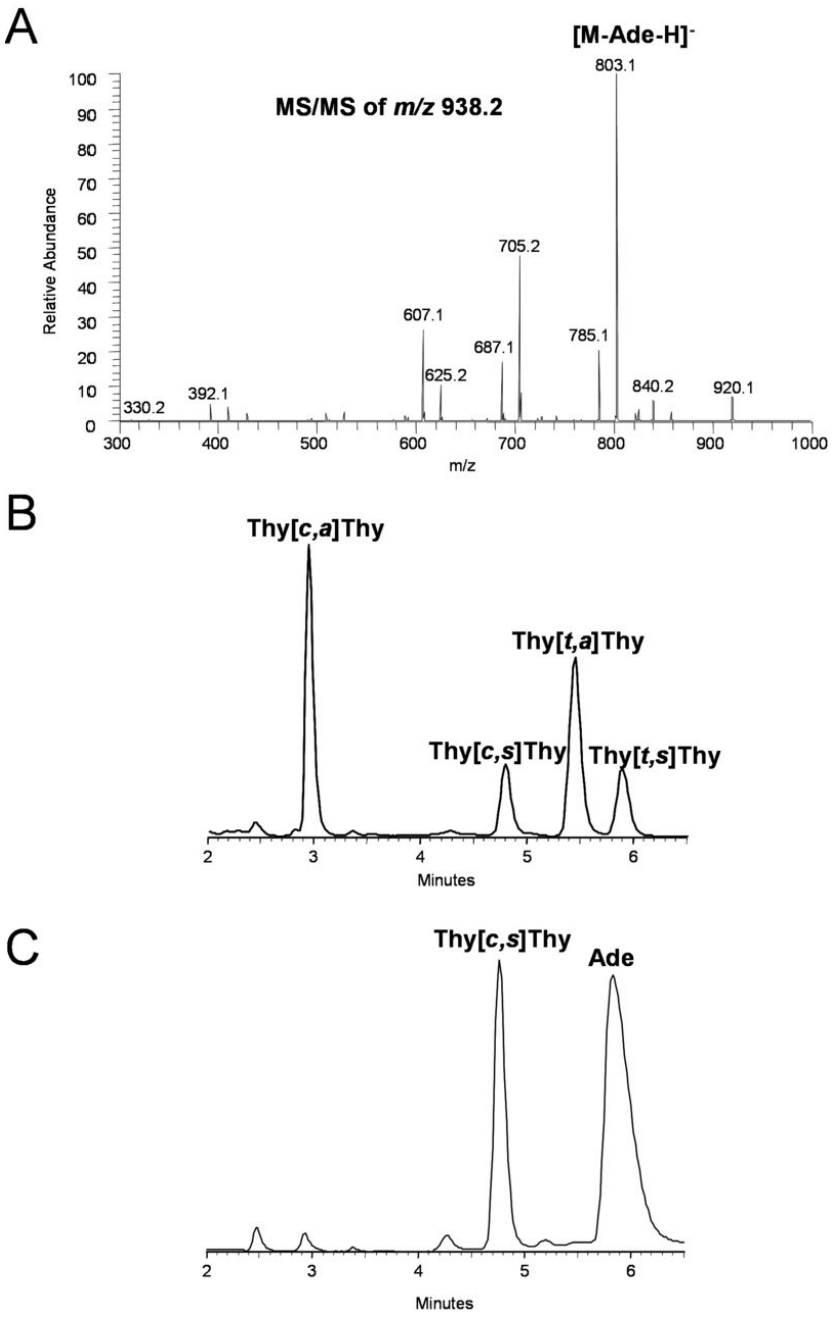


Figure 3.6. ESI-MS/MS and HF/pyridine hydrolysis assay of HPLC peak 3. (A) ESI-MS/MS of peak 3, (B) HPLC of authentic thymine cyclobutane dimers, and (C) HF/pyridine hydrolysis-coupled HPLC assay of peak 3.

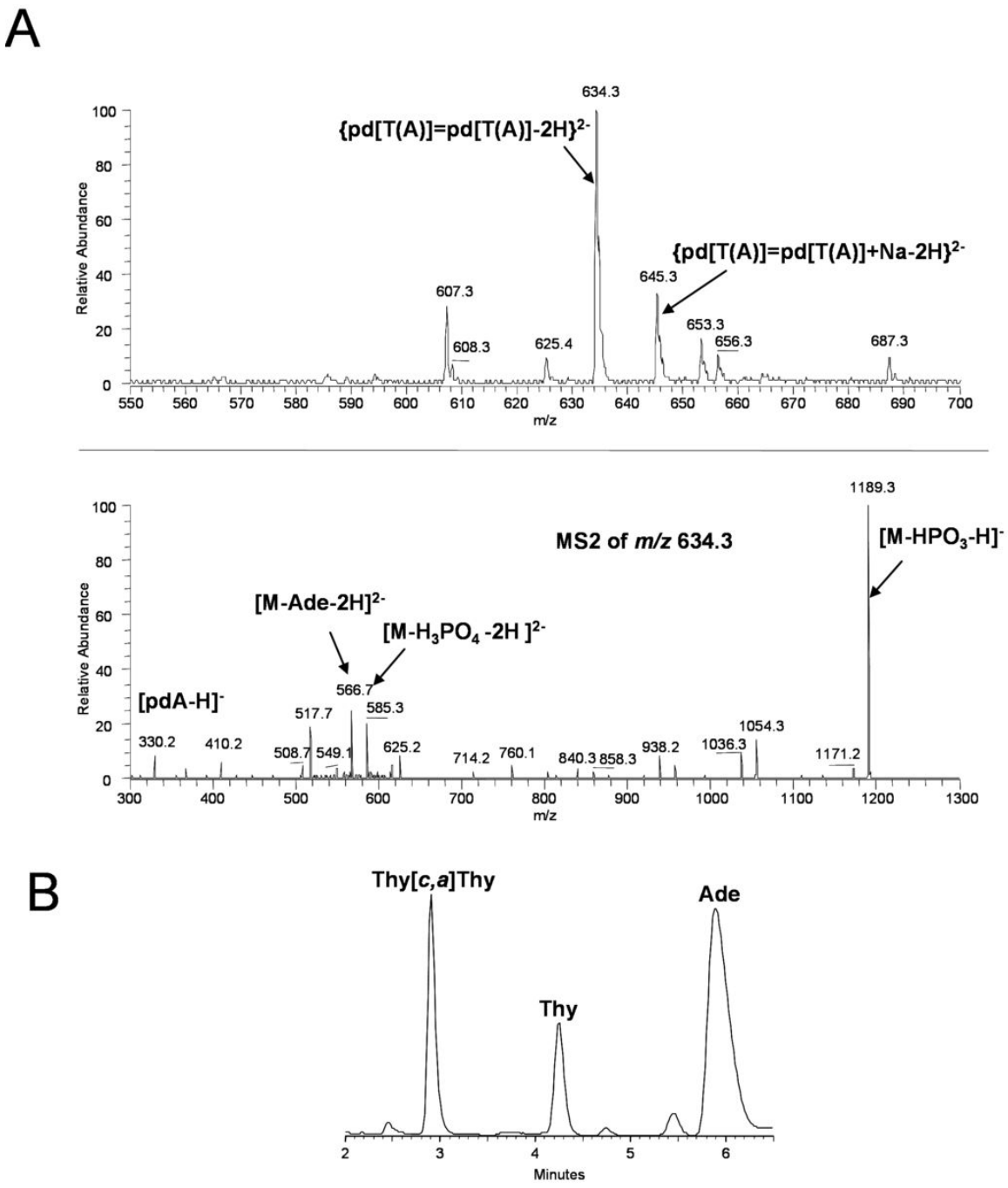


Figure 3.7. ESI-MS/MS of (A) HPLC peak n-1 (n = 1, 2, and 3), pd[T=UA] and (B) HPLC peak n -2 (n = 1, 2, and 3), pd[T=AG] corresponding to the TA* photoproduct.

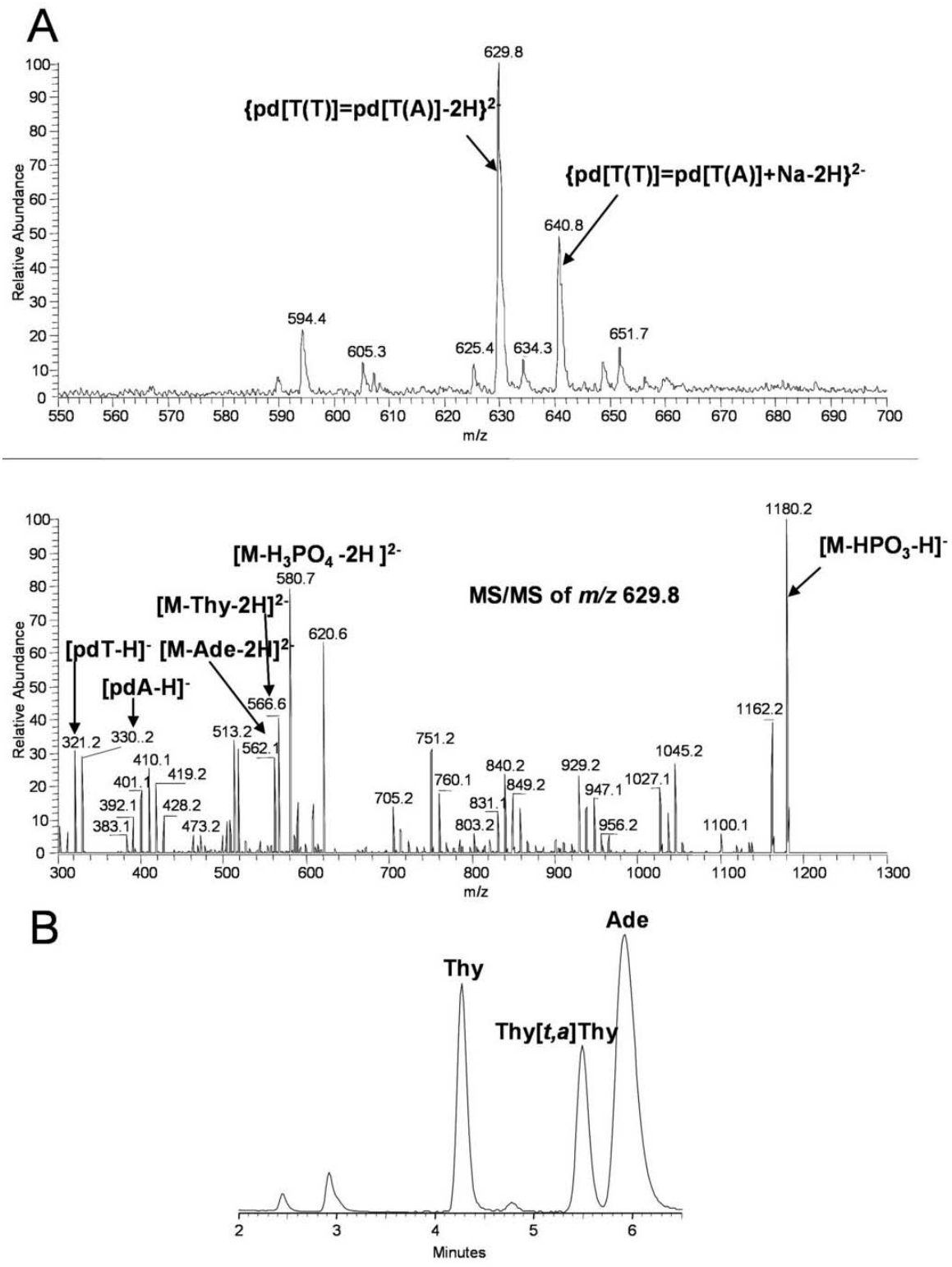


Figure 3.8. ESI-MS, ESI-MS/MS, and HF/pyridine hydrolysis-coupled HPLC assay of HPLC peak 11. (A) ESI-MS and MS/MS. (B) HF/pyridine hydrolysis-coupled HPLC.

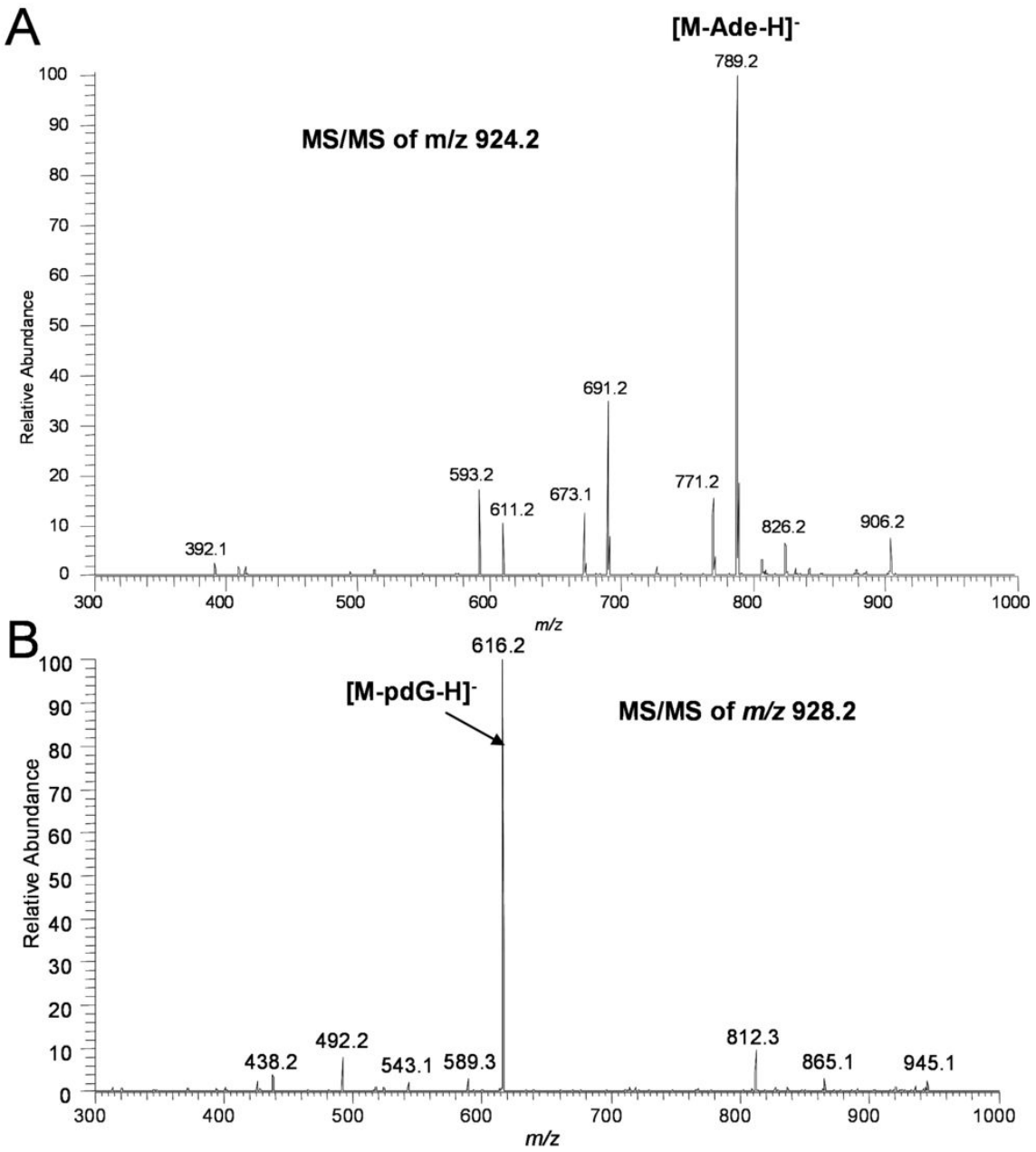


Figure 3.9. ESI-MS, ESI-MS/MS, and HF/pyridine hydrolysis-coupled HPLC assay of HPLC peak 9. (A) ESI-MS and MS/MS. (B) HF/pyridine hydrolysis-coupled HPLC assay.

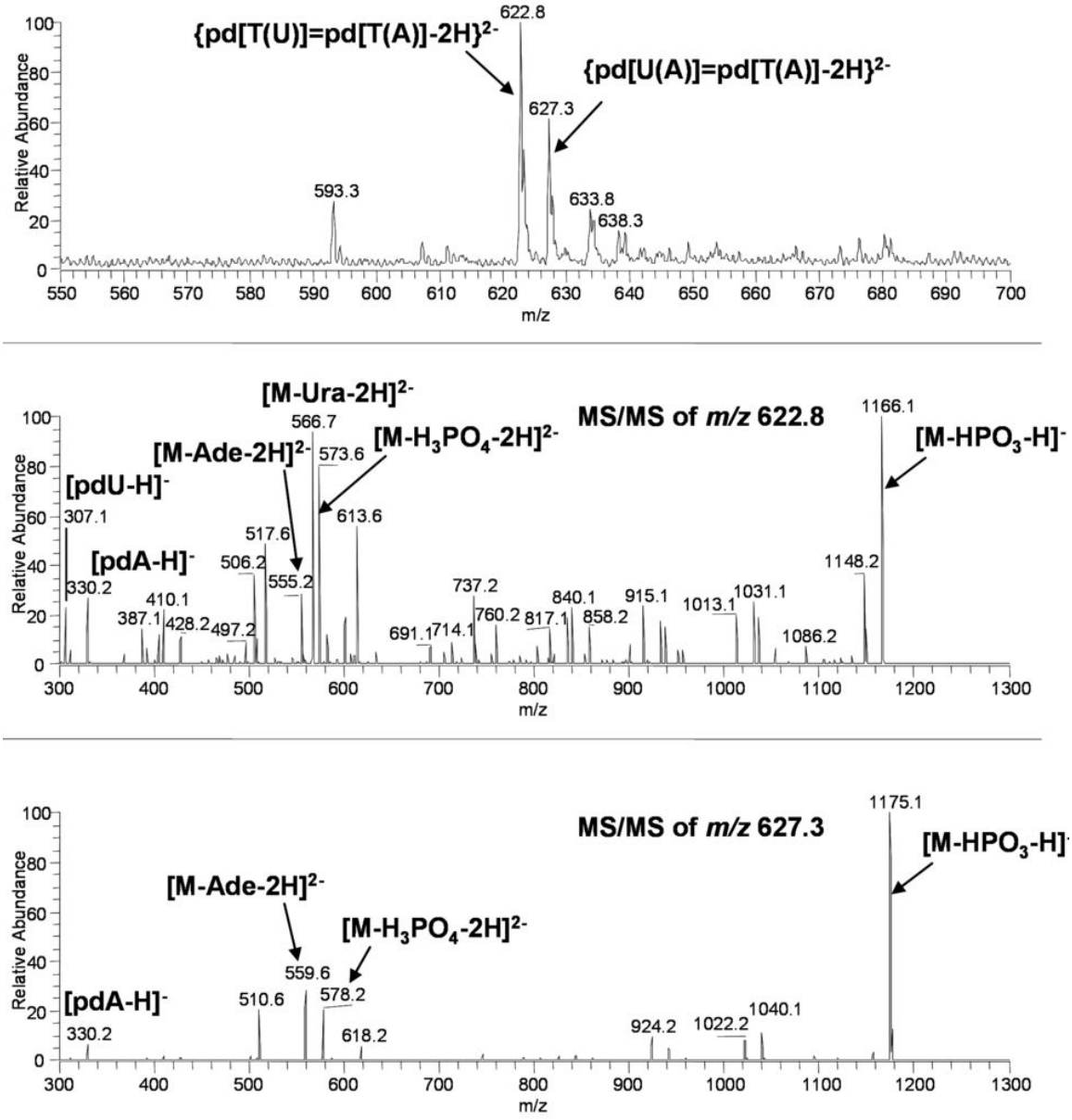


Figure 3.10. ESI-MS and ESI-MS/MS of HPLC peak 1-6.

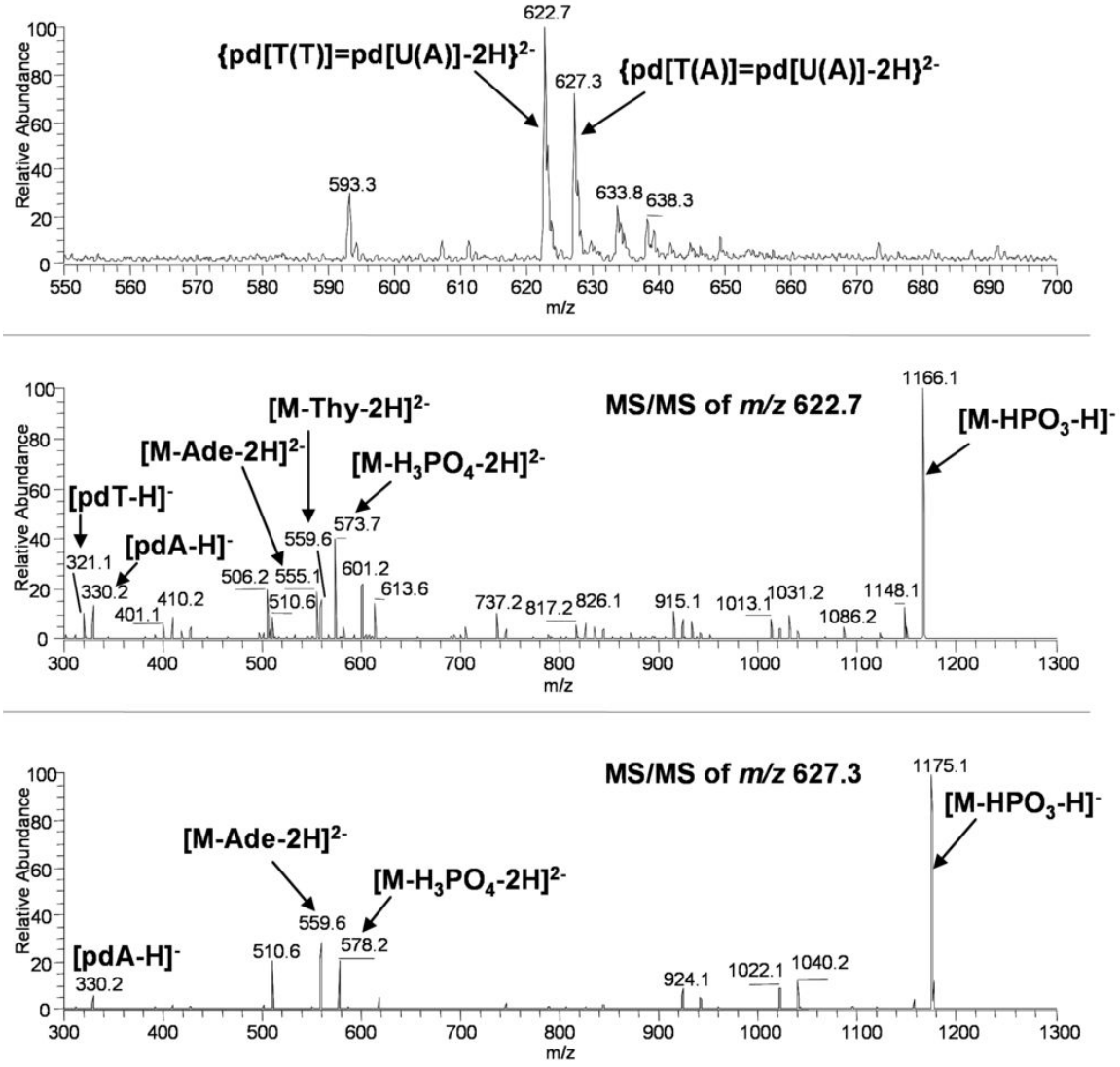


Figure 3.11. ESI-MS and ESI-MS/MS of HPLC peak 3-6.

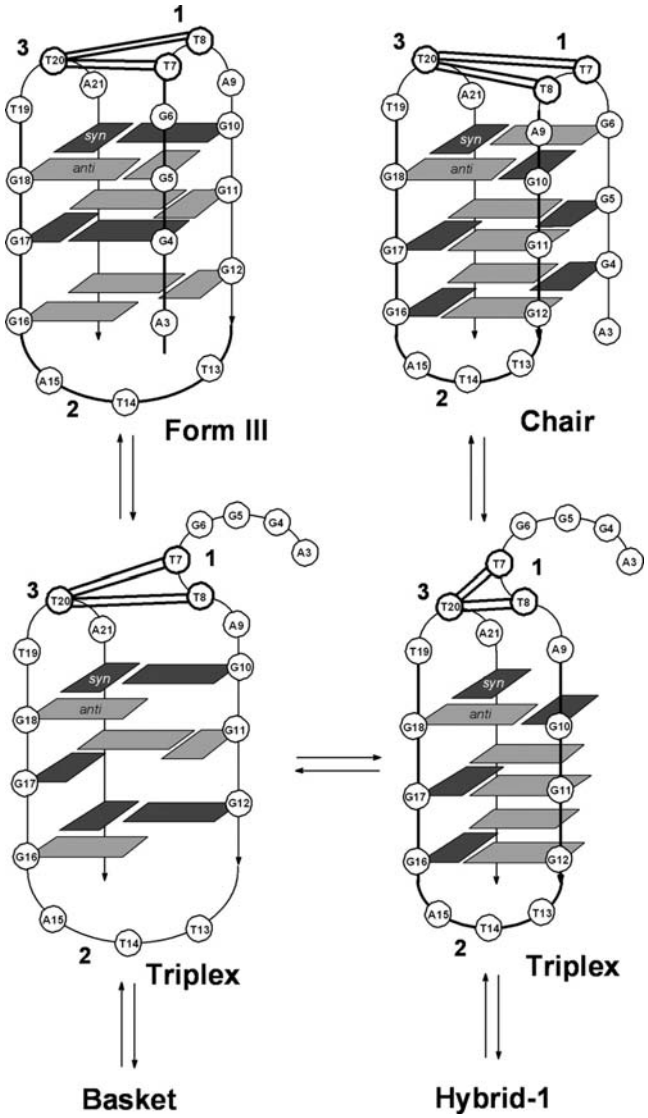


Figure 3.12. Photodimer sites mapped onto chair and Form III G-quadruplex forms of Tel22 and onto triplex intermediates possibly involved in equilibrating between these and other structures. Solid double lines indicate sites of observed *anti* thymine photodimer formation between loops 1 and 3.

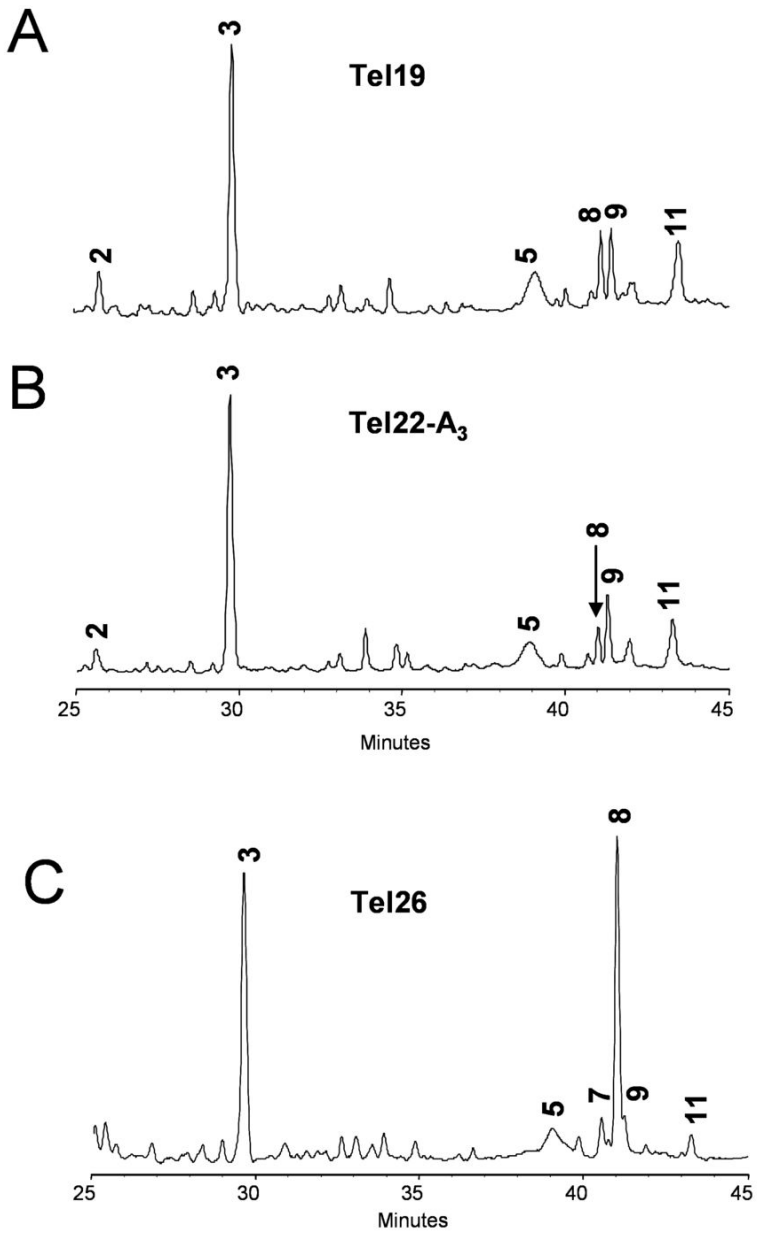


Figure 3.13. NP1-coupled HPLC assay of UV-irradiated Tel19, Tel 22-A_3, and Tel26. (A) Tel19 in K^+ solution after 2.5 h UVB irradiation. (B) Tel 22-A_3 in K^+ solution after 2.5 h UVB irradiation. (C) NP1-coupled HPLC assay of Tel26 in K^+ solution (see Figure 3.3) after 2.5 h UVB irradiation. *Anti* dimers of Tel26 were mapped to loops 1 and 3 by replacement with U.

Chapter 4

^{m}CA*: A New Photoproduct of 5' - Methylcytosine and Adenine Characterized by HPLC and Mass Spectrometry

Abstract

The UV portion of sunlight is mutagenic and can modify DNA by producing various photoproducts. UV photodamage often occurs at dipyrimidine sites, to give cyclobutane, pyrimidine-(6-4)-pyrimidone (6-4), and pyrimidine-(6-4)-Dewar pyrimidone (Dewar) photoproducts, and at TA and AA sites. There is no reported evidence, however, of UV-photoproduct formation between C or 5-methylC (^{m}C) and A. Irradiation of d(GTAT^{m}CATGAGGTGC) with UVB light at physiological pH gives an unexpected photoproduct that undergoes fast thermal deamination but does not revert to its original structure under UVC irradiation. Evidence from nuclease P1 (NP1) digestion coupled with electrospray ionization (ESI)-MS/MS is in accord with product formation between ^{m}C and A. HPLC analysis indicates that deamination gives a T<>A photoproduct that coelutes on reverse-phase chromatography with the well-known TA* photoproduct, formed from an initial [2+2] reaction between C5-C6 and C6-C5 of the adjacent thymine and adenine (as shown by Zhao, X., et al. (1996) *Nucleic Acids Res. 24*, 1554-1560 and Davies, R. J. et al. (2007) *Nucleic Acids Res. 35*, 1048-1053). Furthermore, the deamination product of the unknown ^{m}C<>A photoproduct and the TA* photoproduct undergo nearly identical fragmentation in tandem MS. The evidence, taken together, indicates that the deamination product of the unknown ^{m}CA photoproduct has the same chemical structure as the TA* photoproduct. Therefore, the unknown photoproduct is referred to ^{m}CA* photoproduct, which, upon deamination, gives the TA* photoproduct.

Introduction

Sunlight-induced DNA photo damage is one cause of mutations leading to skin cancer. The toxic, mutagenic and carcinogenic effects of sunlight, however, have been principally attributed to the ultraviolet (UV) portion, which can produce a variety of photoproducts in DNA (*1-5*). If not repaired, these lesions lead to mutations, increasing the risk for cancer. Until now, the most prevalent UV photoproducts are found at pyrimidine sites, giving cyclobutane dimers, cyclobutane, pyrimidine-(6-4)-pyrimidone (6-4)[1], and their pyrimidine-(6-4)-Dewar pyrimidone (Dewar) valence isomers (*6-10*). Other UV photoproducts containing adenine such as TA*, which forms from an initial [2+2] cycloaddition reaction between C5-C6 and C6-C5 of the adjacent thymine and adenine (*11, 12*), A=A and AA* (*13-16*) are also well known. No UV photoproduct formation between other pairs of bases, however, has been reported.

Mutations due to UV irradiation are predominantly C-to-T transition and CC-to-TT tandem transitions at a dipyrimidine sequence (*17, 18*). Most C-to-T mutation hotspots occur at methylated CpG sites (*19-21*). The most common hypothesis is that the C-to-T transition occurs after the deamination of C or 5-methylC (^{m}C), leading to insertion of A instead of a G in the next replication cycle (*22*). The important deamination reaction of ^{m}C-containing photoproducts is affected by flanking bases, pH, and secondary structure (*23*).

In the course of synthesizing T^{m}C photoproducts of d(GTAT^{m}CATGAGGTGC) produced by irradiation with biologically relevant UVB (290-320 nm) light for deamination studies (Scheme 4.1), we found an unknown photoproduct. To elucidate the structure of the unknown, we employed enzyme digestion, correlation of HPLC retention

times of unknown and reference materials, and mass spectrometry (MS). The combined approach allowed us to identify the unknown as a ^{m}CA* photoproduct (photoproduct formed between 5-methylcytosine and adenine that deaminates to give the TA* product), the name of which indicates covalent linkages between C5-C6 and C6-C5 of ^{m}C and A as in TA*. The identification took advantage of its rapid deamination to afford the well-known TA* photoproduct (*11, 12, 16, 24, 25*). In this chapter, we describe the structural determination of this unusual and first known photoproduct between ^{m}C and A.

Materials and Methods

Materials

Oligodeoxynucleotides (ODNs) were purchased from Integrated DNA Technologies, Inc. (IDT) (Coralville, Iowa). Nuclease P1 (NP1) from *Penicillium citrinum* was from Sigma (St. Louis, MO). Milli-Q (18.2 mΩ/ cm) water was prepared by using a Millipore Corporation (Billerica, MA) Milli-Q water purification system. HPLC solvents were from Fisher Scientific (Fair Lawn, NJ).

Instrumentation

UVB irradiation was carried out with a 302 nm transilluminator (Spectroline model TR-302) or a 312 nm lamp with two Spectroline® XX-15B UV 15-watt tubes (peak UV intensity of 1,150 μW/cm² at 25 cm from the source) from Spectronics Corporation (Westbury, New York). UVC (100-290 nm) irradiation was with a model UVG-254 Mineralight lamp (254 nm, Ultra-Violet Products, Inc., San Gabriel, CA). HPLC separation and analysis were carried out with a System Gold HPLC BioEssential HPLC with binary gradient 125 pumps and a diode array Model 168 detector (Beckman Coulter, Inc., Fullerton, CA). An X-Bridge column (C18, 4.6 ×75 mm, 2.5 μm, 135 Å)

from Waters Corporation (Milford, MA) was used for reversed-phase HPLC (RP-HPLC). Mass spectra were obtained with a Thermo Finnigan LTQ-FT mass spectrometer (Thermo Fisher Scientific, Inc., Waltham, MA) or an Applied Biosystems 4700 tandem time-of-flight mass spectrometer (Applied Biosystems, Foster City, CA).

UV Irradiation of ODNs

ODNs from IDT were used without further purification. ODN solutions for UV irradiation were prepared as 50 μM in 10 mM tris(hydroxymethyl)aminomethane HCl (pH 7.6). Samples were then purged with nitrogen gas for 5 min and placed in a polyethylene Ziplock bag. For 302 nm irradiation, the bags were placed on the transilluminator and irradiated for 2.5 h in a cold room (4 ºC). For the 312 nm and UVC irradiations, the bags were placed on ice and irradiated for 2.5 h at a distance of approximately 1 cm from the lamp. All samples from irradiation were immediately transferred to Eppendorf tubes and stored in a -80 ºC freezer to prevent deamination.

RP-HPLC Analysis and Purification of Products

Basic solvents were used for HPLC to prevent any ^{m}C-containing photoproducts from deamination during separation. The flow rate was 1 mL/min with a gradient of 0-20% solvent B in solvent A for 50 min (solvent A: 50 mM triethylammonium acetate, pH 7.8; solvent B: 50% acetonitrile in 50 mM triethylammonium acetate, pH 7.8). Elutes were immediately put on dry ice and then stored in a -80 ºC freezer until MS analyses could be conducted.

NP1 Digestion

Separated photoproducts were collected accompanying the HPLC separation, put on dry ice, and then stored in a -80 ºC freezer. HPLC fractions were thawed on ice

before digestion. A 2-5 μL aliquot was mixed with 0.5 μL, 1 unit/μL NP1 and digested on ice for 3-5 min. The digested samples were then submitted to MS analysis.

MS Analysis

Electrospray ionization (ESI) mass spectra were obtained in the negative-ion mode on a LTQ-FTICR instrument. Typically, samples were introduced to the mass spectrometer by injecting in a 10 μL loop. A solution of 50/50 (v/v) methanol/water was used as the spray solvent at a flow rate of 5 μL /min. The spray voltage was 3.5 kV. The capillary voltage and temperature were 46 V and 250 ºC, respectively. ESI-MS/MS experiments were done by using collision-activated dissociation with helium as the collision gas. The mass window for precursor-ion selection was 2.5 *m/z* units, and the collision energy was 30% of the maximum value, producing product ions of good signal-to-noise ratio. Exonuclease digestion coupled MALDI-MS experiments were done on the 4700 MALDI TOF/TOF instrument by using the procedures previously reported (*26*).

Results and Discussion

HPLC Analysis of the New Photoproduct

After irradiation of d(GTAT^{m}CATGAGGTGC) with 302 nm light, RP-HPLC analysis showed four significant photoproducts (PPs): PP1 (~ 18% yield based on HPLC peak areas), PP2, PP3, and PP4 (Figure 4.1A, upper panel). After incubating the irradiation mixture in a 37 ºC water bath for 1 h to cause deamination, we found that the peaks corresponding to PP1 and PP2 decreased, whereas two new peaks, labeled as PP1* and PP2* increased (Figure 4.1A, lower panel). PP1 deaminates faster than PP2 judging from HPLC peak areas: approximately 60% of PP1 had deaminated whereas only 30% of PP2 had done so. These new peaks may correspond to deamination products of ^{m}C

photoproducts. When 312 nm light was used, we saw only two significant photoproducts PP1 (~ 22% yield based on HPLC peak areas) and PP4 in the HPLC analysis **(**Figure 4.1B, upper panel). Thermal incubation of the irradiation mixture, as described above, caused the peak corresponding to PP1 to decrease and to be replaced by a new peak labeled as PP1* (Figure 4.1B, lower panel). The sum of peak areas corresponding to PP1 and PP1* is roughly that of the original peak PP1, an observation that is consistent with the assignment that PP1* corresponds to deamination of an ^{m}C photoproduct in PP1. Furthermore, PP2* is likely to be the deamination product of PP2, which is also an ^{m}C-containing photoproduct.

NP1-coupled ESI-MS and ESI-MS/MS Assay

When we treated the photodamaged ODN with nuclease P1, we found a trinucleotide triphosphate, which upon mass spectrometric analysis, gave ions of *m/z* of 937 for PP1, PP2, PP3 and PP4, and of *m/z* 938 for PP1* and PP2*. These are the expected digestion products of an ODN containing a photoproduct between adjoining bases (*27, 28*). The trinucleotide that gives an ion of *m/z* 937 has a composition of [pd(T<>^{m}CA) - H]$^{-}$ or [pd(^{m}C<>AT) - H]$^{-}$, and the one of *m/z* 938 is either [pd(T<>TA) - H]$^{-}$ or [pd(T<>AT) - H]$^{-}$. To resolve these possibilities, we carried out MS/MS experiments on selected ions of the tri-nucleotides released from the ODN photoproducts in which only the non-damaged base can be lost while the photochemically coupled bases cannot be fragmented. We found the most abundant product ions of *m/z* 615 and 616 for PP1 and PP1*, respectively, both of which were formed by loss of a 2'-deoxythymidine-5'-phosphate (pdT) (Figure 4.2). These observations allow us to propose that the trinucleotides formed in NP1 digestion of PP1 and PP1* are pd(^{m}C<>AT) and pd(T<>AT), respectively; the latter is a deamination product of the former.

We also observed that the most abundant fragment ions are of *m/z* of 802 and 803 for PP2, PP2*, respectively; these arise from a loss of 2'-deoxyadenosine-5'-phosphate (pdA) (Figure 4.3). We can now propose that the trinucleotides from PP2 and PP2* are pd(T<>^{m}CA) and pd(T<>TA), respectively; the latter is a deamination product of the former. Similarly, we found that the most abundant product ions are of *m/z* of 802 for the digestion products of PP3 (Figure 4.4) and PP4 (Figure 4.5), arising from losses of pdA. The trinucleotides liberated in the digestion of PP3 and PP4 are both assigned as pd(T<>^{m}CA), because they can only lose a pdA species.

Stereochemistry of PP2, PP3 and PP4

Given that PP1 and PP2 deaminate, they may be cyclobutane photoproducts (*6, 29, 30*). It is known that UVC light can revert [2 + 2] cyclobutane photoproducts to their parent ODNs and convert Dewar photoproducts to their (6-4) valence isomers, which are detectable at 325 nm. Deamination does not change the stereochemistry of the photoproducts. We carried out UVC irradiation with 254 nm light on the HPLC-separated PP1*, PP2*, PP3 and PP4. Given that no photoreversion or photoconversion of PP1* occurred, we can exclude PP1 and PP1* as cyclobutane photoproducts. PP2* and PP3, however, reverted to the parent ODN, indicating that they are cyclobutane photoproducts of *cis-syn* (*c,s*) or *trans-syn* (*t,s*) stereochemistry. We assigned PP2 and PP2* as ^{m}C-containing (*c,s*) photoproduct because PP2 deaminates rapidly. We assigned PP3 as a ^{m}C-containing (*t,s*) photoproduct because it undergoes no significant deamination under the conditions used here, as would be expected on the basis of previous work (*6, 10, 29*). PP4 undergoes a transition to its (6-4) isomer, suggesting a Dewar photoproduct (Table 4.1).

Comparison of PP1 and the authentic TA* Photoproduct by HPLC and NP1-coupled ESI-MS^n*

Considering that PP1 is an ^{m}C<>A photoproduct and that it may also contribute to ^{m}C to T mutations via deamination, we were motivated to determine its structure. Interestingly, PP1 deaminates even faster than the T[*c,s*]^{m}C photoproduct. Given that deamination doesn't change the stereochemistry of a photoproduct, we decided to investigate PP1*, the thermally stable deamination product of PP1. Further, given that the deamination product PP1* contains a T<>A photoproduct but is not a cyclobutane adduct, we suspected that it has the structure of the well known TA* photoproduct (*11, 12*). Thus, we compared the RP-HPLC retention time of the T<>A photoproduct formed from deamination of PP1* to that of the authentic TA* photoproduct formed within the same sequence context.

The authentic TA* photoproduct was obtained by UVC (254 nm) of d(GTATTATGAGGTGC). We assigned the reference photoproducts on the basis of photoconversion, photoreversion, and NP1-coupled ESI-MS/MS as described above (Figure 4.6). Particularly, we can distinguish T5A6* from T2A3* with bovine spleen phosphodiesterase-coupled ladder sequencing by MALDI-MS (data not shown) (*26, 31, 32*), taking advantage of exonuclease digestion that goes from 5′ to 3′ and terminates just before T5. The approach was to inject onto the HPLC the same amount of sample individually and then to co-inject them. Co-elution of PP1* and the T5A6*-containing photoproduct 14-mer strongly suggests that the two photoproducts have the same structure (Figure 4.6). This suggestion was then confirmed by NP1-coupled ESI-MS^n analysis, whereby both MS^2 and MS^3 of the NP1 digested PP1* and authentic T5<>A6 photoproduct 14-mer show the same fragmentation pattern (Figure 4.7).

As shown in the product-ion spectra, both compounds undergo the expected major neutral loss of pdT yielding a most abundant product ion of *m/z* 616. An MS^3 experiment shows that the *m/z* 616 ion formed from the precursor ion gives the same major product ion of *m/z* 438. The fragmentation to give the *m/z* 438 ion is a loss of the 2-deoxyribose ring with the attached phosphate group (pdF). The other distinctive product ion is of *m/z* 314, which is most abundant in an MS^4 experiment whereby the *m/z* 438 ion is probed as formed by the sequence *m/z* 938 → *m/z* 616 ion → *m/z* 438 ion. The ion of *m/z* 314 likely arises by the loss of phosphoric acid (H_3PO_4), possibly via an ion-dipole complex, accompanied by the elimination of acetylene (C_2H_2). These fragments are consistent with a photoproduct structure made up of a stable tricyclic center that loses groups on the periphery. More importantly, they add support that the newly discovered ^{m}C<>A photoproduct is an analog of the well-known TA* photoproduct (Scheme 4.2).

Conclusion

We discovered a new photoproduct of d(GTAT^{m}CATGAGGTGC) involving reaction of ^{m}C and A. Deamination gave the well-known TA* photoproduct (*11, 12, 16, 24, 25*) whose structure was previously completely established by NMR (*11*), MS (*16*), and X-ray crystallography (*12*), and its mutation spectrum determined in bacteria (*33*). Comparisons of HPLC retention times and tandem MS show that the properties of the ^{m}C<>A deamination product are nearly identical to those of the authentic TA* photoproduct, substantiating their identical structures. Interestingly, TA* forms under UVC whereas ^{m}CA arises under UVB. This is consistent with the UV absorbance of thymine (λ_{max} = 267 nm), which has greater overlap with UVC whereas ^{m}C (λ_{max} = 278 nm) has greater overlap with the UVB lamp.

The ^{m}C residues in DNA play an important role in regulation of gene expression and chromatin structure (*34, 35*), and UV-induced damage of ^{m}C could be mutagenic. Although methylation of C predominantly occurs at d(CpG) sites in mammals (*35, 36*), it is conceivable that an inter-strand type ^{m}CA* could form in DNA that is folded to bring ^{m}C and A proximate to each other (*26, 37-39*). Non-CG methylation is present primarily in viral or stably integrated plasmid DNA sequences (*40*) and has also been detected in significant abundance in human embryonic stem cells and induced pluripotent stem cells (*41, 42*), although photochemical damage should not be an issue for them, and as well as in plants (*43-45*). Further photochemistry studies could be conducted targeting a series of DNA duplexes containing a ^{m}CA site and folded DNA having ^{m}C and A in close proximity to explore the possibility of intra-and inter-strand type ^{m}CA* formation, respectively. Future studies will also required to define better the biological consequences of ^{m}CA* photoproduct formation.

Acknowledgments

This research was supported by the Washington University NIH Mass Spectrometry Research Resource (Grant No. P41 RR000954) and by NIH (Grant CA40463).

References

(1) Tyrrell, R. M. (1994) The molecular and cellular pathology of solar ultraviolet radiation. *Mol. Aspects Med. 15*, 1-77.

(2) Hemminki, K., Xu, G. and Le Curieux, F. (2001) Ultraviolet radiation-induced photoproducts in human skin DNA as biomarkers of damage and its repair. *IARC Sci. Publ. 154*, 69-79.

(3) Cleaver, J. E. and Crowley, E. (2002) UV damage, DNA repair and skin carcinogenesis. *Front Biosci. 7*, d1024-1043.

(4) Cadet, J., Sage, E. and Douki, T. (2005) Ultraviolet radiation-mediated damage to cellular DNA. *Mutat. Res. 571*, 3-17.

(5) Pfeifer, G. P., You, Y. H. and Besaratinia, A. (2005) Mutations induced by ultraviolet light. *Mutat. Res. 571*, 19-31.

(6) Shetlar, M. D., Basus, V. J., Falick, A. M. and Mujeeb, A. (2004) The cyclobutane dimers of 5-methylcytosine and their deamination products. *Photochem. Photobiol. Sci. 3*, 968-979.

(7) Taylor, J. S. (2006) In *DNA Damage Recognition* (Siede, W., Kow, Y. W. and Doetsch, P. W., Eds.) pp 67-94, Taylor and Francis Group, New York.

(8) Pfeifer, G. P. (1997) Formation and processing of UV photoproducts: effects of DNA sequence and chromatin environment. *Photochem. Photobiol. 65*, 270-283.

(9) Celewicz, L., Mayer, M. and Shetlar, M. D. (2005) The photochemistry of thymidylyl-(3'-5')-5-methyl-2'-deoxycytidine in aqueous solution. *Photochem. Photobiol. 81*, 404-418.

(10) Douki, T. and Cadet, J. (1994) Formation of cyclobutane dimers and (6-4) photoproducts upon far-UV photolysis of 5-methylcytosine-containing dinucleotide monophosphates. *Biochemistry 33*, 11942-11950.

(11) Zhao, X., Nadji, S., Kao, J. L. and Taylor, J. S. (1996) The structure of d(TpA), the major photoproduct of thymidylyl-(3'-5')-deoxyadenosine. *Nucleic Acids Res. 24*, 1554-1560.

(12) Davies, R. J., Malone, J. F., Gan, Y., Cardin, C. J., Lee, M. P. and Neidle, S. (2007) High-resolution crystal structure of the intramolecular d(TpA) thymine-adenine photoadduct and its mechanistic implications. *Nucleic Acids Res. 35*, 1048-1053.

(13) Kumar, S., Sharma, N. D., Davies, R. J., Phillipson, D. W. and McCloskey, J. A. (1987) The isolation and characterisation of a new type of dimeric adenine photoproduct in UV-irradiated deoxyadenylates. *Nucleic Acids Res. 15*, 1199-1216.

(14) Sharma, N. D. and Davies, R. J. (1989) Extent of formation of a dimeric adenine photoproduct in polynucleotides and DNA. *J. Photochem. Photobiol. B 3*, 247-258.

(15) Kumar, S., Joshi, P. C., Sharma, N. D., Bose, S. N., Jeremy, R., Davies, H., Takeda, N. and McCloskey, J. A. (1991) Adenine photodimerization in deoxyadenylate sequences: elucidation of the mechanism through structural studies of a major d(ApA) photoproduct. *Nucleic Acids Res. 19*, 2841-2847.

(16) Wang, Y., Taylor, J. S. and Gross, M. L. (2001) Isolation and mass spectrometric characterization of dimeric adenine photoproducts in oligodeoxynucleotides. *Chem. Res. Toxicol. 14*, 738-745.

(17) Horsfall, M. J., Borden, A. and Lawrence, C. W. (1997) Mutagenic properties of the T-C cyclobutane dimer. *J. Bacteriol. 179*, 2835-2839.

(18) Brash, D. E., Rudolph, J. A., Simon, J. A., Lin, A., McKenna, G. J., Baden, H. P., Halperin, A. J. and Ponten, J. (1991) A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. *Proc. Natl. Acad. Sci. USA 88*, 10124-10128.

(19) You, Y. H., Li, C. and Pfeifer, G. P. (1999) Involvement of 5-methylcytosine in sunlight-induced mutagenesis. *J. Mol. Biol. 293*, 493-503.

(20) Mitchell, D. L. (2000) Effects of cytosine methylation on pyrimidine dimer formation in DNA. *Photochem. Photobiol. 71*, 162-165.

(21) Tommasi, S., Denissenko, M. F. and Pfeifer, G. P. (1997) Sunlight induces pyrimidine dimers preferentially at 5-methylcytosine bases. *Cancer Res. 57*, 4727-4730.

(22) Lee, D. H. and Pfeifer, G. P. (2003) Deamination of 5-methylcytosines within cyclobutane pyrimidine dimers is an important component of UVB mutagenesis. *J. Biol. Chem. 278*, 10314-10321.

(23) Cannistraro, V. J. and Taylor, J. S. (2009) Acceleration of 5-methylcytosine deamination in cyclobutane dimers by G and its implications for UV-induced C-to-T mutation hotspots. *J. Mol. Biol. 392*, 1145-1157.

(24) Bose, S. N., Davies, R. J., Sethi, S. K. and McCloskey, J. A. (1983) Formation of an adenine-thymine photoadduct in the deoxydinucleoside monophosphate d(TpA) and in DNA. *Science 220*, 723-725.

(25) Bose, S. N., Kumar, S., Davies, R. J., Sethi, S. K. and McCloskey, J. A. (1984) The photochemistry of d(T-A) in aqueous solution and in ice. *Nucleic Acids Res. 12*, 7929-7947.

(26) Su, D. G., Kao, J. L., Gross, M. L. and Taylor, J. S. (2008) Structure determination of an interstrand-type cis-anti cyclobutane thymine dimer produced in high yield by UVB light in an oligodeoxynucleotide at acidic pH. *J. Am. Chem. Soc. 130*, 11328-11337.

(27) Wang, Y., Taylor, J. S. and Gross, M. L. (1999) Nuclease P1 digestion combined with tandem mass spectrometry for the structure determination of DNA photoproducts. *Chem. Res. Toxicol. 12*, 1077-1082.

(28) Vu, B., Cannistraro, V. J., Sun, L. and Taylor, J. S. (2006) DNA synthesis past a 5-methylC-containing cis-syn-cyclobutane pyrimidine dimer by yeast pol eta is highly nonmutagenic. *Biochemistry 45*, 9327-9335.

(29) Lemaire, D. G. and Ruzsicska, B. P. (1993) Kinetic analysis of the deamination reactions of cyclobutane dimers of thymidylyl-3',5'-2'-deoxycytidine and 2'-deoxycytidylyl-3',5'-thymidine. *Biochemistry 32*, 2525-2533.

(30) Mosbaugh, D. W. and Bennett, S. E. (1994) Uracil-excision DNA repair. *Prog. Nucleic Acid Res. Mol. Biol. 48*, 315-370.

(31) Zhang, L. K., Rempel, D. and Gross, M. L. (2001) Matrix-assisted laser desorption/ionization mass spectrometry for locating abasic sites and determining the rates of enzymatic hydrolysis of model oligodeoxynucleotides. *Anal. Chem. 73*, 3263-3273.

(32) Zhang, L. K., Ren, Y., Rempel, D., Taylor, J. S. and Gross, M. L. (2001) Determination of photomodified oligodeoxynucleotides by exonuclease digestion,

matrix-assisted laser desorption/ionization and post-source decay mass spectrometry. *J. Am. Soc. Mass Spectrom 12*, 1127-1135.

(33) Zhao, X. and Taylor, J. S. (1996) Mutation spectra of TA*, the major photoproduct of thymidylyl-(3'5')-deoxyadenosine, in Escherichia coli under SOS conditions. *Nucleic Acids Res. 24*, 1561-1565.

(34) Antequera, F. and Bird, A. (1993) CpG islands. *EXS 64*, 169-185.

(35) Bird, A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev. 16*, 6-21.

(36) Illingworth, R. S. and Bird, A. P. (2009) CpG islands--'a rough guide'. *FEBS Lett. 583*, 1713-1720.

(37) Williamson, J. R., Raghuraman, M. K. and Cech, T. R. (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell 59*, 871-880.

(38) Douki, T., Laporte, G. and Cadet, J. (2003) Inter-strand photoproducts are produced in high yield within A-DNA exposed to UVC radiation. *Nucleic Acids Res. 31*, 3134-3142.

(39) Su, D. G., Fang, H., Gross, M. L. and Taylor, J. S. (2009) Photocrosslinking of human telomeric G-quadruplex loops by anti cyclobutane thymine dimer formation. *Proc. Natl. Acad. Sci. USA 106*, 12861-12866.

(40) Lorincz, M. C. and Groudine, M. (2001) C(m)C(a/t)GG methylation: a new epigenetic mark in mammalian DNA? *Proc. Natl. Acad. Sci. USA 98*, 10034-10036.

(41) Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., Tonti-Filippini, J., Nery, J. R., Lee, L., Ye, Z., Ngo, Q. M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A. H., Thomson, J. A., Ren, B. and Ecker, J. R.

(2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature 462*, 315-322.

(42) Costello, J. F., Krzywinski, M. and Marra, M. A. (2009) A first look at entire human methylomes. *Nat. Biotechnol. 27*, 1130-1132.

(43) Finnegan, E. J., Genger, R. K., Peacock, W. J. and Dennis, E. S. (1998) DNA Methylation in Plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol. 49*, 223-247.

(44) Tariq, M. and Paszkowski, J. (2004) DNA and histone methylation in plants. *Trends Genet. 20*, 244-251.

(45) Vanyushin, B. F. (2006) DNA methylation in plants. *Curr. Top. Microbiol. Immunol. 301*, 67-122.

Table 4.1. Identification of UVB induced photoproducts of d(GTAT^{m}CATGAGGTGC) at pH 7.6.

Photoproduct (PP)	PP1	PP1*	PP2	PP2*	PP3	PP4
Photodimer Assignment	^{m}CA*	TA*	T[*c,s*]^{m}C	T[*c,s*]T	T[*t,s*]^{m}C	T[Dewar]^{m}C

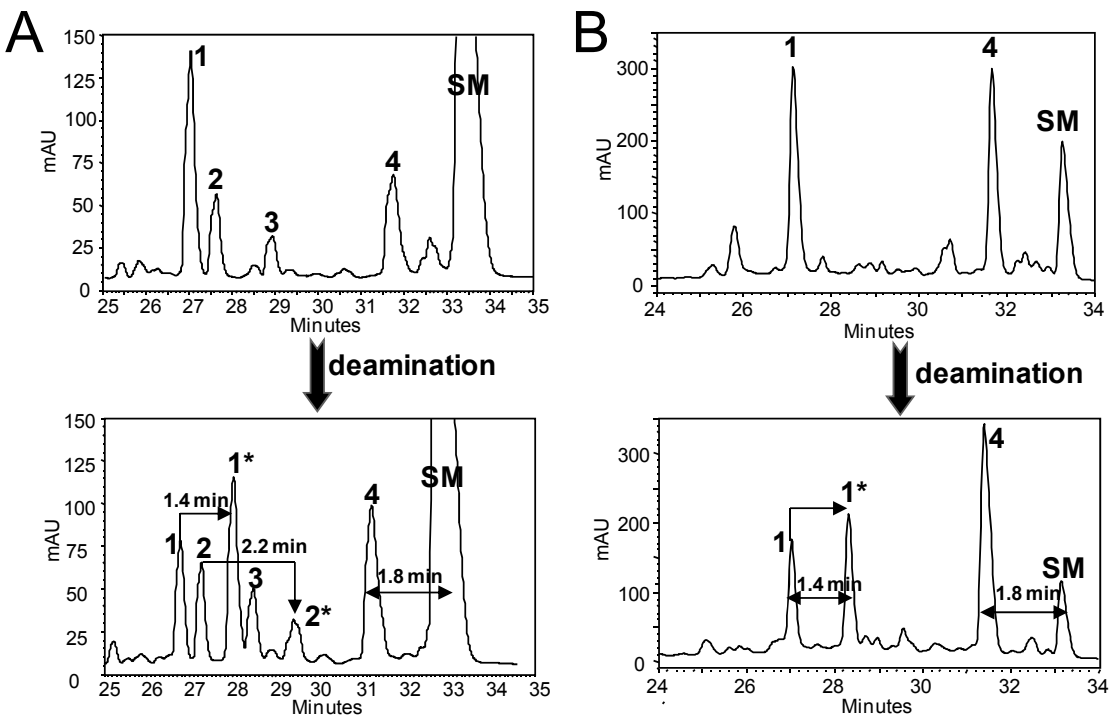


Figure 4.1. RP-HPLC of UVB-induced photoproducts of d(GTAT^{m}CATGAGGTGC) at pH 7.6 (upper panel) and of corresponding deamination products (lower panel) with A) 302 nm light and B) 312 nm light. SM: starting material; The designation * is used in this context for a photoproduct that has undergone deamination.

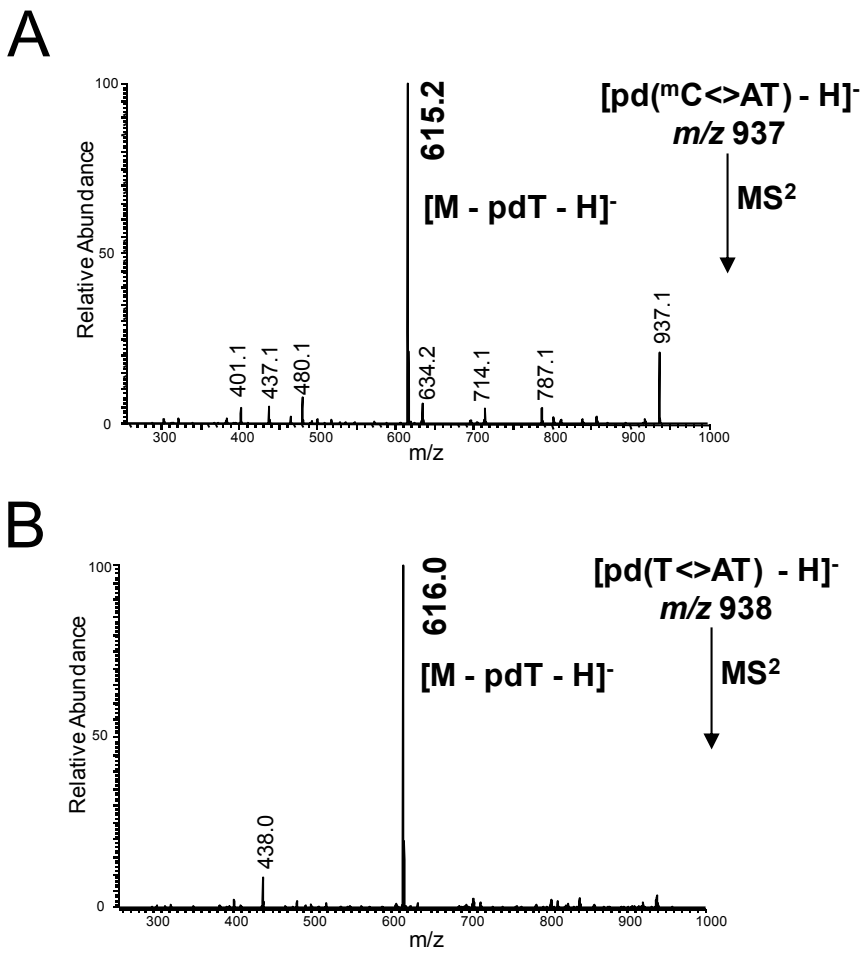


Figure 4.2. Product-ion spectra of the trinucleotides of NP1 digested photoproducts PP1 (A) and PP1* (B), indicating a ^{m}C<>A and T<>A photoproduct, respectively.

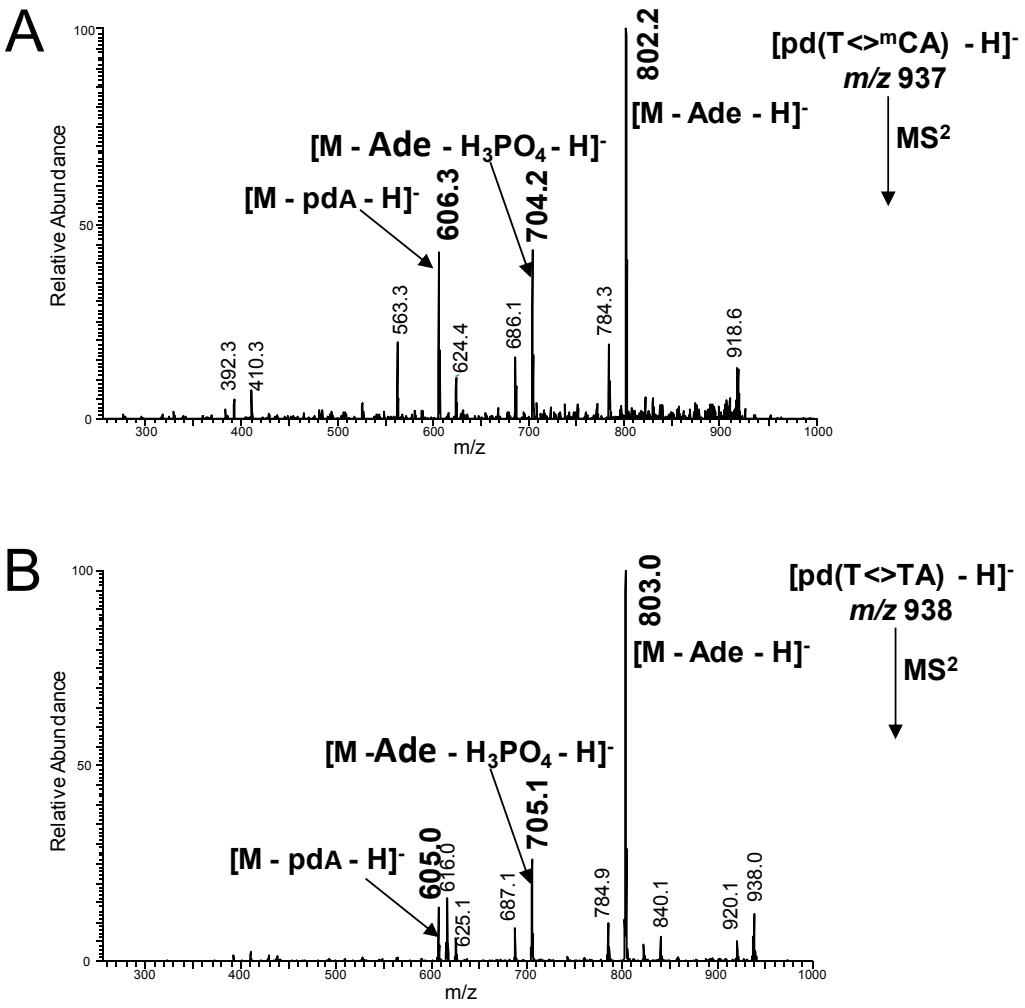


Figure 4.3. Product-ion spectra of the trinuclteotides of NP1 digested photoproducts PP2 (A) and PP2* (B), indicating a T<>^{m}C and T<>T photoproduct, respectively.

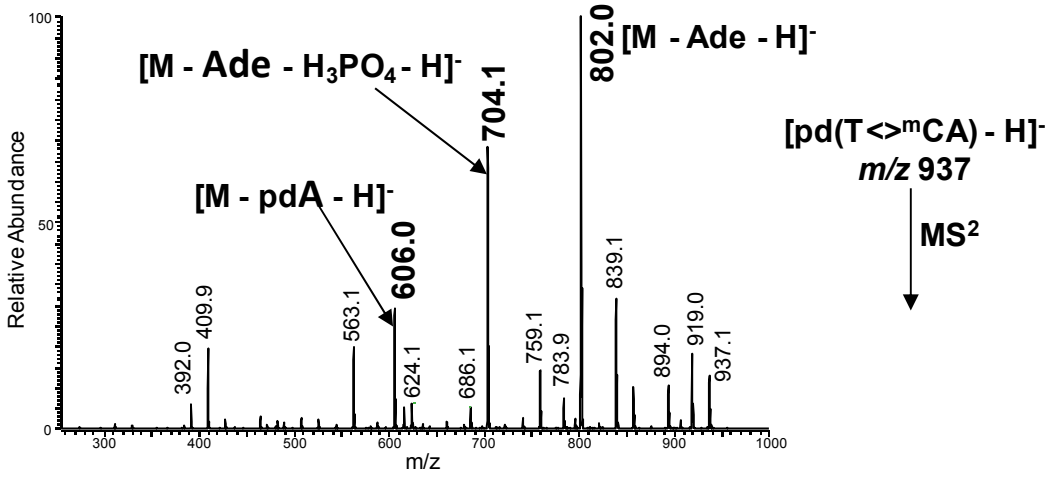


Figure 4.4. Product-ion spectrum of the trinucleotide of NP1 digested photoproduct PP3, indicating a T<>^{m}C photoproduct.

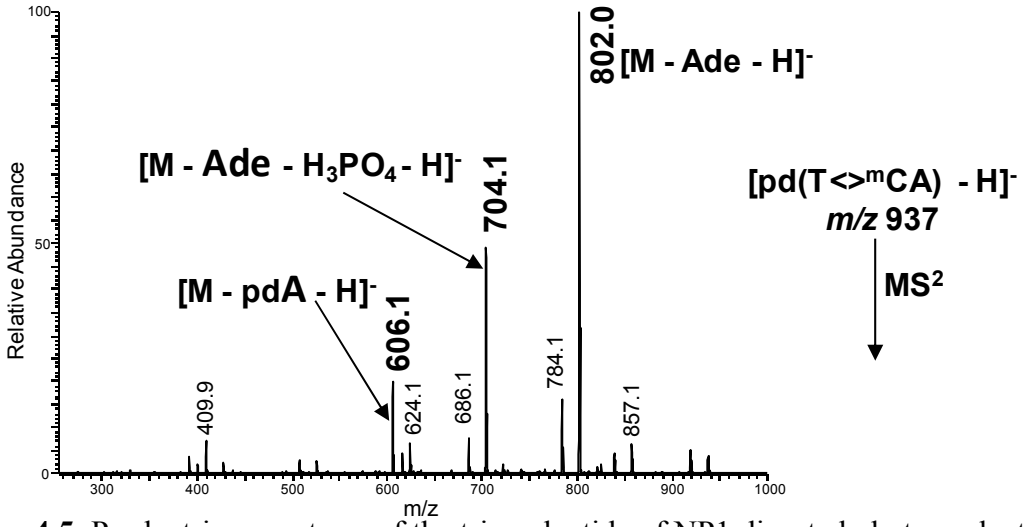


Figure 4.5. Product-ion spectrum of the trimucleotide of NP1 digested photoproduct PP4, indicating a T<>^{m}C photoproduct.

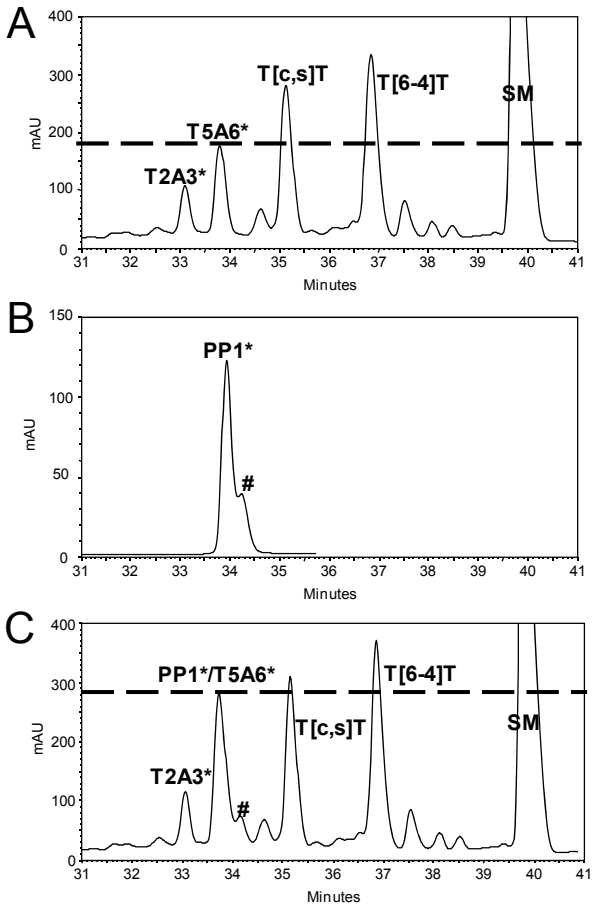


Figure 4.6. RP-HPLC assay of A) ~ 5 nmol UVC irradiation mixture of d(GTATTATGAGGTGC), B) ~ 360 pmol of deamination product of HPLC fraction of d(GTAT^{m}CATGAGGTGC), and C) a coinjected sample of A and B. # is the contamination peak of PP3 from HPLC separation (Figure 4.1A, lower panel)

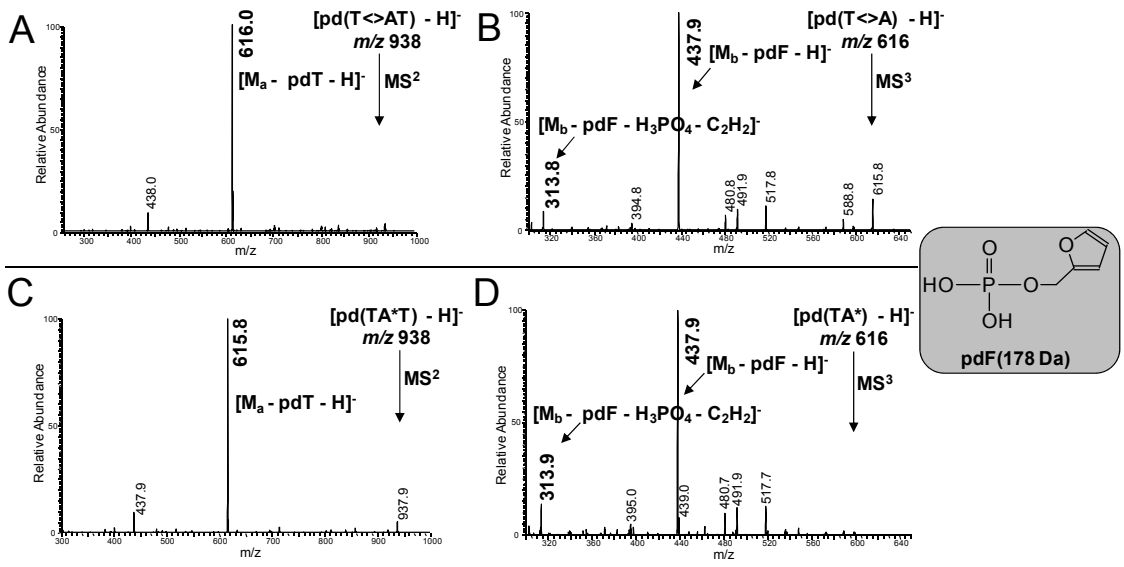
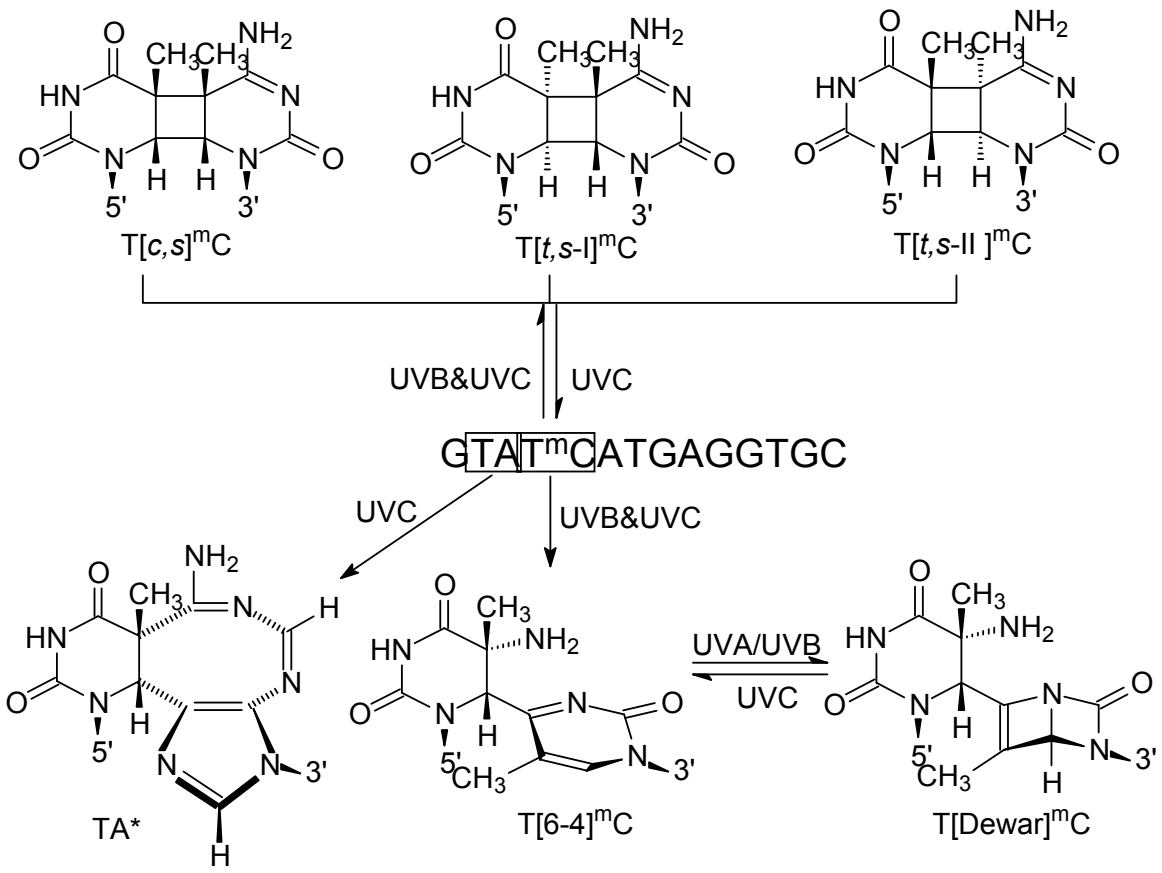
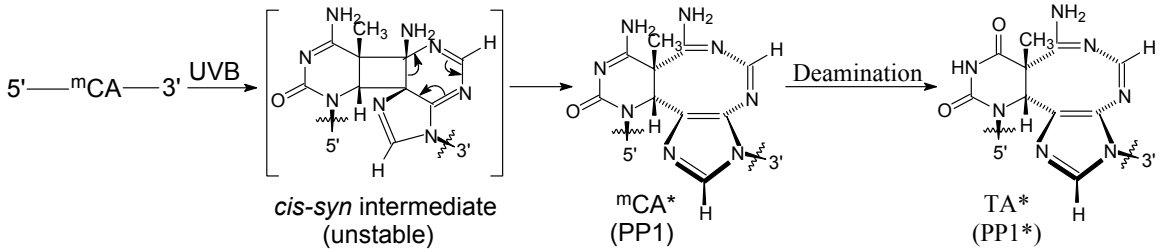


Figure 4.7. NP1 coupled ESI- MS^2 and MS^3 assays of $^{m}C<>A$ photoproduct of d(GTAT^{m}CATGAGGTGC) following deamination to T<>A (A and B, respectively), and TA* photoproduct of T5A6 of d(GTATTATGAGGTGC) (C and D, respectively).



Scheme 4.1. Known photoproducts that could originate from the designed sequence.



Scheme 4.2. Proposed structure of the ^{m}C<>A photoproduct and its mechanism of formation and deamination.

Chapter 5

A Mass Spectrometric Investigation of the Interaction of hApe1 with the Redox Activity Inhibitor, E3330

Abstract

Human apurinic/apyrimidinic endonuclease 1 (hApe1) is a multi-functional protein having two major functions: apurinic/apyrimidinic endonuclease activity for DNA damage repair and redox activity for gene regulation. It was reported that Ape1 reduces a number of important transcription factors that are involved in cancer promotion and progression. The underlying redox mechanism, however, is not yet clear. The crystal structure of hApe1 indicates there are seven cysteines (C65, C93, C99, C138, C208, C296, and C310), among which C99 and C138 are solvent-accessible and C65, C93, and C99 are located in the redox domain. It was hypothesized that one or more cysteines are involved in the redox fuction. A further hypotheis is that hApe1 reduces the down-stream transcription factors via a thiol/disulfide exchange mechanism. E3330, (2E)-3-[5-(2,3-dimethoxy-o-methyl-1,4-benzoquinoyl)]-2-nonyl-2-propenoic acid, is a specific inhibitor for the the redox function of hApe1. The interaction mechanism is not known although a binding constant of 1.6 nM was determined by SPR, indicating a specific and strong interaction between hApe1 and E3330. In our study, hydrogen/deuterium exchange (HDX) experiments show no difference between ApeΔ40 (a truncated form of hApe1 with residues 40-318) without (*apo*) and with E3330 (*holo*) forms, suggesting there is no siginificant interaction between ApeΔ40 and E3330. ESI and nESI-MS analysis indicates that no strongly bound complex formed between ApeΔ40 and E33330 can be introduced to the gas phase. Taken together, there is no evidence for a strong interaction between ApeΔ40 and E3330. In the presence of *N*-ethylmaleimide (NEM), a "+ 7 NEM" adduct was seen for ApeΔ40 with E3330, suggesting that a locally unfolded form coexists with the folded form in an equilibrium that is driven by irreversible NEM labeling. E3330

interacts with and stabilizes this locally unfolded form. Trypsin digestion coupled LC-MS/MS analysis shows there are up to 8% of disulfide bonds induced by E3330 among the cysteines in the redox domain, which are far apart from each other in the folded form, as indicated by the X-ray crystal structures. We also propose that the locally unfolded form interacts with E3330 and allows disulfide bonds to form in the redox domain. The locally unfolded form is thereby proposed to be the potential redox-active form.

Introduction

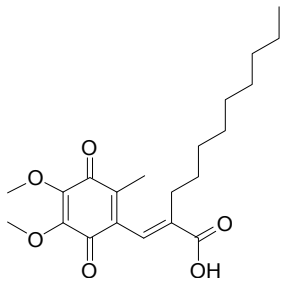
Human apurinic/apyrimidinic endonuclease 1 (hApe1) is a multi-functional nuclear protein (*1-7*). It has two major functions: DNA damage repair via its apurinic/apyrimidinic endonuclease activity and activation of down-stream transcription factors to bind DNA via its redox activity. Ape1 was first identified and characterized as a nuclear factor, termed as a redox effector factor-1 (Ref-1) (*8, 9*), and later identified as Ape1 (*10*). Ape1 reduces a number of important transcription factors including AP-1, P53, NF-κB, HIF-1α, and TTF, thus stimulating their DNA binding activity for gene regulation (*10-20*). Ape1 also participates in the base excision repair (BER) pathway by cleaving the 5' phosphodiester bond of an abasic site, which is generated by DNA glycosylases that excise a base damaged by alkylation or oxidation, and leaving a 3'-OH and 5'-deoxyribose phosphate (*21-23*). The two functional domains of the protein are separable: DNA repair activity is located in the C-terminal, and redox activity in the N-terminal regions (*24-28*). Although the mechanism whereby Ape1 repairs DNA is relatively clear, the mechanism of its redox activity is poorly understood.

Based on the crystal structure (Figure 5.1), C99 and C138, among the seven constituent cysteines (C65, C93, C99, C138, C208, C296, and C310), are solvent-accessible (*28-31*). There are three cysteines (C65, C93, C99) located in the redox domain and buried in the core of the protein; C65 and C93 are positioned on opposite sides of the beta sheet. There are no disulfide bonds according to the X-ray crystal structure. The sulfur atoms of the two closest cysteine residues, C93 and C208, are approximately 3.5 Å apart, which is longer than the length of a typical disulfide bond (2.2 Å). It has been hypothesized that one or more cysteines in the redox domain are involved

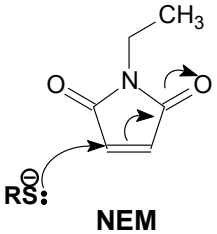
in the redox function of hApe1. Most previous reports suggest that C65, a residue unique to mammalian sequences, is essential for the redox activity (*24, 31*) with one exception that claims Cys 65 is not required (*32*). In a study employing both site-directed mutagenesis and *in vitro* electrophoretic-mobility-shift-assay (EMSA) analysis, each of the seven cysteines in hApe1 was substituted one at a time with Ala. Only hApe1 C65A was redox inactive when reduced, and only hApe1 C93A remains redox active when oxidized (*24*), suggesting that the redox mechanism involves a disulfide bond between C65 and C93. Although the crystal structure shows that C65 is inaccessible to solvent, it is possible that Ape1 can change its conformation to expose C65 so it can react with the target cysteines upon interacting with the down-stream transcription factors. These factors, however, do not share a common structural motif for binding hApe1 (*31*).

hApe1 levels are elevated in some cancers (e.g., ovarian, cervical, and germ cell tumors) and typically associated with aggressive proliferation and enhanced resistance to therapeutic agents (*33-36*). Recently, a novel approach to cancer therapeutic agents targeted the redox activity of hApe1 (*7, 37*). A quinone derivative E3330, (2*E*)-3-[5-(2,3-dimethoxy-*o*-methyl-1,4-benzoquinoyl)]-2-nonyl-2-propenoic acid, was found to inhibit specifically the redox ability of hApe1 (*38-41*). E3330 is able to kill a variety of cancer (i.e., ovarian, colon, lung, breast, brain, pancreatic, prostate, multiple myeloma) cells but does not significantly affect normal cells (*42*). An earlier study with surface plasmon resonance (SPR) reported that the binding constant for hApe1 and E3330 is 1.6 nM, which indicates strong and specific binding between hApe1 and E3330 (*38*). No further studies have been carried on the interaction between hApe1 and E3330 to probe, for example, the location of the binding site and the conformational changes, if any, associated with hApe1 binding to E3330.

We report herein an investigation of the interaction of ApeΔ40, a truncated form of hApe1 (residues 40-318) (Figure 5.2), and E3330 by using amide hydrogen/deuterium exchange (HDX) and *N*-ethylmaleimide (NEM) (see NEM structure) labeling coupled with MS analysis. At the time of this study, ApeΔ40 was more available, and we were able to conduct a complete set of experiments on this protein. Nevertheless, ApeΔ40 is a suitable model because truncations of up to 40 residues on the N-terminus do not affect the redox or DNA repair activity of hApe1 (*24, 25*). HDX and mass spectrometry (MS) combine as a powerful approach for probing protein structure, dynamics, and interactions with ligands. It examines changes in H-bonding at all amide sites, except that of proline (*43-52*). Furthermore, HDX coupled with a titration strategy should verify the high binding affinity of hApe1 and E3330 by using PLIMSTEX (protein-ligand interactions by MS titration and HD exchange) (*53-56*). A complementary approach for determining protein conformational change is amino acid-specific covalent labeling (*57-59*). NEM is a water-soluble, small molecule that specifically reacts with solvent-accessible cysteines via a Michael addition. This reagent is widely used in protein footprinting (*60-65*). A combination of both footprinting strategies should allow a more complete picture of this protein/ligand binding to be elucidated.



E3330

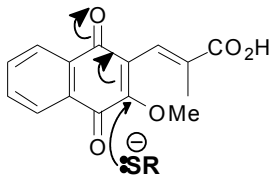


NEM

Material and Methods

Materials

ApeΔ40 was expressed and purified in the Georgiadis lab as pervious described (*66*). The stock solution contained 2.5 mM protein in 10 mM HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) (pH = 7.5) buffer. E3330 and its analog compound 8-51, (*E*)-3-(3-methoxy-1,4-naphthoquinon-2-yl)-2-methylpropenoic acid, were synthesized by Rod Nyland in the Rick Borch Laboratory (Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN 479074) as previously described (*67, 68*). The stock solution contained 100 mM E3330 in DMSO. Both ApeΔ40 and E3330 were stored in a -80 °C freezer until used.



8-51

HDX and Electrospray Ionization Mass Spectrometry **(***ESI-MS) Experiments of ApeΔ40 without and with E3330*

HDX kinetics experiments were conducted with ApeΔ40 without (*apo*) and with E3330 (*holo*). Both apo and holo solutions were prepared by incubating the protein, with or without E3330, in 10 mM HEPES (pH = 7.5) and 150 mM KCl at 25 °C. For the *apo* form, ApeΔ40 was incubated for 1 h. For the holo from, ApeΔ40 and E3330 were incubated at a molar ratio of 1:10 for 1 h. HDX experiments were carried out at 25 °C in a 90% D_2O buffer with 10 mM HEPES (pH = 7.5) and 150 mM KCl. At certain times (0, 0.5, 1, 2, 4, 8, 15, 30, 60, 120, 180, and 240 min), a 10 μL of HDX solution was quenched with cold 1 M HCl to decrease the pH to 2.5. The quenched solution was loaded onto an Opti-Guard C18 column (10 mm x 1 mm i.d., P.J. Cobert Associates, St. Louis, MO) by using a six-port Rheodyne 7725 manual sample injector (P.J. Cobert Associates, St. Louis, MO). The C18 column was pre-equilibrated with 0.1% formic acid (FA) aqueous solution. To minimize back exchange of the D label at the backbone amide

positions, the injector, the Opti-Guard® 1 mm C18 column, and all transfer silica tubing were placed in an ice-bath. The protein was back-exchanged and desalted with ice-cold 0.1% FA solution and eluted with 50% ACN with 0.1% FA at a flow rate of 30 μL/min.

All samples were analyzed by ESI-MS on a Micromass Q-TOF UltimaTM GLOBAL equipped with a Waters CapLC system and a Z-spray ESI source operated in the positive-ion mode. The capillary voltage was 3.0 kV, collision energy was 10 eV, and the source and desolvation temperatures were 80 °C and 150 °C, respectively. All mass spectra were deconvoluted by MaxEnt1 algorithm with MassLynx 4.0, which are part of the software of this system.

All modeling and curve fitting were done by using Mathcad 2001 Professional (Mathsoft, Inc., Cambridge, MA) and a custom-built program (*69*).

NanoESI (nESI) MS

ApeΔ40 was incubated in 1 M ammonium acetate (pH 7.5) with or without E3330 at room temperature (RT) for 4 h. Protein samples were analyzed by nESI-MS in the positive-ion mode on a Bruker MaXis UHR-TOF (ultra-high resolution time-of-flight) at a flow rate of 25 nL/min. The capillary voltage was set at -1000-1200 V. Dry gas and temperature were at 5.0 L/min and 50 °C, respectively. The instrument was externally calibrated by using "Tuning Mix" (Agilent Technologies). The spray tips were made in-house by pulling a 150 μm i.d. × 365 μm o.d. fused silica capillary with a P-2000 laser puller (Sutter Instrument Co., Novato, CA). A four-step program was used with the parameter setup as follows with all other values set to zero: Heat = 290, velocity = 40, delay = 200; Heat = 280, velocity = 30, delay = 200; Heat = 270, velocity = 25, delay = 200; Heat = 260, velocity = 20, delay = 200. Tips were cut accordingly to allow a good

spray under the experimental conditions. For each sample, a new tip was used to avoid cross contamination.

Covalent Labeling and ESI-MS

For compound 8-51 labeling, ApeΔ40 and its mutants (C99A, C138A, C99/C138A), where either or both C99 and C138 were replaced by alanines, were incubated with compound 8-51 in 10 mM HEPES (pH = 7.5) at RT ([protein] = 100 μM; [8-51] = 500 μM). After 4 h, the reactions were quenched by adding 10 μL of 20 mM dithiothreitol (DTT) to 10 μL of the above solution. The DTT-quenched samples were diluted with water before MS analysis. Typically, a solution containing 5 pmol protein was loaded to a C18 column, desalted with water, and then eluted to the Bruker MaXis UHR-TOF instrument with 50% ACN with 0.1% FA at 10 μL/min. The capillary voltage was -3600 V. The nebulizer pressure was 0.4 bar, and the drying gas flow was at 1.0 L/min at a temperature of 180 °C. Deconvolution was done with the MaxEnt1 algorithm supplied by the manufacturer of the mass spectrometer system.

For NEM labeling, 10-20 μL of ApeΔ40/NEM (ApeΔ40:NEM = 1:5, mol/mol) and ApeΔ40/NEM/E3330 (ApeΔ40:NEM:E3330 = 1:5:5, mol/mol/mol) ApeΔ40 samples were incubated in 10 mM HEPES buffer (pH 7.5) at RT; the protein concentration was 100 μM. At a certain time, a 1 μL aliquot was removed and quenched with 1 μL of 20 mM DTT. Mass spectra were collected on the Bruker MaXis UHR-TOF or Waters Micromass Q-TOF instrument as described above.

Data Processing of NEM-Labeling

A number of the equation parameters were extracted from the kinetic data by nonlinear least squares fitting of theoretical signals, computed from the parameter dependent system state trajectories, to experiment data. The system state was a vector

that has the solution chemical species concentrations as the vector coordinates. In each trial of the search, the postulated parameters together with the system state of concentrations permitted the calculation of the time rate of change of the state by computing the fluxes into and out of each species as described by the system equations. This process implemented a vector first-order ordinary differential equation, which was solved by numerical integration for the time interval of reaction initiation to the longest reaction time to give the state time trajectory in each fitting trial. For comparison with the experiment data, the theoretical signal for each Ape species was computed as a fraction of all Ape species concentrations that were first weighted by a relative sensitivity factor that varied linearly with slope g_N with the number of NEMs attached starting with one for ApeΔ40 by itself. The calculations were carried out in the computer application Mathcad 14.0 M010 (Parametric Technology Corporation, Needham, MA 02494 USA). The numerical integration of the differential equation was carried out by the adaptive fourth-order Runge-Kutta function "Rkadapt".

HDX and Electrospray Ionization Mass Spectrometry (ESI-MS) Experiments of the NEM adducts of ApeΔ40

A 30 μL solution of ApeΔ40/E3330/NEM was incubated with 10 mM HEPES (pH 7.5) for 22 h at RT ([protein] = 100 μM; [E3330] = [NEM] = 500 μM). To quench the NEM labeling reaction, 0.5 μL of 1 M DTT was added to the above solution. An aliquot of the 2.8 μL DTT quenched solution was diluted to a final 40 μL of 93% D_2O medium with 10 mM HEPES (pH 7.5) and 150 mM KCl and incubated for various times to obtain the HDX kinetics at 25 °C. The HDX reaction was quenched by adding 1 μL of 1 M cold HCl. ESI-MS and data processing was the same as described above.

LC-MS/MS Assay of Disulfide Bond Formation

A 200 μL solution of ApeΔ40 (10 μM) and E3330 (50 μM) was incubated in 10 mM HEPES (pH = 7.5) at 37 **°C** for 1 h. NEM was then added to block any unreacted Cys and prevent disulfide-bond scrambling in the course of digestion. The sample was diluted with water to a final concentration of 1 μM. A portion of the final diluted solution (50 μL) was submitted to trypsin digestion (protein/trypsin = 50/1, w/w) at 37 °C for 4 h. The solution was then analyzed by LC-MS/MS whereby 5 μL of digestion solution was consumed for each experiment. Reversed-phase capillary LC separations were performed with an Eksigent NanoLC-1D pump (Eksigent Technologies, Inc. Livermore, CA). The reversed-phase capillary column (0.075 mm × 150 mm) was packed in-house by using a PicoFrit^{TM} tip (New Objective, Inc., Woburn, MA) with C18 particles (Magic, 5 μm, 120 Å, Michrom Bioresources, Inc., Auburn, CA). The mobile phase consisted of water, with 0.1% formic acid (A) and acetonitrile with 0.1% formic acid (B). Immediately after sample loading, the mobile phase was held at 98% A for 12 min. A linear gradient was performed by using 2% to 60% solvent B over 60 min, then to 80% solvent B over 10 min at 260 nL/min followed by a 12 min re-equilibration step by 100% solvent A. The flow was directed by PicoView Nanospray Source (PV550, New Objective, Inc., Woburn, MA) to the LTQ Orbitrap (Thermo Fisher Scientific, Inc., San Jose, CA). The spray voltage was 1.8-2.2 kV, and the capillary voltage was 27 V. The LTQ Orbitrap was operated in standard data-dependent MS/MS acquisition mode controlled by its Xcalibur 2.0.7 software, in which a full mass spectral scan was followed by 6 product-ion (MS/MS) scans. The mass spectra of the peptides were acquired at high mass resolving power (60,000 for ions of *m/z* 400) with the FT analyzer over the range of *m/z* 350-2000. The six most abundant precursor ions were dynamically selected in the order of highest to lowest intensity (minimal intensity of 1000 counts) and subjected to

collision-induced dissociation (CID). Precursor activation was performed with an isolation width of 2 Da and activation time 30 ms. The normalized collision energy was 35% of the maximum available. The automatic gain control target value was regulated at 1×10^6 for FT and 3×10^4 for the ion trap with a maximum injection time of 1000 ms for FT and 200 ms for the ion trap. The instrument was externally calibrated by using a standard calibration mixture of caffeine, the peptide MRFA, and Ultramark 1621 (Thermo Fisher Scientific, Inc., San Jose, CA). To identify covalent modifications LC-MS/MS data were searched with Mascot 2.2 (Matrix Science, London, UK) against the NCBI database or MassMatix, an in-house search engine developed by Xu et al. (*70-72*). Parameters used in Mascot were: enzyme, trypsin; maximum missed cleavage, 3; peptide mass tolerance, 10 ppm with one C 13 peak; peptide charge, +1 to +3; product mass tolerance, 0.6 Da; instrument type, default (searching for all types of b and y ions). To locate disulfide bonds, LC-MS/MS data were searched with MassMatrix with the parameter settings as follows: enzyme, trypsin; maximum missed cleavage, 3; variable number of modifications by NEM on cysteines; precursor ion tolerance, 10 ppm; product ion tolerance, 0.8 Da; max # PTM per peptide, 2; minimum peptide length, 4 amino acids; maximum peptide length, 4 amino acids; min pp score, 5.0; min pptag score 1.3; max # match per peptide, 3; max # combination per match, 3; fragmentation method, CID; C13 isotope ions, 1; crosslink, disulfide; crosslink mode, exploratory; crosslink sites cleavability, not applicatble; max # crosslinks per peptide, 2.

Results

The Kinetics of HDX for ApeΔ40 in the Presence and Absence of E3330 Are Nearly Identical

On the basis of a previous SPR study, we presume there is a strong binding interaction between hApe1 and E3330 (binding constant of 1.6 nM) (*38*). We, therefore, conducted HDX kinetics experiments with the ApeΔ40 concentration at 2.5 μM, as a preliminary experiment prior to using PLIMSTEX (*53-56*) to verify the high affinity. For PLIMSTEX to be applicable, there must be a difference in the HDX of the protein with and without the ligand. We did not observe, however, any detectable difference in the HDX kinetics for ApeΔ40 with and without the ligand over a range of protein concentrations from 0.5 to 20 μM (Figure 5.3). With our custom kinetics model executed in Mathcad, we assumed there are three types of exchanging amide protons: fast, intermediate, and slow, and we used the kinetic model to determine the rate constants and the number of amides in each category. The outcome was:

- k_1 (fast exchange) = 5 min^{-1}
- k_2 (intermediate exchange) = 0.16 min^{-1}
- k_3 (slow exchange) = 0.008 min^{-1}.

Based on the model, we were able to assign the number of amide protons of each type. In 90% D_2O, the numbers of fast, intermediate and slow exchanging amide protons are 78 ± 2, 25 ± 4, and 43 ± 1 for ApeΔ40 in the absence of E3330 and 81 ± 3, 21 ± 2, and 42 ± 1 for the protein in the presence of E3330. There is clearly no significant difference between the extents of HDX for ApeΔ40 with and without E3330.

The results can have three explanations: first, the ligand doesn't bind to the protein under the conditions of the experiment. Second, there is binding of the ligand, but it produces no significant change in the protein's conformation (in particular, in the amide backbone). Third, upon ligand binding, some regions of the protein become protected while other regions become compensatorily deprotected, which fortuitously

cancels any global effect. Although the second and third explanations seem improbable, we sought other MS data to confirm or deny the first explanation.

ESI and nESI-MS Show No Strong Evidence of ApeΔ40/E3330 Complex

To test whether there is binding between ApeΔ40 and E3330 in the micromolar and lower range of protein concentrations, we conducted ESI-MS with 50% ACN with 0.3% FA and nESI-MS with aqueous solution to monitor directly the complex formation. Although this experiment is subject to giving false-negative results, we felt that if the complex had such high affinity, it could be successfully introduced into the gas phase by ESI.

We chose to begin this experiment with a "control" ligand quinone, compound 8-51 (see 8-51 structure) that is likely to react with solvent-accessible Cys residues. Indeed, we observed evidence for an interaction between ApeΔ40 and the 8-51 quinone, an analog of E3330, under regular ESI condition with water/acetonitrile/FA as the spray solvent. A mass increase in the "decharged" protein signal of 274 Da indicates the addition of one 8-51 molecule, associated with reduction to its hydroquinone form (Figure 5.4).

To locate the sites of modification, we submitted the complex to trypsin digestion and analyzed the peptide mixture by MS/MS. We were unable to detect any 8-51-modified peptides, however. This suggests that those 8-51 molecules that added to the protein giving detectable complexes in the mass spectrum of the intact protein underwent an elimination reaction, restoring the modified cysteines to free thiols. Thus, the interaction was reversible, which would also be expected of the interaction with E3330.

Although we could not locate the modification sites by digestion and LC/MS/MS, we can reason to their locations by a study of three ApeΔ40 mutants. ESI-MS of the

mutants showed that the proteins added one 8-51 for C99A and C138A and no 8-51 for C99/138A. These results are strong evidence that 8-51 specifically modifies C99 and C138. Together with the evidence that the addition of 8-51 results in a mass increase of 274 Da suggested that 8-51 reacts with C99 and C138 via a Michael addition followed by a reduction to its hydroquinone form. Unlike its analog 8-51, E3330 was only found to form adducts with ApeΔ40 under certain ESI conditions (Figure 5.5). Considering the distribution of adducts with different stoichiometries, we interpret the results as indicators of non-specific interactions.

A typical ESI solvent consisting of water/acetonitrile/FA typically denatures proteins and/or their complexes, so the inability to detect a complex sprayed from such a solvent may be misleading about the formation of a complex in aqueous solution. An alternative is nESI-MS with aqueous solution at near neutral pH; this approach is becoming common for the detection of intact complexes in solution (*73-75*). Use of nanoflow consumes less amount of sample and has gentler ESI conditions, including low capillary voltage and no sheath gas, which are beneficial to maintain solution-state complexes. Meanwhile, the high concentration of ammonium acetate is helpful for increasing the signal-to-noise and resolving the charge states. Therefore, we used 1 M ammonium acetate at pH 7.5 as ESI solvent in accord with the successful approach reported by Robinson and her coworkers (*73*). Under these conditions, a small amount of protein/E3330 adducts in the spectra can be admitted to the gas phase (Figure 5.6). Sometimes, more than one adduct is seen as the ESI experiment (Figure 5.4), but the results are not reproducible. We interpret the results as indicators of non-specific interactions as well. The peaks representing the charge-states distribution of the protein and of the protein/ligand complexes were shifted to higher *m/z*. This shift could be

caused by a small change in protein conformation, allowing fewer charging events or by insulation of the protein by the surrounding ligands, leading to fewer charges on the protein or on the resultant protein/ligand complexes.

NEM Mapping Suggests a Locally Unfolded Conformer in Equilibrium with the Folded Form

NEM specifically reacts with solvent-accessible cysteines in a protein via a Michael addition. In the presence of NEM, ApeΔ40 only adds two NEMs as is indicated by a mass shift of 250 Da (Figure 5.7 left panel). The reaction was complete, consistent with the observation that all the protein was converted to its "+ 2 NEM" adduct after incubation with NEM ([protein] = 100 μM; [NEM] = 500 μM) in approximately 10 min at RT. Further LC-MS/MS analysis of trypsin-digested ApeΔ40/NEM complex confirmed that NEM covalently modified the two cysteines (C99, C138) that X-ray crystallography identifies as solvent-accessible (Figure 5.8) (*28-31*).

Interestingly, a new adduct appears when ApeΔ40 interacts with E3330 in the presence of NEM. That adduct is produced in a reaction whereby all seven cysteines became labeled with NEM. Although the reaction is slow, the yield of the adduct increases slowly with time such that after 6 h, the product yield is approximately 25% (Figure 5.7 right panel). Trypsin digestion coupled with LC-MS/MS analysis conclusively indicates that protein is modified on the five remaining cysteines, which were originally solvent-inaccessible, and not on other reactive amino-acid residues.

We asked whether the "+ 7 NEM" adduct was the result of E3330 denaturing the protein. When we incubated ApeΔ40 (100 μM) and E3330 (500 μM) for 24 h at RT and then added NEM to give a protein:NEM = 1:5 (mol/mol) and allowed any reaction to occur for 0.5 h, we found addition of only two NEMs, not seven. If E3330 were acting as

a denaturant, we would expect that all seven Cys residues would be modified. Therefore, we propose that the "+ 7 NEM" adduct arises in a slow reaction of a locally unfolded form whereby the "+ 2 NEM" state, which still has the native conformation, undergoes a small extent of unfolding that is ultimately trapped by reaction with NEM (Scheme 5.1).

To elucidate more completely this interesting reaction, we conducted a kinetics study to follow the "+ 2 NEM" state in its conversion to the "+ 7 NEM" state. Aliquots were taken from a solution containing ApeΔ40/E3330/NEM (ratio = 100 μM:500 μM: 500 μM) as a function of time, quenched by adding DTT, and infused into a mass spectrometer to monitor the relative abundances of the adducts. Data were fitted by Mathcad to reveal that the rate constant for the rate-determining step in the reaction of ApeΔ40 with NEM is 1.0 min^{-1} M^{-1}, which is approximately 2000 times smaller than that for reaction with a solvent-accessible Cys (Figure 5.9).

The Kinetics Are in Accord with a Locally Unfolded Form

The reactions of ApeΔ40, E3330, and NEM are described by a system of equations (Scheme 5.1) that was worked by D. Rempel in the Gross group. The first equation represents the heart of the unfolding processes: namely a binding equilibrium with E3330 and a reaction with an interior cysteine of the ApeΔ40 molecule to give "+ 3 NEM". The second and third equations show the sequential NEM reactions with the four remaining interior cysteine sites. These NEM products are assumed to remain in equilibrium with E3330 as indicated by the fourth through eighth equations. The reaction of solvent-exposed sites is represented by the ninth and tenth equations. Two exposed cysteines react quickly with NEM as described by k_{NEM} for a single site. Other exposed sites on the ApeΔ40 react slowly to add effectively one NEM with rate constant, k^{slow}_{NEM}. Because the sites are on the solvent-exposed surface, these reactions are postulated to

proceed independently of each other and of the other processes involving the ApeΔ40 molecule. The symbol ApeΔ40 in the ninth and tenth equations represents any species of ApeΔ40 in any other equation.

The correlated reaction of NEM with the five interior cysteines, as indicated by the absence of the "+ 4 NEM", "+ 5 NEM", and "+ 6 NEM" adducts in the mass spectra at all reaction times, indicates that the rate with which the unfolded ApeΔ40/E3330 complex refolds is low compared to the rate that the temporarily exposed cystine reacts with NEM with an assumed rate constant of k_{NEM}, and, as a consequence, that essentially all of unfolded ApeΔ40/E3330 is represented by "+ 7 NEM" or "+ 8 NEM" adducts as the unfolded ApeΔ40/E3330 is formed. A rate-limiting forward step in the first equation is indicated by the long time scale (many hours) for the formation of the "+ 7 NEM" or "+ 8 NEM" adducts of ApeΔ40 in the presence of E3330. This forward step may be due to a unimolecular conformational change (from *ApeΔ40* to *ApeΔ40'* in the first equation), governed by k'_U, that allows E3330 to bind. As a unimolecular forward rate-limiting step, the forward flux at this step should not be a function of the concentration of E3330. An ApeΔ40/NEM mixture titrated with E3330 and observed at a fixed reaction time shows that the "+ 7 NEM" adduct unambiguously increases with the E3330 concentration, thus, ruling out this possibility as the rate-limiting step. A rate-limiting forward step could be possibly due to the rate (k^*_U) at which the E3330 bound ApeΔ40 (*ApeΔ40'E3330*) opens up (*ApeΔ40*E3330*) to expose one or more interior cystines to react with NEM. In this possibility, a steady state for the ApeΔ40/E3330 binding is established. As the E3330 concentration is increased, the concentration of *ApeΔ40'E3330* reaches a maximum when all of Ape becomes bound. The flux to *ApeΔ40*E3330* and, because of the high reactivity of NEM with the now exposed

cysteines, the flux to the "+7 NEM" Ape adduct should show evidence of an asymptotic maximum as E3330 is increased in an ApeΔ40/NEM mixture titrated with E3330. An ApeΔ40-NEM mixture titrated with E3330 and observed at a fixed reaction time shows that the "+7 NEM adduct increased with no trend to lower slope, which would suggest the possibility of an asymptotic limit. This second possibility, thus, is ruled out. One is left with binding rate between ApeΔ40 and E3330 (k_{ON}) as the rate limiting step. The first equation is replaced by the eleventh equation for the system of equations that describes the observable behavior of ApeΔ40. A key feature of this reaction is that it goes to completion without any detectable intermediate adducts (e.g, those adducted to four, five, or six NEMs).

Reaction with Mutant ApeΔ40 Indicates that C65 Is Not the Essential Cys

Others suggested that C65 is the essential residue for the redox function of hApe1, and, therefore, it may be the preferred binding site of E3330 (*24*). Our results indicate that ApeΔ40C65A also undergoes the same local unfolding that is stabilized by E3330 and that reactions with NEM serve as a reporter for this process. Now, however, the NEM product is "ApeΔ40C65A + 6NEM" (Figure 5.10). Therefore, C65 is not the essential binding site for E3330 that stabilizes the local unfolding event.

Production of a "+ 3 NEM" Adduct Does Not Involve Cys

LC-MS/MS results of ApeΔ40/NEM after 6 h incubation show a small amount of "+ 3 NEM" (Figure 5.6); this product comes from the NEM reaction with other, non Cys solvent-accessible sites (e.g., K, H, and the $-NH_2$ at the terminus) and is promoted by excess NEM and long times (*76*). Although the general pK_a values of K, H and the terminal NH_2 are higher than that of Cys, the immediate environment around these residues may be favorable for the Michael addition. Therefore, we conclude that the "+ 3

NEM" adduct arises from the same stable, folded conformer of the protein that gives rise to the "+ 2 NEM" adduct (more evidence for this proposal is in the next section). We propose that the "+ 3 NEM" adduct is a "dead end", and is not an intermediate in forming the "+ 7 NEM" species.

The Extent of HDX Is Different for the "ApeΔ40 + 2 NEM" and "ApeΔ40 + 7 NEM"

Our hypothesis at this point is that E3330 interacts with a locally unfolded form of hApe1 to stabilize it. The reaction to give ultimately the "+ 7 NEM" adduct intercepts this state and preserves it in a more opened form. If this hypothesis is correct, then we should be able to see differences in the extent of HDX for the "+ 7 NEM" adduct in comparison with the "+2 NEM" adduct. Furthermore, if the "+ 3 NEM" adduct is a "deadend" as proposed above, its extent of HDX should match that of the "+2 NEM" adduct. Indeed, the "+ 2 NEM" adduct exchanges 144 amide hydrogens whereas for the "+ 7 NEM" adduct, the corresponding number of exchanges is 188 (Figure 5.11). We used a kinetic model, reported earlier (*77*) to fit the data and "bin" the exchanging amides into the number of fast, intermediate, and slow. For "ApeΔ40 + 2 NEM" adduct, the number of exchanging amide H's are 81± 1, 20 ± 4, 43 ± 2; for the "Ape Δ40 + 3 NEM" adduct, the number of amides are 81 ± 2, 19 ± 2, 41 ± 1, whereas for the "Ape Δ40 + 7 NEM" protein, the numbers are 136 ± 3, 14 ± 1, 38 ± 4. Clearly, the number of fast exchanging adducts has increased substantially for the unfolded conformer that becomes "locked-in" as a result of the reaction with NEM. "ApeΔ40 + 7 NEM" has 40 more exchanging amide H's than ApeΔ40/"ApeΔ40 + 2 NEM"/ "ApeΔ40 + 3 NEM", which is in agreement with the existence of a locally unfolded form.

Considering that there are 318 total amide hydrogens in the protein, we see that only 59% underwent exchange even for the protein modified with seven NEMs, indicating that "+ 7 NEM" form is a locally unfolded form and not a denatured one.

The HDX results for ApeΔ40, "ApeΔ40 + 2 NEM" and "ApeΔ40 + 3 NEM" proteins are similar (compare Figures 5.3 and 5.11), confirming that these three species have the same conformation and that additions of NEM to the solvent accessible Cys and to other reactive sidechains do not change the conformation of ApeΔ40.

The E3330 Interaction with ApeΔ40 Also Allows Disulfide Bond Formation

One hypothesized redox mechanism of hApe1 is thiol/disulfide exchange. That is, the Cys residues that are involved in redox activity become oxidized to disulfides as hApe1 reduces transcription factors. It is possible that E3330 inhibits the redox activity of hApe1 by stabilizing the locally unfolded form, allowing it to be oxidized prematurely, and making it unavailable as a reducing agent for other proteins. When we treated ApeΔ40 (10 μM) with E3330 (50 μM) for 1 h at 37 ºC in the absence of NEM, digested the mixture with trypsin in the presence of NEM (to prevent disulfide bond scrambling), and analyzed the digest by LC-MS/MS, we found that E3330 induced disulfide bond formation among the various Cys especially in the redox domain (C65, C93, and C99). The relative abundance of disulfide bonds in the ApeΔ40 redox domain C65-C93, C65-93 and C93-C99 increased owing to the interaction with E3330 from 0.16, 0.01, and 0.04 to 8.2, 1.6, and 5.8%, respectively (Figure 5.12). MS^2 spectra identifying the disulfide linkages are show in Figures 5.13, 5.14, 5.15, 5.16, and 5.17. As proved by NEM mapping experiments that show "+ 2 NEM" going to "+ 7 NEM", E3330 interacts with and stabilizes the locally unfolded form of ApeΔ40, which then becomes oxidized by

forming disulfide bonds. Thus, this locally unfolded form of ApeΔ40 may be responsible for the redox activity of the protein.

Comparison with Full-length Ape (FL-Ape)

To test if FL-Ape undergoes the same local unfolding process, we tested the FL-Ape /E3330/NEM sample after 12 h incubation by ESI-MS, using an identical approach as that used for ApeΔ40. We also observed the transition from "+2 NEM" to "+ 7 NEM" adduct (Figure 5.18) without any detectable intermediates, suggesting FL-Ape undergoes the same local unfolding as ApeΔ40, and that form is stabilized by the interaction with E3330.

Discussion

Until now, the role of hApe1 in the reduction of transcription factors has remained unclear. It is generally agreed that one or more of the three Cys in the redox domain are involved in the redox activity (*67, 78*), but there are no results that locate the redox-active cysteines. In cell-based experiments, the hApe1 mutant C65A (C65 replaced with A) shows nearly no redox activity, C93A retains approximately 50% activity, and C99A retains approximately 80% activity (*66*). These results are in accord with a conclusion that not only is C65 essential for redox activity but also C93 and C99 are involved.

We suggest that another way to elucidate the redox mechanism of hApe1 is through a study of the interaction of hApe1 and its inhibitor E3330 by using various forms of chemical footprinting with analysis by mass spectrometry. Those residues that become solvent-accessible and interact with E3330 may be those that are redox-active in the reduced form of hApe1. The only study that bears on this prospect is one that

afforded a K_d of 1.6 nM for the binding of hApe1 and E3330, as determined by SPR (*38*), indicating strong and specific binding between hApe1 and E3330. We find no evidence for this strong interaction. ApeΔ40, a truncated form of hApe1 with the full redox and DNA repair activities, was used for our study. Our HDX kinetics results are in accord with no significant difference between ApeΔ40 without and with E3330 at protein concentrations in the range of 0.5-20 μM. It seems highly unlikely that an interaction as strong as that with a K_d = 1.6 nM would cause no changes in the extent of H bonding and in the conformation of ApeΔ40. One explanation of the SPR results is that the E3330 is reacting with the (possibly denatured) protein to give a very small k_{off}.

Efforts to detect the complex directly by ESI also provided no convincing evidence for its existence, either by ESI-MS with 50% ACN with 0.1% FA or by nESI-MS of aqueous solutions at neutral pH. An abundant complex, however, does form when ApeΔ40 interacts with 8-51, another quinone and an analog of E3330. It is likely that the –OMe leaving group of 8-51 makes it a more labile reactant than E3330 for the solvent-accessible Cys residues. Nevertheless, the reaction shows conclusively that covalent bond formation with a quinone ligand can be detected by ESI-MS. Thus, we conclude that E3330 does not interact strongly or react irreversibly with ApeΔ40 via a Michael addition or any other reaction. Nevertheless, E3330 does specifically interact with hApe1 via a process that may be general for other proteins with solvent-accessible cysteines. What are the details of this interaction between ApeΔ40 and E3330?

The NEM mapping experiments surprisingly show formation of a species "ApeΔ40 + 7 NEM" in the presence of E3330. We propose that the ApeΔ40 form that becomes modified with seven NEMs is locally unfolded. Therefore, ApeΔ40 must exist in at least two equilibrating conformers (folded and locally unfolded) in solution. The

locally unfolded form, in the absence of E3330, is too small a fraction to be detectable by MS methods. In the presence of E3330, however, this locally unfolded form is stabilized by E3330, its relative concentration increased, and it is ultimately trapped in a slow reaction with NEM. The first adduct of a normally inaccessible Cys, the "+ 3 NEM" form (3 Cys modified), is sufficiently long-lived and its conformation sufficiently open that the remaining four Cys residues now can react rapidly with NEM in a cascade effect (Scheme 5.1). The outcome is an irreversible uptake of five additional NEMs by Cys residues that are inaccessible in the folded and stable conformation of the protein.

The effects of E3330 are only apparent in the presence of NEM. If NEM were added to ApeΔ40 that was equilibrated for an equal time with E3330, no uptake of seven NEMs can be seen, consistent with the fact that addition of NEM is a slow reaction owing to the low concentration of the locally unfolded form, whereas the reaction is fast for solvent-accessible Cys residues. The "+ 7 NEM" species is viewed as a trapped unfolded form that cannot convert back to its folded state owing to the Cys modifications. This locally unfolded form is not detectable by HDX because its concentration is too small. We propose, however, that HDX of the "+ 7 NEM" is a measure of the locally unfolded form. The locally unfolded conformer that interacts with and is stabilized by E3330 may be the sought-after form that functions as the redox factor.

More confirmation of this picture comes from the measurement of extent of disulfide bond formation in the presence of E3330. Indeed, disulfide bond formation among the cysteines in the redox domain is enhanced in the presence of E3330, providing evidence that the interaction with E3330 is relevant to the redox activity of hApe1. We interpret the results to indicate that, in the locally unfolded form, the cysteines in the redox domain are sufficiently close to form disulfide bonds. Although other disulfide

bonds (C65-138, C93-138) also form, their role in the inhibition activity of E3330 is not yet explained. It is also not clear whether E3330 initiates the disulfide bond formation or whether two thiols simply react with each other spontaneously in the locally unfolded form.

Conclusions

ApeΔ40 exists in solution as an equilibrium between a major folded form and a minor locally unfolded form. The locally unfolded form with at least one formerly solvent-inaccessible cysteine exposed to solvent is proposed to be the one that interacts with E3330. This form can be trapped by NEM reactions and can form disulfide bonds among the various cysteines in the redox domain of ApeΔ40. The FL-Ape underwent the same unfolding process upon interacting with E3330 in the presence of E3330, indicating the same equilibrium between a major folded form and a minor locally unfolded form exists for hApe1. Therefore, we propose that the locally unfolded form is responsible for the redox activity of hApe1 via a thiol/disulfide exchange mechanism. A possible pathway in this mechanism is that C65 acts as a nucleophile to attack a disulfide in the target protein, producing a transient mixed disulfide intermediate. The intermediate is then reduced by C93 and C99 producing a dithiol in the target protein and a disulfide bond in hApe1.

Acknowledgment

We thank Bich Vu, Richard Huang, Hao Zhang, Weidong Cui and Don Rempel for their significant contributions and help with this project. We are also grateful to our collaborators Drs. Millie M. Georgiadis with Sarah Delaplane, Meihua Luo, Mark R. Kelley from Indiana University, and Richard F. Borch and Rodney L. Nyland II from

Purdue University. This research was supported by the Washington University NIH Mass Spectrometry Research Resource (Grant No. P41 RR000954) and by the NIH (Grant R01CA114571).

References

(1) Evans, A. R., Limp-Foster, M. and Kelley, M. R. (2000) Going APE over ref-1. *Mutat. Res. 461*, 83-108.

(2) Fritz, G. (2000) Human APE/Ref-1 protein. *Int J Biochem Cell Biol 32*, 925-929.

(3) Dyrkheeva, N. S., Khodyreva, S. N. and Lavrik, O. I. (2007) Multifunctional human apurinic/apyrimidinic endonuclease 1: the role of additional functions. *Mol. Biol. 41*, 450-466.

(4) Bhakat, K. K., Mantha, A. K. and Mitra, S. (2009) Transcriptional regulatory functions of mammalian AP-endonuclease (APE1/Ref-1), an essential multifunctional protein. *Antioxid. Redox Signal 11*, 621-638.

(5) Jeon, B. H. and Irani, K. (2009) APE1/Ref-1: versatility in progress. *Antioxid. Redox Signal 11*, 571-574.

(6) Tell, G., Quadrifoglio, F., Tiribelli, C. and Kelley, M. R. (2009) The many functions of APE1/Ref-1: not only a DNA repair enzyme. *Antioxid. Redox Signal 11*, 601-620.

(7) Luo, M., He, H., Kelley, M. R. and Georgiadis, M. (2009) Redox Regulation of DNA Repair: Implications for Human Health and Cancer Therapeutic Development. *Antioxid. Redox Signal*.

(8) Abate, C., Luk, D. and Curran, T. (1990) A ubiquitous nuclear protein stimulates the DNA-binding activity of fos and jun indirectly. *Cell Growth Differ. 1*, 455-462.

(9) Xanthoudakis, S. and Curran, T. (1992) Identification and characterization of Ref-1, a nuclear protein that facilitates AP-1 DNA-binding activity. *EMBO J. 11*, 653-665.

(10) Xanthoudakis, S., Miao, G., Wang, F., Pan, Y. C. and Curran, T. (1992) Redox activation of Fos-Jun DNA binding activity is mediated by a DNA repair enzyme. *EMBO J. 11*, 3323-3335.

(11) Huang, R. P. and Adamson, E. D. (1993) Characterization of the DNA-binding properties of the early growth response-1 (Egr-1) transcription factor: evidence for modulation by a redox mechanism. *DNA Cell Biol. 12*, 265-273.

(12) Yao, K. S., Xanthoudakis, S., Curran, T. and O'Dwyer, P. J. (1994) Activation of AP-1 and of a nuclear redox factor, Ref-1, in the response of HT29 colon cancer cells to hypoxia. *Mol. Cell Biol. 14*, 5997-6003.

(13) Huang, L. E., Arany, Z., Livingston, D. M. and Bunn, H. F. (1996) Activation of hypoxia-inducible transcription factor depends primarily upon redox-sensitive stabilization of its alpha subunit. *J. Biol. Chem. 271*, 32253-32259.

(14) Hirota, K., Matsui, M., Iwata, S., Nishiyama, A., Mori, K. and Yodoi, J. (1997) AP-1 transcriptional activity is regulated by a direct association between thioredoxin and Ref-1. *Proc. Natl. Acad. Sci. U S A 94*, 3633-3638.

(15) Ema, M., Hirota, K., Mimura, J., Abe, H., Yodoi, J., Sogawa, K., Poellinger, L. and Fujii-Kuriyama, Y. (1999) Molecular mechanisms of transcription activation by HLF and HIF1alpha in response to hypoxia: their stabilization and redox signal-induced interaction with CBP/p300. *EMBO J. 18*, 1905-1914.

(16) Hirota, K., Murata, M., Sachi, Y., Nakamura, H., Takeuchi, J., Mori, K. and Yodoi, J. (1999) Distinct roles of thioredoxin in the cytoplasm and in the nucleus.

A two-step mechanism of redox regulation of transcription factor NF-kappaB. *J. Biol. Chem. 274*, 27891-27897.

(17) Ueno, M., Masutani, H., Arai, R. J., Yamauchi, A., Hirota, K., Sakai, T., Inamoto, T., Yamaoka, Y., Yodoi, J. and Nikaido, T. (1999) Thioredoxin-dependent redox regulation of p53-mediated p21 activation. *J. Biol. Chem*. *274*, 35809-35815.

(18) Lando, D., Pongratz, I., Poellinger, L. and Whitelaw, M. L. (2000) A redox mechanism controls differential DNA binding activities of hypoxia-inducible factor (HIF) 1alpha and the HIF-like factor. *J. Biol. Chem. 275*, 4618-4627.

(19) Tell, G., Zecca, A., Pellizzari, L., Spessotto, P., Colombatti, A., Kelley, M. R., Damante, G. and Pucillo, C. (2000) An 'environment to nucleus' signaling system operates in B lymphocytes: redox status modulates BSAP/Pax-5 activation through Ref-1 nuclear translocation. *Nucleic Acids Res. 28*, 1099-1105.

(20) Cao, X., Kambe, F., Ohmori, S. and Seo, H. (2002) Oxidoreductive modification of two cysteine residues in paired domain by Ref-1 regulates DNA-binding activity of Pax-8. *Biochem. Biophys. Res. Commun. 297*, 288-293.

(21) Doetsch, P. W. and Cunningham, R. P. (1990) The enzymology of apurinic/apyrimidinic endonucleases. *Mutat. Res. 236*, 173-201.

(22) Demple, B. and Harrison, L. (1994) Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem. 63*, 915-948.

(23) Barzilay, G., Mol, C. D., Robson, C. N., Walker, L. J., Cunningham, R. P., Tainer, J. A. and Hickson, I. D. (1995) Identification of critical active-site residues in the multifunctional human DNA repair enzyme HAP1. *Nat. Struct. Biol. 2*, 561-568.

(24) Walker, L. J., Robson, C. N., Black, E., Gillespie, D. and Hickson, I. D. (1993) Identification of residues in the human DNA repair enzyme HAP1 (Ref-1) that

are essential for redox regulation of Jun DNA binding. *Mol. Cell. Biol. 13*, 5370-5376.

(25) Xanthoudakis, S., Miao, G. G. and Curran, T. (1994) The redox and DNA-repair activities of Ref-1 are encoded by nonoverlapping domains. *Proc. Natl. Acad. Sci. U S A 91*, 23-27.

(26) Jayaraman, L., Murthy, K. G., Zhu, C., Curran, T., Xanthoudakis, S. and Prives, C. (1997) Identification of redox/repair protein Ref-1 as a potent activator of p53. *Genes Dev.11*, 558-570.

(27) Izumi, T. and Mitra, S. (1998) Deletion analysis of human AP-endonuclease: minimum sequence required for the endonuclease activity. *Carcinogenesis 19*, 525-527.

(28) Gorman, M. A., Morera, S., Rothwell, D. G., de La Fortelle, E., Mol, C. D., Tainer, J. A., Hickson, I. D. and Freemont, P. S. (1997) The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. *EMBO J. 16*, 6548-6558.

(29) Mol, C. D., Izumi, T., Mitra, S. and Tainer, J. A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. *Nature 403*, 451-456.

(30) Beernink, P. T., Segelke, B. W., Hadi, M. Z., Erzberger, J. P., Wilson, D. M., 3rd and Rupp, B. (2001) Two divalent metal ions in the active site of a new crystal form of human apurinic/apyrimidinic endonuclease, Ape1: implications for the catalytic mechanism. *J. Mol. Biol. 307*, 1023-1034.

(31) Georgiadis, M. M., Luo, M., Gaur, R. K., Delaplane, S., Li, X. and Kelley, M. R. (2008) Evolution of the redox function in mammalian apurinic/apyrimidinic endonuclease. *Mutat. Res. 643*, 54-63.

(32) Ordway, J. M., Eberhart, D. and Curran, T. (2003) Cysteine 64 of Ref-1 is not essential for redox regulation of AP-1 DNA binding. *Mol. Cell. Biol. 23*, 4257-4266.

(33) Xu, Y., Moore, D. H., Broshears, J., Liu, L., Wilson, T. M. and Kelley, M. R. (1997) The apurinic/apyrimidinic endonuclease (APE/ref-1) DNA repair enzyme is elevated in premalignant and malignant cervical cancer. *Anticancer Res. 17*, 3713-3719.

(34) Kelley, M. R., Xu, Y., Tritt, R. and Robertson, K. A. (1998) The multifunctional DNA base excision repair and redox protein, AP endonuclease (APE/Ref-1), and its role in germ cell tumors, In *Germ Cell Tumors* pp 81-86, Jonh Libbey and Co., London, UK.

(35) Moore, D. H., Michael, H., Tritt, R., Parsons, S. H. and Kelley, M. R. (2000) Alterations in the expression of the DNA repair/redox enzyme APE/ref-1 in epithelial ovarian cancers. *Clin. Cancer Res. 6*, 602-609.

(36) Wang, D., Luo, M. and Kelley, M. R. (2004) Human apurinic endonuclease 1 (APE1) expression and prognostic significance in osteosarcoma: enhanced sensitivity of osteosarcoma to DNA damaging agents using silencing RNA APE1 expression inhibition. *Mol. Cancer Ther. 3*, 679-686.

(37) Kelley, M. R. and Fishel, M. L. (2008) DNA repair proteins as molecular targets for cancer therapeutics. *Anticancer Agents Med. Chem. 8*, 417-425.

(38) Shimizu, N., Sugimoto, K., Tang, J., Nishi, T., Sato, I., Hiramoto, M., Aizawa, S., Hatakeyama, M., Ohba, R., Hatori, H., Yoshikawa, T., Suzuki, F., Oomori, A., Tanaka, H., Kawaguchi, H., Watanabe, H. and Handa, H. (2000) High-performance affinity beads for identifying drug receptors. *Nat. Biotechnol. 18*, 877-881.

(39) Zou, G. M., Luo, M. H., Reed, A., Kelley, M. R. and Yoder, M. C. (2007) Ape1 regulates hematopoietic differentiation of embryonic stem cells through its redox functional domain. *Blood 109*, 1917-1922.

(40) Zou, G. M. and Maitra, A. (2008) Small-molecule inhibitor of the AP endonuclease 1/REF-1 E3330 inhibits pancreatic cancer cell growth and migration. *Mol. Cancer Ther. 7*, 2012-2021.

(41) Luo, M., Delaplane, S., Jiang, A., Reed, A., He, Y., Fishel, M., Nyland, R. L., 2nd, Borch, R. F., Qiao, X., Georgiadis, M. M. and Kelley, M. R. (2008) Role of the multifunctional DNA repair and redox signaling protein Ape1/Ref-1 in cancer and endothelial cells: small-molecule inhibition of the redox function of Ape1. *Antioxid. Redox Signal 10*, 1853-1867.

(42) Luo, M., He, H., Kelley, M. R. and Georgiadis, M. (2009) Redox Regulation of DNA Repair: Implications for Human Health and Cancer Therapeutic Development. *Antioxid. Redox Signal*, in press.

(43) Katta, V. and Chait, B. T. (1991) Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Commun Mass Spectrom 5*, 214-217.

(44) Engen, J. R. and Smith, D. L. (2001) Investigating protein structure and dynamics by hydrogen exchange MS. *Anal. Chem. 73*, 256A-265A.

(45) Hoofnagle, A. N., Resing, K. A. and Ahn, N. G. (2003) Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct. 32*, 1-25.

(46) Garcia, R. A., Pantazatos, D. and Villarreal, F. J. (2004) Hydrogen/deuterium exchange mass spectrometry for investigating protein-ligand interactions. *Assay Drug Dev. Technol 2*, 81-91.

(47) Hamuro, Y., Weber, P. C. and Griffin, P. R. (2005) High-throughput analysis of protein structure by hydrogen/deuterium exchange mass spectrometry. *Methods Biochem. Anal. 45*, 131-157.

(48) Maier, C. S. and Deinzer, M. L. (2005) Protein conformations, interactions, and H/D exchange. *Methods Enzymol. 402*, 312-360.

(49) Tsutsui, Y. and Wintrode, P. L. (2007) Hydrogen/deuterium exchange-mass spectrometry: a powerful tool for probing protein structure, dynamics and interactions. *Curr. Med. Chem. 14*, 2344-2358.

(50) Weis, D. D., Kaveti, S., Wu, Y. and Engen, J. R. (2007) Probing protein interactions using hydrogen-deuterium exchange mass spectrometry In *Mass Spectrometry of Protein Interactions* (Downard, K. M., Ed.) pp 45-61, John Wiley & Sons, Inc., Hoboken, NJ.

(51) Zheng, X. and Wintrode, P. K. (2008) Complementary methods for structure determination: hydroxy-radical-mediated footprinting and deuterium exchange mass spectrometry as applied to serpin structure In *Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics* (Chance, M., Ed.) pp 69-90, John Wiley & Sons, Inc., Hoboken, NJ.

(52) Engen, J. R. (2009) Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS. *Anal. Chem. 81*, 7870-7875.

(53) Zhu, M. M., Rempel, D. L., Du, Z. and Gross, M. L. (2003) Quantification of protein-ligand interactions by mass spectrometry, titration, and H/D exchange: PLIMSTEX. *J. Am. Chem. Soc. 125*, 5252-5253.

(54) Zhu, M. M., Rempel, D. L. and Gross, M. L. (2004) Modeling data from titration, amide H/D exchange, and mass spectrometry to obtain protein-ligand binding constants. *J. Am. Soc. Mass Spectrom 15*, 388-397.

(55) Zhu, M. M., Chitta, R. and Gross, M. L. (2005) PLIMSTEX: a novel mass spectrometric method for the quantification of protein-ligand interactions in solution. *Int. J. Mass Spectrom 240*, 213-220.

(56) Zhu, M. M., Hambly, D. and Gross, M. L. (2007) Quantification of Protein-Ligand Interactions in Solution by Hydrogen/Deuterium Exchange (PLIMSTEX) In *Methods and Principles in Medicinal Chemistry* (Wanner, K. and Höfner, G., Eds.) pp 341-376, Wiley-VCH.

(57) Takamoto, K. and Kiselar, J. (2008) Covalent labeling methods for examining protein structure and protein interactions In *Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics* (Chance, M., Ed.) pp 45-68, John Wiley & Sons, Inc., Hoboken, NJ.

(58) Fitzgerald, M. C. and West, G. M. (2009) Painting proteins with covalent labels: what's in the picture? *J. Am. Soc. Mass Spectrom 20*, 1193-1206.

(59) Mendoza, V. L. and Vachet, R. W. (2009) Probing protein structure by amino acid-specific covalent labeling and mass spectrometry. *Mass Spectrom Rev. 28*, 785-815.

(60) Titani, Y. and Tsuruta, Y. (1974) Some chemical and biological characteristics of showdomycin. *J. Antibiot. (Tokyo) 27*, 956-962.

(61) Kim, Y. J., Pannell, L. K. and Sackett, D. L. (2004) Mass spectrometric measurement of differential reactivity of cysteine to localize protein-ligand binding sites. Application to tubulin-binding drugs. *Anal. Biochem. 332*, 376-383.

(62) Rishavy, M. A., Pudota, B. N., Hallgren, K. W., Qian, W., Yakubenko, A. V., Song, J. H., Runge, K. W. and Berkner, K. L. (2004) A new model for vitamin K-dependent carboxylation: the catalytic base that deprotonates vitamin K hydroquinone is not Cys but an activated amine. *Proc. Natl. Acad. Sci. U S A 101*, 13732-13737.

(63) Schilling, B., Yoo, C. B., Collins, C. J. and Gibson, B. W. (2004) Determining cysteine oxidation status using differential alkylation. *International Journal of Mass Spectrometry 236*, 117–127.

(64) Kurono, S., Kurono, T., Komori, N., Niwayama, S. and Matsumoto, H. (2006) Quantitative proteome analysis using D-labeled N-ethylmaleimide and 13C-labeled iodoacetanilide by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Bioorg. Med. Chem. 14*, 8197-8209.

(65) Guan, L. and Kaback, H. R. (2007) Site-directed alkylation of cysteine to test solvent accessibility of membrane proteins. *Nat. Protoc. 2*, 2012-2017.

(66) Georgiadis, M. M., Luo, M., Gaur, R. K., Delaplane, S., Li, X. and Kelley, M. R. (2008) Evolution of the redox function in mammalian apurinic/apyrimidinic endonuclease. *Mutat. Res. 643*, 54-63.

(67) Luo, M., Delaplane, S., Jiang, A., Reed, A., He, Y., Fishel, M., Nyland, R. L., 2nd, Borch, R. F., Qiao, X., Georgiadis, M. M. and Kelley, M. R. (2008) Role of the

multifunctional DNA repair and redox signaling protein Ape1/Ref-1 in cancer and endothelial cells: small-molecule inhibition of the redox function of Ape1. *Antioxid. Redox Signal 10*, 1853-1867.

(68) Nyland, R. L., Luo, M., Kelley, M. R. and Borch, R. F. Design and synthesis of novel quinone inhibitors targeted to the redox function of apurinic/apyrimidinic endonuclease 1/redox enhancing factor-1 (Ape1/ref-1). *J. Med. Chem. 53*, 1200-1210.

(69) Zhu, M. M. (2004) PLIMSTEX (Protein-Ligand Interactions in Solution by Mass Spectrometry, Titration and H/D Exchange), In *Chemistry*, Washington University in St. Louis, St. Louis.

(70) Xu, H. and Freitas, M. A. (2007) A mass accuracy sensitive probability based scoring algorithm for database searching of tandem mass spectrometry data. *BMC Bioinformatics 8*, 133.

(71) Xu, H., Yang, L. and Freitas, M. A. (2008) A robust linear regression based algorithm for automated evaluation of peptide identifications from shotgun proteomics by use of reversed-phase liquid chromatography retention time. *BMC Bioinformatics 9*, 347.

(72) Xu, H., Zhang, L. and Freitas, M. A. (2008) Identification and characterization of disulfide bonds in proteins and peptides from tandem MS data by use of the MassMatrix MS/MS search engine. *J. Proteome Res. 7*, 138-144.

(73) Hernandez, H. and Robinson, C. V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc. 2*, 715-726.

(74) Sandercock, A. M. and Robinson, C. V. (2007) Electrospray Ionization Mass Spectrometry and the Study of Protein Complexes, In *Protein Interactions* (Schuck, P., Ed.) pp 447-468, Springer US.

(75) Sharon, M. and Robinson, C. V. (2007) The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annu. Rev. Biochem. 76*, 167-193.

(76) Brewer, C. F. and Riehm, J. P. (1967) Evidence for Possible Nonspecific Reactions between N-Ethylmaleimide and Proteins. *Anal. Biochem. 18*, 248-255.

(77) Zhu, M. M., Rempel, D. L., Zhao, J., Giblin, D. E. and Gross, M. L. (2003) Probing Ca2+-induced conformational changes in porcine calmodulin by H/D exchange and ESI-MS: effect of cations and ionic strength. *Biochemistry 42*, 15388-15397.

(78) Walker, L. J., Robson, C. N., Black, E., Gillespie, D. and Hickson, I. D. (1993) Identification of residues in the human DNA repair enzyme HAP1 (Ref-1) that are essential for redox regulation of Jun DNA binding. *Mol. Cell Biol. 13*, 5370-5376.

Figure 5.1. X-ray crystal structure (PDB: 1BIX) of hApe1 with seven cysteines shown as sticks (red). C99 and C138 are surface accessible, whereas C65, C93, C208, C296, and C310 are buried. C65, C93, and C99 are located in redox domain. C65 is thought to be essential in the redox chemistry of the protein.

```
        40          50          60          70          80          90
   GSHMASE GPALYEDPPD QKTSPSGKPA TLKICSWNVD GLRAWIKKKG LDWVKEEAPD

       100         110         120         130         140         150
 ILCLQETKCS ENKLPAELQE LPGLSHQYWS APSDKEGYSG VGLLSRQCPL KVSYGIGDEE

       160         170         180         190         200         210
 HDQEGRVIVA EFDSFVLVTA YVPNAGRGLV RLEYRQRWDE AFRKFLKGLA SRKPLVLCGD

       220         230         240         250         260         270
 LNVAHEEIDL RNPKGNKKNA GFTPQERQGF GELLQAVPLA DSFRHLYPNT PYAYTFWTYM

       280         290         300         310
 MNARSKNVGW RLDYFLLSHS LLPALCDSKI RSKALGSDHC PITLYLAL
```

Figure 5.2. Amino acid sequence of the ApeΔ40 (cysteine residues are underscored).

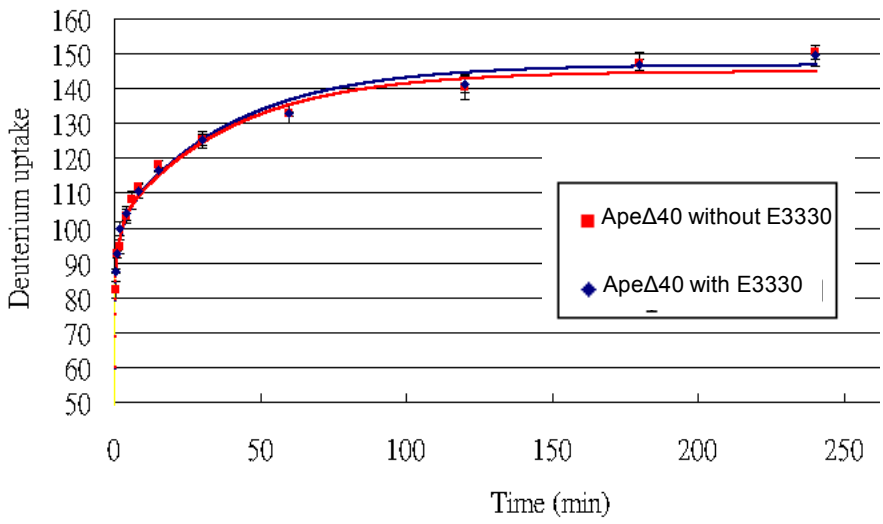


Figure 5.3. HDX kinetics of ApeΔ40 alone and in the presence of E3330. The curves were fitted with a custom program operating in MathCAD. It used the D uptakes of ApeΔ40 after 1 h incubation without and with E3330. The protein concentration was 12.5 μM, and for the experiment in the presence of E3330, the [protein]/[E3330] was 1/10. The samples were incubated in 10 mM HEPES with 150 mM KCl at 25 ºC for 1 h. HDX was done in a 90% D_2O medium (pH 7.5) with 10 mM HEPES and 150 mM KCl at 25 ºC. The HDX reaction was quenched by adding sufficient 1.0 M cold HCl to give a pH of 2.5.

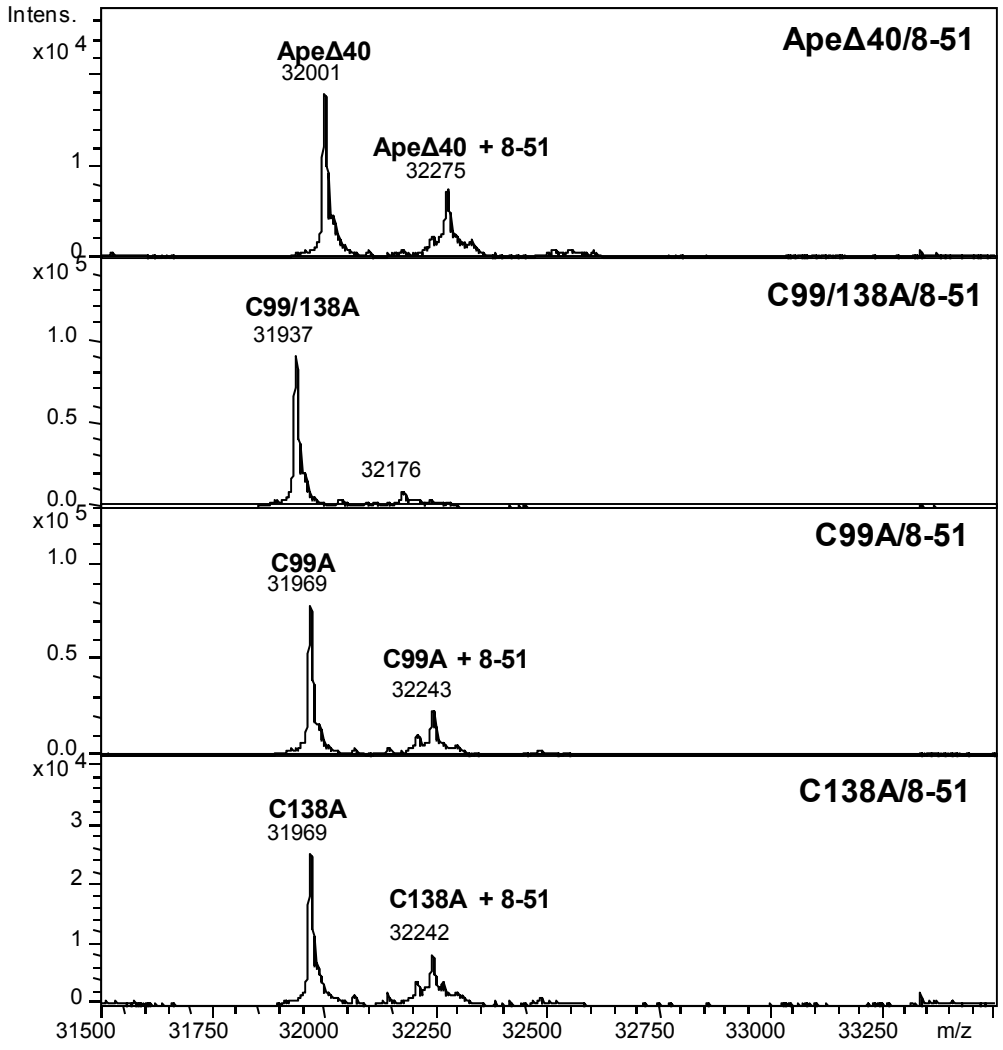


Figure 5.4. ESI mass spectra of ApeΔ40 (A) and its mutants C99A (B), C138A (C), C99/138A (D) after interaction with compound 8-51. Mass spectra were collected on a Bruker MaXis UHR-TOF instrument and deconvolution was done with MaxEnt1 algorithm.

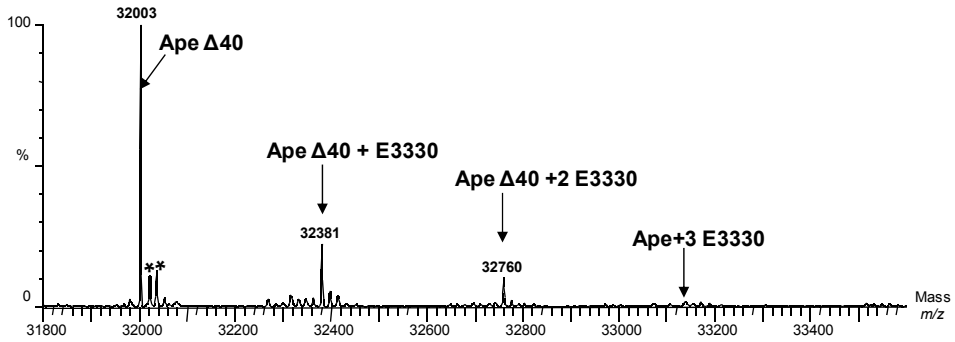


Figure 5.5. ESImass spectrum of ApeΔ40 and E3330 with 50% ACN with 0.1% FA at 10 μL/min. Sample was incubated in 100 mM ammonium bicarbonate at pH 7.5 for 4 h at RT and before MS analysis: [protein] = 100 μM; [E3330] = 500 μM. The symbol * denotes peaks for the water and salt adducts. Data were collected on a Waters Micromass Q-TOF instrument, and deconvolution was done with MaxEnt1 algorithm provided with that system. Instrument parameters were the same as described in the HDX experiments except theflow rate was 10 μL/min, capillary voltage was 2.5 kV, and the collision energy was 5 eV.

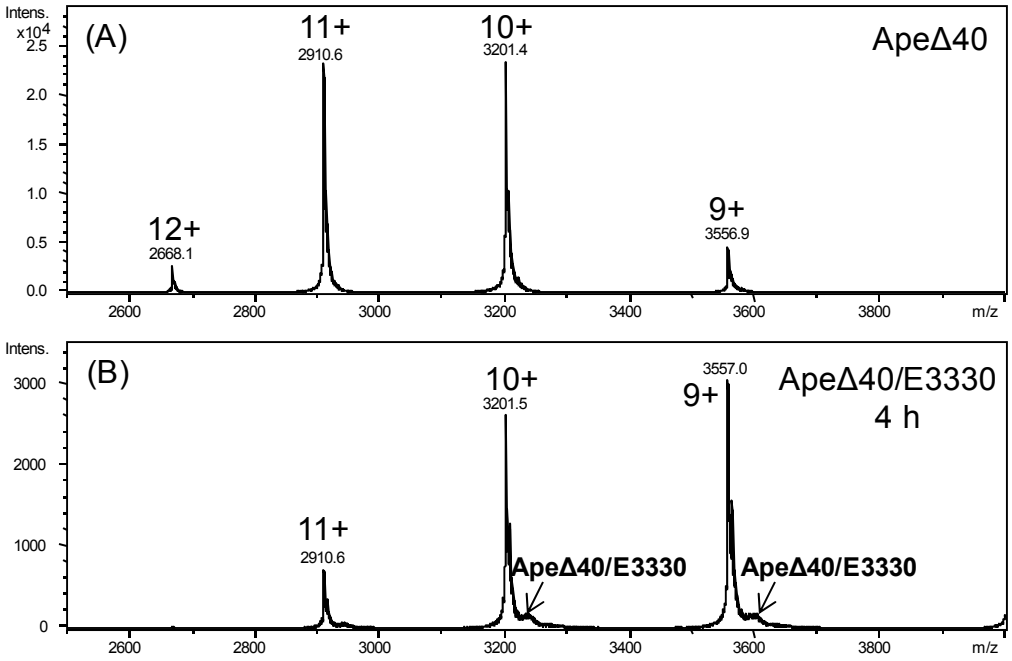


Figure 5.6. NanoESI mass spectra of ApeΔ40 without (A) and with E3330 (B). Samples were incubated in 1 M ammonium acetate at pH 7.5 for 4 h at RT and then directly analyzed by nESI-MS ([protein] = 100 μM; [E3330] = 500 μM). Mass spectra were collected on a Bruker MaXis UHR-TOF instrument.

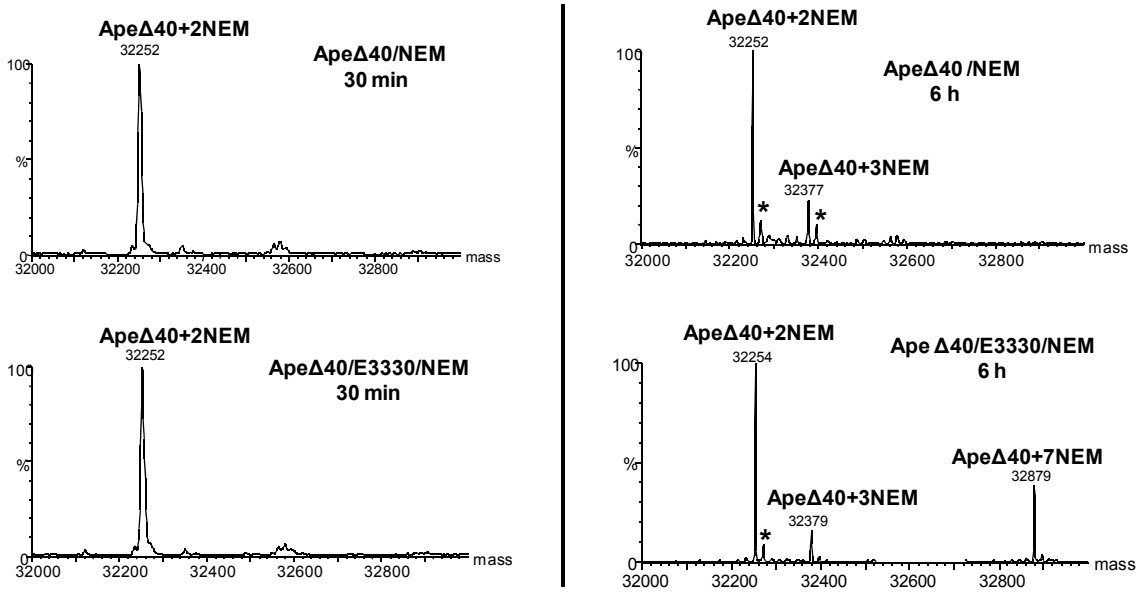


Figure 5.7. ESI mass spectra of ApeΔ40 after incubation without and with E3330 in the presence of NEM for 30 min (left panel) and 6 h (right panel). Samples were incubated in 10 mM HEPES with 150 mM KCl at pH 7.5([protein] = 100 μM; [E3330] = [NEM] = 500 μM). The symbol * denotes peaks for the water adducts. Mass spectra were collected on a Waters Micromass Q-TOF instrument and deconvolution was done with MaxEnt1 algorithm provided with that system.

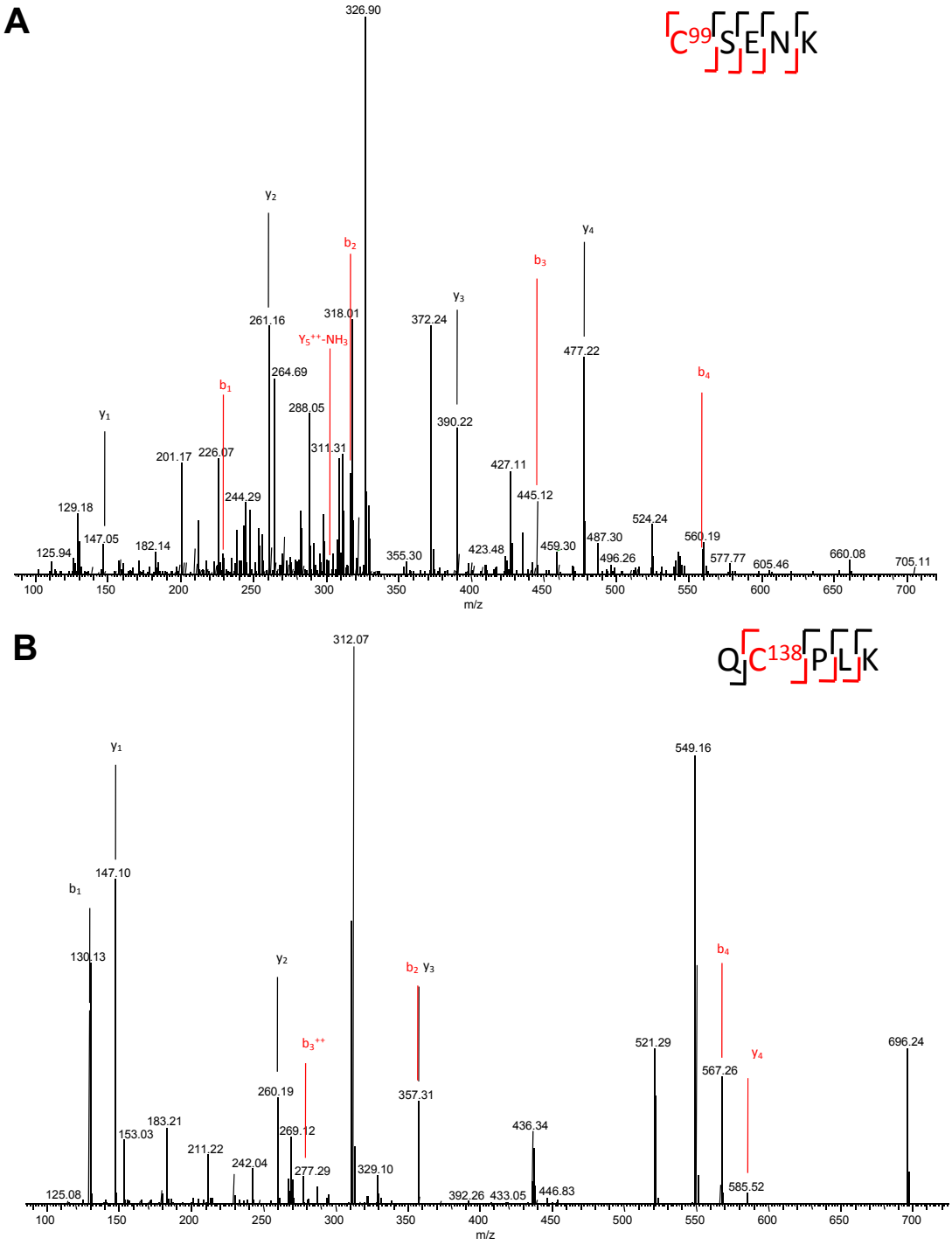


Figure 5.8. MS/MS spectra that show C99 and C138 were modified by NEM. Product ions labeled in read contain the NEM modified cysteines C99 (A) and C138 (B). The symbol ++ denotes doubly charged ions.

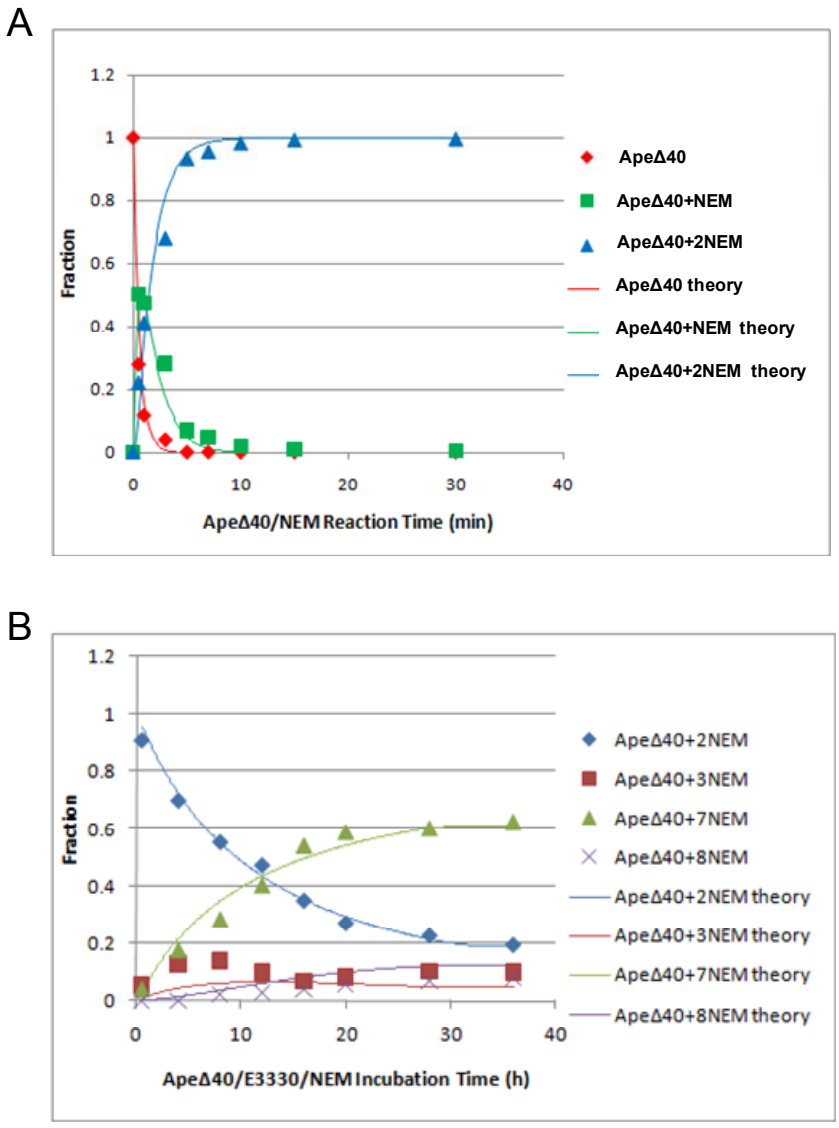


Figure 5.9. Kinetics of Ape and NEM reaction in (A) Ape Δ40/NEM sample and (B) ApeΔ40/E3330/NEM sample. Both samples were prepared in 10 mM HEPES at pH 7.5 at RT ([protein] = 100 μM; [E3330] = [NEM] = 500 μM). Aliquots were quenched with DTT at various times. Mass spectra were collected on a Bruker MaXis UHR-TOF instrument, and deconvolution was done with MaxEnt1 algorithm provided with that system. The sums of intensities of different NEM adducts of ApeΔ40 were normalized to 1. Data were fitted with Mathcad. The value of k_{NEM} was determined in a relatively

straight forward one-parameter fit of the "+ 2 NEM" adduct kinetic curves. A value of 2100 min^{-1} M^{-1} was obtained when the relative sensitivity factor was fixed as a constant. When g_N was fixed at the more extreme value of 1/NEM, a better fit was obtained, and a value for k_{NEM} of 1500 min^{-1} M^{-1} was obtained. This variation is only 29% despite the uncertainty of ESI sensitivity as a function of the number of adducted NEMs.

The "+ 7 NEM" kinetic curves where fit by allowing three parameters to vary. They are listed with their solution values in the following equations.

$k^{slow}_{NEM} = 0.741932547$ min^{-1} M^{-1}

$k_{ON} = 1.010476359$ min^{-1} M^{-1}

$g_N = 0.6519354671$

For this calculation, it was assumed that effect of the dissociation rate of the ApeΔ40/E3330 complexis small. This is consistent with the hypothesis that the association is the rate limiting step so that the concentration $Ape\Delta 40^{*}E3330$ remains small because of the relatively rapid reaction with NEM. The value of k_{OFF} must approach the value of k_{NEM} in order for there to be an effect. The value of k_{NEM} was fixed at 1740 for the fit.

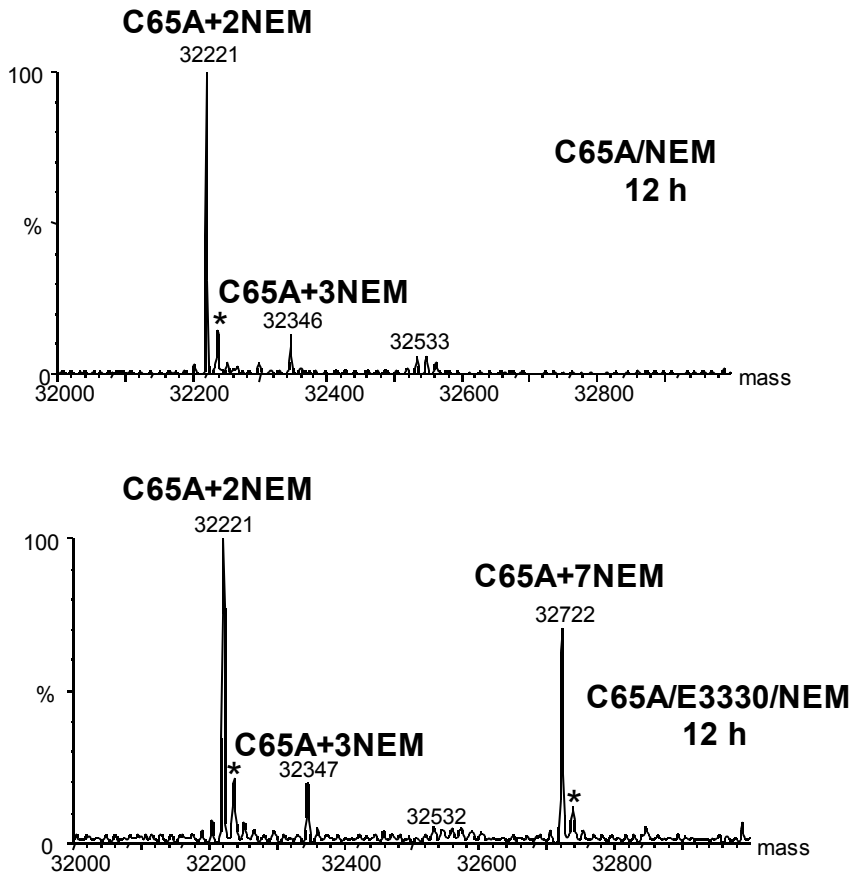


Figure 5.10. ESI mass spectra of ApeΔ40 mutant C65A without and with incubation with E3330 in the presence of NEM for 12 h. Samples were incubated in 10 mM HEPES at pH 7.5([protein] = 100 μM; [E3330] = [NEM] = 500 μM). The symbol * denotes water adducts. Mass spectra were collected on a Waters Micromass Q-TOF instrument and deconvolution was done with MaxEnt1 algorithm that was part of that system.

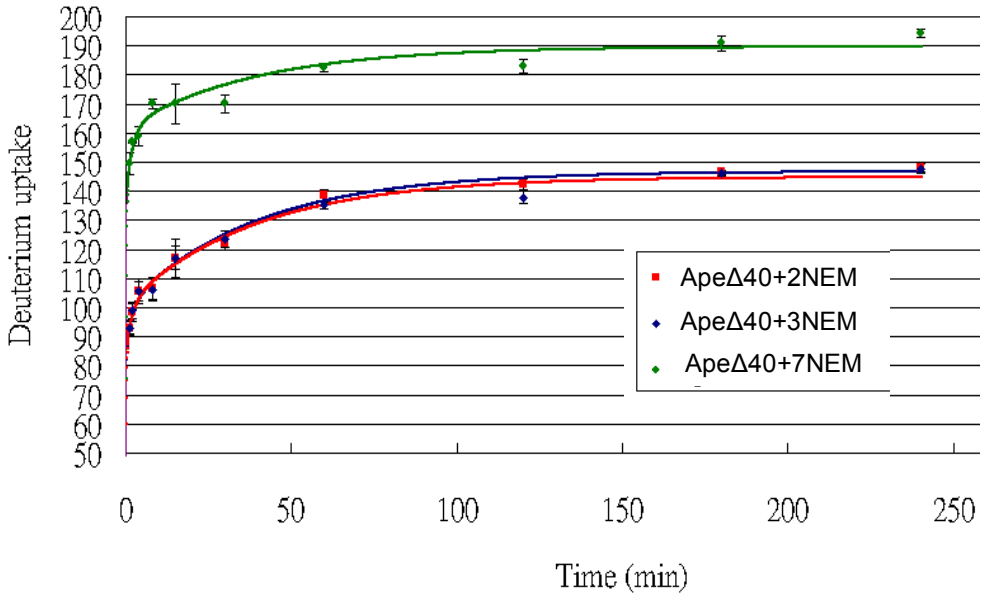


Figure 5.11. HDX kinetics results of "ApeΔ40 + 2 NEM", "ApeΔ40 + 3 NEM" and "ApeΔ40 + 7 NEM". ApeΔ40/E3330/NEM sample was incubated in 10 mM HEPES with 150mM KCl (pH 7.5) for 22 h at RT ([protein] = 100 μM; [E3330] = [NEM] = 500 μM). Reaction was quenched by adding DTT. HDX was conducted in a 93% D_2O medium with 10 mM HEPES (pH 7.5) and 150 mM KCl at 25 ºC and quenched by adding sufficient 1 M cold HCl to give a pH of 2.5.

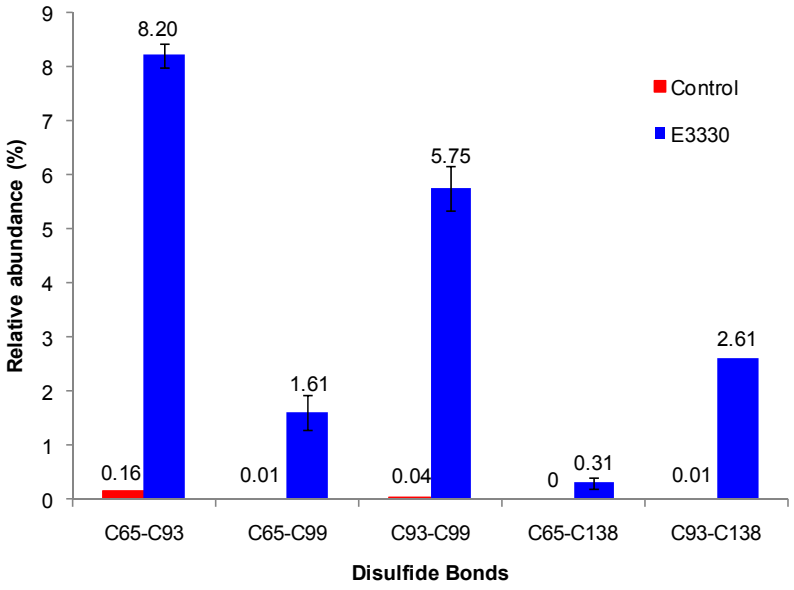


Figure 5.12. Normalized percentages of disulfide linkages in the ApeΔ40 control and ApeΔ40/E3330 samples. Samples were prepared by incubating the protein without (control) and with E3330 in 10 mM HEPES at pH 7.5 for 1 h at 37 ºC ([protein] = 10 µM; [E3330] = 50 µM). An excess of NEM (protein/NEM = 1/50, mol/mol) was immediately added to quench the disulfide crosslinking reaction and prevent disulfide scrambling during the subsequent digestion process. Data base searching was done with MassMatrix, an algorithm with the ability to identify disulfide-linked and crosslinked peptides (70-72). The amounts of peptides were estimated by dividing the peak areas in the LC chromatogram of the disulfide-linked peptides by that of a "standard" peptide "WDEAFR" in the sequence, which does not become modified during the sample preparation and LC-MS/MS analysis process. The assumption for this method is that there is no or little ionization discrimination between the "standard" peptide and the didulfide linked peptides. No C65-C138 linkage was observed in the control sample. Bars represent "mean ± standard derivation" taken from triplicate measurements.

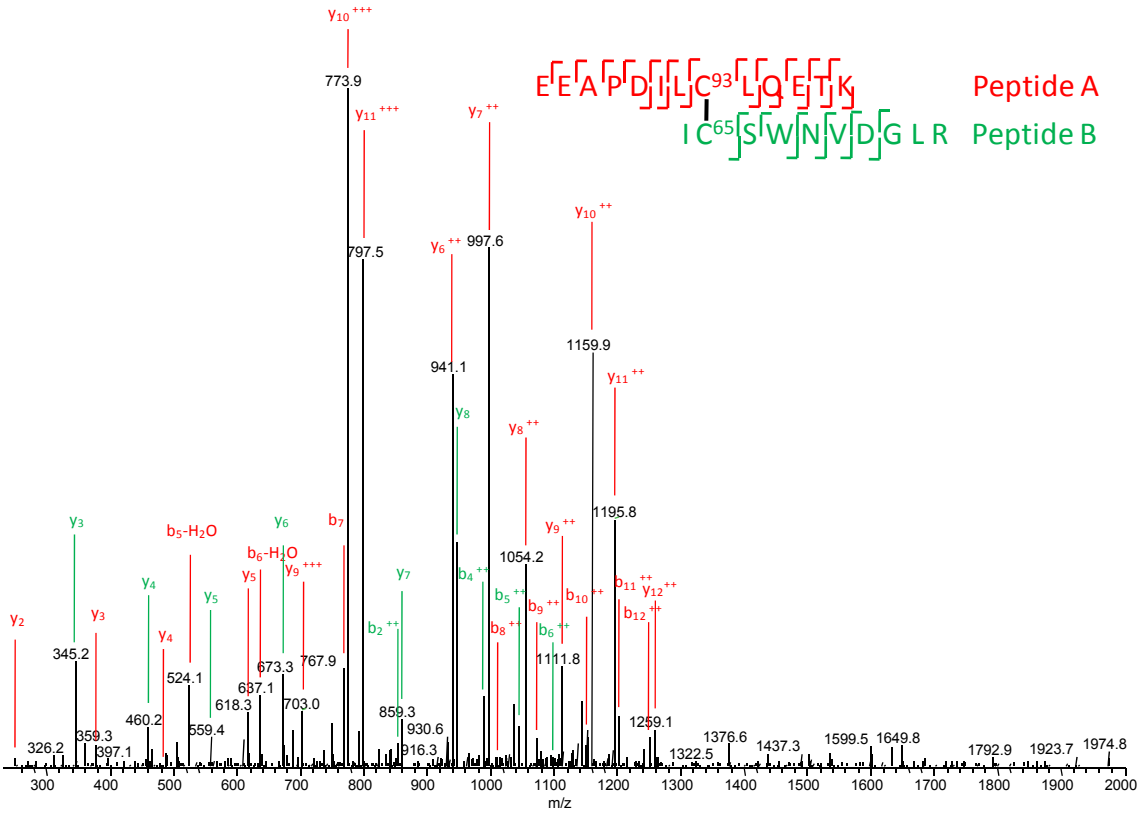


Figure 5.13. MS/MS spectrum that shows disulfide bond formation between C65 and C93. Product ions generated from peptide A are labeled in read and those from peptide B in green. The symbol ++ denotes doubly charged ions; the symbol +++ denotes triply charged ions.

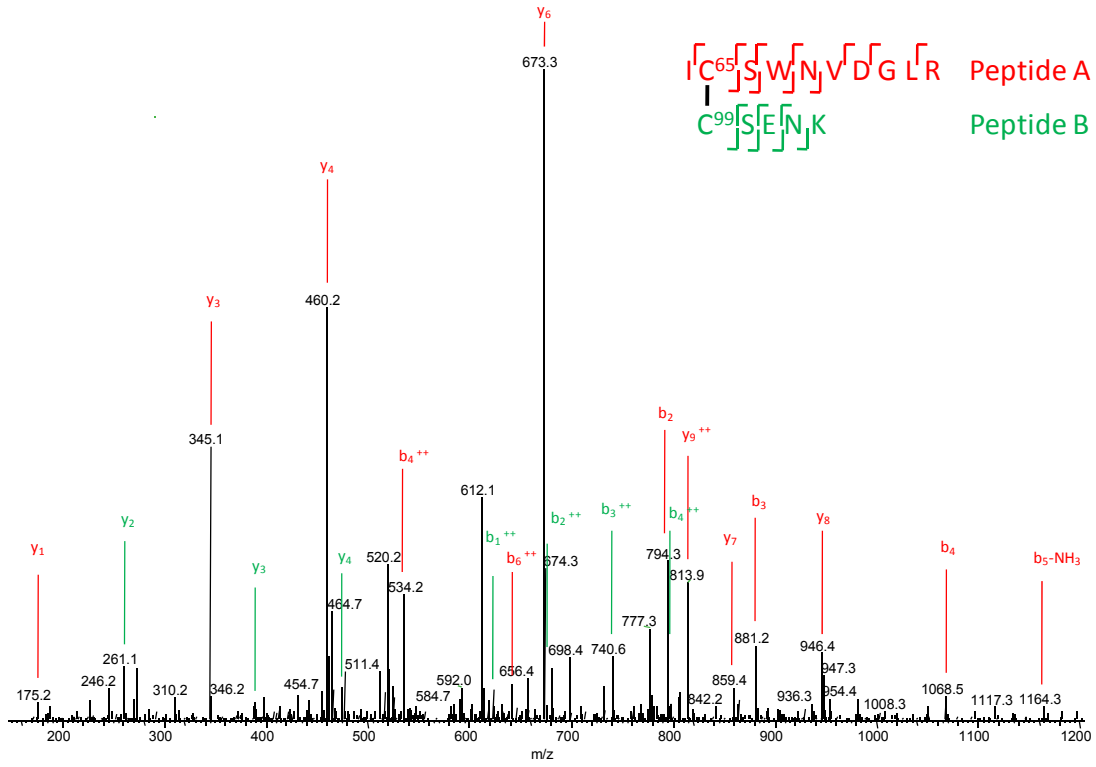


Figure 5.14. MS/MS spectrum that shows disulfide bond formation between C65 and C99. Product ions generated from peptide A are labeled in read and those from peptide B in green. The symbol ++ denotes doubly charged ions; the symbol +++ denotes triply charged ions.

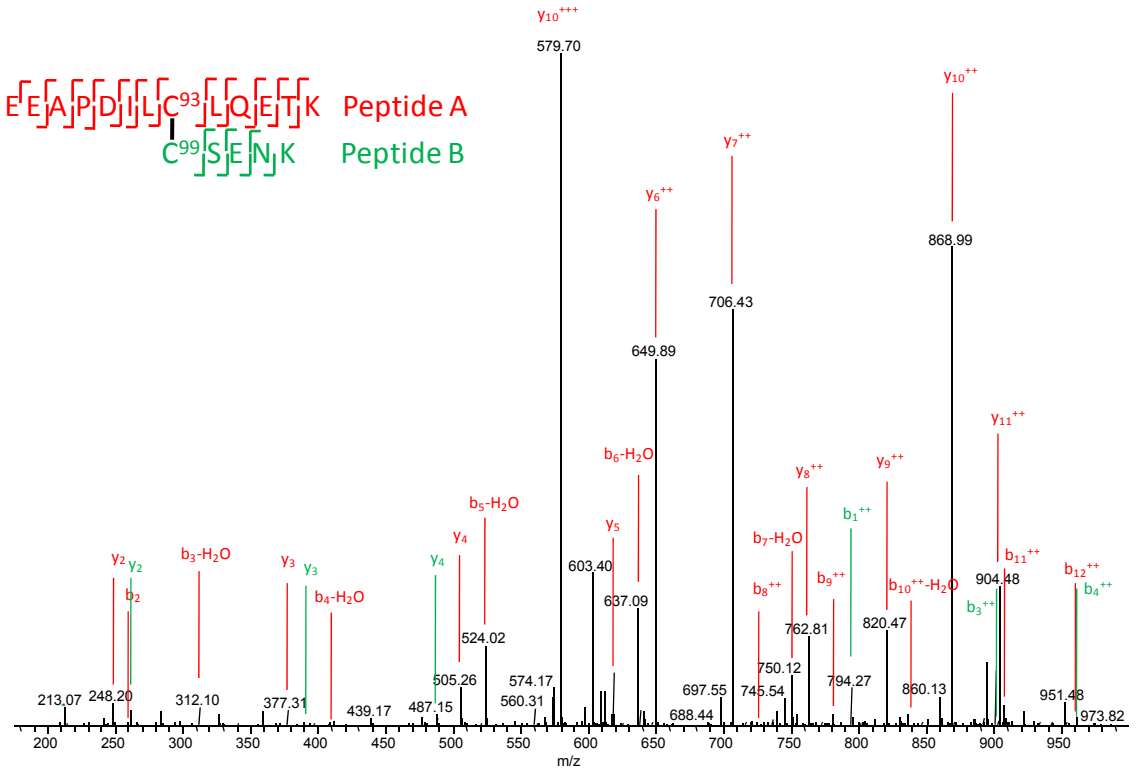


Figure 5.15. Product-ion spectrum (MS/MS) showing disulfide bond formation between C93 and C99. Product ions generated from peptide A are labeled in read and those from peptide B in green. The symbol ++ denotes doubly charged ions; the symbol +++ denotes triply charged ions.

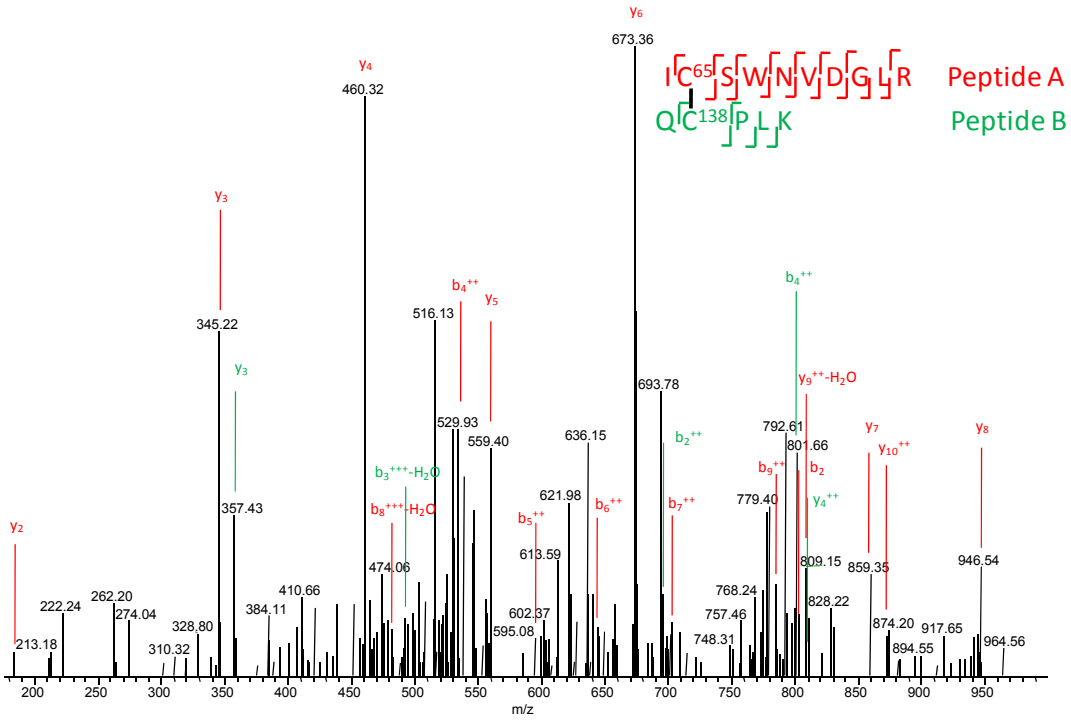


Figure 5.16. MS/MS spectrum that shows disulfide bond formation between C65 and C138. Product ions generated from peptide A are labeled in read and those from peptide B in green. The symbol ++ denotes doubly charged ions; the symbol +++ denotes triply charged ions.

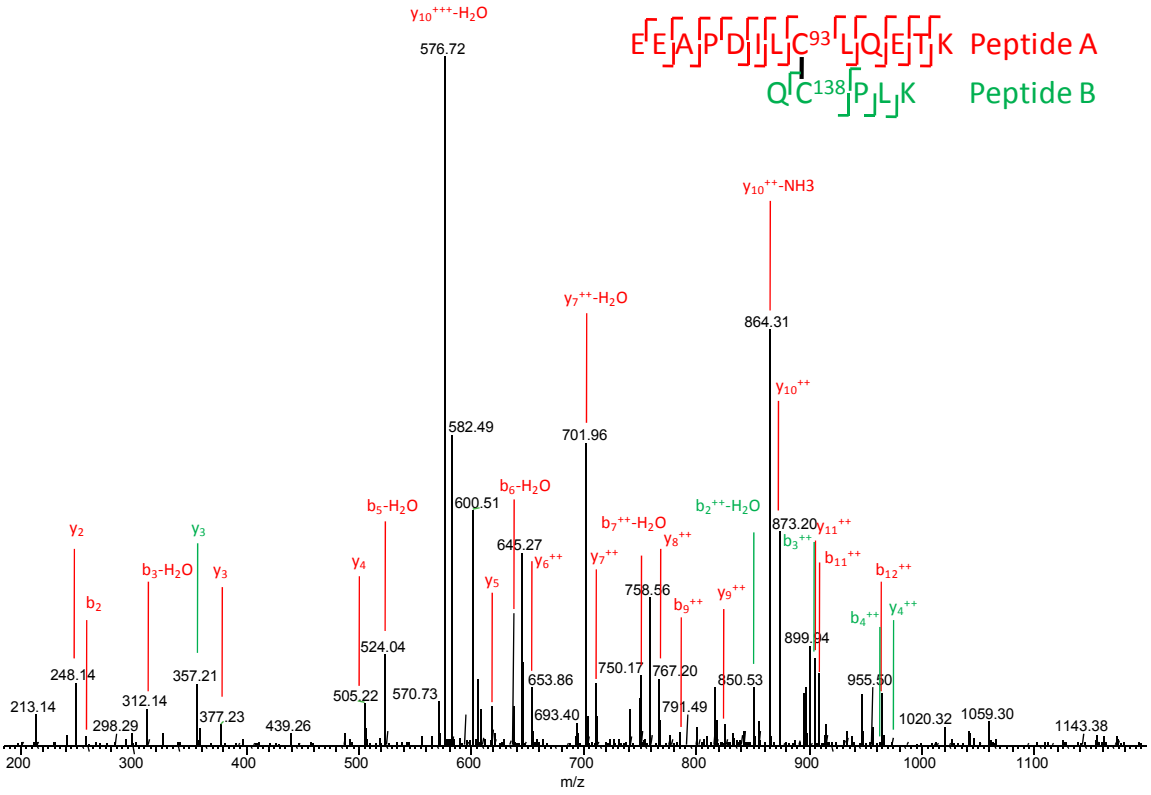


Figure 5.17. MS/MS spectrum that shows disulfide bond formation between C93 and C138. Product ions generated from peptide A are labeled in read and those from peptide B in green. The symbol ++ denotes doubly charged ions; the symbol +++ denotes triply charged ions.

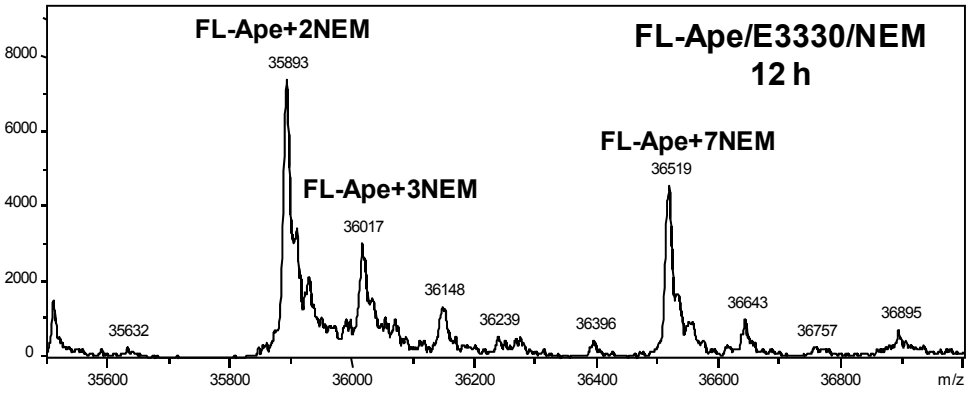
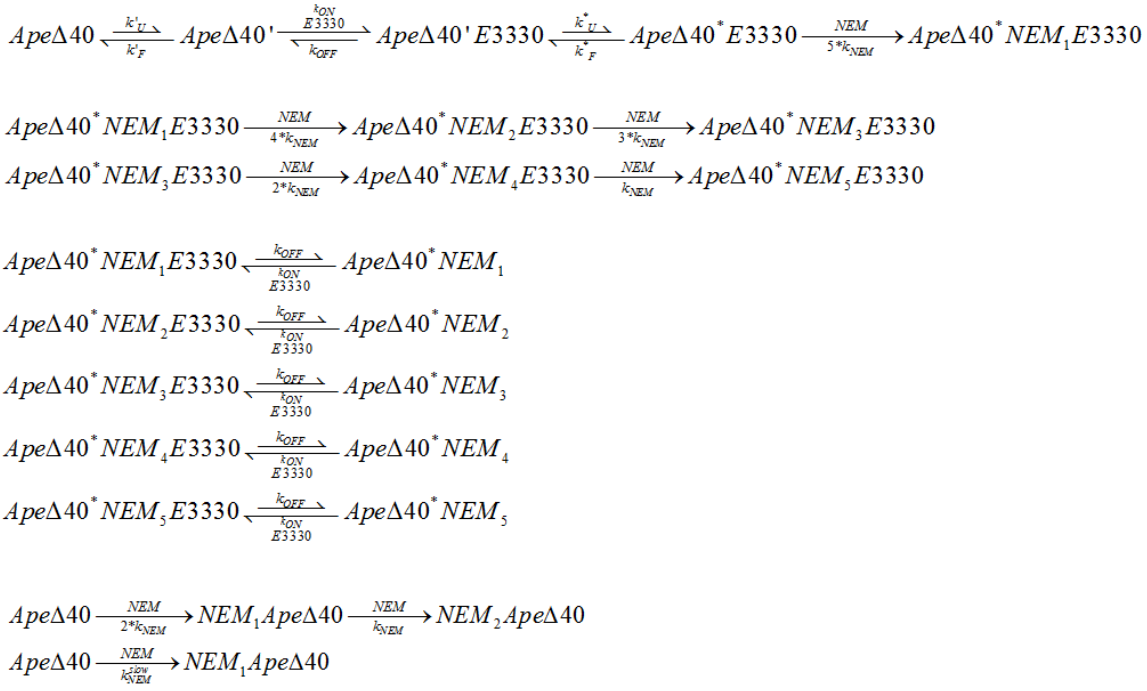


Figure 5.18. ESI mass spectrum of FL-Ape with E3330 in the presence of NEM for 12 h. Sample was incubated in 10 mM HEPES at pH 7.5([protein] = 100 μM; [E3330] = [NEM] = 500 μM). Mass spectrum was collected on a Bruker MaXis UHR-TOF instrument and deconvolution was done with MaxEnt1 algorithm.



Scheme 5.1. Proposed system equations explaining the NEM labeling of ApeΔ40 in the presence of E3330.