UNIVERSITY OF CALIFORNIA
RIVERSIDE

Mass Spectrometry and Density Functional Theory Characterizations of DNA Modifications

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

Doctor of Philosophy

in

Chemistry

by

Renee Therese Williams

March 2013

Dissertation Committee:
Dr. Yinsheng Wang, Chairperson
Dr. Cynthia Larive
Dr. Dallas Rabenstein

The Dissertation of Renee Therese Williams is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

COPYRIGHT ACKNOWLEDGEMENTS

The text and figures in Chapter 2, in part or in full, are a reprint of the material as it appears in *Clin. Chim. Acta* **2012**, X, XX-XX and the supporting information therein. The co-author, Dr. Yinsheng Wang, listed in that publication directed and supervised the research, which forms the basis of this chapter.

The text and figures in Chapter 4, in part or in full, are a reprint of the material as it appears in *Biochemistry* **2012**, 51, 6458-6462. The co-author, Dr. Yinsheng Wang, listed in that publication directed and supervised the research, which forms the basis of this chapter.

DEDICATION

This dissertation is dedicated to my parents, Derrick and Wanda, and to my sister, Erricka.

With family, strength, and perseverance anything is possible.

ABSTRACT OF THE DISSERTATION

Mass Spectrometry and Density Functional Theory Characterizations of DNA Modifications

by

Renee Therese Williams

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, March 2013
Dr. Yinsheng Wang, Chairperson

Modifications to deoxyribonucleic acids (DNA) may arise endogenously or be initiated by exogenous sources. In this dissertation, we focus on utilizing mass spectrometry (MS), liquid chromatography (LC), and density functional theory (DFT) methods to characterize DNA modifications induced by chemotherapeutic agents, and to assess the roles kinetics and thermodynamics play in the removal of oxidized pyrimidines by human thymine DNA glycosylase (hTDG).

In part I, we conducted an enzymatic digestion study coupled with detailed MS, tandem MS (MS/MS), and LC-MS/MS analyses of oligodeoxyribonucleotides (ODNs) containing cisplatin adducts (Chapter 2). We identified and characterized several digestion products for ODNs containing a 1,2-GpG, 1,2-ApG or 1,3-GpXpG intrastrand cross-link. These results built a solid foundation for the future LC-MS/MS quantification of these DNA lesions.

Additionally, we characterized the threshold energy of cross-link formation based upon the UVB irradiation of ODNs harboring a 5-halouracil or 5-halocytosine (Chapter 3). We found that the calculated threshold energies for 5-halocytosines were more exothermic than the corresponding 5-halouracils, and more exothermic for 5-iodopyrimidines than for the 5-bromopyrimidine counterparts, suggesting greater cross-link formation for the former pyrimidine derivatives. These calculation results were in excellent agreement with reported experimentally determined yields of the cross-link products and supported the mechanistic proposal that cross-link products form from UV-induced electron transfer followed by heterolytic cleavage of the C5-X bond.

In part II, we focused on the characterization of hTDG activity towards oxidized pyrimidine base modifications formed at methyl-CpG site. We used DFT methods to assess the kinetic and thermodynamic parameters for the cleavage of the *N*-glycosidic bond of 5-methyl-2'-deoxycytidine (mdC) and its, (1) oxidation products as it pertains to a newly proposed mechanism for active cytosine demethylation (Chapter 4); and (2) oxidation and subsequent deamination products [i.e., the 2'-deoxyuridine (dU) derivatives] (Chapter 5). Our studies suggest that the inherent chemistry associated with the nucleophilic cleavage of the *N*-glycosidic bond constitutes a major factor contributing to the selectivity of hTDG towards 5-substituted dC and dU derivatives. These findings provided novel chemical insights into the role TDG may play in maintaining the epigenetic integrity of the CpG site.

Table of Contents

List of Figures

List of Tables

List of Schemes

Chapter 1

Introduction

Modifications to deoxyribonucleic acids (DNA) may arise from both endogenous (i.e. epigenetic regulation, cellular metabolism, and replication errors) and exogenous (i.e., ultraviolet light, ionizing radiation, and environmental chemical mutagens) sources. Irrespective of the origin, it is important to characterize these modifications in order to understand better their adverse human health consequences. In this dissertation, we focus on utilizing mass spectrometry (MS) and liquid chromatography (LC)-based experimental methods, as well as computational approaches, to investigate how various chemotherapeutic agents damage DNA *in vitro*, specifically intrastrand cross-link formation, and to assess the role kinetics and thermodynamics play in the excision of 5-substituted pyrimidine nucleobases by thymine DNA glycosylase (TDG).

Background

Cancer and Chemotherapeutic Agents

The first known documentation of cancer was reported *circa* 3000-1500 BC by the ancient Egyptians in a series of medical papyri that described eight cases of cauterized breast tumors because "there [was] no treatment." Now, despite more than five millennia of acquired knowledge accredited to this disease, cancer still remains one of the leading causes of death.

Many factors contribute to the challenging nature of cancer treatment, and ultimately its cure. Perhaps most importantly, cancer is a disease that affects the entire collection of genetic material, thus rendering its understanding highly complex. In most cases, the how, why, and/or where of the onset and progression of this disease is poorly if not at all defined. Secondly, there are more than 200 distinct types of tumors, all of which exhibit different alterations in the genome. Thirdly, for a given type of cancer, there may be hundreds or even thousands of mutations, meaning no singular course of treatment will suffice. Instead, a variety of drugs (cancer chemotherapeutic agents) that target specific types of cancers are needed.

There are dozens of chemotherapeutic agents available on the market today. Their functions range from chemically altering DNA by the addition of alkyl groups (alkylating agents), imitating the role of DNA bases (nucleoside analogs), plant-derived cell division inhibitors (alkaloids and terpenoids; podophyllotoxin; and taxanes), enzyme inhibitors (topoisomerase), as well as hormones, and monoclonal antibodies. In order to discern the mechanism(s) of action for these drugs and for the development of future treatments, it is necessary to understand how they interact with biological molecules (i.e. DNA, proteins, and RNA). In this dissertation, we considered two classes of chemotherapy drugs, alkylating agents (i.e. cisplatin) and nucleoside analogs (i.e. 5-halopyrimidines), as it relates to their effect on DNA.

Cisplatin, $[PtCl_2(NH_3)_2]$, was first described in 1845 and aptly named Peyrone's salts given the origin of its discovery. In 1965, Barnett Rosenberg fortuitously reported its usefulness as an alkylating anti-tumor agent. Since its approval by the Food and Drug

Administration in 1978, cisplatin, or Platinol ®, is responsible for 90% of all cases of cured testicular cancer, and plays a pivotal role in the treatment of ovarian, head and neck, bladder, and cervical cancers, as well as melanoma and lymphomas.[1] However, its administration is not met without difficulties. The downside for patients treated with cisplatin are the severe side effects (i.e., nausea, vomiting, ototoxicity, and nephrotoxicity) and a strong propensity to develop drug resistance.[2] Given that the most prominent mode of action for this drug involves the generation of cross-link DNA adducts, it is necessary to have an analytical method that can identify, characterize, and quantify the DNA modifications in order to address these challenges.[3, 4] For this reason, we used MS and MS/MS to fully characterize synthetic DNA bearing cisplatin cross-link adducts (i.e. 1,2-GpG, 1,2-ApG, and 1,3-GpXpG), as well as liquid chromatography-tandem mass spectrometry (LC-MS/MS) to characterize in detail their respective enzymatic hydrolysis products. In doing so, this sets the stage for future LC-MS/MS quantification studies of these DNA lesions in cultured human cells or in patients undergoing cisplatin therapy.

5-Halopyrimidines, e.g. 5-fluorouracil (Adrucil®), has been used since the 1950s to treat breast, colon, esophageal, pancreatic, stomach, head and neck cancers. Similar to cisplatin, treatment with 5-halopyrimidines, when coupled with UV irradiation or ionizing radiation, yield intrastrand cross-link lesions between adjacent nucleobases.[5-11] These DNA modification, as well as others, have been implicated in the observed photosensitizing effect of halogenated nucleosides after their incorporation into cellular DNA.[12-37] Previously, Zeng et al.[10] proposed that cross-link formation occurs via

dissociative electron attachment (DEA) mechanism, which initiates from the UV-induced transfer of an electron from the neighboring purine base to the 5-halopyrimidine. The resulting radical anion of the halogenated pyrimidine base can then eliminate a halide to yield a pyrimidin-5-yl radical. The radical can subsequently couple with the C8 of the neighboring purine base, and the resultant products can lose an electron and deprotonate to yield the intrastrand cross-link d(Y[8-5]Z), where Y = G or A, and Z = U or C. In this dissertation, we employed density functional theory (DFT)[38] methods to assess the thermodynamics (threshold energy) of the photochemically induced DEA reaction. The work herein strengthens our mechanistic understanding of how intrastrand cross-links are formed.

Thymine DNA Glycosylase

Thymine DNA glycosylase (TDG) is an enzyme that functions as part of the base excision repair (BER) machinery to correct mismatches with guanine (G) in duplex DNA specifically at 5'-CpG-3' sites. Crystal structure analyses revealed that the sequence-specific mismatch recognition of TDG may be attributed to electrostatic interactions of the N1H and $N2H_2$ of the opposing G, relative to the damaged base, with the backbone amides of A274 and P280 within the enzyme.[39, 40] It was also demonstrated that human thymine DNA glycosylase (hTDG) has a relatively large, non-specific active site that will accommodate a wide variety of substrates; however, it exhibits remarkable selectivity for excision of thymine (T), uracil (U), and a number of 5-substituted pyrimidines.[41-44] In this vein, it was found that hTDG could readily excise 5-formylcytosine (fmC) and 5-carboxycytosine (caC) from mammalian DNA, but not 5-hydroxymethylcytosine (hmC)

or 5-methylcytosine (mC). [42, 45-47] These findings provided important evidence to support a newly proposed mechanism of active cytosine demethylation, which involves iterative oxidation of meC to hmC,[48, 49] fmC,[46] and caC[45, 46] by ten-eleven translocation (TET) 1-3 proteins, TDG-mediated excision of fmC and caC, and subsequent employment of BER machinery to yield unmethylated cytosine at CpG sites. Given the novelty of these findings and the general uncertainty concerning DNA replication-independent demethylation of cytosine in mammals, we used DFT[38] methods to provide the first computational study of the inherent chemistry associated with the selective cleavage of the *N*-glycosidic bond by hTDG in 5-substituted cytosine derivatives. Moreover, given that deamination of mC and its oxidation products (i.e. hmC, fmC, caC) can lead to the formation of corresponding thymine derivatives [i.e. thymine (T), 5-formyluracil (fmU), 5-carboxyluracil (caU)] at CpG sites,[50] we also explored the kinetics (activation energy) and thermodynamics (reaction energy) for the cleavage of the *N*-glycosidic bond in 5-substituted uracil derivatives. Together, these calculations help to shape a predictive model of varied hTDG activity towards a host of 5-substituted pyrimidine derivatives.

Mass Spectrometry-based Assay

The specificity, versatility, and sensitivity of mass spectrometry make it the ideal analytical tool for characterizing modified DNA bases. Modification notwithstanding, there are a number of factors that must be considered in order to construct an assay that yields useful results (Figure 1.1). In this section, we address two key features in sample preparation that were pertinent to the MS measurements made in Chapters 2.

Incorporation of Modified DNA Bases

When studying modified DNA by MS, it is routine to generate a model system *in vitro* in order to assess its behavior within the instrument. Additionally, the controlled nature of these experiments allows one to investigate the mechanism(s) of formation in an environment free from influences that are otherwise present in cellular and animal models. To this end, we introduced *cis*-diamminedichloroplatinum (II) (DDP) modifications into oligodeoxyribonucleotides (ODNs), where we incubated our ODN substrates with DDP and allowed the drug to selectively complex with the DNA to form cisplatin intrastrand cross-links, $[Pt(NH_3)_2]^{2+}/[Pt(^{15}NH_3)_2]^{2+}$. After incubation, it is necessary to purify the ODN by high performance liquid chromatography (HPLC) before verification of the modified sequence by MS and MS/MS, and quantification via UV absorption spectrophotometry.

Enzymatic Digestion

Hydrolysis of phosphodiester bonds in DNA by enzymatic digestion is a salient feature of these assays. The type and quantity of enzymes used affect directly the structure of the released modification-containing nucleotide (or nucleoside). Consequently, the digestion conditions also dictate which particular product(s) to monitor during LC-MS/MS analysis.

For the ODN digestion, we employed a four-enzyme cocktail consisting of nuclease P1 (NP1), calf spleen phosphodiesterase (exoPII), snake venom phosphodiesterase (exoPI), and calf intestinal alkaline phosphatase (CIP). We chose

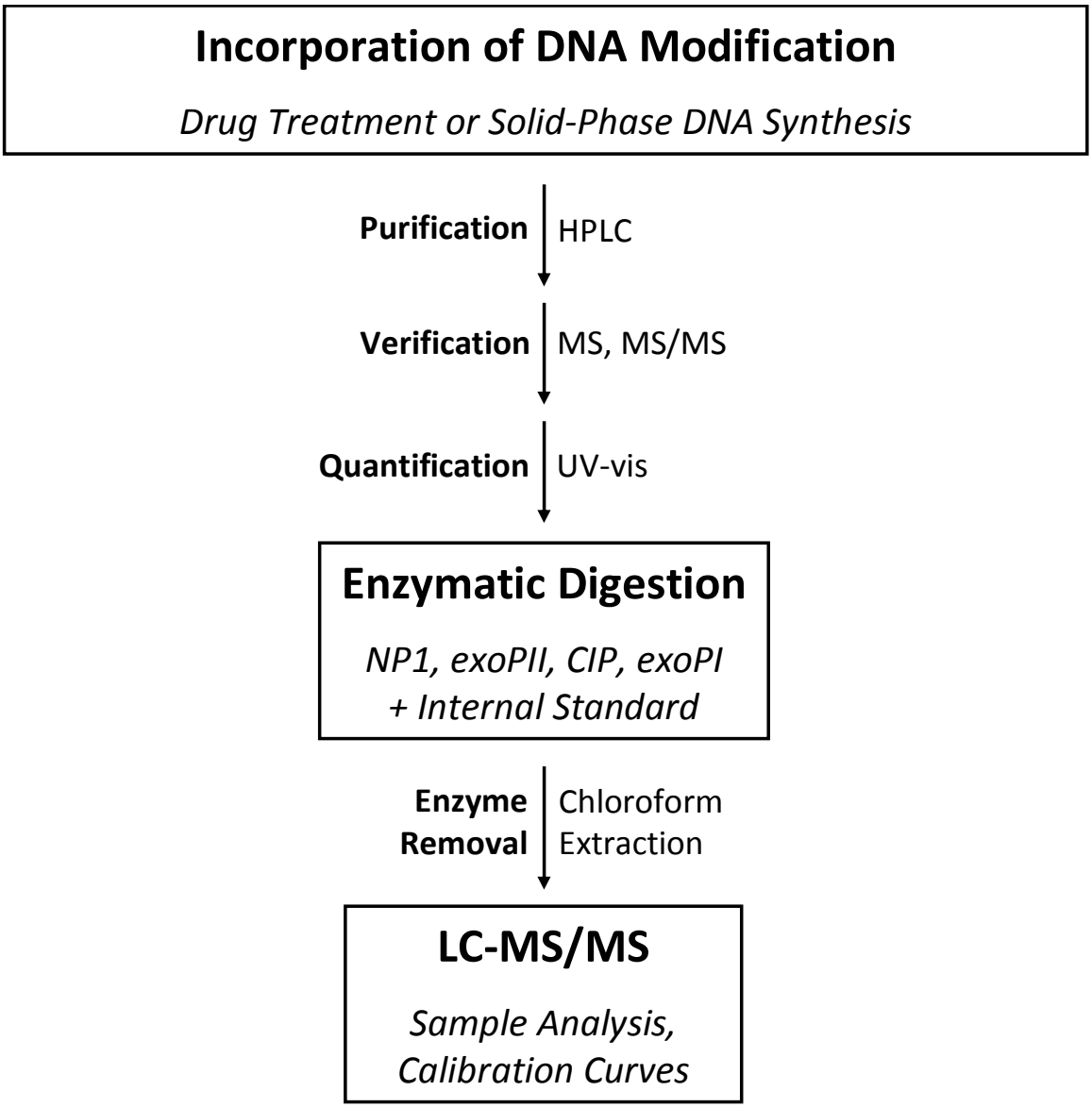


Figure 1.1 Work flow employed for the characterization of modified DNA bases by liquid chromatography tandem mass spectrometry method (LC-MS/MS).

these four enzymes because they have been used successfully for the release of cross-link product analogues to those studied in Chapter 3.[5, 10, 11, 51] However, this combination of digestion enzymes was unique to the investigation of cisplatin-modified DNA as described in Chapter 2.

Briefly, NP1, exoPII, and DNA were incubated together in an acidic environment to initiate the digestion reaction. NP1 is a nonspecific endonuclease that cleaves on the 3'-side (*a* side) of any internal phosphodiester bond, which results in a phosphate attached to the 5'-position of the adjacent nucleotide. ExoPII is a *b* specific exonuclease that works in the 5' → 3' direction. It hydrolytically removes terminal nucleotide residues by cleaving on the 5'-side of the phosphodiester bond, which leaves a phosphate attached to the 3'-position of the adjacent nucleotide. Upon treatment with NP1 and exoPII, the pH of the solution is then adjusted to basic conditions, followed with the addition of exoPI and CIP. Similar to NPI, exoPI creates nucleotide-5'-phosphates; however, it is an exonuclease that specifically works in the 3' → 5' direction. Ideally, the combination of NP1, exoP1, and exoPII results in a mixture of 3'- and 5'-nucleoside monophosphates that can be converted to nucleosides by CIP (Figure 1.2).

As discussed in Chapter 2, intrastrand cross-link modifications can block enzyme activity and yield nucleotide dimers that retain an internal phosphate. Cisplatin-containing modification (Chapter 2) are particularly disruptive to the digestion efficiency of the enzymes as evidenced by the variety of digestion products that have been reported.[52-63] For this reason, we also added $^{15}N_2$-labeled cisplatin-containing ODNs (10-mers) to the digestion mixture to serve as the internal standards that can correct for

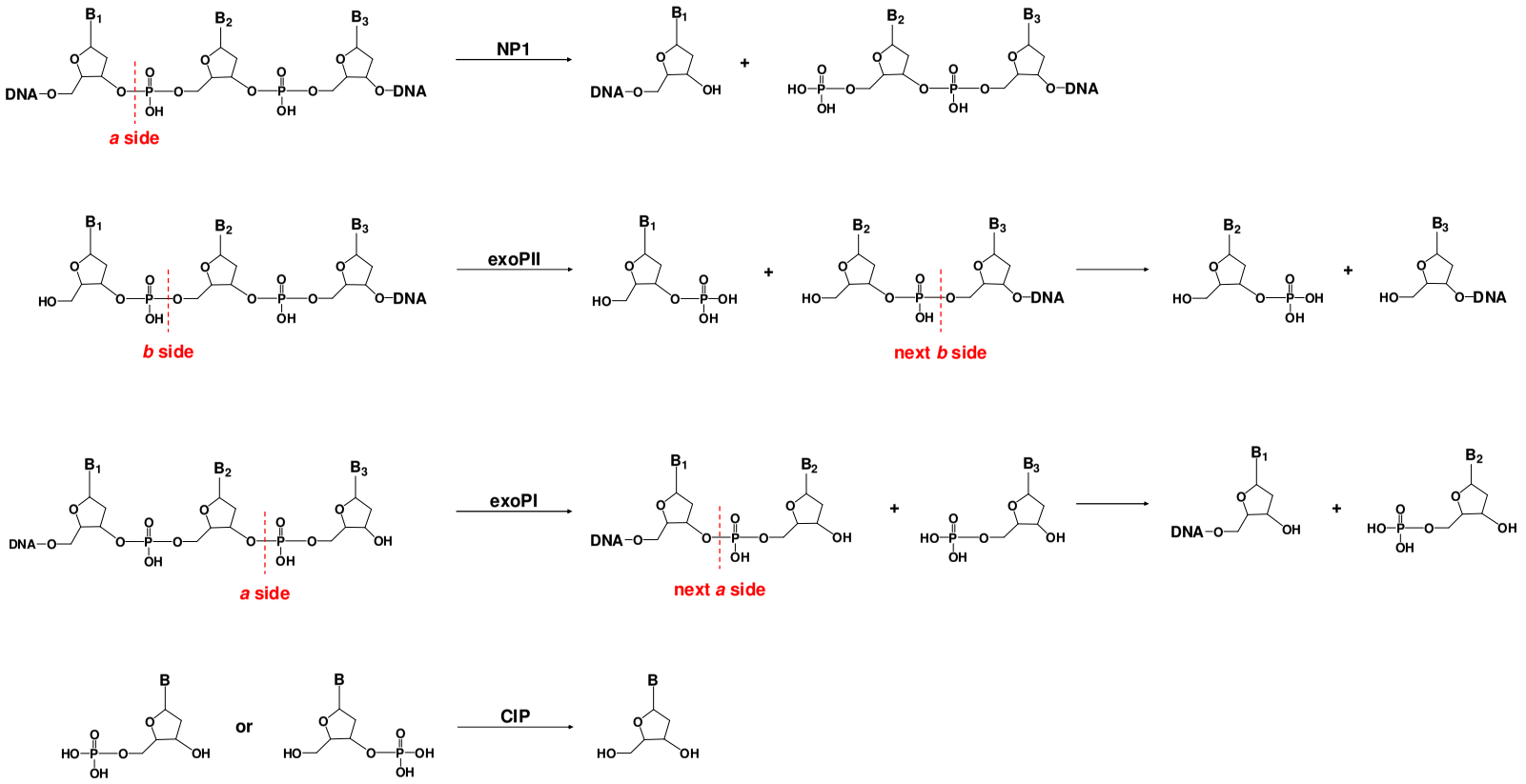


Figure 1.2 Four enzymes used in the hydrolysis of synthetic DNA as part of the MS-based assay for the characterization of modified DNA bases.

changes in digestion behavior across samples, as well as account for losses that occur during the removal of the digestion enzymes (chloroform extraction).

Density Functional Theory Methods

In recent decades, the application of computational methods, particularly when combined with laboratory experiments, has emerged as a crucial component of drug discovery and development. In this dissertation, we employed a variety of theoretical calculations as a means to explain findings made using cell models by our and other research groups (Chapters 3-5). In the following sections, we provide a general overview of the rationale behind each type of calculation.

Functional and Basis Set

All calculations herein utilized density functional theory (DFT) methods included in the Gaussian 09W package.[38] The fundamental validity of any DFT calculation lies in selecting the appropriate level of theory, or approximation methods, to describe a particular system. For DFT, this includes a functional and basis set.

A functional is defined as a function of electron density, which itself is a function of coordinates in real space. Specifically, a DFT functional partitions the electronic energy into components including kinetic energy, electron-nuclear interaction, Coulombic repulsion, and the exchange-correlation term, which are all computed separately.[64] Functionals are typically distinguished by the way they treat the exchange and correlation components. Local exchange and correlation functionals involve only the values of the electron spin densities, whereas gradient-corrected (non-local) functionals

involve both the values of the electron spin densities and their gradients. In our calculations, we used the well-proven Becke, 3-parameter, Lee-Yang-Parr (B3LYP), which is a hybrid of local and non-local exchange and correlation functionals.[38]

Our selection of basis set, that is, the mathematical representation of the molecular orbitals within the molecule, was sensitive to the individual nuclei represented in the calculations. In Chapter 3, the measurements involved anions as well as iodine-containing molecules, both of which require specific parameters within the basis set to generate a good approximation for the resulting energy. For anions, it is necessary to incorporate a diffuse function into the basis set to create large-size versions of s- and p-type functions that can accommodate the negative charge brought about by the addition of an electron. The diffuse function also benefits the iodine- (and bromine-) containing molecules whose valence electrons are relatively far from the nucleus. Additionally, large charged molecules require a polarized basis set that increases the angular momentum beyond what is required for the ground state in order to allow an orbital to change size and shape as needed. Due to its size, very few basis sets are available that include iodine, and even less that contain diffuse and polarized functions. Based on these needs, we selected the LANL2DZdp basis set which works well with the B3LYP functional.[65]

In Chapters 4 and 5, we again included anions in our calculations, which require the same basis set parameters as previously discussed. However, the atomic nuclei in these experiments were relatively small (i.e. hydrogen, carbon, nitrogen, and oxygen); therefore, the availability of appropriate basis sets was greatly improved. Based on previous reports,[66, 67] we selected the 6-31+G(d) and 6-311+G(2d,p) basis sets for

geometry optimization and energy calculations, respectively. 6-31+(d) is a small, polarized, (d), basis set with a diffuse function, +, applied to heavy atoms. For this system, it was shown to yield relative geometries in excellent agreement with those obtained using larger basis sets. When calculating basis set superposition error (BSSE) and single-point energies, the larger 6-311+G(2d,p) basis set with higher angular momentum, (2d,p), was employed.

Geometry Optimization

All theoretical calculations (Chapters 3-5) began with locating the stationary point, that is, the optimized geometry for the molecules of interest. For the majority of our calculations, the stationary point represents a minimum where all of the second derivatives are positive. A noted exception is when we locate transition states (TS, Chapters 4 and 5), or in other words, first-order saddle points along the potential energy surface in which the second derivative is negative in one direction and positive in all others. By comparison, locating the saddle point(s) for a function is far more difficult than determining its minima.

We employed Synchronous Transit-guided Quasi-Newton (STQN) methods to optimize the TS structures (Figure 1.3). STQN first initiates the linear synchronous transit (LST) method, which linearly varies the coordinates between reactant (R) and product (P) without geometry optimization. This assumes that all variables change in the same way following the reaction path; thus, the transition state is simply the highest energy structure along the interpolation line. Once this maximum is found, optimization

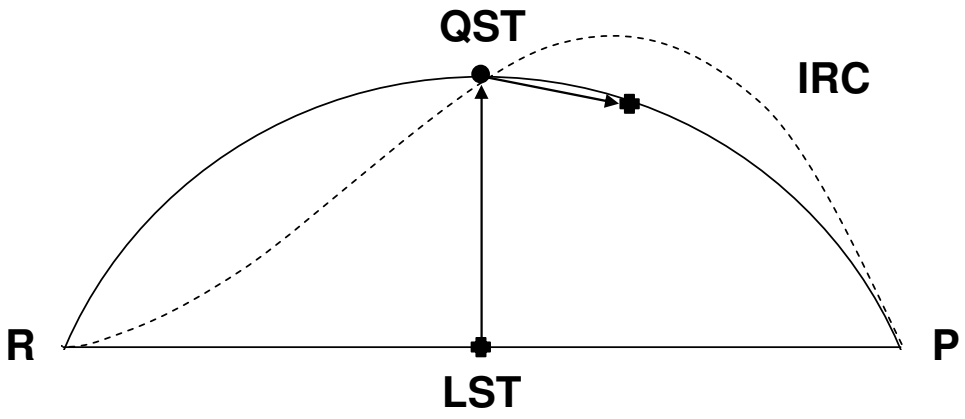


Figure 1.3 Illustration of the linear synchronous transit (LST), quadratic synchronous transit (QST), and intrinsic reaction coordinate (IRC) methods, where R represents the reactant and P is the product. Energy maxima and minima are denoted by a ✚ and •, respectively.

Image recreated from Frank Jensen, Introduction to Computational Chemistry, 2^{ed}, Wiley & Sons, West Sussex, England, page396.

switches to quadratic synchronous transit (QST) in which the current structure is minimized in the direction perpendicular to the LST path. When the structure is near the TS region, indicated by one negative eigenvalue (negative frequency), the method is switched to STQN to complete the optimization. Interpolation now follows a circle arc oppose to a straight line (LST) or a parabola (QST) by using tangents to the circle for guiding the search towards the TS. It should be noted that interpolation methods such as LST, QST, and STQN, in which the reactant and product geometries are known, do not actually find the TS, but rather, a point proximal to it. An intrinsic reaction coordinate (IRC) calculation may be used to find the exact TS; however, a local method such as this requires an excellent approximation of the TS structure, and the exact reactant and product structures are not known. Use of the interpolation methods better served our needs.

Frequency

Whether a minimum or TS, when an optimization calculation has converged, the resulting geometry yields a frozen nuclei energy located at the bottom of the potential well where the temperature is equal to 0 K. Frequency calculations must be performed, at the same level of theory as geometry optimization, in order to raise this energy to where the vibrational frequency (ν) equals zero. The difference between the frozen nuclei energy and $\nu = 0$ is termed the zero point energy (ZPE).

Basis Set Superposition Error

When two molecules are infinitely far apart, as is the case with the reactants (R) and products (P) shown in Figure 1.4, there is no interaction between their basis functions (atomic orbitals). However, as the molecules move closer, their atomic orbitals become available to each other thus creating a larger basis set, which can have a profound effect on the calculated energy. As illustrated in Figure 1.4, this phenomenon, termed basis set superposition error (BSSE), must be accounted for when locating the minima of a reactant- (R_{TS}) or product-complex (P_{TS}), and the saddle point of a transition state structure (TS). In theory, one could use a sufficiently large basis set so that extra functions from nearby molecules would have no significant effect. In practice, the basis set would need to be so large that it would be computationally expensive and unrealistic. Instead, counterpoise may be used to generate approximate correction values for BSSE.

Counterpoise corrects for weak interactions (i.e. van der Waals and hydrogen bonding) that contribute to BSSE by creating "ghost orbitals" for each molecule. These orbitals have no effect from atomic nuclei or additional electrons because they do not contain any atoms. The ghost orbitals are defined in fragments and positioned where they would be accompanied by atoms of other molecules in proximity. Ultimately, counterpoise raises the energy of the complex structures.

Polarizable Continuum Model

Polarizable continuum model (PCM) is used to predict the solvent effects on a given system. This is achieved by creating a solute (van der Waals) cavity of interlocking spheres around each nucleus, which excludes solvent molecules. The

molecular free energy of the solute is then calculated as the sum of the electrostatic and dispersion-repulsion contributions, and cavitation energy, using the dielectric constant (ε) of the solvent of interest. In this dissertation, PCM calculations for water (ε = 78; Chapter 3) and diethylether (ε = 4; Chapter 5) were performed.

Scope of this Dissertation

In part I of this dissertation (Chapters 2 and 3), we focused on chemotherapeutic agent-induced modifications of DNA. Specifically, we implement a MS-based assay for the characterization of intrastrand cross-linked ODNs (Chapter 2) and DFT methods for the prediction of the tendency for different halogenated pyrimidine bases to induce the formation intrastrand cross-link products (Chapter 3). In Chapter 2, we provided the detailed MS and MS/MS analyses of ODNs housing three cisplatin-DNA intrastrand cross-link lesions, and their respective digestion products by LC-MS/MS. To our knowledge, this represents the first complete MS/MS characterization of the intact ODNs harboring an adjacent adenine-guanine- and a nonadjacent guanine-guanine-cross-link, as well as the majority of the hydrolysis products.

In Chapter 3, we employ DFT methods to assess the thermodynamics (threshold energy) for the dissociative electron attachment of 5-X-pyrimidines where X = F, Cl, Br, and I, as it pertains to the mechanism of formation. The combined theoretical (this dissertation) and experimental (other published studies) data support the proposed mechanism that cross-link formation involves electron transfer from a nearby purine base followed by heterolytic cleavage of the C5-X bond, and thus strengthens our mechanistic

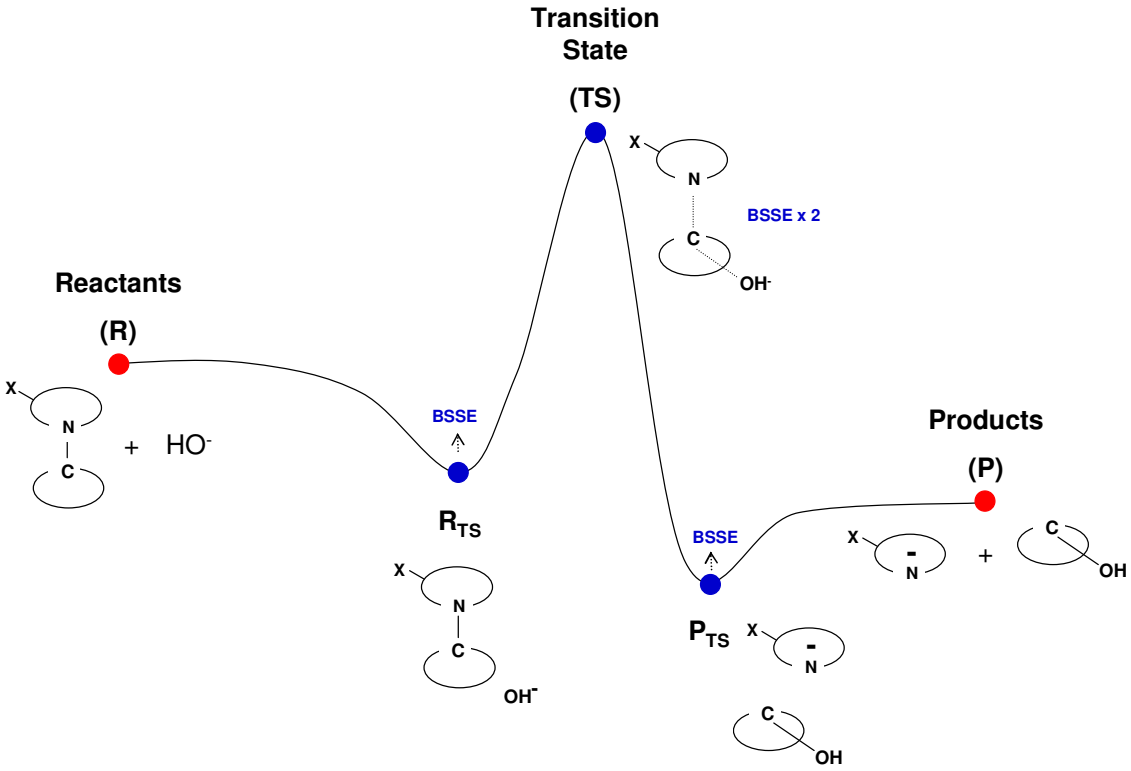


Figure 1.4 Schematic of the reaction coordinates calculated in Chapters 4 and 5, which illustrates the use of basis set superposition error (BSSE).

understanding of how intrastrand cross-links are formed.

In Part II of this dissertation (Chapters 4 and 5), we focused on employing DFT methods to characterize the removal of endogenously produced DNA modifications by hTDG at the epigenetically relevant CpG site. In Chapter 4, we assessed the kinetics (activation energy), thermodynamics (reaction energy), molecular geometry (*N*-glycosidic bond length), and enthalpy of deprotonation (acidity) associated with the hydroxide-mediated cleavage of the *N*-glycosidic bond from thymidine (dT) and 5-X-2'-deoxycytidines (XdC), where X = -H (dC), -CH_3 (medC), -CH_2OH (hmdC), -CHO (fmdC), -COOH (cadC), -F (FdC), and -Br (BrdC), as it relates to a newly proposed mechanism for active cytosine demethylation.[42, 45-49, 68] This work is the first to provide computational evidence to suggest active cytosine demethylation may occur through iterative oxidation of meC to fmC and caC by the ten-eleven translocation 1-3 proteins, followed by TDG -mediated excision of fmC and caC, and subsequent employment of base excision repair machinery to yield unmethylated cytosine.

In Chapter 5, we applied the same methods to 5-substituted 2'-deoxyuridine derivatives [XdU, where X = -H (dU), -CH_3 (dT), -CH_2OH (hmdU), -CHO (fmdU), and -COOH (cadU)]. We demonstrated that fmdU and cadU could be highly favorable substrates for hTDG and potentially excised from duplex DNA more readily than hmdU. Moreover, we establish that the assessment of the kinetics, thermodynamics, and molecular geometry involved in the cleavage of the *N*-glycosidic bond provides a more accurate prediction of varied hTDG activity towards a host of 5-substituted pyrimidine derivatives.

Finally, concluding remarks and future directions for the research presented herein are discussed in Chapter 6.

References

(1) Alderden, R. A.; Hall, M. D.; Hambley, T. W. *J. Chem. Educ.* **2006**, *83*, 728.

(2) Kartalou, M.; Essigmann, J. M. *Mutat. Res.* **2001**, *478*, 23-43.

(3) Jamieson, E. R.; Lippard, S. J. *Chem. Rev.* **1999**, *99*, 2467-2498.

(4) Siddik, Z. H. *Oncogene* **2003**, *22*, 7265-7279.

(5) Hong, H.; Wang, Y. *J. Am. Chem. Soc.* **2005**, *127*, 13969-13977.

(6) Saito, I. *Pure Appl. Chem.* **1992**, *64*, 1305-1310.

(7) Watanabe, T.; Bando, T.; Xu, Y.; Tashiro, R.; Sugiyama, H. *J. Am. Chem. Soc.* **2005**, *127*, 44-45.

(8) Watanabe, T.; Tashiro, R.; Sugiyama, H. *J. Am. Chem. Soc.* **2007**, *129*, 8163-8168.

(9) Zeng, Y.; Wang, Y. *J. Am. Chem. Soc.* **2004**, *126*, 6552-6553.

(10) Zeng, Y.; Wang, Y. *Nucleic Acids Res.* **2006**, *34*, 6521-6529.

(11) Zeng, Y.; Wang, Y. *Biochemistry* **2007**, *46*, 8189-8195.

(12) Brust, D.; Feden, J.; Farnsworth, J.; Amir, C.; Broaddus, W. C.; Valerie, K. *Cancer Gene Ther.* **2000**, *7*, 778-788.

(13) Cadet, J.; Vigny, P. *Bioinorganic photochemistry*, In Morrison, H. ed.; John Wiley: New York, 1990.

(14) Chu, E. H. Y. *Mutat. Res. Fund. Mol. Mech. Mut.* **1965**, *2*, 75-94.

(15) Dewey, W. C.; Humphrey, R. M. *Radiat. Res.* **1965**, *26*, 538-553.

(16) Erikson, R. L.; Szybalski, W. *Cancer Res.* **1963**, *23*, 122-130.

(17) Erikson, R. L.; Szybalski, W. *Radiat. Res.* **1963**, *20*, 252-262.

(18) Hale, J. T.; Bigelow, J. C.; Mathews, L. A.; McCormack, J. J. *Biochem. Pharmacol.* **2002**, *64*, 1493-1502.

(19) Heidelberger, C.; Chaudhuri, N. K.; Danneberg, P.; Mooren, D.; Griesbach, L.; Duschinsky, R.; Schnitzer, R. J.; Pleven, E.; Scheiner, J. *Nature* **1957**, *179*, 663-666.

(20) Hutchinson .F *Q. Rev. Biophys.* **1973**, *6*, 201-246.

(21) Hutchinson F. and Kohnlein, W. *Prog. Mol. Subcell. Biol.* **1980**, *7*, 1-42.

(22) Ikushima, T.; Wolff, S. *Exp. Cell Res.* **1974**, *87*, 15-19.

(23) Kaysen, J.; Spriggs, D.; Kufe, D. *Cancer Res.* **1986**, *46*, 4534-4538.

(24) Lett, J. T.; Parkins, G.; Alexader, P.; Ormerod, M. G. *Nature* **1964**, *203*, 593-596.

(25) Ling, L. L.; Ward, J. F. *Radiat. Res.* **1990**, *121*, 76-83.

(26) Mekras, J. A.; Boothman, D. A.; Perez, L. M.; Greer, S. *Cancer Res.* **1984**, *44*, 2551-2560.

(27) Perez, L. M.; Greer, S. *Int. J. Radiat. Oncol., Biol., Phys.* **1986**, *12*, 1523-1527.

(28) Perez, L. M.; Mekras, J. A.; Briggle, T. V.; Greer, S. *Int. J. Radiat. Oncol., Biol., Phys.* **1984**, *10*, 1453-1458.

(29) Ribas, M.; Korenberg, J. R.; Peretti, D.; Pichiri, G.; Stockert, J. C.; Gosalvez, J.; Mezzanotte, R. *Chromosome Res.* **1994**, *2*, 428-438.

(30) Russell, K. J.; Rice, G. C.; Brown, J. M. *Cancer Res.* **1986**, *46*, 2883-2887.

(31) Sano, K.; Hoshino, T.; Nagai, M. *J. Neurosurg.* **1968**, *28*, 530-538.

(32) Santos, O.; Perez, L. M.; Briggle, T. V.; Boothman, D. A.; Greer, S. B. *Int. J. Radiat. Oncol., Biol., Phys.* **1990**, *19*, 357-365.

(33) Stahl, F. W.; Crasemann, J. M.; Okun, L.; Fox, E.; Laird, C. *Virology* **1961**, *13*, 98-104.

(34) Taguchi, T.; Shiraishi, Y. *Mutat. Res.* **1989**, *211*, 43-49.

(35) Webb, C. F.; Jones, G. D. D.; Ward, J. F.; Moyer, D. J.; Aguilera, J. A.; Ling, L. L. *Int. J. Radiat. Biol.* **1993**, *64*, 695-705.

(36) Wojcik, A.; von Sonntag, C.; Obe, G. *J. Photochem. Photobiol. B* **2003**, *69*, 139-144.

(37) Zamenhof, S.; De Giovanni, R.; Greer, S. *Nature* **1958**, *181*, 827-829.

(38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A. ; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford CT, 2010.

(39) Maiti, A.; Morgan, M. T.; Drohat, A. C. *J. Biol. Chem.* **2009**, *284*, 36680-36688.

(40) Maiti, A.; Morgan, M. T.; Pozharski, E.; Drohat, A. C. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 8890-8895.

(41) Bennett, M. T.; Rodgers, M. T.; Hebert, A. S.; Ruslander, L. E.; Eisele, L.; Drohat, A. C. *J. Am. Chem. Soc.* **2006**, *128*, 12510-12519.

(42) Maiti, A.; Drohat, A. C. *J. Biol. Chem.* **2011**, *286*, 35334-35338.

(43) Neddermann, P.; Jiricny, J. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 1642-1646.

(44) Wiebauer, K.; Jiricny, J. *Nature* **1989**, *339*, 234-236.

(45) He, Y.-F.; Li, B.-Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L.; Sun, Y.; Li, X.; Dai, Q.; Song, C.-X.; Zhang, K.; He, C.; Xu, G.-L. *Science* **2011**, *333*, 1303-1307.

(46) Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. *Science* **2011**, *333*, 1300-1303.

(47) Pfaffeneder, T.; Hackner, B.; Truß, M.; Münzel, M.; Müller, M.; Deiml, C. A.; Hagemeier, C.; Carell, T. *Angew. Chem., Int. Ed.* **2011**, *50*, 7008-7012.

(48) Ito, S.; D'Alessio, A. C.; Taranova, O. V.; Hong, K.; Sowers, L. C.; Zhang, Y. *Nature* **2010**, *466*, 1129-1133.

(49) Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A. *Science* **2009**, *324*, 930-935.

(50) Privat, E.; Sowers, L. C. *Chemical Research in Toxicology* **1996**, *9*, 745-750.

(51) Bellon, S.; Gasparutto, D.; Saint-Pierre, C.; Cadet, J. *Org. Biomol. Chem.* **2006**, *4*, 3831-3837.

(52) Da Col, R.; Silvestro, L.; Baiocchi, C.; Giacosa, D.; Viano, I. *J. Chromatogr., A* **1993**, *633*, 119-128.

(53) Eastman, A. *Biochemistry* **1983**, *22*, 3927-3933.

(54) Eastman, A. *Biochemistry* **1986**, *25*, 3912-3915.

(55) Fichtinger-Schepman, A. M. J.; Van der Veer, J. L.; Den Hartog, J. H. J.; Lohman, P. H. M.; Reedijk, J. *Biochemistry* **1985**, *24*, 707-713.

(56) Fichtinger-Schepman, A. M. J.; van Oosterom, A. T.; Lohman, P. H. M.; Berends, F. *Cancer Res.* **1987**, *47*, 3000-3004.

(57) Gupta, R.; Beck, J. L.; Sheil, M. M.; Ralph, S. F. *J. Inorg. Biochem.* **2005**, *99*, 552-559.

(58) Harrington, C. F.; Le Pla, R. C.; Jones, G. D. D.; Thomas, A. L.; Farmer, P. B. *Chem. Res. Toxicol.* **2010**, *23*, 1313-1321.

(59) Iijima, H.; Patrzyc, H. B.; Dawidzik, J. B.; Budzinski, E. E.; Cheng, H. C.; Freund, H. G.; Box, H. C. *Anal. Biochem.* **2004**, *333*, 65-71.

(60) Morrison, J. G.; Bissett, D.; Stephens, I. F. D.; McKay, K.; Brown, R.; Graham, M. A.; Fichtingerschepman, A. M.; Kerr, D. J. *Int. J. Oncol.* **1993**, *2*, 33-37.

(61) Plooy, A. C. M.; Fichtinger-Schepman, A. M. J.; Schutte, H. H.; van Dijk, M.; Lohman, P. H. M. *Carcinogenesis* **1985**, *6*, 561-566.

(62) Pluim, D.; Maliepaard, M.; van Waardenburg, R.; Beijnen, J. H.; Schellens, J. H. M. *Anal. Biochem.* **1999**, *275*, 30-38.

(63) Schneider, S.; Reißner, T.; Ziv, O.; Livneh, Z.; Carell, T. *ChemBioChem* **2010**, *11*, 1521-1524.

(64) Foresman, J. B.; Frisch, Æ. *Exploring chemistry with electronic structure methods*, 2 ed.; Gaussian, Inc.: Pittsburgh, 1996.

(65) Check, C. E.; Faust, T. O.; Bailey, J. M.; Wright, B. J.; Gilbert, T. M.; Sunderlin, L. S. *J. Phys. Chem. A* **2001**, *105*, 8111-8116.

(66) Millen, A. L.; Archibald, L. A. B.; Hunter, K. C.; Wetmore, S. D. *J. Phys. Chem. B* **2007**, *111*, 3800-3812.

(67) Millen, A. L.; Wetmore, S. D. *Can. J. Chem.* **2009**, *87*, 850-863.

(68) Wu, H.; Zhang, Y. *Genes Dev.* **2011**, *25*, 2436-2452.

Chapter 2

Nuclease Digestion and Mass Spectrometric Characterization of Oligodeoxyribonucleotides Containing 1,2-GpG, 1,2-ApG, and 1,3-GpXpG Cisplatin Intrastrand Cross-links

Abstract

The primary mode of action for *cis*-diamminedichloroplatinum (II), referred to as cisplatin, towards the treatment of solid malignancies is through formation of cross-links with DNA at purine sites, especially guanines. We prepared oligodeoxyribonucleotides (ODNs) containing a 1,2-GpG, 1,2-ApG, or 1,3-GpXpG cisplatin intrastrand cross-link and the corresponding ODNs modified with $^{15}N_2$-labeled cisplatin, and characterized these ODNs with electrospray ionization mass spectrometry (ESI-MS) and tandem MS (MS/MS). We also employed LC-MS/MS to characterize the digestion products of these ODNs after treatment with a cocktail of 4 enzymes (nuclease P1, phosphodiesterases I and II, and alkaline phosphatase). Our results showed that 1,2-GpG was released from the ODNs as a dinucleoside monophosphate or a dinucleotide. Analyses of the digestion products of ODNs containing a 1,2-GpG cross-link on the 5' or 3' terminus revealed that the dinucleotide carries a terminal 5' phosphate. On the other hand, digestion of the 1,3-GpXpG intrastrand cross-link yielded 3 dinucleoside products with 0, 1, or 2 phosphate groups. The availability of the ODNs carrying the stable isotope-labeled lesions, MS/MS analyses of the cisplatin-modified ODNs, and the characterization of the enzymatic

digestion products of these ODNs set the stage for the future LC-MS/MS quantification of the 1,2-GpG, 1,2-ApG, and 1,3-GpXpG lesions in cellular DNA.

Introduction

Cancer chemotherapy was revolutionized with the serendipitous discovery of the antitumor properties of *cis*-diamminedichloroplatinum (II) (cisplatin, DDP). Since its approval for clinical use by the Food and Drug Administration in 1978, DDP is the most effective, and therefore, widely used therapeutic agent against epithelial malignancies including lung, neck, ovarian, bladder, and testicular cancers.[1] Cisplatin exerts its anticancer effects by inhibiting transcription, as well as targeting proteins and phospholipids; however, the most prominent mode of action involves the generation of DNA adducts which can inhibit DNA replication, and ultimately lead to cell cycle arrest and programmed cell death.[1, 2]

Cisplatin is intravenously administered in its neutral (inactive) form, $[Pt(NH_3)_2Cl_2]$. The concentration of chloride ions $[Cl^-]$ within the bloodstream (~100 mM) is sufficiently high that the drug remains neutral and is able, by and large, to passively diffuse through the cell membrane.[3] Within the cell the $[Cl^-]$ decreases to 4-12 mM, thus shifting the equilibrium towards aquation of DDP.[3] Water displaces the chloride ions to yield the monoaqua-, $[Pt(NH_3)_2(H_2O)Cl]^+$, and diaqua-, $[Pt(NH_3)_2(H_2O)_2]^{2+}$, intermediate complexes, otherwise referred to as the active forms of cisplatin.[3] Both the monoaqua- and diaqua-complexes have a strong kinetic preference for binding with DNA at the N7 position of guanine to form DNA cross-link lesions.[4-6]

Specifically, activated cisplatin forms stable intrastrand cross-links between a guanine and: (1) an adjacent guanine (1,2-GpG); (2) a non-adjacent guanine separated by one nucleobase (1,3-GpXpG, where X = C or T); or (3) a 5'-neighboring adenine (1,2-ApG), at frequencies of 65%, 5-10%, and 25%, respectively.[7] An interstrand cross-link can also form between complementary strands of DNA on guanines at 5'-GpC/GpC-3' sites, as well as mono-adducts at the N7 position of the purine bases; however, these represent less than 1% of all cisplatin-DNA adducts.[7]

The current challenges to cisplatin therapy are to minimize the side effects (i.e., nausea, vomiting, ototoxicity, and nephrotoxicity) of the drug while maintaining its potency against cancer cells, to extend successful treatment to a wider range of human cancers, and to determine the molecular mechanisms underlying drug resistance.[8] To address these challenges, it is necessary to have an analytical method that can simultaneously quantify all major cisplatin adducts in DNA. Given its inherent selectivity and sensitivity, liquid chromatography-tandem mass spectrometry (LC-MS/MS) is well suited to meet this need. LC-MS/MS was employed successfully for the quantification of the 1,2-GpG adduct using a stable isotope-labeled internal standard at a detection limit of 3 fmol/μg DNA,[9] however, the described protocol was limited in that it could not be extended to include other cisplatin-DNA adducts. Moreover, the use of a single internal standard did not correct for the variation in enzymatic digestion efficiency that occurs due to the blocking nature of the different cisplatin adducts.[10-22] Thus, development of a quantitative LC-MS/MS-based method for the detection of cisplatin-

bearing DNA necessitates optimization of the enzymatic digestion procedures for the release of the aforementioned cross-link lesions from DNA.

In the present study, we prepared oligodeoxyribonucleotides (ODNs) containing $^{15}N_2$-labeled cisplatin-modified 1,2-GpG, 1,2-ApG, and 1,3-GpXpG adducts, and provided detailed MS/MS characterization of both the light and heavy cisplatin-containing ODNs. We also analyzed the enzymatic digestion products formed from 1,2-GpG, 1,2-ApG, and 1,3-GpXpG cisplatin-containing ODNs using a 4-enzyme cocktail for digestion and LC-MS/MS for monitoring the digestion products. Our results led to the identification of two types of digestion products for the 1,2-GpG- and 1,2-ApG-bearing sequences, and three for the 1,3-GpXpG sequence.

Materials and Methods

Chemicals

Cisplatin, nuclease P1, snake venom phosphodiesterase (exoPI), calf spleen phosphodiesterase (exoPII), calf intestinal alkaline phosphatase (CIP), and 3-(6-aminopurin-9-yl)nonan-2-ol hydrochloride (EHNA) were purchased from Sigma-Aldrich (St. Louis, MO). ^{15}N-labeled ammonium acetate was obtained from Cambridge Isotope Laboratories, Inc. (Andover, MA). Unmodified ODNs were purchased from Integrated DNA Technologies (San Diego, CA).

Preparation of cisplatin-modified ODNs

Cisplatin-modified ODNs were prepared according to previously published protocols with modifications.[21] To activate cisplatin (conversion of *cis*-$[Pt(NH_3)_2Cl_2]$ to

cis-$[Pt(H_2O)(NH_3)_2Cl]^+$ or *cis*-$[Pt(H_2O)_2(NH_3)_2]^{2+}$), 4.5 mg cisplatin (15 nmol/μL) was treated with 12 μL of $AgNO_3$ stock solution (2.5 μmol/μL) and diluted to 1 mL with Millipore water. The samples were allowed to shake vigorously in the dark at 37ºC for 24 hrs. The silver chloride precipitate was removed by centrifugation and the activated cisplatin (supernatant) was recovered. The single-stranded ODN sequences **1**-**5** (100 nmol) were treated with activated cisplatin (300 nmol, 20 μL) in a 1-mL reaction mixture containing 8.4 mM sodium perchlorate. The reaction mixture was allowed to shake vigorously in the dark at 37ºC for 4-5 hrs. The samples were concentrated by Speed-vac before immediate HPLC purification.

The $^{15}N_2$-labeled cisplatin, $[Pt(^{15}NH_3)_2Cl_2]$, was synthesized following published protocols.[23] After recrystallization, we collected 282 mg of $[Pt(^{15}NH_3)_2Cl_2]$ at a yield of 33%. The isotope purity was confirmed using the higher resolution "zoom scan" ESI-MS analyses of treated ODN as described in the following sections (see **Results**). $[Pt(^{15}NH_3)_2Cl_2]$ was activated with $AgNO_3$ before treating the ODN sequences as described above for the unlabeled cisplatin.

Enzymatic Digestion

To 100 μg (~35 nmol) of dried, purified ODN samples (**1**-**3**) were added 0.05 or 0.2 units/μg DNA of nuclease P1, 6.3×10^{-5} or 2.5×10^{-4} units/μg DNA of calf spleen phosphodiesterase in a 20-μL buffer containing 300 mM sodium acetate, 10 mM Zn^{2+} (pH 5.0) and water to make the final volume of the solution 150 μL. The quantity of enzymes used represents ½- and 2-times the number of units required per μg of DNA

based on vendor's specifications, respectively. The digestion was carried out at 37°C for 6 hrs, and the enzymes were deactivated by heating to 95°C for 5 min. The resulting solution was dried by Speed-vac, and to the dried residues were added a 150-μL solution containing 40 μL of a 50 mM Tris-HCl buffer (pH 8.6), 0.1 units/μg DNA of calf intestinal phosphatase, 2.5×10^{-4} units/μg DNA of snake venom phosphodiesterase, and water. EHNA was also added to the samples (1 nmol EHNA/4 μg DNA) to prevent deamination of 2'-deoxyadenosine to 2'-deoxyinosine (dI), which is important for off-line HPLC analysis because dI and the d(pGpG) digestion product coelute under the conditions used. The digestion mixture was incubated at 37°C for 4 hrs before an equal volume of chloroform was added to remove the enzymes. The pH of the aqueous layer was adjusted to neutral or slightly acidic conditions with 5% formic acid, and the aqueous layer was dried and redissolved in water prior to LC-MS/MS and HPLC analysis. Additionally, to 20 μg (~13 nmol) of dried, purified, cisplatinated ODN sequences **4** and **5** were added 2 units NP1, 0.003 units exoPII, 2 units CIP, and 0.005 units exoPI using similar procedures as described above.

HPLC

All cisplatin-modified ODNs were purified by reversed-phase HPLC using a Waters 2695 Separation Module coupled to a Linear^{TM} UVIS 200 detector (set at 260 nm) and a Grace Apollo C18 column (4.6×250 mm, 5 μm in particle size, Part # 36511). A gradient of 10% B for 9 min (desalting), 10-20% B for 30 min, 20-75% B for 3 min, and 75% B for 6 min at a flow rate of 0.7 mL/min was employed, where mobile phase A

contained 0.1 M triethylammonium acetate (TEAA) in water and mobile phase B contained 0.1 M TEAA in 80% acetonitrile. In most cases, a second purification was conducted to isolate the desired singly platinated ODN from the unmodified and multiply platinated ODNs. In this case, a gradient of 15% B for 9 min (desalting), 15-20% B for 30 min, 20-75% B for 3 min, and 75% B for 6 min at a flow rate of 0.7 mL/min was used. All purified samples were dried by Speed-vac, redissolved in water for ESI-MS and MS/MS analyses, or for enzymatic digestion followed by LC-MS/MS analysis.

Off-line analysis of the digestion mixture was performed on an HPLC system as described above. A gradient of 5-15% B for 5 min, 15-30% B for 60 min, 30-98% B for 2 min, and 98% B for 10 min at a flow rate of 400 μL/min was used, where buffer A contained 10 mM ammonium formate in water and buffer B contained 10 mM ammonium formate in 80% acetonitrile.

Mass Spectrometry

The ESI-MS and MS/MS analyses of cisplatin-modified ODNs were conducted on an LCQ Deca XP ion-trap mass spectrometer (ThermoFinnigan, San Jose, CA). An equal-volume solvent mixture of methanol and water was used as the carrier and electrospray solvent. A 2-5 μL aliquot of each purified HPLC fraction was injected for sequence verification. The spray voltage was 3.6 kV and the mass width for precursor ion selection in MS/MS mode was 3 *m/z* units. Each spectrum was obtained by averaging 20-40 scans, and the time for each scan was 0.1 s.

LC-MS/MS

A Zorbax SB-C18 column (0.5×150 mm, 5 μm in particle size, Agilent Technologies, Palo Alto, CA) and an Agilent 1200 series capillary HPLC pump were employed for all LC-MS/MS experiments. A gradient of 5-10% B (5 min) and 10-35% B (40 min) was used at a flow rate of 6 μL/min, where the mobile phases were 20 mM ammonium acetate in water (A) and 20 mM ammonium acetate in 40% acetonitrile (B). The effluent from the capillary column was coupled to an LTQ linear ion-trap mass spectrometer (Thermo Electron, San Jose, CA) operated in positive-ion mode in order to identify the cisplatin-containing products released upon enzymatic digestion.

Results and Discussion

Formation of cisplatin intrastrand cross-links

We begin by verifying the formation of the 1,2-GpG, 1,2-ApG, and 1,3-GpXpG intrastrand cross-links in the 10-mer ODNs that were treated with activated cisplatin. To this end, we subjected sequences **1**, **2**, and **3** (Table 2.1) to MS and MS/MS analyses. The negative-ion ESI-MS for the ODNs modified by unlabeled (light) cisplatin $[Pt(NH_3)_2]^{2+}$ (Figure 2.1a-2.3a) reflects a mass increase of 227 Da, relative to the unmodified ODN, which corresponds to the monoisotopic mass of the $[Pt(NH_3)_2]^{2+}$ less two protons. Similarly, the MS for the ODN treated with labeled (heavy) cisplatin $[Pt(^{15}NH_3)_2]^{2+}$ (Figure 2.1c-2.3c) shows a mass increase of 229 Da, which is consistent with the monoisotopic mass of $[Pt(^{15}NH_3)_2]^{2+}$ less two protons. To further demonstrate the presence of a cisplatin adduct and to unambiguously distinguish between the light and

Table 2.1 The ODN sequences employed for enzymatic digestion experiments.

Sequence Number	Abbreviation	Sequences
1	1,2-GpG	5'-d(ATCCGGCCTA)-3'
2	1,2-ApG	5'-d(ATCCAGCCTA)-3'
3	1,3-GpXpG	5'-d(ATCGCGCCTA)-3'
4	***3'***-1,2-GpG	5'-d(TCTGG)-3'
5	***5'***-1,2-GpG	5'-d(GGTCT)-3'

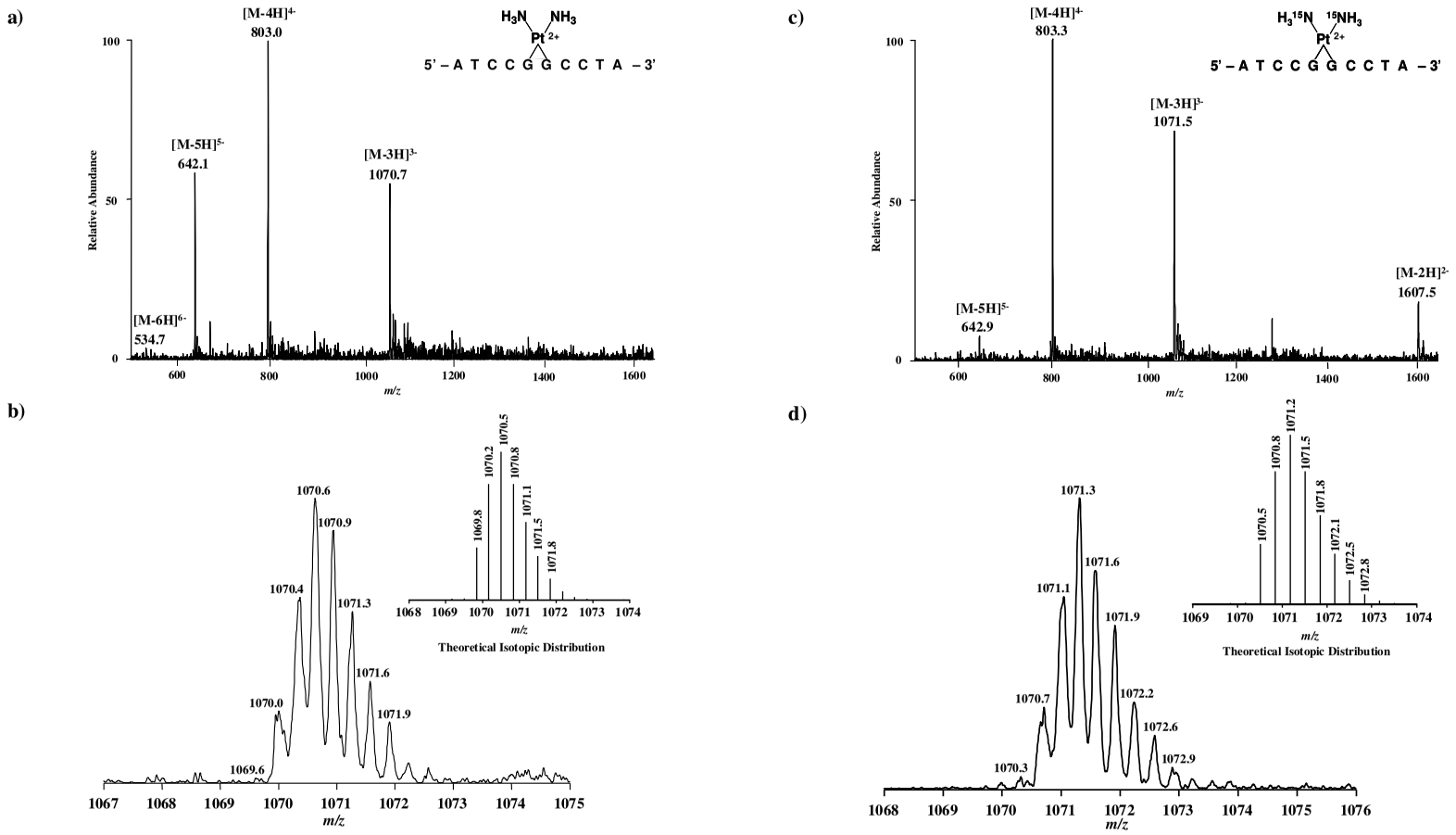


Figure 2.1 Negative-ion ESI-MS (a,c) and high-resolution "zoom-scan" analysis showing the experimental isotopic distribution of the $[M-3H]^{3-}$ ion with an inset depicting the corresponding theoretical isotopic distribution (b,d) of unlabeled (a,b) and $^{15}N_2$-labeled (c,d) d(ATCCG*G*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$ cisplatin coordination, respectively.

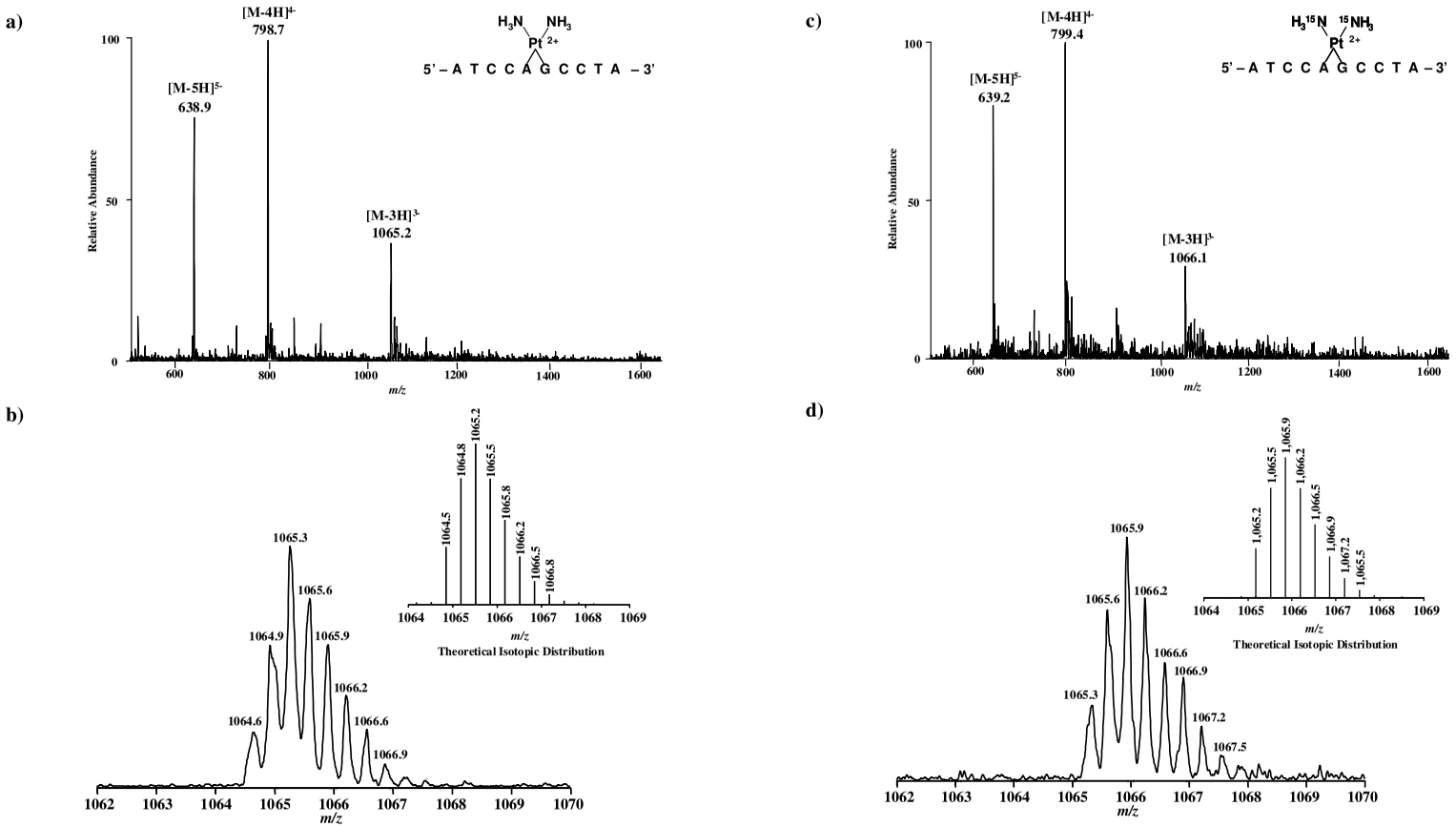


Figure 2.2 Negative-ion ESI-MS (a,c) and high-resolution "zoom-scan" analysis showing the experimental isotopic distribution of the $[M-3H]^{3-}$ ion with an inset depicting the corresponding theoretical isotopic distribution (b,d) of unlabeled (a,b) and $^{15}N_2$-labeled (c,d) d(ATCCA*G*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$ cisplatin coordination, respectively.

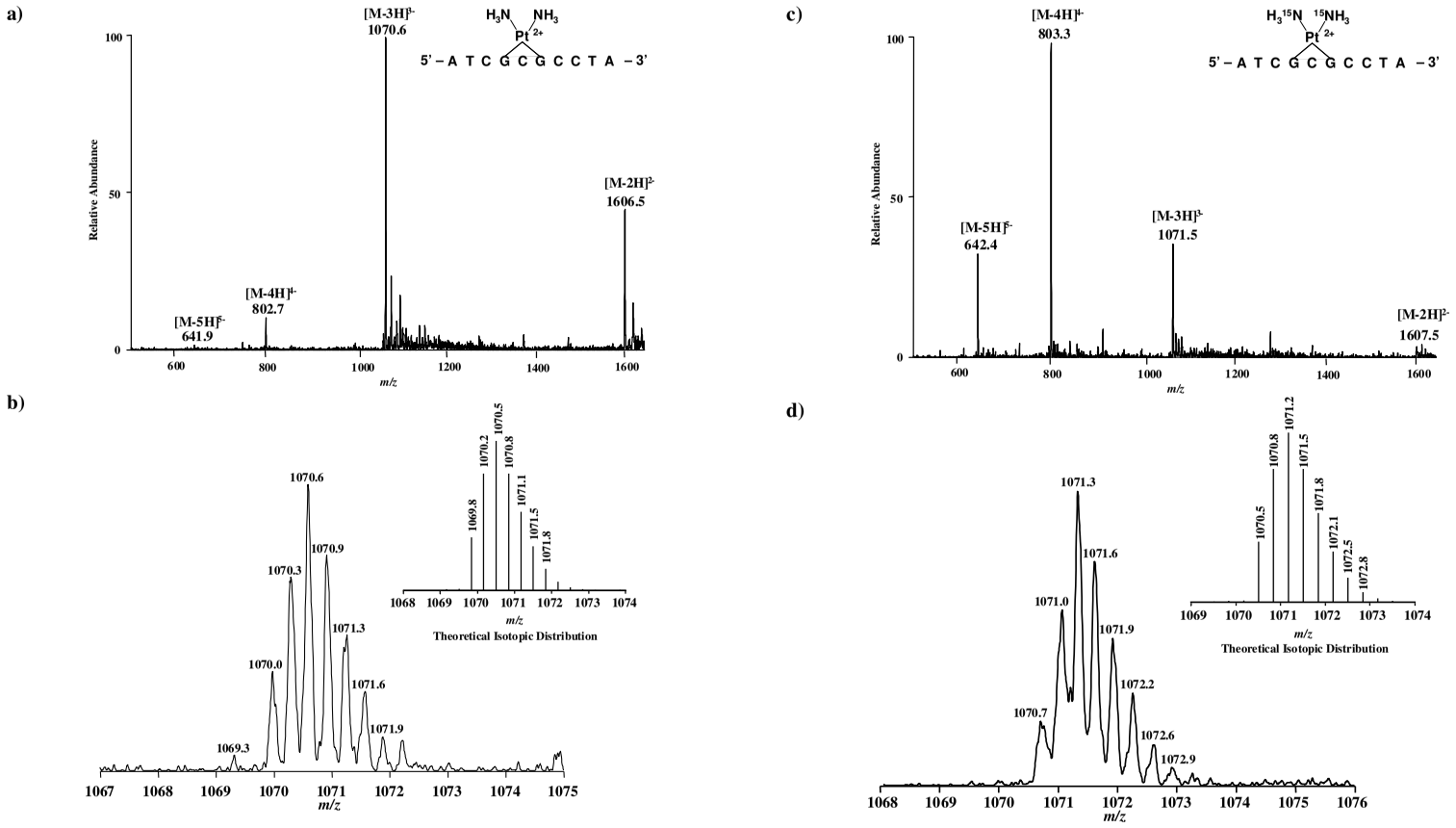


Figure 2.3 Negative-ion ESI-MS (a,c) and high-resolution "zoom-scan" analysis showing the experimental isotopic distribution of the $[M-3H]^{3-}$ ion with an inset depicting the corresponding theoretical isotopic distribution (b,d) of unlabeled (a,b) and $^{15}N_2$-labeled (c,d) d(ATCG*CG*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$ cisplatin coordination, respectively.

heavy sequences, we also acquired higher-resolution "zoom-scan" MS to reveal the experimental isotopic distribution of the deprotontated ions of the platinated ODNs (Figures 2.1 b, c – 2.3 b, c). For all sequences, the experimental isotopic distributions are in excellent agreement with the theoretical (calculated) ones. Therefore, we are confident that sequences **1**, **2**, and **3** were modified by the addition of one cisplatin adduct.

MS/MS analysis was implemented to confirm the location of the cisplatin intrastrand cross-link within the ODN sequences. Fragmentation of unmodified ODN by collision-induced dissociation (CID) typically gives rise to the a-B/w fragment ion pairs.[24] However, the coordination of $[Pt(NH_3)_2]^{2+}$ or $[Pt(^{15}NH_3)_2]^{2+}$ to the ODN resulted in unique fragment ions, as illustrated in Figures 2.4, 2.5, and 2.6. Beginning with the 1,2-GpG-containing ODN (sequence **1**, Figure 2.4), the a_5-G/w_5*+G ion pair is readily detectable, which arises from the release of the 5th nucleobase (counting from the 5' end) from the backbone of the ODN, and cisplatin (*) coordination to the 5th and 6th guanine residues, which are retained on the fragment ion carrying the 3' terminus (i.e., w_5* ion). Additionally, the a_5-G/w_5* ion pair was identified, which reflects the loss of the 5th guanine and retention of only cisplatin on the 6th nucleobase. To a lesser degree, we found the a_5*/w_5 fragment-ion pair in which the cisplatin is coordinated with the 5th guanine base that remains bound to the deoxyribose on the 5' fragment ion.

We also observed unique cleavage between the 6th and 7th nucleotides to yield the a_6*-G/w_4 and a_6*/w_4 fragment ion pairs. The a_6*-G and a_6* ions form from the loss of the 6th nucleobase with cisplatin remaining on the 5th guanine, and retention of both the 6th guanine, albeit cleaved from the deoxyribose, and cisplatin on the fragment ion

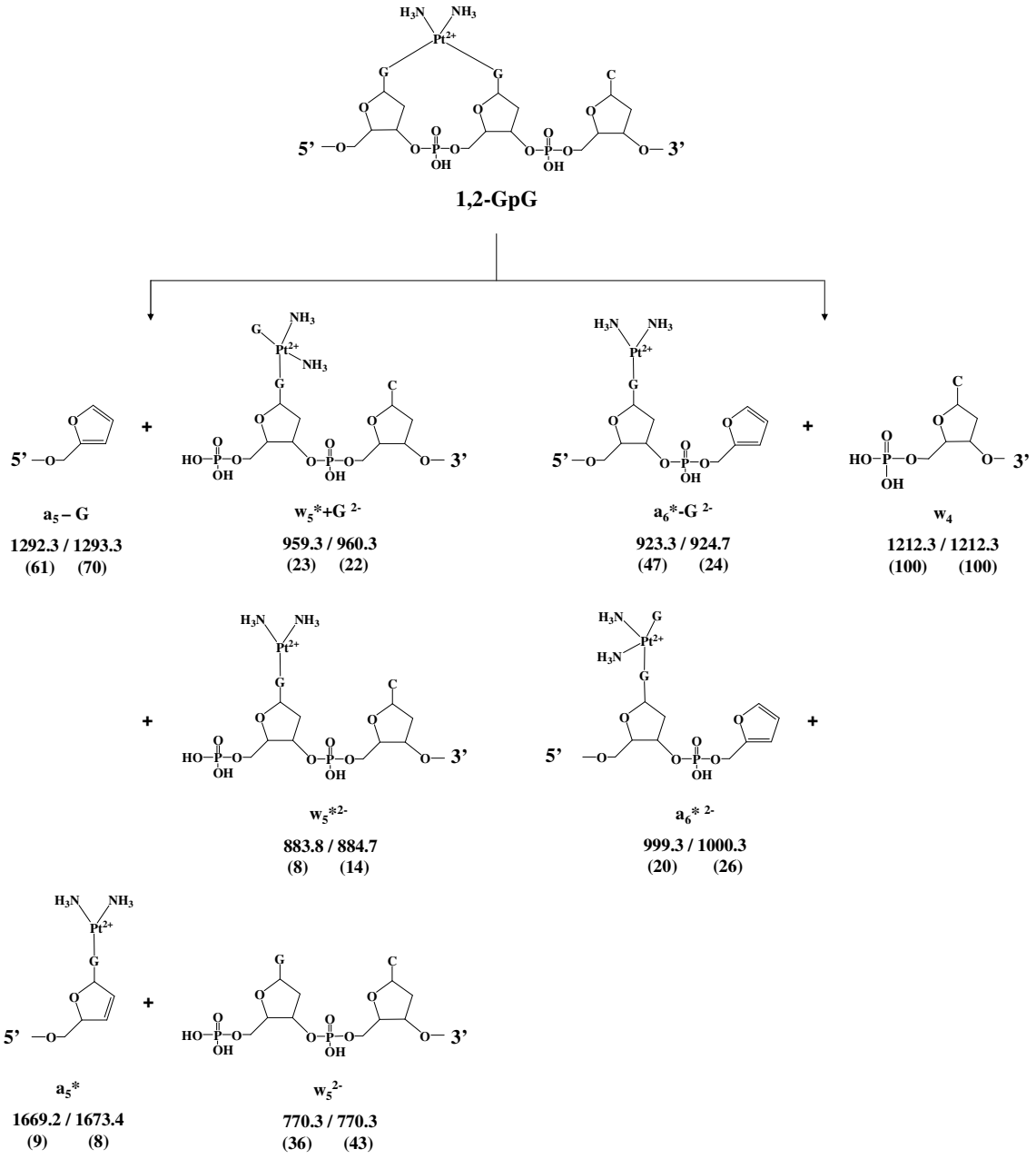


Figure 2.4 Unique fragment ion pairs and the corresponding *m/z* values $\{[Pt(NH_3)_2]^{2+}/[Pt(^{15}NH_3)_2]^{2+}\}$ generated during the CID of ODN sequence **1**, which harbors a 1,2-GpG cisplatin (*) intrastrand cross-link. Numbers in parenthesis represents relative ion abundance.

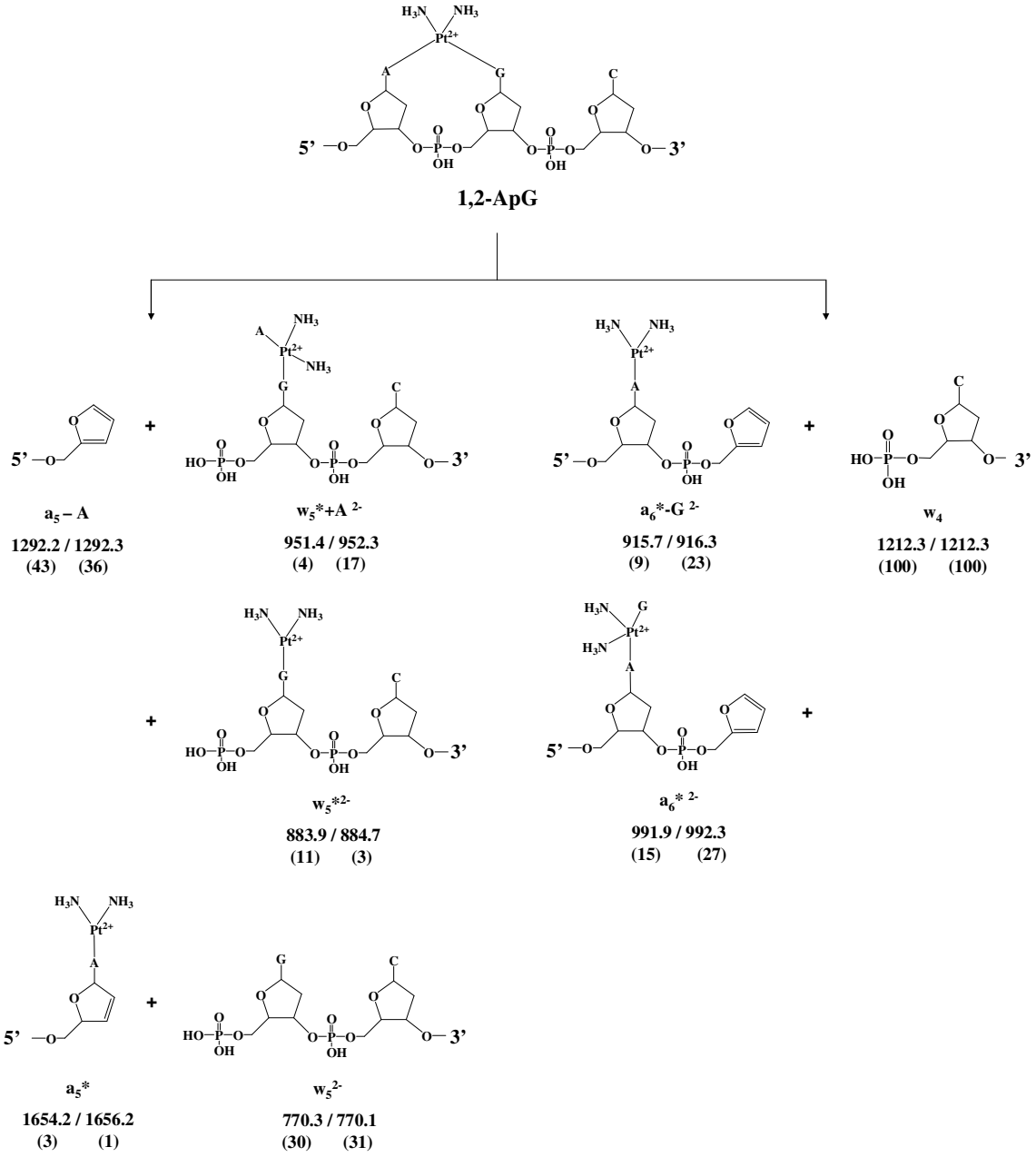


Figure 2.5 Unique fragment ion pairs and the corresponding *m/z* values $\{[Pt(NH_3)_2]^{2+}/[Pt(^{15}NH_3)_2]^{2+}\}$ generated during the CID of ODN sequence **2,** which harbors a 1,2-ApG cisplatin (*) intrastrand cross-link. Numbers in parenthesis represents relative ion abundance.

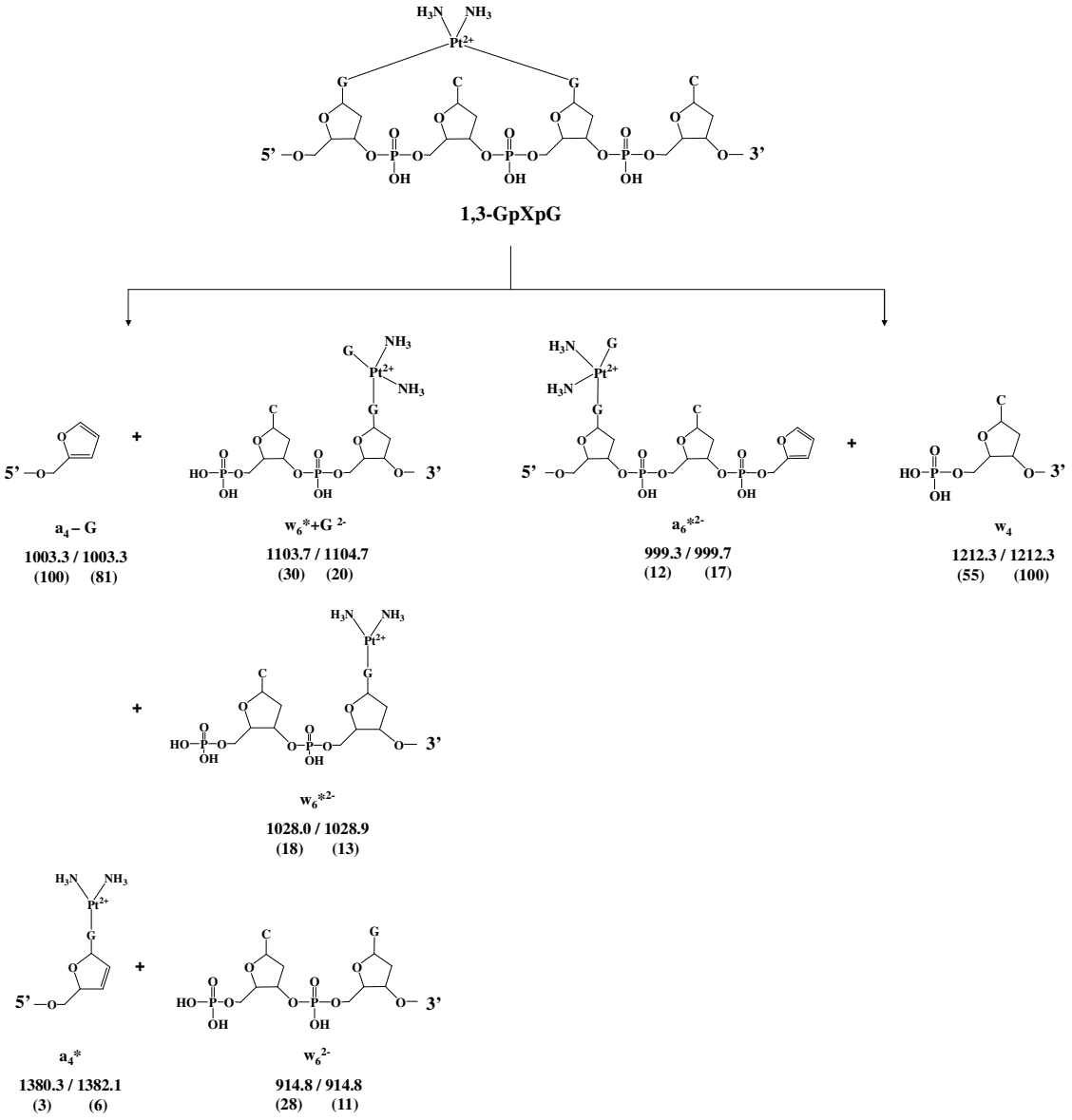


Figure 7. Unique fragment ion pairs and the corresponding *m/z* values {$[Pt(NH_3)_2]^{2+}$/ $[Pt(^{15}NH_3)_2]^{2+}$} generated during the CID of ODN sequence **3,** which harbors a 1,2-GpXpG cisplatin (*) intrastrand cross-link. Numbers in parenthesis represents relative ion abundance.

carrying the 5' terminus, respectively. Ions arising from the loss of one, and in many cases, both ammine groups for the light and heavy cisplatin-containing fragments were also identified. The presence of fragment ions with cisplatin bonded with the 5^{th} or 6^{th} guanine, or both simultaneously, may be explained by the metal complex nature of the adduct coordination to the N7 position of guanine, as opposed to covalent bonds, which allows for mobility of platinum moiety between the two nucleobases. Based on the ions generated upon cleavage at the a_5-B/w_5 and a_6-B/w_4 sites, coupled with a previously published report describing a comparable fragmentation pattern for cisplatin-harboring ODN with a 1,2-GpG motif,[25] it is evident that the $[Pt(NH_3)_2]^{2+}$ and $[Pt(^{15}NH_3)_2]^{2+}$ adducts are complexed with both the 5^{th} and 6^{th} guanines, thus confirming the presence of a bifunctional intrastrand cross-link.

Although analogous in trend, to our knowledge we are the first to fully characterize the fragmentation pattern of ODNs carrying a 1,2-ApG (sequence **2**, Figure 2.5) or 1,3-GpXpG (sequence **3**, Figure 2.6) cisplatin intrastrand cross-link. For sequence **2**, we observed the a_5-A/w_5*+A ion pair, which arises from the cleavage of *N*-glycosidic bond of the 5^{th} nucleoside, and the retention of the 5^{th} nucleobase (i.e., adenine) along with the cisplatin on the 6^{th} guanine of the 3' fragment ion. The a_5-A/w_5* and a_5*/w_5 fragments are also present, indicating the loss of the 5^{th} adenine with only the cisplatin remaining on the 6^{th} guanine, and the retention of the 5^{th} adenine as well as cisplatin on the 5'-fragment ion, respectively. As with the 1,2-GpG-containing sequence, the a_5* fragment is significantly less abundant than w_5*+A and w_5*.

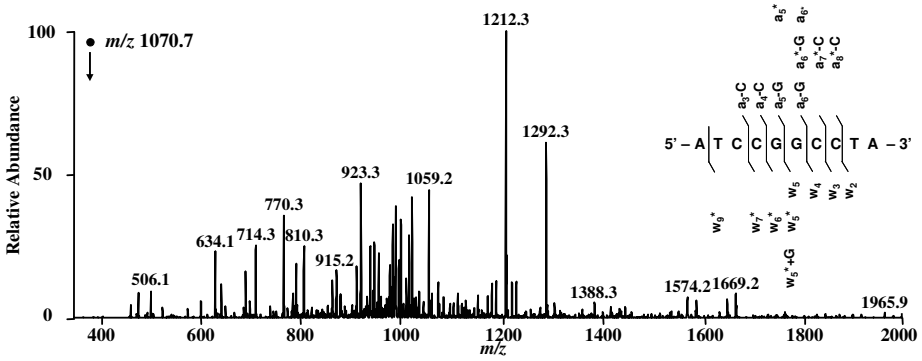
In the case of sequence **3**, we again identified ion pairs where the cisplatin adduct was present on the 5' fragment ions (a_4*/w_6 and a_6*/w_4), or 3' fragment ions either with (a_4-G/w_6*+G) or without (a_4-G/w_6*) the guanine being cleaved from the 5' deoxyribose. As was observed with sequences **1** and **2**, ions representing the loss of one or both of the cisplatin ammines were abundantly present in the product-ion spectra for sequence **3,** which was described in part previously.[25, 26] Complete assignment of the MS/MS for all sequences may be found in Figures 2.7-2.12. Additionally, we verified the coordination of one cisplatin adduct to the 5-mer sequences (**4** and **5**) by MS and zoom-scan MS (Figure 2.13).

Identification of Enzymatic Digestion Products

We next digested the aforementioned cisplatin-modified ODNs with a 4-enzyme cocktail of NP1, exoPII, CIP, and exoPI and analyzed the enzymatic digestion products using LC-MS/MS (see **Methods and Materials** for reaction conditions). We chose these four enzymes because they have been employed successfully for the efficient release of other intrastrand cross-links from isolated and cellular DNA.[27-30] Moreover, this combination of enzymes has not been previously reported for digesting cisplatin-modified DNA, though its usefulness towards developing a quantitative assay of $Pt(NH_3)_2$-bearing DNA was suggested.[22]

Figure 2.14 displays the products housing the cisplatin remnant that we identified from LC-MS/MS analysis (Figures 2.15-2.21). As was previously reported with unlabeled cisplatin, we observed *m/z* values corresponding to cisplatin-chelated (*)

a)



b)

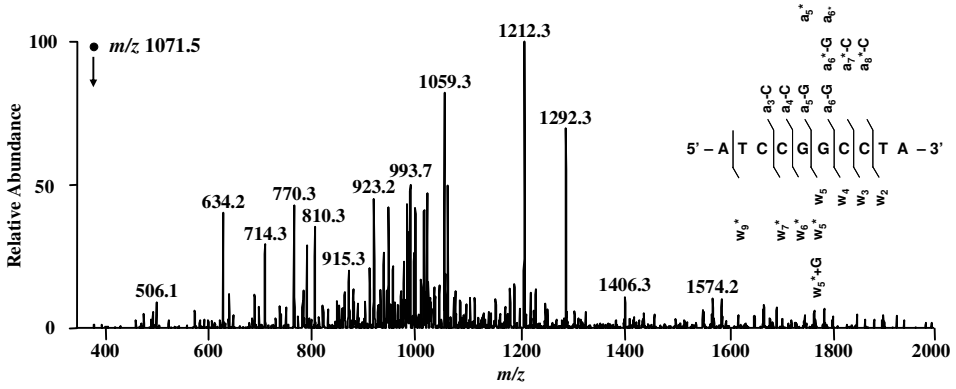
a-B Fragment Ions						w Fragment Ions		
3-	**2-**	**1-**		**5'**		**1-**	**2-**	**3-**
			1	**A**	10			
			2	**T**	9			**992.8 (39)**
		714.2 (24)	3	**C**	8			
		1003.3 (33)	4	**C**	7		**1192.8 (13)**	**795.3 (19)**
	645.7 (12)	1292.3 (61)	5	**G**	6		**1048.2 (6)**	
	810.3 (25)		6	**G**	5	1541.1 (2)	770.3 (36)	
	¥**923.3 (47)**					**1768.1 (1)**	**883.8 (8)**	
	1088.0 (7)		7	**C**	4	1212.3 (100)	605.7 (6)	
	1233.2 (13)		8	**C**	3	¥923.3 (47)		
			9	**T**	2	634.1 (23)		
			10	**A**	1			
				3'				

Unique Fragment Ions							
3-	**2-**	**1-**			**1-**	**2-**	**3-**
	834.2 (3)	**1669.2 (9)**	a_5	w_5-NH_3		**875.3 (17)**	
	826.2 (4)	**1652.3 (7)**	a_5-NH_3	w_5-$2NH_3$		**866.2 (13)**	
		1636.1 (2)	a_5-$2NH_3$	w_5+G		**959.3 (23)**	
				w_5+G-NH_3		**950.4 (27)**	
				w_5+G-$2NH_3$		**942.2 (25)**	
	915.2 (18)		a_6-G-NH_3	w_6-NH_3		**1039.8 (9)**	
	907.1 (4)		a_6-G-$2NH_3$	w_6-$2NH_3$		**1031.2 (7)**	
	999.3 (20)	**1998.9 (2)**	a_6				
	990.3 (13)		a_6-NH_3				
	981.3 (14)	**1965.9 (2)**	a_6-$2NH_3$				
	1079.9 (5)		a_7-C-NH_3	w_7-NH_3		**1184.2 (12)**	**789.3 (4)**
				w_7-$2NH_3$		**1175.7 (8)**	
	1224.2 (12)		a_8-C-NH_3				
				w_9-NH_3			**987.0 (33)**
				w_9-$2NH_3$			**981.3 (19)**
1025.3 (42)			M-A	M-NH_3			**1065.1 (9)**
1019.5 (29)			M-A-NH_3	M-$2NH_3$			**1059.2 (45)**
1013.4 (14)			M-A-$2NH_3$				

¥ Shared *m/z* values

Figure 2.7 Negative-ion ESI-MS/MS of the unlabeled $[M-3H]^{3-}$ ion with an inset showing a list of the observed fragment ions (a) and the peak assignments (with relative ion abundance in parenthesis) for the product-ion spectrum (b) of d(ATCCG*G*CCTA) where * (a) and bold *m/z* values (b) represent $[Pt(NH_3)_2]^{2+}$ coordination.

a)



b)

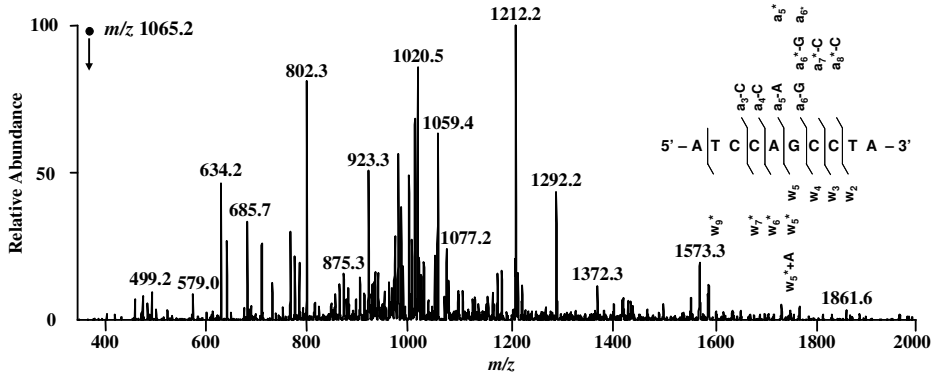
a-B Fragment Ions						w Fragment Ions		
3-	**2-**	**1-**		**5'**		**1-**	**2-**	**3-**
			1	**A**	10			
			2	**T**	9			**993.5 (50)**
		714.3 (29)	3	**C**	8			
		1003.3 (42)	4	**C**	7		**1194.7 (15)**	**795.4 (1)**
	645.8 (12)	1292.3 (70)	5	**G**	6		**1049.5 (15)**	
	810.3 (35)		6	**G**	5		770.3 (43)	
	924.7 (24)						**884.7 (14)**	
	1089.9 (10)		7	**C**	4	1212.3 (100)	605.7 (3)	
	1234.1 (14)		8	**C**	3	923.3 (45)		
			9	**T**	2	634.2 (40)		
			10	**A**	1			
				3'				

Unique Fragment Ions

3-	**2-**	**1-**			**1-**	**2-**	**3-**
	835.5 (7)	**1673.4 (8)**	a_5	w_5-NH_3		**875.7 (20)**	
	826.7 (7)	**1654.6 (4)**	a_5-NH_3	w_5+G		**960.3 (22)**	
	817.8 (4)	**1637.3 (2)**	a_5-$2NH_3$	w_5+G-NH_3		**951.0 (42)**	
				w_5+G-$2NH_3$		**942.4 (26)**	
	915.3 (21)		a_6-G-NH_3	w_6-NH_3		**1039.7 (14)**	
	907.2 (9)		a_6-G-$2NH_3$	w_6-$2NH_3$		**1031.3 (11)**	
	1000.3 (26)		a_6				
	991.3 (34)		a_6-NH_3				
	1080.3 (13)		a_7-C-NH_3	w_7-NH_3		**1185.2 (14)**	**788.1 (13)**
				w_7-$2NH_3$		**1176.2 (8)**	
	1225.7 (13)		a_8-C-NH_3				
				w_9-NH_3			**987.3 (43)**
1026.4 (47)			M-A	M-NH_3			**1065.5 (50)**
1020.1 (41)			M-A-NH_3	M-$2NH_3$			**1059.3 (82)**
1013.8 (7)			M-A-$2NH_3$				

Figure 2.8 Negative-ion ESI-MS/MS of the $^{15}N_2$-labeled $[M-3H]^{3-}$ ion with an inset showing a list of the observed fragment ions (a) and the peak assignments (with relative ion abundance in parenthesis) for the product-ion spectrum (b) of d(ATCCG*G*CCTA) where * (a) and bold *m/z* values (b) represent $[Pt(^{15}NH_3)_2]^{2+}$ coordination.

a)



b)

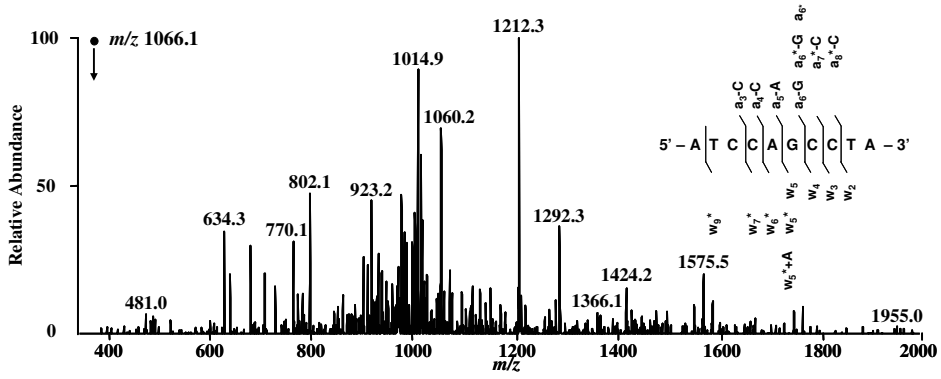
a-B Fragment Ions						w Fragment Ions		
3-	**2-**	**1-**		**5'**		**1-**	**2-**	**3-**
			1	**A**	10			
			2	**T**	9			**987.7 (35)**
		714.1 (20)	3	**C**	8			
		1003.3 (29)	4	**C**	7			**788.2 (19)**
	645.7 (27)	1292.2 (43)	5	**A**	6		**1040.7 (6)**	
	802.3 (81)	1605.4 (2)	6	**G**	5	1541.5 (2)	770.3 (30)	
	915.7 (9)	**1832.9 (2)**				**1768.3 (5)**	**883.9 (11)**	
	1080.9 (11)		7	**C**	4	1212.3 (100)		
	1225.2 (12)		8	**C**	3	923.2 (49)		
			9	**T**	2	634.2 (47)		
			10	**A**	1			
				3'				

Unique Fragment Ions

3-	**2-**	**1-**			**1-**	**2-**	**3-**
	826.2 (5)	**1654.2 (3)**	a_5	w_5-NH_3	**1751.3 (3)**	**875.2 (15)**	
	819.4 (4)		a_5-NH_3	w_5-$2NH_3$	**1734.2 (5)**	**866.7 (12)**	
		1620.5 (2)	a_5-$2NH_3$	w_5+A		**951.4 (4)**	
				w_5+A-NH_3		**943.0 (14)**	
				w_5+A-$2NH_3$		**934.3 (12)**	
	907.7 (11)	**1815.3 (2)**	a_6-G-NH_3	w_6-NH_3		**1032.4 (18)**	
	898.7 (7)		a_6-G-$2NH_3$	w_6-$2NH_3$		**1023.3 (18)**	
	991.9 (15)		a_6				
	982.2 (51)	**1967.1 (2)**	a_6-NH_3				
	974.9 (18)		a_6-$2NH_3$				
	1072.2 (8)		a_7-G-NH_3	w_7-NH_3		**1176.3 (16)**	
				w_7-$2NH_3$		**1167.7 (9)**	
	1217.1 (14)		a_8-C-NH_3				
	1208.4 (3)		a_8-C-$2NH_3$				
				w_9-NH_3			**982.2 (51)**
				w_9-$2NH_3$			**976.5 (24)**
1020.5 (83)			M-A	M-NH_3			**1059.3 (62)**
1014.8 (68)			M-A-NH_3	M-$2NH_3$			**1053.8 (20)**
1008.7 (27)			M-A-$2NH_3$				

Figure 2.9 Negative-ion ESI-MS/MS of the unlabeled $[M-3H]^{3-}$ ion with an inset showing a list of the observed fragment ions (a) and the peak assignments (with relative ion abundance in parenthesis) for the product-ion spectrum (b) of d(ATCCA*G*CCTA) where * (a) and bold *m/z* values (b) represent $[Pt(NH_3)_2]^{2+}$ coordination.

a)



b)

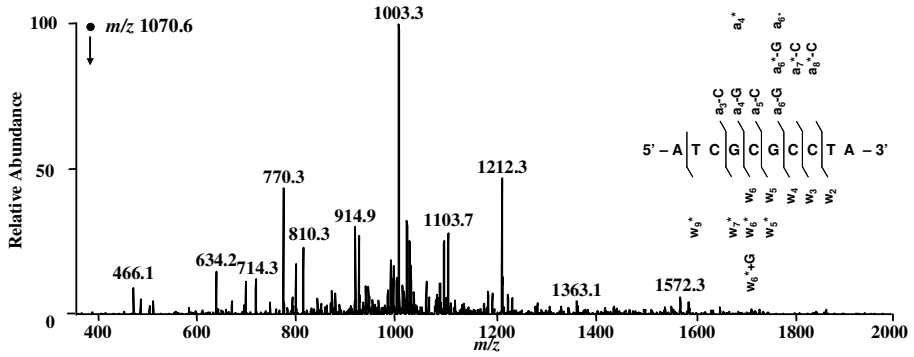
a-B Fragment Ions						w Fragment Ions		
3-	**2-**	**1-**		**5'**		**1-**	**2-**	**3-**
			1	**A**	10			
			2	**T**	9		**1482.3 (8)**	**988.3 (33)**
		714.2 (20)	3	**C**	8			
		1003.3 (31)	4	**C**	7		**1185.8 (7)**	
	645.7 (20)	1292.3 (36)	5	**A**	6		**1041.7 (7)**	
	802.1 (47)		6	**G**	5		770.1 (31)	
	916.3 (23)					**1770.2 (9)**	**884.7 (3)**	
	1081.9 (14)		7	**C**	4	1212.3 (100)		
	1225.5 (10)		8	**C**	3	923.2 (45)		
			9	**T**	2	634.3 (35)		
			10	**A**	1			
				3'				

Unique Fragment Ions

3-	**2-**	**1-**			**1-**	**2-**	**3-**
	819.4 (2)	**1638.3 (2)**	a_5-NH_3	w_5-NH_3	**1752.5 (8)**	**875.4 (2)**	
	809.3 (3)	**1620.1 (3)**	a_5-$2NH_3$	w_5-$2NH_3$	**1734.1 (3)**	**866.7 (7)**	
				w_5+A		**952.3 (17)**	
				w_5+A-NH_3		**943.6 (21)**	
				w_5+A-$2NH_3$		**934.6 (13)**	
	908.2 (26)		a_6-G-NH_3	w_6-NH_3		**1032.5 (20)**	
	899.1 (5)		a_6-G-$2NH_3$	w_6-$2NH_3$		**1023.0 (38)**	
	992.3 (27)		a_6				
	983.2 (25)		a_6-NH_3				
	974.2 (16)		a_6-$2NH_3$				
	1072.7 (4)		a_7-C-NH_3	w_7-NH_3		**1176.7 (10)**	
				w_7-$2NH_3$		**1168.4 (4)**	
	1216.9 (13)		a_8-C-NH_3				
	1207.9 (4)		a_8-C-$2NH_3$				
				w_9-NH_3			**982.1 (16)**
				w_9-$2NH_3$			**976.2 (23)**
1020.9 (60)			M-A	M-NH_3			**1060.2 (70)**
1014.9 (88)			M-A-NH_3	M-$2NH_3$			**1053.7 (14)**
1008.8 (41)			M-A-$2NH_3$				

Figure 2.10 Negative-ion ESI-MS/MS of the unlabeled $[M-3H]^{3-}$ ion with an inset showing a list of the observed fragment ions (a) and the peak assignments (with relative ion abundance in parenthesis) for the product-ion spectrum (b) of d(ATCCA*G*CCTA) where * (a) and bold *m/z* values (b) represent $[Pt(^{15}NH_3)_2]^{2+}$ coordination.

a)



b)

a-B Fragment Ions						w Fragment Ions		
3-	**2-**	**1-**		**5'**		**1-**	**2-**	**3-**
			1	**A**	10			
			2	**T**	9			**992.5 (17)**
		714.3 (13)	3	**C**	8			
		1003.3 (100)	4	**G**	7		**1193.2 (7)**	**795.3 (17)**
	665.7 (5)	1332.2 (3)	5	**C**	6		¥914.8 (28)	
							1028.0 (18)	
	810.3 (22)		6	**G**	5		770.3 (42)	
	¥923.3 (30)						**883.1 (3)**	
	1088.9 (3)		7	**C**	4	1212.3 (55)		
	1232.9 (5)		8	**C**	3	¥923.3 (30)		
			9	**T**	2	634.2 (16)		
			10	**A**	1			
				3'				

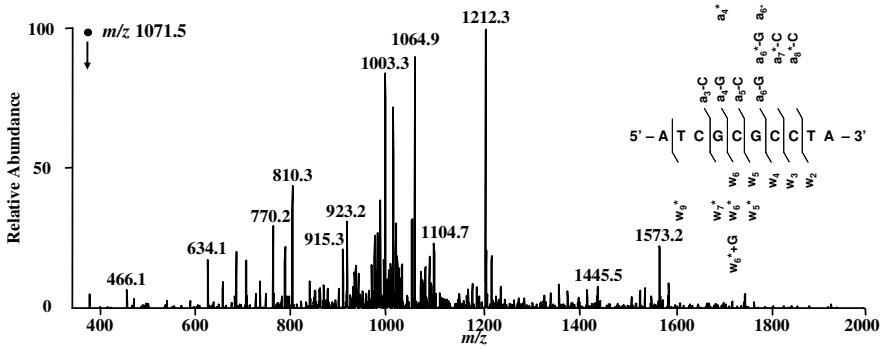
Unique Fragment Ions

3-	**2-**	**1-**			**1-**	**2-**	**3-**
		1380.3 (3)	a_4				
		1363.2 (4)	a_4-NH_3				
				w_5-NH_3		**875.1 (7)**	
				w_5-$2NH_3$		**866.8 (4)**	
	¥914.8 (28)		a_6-G-NH_3	w_6-NH_3		**1019.3 (30)**	
	906.9 (3)		a_6-G-$2NH_3$	w_6-$2NH_3$		**1010.8 (9)**	
	999.3 (12)		a_6	w_6+G		**1103.7 (30)**	
	990.8 (13)		a_6-NH_3	w_6+G-NH_3		**1095.2 (25)**	
	981.5 (9)		a_6-$2NH_3$	w_6+G-$2NH_3$		**1086.7 (10)**	
	1080.1 (6)		a_7-C-NH_3	w_7-NH_3		**1183.7 (9)**	
				w_7-$2NH_3$		**1175.7 (5)**	
	1224.3 (6)		a_8-C-NH_3				
	1215.2 (2)		a_8-C-$2NH_3$				
				w_9-NH_3			**987.1 (20)**
				w_9-$2NH_3$			**981.5 (9)**
1025.5 (25)			M-A	M-NH_3			**1064.9 (7)**
1019.9 (36)			M-A-NH_3	M-$2NH_3$			**1059.3 (11)**
1013.8 (8)			M-A-$2NH_3$				

¥ Shared *m/z* values

Figure 2.11 Negative-ion ESI-MS/MS of the unlabeled $[M-3H]^{3-}$ ion with an inset showing a list of the observed fragment ions (a) and the peak assignments (with relative ion abundance in parenthesis) for the product-ion spectrum (b) of d(ATCG*CG*CCTA) where * (a) and bold *m/z* values (b) represent $[Pt(NH_3)_2]^{2+}$ coordination.

a)



b)

a-B Fragment Ions						w Fragment Ions		
3-	**2-**	**1-**		**5'**		**1-**	**2-**	**3-**
			1	**A**	10			
			2	**T**	9			**993.2 (40)**
		714.2 (18)	3	**C**	8			
		1003.3 (81)	4	**G**	7		**1193.7 (8)**	**795.1 (23)**
	665.7 (10)	1332.5 (3)	5	**C**	6		914.8 (11)	
							1028.9 (13)	
	810.3 (45)		6	**G**	5		770.2 (27)	
	924.3 (11)						**884.7 (7)**	
	1088.8 (4)		7	**C**	4	1212.3 (100)		
	1233.7 (5)		8	**C**	3	923.2 (31)		
			9	**T**	2	634.1 (18)		
			10	**A**	1			
				3'				

Unique Fragment Ions

3-	**2-**	**1-**			**1-**	**2-**	**3-**
		1382.1 (6)	a_4				
		1364.1 (9)	a_4-NH_3				
				w_5-NH_3		**875.4 (9)**	
				w_5-$2NH_3$		**866.7 (7)**	
	915.3 (21)		a_6-G-NH_3	w_6-NH_3		**¥1019.9 (70)**	
	906.5 (7)		a_6-G-$2NH_3$	w_6-$2NH_3$		**1010.7 (14)**	
	999.7 (17)		a_6	w_6+G		**1104.7 (20)**	
	991.3 (26)		a_6-NH_3	w_6+G-NH_3		**1095.2 (19)**	
	981.4 (27)		a_6-$2NH_3$	w_6+G-$2NH_3$		**1086.7 (15)**	
	1079.9 (10)		a_7-C-NH_3	w_7-$2NH_3$		**1175.7 (6)**	
	1224.4 (19)		a_8-C-NH_3				
	1215.7 (7)		a_8-C-$2NH_3$				
				w_9-NH_3			**987.1 (28)**
				w_9-$2NH_3$			**981.5 (27)**
1025.8 (30)			M-A	M-NH_3			**1064.9 (92)**
¥1019.9 (70)			M-A-NH_3	M-$2NH_3$			**1059.3 (33)**
1014.3 (16)			M-A-$2NH_3$				

¥ Shared *m/z* values

Figure 2.12 Negative-ion ESI-MS/MS of the unlabeled $[M-3H]^{3-}$ ion with an inset showing a list of the observed fragment ions (a) and the peak assignments (with relative ion abundance in parenthesis) for the product-ion spectrum (b) of d(ATCG*CG*CCTA) where * (a) and bold *m/z* values (b) represent $[Pt(^{15}NH_3)_2]^{2+}$ coordination.

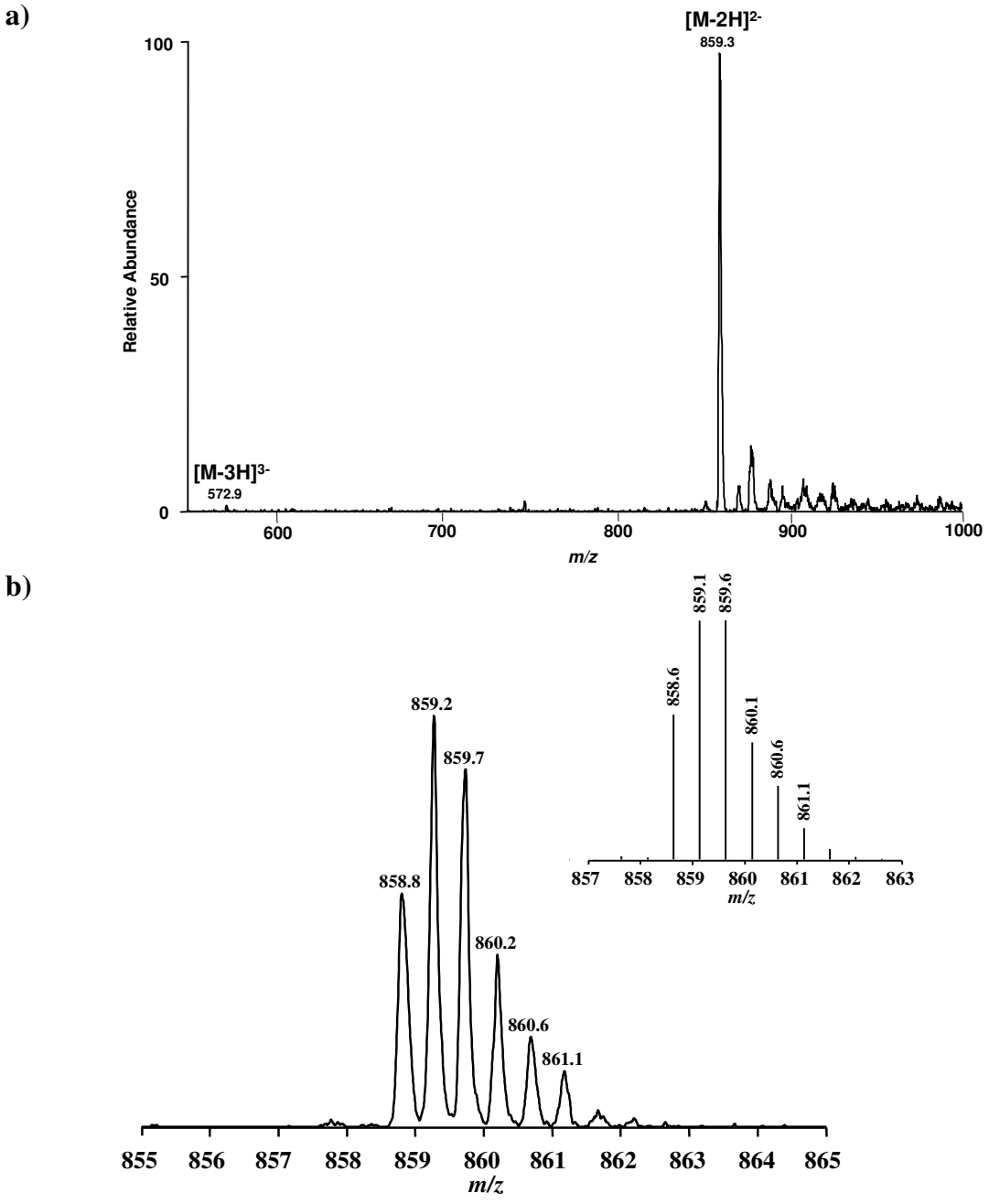


Figure 2.13 Negative-ion ESI-MS (a), and zoom scan of the experimental isotopic distribution of the $[M-2H]^{2-}$ ion with an inset of the corresponding theoretical isotopic distribution (b) for sequences **4** and **5**.

1,2-GpG dinucleoside digestion products harboring one [structure **II**, d(G*pG*)] [9, 11, 12, 16, 17, 20, 22, 31] or two [structure **I**, d(pG*pG*) or d(G*pG*p)] [13, 15, 18, 19, 22, 31-35] phosphate groups. Similarly, cisplatin-containing 1,2-ApG dinucleoside digestion products bearing one [structure **IV**, d(A*pG*)] or two [structure **III**, d(pA*pG*) or d(A*pG*p)] phosphate groups were observed.[11-13, 15, 17-20] The digestion products for the non-adjacent 1,3-GpXpG-harboring motif differed from the adjacent cross-links in that the phosphodiester bonds located in between the cisplatin adduct were cleaved. This gave rise to dinucleoside digestion products with zero [structure **VII**, d(G*G*)],[11, 12, 20] one [structure **VI**, d(pG*G*) or d(G*G*p)], or two [structure **V**, $(dG*MP)_2$][13, 18, 19] attached phosphate groups, where, to our knowledge, structure **VI** has not been previously reported. The accessibility of the phosphodiester bonds in the non-adjacent intrastrand cross-link may be explained by its distortion of DNA. NMR [36] and molecular dynamic studies [37] have shown that the central nucleotide (X = T [36] or C [37]) is looped out away from the other nucleotides, thus exposing the phosphodiester bonds and rendering them available for nuclease cleavage. In the 1,2-GpG and 1,2-ApG cross-links the interposed phosphodiester bonds do not protrude in this way.

Structure Verification of Dinucleotide Digestion Products

Although structures **I** and **III** were previously reported, we found some ambiguity in the literature regarding whether the terminal phosphate resides on the 5' or 3' end of the dinucleotide.[15, 22, 31] To address this question directly, we synthesized cisplatin-bearing 5-mers in which the GG motif was located on either the 3'- (sequence **4**) or 5'- (sequence **5**) terminus, and then subjected both to enzymatic hydrolysis (see **Materials**

and Methods). The HPLC trace of the digestion mixture for each sequence is provided in Figures 2.22a and 2.22b. For sequence **4**, the peak at ~ 25 min represents the major digestion product, d(pG*pG*), in which the phosphate resides on the 5'-side of the cisplatin adduct. A trace amount of the complete digestion product, d(G*pG*), was also detected at ~ 46 min. Conversely, sequence **5** yielded only the d(G*pG*) dinucleoside at ~ 46 min. Together, these findings unequivocally demonstrate that the terminal phosphate observed in structure **I** resides solely on the 5'-side of the cisplatin adduct. The nature of both digestion products were confirmed by MS analysis (Figure 2.22c and 2.22d). Given that the 1,2-ApG and 1,2-GpG intrastrand cross-links form in the same manner,[4, 38, 39] we may deduce that the terminal phosphate for structure **III** is also located on the 5'-end.

Conclusions

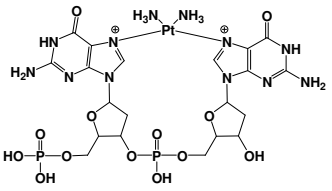
Cisplatin is the most widely used chemotherapeutic agent for the treatment of epithelial malignancies. However, its administration is met with challenges that include severe side effects, a limited breadth of effectiveness against certain common tumors, and an acquired or intrinsic drug resistance. To overcome these limitations, it is necessary to have a quantitative analytical method that can characterize the interactions between cisplatin and biomolecules, especially DNA. The selectivity and sensitivity of LC-MS/MS makes it an ideal tool for this task. In the present study, we prepared ODN substrates housing the light or heavy ($^{15}N_2$) cisplatin-modified 1,2-GpG, 1,2-ApG, or 1,3-GpXpG intrastrand cross-link, and characterized these modified ODNs with ESI-MS and

MS/MS analyses. The MS/MS results confirmed the sites of modification and also revealed that the cisplatin cross-link led to formation of unique fragment ions arising from cleavage at the modification sites.

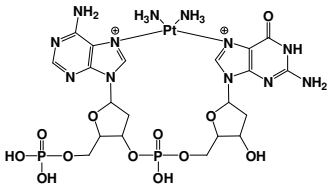
We also digested the cisplatin-modified ODNs with a cocktail of four enzymes and characterized the digestion products using LC-MS/MS. Similar as previously reported,[9, 11-13, 15-20, 22, 31-35] digestion of the 1,2-GpG and 1,2-ApG sequences gave rise to dinucleoside products carrying an internal phosphate (Figure 2.14, structures **I** and **III**, respectively), or containing both an internal and a terminal phosphate (Figure 2.14, structures **II** and **IV**, respectively). By analyzing digestion products of ODNs containing a 1,2-GpG on the 5' or 3' terminus, we demonstrated unambiguously that the terminal phosphate group was located on the 5'-side of the cisplatin adduct. Similar as what were described in the literature,[11-13, 18-20] we also detected the d(G*G*) and $(dG*MP)_2$ products (Figure 2.14, structures **V** and **VII**, respectively) upon digestion of the 1,3-GpXpG cisplatin-containing DNA. Additionally, we reported a new digestion product (structure **VI,** Figure 2.14) for the 1,3-GpXpG cross-link.

The formation of multiple digestion products for each of the three intrastrand cross-link products supports that the accurate quantification of these cisplatin DNA adducts requires the use of ODN substrates housing the stable isotope-labeled lesions. Thus, the availability of ODNs containing stable isotope-labeled standard 1,2-GpG, 1,2-ApG, and 1,3-GpXpG cisplatin-modified DNA adducts and the examination of the enzymatic digestion products facilitated the future LC-MS/MS quantification studies of these DNA lesions in cultured human cells or in patients undergoing cisplatin therapy.

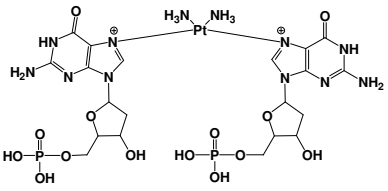
1,2-GpG **1,2-ApG** **1,3-GpXpG**



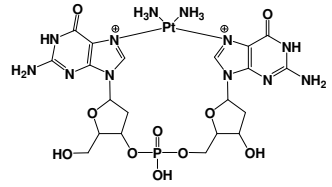
I
***m/z* 904 / 906**



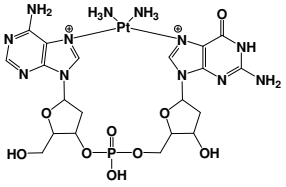
III
***m/z* 888 / 890**



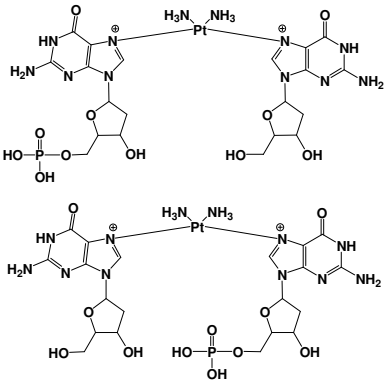
V
***m/z* 922 / 924**



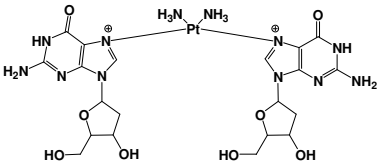
II
***m/z* 824 / 826**



IV
***m/z* 808 / 810**



VI
***m/z* 842 / 844**



VII
***m/z* 762 / 764**

Figure 2.14 Structures and *m/z* values for the $[M+H]^+$ ions for light $[Pt(NH_3)]^{2+}$ / heavy $[Pt(^{15}NH_3)]^{2+}$ cisplatin adducts of the enzymatic digestion products.

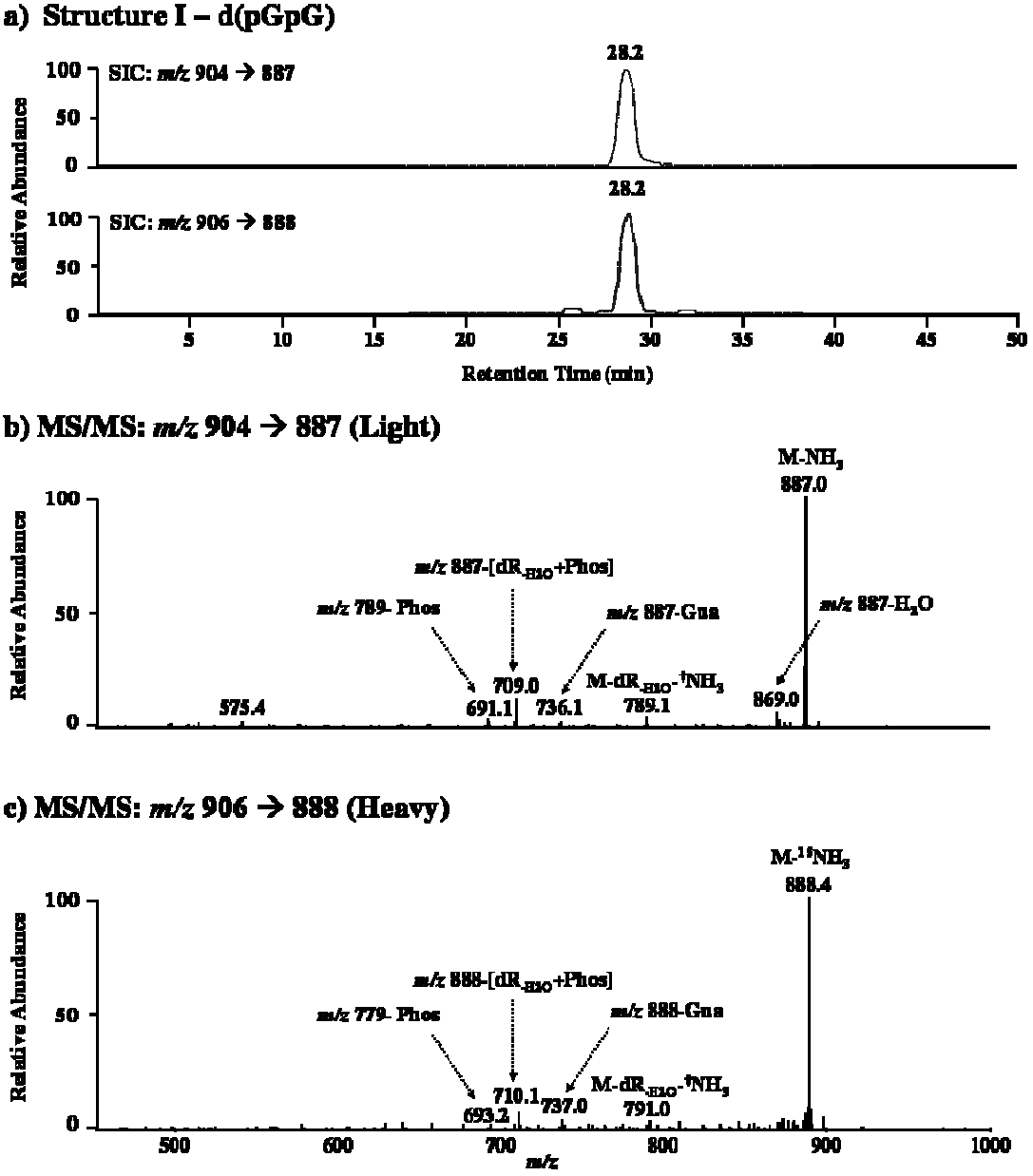
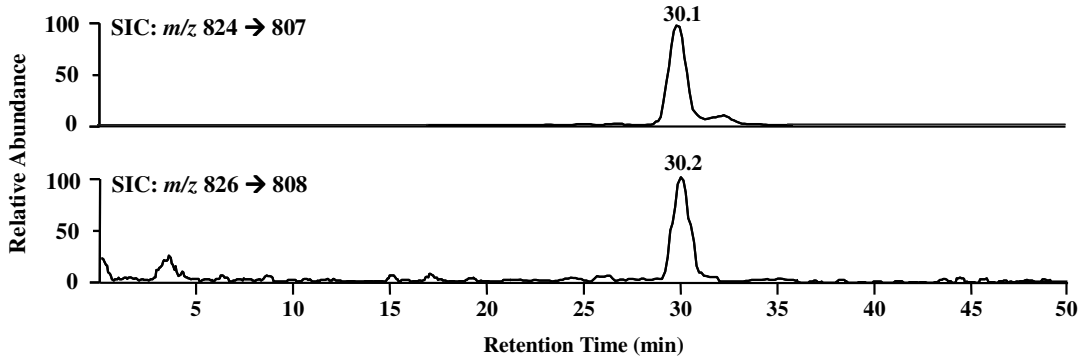
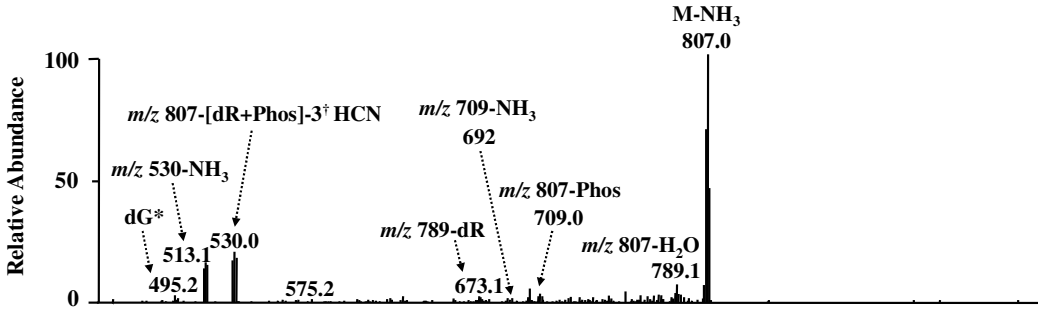


Figure 2.15 Selected-ion chromatograms (a) and MS/MS spectra where † indicates a loss from the nucleobase ring (b,c) of unlabelled (top panel of a, b) and $^{15}N_2$-labelled (bottom panel of a, c) Structure **I** generated from the enzymatic digestion of (ATCCG*G*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$.

a) Structure II – d(GpG)



b) MS/MS: *m/z* 824 → 807 (Light)



c) MS/MS: *m/z* 826 → 808 (Heavy)

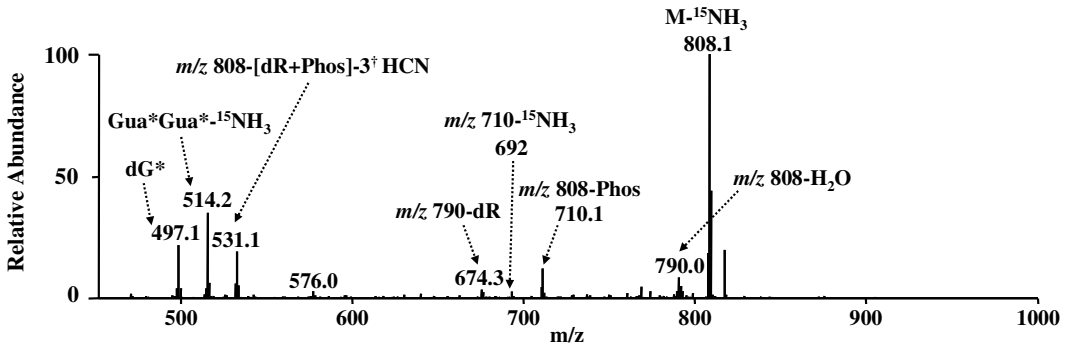


Figure 2.16 Selected-ion chromatograms (a) and MS/MS spectra where † indicates a loss from the nucleobase ring (b,c) of unlabelled (top panel of a, b) and $^{15}N_2$-labelled (bottom panel of a, c) Structure **II** generated from the enzymatic digestion of (ATCCG*G*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$ cisplatin coordination.

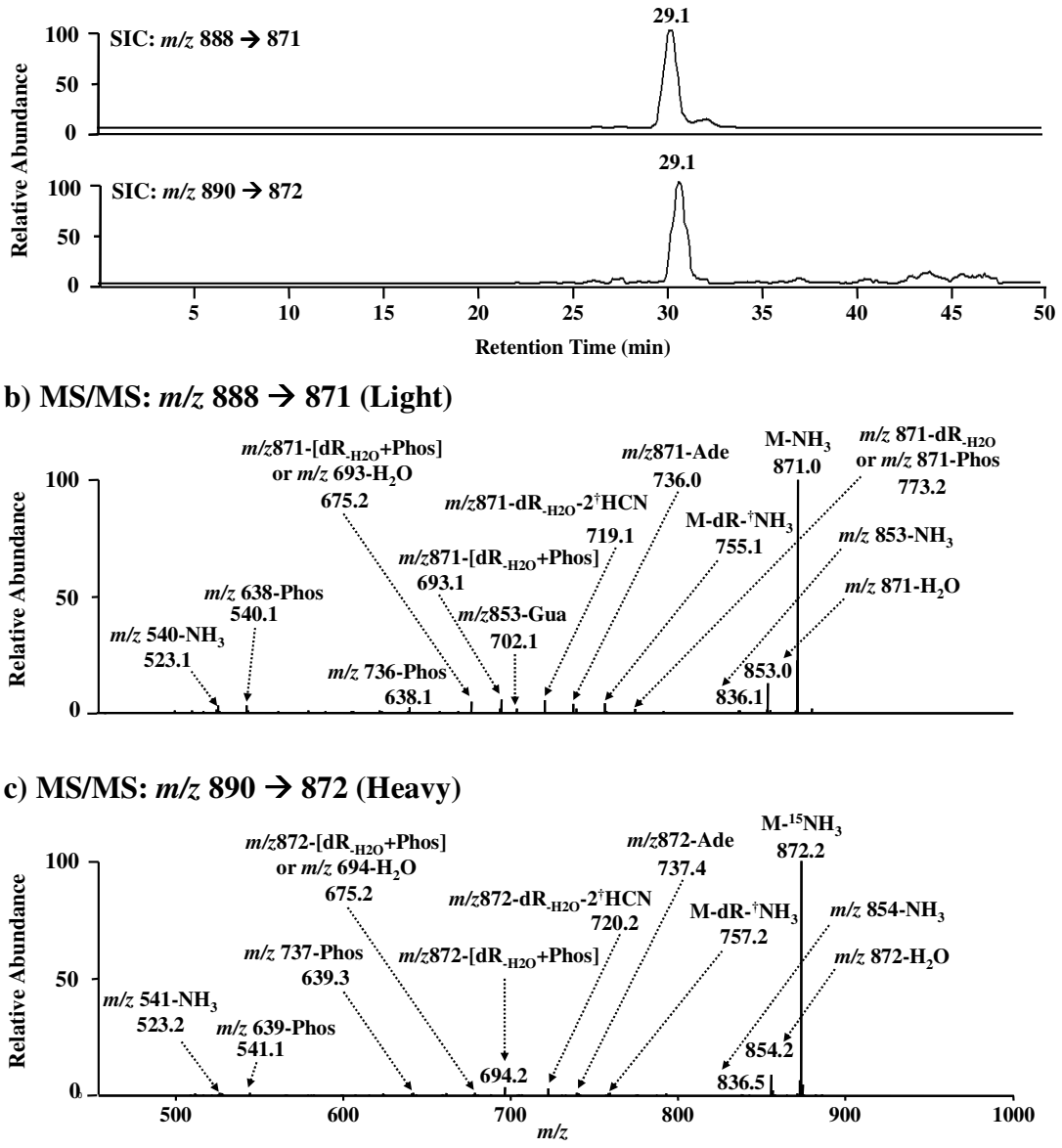


Figure 2.17 Selected-ion chromatograms (a) and MS/MS spectra where † indicates a loss from the nucleobase ring (b,c) of unlabelled (top panel of a, b) and $^{15}N_2$-labelled (bottom panel of a, c) Structure **III** generated from the enzymatic digestion of (ATCCA*G*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$ cisplatin coordination.

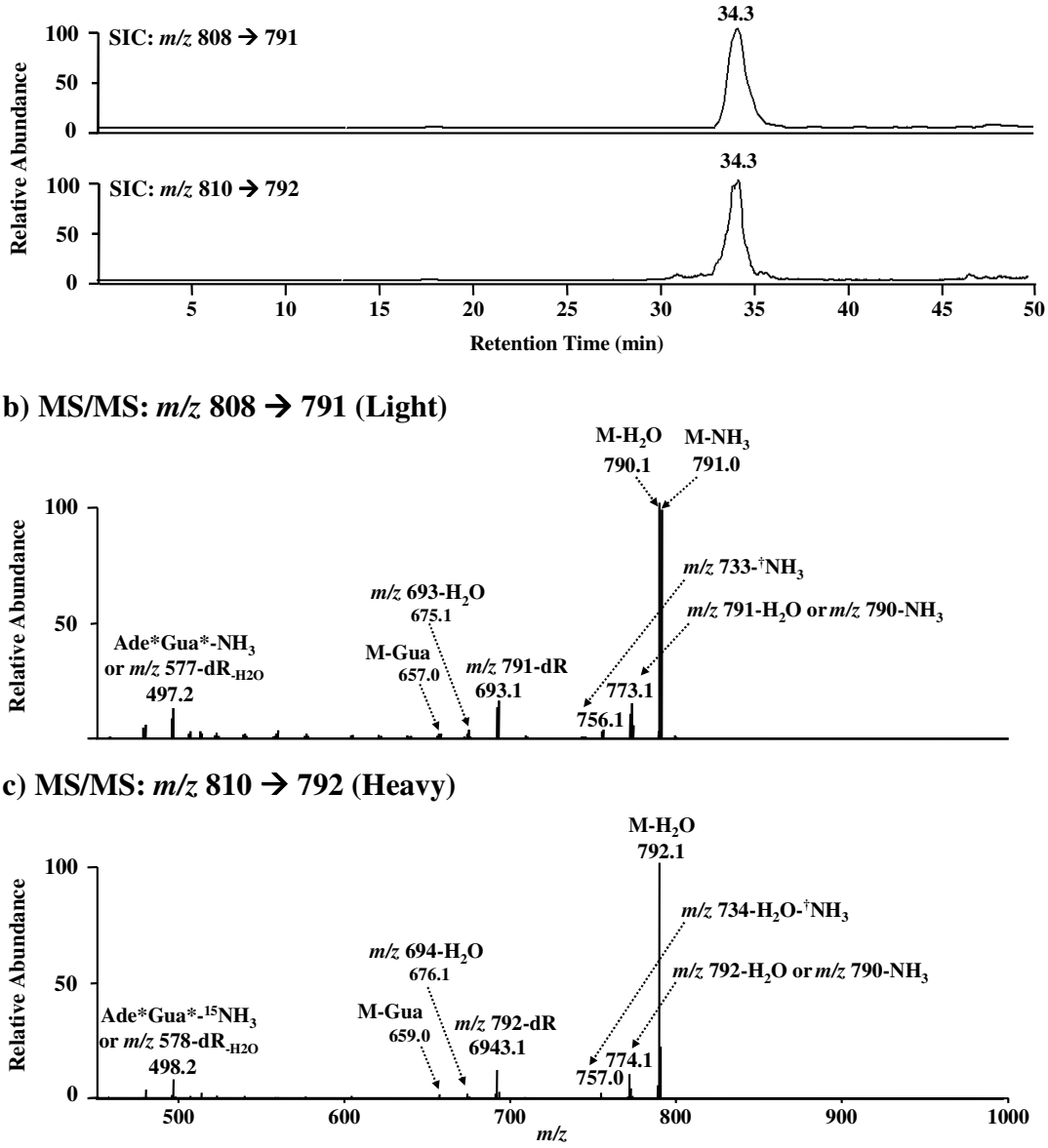


Figure 2.18 Selected-ion chromatograms (a) and MS/MS spectra where † indicates a loss from the nucleobase ring (b,c) of unlabelled (top panel of a, b) and $^{15}N_2$-labelled (bottom panel of a, c) Structure **IV** generated from the enzymatic digestion of (ATCCA*G*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$ cisplatin coordination.

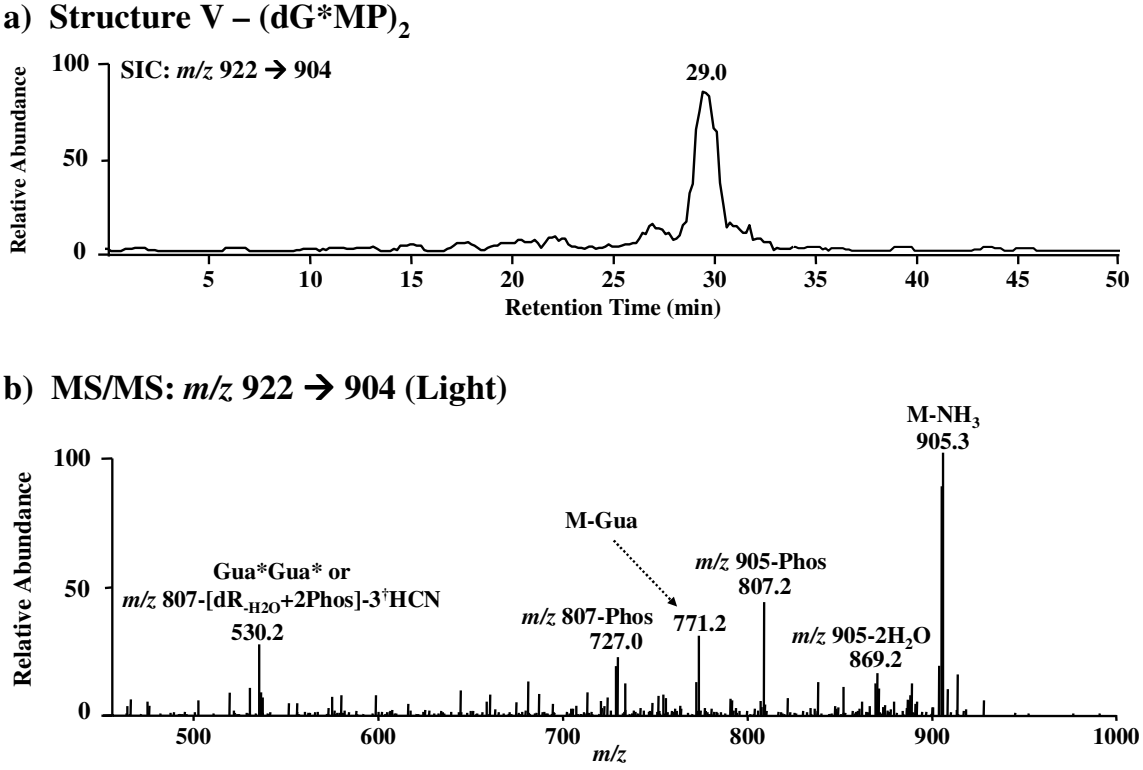


Figure 2.19 Selected-ion chromatogram (a) and MS/MS spectrum where † indicates a loss from the nucleobase ring (b) of unlabelled Structure **V** generated from the enzymatic digestion of (ATCG*CG*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ cisplatin coordination.

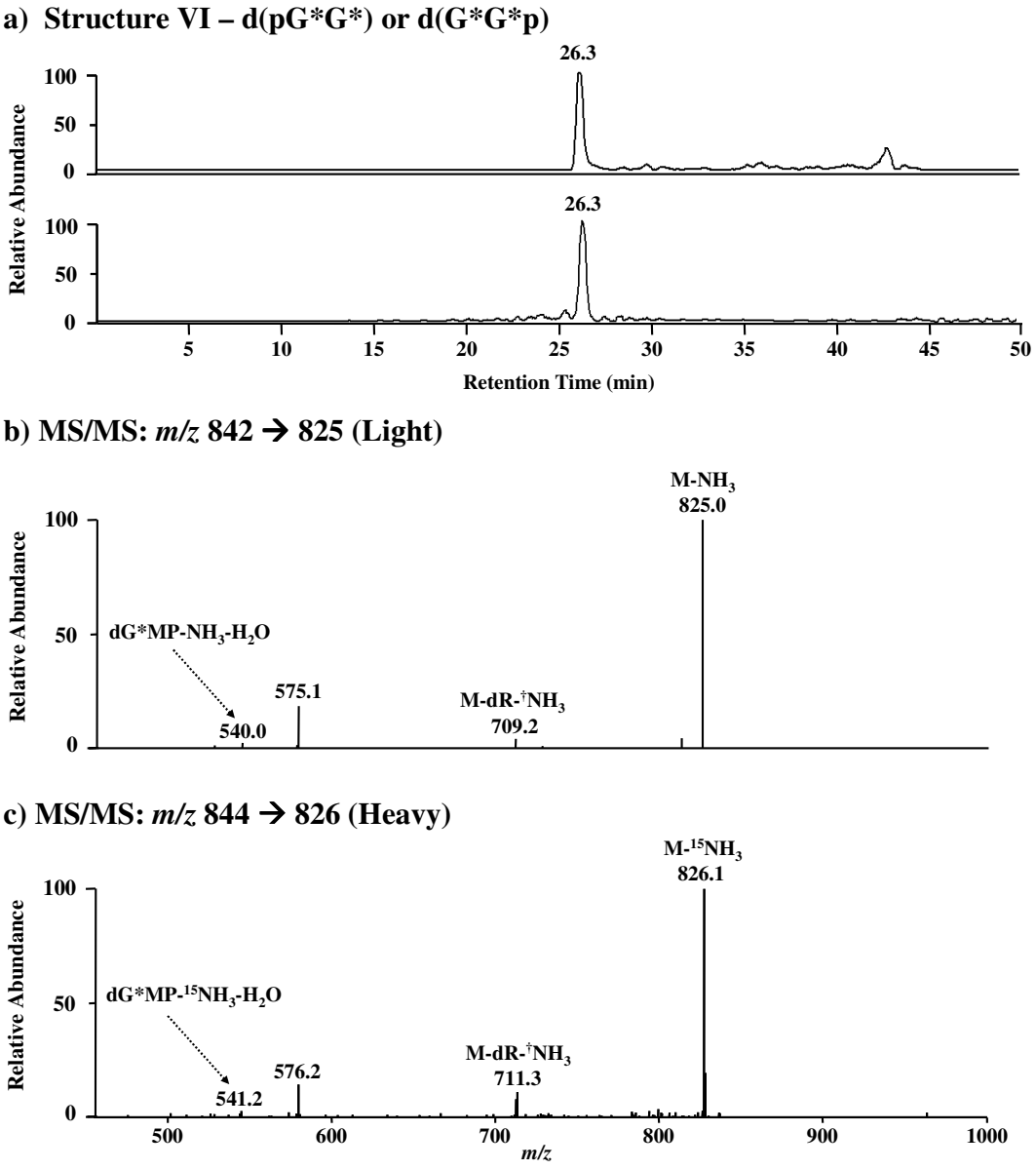


Figure 2.20 Selected-ion chromatograms (a) and MS/MS spectra where † indicates a loss from the nucleobase ring (b,c) of unlabelled (top panel of a, b) and $^{15}N_2$-labelled (bottom panel of a, c) Structure **VI** generated from the enzymatic digestion of (ATCG*CG*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$ cisplatin coordination.

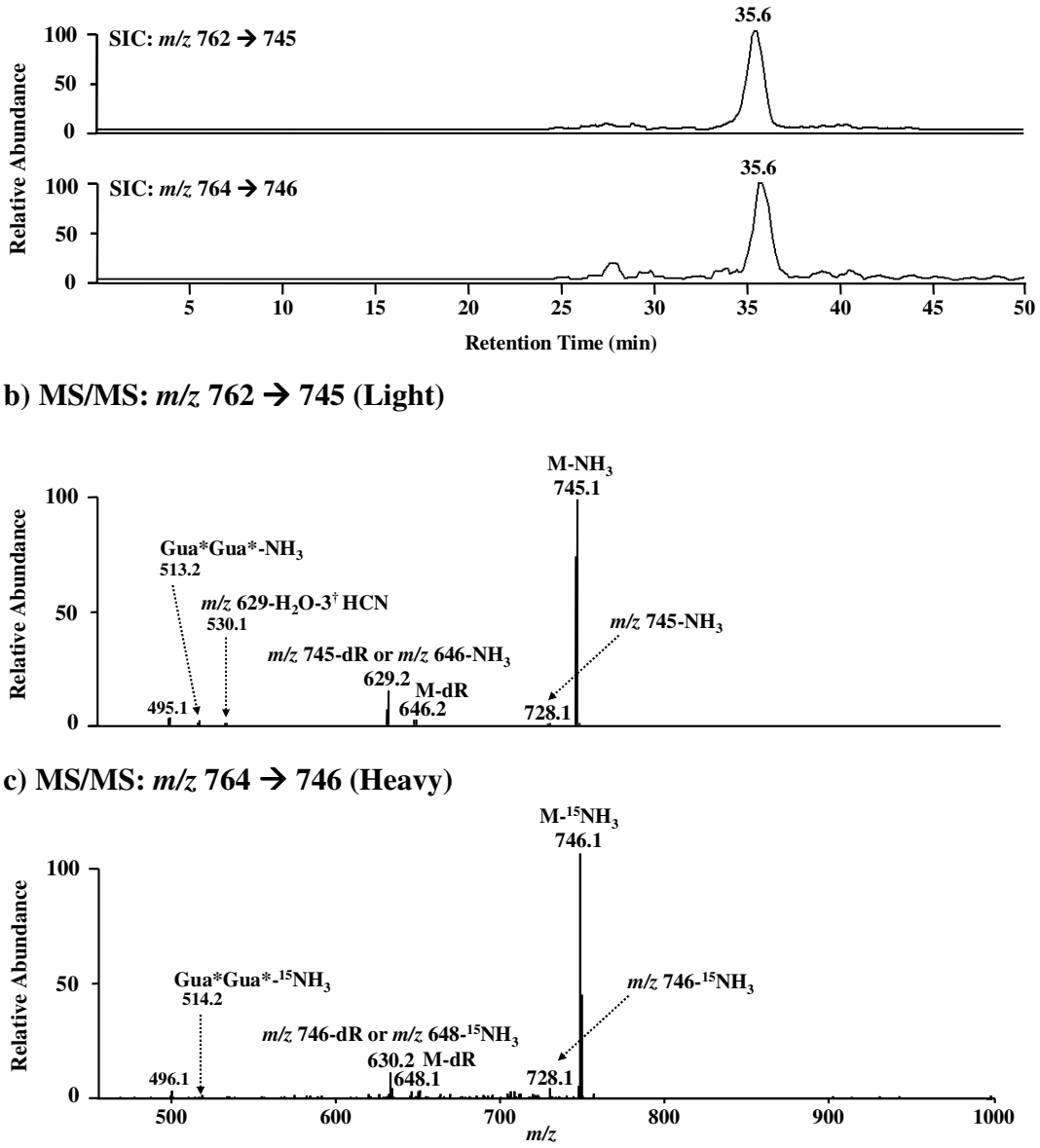


Figure 2.21 Selected-ion chromatograms (a) and MS/MS spectra where † indicates a loss from the nucleobase ring (b,c) of unlabelled (top panel of a, b) and $^{15}N_2$-labelled (bottom panel of a, c) Structure **VII** generated from the enzymatic digestion of (ATCG*CG*CCTA), where * represents light $[Pt(NH_3)_2]^{2+}$ or heavy $[Pt(^{15}NH_3)_2]^{2+}$ cisplatin coordination.

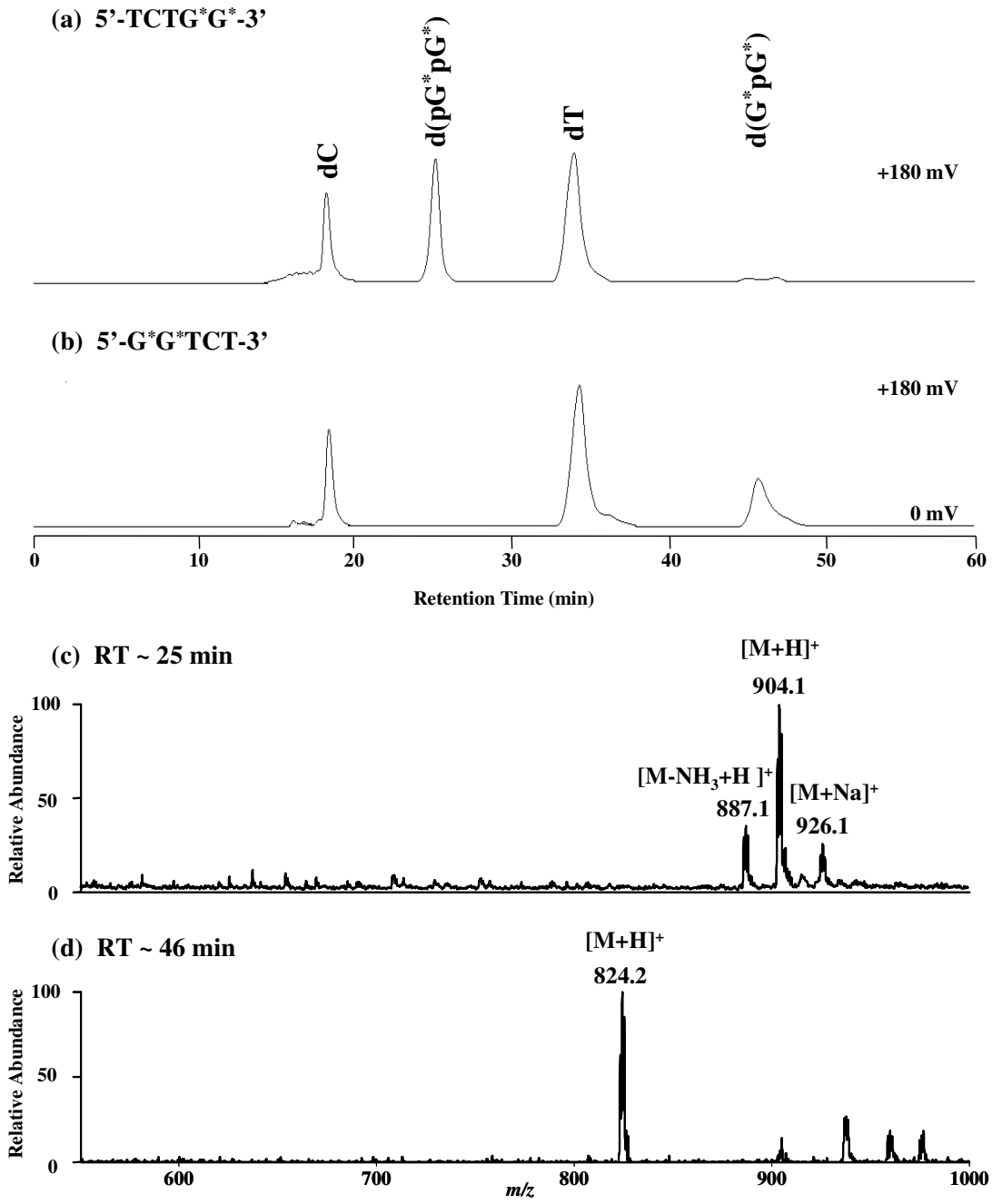


Figure 2.22 HPLC traces for the separation of the enzymatic digestion mixture of synthetic sequences **4** (a) and **5** (b) upon cisplatin treatment; and the corresponding positive-ion ESI-MS of the hydrolysis products d(pG*pG*) (c) and d(G*pG*) (d), where * indicates the coordination of a cisplatin adduct.

References

(1) Jamieson, E. R.; Lippard, S. J. *Chem. Rev.* **1999**, *99*, 2467-2498.

(2) Siddik, Z. H. *Oncogene* **2003**, *22*, 7265-7279.

(3) Jung, Y.; Lippard, S. J. *Chem. Rev.* **2007**, *107*, 1387-1407.

(4) Kozelka, J.; Legendre, F.; Reeder, F.; Chottard, J.-C. *Coord. Chem. Rev.* **1999**, *190-192*, 61-82.

(5) Robertazzi, A.; Platts, J. A. *Chem. Eur. J.* **2006**, *12*, 5747-5756.

(6) Baik, M.-H.; Friesner, R. A.; Lippard, S. J. *J. Am. Chem. Soc.* **2003**, *125*, 14082-14092.

(7) Kartalou, M.; Essigmann, J. M. *Mutat. Res.* **2001**, *478*, 1-21.

(8) Kartalou, M.; Essigmann, J. M. *Mutat. Res.* **2001**, *478*, 23-43.

(9) Baskerville-Abraham, I. M.; Boysen, G.; Troutman, J. M.; Mutlu, E.; Collins, L.; deKrafft, K. E.; Lin, W.; King, C.; Chaney, S. G.; Swenberg, J. A. *Chem. Res. Toxicol.* **2009**, *22*, 905-912.

(10) Da Col, R.; Silvestro, L.; Baiocchi, C.; Giacosa, D.; Viano, I. *J. Chromatogr., A* **1993**, *633*, 119-128.

(11) Eastman, A. *Biochemistry* **1983**, *22*, 3927-3933.

(12) Eastman, A. *Biochemistry* **1986**, *25*, 3912-3915.

(13) Fichtinger-Schepman, A. M. J.; Van der Veer, J. L.; Den Hartog, J. H. J.; Lohman, P. H. M.; Reedijk, J. *Biochemistry* **1985**, *24*, 707-713.

(14) Fichtinger-Schepman, A. M. J.; van Oosterom, A. T.; Lohman, P. H. M.; Berends, F. *Cancer Res.* **1987**, *47*, 3000-3004.

(15) Gupta, R.; Beck, J. L.; Sheil, M. M.; Ralph, S. F. *J. Inorg. Biochem.* **2005**, *99*, 552-559.

(16) Harrington, C. F.; Le Pla, R. C.; Jones, G. D. D.; Thomas, A. L.; Farmer, P. B. *Chem. Res. Toxicol.* **2010**, *23*, 1313-1321.

(17) Iijima, H.; Patrzyc, H. B.; Dawidzik, J. B.; Budzinski, E. E.; Cheng, H. C.; Freund, H. G.; Box, H. C. *Anal. Biochem.* **2004**, *333*, 65-71.

(18) Morrison, J. G.; Bissett, D.; Stephens, I. F. D.; McKay, K.; Brown, R.; Graham, M. A.; Fichtingerschepman, A. M.; Kerr, D. J. *Int. J. Oncol.* **1993**, *2*, 33-37.

(19) Plooy, A. C. M.; Fichtinger-Schepman, A. M. J.; Schutte, H. H.; van Dijk, M.; Lohman, P. H. M. *Carcinogenesis* **1985**, *6*, 561-566.

(20) Pluim, D.; Maliepaard, M.; van Waardenburg, R.; Beijnen, J. H.; Schellens, J. H. M. *Anal. Biochem.* **1999**, *275*, 30-38.

(21) Schneider, S.; Reißner, T.; Ziv, O.; Livneh, Z.; Carell, T. *ChemBioChem* **2010**, *11*, 1521-1524.

(22) Inagaki, K.; kidani, Y. *Inorg. Chim. Acta* **1985**, *106*, 187-191.

(23) Kukushikin, V. Y.; Oskarsson, Å.; Elding, L. I.; Farrell, N. In *Inorganic syntheses*; Dunham, S., Lippard, S. J., Eds.; John Wiley & Sons, Inc., 1998; Vol. 32, pp 141-144.

(24) McLuckey, S. A.; Van Berker, G. J.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 60-70.

(25) Nyakas, A.; Eymann, M.; Schürch, S. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 792-804.

(26) Costello, C. E.; Comess, K. M.; Plaziak, A. S.; Bancroft, D. P.; Lippard, S. J. *Int. J. Mass Spectrom. Ion Processes* **1992**, *122*, 255-279.

(27) Bellon, S.; Gasparutto, D.; Saint-Pierre, C.; Cadet, J. *Org. Biomol. Chem.* **2006**, *4*, 3831-3837.

(28) Hong, H.; Wang, Y. *J. Am. Chem. Soc.* **2005**, *127*, 13969-13977.

(29) Zeng, Y.; Wang, Y. *Nucleic Acids Res.* **2006**, *34*, 6521-6529.

(30) Zeng, Y.; Wang, Y. *Biochemistry* **2007**, *46*, 8189-8195.

(31) Inagaki, K.; Kasuya, K.; Kidani, Y. *Inorg. Chim. Acta* **1984**, *91*, L13-L15.

(32) Garcia Sar, D.; Aguado, L.; Montes Bayon, M.; Comendador, M. A.; Blanco Gonzalez, E.; Sanz-Medel, A.; Sierra, L. M. *Mutat. Res.* **2012**, *741*, 81-88.

(33) Sar, D. G.; Montes-Bayon, M.; Gonzalez, E. B.; Sanz-Medel, A. *J. Anal. At. Spectrom.* **2006**, *21*, 861-868.

(34) Sar, D. G.; Montes-Bayon, M.; Gonzalez, E. B.; Sierra, L. M.; Aguado, L.; Comendador, M. A.; Koellensperger, G.; Hann, S.; Sanz-Medel, A. *Anal. Chem.* **2009**, *81*, 9553-9560.

(35) Sar, D. G.; Montes-Bayon, M.; Ortiz, L. A.; Blanco-Gonzalez, E.; Sierra, L. M.; Sanz-Medel, A. *Anal. Bioanal. Chem.* **2008**, *390*, 37-44.

(36) Denhartog, J. H. J.; Altona, C.; Vandenelst, H.; Vandermarel, G. A.; Reedijk, J. *Inorg. Chem.* **1985**, *24*, 983-986.

(37) Mazeau, K.; Vovelle, F.; Rahmouni, A.; Leng, M.; Ptak, M. *Anti-Cancer Drug Design* **1989**, *4*, 63-78.

(38) Kozelka, J.; Petsko, G. A.; Lippard, S. J.; Quigley, G. J. *Journal of the American Chemical Society* **1985**, *107*, 4079-4081.

(39) Mantri, Y.; Lippard, S. J.; Baik, M.-H. *J. Am. Chem. Soc.* **2007**, *129*, 5023-5030.

Chapter 3

Theoretical Study on the Formation of Intrastrand Cross-link Products from the UV Irradiation of 5-Halopyrimidine-containing DNA

Abstract

Halogenated nucleosides 5-bromo-2'-deoxyuridine (^{Br}dU) and 5-iodo-2'-deoxyuridine (^{I}dU), after being incorporated into cellular DNA, are well known to sensitize cells to UV irradiation. Density functional theory (DFT) calculations at B3LYP/LANL2DZdp level of theory predicted that the gas-phase threshold energy (ΔH in kJ/mol) for dissociative electron attachment (DEA) follows the order of 5-bromouracil (-11.9) > 5-bromocytosine (-17.9) > 5-iodouracil (-36.8), and > 5-iodocytosine (-45.8), suggesting 5-iodo-pyrimidines better facilitate the formation of intrastrand cross-links compared to the 5-bromo-pyrimidine counterparts. Moreover, the generation of cross-links is more favorable (exothermic) with 5-halocytosines. These calculations are in excellent agreement with *in vitro* studies of the photoreactivity and cross-link yield from synthetic DNA harboring 5-X-pyrimidines, where X = Br or I. The DFT calculation results support the mechanistic proposal that cross-link products form from UV-induced electron transfer followed by heterolytic cleavage of the C5-X bond.

Introduction

The halogenated pyrimidine nucleosides, 5-bromo-2'-deoxyuridine (^{Br}dU) and 5-iodo-2'-deoxyuridine (^{I}dU), have been known for decades to be capable of sensitizing

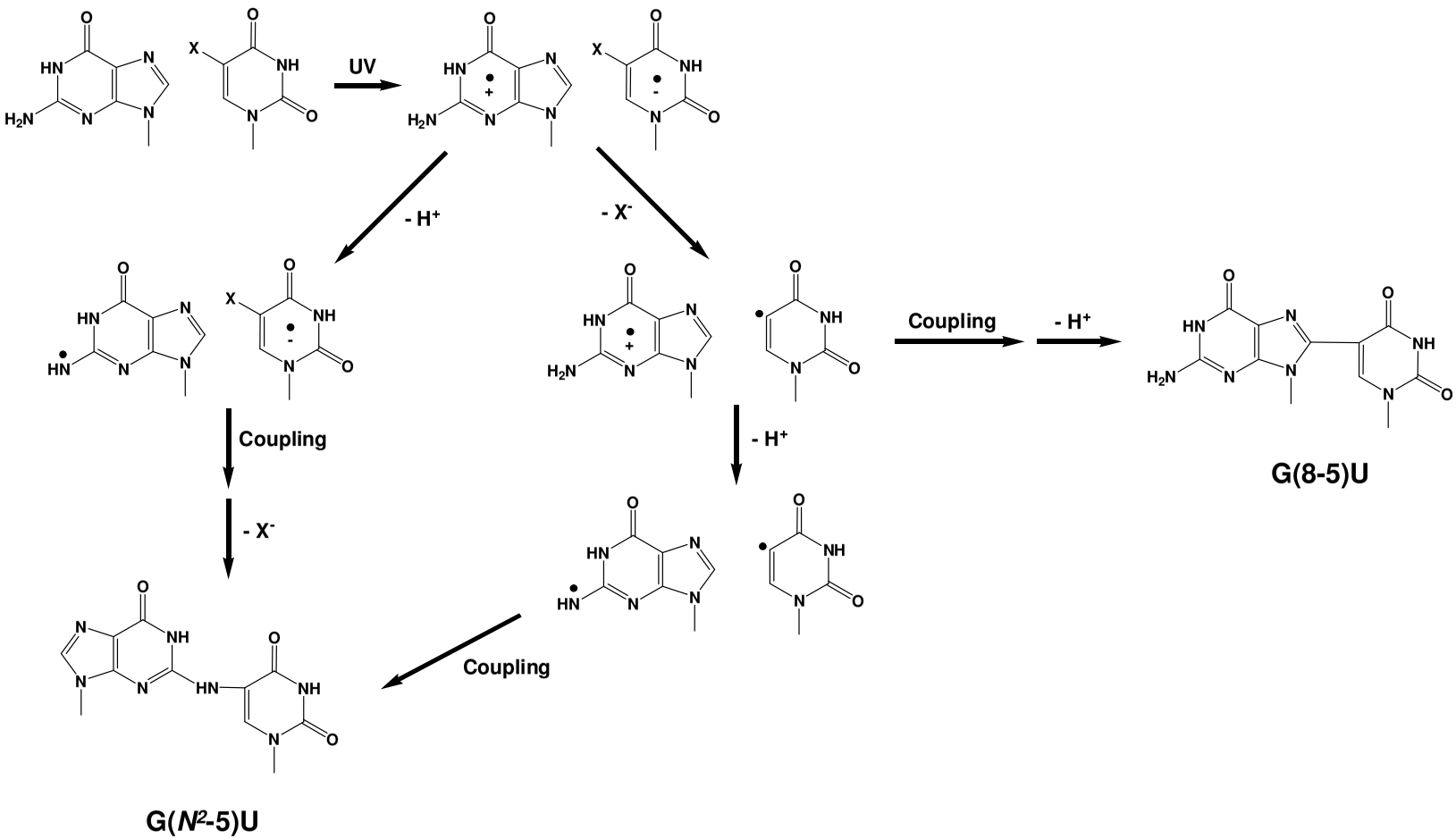
cells to both ionizing radiation[40-47] and photoirradiation.[48-55] In this respect, partial replacement of thymidine with ^{Br}dU or ^{I}dU in DNA increases considerably the amount of strand breaks and alkali-labile sites emanating from exposure to UV light and γ rays.[48, 56-59] Mechanistic studies suggest that the strand breaks arise from the formation of a uracil-5-yl radical and subsequent hydrogen abstraction from the 5'-neighboring 2-deoxyribose as a result of photoirradiation or ionizing radiation.[56, 59, 60] Similarly, it has been demonstrated that the combination of 5-bromo-2'-deoxycytidine (^{Br}dC) with the infection of an adenovirus expressing the Herpes simplex virus thymidine kinase gene could result in significant radiosensitization of rat RT2 glioma cells.[61] Studies have also been extended to other 5-halopyrimidine nucleosides including 5-fluoro-2'-deoxyuridine (^{F}dU),[62] 5-fluoro-2'-deoxycytidine (^{F}dC),[63, 64] and 5-chloro-2'-deoxycytidine (^{Cl}dC),[65-69] where halogenated dC was shown to afford similar degrees of radiosensitization as the corresponding halogenated dU when incorporated into cellular DNA.[70]

Recently, we examined the photoproduct formation from the UVB irradiation of ^{Br}dU- and ^{Br}dC-harboring duplex DNA *in vitro*.[28, 29, 71] We demonstrated that, in isolated DNA, there exists a sequence-dependant formation of the intrastrand cross-links d(U[5-8]G), d(U[5-N^2]G), d(G[8-5]U), d(C[5-8]G), d(C[5-N^2]G), and d(G[8-5]C), along with the presence of other cross-link products including adenine being conjugated with uracil or cytosine.[28, 29, 71] Similar cross-link products with adenine, and to a greater extent, the loss of adenine to form 2'-deoxyribonolactone (dL) have also been reported.[72-75] In addition, our laboratory showed that many of the same cross-links are produced in cultured human cells treated with ^{Br}dU and exposed to UVB light.[30] Replication studies

revealed that intrastrand cross-links are capable of blocking DNA synthesis by replicative DNA polymerases *in vitro*[27, 76] and lead to error-prone replication by a translesion synthesis DNA polymerase (yeast pol η)[77] and in *Escherichia coli* (*E. coli*) cells.[78] Moreover, studies demonstrated that other structurally related intrastrand cross-link lesions could be recognized readily by *E. coli* nucleotide excision repair proteins, i.e., the UvrABC excision nuclease.[79, 80]

Previously, Zeng et al. [29] proposed that cross-link formation initiates from the transferring of an electron from the neighboring purine base to the 5-halopyrimidine (Scheme 3.1). The resulting radical anion of the halogenated pyrimidine base can then eliminate a bromide to yield a pyrimidin-5-yl radical. The radical can subsequently couple with the C8 of the neighboring purine base, and the resultant products can lose an electron and deprotonate to yield the intrastrand cross-link d(Y[8-5]Z), where Y = G or A, and Z = U or C. Assuming that the proposed mechanism is valid, it follows that the formation of the cross-link product would depend on the ease of electron transfer and the dissociation energy of the carbon-halogen bond. Together, these two parameters represent the threshold energy (E_{Th}) for the dissociative electron attachment (DEA).

Herein, we use B3LYP/LANL2DZdp density functional theory (DFT) calculations to assess the thermodynamics of the photochemically induced DEA reaction. We place our emphasis on a comparison of reported photoproduct yields from various 5-halopyrimidines to our calculated values as it pertains to the mechanisms of cross-link formation.



Scheme 3.1 Proposed dissociative electron attachment mechanism for the formation of G^U intrastrand cross-link products.

Computational Methods

A B3LYP functional with the LANL2DZdp basis set[81] was employed for all calculations using DFT methods included in the Gaussian 09W package.[82] LANL2DZdp contains diffuse and polarization functions that have been optimized for the LANL2DZ basis set for elements in groups 14-17, and it is among a limited number of available basis sets that can be used for both anionic and iodine-containing species. We found that B3LYP/LANL2DZdp yielded results in good agreement with B3LYP/6-31+G(d)[83] and B3LYP/6-311+G(2df,p) with 6-31+G(d) geometry optimizations[84] in the gas phase calculations on 5-substituted pyrimidine bases analogous to those being examined herein (Table 3.1).

Geometry optimizations and frequency (without scaling) were performed in the gas phase at the B3LYP/LANL2DZdp level of theory. Polarizable continuum model (PCM) frequency calculations were also completed on gas-phase geometries in water (ε=78) to mimic an *in vitro* environment. Threshold energy (E_{Th}, Equation 3.1), which thermodynamically describes the dissociative electron attachment (DEA) for the 5-substituted pyrimidines (5-X-Pyr), was determined as the difference between the adiabatic heterolytic bond dissociation (D) and adiabatic electron attachment (EA),

$$\mathbf{E_{Th} = D - EA = \left[\left(E_{Pyr^{\bullet}} + E_{X^-} \right) - E_{5\text{-}X\text{-}Pyr^{\bullet -}} \right] - \left[E_{5\text{-}X\text{-}Pyr} - E_{5\text{-}X\text{-}Pyr^{\bullet -}} \right]} \qquad \textbf{Equation 3.1}$$

where E was calculated in terms of ΔH (sum of electronic and thermal enthalpies + thermal correction to enthalphy; kJ/mol), ΔG (sum of electronic and thermal free energies + thermal correction to Gibbs free energy; kJ/mol), and $\Delta S = -[(\Delta G - \Delta H)/T$; kJ/mol-K].

Results

Validation of Functional and Basis Set

Only a handful of studies have previously examined the theoretical electron affinity, and to a lesser degree, the heterolytic bond dissociation energy of 5-X-uracils, where X = F, Cl, or Br.[83-86] Given the inclusion of 5-iodouracil in this work, it was necessary to use the B3LYP/LANL2DZdp level of theory. In order to assess the validity of applying the LANL2DZdp basis set to the molecules described herein, we calculated the E_{Th} of 5-substituted uracils using the EA and D values generated from *ab initio* [B3LYP/6-31+G(d) and B3LYP/6-311+G(2df,p) with 6-31+G(d) geometry optimization (referred to as B3LYP/6-311+G(2df,p) in text)] and semi-empirical [CURES-EC (Configuration interaction or Unrestricted orbitals to Related Experimental quantities to Self consistent field values by estimating Electron Correlation)] calculations reported in the literature (Table 3.1).[83-85] To our knowledge, no analogous computational data are available for 5-halocytosines. Although B3LYP/LANL2DZdp yields EA values notably larger than both of the *ab initio* methods considered, the E_{Th} values are comparable to B3LYP/6-311+G(2df,p), and the overall trend to B3LYP/6-311+G(2df,p) and CURES-EC (Table 3.1). As such, we believe B3LYP/LANL2DZdp to be a reasonable level of

Table 3.1 Comparison of the adiabatic electron affinity (EA), heterolytic bond dissociation energy (D), and threshold energy (E_{Th}) of 5-substituted uracils in the gas-phase (in eV).

	Electron Affinity *EA* (eV)				**Heterolytic Bond Dissociation** *D* (eV)				**Threshhold Energy** E_{Th} (eV)			
Method	B3LYP	B3LYP	B3LYP	CURES-EC	B3LYP	B3LYP	B3LYP	CURES-EC	B3LYP	B3LYP	B3LYP	CURES-EC
basis set:	LANL2DZdp	6-311+G(2df,p)	6-31+G(d)	(adiabatic)	LANL2DZdp	6-311+G(2df,p)	6-31+G(d)	(adiabatic)	LANL2DZdp	6-311+G(2df,p)	6-31+G(d)	(adiabatic)
Uracil	**0.26**	**0.18**	**0.18**	**(0.82)**	**4.2**				**3.9**			
5-Fluorouracil	*0.54*	*0.45*	*0.48*	*3.4 (1.0)*	*2.2*	*2.2*	*2.2*	*5.2*	*1.7*	*1.8*	*1.7*	*1.8*
5-Chlorouracil	*0.67*	*0.58*	*0.60*	*3.6 (1.1)*	*0.85*	*0.87*	*0.78*	*4.0*	*0.18*	*0.29*	*0.18*	*0.38*
5-Bromouracil	*0.69*	*0.64*	*0.63*	*3.4 (1.2)*	*0.60*	*0.55*	*0.91*	*3.4*	*-0.09*	*-0.09*	*0.28*	*0.06*
5-Iodouracil	*0.72*			*3.1 (1.2)*	*0.36*			*2.9*	*-0.36*			*-0.16*
Thymine	**0.22**	**0.14**			**4.6**				**4.38**			

theory for calculating the E_{Th} of 5-X-uracils and 5-X-cytosines where X =F, Cl, Br, and I.

5-X-Uracils

We explored the thermodynamic properties of cross-link product formation using a series of B3LYP/LANL2DZdp calculations. Specifically, we focused on the threshold energy, E_{Th}, for the DEA reaction mechanism as a function of C5-X substitution. As shown in Table 3.2a, E_{Th} values for 5-X-uracil in terms of ΔH (in kJ/mol) in the gas phase are CH_3 (399) > H (384) > F (159) >> Cl (13.8) > Br (-11.9) > I (-36.8). These calculations suggest that the DEA reaction with ^{I}U is three times as exothermic as that with ^{Br}U. In solution (water), we observed an analogous trend; however, due to solvation effects, ^{I}U (-281 kJ/mol) was only slightly more exothermic than ^{Br}U (-278 kJ/mol). In addition, these two nucleosides shared comparable ΔG and ΔS values (Table 3.2b).

5-X-Cytosines

Although no experimental electron affinity or bond dissociation energy data are available for comparison, we extended our calculations to include 5-X-cytosines given the effectiveness of B3LYP/LANL2DZdp in describing 5-X-uracils compared to other levels of theory (Table 3.1). We report that E_{Th} for 5-X-cytosine in terms of ΔH (in kJ/mol) in the gas phase are CH_3 (377) > H (369) > F (154) >> Cl (8.8) > Br (-17.9) > I (-45.8), again suggesting that DEA and cross-link formation are thermodynamically favorable with ^{Br}C and ^{I}C. Similar to the 5-X-uracil calculations, the reaction with ^{I}C is approximately 2.5 times as exothermic than with ^{Br}C. In solution, we again noted a

similar trend among C5-substituents as was observed with gas-phase calculations, albeit, ^{I}C (-291 kJ/mol) was only slightly more exothermic than ^{Br}C (-285 kJ/mol) with comparable ΔG and ΔS values (Table 3.2b).

Discussion

In our earlier work, we proposed that cross-link products form via a DEA reaction that involves UV-induced electron transfer followed by heterolytic cleavage of the C5-X bond (Scheme 3.1).[29] In this study, we calculated the E_{Th} for the DEA of 5-X-pyrimidines where X = F, Cl, Br, and I. We found that the E_{Th} for ^{I}dU and ^{I}dC have a greater exothermicity than ^{Br}dU and ^{Br}dC, respectively. Based on a report by Abdoul-Carime et al.,[87] in which the cross-sections of halide ions released from 5-X-uracils, where X = F, Cl, Br, or I, upon electron transfer were measured, there was approximately two-fold more iodide formation than that of bromide, suggesting that ^{I}dU, and by extension ^{I}dC, are more efficient than ^{Br}dU and ^{Br}dC at facilitating intrastrand cross-link formation, respectively. Additionally, it was demonstrated that upon UVB irradiation (λ = 302-315 nm), heterolysis of the C5-X bond occurs in ^{Br}dU- and ^{I}dU- bearing DNA.[59, 73, 88] It was found that when guanine, which serves as an electron donor, is adjacently located 5' upstream from ^{X}dU (X= Br or I), ^{I}dU exhibited the greatest photoreactivity. This suggests that the rate of photocrosslinking is also greatest with ^{I}dU-containing DNA. Again, assuming cross-link formation is a function of electron transfer and heterolysis of the C5-X bond, the computational values reported herein correlate well with these studies.

Table 3.2 Adiabatic electron affinity (*EA*), heterolytic bond dissociation energy (*D*), and threshold energy (E_{Th}) of 5-substituted pyrimidines in the gas-phase (a), and water environment (b).

(a) Gas-phase

	Electron Affinity *EA*			Heterolytic Bond Dissociation *D*			Threshold Energy E_{Th}		
Method: B3LYP basis set: lan12dzdp	ΔH kJ/mol	ΔG kJ/mol	ΔS kJ/mol-K	ΔH kJ/mol	ΔG kJ/mol	ΔS kJ/mol-K	ΔH kJ/mol	ΔG kJ/mol	ΔS kJ/mol-K
Uracil	**41.9**	**43.9**	**-2.1**	**425.9**	**358.6**	**67.3**	**384.1**	**314.7**	**69.4**
5-Fluorouracil	*66.9*	*70.8*	*-3.9*	*226.0*	*150.1*	*75.9*	*159.1*	*79.3*	*79.8*
5-Chlorouracil	*79.2*	*82.5*	*-3.3*	*93.1*	*18.7*	*74.4*	*13.8*	*-63.9*	*77.7*
5-Bromouracil	*81.4*	*85.3*	*-3.8*	*69.5*	*-3.4*	*72.9*	*-11.9*	*-88.7*	*76.8*
5-Iodouracil	*83.7*	*88.3*	*-4.7*	*46.9*	*-23.7*	*70.6*	*-36.8*	*-112.0*	*75.2*
Thymine	**37.1**	**42.3**	**-5.2**	**436.2**	**340.9**	**95.3**	**399.1**	**298.6**	**100.5**
Cytosine	**282.4**	**187.9**	**94.5**	**651.0**	**486.6**	**164.4**	**368.6**	**298.7**	**69.9**
5-Methylcytosine	*16.4*	*30.3*	*-13.9*	*393.7*	*304.6*	*89.1*	*377.3*	*274.3*	*103.0*
5-Fluorocytosine	*42.6*	*51.7*	*-9.1*	*196.7*	*126.6*	*70.1*	*154.1*	*74.9*	*79.2*
5-Chlorocytosine	*51.3*	*60.9*	*-9.6*	*60.2*	*-6.8*	*67.0*	*8.8*	*-67.7*	*76.6*
5-Bromocytosine	*116.3*	*133.8*	*-17.5*	*98.3*	*40.2*	*58.1*	*-17.9*	*-93.6*	*75.7*
5-Iodocytosine	*138.2*	*155.7*	*-17.5*	*92.4*	*35.5*	*56.9*	*-45.8*	*-120.2*	*74.4*

(b) Solution-phase (Water)

	Electron Affinity *EA*			Heterolytic Bond Dissociation *D*			Threshold Energy E_{Th}		
Method: B3LYP basis set: lan12dzdp	ΔH kJ/mol	ΔG kJ/mol	ΔS kJ/mol-K	ΔH kJ/mol	ΔG kJ/mol	ΔS kJ/mol-K	ΔH kJ/mol	ΔG kJ/mol	ΔS kJ/mol-K
Uracil	**191.9**	**185.4**	**6.5**	**42.2**	**-33.7**	**75.9**	**-149.7**	**-219.0**	**69.3**
5-Fluorouracil	*250.5*	*254.2*	*-3.7*	*66.2*	*-10.3*	*76.5*	*-184.3*	*-264.5*	*80.2*
5-Chlorouracil	*256.3*	*259.4*	*-3.1*	*-15.1*	*-90.1*	*75.0*	*-271.4*	*-349.5*	*78.0*
5-Bromouracil	*256.3*	*260.0*	*-3.7*	*-21.3*	*-94.8*	*73.6*	*-277.6*	*-354.9*	*77.3*
5-Iodouracil	*254.7*	*258.8*	*-4.1*	*-26.3*	*-98.0*	*71.7*	*-281.0*	*-356.8*	*75.8*
Thymine	**224.0**	**249.4**	**-25.4**	**352.6**	**271.1**	**81.5**	**128.7**	**21.7**	**106.9**
Cytosine	**203.0**	**213.8**	**-10.8**	**244.1**	**184.4**	**59.7**	**41.1**	**-29.4**	**70.5**
5-Methylcytosine	*199.5*	*211.4*	*-11.9*	*321.8*	*229.7*	*92.1*	*122.3*	*18.3*	*104.0*
5-Fluorocytosine	*225.8*	*232.6*	*-6.9*	*36.0*	*-35.3*	*71.3*	*-189.8*	*-267.9*	*78.2*
5-Chlorocytosine	*229.2*	*234.1*	*-4.9*	*-49.2*	*-117.6*	*68.4*	*-278.4*	*-351.7*	*73.3*
5-Bromocytosine	*296.2*	*310.6*	*-14.4*	*11.0*	*-47.5*	*58.5*	*-285.2*	*-358.1*	*72.9*
5-Iodocytosine	*305.7*	*319.6*	*-14.0*	*14.8*	*-42.7*	*57.6*	*-290.8*	*-362.4*	*71.5*

We also found that the E_{Th} for 5-halocytosines were 5-9 kJ/mol lower in energy than the corresponding 5-halouracils, suggesting that cross-link formation is more favorable with 5-halocytosines. In an earlier study, in which the yields for the cross-link products generated from the UVB irradiation (λ=313 nm) of ^{Br}dU- and ^{Br}dC-containing synthetic DNA were reported, it was found that the major cross-link product, d(Y[8-5]Z) (Y = G or A and Z = U or C), where the C8 of the purine is covalently bonded with the C5 of its neighboring 3' pyrimidine base, was 13- (Y = G) and 20- (Y = A) times more abundant with ^{Br}dC-harboring sequences.[29] Moreover, 5'-G^{Br}Z-3' sequences yielded 20- (Z = C) to 30- (Z = U) times more photoproduct than 5'-A^{Br}Z-3', which follows given that the ionization potential of guanine [7.77 eV (experimental)[89]; 7.64 eV (theoretical)[86]] is lower than adenine [8.26 eV (experimental)[89]; 8.09 eV (theoretical)[86]]. Together, these findings correlate well with the observed trends reported in this work, and support the mechanistic proposal that photocrosslinking occurs through a dissociative electron attachment reaction.

Conclusions

In this chapter, we investigated the thermodynamics of photocrosslink products of 5-halopyrimidines as it pertains to their respective mechanisms of formation. We performed B3LYP/LANL2DZdp threshold energy calculations to rationalize reported experimental findings, and support the proposed mechanism that cross-link formation involves electron transfer from a nearby purine base followed by heterolytic cleavage of the C5-X bond. To the extent that various intrastrand cross-links can block DNA

replication and introduce errors during translesion synthesis,[76, 77] the formation of such photoproducts may account, at least in part, for the enhanced sensitivity of living cells to UV irradiation when thymidine and 2'-deoxycytidine in DNA are substituted with ^{X}dU and ^{X}dC (X = Br or I), respectively.[2,9,14-16,31-35] The work herein strengthens our mechanistic understanding of how intrastrand cross-links are formed.

References

(1) Dewey, W. C.; Humphrey, R. M. *Radiat. Res.* **1965**, *26*, 538-553.

(2) Erikson, R. L.; Szybalski, W. *Cancer Res.* **1963**, *23*, 122-130.

(3) Lett, J. T.; Parkins, G.; Alexader, P.; Ormerod, M. G. *Nature* **1964**, *203*, 593-596.

(4) Ling, L. L.; Ward, J. F. *Radiat. Res.* **1990**, *121*, 76-83.

(5) Sano, K.; Hoshino, T.; Nagai, M. *J. Neurosurg.* **1968**, *28*, 530-538.

(6) Stahl, F. W.; Crasemann, J. M.; Okun, L.; Fox, E.; Laird, C. *Virology* **1961**, *13*, 98-104.

(7) Webb, C. F.; Jones, G. D. D.; Ward, J. F.; Moyer, D. J.; Aguilera, J. A.; Ling, L. L. *Int. J. Radiat. Biol.* **1993**, *64*, 695-705.

(8) Zamenhof, S.; De Giovanni, R.; Greer, S. *Nature* **1958**, *181*, 827-829.

(9) Cadet, J.; Vigny, P. *Bioinorganic photochemistry*, In Morrison, H. ed.; John Wiley: New York, 1990.

(10) Chu, E. H. Y. *Mutat. Res. Fund. Mol. Mech. Mut.* **1965**, *2*, 75-94.

(11) Hutchinson .F *Q. Rev. Biophys.* **1973**, *6*, 201-246.

(12) Hutchinson F. and Kohnlein, W. *Prog. Mol. Subcell. Biol.* **1980**, *7*, 1-42.

(13) Ikushima, T.; Wolff, S. *Exp. Cell Res.* **1974**, *87*, 15-19.

(14) Ribas, M.; Korenberg, J. R.; Peretti, D.; Pichiri, G.; Stockert, J. C.; Gosalvez, J.; Mezzanotte, R. *Chromosome Res.* **1994**, *2*, 428-438.

(15) Taguchi, T.; Shiraishi, Y. *Mutat. Res.* **1989**, *211*, 43-49.

(16) Wojcik, A.; von Sonntag, C.; Obe, G. *J. Photochem. Photobiol. B* **2003**, *69*, 139-144.

(17) Chen, T. Q.; Cook, G. P.; Koppisch, A. T.; Greenberg, M. M. *J. Am. Chem. Soc.* **2000**, *122*, 3861-3866.

(18) Fuciarelli, A. F.; Sisk, E. C.; Zimbrick, J. D. *Int. J. Radiat. Biol.* **1994**, *65*, 409-418.

(19) Limoli, C. L.; Wu, C. C. L.; Milligan, J. R.; Ward, J. F. *Mutagenesis* **1997**, *12*, 443-447.

(20) Sugiyama, H.; Tsutsumi, Y.; Saito, I. *J. Am. Chem. Soc.* **1990**, *112*, 6720-6721.

(21) Cook, G. P.; Chen, T.; Koppisch, A. T.; Greenberg, M. M. *Chem. Biol.* **1999**, *6*, 451-459.

(22) Brust, D.; Feden, J.; Farnsworth, J.; Amir, C.; Broaddus, W. C.; Valerie, K. *Cancer Gene Ther.* **2000**, *7*, 778-788.

(23) Heidelberger, C.; Chaudhuri, N. K.; Danneberg, P.; Mooren, D.; Griesbach, L.; Duschinsky, R.; Schnitzer, R. J.; Pleven, E.; Scheiner, J. *Nature* **1957**, *179*, 663-666.

(24) Mekras, J. A.; Boothman, D. A.; Perez, L. M.; Greer, S. *Cancer Res.* **1984**, *44*, 2551-2560.

(25) Kaysen, J.; Spriggs, D.; Kufe, D. *Cancer Res.* **1986**, *46*, 4534-4538.

(26) Hale, J. T.; Bigelow, J. C.; Mathews, L. A.; McCormack, J. J. *Biochem. Pharmacol.* **2002**, *64*, 1493-1502.

(27) Perez, L. M.; Greer, S. *Int. J. Radiat. Oncol., Biol., Phys.* **1986**, *12*, 1523-1527.

(28) Perez, L. M.; Mekras, J. A.; Briggle, T. V.; Greer, S. *Int. J. Radiat. Oncol., Biol., Phys.* **1984**, *10*, 1453-1458.

(29) Russell, K. J.; Rice, G. C.; Brown, J. M. *Cancer Res.* **1986**, *46*, 2883-2887.

(30) Santos, O.; Perez, L. M.; Briggle, T. V.; Boothman, D. A.; Greer, S. B. *Int. J. Radiat. Oncol., Biol., Phys.* **1990**, *19*, 357-365.

(31) Erikson, R. L.; Szybalski, W. *Radiat. Res.* **1963**, *20*, 252-262.

(32) Zeng, Y.; Wang, Y. *J. Am. Chem. Soc.* **2004**, *126*, 6552-6553.

(33) Zeng, Y.; Wang, Y. *Nucleic Acids Res.* **2006**, *34*, 6521-6529.

(34) Hong, H.; Wang, Y. *J. Am. Chem. Soc.* **2005**, *127*, 13969-13977.

(35) Saito, I. *Pure Appl. Chem.* **1992**, *64*, 1305-1310.

(36) Tashiro, R.; Nakamura, K.; Sugiyama, H. *Tetrahedron Lett.* **2008**, *49*, 428-431.

(37) Watanabe, T.; Tashiro, R.; Sugiyama, H. *J. Am. Chem. Soc.* **2007**, *129*, 8163-8168.

(38) Watanabe, T.; Bando, T.; Xu, Y.; Tashiro, R.; Sugiyama, H. *J. Am. Chem. Soc.* **2005**, *127*, 44-45.

(39) Zeng, Y.; Wang, Y. *Biochemistry* **2007**, *46*, 8189-8195.

(40) Gu, C.; Wang, Y. *Biochemistry* **2005**, *44*, 8883-8889.

(41) Bellon, S.; Gasparutto, D.; Saint-Pierre, C.; Cadet, J. *Org. Biomol. Chem.* **2006**, *4*, 3831-3837.

(42) Gu, C.; Wang, Y. *Biochemistry* **2004**, *43*, 6745-6750.

(43) Raychaudhury, P.; Basu, A. K. *Biochemistry* **2011**, *50*, 2330-2338.

(44) Gu, C. N.; Zhang, Q. B.; Yang, Z. G.; Wang, Y. S.; Zou, Y. *Biochemistry* **2006**, *45*, 10739-10746.

(45) Yang, Z. G.; Colis, L. C.; Basu, A. K.; Zou, Y. *Chem. Res. Toxicol.* **2005**, *18*, 1339-1346.

(46) Check, C. E.; Faust, T. O.; Bailey, J. M.; Wright, B. J.; Gilbert, T. M.; Sunderlin, L. S. *J. Phys. Chem. A* **2001**, *105*, 8111-8116.

(47) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A. ; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford CT, 2010.

(48) Li, X.; Sanche, L.; Sevilla, M. D. *J. Phys. Chem. A* **2002**, *106*, 11248-11253.

(49) Wetmore, S. D.; Boyd, R. J.; Eriksson, L. A. *Chem. Phys. Lett.* **2001**, *343*, 151-158.

(50) Chen, E. C. M.; Herder, C.; Chen, E. S. *J. Mol. Struct.* **2006**, *798*, 126-133.

(51) Wetmore, S. D.; Boyd, R. J.; Eriksson, L. A. *Chem. Phys. Lett.* **2000**, *322*, 129-135.

(52) Abdoul-Carime, H.; Huels, M. A.; Illenberger, E.; Sanche, L. *Int. J. Mass Spectrom.* **2003**, *228*, 703-716.

(53) Sugiyama, H.; Tsutsumi, Y.; Fujimoto, K.; Saito, I. *J. Am. Chem. Soc.* **1993**, *115*, 4443-4448.

(54) Orlov, V. M.; Smirnov, A. N.; Varshavsky, Y. M. *Tetrahedron Lett.* **1976**, *17*, 4377-4378.

Chapter 4

A Density Functional Theory Study on the Kinetics and Thermodynamics of *N*-glycosidic Bond Cleavage in 5-Substituted 2'-Deoxycytidines

Abstract

B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) density functional theory calculations were employed to explore the kinetics and thermodynamics of gas-phase *N*-glycosidic bond cleavage induced by nucleophilic attack of the C1' carbon with a hydroxide ion in 5-substituted 2'-deoxycytidines. The results showed that, among the 5-substituted 2'-deoxycytidine derivatives examined [XdC, where X = -H (dC), -CH_3 (medC), -CH_2OH (hmdC), -CHO (fmdC), -COOH (cadC), -F (FdC), and -Br (BrdC)], fmdC and cadC exhibited the lowest energy barrier and largest exothermicity for *N*-glycosidic bond cleavage. These results paralleled previously reported nucleobase excision activities of human thymine DNA glycosylase (hTDG) toward duplex DNA substrates harboring a thymine and 5-substituted cytosine derivatives when paired with a guanine. Our study suggests that the inherent chemistry associated with the nucleophilic cleavage of *N*-glycosidic bond constitutes a major factor contributing to the selectivity of hTDG towards 5-substituted dC derivatives. These findings provided novel insights into the role of TDG in active cytosine demethylation.

Introduction

Cytosine methylation at CpG sites is one of the best-characterized epigenetic modifications and is central to a variety of cellular processes including retrotransposon silencing, genomic imprinting, X-chromosome inactivation, regulation of gene expression, and maintenance of epigenetic memory.[1] Aberrant cytosine methylation is linked to imprinting disorders such as Prader-Willi and Angelman syndromes and is implicated in diseases including cancer.[2, 3] In mammalian cells, DNA (cytosine-5)-methyltransferase 1 is known to convert nascent, replication-produced hemimethylated CpG sites to fully methylated ones,[4] and loss or inhibition of this process can give rise to passive DNA demethylation. However, the mechanism for active, DNA replication-independent demethylation in mammals remains poorly defined.[4] It was shown very recently that the ten-eleven translocation 1-3 (Tet 1-3) proteins can successively oxidize meC to hmC,[5, 6] fmC,[7] and caC (Figure 4.1).[7, 8] Additionally, hmC, fmC and caC are found in mammalian DNA,[7-9] and human thymine DNA glycosylase (TDG) can readily excise fmC and caC, but not C, meC or hmC from duplex DNA.[10] These findings provide important evidence to support a newly proposed mechanism of active cytosine demethylation, which involves iterative oxidation of meC to fmC and caC by Tet 1-3, TDG-mediated excision of fmC and caC, and subsequent employment of base excision repair (BER) machinery to yield unmethylated cytosine at CpG sites, though the involvement of direct decarboxylation of caC in active cytosine demethylation cannot be excluded (Figure 4.1A).[11]

TDG is an enzyme that removes thymine (T)[12] or uracil[13] when paired with guanine as well as a variety of cytosine derivatives[14] located at 5'-CpG-3' site in duplex

DNA. Crystal structure analysis revealed that the sequence-specific mismatch recognition of TDG may be attributed to electrostatic interactions of the N1H and $N2H_2$ of the opposing G, relative to the damaged base, with the backbone amides of A274 and P280 within the enzyme.[15, 16] These initial structures were solved for the complexes formed between the catalytic domain of human TDG and double-stranded DNA (dsDNA) containing an abasic site; thus, they did not provide insight into the mechanism of nucleobase recognition. Only very recently were X-ray crystal structures solved for the catalytic domain of human TDG in complex with dsDNA harboring a non-hydrolyzable cadC analog or the corresponding catalytically inactive mutant with cadC-containing DNA.[17] It was found that the active site of human TDG recognizes the 5-carboxylate group of cadC through a network of hydrogen bonding in a small pocket formed by the side chains of A145 and N157, and the backbone atoms of H150, H151, and Y152.[17]

To our knowledge, no computational analysis of the kinetics and thermodynamics associated with the cleavage of the *N*-glycosidic bond in the context of TDG-mediated selective cleavage of caC and fmC has been reported. Thus, we used density functional theory (DFT) calculations to computationally assess the activation energy (E_a) and thermochemistry (E_{rxn}) of nucleobase excision from thymidine (dT) and 5-X-2'-deoxycytidines (XdC), where X = -H (dC), $-CH_3$ (medC), $-CH_2OH$ (hmdC), -CHO (fmdC), -COOH (cadC), -F (FdC), and -Br (BrdC), via hydroxide ion-mediated nucleophilic attack on C1' of 2'-deoxyribose in the gas phase. Results from DFT calculations at the B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) level of theory predicted that, among all the dC derivatives examined, cleavage of the *N*-glycosidic bond in fmdC and cadC exhibits the lowest energy barriers and highest exothermicity. These results are

consistent with the largest k_{max} (maximum rate) for nucleobase excision observed for duplex-DNA harboring these two modified nucleosides measured under single-turnover conditions.[10]

Computational Methods

All calculations were carried out using DFT methods included in the Gaussian 09W package.[18] Geometry optimizations (including transition states) and frequency calculations (without scaling) were performed at the B3LYP/6-31+G(d) level of theory in the gas phase. The transition state structures were confirmed by the presence of only one imaginary frequency. The initial reactant- and product-complexes (R_{TS} and P_{TS}), respectively, were determined based on manual displacements of the optimized transition state (TS) structure. Those initial structures were then optimized [B3LYP/6-31+G(d)] to ensure a minimum, where no imaginary frequency was identified. Basis set superposition error (BSSE) corrections were calculated for $R_{TS,}$ TS, and P_{TS} at the B3LYP/6-311+G(2d,p) level of theory on B3LYP/6-31+G(d) optimized geometries [referred to as B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) within the text] using counterpoise (fragments=2) to account for molecular interactions between basis functions that are within proximity. Single-point energies (SPE) were also determined from B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) calculations.

Our choice of basis sets stemmed from earlier studies showing that 6-31G(d) geometries yield final relative energies in excellent agreement with geometries obtained using the larger basis set, 6-31+G(d,p).[19, 20] These studies also showed that the C3' and

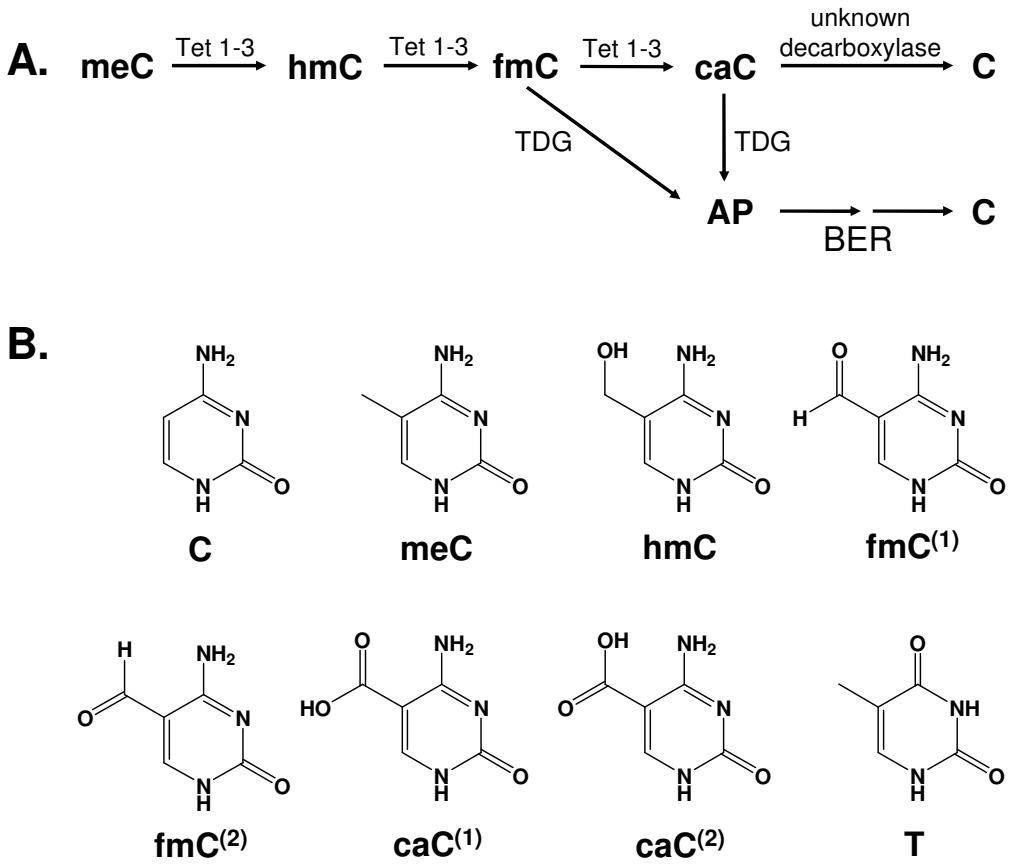


Figure 4.1 Proposed mechanism for active DNA demethylation of meC to C (a); Chemical structures of nucleobases examined in this work (b).

C5' hydroxyl groups of the 2'-deoxyribose interact with the nucleophile and/or nucleobase, which cannot occur in duplex DNA. As such, all C3' and C5' hydroxyl groups in the nucleosides and deoxyribose structures modeled herein were replaced with methoxyl groups for more accurate calculations. Although modified, these structures are still referred to as "2'-deoxy-" within the text. On the grounds that an Asn140-activated H_2O molecule serves as a nucleophile attacking the C1' carbon to induce the *N*-glycosidic bond cleavage,[17] we employed hydroxide ion as the nucleophile for modeling the TDG-mediated cleavage of the pyrimidine nucleosides.

The total energy (in kJ/mol) for R_{TS}, TS, and P_{TS} were calculated as follows, $E = (E_{SPE} + E_{BSSE})_{B3LYP/6\text{-}311+G(2d,p)} + (E_{ZPEC})_{B3LYP/6\text{-}31+G(d)}$, where E_{ZPEC} is the zero-point energy correction (ZPEC) extracted from the frequency calculations. Additionally, the B3LYP/6-31+G(d) optimized TS complex was used to generate electrostatic potential maps with an isosurface of 0.02 electrons/Å^3. Activation energy (E_A) and reaction energy (E_{rxn}) were determined by taking the difference in total energy between TS and R_{TS}, and P_{TS} and R_{TS}, respectively. We also calculated acidity (kJ/mol) in terms of the change in enthalpy (ΔH) at the N1 position of thymine and 5-X-cytosine substituted derivatives at the B3LYP/6-31+G(d) level of theory as follows, $\Delta H = E_{H+} + E_{XC-} - E_{XC}$, where H+, XC, and XC represent proton, deprotonated (N1 position) nucleobase, and neutral nucleobase, respectively.

Results and Discussion

Activation Energy and Reaction Energy

Results from DFT calculations at the B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) level of theory predicted that, among all the dC derivatives examined, the cleavages of *N*-glycosidic bond in fmdC and cadC exhibits the lowest energy barriers and highest exothermicity (Table 4.1 and Figure 4.2). These results are consistent with the largest k_{max} (maximum rate) for nucleobase excision observed for these two modified nucleosides measured under single-turnover conditions, whereas TDG exhibited no activities toward dC, medC or hmdC under physiological conditions.[14]

Using E_a as a benchmark, dT (E_a = 43.1 kJ/mol), fmdC (E_a = 46.8 kJ/mol, for the conformer with lower energy barrier) and cadC (E_a = 46.1 kJ/mol, for the conformer with lower energy barrier) display significantly lower energy barriers for *N*-glycosidic bond cleavage than all the other nucleosides assessed. This result correlates well with the largest k_{max} values obtained for these three nucleosides under single-turnover conditions (Table 4.1). In this vein, the electrostatic potential maps of the transition states obtained for all nucleosides assessed herein showed slightly less electron density on the N1 position of thymine, fmC and caC than the other nucleosides (Figure 4.3). This result suggests that the transition state is more stabilized for these three nucleosides via better delocalization of the negative charge formed on N1 arising from partial glycosidic bond cleavage. Such stabilization of the transition state could be attributed, in part, to the electron-withdrawing nature of the C5 substituents (see more discussion about this below).[10, 14]

Table 4.1 Properties for the thymidine and C5-substituted 2'-deoxycytidine derivatives.

Nucleoside	E_A[a] *kJ/mol*	E_{Rxn}[a] *kJ/mol*	Acidity[a] *kJ/mol*	k_{max} (37°C)[b] *min^{-1}*	k_{max} (22°C)[c] *min^{-1}*
Thymidine	43.1	-194.3	1390	1.83 ± 0.04	0.32 ± 0.01
2'-deoxycytidine (dC)	58.3	-174.9	1436		$1.2\text{x}10^{-5} \pm 0.05\text{x}10^{-5}$
5-methyl-dC	59.6	-173.8	1440		
5-hydroxymethyl-dC	53.7	-178.9	1434		$< 1.4 \text{ x } 10^{-5}$
5-formyl-dC (1)	49.6	-193.6	1370	2.64 ± 0.09	0.61 ± 0.04
5-formyl-dC (2)	46.8	-197.9	1372		
5-carboxy-dC (1)	48.5	-193.4	1383	0.47 ± 0.01	0.14 ± 0.01
5-carboxy-dC (2)	46.1	-196.4	1384		
5-fluoro-dC	54.7	-188.2	1416		0.035 ± 0.004
5-bromo-dC	56.7	-185.5	1402		0.008 ± 0.001

a. this work
b. Reference 10
c. Reference 14

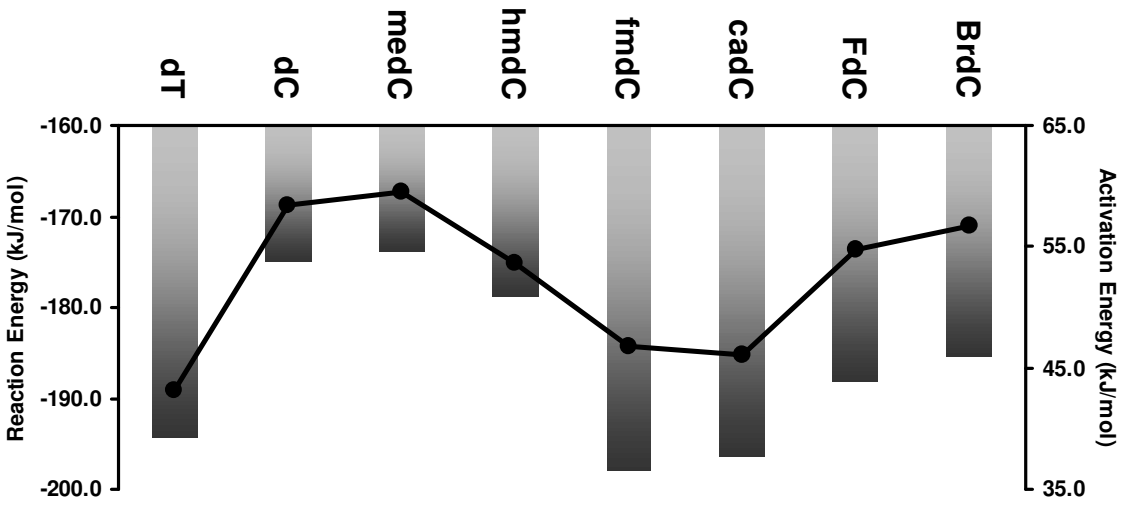


Figure 4.2 The reaction energy (E_{rxn}, bars) and activation energy (E_a, points) for cleavage of the *N*-glycosidic bond in dT and various 5-X-dC derivatives.

The DFT calculation results also predict that the hydroxide-mediated cleavages of the *N*-glycosidic bond are thermodynamically more favorable for dT (E_{rxn} = -194.3 kJ/mol), and the lower-energy conformers of fmdC (E_{rxn} = -197.9 kJ/mol) and cadC (E_{rxn} = -196.4 kJ/mol), than the other nucleosides examined (Table 4.1 and Figure 4.2).

Enthalpy of Deprotonation

We also calculated the enthalpy of deprotonation, that is, the acidity of N1 for each nucleobase of interest, and we found that T, fmC, and caC had enthalpies that were lower than those of other 5-substituted cytosine derivatives examined (Table 4.1). Given that a decrease in deprotonation enthalpy reflects a rise in acidity, and thus, a greater stability of the monoanion that forms upon cleavage of the *N*-glycosidic bond,[20] we report a strong correlation between higher N1 acidity, lower barrier (E_a), greater exothermicity (E_{rxn}), and larger observed k_{max} values.

Molecular Geometry

The glycosidic bond length (N1-C1'), the distance between C1' of the 2'-deoxyribose and the nucleophile (C1'-OH), and the dihedral angle about the glycosidic bond (∠O1'C1'N1C2) are listed in Table 4.2. As expected, there is a decrease in the C1'-OH distance, which is accompanied by a rise in the N1-C1' bond length from the reactant complex to the transition state; however, dT, fmdC, and cadC bear noticeably shorter *N*-glycosidic bond lengths in the transition state structure relative to the other dC-derivatives. In the product complexes, both the C1'-OH and N1-C1' bond lengths are comparable among all nucleosides.

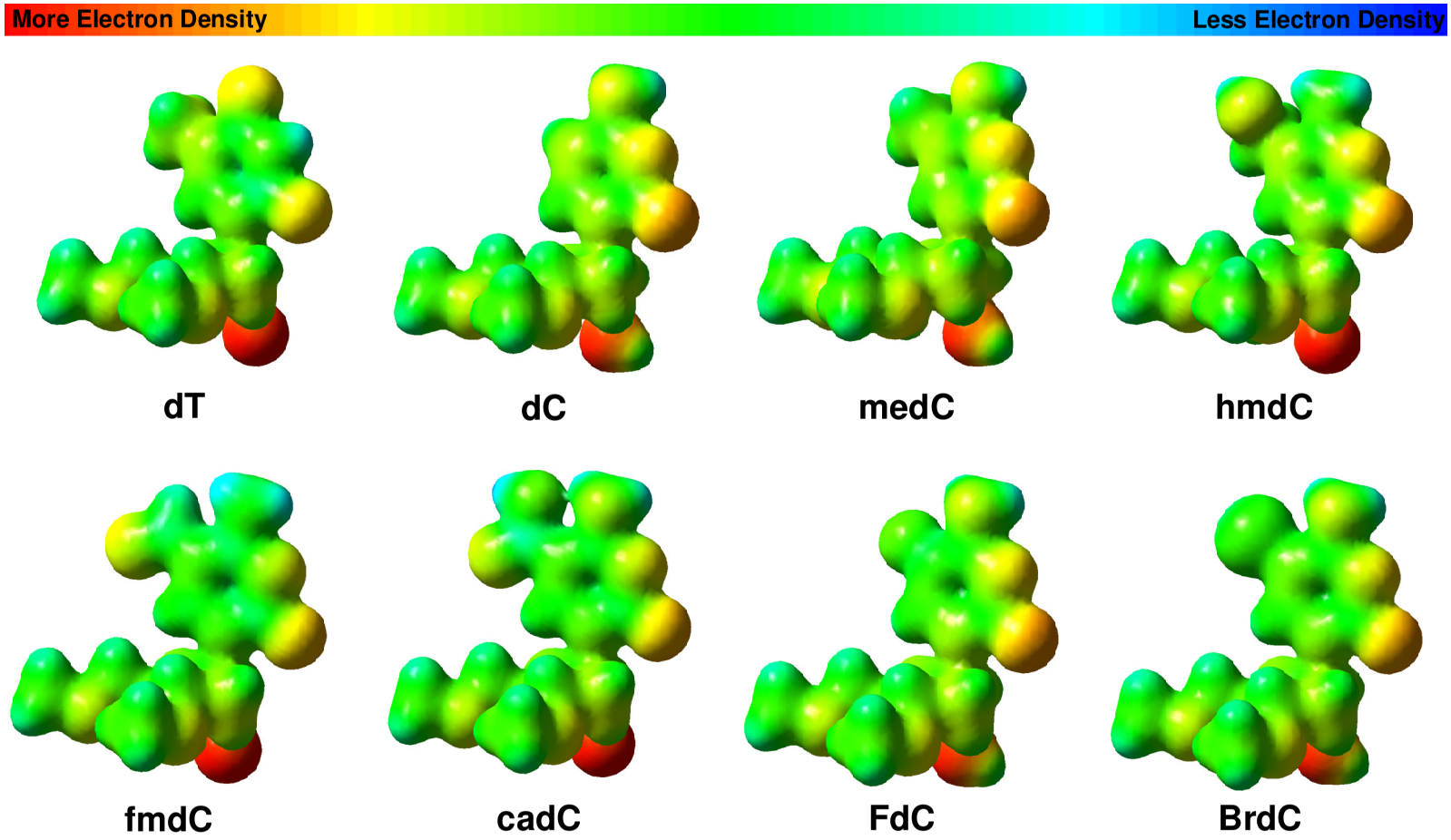


Figure 4.3 The elestrostatic potential maps of transition state structures for the hydroxide-mediated nucleophilic cleavage of dT and 5 substituted 2'-deoxycytidines examined in this study.

Table 4.2 The bond length and dihedral angles determined for the transition state as well as the reactant and product complexes.

	Reactant Complex			Transition Structure			Product Complexes		
	Bond Length (Å)		Dihedral (°)	Bond Length (Å)		Dihedral (°)	Bond Length (Å)		Dihedral (°)
Nucleoside	N1-C1'	C1'-OH	∠(O1' C1' N1 C2)	N1-C1'	C1'-OH	∠(O1' C1' N1 C2)	N1-C1'	C1'-OH	∠(O1' C1' N1 C2)
Thymidine	**1.519**	**3.091**	**197.5**	**1.988**	**2.364**	**175.8**	**4.413**	**1.428**	**27.1**
2'-deoxycytidine (dC)	**1.511**	**3.000**	**193.3**	**2.015**	**2.301**	**180.3**	**4.413**	**1.431**	**35.8**
5-methyl-dC	**1.511**	**2.993**	**192.7**	**2.017**	**2.302**	**179.0**	**4.473**	**1.431**	**28.8**
5-hydroxymethyl-dC	**1.513**	**3.000**	**191.5**	**2.005**	**2.329**	**178.6**	**4.497**	**1.430**	**27.2**
5-formyl-dC (1)	**1.477**	**2.767**	**168.2**	**1.972**	**2.360**	**180.7**	**4.533**	**1.428**	**27.3**
5-formyl-dC (2)	**1.473**	**2.743**	**176.0**	**1.965**	**2.372**	**185.3**	**4.549**	**1.428**	**29.9**
5-carboxy-dC (1)	**1.477**	**2.773**	**171.6**	**1.975**	**2.309**	**182.4**	**4.537**	**1.415**	**31.6**
5-carboxy-dC (2)	**1.476**	**2.765**	**174.4**	**1.971**	**2.368**	**184.1**	**4.538**	**1.485**	**30.8**
5-fluoro-dC	**1.513**	**3.000**	**192.4**	**1.999**	**2.311**	**180.1**	**4.493**	**1.430**	**30.5**
5-bromo-dC	**1.516**	**3.000**	**194.6**	**1.998**	**2.310**	**186.3**	**4.441**	**1.429**	**31.0**

It has been shown that, in order for cleavage of the *N*-glycosidic bond to occur, the nucleophile positioned underneath the 2'-deoxyribose ring in the reactant complex must first move closer to the C1' of the *N*-glycosidic bond.[20] In doing so, a transition state forms where the partial cation generated on the 2'-deoxyribose is stabilized by charges on both the nucleophile and the departing nucleobase. Thus, the net charge is smaller when the *N*-glycosidic bond length in the transition state structure is shorter. This results in a greater stabilizing effect of the monoanion, which may lead to elevated TDG activity.[20]

Conclusions

In this chapter, we employed density functional theory to investigate the kinetics (E_a) and thermodynamics (E_{rxn}) for the hydroxide-mediated nucleophilic cleavage of the *N*-glycosidic bond in dT and 5-substituted dC derivatives. We provided computational evidence in support of a newly proposed mechanism for active cytosine demethylation where meC is iteratively oxidized to hmC, fmC and then caC, followed by TDG-mediated nucleobase excision, and subsequent employment of the BER machinery to generate the unmethylated cytosine.[5-11] Specifically, we demonstrated that selective TDG activity towards fmC and caC compared to C, meC, and hmC might arise partly from the inherent chemistry associated with nucleophilic cleavage of *N*-glycosidic bonds in these nucleosides. These cleavages exhibit lower energy barrier and greater exothermicity than the corresponding cleavages of other dC derivatives examined. We further showed that this dependence might be linked to the length of the *N*-glycosidic bond in the transition state structure, such that cleavage of the departing anionic nucleobase and approaching of

the nucleophile stabilize the partial cation formed on the 2'-deoxyribose, thus rendering the reaction more energetically favorable. This is in contrast to previous reports suggesting that k_{max} for forming the monoanion depends on the electronic substituent constant (σ_m) of -X at the C5 position of the pyrimidine ring.[10, 14] However, this theory has its shortcomings when comparing nucleobases such as fmC (k_{max} = 0.61 min^{-1}) and FC (k_{max} = 0.035 min^{-1}), which have similar σ_m values, 0.35 and 0.34, respectively, but markedly different activity towards TDG. Instead, we reason that assessment of the kinetics, thermodynamics, and molecular geometry involved in the cleavage of the *N*-glycosidic bond provides a more accurate prediction of varied hTDG activity towards a host of 5-substituted dC derivatives.

References

(1) Bird, A. *Genes Dev.* **2002**, *16*, 6-21.

(2) Lindahl, T. *Nature* **1993**, *362*, 709-715.

(3) Robertson, K. D. *Nat. Rev. Genet.* **2005**, *6*, 597-610.

(4) Chen, Z.-x.; Riggs, A. D. *J. Biol. Chem.* **2011**, *286*, 18347-18353.

(5) Ito, S.; D'Alessio, A. C.; Taranova, O. V.; Hong, K.; Sowers, L. C.; Zhang, Y. *Nature* **2010**, *466*, 1129-1133.

(6) Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A. *Science* **2009**, *324*, 930-935.

(7) Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. *Science* **2011**, *333*, 1300-1303.

(8) He, Y.-F.; Li, B.-Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L.; Sun, Y.; Li, X.; Dai, Q.; Song, C.-X.; Zhang, K.; He, C.; Xu, G.-L. *Science* **2011**, *333*, 1303-1307.

(9) Pfaffeneder, T.; Hackner, B.; Truß, M.; Münzel, M.; Müller, M.; Deiml, C. A.; Hagemeier, C.; Carell, T. *Angew. Chem., Int. Ed.* **2011**, *50*, 7008-7012.

(10) Maiti, A.; Drohat, A. C. *J. Biol. Chem.* **2011**, *286*, 35334-35338.

(11) Wu, H.; Zhang, Y. *Genes Dev.* **2011**, *25*, 2436-2452.

(12) Wiebauer, K.; Jiricny, J. *Nature* **1989**, *339*, 234-236.

(13) Neddermann, P.; Jiricny, J. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 1642-1646.

(14) Bennett, M. T.; Rodgers, M. T.; Hebert, A. S.; Ruslander, L. E.; Eisele, L.; Drohat, A. C. *J. Am. Chem. Soc.* **2006**, *128*, 12510-12519.

(15) Maiti, A.; Morgan, M. T.; Drohat, A. C. *J. Biol. Chem.* **2009**, *284*, 36680-36688.

(16) Maiti, A.; Morgan, M. T.; Pozharski, E.; Drohat, A. C. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 8890-8895.

(17) Zhang, L.; Lu, X.; Lu, J.; Liang, H.; Dai, Q.; Xu, G.-L.; Luo, C.; Jiang, H.; He, C. *Nat Chem Biol* **2012**, *8*, 328-330.

(18) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A. ; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford CT, 2010.

(19) Millen, A. L.; Archibald, L. A. B.; Hunter, K. C.; Wetmore, S. D. *J. Phys. Chem. B* **2007**, *111*, 3800-3812.

(20) Millen, A. L.; Wetmore, S. D. *Can. J. Chem.* **2009**, *87*, 850-863.

Chapter 5

A Density Functional Theory Study on the Kinetics and Thermodynamics of *N*-glycosidic Bond Cleavage in 5-Substituted 2'-Deoxyuridines

Abstract

B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) density functional theory calculations were employed to explore the kinetics and thermodynamics of gas- and solution-phase *N*-glycosidic bond cleavage induced by nucleophilic attack of the C1' carbon with a hydroxide ion in 5-substituted 2'-deoxyuridines. The results showed that, among the 5-substituted 2'-deoxyuridine derivatives assessed [XdU, where X = -H (dU), -CH_3 (dT), -CH_2OH (hmdU), -CHO (fmdU), and -COOH (cadU)], fmdU and cadU exhibited the lowest energy barrier and largest exothermicity for *N*-glycosidic bond cleavage. Despite the absence of experimental data for thymine DNA glycosylase (TDG)-mediated cleavage of substrates containing fmdU and cadU, our results for the remaining nucleosides paralleled previously reported nucleobase excision activities of human TDG toward duplex DNA harboring a dU, dT, or hmU when paired with a guanine. Our study predicts that fmdU and cadU are highly active substrates for hTDG and suggests that the inherent chemistry associated with the nucleophilic cleavage of *N*-glycosidic bond may constitute a major factor contributing to the activity of hTDG towards 5-substituted dU derivatives. These findings provided insights into the role of hTDG in removing thymine oxidation products from CpG sites.

Introduction

Oxidative DNA damage induced by both endogenous (i.e., normal aerobic metabolism) and exogenous (i.e., ionizing radiation and UV light) sources has been implicated in aging, cancer, and a host of other degenerative diseases [1, 2]. 5-Hydroxymethyluracil (hmU) is a common product that may arise from the oxidation of the methyl group at the C5 position of thymine (T) [3] or the oxidation of 5-methylcytosine (mC) to 5-hydroxymethylcytosine (hmC) followed by deamination of the latter [4]. It was shown that hmU does not inhibit DNA replication [5], nor does it direct nucleotide misincorporation [6]. In certain bacteriophage, it was found that hmU completely replaces T [7]. In this respect, the formation of hmU from T appears benign; however, the fact that hmU can be repaired in murine cells and tissue, Chinese hamster cells, and human cell lines suggests its presence may have some adverse effects [8-10]. The formation of hmU from mC may have more serious consequences. DNA polymerase recognizes hmU as T and inserts dAMP opposite the lesion, thereby creating a mutation. When oxidative damage occurs specifically at a mCpG site, it may also lead to the disruption of normal cellular processes such as retrotransposon silencing, genomic imprinting, X-chromosome inactivation, and regulation of gene expression [11]. Interestingly, Sowers et al. [12] showed that hmU is removed 60 times faster from hmU:G pair than from hmU:A pair in HeLa cell nuclear extract and at least 10^4 times faster by cloned thymine DNA glycosylase. Along this line, HmU:G pair, which originates from the oxidation and deamination of mC at CpG site, is expected to occur at a frequency that is 10^7 times lower than hmU:A pair, which derives from the oxidation of thymine. On the basis that DNA repair activity

reflects a need to correct harmful genomic perturbations, conversion of mC to hmU may be more frequent than previously considered.

Irrespective of the origin, hmU may be further oxidized to 5-formyluracil (fmU), which was also found to be mutagenic given that it directs DNA polymerase(s) to misincorporate dGMP [13]. It has been suggested that the misincorporation results from an increased acidity of the N3 proton due to the strongly electron-withdrawing nature of the 5-formyl substituent, which may alter the hydrogen bonding interactions on the Watson-Crick base-pairing face [14-17]. It was demonstrated that fmU can be further oxidized to 5-carboxyluracil (caU), which was found to be an inhibitor of *de novo* pyrimidine biosynthesis pathway involving orotate phosphoribosyl transferase and orotidine 5'-phosphate decarboxylase.[18-20] Additionally, it was shown that caU is the primary photo-induced deamination product of 5-carboxylcytosine that was formed from the sequential oxidation of mC.[21] Thus, to maintain genomic integrity, it is important for these oxidation products (i.e., hmU, fmU and cadU) to be removed from DNA.

Thymine DNA glycosylase (TDG) is an enzyme that removes thymine (T) [22], uracil (U) [23], and a variety of cytosine (C) and uracil derivatives [24-27] when paired with guanine located specifically at 5'-CpG-3' site in duplex DNA. In a previous study, we provided computational evidence to suggest that the selective activity of human TDG (hTDG) towards 5-substituted cytosine derivatives may arise from the inherent chemistry associated with the nucleophilic cleavage of the *N*-glycosidic bond [28]. We demonstrated that the assessment of the kinetics, thermodynamics, and molecular geometry involved with the cleavage of the *N*-glycosidic bond provides a more accurate prediction of the observed hTDG activity towards a host of 5-substituted dC substrates.

To our knowledge, no computational analysis of the kinetics and thermodynamics associated with the cleavage of the *N*-glycosidic bond of the thymine oxidation products (hmdU, fmdU and cadU) has been reported. Thus, we used density functional theory (DFT) calculations to computationally assess the activation energy (E_a) and thermochemistry (E_{rxn}) of nucleobase excision from 5-X-2'-deoxyuridines (XdU), where X = -H (dU), -CH_3 (dT), -CH_2OH (hmdU), -CHO (fmdU), and -COOH (cadU) (Figure 5.1), via hydroxide ion-mediated nucleophilic attack on C1' of 2'-deoxyribose in the gas- and solution-phases. Results from DFT calculations at the B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) level of theory predicted that, among all the dU derivatives examined, cleavage of the *N*-glycosidic bond in fmdU and cadU exhibits the lowest energy barriers and highest exothermicity. Despite the absence of experimental k_{max} (maximum rate for nucleobase excision) data for the activity of isolated hTDG towards fmdU and cadU, our results for dU, dT, and hmdU are consistent with reported trends. Thus, we predict that fmdU and cadU are highly favorable substrates for hTDG and would be excised from duplex DNA more readily than hmdU.

Computational Methods

All calculations were carried out using DFT methods included in the Gaussian 09W package [29]. Geometry optimizations (including transition states) and frequency calculations (without scaling) were performed at the B3LYP/6-31+G(d) level of theory in the gas phase. The transition state (TS) structures were confirmed by the presence of only one imaginary frequency. The initial reactant- and product-complexes (R_{TS} and P_{TS}, respectively) were determined based on manual displacements of the optimized TS

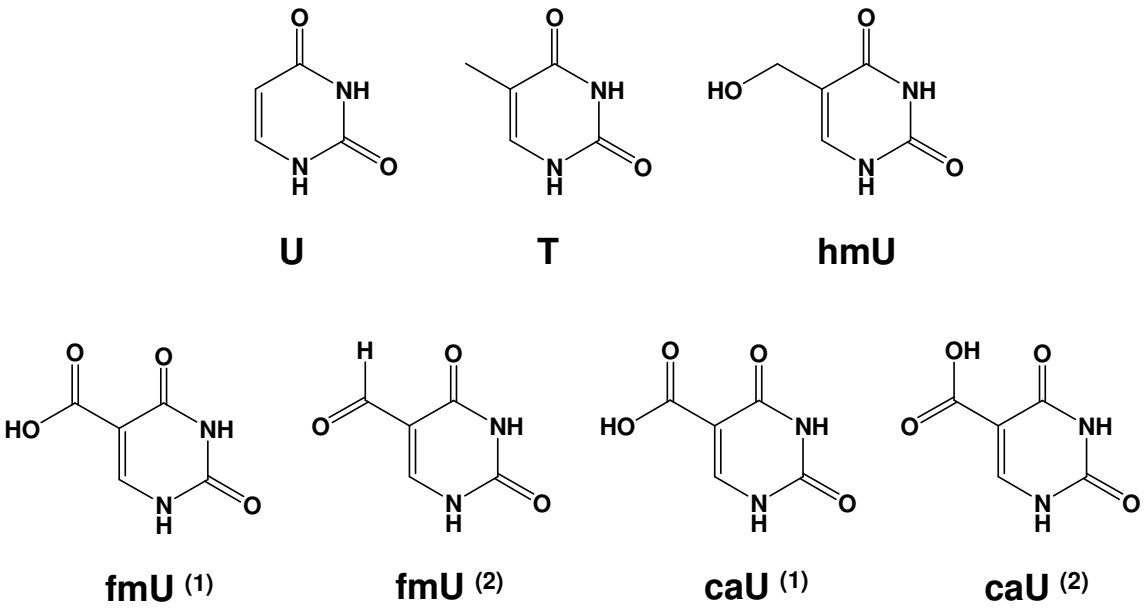


Figure 5.1 Chemical structures of nucleobases examined in this work.

structure. These initial structures were then optimized [B3LYP/6-31+G(d)] to ensure a minimum, where no imaginary frequency was identified. Basis set superposition error (BSSE) corrections were calculated for $R_{TS,}$ TS, and P_{TS} at the B3LYP/6-311+G(2d,p) level of theory on B3LYP/6-31+G(d)-optimized geometries [referred to as B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) within the text] using counterpoise (fragments=2) to account for molecular interactions between basis functions that are within proximity. B3LYP/6-311+G(2d,p) polarizable continuum model (PCM) single-point energy (SPE) calculations were performed on gas-phase geometries in ether (ε=4) to describe the active site environment of the enzyme. Additionally, natural population analyses were conducted on optimized R_{TS} and TS structures at the B3LYP/6-31+G(d) level of theory using PCM methods in ether.

Our choice of basis sets stemmed from earlier studies showing that 6-31G(d) geometries yield final relative energies in excellent agreement with geometries obtained using the larger basis set, 6-31+G(d,p).[30, 31] These studies also showed that the C3' and C5' hydroxyl groups of the 2'-deoxyribose interact with the nucleophile and/or nucleobase, which cannot occur in duplex DNA. Therefore, all C3' and C5' hydroxyl groups in the nucleosides and deoxyribose structures modeled herein were replaced with methoxyl groups for more accurate calculations. Although modified, these structures are still referred to as "2'-deoxy-" within the text. Viewing that an Asn140-activated H_2O molecule serves as a nucleophile attacking the C1' carbon to induce the *N*-glycosidic bond cleavage [32], we employed hydroxide ion as the nucleophile for modeling the TDG-mediated cleavage of the pyrimidine nucleosides.

The total energy (in kJ/mol) for R_{TS}, TS, and P_{TS} were calculated as follows, E = $(E_{SPE} + E_{BSSE})_{B3LYP/6\text{-}311+G(2d,p)} + (E_{ZPEC})_{B3LYP/6\text{-}31+G(d)}$, where E_{ZPEC} is the zero-point energy correction (ZPEC) extracted from frequency calculations. We also calculated acidity (kJ/mol) in terms of the change in enthalpy (ΔH) at the N1 and N3 position of 5-substituted uracil derivatives at the B3LYP/6-31+G(d) level of theory as follows, $\Delta H = E_{H+} + E_{XU^-} - E_{XU}$, where H^+, XU^-, and XU represent proton, deprotonated nucleobase (N1 or N3 position), and neutral nucleobase, respectively.

Results and Discussion

Activation Energy and Reaction Energy

Results from DFT calculations at the B3LYP/6-311+G(2d,p)//B3LYP/6-31+G(d) level of theory predicted that, among all the dU derivatives examined, cleavage of the *N*-glycosidic bond in fmdU and cadU exhibits the lowest energy barriers and highest exothermicity (Table 5.1, Figures 5.2). Although no k_{max} (maximum rate) data for nucleobase excision of fmdU or cadU by hTDG are available, our calculations are consistent with published data of hTDG activity towards dU, dT, and hmdU (Table 5.1) [24]. Moreover, it has been shown that mouse TDG and thermostable TDG from archaeon *Methanobacterium thermoautotrophicum* have relatively large cleavage rates for the fmU:G mispair compared to other uracil derivatives [25]. As such, we predict that fmdU and cadU are highly favorable substrates for hTDG and, when base paired with guanine, they would be excised from duplex-DNA at greater efficiencies than dU, dT, and hmdU.

Upon examination of activation energy, fmdU [E_a = 33.4 kJ/mol (gas) and E_a = 26.1 kJ/mol (ether) for the lowest energy conformer] and cadU [E_a = 33.5 kJ/mol (gas)

and E_a = 46.0 kJ/mol (ether) for the lowest energy conformer] display significantly lower energy barriers for *N*-glycosidic bond cleavage than all the other nucleosides assessed (Table 5.1, Figures 5.2). In our previous study of *N*-glycosidic bond cleavage of 5-substituted 2'-deoxycytidines [28], we found an inverse correlation between E_a and k_{max} values. Despite a lack of experimental data for fmdU and cadU, we note a similar relationship between E_a and k_{max} for dU, dT, and hmdU, particularly in ether, which was used to mimic the non-aqueous environment of the active site of TDG.

The DFT calculation results also predict that the hydroxide-mediated cleavages of the *N*-glycosidic bond are thermodynamically more favorable for fmdU [E_{rxn} = -218.4 kJ/mol (gas) and E_{rxn} = -176.4 kJ/mol (ether) for the lowest energy conformer] and cadU [E_{rxn} = -217.3 kJ/mol (gas) and E_{rxn} = -173.5 kJ/mol (ether) for the lowest energy conformer] than the other nucleosides examined (Table 5.1, Figures 5.2). We again found a strong correlation between lower calculated E_{rxn} (in ether) and increased k_{max} for dU, dT, and hmdU, suggesting that hTDG would exhibit the greatest activity towards fmdU and cadU.

Enthalpy of Deprotonation

We calculated the enthalpy of deprotonation (acidity) at the N1 position of each nucleobase. We found fmU and caU to have notably greater acidities at N1 relative to the other derivatives examined (Table 5.1). Given that a decrease in deprotonation enthalpy reflects a rise in acidity, and thus, a greater stability of the monoanion that forms upon cleavage of the *N*-glycosidic bond [31], we report a relationship between higher

Table 5.1 Properties for the C5-substituted 2'-deoxyuridine derivatives

		Gas-Phase				Ether			
				Acidity				Natural Charge	
Nucleoside	k_{max} [a] *min* $^{-1}$	E_a *kJ/mol*	E_{rxn} *kJ/mol*	N1 *kJ/mol*	N3 *kJ/mol*	E_a *kJ/mol*	E_{rxn} *kJ/mol*	N3H R_{TS}	C1' *TS*
2'-deoxyuridine (dU)	**2.6 ± 0.3**	**42.2**	**-195.9**	**1383**	**1440**	**59.8**	**-149.9**	**0.459**	**0.379**
Thymidine	**0.22 ± 0.04**	**43.1**	**-194.3**	**1390**	**1442**	**62.3**	**-144.3**	**0.458**	**0.380**
5-hydroxymethyl-dU	**1.9 ± 0.2**	**47.0**	**-194.3**	**1392**	**1438**	**61.3**	**-148.0**	**0.458**	**0.380**
5-formyl-dU [(1)]		**36.3**	**-216.6**	**1321**	**1397**	**63.8**	**-177.7**	**0.462**	**0.375**
5-formyl-dU [(2)]		**33.4**	**-218.4**	**1325**	**1390**	**26.1**	**-176.4**	**0.463**	**0.375**
5-carboxy-dU [(1)]		**34.8**	**-216.1**	**1332**	**1403**	**46.2**	**-174.1**	**0.462**	**0.375**
5-carboxy-dU [(2)]		**33.5**	**-217.3**	**1334**	**1404**	**46.0**	**-173.5**	**0.462**	**0.376**

a. Reference 24

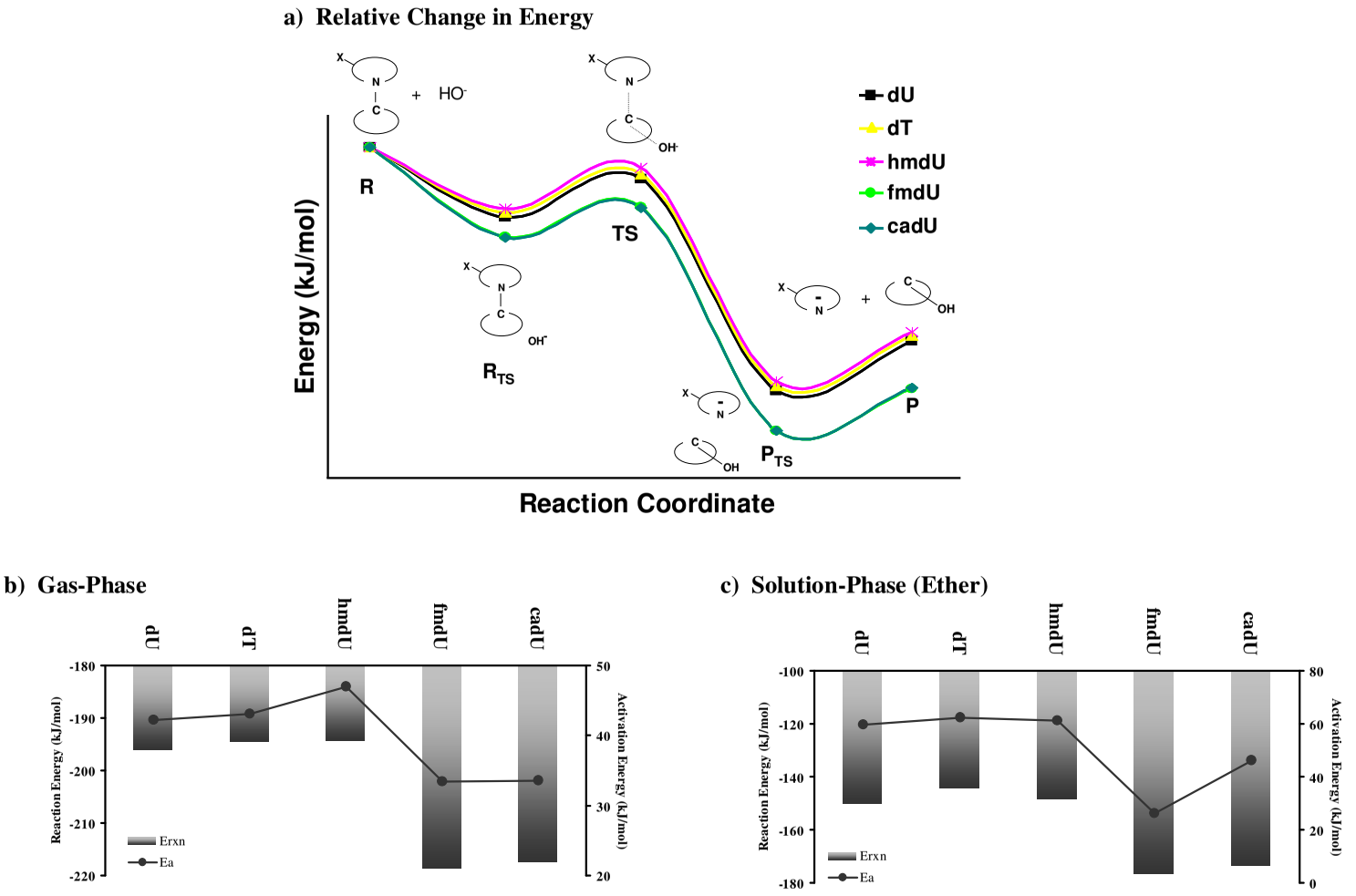


Figure 5.2 Reaction coordinate depicting the relative change in energy (a), the gas-phase (b), and solution (ether) (c) of the reaction energy (E_{rxn}, bars) and activation energy (E_a, points) for the nucleophilic-driven cleavage of the *N*-glycosidic bond in various 5-X-dU derivatives.

N1 acidity, lower barrier (E_a), greater exothermicity (E_{rxn}), and larger observed k_{max} values, which is in good agreement with our previous report [28].

We also calculated the enthalpy of deprotonation at the N3 position of each nucleobase of interest and found that fmU and caU had enthalpies that were significantly lower than U, T, and hmU (Table 5.1). A population analysis of the reactant complexes (R_{TS}) showed that the natural charge on the N3 proton (N3H) is considerably more positive for fmdU and cadU compared to the other nucleosides examined in both gas and ether environments (Table 5.1). Together, these findings are in line with previous reports suggesting that the misincorporation of dGMP opposite fmU results from a change in acidity of N3H due to the strong electron-withdrawing nature of the 5-formyl substituent (σ_m= 0.35) [13, 33]. Additionally, we predict that caU may also direct DNA polymerase to misincorporate dGMP during DNA replication. Interestingly, dT and hmdU have identical N3H charges when calculated in ether (Table 5.1), similar electronic substituent constants (σ_m= -0.07 and σ_m= 0.0, respectively) [33], and are recognized as the same nucleobase by replicative DNA polymerase [5, 6].

Although a trend similar to fmU is predicted for caU, it should be noted that X-ray crystal structures were recently solved for the catalytic domain of hTDG in complex with dsDNA harboring a non-hydrolyzable 5-carboxyl-2'-deoxycytidine (cadC) analog or the corresponding catalytically inactive mutant with cadC-containing DNA [32]. It was found that the carboxyl group is deprotonated within the active site (COO-) and thus lowers the pK_a of N3 of cytosine [34]. It is quite likely that cadU is ionized in the same manner when incorporated into the active site and thus the acidity of N3H may be significantly different than predicted.

Molecular Geometry

The glycosidic bond length (N1-C1'), the distance between C1' of the 2-deoxyribose and the nucleophile (C1'-OH), and the dihedral angle about the glycosidic bond (∠O1'C1'N1C2) are listed in Table 5.2. FmdU and cadU bear noticeably shorter *N*-glycosidic bond lengths in the transition state (TS) structure relative to other dU-derivatives. Although no experimental data are available, comparison of N1-C1' (Table 5.2) and k_{max} (Table 5.1) for dU, dT, and hmdU reveals a relationship between shorter *N*-glycosidic bond lengths and larger k_{max} values. From previous reports [28, 31], this relationship may arise from a stabilizing effect of the partial cation generated on the 2'-deoxyribose by the departing monoanionic nucleobase and the approaching nucleophile. Population analyses in ether confirms that the natural charge on C1' in the TS structure is nearest to zero for fmdU and cadU. Therefore, we report that shorter N1-C1' bond lengths and reduced positive charge on C1' are in excellent agreement with increased hTDG activity (k_{max}), and in turn, smaller E_a, more exothermic E_{rxn}, and lower enthalpy of deprotonation (greater acidity) at N1.

Conclusions

In this chaper, we employed DFT calculations to investigate the kinetics (E_a) and thermodynamics (E_{rxn}) for the hydroxide-mediated nucleophilic cleavage of the *N*-glycosidic bond in 5-substituted dU derivatives as it pertains to oxidative DNA damage. Specifically, we demonstrated that hTDG activity towards dU, dT, and hmdU might arise partly from the inherent chemistry associated with nucleophilic cleavage of *N*-glycosidic bonds in these nucleosides. These findings are in line with our earlier report [28]. Our

Table 5.2 The bond length and dihedral angles determined for the transition state as well as the reactant and product complexes.

	Reactant Complexes			Transition Structure			Product Complexes		
	Bond Length (Å)		Dihedral (°)	Bond Length (Å)		Dihedral (°)	Bond Length (Å)		Dihedral (°)
Nucleoside	N1-C1'	C1'-OH	∠(O1' C1' N1 C2)	N1-C1'	C1'-OH	∠(O1' C1' N1 C2)	N1-C1'	C1'-OH	∠(O1' C1' N1 C2)
2'-deoxyuridine (dU)	1.518	3.000	196.3	1.983	2.354	179.9	4.397	1.428	29.3
Thymidine	1.519	3.091	197.5	1.988	2.364	175.8	4.413	1.428	27.1
5-hydroxymethyl-dU	1.490	2.827	175.0	1.986	2.344	182.5	4.420	1.428	33.0
5-formyl-dU [(1)]	1.486	2.753	172.8	1.942	2.391	181.5	4.474	1.426	23.1
5-formyl-dU [(2)]	1.483	2.739	178.2	1.938	2.405	185.7	4.434	1.426	39.6
5-carboxy-dU [(1)]	1.487	2.758	175.0	1.944	2.393	183.9	4.480	1.427	38.5
5-carboxy-dU [(2)]	1.488	2.743	178.6	1.942	2.398	186.6	4.503	1.426	32.6

calculation results also allow us to predict that hTDG may possess greater activity towards fmU and caU than dU, dT, or hmdU given that the cleavage of *N*-glycosidic bond in these two nucleosides exhibited the lowest energy barriers and greatest exothermicity compared to the other nucleosides examined herein. This suggests that the removal of hmU by hTDG would be much slower than its subsequent oxidative products (i.e., fmU and caU). Although fmU and caU have not been reported as oxidation products of meC at 5'-CpG-3' sites in cellular DNA, the dramatic increase in HeLa cell extract- [12], cloned TDG- [12], and hSMUG- [35, 36] activity towards the hmU:G mispair, compared to hmU:A, suggests the presence of fmU:G and caU:G is highly plausible. The absence of evidence to date may be attributed to lack of effort in the identification of these bases at 5'-CpG-3' sites, as well as their rapid removal as predicted in this work, and demonstrated *in vitro* with ODN [25, 36].

We further established that this dependence might be linked to the length of the *N*-glycosidic bond in the TS structure, such that cleavage of the departing anionic nucleobase and approaching of the nucleophile stabilize the partial cation formed on C1' of the 2'-deoxyribose, thus rendering the reaction more energetically favorable. This is in contrast to previous reports suggesting that k_{max} for forming the monoanion depends on the electronic substituent constant (σ_m) of -X at the C5 position of the pyrimidine ring.[24, 26] Additionally, we provide evidence to support the notion that miscoding of fmU may result from increased acidity of the N3 proton (N3H) due to the electron-withdrawing nature of the formyl group. Based on our previous work [28] and the data presented herein, we maintain our position that assessment of the kinetics, thermodynamics, and molecular

geometry involved in the cleavage of the *N*-glycosidic bond provides a more accurate prediction of varied hTDG activity towards a host of 5-substituted pyrimidine derivatives.

References

(1) Lindahl, T. *Nature* **1993**, *362*, 709-715.

(2) Loeb, L. A.; Christians, F. C. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **1996**, *350*, 279-286.

(3) Teebor, G. W.; Frenkel, K.; Goldstein, M. S. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 318-321.

(4) Cannon-Carlson, S. V.; Gokhale, H.; Teebor, G. W. *Journal of Biological Chemistry* **1989**, *264*, 13306-13312.

(5) Levy, D. D.; Teebor, G. W. *Nucleic Acids Res.* **1991**, *19*, 3337-3343.

(6) Zhang, Q. M. *Int. J. Radiat. Biol.* **1999**, *75*, 59-65.

(7) Kallen, R. G.; Marmur, J.; Simon, M. *J. Mol. Biol.* **1962**, *5*, 248-&.

(8) Hollstein, M. C.; Brooks, P.; Linn, S.; Ames, B. N. *Proc. Natl. Acad. Sci. U. S. A.* **1984**, *81*, 4003-4007.

(9) Boorstein, R. J.; Levy, D. D.; Teebor, G. W. *Mutat. Res.* **1987**, *183*, 257-263.

(10) Boorstein, R. J.; Chiu, L.-N.; Teebor, G. W. *Nucleic Acids Res.* **1989**, *17*, 7653-7661.

(11) Bird, A. *Genes Dev.* **2002**, *16*, 6-21.

(12) Rusmintratip, V.; Sowers, L. C. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 14183-14187.

(13) Kasai, H.; Iida, A.; Yamaizumi, Z.; Nishimura, S.; Tanooka, H. *Mutat. Res. Lett.* **1990**, *243*, 249-253.

(14) Masaoka, A.; Terato, H.; Kobayashi, M.; Ohyama, Y.; Ide, H. *J. Biol. Chem.* **2001**, *276*, 16501-16510.

(15) Privat, E. J.; Sowers, L. C. *Mutat. Res.* **1996**, *354*, 151-156.

(16) Terato, H.; Masaoka, A.; Kobayashi, M.; Fukushima, S.; Ohyama, Y.; Yoshida, M.; Ide, H. *J. Biol. Chem.* **1999**, *274*, 25144-25150.

(17) Yoshida, M.; Makino, K.; Morita, H.; Terato, H.; Ohyama, Y.; Ide, H. *Nucleic Acids Res.* **1997**, *25*, 1570-1577.

(18) Berthod, T.; Cadet, J.; Molko, D. *J. Photochem. Photobiol., A* **1997**, *104*, 97-104.

(19) Clough, D. W.; Wigdahl, B. L.; Parkhurst, J. R. *Antimicrob. Agents Chemother.* **1978**, *14*, 126-131.

(20) Thornburg, L. D.; Lai, M. T.; Wishnok, J. S.; Stubbe, J. *Biochemistry* **1993**, *32*, 14023-14033.

(21) Privat, E.; Sowers, L. C. *Chemical Research in Toxicology* **1996**, *9*, 745-750.

(22) Wiebauer, K.; Jiricny, J. *Nature* **1989**, *339*, 234-236.

(23) Neddermann, P.; Jiricny, J. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 1642-1646.

(24) Bennett, M. T.; Rodgers, M. T.; Hebert, A. S.; Ruslander, L. E.; Eisele, L.; Drohat, A. C. *J. Am. Chem. Soc.* **2006**, *128*, 12510-12519.

(25) Liu, P.; Burdzy, A.; Sowers, L. C. *DNA Repair* **2003**, *2*, 199-210.

(26) Maiti, A.; Drohat, A. C. *J. Biol. Chem.* **2011**, *286*, 35334-35338.

(27) Morgan, M. T.; Bennett, M. T.; Drohat, A. C. *J. Biol. Chem.* **2007**, *282*, 27578-27586.

(28) Williams, R. T.; Wang, Y. *Biochemistry* **2012**, *51*, 6458-6462.

(29) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J., J. A. ; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Keith, T.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford CT, 2010.

(30) Millen, A. L.; Archibald, L. A. B.; Hunter, K. C.; Wetmore, S. D. *J. Phys. Chem. B* **2007**, *111*, 3800-3812.

(31) Millen, A. L.; Wetmore, S. D. *Can. J. Chem.* **2009**, *87*, 850-863.

(32) Zhang, L.; Lu, X.; Lu, J.; Liang, H.; Dai, Q.; Xu, G.-L.; Luo, C.; Jiang, H.; He, C. *Nat Chem Biol* **2012**, *8*, 328-330.

(33) Hansch, C.; Leo, A.; Unger, S. H.; Kim, K. H.; Nikaitani, D.; Lien, E. J. *J. Med. Chem.* **1973**, *16*, 1207-1216.

(34) Sumino, M.; Ohkubo, A.; Taguchi, H.; Seio, K.; Sekine, M. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 274-277.

(35) Baker, D.; Liu, P.; Burdzy, A.; Sowers, L. C. *Chem. Res. Toxicol.* **2001**, *15*, 33-39.

(36) Masaoka, A.; Matsubara, M.; Hasegawa, R.; Tanaka, T.; Kurisu, S.; Terato, H.; Ohyama, Y.; Karino, N.; Matsuda, A.; Ide, H. *Biochemistry* **2003**, *42*, 5003-5012.

Chapter 6

Summary and Future Direction

In this dissertation, mass spectrometry and density functional theory methods were used to characterize a variety of DNA modifications, including 1,2-GpG, 1,2-ApG, and 1,3-GpXpG cisplatin-containing adducts (Chapter 2); 5-X-uracils and 5-X-cytosines, where X = H, CH_3, F, Cl, Br, and I (Chapter 3); 5-X-cytidines, where X = -H (dC), -CH_3 (medC), -CH_2OH (hmdC), -CHO (fmdC), -COOH (cadC), -F (FdC), and -Br (BrdC) (Chapter 4); and 5-X-uridines, where X = -H (dC), -CH_3 (dT), -CH_2OH (hmdU), -CHO (fmdU), and -COOH (cadU) (Chapter 5).

In Chapter 2, we used MS and MS/MS to characterize cisplatin-bearing intrastrand cross-links in DNA and the corresponding ODNs modified with $^{15}N_2$-labeled cisplatin. We also employed LC-MS/MS to assess the digestion products of these ODNs upon enzymatic hydrolysis. The availability of the ODNs carrying the stable isotope-labeled lesions, MS/MS analyses of the cisplatin-modified ODNs, and the characterization of the enzymatic digestion products of these ODNs set the stage for the LC-MS/MS quantification of the 1,2-GpG, 1,2-ApG, and 1,3-GpXpG adducts. Future work should include quantification of these lesions in a variety of cell lines for the purpose of monitoring adduct levels as it pertains to initial treatment, repair, and its implications in cellular drug resistance. Additionally, these characterizations may be extended to include the detection of the 5'-GpC/GpC-3' interstrand cross-link. Although this lesion represents less than 1% of all adducts, interstrand cross-links are considered

absolute blocks to DNA replication and transcription machineries, which may have a profound affect on the drug's mechanism of action.

In Chapter 3, we utilized DFT methods to assess the mechanism of formation of guanine-pyrimidine intrastrand cross-links that formed upon UV irradiation of DNA containing a 5-X-uracil or 5-X-cytosine, where X = Br or I. Here, we provided theoretical evidence that suggests these cross-link products result upon electron transfer and subsequent heterolytic cleavage of the C5-X bond. Although much work has been done in this area, the mechanistic aspects of intrastrand crosslinking photoreactions have not been well studied. As such, further mechanistic insights could be gained from *in vitro* experiments comparing single- and double-stranded ODN harboring 5-halo-dU or 5-halo-dC moieties; and a computational assessment of the kinetics and thermodynamics of the transition state structure(s) of each cross-link dimer.

In Chapter 4 we implemented DFT methods to explore the energetics associated with the nucleophilic-driven cleavage of the *N*-glycosidic bond in potential thymine DNA glycosylase (TDG) substrates as it pertains to a newly proposed mechanism for active cytosine demethylation. Our study suggests that the inherent chemistry (i.e. kinetics, thermodynamics, and molecular geometry) associated with this cleavage constitutes a major factor contributing to the observed selectivity of TDG towards 5-substituted dC derivatives. Given that this is a topic of intense current interest, many opportunities exist to further this line of research. To this end, ongoing experimental research in our lab includes assessing the formation of these 5-substituted cytosine derivatives, namely hmC, fmC, and caC, in cells and tissues. Calculations surveying the kinetics and

thermodynamics of 5-X-uracil nucleobases, where X = -H (dU), -CH_2OH (hmdU), -CHO (fmdU), -COOH (cadU), -F (FdU), and -Br (BrdU), would serve as a value tool for rationalizing the observations made *in vivo*.

Based on our previous finding, we extended our calculations in Chapter 5 to include a survey of the kinetics and thermodynamics associated with the hydroxide-mediated nucleophilic cleavage of the *N*-glycosidic bond in 5-substituted dU derivatives as it pertains to oxidative DNA damage. Our study predicts that fmdU and cadU are highly active substrates for human TDG (hTDG), which again suggests that the inherent chemistry associated with the nucleophilic cleavage of *N*-glycosidic bond may constitute a major factor contributing to the activity of hTDG towards these and other substrates. Given the absence of experimental data to support these claims, ongoing work in our lab currently involves the following: (1) organic synthesis of 5-X-uracil phosphoramidites, where X = -CH_2OH (hmdU), -CHO (fmdU), -COOH (cadU); (2) incorporation of the phosphoramidites into synthetic DNA at a 5'-CpG-3' site; and (3) examination of the glycosylase activity of hTDG towards these oxidative modifications in synthetic DNA. Together, the theoretical and experimental findings will strengthen our understanding of the selective activity of hTDG towards a host of substrates.